INFORMATION RESOURCE GUIDE

Computer, Internet and Network Systems Security

An Introduction to Security

Security Manual

Compiled By:

S.K.PARMAR, Cst

N.Cowichan Duncan RCMP Det
6060 Canada Ave., Duncan, BC
250-748-5522

sunny@seaside.net

Table of Content

Forward

This manual is an effort to assist law enforcement agencies and other computer crime investigators by providing a resource guide compiled from the vast pool of information on the Internet. This manual is not intended to replace any formal training or education. This manual should be used as a supplemental guide to reference too. It was not my intention to compile this manual to provide a specific solution for investigators. This was intended to provide a general overview, which would assist in helping to developing a solution. This solution does not have to be hardware or software based. Today policy-based protection can also be incorporated into hardware and software systems.

I would like to thank all the authors, and organizations that have provided me with materials to compile this manual. Some of the material contained in this manual were a part of a larger document. It is strongly recommended that if anyone has an interest in learning more about a particular topic to find these documents on the Internet and read them.

A very special thanks to:

Dr. Bill Hancock Network-1 Security Solutions, Inc.
(hancock@network-1.com)

who played an active role in the modeling of this manual.

Finally, please respect the copyrights of the original authors and organizations and give them credit for their work.

Any questions or concerns can be directed to me c/o

RCMP Duncan Detachment
6060 Canada Ave., Duncan, BC
CANADA V9L 1V3
ATN: Cst. S.K.PARMAR

Telephone number 250-748-5522
Email: sunny@seaside.net

SUNNY

1.1 Basic Internet Technical Details

The Internet utilizes a set of networking protocols called TCP/IP. The applications protocols that can be used with TCP/IP are described in a set of Internet Engineering Task Force (IETF) RFCs (Request For Comment). These documents describe the "standard" protocols and applications that have been developed to support these protocols. Protocols provide a standard method for passing messages. They define the message formats and how to handle error conditions. Protocols are independent of vendor network hardware, this allows communication between various networks with different hardware as long as they communicate (understand) the same protocol. The following diagram provides a conceptual layering diagram of the protocols.

1.1.1 TCP/IP : Transmission Control Protocol/Internet Protocol

TCP/IP is used to facilitate communication within a network of diverse hardware technology. Information is broken into packets (usually in the range of 1-1500 characters long) to prevent monopolizing of the network. TCP is a transport level protocol which allows a process on one computer to send data to a process on another computer. It is a connection oriented protocol which means that a path must be established between the two computers. IP defines the datagram, the format of the data being transferred throughout the network and performs connectionless delivery. Connectionless delivery requires each datagram to contain the source and destination address and each datagram is processed separately. TCP takes the information, and breaks it into pieces called packets, numbers the packets, and then sends them.

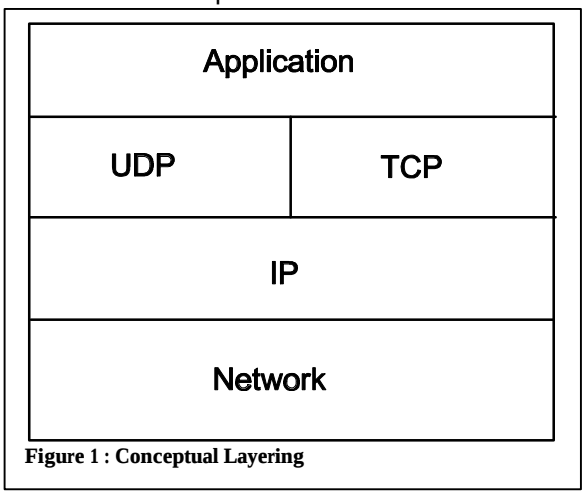


Figure 1 : Conceptual Layering

The receiving computer collects the packets, takes out the data and puts them in the proper order. If something is missing, the receiving computer asks the sender to retransmit. The packet sent also contains a checksum which is used to find errors that may have occurred during transmission. If the receiving computer notices that an error has occurred when it computes and compares the checksum, it throws that packet away and asks for a retransmission. Once everything is received, the data is passed to the proper application (e.g. e-mail).

1.1.2 UDP:User Datagram Protocol

The UDP has less overhead and is simpler than TCP. The concept is basically the same except that UDP is not concerned about lost packets or keeping things in order. It is used for short messages. If it does not receive a response, it just resends the request. Thjs type of protocol transfer method is called a “connectionless protocol.”

1.1.3 Internet Addressing

All computers on the Internet must have a distinct network address to be able to efficiently communicate with each other. The addressing scheme used within the Internet is a 32 - bit address segmented into a hierarchical structure. IP addresses consist of four numbers, each less than 256 which are separated by periods. (#.#.#.#) At the lowest level, computers communicate with each other using a hardware address (on LANs, this is called the Medium Access Control or MAC address). Computer users, however, deal with 2 higher levels of abstraction in order to help visualize and remember computers within the network. The first level of abstraction is the IP address of the computer (e.g. 131.136.196.2) and the second level is the human readable form of this address (e.g. manitou.cse.dnd.ca). This address scheme is currently under review as the address space is running out. Address Resolution Protocol (ARP) can be used by the computer to resolve IP addresses into the corresponding hardware addresses.

1.1.4 Types of Connections and Connectors

There are two types of computer hosts connected to the Internet: server hosts and client hosts. The server host can be described as an “information provider”. This type of host contains some type of resource or data which is available to other hosts on the Internet. The second type of host connected to the Internet is the client host which can be described as an “information retriever”. The client host will access resources and data located on the server hosts, but usually will not provide any resources back to the server host.

Both server and client host computers can be connected to the Internet by various methods that offer different communication capabilities dependent on varied communications surcharges.

Direct Internet Connections: A computer connected directly to the Internet via a network interface will allow the user the highest internetwork functionality. Each computer connected in this manner must also have a unique Internet (IP) address. This type of connection is also the most expensive.

Serial Internet Connections: Another type of connection offering most communications capabilities is a SLIP (Serial Line Internet Protocol) or PPP (Point to Point Protocol) connection. These two connection schemes offer similar services: full network and application capability over a serial (modem) line. Since this connection offers full TCP/IP and ICMP functionality each computer configured in this manner requires its own IP address. This type of connection is an on-demand service, at slower speeds, that therefore reduces communications charges, however all TCP/IP and Internet vulnerabilities remain when the connection is "live".

An important point for the network security investigator to remember is that most dial-up TCP connections, either SLIP or PPP, assign the IP address to a connected machine dynamically. This means that when a system dials-up to the Internet Service Provider (ISP), the ISP assigns an IP address at that point. It also means that the address for the dialer may change each and every time the system connects. This can cause serious problems for the investigator when attempting to trace access back through firewall and router logs for specific IP addresses. You will need to work closely with the victim and the ISP to properly track which system was assigned a particular IP address when the system connected to the ISP at a particular point in time.

Host Access Connections: The most limited type of network access is available as a user account on a host which is directly connected to the Internet. The user will then use a terminal to access that host using a standard serial connection. This type of connection is usually the most inexpensive form of access.

Sneaker-Net Connections: This type of connection is by far the most limiting, since the computer has no electrical connection to the Internet at all. This type of connection is the most secure because there is no direct access to the user's computer by a hacker. If information and programs are required on the computer they must be transferred from a networked computer to the user's computer via magnetic media or manually.

All computers with direct, SLIP, and PPP connections must have their own IP address, and their security administrators must be aware of the vulnerability concerns associated with these connections. Communications channels work both ways: a user having access to the Internet implies that the Internet also has access to that user. Therefore, these computers must be protected and secured to ensure the Internet has limited access. A terminal user calling using an Internet host has fewer concerns since the host is where the Internet interface lies. In this situation the host must take all necessary security precautions.

To connect the various sub-networks and pieces of the Internet together, hardware equipment is required. The following are definitions of the various terms which are use to describe this equipment.

Repeater	A repeater is a hardware device which is used to connect two Local Area Segments that use the same physical level protocol. The repeater will copy all bits from one network segment to another network segment. This device will not make any routing decisions at all, and will not modify the packets. This device operates at layer 1 (Physical) of the OSI Network Model. A repeater may also be used to connect specific workstations in a physically local area to each other. All units connected to a repeater “see” each other’s traffic on the network. Repeaters are very often used on networks like Ethernet/802.3 networks and very commonly available at most computer stores at a low price.
Modem	A modem is a device which will convert between the digital signal structures that computers require and the analog voltage levels that are used by telephone services. The term MODEM stands for MOdulator DEModulator. A modem operates at level 1 (Physical) of the OSI Network Model and therefore does not modify the data packets or make any routing decisions. Modems are used to connect two computers together over standard phone lines (usually for on-demand services). Current MODEM speeds range from 50 bits per second to over 56 thousand bits per second (56kbps).
Bridge	A bridge is a device which is used to connect two Local Area Networks that use the same LAN framing protocol (such as Ethernet or token ring). The bridge acts as an address filter by picking up packets from one LAN segment and transferring them to another IF the bridge recognizes that the packets need to travel from one LAN to the other. If

	the communicating source system and destination system are on the same side of the bridge, the bridge will not forward the frame to the other side of the bridge.. The bridge makes no modification to any packets it forwards, and the bridge operates at layer 2 (data-link) of the OSI Network Model.
Router	A router is a device that is used to connect two or more LAN, MAN or WANsegments that may or may not use the framing protocols. Since the router operates at level 3 (Network) of the OSI Network Model it is able to make routing decisions based on the destination network address (IP address for the Internet). Routers will sometimes have filtering capability included. In this case a router might be used as a packet filter to enhance security and/or reduce traffic flow throughout the network that does not need to traverse all locations on the network (described below). Some very large routers at larger network sites can interconnect dozens of different types of network framing formats.
Gateway	A gateway is a device which will interconnect two network segments which utilize different communications architectures. Gateways typically function on a program-type by program-type (application) basis.The gateway maps (or translates) data from one application to another application and as such operates at level 7 (Application) of the OSI Network Model.
Packet filter	Packet filtering is a capability usually added to routers, but can be implemented in host or firewall systems as well. Packet filtering applies a set of filters (or rules of traversal) to all packets entering or leaving the filtering mechanism that enable the router to decide whether the packet should be forwarded or disregarded. For instance, security configurations may add address filters for certain ranges of addresses to keep traffic from roaming all over a network or to keep undesireable addresses from accessing resources that are restricted in access.
Firewall	A firewall is a description of a system (one or more pieces of hardware) that acts as a barrier between two or more network segments. A firewall can be used to provide a barrier between an internal network and the Internet. A firewall can be considered the technical implementation of a security policy. The firewall upholds the security policy of a network when connecting that network to a second network which has a less stringent security policy.
Cyberwall	A cyberwall is similar in scope to a firewall, but instead of offering perimeter defense filtering between two or more networks, cyberwalls are typically installed on desktop and server systems on the inside network at a corporate site. Cyberwalls provide a defensive barrier to attacks on mission critical systems on internal networks and help “harden” the operating system environment from a network

attack. Some cyberwalls also include intrusion detection software to allow the system to detect an attack of specific types in progress and effect some levels of defense against them.

Readers are cautioned that these terms are not always used in a consistent manner in publications which can cause confusion or misconceptions.

1.1.5 Routing

There are two types of routing used by the Internet: source routing and dynamic routing. The Internet is a very robust networking system. The network routers will automatically (dynamically) send out messages to other routers broadcasting routes to known domains and addresses. If a network or router goes down, packets can be dynamically rerouted to the destination. The user does not usually know how a packet will be routed to the destination. The packet could be rerouted through an untrusted network and intercepted. A router connected to the Internet should be configured to ignore dynamic routing changes and the routing tables should remain static. If the routing tables must be changed, then they should be changed by the network administrator after understanding the reasons for the changes. Unfortunately this is not usually convenient for Internet connected routers. This is another example of when a tradeoff must be made. If the router is configured in this manner then the dynamic routing that the Internet depends on would be disabled. In this situation your network could be cut off (completely or partially) until the Network Administrator makes the required changes in the routing tables.

The second type of routing is known as source routing. In this method of routing a user is able to define a route for the packet between the source and destination. All packets returning to the destination will follow the route information given. A hacker can use a source routed packet to spoof another address. Computers and routers connected to external networks should be configured to ignore source routed packets.

1.2 Internet Applications and Protocols

The Internet is a global collection of networks all using the TCP/IP network protocol suite to communicate. The TCP/IP protocols allow data packets to be transmitted, and routed from a source computer to a destination computer. Above this set of protocols reside the applications that allow users to generate data packets. The following sections describe some of the more common applications as well as some security vulnerabilities and concerns.

1.2.1 ARCHIE

Archie is a system for locating public files available via anonymous ftp (see ftp for vulnerability information). A program is run by an Archie site to contact servers with public files and the program builds a directory of all the files on the servers. Archie can then be used to search the merged directories for a filename and will provide a list of all the files that match and the servers on which the files reside. Public Archie servers are available and can be accessed using telnet, e-mail or an Archie client. Once the filename/server pair has been found using Archie, ftp can be used to get the file from the server. Archie can be used to find security related information(e.g. if one looks up *firewall,* Archie will give all the matches and locations for information on firewalls). Archie is limited in that it can only match on filenames exactly (e.g. if the file contains information on firewalls but the author named it *burnbarrier*, Archie will not find it if the search was for firewalls).

Archie can be exploited to locate anonymous ftp sites that provide world writable areas that can then be used to store and disseminate illegal versions of software. In this case, a hacker uses the Internet tool to gain legitimate access to the database and then misuse the information.

1.2.2 DNS — DOMAIN NAME SYSTEM

DNS is a hierarchical, distributed method or organizing the name space of the Internet. It is used to map human readable host names into IP addresses and vice-versa. A host sends a User Datagram Protocol (UDP) query to a DNS server which either provides the IP address or information about a smarter server than itself. Different groups are given the responsibility for a subset or subsets of names. The number of names in each group gets larger from left to right. For example: cse.dnd.ca, each level of the system is called a domain, cse represents the domain of the Communications Security Establishment which is smaller and within the dnd - Department of National Defense domain. The dnd domain is within the ca - Canada domain. The elements of the domain are separated by periods. Queries can also be made using TCP (port 53) and are called zone transfers. Zone transfers are used by backup servers to obtain a copy of their portion of the name space. Zone transfers can also be used by hackers to obtain lists of targets. The Computer Emergency Response Team (CERT) advises that access to this port be only permitted from known secondary domain servers. This prevents intruders from gaining additional information about the system connected to the local network.

1.2.3 E-MAIL — ELECTRONIC MAIL

Electronic mail is probably the most widely used application on the Internet. Messages are transported using a specific message format and the simple mail transport protocol (SMTP). This protocol offers no security features at all. E-mail messages can be read by a hacker residing on the network between the source and destination of the message. As well, SMTP e-mail messages can be forged or modified very easily. The SMTP protocol offers no message integrity or sender authentication mechanisms.

Some security and a higher level of trust can be provided to SMTP by applying some cryptographic measures to the message. If message integrity or sender authentication are required then the application of a digital signature is called for. A digital signature allows a user to authenticate the e-mail message just as a written signature authenticates a document in today's paper world. Message confidentiality can be obtained by applying an encryption algorithm to the message prior to sending it.

1.2.4 SMTP — SIMPLE MAIL TRANSPORT PROTOCOL

SMTP is an application level protocol used to distribute e-mail messages between computers. This protocol is very simple and understands only simple text based messages and commands. All messages transferred between computers are in ASCII form and are unencrypted. The message is available to everyone in the path that the message takes. There is no method of verifying the message source or ensuring the message integrity, this must be done at a higher level using another protocol such as PEM.

A common implementation of the SMTP protocol is found in the UNIX sendmail facility. This program has a very colourful security history. Sendmail is an extensive

program which allows remote computers more access than required to drop off e-mail.

SMTP is also commonly implemented in Post Office Protocol version 3 servers (also known as POP3) and the new IMAP4 protocol used on newer e-mail servers on Internet.

1.2.5 PEM — PRIVACY ENHANCED MAIL

PEM is a set of standards for adding a security overlay to Internet e-mail providing message confidentiality and integrity. This set of standards describes a security protocol that can be used above the common Simple Mail Transport Protocol (SMTP) or the UNIX-to-UNIX Copy Protocol (UUCP). The PEM security enhancements provide three security services: message integrity, message origin authentication, and message confidentiality. The PEM enhancements can be used as a foundation to provide non-repudiation for electronic commerce applications.

Currently the PEM standard defines the use of the RSA public key algorithm to be used for key management and digital signature operations, and the DES algorithm is included for message confidentiality encryption.

The PEM protocols rely on the trusted distribution of the public keys. PEM public keys are distributed within an X.509 certificate. These certificates are digitally signed by a certification authority. The PEM user trusts a certification authority to provide public key certificates. The certification authorities can also cross certify public key certificates from another certification authority. The certification authorities are distributed in a hierarchical structure with the Internet Policy Registration Authority (IPRA) at the top. The IPRA will certify the certification authorities. The IPRA is a non-government, private agency and may or may not be trusted by an organization.

1.2.6 ENTRUST AND ENTRUST-LITE

Entrust is an cryptographic module that is being developed by Bell Northern Research (BNR). This module will be available for multiple computer platforms and operating systems. The module provides an Application Interface for user applications to utilize the cryptographic functions. This module will provide the cryptographic functionality required for both message and document integrity (Digital Signatures) as well as message/document confidentiality.

This cryptographic module is being validated by the Communications Security Establishment against the FIPS 140-1 standards.

1.2.7 PGP — PRETTY GOOD PRIVACY

PGP is a public key encryption package to protect e-mail and data files. It lets you communicate securely with people you've never met, with no secure channels needed for prior exchange of keys. It's well featured and fast, with sophisticated key management, digital signatures, data compression, and good ergonomic design.

This program provides the RSA algorithm for key management and digital signatures, and uses the IDEA algorithm to provide confidentiality. The program is available for non-commercial use to Canadian citizens from the site ftp://ftp.wimsey.bc.ca. There is commercial version of this program for sale from ViaCrypt, and an international version available as well. The international version has the message encryption (IDEA algorithm) functionality removed.

1.2.8 RIPEM — RIORDAN'S INTERNET PRIVACY-ENHANCED MAIL

RIPEM (pronounced RYE-pehm) is a public key encryption program oriented toward use with electronic mail. It allows you to generate your own public keypairs, and to encrypt and decrypt messages based on your key and the keys of your correspondents. RIPEM is free, but each user is required to agree to a license agreement which places some limitations on its use.

RIPEM is available on Internet at ftp://ftp.rsa.com. This program is a public domain implementation of the PEM standard. The RIPEM application is available for a variety of computer platforms and operating systems.

1.2.9 MIME — MULTIPURPOSE INTERNET MAIL EXTENSIONS

MIME is an Internet Engineering Task Force (IETF) solution that allows users to attach non-text objects to Internet messages. A MIME-capable e-mail client can be configured to automatically retrieve and execute data files that are attached to an e-mail message. The MIME standard provides a standard method of providing attachments to e-mail messages. Some of the MIME e-mail programs allow the user to configure what type of attachments are accepted and how they are interpreted, other programs are not configurable. Users are cautioned to disable the automatic execution and interpretation of mail attachments. The attachments can be examined and processed after the user responds to prompt. In this configuration the user is warned that an attachment is going to be processed and the user has the option of cancelling that processing if they are unsure of the consequences.

There is a system in development called *atomicmail*. *Atomicmail* is described as a language for interactive and computational e-mail. This language is being developed to provide portability between computer systems for the advanced e-mail attachments as well as to address security concerns. The *atomicmail* language is being designed with the constraints that processing does no harm and that access to the operating system, CPU, files and other resources is tightly controlled.

1.3 File Systems

1.3.1 AFS — ANDREW FILE SYSTEM

AFS is a networked file system with similar functionality to NFS. This file system is newer in design and can interoperate (to some degree) with NFS file systems. Unlike NFS, the AFS designers placed security in the protocol and incorporated the Kerberos authentication system into the file protocol.

1.3.2 NFS — NETWORK FILE SYSTEM

NFS is a Remote Procedure Call (RPC) based facility which utilizes port 2049. This facility allows NFS-capable clients to mount a file system on a NFS server located on the network. Once the NFS file system has been mounted it is treated like a local file system. If an internal system exports a file system to external systems, then the file system is available to a hacker across the network. Even if the file system is exported to only a select set of clients the possibility of a hacker spoofing one of those clients is possible. As well, it might be possible for a hacker to hijack an existing NFS connection. NFS should never be allowed across a firewall to an external network such as the Internet.

1.3.3 FTP — FILE TRANSFER PROTOCOL

FTP allows a user to transfer text or binary files between two networked computers using ports 20 and 21. The ftp protocol uses a client-server structure with a client program opening a session on a server. There are many "anonymous ftp servers" located across the Internet. An anonymous server allows anyone to log on and retrieve information without any user identification and authentication (the user gives the username "anonymous" or "ftp").

If an anonymous ftp server allows world writable areas then the server could be used to distribute malicious or illegal software. A server could also be the source of computer viruses, trojan horses or other malicious software.

CERT provides a document on setting up an anonymous ftp server which is available via anonymous ftp from:

```
ftp://info.cert.org/pub/tech_tips/anonymous_ftp
```

This document describes the procedures of configuring an anonymous server, with restricted access. The procedures for restricting access to incoming files are also provided. Even though access to incoming files is restricted, a hacker is able to deposit corrupt, malicious, or illegal software on a server; it is unavailable however, until the server administrator reviews the software and moves it to the archive of retrievable software.

1.3.4 GOPHER

Gopher is a client-server system designed to locate and retrieve files or information from servers, "gopher holes", across the Internet. When a user initiates a connection to a Gopher server, the user is presented with a menu of data topics to choose from. When a user selects a topic, Gopher returns access information and a data type description. The access information tells the client program what IP address, port and filename to access. The data type description informs the client program how to interpret the raw information that is being retrieved. The data types include text and graphic files, script programs and binary executable files. If software is retrieved and executed automatically without user intervention then malicious code (e.g. viruses or trojan horses) could be obtained and executed without prior screening. Therefore, software should not be executed until it has been screened by a virus checker.

For those trivia hounds, it was originally developed at a U.S. university whose mascot was a gopher…

1.3.5 ICMP — INTERNET CONTROL MESSAGE PROTOCOL

The ICMP protocol is used to determine routing information and host status. An ICMP redirect packet is used to inform a router or computer about "new and improved" routes to a destination. These packets can be forged providing false routes to a destination to allow an attacker to spoof another system.

Another common ICMP packet is known as the ICMP unreachable message. These packets indicate problems with a route to a destination address. A false ICMP unreachable message could be used to deny access to another network or host. If this type of vulnerability is of concern to your organization then the routing server or firewall can be configured to ignore ICMP unreachable messages. The drawback of this configuration is that if the packet is genuine and a host is actually unreachable, the network routing tables will still not be updated and users will not know that the host is not available. They will simply be denied access.

Ping is a common ICMP based service. Ping sends a packet to a given destination which in effect says "Are you alive?" The destination returns an acknowledgement to the ping or an ICMP unreachable message may be returned by a routing system in the path. PING also has an ugly and sordid history in its use in network attacks and in network infiltrations.

ICMP packets should be filtered and not allowed across network boundaries.

1.3.6 LPD — LINE PRINTER DAEMON

LPD allows networked computers to access printing services on another computer. If lpd packets (destined for port 515) are allowed to be printed on an internal print server from external sources, a hacker could deny printing services to internal users by monopolizing the printer. This can be prevented by applying quotas, such as, limiting amount of time the printer can be used, time of day it can be used, etc. This can also be prevented by denying external network access to the printer.

1.3.7 NNTP — NETWORK NEWS TRANSFER PROTOCOL

NNTP is an application level protocol which is used to distribute news groups. This protocol provides an unauthenticated and unsecured transfer service. The information passed between computers using this protocol is not encrypted and can be read by anyone with a network monitoring device located in the information pathway. Since there is no authentication, neither the integrity nor the source of the information can be guaranteed.

To provide some sort of information integrity or confidentiality, a higher level of security protocol must be applied to the news messages. One example of this type of security service is the PEM protocol.

1.3.8 NEWS READERS

Network news readers are applications which provide the user with access to NNTP. The news readers usually do not require privileges to run and therefore can only get access to the files owned by the user running the news reader. One concern with these applications is that they do not control the flow of information. An organization cannot control the content of the message; the news reader will not screen information.

1.3.9 NIS — NETWORK INFORMATION SERVICES

NIS was originally developed and known as "yp or yellow pages". The NIS protocol acts in a client server type of fashion where the server provides user and host information to a client. The NIS system provides a central password and host file system for networks of computers. It is possible for a hacker to inform an NIS client to use another NIS server to authenticate logins. If this was successful then a hacker could gain unauthorized access to the client computer.

A hacker can use the NIS protocol to gain information about the network configuration including host and usernames. The more information that a hacker has available, the easier it is to break into a system. NIS should never be allowed across a firewall to an external network such as the Internet.

1.3.10 RPC — REMOTE PROCEDURE CALL

A RPC is similar to a procedure call in the C programming language. The difference is that the procedure call includes a remote IP address and port. The procedure is called from one computer and is executed on another computer across the network. The network file system (NFS) works in this manner. These procedure calls and ports can be used by a hacker to obtain unauthorized access to resources and information on a system. RPC calls should be filtered and not allowed across network boundaries.

The unfortunate thing about RPC's is that programs, such as certain Windows 32 bit applications, require RPCs to operate. Because so many ports must be opened to support the RPC functionality, the additional application flexibility also causes major and serious security problems.

1.3.11 R-UTILS (RLOGIN, RCP, RSH)

These utilities came with the original Berkly version of UNIX. These utilities allow a "trusted" user from a known host to login or execute commands on another network computer. No user identification and authentication is required, since these systems assume a trusted user and host. If a hacker was to spoof one of the trusted hosts, then unauthorized access could be possible. These utilities should never be allowed across a firewall to the Internet.

1.3.12 SNMP — SIMPLE NETWORK MANAGEMENT PROTOCOL

The SNMP protocol allows a network administrator to manage network resources from a remote node. This protocol should never be allowed through a firewall connected to the Internet. A hacker would have the ability to remotely manage and change the configuration of network systems. It would also allow a hacker to rewrite the security policy of the internal network.

1.3.13 TELNET

Telnet is an application which allows a user to log in to a remote computer. Telnet transmits all data between computers in an unencrypted fashion (including the username and password pair). A hacker located on the routing path could monitor all information transferred and pick up sensitive data or the username-password that was used. As well, an ambitious hacker could possibly hijack an existing telnet session. If a hacker gained access to a telnet session then all system resources available to the authorized user would be compromised. A possible solution for this is to use an encryption scheme with telnet.

Telnet is also used as the connection method for most network infrastructure devices such as routers, bridges and lower-level hardware such as CSU/DSU facilities on leased lines and frame relay connections. It has great potential to allow a hacker access to a great deal of very sensitive hardware that can cripple a network if compromised.

1.3.14 TFTP ? TRIVIAL FILE TRANSFER PROTOCOL

TFTP is mainly used for remotely booting another networked computer and operates on port 69. A computer can initiate a tftp session to a boot server and transfer the system boot information it requires to start up. This protocol should be disabled if not required and should never be allowed across a firewall to the Internet. TFTP can also be used to transfer and deposit information to a networked

computer. An attacker could use this protocol to grab sensitive data, password files or to deposit compromised system files. TFTP should not be allowed.

TFTP is also the most common protocol used to download bootstrap kernel software for diskless systems such as routers. Compromise of TFTP host systems on a network can cause a great deal of security problems for a customer network.

1.3.15 MOTIF

Motif is a graphical environment developed by the Open Software Foundation (OSF) as a front end for the X11 X-windows interface. The vulnerabilities of the X-Windows system are described below.

1.3.16 OPENWINDOWS

Openwindows is a graphical environment developed by Sun for its SunOS and Solaris operating systems. This system is now publicly available within other versions of the UNIX operating system. This graphical environment is similar to the Xwindows system, however, it connects to port number 2000.

1.3.17 WINSOCK

Winsock is a Microsoft Windows dynamic link library providing TCP/IP port services to windows applications. These services allow users to run many Internet tools, such as Archie, Cello, ftp, Gopher, Mosaic and telnet on an MS-DOS/MS-Windows computer.

1.3.18 WINDOWS — X11

X windows is a graphical environment for user application software. This environment supports distributed services using TCP ports numbered 6000+. This system is designed to remotely control and display processes across the network. It is possible for a malicious process to monitor or take control of the screen, mouse and keyboard devices. The opening of so many ports also allows the intruder an opportunity to use an open port to compromise a trusted network from an untrusted connection.

1.3.19 WAIS — WIDE AREA INFORMATION SERVERS

This is another of the WWW family of applications and protocols. (see http for vulnerability information)

1.3.20 WWW — WORLD WIDE WEB

WWW is a new family of applications and protocols developed to provide users with a convenient method of accessing information across the Internet. (see http for vulnerability information)

1.3.21 HTTP — HYPERTEXT TRANSFER PROTOCOL

HTTP is the application level protocol used to access world wide web (WWW) servers and information. Http is similar to the Gopher protocol; it transfers an information block and a data type description to the client. The client program (Internet Explorer, Mosaic, Lynx, and Netscape Navigator are common client applications) is responsible for interpreting the information and presenting it to the user in the correct form. As with the Gopher protocol, executable code is a valid data type to be retrieved. Some client programs can be configured to automatically

interpret and process the information that is retrieved. If this protocol is supported care should be taken to configure client programs to prompt prior to executing any script or executable programs. Any executable code retrieved should be scanned for viruses, trojan horses or other malicious activities before being executed.

A potential solution is s-http, which is intended to be a secure version of the http protocol. The s-http protocol is still in development and further information will be sent automatically if an e-mail message is sent to: *info@commerce.net*. This protocol uses the PEM standard for mail and data exchange and provides the PEM capabilities above the http protocol. In this manner all data exchanged between an http server and client can be both authenticated and/or encrypted as required.

Another standard in progress is the SSL or Secure Sockets Layer activity. This standard provides a security layer between the TCP and application protocol layers. SSL can be used to provide integrity (proof of sender) and confidentiality for any TCP data stream. This security protocol can be used with all applications level protocols not just http.

Section References

1. 0 INFOSEC Services, Communications Security Establishment, *An Introduction to the Internet and Internet Security*. Ottawa, Canada, September 1995.

2.1 Security Policy

2.1.0 What is a Security Policy and Why Have One?

The security-related decisions you make, or fail to make, as administrator largely determines how secure or insecure your network is, how much functionality your network offers, and how easy your network is to use. However, you cannot make good decisions about security without first determining what your security goals are. Until you determine what your security goals are, you cannot make effective use of any collection of security tools because you simply will not know what to check for and what restrictions to impose. For example, your goals will probably be very different from the goals of a product vendor. Vendors are trying to make configuration and operation of their products as simple as possible, which implies that the default configurations will often be as open (i.e., insecure) as possible. While this does make it easier to install new products, it also leaves access to those systems, and other systems through them, open to any user who wanders by.

Your goals will be largely determined by the following key tradeoffs:

1. **services offered versus security provided -**
   Each service offered to users carries its own security risks. For some services the risk outweighs the benefit of the service and the administrator may choose to eliminate the service rather than try to secure it.
2. **ease of use versus security -**
   The easiest system to use would allow access to any user and require no passwords; that is, there would be no security. Requiring passwords makes the system a little less convenient, but more secure. Requiring device-generated one-time passwords makes the system even more difficult to use, but much more secure.
3. **cost of security versus risk of loss -**
   There are many different costs to security: monetary (i.e., the cost of purchasing security hardware and software like firewalls and one-time password generators), performance (i.e., encryption and decryption take time), and ease of use (as mentioned above). There are also many levels of risk: loss of privacy (i.e., the reading of information by unauthorized individuals), loss of data (i.e., the corruption or erasure of information), and the loss of service (e.g., the filling of data storage space, usage of computational resources, and denial of network access). Each type of cost must be weighed against each type of loss.

Your goals should be communicated to all users, operations staff, and managers through a set of security rules, called a "security policy." We are using this term, rather than the narrower "computer security policy" since the scope includes all types of information technology and the information stored and manipulated by the technology.

2.1.1 Definition of a Security Policy

A security policy is a formal statement of the rules by which people who are given access to an organization's technology and information assets must abide.

2.1.2 Purposes of a Security Policy

The main purpose of a security policy is to inform users, staff and managers of their obligatory requirements for protecting technology and information assets. The policy should specify the mechanisms through which these requirements can be met. Another purpose is to provide a baseline from which to acquire, configure and audit computer systems and networks for compliance with the policy. Therefore, an attempt to use a set of security tools in the absence of at least an implied security policy is meaningless.

Another major use of an AUP is to spell out, exactly, the corporate position on privacy issues and intellectual property issues. In some countries, if the company does not explicitly state that e-mail is not secure, it is considered to be so and any breach could cause privacy and confidentiality liabilities. It is very important to spell out what is and is not acceptable in intellectual transfers and storage and what the corporate privacy policies are to prevent litigation about same.

An Appropriate Use Policy (AUP) may also be part of a security policy. It should spell out what users shall and shall not do on the various components of the system, including the type of traffic allowed on the networks. The AUP should be as explicit as possible to avoid ambiguity or misunderstanding. For example, an AUP might list any prohibited USENET newsgroups. (Note: Appropriate Use Policy is referred to as Acceptable Use Policy by some sites.)

2.1.3 Who Should be Involved When Forming Policy?

In order for a security policy to be appropriate and effective, it needs to have the acceptance and support of all levels of employees within the organization. It is especially important that corporate management fully support the security policy process otherwise there is little chance that they will have the intended impact. The following is a list of individuals who should be involved in the creation and review of security policy documents:

- site security administrator
- information technology technical staff (e.g., staff from computing center)
- administrators of large user groups within the organization (e.g., business divisions, computer science department within a university, etc.)
- security incident response team
- representatives of the user groups affected by the security policy
- responsible management
- legal counsel (if appropriate)

The list above is representative of many organizations, but is not necessarily comprehensive. The idea is to bring in representation from key stakeholders, management who have budget and policy authority, technical staff who know what can and cannot be supported, and legal counsel who know the legal ramifications of various policy choices. In some organizations, it may be appropriate to include EDP

audit personnel. Involving this group is important if resulting policy statements are to reach the broadest possible acceptance. It
is also relevant to mention that the role of legal counsel will also vary from country to country.

2.1.4 What Makes a Good Security Policy?

The characteristics of a good security policy are:

1. It must be implementable through system administration procedures, publishing of acceptable use guidelines, or other appropriate methods.

2. It must be enforceable with security tools, where appropriate, and with sanctions, where actual prevention is not technically feasible.

3. It must clearly define the areas of responsibility for the users, administrators, and management.

The components of a good security policy include:

1. Computer Technology Purchasing Guidelines which specify required, or preferred, security features. These should supplement existing purchasing policies and guidelines.

2. A Privacy Policy which defines reasonable expectations of privacy regarding such issues as monitoring of electronic mail, logging of keystrokes, and access to users' files.

3. An Access Policy which defines access rights and privileges to protect assets from loss or disclosure by specifying acceptable use guidelines for users, operations staff, and management. It should provide guidelines for external connections, data communications, connecting devices to a network, and adding new software to systems. It should also specify any required notification messages (e.g., connect messages should provide warnings about authorized usage and line monitoring, and not simply say "Welcome").

4. An Accountability Policy which defines the responsibilities of users, operations staff, and management. It should specify an audit capability, and provide incident handling guidelines (i.e., what to do and who to contact if a possible intrusion is detected).

5. An Authentication Policy which establishes trust through an effective password policy, and by setting guidelines for remote location authentication and the use of authentication devices (e.g., one-time passwords and the devices that generate them).

6. An Availability statement which sets users' expectations for the availability of resources. It should address redundancy and recovery issues, as well as specify operating hours and maintenance downtime periods. It should also include contact information for reporting system and network failures.

7. An Information Technology System & Network Maintenance Policy which describes how both internal and external maintenance people are allowed to handle and access technology. One important topic to be addressed here is whether remote maintenance is allowed and how such access is controlled. Another area for consideration here is outsourcing and how it is managed.

8. A Violations Reporting Policy that indicates which types of violations (e.g., privacy and security, internal and external) must be reported and to whom the reports are made. A non-threatening atmosphere and the possibility of anonymous reporting will result in a greater probability that a violation will be reported if it is detected.

9. Supporting Information which provides users, staff, and management with contact information for each type of policy violation; guidelines on how to handle outside queries about a security incident, or information which may be considered confidential or proprietary; and cross-references to security procedures and related information, such as company policies and governmental laws and regulations.

There may be regulatory requirements that affect some aspects of your security policy (e.g., line monitoring). The creators of the security policy should consider seeking legal assistance in the creation of the policy. At a minimum, the policy should be reviewed by legal counsel.

Once your security policy has been established it should be clearly communicated to users, staff, and management. Having all personnel sign a statement indicating that they have read, understood, and agreed to abide by the policy is an important part of the process. Finally, your policy should be reviewed on a regular basis to see if it is successfully supporting your security needs.

2.1.5 Keeping the Policy Flexible

In order for a security policy to be viable for the long term, it requires a lot of flexibility based upon an architectural security concept. A security policy should be (largely) independent from specific hardware and software situations (as specific systems tend to be replaced or moved overnight). The mechanisms for updating the policy should be clearly spelled out. This includes the process, the people involved, and the people who must sign-off on the changes. It is also important to recognize that there are exceptions to every rule. Whenever possible, the policy should spell out what exceptions to the general policy exist. For example, under what conditions is a system administrator allowed to go through a user's files. Also, there may be some cases when multiple users will have access to the same userid. For example, on systems with a "root" user, multiple system administrators may know the password and use the root account.

2.2 Threats

A threat can be any person, object, or event that, if realized, could potentially cause damage to the LAN. Threats can be malicious, such as the intentional modification of sensitive information, or can be accidental, such as an error in a calculation, or the accidental deletion of a file. Threats can also be acts of nature,

i.e. flooding, wind, lightning, etc. ***The immediate damage caused by a threat is referred to as an impact.***

Vulnerabilities are weaknesses in a LAN that can be exploited by a threat. For example, unauthorized access (the threat) to the LAN could occur by an outsider guessing an obvious password. The vulnerability exploited is the poor password choice made by a user. Reducing or eliminating the vulnerabilities of the LAN can reduce or eliminate the risk of threats to the LAN. For example, a tool that can help users choose robust passwords may reduce the chance that users will utilize poor passwords, and thus reduce the threat of unauthorized LAN access.

A security service is the collection of security mechanisms, supporting data files, and procedures that help protect the LAN from specific threats. For example, the identification and authentication service helps protect the LAN from unauthorized LAN access by requiring that a user identify himself, as well as verifying that identity. The security service is only as robust as the mechanisms, procedures, etc. that make up the service.

Security mechanisms are the controls implemented to provide the security services needed to protect the LAN. For example, a token based authentication system (which requires that the user be in possession of a required token) may be the mechanism implemented to provide the identification and authentication service. Other mechanisms that help maintain the confidentiality of the authentication information can also be considered as part of the identification and authentication service.

Threats and Vulnerabilities

Identifying threats requires one to look at the impact and consequence of the threat if it is realized. The impact of the threat, which usually points to the immediate near-term problems, results in disclosure, modification, destruction, or denial of service. The more significant long-term consequences of the threat being realized are the result of lost business, violation of privacy, civil law suits, fines, loss of human life or other long term effects. The approach taken here is to categorize the types of impacts that can occur on a LAN so that specific technical threats can be grouped by the impacts and examined in a meaningful manner. For example, the technical threats that can lead to the impact 'LAN traffic compromise' in general can be distinguished from those threats that can lead to the impact 'disruption of LAN functionalities'. It should be recognized that many threats may result in more than one impact; however, for this discussion a particular threat will be discussed only in conjunction with one impact. The impacts that will be used to categorize and discuss the threats to a LAN environment are:

- **Unauthorized LAN access** - results from an unauthorized individual gaining access to the LAN.
- **Inappropriate access to LAN resources** - results from an individual, authorized or unauthorized, gaining access to LAN resources in an unauthorized manner.
- **Disclosure of data** - results from an individual accessing or reading information and possibly revealing the information in an accidental or unauthorized intentional manner.
- **Unauthorized Modification to data and software** - results from an individual modifying, deleting or destroying LAN data and software in an unauthorized or accidental manner.

- **Disclosure of LAN traffic** - results from an individual accessing or reading information and possibly revealing the information in an accidental or unauthorized intentional manner as it moves through the LAN.
- **Spoofing of LAN traffic** - results when a message appears to have been sent from a legitimate, named sender, when actually the message had not been.
- **Disruption of LAN functions** - results from threats that block LAN resources from being available in a timely manner.

2.2.0 Unauthorized LAN Access

LANs provide file sharing, printer sharing, file storage sharing, etc. Because resources are shared and not used solely by one individual there is need for control of the resources and accountability for use of the resources. *Unauthorized LAN access occurs when someone, who is not authorized to use the LAN, gains access to the LAN (usually by acting as a legitimate user of LAN).* Three common methods used to gain unauthorized access are password sharing, general password guessing and password capturing. Password sharing allows an unauthorized user to have the LAN access and privileges of a legitimate user; with the legitimate user's knowledge and acceptance. General password guessing is not a new means of unauthorized access. Password capturing is a process in which a legitimate user unknowingly reveals the user's login ID and password. This may be done through the use of a trojan horse program that appears to the user as a legitimate login program; however, the trojan horse program is designed to capture passwords. Capturing a login ID and password as it is transmitted across the LAN unencrypted is another method used to ultimately gain access. The methods to capture cleartext LAN traffic, including passwords, is readily available today. Unauthorized LAN access can occur by exploiting the following types of vulnerabilities:

- lack of, or insufficient, identification and authentication scheme,
- password sharing,
- poor password management or easy to guess passwords,
- using known system holes and vulnerabilities that have not been patched,
- single-user PCs that are not password protected at boot time,
- underutilized use of PC locking mechanisms,
- LAN access passwords that are stored in batch files on PCs,
- poor physical control of network devices,
- unprotected modems,
- lack of a time-out for login time period and log of attempts,
- lack of disconnect for multiple login failures and log of attempts,
- lack of 'last successful login date/time' and 'unsuccessful login attempt' notification and log,
- lack of real-time user verification (to detect masquerading).

2.2.1 Inappropriate Access to LAN Resources

One of the benefits of using a LAN is that many resources are readily available to many users, rather than each user having limited dedicated resources. These resources may include file stores, applications, printers, data, etc. However, not all resources need to be made available to each user. To prevent compromising the security of the resource (i.e. corrupting the resource, or lessening the availability of the resource), only those who require the use of the resource should be permitted to utilize that resource. *Unauthorized access occurs when a user, legitimate or unauthorized, accesses a resource that the user is not permitted to use.* Unauthorized access may occur simply because the access rights assigned to the resource are not assigned properly. However, unauthorized access may also occur

because the access control mechanism or the privilege mechanism is not granular enough. In these cases, the only way to grant the user the needed access rights or privileges to perform a specific function is to grant the user more access than is needed, or more privileges than are needed. Unauthorized access to LAN resources can occur by exploiting the following types of vulnerabilities:

- use of system default permission settings that are too permissive to users,
- improper use of administrator or LAN manager privileges,
- data that is stored with an inadequate level or no protection assigned,
- lack of or the improper use of the privilege mechanism for users,
- PCs that utilize no access control on a file level basis.

Disclosure of Data

As LANs are utilized throughout an agency or department, some of the data stored or processed on a LAN may require some level of confidentiality. *The disclosure of LAN data or software occurs when the data or software is accessed, read and possibly released to an individual who is not authorized for the data.* This can occur by someone gaining access to information that is not encrypted, or by viewing monitors or printouts of the information. The compromise of LAN data can occur by exploiting the following types of vulnerabilities:

- improper access control settings,
- data, that has been deemed sensitive enough to warrant encryption, stored in unencrypted form,
- application source code stored in unencrypted form,
- monitors viewable in high traffic areas,
- printer stations placed in high traffic areas,
- data and software backup copies stored in open areas.

Unauthorized Modification of Data and Software

Because LAN users share data and applications, changes to those resources must be controlled. *Unauthorized modification of data or software occurs when unauthorized changes (additions, deletions or modifications) are made to a file or program.*

When undetected modifications to data are present for long periods of time, the modified data may be spread through the LAN, possibly corrupting databases, spreadsheet calculations, and other various application data. This can damage the integrity of most application information.

When undetected software changes are made, all system software can become suspect, warranting a thorough review (and perhaps reinstallation) of all related software and applications. These unauthorized changes can be made in simple command programs (for example in PC batch files), in utility programs used on multi-user systems, in major application programs, or any other type of software. They can be made by unauthorized outsiders, as well as those who are authorized to make software changes (although the changes they make are not authorized). These changes can divert information (or copies of the information) to other destinations, corrupt the data as it is processed, or harm the availability of system or LAN services.

PC viruses can be a nuisance to any organization that does not choose to provide LAN users the tools to effectively detect and prevent virus introduction to the LAN. Currently viruses have been limited to corrupting PCs, and generally do not corrupt

LAN servers (although viruses can use the LAN to infect PCs). [WACK89] provides guidance on detecting and preventing viruses.

The unauthorized modification of data and software can occur by exploiting the following types of vulnerabilities:

- write permission granted to users who only require read permission to access,
- undetected changes made to software, including the addition of code to create a trojan horse program,
- lack of a cryptographic checksum on sensitive data,
- privilege mechanism that allow unnecessary write permission,
- lack of virus protection and detection tools.

Disclosure of LAN Traffic

The disclosure of LAN traffic occurs when someone who is unauthorized reads, or otherwise obtains, information as it is moved through the LAN. LAN traffic can be compromised by listening and capturing traffic transmitted over the LAN transport media (tapping into a network cable, listening to traffic transmitted over the air, misusing a provided network connection by attaching an analysis device, etc.). Many users realize the importance of confidential information when it is stored on their workstations or servers; however, it is also important to maintain that confidentiality as the information travels through the LAN. Information that can be compromised in this way includes system and user names, passwords, electronic mail messages, application data, etc. For example, even though passwords may be in an encrypted form when stored on a system, they can be captured in plaintext as they are sent from a workstation or PC to a file server. Electronic mail message files, which usually have very strict access rights when stored on a system, are often sent in plaintext across a wire, making them an easy target for capturing. The compromise of LAN traffic can occur by exploiting the following types of vulnerabilities:

- inadequate physical protection of LAN devices and medium,
- transmitting plaintext data using broadcast protocols,
- transmitting plaintext data (unencrypted) over the LAN medium,

2.2.2 Spoofing of LAN Traffic

Data that is transmitted over a LAN should not be altered in an unauthorized manner as a result of that transmission, either by the LAN itself, or by an intruder. LAN users should be able to have a reasonable expectation that the message sent, is received unmodified. *A modification occurs when an intentional or unintentional change is made to any part of the message including the contents and addressing information.*

Messages transmitted over the LAN need to contain some sort of addressing information that reports the sending address of the message and the receiving address of the message (along with other pieces of information). *Spoofing of LAN traffic involves (1) the ability to receive a message by masquerading as the legitimate receiving destination, or (2) masquerading as the sending machine and sending a message to a destination.* To masquerade as a receiving machine, the LAN must be persuaded into believing that the destination address is the legitimate address of the machine. (Receiving LAN traffic can also be done by listening to messages as they are broadcast to all nodes.) Masquerading as the sending machine to deceive a receiver into believing the message was legitimately sent can

be done by masquerading the address, or by means of a playback. A playback involves capturing a session between a sender and receiver, and then retransmitting that message (either with the header only, and new message contents, or the whole message). The spoofing of LAN traffic or the modification of LAN traffic can occur by exploiting the following types of vulnerabilities:

- transmitting LAN traffic in plaintext,
- lack of a date/time stamp (showing sending time and receiving time),
- lack of message authentication code mechanism or digital signature,
- lack of real-time verification mechanism (to use against playback).

2.2.3 Disruption of LAN Functions

A LAN is a tool, used by an organization, to share information and transmit it from one location to another. *A disruption of functionality occurs when the LAN cannot provide the needed functionality in an acceptable, timely manner.* A disruption can interrupt one type of functionality or many. A disruption of LAN functionalities can occur by exploiting the following types of vulnerabilities:

- inability to detect unusual traffic patterns (i.e. intentional flooding),
- inability to reroute traffic, handle hardware failures, etc,
- configuration of LAN that allows for a single point of failure,
- unauthorized changes made to hardware components (reconfiguring addresses on workstations, modifying router or hub configurations, etc.),
- improper maintenance of LAN hardware,
- improper physical security of LAN hardware.

2.2.4 Common Threats

A variety of threats face today's computer systems and the information they process. In order to control the risks of operating an information system, managers and users must know the vulnerabilities of the system and the threats, which may exploit them. Knowledge of the threat environment allows the system manager to implement the most cost-effective security measures. In some cases, managers may find it most cost-effective to simply tolerate the expected losses.

The following threats and associated losses are based on their prevalence and significance in the current computing environment and their expected growth. The list is not exhaustive; some threats may combine elements from more than one area.

2.2.4.0 ERRORS AND OMISSIONS

Users, data entry clerks, system operators, and programmers frequently make unintentional errors, which contribute to security problems, directly and indirectly. Sometimes the error is the threat, such as a data entry error or a programming error that crashes a system. In other cases, errors create vulnerabilities. Errors can occur in all phases of the system life cycle. Programming and development errors, often called bugs, range in severity from benign to catastrophic. In the past decade, software quality has improved measurably to reduce this threat, yet software "horror stories" still abound. Installation and maintenance errors also cause security problems. Errors and omissions are important threats to data integrity. Errors are caused not only by data entry clerks processing hundreds of transactions per day, but also by all users who create and edit data. Many programs, especially those designed by users for personal computers, lack quality control measures. However,

even the most sophisticated programs cannot detect all types of input errors or omissions.

The computer age saying "garbage in, gospel out" contains a large measure of truth. People often assume that the information they receive from a computer system is more accurate than it really is. Many organizations address errors and omissions in their computer security, software quality, and data quality programs.

2.2.4.1 FRAUD AND THEFT

Information technology is increasingly used to commit fraud and theft. Computer systems are exploited in numerous ways, both by automating traditional methods of fraud and by using new methods. For example, individuals may use a computer to skim small amounts of money from a large number of financial accounts, thus generating a significant sum for their own use. In addition, deposits may be intentionally misdirected. Financial systems are not the only ones subject to fraud. Systems, which control access to any resource, are targets, such as time and attendance systems, inventory systems, school grading systems, or long-distance telephone systems.

Fraud can be committed by insiders or outsiders. The majority of fraud uncovered on computer systems is perpetrated by insiders who are authorized users of a system. Since insiders have both access to and familiarity with the victim computer system, including what resources it controls and where the flaws are, authorized system users are in a better position to commit crimes. An organization's former employees may also pose threats, particularly if their access is not terminated promptly.

2.2.4.2 DISGRUNTLED EMPLOYEES

Disgruntled employees can create both mischief and sabotage on a computer system. Employees are the group most familiar with their employer's computers and applications, including knowing what actions might cause the most damage. Organizational downsizing in both public and private sectors has created a group of individuals with organizational knowledge who may retain potential system access. System managers can limit this threat by invalidating passwords and deleting system accounts in a timely manner. However, disgruntled current employees actually cause more damage than former employees do.

Common examples of computer-related employee sabotage include:

- Entering data incorrectly
- Changing data
- Deleting data
- Destroying data or programs with logic bombs
- "Crashing" systems
- Holding data hostage
- Destroying hardware or facilities

2.2.4.3 PHYSICAL AND INFRASTRUCTURE

The loss of supporting infrastructure includes power failures (including outages, spikes and brownouts), loss of communications, water outages and leaks, sewer problems, lack of transportation services, fire, flood, civil unrest, strikes, and so forth. These losses include dramatic events such as the explosion at the World Trade Center and the Chicago tunnel flood as well as more common events such as

a broken water pipe. System owners must realize that more loss is associated with fires and floods than with viruses and other more widely publicized threats. A loss of infrastructure often results in system downtime, sometimes in unexpected ways. For example, employees may not be able to get to work during a winter storm, although the computer system may be functional.

2.2.4.4 MALICIOUS HACKERS

Hackers, sometimes called crackers, are a real and present danger to most organizational computer systems linked by networks. From outside the organization, sometimes from another continent, hackers break into computer systems and compromise the privacy and integrity of data before the unauthorized access is even detected. Although insiders cause more damage than hackers do, the hacker problem remains serious and widespread.

The effect of hacker activity on the public switched telephone network has been studied in depth. Studies by the National Research Council and the National Security Telecommunications Advisory Committee show that hacker activity is not limited to toll fraud. It also includes the ability to break into telecommunications systems (such as switches) resulting in the degradation or disruption of system availability. While unable to reach a conclusion about the degree of threat or risk, these studies underscore the ability of hackers to cause serious damage.

The hacker threat often receives more attention than more common and dangerous threats. The U.S. Department of Justice's Computer Crime Unit suggests three reasons. First, the hacker threat is a more recently encountered threat. Organizations have always had to worry about the actions of their own employees and could use disciplinary measures to reduce that threat. However, these controls are ineffective against outsiders who are not subject to the rules and regulations of the employer.

Secondly, organizations do not know the purposes of a hacker; some hackers only browse, some steal, some damage. This inability to identify purposes can suggest that hacker attacks have no limitations. Finally, hacker attacks make people feel vulnerable because the perpetrators are unknown.

2.2.4.5 INDUSTRIAL ESPIONAGE

Industrial espionage involves collecting proprietary data from private corporations or government agencies for the benefit of another company or organization. Industrial espionage can be perpetrated either by companies seeking to improve their competitive advantage or by governments seeking to aid their domestic industries. Foreign industrial espionage carried out by a government is known as economic espionage.

Industrial espionage is on the rise. The most damaging types of stolen information include manufacturing and product development information. Other types of information stolen include sales and cost data, client lists, and research and planning information.

Within the area of economic espionage, the Central Intelligence Agency states that the main objective is obtaining information related to technology, but that information on U.S. government policy deliberations concerning foreign affairs and information on commodities, interest rates, and other economic factors is also a target. The Federal Bureau of Investigation concurs that technology-related information is the

main target, but also cites corporate proprietary information such as negotiating positions and other contracting data as a target.

2.2.4.6 MALICIOUS CODE

Malicious code refers to viruses, worms, Trojan horses, logic bombs, and other "uninvited" software. Malicious code is sometimes mistakenly associated only with personal computers, but can also attack systems that are more sophisticated. However, actual costs attributed to the presence of malicious code have resulted primarily from system outages and staff time involved in repairing the systems. Nonetheless, these costs can be significant.

2.2.4.7 MALICIOUS SOFTWARE: TERMS

Virus: A code segment, which replicates by attaching copies of itself to existing executables. The new copy of the virus is executed when a user executes the new host program. The virus may include an additional "payload" that triggers when specific conditions are met. For example, some viruses display a text string on a particular date. There are many types of viruses including variants, overwriting, resident, stealth, and polymorphic.

Trojan Horse: A program that performs a desired task, but also includes unexpected (and undesirable) functions. Consider as an example an editing program for a multi-user system. This program could be modified to randomly delete one of the users' files each time they perform a useful function (editing) but the deletions are unexpected and definitely undesired!

Worm: A self-replicating program, which is self-contained and does not require a host program. The program creates a copy of itself and causes it to execute; no user intervention is required. Worms commonly utilize network services to propagate to other host systems.

The number of known viruses is increasing, and the rate of virus incidents is growing moderately. Most organizations use anti-virus software and other protective measures to limit the risk of virus infection.

2.2.4.8 FOREIGN GOVERNMENT ESPIONAGE

In some instances, threats posed by foreign government intelligence services may be present. In addition to possible economic espionage, foreign intelligence services may target unclassified systems to further their intelligence missions.

2.3 Security Services and Mechanisms Introduction

A security service is the collection of mechanisms, procedures and other controls that are implemented to help reduce the risk associated with threat. For example, the identification and authentication service helps reduce the risk of the unauthorized user threat. Some services provide protection from threats, while other services provide for detection of the threat occurrence. An example of this would be a logging or monitoring service. The following services will be discussed in this section:

Identification and authentication - is the security service that helps ensure that the LAN is accessed by only authorized individuals.

Access control - is the security service that helps ensure that LAN resources are being utilized in an authorized manner.
Data and message confidentiality - is the security service that helps ensure that LAN data, software and messages are not disclosed to unauthorized parties.
Data and message integrity - is the security service that helps ensure that LAN data, software and messages are not modified by unauthorized parties.
Non-repudiation - is the security service by which the entities involved in a communication cannot deny having participated. Specifically the sending entity cannot deny having sent a message (non-repudiation with proof of origin) and the receiving entity cannot deny having received a message (non-repudiation with proof of delivery).
Logging and Monitoring - is the security service by which uses of LAN resources can be traced throughout the LAN.

Determining the appropriate controls and procedures to use in any LAN environment is the responsibility of those in each organization charged with providing adequate LAN protection.

2.3.0 Identification and Authentication

The first step toward securing the resources of a LAN is the ability to verify the identities of users [BNOV91]. The process of verifying a user's identity is referred to as authentication. Authentication provides the basis for the effectiveness of other controls used on the LAN. For example the logging mechanism provides usage information based on the userid. The access control mechanism permits access to LAN resources based on the userid. Both these controls are only effective under the assumption that the requestor of a LAN service is the valid user assigned to that specific userid.

Identification requires the user to be known by the LAN in some manner. This is usually based on an assigned userid. However the LAN cannot trust the validity that the user is in fact, who the user claims to be, without being authenticated. The authentication is done by having the user supply something that only the user has, such as a token, something that only the user knows, such as a password, or something that makes the user unique, such as a fingerprint. The more of these that the user has to supply, the less risk in someone masquerading as the legitimate user.

A requirement specifying the need for authentication should exist in most LAN policies. The requirement may be directed implicitly in a program level policy stressing the need to effectively control access to information and LAN resources, or may be explicitly stated in a LAN specific policy that states that all users must be uniquely identified and authenticated.

On most LANs, the identification and authentication mechanism is a userid/password scheme. [BNOV91] states that "password systems can be effective if managed properly [FIPS112], but seldom are. Authentication which relies solely on passwords has often failed to provide adequate protection for systems for a number of reasons. Users tend to create passwords that are easy to remember and hence easy to guess. On the other hand users that must use passwords generated from random characters, while difficult to guess, are also difficult to be remembered by users. This forces the user to write the password down, most likely in an area easy accessible in the work area". Research work such as [KLEIN] detail the ease at which passwords can be guessed. Proper password selection (striking a balance between being easy-to-remember for the user but difficult-to-guess for everyone else) has always been an issue. Password generators that produce passwords

consisting of pronounceable syllables have more potential of being remembered than generators that produce purely random characters. [FIPS180] specifies an algorithm that can be used to produce random pronounceable passwords. Password checkers are programs that enable a user to determine whether a new passwords is considered easy-to-guess, and thus unacceptable.

Password-only mechanisms, especially those that transmit the password in the clear (in an unencrypted form) are susceptible to being monitored and captured. This can become a serious problem if the LAN has any uncontrolled connections to outside networks. Agencies that are considering connecting their LANs to outside networks, particularly the Internet, should examine [BJUL93] before doing so. If, after considering all authentication options, LAN policy determines that password-only systems are acceptable, the proper management of password creation, storage, expiration and destruction become all the more important. [FIPS 112] provides guidance on password management. [NCSC85] provides additional guidance that may be considered appropriate.

Because of the vulnerabilities that still exist with the use of password-only mechanisms, more robust mechanisms can be used. [BNOV91] discusses advances that have been made in the areas of token-based authentication and the use of biometrics. A smartcard based or token based mechanism requires that a user be in possession of the token and additionally may require the user to know a PIN or password. These devices then perform a challenge/response authentication scheme using realtime parameters. Using realtime parameters helps prevent an intruder from gaining unauthorized access through a login session playback. These devices may also encrypt the authentication session, preventing the compromise of the authentication information through monitoring and capturing.

Locking mechanisms for LAN devices, workstations, or PCs that require user authentication to unlock can be useful to users who must leave their work areas frequently. These locks allow users to remain logged into the LAN and leave their work areas (for an acceptable short period of time) without exposing an entry point into the LAN.

Modems that provide users with LAN access may require additional protection. An intruder that can access the modem may gain access by successfully guessing a user password. The availability of modem use to legitimate users may also become an issue if an intruder is allowed continual access to the modem.

Mechanisms that provide a user with his or her account usage information may alert the user that the account was used in an abnormal manner (e.g. multiple login failures). These mechanisms include notifications such as date, time, and location of last successful login, and number of previous login failures. The type of security mechanisms that could be implemented to provide the identification and authentication service are listed below.

- password based mechanism,
- smartcards/smart tokens based mechanism,
- biometrics based mechanism,
- password generator,
- password locking,
- keyboard locking,
- PC or workstation locking,
- termination of connection after multiple failed logins
- user notification of ‘last successful login’ and ‘number of login failures’,

- real-time user verification mechanism,
- cryptography with unique user keys.

2.3.1 Access Control

This service protects against the unauthorized use of LAN resources, and can be provided by the use of access control mechanisms and privilege mechanisms. Most file servers and multi-user workstations provide this service to some extent. However, PCs which mount drives from the file servers usually do not. Users must recognize that files used locally from a mounted drive are under the access control of the PC. For this reason it may be important to incorporate access control, confidentiality and integrity services on PCs to whatever extent possible.

According to [NCSC87], access control can be achieved by using discretionary access control or mandatory access control. Discretionary access control is the most common type of access control used by LANs. The basis of this kind of security is that an individual user, or program operating on the user's behalf is allowed to specify explicitly the types of access other users (or programs executing on their behalf) may have to information under the user's control.

Discretionary security differs from mandatory security in that it implements the access control decisions of the user. Mandatory controls are driven by the results of a comparison between the user's trust level or clearance and the sensitivity designation of the information.

Access control mechanisms exist that support access granularity for acknowledging an owner, a specified group of users, and the world (all other authorized users). This allows the owner of the file (or directory) to have different access rights than all other users, and allows the owner to specify different access rights for a specified group of people, and also for the world. Generally access rights allow read access, write access, and execute access. Some LAN operating systems provide additional access rights that allow updates, append only, etc.

A LAN operating system may implement user profiles, capability lists or access control lists to specify access rights for many individual users and many different groups. Using these mechanisms allows more flexibility in granting different access rights to different users, which may provide more stringent access control for the file (or directory). (These more flexible mechanisms prevent having to give a user more access than necessary, a common problem with the three level approach.) Access control lists assign the access rights of named users and named groups to a file or directory. Capability lists and user profiles assign the files and directories that can be accessed by a named user.

User access may exist at the directory level, or the file level. Access control at the directory level places the same access rights on all the files in the directory. For example, a user that has read access to the directory can read (and perhaps copy) any file in that directory. Directory access rights may also provide an explicit negative access that prevents the user from any access to the files in the directory.

Some LAN implementations control how a file can be accessed. (This is in addition to controlling who can access the file.) Implementations may provide a parameter that allows an owner to mark a file sharable, or locked. Sharable files accept multiple accesses to the file at the same time. A locked file will permit only one user to access it. If a file is a read only file, making it sharable allows many users to read it at the same time.

These access controls can also be used to restrict usage between servers on the LAN. Many LAN operating systems can restrict the type of traffic sent between servers. There may be no restrictions, which implies that all users may be able to access resources on all servers (depending on the users access rights on a particular server). Some restrictions may be in place that allow only certain types of traffic, for example only electronic mail messages, and further restrictions may allow no exchange of traffic from server to server. The LAN policy should determine what types of information need to be exchanged between servers. Information that is not necessary to be shared between servers should then be restricted.

Privilege mechanisms enable authorized users to override the access permissions, or in some manner legally bypass controls to perform a function, access a file, etc. A privilege mechanism should incorporate the concept of least privilege. [ROBA91] defines least privilege as "a principle where each subject in a system be granted the most restrictive set or privileges needed for the performance of an authorized task." For example, the principle of least privilege should be implemented to perform the backup function. A user who is authorized to perform the backup function needs to have read access to all files in order to copy them to the backup media. (However the user should not be given read access to all files through the access control mechanism.) The user is granted a 'privilege' to override the read restrictions (enforced by the access control mechanism) on all files in order to perform the backup function. The more granular the privileges that can be granted, the more control there is not having to grant excessive privilege to perform an authorized function. For example, the user who has to perform the backup function does not need to have a write override privilege, but for privilege mechanisms that are less granular, this may occur. The types of security mechanisms that could be implemented to provide the access control service are listed below.

- access control mechanism using access rights (defining owner, group, world permissions),
- access control mechanism using access control lists, user profiles, capability lists,
- access control using mandatory access control mechanisms (labels),
- granular privilege mechanism,

2.3.2 Data and Message Confidentiality

The data and message confidentiality service can be used when the secrecy of information is necessary. As a front line protection, this service may incorporate mechanisms associated with the access control service, but can also rely on encryption to provide further secrecy protection. Encrypting information converts it to an unintelligible form called ciphertext, decrypting converts the information back to its original form. Sensitive information can be stored in the encrypted, ciphertext, form. In this way if the access control service is circumvented, the file may be accessed but the information is still protected by being in encrypted form. (The use of encryption may be critical on PCs that do not provide an access control service as a front line protection.)

It is very difficult to control unauthorized access to LAN traffic as it is moved through the LAN. For most LAN users, this is a realized and accepted problem. The use of encryption reduces the risk of someone capturing and reading LAN messages in transit by making the message unreadable to those who may capture it. Only the authorized user who has the correct key can decrypt the message once it is received.

A strong policy statement should dictate to users the types of information that are deemed sensitive enough to warrant encryption. A program level policy may dictate the broad categories of information that need to be stringently protected, while a system level policy may detail the specific types of information and the specific environments that warrant encryption protection. At whatever level the policy is dictated, the decision to use encryption should be made by the authority within the organization charged with ensuring protection of sensitive information. If a strong policy does not exist that defines what information to encrypt, then the data owner should ultimately make this decision.

Cryptography can be categorized as either secret key or public key. Secret key cryptography is based on the use of a single cryptographic key shared between two parties . The same key is used to encrypt and decrypt data. This key is kept secret by the two parties. If encryption of sensitive but unclassified information (except Warner Amendment information) is needed, the use of the *Data Encryption Standard* (DES), FIPS 46-2, is required unless a waiver is granted by the head of the federal agency. The DES is a secret key algorithm used in a cryptographic system that can provide confidentiality. FIPS 46-2 provides for the implementation of the DES algorithm in hardware, software, firmware or some combination. This is a change from 46-1 which only provided for the use of hardware implementations. For an overview of DES, information addressing the applicability of DES, and waiver procedures see [NCSL90].

Public key cryptography is a form of cryptography which make use of two keys: a public key and a private key. The two keys are related but have the property that, given the public key, it is computationally infeasible to derive the private key [FIPS 140-1]. In a public key cryptosystem, each party has its own public/private key pair. The public key can be known by anyone; the private key is kept secret. An example for providing confidentiality is as follows: two users, Scott and Jeff, wish to exchange sensitive information, and maintain the confidentiality of that information. Scott can encrypt the information with Jeff's public key. The confidentiality of the information is maintained since only Jeff can decrypt the information using his private key. There is currently no FIPS approved public-key encryption algorithm for confidentiality. Agencies must waive FIPS 46-2 to use a public-key encryption algorithm for confidentiality. Public key technology, in the form of digital signatures, can also provide integrity and non-repudiation.

FIPS 140-1, *Security Requirements for Cryptographic Modules*, should be used by agencies to specify the security requirements needed to protect the equipment that is used encryption. This standard specifies requirements such as authentication, physical controls and proper key management for all equipment that is used for encryption. Systems that implement encryption in software have additional requirements placed on them by FIPS 140-1. LAN servers, PCs, encryption boards, encryption modems, and all other LAN and data communication equipment that has an encryption capability should conform to the requirements of FIPS 140-1. The types of security mechanisms that could be implemented to provide the message and data confidentiality service are listed below.

- file and message encryption technology,
- protection for backup copies on tapes, diskettes, etc,
- physical protection of physical LAN medium and devices,
- use of routers that provide filtering to limit broadcasting (either by blocking or by masking message contents).

2.3.3 Data and Message Integrity

The data and message integrity service helps to protect data and software on workstations, file servers, and other LAN components from unauthorized modification. The unauthorized modification can be intentional or accidental. This service can be provided by the use of cryptographic checksums, and very granular access control and privilege mechanisms. The more granular the access control or privilege mechanism, the less likely an unauthorized or accidental modification can occur.

The data and message integrity service also helps to ensure that a message is not altered, deleted or added to in any manner during transmission. (The inadvertent modification of a message packet is handled through the media access control implemented within the LAN protocol.) Most of the security techniques available today cannot prevent the modification of a message, but they can detect the modification of a message (unless the message is deleted altogether).

The use of checksums provide a modification detection capability. A Message Authentication Code (MAC), a type of cryptographic checksum, can protect against both accidental and intentional, but unauthorized, data modification. A MAC is initially calculated by applying a cryptographic algorithm and a secret value, called the key, to the data. The initial MAC is retained. The data is later verified by applying the cryptographic algorithm and the same secret key to the data to produce another MAC; this MAC is then compared to the initial MAC. If the two MACs are equal, then the data is considered authentic. Otherwise, an unauthorized modification is assumed. Any party trying to modify the data without knowing the key would not know how to calculate the appropriate MAC corresponding to the altered data. FIPS 113, *Computer Data Authentication*, defines the Data Authentication Algorithm, based on the DES, which is used to calculate the MAC. See [SMID88] for more information regarding the use of MACs.

The use of electronic signatures can also be used to detect the modification of data or messages. An electronic signature can be generated using public key or private key cryptography. Using a public key system, documents in a computer system are electronically signed by applying the originator's private key to the document. The resulting digital signature and document can then be stored or transmitted. The signature can be verified using the public key of the originator.
If the signature verifies properly, the receiver has confidence that the document was signed using the private key of the originator and that the message had not been altered after it was signed. Because private keys are known only to their owner, it may also possible to verify the originator of the information to a third party. A digital signature, therefore, provides two distinct services: nonrepudiation and message integrity. FIPS PUB 186, Digital Signature Standard, specifies a digital signature algorithm that should be used when message and data integrity are required.

The message authentication code (MAC) described above can also be used to provide an electronic signature capability. The MAC is calculated based on the contents of the message. After transmission another MAC is calculated on the contents of the received message. If the MAC associated with the message that was sent is not the same as the MAC associated with the message that was received, then there is proof that the message received does not exactly match the message sent. A MAC can be used to identify the signer of the information to the receiver. However, the implementations of this technology do not inherently provide nonrepudiation because both the sender of the information and the receiver of the information share the same key. The types of security mechanisms that could be implemented to provide the data and message integrity service are listed below.

- message authentication codes used for software or files,
- use of secret key based electronic signature,
- use of public key digital signature,
- granular privilege mechanism,
- appropriate access control settings (i.e. no unnecessary write permissions),
- virus detection software,
- workstations with no local storage (to prevent local storage of software and files),
- workstations with no diskette drive/tape drive to prevent introduction of uspect software.
- use of public key digital signatures.

2.3.4 Non-repudiation

Non-repudiation helps ensure that the entities in a communication cannot deny having participated in all or part of the communication. When a major function of the LAN is electronic mail, this service becomes very important. Non-repudiation with proof of origin gives the receiver some confidence that the message indeed came from the named originator. The nonrepudiation service can be provided through the use of public key cryptographic techniques
using digital signatures. The security mechanism that could be implemented to provide the non-repudiation service is listed below.

- use of public key digital signatures.

2.3.5 Logging and Monitoring

This service performs two functions. The first is the detection of the occurrence of a threat. (However, the detection does not occur in real time unless some type of real-time monitoring capability is utilized.) Depending on the extensiveness of the logging, the detected event should be traceable throughout the system. For example, when an intruder breaks into the system, the log should indicate who was logged on to the system at the time, all sensitive files that had failed accesses, all programs that had attempted executions, etc. It should also indicate sensitive files and programs that were successfully accessed in this time period. It may be appropriate that some areas of the LAN (workstations, fileservers, etc.) have some type of logging service.

The second function of this service is to provide system and network managers with statistics that indicate that systems and the network as a whole are functioning properly. This can be done by an audit mechanism that uses the log file as input and processes the file into meaningful information regarding system usage and security. A monitoring capability can also be used to detect LAN availability problems as they develop. The types of security mechanisms that could be used to provide the logging and monitoring service are listed below.

- logging of I&A information (including source machine, modem, etc.),
- logging of changes to access control information,
- logging of use of sensitive files,
- logging of modifications made to critical software,
- utilizing LAN traffic management tools,
- use of auditing tools.

2.4 Architecture Objectives

2.4.0 Separation of Services

There are many services which a site may wish to provide for its users, some of which may be external. There are a variety of security reasons to attempt to isolate services onto dedicated host computers. There are also performance reasons in most cases, but a detailed discussion is beyond to scope of this document.

The services which a site may provide will, in most cases, have different levels of access needs and models of trust. Services which are essential to the security or smooth operation of a site would be better off being placed on a dedicated machine with very limited access (see "deny all" model), rather than on a machine that provides a service (or services) which has traditionally been less secure, or requires greater accessibility by users who may accidentally suborn security.

It is also important to distinguish between hosts which operate within different models of trust (e.g., all the hosts inside of a firewall and any host on an exposed network).

Some of the services which should be examined for potential separation are outlined in the section on service protection. It is important to remember that security is only as strong as the weakest link in the chain. Several of the most publicized penetrations in recent years have been through the exploitation of vulnerabilities in electronic mail systems. The intruders were not trying to steal electronic mail, but they used the vulnerability in that service to gain access to other systems.

If possible, each service should be running on a different machine whose only duty is to provide a specific service. This helps to isolate intruders and limit potential harm.

2.4.0.1 DENY ALL/ ALLOW ALL

There are two diametrically opposed underlying philosophies which can be adopted when defining a security plan. Both alternatives are legitimate models to adopt, and the choice between them will depend on the site and its needs for security.

The first option is to turn off all services and then selectively enable services on a case by case basis as they are needed. This can be done at the host or network level as appropriate. This model, which will here after be referred to as the "deny all" model, is generally more secure than the other model described in the next paragraph. More work is required to successfully implement a "deny all" configuration as well as a better understanding of services. Allowing only known services provides for a better analysis of a particular service/protocol and the design of a security mechanism suited to the security level of the site.

The other model, which will here after be referred to as the "allow all" model, is much easier to implement, but is generally less secure than the "deny all" model. Simply turn on all services, usually the default at the host level, and allow all protocols to travel across network boundaries, usually the default at the router level. As security holes become apparent, they are restricted or patched at either the host or network level.

Each of these models can be applied to different portions of the site, depending on functionality requirements, administrative control, site policy, etc. For example, the

policy may be to use the "allow all" model when setting up workstations for general use, but adopt a "deny all" model when setting up information servers, like an email hub. Likewise, an "allow all" policy may be adopted for traffic between LAN's internal to the site, but a "deny all" policy can be adopted between the site and the Internet.

Be careful when mixing philosophies as in the examples above. Many sites adopt the theory of a hard "crunchy" shell and a soft "squishy" middle. They are willing to pay the cost of security for their external traffic and require strong security measures, but are unwilling or unable to provide similar protections internally. This works fine as long as the outer defenses are never breached and the internal users can be trusted. Once the outer shell (firewall) is breached, subverting the internal network is trivial.

2.4.1 Protecting Services

2.4.1.0 NAME SERVERS (DNS AND NIS(+))

The Internet uses the Domain Name System (DNS) to perform address resolution for host and network names. The Network Information Service (NIS) and NIS+ are not used on the global Internet, but are subject to the same risks as a DNS server. Name-to-address resolution is critical to the secure operation of any network. An attacker who can successfully control or impersonate a DNS server can re-route traffic to subvert security protections. For example, routine traffic can be diverted to a compromised system to be monitored; or, users can be tricked into providing authentication secrets. An organization should create well known, protected sites to act as secondary name servers and protect their DNS masters from denial of service attacks using filtering routers.

Traditionally, DNS has had no security capabilities. In particular, the information returned from a query could not be checked for modification or verified that it had come from the name server in question. Work has been done to incorporate digital signatures into the protocol which, when deployed, will allow the integrity of the information to be cryptographically verified.

2.4.1.1 PASSWORD/KEY SERVERS (NIS(+) AND KDC)

Password and key servers generally protect their vital information (i.e., the passwords and keys) with encryption algorithms. However, even a one-way encrypted password can be determined by a dictionary attack (wherein common words are encrypted to see if they match the stored encryption). It is therefore necessary to ensure that these servers are not accessible by hosts which do not plan to use them for the service, and even those hosts should only be able to access the service (i.e., general services, such as Telnet and FTP, should not be allowed by anyone other than administrators).

2.4.1.2 AUTHENTICATION/PROXY SERVERS (SOCKS, FWTK)

A proxy server provides a number of security enhancements. It allows sites to concentrate services through a specific host to allow monitoring, hiding of internal structure, etc. This funneling of services creates an attractive target for a potential intruder. The type of protection required for a proxy server depends greatly on the proxy protocol in use and the services being proxied. The general rule of limiting access only to those hosts which need the services, and limiting access by those hosts to only those services, is a good starting point.

2.4.1.3 ELECTRONIC MAIL

Electronic mail (email) systems have long been a source for intruder break-ins because email protocols are among the oldest and most widely deployed services. Also, by it's very nature, an email server requires access to the outside world; most email servers accept input from any source. An email server generally consists of two parts: a receiving/sending agent and a processing agent. Since email is delivered to all users, and is usually private, the processing agent typically requires system (root) privileges to deliver the mail. Most email implementations perform both portions of the service, which means the receiving agent also has system privileges. This opens several security holes which this document will not describe. There are some implementations available which allow a separation of the two agents. Such implementations are generally considered more secure, but still require careful installation to avoid creating a security problem.

2.4.1.4 WORLD WIDE WEB (WWW)

The Web is growing in popularity exponentially because of its ease of use and the powerful ability to concentrate information services. Most WWW servers accept some type of direction and action from the persons accessing their services. The most common example is taking a request from a remote user and passing the provided information to a program running on the server to process the request. Some of these programs are not written with security in mind and can create security holes. If a Web server is available to the Internet community, it is especially important that confidential information not be co-located on the same host as that server. In fact, it is recommended that the server have a dedicated host which is not "trusted" by other internal hosts.

Many sites may want to co-locate FTP service with their WWW service. But this should only occur for anon-ftp servers that only provide information (ftp-get). Anon-ftp puts, in combination with WWW, might be dangerous (e.g., they could result in modifications to the information your site is publishing to the web) and in themselves make the security considerations for each service different.

2.4.1.5 FILE TRANSFER (FTP, TFTP)

FTP and TFTP both allow users to receive and send electronic files in a point-to-point manner. However, FTP requires authentication while TFTP requires none. For this reason, TFTP should be avoided as much as possible.

Improperly configured FTP servers can allow intruders to copy, replace and delete files at will, anywhere on a host, so it is very important to configure this service correctly. Access to encrypted passwords and proprietary data, and the introduction of Trojan horses are just a few of the potential security holes that can occur when the service is configured incorrectly. FTP servers should reside on their own host. Some sites choose to co-locate FTP with a Web server, since the two protocols share common security considerations However, the practice isn't recommended, especially when the FTP service allows the deposit of files (see section on WWW above). Services offered internally to your site should not be co-located with services offered externally. Each should have its own host.

TFTP does not support the same range of functions as FTP, and has no security whatsoever. This service should only be considered for internal use, and then it should be configured in a restricted way so that the server only has access to a set of predetermined files (instead of every world-readable file on the system). Probably the most common usage of TFTP is for downloading router configuration

files to a router. TFTP should reside on its own host, and should not be installed on hosts supporting external FTP or Web access.

2.4.1.6 NFS

The Network File Service allows hosts to share common disks. NFS is frequently used by diskless hosts who depend on a disk server for all of their storage needs. Unfortunately, NFS has no built-in security. It is therefore necessary that the NFS server be accessible only by those hosts which are using it for service. This is achieved by specifying which hosts the file system is being exported to and in what manner (e.g., read-only, read-write, etc.). Filesystems should not be exported to any hosts outside the local network since this will require that the NFS service be accessible externally. Ideally, external access to NFS service should be stopped by a firewall.

2.4.2 Protecting the Protection

It is amazing how often a site will overlook the most obvious weakness in its security by leaving the security server itself open to attack. Based on considerations previously discussed, it should be clear that: the security server should not be accessible from off-site; should offer minimum access, except for the authentication function, to users on-site; and should not be co-located with any other servers. Further, all access to the node, including access to the service itself, should be logged to provide a "paper trail" in the event of a security breach.

2.5 Auditing

This section covers the procedures for collecting data generated by network activity, which may be useful in analyzing the security of a network and responding to security incidents.

2.5.1 What to Collect

Audit data should include any attempt to achieve a different security level by any person, process, or other entity in the network. This includes login and logout, super user access (or the non-UNIX equivalent), ticket generation (for Kerberos, for example), and any
other change of access or status. It is especially important to note "anonymous" or "guest" access to public servers.

The actual data to collect will differ for different sites and for different types of access changes within a site. In general, the information you want to collect includes: username and hostname, for login and logout; previous and new access rights, for a change of access rights; and a timestamp. Of course, there is much more information which might be gathered, depending on what the system makes available and how much space is available to store that information.

One very important note: do not gather passwords. This creates an enormous potential security breach if the audit records should be improperly accessed. Do not gather incorrect passwords either, as they often differ from valid passwords by only a single character or transposition.

2.5.2 Collection Process

The collection process should be enacted by the host or resource being accessed. Depending on the importance of the data and the need to have it local in instances

in which services are being denied, data could be kept local to the resource until needed or be transmitted to storage after each event.

There are basically three ways to store audit records: in a read/write file on a host, on a write-once/read-many device (e.g., a CD-ROM or a specially configured tape drive), or on a write-only device (e.g., a line printer). Each method has advantages and disadvantages.

File system logging is the least resource intensive of the three methods and the easiest to configure. It allows instant access to the records for analysis, which may be important if an attack is in progress. File system logging is also the least reliable method. If the logging host has been compromised, the file system is usually the first thing to go; an intruder could easily cover up traces of the intrusion.

Collecting audit data on a write-once device is slightly more effort to configure than a simple file, but it has the significant advantage of greatly increased security because an intruder could not alter the data showing that an intrusion has occurred. The disadvantage of this method is the need to maintain a supply of storage media and the
cost of that media. Also, the data may not be instantly available.

Line printer logging is useful in system where permanent and immediate logs are required. A real time system is an example of this, where the exact point of a failure or attack must be recorded. A laser printer, or other device which buffers data (e.g., a print server), may suffer from lost data if buffers contain the needed data at a critical instant. The disadvantage of, literally, "paper trails" is the need to keep the printer fed and the need to scan records by hand. There is also the issue of where to store the, potentially, enormous volume of paper which may be generated.

2.5.3 Collection Load

Collecting audit data may result in a rapid accumulation of bytes so storage availability for this information must be considered in advance. There are a few ways to reduce the required storage space. First, data can be compressed, using one of many methods. Or, the required space can be minimized by keeping data for a shorter period of time with only summaries of that data kept in long-term archives. One major drawback to the latter method involves incident response. Often, an incident has been ongoing for some period of time when a site notices it and begins to investigate. At that point in time, it's very helpful to have detailed audit logs available. If these are just summaries, there may not be sufficient detail to fully handle the incident.

2.5.4 Handling and Preserving Audit Data

Audit data should be some of the most carefully secured data at the site and in the backups. If an intruder were to gain access to audit logs, the systems themselves, in addition to the data, would be at risk.

Audit data may also become key to the investigation, apprehension, and prosecution of the perpetrator of an incident. For this reason, it is advisable to seek the advice of legal council when deciding how audit data should be treated. This should happen before an incident occurs.

If a data handling plan is not adequately defined prior to an incident, it may mean that there is no recourse in the aftermath of an event, and it may create liability resulting from improper treatment of the data.

2.5.5 Legal Considerations

One area concerns the privacy of individuals. In certain instances, audit data may contain personal information. Searching through the data, even for a routine check of the system's security, could represent an invasion of privacy.

A second area of concern involves knowledge of intrusive behavior originating from your site. If an organization keeps audit data, is it responsible for examining it to search for incidents? If a host in one organization is used as a launching point for an attack against another organization, can the second organization use the audit data of the first organization to prove negligence on the part of that organization?

The above examples are meant to be comprehensive, but should motivate your organization to consider the legal issues involved with audit data.

2.5.6 Securing Backups

The procedure of creating backups is a classic part of operating a computer system. Within the context of this document, backups are addressed as part of the overall security plan of a site. There are several aspects to backups that are important within this context:

1. Make sure your site is creating backups
2. Make sure your site is using offsite storage for backups. The storage site should be carefully selected for both its security and its availability.
3. Consider encrypting your backups to provide additional protection of the information once it is off-site. However, be aware that you will need a good key management scheme so that you'll be able to recover data at any point in the future. Also, make sure you will have access to the necessary decryption programs at such time in the future as you need to perform the decryption.
4. Don't always assume that your backups are good. There have been many instances of computer security incidents that have gone on for long periods of time before a site has noticed the incident. In such cases, backups of the affected systems are also tainted.
5. Periodically verify the correctness and completeness of your backups.

2.6 Incidents

2.6.0 Preparing and Planning for Incident Handling

Part of handling an incident is being prepared to respond to an incident before the incident occurs in the first place. This includes establishing a suitable level of protections as explained in the preceding chapters. Doing this should help your site prevent incidents as well as limit potential damage resulting from them when they do occur. Protection also includes preparing incident handling guidelines as part of a contingency plan for your organization or site. Having written plans eliminates much of the ambiguity which occurs during an incident, and will lead to a more appropriate and thorough set of responses. It is vitally important to test the proposed plan before an incident occurs through "dry runs". A team might even consider hiring a tiger team to act in parallel with the dry run. (Note: a tiger team is a team of specialists that try to penetrate the security of a system.)

Learning to respond efficiently to an incident is important for a number of reasons:

1. Protecting the assets which could be compromised
2. Protecting resources which could be utilized more profitably if an incident did not require their services
3. Complying with (government or other) regulations
4. Preventing the use of your systems in attacks against other systems (which could cause you to incur legal liability)
1. Minimizing the potential for negative exposure

As in any set of pre-planned procedures, attention must be paid to a set of goals for handling an incident. These goals will be prioritized differently depending on the site. A specific set of objectives can be identified for dealing with incidents:

1. Figure out how it happened.
2. Find out how to avoid further exploitation of the same vulnerability.
3. Avoid escalation and further incidents.
4. Assess the impact and damage of the incident.
5. Recover from the incident.
6. Update policies and procedures as needed.
7. Find out who did it (if appropriate and possible).

Due to the nature of the incident, there might be a conflict between analyzing the original source of a problem and restoring systems and services. Overall goals (like assuring the integrity of critical systems) might be the reason for not analyzing an incident. Of course, this is an important management decision; but all involved parties must be aware that without analysis the same incident may happen again.

It is also important to prioritize the actions to be taken during an incident well in advance of the time an incident occurs. Sometimes an incident may be so complex that it is impossible to do everything at once to respond to it; priorities are essential. Although priorities will vary from institution to institution, the following suggested priorities may serve as a starting point for defining your organization's response:

1. **Priority one** -- protect human life and people's safety; human life always has precedence over all other considerations.

2. **Priority two** -- protect classified and/or sensitive data. Prevent exploitation of classified and/or sensitive systems, networks or sites. Inform affected classified and/or sensitive systems, networks or sites about already occurred penetrations. (Be aware of regulations by your site or by government)

3. **Priority three** -- protect other data, including proprietary, scientific, managerial and other data, because loss of data is costly in terms of resources. Prevent exploitations of other systems, networks or sites and inform already affected systems, networks or sites about successful penetrations.

4. **Priority four** -- prevent damage to systems (e.g., loss or alteration of system files, damage to disk drives, etc.). Damage to systems can result in costly down

time and recovery.

5. **Priority five** -- minimize disruption of computing resources (including processes). It is better in many cases to shut a system down or disconnect from a network than to risk damage to data or systems. Sites will have to evaluate the trade-offs between shutting down and disconnecting, and staying up. There may be service agreements in place that may require keeping systems up even in light of further damage occurring. However, the damage and scope of an incident may be so extensive that service agreements may have to be over-ridden.

2.6.1 Notification and Points of Contact

It is important to establish contacts with various personnel before a real incident occurs. Many times, incidents are not real emergencies. Indeed, often you will be able to handle the activities internally. However, there will also be many times when others outside your immediate department will need to be included in the incident handling. These additional contacts include local managers and system administrators, administrative contacts for other sites on the Internet, and various investigative organizations. Getting to know these contacts before incidents occurs will help to make your incident handling process more efficient.

For each type of communication contact, specific "Points of Contact" (POC) should be defined. These may be technical or administrative in nature and may include legal or investigative agencies as well as service providers and vendors. When establishing these contact, it is important to decide how much information will be shared with each class of contact. It is especially important to define, ahead of time, what information will be shared with the users at a site, with the public (including the press), and with other sites.

2.6.2 Law Enforcement and Investigative Agencies

In the event of an incident that has legal consequences, it is important to establish contact with investigative agencies (e.g, the FBI and Secret Service in the U.S. and the RCMP in Canada) as soon as possible. Local law enforcement, local security offices, and campus police departments should also be informed as appropriate. This section describes many
of the issues that will be confronted, but it is acknowledged that each organization will have its own local and governmental laws and regulations that will impact how they interact with law enforcement and investigative agencies. The most important point to make is that each site needs to work through these issues.

A primary reason for determining these point of contact well in advance of an incident is that once a major attack is in progress, there is little time to call these agencies to determine exactly who the correct point of contact is. Another reason is that it is important to cooperate with these agencies in a manner that will foster a good working relationship, and that will be in accordance with the working procedures of these agencies. Knowing the working procedures in advance, and the expectations of your point of contact is a big step in this direction. For example, it is important to gather evidence that will be admissible in any subsequent legal proceedings, and this will require prior knowledge of how to gather such evidence. A final reason for establishing contacts as soon as possible is that it is impossible to know the particular agency that will assume jurisdiction in any given incident.

Making contacts and finding the proper channels early on will make responding to an incident go considerably more smoothly.

If your organization or site has a legal counsel, you need to notify this office soon after you learn that an incident is in progress. At a minimum, your legal counsel needs to be involved to protect the legal and financial interests of your site or organization. There are many legal and practical issues, a few of which are:

1. Whether your site or organization is willing to risk negative publicity or exposure to cooperate with legal prosecution efforts.

2. Downstream liability--if you leave a compromised system as is so it can be monitored and another computer is damaged because the attack originated from your system, your site or organization may be liable for damages incurred.

3. Distribution of information--if your site or organization distributes information about an attack in which another site or organization may be involved or the vulnerability in a product that may affect ability to market that product, your site or organization may again be liable for any damages (including damage of reputation).

4. Liabilities due to monitoring--your site or organization may be sued if users at your site or elsewhere discover that your site is monitoring account activity without informing users.

Unfortunately, there are no clear precedents yet on the liabilities or responsibilities of organizations involved in a security incident or who might be involved in supporting an investigative effort. Investigators will often encourage organizations to help trace and monitor intruders. Indeed, most investigators cannot pursue computer intrusions without extensive support from the organizations involved. However, investigators cannot provide protection from liability claims, and these kinds of efforts may drag out for months and may take a lot of effort.

On the other hand, an organization's legal council may advise extreme caution and suggest that tracing activities be halted and an intruder shut out of the system. This, in itself, may not provide protection from liability, and may prevent investigators from identifying the perpetrator.

The balance between supporting investigative activity and limiting liability is tricky. You'll need to consider the advice of your legal counsel and the damage the intruder is causing (if any) when making your decision about what to do during any particular incident.

Your legal counsel should also be involved in any decision to contact investigative agencies when an incident occurs at your site. The decision to coordinate efforts with investigative agencies is most properly that of your site or organization. Involving your legal counsel will also foster the multi-level coordination between your site and the particular investigative agency involved, which in turn results in an efficient division of labor. Another result is that you are likely to obtain guidance that will help you avoid future legal mistakes.

Finally, your legal counsel should evaluate your site's written procedures for responding to incidents. It is essential to obtain a "clean bill of health" from a legal perspective before you actually carry out these procedures.

It is vital, when dealing with investigative agencies, to verify that the person who calls asking for information is a legitimate representative from the agency in question. Unfortunately, many well intentioned people have unknowingly leaked sensitive details about incidents, allowed unauthorized people into their systems, etc., because a caller has masqueraded as a representative of a government agency. (Note: this word of caution actually applies to all external contacts.)

A similar consideration is using a secure means of communication. Because many network attackers can easily re-route electronic mail, avoid using electronic mail to communicate with other agencies (as well as others dealing with the incident at hand). Non-secured phone lines (the phones normally used in the business world) are also
frequent targets for tapping by network intruders, so be careful!

There is no one established set of rules for responding to an incident when the local government becomes involved. Normally (in the U.S.), except by legal order, no agency can force you to monitor, to disconnect from the network, to avoid telephone contact with the suspected attackers, etc. Each organization will have a set of local and national laws and regulations that must be adhered to when handling incidents. It is recommended that each site be familiar with those laws and regulations, and identify and get know the contacts for agencies with jurisdiction well in advance of handling an incident.

2.6.3 Internal Communications

It is crucial during a major incident to communicate why certain actions are being taken, and how the users (or departments) are expected to behave. In particular, it should be made very clear to users what they are allowed to say (and not say) to the outside world (including other departments). For example, it wouldn't be good for an organization if users replied to customers with something like, "I'm sorry the systems are down, we've had an intruder and we are trying to clean things up." It would be much better if they were instructed to respond with a prepared statement like, "I'm sorry our systems are unavailable, they are being maintained for better service in the future."

Communications with customers and contract partners should be handled in a sensible, but sensitive way. One can prepare for the main issues by preparing a checklist. When an incident occurs, the checklist can be used with the addition of a sentence or two for the specific circumstances of the incident.

Public relations departments can be very helpful during incidents. They should be involved in all planning and can provide well constructed responses for use when contact with outside departments and organizations is necessary.

2.6.4 Public Relations - Press Releases

One of the most important issues to consider is when, who, and how much to release to the general public through the press. There are many issues to consider when deciding this particular issue. First and foremost, if a public relations office exists for the site, it is important to use this office as liaison to the press. The public relations office is trained in the type and wording of information released, and will help to assure that the image of the site is protected during and after the incident (if

possible). A public relations office has the advantage that you can communicate candidly with them, and provide a buffer between the constant press attention and the need of the POC to maintain control over the incident.

If a public relations office is not available, the information released to the press must be carefully considered. If the information is sensitive, it may be advantageous to provide only minimal or overview information to the press. It is quite possible that any information provided to the press will be quickly reviewed by the perpetrator of the incident. Also note that misleading the press can often backfire and cause more damage than releasing sensitive information.

While it is difficult to determine in advance what level of detail to provide to the press, some guidelines to keep in mind are:

1. Keep the technical level of detail low. Detailed information about the incident may provide enough information for others to launch similar attacks on other sites, or even damage the site's ability to prosecute the guilty party once the event is over.

2. Keep the speculation out of press statements. Speculation of who is causing the incident or the motives are very likely to be in error and may cause an inflamed view of the incident.

3. Work with law enforcement professionals to assure that evidence is protected. If prosecution is involved, assure that the evidence collected is not divulged to the press.

4. Try not to be forced into a press interview before you are prepared. The popular press is famous for the "2 am" interview, where the hope is to catch the interviewee off guard and obtain information otherwise not available.

5. Do not allow the press attention to detract from the handling of the event. Always remember that the successful closure of an incident is of primary importance.

2.6.5 Identifying an Incident

2.6.5.1 IS IT REAL?

This stage involves determining if a problem really exists. Of course many if not most signs often associated with virus infection, system intrusions, malicious users, etc., are simply anomalies such as hardware failures or suspicious system/user behavior. To assist in identifying whether there really is an incident, it is usually helpful to obtain and use any detection software which may be available. Audit information is also extremely useful, especially in determining whether there is a network attack. It is extremely important to obtain a system snapshot as soon as one suspects that something is wrong. Many incidents cause a dynamic chain of events to occur, and an initial system snapshot may be the most valuable tool for identifying the problem and any source of attack. Finally, it is important to start a log book. Recording system events, telephone conversations, time stamps, etc., can lead to a more rapid and systematic identification of the problem, and is the basis for subsequent stages of incident handling.

There are certain indications or "symptoms" of an incident that deserve special attention:

1. System crashes.
2. New user accounts (the account RUMPLESTILTSKIN has been unexpectedly created), or high activity on a previously low usage account.
3. New files (usually with novel or strange file names, such as data.xx or k or .xx).
4. Accounting discrepancies (in a UNIX system you might notice the shrinking of an accounting file called /usr/admin/lastlog, something that should make you very suspicious that there may be an intruder).
5. Changes in file lengths or dates (a user should be suspicious if .EXE files in an MS DOS computer have unexplainedly grown by over 1800 bytes).
6. Attempts to write to system (a system manager notices that a privileged user in a VMS system is attempting to alter RIGHTSLIST.DAT).
7. (Data modification or deletion (files start to disappear).
8. Denial of service (a system manager and all other users become locked out of a UNIX system, now in single user mode).
9. Unexplained, poor system performance
10. Anomalies ("GOTCHA" is displayed on the console or there are frequent unexplained "beeps").
11. Suspicious probes (there are numerous unsuccessful login attempts from another node).
12. Suspicious browsing (someone becomes a root user on a UNIX system and accesses file after file on many user accounts.)
13. Inability of a user to log in due to modifications of his/her account.

By no means is this list comprehensive; we have just listed a number of common indicators. It is best to collaborate with other technical and computer security personnel to make a decision as a group about whether an incident is occurring.

2.6.6 Types and Scope of Incidents

Along with the identification of the incident is the evaluation of the scope and impact of the problem. It is important to correctly identify the boundaries of the incident in order to effectively deal with it and prioritize responses.

In order to identify the scope and impact a set of criteria should be defined which is appropriate to the site and to the type of connections available. Some of the issues include:

1. Is this a multi-site incident?
2. Are many computers at your site affected by this incident?
3. Is sensitive information involved?
4. What is the entry point of the incident (network, phone line, local terminal, etc.)?
5. Is the press involved?
6. What is the potential damage of the incident?
7. What is the estimated time to close out the incident?
8. What resources could be required to handle the incident?

9. Is law enforcement involved?

2.6.7 Assessing the Damage and Extent

The analysis of the damage and extent of the incident can be quite time consuming, but should lead to some insight into the nature of the incident, and aid investigation and prosecution. As soon as the breach has occurred, the entire system and all of its components should be considered suspect. System software is the most probable target. Preparation is key to be able to detect all changes for a possibly tainted system. This includes checksumming all media from the vendor using a algorithm which is resistant to tampering.

Assuming original vendor distribution media are available, an analysis of all system files should commence, and any irregularities should be noted and referred to all parties involved in handling the incident. It can be very difficult, in some cases, to decide which backup media are showing a correct system status. Consider, for example, that the incident may have continued for months or years before discovery, and the suspect may be an employee of the site, or otherwise have intimate knowledge or access to the systems. In all cases, the pre-incident preparation will determine what recovery is
possible.

If the system supports centralized logging (most do), go back over the logs and look for abnormalities. If process accounting and connect time accounting is enabled, look for patterns of system usage. To a lesser extent, disk usage may shed light on the incident. Accounting can provide much helpful information in an analysis of an incident and subsequent prosecution. Your ability to address all aspects of a specific incident strongly depends on the success of this analysis.

2.6.8 Handling an Incident

Certain steps are necessary to take during the handling of an incident. In all security related activities, the most important point to be made is that all sites should have policies in place. Without defined policies and goals, activities undertaken will remain without focus. The goals should be defined by management and legal counsel in advance.

One of the most fundamental objectives is to restore control of the affected systems and to limit the impact and damage. In the worst case scenario, shutting down the system, or disconnecting the system from the network, may the only practical solution.

As the activities involved are complex, try to get as much help as necessary. While trying to solve the problem alone, real damage might occur due to delays or missing information. Most administrators take the discovery of an intruder as a personal challenge. By proceeding this way, other objectives as outlined in the local policies may not always be considered. Trying to catch intruders may be a very low priority, compared to system integrity, for example. Monitoring a hacker's activity is useful, but it might not be considered worth the risk to allow the continued access.

2.6.9 Protecting Evidence and Activity Logs

When you respond to an incident, document all details related to the incident. This will provide valuable information to yourself and others as you try to unravel the course of events. Documenting all details will ultimately save you time. If you don't document every relevant phone call, for example, you are likely to forget a

significant portion of information you obtain, requiring you to contact the source of information again. At the same time, recording details will provide evidence for prosecution efforts, providing the case moves in that direction. Documenting an incident will also help you perform a final assessment of damage (something your management, as well as law enforcement officers, will want to know), and will provide the basis for later phases of the handling process: eradication, recovery, and follow-up "lessons learned."

During the initial stages of an incident, it is often infeasible to determine whether prosecution is viable, so you should document as if you are gathering evidence for a court case. At a minimum, you should record:

- all system events (audit records)
- all actions you take (time tagged)
- all external conversations (including the person with whom you talked, the date and time, and the content of the conversation)

The most straightforward way to maintain documentation is keeping a log book. This allows you to go to a centralized, chronological source of information when you need it, instead of requiring you to page through individual sheets of paper. Much of this information is potential evidence in a court of law. Thus, when a legal follow-up is a possibility, one should follow the prepared procedures and avoid jeopardizing the legal follow-up by improper handling of possible evidence. If appropriate, the following steps may be taken.

- Regularly (e.g., every day) turn in photocopied, signed copies of your logbook (as well as media you use to record system events) to a document custodian.
- The custodian should store these copied pages in a secure place (e.g., a safe).
- When you submit information for storage, you should receive a signed, dated receipt from the document custodian.

Failure to observe these procedures can result in invalidation of any evidence you obtain in a court of law.

2.6.10 Containment

The purpose of containment is to limit the extent of an attack. An essential part of containment is decision making (e.g., determining whether to shut a system down, disconnect from a network, monitor system or network activity, set traps, disable functions such as remote file transfer, etc.).

Sometimes this decision is trivial; shut the system down if the information is classified, sensitive, or proprietary. Bear in mind that removing all access while an incident is in progress obviously notifies all users, including the alleged problem users, that the administrators are aware of a problem; this may have a deleterious effect on an investigation. In some cases, it is prudent to remove all access or functionality as soon as possible, then restore normal operation in limited stages. In other cases, it is worthwhile to risk some damage to the system if keeping the system up might enable you to identify an intruder.

This stage should involve carrying out predetermined procedures. Your organization or site should, for example, define acceptable risks in dealing with an incident, and

should prescribe specific actions and strategies accordingly. This is especially important when a quick decision is necessary and it is not possible to first contact all involved parties to discuss the decision. In the absence of predefined procedures, the person in charge of the incident will often not have the power to make difficult management decisions (like to lose the results of a costly experiment by shutting down a system). A final activity that should occur during this stage of incident handling is the notification of appropriate authorities.

2.6.11 Eradication

Once the incident has been contained, it is time to eradicate the cause. But before eradicating the cause, great care should be taken to collect all necessary information about the compromised system(s) and the cause of the incident as they will likely be lost when cleaning up the system.

Software may be available to help you in the eradication process, such as anti-virus software. If any bogus files have been created, archive them before deleting them. In the case of virus infections, it is important to clean and reformat any media containing infected files. Finally, ensure that all backups are clean. Many systems infected with viruses become periodically re-infected simply because people do not systematically eradicate the virus from backups. After eradication, a new backup should be taken.

Removing all vulnerabilities once an incident has occurred is difficult. The key to removing vulnerabilities is knowledge and understanding of the breach.

It may be necessary to go back to the original distribution media and re-customize the system. To facilitate this worst case scenario, a record of the original system setup and each customization change should be maintained. In the case of a network-based attack, it is important to install patches for each operating system vulnerability which was exploited.

If a particular vulnerability is isolated as having been exploited, the next step is to find a mechanism to protect your system. The security mailing lists and bulletins would be a good place to search for this information, and you can get advice from incident response
teams.

2.6.12 Recovery

Once the cause of an incident has been eradicated, the recovery phase defines the next stage of action. The goal of recovery is to return the system to normal. In general, bringing up services in the order of demand to allow a minimum of user inconvenience is the best practice. Understand that the proper recovery procedures for the system are extremely important and should be specific to the site.

2.6.13 Follow-Up

Once you believe that a system has been restored to a "safe" state, it is still possible that holes, and even traps, could be lurking in the system. One of the most important stages of responding to incidents is also the most often omitted, the follow-up stage. In
the follow-up stage, the system should be monitored for items that may have been missed during the cleanup stage. It would be prudent to utilize some of the tools mentioned in chapter 7 as a start. Remember, these tools don't replace continual system monitoring and good systems administration practices.

The most important element of the follow-up stage is performing a postmortem analysis. Exactly what happened, and at what times? How well did the staff involved with the incident perform? What kind of information did the staff need quickly, and how could they have gotten that information as soon as possible? What would the staff do differently next time?

After an incident, it is prudent to write a report describing the exact sequence of events: the method of discovery, correction procedure, monitoring procedure, and a summary of lesson learned. This will aid in the clear understanding of the problem. Creating a formal chronology of events (including time stamps) is also important for legal reasons.

2.6.14 Aftermath of an Incident

In the wake of an incident, several actions should take place. These actions can be summarized as follows:

1. An inventory should be taken of the systems' assets, (i.e., a careful examination should determine how the system was affected by the incident).

2. The lessons learned as a result of the incident should be included in revised security plan to prevent the incident from re-occurring.

3. A new risk analysis should be developed in light of the incident.

4. An investigation and prosecution of the individuals who caused the incident should commence, if it is deemed desirable.

If an incident is based on poor policy, and unless the policy is changed, then one is doomed to repeat the past. Once a site has recovered from and incident, site policy and procedures should be reviewed to encompass changes to prevent similar incidents. Even without an incident, it would be prudent to review policies and procedures on a regular basis. Reviews are imperative due to today's changing computing environments.

The whole purpose of this post mortem process is to improve all security measures to protect the site against future attacks. As a result of an incident, a site or organization should gain practical knowledge from the experience. A concrete goal of the post mortem is to develop new proactive methods. Another important facet of the aftermath may be end user and administrator education to prevent are occurrence of the security problem.

2.7 Intrusion Management Summary

Intrusion management is a four-step process. The steps are avoidance, assurance, detection and investigation. Intrusion management has as its objective:

> *Limiting the possibility of a successful intrusion through effective preventative, quality management and detective processes, and facilitating successful investigation of an intrusion should one occur.*

The primary goal of intrusion management is to prevent intrusions entirely. We can address that goal by implementing a program of effective security controls. Those controls should be present at every interface point within an information management system. Effective controls grow out of effective information security policies, standards and practices. Organizations should impose controls aimed at mitigating threats against functional areas of vulnerability at each interface point. There are six such functional areas of vulnerability:

1. Identification and Authentication: Functions intended to establish and verify the identity of the user or using process.
2. Access Control: Functions intended to control the flow of data between, and the use of resources by, users, processes
3. and objects. This includes administration and verification of access rights.
4. Accountability: Functions intended to record exercising of rights to perform security-relevant actions.
5. Object Reuse: Functions intended to control reuse or scavenging of data objects.
6. Accuracy: Functions intended to insure correctness and consistency of security-relevant information.
7. Reliability of Service: Functions intended to insure security of data over communication links.

2.7.0 Avoidance

The first step in the Intrusion Management process is Avoidance. Avoidance includes all of those underlying processes that seek to create a secure environment. Some examples of Avoidance are:

- Security policy
- Standards and practices
- Security Awareness
- Incident response planning
- Disaster planning
- Training of security and IT Audit personnel
- Evaluating the results of a successful intrusion ("lessons learned" feedback)

2.7.1 Assurance

The second step is Assurance. Assurance includes everything we do to ensure that policies, standards and practices are being followed. These processes include:

- IT audits
- Intrusion testing
- Vulnerability testing
- Security reviews
- Risk assessments on new systems

Using appropriate tools, we can test our systems for these vulnerabilities and through proper configuration or use of third party products we can ensure that appropriate steps are taken to reduce or eliminate them. Tools that we should use are of two types: preventative and detective. Preventative tools include those that we use to perform initial evaluation and configuration. Detective tools are intended to ensure that any change to the configuration is detected.

In broad terms, we may consider that type of monitoring to be an audit function. Thus, we see that auditing is an important part of the intrusion management process. However, many organizations have subdivided the monitoring function between Information Security and IT Auditing. The security personnel monitor on a full time basis, while audits occur periodically to ensure that monitoring is effective. How your organization splits these tasks, or if they split them at all, is probably a function of organization size and resources.

2.7.2 Detection

The third step is Detection. This is somewhat different from the detective controls present during the avoidance and testing steps. In this case we are talking about detecting an intrusion attempt in real time. The real time aspect of detection is important. Knowing that an attack is in progress and being able to take immediate action greatly improves the odds of successfully terminating the intrusion and apprehending the perpetrator.

Real time detection depends upon having a "watch dog" system that sits in the background and watches all activities involving the device under surveillance. The watch dog also must be able to interpret what constitutes an attack.

2.7.3 Investigation

Finally, intrusion management defaults to Investigation when all other measures have failed to prevent an attack. However, investigation, as you may have already gathered, may be futile unless luck and circumstances are with you. By integrating your investigation process into the intrusion management process you improve your odds markedly because you have gathered significant important information and make critical preparations along the way.

Attacks often are not discovered until well after the fact. That problem constitutes strike one in the intrusion management ball game. Strike two comes when the attacker has been clever enough to cover his or her tracks effectively. If the logs are not complete, protected from tampering and retained long enough, it's strike three and your investigation never gets to first base.

Investigations of security incidents, whether they are successful or simply strong attempts, should be undertaken by the organization's Computer Incident Response Team (CIRT). The CIRT should be trained and prepared to initiate a formal investigation, present results to management, support litigation or criminal prosecution if necessary, and ensure that lessons learned are fed back into the Intrusion Management process.

Good intrusion management mitigates against all of the problems surrounding an investigation and ensures you a chance to start around the bases. Whether you get an eventual home run, of course, depends upon many other factors. If you think of the Intrusion Management process as a circle, the results of investigations feed back into the start of the process: Avoidance. By taking advantage of "lessons learned" we help improve the odds against additional successful attacks.

2.8 Modems

2.8.0 Modem Lines Must Be Managed
Although they provide convenient access to a site for its users, they can also provide an effective detour around the site's firewalls. For this reason it is essential to maintain proper control of modems.

Don't allow users to install a modem line without proper authorization. This includes temporary installations (e.g., plugging a modem into a facsimile or telephone line overnight). Maintain a register of all your modem lines and keep your register up to date. Conduct regular (ideally automated) site checks for unauthorized modems.

The reality at most companies is that there are more and more laptop computers being used on desktops. Practically every one of them has a MODEM either built-in or in a PCCARD slot. This means the number of actual MODEMs in corporate networks is growing dramatically without the care and security require of such installations. It is ludicrous to think that they can be eliminated, but it is also important to set policies and technology in place to ensure that the desktop MODEM installed on a laptop does not become the back-door entry point into a corporate network.

2.8.1 Dial-in Users Must Be Authenticated

A username and password check should be completed before a user can access anything on your network. Normal password security considerations are particularly important.

Remember that telephone lines can be tapped, and that it is quite easy to intercept messages to cellular phones. Modern high-speed modems use more sophisticated modulation techniques, which makes them somewhat more difficult to monitor, but it is prudent to assume that hackers know how to eavesdrop on your lines. For this reason, you should use one-time passwords if at all possible.

It is helpful to have a single dial-in point (e.g., a single large modem pool) so that all users are authenticated in the same way.

Users will occasionally mis-type a password. Set a short delay – say two seconds - after the first and second failed logins, and force a disconnect after the third. This will slow down automated password attacks. Don't tell the user whether the username, the password, or both, were incorrect.

Remember that MODEMs on inside computers can be accessed from outside. Often, users will install remote access software on the office system that allows the user to remotely connect to the desktop system from a remote site. Packages like Rapid Remote, PC Anywhere, Carbon Copy and many others allow this capability. These packages, while capable of implementing passwords, typically do NOT as users find them irritating. Programs such as war dialers allow the hacker to find these active systems and attempt connection with compatible programs and take over the internal system. With this method, it’s not difficult to infitrate a network and gain access to valuable corporate computing assets.

2.8.2 Call-back Capability

Some dial-in servers offer call-back facilities (i.e., the user dials in and is authenticated, then the system disconnects the call and calls back on a specified number). Call-back is useful since if someone were to guess a username and password, they are disconnected, and the system then calls back the actual user whose password was cracked; random calls from a server are suspicious, at best. This does mean users may only log in from one location (where the server is configured to dial them back), and of course there may be phone charges associated with there call-back location.

This feature should be used with caution; it can easily be bypassed. At a minimum, make sure that the return call is never made from the same modem as the incoming one. Overall, although call-back can improve modem security, you should not depend on it alone.

2.8.3 All Logins Should Be Logged

All logins, whether successful or unsuccessful should be logged. However, do not keep correct passwords in the log. Rather, log them simply as a successful login attempt. Since most bad passwords are mistyped by authorized users, they only vary by a single character
from the actual password. Therefore if you can't keep such a log secure, don't log it at all.

If Calling Line Identification is available, take advantage of it by recording the calling number for each login attempt. Be sensitive to the privacy issues raised by Calling Line Identification. Also be aware that Calling Line Identification is not to be trusted (since intruders have been known to break into phone switches and forward phone numbers or make other changes); use the data for informational purposes only, not for authentication.

2.8.4 Choose Your Opening Banner Carefully

Many sites use a system default contained in a message of the day file for their opening banner. Unfortunately, this often includes the type of host hardware or operating system present on the host. This can provide valuable information to a would-be intruder. Instead, each site should create its own specific login banner, taking care to only include necessary information.

Display a short banner, but don't offer an "inviting" name (e.g., University of XYZ, Student Records System). Instead, give your site name, a short warning that sessions may be monitored, and a username/password prompt. Verify possible legal issues related to
the text you put into the banner.

For high-security applications, consider using a "blind" password (i.e., give no response to an incoming call until the user has typed in a password). This effectively simulates a dead modem.

2.8.5 Dial-out Authentication

Dial-out users should also be authenticated, particularly since your site will have to pay their telephone charges.

Never allow dial-out from an unauthenticated dial-in call, and consider whether you will allow it from an authenticated one. The goal here is to prevent callers using your modem pool as part of a chain of logins. This can be hard to detect, particularly if a hacker sets up a path through several hosts on your site.

At a minimum, don't allow the same modems and phone lines to be used for both dial-in and dial-out. This can be implemented easily if you run separate dial-in and dial-out modem pools.

2.8.6 Make Your Modem Programming as "Bullet-proof" as Possible

Be sure modems can't be reprogrammed while they're in service. At a minimum, make sure that three plus signs won't put your dial-in modems into command mode!

Program your modems to reset to your standard configuration at the start of each new call. Failing this, make them reset at the end of each call. This precaution will protect you against accidental reprogramming of your modems. Resetting at both the end and the beginning of each call will assure an even higher level of confidence that a new caller will not inherit a previous caller's session.

Check that your modems terminate calls cleanly. When a user logs out from an access server, verify that the server hangs up the phone line properly. It is equally important that the server forces logouts from whatever sessions were active if the user hangs up unexpectedly.

2.9 Dial Up Security Issues

Customer organizations require that scientists, management, sales and other personnel be able to remotely access systems on the customer network. This is frequently done via dial-up phone lines through MODEMs and via X.25 public data network (PDN) terminal connection from other locations in the U.S. and around the world.

Because of the common use of MODEMs and the relatively low expense, it is simple for anyone to acquire and use MODEM technology to dial-up and connect to any system that supports MODEM connectivity. This supports the ability for criminals and competitive elements to acquire technology to infiltrate computer and network systems and engage in illegal or, at a minimum, highly annoying activities centered around the ability to disrupt operations, steal information , etc.

2.9.0 Classes of Security Access Packaged for MODEM Access

In the security access business, there are the following types of systems available for providing differing levels of security facilities for dial-up MODEM access:

- MODEMs with internal security facilities. These systems provide some levels of password authorization and access methods and are fairly inexpensive. They are also vendor proprietary and provide limited flexibility in multiple protocol remote access environments.
- MODEM pool management systems. These tools usually run on a PC or equivalent system and provide specific numbers of dial-up ports with some management software to provide minimal security facilities as well as user accounting for very specific protocols or terminal access facilities. These are useful for small sites or sites where all access is always the same protocol method and high levels of security and reporting are not required.
- Asynchronous access software/hardware facilities. Many vendors of specific protocol server solutions provide MODEM dial-up asynchronous protocol access methods which allow dial-up to file and mail servers utilizing their own proprietary methods. This allows ease of user access, but these packages limit their security facilities to whatever the server provides, which is usually password-only authentication, and are limited to whatever user tracking facilities the server technology allows which is also limited at best.
- Multiprotocol server facilities. Vendors of asynchronous connection hardware and software are providing all-in-a-box units that allow customers to dial-up to the box and the user select a connection protocol such as AppleTalk, IP, IPX, DECnet or a proprietary protocol of choice. These solutions are very versatile, but frequently provide single password facilities for users and usually provide no tracking software of user activities whatsoever. Also, these boxes are limited to the number of supported dial-up ports and usually do not support security

facilities for simple terminal emulation to systems such as IBM's MVS/XA and OS/400, UNIX, OpenVMS, etc.

- Terminal servers. Many vendors of terminal servers allow MODEM connection facilities which allow many dial-up user connections. These devices are becoming more flexible as they not only offer the traditional terminal access facilities for terminal emulation to mini's, supermini's, mainframes and supercomputers, they also are supporting asynchronous access to TCP/IP's SLIP and PPP protocols, AppleTalk, IPX, etc. The problem with this approach is an extremely limited security access facility (it is frequently limited to a terminal server-wide password which everyone has access to use), limited access speeds, non-flexibility of hardware and limited user tracking and reporting.
- "Small" routers. Many of the major router vendors are building small, inexpensive router systems that provide asynchronous access facilities as well as router access software to existing LAN and WAN resources. These provide extremely limited security facilities, if any at all, but are useful due to their inexpensiveness and ease of integration in to existing networks.
- All-inclusive MODEM and remote access control systems. This is a relatively new class of MODEM access security system that allows terminal emulation facilities, remote protocol access capabilities, user authentication methods, security facilities (passwords, accounting, session tracking, live monitoring, exception handling, alarms, etc.), user menu facilities, user profile tracking and multiple hardware facility access (Ethernet/802.3, token ring/802.5, FDDI, ISDN, ISDN-B, ATM, etc.) all at the same time from the same facility. These types of systems are complex and very capable and are rapidly becoming the system of choice for sites with many differing types of dial-up requirements for many different types of systems.

While this does not provide an all-inclusive list of access facilities, it serves as an illustration of what has traditionally been available. Most of these tools are limited to either a traditional RS-232, RS449, RJ11 or RJ45 interface to a given system. In some of the server access facilities, Ethernet/802.3 or token ring/802.5 LAN access are also supported for access to remote servers as well as local resources.

2.9.1 Tactical and Strategic Issues in Selecting a MODEM Connection Solution

In most sites considering dial-up facilities, the need is real and is not going away. Many companies are becoming more mobile and the need for remote dial-up access is becming critical. It is estimated in 1999 that over 60% of all computers that will be sold will be notebook sized or smaller. This, coupled with the trend towards docking-station systems that can be moved at will, provides a market for remote access that is growing dramatically and does not show any signs of diminishing. Further, practically all consumer-level computers come equipped with a 56kbps V.90 MODEM.

Where most sites fail in their tactical and strategic planning for such facilities is in the expectation that they can contain the requirement for dial-up and that they can dictate the user's options. What happens in many situations is the users will implement their own solutions and not provide any feedback to IT facilities until it has become firmly entrenched in the deliverable solutions for management. As a result, the opportunity to control the unauthorized facilities is reduced to nil and the IT groups must deal with a myriad of dial-up options based upon what was planned and what happened "on its own."

From a tactical perspective, it is better to provide the solution in a manner that is acceptable to the users before they have the opportunity to circumvent the dial-up solution with a substandard solution that will be incorporated due to default access.

If dial-up solutions are in place, it is tactically wise to implement substitute solutions that provide the following features:

- Does not affect the user's computing budget. People always like something they feel is "free."
- Does not impose too much more additional effort to use
- Provides a substantial improvement over the current method of dial-up such that the new method is immediately attractive regardless of new user effort required to use it
- Allows greater user flexibility, speed and access facilities

While most of this is common sense, it is interesting how many companies provide an inferior solution to current user access methods or a one-for-one solution which irritates users with new procedures and facilities. No one wants to deal with a step-back in productivity or technology. Stepping forward, however, has to show a reasonable increase in productivity or user-desired features or it will be unacceptable as well.

From a strategic perspective, companies need to consider what dial-up protocols will be required, speed of access to remote facilities and eventual hardware facilities that will be used on internal and external networks. Many companies will start off with LAN technologies such as Ethernet/802.3 and token ring/802.5 networks and eventually implement 100mbps LAN/MAN technologies such as FDDI. This eventually leads to the inevitable implementation of ISDN-B, ATM and SONET access. Any remote access facility needs to be upgradeable to these environments as the company requirement grow.

Of importance in the selection of any solution is the realization that MODEMs are, technologically, on the way out as digital communications replace analog facilities in the phone systems of the world. Some telecommunications providers already provide direct ISDN and ISDN-B facilities which allow a technology called unbundled ISDN services. In this offering, the local equipment company (the LEC), provides a T1 connection to the customer site, divided into 24 separate 56kbps digital channels. At the LEC, MODEM emulation is provided to a dial-up user which is converted to a digital channel access to one of the channels to the customer. The effect is that the customer does not need to purchase any MODEMs, the user population can use existing MODEM technologies and when the phone system goes pure digital in the future, there are no corporate MODEM banks to replace. Since the trend is to go digital, the need to support ISDN, ISDN-B and ATM is crucial for long term user satisfaction and in the support of alternate connection technologies in the future.

2.9.2 Background on User Access Methods and Security

To access any system via terminal, a user is expected to enter, as a minimum, some type of user identification (such as as user ID, username, or some other identifier), a password, and other optional login information as may be required by the systems or network manager. In some situations, an additional “system” password is used before the user ID to allow the system to automatically detect access baud rate as well as provide the user the opportunity to enter a general access password in order to gain entry in to the system or front-end being used. To enhance system security for dial-up access, other methods may also be added such as digital ID cards, dial-back MODEMs that reconnect the user to the system after the system dials the user back, and other types of electronic equipment security denial or restricted access methods.

Some of the security flaws with this level of access in the general systems area are:

- The steps above allow the opportunity to exploit flaws in the access method as it is by rote, mechanical in nature, and easily analyzed
- Simple access methods simplify user access efforts, but do not keep general security intact. Because users share information and also leave security access information in compromising locations, the information must change or be generally compromised
- Most system access methods are highly susceptible to an exhaustive attack from the terminal access methods (dial-up, X.29, and others) via something as small as a personal computer
- Many users are never physically seen by the systems personnel and their login information is frequently transmitted to them via phone call or facsimile, which is highly subject to be compromised

Few operating systems provide intensive monitoring and activity recording facilities to help trace sources of intrusion and to also detect unauthorized usage

- Few companies trace employees who have left the firm and properly clean up access methods for employees. The result are accounts that exist, sometimes for years, before they are deleted or even changed.
- For companies with highly mobile employees or employees that travel extensively, dial-back MODEM management is extensive and time consuming. Further, within the next 12-24 months from this writing, many MODEM devices will be rendered in-effective due to pure digital phone systems such as ISDN coming on-line and replacing current analog offerings
- Dial-back MODEM units are not compatible, in some cases, with foreign system access due to CEPT or ITU-T incompatibilities with phone systems (ITU-T E.163 POTS and V series standards), carrier frequencies, DTMF tone levels, and other electronic incompatibilities. As such, some dial-back systems will not work with some foreign phone systems which can cause problems for a multinational corporation.
- None of the current systems direct user logins to a specific destination; they only restrict access to “a” system of some sort
- No current user interface logins allow for protocol security for asynchronous connections via DECnet Phase IV, TCP/IP PPP or SLIP links, asynchronous AppleTalk or other types of protocols that support an asynchronous interface
- Security encryption cards and other electromechanical interface devices are frequently lost and are expensive to replace and manage
- Dial-back modems are subject to abuse by use of phone system features such as call forwarding

For these reasons and others too numerous to mention in a short summary, the author, Dr. Hancock, believes that many currently available commercial dial-up access security products are inadequate for a secure information access method to systems on a computer network.

With the rise of computer crime via dial-up access, there is a natural paranoia that systems professionals are required to recognize: dial-up access makes system access possible for non-authorized individuals and this exposure must be minimized. The reasons for keeping non-authorized individuals out of customer systems include:

- Potential discovery and publication of sensitive internal memoranda
- Industrial espionage
- Destructive systems interference (”hacking”) by unauthorized individuals
- Potential virus infestation from external sources

- Isolation of company proprietary data from unauthorized individuals (such as food and drug filings, patent data, primary research data, market information, demographics, corporate financial data, test and research results, etc.)
- Potential for external sources to “taint” valid data, causing the data to appear valid and cause irreparable harm
- Potential safety hazards if manufacturing or other production systems were accessed from external sources and process control software were changed or modified in some way

There are many other examples, but these give the general issues on why restrictive connectivity is required at customer sites. Also, as recent as late 1993, customer research centers have experienced multiple attempts at system compromise from external sources via dial-up and X.29 terminal pad connection. While no specific break-in was detected, the attempts have been numerous and getting more creative with time. It was deemed necessary to improve terminal connectivity security procedures.

Some customers have used dial-back MODEMs and hardware security cards for user terminal access.

The dial-back MODEMs, while previously useful, are now easier to violate due to new phone system facilities offered by regional telephone companies. Facilities such as call forwarding, call conferencing and other facilities that will be offered via Signaling System 7 (SS7) and Integrated Services Digital Network (ISDN) connectivity facilities make the general functionality of dial-back MODEMs easier to violate (dial-back facilities could be re-routed via the phone system to other locations other than the phone number expected and desired) and a total lack of security on the phone network itself helps to propagate this effort.

In recent months, the hackers magazine *2600* has published articles on how to provide remote call-forwarding and how to “hack” public phone switching systems and access a variety of information including call routing tables. With this type of information, potential disruptors of corporate dial-up methods can forward calls to any desired location.

A recent example is that of Kevin Poulsen in California, who successfully "hacked" the local phone switch over a period of two years. The result was interesting. He successfully made his personal phone line the only one able to gain access to radio station lines and busy-ed out all other lines to make himself the winner of numerous phone offers. His winnings included two Porches, two trips to Hawaii and over $22,000.00 in cash. Investigation by the FBI showed that Poulsen accessed much, much more than the stated "hacks" and was charged with a long list of crimes including computer fraud, interception of wire communications, mail fraud, money laundering, obstruction of justice, telecommunications fraud and others. His primary vehicle was access to the telephone switching system, which effectively defeats any type of dial-back facility which depends on the phone system to be "untouched."

Devices such as security identification cards, approximately the size of a credit card and possessing verification algorithms that allow exact identification of a user, are very secure provided that they are not shared between users. They are also somewhat expensive (est. $60.00 per user) and are easily destroyed (sat upon, placed in washing machines, etc.) or lost. Because of accounting problems and the size of the dial-up population, some former employees have left customer’s employ and taken their cards with them making recovery virtually impossible. There are also some terminal connection facilities in which security identification cards will not work and this requires another approach to the problem.

Such cards work by the user entering a number when prompted by the destination system, in a specified amount of time, that is visible in an LCD window in the card. This number is synchronized with the destination system and, algorithmically, the number should decypher to a valid combination the system will accept.

Another type of security access method, called a token card, works on the concept that the card cannot possibly be in any one else's possession. This is accomplished by installation of token hardware and software in notebook computers and, in some cases, in the inclusion in operating system ROMs on the motherboard of the remote system. While secure and the loss levels are low, the costs are serious and severely restrict the types of remote systems that may access a centralized dial-up method as well as the type of dial-up or remote access method available.

In many circumstances there is the problem of identifying who has left the firm (and when) so that their security card information may be removed from the access database. At present, there are former customer employees that have left their firms some time ago and are still identified as being active users in the security card database. While this is mostly an accounting and tracking problem, there is no automated “user X has not logged in via dial-up in Y amount of time” facilities to allow tracking of user activity levels.

Even with proper accounting and user tracking, there is a recurring expense required for the use of security identification cards (replacements, failed units, damaged units, etc.) and this is growing due to the number of people desiring access to the system resources at customer sites.

A major problem with security cards and token cards is the problem of user accounting and session tracking. Many products provide a method by which users may be accounted for in terms of access time and line identification, but that is about it. There are no investigative tracking facilities, session tracking facilities, session capture (for the extreme cases), user profiling and many other required features for proper investigation of penetrations or improper activities.

What consumers require is an easy-to-use secure dial-up access method that allows different types of terminal connection platforms (dial-up async, sync, X.29 dynamic PAD access, etc.) to customer system resources. Further, the system must use off-the-shelf hardware to keep the short and long term costs of dial-up low and support multiple terminal protocol facilities. Finally, the interface must have logging and auditing facilities useful in user tracking and user access abnormality detection by monitoring user activity profiles and reporting such information to systems personnel for action.

2.9.3 Session Tracking and User Accounting Issues

In any dial-up solution, there is the need to provide reports on user access, where the user connected and rudimentary reporting of times, activity levels and dates of access for accounting facilities.

Where many companies find problems *after* implementation are the issues of tracking down breaches of security or monitoring specific user activities for users performing activities that are considered counterproductive to corporate goals or illegal. Even if the system is successful in keeping out unwanted intruders, many company security breaches are from employees or contractors working within the company facilities. Tracking of activities is important when attempting to isolate

internal breaches, the most common type, and when trying to isolate illegal activities.

Tracking may be done in a variety of manners. The easiest is when the system is set up to detect deviations from established access and activity patterns and reports alarms on deviations. Unfortunately, setting up such facilities is non-trivial in larger dial-up environments where there may be hundreds or thousands of accounts. What is needed is software facilities that will establish a normalization baseline on a user-by-user basis and then provide a method to report anomalies and deviations from established operations.

Once the dial-up system has detected deviations, reporting and session management/capture facilities need to be activated to properly identify user actions and track activities to the keystroke level. This provides a chain of evidence of malfeasance and can be used to procecute a malicious user or to prove the innocence of falsely accused users. Evidence is essential in any security breach or suspected misuse of system and network resources. Keeping people off of systems is not terribly difficult and there are well established manners in which this is done. Tracking them, developing a reliable trail of activity patterns and evidence that may be used for procecution is difficult and the system has to be designed from the start to provide this level of information.

Reporting for user access needs to be very dynamic for the production of accounting report for chargeback and also

2.9.4 Description of Proposed Solution to Dial-Up Problem

The author, has implemented various types of secure access systems for various types of customers requiring dial-up network access without using dial-back MODEMs. The most productive and flexible method to do this is to use an intermediate network connection to provide connectivity and access services. This may be accomplished through the use of a local Ethernet, terminal servers, and a small 32-bit or 64-bit system to provide dial-up connection authorization. Graphically, the connection path would appear as follows:

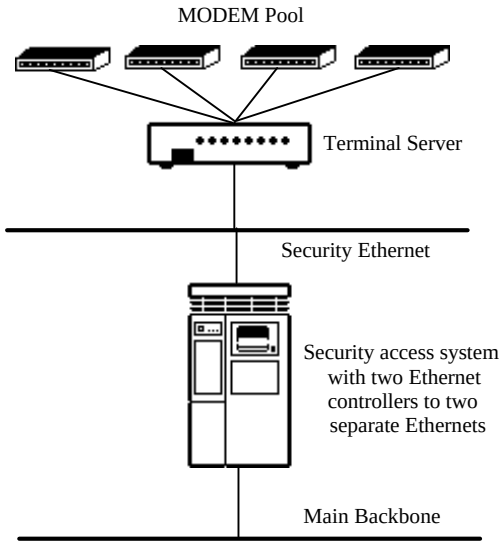


Figure 1: Architectural Drawing of Secure Front-End Simple Configuration

In a typical usage scenario, users dial up to a customer specified phone number pool with V.32bis, V.34, V.90 or similar MODEMs (this allows 300 through 56Kbps async dial-up). The number pool, due to the nature of the software, could be a toll-free access number (800-type in the U.S. and Canada) or a connection number and ID on a public data network (X.25/X.29). The security access server(s) would then automatically connect the user to special login security software that would ask for a username, password, and any other type of required information. In this manner, should it be necessary, a terminal emulation request, an asynchronous protocol connection (such as PPP, SLIP or async AppleTalk) could be authorized or other type of connection protocol. Following authorization and authentication of the user over the dial-up connection, the security system software would connect the dialed-up user to a system on the main Ethernet backbone at the customer's site. This would allow the secure access server system to provide very specific connection facilities on a user-by-user basis and at the system and network manager's discretion. Based upon previous implementations at other facilities, this type of connectivity would prove useful to customers where security is a serious concern and yet remote access to the network and systems thereon is essential to fulfilling corporate needs and goals.

Positive-acknowledgement systems, also sometimes called extended user authorization systems (EUAS), are those that require user action to initiate connection to or from a system. In the case of most customer sites, the system will require the user to provide positive identification via the following methods:

- Access password upon initial MODEM or system connection to the secure front-end in a manner similar (but not the same as) to many pre-user password security methods. This allows connection but does not divulge the corporate identity, which is usually the first place that a "hacker" would receive information on what company is being attacked.

- Specific pre-defined user ID and password through a special front-end system on the dial-up Ethernet segment. This is designed in such a way as the user will not be able to tell that he/she is actually connected to a security screening system. This is provided to simplify the user access and not divulge system identity or corporate identity as well as provide a highly secure access method.

- Following identification look-up and acknowledgement (which will be done via secure cryptography, not a hashing mechanism as used in most operating systems or suggested in ITU-T X.509), the user will either be presented with a menu of services he/she is allowed to access or connected to the only network service he/she may be allowed to access. Since the menus are customizable, the user will not be allowed to roam the network looking for connection points.

- The user would then be required to log in to the destination system via normal log-in procedures for that system.

An additional alternative is to use personal access cards on the remote systems prior to connection. While user card access at the remote facility is desirable, the ISO standard for such access is being experimented with at this time in X.72 and X.75 standards (and, by default, X.25) and is having great difficulty in properly forwarding the ID values. It is the opinion of the author that card access is definitely desirable in the future but is much too immature for the variety of dial-up connections and remote facilities that customer sites are expected to support. Further, the ISO standard will most likely change in the next year which would cause a re-write of any card access programming (this could get costly and delay any

upgrades for a considerable time). At a meeting of the ISO group working on the X.75 test, serious problems were raised with the issues of secure cards and credit card authorization facilities in public access networks and it was decided that a considerable amount of additional work is required before these can effectively be used for secure access.

As a side issue, a successful network break-in in France's PTT Minitel videotex system was accomplished by using a PC to emulate card key access. The PC was a portable laptop and the program was written in Turbo C, a common and inexpensive compiler. This has caused proponents of card and digital signature access to re-think how the formats of data are provided from the card access method.

2.9.5 Dissimilar Connection Protocols Support

One feature of remote access facilities are their ability to connect to remote systems via network or async connection(s). The user may log in to the remote access system and then be connected to a networked system on the corporate network in a variety of ways.

Because of the manner in which terminal session management is done, some remote access systems are capable of acting similar to a terminal "gateway" between protocol types. This means that a user may connect via dial-up to the remote access system and then request an SNA terminal connection to a mainframe. A user from a remote UNIX system may connect with Telnet via the network to the remote access system and then be re-connected by the system to an Alpha AXP system using DECnet's CTERM protocol.

2.9.6 Encryption/Decryption Facilities

Some remote access systems use the ANSI Data Encryption Standard (DES) for encryption and decryption of files in U.S. installations and an exportable hashing algorithm for installations outside the U.S. This is due to exportation of encryption technologies laws in the U.S. and is not a reflection on the vendor's desire for customers in the international marketplace to have less secure installations than those in the U.S. The vendors in the U.S. have no control over this law and must comply.

Some remote access products do not store sensitive files on disk in an unencrypted manner. All screen captures, user information and other files that are sensitive in nature are encrypted in real-time and stored on disk in an encrypted form. Should files be backed-up and moved to another system, the files will be unintelligible when printed or sent to a terminal screen.

Remote access products with session and information capturing facilities have the ability for a system manager to store captured data for a user in a file. When stored, the file buffers are encrypted prior to being written to disk. If the system manager wishes to view the file, the file is retrieved from disk and decrypted "on-the-fly" and viewed with a special encrypt/decrypt editor.

2.9.7 Asynchronous Protocol Facilities

Secure remote access servers often provide the ability for the system manager to set up specific user accounts for asynchronous DECnet access, TCP/IP's SLIP protocol, asynchronous AppleTalk and others. The user must go through the standard security login dialog and, when the user has been authenticated, the line is automatically modified and converted to an asynchronous protocol port. Some

systems allow multiple protocol access and a user menu may be provided for access to various protocol services.

2.9.8 Report Item Prioritization

One of the more aggravating items in generation of reports is having to wade through the amount of paper generated to find truly significant events and take appropriate action.

Some remote access servers allow the system manager to set priorities (critical, urgent and routine) on various data items in the system. In this manner, as security exception reports are generated they may be printed in priority order. When a security exception report is read by the systems or security manager, the report may be organized such that high-priority items are at the beginning of the report, precluding a search operation to find what is truly important in the report.

2.9.9 User Profile "Learning" Facility

When designing secure remote access servers, the author found that one of the worst situations was the lack of knowledge of who logged in to systems "when." While some operating system environments could allow the system manager the flexibility to specify login times to be at specific times of the day, these facilities are very rarely used as it was deemed too difficult to set up and figure out what times of the day the user is active.

Some systems now have an autoprofiling feature, which may be enabled for the entire system or on a user-by-user basis. This allows the secure access server to "learn" how a user interacts with systems on the network. The secure access server collects activity levels and time of day parameters, stores them and sets up, automatically, an activity profile for the user. If the user attempts to log in to the secure access system at times not specified by the profile, access is denied. Further, if operating parameters during a login session exceed the learned "norm," the user may be disconnected. Obviously, there are user-by-user overrides available to the system manager that may be set-up to allow individual user flexibility. For large user count sites, this feature has proven to be very valuable and allows establishment of activity patterns and detection of abnormalities (this is the first step to detecting illicit connectivity).

2.10 Network Security

1. Ensure that any message sent arrives at the proper destination.
2. Ensure that any message received was in fact the one that was sent. (nothing added or deleted)
3. Control access to your network and all its related parts. (this means terminals, switches, modems, gateways, bridges, routers, and even printers)
4. Protect information in-transit, from being seen, altered, or removed by an unauthorized person or device.
5. Any breaches of security that occur on the network should be revealed, reported and receive the appropriate response.
6. Have a recovery plan, should both your primary and backup communications avenues fail.

Things to consider in designing a network security policy (as covered earlier).

1. Who should be involved in this process?
2. What resources are you trying to protect? (Identify your assets)

3. Which people do you need to protect the resources from?
4. What are the possible threats? (Risk assessment)
5. How important is each resource?

Unless your local network is completely isolated, (standalone) Your will need to address the issue of how to handle local security problems that result from a remote site. As well as problems that occur on remote systems as a result of a local host or user.

What security measures can you implement today? and further down the road?

*Always re-examine your network security policy to see if your objectives and network circumstances have changed. (every 6 months is ideal.)

2.10.0 NIST Check List

NIST Checklist for functions to consider when developing a security system The National Institute for Standards and Technology (NIST) has developed a list for what they refer to as Minimal Security Functional Requirements for Multi-User Operational Systems. The major functions are listed below.

1. Identification and authentication - Use of a password or some other form of identification to screen users and check their authorization.
2. Access Control - Keeping authorized and unauthorized users from gaining access to material they should not see.
3. Accountability - Links all of the activities on the network to the users identity.
4. Audit Trails - Means by which to determine whether a security breach has occurred and what if anything was lost.
5. Object Reuse - Securing resources for the use of multiple users.
6. Accuracy - Guarding against errors and unauthorized modifications.
7. Reliability - Protection against the monopolization by any user.
8. Data Exchange - Securing transmissions over communication channels.

2.10.0.0 BASIC LEVELS OF NETWORK ACCESS:

1. Network Supervisor- has access to all functions including security.
2. Administrative Users- a small group given adequate rights to maintain and support the network.
3. Trusted Users- users that need access to sensitive information.
4. Vulnerable Users- users that only need access to information within
5. their job responsibilities.

2.10.1 Auditing the Process

Making sure your security measures work is imperative to successfully securing your data and users. You have to make sure you know who is doing what on the network. Components of a good audit will include;

1. A log of all attempts to gain access to the system.
2. A chronological log of all network activity.
3. Flags to identify unusual activity and variations from established procedures.

2.10.2 Evaluating your security policy

1. Does your policy comply with law and with duties to third parties?
2. Does your policy compromise the interest of your employees, your company or third parties?
3. Is your policy practical, workable and likely to be enforced?
4. Does your policy address all of the different forms of communication and record keeping within your organization?
5. Has your policy been properly presented and agreed to by all concerned parties?

With adequate policies, passwords, and precautions in place, the next step is to insist that every vender, supplier, and consultants with access to your system secure their computers as adequately as you secure yours. Also, work with your legal department or legal advisors to draft a document that upon signing it would recognize that the data they are in contact with is yours.

2.11 PC Security

One of the most critical security issues, one that has been compounded by the micro and LAN/WAN revolution, is a lack of awareness, by executives and users, to the vulnerability of their critical and sensitive information. Microcomputers have unique security problems that must be understood for effective implementation of security measures. These problems include;

- Physical Accessibility
- Hardware
- Software
- Data Communications
- Networking
- Disaster Recovery

Physical Accessibility

Several approaches need implementing in order to provide the necessary security for microcomputers.

- Hardware Solutions
- Locks
- Desk Mounts
- Enclosures
- Steel Cables

Disk locks are also available to prevent access to hard drives and diskette drives. Planning and diligent administration are the keys to securing microcomputers and the information they process.

An increasing problem in most organizations is microcomputer and/or component theft involving personnel within the company as well as outsiders. Some of these components are easy to carry away in a purse, briefcase, or coat pocket. Organizations that lack accurate or current inventories of their PC equipment, components and peripherals are the most vulnerable.

A situation similar to automobile "chop shops" has become prevalent in the PC industry. Black market sales of "hot" PC parts are costing corporate America over $8 billion a year.

Things to consider in regards to system security

1. Can the Casing on the equipment be removed by unauthorized personnel.
2. Are notebook and laptop computers secured to desktops.
3. Is peripheral equipment such as CD ROM readers, tape back up units and speakers secured to desktops.
4. Are floppy drives secure from the introduction of unauthorized software, viruses or the removal of confidential corporate information.

Software Solutions

Viruses have left a number of corporations sadder but all the wiser. A virus can change data within a file, erase a disk, or direct a computer to perform system-slowing calculations. Viruses may be spread by downloading programs off of a bulletin board, sharing floppy diskettes, or communicating with an infected computer through a network, by telephone or through the Internet. Anti-virus products are a necessity for the detection, eradication and prevention of viruses. In addition, micro security policy should define permissible software sources, bulletin board use, and the types of applications that can be run on company computers. The policy should also provide standards for testing unknown applications and limit diskette sharing.

Data Residue is data that is stored on erased media. Such data can often be read by subsequent users of that media. This presents a danger in sharing files on diskettes that once contained sensitive or confidential data. This problem also exists for hard drives. One solution available to companies is the use of degausser products. Primarily used by the US government, corporate America is now finding these effective tools for preventing the disclosure of sensitive information.

2.12 Access

2.12.0 Physical Access

Restrict physical access to hosts, allowing access only to those people who are supposed to use the hosts. Hosts include "trusted" terminals (i.e., terminals which allow unauthenticated use such as system consoles, operator terminals and terminals dedicated to special tasks), and individual microcomputers and workstations, especially those connected to your network. Make sure people's work areas mesh well with access restrictions; otherwise they will find ways to circumvent your physical security (e.g., jamming doors open).

Keep original and backup copies of data and programs safe. Apart from keeping them in good condition for backup purposes, they must be protected from theft. It is important to keep backups in a separate location from the originals, not only for damage considerations, but also to guard against thefts.

Portable hosts are a particular risk. Make sure it won't cause problems if one of your staff's portable computer is stolen. Consider developing guidelines for the kinds of data that should be allowed to reside on the disks of portable computers as well as how the data should be protected (e.g., encryption) when it is on a portable computer.

Other areas where physical access should be restricted is the wiring closets and important network elements like file servers, name server hosts, and routers.

2.12.1 Walk-up Network Connections

By "walk-up" connections, we mean network connection points located to provide a convenient way for users to connect a portable host to your network.

Consider whether you need to provide this service, bearing in mind that it allows any user to attach an unauthorized host to your network. This increases the risk of attacks via techniques such as IP address spoofing, packet sniffing, etc. Users and site management must appreciate the risks involved. If you decide to provide walk-up connections, plan the service carefully and define precisely where you will provide it so that you can ensure the necessary physical access security.

A walk-up host should be authenticated before its user is permitted to access resources on your network. As an alternative, it may be possible to control physical access. For example, if the service is to be used by students, you might only provide walk-up connection sockets in student laboratories.

If you are providing walk-up access for visitors to connect back to their home networks (e.g., to read e-mail, etc.) in your facility, consider using a separate subnet that has no connectivity to the internal network.

Keep an eye on any area that contains unmonitored access to the network, such as vacant offices. It may be sensible to disconnect such areas at the wiring closet, and consider using secure hubs and monitoring attempts to connect unauthorized hosts.

2.13 RCMP Guide to Minimizing Computer Theft

2.13.0 Introduction

Increasingly, media reports bring to light incidents of thefts occurring in offices at any time of the day or night. Victims include government departments, the private sector and universities in Canada and in the United States. The targets: computers and computer components. Perpetrators include opportunists, petty thieves, career criminals, organized gangs, people legally in contact with the products, e.g. transportation and warehouse workers, as well as individuals working in the targeted environment.

While incidents of this nature have increased dramatically in the last few years, the number of reported incidents reflect only a portion of the total number of occurrences. One reason for this is that government institutions, the private sector and universities alike are often reluctant to report such incidents, for fear they'll be ridiculed or that their operations will be negatively affected.

Advances in electronics and the miniaturization of components have provided thieves with ideal targets — expensive items that are easily concealable, readily marketable and hard to trace. Components can be transferred from thief to middleman to a distributor without anyone knowing they are stolen. Items such as cellular phones, laptops, integrated circuits, electronic cards, disk drives and CD-ROMs have become the target of choice of both novice thieves and career criminals.

This publication identifies the primary areas of vulnerability that may lead to loss of assets (computer components) and proposes safeguards designed to minimize the risks of losing these components. Samples of physical security devices are described, and strategies are offered for minimizing computer and component theft.

2.13.1 Areas of Vulnerability and Safeguards.

2.13.1.0 PERIMETER SECURITY

Minimizing Perimeter Security Vulnerabilities
Examining the perimeter security of a building is the first step and involves establishing appropriate safeguards, through target hardening. Target hardening is the process of setting up a series of physical barriers (protection) to discourage an adversary's progress. The objective is to have an adversary either give up the idea of an attack, give up during the attack, or take enough time for a response force to react to the attack before its completion. A building's entrances exits and trade entrances are vulnerable areas that should be the focal point for enhanced perimeter security.

The following checklist can help determine the security posture of the perimeter:

- Is the building secured at ground or grade level by locked doors, using heavy-duty commercial hardware (locks, hinges)?
- Are the windows at ground level either fixed or locked with heavy-duty commercial hardware?
- Are trade entrances locked or controlled or are they wide open to strangers?
- Are rooftop openings locked with heavy-duty commercial hardware if accessible from outside the building?
- Does the building have an outside ladder? If so, is the ladder secure?
- Is it protected with a ladder barrier to prevent unauthorized access to the roof?
- Do employees work during the evening?
- Is there sufficient lighting surrounding the building, including the parking lot and service entrances?

Examples of Enhanced Perimeter Security Safeguards

- Alarm grade level doors and windows against opening and breakage.
- Ensure day and night security patrols are conducted by security personnel.
- Monitor the building perimeter by CCTV.
- Install entry security controls for single-tenant facilities, or in facilities shared with other government departments requiring the same level of security.
- Whenever possible, avoid multi-tenant buildings where private tenants do not want entry controls.
- Surround the building with tamper-proof lighting fixtures. Position the security lighting to prevent deep shadows from the building or vegetation, so intruders can be noticed.

2.13.1.1 SECURITY INSIDE THE FACILITY

Minimizing Vulnerabilities Inside the Facility
Once the building perimeter has been secured, the next important step is controlling personnel, visitors and equipment entering and exiting the building. One effective method to maximize the control and usefulness of security staff is to have all employees and visitors enter the facility through one entry point, with material entering at another identified entry point. It is recognized that with high-occupancy or multi-tenant buildings it may not be practical to have a single entry point. Departments providing services to the public should be located on the main floor, to limit access to working areas. Only authorized employees and supervised visitors should have access to operational areas. All service vehicles should enter the site through a single vehicle control point. Canteens, lunch rooms and stores should be designed and situated such that deliveries to and from

such areas do not have to enter the secure perimeter. Every facility should have a reception zone, accessed directly from the public-access zone, where visitors, if necessary, wait for service or for permission to proceed to an operational or secure zone. If this process cannot be accommodated then each floor must be secured. Other security vulnerabilities include the improper use of a guard force and granting unlimited access to all areas of the building's working or technical areas, e.g, electrical and telephone rooms.

Examples of Enhanced Safeguards Inside a Facility

- Establish reception points at interface points between functional groups or secure zones.
- Do not use stairs forming part of a means of egress to enter office environment.
- Establish access controls, either manually, mechanically or electronically.
- Establish different public access zones, operational zones and security zones.
- Clearly define the limits to which public access is permitted, through signage.
- Control access to floors through short distance stairs (i.e. circulation stairs) running between floors.
- Do not allow elevators to stop on all floors during silent hours, unless persons have been granted access by key, access card or the entry control desk.

2.13.2 Physical Security Devices

Minimizing Vulnerabilities Using Physical Security Devices
Physical security devices are another method of preventing unauthorized use, intentional damage or destruction, or theft of computer equipment and components.

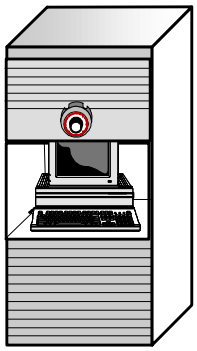
Many different devices are available on the market, including alarms, locks, cabinets, cable kits, lock-down plates and special security screws. One company has marketed theft retrieval software that notifies police of a stolen PC's whereabouts. The use of security seals tamper-evident labels and ultraviolet detection lamps is also being implemented.

The RCMP has not endorsed these products, other than containers, because the majority have not been tested to evaluate their effectiveness. Some of the products may be useful, but may not be cost-effective. In many instances, it is more cost-effective to protect the working area than it is to tie down or alarm each PC.

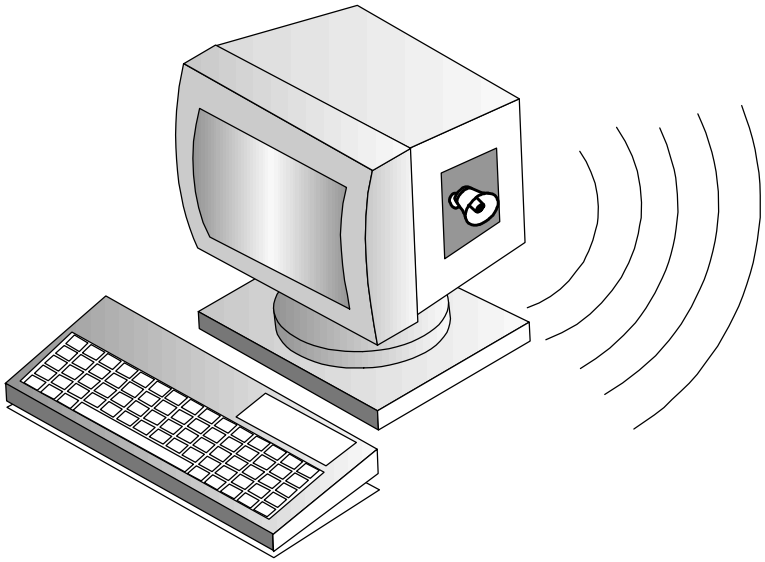
Labelling, engraving and ultraviolet detection is time-consuming to implement; and inventory has to be kept up-to-date. In addition, there is little to indicate that these methods will reduce thefts. Laptops and portable computers are usually stolen for personal use or for resale. The buyer knows the item has been stolen but is willing to take the chance of receiving stolen goods because of the low price and the improbability of being caught.

2.13.2.0 EXAMPLES OF SAFEGUARDS

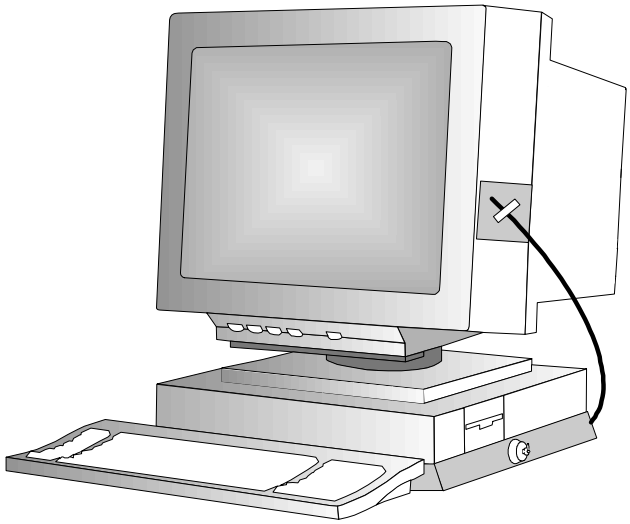
Cabinets enclose the entire computer, including the monitor, keyboard, printer and CPU. Cabinets are usually metal or composite materials, making them difficult to break into. Information on approved cabinets is available from Public Works and Government Services Canada.



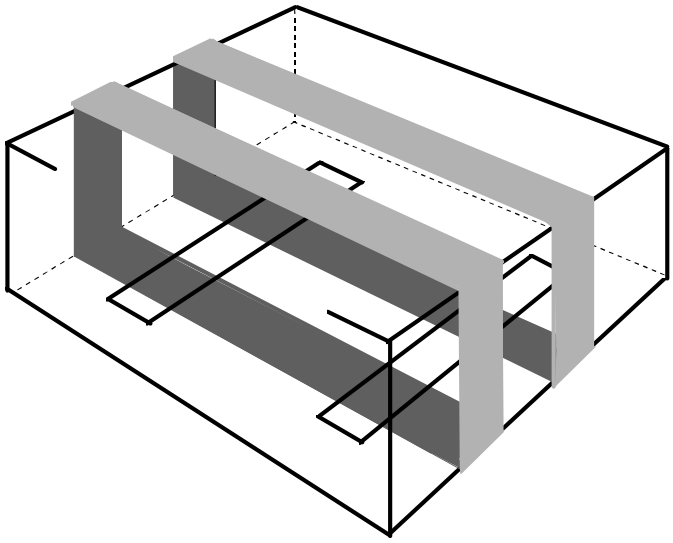
Alarms are installed either inside or outside each CPU unit. The alarms do not prevent the theft of computer equipment but they usually act as a deterrent. In addition, people in the vicinity or at a central location are alerted by a loud piercing sound if the equipment is moved or if the alarm is tampered with.



Anchoring pads and cables are used to anchor devices to desks and tabletops, using high-strength adhesive pads or cables. Once the pad is installed on the table or desk, it is very difficult to remove, and the adhesive usually ruins the finish. Cables are probably the most common physical securing devices, and the least expensive. Steel cables are passed through metal rings that are attached to the equipment and a desk or table. Although cables prevent anyone from quickly walking away with a piece of equipment, they can be cut. Another anchoring method is the use of steel locking plates and cables to secure a variety of computer components and office equipment to desks or tables. The bottom plate is either bolted to the desk or fastened with adhesive. The top and bottom plates slide together and are secured with a high-security lock.



Secure lid locks help prevent intrusion into PC servers and routers and protect microprocessors and memory chips. The metal construction is crushproof, with no adhesive or cables to damage the equipment.



Secure drive locks prevent the introduction of external viruses to PCs and networks, avert the removal of sensitive corporate files by unauthorized individuals, deter the introduction of unauthorized software to PCs and networks and prevent booting from the floppy drive.

Security software uses anti-theft retrieval encryption stealth technology to locate stolen computers. Upon a customer's report of computer theft, the company initiates its tracking feature. As soon as the stolen computer is connected to a telephone line, the software turns off the modem's speaker and silently dials the company's tracking line, giving the PC's current location. The company then informs law enforcement officials, who can obtain a search warrant and retrieve the computer.

2.13.3 Strategies to Minimize Computer Theft

Computer theft cannot be eliminated, but can be reduced by implementing a few simple strategies.

2.13.3.0 APPOINTMENT OF SECURITY PERSONNEL

Departments must appoint a departmental security officer (DSO). The DSO should have direct access to the deputy head to report probable security breaches and illegal acts, as warranted and in accordance with the DSO's mandate. The DSO is responsible for developing, implementing, maintaining, coordinating and monitoring a departmental security program.

2.13.3.1 MASTER KEY SYSTEM

An appropriate master key system must be developed, and comply with the following guidelines:

- All perimeter doors should be keyed alike and not placed on the master key system.
- Restricted access areas should be keyed differently and not placed on the master key system.
- All utility rooms should be keyed alike, in groups.

2.13.3.2 TARGET HARDENING

Minimizing Vulnerabilities Through Target Hardening

Target hardening creates an environment, which makes it difficult for the aggressor to reach a target. The goal of target hardening is to prevent a successful attack through the use of barriers to reduce the adversary's speed of progress, leading to the adversary either giving up the idea of an attack, or taking enough time that a response force can react.

Examples of Enhanced Target Hardening Safeguards

- Increase the number of barriers.
- Increase penetration delay time by strengthening barriers, e.g., doors. The adversary loses speed moving from one barrier to the next due to the weight of the equipment necessary for penetration.
- Increase the time needed to reach an asset, to augment the chances of detection and response. To get full delay time from any barrier, a detection device must detect suspicious activity at first contact with the barrier, rather than after it has been breached.
- Compartmentalize facilities to develop progressively restrictive zones. Every facility should have a reception area where visitors wait for service or permission to proceed to a more restricted area.
- Control circulation of persons and equipment by having all individuals and materials enter through two distinct control points; one for employees and visitors and the other for service vehicles and trade personnel.
- Physically separate zones with a wall extending from the true floor to the true ceiling, including a door equipped with an approved auxiliary deadbolt for use during silent hours.
- Ensure elevators open in a public reception area. Uncontrolled opening of an elevator on a floor is permissible if access to the floor is continuously monitored or if the floor is secure at all times. After business hours, elevators should be controlled by the entry control desk. To further enhance security, elevators should not stop on floors unless persons have been granted access by the entry control desk, or have a key, card or other access device.

2.13.4 PERSONNEL RECOGNITION SYSTEM

2.13.4.0 MINIMIZING VULNERABILITIES THROUGH PERSONNEL RECOGNITION

A personnel recognition system is based on the visual identification of individuals known to authorized personnel or control staff. This system depends solely on personal knowledge of the individuals having access to a particular facility or zone. For this system to be effective, it is necessary to comply with the following guidelines:

- For ease of recognition, the number of employees should not exceed 100 per shift, unless the personnel recognition system has dedicated control staff, i.e., the same guard works the day shift from Monday to Friday.
- There must not be a high turnover of control staff.
- The control staff must recognize all the personnel they will be required to identify prior to assuming control functions.
- The control staff must be advised immediately upon resignation or termination of an employee, to prevent former employees from entering at any time except under escort.

- Identification cards must be available for presentation, if necessary.

Examples of Personnel Recognition System Safeguards

- Issue an identification (ID) card to all employees. An ID card should contain the individual's photograph, name and signature, the name of the issuing department, a card number and an expiry date. The individual's screening level can also be displayed, if desired, unless a Threat and Risk Assessment (TRA) recommends otherwise.
- Issue a building pass or access badge to employees who require regular access to restricted areas, indicating their authorization to enter specific zones.
- Allow for additional processes to verify identity, where warranted.

Procedures for ID Card or Authorization Badge Use

Departments using ID cards or authorization badges must develop procedures for their use, including:

- Establishing a log for the issuance and recovery of both identification cards and access badges, in which is recorded the date of issue, the identity of the bearer, the number of the card or badge, reliability level of the bearer, expiry date and the recovery date of the card or badge;
- Establishing a process for verifying the authenticity of cards or badges held by personnel;
- Providing guidelines for the withdrawal of either cards or badges for cause;
- Indicating how to report improper use, damage, loss or theft of cards or badges;
- Ensuring retrieval of employee cards or badges upon termination of employment;
- Ensuring all blank inserts and equipment necessary for issuing cards and badges are physically protected. The protection should be at a level equal to that of the classified or designated information and assets to which they will indicate authorized access; and
- Ensuring the destruction of all expired or damaged cards and badges.

2.13.5 SECURITY AWARENESS PROGRAM

2.13.5.0 POLICY REQUIREMENTS

The Security Policy of the Government of Canada (GSP) requires that departments implement a security awareness program for all personnel, to define their security responsibilities. Security awareness training is an essential element of a comprehensive and effective security program. Such training is a continuing series of activities, with two overall objectives:

- Keep staff aware of their responsibilities and role in implementing and maintaining security within the department; and
- Obtain and maintain the commitment of staff to those responsibilities and actions. To be effective, security awareness training must be continually reinforced, through the use of periodical newsletters, bulletins and lectures to all personnel.

Without the full cooperation of management, the security awareness program will not succeed and the employees will not cooperate. In these times of restraint, the

security staff needs the cooperation of all employees. Managers must get involved and show leadership to enhance awareness in their departments. Building badges distinguish employees from visitors, contractors or trade persons, and have shown good results in reducing crime. When building badges were implemented during the Gulf War and every government employee was required to wear an ID badge or a building badge, computer theft was almost non-existent. Once the Gulf War ended, some government departments discontinued the use of badges. Had the badge process been continued, theft in the federal government would have been kept to a minimum. It should be impressed upon staff at all levels that security is part of their every day duties, and not an option or someone else's job.

2.13.5.1 SECURITY AWARENESS SAFEGUARDS

- Inform management and all employees, new and old, of the operations of the building during working and silent hours.
- Instruct employees to alert security staff whenever they notice unescorted strangers or visitors without identification badges around their area.
- Lock up laptops at all times when not in use, during coffee breaks, at lunch time and even when at home, because of the value of the asset and of the information they contain.

2.13.6 Conclusion

Computer theft cannot be eliminated, but departments can greatly reduce it by following these simple rules:

- Implement an identification system for employees, visitors and trade persons,
- Provide adequate security for the facility and ensure that barriers exist for the protection of computers, through the use of physical security devices, electronic intrusion detection or a security-cleared guard force,
- Implement a security awareness program that suits the department, and
- Inform employees they will be held responsible for government assets lost or stolen because of carelessness.

Although there are no simple solutions, computer theft can be controlled in a cost-effective manner through a team effort from everyone in the workplace — ministers, directors, managers and all employees.

2.14 Physical and Environmental Security

Physical and environmental security controls are implemented to protect the facility housing system resources, the system resources themselves, and the facilities used to support their operation.

The term *physical and environmental security*, as used in this chapter, refers to measures taken to protect systems, buildings, and related supporting infrastructure against threats associated with their physical environment. Physical and environmental security controls include the following three broad areas:

1. The physical facility is usually the building, other structure, or vehicle housing the system and network components. Systems can be characterized, based upon their operating location, as static, mobile, or portable. Static systems are installed in structures at fixed locations. Mobile systems are installed in vehicles that perform the function of a structure, but not at a fixed location. Portable systems are not installed in fixed operating locations. They may be operated in wide variety of locations, including buildings or vehicles, or in the open. The physical characteristics of these structures and vehicles determine the level of such physical threats as fire, roof leaks, or unauthorized access.

2. The facility's general geographic operating location determines the characteristics of *natural threats*, which include earthquakes and flooding; *man-made threats* such as burglary, civil disorders, or interception of transmissions and emanations; and *damaging nearby activities*, including toxic chemical spills, explosions, fires, and electromagnetic interference from emitters, such as radars.
3. Supporting facilities are those services (both technical and human) that underpin the operation of the system. The system's operation usually depends on supporting facilities such as electric power, heating and air conditioning, and telecommunications. The failure or substandard performance of these facilities may interrupt operation of the system and may cause physical damage to system hardware or stored data.

This section first discusses the benefits of physical security measures, and then presents an overview of common physical and environmental security controls. Physical and environmental security measures result in many benefits, such as protecting employees. This chapter focuses on the protection of computer systems from the following:

- *Interruptions in Providing Computer Services*. An external threat may interrupt the scheduled operation of a system. The magnitude of the losses depends on the duration and timing of the service interruption and the characteristics of the operations end users perform.
- *Physical Damage*. If a system's hardware is damaged or destroyed, it usually has to be repaired or replaced. Data may be destroyed as an act of sabotage by a physical attack on data storage media (e.g., rendering the data unreadable or only partly readable). If data stored by a system for operational use is destroyed or corrupted, the data needs to be restored from back-up copies or from the original sources before the system can be used. The magnitude of loss from physical damage depends on the cost to repair or replace the damaged hardware *and* data, as well as costs arising from service interruptions.
- *Unauthorized Disclosure of Information*. The physical characteristics of the facility housing a system may permit an intruder to gain access both to media external to system hardware (such as diskettes, tapes and printouts) and to media within system components (such as fixed disks), transmission lines or display screens. All may result in loss of disclosure-sensitive information.
- *Loss of Control over System Integrity*. If an intruder gains access to the central processing unit, it is usually possible to reboot the system and *bypass* logical access controls. This can lead to information disclosure, fraud, replacement of system and application software, introduction of a Trojan horse, and more.

Life Safety

It is important to understand that the objectives of physical access controls may be in conflict with those of *life safety*. Simply stated, life safety focuses on providing easy exit from a facility, particularly in an emergency, while physical security strives to control entry. In general, life safety must be given first consideration, but it is usually possible to achieve an effective balance between the two goals. For example, it is often possible to equip emergency exit doors with a time delay. When one pushes on the panic bar, a loud alarm sounds, and the door is released after a brief delay. The expectation is that people will be deterred from using such exits improperly, but will not be significantly endangered during an emergency evacuation.

There are many types of physical access controls, including badges, memory cards, guards, keys, true-floor-to-true-ceiling wall construction, fences, and locks.

Moreover, if such access is gained, it may be very difficult to determine what has been modified, lost, or corrupted.

- *Physical Theft*. System hardware may be stolen. The magnitude of the loss is determined by the costs to replace the stolen hardware and restore data stored on stolen media. Theft may also result in service interruptions.

This section discusses seven major areas of physical and environmental security controls:

- physical access controls,
- fire safety,
- supporting utilities,
- structural collapse,
- plumbing leaks,
- interception of data, and
- mobile and portable systems.

2.14.0 Physical Access Controls

Physical access controls restrict the entry and exit of personnel (and often equipment and media) from an area, such as an office building, suite, data center, or room containing a LAN server.

The controls over physical access to the elements of a system can include controlled areas, barriers that isolate each area, entry points in the barriers, and screening measures at each of the entry points. In addition, staff members who work in a restricted area serve an important role in providing physical security, as they can be trained to challenge people they do not recognize.

Physical access controls should address not only the area containing system hardware, but also locations of wiring used to connect elements of the system, the electric power service, the air conditioning and heating plant, telephone and data lines, backup media and source documents, and any other elements required system's operation. This means that all the areas in the building(s) that contain system elements must be identified.

It is also important to review the effectiveness of physical access controls in each area, both during normal business hours, and at other times particularly when an area may be unoccupied. Effectiveness depends on both the characteristics of the control devices used (e.g., keycard-controlled doors) and the implementation and operation. Statements to the effect that "only authorized persons may enter this area" are not particularly effective. Organizations should determine whether intruders can easily defeat the controls, the extent to which strangers are challenged, and the effectiveness of other control procedures. Factors like these modify the effectiveness of physical controls.

The feasibility of surreptitious entry also needs to be considered. For example, it may be possible to go over the top of a partition that stops at the underside of a suspended ceiling or to cut a hole in a plasterboard partition in a location hidden by furniture. If a door is controlled by a combination lock, it may be possible to observe an authorized person entering the lock combination. If keycards are not carefully controlled, an intruder may be able to steal a card left on a desk or use a card passed back by an accomplice.

Corrective actions can address any of the factors listed above. Adding an additional barrier reduces the risk to the areas behind the barrier. Enhancing the screening at an entry point can reduce the number of penetrations. For example, a guard may

provide a higher level of screening than a keycard-controlled door, or an anti-passback feature can be added. Reorganizing traffic patterns, work flow, and work areas may reduce the number of people who need access to a restricted area. Physical modifications to barriers can reduce the vulnerability to surreptitious entry. Intrusion detectors, such as closed-circuit television cameras, motion detectors, and other devices, can detect intruders in unoccupied spaces.

2.14.1 Fire Safety Factors

Building fires are a particularly important security threat because of the potential for complete destruction of both hardware and data, the risk to human life, and the pervasiveness of the damage. Smoke, corrosive gases, and high humidity from a localized fire can damage systems throughout an entire building. Consequently, it is important to evaluate the fire safety of buildings that house systems. Following are important factors in determining the risks from fire.

- ***Ignition Sources***. Fires begin because something supplies enough heat to cause other materials to burn. Typical ignition sources are failures of electric devices and wiring, carelessly discarded cigarettes, improper storage of materials subject to spontaneous combustion, improper operation of heating devices, and, of course, arson.
- ***Fuel Sources***. If a fire is to grow, it must have a supply of fuel, material that will burn to support its growth, and an adequate supply of oxygen. Once a fire becomes established, it depends on the combustible materials in the building (referred to as the fire load) to support its further growth. The more fuel per square meter, the more intense the fire will be.
- ***Building Operation***. If a building is well maintained and operated so as to minimize the accumulation of fuel (such as maintaining the integrity of fire barriers), the fire risk will be minimized.
- ***Building Occupancy***. Some occupancies are inherently more dangerous than others because of an above-average number of potential ignition sources. For example, a chemical warehouse may contain an above-average fuel load.

Types of Building Construction

There are four basic kinds of building construction:
(a)light frame, (b) heavy timber, (c) incombustible, and (d) fire resistant.

Note that the term *fireproof* is not used because no structure can resist a fire indefinitely. Most houses are light frame, and cannot survive more than about thirty minutes in a fire. Heavy timber means that the basic structural elements have a minimum thickness of four inches. When such structures burn, the char that forms tends to insulate the interior of the timber and the structure may survive for an hour or more depending on the details.

Incombustible means that the structure members will not burn. This almost always means that the members are steel. Note, however, that steel loses it strength at high temperatures, at which point the structure collapses. Fire resistant means that the structural members are incombustible and are insulated. Typically, the insulation is either concrete that encases steel members, or is a mineral wool that is sprayed onto the members. Of course, the heavier the insulation, the longer the structure will resist a fire. Note that a building constructed of reinforced concrete can still be destroyed in a fire if there is sufficient fuel present and fire fighting is ineffective. The prolonged heat of a fire can cause differential expansion of the concrete which causes *spalling*. Portions of the concrete split off, exposing the reinforcing, and the interior of the concrete is subject to additional spalling. Furthermore, as heated floor slabs expand outward, they deform supporting columns. Thus, a reinforced concrete parking garage with open exterior walls and a relatively low fire load has a low fire risk, but a similar archival record storage facility with closed exterior walls and a high fire load has a higher risk even though the basic building material is incombustible.

- ***Fire Detection***. The more quickly a fire is detected, all other things being equal, the more easily it can be extinguished, minimizing damage. It is also important to accurately pinpoint the location of the fire.
- ***Fire Extinguishment***. A fire will burn until it consumes all of the fuel in the building or until it is extinguished. Fire extinguishment may be automatic, as with an automatic sprinkler system or a HALON discharge system, or it may be performed by people using portable extinguishers, cooling the fire site with a stream of water, by limiting the supply of oxygen with a blanket of foam or powder, or by breaking the combustion chemical reaction chain.

When properly installed, maintained, and provided with an adequate supply of water, automatic sprinkler systems are highly effective in protecting buildings and their contents. Nonetheless, one often hears uninformed persons speak of the *water damage* done by sprinkler systems as a disadvantage. *Fires that trigger sprinkler systems* cause the water damage. In short, sprinkler systems reduce fire damage, protect the lives of building occupants, and limit the fire damage to the building itself. All these factors contribute to more rapid recovery of systems following a fire.

Each of these factors is important when estimating the occurrence rate of fires and the amount of damage that will result. The objective of a fire-safety program is to optimize these factors to minimize the risk of fire.

2.14.2 Failure of Supporting Utilities

Systems and the people who operate them need to have a reasonably well controlled operating environment. Consequently, failures of heating and air-conditioning systems will usually cause a service interruption and may damage hardware. These utilities are composed of many elements, each of which must function properly.

For example, the typical air-conditioning system consists of:

1. air handlers that cool and humidify room air,
2. circulating pumps that send chilled water to the air handlers,
3. chillers that extract heat from the water, and
4. cooling towers that discharge the heat to the outside air.

Each of these elements has a mean-time-between-failures (MTBF) and a mean-time-to-repair (MTTR). Using the MTBF and MTTR values for each of the elements of a system, one can estimate the occurrence rate of system failures and the range of resulting service interruptions.

This same line of reasoning applies to electric power distribution, heating plants, water, sewage, and other utilities required for system operation or staff comfort. By identifying the failure modes of each utility and estimating the MTBF and MTTR, necessary failure threat parameters can be developed to calculate the resulting risk. The risk of utility failure can be reduced by substituting units with lower MTBF values. MTTR can be reduced by stocking spare parts on site and training maintenance personnel. And the outages resulting from a given MTBF can be reduced by installing redundant units under the assumption that failures are distributed randomly in time. Each of these strategies can be evaluated by comparing the reduction in risk with the cost to achieve it.

2.14.3 Structural Collapse

A building may be subjected to a load greater than it can support. Most commonly this is a result of an earthquake, a snow load on the roof beyond design criteria, an explosion that displaces or cuts structural members, or a fire that weakens structural members. Even if the structure is not completely demolished, the authorities may decide to ban its further use, sometimes even banning entry to remove materials. This threat applies primarily to high-rise buildings and those with large interior spaces without supporting columns.

2.14.4 Plumbing Leaks

While plumbing leaks do not occur every day, they can be seriously disruptive. The building's plumbing drawings can help locate plumbing lines that might endanger system hardware. These lines include hot and cold water, chilled water supply and return lines, steam lines, automatic sprinkler lines, fire hose standpipes, and drains. If a building includes a laboratory or manufacturing spaces, there may be other lines that conduct water, corrosive or toxic chemicals, or gases.

As a rule, analysis often shows that the cost to relocate threatening lines is difficult to justify. However, the location of shutoff valves and procedures that should be followed in the event of a failure must be specified. Operating and security personnel should have this information immediately available for use in an emergency. In some cases, it may be possible to relocate system hardware, particularly distributed LAN hardware.

2.14.5 Interception of Data

Depending on the type of data a system processes, there may be a significant risk if the data is intercepted. There are three routes of data interception: direct observation, interception of data transmission, and electromagnetic interception.

- *Direct Observation*. System terminal and workstation display screens may be observed by unauthorized persons. In most cases, it is relatively easy to relocate the display to eliminate the exposure.
- *Interception of Data Transmissions*. If an interceptor can gain access to data transmission lines, it may be feasible to tap into the lines and read the data being transmitted. Network monitoring tools can be used to capture data packets. Of course, the interceptor cannot control what is transmitted, and so may not be able to immediately observe data of interest. However, over a period of time there may be a serious level of disclosure. Local area networks typically broadcast messages. Consequently, all traffic, including passwords, could be retrieved. Interceptors could also transmit spurious data on tapped lines, either for purposes of disruption or for fraud.
- *Electromagnetic Interception*. Systems routinely radiate electromagnetic energy that can be detected with special-purpose radio receivers. Successful interception will depend on the signal strength at the receiver location; the greater the separation between the system and the receiver, the lower the success rate. TEMPEST shielding, of either equipment or rooms, can be used to minimize the spread of electromagnetic signals. The signal-to-noise ratio at the receiver, determined in part by the number of competing emitters will also affect the success rate. The more workstations of the same type in the same location performing "random" activity, the more difficult it is to intercept a given workstation's radiation. On the other hand, the trend toward wireless (i.e., deliberate radiation) LAN connections may increase the likelihood of successful interception.

2.14.6 Mobile and Portable Systems

The analysis and management of risk usually has to be modified if a system is installed in a vehicle or is portable, such as a laptop computer. The system in a vehicle will share the risks of the vehicle, including accidents and theft, as well as regional and local risks.

Encryption of data files on stored media may also be a cost-effective precaution against disclosure of confidential information if a laptop computer is lost or stolen.

Portable and mobile systems share an increased risk of theft and physical damage. In addition, portable systems can be "misplaced" or left unattended by careless users. Secure storage of laptop computers is often required when they are not in use. If a mobile or portable system uses particularly valuable or important data, it may be appropriate to either store its data on a medium that can be removed from the system when it is unattended or to encrypt the data. In any case, the issue of how custody of mobile and portable computers are to be controlled should be addressed. Depending on the sensitivity of the system and its application, it may be appropriate to require briefings of users and signed briefing acknowledgments.

2.14.7 Approach to Implementation

Like other security measures, physical and environmental security controls are selected because they are cost-beneficial. This does not mean that a user must conduct a detailed cost-benefit analysis for the selection of every control. There are four general ways to justify the selection of controls:

1. ***They are required by law or regulation***. Fire exit doors with panic bars and exit lights are examples of security measures required by law or regulation. Presumably, the regulatory authority has considered the costs and benefits and has determined that it is in the public interest to require the security measure. A lawfully conducted organization has no option but to implement all required security measures.
2. ***The cost is insignificant, but the benefit is material***. A good example of this is a facility with a key-locked low-traffic door to a restricted access. The cost of keeping the door locked is minimal, but there is a significant benefit. Once a significant benefit/minimal cost security measure has been identified, no further analysis is required to justify its implementation.
3. ***The security measure addresses a potentially "fatal" security exposure but has a reasonable cost*.** Backing up system software and data is an example of this justification. For most systems, the cost of making regular backup copies is modest (compared to the costs of operating the system), the organization would not be able to function if the stored data were lost, and the cost impact of the failure would be material. In such cases, it would not be necessary to develop any further cost justification for the backup of software and data. However, this justification depends on what constitutes a *modest* cost, and it does not identify the optimum backup schedule. Broadly speaking, a cost that does not require budgeting of additional funds would qualify.
4. ***The security measure is estimated to be cost-beneficial***. If the cost of a potential security measure is significant, and it cannot be justified by any of the first three reasons listed above, then its cost (both implementation and ongoing operation) and its benefit (reduction in future expected losses) need to be analyzed to determine if it is cost-beneficial. In this context, *cost-beneficial*

means that the reduction in expected loss is significantly greater than the cost of implementing the security measure.

Arriving at the fourth justification requires a detailed analysis. Simple rules of thumb do not apply. Consider, for example, the threat of electric power failure and the security measures that can protect against such an event. The threat parameters, rate of occurrence, and range of outage durations depend on the location of the system, the details of its connection to the local electric power utility, the details of the internal power distribution system, and the character of other activities in the building that use electric power. The system's potential losses from service interruption depends on the details of the functions it performs. Two systems that are otherwise identical can support functions that have quite different degrees of urgency. Thus, two systems may have the same electric power failure threat and vulnerability parameters, yet entirely different loss potential parameters.

Furthermore, a number of different security measures are available to address electric power failures. These measures differ in both cost and performance. For example, the cost of an uninterruptible power supply (UPS) depends on the size of the electric load it can support, the number of minutes it can support the load, and the speed with which it assumes the load when the primary power source fails. An on-site power generator could also be installed either in place of a UPS (accepting the fact that a power failure will cause a brief service interruption) or in order to provide long-term backup to a UPS system. Design decisions include the magnitude of the load the generator will support, the size of the on-site fuel supply, and the details of the facilities to switch the load from the primary source or the UPS to the on-site generator.

This example shows systems with a wide range of risks and a wide range of available security measures (including, of course, no action), each with its own cost factors and performance parameters.

2.14.8 Interdependencies

Physical and environmental security measures rely on and support the proper functioning of many of the other areas discussed in this handbook. Among the most important are the following:

- *Logical Access Controls*. Physical security controls augment technical means for controlling access to information and processing. Even if the most advanced and best-implemented logical access controls are in place, if physical security measures are inadequate, logical access controls may be circumvented by directly accessing the hardware and storage media. For example, a computer system may be rebooted using different software.
- *Contingency Planning*. A large portion of the contingency planning process involves the failure of physical and environmental controls. Having sound controls, therefore, can help minimize losses from such contingencies.
- *Identification and Authentication* (I&A). Many physical access control systems require that people be identified and authenticated. Automated physical security access controls can use the same types of I&A as other computer systems. In addition, it is possible to use the same tokens (e.g., badges) as those used for other computer-based I&A.
- *Other.* Physical and environmental controls are also closely linked to the activities of the local guard force, fire house, life safety office, and medical office. These organizations should be consulted for their expertise in planning controls for the systems environment.

2.14.9 Cost Considerations

Costs associated with physical security measures range greatly. Useful generalizations about costs, therefore, are difficult make. Some measures, such as keeping a door locked, may be a trivial expense. Other features, such as fire-detection and -suppression systems, can be far more costly. Cost considerations should include operation. For example, adding controlled-entry doors requires persons using the door to stop and unlock it. Locks also require physical key management and accounting (and rekeying when keys are lost or stolen). Often these effects will be inconsequential, but they should be fully considered. As with other security measures, the objective is to select those that are cost-beneficial.

2.15 Class C2: Controlled Access Protection –An Introduction

There has been a fair amount of confusion about what is meant by "C2 compatibility". Windows NT has, for example, been tested and rated C2 compliant, but only under a very specific set of circumstances. However, simply because a system or system component is "C2 compliant" doesn't mean that it may be considered completely secure under all conditions. Unfortunately, the "C2" label has come to be a catch-all designation appearing to encompass many security features which, in fact, it does not.

Using the above example as a starting point, Windows NT workstations are only C2 compliant when they are not connected to a multi-user system (network) of any kind and when they have their A: drives disconnected at the hardware level. There are a few other restrictions that, if violated, negate the C2 compliance.

The below discussion describes what is meant by C2 compliance. This is paraphrased from the Department of Defense "Orange Book", which is the authoritative document, and the National Computer Security Center's "Red Book" which is the official network interpretation of the Orange book. Systems in this class make 'users individually accountable for their actions through login procedures, auditing of security-relevant events and resource isolation.' There are only four top level criteria for C2 systems:

1. Security Policy
2. Accountability
3. Assurance (operational and life cycle)
4. Documentation

2.15.0 C2 Criteria Simplified

Security Policy: Discretionary Access Control - The system defines and controls access between named users and named objects. The enforcement mechanism (self/group/public controls, access control lists, etc.) allows users to specify and control sharing of these objects by named individuals, groups or both and provides controls to limit propagation of access rights. Objects must be protected from unauthorized access either by explicit user action or default. The controls are capable of including or excluding access to the granularity of a single user.

Security Policy: Object Reuse - All authorizations to a storage object must be revoked prior to initial assignment, allocation or reallocation. No information including encrypted representations of information is available to any user that obtains access to an object that has been released back to the system.

Accountability: Identification and Authentication - All users must identify themselves before performing any other actions that the system is expected to mediate. Some protected mechanism such as passwords must be used to authenticate the user's identity. The system must protect authentication data so that it cannot be accessed by any unauthorized user. The system must be able to enforce individual accountability by uniquely identifying each user and providing the capability of associating the identity with all auditable actions taken by the individual.

Accountability: Audit - The system must be able to create, maintain and protect from modification or unauthorized access or destruction an audit trail of access to the objects it protects. It must record the following types of events: use of identification and authentication mechanism, introduction of objects into a user's address space, deletions of objects, actions taken by system administrators and security officers, and other security-relevant events. For each recorded event the system must record the date and time, user, type of event and success or failure of the event. If the event is the introduction of an object into the user's address space (such as file open or program execution) the name of the object must be included.

Operational Assurance: System Architecture - The system must maintain a domain for its own execution that protects it from external interference or tampering such as modification of its code or data structures. The system resources to be protected must be isolated and subject to access control and auditing requirements.

Operational Assurance: System Integrity - Hardware and/or software features must be provided to validate the correct operation of the system resources.

Life Cycle Assurance: Security Testing - The security mechanisms of the system must be tested and found to work as presented in system documentation. Testing must insure that there are no obvious ways for an unauthorized user to bypass or otherwise defeat the security protection mechanisms of the system. This must include a search for obvious flaws that would allow a violation of resource isolation or that would permit unauthorized access to the audit or authentication data.

Documentation: Security Features User's Guide - The protection mechanisms provided by the system, their use and how they interact with each other must be provided.

Documentation: Trusted Facility Manual - A manual addressed to the system administrator must present cautions about functions and privileges that should be controlled when running a secure facility. The procedures for examining and maintaining audit files as well as the detailed audit record structure for each type of audit event must be given.

Documentation: Test Documentation -The system developer must provide a document that describes the test plan and procedures that shows how the security mechanisms were tested and the results of the testing.

Documentation: Design Documentation - A document must be provided that describes the developer's philosophy of protection and how the philosophy was implemented in the system. If the system is modular, the interfaces between the modules must be described.

2.15.1 The Red Book

By extension, the Red Book applies the C2 standard criteria to the network environment. The Red Book's purpose is to interpret the Orange Book's standalone

standards as the may be applied to a network. The Red Book places responsibility for the security of the network not on a manufacturer, but on the network's sponsor. This differentiation acknowledges that a device, regardless of its Orange Book rating as delivered from the manufacturer, becomes a component of the multivendor network when it is interconnected. As such, it takes on some of the characteristics of the network, interacts with other components of the network and may have its own characteristics altered by these interactions. The manufacturer no longer controls the security mechanisms of the device for these reasons. The Red Book places the system-wide responsibility with the network sponsor.

Policy - The Red Book defines two types of policy within the network environment: Secrecy and Data Integrity. The document defines them as follows:

Secrecy Policy: The network sponsor shall define the form of the discretionary secrecy policy that is enforced in the network to prevent unauthorized users from reading the sensitive information entrusted to the network.

Data Integrity Policy: The network sponsor shall define the discretionary integrity policy to prevent unauthorized users from modifying, viz., writing, sensitive information. The definition of data integrity presented by the network sponsor refers to the requirement that the information has not been subjected to unauthorized modification in the
network.

Accountability: Identification and Authentication - The requirement for identification and authentication of users is the same for a network system as for an ADP system. The identification and authentication may be done by the component to which the user is directly connected or some other component, such as an identification and authentication server.

Accountability: Audit - This criterion applies as stated in the Orange Book.

Assurance: Architecture -The system architecture criterion must be met individually by all NTCB (Network Trusted Computing Base) partitions. Implementation of the requirement that the NTCB maintain a domain for its own execution is achieved by having each NTCB partition maintain a domain for its own execution.

Assurance: Integrity - Implementation of the requirement is partly achieved by having hardware and/or software features that can be used to periodically validate the correct operation of the hardware and firmware elements of each component's NTCB partition. Features shall also be provided to validate the identity and correct operation of a component prior to its incorporation in the network system and throughout system operation.

Assurance: Security Testing - The testing of a security mechanism of the network system for meeting this criterion shall be an integrated testing procedure involving all components containing an NTCB partition that implement the given mechanism. This integrated testing is additional to any individual component tests involved in the evaluation of the network system. The sponsor should identify the allowable set of configurations including the sizes of the networks.

Documentation - This user documentation describes user visible protection mechanisms at the global (network system) level and at the user interface of each component, and the interaction among these. The Trusted Facility Manual contains

specifications and procedures to assist the system administrator(s) maintain cognizance of
the network configuration. These specifications and procedures address the following:

1. The hardware configuration of the network itself;
2. The implications of attaching new components to the network;
3. The case where certain components may periodically leave the network (e.g., by crashing, or by being disconnected) and then rejoin;
4. Network configuration aspects that can impact the security of the network system; (For example, the manual should describe for the network system administrator the interconnections among components that are consistent with the overall network system architecture.)
5. Loading or modifying NTCB software or firmware (e.g., down-line loading). The physical and administrative environmental controls must be specified. Any assumptions about security of a given network should be clearly stated (e.g., the fact that all communications links must be physically protected to a certain level). As in the Orange Book criteria, there must also be a documented test plan to ensure that the system meets its published standards. However, in the case of a network, the "system developer" is interpreted as "the network system sponsor". The description of the test plan should establish the context in which the testing was or should be conducted. The description should identify any additional test components that are not part of the system being evaluated. This includes a description of the test-relevant functions of such test components and a description of the interfacing of those test components to the system being evaluated. The description of the test plan should also demonstrate that the tests adequately cover the network security policy.

2.15.2 Summary

The application of C2 standards in both the standalone and network environment has specific implications. However, to extend those implications, either beyond the standard or into an altered environment requires some form of supporting guidelines. It is not enough to say, for example, that a device is "C2 certified" and assume that, however it is employed, it will continue to meet C2 (or, even, its own published) specifications.

Finally, vendors often represent products as "C2 compliant" implying that the product has undergone and passed the rigid C2 testing process employed by the NCSC. In most cases this is misleading. Very few PC-based systems have undergone and completed such testing. In fact, there are no current (July, 1996) network systems that are officially rated C2.

When a vendor represents a device as C2 compliant, it usually means that the device was designed to C2 standards, not that it has been tested and certified. While this may certainly be adequate for many implementations, it is important to understand that 1) C2 certification probably does not exist for that product, and 2) the ability of the product to maintain its C2-type architecture may change materially as its network environment changes.

C2 compliance, per se, is usually not a requirement for commercial systems. In fact, there are a number of "standards" that purport to be taking the place of C2 for the commercial world. These new criteria, such as "Extended C2" and "Commercial C2" are, generally, simply extensions of the original C2 standard. We do not recommend slavish adherence to C2 as a method of securing today's commercial networks. Rather, we believe that the principles upon which the C2 standard is built offer an

excellent measuring stick for the over-all security of the corporate computing environment.

However, as many security and audit professionals point out, the architecture of the system is only the beginning. It is at least as important to ensure that the policies, standards and practices which the C2 environment enforces are current and appropriate. The system administrators must be well-trained and empowered to do their jobs properly. There must be periodic risk assessments and formal audits to ensure compliance with policies. Finally, there must be a firm system of enforcement, both at the system and administrative levels.

Good security is not a single layer of protection. It consists of proper policies, standards and practices, adequate architecture, compliance testing and auditing, and appropriate administration. Most important, good information security requires awareness at all levels of the organization and solid, visible support from the highest management. Only when these other criteria are met will the application of C2 principles to the computing system be effective.

Section References

2.1 Fraser, B. ed. *RFC 2196. Site Security Handbook.* Network Working Group, September 1997.
Chapter 2.

2.2 Guideline for the Analysis Local Area Network Security.,
Federal Information Processing Standards Publication 191,
November 1994. Chapter 2.

2.3 NIST. *An Introduction to Security: The NIST Handbook, Special Publication 800-12*. US Dept. of Commerce. Chapter 5.

Howe, D. "Information System Security Engineering: Cornerstone to the Future." *Proceedings of the 15th National Computer Security Conference*. Baltimore, MD, Vol. 1, October 15, 1992. pp.244-251.

Fites, P., and M. Kratz. "Policy Development." *Information Systems Security: A Practitioner's Reference*. New York, NY: Van Nostrand Reinhold, 1993. pp. 411-427.

Lobel, J. "Establishing a System Security Policy." *Foiling the System Breakers*. New York, NY:McGraw-Hill, 1986. pp. 57-95.

Menkus, B. "Concerns in Computer Security." *Computers and Security*. 11(3), 1992. pp.211-215.

Office of Technology Assessment. "Federal Policy Issues and Options." *Defending Secrets, Sharing Data: New Locks for Electronic Information*. Washington, DC: U.S Congress, Office of Technology Assessment, 1987. pp. 151-160.

Office of Technology Assessment. "Major Trends in Policy Development." *Defending Secrets, Sharing Data: New Locks and Keys for Electronic Information*. Washington, DC: U.S. Congress,

Office of Technology Assessment, 1987. p. 131-148.
O'Neill, M., and F. Henninge, Jr. "Understanding ADP System and Network Security Considerations and Risk Analysis." *ISSA Access*. 5(4), 1992. pp. 14-17.

Peltier, Thomas. "Designing Information Security Policies That Get Results." *Infosecurity News*.4(2), 1993. pp. 30-31.

President's Council on Management Improvement and the President's Council on Integrity and Efficiency. *Model Framework for Management Control Over Automated Information System*.

Washington, DC: President's Council on Management Improvement, January 1988. Smith, J. "Privacy Policies and Practices: Inside the Organizational Maze." *Communications of the ACM*. 36(12), 1993. pp. 104-120.
Sterne, D. F. "On the Buzzword `Computer Security Policy.'" In *Proceedings of the 1991 IEEE Symposium on Security and Privacy*, Oakland, CA: May 1991. pp. 219-230.

Wood, Charles Cresson. "Designing Corporate Information Security Policies." *DATAPRO Reports on Information Security*, April 1992.

2.4 Guideline for the Analysis Local Area Network Security.,
Federal Information Processing Standards Publication 191,
November 1994. Chapter 2.2.

[MART89] Martin, James, and K. K. Chapman, The Arben Group, Inc.; Local Area Networks, Architectures and Implementations, Prentice Hall, 1989.

[BARK89] Barkley, John F., and K. Olsen; Introduction to Heterogenous Computing Environments, NIST Special Publication 500-176, November, 1989.

[NCSC87] A Guide to Understanding Discretionary Access Control in Trusted Systems, NCSC-TG-003, Version 1, September 30, 1987

[NCSL90] National Computer Systems Laboratory (NCSL) Bulletin, Data Encryption Standard, June, 1990.

[SMID88] Smid, Miles, E. Barker, D. Balenson, and M. Haykin; Message Authentication Code (MAC) Validation System: Requirements and Procedures, NIST Special Publication 500-156, May, 1988.

[OLDE92] Oldehoeft, Arthur E.; Foundations of a Security Policy for Use of the National Research and Educational Network, NIST Interagency Report, NISTIR 4734, February 1992.

[COMM91] U.S. Department of Commerce Information Technology Management Handbook, Attachment 13-D: Malicious Software Policy and Guidelines, November 8, 1991.

[WACK89] Wack, John P., and L. Carnahan; Computer Viruses and Related Threats: A Management Guide, NIST Special Publication 500-166, August 1989.

[X9F292] Information Security Guideline for Financial Institutions, X9/TG-5, Accredited Committee X9F2, March 1992.

[BJUL93] National Computer Systems Laboratory (NCSL) Bulletin, Connecting to the Internet: Security Considerations, July 1993.

[BNOV91] National Computer Systems Laboratory (NCSL) Bulletin, Advanced Authentication Technology, November 1991.

[KLEIN] Daniel V. Klein, "Foiling the Cracker: A Survey of, and Improvements to, Password Security", Software Engineering Institute. (This work was sponsored in part by the Department of Defense.)

[GILB89] Gilbert, Irene; Guide for Selecting Automated Risk Analysis Tools, NIST Special Publication 500-174, October, 1989.

[KATZ92] Katzke, Stuart W. ,Phd., "A Framework for Computer Security Risk Management", NIST, October, 1992.

[NCSC85] Department of Defense Password Management Guideline, National Computer Security Center, April, 1985.

[NIST85] Federal Information Processing Standard (FIPS PUB) 112, Password Usage, May, 1985.

[ROBA91] Roback Edward, NIST Coordinator, Glossary of Computer Security Terminology, NISTIR 4659, September, 1991.

[TODD89] Todd, Mary Anne and Constance Guitian, Computer Security Training Guidelines,NIST Special Publication 500-172, November, 1989.

[STIE85] Steinauer, Dennis D.; Security of Personal Computer Systems: A Management Guide, NBS Special Publication 500-120, January, 1985.

[WACK91] Wack, John P.; Establishing a Computer Security Incident Response Capability (CSIRC), NIST Special Publication 800-3, November, 1991.

[NIST74] Federal Information Processing Standard (FIPS PUB) 31, Guidelines for Automatic Data Processing Physical Security and Risk Management, June, 1974.

2.5 Fraser, B. ed. *RFC 2196. Site Security Handbook.* Network Working Group, September 1997. Chapter 3.

2.6. Fraser, B. ed. *RFC 2196. Site Security Handbook.* Network Working Group, September 1997. Chapter 4.6.

2.7 Fraser, B. ed. *RFC 2196. Site Security Handbook.* Network Working Group, September 1997. Chapter 5.

2.8 Fraser, B. ed. *RFC 2196. Site Security Handbook.* Network Working Group, September 1997. Chapter 4.5.4

2.9 Hancock, William M. *Dial-Up MODEM Protection Schemes: A Case Study in Secure Dial-Up Implementation.* Network-1 Software and Technology, Inc.1995.

2.10 Innovative Security Products. *Security White Paper Series: Securing Your Companies Network*. Prairie Village, KS, 1998.

2.11 Innovative Security Products. *Security White Paper Series: Microcomputer Security*. Prairie Village, KS, 1998.

2.12 Fraser, B. ed. *RFC 2196. Site Security Handbook.* Network Working Group, September 1997. Chapter 4.5.

2.13 Royal Canadian Mounted Police Technical Operations Directorate. Information Technology Security Branch. *Guide to Minimizing Computer Theft.* Security Information Publications June 1997

2.14 NIST. *An Introduction to Security: The NIST Handbook, Special Publication 800-12*. US Dept. of Commerce. Chapter 15.

Alexander, M., ed. "Secure Your Computers and Lock Your Doors." *Infosecurity News*. 4(6),1993. pp. 80-85.

Archer, R. "Testing: Following Strict Criteria." *Security Dealer*. 15(5), 1993. pp. 32-35.

Breese, H., ed. *The Handbook of Property Conservation*. Norwood, MA: Factory Mutual Engineering Corp.

Chanaud, R. "Keeping Conversations Confidential." *Security Management*. 37(3), 1993.pp. 43-48.

Miehl, F. "The Ins and Outs of Door Locks." *Security Management*. 37(2), 1993. pp. 48-53.

National Bureau of Standards. *Guidelines for ADP Physical Security and Risk Management*.

Federal Information Processing Standard Publication 31. June 1974.

Peterson, P. "Infosecurity and Shrinking Media." *ISSA Access*. 5(2), 1992. pp. 19-22.
Roenne, G. "Devising a Strategy Keyed to Locks." *Security Management*. 38(4), 1994.pp. 55-56.

Zimmerman, J. "Using Smart Cards - A Smart Move." *Security Management*. 36(1), 1992. pp. 32-36.

2.15 Stephenson, Peter. *CLASS C2: CONTROLLED ACCESS PROTECTION - A Simplified Description.* Sanda International Corp. 1997

3.1 Introduction

For most systems, identification and authentication (I&A) is the first line of defense. I&A is a technical measure that prevents unauthorized people (or unauthorized processes) from entering a computer system.

I&A is a critical building block of computer security since it is the basis for most types of access control and for establishing user accountability. Access control often requires that the system be able to identify and differentiate among users. For example, access control is often based on *least privilege*, which refers to the granting to users of only those accesses required to perform their duties. User accountability requires the linking of activities on a computer system to specific individuals and, therefore, requires the system to identify users.

- ***Identification*** is the means by which a user *provides* a claimed identity to the system.
- ***Authentication*** is the means of establishing the *validity* of this claim.

A typical user identification could be JSMITH (for Jane Smith). This information can be known by system administrators and other system users. A typical user authentication could be Jane Smith's password, which is kept secret. This way system administrators can set up Jane's access and see her activity on the audit trail, and system users can send her e-mail, but no one can pretend to be Jane.

Computer systems recognize people based on the authentication data the systems *receive.* Authentication presents several challenges: collecting authentication data, transmitting the data securely, and knowing whether the person who was originally authenticated is *still* the person using the computer system. For example, a user may walk away from a terminal while still logged on, and another person may start using it. There are three means of authenticating a user's identity *which can be used alone or in combination*:

For most applications, trade-offs will have to be made among security, ease of use, and ease of administration, especially in modern networked environments.

- something the individual *knows* (a secret e.g., a password, Personal Identification Number (PIN), or cryptographic key)
- something the individual *possesses* (a token e.g., an ATM card or a smart card)
- and something the individual *is* (a biometric e.g., such characteristics as a voice pattern, handwriting dynamics, or a fingerprint).

While it may appear that any of these means could provide strong authentication, there are problems associated with each. If people wanted to pretend to be someone else on a computer system, they can guess or learn that individual's password; they can also steal or fabricate tokens. Each method also has drawbacks for legitimate users and system administrators: users forget passwords and may lose tokens, and administrative overhead for keeping track of I&A data and tokens can be substantial. Biometric systems have significant technical, user acceptance, and cost problems as well.

This section explains current I&A technologies and their benefits and drawbacks as they relate to the three means of authentication. Although some of the technologies make use of cryptography because it can significantly strengthen authentication.

3.1.0 I&A Based on Something the User Knows

The most common form of I&A is a user ID coupled with a password. This technique is based solely on something the user knows. There are other techniques besides *conventional* passwords that are based on knowledge, such as knowledge of a cryptographic key.

3.1.0.1 PASSWORDS

In general, password systems work by requiring the user to enter a user ID and password (or passphrase or personal identification number). The system compares the password to a previously stored password for that user ID. If there is a match, the user is authenticated and granted access.

Benefits of Passwords. Passwords have been successfully providing security for computer systems for a long time. They are integrated into many operating systems, and users and system administrators are familiar with them. When properly managed in a controlled environment, they can provide effective security.

Problems With Passwords. The security of a password system is dependent upon keeping passwords secret. Unfortunately, there are many ways that the secret may be divulged. All of the problems discussed below can be significantly mitigated by improving password security, as discussed in the sidebar. However, there is no fix for the problem of electronic monitoring, except to use more advanced authentication (e.g., based on cryptographic techniques or tokens).

1. *Guessing or finding passwords*. If users select their own passwords, they tend to make them easy to remember. That often makes them easy to guess. The names of people's children, pets, or favorite sports teams are common examples. On the other hand, assigned passwords may be difficult to remember, so users are more likely to write them down. Many computer systems are shipped with administrative accounts that have preset passwords. Because these passwords are standard, they are easily "guessed." Although security practitioners have been warning about this problem for years, many system administrators still do not change default passwords. Another method of learning passwords is to observe someone entering a password or PIN. The observation can be done by someone in the same room or by someone some distance away using binoculars. This is often referred to as *shoulder surfing*.

Improving Password Security

Password generators. If users are not allowed to generate their own passwords, they cannot pick easy-to-guess passwords. Some generators create only pronounceable nonwords to help users remember them. However, users tend to write down hard-to-remember passwords.

Limits on log-in attempts. Many operating systems can be configured to lock a user ID after a set number of failed log-in attempts. This helps to prevent guessing of passwords.

Password attributes. Users can be instructed, or the system can force them, to select passwords (1) with a certain minimum length, (2) with special characters, (3) that are unrelated to their user ID, or (4) to pick passwords which are not in an on-line dictionary. This makes passwords more difficult to guess (but more likely to be written down).

Changing passwords. Periodic changing of passwords can reduce the damage done by stolen passwords and can make brute-force attempts to break into systems more difficult. Too frequent changes, however, can be irritating to users.

Technical protection of the password file. Access control and one-way encryption can be used to protect the password file itself.

Note: Many of these techniques are discussed in FIPS 112, *Password Usage* and FIPS 181, *Automated Password Generator*.

2. *Giving passwords away*. Users may share their passwords. They may give their password to a co-worker in order to share files. In addition, people can be tricked into divulging their passwords. This process is referred to as *social engineering*.
3. *Electronic monitoring*. When passwords are transmitted to a computer system, they can be electronically monitored. This can happen on the network used to transmit the password or on the computer system itself. Simple encryption of a password that will be used again does not solve this problem because encrypting the same password will create the same ciphertext; the ciphertext becomes the password.
4. *Accessing the password file*. If the password file is not protected by strong access controls, the file can be downloaded. Password files are often protected with one-way encryption so that plain-text passwords are not available to system administrators or hackers (if they successfully bypass access controls). Even if the file is encrypted, brute force can be used to learn passwords if the file is downloaded (e.g., by encrypting English words and comparing them to the file).

Passwords Used as Access Control. Some mainframe operating systems and many PC applications use passwords as a means of restricting access to specific resources within a system. Instead of using mechanisms such as access control lists, access is granted by entering a password. The result is a proliferation of passwords that can reduce the overall security of a system. While the use of passwords as a means of access control is common, it is an approach that is often less than optimal and not cost-effective.

3.1.0.2 CRYPTOGRAPHIC KEYS

Although the authentication derived from the knowledge of a cryptographic key may be based entirely on something the user knows, it is necessary for the user to also possess (or have access to) something that can perform the cryptographic computations, such as a PC or a smart card. For this reason, the protocols used are discussed in the Smart Tokens section of this chapter. However, it is possible to implement these types of protocols without using a smart token. Additional discussion is also provided under the Single Log-in section.

3.1.1 I&A Based on Something the User Possesses

Although some techniques are based solely on something the user possesses, most of the techniques described in this section are combined with something the user knows. This combination can provide significantly stronger security than either something the user knows or possesses alone. Objects that a user possesses for the purpose of I&A are called *tokens*. This section divides tokens into two categories: *memory tokens* and *smart tokens*.

3.1.1.0 MEMORY TOKENS

Memory tokens store, but do not process, information. Special reader/writer devices control the writing and reading of data to and from the tokens. The most common type of memory token is a magnetic striped card, in which a thin stripe of magnetic material is affixed to the surface of a card (e.g., as on the back of credit cards). A common application of memory tokens for authentication to computer systems is the automatic teller machine (ATM) card. This uses a combination of something the user possesses (the card) with something the user knows (the PIN). Some computer systems authentication technologies are based solely on possession of a token, but

they are less common. Token-only systems are more likely to be used in other applications, such as for physical access.

Benefits of Memory Token Systems. Memory tokens when used with PINs provide significantly more security than passwords. In addition, memory cards are inexpensive to produce. For a hacker or other would-be masquerader to pretend to be someone else, the hacker must have both a valid token *and* the corresponding PIN. This is much more difficult than obtaining a valid password and user ID combination (especially since most user IDs are common knowledge).

Another benefit of tokens is that they can be used in support of log generation without the need for the employee to key in a user ID for each transaction or other logged event since the token can be scanned repeatedly. If the token is required for physical entry and exit, then people will be forced to remove the token when they leave the computer. This can help maintain authentication.

Problems With Memory Token Systems. Although sophisticated technical attacks are possible against memory token systems, most of the problems associated with them relate to their cost, administration, token loss, user dissatisfaction, and the compromise of PINs. Most of the techniques for increasing the security of memory token systems relate to the protection of PINs. Many of the techniques discussed in the sidebar on Improving Password Security apply to PINs.

Attacks on memory-card systems have sometimes been quite creative. One group stole an ATM machine that they installed at a local shopping mall. The machine collected valid account numbers and corresponding PINs, which the thieves used to forge cards. The forged cards were then used to withdraw money from legitimate ATMs.

1. *Requires special reader*. The need for a special reader increases the cost of using memory tokens. The readers used for memory tokens must include both the physical unit that reads the card and a processor that determines whether the card and/or the PIN entered with the card is valid. If the PIN or token is validated by a processor that is not physically located with the reader, then the authentication data is vulnerable to electronic monitoring (although cryptography can be used to solve this problem).
2. *Token loss*. A lost token may prevent the user from being able to log in until a replacement is provided. This can increase administrative overhead costs. The lost token could be found by someone who wants to break into the system, or could be stolen or forged. If the token is also used with a PIN, any of the methods described above in password problems can be used to obtain the PIN. Common methods are finding the PIN taped to the card or observing the PIN being entered by the legitimate user. In addition, any information stored on the magnetic stripe that has not been encrypted can be read.
3. *User Dissatisfaction*. In general, users want computers to be easy to use. Many users find it inconvenient to carry and present a token. However, their dissatisfaction may be reduced if they see the need for increased security.

3.1.1.1 SMART TOKENS

A smart token expands the functionality of a memory token by incorporating one or more integrated circuits into the token itself. When used for authentication, a smart token is another example of authentication based on something a user possesses (i.e., the token itself). A smart token typically requires a user also to provide something the user knows (i.e., a PIN or password) in order to "unlock" the smart token for use.

There are many different types of smart tokens. In general, smart tokens can be divided three different ways based on physical characteristics, interface, and protocols used. These three divisions are not mutually exclusive.

- *Physical Characteristics*. Smart tokens can be divided into two groups: smart cards and other types of tokens. A smart card looks like a credit card, but incorporates an embedded microprocessor. Smart cards are defined by an International Standards Organization (ISO) standard. Smart tokens that are not smart cards can look like calculators, keys, or other small portable objects.
- *Interface*. Smart tokens have either a manual or an electronic interface. Manual or human interface tokens have displays and/or keypads to allow humans to communicate with the card. Smart tokens with electronic interfaces must be read by special reader/writers. Smart cards, described above, have an electronic interface. Smart tokens that look like calculators usually have a manual interface.
- *Protocol*. There are many possible protocols a smart token can use for authentication. In general, they can be divided into three categories: static password exchange, dynamic password generators, and challenge-response.
- *Static* tokens work similarly to memory tokens, except that the users authenticate themselves *to the token* and then the token authenticates the user to the computer.
- A token that uses a *dynamic password generator* protocol creates a unique value, for example, an eight-digit number, that changes periodically (e.g., every minute). If the token has a manual interface, the user simply reads the current value and then types it into the computer system for authentication. If the token has an electronic interface, the transfer is done automatically. If the correct value is provided, the log-in is permitted, and the user is granted access to the system.
- Tokens that use a *challenge-response* protocol work by having the computer generate a challenge, such as a random string of numbers. The smart token then generates a response based on the challenge. This is sent back to the computer, which authenticates the user based on the response. The challenge-response protocol is based on cryptography. Challenge-response tokens can use either electronic or manual interfaces.

There are other types of protocols, some more sophisticated and some less so. The three types described above are the most common.

Benefits of Smart Tokens

Smart tokens offer great flexibility and can be used to solve many authentication problems. The benefits of smart tokens vary, depending on the type used. In general, they provide greater security than memory cards. Smart tokens can solve the problem of electronic monitoring even if the authentication is done across an open network by using *one-time passwords*.

1. *One-time passwords*. Smart tokens that use either dynamic password generation or challenge-response protocols can create one-time passwords. Electronic monitoring is not a problem with one-time passwords because each time the user is authenticated to the computer, a different "password" is used. (A hacker could learn the one-time password through electronic monitoring, but would be of no value.)
2. *Reduced risk of forgery*. Generally, the memory on a smart token is not readable unless the PIN is entered. In addition, the tokens are more complex and, therefore, more difficult to forge.

3. *Multi-application*. Smart tokens with electronic interfaces, such as smart cards, provide a way for users to access many computers using many networks with only one log-in. This is further discussed in the Single Log-in section of this chapter. In addition, a single smart card can be used for multiple functions, such as physical access or as a debit card.

Problems with Smart Tokens

Like memory tokens, most of the problems associated with smart tokens relate to their cost, the administration of the system, and user dissatisfaction. Smart tokens are generally less vulnerable to the compromise of PINs because authentication usually takes place on the card. (It is possible, of course, for someone to watch a PIN being entered and steal that card.) Smart tokens cost more than memory cards because they are more complex, particularly challenge-response calculators.

Electronic reader/writers can take many forms, such as a slot in a PC or a separate external device. Most human interfaces consist of a keypad and display.

1. *Need reader/writers or human intervention*. Smart tokens can use either an electronic or a human interface. An electronic interface requires a reader, which creates additional expense. Human interfaces require more actions from the user. This is especially true for challenge-response tokens with a manual interface, which require the user to type the challenge into the smart token and the response into the computer. This can increase user dissatisfaction.
2. *Substantial Administration*. Smart tokens, like passwords and memory tokens, require strong administration. For tokens that use cryptography, this includes key management.

3.1.2 I&A Based on Something the User Is

Biometric authentication technologies use the unique characteristics (or attributes) of an individual to authenticate that person's identity. These include physiological attributes (such as fingerprints, hand geometry, or retina patterns) or behavioral attributes (such as voice patterns and hand-written signatures). Biometric authentication technologies based upon these attributes have been developed for computer log-in applications.

Biometric authentication generally operates in the following manner:
Before any authentication attempts, a user is "enrolled" by creating a reference profile (or template) based on the desired physical attribute. The resulting template is associated with the identity of the user and stored for later use.
When attempting authentication, the user's biometric attribute is measured. The previously stored reference profile of the biometric attribute is compared with the measured profile of the attribute taken from the user. The result of the comparison is then used to either accept or reject the user.

Biometric authentication is technically complex and expensive, and user acceptance can be difficult. However, advances continue to be made to make the technology more reliable, less costly, and more user-friendly. Biometric systems can provide an increased level of security for computer systems, but the technology is still less mature than that of memory tokens or smart tokens. Imperfections in biometric authentication devices arise from technical difficulties in measuring and profiling physical attributes as well as from the somewhat variable nature of physical attributes. These may change, depending on various conditions. For example, a

person's speech pattern may change under stressful conditions or when suffering from a sore throat or cold.

Due to their relatively high cost, biometric systems are typically used with other authentication means in environments requiring high security.

3.1.3 Implementing I&A Systems

Some of the important implementation issues for I&A systems include administration, maintaining authentication, and single log-in.

3.1.3.0 ADMINISTRATION

Administration of authentication data is a critical element for all types of authentication systems. The administrative overhead associated with I&A can be significant. I&A systems need to create, distribute, and store authentication data. For passwords, this includes creating passwords, issuing them to users, and maintaining a password file. Token systems involve the creation and distribution of tokens/PINs and data that tell the computer how to recognize valid tokens/PINs.

> One method of looking for improperly used accounts is for the computer to inform users when they last logged on. This allows users to check if someone else used their account.

For biometric systems, this includes creating and storing profiles. The administrative tasks of creating and distributing authentication data and tokens can be a substantial. Identification data has to be kept current by adding new users and deleting former users. If the distribution of passwords or tokens is not controlled, system administrators will not know if they have been given to someone other than the legitimate user. It is critical that the distribution system ensure that authentication data is firmly linked with a given individual.

In addition, I&A administrative tasks should address lost or stolen passwords or tokens. It is often necessary to monitor systems to look for stolen or shared accounts.

Authentication data needs to be stored securely, as discussed with regard to accessing password files. The value of authentication data lies in the data's confidentiality, integrity, and availability. If confidentiality is compromised, someone may be able to use the information to masquerade as a legitimate user. If system administrators can read the authentication file, they can masquerade as another user. Many systems use encryption to hide the authentication data from the system administrators. If integrity is compromised, authentication data can be added or the system can be disrupted. If availability is compromised, the system cannot authenticate users, and the users may not be able to work.

3.1.3.1 MAINTAINING AUTHENTICATION

So far, this chapter has discussed initial authentication only. It is also possible for someone to use a legitimate user's account after log-in. Many computer systems handle this problem by logging a user out or locking their display or session after a certain period of inactivity. However, these methods can affect productivity and can make the computer less user-friendly.

3.1.3.2 SINGLE LOG-IN

From an efficiency viewpoint, it is desirable for users to authenticate themselves only once and then to be able to access a wide variety of applications and data available on local and remote systems, even if those systems require users to authenticate themselves. This is known as *single log-in*. If the access is within the same host computer, then the use of a modern access control system (such as an access control list) should allow for a single log-in. If the access is across multiple platforms, then the issue is more complicated, as discussed below. There are three main techniques that can provide single log-in across multiple computers: host-to-host authentication, authentication servers, and user-to-host authentication.

- *Host-to-Host Authentication.* Under a host-to-host authentication approach, users authenticate themselves once to a host computer. That computer then authenticates itself to other computers and vouches for the specific user. Host-to-host authentication can be done by passing an identification, a password, or by a challenge-response mechanism or other one-time password scheme. Under this approach, it is necessary for the computers to recognize each other and to trust each other.
- *Authentication Servers.* When using authentication server, the users authenticate themselves to a special host computer (the authentication server). This computer then authenticates the user to other host computers the user wants to access. Under this approach, it is necessary for the computers to trust the authentication server. (The authentication server need not be a separate computer, although in some environments this may be a cost-effective way to increase the security of the server.) Authentication servers can be distributed geographically or logically, as needed, to reduce workload.
- *User-to-Host.* A user-to-host authentication approach requires the user to log-in to each host computer. However, a smart token (such as a smart card) can contain all authentication data and perform that service for the user. To users, it looks as though they were only authenticated once.

> Kerberos and SPX are examples of network authentication server protocols. They both use cryptography to authenticate users to computers on networks.

3.1.3.3 INTERDEPENDENCIES

There are many interdependencies among I&A and other controls. Several of them have been discussed in the section.

- *Logical Access Controls*. Access controls are needed to protect the authentication database. I&A is often the basis for access controls. Dial-back modems and firewalls, can help prevent hackers from trying to log-in.
- *Audit*. I&A is necessary if an audit log is going to be used for individual accountability.
- *Cryptography*. Cryptography provides two basic services to I&A: it protects the confidentiality of authentication data, and it provides protocols for proving knowledge and/or possession of a token without having to transmit data that could be replayed to gain access to a computer system.

3.1.3.4 COST CONSIDERATIONS

In general, passwords are the least expensive authentication technique and generally the least secure. They are already embedded in many systems. Memory tokens are less expensive than smart tokens, but have less functionality. Smart tokens with a human

interface do not require readers, but are more inconvenient to use. Biometrics tend to be the most expensive.

For I&A systems, the cost of administration is often underestimated. Just because a system comes with a password system does not mean that using it is free. For example, there is significant overhead to administering the I&A system.

3.1.4 Authentication

Identification is the means by which a user *provides* a claimed identity to the system. The most common form of identification is the user ID. In this section of the plan, describe how the major application identifies access to the system. Note: the explanation provided below is an excerpt from NIST Special Publication, *Generally Accepted Principles and Practices for Securing Information Technology Systems.*

Authentication is the means of establishing the *validity* of this claim. There are three means of authenticating a user's identity *which can be used alone or in combination*: something the individual *knows* (a secret -- e.g., a password, Personal Identification Number (PIN), or cryptographic key); something the individual *possesses* (a token -- e.g., an ATM card or a smart card); and something the individual *is* (a biometrics -- e.g., characteristics such as a voice pattern, handwriting dynamics, or a fingerprint).

In this section, describe the major application's authentication control mechanisms. Below is a list of items that should be considered in the description:

- Describe the method of user authentication (password, token, and biometrics).
- If a password system is used, provide the following specific information:
- Allowable character set,
- Password length (minimum, maximum),
- Password aging time frames and enforcement approach,
- Number of generations of expired passwords disallowed for use,
- Procedures for password changes,
- Procedures for handling lost passwords, and
- Procedures for handling password compromise.
- Procedures for training users and the materials covered.

Note: The recommended minimum number of characters in a password is six to eight characters in a combination of alpha, numeric, or special characters.

- Indicate the frequency of password changes, describe how password changes are enforced (e.g., by the software or System Administrator), and identify who changes the passwords (the user, the system, or the System Administrator).
- Describe any biometrics controls used. Include a description of how the biometrics controls are implemented on the system.
- Describe any token controls used on the system and how they are implemented. Are special hardware readers required?
- Are users required to use a unique Personal Identification Number (PIN)?
- Who selects the PIN, the user or System Administrator?
- Does the token use a password generator to create a one-time password?

- Is a challenge-response protocol used to create a one-time password?
- Describe the level of enforcement of the access control mechanism (network, operating system, and application).

- Describe how the access control mechanism supports individual accountability and audit trails (e.g., passwords are associated with a user identifier that is assigned to a single individual).
- Describe the self-protection techniques for the user authentication mechanism (e.g., passwords are stored with one-way encryption to prevent anyone [including the System Administrator] from reading the clear-text passwords, passwords are automatically generated, passwords are checked against a dictionary of disallowed passwords, passwords are encrypted while in transmission).
- State the number of invalid access attempts that may occur for a given user identifier or access location (terminal or port) and describe the actions taken when that limit is exceeded.
- Describe the procedures for verifying that all system-provided administrative default passwords have been changed.
- Describe the procedures for limiting access scripts with embedded passwords (e.g., scripts with embedded passwords are prohibited, scripts with embedded passwords are only allowed for batch applications).
- Describe any policies that provide for bypassing user authentication requirements, single-sign-on technologies (e.g., host-to-host, authentication servers, user-to-host identifier, and group user identifiers) and any compensating controls.
    - If digital signatures are used, the technology must conforms with FIPS 186, *(Digital Signature Standard)* and FIPS 180, *(Secure Hash Standard)* issued by NIST, unless a waiver has been granted. Describe any use of digital or electronic signatures. Address the following specific issues:State the digital signature standards used. If the standards used are not NIST standards, please state the date the waiver was granted and the name and title of the official granting the waiver.

    - Describe the use of electronic signatures and the security control provided.

    - Discuss cryptographic key management procedures for key generation, distribution, storage, entry, use, destruction and archiving.

For many years, the prescribed method for authenticating users has been through the use of standard, reusable passwords. Originally, these passwords were used by users at terminals to authenticate themselves to a central computer. At the time, there were no networks (internally or externally), so the risk of disclosure of the clear text password was minimal. Today, systems are connected together through local networks, and these local networks are further connected together and to the Internet. Users are logging in from all over the globe; their reusable passwords are often transmitted across those same networks in clear text, ripe for anyone in-between to capture. And indeed, the CERT* Coordination Center and other response teams are seeing a tremendous number of incidents involving packet sniffers which are capturing the clear text passwords.

With the advent of newer technologies like one-time passwords (e.g., S/Key), PGP, and token-based authentication devices, people are using password-like strings as secret tokens and pins. If these secret tokens and pins are not properly selected and protected, the authentication will be easily subverted.

3.1.4.0 ONE-TIME PASSWORDS

As mentioned above, given today's networked environments, it is recommended that sites concerned about the security and integrity of their systems and networks consider moving away from standard, reusable passwords. There have been many incidents involving Trojan network programs (e.g., telnet and rlogin) and network packet sniffing programs. These programs capture clear text hostname/account name/password triplets. Intruders can use the captured information for subsequent access to those hosts and accounts. This is possible because:

- the password is used over and over (hence the term "reusable"), and
- the password passes across the network in clear text.

Several authentication techniques have been developed that address this problem. Among these techniques are challenge-response technologies that provide passwords that are only used once (commonly called one-time passwords). There are a number of products available that sites should consider using. The decision to use a product is the responsibility of each organization, and each organization should perform its own evaluation and selection.

3.1.4.1 KERBEROS

Kerberos is a distributed network security system, which provides for authentication across unsecured networks. If requested by the application, integrity and encryption can also be provided. Kerberos was originally developed at the Massachusetts Institute of Technology (MIT) in the mid 1980s. There are two major releases of Kerberos, version 4 and 5, which are for practical purposes, incompatible.

Kerberos relies on a symmetric key database using a key distribution center (KDC) which is known as the Kerberos server. A user or service (known as "principals") are granted electronic "tickets" after properly communicating with the KDC. These tickets are used for authentication between principals. All tickets include a time stamp, which limits the time period for which the ticket is valid. Therefore, Kerberos clients and server must have a secure time source, and be able to keep time accurately.

The practical side of Kerberos is its integration with the application level. Typical applications like FTP, telnet, POP, and NFS have been integrated with the Kerberos system. There are a variety of implementations which have varying levels of integration. Please see the Kerberos FAQ available at http://www.ov.com/misc/krb-faq.html for the latest information.

3.1.4.2 CHOOSING AND PROTECTING SECRET TOKENS AND PINS

When selecting secret tokens, take care to choose them carefully. Like the selection of passwords, they should be robust against brute force efforts to guess them. That is, they should not be single words in any language, any common, industry, or cultural acronyms, etc. Ideally, they will be longer rather than shorter and consist of pass phrases that combine upper and lower case character, digits, and other characters.

Once chosen, the protection of these secret tokens is very important. Some are used as pins to hardware devices (like token cards) and these should not be written down or placed in the same location as the device with which they are associated. Others, such as a secret Pretty Good Privacy (PGP) key, should be protected from unauthorized access.

One final word on this subject. When using cryptography products, like PGP, take care to determine the proper key length and ensure that your users are trained to do likewise. As technology advances, the minimum safe key length continues to grow. Make sure your site keeps up with the latest knowledge on the technology so that you can ensure that any cryptography in use is providing the protection you believe it is.

3.1.4.3 PASSWORD ASSURANCE

While the need to eliminate the use of standard, reusable passwords cannot be overstated, it is recognized that some organizations may still be using them. While it's recommended that these organizations transition to the use of better technology, in the mean time, we have the following advice to help with the selection and maintenance of traditional passwords. But remember, none of these measures provides protection against disclosure due to sniffer programs.

1. The importance of robust passwords - In many (if not most) cases of system penetration, the intruder needs to gain access to an account on the system. One way that goal is typically accomplished is through guessing the password of a legitimate user. This is often accomplished by running an automated password cracking program, which utilizes a very large dictionary, against the system's password file. The only way to guard against passwords being disclosed in this manner is through the careful selection of passwords which cannot be easily guessed (i.e., combinations of numbers, letters, and punctuation characters). Passwords should also be as long as the system supports and users can tolerate.

2. Changing default passwords - Many operating systems and application programs are installed with default accounts and passwords. These must be changed immediately to something that cannot be guessed or cracked.

3. Restricting access to the password file - In particular, a site wants to protect the encrypted password portion of the file so that would-be intruders don't have them available for cracking. One effective technique is to use shadow passwords where the password field of the standard file contains a dummy or false password. The file containing the legitimate passwords are protected elsewhere on the system.

4. Password aging - When and how to expire passwords is still a subject of controversy among the security community. It is generally accepted that a password should not be maintained once an account is no longer in use, but it is hotly debated whether a user should be forced to change a good password that's in active use. The arguments for changing passwords relate to the prevention of the continued use of penetrated accounts. However, the opposition claims that frequent password changes lead to users writing down their passwords in visible areas (such as pasting them to a terminal), or to users selecting very simple passwords that are easy to guess. It should also be stated that an intruder will probably use a captured or guessed password sooner rather than later, in which case password aging provides little if any protection.

While there is no definitive answer to this dilemma, a password policy should directly address the issue and provide guidelines for how often a user should change the password. Certainly, an annual change in their password is usually not difficult for most users, and you should consider requiring it. It is recommended that passwords be changed at least whenever a privileged account is compromised, there is a critical change in personnel (especially if it is an administrator!), or when an account has been compromised. In addition, if a privileged account password is compromised, all passwords on the system should be changed.

5. Password/account blocking - Some sites find it useful to disable accounts after a predefined number of failed attempts to authenticate. If your site decides to employ this mechanism, it is recommended that the mechanism not "advertise" itself. After disabling, even if the correct password is presented, the message displayed should remain that of a failed login attempt. Implementing this mechanism will require that legitimate users contact their system administrator to request that their account be reactivated.

6. A word about the finger daemon - By default, the finger daemon displays considerable system and user information. For example, it can display a list of all users currently using a system, or all the contents of a specific user's .plan file. This information can be used by would-be intruders to identify usernames and guess their passwords. It is recommended that sites consider modifying finger to restrict the information displayed.

3.1.4.4 CONFIDENTIALITY

There will be information assets that your site will want to protect from disclosure to unauthorized entities. Operating systems often have built-in file protection mechanisms that allow an administrator to control who on the system can access, or "see," the contents of a given file. A stronger way to provide confidentiality is through encryption. Encryption is accomplished by scrambling data so that it is very difficult and time consuming for anyone other than the authorized recipients or owners to obtain the plain text. Authorized recipients and the owner of the information will possess the corresponding decryption keys that allow them to easily unscramble the text to a readable (clear text) form. We recommend that sites use encryption to provide confidentiality and protect valuable information.

The use of encryption is sometimes controlled by governmental and site regulations, so we encourage administrators to become informed of laws or policies that regulate its use before employing it. It is outside the scope of this document to discuss the various algorithms and programs available for this purpose, but we do caution against the casual use of the UNIX crypt program as it has been found to be easily broken. We also encourage everyone to take time to understand the strength of the encryption in any given algorithm/product before using it. Most well-known products are well-documented in the literature, so this should be a fairly easy task.

3.1.4.5 INTEGRITY

As an administrator, you will want to make sure that information (e.g., operating system files, company data, etc.) has not been altered in an unauthorized fashion. This means you will want to provide some assurance as to the integrity of the information on your systems. One way to provide this is to produce a checksum of the unaltered file, store that checksum offline, and periodically (or when desired) check to make sure the checksum of the online file hasn't changed (which would indicate the data has been modified).

Some operating systems come with checksumming programs, such as the UNIX sum program. However, these may not provide the protection you actually need. Files can be modified in such a way as to preserve the result of the UNIX sum program! Therefore, we suggest that you use a cryptographically strong program, such as the message digesting program MD5, to produce the checksums you will be using to assure integrity.

There are other applications where integrity will need to be assured, such as when transmitting an email message between two parties. There are products available that can provide this capability. Once you identify that this is a capability you need, you can go about identifying technologies that will provide it.

3.1.4.6 AUTHORIZATION

Authorization refers to the process of granting privileges to processes and, ultimately, users. This differs from authentication in that authentication is the process used to identify a user. Once identified (reliably), the privileges, rights, property, and permissible actions of the user are determined by authorization. Explicitly listing the authorized activities of each user (and user process) with respect to all resources (objects) is impossible in a reasonable system. In a real system certain techniques are used to simplify the process of granting and checking authorization(s).

One approach, popularized in UNIX systems, is to assign to each object three classes of user: owner, group and world. The owner is either the creator of the object or the user assigned as owner by the super-user. The owner permissions (read, write and execute) apply only to the owner. A group is a collection of users, which share access rights to an object. The group permissions (read, write and execute) apply to all users in the group (except the owner). The world refers to everybody else with access to the system. The world permissions (read, write and execute) apply to all users (except the owner and members of the group).

Another approach is to attach to an object a list, which explicitly contains the identity of all, permitted users (or groups). This is an Access Control List (ACL). The advantage of ACLs are that they are easily maintained (one central list per object) and it's very easy to visually check who has access to what. The disadvantages are the extra resources required to store such lists, as well as the vast number of such lists required for large systems.

Section References

3.1 NIST. *An Introduction to Security: The NIST Handbook, Special Publication 800-12*. US Dept. of Commerce. Chapter 16.

Alexander, M., ed. "Keeping the Bad Guys Off-Line." *Infosecurity News.* 4(6), 1993. pp. 54-65.

American Bankers Association. *American National Standard for Financial Institution Sign-On Authentication for Wholesale Financial Transactions*. ANSI X9.26-1990. Washington, DC,February 28, 1990.

CCITT Recommendation X.509. The Directory - Authentication Framework. November 1988

(Developed in collaboration, and technically aligned, with ISO 9594-8).
Department of Defense. Password Management Guideline. CSC-STD-002-85. April 12, 1985.

Feldmeier, David C., and Philip R. Kam. "UNIX Password Security - Ten Years Later." *Crypto'89 Abstracts*. Santa Barbara, CA: Crypto '89 Conference, August 20-24, 1989.

Haykin, Martha E., and Robert B. J. Warnar. *Smart Card Technology: New Methods for Computer Access Control.* Special Publication 500-157. Gaithersburg, MD: National Institute of Standards and Technology, September 1988.

Kay, R. "Whatever Happened to Biometrics?" *Infosecurity News*. 4(5), 1993. pp. 60-62.
National Bureau of Standards. *Password Usage.* Federal Information Processing Standard Publication 112. May 30, 1985.

National Institute of Standards and Technology. *Automated Password Generator*. Federal Information Processing Standard Publication 181. October, 1993.

National Institute of Standards and Technology. *Guideline for the Use of Advanced Authentication Technology Alternatives*. Federal Information Processing Standard Publication

Salamone, S. "Internetwork Security: Unsafe at Any Node?" *Data Communications.* 22(12), 1993. pp. 61-68.

Sherman, R. "Biometric Futures." *Computers and Security*. 11(2), 1992. pp. 128-133.
Smid, Miles, James Dray, and Robert B. J. Warnar. "A Token-Based Access Control System for Computer Networks." *Proceedings of the 12th National Commuter Security Conference*. National Institute of Standards and Technology, October 1989.

Steiner, J.O., C. Neuman, and J. Schiller. "Kerberos: An Authentication Service for Open Network Systems." *Proceedings Winter USENIX*. Dallas, Texas, February 1988. pp. 191-202.

Troy, Eugene F. *Security for Dial-Up Lines*. Special Publication 500-137, Gaithersburg, MD:National Bureau of Standards, May 1986.

NIST Computer Security Resource Clearinghouse Web site URL: http://csrc.nist.gov

Office of Management and Budget. Circular A-130, “Management of Federal Information Resources,” Appendix III, “Security of Federal Automated Information Resources.” 1996.

Public Law 100-235, “Computer Security Act of 1987.”

[Schultz90] Schultz, Eugene. Project Leader, Lawrence Livermore National Laboratory.

CERT Workshop, Pleasanton, CA, 1990.

Swanson, Marianne and Guttman, Barbara . *Generally Accepted Principles and Practices for Securing Information Technology Systems.* Special Publication 800-14. Gaithersburg, MD: National Institute of Standards and Technology, September 1996.

3.1.3 Swanson, Marianne . Guide for Developing Security Plans for Unclassified Systems, Special Publication 800-18. US Dept. of Commerce. Chapter 6 1997

3.1.4 Fraser, B. ed. *RFC 2196. Site Security Handbook.* Network Working Group, September 1997. Chapter 4.1.

4.1 The 7 Processes

4.1.0 Process 1 - Define the Scope and Boundary, and Methodology

This process determines the direction that the risk management effort will take. It defines how much of the LAN (the boundary) and in how much detail (the scope) the risk management process should entail. The boundary will define those parts of the LAN that will be considered. The boundary may include the LAN as a whole or parts of the LAN, such as the data communications function, the server function, the applications, etc. Factors that determine the boundary may be based on LAN ownership, management or control. Placing the boundary around a part of the LAN controlled elsewhere may result in cooperation problems that may lead to inaccurate results. This problem stresses the need for cooperation among those involved with the ownership and management of the different parts of the LAN, as well as the applications and information processed on it.

Figure 4.1 Risk Management Process

1. Define the Scope and Boundary and Methodology
2. Identify and Value Assets,
3. Identify Threats and Determine Likelihood,
4. Measure Risk,
5. Select Appropriate Safeguards,
6. Implement and Test Safeguards,
7. Accept Residual Risk.

The scope of the risk management effort must also be defined. The scope can be thought of as a logical outline showing, within the boundary, the depth of the risk management process. The scope distinguishes the different areas of the LAN (within the boundary) and the different levels of detail used during the risk management process. For example some areas may be considered at a higher or broader level, while other areas may be treated in depth and with a narrow focus.

For smaller LANs, the boundary may be the LAN as a whole, and the scope may define a consistent level of detail throughout the LAN. For larger LANs, an organization may decide to place the boundary around those areas that it controls and to define the scope to consider all areas within the boundary. However the focus on data communications, external connections, and certain applications might be more narrow. Changes in the LAN configuration, the addition of external connections, or updates or upgrades to LAN software or applications may influence the scope.

Figure 4.2 - Simple Asset Valuation

The value of the asset can be represented in terms of the potential loss. This loss can be based on the replacement value, the immediate impact of the loss, and the consequence. One of the simplest valuing techniques to indicate the loss of an asset is to use a qualitative ranking of high, medium and low.
Assigning values to these rankings (3=high, 2=medium, and 1=low) can assist in the risk measure process.

The appropriate risk management methodology for the LAN may have been determined prior to defining the boundary and scope. If the methodology has already been determined, then it may be useful to scrutinize the chosen methodology given the defined boundary and scope. If a methodology has not been chosen, the boundary and scope information may be useful in selecting a methodology that produces the most effective results.

4.1.0.1 Process 2 - Identify and Value Assets

Asset valuation identifies and assigns value to the assets of the LAN. All parts of the LAN have value although some assets are definitely more valuable than others. This step gives the first indication of those areas where focus should be placed. For

LANs that produce large amounts of information that cannot be reasonably analyzed, initial screening may need to be done. Defining and valuing assets may allow the organization to initially decide those areas that can be filtered downward and those areas that should be flagged as a high priority.

Different methods can be used to identify and value assets. The risk methodology that an organization chooses may provide guidance in identifying assets and should provide a technique for valuing assets. Generally assets can be valued based on the impact and consequence to the organization. This would include not only the replacement cost of the asset, but also the effect on the organization if the asset is disclosed, modified, destroyed or misused in any other way.

Because the value of an asset should be based on more than just the replacement cost, valuing assets is one of the most subjective of the processes. However, if asset valuation is done with the goal of the process in mind, that is, to define assets in terms of a hierarchy of importance or criticality, the relativeness of the assets becomes more important than placing the "correct" value on them.

The risk assessment methodology should define the representation of the asset values. Purely quantitative methodologies such as FIPS 65 may use dollar values. However having to place a dollar value on some of the consequences that may occur in today's environments may be sufficient to change the perception of the risk management process from being challenging to being unreasonable.

Many risk assessment methodologies in use today require asset valuation in more qualitative terms. While this type of valuation may be considered more subjective than a quantitative approach, if the scale used to value assets is utilized consistently throughout the risk management process, the results produced should be useful. Figure 4.2 shows one of the simplest methods for valuing assets. Throughout this discussion of the risk management process, a simple technique for valuing assets (as shown in Figure 4.2), determining risk measure, estimating safeguard cost, and determining risk mitigation will be presented. This technique is a simple, yet valid technique; it is being used here to show the relationship between the processes involved in risk management. The technique is not very granular and may not be appropriate for environments where replacement costs, sensitivities of information and consequences vary widely.

Figure 4.3 - Defining the LAN Configuration

Hardware configuration - includes servers, workstations, PCs, peripheral devices, external connections, cabling maps, bridges or gateway connections, etc. Software configuration - includes server operating systems, workstation and PC operating systems, the LAN operating system, major application software, software tools, LAN management tools, and software under development. This should also include the location of the software on the LAN and from where it is commonly accessed.
Data - Includes a meaningful typing of the dataprocessed and communicated through the LAN, as well as the types of users who generally access the data. Indications of where the data is accessed, stored and processed on the LAN is important. Attention to the sensitivity of the data should also be considered.

One of the implicit outcomes of this process is that a detailed configuration of the LAN, as well as its uses is produced. This configuration should indicate the hardware incorporated, major software applications used, significant information processed on the LAN, as well as how that information flows through the LAN. The degree of knowledge of the LAN configuration will depend on the defined boundary and scope. Figure 4.3 exemplifies some of the areas that should be included.

After the LAN configuration is completed, and the assets are determined and valued, the organization should have a reasonably correct view of what the LAN consists of and what areas of the LAN need to be protected.

4.1.0.2 Process 3 - Identify Threats and Determine Likelihood

The outcome of this process should be a strong indication of the adverse actions that could harm the LAN, the likelihood that these actions could occur, and the weaknesses of the LAN that can be exploited to cause the adverse action. To reach this outcome, threats and vulnerabilities need to be identified and the likelihood that a threat will occur needs to be determined.

Large amounts of information on various threats and vulnerabilities exist. The Reference and Further Reading Sections of this document provide some information on LAN threats and vulnerabilities. Some risk management methodologies also provide information on potential threats and vulnerabilities. User experience and LAN management experience also provide insight into threats and vulnerabilities.

The degree to which threats are considered will depend on the defined boundary and scope defined for the risk management process. A high level analysis may point to threats and vulnerabilities in general terms; a more focused analysis may tie a threat to a specific component or usage of the LAN. For example a high level analysis may indicate that the consequence due to loss of data confidentiality through disclosure of information on the LAN is too great a risk. A more narrowly focused analysis may indicate that the consequence due to disclosure of personnel data captured and read through LAN transmission is too great a risk. More than likely, the generality of the threats produced in the high level analysis, will, in the end, produce safeguard recommendations that will also be high level. This is acceptable if the risk assessment
was scoped at a high level. The more narrowly focused assessment will produce a safeguard that can specifically reduce a given risk, such as the disclosure of personnel data.

The threats and vulnerabilities introduced in Section 2 may be used as a starting point, with other sources included where appropriate. New threats and vulnerabilities should be addressed when they are encountered. Any asset of the LAN that was determined to be important enough (i.e., was not filtered through the screening process) should be examined to determine those threats that could potentially harm it. For more focused assessments, particular attention should be paid to detailing the ways that these threats could occur. For example, methods of attack that result in unauthorized access may be from a login session playback, password cracking, the attachment of unauthorized equipment to the LAN, etc. These specifics provide more information in determining LAN vulnerabilities and will provide more information for proposing safeguards.

This process may uncover some vulnerabilities that can be corrected by improving LAN management and operational controls immediately. These improved controls will usually reduce the risk of the threat by some degree, until such time that more thorough improvements are planned and implemented. For example, increasing the length and composition of the password for authentication may be one way to reduce a vulnerability to guessing passwords. Using more robust passwords is a measure that can be quickly implemented to increases the security of the LAN. Concurrently, the planning and implementation of a more advanced authentication mechanism can occur.

Figure 4.4 Assigning Likelihood Measure

The likelihood of the threat occurring can be normalized as a value that ranges from 1 to 3. A 1 will indicate a low likelihood, a 2 will indicate a moderate likelihood and a 3 will indicate a high likelihood.

Existing LAN security controls should be analyzed to determine if they are currently providing adequate protection. These controls may be technical, procedural, etc. If a

control is not providing adequate protection, it can be considered a vulnerability. For example, a LAN operating system may provide access control to the directory level, rather than the file level. For some users, the threat of compromise of information may be too great not to have file level protection. In this example, the lack of granularity in the access control could be considered a vulnerability.

As specific threats and related vulnerabilities are identified, a likelihood measure needs to be associated with the threat/vulnerability pair (i.e. What is the likelihood that a threat will be realized, given that the vulnerability is exploited?). The risk methodology chosen by the organization should provide the technique used to measure likelihood. Along with asset valuation, assigning likelihood measures can also be a subjective process. Threat data for traditional threats (mostly physical threats) does exist and may aid in determining likelihood. However experience regarding the technical aspects of the LAN and knowledge of operational aspects of the organization may prove more valuable to decide likelihood measure. Figure 4.4 defines a simple likelihood measure. This likelihood measure coincides with the asset valuation measure defined in Figure 4.1. Although the asset valuation and the likelihood measures provided in this example appear to be weighted equally for each threat/vulnerability pair, it is a user determination regarding which measure should be emphasized during the risk measurement process.

4.1.0.3 Process 4 - Measure Risk

In its broadest sense the risk measure can be considered the representation of the kinds of adverse actions that may happen to a system or organization and the degree of likelihood that these actions may occur. The outcome of this process should indicate to the organization the degree of risk associated with the defined assets. This outcome is important because it its the basis for making safeguard selection and risk mitigation decisions. There are many ways to measure and represent risk. [KATZ92] points out that depending on the particular methodology or approach, the measure could be defined in qualitative terms, quantitative terms, one dimensional, multidimensional, or some combination of these. The risk measure process should be consistent with (and more than likely defined by) the risk assessment methodology being used by the organization. Quantitative approaches are often associated with measuring risk in terms of dollar losses (e.g. FIPS 65). Qualitative approaches are often associated with measuring risk in terms of quality as indicated through a scale or ranking. One dimensional approaches consider only limited components (e.g. risk = magnitude of loss X frequency of loss). Multidimensional approaches consider additional components in the risk measurement such as reliability, safety, or performance. One of the most important aspects of risk measure is that the representation be understandable and meaningful to those who need to make the safeguard selection and risk mitigation decisions.

Figure 4.5 - One Dimensional Approach to Calculate Risk

The risk associated with a threat can be considered as a function of the relative likelihood that the threat can occur, and the expected loss incurred given that the threat occurred. The risk is calculated as follows:

risk = likelihood of threat occurring (given the specific vulnerability) x loss incurred

The value estimated for loss is determined to be a value that ranges from 1 to 3. Therefor risk may be calculated as a number ranging from 1 to 9 meaning a risk of 1 or 2 is considered a low risk, a risk of 3 or 4 would be a moderate risk, and a risk of 6 or 9 would be considered a high risk.

LIKELIHOOD	LOSS	RISK
1	1	1 - LOW
1	2	2 - LOW
1	3	3 - MODERATE
2	1	2 - LOW
2	2	4 - MODERATE
2	3	6 - HIGH
3	1	3 - MODERATE
3	2	6 - HIGH
3	3	9 - HIGH

Figure 4.5 provides an example of a one dimensional approach for calculating risk. In this example, the levels of risk are now normalized (i.e. low, medium and high) and can be used to compare risks associated with each threat. The comparison of risk measures should factor in the criticality of the components used to determine the risk measure. For simple methodologies that only look at loss and likelihood, a risk measure that was derived from a high loss and low likelihood may result in the same risk measure as one that resulted from a low loss and high likelihood. In these cases, the user needs to decide which risk measure to consider more critical, even though the risk measures may be equal. In this case, a user may decide that the risk measure derived from the high loss is more critical than the risk measure derived from the high likelihood.

Figure 4.6 - Calculating Cost Measure

In this example cost measure, the cost of the safeguard is the amount needed to purchase or develop and implement each of the mechanisms. The cost can be normalized in the same manner as was the value for potential loss incurred. A 1 will indicate a mechanism with a low cost, a 2 will indicate a mechanism with a moderate cost, and a 3 will indicate a mechanism with a high cost.

With a list of potential threats, vulnerabilities and related risks, an assessment of the current security situation for the LAN can be determined. Areas that have adequate protection will not surface as contributing to the risk of the LAN (since adequate protection should lead to low likelihood) whereas those areas that have weaker protection do surface as needing attention.

4.1.0.4 Process 5 - Select Appropriate Safeguards

The purpose of this process is to select appropriate safeguards. This process can be done using risk acceptance testing.

Risk acceptance testing is described by [KATZ92] as an activity that compares the current risk measure with acceptance criteria and results in a determination of whether the current risk level is acceptable. While effective security and cost considerations are important factors, there may be other factors to consider such as: organizational policy, legislation and regulation, safety and reliability requirements, performance requirements, and technical requirements.

The relationship between risk acceptance testing and safeguard selection can be iterative. Initially, the organization needs to order the different risk levels that were determined during the risk assessment. Along with this the organization needs to decide the amount of residual risk that it will be willing to accept after the selected safeguards are implemented. These initial risk acceptance decisions can be factored into the safeguard selection equation. When the properties of the candidate safeguards are known, the organization can reexamine the risk acceptance test measures and determine if the residual risk is achieved, or alter the risk acceptance decisions to reflect the known properties of the safeguards. For example there may be risks that are determined to be too high. However after reviewing the available safeguards, it may be realized that the currently offered solutions are very costly and cannot be easily implemented into the current configuration and network software. This may force the organization into either expending the resources to

Figure 4.7 - Comparing Risk and Cost

To calculate risk/cost relationships use the risk measure and the cost measure associated with each threat/mechanism relationship and create a ratio of the risk to the cost (i.e., risk/cost). A ratio that is less than 1 will indicate that the cost of the mechanism is greater than the risk associated with the threat. This is generally not an acceptable situation (and may be hard to justify) but should not be automatically dismissed. Consider that the risk value is a function of both the loss measure and the likelihood measure. One or both of these may represent something so critical about the asset that the costly mechanism is justified. This situation may occur when using simple methodologies such as this one.

reduce the risk, or deciding through risk acceptance that the risk will have to be accepted because it is currently too costly to mitigate.

Many sources exist that can provide information on potential safeguards. The methodology discussed here defines safeguards in terms of security services and mechanisms. A security service is the sum of mechanisms, procedures, etc. that are implemented on the LAN to provide protection. The security services (and mechanisms) provided in Section 2 can be used as a starting point. The security services should be related to the threats defined in the risk assessment.

In most cases the need for a specific service should be readily apparent. If the risk acceptance results indicate that a risk is acceptable, (i.e., existing mechanisms are adequate) then there is no need to apply additional mechanisms to the service that already exists.

After the needed security services are determined, consider the list of security mechanisms for each service. For each security service selected, determine the candidate mechanisms that would best provide that service. Using the threat/vulnerability/risk relationships developed in the previous processes, choose those mechanisms that could potentially reduce or eliminate the vulnerability and thus reduce the risk of the threat. In many cases, a threat/vulnerability relationship will yield more than one candidate mechanism. For example the vulnerability of using weak passwords could be reduced by using a password generator mechanism, by using a token based mechanism, etc. Choosing the candidate mechanisms is a subjective process that will vary from one LAN implementation to another. Not every mechanism presented in Section 2 is feasible for use in every LAN. In order for this process to be beneficial, some filtering of the mechanisms presented needs to be made during this step.

Selecting appropriate safeguards is a subjective process. When considering the cost measure of the mechanism, it is important that the cost of the safeguard be related to the risk measure to determine if the safeguard will be cost-effective. The methodology chosen by the organization should provide a measure for representing costs that is consistent with the measures used for representing the other variables determined so far. Figure 4.6 shows a cost measure that is consistent with the other measuring examples presented. This cost measuring method, while appearing to only consider the cost of the safeguard, can have the other factors mentioned above factored in.

When a measure (or cost) is assigned to the safeguard, it can be compared to the other measures in the process. The safeguard measure can be compared to the risk measure (if it consists of one value, as shown in Figure 4.7) or the components of the risk measure. There are different ways to compare the safeguard measure to the risk measure. The risk management methodology chosen by the organization should provide a method to select those effective safeguards that will reduce the risk to the LAN to an acceptable level.

4.1.0.5 Process 6 - Implement And Test Safeguards

The implementation and testing of safeguards should be done in a structured manner. The goal of this process is to ensure that the safeguards are implemented correctly, are compatible with other LAN functionalities and safeguards, and provide expected protection. This process begins by developing a plan to implement the safeguards. This plan should consider factors such as available funding, users' learning curve, etc. A testing schedule for each safeguard should be incorporated into this plan. This schedule should show how each safeguard interacts or effects other safeguards (or mechanisms of some other functionality). The expected results

(or the assumption of no conflict) of the interaction should be detailed. It should be recognized that not only is it important that the safeguard perform functionally as expected and provide the expected protections, but that the safeguard does not contribute to the risk of the LAN through a conflict with some other safeguard or functionality.

Each safeguard should first be tested independently of other safeguards to ensure that it provides the expected protection. This may not be relevant to do if the safeguard is designed to interwork with other safeguards. After testing the safeguard independently, the safeguard should be tested with other safeguards to ensure that it does not disrupt the normal functioning of those existing safeguards. The implementation plan should account for all these tests and should reflect any problems or special conditions as a result of the testing.

4.1.0.6 Process 7 - Accept Residual Risk
After all safeguards are implemented, tested and found acceptable, the results of the risk acceptance test should be reexamined. The risk associated with the threat/vulnerability relationships should now be reduced to an acceptable level or eliminated. If this is not the case, then the decisions made in the previous steps should be reconsidered to determine what the proper protections should be.

4.2 RCMP Guide to Threat and Risk Assessment For Information Technology

4.2.1 Introduction

This guide is intended to assist practitioners in assessing the threats and risks to Information Technology (IT) assets held within their organizations, and in making recommendations related to IT security. The objective of a threat and risk assessment (TRA) is to involve the various players and gain their support, to enable management to make informed decisions about security and to recommend appropriate and cost-effective safeguards. An assessment of the adequacy of existing safeguards also forms part of the TRA process. Where this assessment indicates that safeguards are inadequate to offset vulnerabilities, additional safeguards are recommended. Also, where the TRA indicates that certain safeguards are no longer needed, the elimination of those safeguards is recommended. A TRA does not result in the selection of mechanisms of prevention, detection and response to reduce risks; instead, it simply indicates the areas where these mechanisms should be applied, and the priorities, which should be assigned to the development of such mechanisms. Within the context of risk management, the TRA will recommend how to minimize, avoid, and accept risk.

Planning for the TRA process encompasses establishing the scope of the project, determining the appropriate methodology, setting the time frame, identifying the key players and allocating resources to perform the assessment. Those involved in the TRA process must be cautioned to protect the sensitivity of working papers produced during the process. These working papers often contain information related to the vulnerability of systems and environments, and should be protected at a level commensurate with the most sensitive information available on those systems.

Consideration must be given to specific organizational characteristics that might indicate the need for a strengthened security profile. Such characteristics might include the organization's mandate, the location (i.e. remoteness) and the organization's composition in terms of environment ("hostile", public access) and resources.

4.2.2 Process

To conduct a TRA, the following four-step process is typically followed.

1. Preparation:	determining what to protect;
2. Threat Assessment:	determining what to protect against, consequences of a threat;
3. Risk Assessment:	determining whether existing or proposed safeguards are satisfactory; and
4. Recommendations:	identifying what should be done to reduce the risk to a level acceptable to senior management.

Each of these steps is described in detail in subsequent sections.

4.2.2.0 PREPARATION

Defining the Environment

1. Determining the Scope of the Threat and Risk Assessment

Prior to the actual conduct of the TRA, it is necessary to establish its scope, which will include the systems under consideration, the interconnectivity with other systems and the profile of the user community. The entire TRA process will often span a number of systems and environments. Thus, in determining the scope, care must be taken to ensure that priorities are set to determine an appropriate order of assessment, i.e. that areas of primary concern or sensitivity are assessed first.

2. Identifying Team Participants

Once the scope of the TRA has been established, the practitioner can establish a representative team of users of the system under consideration. For example, let us suppose that the system contains several applications used by a variety of groups within the institution. To provide a valid cross section of the information required to conduct the TRA, users, developers, and telecommunications and operations staff must be selected for the team. This team will (at a later step) provide the practitioner with the information required to identify known threats and their potential impact.

3. Determining Intrinsic Concerns

All organizations have certain security concerns that are directly related to the nature of their business. The practitioner should document these special concerns, as they will be instrumental in determining the appropriateness of existing security measures and in making recommendations for improvements.

4. Developing the Baseline

Once the preliminary work is completed, the practitioner can establish the current profile of the organization's security posture. These parameters establish what is known as the security baseline for the TRA process. It is from this baseline that the risks are assessed, and any updates to the TRA are prepared. For example, when a particular safeguard is recommended, that safeguard and its defining recommendation are referred directly to the baseline. A baseline against which recommendations can be made is necessary for two reasons:

- The baseline provides a starting point for any measurement of progress.
- The environment is subject to continual change.

The first point provides the practitioner with a means of determining what changes have been made to the environment and how security has been impacted by those changes. The second point allows the practitioner to identify the difference between the current security profile and any future requirements for security, given the changes to the environment, which have taken place since the baseline was established.

Assets Identification and Valuation

Identifying IT assets according to their physical and logical groupings can be a difficult task, depending on the size of the organization and the soundness of supporting activities such as materiel management and the availability of comprehensive inventories. The practitioner must identify those assets that form the IT environment, and then assign a value to them. The participants identified in the preparation stage will be instrumental in identifying and assigning value to assets. In the case of IT applications, the "owners" of the information processed by those applications are responsible for preparing the statement of sensitivity which will detail the specific sensitivity requirements for each application in terms of confidentiality, integrity and availability.

The practitioner must consider several aspects contributing to the worth of an asset including, but not limited to, the initial cost of the item. An asset may have an **acquired** value that far outweighs the initial cash outlay. Consider the example of the data collected by geologists during a summer survey of a remote northern area. The project objective may be to collect the data while the area is accessible and interpret and analyze the data over the winter months. The value could be considered to be equal to the cost of the survey in terms of scientists' time, support and travel costs. However, suppose the data is lost in September (therefore not available) and the area is inaccessible until spring. The geologists will have lost an entire year's work plus the cost of the initial survey in that the data must be gathered again the following summer. The asset value must be increased by the costs associated with an additional year's support, time and travel costs as well as any uniqueness in time, conditions and opportunity.

The question of using qualitative versus quantitative methods in the determination of asset value must also be addressed. When considering the acquired value of certain assets, it may be more meaningful (than assigning a dollar value) to establish the relative value of an asset within the context of the organizational objectives and mandate. This relative value can be expressed in terms of the confidentiality, integrity and availability requirements for that asset.

Confidentiality

Confidentiality is used in the context of sensitivity to disclosure. In some instances, the sensitivity involves a degree of time dependency. For example, some research is sensitive as data is being gathered and processed; but once published it becomes a matter of public record and therefore no longer possesses the same degree of confidentiality. In some instances, data may acquire a higher level of confidentiality when put together in an aggregate form; e.g. army movement logistics may be derived from an aggregate of supply data to individual units.

To assess the impact of loss of confidentiality, practitioners must relate the level of sensitivity of the data to the consequences of its untimely release. The data must be appropriately classified or designated according to the following levels:

UNCLASSIFIED OR UNDESIGNATED	basic information
DESIGNATED	varying levels, personal information, sensitive business information
CONFIDENTIAL	compromise could cause injury to the national interest
SECRET	compromise could cause serious injury to the national interest
TOP SECRET	compromise could cause exceptionally grave injury to the national interest

The confidentiality considerations checklist (Table 1) stipulates some questions to be answered in the assessment of the confidentiality requirements of the system or of the information it contains.

CONFIDENTIALITY CONSIDERATIONS CHECKLIST	
	Is the information sensitive in the national interest, i.e. classified?
	Is the information personal?
	What is the consequence of loss of confidentiality of this information?

TABLE 1 - Confidentiality

Integrity

Integrity is used in the context of accuracy and completeness of the information accessible on the system and of the system itself. Where integrity requirements are high, as is the case with financial transactions in banking systems, the potential financial losses will indicate the appropriate levels of investment in safeguards.

The integrity considerations checklist (Table 2) stipulates some aspects to be addressed in the assessment of the integrity requirements of the system or of the information it contains.

INTEGRITY CONSIDERATIONS CHECKLIST	
	Impact of inaccurate data.
	Impact of incomplete data.

TABLE 2 - Integrity

Availability

The system, to be considered available, must be in place and useable for the intended purpose. While the complete loss of data processing capability is unlikely, it could occur. Unscheduled downtimes of varying degrees of severity are certain. The practitioner must assist the users in establishing how much they rely on the system's being available to provide the expected service. The users must clearly define for the systems staff the maximum acceptable levels of downtime. In this context, the term "availability" relates to continuity of service.

To the practitioner, establishing processing priority based on availability requirements often involves mediating between user groups and reaching agreement on the relative

importance of applications to each group. The practitioner must also recognize that availability requirements often change during the lifespan of the application. The user community should document for the systems staff the impact of the loss of availability of the IT systems, support personnel and data.

Those services that are considered to be ***essential or mission-critical services*** must be identified. Such services have a high availability requirement and, as a result, special consideration must be given to the support resources and environmental aspects, which affect the provision of service.

The practitioner must determine all critical components involved in the provision of essential service that could be vulnerable to threats. These critical components are also considered to be "assets" for the purposes of the TRA.

The availability considerations checklist (Table 3) stipulates some aspects, which should be addressed in the assessment of the availability requirements.

AVAILABILITY CONSIDERATIONS CHECKLIST	
	Changes in availability requirements within the system's life cycle
	Documented impact of loss of availability
	Documented maximum acceptable periods of downtime

TABLE 3 - Availability

Statements of Sensitivity

The CIA requirements are documented in the statements of sensitivity (SOSs). The preparation of a statement of sensitivity should be a prerequisite to the implementation of a new application or changes to existing ones. Applications developed and implemented without statements of sensitivity often do not allow for the necessary security requirements to adequately protect the information available on the system. The statement of sensitivity should be prepared by the responsibility centre, which provides data to, and uses or has ownership of, the application. The analysis that leads to the preparation of the statement of sensitivity is sometimes conducted by a number of different people each of whom has some interest in the system or data under consideration.

The user representation for completing the statement of sensitivity could be one person or several, depending on the size and complexity of the application being assessed.

A separate statement of sensitivity is required for each major application used on the computer system or anticipated for installation. For example, payroll and inventory would each require a statement of sensitivity, even if they are to be run on the same system. The sensitivity-related valuation of assets is not necessarily linked to numerical values associated with initial or replacement costs; but rather is linked to a relative value associated with the application's requirements for confidentiality, integrity and availability.

4.2.2.1 THREAT ASSESSMENT

The second step of the TRA process is the **Threat Assessment.** The threat concepts of class, likelihood, consequence, impact and exposure are highlighted. Specific threat

events such as earthquakes, hacker attempts, virus attacks etc. fall into a particular threat class, depending on the nature of the compromise. Examples of threats within each class can be found in Figure 3.

THREAT CLASS	SAMPLE THREATS
DISCLOSURE	• Compromising Emanations • Interception • Improper Maintenance Procedures • Hackers
INTERRUPTION	• Earthquake • Fire • Flood • Malicious Code • Power Failure
MODIFICATION	• Data Entry Errors • Hackers • Malicious Code
DESTRUCTION	• Earthquake • Fire • Flood • Power Spikes
REMOVAL	• Theft of Data • Theft of Systems

FIGURE 3 - Sample Threats

Description of Threat
The threats that may target the assets under consideration must be described by the practitioner. These threats may originate from either deliberate or accidental events.

Classes of Threats
The practitioner will classify the threats into one of the five main classes of threats: disclosure, interruption, modification, destruction and removal or loss.

Disclosure
Assets that have a high confidentiality requirement are sensitive to disclosure. This class of threats compromises sensitive assets through unauthorized disclosure of the sensitive information.

Interruption
Interruption relates primarily to service assets. Interruption impacts the availability of the asset or service. A power outage is an example of a threat, which falls into the interruption class.

Modification
The primary impact of this class of threats is on the integrity requirement. Recall that integrity, as defined in the GSP, includes both accuracy and completeness of the information. A hacker attempt would fall into this class of threat if changes were made.

Destruction
A threat, which destroys the asset, falls into the destruction class. Assets that have a high availability requirement are particularly sensitive to destruction. Threats such as earthquake, flood, fire and vandalism are within the destruction class.

Removal or Loss
When an asset is subject to theft or has been misplaced or lost, the impact is primarily on the confidentiality and availability of the asset. Portable computers or laptops are particularly vulnerable to the threat of removal or loss.

Threat Likelihood
The practitioner must consider, on a per-asset basis, both the type of threat that the asset may be subjected to and the likelihood of the threat. The likelihood of threat can be estimated from past experience, from threat information provided by lead agencies and from sources such as other organizations or services.

Likelihood levels of low, medium and high are used according to the following definitions (Source: Government of Canada Security Policy):

- **Not Applicable** may be used to indicate that a threat is considered not to be relevant to the situation under review.
- **Low** means there is no history and the threat is considered unlikely to occur.
- **Medium** means there is some history and an assessment that the threat may occur.
- **High** means there is a significant history and an assessment that the threat is quite likely to occur.

Consequences, Impact and Exposure
Once the assets are listed and the threats are categorized according to the five major classes, the practitioner must assess the impact of a threat occurring in the absence of any safeguards. In order to assess the impact, the practitioner must be able to understand and describe the business of the organization. The practitioner must consider what the effect would be on the work being done, on the organization itself, and on those elements of the business that rely on the information or service provided by the specific asset under threat.

During this process, the practitioner seeks to answer the question "What is the consequence of each particular threat?" This consequence is related to the losses or other consequences (both real and perceived) which could result from a specific threat being successful.

The Government of Canada Security policy identifies an impact- reporting mechanism based on an injury assessment. In the case of classified or designated assets or information, group impact into levels of less serious injury, serious injury and exceptionally grave injury. Consequences could be expressed in such terms as "loss of trust", "loss of privacy", "loss of asset" or "loss of service". The practitioner could add other similarly phrased consequences as needed.

The mapping of the consequence onto one of the three impact ratings (exceptionally grave, serious, less serious) would vary according to departmental priorities. For example, in one department a loss of trust might be regarded as serious injury in terms

of impact, while in another department, the same loss of trust might be considered to be exceptionally grave injury. The impact assessment allows the practitioner to determine the impact to the organization in terms of the real and perceived costs associated with the loss of confidentiality, integrity, and availability.

The identification of exposure allows the organization to rank the risk scenario according to the likelihood and impact, and thus assign a priority.

This general exposure rating for data and assets is outlined in Table 4 where impact takes precedence over likelihood. This table provides a means of prioritizing the impact through a rating that considers only the likelihood of a particular threat and the associated impact on the organization should the threat materialize. Table 4 does not consider the safeguards employed to counterbalance a particular threat.

		IMPACT (INJURY)		
		Exceptionally Grave	Serious	Less Serious
Likeli-hood		9	8	5
	MEDIUM	7	6	3
	LOW	4	2	1

TABLE 4 - Exposure Ratings for Data and Assets

Summarizing Threat Assessment
Threat Assessment as described in this section encompasses:

a) Describing threats in terms of who, how and when.
b) Establishing into which threat class a threat falls.
c) Determining the threat likelihood.
d) Determining the consequences on the business operations should a threat be successful.
e) Assessing the impact of the consequences as less serious, serious or exceptionally grave injury.
f) Assigning an exposure rating to each threat, in terms of the relative severity to the organization.
g) Prioritising the impacts/likelihood pairs, according to the ratings determined in (f).

Table 5 provides a sample summary sheet on which the threat assessment information may be entered on a per-asset basis.

ASSET	THREAT ASSESSMENT					
	AGENT/ EVENT	CLASS OF THREAT	LIKELIHOOD OF OCCURRENCE	CONSEQUENCE OF OCCURRENCE	IMPACT (INJURY)	EXPOSURE RATING
Describe the Asset.	Describe the threat event.	Disclosure Interruption Modification Destruction Removal	Low Medium High	List the consequences to the organization of the threat occurring.	Exceptionally grave, serious, less serious.	Numerical Value 1 to 9

TABLE 5 - Generic Threat Assessment

4.2.2.2 RISK ASSESSMENT

Risk assessment is necessary to determine risk assumed by the organization where existing or proposed safeguards are deemed inadequate to protect the asset against an identified threat. Where existing safeguards are not adequate, a vulnerability is noted and analyzed.

Risk assessment is *"an evaluation of the chance of vulnerabilities being exploited, based on the effectiveness of existing or proposed security safeguards".*

This definition leads the risk assessment process into an evaluation of the vulnerabilities and the likelihood that a vulnerability would be exploited by a threat in the presence of either existing or proposed security measures.

Evaluating Existing Safeguards
Determining what existing safeguards could counter the identified threats is the next logical step in the process of TRA. Once the existing safeguards are grouped on a per-threat basis, the practitioner can assess the security posture of the business or facility relative to each threat, and determine whether any residual vulnerability or weakness exists.

Vulnerabilities
Attention should be paid to times during which the asset is most vulnerable, for example, during periods of public access and unrestricted access or while in transit. In some instances, an asset has an associated time sensitivity. For example, the information may be sensitive while under review or development (e.g. budget) and then may lose its sensitivity upon release to the public.

There are three possible security posture scenarios in the threat and safeguards environment. The first is identified in Figure 2 as an *equilibrium* state. This state of equilibrium is the most desirable security posture. In this environment, threats are

identified and appropriate safeguards are in place to reduce the associated risks to a level, which is acceptable to the organization's senior management.

The second security posture, which an organization might experience, is referred to as a *vulnerable* state (Figure 3), since the threats outweigh the safeguards. The insecurity produced can result in a variety of IT - related losses, which compromise the confidentiality, integrity and availability of the information.

The third security posture is referred to as an *excessive* state (Figure 4) since the safeguards employed exceed the threats. The result is an overspending in the area of security measures, which is not commensurate with the threat; and thus is not justifiable.

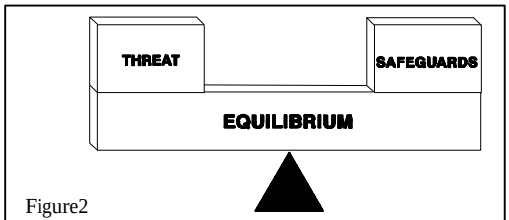


Figure2

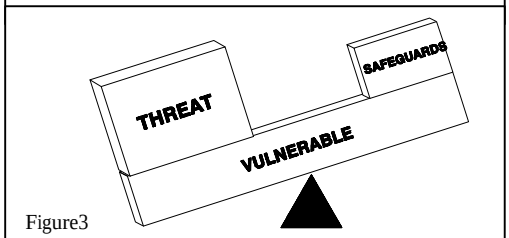


Figure3

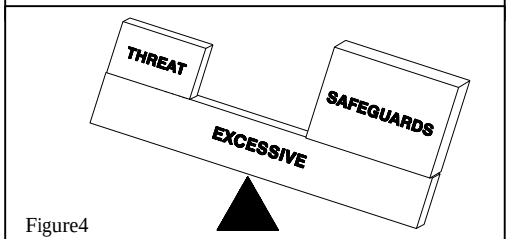


Figure4

When it is determined that the security posture matches Figure 3 - Vulnerable, the practitioner must consider the possibility that a vulnerability would be exploited. This depends on a number of factors, some of which were explored in the Threat Assessment:

- likelihood of threat,
- possible motive for exploiting the vulnerability,
- value of the asset to the organization and to the threat agent, and
- effort required to exploit the vulnerability.

For example, a vulnerability could exist but, in the absence of one or more of the above factors, it may never be exploited.

Risk

Risk is defined as, *"the chance of vulnerabilities being exploited".*

The level of risk existing in the organization can be categorized as:

- **high:** requiring immediate attention and safeguard implementation,
- **medium:** requiring attention and safeguard implementation in the near future, or
- **low:** requiring some attention and consideration for safeguard implementation as good business practice.

The practitioner will be able to decide the priority for each component of the risk management program based on items such as the nature of identified threats and the impact on the organization. Having reviewed the existing safeguards and vulnerabilities, the practitioner establishes the adequacy of safeguards and recommends change. For an example of establishing risk for deliberate threat scenarios, refer to Annex E.

Summarizing Risk Assessment

Risk Assessment as described in this section encompasses:

- examining existing safeguards,

- establishing vulnerabilities, and
- determining the level of risk based on a number of factors.

Table 6 provides a sample summary sheet for entering the risk assessment information on a per-asset basis.

ASSET	THREAT	Risk Assessment		
		Existing Safeguards	**Vulnerability**	**RISK**
Describe the Asset	Describe the specific threat against it	Describe existing safeguards to protect the asset against the threat	Describe any vulnerabilities that may be observed	Establish risk level

TABLE 6 - Generic Risk Assessment

4.2.2.3 RECOMMENDATIONS

The closing phase of the TRA process includes the proposal of recommendations. These recommendations are intended to improve the security posture of the organization through risk reduction, provide considerations for business recovery activities should a threat cause damage, and identify implementation constraints. Once safeguards that would augment the existing safeguards and improve the security profile are proposed, the risk posture can be re-evaluated as low, medium or high.

Proposed Safeguards

At this point in the process, the practitioner has analyzed the nature of the threats, the impact of successful threats, and the organization's vulnerability to these threats and has subsequently judged the risk to be low, medium, or high. Where the practitioner perceives that the risk can be reduced, appropriate recommendations are made. The practitioner may recommend a number of scenarios, each with an associated effect and cost, from which senior management will make an appropriate selection.

Where the assessment of threats and associated risks leads to specific recommendations, the practitioner must also consider the feasibility of such recommendations.

Projected Risk

In some instances, proposed safeguards will reduce or eliminate some, but not all, risks. For such instances, the resulting projected risk should be documented and signed off by senior management. For example, the initial risk assessment indicated a high risk situation, and several safeguards were recommended by the TRA team. In the presence of these additional safeguards, the risk is re-evaluated as being moderate to low. Thus the priority level of this scenario is reduced but not eliminated, and senior management should acknowledge and accept or reject the

projected risk levels. Rejecting the risk implies that other safeguards must be sought to further reduce or eliminate the risk.

Ranking of the implemented safeguards can be accomplished in a number of ways, for example:

- Refer to the impact-rating column of the threat assessment phase
- Compare the change in risk level before a proposed safeguard is implemented, in the risk assessment phase risk column to after, in the recommendations phase risk column.

Impact ratings of 9 should be looked at first because they represent events that have high likelihood and very serious impact. In some instances the change in risk level from high to low is desirable, in particular where the exposure rating is high.

Overall Assessment of Safeguards

Safeguards and associated risk should be evaluated based on the following categories:

- completely satisfactory;
- satisfactory in most aspects;
- needs improvement.

The risks of deliberate threats to the organization have been established by way of the Risk Assessment Grid described in Appendix E. For accidental threats, the risk will be assessed according to their history within the organization or similar institutions and the observed effectiveness of associated safeguards in each comparable environment. The highest priority must be assigned to those threats posing a high risk to the organization. For each of these threats, the practitioner will propose safeguards to eliminate the risk or reduce it to a level acceptable to senior management. The adequacy of each of these proposed safeguards must be evaluated as completely satisfactory, satisfactory in most aspects, or needs improvement.
The practitioner establishes the appropriateness and interdependencies of safeguards, and answers such questions as: Are safeguards in conflict? Does one safeguard offset the usefulness of another? Does the safeguard overcompensate the threat? What threats have not been fully compensated for? What is the risk that vulnerabilities which are not fully compensated for are likely to be exploited and by whom?

4.2.3 ***Updates***

The TRA is considered to be a vital, living document, which is essential to meeting the security objectives of the organization. The TRA must be updated at least annually, or whenever an occurrence reveals a deficiency in the existing assessment. The TRA should also be updated whenever changes are planned to the systems or environments in which the IT processing occurs, which could create new risks or redundant safeguards.

Regular Review

Regular reviews allow the practitioner to revisit the TRA document and assess whether the IT security requirements within the organization have changed. These regular reviews are necessary in light of both the dynamics of the technologies in place to support IT and the dynamics of technologies available to threat agents to help them attack the IT systems of the organization.

Systems Changes
Changes to systems can greatly impact the security profile; therefore, every change must be assessed. The TRA document provides the practitioner with a baseline against which the effects of these changes can be measured. Examples of changes include the move of an organization from stand-alone PCs to a Local Area Network environment, the introduction of new applications to existing systems, the introduction of Wide Area Network capability to existing IT environments, a change in communications links or protocols used to move information between departmental units, or a change in the level of the most sensitive information on the system.

Threat Profile Changes
Changes in the threat profile will also have a potential impact on the TRA. For example, when threat agent motivation diminishes or the effort expended by the threat agent increases, the threat from that source may be reduced. Since changes in the threat profile do not always follow a cyclical pattern, the practitioner must stay in touch with the current threat levels and update the TRA accordingly.

4.2.4 Advice and Guidance

Threats
Sources of historical threat information vary, depending on the type of information sought. For threat information based on events that have already occurred within the organization, the practitioner should consult the Departmental Security Officer. For threat information related to investigations under the Criminal Code of Canada involving IT assets, the practitioner should consult the OIC, Information Technology (IT) Security Branch of the RCMP. Where threat information relates to COMSEC, the practitioner should consult the Communications Security Establishment. The Canadian Security Intelligence Service (CSIS) provides threat information and advice on threat assessment when requested.

TRA Process

Advice and guidance on the TRA process as described in this document are available through the OIC,IT Security Branch of the RCMP.

4.2.5 *Glossary of Terms*

1. **Analyse**: to study or determine the nature and relationship of the parts.
2. **Assess**: to evaluate the extent to which certain factors (Threats, Vulnerabilities and Risks) affect the IT environment.
3. **Asset**: any item that has value.
4. **Availability**: the condition of being usable on demand to support business functions.
5. **Compromise**: unauthorized disclosure, destruction, removal, modification or interruption.
6. **Confidentiality**: the sensitivity of information or assets to unauthorized disclosure, recorded as classification or designation, each of which implies a degree of injury should unauthorized disclosure occur.
7. **Consequence**: outcome, effect.
8. **Critical**: crucial, decisive.
9. **Equilibrium**: a state of balance existing between two or more opposing forces.
10. **Evaluate**: to determine the amount or worth of, or to appraise.
11. **Exposure**: the state of being vulnerable to criticism or attack.
12. **Impact**: effect of one thing on another.
13. **Information technology**: The scientific, technological and engineering disciplines and the management technologies used in information handling, communication and processing; the fields of electronic data processing, telecommunications, networks, and their convergence in systems; applications and associated software and equipment together with their interaction with humans and machines.
14. **Intangible**: incapable of being perceived by touch.
15. **Integrity**: the accuracy and completeness of information and assets and the authenticity of transactions.
16. **Likelihood**: the state or quality of being probable, probability.
17. **Practitioner**: one who practises within an area of expertise.
18. **Process**: a series of continuous actions to bring about a result.
19. **Qualitative**: of or pertaining to quality, describable.
20. **Quantitative**: of or pertaining to quantity, measurable.
21. **Risk assessment**: an evaluation of the chance of vulnerabilities being exploited, based on the effectiveness of existing or proposed safeguards.
22. **Safeguards**: actions or measures taken to offset a particular security concern or threat.
23. **Security baseline**: an established security profile or posture, which has been determined at an established point in time.
24. **Tangible**: perceptible by touch.
25. **Threat assessment**: an evaluation of the nature, likelihood and consequence of acts or events that could place sensitive information and assets as risk.
26. **Threat**: any potential event or act that could cause one or more of the following to occur: unauthorized disclosure, destruction, removal, modification or interruption of sensitive information, assets or services, or injury to people. A threat may be deliberate or accidental.

4.1 Guideline for the Analysis Local Area Network Security., Federal Information Processing Standards Publication 191, November 1994. Chapter 3.4.

[MART89] Martin, James, and K. K. Chapman, The Arben Group, Inc.; Local Area Networks, Architectures and Implementations, Prentice Hall, 1989.

[BARK89] Barkley, John F., and K. Olsen; Introduction to Heterogenous Computing Environments, NIST Special Publication 500-176, November, 1989.

[NCSC87] A Guide to Understanding Discretionary Access Control in Trusted Systems, NCSC-TG-003, Version 1, September 30, 1987

[NCSL90] National Computer Systems Laboratory (NCSL) Bulletin, Data Encryption Standard, June, 1990.

[SMID88] Smid, Miles, E. Barker, D. Balenson, and M. Haykin; Message Authentication Code (MAC) Validation System: Requirements and Procedures, NIST Special Publication 500-156, May, 1988.

[OLDE92] Oldehoeft, Arthur E.; Foundations of a Security Policy for Use of the National Research and Educational Network, NIST Interagency Report, NISTIR 4734, February 1992.

[COMM91] U.S. Department of Commerce Information Technology Management Handbook, Attachment 13-D: Malicious Software Policy and Guidelines, November 8, 1991.

[WACK89] Wack, John P., and L. Carnahan; Computer Viruses and Related Threats: A Management Guide, NIST Special Publication 500-166, August 1989.

[X9F292] Information Security Guideline for Financial Institutions, X9/TG-5, Accredited Committee X9F2, March 1992.

[BJUL93] National Computer Systems Laboratory (NCSL) Bulletin, Connecting to the Internet: Security Considerations, July 1993.

[BNOV91] National Computer Systems Laboratory (NCSL) Bulletin, Advanced Authentication Technology, November 1991.

[KLEIN] Daniel V. Klein, "Foiling the Cracker: A Survey of, and Improvements to, Password Security", Software Engineering Institute. (This work was sponsored in part by the Department of Defense.)

[GILB89] Gilbert, Irene; Guide for Selecting Automated Risk Analysis Tools, NIST Special Publication 500-174, October, 1989.

[KATZ92] Katzke, Stuart W. ,Phd., "A Framework for Computer Security Risk Management", NIST, October, 1992.

[NCSC85] Department of Defense Password Management Guideline, National Computer Security Center, April, 1985.

[NIST85] Federal Information Processing Standard (FIPS PUB) 112, Password Usage, May,1985.

[ROBA91] Roback Edward, NIST Coordinator, Glossary of Computer Security Terminology,NISTIR 4659, September, 1991.

[TODD89] Todd, Mary Anne and Constance Guitian, Computer Security Training Guidelines,NIST Special Publication 500-172, November, 1989.

[STIE85] Steinauer, Dennis D.; Security of Personal Computer Systems: A

Management Guide, NBS Special Publication 500-120, January, 1985.

[WACK91] Wack, John P.; Establishing a Computer Security Incident Response Capability (CSIRC), NIST Special Publication 800-3, November, 1991.

[NIST74] Federal Information Processing Standard (FIPS PUB) 31, Guidelines for Automatic Data Processing Physical Security and Risk Management, June, 1974.

4.2 Royal Canadian Mounted Police Technical Operations Directorate. Information Technology Security Branch. *Guide to Threat and Risk Assessment. For Information Technology.* Security Information Publications, November 1994.

5.1 Introduction

Perhaps it is best to describe first what a firewall is not: A firewall is not simply a router, host system, or collection of systems that provides security to a network. Rather, a firewall is an approach to security; it helps implement a larger security policy that defines the services and access to be permitted, and it is an implementation of that policy in terms of a network configuration, one or more host systems and routers, and other security measures such as advanced authentication in place of static passwords. The main purpose of a firewall system is to control access to or from a protected network (i.e., a site). It implements a network access policy by forcing connections to pass through the firewall, where they can be examined and evaluated. A firewall system can be a router, a personal computer, a host, or a collection of hosts, set up specifically to shield a site or subnet from protocols and services that can be abused from hosts outside the subnet. A firewall system is usually located at a higher level gateway, such as a site's connection to the Internet, however firewall systems can be located at lower-level gateways to provide protection for some smaller collection of hosts or subnets.

The main function of a firewall is to centralize access control. A firewall serves as the gatekeeper between the untrusted Internet and the more trusted internal networks. If outsiders or remote users can access the internal networks without going through the firewall, its effectiveness is diluted. For example, if a traveling manager has a modem connected to his office PC that he or she can dial into while traveling, and that PC is also on the protected internal network, an attacker who can dial into that PC has circumvented the firewall. Similarly, if a user has a dial-up Internet account with a commercial ISP, and sometimes connects to the Internet from their office PC via modem, he or she is opening an unsecured connection to the Internet that circumvents the firewall.

What is being protected by firewalls?

- **Your data**
  - Secrecy - what others should not know
  - Integrity - what others should not change
  - Availability - your ability to use your own systems
- **Your resources**
  - Your systems and their computational capabilities
- **Your reputation**
  - Confidence is shaken in your organization
  - Your site can be used as a launching point for crime
  - You may be used as a distribution site for unwanted data
  - You may be used by impostors to cause serious problems
  - You may be viewed as “untrusted” by customers and peers

Firewalls provide several types of protection:

- They can block unwanted traffic.
- They can direct incoming traffic to more trustworthy internal systems.
- They hide vulnerable systems, which can’t easily be secured from the Internet.
- They can log traffic to and from the private network.
- They can hide information like system names, network topology, network device types, and internal user ID’s from the Internet.
- They can provide more robust authentication than standard applications might be able to do.

As with any safeguard, there are trade-offs between convenience and security. Transparency is the visibility of the firewall to both inside users and outsiders going through a firewall. A firewall is transparent to users if they do not notice or stop at the firewall in order to access a network. Firewalls are typically configured to be transparent to internal network users (while going outside the firewall); on the other hand, firewalls are configured to be non-transparent for outside network coming through the firewall. This generally provides the highest level of security without placing an undue burden on internal users.

5.2 Firewall Security and Concepts

- The amount of security required for an entity is based on the security threat
- If you do not know what your threat is to the Intranet systems, it is extremely difficult to properly secure the environment and all systems interconnected
- Network compartmentalization is the buzzword for this type of effort
- Switching technology is a big help, but it does not tell you who is going where and why - that's what analysis is all about
- Not knowing the threat causes false security to be deployed and money spent in the wrong places

The main reasons for systems and computers not being secure are

- Lack of password encryption
- Lack of personnel with experience
- Lack of management backing
  - Authority
  - Responsibility
- Legal and political issues
- Lack of recurring effort
- Budget

5.2.0 Firewall Components

The primary components (or aspects) of a firewall are:

- Network policy,
- Advanced authentication mechanisms,
- Packet filtering, and Application gateways.

The following sections describe each of these components more fully.

5.2.0.0 NETWORK POLICY

There are two levels of network policy that directly influence the design, installation and use of a firewall system. The higher-level policy is an issue-specific, network access policy that defines those services that will be allowed or explicitly denied from the restricted network, how these services will be used, and the conditions for exceptions to this policy. The lower-level policy describes how the firewall will actually go about restricting the access and filtering the services that were defined in the higher level policy. The following sections describe these policies in brief.

5.2.0.1 SERVICE ACCESS POLICY

The service access policy should focus on Internet-specific use issues as defined above, and perhaps all outside network access (i.e., dial-in policy, and SLIP and PPP connections) as well. This policy should be an extension of an overall organizational policy regarding the protection of information resources in the organization. For a firewall to be successful, the service access policy must be realistic and sound and should be

drafted before implementing a firewall. A realistic policy is one that provides a balance between protecting the network from known risks, while still providing users access to network resources. If a firewall system denies or restricts services, it usually requires the strength of the service access policy to prevent the firewall's access controls from being modified on an ad hoc basis. Only a management-backed, sound policy can provide this.

A firewall can implement a number of service access policies, however a typical policy may be to allow no access to a site from the Internet, but allow access from the site to the Internet. Another typical policy would be to allow some access from the Internet, but perhaps only to selected systems such as information servers and e-mail servers. Firewalls often implement service access policies that allow some user access from the Internet to selected internal hosts, but this access would be granted only if necessary and only if it could be combined with advanced authentication.

5.2.0.2 FIREWALL DESIGN POLICY

The firewall design policy is specific to the firewall. It defines the rules used to implement the service access policy. One cannot design this policy in a vacuum isolated from understanding issues such as firewall capabilities and limitations, and threats and vulnerabilities associated with TCP/IP. Firewalls generally implement one of two basic design policies:

- permit any service unless it is expressly denied, and
- deny any service unless it is expressly permitted.

A firewall that implements the first policy allows all services to pass into the site by default, with the exception of those services that the service access policy has identified as disallowed. A firewall that implements the second policy denies all services by default, but then passes those services that have been identified as allowed. This second policy follows the classic access model used in all areas of information security.

The first policy is less desirable, since it offers more avenues for getting around the firewall, e.g., users could access new services currently not denied by the policy (or even addressed by the policy) or run denied services at non-standard TCP/UDP ports that aren't denied by the policy. Certain services such as X Windows, FTP, Archie, and RPC cannot be filtered easily [Chap92], [Ches94], and are better accommodated by a firewall that implements the first policy. The second policy is stronger and safer, but it is more difficult to implement and may impact users more in that certain services such as those just mentioned may have to be blocked or restricted more heavily.

The relationship between the high level service access policy and its lower level counterpart is reflected in the discussion above. This relationship exists because the implementation of the service access policy is so heavily dependent upon the capabilities and limitations of the firewall system, as well as the inherent security problems associated with the wanted Internet services. For example, wanted services defined in the service access policy may have to be denied if the inherent security problems in these services cannot be effectively controlled by the lower level policy and if the security of the network takes precedence over other factors. On the other hand, an organization that is heavily dependent on these services to meet its mission may have to accept higher risk and allow access to these services. This relationship between the service access policy and its lower level counterpart allows for an iterative process in defining both, thus producing the realistic and sound policy initially described.

The service access policy is the most significant component of the four described here. The other three components are used to implement and enforce the policy. (And as

noted above, the service access policy should be a reflection of a strong overall organization security policy.) The effectiveness of the firewall system in protecting the network depends on the type of firewall implementation used, the use of proper firewall procedures, and the service access policy.

5.2.1 Advanced Authentication

Sections 1.3,1.3.1, and 1.3.2 describe incidents on the Internet that have occurred in part due to the weaknesses associated with traditional passwords. For years, users have been advised to choose passwords that would be difficult to guess and to not reveal their passwords. However, even if users follow this advice (and many do not), the fact that intruders can and do monitor the Internet for passwords that are transmitted in the clear has rendered traditional passwords obsolete.

Advanced authentication measures such as smartcards, authentication tokens, biometrics, and software-based mechanisms are designed to counter the weaknesses of traditional passwords. While the authentication techniques vary, they are similar in that the passwords generated by advanced authentication devices cannot be reused by an attacker who has monitored a connection. Given the inherent problems with passwords on the Internet, an Internet-accessible firewall that does not use or does not contain the hooks to use advanced authentication makes little sense.

Some of the more popular advanced authentication devices in use today are called one-time password systems. A smartcard or authentication token, for example, generates a response that the host system can use in place of a traditional password. Because the token or card works in conjunction with software or hardware on the host, the generated response is unique for every login. The result is a one-time password that, if monitored, cannot be reused by an intruder to gain access to an account. [NIST94a] and [NIST91a] contain more detail on advanced authentication devices and measures.

Since firewalls can centralize and control site access, the firewall is the logical place for the advanced authentication software or hardware to be located. Although advanced authentication measures could be used at each host, it is more practical and manageable to centralize the measures at the firewall. Figure above illustrates that a site without a firewall using advanced authentication permits unauthenticated application traffic such as TELNET or FTP directly to site systems. If the hosts do not use advanced authentication, then intruders could attempt to crack passwords or could monitor the network for login sessions that would include the passwords. Figure above also shows a site with a firewall using advanced authentication, such that TELNET or FTP sessions originating from the Internet to site systems must pass the advanced authentication before being permitted to the site systems. The site systems may still require static passwords before permitting access, however these passwords would be immune from exploitation, even if the passwords are monitored, as long as the advanced authentication measures and other firewall components prevent intruders from penetrating or bypassing the firewall.

5.3 Packet Filtering

IP packet filtering is done usually using a packet filtering router designed for filtering packets as they pass between the router's interfaces. A packet filtering router usually can filter IP packets based on some or all of the following fields:

- source IP address,
- destination IP address,
- TCP/UDP source port, and
- TCP/UDP destination port.

Not all packet filtering routers currently filter the source TCP/UDP port, however more vendors are starting to incorporate this capability. Some routers examine which of the router's network interfaces a packet arrived at, and then use this as an additional filtering criterion. Some UNIX hosts provide packet filtering capability, although most do not.

Filtering can be used in a variety of ways to block connections from or to specific hosts or networks, and to block connections to specific ports. A site might wish to block connections from certain addresses, such as from hosts or sites that it considers to be hostile or untrustworthy. Alternatively, a site may wish to block connections from all addresses external to the site (with certain exceptions, such as with SMTP for receiving e-mail).

Adding TCP or UDP port filtering to IP address filtering results in a great deal of flexibility. Recall from Chapter 1 that servers such as the TELNET daemon reside usually at specific ports, such as port 23 for TELNET. If a firewall can block TCP or UDP connections to or from specific ports, then one can implement policies that call for certain types of connections to be made to specific hosts, but not other hosts. For example, a site may wish to block all incoming connections to all hosts except for several firewalls-related systems. At those systems, the site may wish to allow only specific services, such as SMTP for one system and TELNET or FTP connections to another system. With filtering on TCP or UDP ports, this policy can be implemented in a straightforward fashion by a packet filtering router or by a host with packet filtering capability.

As an example of packet filtering, consider a policy to allow only certain connections to a network of address 123.4.*.*. TELNET connections will be allowed to only one host, 123.4.5.6, which may be the site's TELNET application gateway, and SMTP connections will be allowed to two hosts, 123.4.5.7 and 123.4.5.8, which may be the site's two electronic mail gateways. NNTP (Network News Transfer Protocol) is allowed only from the site's NNTP feed system, 129.6.48.254, and only to the site's NNTP server, 123.4.5.9, and NTP (Network Time Protocol) is allowed to all hosts. All other services and packets are to be blocked. An example of the rule set would be as follows:

he first rule allows TCP packets from any source address and port greater than 1023 on the Internet to the destination address of 123.4.5.6 and port of 23 at the site. Port 23 is the port associated with the TELNET server, and all TELNET clients should have unprivileged source ports of 1024 or higher. The second and third rules work in a similar fashion, except packets to destination addresses 123.4.5.7 and 123.4.5.8, and port 25 for SMTP, are permitted. The fourth rule permits packets to the site's NNTP server, but only from source address 129.6.48.254 to destination address 123.4.5.9 and port 119 (129.6.48.254 is the only NNTP server that the site should receive news from, thus access to the site for NNTP is restricted to only that system). The fifth rule permits NTP traffic, which uses UDP as opposed to TCP, from any source to any destination address at the site. Finally, the sixth rule denies all other packets - if this rule weren't present, the router may or may not deny all subsequent packets. This is a very basic example of packet filtering. Actual rules permit more complex filtering and greater flexibility.

5.3.0 Which Protocols to Filter

The decision to filter certain protocols and fields depends on the network access policy, i.e., which systems should have Internet access and the type of access to permit. The following services are inherently vulnerable to abuse and are usually blocked at a firewall from entering or leaving the site [Chap92], [Garf92]:

- tftp, port 69, trivial FTP, used for booting diskless workstations, terminal servers and routers, can also be used to read any file on the system if set up incorrectly,

- X Windows, OpenWindows, ports 6000+, port 2000, can leak information from X window displays including all keystrokes,
- RPC, port 111, Remote Procedure Call services including NIS and NFS, which can be used to steal system information such as passwords and read and write to files, and
- rlogin, rsh, and rexec, ports 513, 514, and 512, services that if improperly configured can permit unauthorized access to accounts and commands.

Other services, whether inherently dangerous or not, are usually filtered and possibly restricted to only those systems that need them. These would include:

- TELNET, port 23, often restricted to only certain systems,
- FTP, ports 20 and 21, like TELNET, often restricted to only certain systems,
- SMTP, port 25, often restricted to a central e-mail server,
- RIP, port 520, routing information protocol, can be spoofed to redirect packet routing,
- DNS, port 53, domain names service zone transfers, contains names of hosts and information about hosts that could be helpful to attackers, could be spoofed,
- UUCP, port 540, UNIX-to-UNIX CoPy, if improperly configured can be used for unauthorized access,
- NNTP, port 119, Network News Transfer Protocol, for accessing and reading network news, and gopher, http (for Mosaic), ports 70 and 80, information servers and client programs for gopher and WWW clients, should be restricted to an application gateway that contains proxy services.

While some of these services such as TELNET or FTP are inherently risky, blocking access to these services completely may be too drastic a policy for many sites. Not all systems, though, generally require access to all services. For example, restricting TELNET or FTP access from the Internet to only those systems that require the access can improve security at no cost to user convenience. Services such as NNTP may seem to pose little threat, but restricting these services to only those systems that need them helps to create a cleaner network environment and reduces the likelihood of exploitation from yet-to-be-discovered vulnerabilities and threats.

5.3.1 Problems with Packet Filtering Routers

Packet filtering routers suffer from a number of weaknesses, as described in [Chap92]. Packet filtering rules are complex to specify and usually no testing facility exists for verifying the correctness of the rules (other than by exhaustive testing by hand). Some routers do not provide any logging capability, so that if a router's rules still let dangerous packets through, the packets may not be detected until a break-in has occurred.

Often times, exceptions to rules need to be made to allow certain types of access that normally would be blocked. But, exceptions to packet filtering rules sometimes can make the filtering rules so complex as to be unmanageable. For example, it is relatively straightforward to specify a rule to block all inbound connections to port 23 (the TELNET server). If exceptions are made, i.e., if certain site systems need to accept TELNET connections directly, then a rule for each system must be added. Sometimes the addition of certain rules may complicate the entire filtering scheme. As noted previously, testing a complex set of rules for correctness may be so difficult as to be impractical.

Some packet filtering routers do not filter on the TCP/UDP source port, which can make the filtering rule set more complex and can open up ``holes'' in the filtering scheme. [Chap92] describes such a problem with sites that wish to allow inbound and outbound SMTP connections. As described in section , TCP connections include a source and destination port. In the case of a system initiating an SMTP connection to a server, the source port would be a randomly chosen port at or above 1024 and the destination port

would be 25, the port that the SMTP server “listens” at. The server would return packets with source port of 25 and destination port equal to the randomly-chosen port at the client. If a site permits both inbound and outbound SMTP connections, the router must allow destination ports and source ports > 1023 in both directions. If the router can filter on source port, it can block all packets coming into the site that have a destination port > 1023 and a source port other than 25. Without the ability to filter on source port, the router must permit connections that use source and destination ports > 1024.

Users could conceivably run servers at ports > 1023 and thus get “around” the filtering policy (i.e., a site system's telnet server that normally listens at port 23 could be told to listen at port 9876 instead; users on the Internet could then telnet to this server even if the router blocks destination port 23).

Another problem is that a number of RPC (Remote Procedure Call) services are very difficult to filter effectively because the associated servers listen at ports that are assigned randomly at system startup. A service known as portmapper maps initial calls to RPC services to the assigned service numbers, but there is no such equivalent for a packet filtering router. Since the router cannot be told which ports the services reside at, it isn't possible to block completely these services unless one blocks all UDP packets (RPC services mostly use UDP). Blocking all UDP would block potentially necessary services such as DNS. Thus, blocking RPC results in a dilemma.

Packet filtering routers with more than two interfaces sometimes do not have the capability to filter packets according to which interface the packets arrived at and which interface the packet is bound for. Filtering inbound and outbound packets simplifies the packet filtering rules and permits the router to more easily determine whether an IP address is valid or being spoofed. Routers without this capability offer more impediments to implementing filtering strategies.

Related to this, packet filtering routers can implement both of the design policies discussed in section 2.4.1. A rule set that is less flexible, i.e., that does not filter on source port or on inbound and outbound interfaces, reduces the ability of the router to implement the second and more stringent policy, deny all services except those expressly permitted, without having to curtail the

types of services permitted through the router. For example, problematic services such as those that are RPC-based become even more difficult to filter with a less-flexible rule set; no filtering on source port forces one to permit connections between ports > 1023. With a less-flexible rule set, the router is less able to express a stringent policy, and the first policy, permit all services except those expressly permitted, is usually followed.

Readers are advised to consult [Chap92], which provides a concise overview of packet filtering and associated problems. While packet filtering is a vital and important tool, it is very important to understand the problems and how they can be addressed.

5.3.1.0 APPLICATION GATEWAYS

To counter some of the weaknesses associated with packet filtering routers, firewalls need to use software applications to forward and filter connections for services such as TELNET and FTP. Such an application is referred to as a proxy service, while the host running the proxy service is referred to as an application gateway. Application gateways and packet filtering routers can be combined to provide higher levels of security and flexibility than if either were used alone.

As an example, consider a site that blocks all incoming TELNET and FTP connections using a packet filtering router. The router allows TELNET and FTP packets to go to one

host only, the TELNET/FTP application gateway. A user who wishes to connect inbound to a site system would have to connect first to the application gateway, and then to the destination host, as follows:

- a user first telnets to the application gateway and enters the name of an internal host,
- the gateway checks the user's source IP address and accepts or rejects it according to any access criteria in place,
- the user may need to authenticate herself (possibly using a one-time password device),
- the proxy service creates a TELNET connection between the gateway and the internal host,
- the proxy service then passes bytes between the two connections, and
- the application gateway logs the connection.

This example points out several benefits to using proxy services. First, proxy services allow only those services through for which there is a proxy. In other words, if an application gateway contains proxies for FTP and TELNET, then only FTP and TELNET may be allowed into the protected subnet, and all other services are completely blocked. For some sites, this degree

of security is important, as it guarantees that only those services that are deemed ``trustworthy'' are allowed through the firewall. It also prevents other untrusted services from being implemented behind the backs of the firewall administrators.

Another benefit to using proxy services is that the protocol can be filtered. Some firewalls, for example, can filter FTP connections and deny use of the FTP put command, which is useful if one wants to guarantee that users cannot write to, say, an anonymous FTP server.

Application gateways have a number of general advantages over the default mode of permitting application traffic directly to internal hosts. These include:

- information hiding, in which the names of internal systems need not necessarily be made known via DNS to outside systems, since the application gateway may be the only host whose name must be made known to outside systems,
- robust authentication and logging, in which the application traffic can be pre-authenticated before it reaches internal hosts and can be logged more effectively than if logged with standard host logging,
- cost-effectiveness, because third-party software or hardware for authentication or logging need be located only at the application gateway, and
- less-complex filtering rules, in which the rules at the packet filtering router will be less complex than they would if the router needed to filter application traffic and direct it to a number of specific systems. The router need only allow application traffic destined for the application gateway and reject the rest.

A disadvantage of application gateways is that, in the case of client-server protocols such as TELNET, two steps are required to connect inbound or outbound. Some application gateways require modified clients, which can be viewed as a disadvantage or an advantage, depending on whether the modified clients make it easier to use the firewall. A TELNET application gateway would not necessarily require a modified TELNET client, however it would require a modification in user behavior: the user has to connect (but not login) to the firewall as opposed to connecting directly to the host. But a modified TELNET client could make the firewall transparent by permitting a user to specify the destination system (as opposed to the firewall) in the TELNET command. The firewall

would serve as the route to the destination system and thereby intercept the connection, and then perform additional steps as necessary such as querying for a one-time password. User behavior stays the same, however at the price of requiring a modified client on each system.

In addition to TELNET, application gateways are used generally for FTP and e-mail, as well as for X Windows and some other services. Some FTP application gateways include the capability to deny put and get command to specific hosts. For example, an outside user who has established an FTP session (via the FTP application gateway) to an internal system such as an anonymous FTP server might try to upload files to the server. The application gateway can filter the FTP protocol and deny all puts to the anonymous FTP server; this would ensure that nothing can be uploaded to the server and would provide a higher degree of assurance than relying only on file permissions at the anonymous FTP server to be set correctly.

An e-mail application gateway serves to centralize e-mail collection and distribution to internal hosts and users. To outside users, all internal users would have e-mail addresses of the form:

user@emailhost

where emailhost is the name of the e-mail gateway. The gateway would accept mail from outside users and then forward mail along to other internal systems as necessary. Users sending e-mail from internal systems could send it directly from their hosts, or in the case where internal system names are not known outside the protected subnet, the mail would be sent to the application gateway, which could then forward the mail to the destination host. Some e-mail gateways use a more secure version of the sendmail program to accept e-mail.

5.3.1.1 CIRCUIT-LEVEL GATEWAYS

[Ches94] defines another firewall component that other authors sometimes include under the category of application gateway. A circuit-level gateway relays TCP connections but does no extra processing or filtering of the protocol. For example, the TELNET application gateway example provided here would be an example of a circuit-level gateway, since once the connection between the source and destination is established, the firewall simply passes bytes between the systems. Another example of a circuit-level gateway would be for NNTP, in which the NNTP server would connect to the firewall, and then internal systems' NNTP clients would connect to the firewall. The firewall would, again, simply pass bytes.

5.4 Firewall Architectures

Firewalls can be configured in a number of different architectures, provided various levels of security at different costs of installation and operation. Organizations should match their risk profile to the type of firewall architecture selected. The following sections describe typical firewall architectures and sample policy statements.

5.4.1 Multi-homed host

A multi-homed host is a host (a firewall in this case) that has more than one network interface, with each interface connected to logically and physically separate network segments. A dual-homed host (host with two interfaces) is the most common instance of a multi-homed host.

A dual-homed firewall is a firewall with two network interfaces cards (NICs) with each interface connected to a different network. For instance, one network interface is typically connected to the external or untrusted network, while the other interface is connected to

the internal or trusted network. In this configuration, an important security tenet is not to allow traffic coming in from the untrusted network to be directly routed to the trusted network - the firewall must always act as an intermediary.

Routing by the firewall shall be disabled for a dual-homed firewall so that IP packets from one network are not directly routed from one network to the other.

5.4.2 Screened host
A screened host firewall architecture uses a host (called a bastion host) to which all outside hosts connect, rather than allow direct connection to other, less secure internal hosts. To achieve this, a filtering router is configured so that all connections to the internal network from the outside network are directed towards the bastion host.

If a packet-filtering gateway is to be deployed, then a bastion host should be set up so that all connections from the outside network go through the bastion host to prevent direct Internet connection between the ORGANIZATION network and the outside world.

5.4.3 Screened subnet
The screened subnet architecture is essentially the same as the screened host architecture, but adds an extra strata of security by creating a network which the bastion host resides (often called a perimeter network) which is separated from the internal network.

A screened subnet will be deployed by adding a perimeter network in order to separate the internal network from the external. This assures that if there is a successful attack on the bastion host, the attacker is restricted to the perimeter network by the screening router that is connected between the internal and perimeter network.

5.5 Types of Firewalls
There are different implementations of firewalls, which can be arranged in different ways. The various firewall implementations are discussed below.

5.5.0 Packet Filtering Gateways
Packet filtering firewalls use routers with packet filtering rules to grant or deny access based on source address, destination address and port. They offer minimum security but at a very low cost, and can be an appropriate choice for a low risk environment. They are fast, flexible, and transparent. Filtering rules are not often easily maintained on a router, but there are tools available to simplify the tasks of creating and maintaining the rules.

Filtering gateways do have inherent risks including:

- The source and destination addresses and ports contained in the IP packet header are the only information that is available to the router in making decision whether or not to permit traffic access to an internal network.
- They don't protect against IP or DNS address spoofing.
- An attacker will have a direct access to any host on the internal network once access has been granted by the firewall.
- Strong user authentication isn't supported with some packet filtering gateways.
- They provide little or no useful logging.

5.5.1 Application Gateways
An application gateway uses server programs (called proxies) that run on the firewall. These proxies take external requests, examine them, and forward legitimate requests to the internal host that provides the appropriate service. Application gateways can support functions such as user authentication and logging.

Because an application gateway is considered as the most secure type of firewall, this configuration provides a number of advantages to the medium-high risk site:

- The firewall can be configured as the only host address that is visible to the outside network, requiring all connections to and from the internal network to go through the firewall.
- The use of proxies for different services prevents direct access to services on the internal network, protecting the enterprise against insecure or misconfigured internal hosts.
- Strong user authentication can be enforced with application gateways.
- Proxies can provide detailed logging at the application level.

Application level firewalls should be configured such that out-bound network traffic appears as if the traffic had originated from the firewall (i.e. only the firewall is visible to outside networks). In this manner, direct access to network services on the internal network is not allowed. All incoming requests for different network services such as Telnet, FTP, HTTP, RLOGIN, etc., regardless of which host on the internal network will be the final destination, must go through the appropriate proxy on the firewall.

Applications gateways require a proxy for each service, such as FTP, HTTP, etc., to be supported through the firewall. When a service is required that is not supported by a proxy, an organization has three choices:

Deny the service until the firewall vendor has developed a secure proxy - This is the preferred approach, as many newly introduced Internet services have unacceptable vulnerabilities.

Develop a custom proxy - This is a difficult task and should be undertaken only by very sophisticated technical organizations.

Pass the service through the firewall - Using what are typically called “plugs,” most application gateway firewalls allow services to be passed directly through the firewall with only a minimum of packet filtering. This can limit some of the vulnerability but can result in compromising the security of systems behind the firewall.

Low Risk

When an in-bound Internet service not supported by a proxy is required to pass through the firewall, the firewall administrator shall define the configuration or plug that will allow the required service. When a proxy is available from the firewall vendor, the plug must be disabled and the proxy made operative.

Medium-high Risk

All in-bound Internet services must be processed by proxy software on the firewall. If a new service is requested, that service will not be made available until a proxy is available from the firewall vendor and tested by the firewall administrator. A custom proxy can be developed in-house or by other vendors only when approved by the CIO.

5.5.2 Hybrid or Complex Gateways

Hybrid gateways combine two or more of the above firewall types and implement them in series rather than in parallel. If they are connected in series, then the overall security is enhanced; on the other hand, if they are connected in parallel, then the network security perimeter will be only as secure as the least secure of all methods used. In medium to high-risk environments, a hybrid gateway may be the ideal firewall implementation.

4	recommended choice
3	effective option
2	acceptable
1	minimal security
0	unacceptable

Firewall Architecture (if any one of these is being implemented)	High Risk Environment e.g. Hospital	Medium Risk Environment e.g. University	Low Risk Environment e.g. florist shop
Packet filtering	0	1	4
Application Gateways	3	4	2
Hybrid Gateways	4	3	2

5.5.3 Firewall Issues

5.5.3.0 AUTHENTICATION

Router-based firewalls don't provide user authentication. Host-based firewalls can provide these kinds of authentication:

Username/password: This provides the lowest level of protection, because the information can be sniffed off the network or shoulder-surfed.

One-time passwords: One-time passwords using software or hardware tokens, generate a new password for each session. This means that old passwords cannot be reused if they are sniffed or otherwise borrowed or stolen.

Digital Certificates: Digital certificates use a certificate generated using public key encryption.

5.5.3.1 ROUTING VERSUS FORWARDING

A clearly defined policy has to be written as to whether or not the firewall will act as a router or a forwarder of Internet packets. This is trivial in the case of a router that acts as a packet filtering gateway: the firewall (router in this case) has no option but to route packets. Applications gateway firewalls should generally not be configured to route any traffic between the external interface and the internal network interface, since this could bypass security controls. All external to internal connections should go through the application proxies.

5.5.3.2 SOURCE ROUTING

Source routing is a routing mechanism whereby the path to a target machine is determined by the source, rather than by intermediate routers. Source routing is mostly used for debugging network problems but could also be used to attack a host. If an attacker has knowledge of some trust relationship between your hosts, source routing can be used to make it appear that the malicious packets are coming from a trusted host. Therefore, because of this security threat, a packet filtering router can easily be configured to reject packets containing source route option. Thus, a site that wishes to avoid the problem of source routing entirely would write a policy similar to the following:

5.5.3.3 IP SPOOFING

IP spoofing is when an attacker masquerades his machine as a host on the target's network (i.e. fooling a target machine that packets are coming from a trusted machine on the target's internal network). Policy regarding packet routing has to be clearly written so that they will be handled accordingly if there is a security problem. It is necessary that authentication based on source address be combined with other security scheme to protect against IP spoofing attacks.

5.5.3.4 PASSWORD SNIFFING

Password sniffing can be very simple and done with considerable ease. Encryption of passwords is a system-to-system capability IF the operating systems are a matched pair (same OS, usually same version) then the passwords are usually encrypted in a session. Perhaps the biggest problem is that users tend to use the same user ID and password on all systems they may use in a network. If one system is compromised then this leads to a compromise of all systems.

There are many public domain password "grabbers" available on the Internet. These programs are free and readily available.
Example:

> http://www.geocities.com/SiliconValley/Bay/4854/snoopie.zip
> DOS-based TCP-specific password grabber and filter software
> Download, unpack, load drivers, grab passwords off internal nets

Any protocol analyzer that is used by system administrators can grab passwords (there are over 75 of them on the market, many software-only). A good programmer can write their own in about 45 minutes. Password crypto is frequently no help: it only works when OS-to-OS are the same OS's (e.g. NT-to-NT, UNIX-to-UNIX)

```
Flags:        0x00
  Status:       0x00
  Packet Length:85
  Timestamp:    15:35:27.247
  Filter:       IP
Ethernet Header
  Destination:  00:00:0c:19:99:49
  Source:       00:05:a8:00:84:3b
  Protocol Type:0x0800  IP
IP Header - Internet Protocol Datagram
  Version:              4
  Header Length:        5
  Precedence:           0
  Type of Service:      %000
  Unused:               %00
  Total Length:         67
  Identifier:           19528
  Fragmentation Flags:  %010  Do Not Fragment
  Fragment Offset:      0
  Time To Live:         255
  IP Type:              0x06  TCP
  Header Checksum:      0xdde2
  Source IP Address:    192.246.254.153
  Dest. IP Address:     129.170.16.79
  No Internet Datagram Options
TCP - Transport Control Protocol
  Source Port:      2050
  Destination Port: 21  FTP - File Transfer Protocol
  Sequence Number:  1241405969
  Ack Number:       1629760546
  Offset:           5
  Reserved:         %000000
```

```
  Code:             %011000
                Ack is valid
                Push Request
  Window:           17688
  Checksum:         0xf86c
  Urgent Pointer:   0
  No TCP Options
FTP Control - File Transfer Protocol
  FTP Command:      0x50415353  (PASS)  Password
  Password:
  rmasey@network-1   72 6d 61 73 65 79 40 6e 65 74 77 6f 72 6b 2d 31
  .com               2e 63 6f 6d
  Newline Sequence: 0x0d0a
Frame Check Sequence:  0x06c1fd4a
```

5.5.3.5 DNS AND MAIL RESOLUTION

On the Internet, the Domain Name Service provides the mapping and translation of domain names to IP addresses, such as mapping server1.acme.com to 123.45.67.8. Some firewalls can be configured to run as a primary, secondary, or caching DNS server.

Deciding how to manage DNS services is generally not a security decision. Many organizations use a third party, such as an Internet Service Provider, to manage their DNS. In this case, the firewall can be used as a DNS caching server, improving performance but not requiring your organization to maintain its own DNS database.

If the organization decides to manage its own DNS database, the firewall can (but doesn't have to) act as the DNS server. If the firewall is to be configured as a DNS server (primary, secondary, or caching), it is necessary that other security precautions be in place. One advantage of implementing the firewall as a DNS server is that it can be configured to hide the internal host information of a site. In other words, with the firewall acting as a DNS server, internal hosts get an unrestricted view of both internal and external DNS data. External hosts, on the other hand, do not have access to information about internal host machines. To the outside world, all connections to any host in the internal network will appear to have originated from the firewall. With the host information hidden from the outside, an attacker will not know the host names and addresses of internal hosts that offer service to the Internet.

A security policy for DNS hiding might state:

If the firewall is to run as a DNS server, then the firewall must be configured to hide information about the network so that internal host data are not advertised to the outside world.

The best type of a network security setup is one that is multi tiered or layered. This type of a setup allows for built in redundancy.

5.5.4 Firewall Administration

A firewall, like any other network device, has to be managed by someone. Security policy should state who is responsible for managing the firewall.

Two firewall administrators (one primary and secondary) shall be designated by the Chief Information Security Officer (or other manager,) and shall be responsible for the upkeep of the firewall. The primary administrator shall make changes to the firewall and the secondary shall only do so in the absence of the former so that there is no simultaneous or contradictory access to the firewall.

Each firewall administrator shall provide their home phone number, pager number, cellular phone number and other numbers or codes in which they can be contacted when support is required.

5.5.4.0 Qualification of the Firewall Administrator
Two experienced people are generally recommended for the day to day administration of the firewall. In this manner availability of the firewall administrative function is largely insured. It should be required that information about each firewall administrator be written down so that he may contacting is possible in case of a problem.

Security of a site is crucial to the day to day business activity of an organization. It is therefore required that the administrator of the firewall have a sound understanding of network concepts and implementation. For instance, since most firewalls are TCP/IP based, a thorough understanding of this protocol is compulsory.

An individual that is assigned the task of firewall administration must have a good hands-on experience with networking concepts, design, and implementation so that the firewall is configured correctly and administered properly. Firewall administrators should receive periodic training on the firewalls in use and in network security principals and practices.

5.5.4.1 Remote Firewall Administration
Firewalls are the first line of defense visible to an attacker. By design, firewalls are generally difficult to attack directly, causing attackers to often target the administrative accounts on a firewall. The username/password of administrative accounts must be strongly protected.

The most secure method of protecting against this form of attack is to have strong physical security around the firewall host and to only allow firewall administration from an attached terminal. However, operational concerns often dictate that some form of remote access for firewall administration be supported. In no case should remote access to the firewall be supported over untrusted networks without some form of strong authentication. In addition, to prevent eavesdropping, session encryption should be used for remote firewall connections.

Low

Any remote access over untrusted networks to the firewall for administration must use strong authentication, such as one time passwords and/or hardware tokens.

Medium

The preferred method for firewall administration is directly from the attached terminal. Physical access to the firewall terminal is limited to the firewall administrator and backup administrator.

Where remote access for firewall administration must be allowed, it should be limited to access from other hosts on the ORGANIZATION internal network. Such internal remote access requires the use of strong authentication, such as one time passwords and/or hardware tokens. Remote access over untrusted networks such as the Internet requires end to end encryption and strong authentication to be employed.

High

All firewall administration must be performed from the local terminal - no access to the firewall operating software is permitted via remote access. Physical access to the firewall terminal is limited to the firewall administrator and backup administrator.

5.5.4.2 User Accounts
Firewalls should never be used as general purpose servers. The only user accounts on the firewall should be those of the firewall administrator and any backup administrators. In addition, only these administrators should have privileges for updating system executables or other system software.

> *Only the firewall administrator and backup administrators will be given user accounts on the ORGANIZATION firewall. Any modification of the firewall system software must be done by the firewall administrator or backup administrator and requires approval of the Network Services Manager*

5.5.4.3 Firewall Backup
To support recovery after failure or natural disaster, a firewall like any other network host has to have some policy defining system backup. Data files as well as system configuration files need to be have some backup plan in case of firewall failure.

The firewall (system software, configuration data, database files, etc.) must be backed up daily, weekly, and monthly so that in case of system failure, data and configuration files can be recovered. Backup files should be stored securely on a read-only media so that data in storage is not over-written inadvertently and locked up so that the media is only accessible to the appropriate personnel.

Another backup alternative would be to have another firewall configured as one already deployed and kept safely so that in case there is a failure of the current one, this backup firewall would simply be turned on and used as the firewall while the previous is undergoing a repair.

> *At least one firewall shall be configured and reserved (not-in-use) so that in case of a firewall failure, this backup firewall can be switched in to protect the network.*

5.5.4.4 System Integrity
To prevent unauthorized modifications of the firewall configuration, some form of integrity assurance process should be used. Typically, checksums, cyclic redundancy checks, or cryptographic hashes are made from the runtime image and saved on protected media. Each time the firewall configuration has been modified by an authorized individual (usually the firewall administrator), it is necessary that the system integrity online database be updated and saved onto a file system on the network or removable media. If the system integrity check shows that the firewall configuration files have been modified, it will be known that the system has been compromised.

> *The firewall's system integrity database shall be updated each time the firewall is configuration is modified. System integrity files must be stored on read only media or off-line storage. System integrity shall be checked on a regular basis on the firewall in order for the administrator to generate a listing of all files that may have been modified, replaced, or deleted.*

5.5.4.5 Documentation

It is important that the operational procedures for a firewall and its configurable parameters be well documented, updated, and kept in a safe and secure place. This assures that if a firewall administrator resigns or is otherwise unavailable, an experienced individual can read the documentation and rapidly pick up the administration of the firewall. In the event of a break-in such documentation also supports trying to recreate the events that caused the security incident.

5.5.4.6 Physical Firewall Security

Physical access to the firewall must be tightly controlled to preclude any authorized changes to the firewall configuration or operational status, and to eliminate any potential for monitoring firewall activity. In addition, precautions should be taken to assure that proper environment alarms and backup systems are available to assure the firewall remains online.

> *The ORGANIZATION firewall should be located in an controlled environment, with access limited to the Network Services Manager, the firewall administrator, and the backup firewall administrator.*

The room in which the firewall is to be physically located must be equipped with heat, air-conditioner, and smoke alarms to assure the proper working order of the room. The placement and recharge status of the fire extinguishers shall be checked on a regular basis. If uninterruptible power service is available to any Internet-connected systems, such service should be provided to the firewall as well.

5.5.4.7 Firewall Incident Handling

Incident reporting is the process whereby certain anomalies are reported or logged on the firewall. A policy is required to determine what type of report to log and what to do with the generated log report. This should be consistent with Incident Handling policies. The following policies are appropriate to all risk environments.

> *The firewall shall be configured to log all reports on daily, weekly, and monthly bases so that the network activity can be analyzed when needed.*
>
> *Firewall logs should be examined on a weekly basis to determine if attacks have been detected.*
>
> *The firewall administrator shall be notified at anytime of any security alarm by email, pager, or other means so that he may immediately respond to such alarm.*
>
> *The firewall shall reject any kind of probing or scanning tool that is directed to it so that information being protected is not leaked out by the firewall. In a similar fashion, the firewall shall block all software types that are known to present security threats to a network (such as Active X and Java) to better tighten the security of the network.*

5.5.4.8 Restoration of Services

Once an incident has been detected, the firewall may need to be brought down and reconfigured. If it is necessary to bring down the firewall, Internet service should be disabled or a secondary firewall should be made operational - internal systems should not be connected to the Internet without a firewall. After being reconfigured, the firewall must be brought back into an operational and reliable state. Policies for restoring the firewall to a working state when a break-in occurs are needed.

> *In case of a firewall break-in, the firewall administrator(s) are responsible for reconfiguring the firewall to address any vulnerabilities that were exploited. The firewall shall be restored to the state it was before the break-in so that the*

network is not left wide open. While the restoration is going on, the backup firewall shall be deployed.

5.5.4.9 Upgrading the firewall

It is often necessary that the firewall software and hardware components be upgraded with the necessary modules to assure optimal firewall performance. The firewall administrator should be aware of any hardware and software bugs, as well as firewall software upgrades that may be issued by the vendor. If an upgrade of any sort is necessary, certain precautions must be taken to continue to maintain a high level of operational security. Sample policies that should be written for upgrades may include:

To optimize the performance of the firewall, all vendor recommendations for processor and memory capacities shall be followed.

The firewall administrator must evaluate each new release of the firewall software to determine if an upgrade is required. All security patches recommended by the firewall vendor should be implemented in a timely manner.

Hardware and software components shall be obtained from a list of vendor-recommended sources. Any firewall specific upgrades shall be obtained from the vendor. NFS shall not be used as a means of obtaining hardware and software components. The use of virus checked CDROM or FTP to a vendor's site is an appropriate method.

The firewall administrator(s) shall monitor the vendor's firewall mailing list or maintain some other form of contact with the vendor to be aware of all required upgrades. Before an upgrade of any of the firewall component, the firewall administrator must verify with the vendor that an upgrade is required. After any upgrade the firewall shall be tested to verify proper operation prior to going operational.

5.5.4.10 Logs and Audit Trails

Most firewalls provide a wide range of capabilities for logging traffic and network events. Some security-relevant event that should be recorded on the firewall's audit trail logs are: hardware and disk media errors, login/logout activity, connect time, use of system administrator privileges, inbound and outbound e-mail traffic, TCP network connect attempts, in-bound and out-bound proxy traffic type.

5.5.4.11 Revision/Update of Firewall Policy

Given the rapid introduction of new technologies, and the tendency for organizations to continually introduce new services, firewall security policies should be reviewed on a regular basis. As network requirements changes, so should security policy.

5.5.4.12 Example General Policies

The following policy statements are only examples. They do not constitute a complete firewall policy, and even if they did, they would not necessarily apply to your organization's environment. The statements are grouped into those applicable to Low-, Medium- and High-Risk environments. Within each category, they are divided into statements targeted toward users, managers and technicians. In general, all organizations would employ at least the Low-Risk policies.

5.5.4.12.0 LOW-RISK ENVIRONMENT POLICIES

User

All users who require access to Internet services must do so by using ORGANIZATION-approved software and Internet gateways.

A firewall has been placed between our private networks and the Internet to protect our systems. Employees must not circumvent the firewall by using modems or network tunneling software to connect to the Internet.

Some protocols have been blocked or redirected. If you have a business need for a particular protocol, you must raise the issue with your manager and the Internet security officer.

Manager

A firewall shall be placed between the ORGANIZATION's network and the Internet to prevent untrusted networks from accessing the ORGANIZATION network. The firewall will be selected by and maintained by the Network Services Manager.

All other forms of Internet access (such as via dial-out modems) from sites connected to the ORGANIZATION wide-area network are prohibited.

All users who require access to Internet services must do so by using ORGANIZATION-approved software and Internet gateways.

Technician

All firewalls should fail to a configuration that denies all services, and require a firewall administrator to re-enable services after a failure.

Source routing shall be disabled on all firewalls and external routers (see section 0).

The firewall shall not accept traffic on its external interfaces that appear to be coming from internal network addresses (see section 0).

The firewall shall provide detailed audit logs of all sessions so that these logs can be reviewed for any anomalies.

Secure media shall be used to store log reports such that access to this media is restricted to only authorized personnel.

Firewalls shall be tested off-line and the proper configuration verified.

The firewall shall be configured to implement transparency for all outbound services. Unless approved by the Network Services manager, all in-bound services shall be intercepted and processed by the firewall.

Appropriate firewall documentation will be maintained on off-line storage at all times. Such information shall include but not be limited to the network diagram, including all IP addresses of all network devices, the IP addresses of relevant hosts of the Internet Service Provider (ISP) such as external news server, router, DNS server, etc. and all other configuration parameters such as packet filter rules, etc. Such documentation shall be updated any time the firewall configuration is changed.

5.5.4.12.1 MEDIUM-RISK ENVIRONMENT POLICIES

User

When you are off-site, you may only access internal systems by using ORGANIZATION-approved one-time passwords and hardware tokens to authenticate yourself to the firewall. Any other means of accessing internal systems is prohibited.

Manager

Strong authentication using ORGANIZATION-approved one-time passwords and hardware tokens is required all remote access to internal systems through the firewall.

The network security policy shall be reviewed on a regular basis (every three months minimum) by the firewall administrator(s) and other top information (security) managers. Where requirements for network connections and services have changed, the security policy shall be updated and approved. If a change is to be made, the firewall administrator shall ensure that the change is implemented and the policy modified.

The details of the ORGANIZATION internal trusted network should not be visible from outside the firewall.

Technician

The firewall will be configured to deny all services not expressly permitted and will be regularly audited and monitored to detect intrusions or misuse.

The firewall shall notify the system administrator in near-real-time of any item that may need immediate attention such as a break-in into the network, little disk space available, or other related messages so that an immediate action could be taken.

The firewall software will run on a dedicated computer - all non-firewall related software, such as compilers, editors, communications software, etc., will be deleted or disabled.

The firewall will be configured to deny all services not expressly permitted and will be regularly audited and monitored to detect intrusions or misuse.

5.5.4.12.2 HIGH-RISK ENVIRONMENT POLICIES

User

All non-business use of the Internet from ORGANIZATION systems is forbidden. All access to Internet services is logged. Employees who violate this policy are subject to disciplinary action.

Your browser has been configured with a list of forbidden sites. Any attempts to access those sites will be reported to your manager.

Manager

All non-business use of the Internet from ORGANIZATION systems is forbidden. All access to Internet services is logged. Employees who violate this policy are subject to disciplinary action.

Technician

All access to Internet services is logged. Summary and exception reports will be prepared for the network and security managers.

5.5.4.13 Firewall Concerns: Management

Purpose	Protocols	What	Why
Email		Users have a single external email address	Does not reveal business info.
	SMTP	A single server or cluster of servers provides email service for organization	Centralized email is easier to maintain. SMTP servers are difficult to configure securely.
	POP3	POP users must use AUTH identification.	Prevents password sniffing.
	IMAP	Groups are encouraged to transition to IMAP.	Better support for travel, encryption.
USENET news	NTTP	blocked at firewall	no business need
WWW	HTTP	directed to www.my.org	Centralized WWW is easier to maintain. WWW servers are difficult to configure securely.
*	all others	routed	

5.5.4.14 Service Policies Examples

Service	*Policy*				*Sample Policy*
	Inside to Outside		*Outside to Inside*		
	Status	*Auth*	*Status*	*Auth*	
FTP	*y*	*n*	*y*	*y*	*FTP access shall be allowed from the internal network to the external. Strong authentication shall be required for FTP access from the outside to the inside.*
Telnet	*y*	*n*	*y*	*y*	*Telnet access shall be allowed from the inside network to the outside network. For the telnet from the outside to the inside network, authentication shall be required.*
Rlogin	*y*	*n*	*y*	*y*	*rlogin to ORGANIZATION hosts from external networks requires written approval from the Network Services Manager and the use of strong authentication.*
HTTP	*y*	*n*	*n*	*n*	*All WWW servers intended for access by external users will be hosted outside the ORGANIZATION firewall. No inbound HTTP will be allowed through the ORGANIZATION firewall.*
SSL	*y*	*n*	*y*	*y*	*Secure Sockets Layer sessions using client side certificates is required when SSL sessions are to be passed through the ORGANIZATION firewall.*
POP3	*n*	*n*	*y*	*n*	*The ORGANIZATION Post Office Protocol server is to be hosted inside the ORGANIZATION firewall. The firewall will pass POP traffic only to the POP server. The use of APOP is required.*
NNTP	*y*	*n*	*n*	*n*	*No external access will be allowed to the NNTP server.*

Real Audio	*n*	*n*	*n*	*n*	*There is currently no business requirement for supporting streaming audio sessions through the ORGANIZATION firewall. Any business units requiring such support should contact the Network Services Manager.*
Lp	*y*	*n*	*n*	*n*	*Inbound lp services are to be disabled at the ORGANIZATION firewall*
finger	*y*	*n*	*n*	*n*	*Inbound finger services are to be disabled at the ORGANIZATION firewall*
gopher	*y*	*n*	*n*	*n*	*Inbound gopher services are to be disabled at the ORGANIZATION firewall*
whois	*y*	*n*	*n*	*n*	*Inbound whois services are to be disabled at the ORGANIZATION firewall*
SQL	*y*	*n*	*n*	*n*	*Connections from external hosts to internal databases must be approved by the Network Services Manager and used approved SQL proxy services.*
Rsh	*y*	*n*	*n*	*n*	*Inbound rsh services are to be disabled at the ORGANIZATION firewall*
Other, such as NFS	*n*	*n*	*n*	*n*	*Access to any other service not mentioned above shall be denied in both direction so that only Internet services we have the need for and we know about are allowed and all others are denied.*

An organization may wish to support some services without using strong authentication. For example, an anonymous FTP server may be used to allow all external users to download open information. In this case, such services should be hosted outside the firewall or on a service network not connected to corporate

networks that contain sensitive data. The table that follows summarizes a method of describing such policy for a service such as FTP.

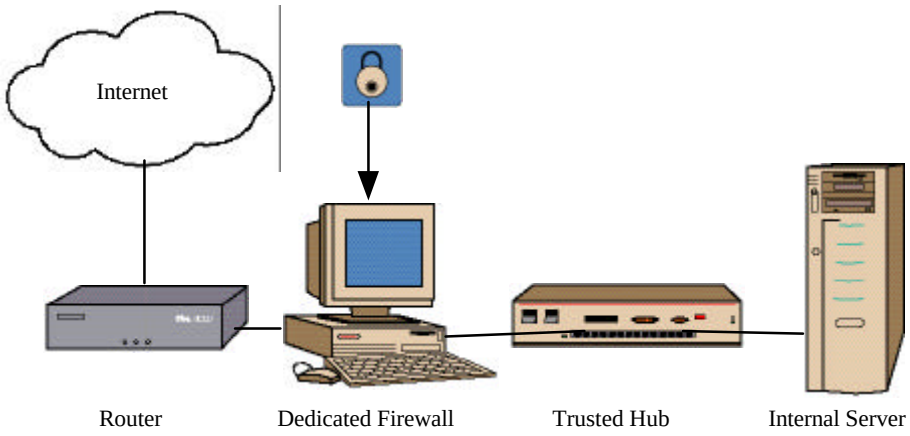
Table 1 - Summarized Security Policy

Policy	*Non-Anonymous FTP service*	*Anonymous FTP service*
Put server machine outside the firewall	*N*	*Y*
Put server machine on the service network	*N*	*Y*
Put server machine on protected network	*Y*	*N*
Put server machine on the firewall itself	*N*	*N*
Server will be accessed by everyone on the Internet	*N*	*Y*

5.5.5 Client and Server Security in Enterprise Networks

5.5.5.0 Historical Configuration of Dedicated Firewall Products

In today's network security firewall marketplace, the most common firewall configuration is the use of a dedicated firewall system between an "untrusted" network and the corporate network, usually referred to as the "trusted" side of the firewall.



5.5.5.1 Advantages and Disadvantages of Dedicated Firewall Systems

A dedicated firewall has distinct performance and security advantages. First off, you gain total performance of the system dedicated to the function of firewall services (if nothing else is on the system, there is nothing else for the firewall software to compete with for CPU access). Second, a dedicated firewall system helps increase security of the firewall itself as the number of privileged users who have access to the firewall system are much less than other systems and are usually carefully screened so that those individuals who do have access to the firewall are in positions of trust within the company. Finally, any other software which runs on a firewall that is NOT the firewall software or the operating environment puts the firewall at risk simply due to failures of the software "killing" the firewall, other software creating system security holes, software bugs and errors in non-firewall

software “opening” up the system in some manner or other such problems. The less amount of software on a firewall, the better for performance and firewall security.

Dedicated firewalls have their disadvantages as well. Many are based on the UNIX operating system or its variants which are not known for their “user friendliness.” While many vendors have strived to put a graphical interface on their firewall products when running under the UNIX environments, most still rely on UNIX properties to help make the firewall work and this requires anywhere from minimal UNIX skills to expert-level UNIX skills to configure and manage the firewall system.

Another problem with UNIX systems as firewalls is the availability of source code for the UNIX environment. While there are valid arguments for such availability, there are as many arguments against as if a “good” consumer can read the source code and discover how something works, so can an “evil” attacker who wants to attack a UNIX-based firewall system or systems being protected in the UNIX environments.

Some of the problems associated with a UNIX firewall have to do with the availability of in-house expertise and the logistics of getting a UNIX system set-up properly to be a firewall system. It is no coincidence that most UNIX-based firewalls require a customized version of the UNIX environment being used to patch and control system security “holes” that may be used by an attacker to gain access. Then there is the definition and management of the UNIX system for firewall operations which usually require UNIX-specific management commands and facilities as well as the “tightening up” of the UNIX environment to close commonly used network and system interfaces. In many UNIX-based firewalls, firewall rule bases require the writing of either UNIX shell scripts or scripts in the *perl* language to provide firewall functionality. While companies who make such products will argue towards their approach, and there is nothing wrong with that, there is a certain amount of UNIX-based work that must happen on any UNIX-based firewall to make it work correctly and to manage the computational environment properly.

Even in the case of non-UNIX dedicated firewall systems, such as FireWall/Plus™ for MS-DOS, there is the non-flexibility of using the system for other system functions. This is a double-edged sword as there is the conflict between the “don’t put anything on the firewall but firewall software” crowd and the “we have to use all equipment to its fullest potential as this is a small site and we can’t afford a dedicated firewall box” crowd. Both have valid points, but true firewall functionality means security first - not last.

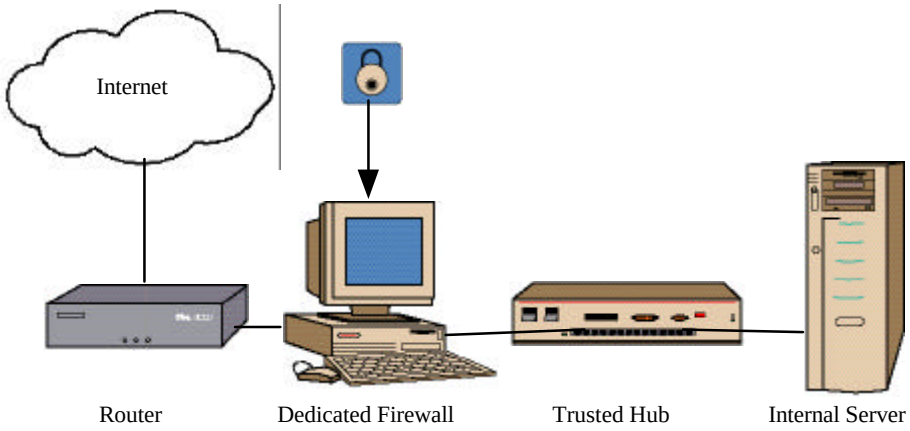
Dedicated firewalls which are, in fact, router systems with filters in them have many of the same concerns as a dedicated firewall running other applications at the same time. Firewall functions are different than routing functions. By putting both functions in the same hardware processor system, either function could “kill” the other function at a maximum or cause problems and security holes at a minimum - just like a firewall which runs other applications at the same time. There are plenty of CERT and CIAC alerts issued over the last few years on router vendors for their firewall filtering failures which were due to bugs or problems in the routing facilities which allowed the firewall function in the router to either be bypassed or breached. Having a dedicated router with screening functions is ONE layer in a properly defined network security set up. Network security means multiple layers of protection and putting all the protection facilities in a singular router/firewall combination means that if the unit is breached, there is an entire trusted network to attack with no other warning or security mechanism.

5.5.5.2 Are Dedicated Firewalls A Good Idea?

Security wise, an emphatic *yes* - for the reasons previously mentioned and plenty more. But, to satisfy tight budgets and management who do not understand the true requirements for security systems, it is more and more common to use a firewall system as a multi-function computer where firewall functionality is one component of the system. But even dedicated security firewalls are not a total network solution - they remain a single level in security management of network environments. True, functional network security must be a layered approach and use different types of security technologies to ensure proper control over data as it moves around any network between systems.

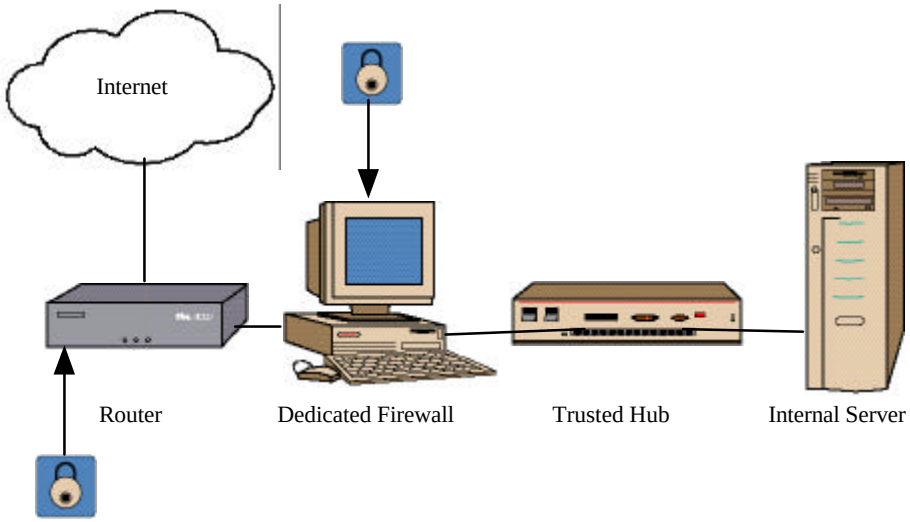
5.5.5.3 Layered Approach to Network Security - How To Do It

As an example, system vulnerability to attack is greater when only a firewall is used with no router filters on an Internet connection (the padlock symbol indicates a security layer function).:



In the above configuration, if an attacker were to get “around” the firewall system, the server is vulnerable to attack from the network.

Adding screening filters for incoming packets into a router adds another layer to the network security architecture:

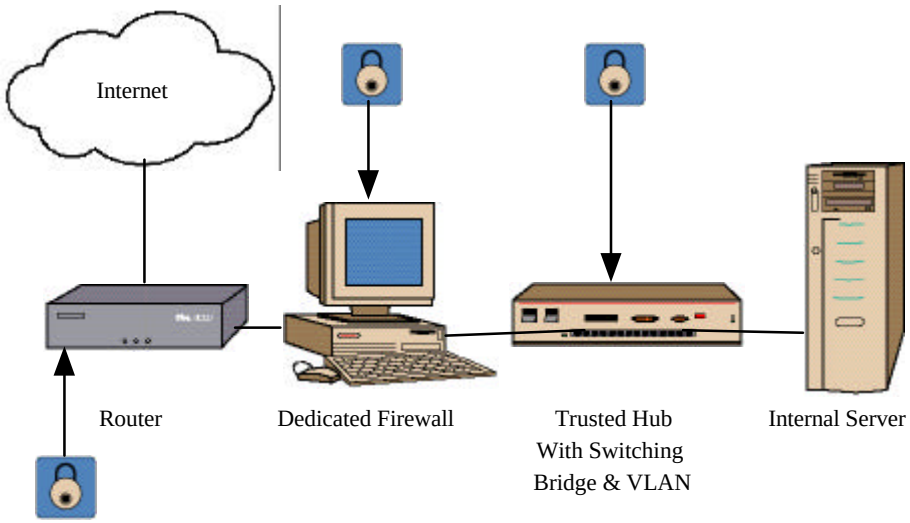


At this point, the security manager would be wise to insert some duplicate security rules into the router filter rule base and the firewall security rule base for some of the more important security functions. This would allow detection of a first-layer breach of the router by security facilities in the firewall. For instance, if a TELNET filter were placed in the router that denied all TELNET access, this would supposedly stop TELNET functions from arriving to the firewall system. If the firewall also had filters in it denying a TELNET connection from the untrusted Internet side of its connections, then if a TELNET connection should arrive, the security manager knows immediately that something very ugly has happened in the router for the TELNET attempt to even reach the firewall and it's time to find out what is going on in the router.

Putting filters in a screening router has the following effects to the security hierarchy:

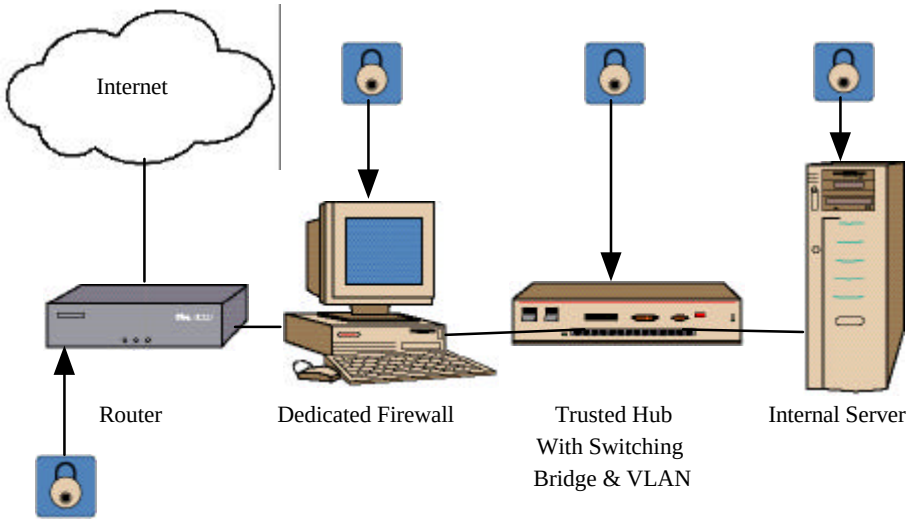
- Pre-screens security threats and dismisses them from the connection path
- Offloads security checking from the firewall except in the case of a failure by the router to properly screen the attempted function
- Offloads packet filtering functions from the firewall
- Allows secondary security exception failure detection by the firewall of a router where the security filter in the router has failed for some reason and still does not allow the security exception condition to reach the trusted network side

Another layer of security is possible by using a switching bridge in the hub to control traffic directions and provide additional layers of packet filtering. By using hub-based virtual local area network (VLAN) software in the switching bridge (this is available from some switching bridge vendors - but not all), the network path is further protected from attackers. This might be configured as follows:



There are situations where using network security firewall software on an active client or server system acts as another security layer in the implementation of a layered network security architecture. This concept, while functionally similar in implementation to the shared system-firewall concepts previously explored, is not the same from a security rule base situation and from a performance situation. Further, this concept is different in that the security threat is lesser in this configuration as it is predisposed that there is a real firewall in the network path BEFORE the system being accessed (running network security firewall software)

that has pre-screened connection facilities coming towards the client or server. Adding server-based network security firewall software allows a final layer of network security prior to reaching the server operating environment:



By putting network security into the corporate environment as a layered methodology, different levels of security (depending on the criticality of a component to the company) are possible throughout the network. Further, while external security is indeed needed and essential, the bulk of network attacks actually happen from internal entities (over 80% in some studies) that actually are a part of the corporate resource list.

In the above configuration, there are at least four layers of network security before the server's operating assets are accessed. This is far superior to a singular network layer solution as is usually implemented via a singular dedicated firewall or through the use of a screening router as the firewall. Additional network security layers may be added via authentication facilities, encryption, digital signatures and other security methods that are used in the various layers of network protocols (including applications). Oddly enough, properly implemented many network security methods may be added in such a manner as to be transparent to the user's activities as long as the user is attempting to access authorized systems and facilities.

With a layered network defense environment the chances of actual network attacks getting to sensitive data are greatly minimized and the opportunities to detect inappropriate security behavior before it reaches a significant asset are greatly improved.

5.5.5.4 Improving Network Security in Layers - From Inside to Outside

Another improvement in the layered network security approach is that of keeping sensitive assets "in" instead of just keeping attackers "out" of asset collections (such as file or database servers). Firewalls and security filtering facilities work not only with incoming requests, but also with outgoing requests. A typical "trusted" attack on a server might be to set up a program which initiates a file transfer from the server to an untrusted entity during off-hours. In this case, many companies might not think anything of the activity as a) they probably are not monitoring for it and b) not many companies think of their systems as voluntarily moving data from the trusted side unassisted by a connection from the untrusted side of a network connection

hierarchy. Proper network security is a bi-directional effort - not just from outside to inside, but inside to outside as well.

5.5.5.5 Operating Systems and Network Software - Implementing Client and Server Security

System security on a client or server system is the function of the following general items:

- ***Operating system security reference monitor*.** The security reference monitor is the main security "traffic cop" for the operating system. It is responsible for taking the defined security rule base in an operating system and providing methods to enforce the security decisions made by the systems and security personnel. For instance, file access may be controlled by disk security facilities, access control lists to directories and files, disk "vaulting" facilities, file encryption, file size constraints, disk "area" security mapping and many other concepts and facilities. These concepts exist for device access, memory access, CPU utilization and, in some operating environments, network protocol access.
- ***Application security facilities.*** In the writing of applications for user access, programmers may implement a variety of security facilities for user and remote system access. These may include user authentication facilities, time-based access modes, implementation of external security packages within the application and many other concepts and facilities. Specific "commercial" packages may implement very sophisticated security facilities, such as major database systems, to control access to data entities stored or accessed by applications.
- ***Physical security.*** On many operating systems, physical access is a method of controlling security facilities. For instance, only access to a specific physical systems console keyboard will allow certain very sensitive actions to take place. Further protection at a physical level might include a console key (made of metal or plastic), locked system access, physical environment (locked room, security facilities via physical room access, electronic cryptolocks, card-key access, console card access, etc.), etc...
- ***Key certificates.*** Many applications and operating systems are starting to implement key certificates in software. These are special license keys that are installed at product installation time that are also locked down to some physical attribute of the computer system to specifically identify a machine. For instance, key certificates may be used for database access programs where the program on a server requires the program on the client to forward its key certification information before any application access to the database can begin.
- ***Network protocols.*** While network protocols do not implement security facilities, as a pretty standard rule, their presence on a system dictate the potential of attack on the system from a network. For instance, if the bulk of network attacks at a site are based on TCP/IP and the only protocol on the system is Novell's IPX, it's pretty hard to attack a system without the protocol the attacker would use and the system being attacked does not have. If the system implements multiple basic protocols (as does Windows-NT with IP, IPX and NetBEUI with the shipped standard versions for clients and servers), then security becomes a greater problem as there are more methods to access the system and, therefore, the greater the chance of a network attack in some form.
- ***System accounting.*** Oddly enough, one of the main detection facilities in security analysis are statistics generated by users, applications, devices, etc. Great security features may be implemented at all levels of an operating system environment, but accounting provides statistical tracking over time. Very good system attacks may be launched "looking" like valid logins or accesses to data.

Using accounting statistics and averaging methods for individual functions will tip off the security professional that someone or something is acting outside the normal operating pattern and deserves attention. Also, attempts to modify the accounting facilities are a sure sign that someone wants to cover their tracks and this should tip off the security team that something unusual and unwanted is going on.

- ***Security Add-ons*.** One item often overlooked are system additions by 3rd party companies that provide additional security facilities to an operating environment. These might include system security management software, encryption systems, key exchange facilities, authentication facilities (such as token card and key certificate management software) and many other items. All of these items still do not address the issues of protocol security, but they do increase the difficulty to attack the operating system environment being protected.

Implementing all these facilities on an operating environment is not without penalty. System performance is degraded as more items are activated. File services are degraded as more information is logged, sorted, alarmed and accessed. Network facilities are degraded as packets are examined for content and connection types. In all, proper system security is a great deal of work, done correctly, and checks and crosschecks are required to ensure system and application integrity. And, system security requires CPU and I/O horsepower - a lot of it when done properly.

5.5.5.6 Operating System Attacks From the Network Resource(s) - More Protocols Are The Norm - and They Are Not Just IP

Network security firewalls provide a “bottleneck” facility at strategic points on the network to prevent wholesale attacks of systems on a network. It’s pretty common practice to put a firewall facility between known troublesome networks such as Internet. Oddly enough, most companies do not implement firewall facilities between different company divisions, “sister” company network connections, customer network connections and other 3rd party or vendor supplied network connections. The funny part is that most of the documented network break-ins are from the non-Internet connections (although the Internet break-ins are accelerating). The other problem is that on practically all corporate networks, the protocol environment is multi-protocol; IP is not all that is used by any stretch of reality. In most established networks, the predominant protocols are Novell’s IPX, LAN Manager/LAN Server/Pathworks NetBEUI and Apple Computer’s AppleTalk. In mainframe environments there is a predominance of SNA-related protocols and in the mid-range environment other protocols such as DECnet, LAT, various hardware-specific protocols and many non-IP protocols. In short, the standard company environment most operating environments must function within are not just IP - they’re a lot of every type of protocol you can find. Most corporate networks operate between 6-8 protocol suites in addition to an IP environment.

Preventing a network attack to an operating system resource, especially with the fact that most attacks are inside jobs, requires security for ALL protocols, not just IP. In a trusted network environment on most non-UNIX servers, IPX and NetBEUI reign supreme as do other non-IP protocols and any of these may be used to gain access to a server and thusly attack the server.

5.5.5.7 Client Attacks - A New Threat

For a while, network security defenses have concentrated on keeping attackers at bay from servers of various shapes and sorts. The problem, especially in the last three years, has shifted towards client-side connections as well.

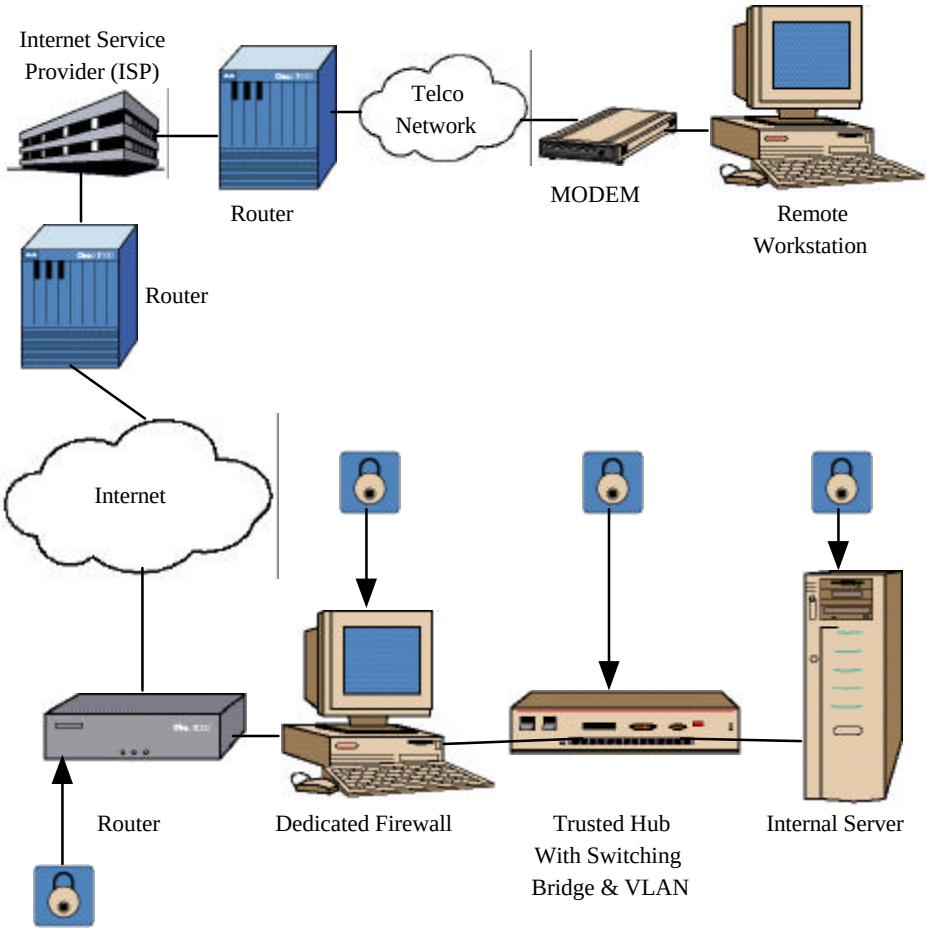
With Apple Computer's MacOS V7.1 and later versions, AppleTalk protocol was included in all versions of the operating system with functionality to not only access servers, but also to allow the client to publish itself as a disk service in a network and allow other clients to access the disk services. This is called peer-to-peer access as there is no intermediary system required for the connection to be made and maintained. Other vendors, noticeably Microsoft, have followed suit and included peer-to-peer services in their operating systems when shipped for consumption.

In Windows-95 and Windows-NT, protocol stacks for NetBEUI (a connection-less protocol which was originally used in LAN Manager), IPX (for accessing Novell NetWare servers) and IP (for use with TCP/IP savvy applications) are included at no extra charge as are various popular applications, such as web browsers and file sharing software, to make use of the various protocols. It is, therefore, very common and normal to find many protocols active on a trusted intranet. Now, however, many of the disk services or printer sharing services may well be based on a client system and not a dedicated server.

In the very near future (beginning in late 1996), high-speed residential connections will be more and more popular. The author has been directly involved in using a 7mbps connection from his home to the Internet for $19.95 per month via the local cable television network. This connection "looks" like a standard Ethernet connection (it even provides a standard RJ45 UTP connection on the set-top box connection to the cable broadband network) and even works like one with the client software. It also means that it was a trivial matter for the author to load up protocol analysis software on his workstation client and see, quite literally, activity on the cable television network by other persons in the neighborhood including Internet Service Provider (ISP) passwords by other users, files being transferred and popular locations that other neighbors access on the network. Therefore, there is basically NO security when all traffic can be seen in the clear on the network by nodes using the network.

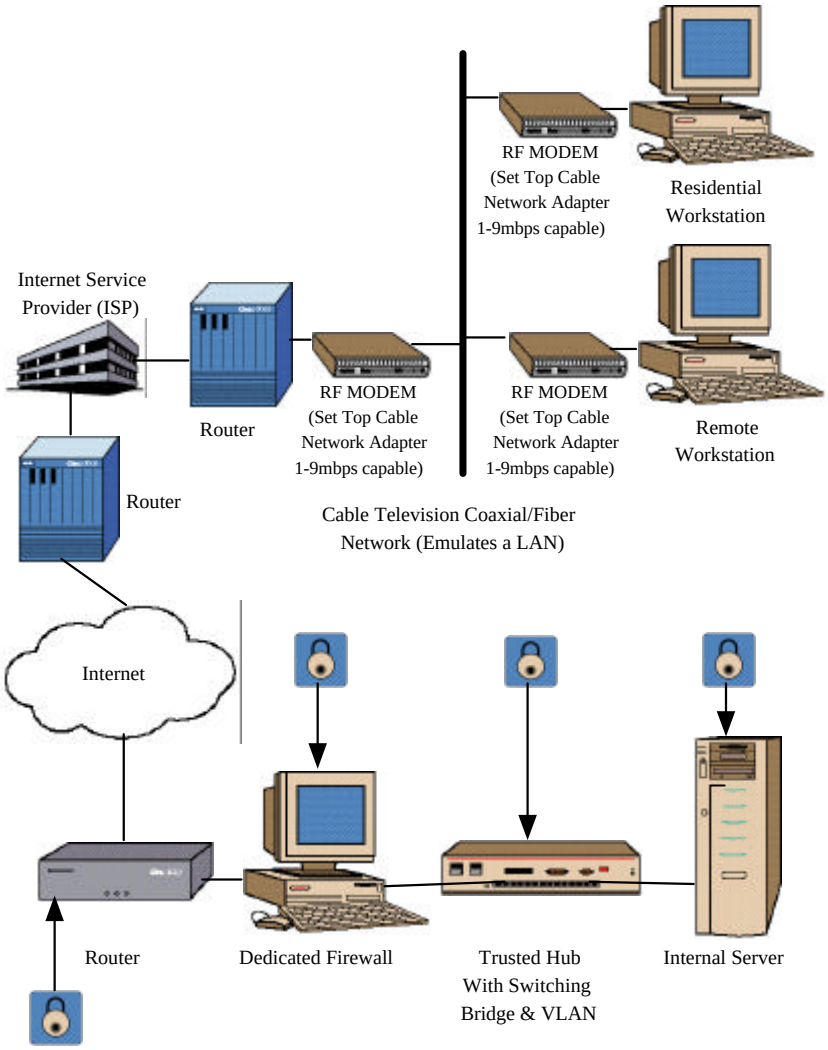
5.5.5.8 Telecommuting Client Security Problems - Coming to Your Company Soon

Obviously, this is a considerable security problem brewing considering that telecommuting is rapidly becoming the norm and high-speed network connections via cable television networks, Asymmetric High Speed Links (ADSL) and other technologies will be the normal mode of connection in the future. Some studies suggest that over 60% of information-related jobs will telecommute over 40% of the week by the year 2000, so this is a problem that will accelerate - rather quickly. A typical dial-in or ISDN telecommuter connection path is as follows:



For telecommuters, the need to support more than IP will also be the norm. Companies are adding IP generously to their internal systems, but they are also keeping protocols they have invested in for some time such as IPX, AppleTalk and NetBEUI. Therefore, for some considerable timeframe, the need to support IP and other protocols for telecommuting will be required in most corporate environments.

As telecommuting becomes more prevalent, telecommuters will keep more sensitive corporate information at their residences. This increases the overall security threat to the company as information deemed sensitive can now be accessed outside the physical perimeter of the corporate campus and the handful of allowed remote access facilities currently in place. Since client computers hooked to networks, like cable television, become “information appliances” due to their being continually network connected, they will be subjected to systematic network attacks no differently than corporate networks connected to any untrusted network. A typical cable TV connection methodology would appear as:



Since most client computers do not include the ability to provide a firewall facility in the client remote or residential computer, the chances of being attacked when connected to public high-speed networks is extremely good as well as having a high potential for success. A 1996 U.S. General Accounting Office report showed over 240,000 attempts at attacking the U.S. Department of Defense (DoD) unclassified networks and they suggested that over 64% of the attacks were successful. It is well known that the DoD takes security very seriously. So, what is going to happen to the potential millions of telecommuters who connect to their office facilities with no network security facilities and who leave their home-based systems on all day while at the office and also while connected to the high-speed network provided by the cable television vendor? Free-lance attacks will be the norm and easily accomplished.

5.5.5.9 Compromising Network Traffic - On LANs and Cable Television It's Easy

To simplify the matter, the chances of collecting data on in-path transactions on the Internet via a dial-up connection requires some specific levels of expertise. In the case of connections to cable television, very inexpensive or "free" network analysis software is available for PC and Macintosh systems and can allow the connection's data to be viewed in ASCII and sensitive information freely seen.

It should be noted that on intranets, most other protocols do not have encryption as well and those who do usually only use the encryption function for session establishment or, in the case of Novell Netware, for password security. The problem is that for some devices, such as Netware-aware printers, encryption is not always supported for passwords so it is commonly disabled to allow users access to printers. Just because a security feature exists does not mean that it is used properly or at all.

On corporate enterprise networks, it is the norm for the users to have a common format for user ID's and passwords to keep them from being too confused when accessing many different systems and servers. Therefore securing one protocol is not good enough. If the user accesses another network system using the same user ID and password as is used on an encrypted protocol session and the second protocol is unencrypted, then the password is compromised even for the encrypted session. To properly protect network connectivity, all protocols must be encrypted for all transactions and then all packets must be controlled (firewalled) when they arrive at the destination to keep users from accessing sensitive information and to protect the user's client system integrity.

5.5.5.10 Encryption is Not Enough - Firewall Services Are Needed As Well

Even in those situations where encryption capabilities have been introduced into client systems via encryption MODEMs or via software facilities in a specific protocol, this does not solve the end-to-end network security problem. Encryption is very good for authentication of a specific remote entity and is also very good for "hiding" any transaction over the network from observers of the traffic being transferred. The problem is that encryption is very much like giving someone you trust the keys to your house in such a manner that no one can see your friend accessing your house and no one can see what your friend is doing between his/her house and your house. This is good. What is not so good is that encryption does not stop a trusted user from still attacking the destination system's services that are offered. For instance, encryption may ensure that only corporate users get access to a system but encryption does not restrict, to a very fine degree, what a trusted user may be allowed to access and extract from the server. It's very much like letting someone you trust in the front door and not placing any restrictions on where someone is allowed to go in the house and what they are not allowed to deliver or remove from the house.

Firewall facilities, at the destination or the source of a network session, when used with encryption facilities add the additional filtering and security controls that are needed for network security on a client or a server. Encryption ensures that the connection is allowed and protected from observation. Firewall facilities on the client or server restrict where incoming or outgoing connections can access data on entities on the client or server. By setting up specific firewall rule bases on the client and server in addition to encryption software, the security manager can properly protect system resources from systematic and asymmetric network attacks.

5.5.5.11 Multiprotocol Security Requirements are the Norm - Not the Exception. Even for Singular Protocol Suites...

On corporate intranets, IP is not the only protocol used. Therefore, any network security solution that is used must include support for any corporate protocol. Further, any remote solutions must provide support for whatever protocol is required

to access the corporate facilities plus supply facilities for any cooperative protocol to be passed over the connection link (this is typically called “tunneling”).

Even if IP is decided to be the main corporate protocol now and in the future, it is a known fact that IP will get periodic lobotomies to support additional network types, addressing types, applications and other technological changes. This means that the need to run the “old” version of IP and the “new” version of IP at the same time on the same systems is highly likely while conversions are in progress on any network. Any network manager can tell you horror stories about converting from one version to another version of practically any protocol. And, practically without exception, most companies want to run the new version and the old version at the same time during testing before going to the new version due to potential problems and outages that happen with any new protocol environment. Therefore, any protocol security solution must be multiple protocol capable - even if it is only for the same protocol suite and is required to run multiple versions of the same protocol suite.

5.5.5.12 Protecting Clients and Servers on Multiprotocol Networks - How to Do It

So, how do you protect a server or client from network attack on the trusted, multiprotocol network? How do you protect remote clients that are used by telecommuters from localized attack or asymmetric attacks from other sources on a public-accessible network?

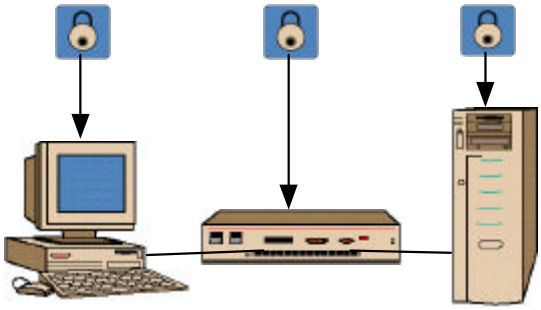
With the proper network security architecture, there are some basic, *major* elements required on each and every system to make such a feat work:

- Encryption software on each system which will allow multiprotocol support (client and server)
- Firewall software on *each* system which contains frame, packet, application filtering as well as “stateful” inspection facilities - for multiple protocols that are used or multiple versions of the same protocol suite (e.g. IP and IPV6 at the same time)
- Support for the proper network hardware being used by the client or server
- Virtual Private Networking (VPN) facilities for client-to-server, server-to-server and client-to-client (peer-to-peer) connections

There are a lot of other items which make life easier (like remote management) that are not critical to the security function but certainly very useful. Without the four major facilities listed above, there is not much likelihood of providing a useful set of network security facilities for end-to-end connections.

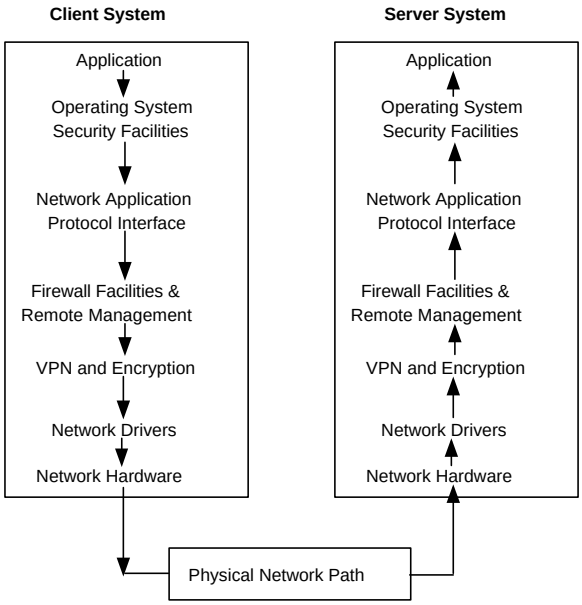
5.5.5.13 New Firewall Concepts - Firewalls with One Network Connection

Historically, firewall systems filter data from an untrusted network to/from a trusted network. With the need for end-to-end security, there is a need to provide the functionality of a firewall with VPNs at the workstation and singly-connected server level. In this scenario, the firewall software treats the singular network connection on a node as the untrusted side of the network and the node itself as the trusted side of the network. Any connection going out of the client or server is considered to be a trusted connection. A general hardware connection diagram would be as follows:



Client System with Firewall & VPN Software | Trusted Hub With Switching Bridge & VLAN | Internal Server with Firewall and VPN Software

Architecturally, the connection path for applications utilizing a singular network interface firewall system would appear as follows:



In the above architecture, both the client and the server treat all incoming connections through their internal firewall facilities as “untrusted.” All outgoing connections are considered as sourced from the “trusted” side.

Section References

5.0 Wack, John P. and Carnahan Lisa J., *Keeping Your Site Comfortably Secure: An Introduction to Internet Firewalls. NIST Special Publication 800-10*, U.S. Dept of Commerce.

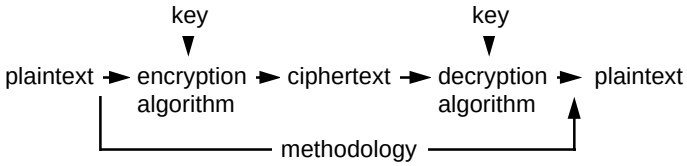
5.5.4 Guttman, Barbara and Bagwill, Robert. *Implementing Internet Firewall Security policy. Nist Special Publication 800-XX*. U.S Dept of Commerce. April 1998.

5.5.5 Hancock, William M. *Intranet Firewalls (Presentation).* Network-1 Software and Technology, Inc.1997-8.

Cryptography is the science of securing data. It addresses four major concerns—confidentiality, authentication, integrity and non-repudiation. Encryption is the transformation of data into an unreadable form, using an encryption/decryption key. Encryption ensures privacy and confidentiality, keeping information hidden from anyone for whom it is not intended including those who can see the encrypted data.

6.1 Cryptosystems

A cryptosystem obeys a methodology (procedure). It includes: one or more encryption algorithms (mathematical formulae); keys used with the encryption algorithms; a key management system; plain text (the original text); and, ciphertext (the original text that has been obscured).



The methodology first applies the encryption algorithm and key to the plaintext to produce ciphertext. The ciphertext is transmitted to a destination where the same algorithm is used to decrypt it to produce the plaintext. The procedure (included in the methodology) to support key creation and distribution is not shown in the diagram.

6.1.0 Key-Based Methodology

In this methodology, the encryption algorithm combines with a key and plaintext to create ciphertext. The security of a strong key-based system resides with the secrecy of the key used with the encryption algorithm rather than the supposed secrecy of the algorithm. Many encryption algorithms are publicly available and have been well tested (e.g. Data Encryption Standard).

However, the main problem with any key-based methodology is how to create and move the keys securely among communicating parties. How does one establish a secure channel between the parties prior to transmitting keys?

Another problem is authentication. There are two potential areas of concern here:

- The message is encrypted by whomever holds the key at a given moment. This should be the owner of the key; but if the system has been compromised, it could be a spoofer.

- When the communicating parties receive the keys, how do those parties know that the keys were actually created and sent by the proper authority?

There are two types of key-based methodologies—symmetric (private-key) and asymmetric (public-key). Each methodology uses its own procedures, key distribution, types of keys and encryption/decryption algorithms. The terminology employed in discussing these methodologies can be very confusing. The following terms are used:

TERM	MEANING	POTENTIAL CONFUSION
Symmetric methodology	Uses one key which both encrypts and decrypts using the same symmetric encryption algorithm The key is distributed to the two communicating parties in a secure manner prior to transfer of encrypted data	Often called private or private-key methodology
Asymmetric methodology	Uses symmetric encryption algorithms and symmetric keys to encrypt data Uses asymmetric encryption algorithms and asymmetric keys to encrypt the symmetric key. The two keys are created and are linked together. The symmetric key encrypted with one must be decrypted by the other (in either direction) using the same asymmetric encryption algorithm. The two linked asymmetric keys are created together. One must be distributed to the owner, and the other to the party which is keeping these keys (often called the CA) in a secure manner prior to transfer of data	Often called public or public-key methodology
Private key (1)	Symmetric methodology	Uses a single key which can both encrypt and decrypt. See above.
Private key (2)	Symmetric (private) encryption key	Symmetric private key
Private key (3)	Asymmetric private encryption key	Asymmetric private key Asymmetric keys are created as pairs that are linked together. The words private key often mean the half of the asymmetric key pair that is kept private. The asymmetric private key is a totally different thing from the symmetric private key.
Public key (1)	Asymmetric methodology	Uses a pair of keys, both of which are created together and are linked. Anything encrypted by one must be decrypted by the other.
Public key (2)	Asymmetric (public) encryption key	Asymmetric keys are created as pairs that are linked together.

		The words public key often mean the half of the asymmetric key pair which is made publicly available.
Session key	Symmetric (private) encryption key	Used by asymmetric methodology for the actual data encryption of data using symmetric methodologies Simply a symmetric private key (see above)
Encryption algorithm	Mathematical formula	Symmetric keys are required for symmetric algorithms Asymmetric keys are required for asymmetric algorithms You cannot use symmetric keys with asymmetric algorithms, and vice versa
Private cryptosystems	Use symmetric algorithms and symmetric (private) keys to encrypt data	Used by symmetric (private) cryptosystems
Public cryptosystems	Use asymmetric algorithms and asymmetric keys to encrypt session keys uses symmetric algorithms and symmetric keys to encrypt data	Used by asymmetric (public) cryptosystems only
Public/private	Many asymmetric cryptosystem vendors define their methodologies as public/private	Usually not clarified that asymmetric methodologies use symmetric methodologies to actually encrypt data

6.1.1 Symmetric (Private) Methodology

In this methodology, both encryption and decryption operations use the same key with the sender and receiver agreeing on the key before they can communicate. Provided the keys have not been compromised, authentication is implicitly resolved because only the sender has a key capable of encrypting and only the receiver has the same key capable of decrypting. Because the sender and the receiver are the only people who know this symmetric key, if the key is compromised, only these two users' communication is compromised. The problem, which is the same for all types of cryptosystems, is how to distribute the symmetric (private) key securely.

Symmetric key encryption algorithms use small-length keys and can quickly encrypt large quantities of data.

The process involved with symmetric key systems is:

1. Create, distribute and store the symmetric private key securely
2. Sender creates a digital signature by hashing the plaintext and attaching the resulting string to the plaintext

3. Sender applies the fast symmetric encryption/decryption algorithm with the symmetric private key to the package (plaintext and attached digital signature) to produce the ciphertext. Authentication happens inherently because only the sender has the symmetric private key and can encrypt the package. Only the receiver holding the symmetric private key and can decrypt this package
4. Sender transfers the ciphertext. The private symmetric key is never transmitted over the unsecured communication lines.
5. Receiver applies the same symmetric encryption/decryption algorithm with the same symmetric key (which the receiver already has) to the ciphertext to produce the original plaintext and digital signature. This authenticates whoever holds the private key.
6. Receiver detaches the digital signature from the plaintext
7. Receiver creates a digital signature by hashing the plaintext
8. Receiver compares the two digital signatures to prove message integrity (unaltered data)

Services available today that use symmetric methodologies include:

- Kerberos, which is designed to authenticate access to network resources rather than to verify data. It uses a central database that generates and keeps copies of the secret keys of all users.
- ATM Banking Networks (automated teller machines). These systems are proprietary and are not for resale, although they use symmetric methodologies.

6.1.2 Asymmetric (Public) Methodology

Here, the encryption and decryption keys are different from each other, although they are produced together. One key is made public; the other key is kept private. While both keys can encrypt and decrypt, data encrypted by one can only be decrypted by the other.

All asymmetric cryptosystems are subject to shortcut attacks as well as brute force, and therefore, must use much larger keys than symmetric cryptosystems to provide equivalent levels of security. This immediately impacts computing cost, although using elliptic curve algorithms may reduce this problem. Bruce Schneier in his book "Applied Cryptography: Protocols, Algorithms, and Source Code in C" provides the following table comparing equivalent key lengths:

SYMMETRIC KEY LENGTH	PUBLIC-KEY KEY LENGTH
56 bits	384 bits
64 bits	512 bits
80 bits	768 bits
112 bits	1792 bits
128 bits	2304 bits

In order to circumvent the slowness of the asymmetric encryption algorithms, a temporary, random, small, symmetric session key is generated for each message and is the only part encrypted by the asymmetric algorithm. The message itself is encrypted using this session key and an encryption/decryption algorithm. The small session key is then encrypted using the sender's asymmetric private key and encryption/decryption algorithm. This encrypted session key along with the encrypted message is then transmitted to the receiver. The receiver uses the same asymmetric algorithm and the sender's asymmetric public key to decrypt the session key, and the recovered plaintext session key is used to finally decrypt the message.

It is important in asymmetric cryptosystems that the session and asymmetric keys must be comparable in terms of the security they produce. If a short session key is used (e.g. 40 bit DES), it does not matter how large the asymmetric keys are. Hackers will attack the session key instead. The asymmetric public keys are susceptible to brute-force attacks partly because it is difficult to change them. Once broken, all current and future communication is compromised, often without anyone knowing.

The process involved with asymmetric-key systems is:

1. Create and distribute the asymmetric public and private keys securely. The asymmetric private key is delivered to the owner. The asymmetric public key is stored in an X.500 database and managed by the Certification Authority (CA). Users must implicitly trust the secure creation, distribution and management of the keys. Further, if the creator and the person or system managing the keys are different, then the end user must implicitly trust that the creator of the keys has actually deleted his copies.
2. Create a digital signature by hashing the plaintext. Encrypt the resulting digital signature using the sender's asymmetric private key and attach the resulting string to the plaintext (only the sender has created the digital signature).
3. Create a private symmetric key used only for this transmission (the session key), and apply it and the symmetric encryption/decryption algorithm to the plaintext and attached encrypted digital signature to produce the ciphertext.
4. The problem of sending the session key to the receiver must now be addressed
5. Make certain the sender has the Certification Authority's (CA) asymmetric public key. Interception of unencrypted requests for the public key is a common form of attack. There may be a whole hierarchy of certificates attesting to the validity of the CA's public key. X.509 describes different methods for establishing user access to the CA public keys, all of which provide an entry point to spoofers, and show that there is no system that guarantees the identity of the CA.
6. Ask the CA for the receiver's asymmetric public key. This process is vulnerable to the man-in-the-middle attack. The receiver's asymmetric public key has been 'digitally signed' by the CA. This means that the CA has used the CA's asymmetric private key to encrypt the receiver's asymmetric public key. Since only the CA holds the CA's asymmetric private key, then the receiver's asymmetric public key came from the CA
7. Once received, decrypt the receiver's asymmetric public key using the CA's asymmetric public key and an asymmetric encryption/decryption algorithm. Implicit trust in the CA and that the CA is not compromised are required. If the CA is compromised, the entire infrastructure is unusable. Those holding the public key can encrypt, but there is no way of knowing if the key has been compromised. (When you requested the CA's public key, did you actually receive the CA's public key or something else?)
8. Using the receiver's asymmetric public key (now received from the CA and decrypted) and an asymmetric encryption/decryption algorithm, encrypt the session key. Only those holding the receiver's public key can encrypt, but there is no way of knowing if the key has been compromised
9. Attach the encrypted session key to the ciphertext (which includes the previously encrypted digital signature
10. Transfer the package (ciphertext that includes the digital signature and the attached encrypted session key). The encrypted session key is transmitted across the unsecured network and is an obvious target for various types of attacks.
11. Receiver detaches the encrypted session key from the ciphertext
12. The problem of decrypting the session key by the receiver must now be addressed

13. Make certain the receiver has the CA's asymmetric public key. The same comments as above can be made here.
14. Using the receiver's asymmetric private key and the same asymmetric encryption/decryption algorithm, receiver decrypts the session key
15. Receiver applies the same symmetric encryption/decryption algorithm with the now unencrypted symmetric key (session key) to the ciphertext to produce the plaintext and attached hash or digital signature
16. Receiver detaches the hash from the plaintext
17. Receiver asks the CA for the sender's asymmetric public key
18. Once received, receiver decrypts the sender's asymmetric public key using the CA's public key and the correct asymmetric encryption/decryption algorithm. The same comments as above can be made here.
19. Using the sender's asymmetric public key and an asymmetric encryption/decryption algorithm, receiver decrypts the hash string
20. Create a digital signature by hashing the plaintext
21. Compare the two hashes to prove that the data has not been altered

6.1.3 Key Distribution

It is obvious that both types of cryptosystems have a problem distributing the keys.

Symmetric methodologies squarely face up to this fact and define how keys are to be moved between the parties before communication can take place. How this is done depends upon the security required. For lower security requirements, sending keys by a delivery mechanism of some kind (such as postal mail or a parcel delivery service) may be adequate. Banks use the postal service to deliver PINs, which are, in essence, easily crackable symmetric keys that may or may not unlock other keys, or your money! Very high security requirements may require hand delivery of keys, possibly in parts by several people.

Asymmetric methodologies try to get around the problem by encrypting the symmetric key and attaching it to the encrypted data. They then try to make it possible to distribute the asymmetric keys used to encrypt the symmetric key by employing a CA to store the public asymmetric key. The CA in turn digitally signs the keys with the CA's private asymmetric key. Users of the system must also have a copy of the CA's public key. In theory, this means that the communicating parties do not need to know about each other ahead of secure communication.

Proponents of asymmetric cryptosystems maintain that this mechanism proves authenticity and is sufficient.

The problem still remains, however. The asymmetric key pair must be created together. Both keys, whether they can be made publicly available or not, must be sent securely to the owner of the key, as well as to the Certification Authority. The only way to do this is by some kind of delivery mechanism for low security requirements, and hand-delivery for high security requirements.

The problems of the asymmetric mechanism include the following:

- X.509 assumes that keys are securely distributed and does not address the issue other than identifying it. There are no standards covering this area. To be safe, keys (whether symmetric or asymmetric) must be hand-delivered. Even then, people could be intimidated or bribed.
- There is no mechanism in place to reliably validate what system is actually talking to what system. The man-in-the-middle attack is an attack by a spoofer masquerading as the CA and getting the data before it is realized the spoofer

was actually in the picture. All the spoofer has to do is to capture the request to the CA and substitute his own keys in its place. This type of spoofer has come and gone long before users become aware that something might be wrong.

- The digital signing by the CA of a key still does not prove the authenticity of the key because the CA's own key could have been compromised. X.509 describes digital signing of CA keys by higher level CAs, and describes this as a certification path (a hierarchy of CA public keys). X.509 discusses the problems associated with verifying the correctness of a public key, suggesting that it can only operate if there is an unbroken chain of trusted points in the directory between the users required to authenticate. The standards do not offer any mechanism to get around this.
- X.509 assumes the user has prior access to the CA's public key. How this is achieved is not defined in the standards document.
- Compromise of the Certification Authority is a very real threat. Compromise of the CA means that ALL users of the system are compromised. And no one might ever know. X.509 assumes that storage of all keys, including the CA keys, is secure. The deployment of X.500 directory systems (where X.509 keys are stored) is difficult and prone to misconfiguration. There are very few people available today with the technical knowledge required to manage these systems properly. Further, it is a well-known fact that people in trusted positions can be subverted—kidnapped or bribed.
- The CA may become a bottleneck. To provide for fault tolerance, X.509 suggests that the CA database be replicated or shadowed by using the X.500 standard directory services; this considerably raises the cost of the cryptosystem. If spoofing occurs, it is difficult to identify which system was attacked. Furthermore all the data must be sent across communication lines somehow when the data is being distributed.
- An X.500 directory system is costly to install, configure and maintain. Access to this directory is either by using an outside subscription service or by an organization providing its own. The X.509 certificate is based on each individual possessing a unique name. The allocation of names is the responsibility of yet another trusted authority, the naming authority.
- Full keys, even though encrypted, are transmitted across the unsecured communications medium.

In spite of these major drawbacks, users must blindly trust the asymmetric cryptosystem.

Key management refers to the distribution, authentication and handling of keys. No matter what kind of cryptosystem is used, keys must be managed. Secure methods of management are very important as many attacks on key-based cryptosystems are aimed at key management procedures.

PROCEDURE	COMMENTS
Physically distribute the keys	Couriers and hand delivery are two examples. Of the two, hand delivery is better. Secure organizations have written procedures surrounding key distribution Can be audited and logged, although open to compromise by individuals Used by both symmetric and asymmetric cryptosystems. In spite of claims that asymmetric cryptosystems avoid the problem of physical delivery of keys, the problem actually exists. X.509 assumes that the creator will release the asymmetric private key to the user (and/or the asymmetric public key to the CA) in a physically secure manner, and that suitable physical security measures are in place so that the creator and data operations are free from tampering.
Issue a common key from a central issuing authority	Could be used by both symmetric and asymmetric cryptosystems As each user must be able to communicate with the central authority securely in the first place, this is yet another situation where initial key exchange is a problem If the central authority is compromised, further requests for keys are at risk; keys already in place may be safe depending on the cryptosystem
Allow access to public keys from a centralized certification authority and provide private keys to users	Used by asymmetric cryptosystems Users must blindly trust the entire system A single security breach compromises the entire system Hierarchical system of attestation leads to more potential intruder entry points—a CA must publicize its asymmetric public key and provide a certificate from a higher-level CA validating it. This sets up a hierarchy of CAs. CA asymmetric private keys must be stored securely because compromise could result in undetectable forgeries
Web of trust	Used by asymmetric cryptosystems Users distribute and track each other's keys, and trust in an informal, distributed fashion
Diffie-Hellman	Exchange of a secret key over an insecure medium by two users without any prior secrets Cannot be used to encrypt or decrypt messages Based on the difficulty of taking logarithms in finite fields. If the elements are carefully chosen, and are large, then the discrete logarithm problem is computationally infeasible. Vulnerable to man-in-the-middle attacks Patented by PKP (Public Key Partners)

6.1.4 Encryption Ciphers or Algorithms

Key-based algorithms disguise data so that it cannot be read by anyone without a decryption key. They are divided into two classes depending on the cryptography methodology they directly support. Please read Schneier's *Applied Cryptography* for a full description of the algorithms.

6.1.5 Symmetric Algorithms

The same private key is used to encrypt and decrypt. This type of algorithm is used by both symmetric and asymmetric methodologies to encrypt data.

TYPE	DESCRIPTION
DES (Data Encryption Standard)	Popular, product cipher used by the Data Encryption Standard of the US Government 64-bit block cipher, 64-bit key (only 56 are needed), 16 rounds Operates in four modes: • ECB—Electronic Code Book (native DES), using two distinct algorithms • CBC—Cipher Block Chaining in which the encryption of each block depends upon the encryption of the previous block • OFB—Output Feedback, used as a random number generator • CFB—Cipher Feedback, used for message authentication codes
3-DES or Triple DES	64-bit block cipher, using the DES cipher 3 times, three distinct 56-bit keys Strong under all attacks
Chained 3-DES	Standard Triple-DES with the addition of a feedback mechanism such as CBC, OFB or CFB Very strong under all attacks
FEAL (Fast Encryption Algorithm)	Block cipher, used as an alternative to DES Broken, although new versions have been proposed
IDEA (International Data Encryption Algorithm)	64-bit block cipher, 128-bit key, 8 rounds Recently proposed; although it has not yet received enough scrutiny for full confidence, it is considered superior to DES
Skipjack	Developed by NSA as part of the US Government Clipper and Capstone projects Classified as secret, although its strength does not depend only on the secrecy of the algorithm 64-bit block cipher, 80-bit keys used in ECB, CFB, OFB or CBC modes, 32 rounds
RC2	64-bit block cipher, variable key sizes Approximately twice as fast as DES Can be used in same modes as DES including triple encryption Confidential algorithm proprietary to RSA Data Security
RC4	Stream cipher, byte-oriented, variable key size Approximately 10 times as fast as DES

	Confidential algorithm proprietary to RSA Data Security
RC5	32, 64 or 128-bit variable block size, 0 to 2048 variable key size, 0 to 255 rounds A fast block cipher Proprietary to RSA Data Security
CAST	64-bit block cipher, 40 to 64 bit keys, 8 rounds No known way to break other than brute force Generally, the particular S-boxes used (which form the strength of the algorithm) are not made public
Blowfish	64-bit block cipher, variable, up to 448-bit key, 16 rounds, each consisting of a key-dependent permutation and a key-and-data-dependent substitution Faster than DES Designed for 32-bit machines
One-time pad	A proven unbreakable cipher The key (same length as the text) is the next ‘n’ bits of randomly created bits found on a pad to which both the sender and the receiver have access. As soon as the bits are used, they are destroyed and the next bits on the pad are used for the next encryption
Stream ciphers	Fast, symmetric encryption algorithms, usually operating on bits (not blocks) of data Developed as an approximation of the one-time pad which, while not as secure as the one-time pad, are at least practical

6.1.6 Asymmetric Algorithms

Asymmetric algorithms are used by asymmetric cryptosystem methodologies in order to encrypt a symmetric session key (which is actually used to encrypt the data).

Two distinct keys are used—one that is publicly available, and the other that is kept private and secret. Usually both keys perform encryption and decryption functions. However, data encrypted by one can only be decrypted by the companion key.

TYPE	DESCRIPTION
RSA	Popular asymmetric encryption algorithm, whose security depends on the difficulty in factoring large integers
ECC (Elliptic Curve Cryptosystem)	Uses the algebraic system defined on the points of an elliptic curve to provide asymmetric cryptographic algorithms Emerging as competition to other asymmetric algorithms because it offers equivalent security using shorter key lengths and faster performance. Current implementations indicate that these systems are far more efficient than other public-key systems. Performance figures show an order of magnitude improvement in efficiency over RSA, Diffie-Hellman and DSA.
ElGamal	Variant of the Diffie-Hellman which can be used for both digital signatures and encryption

6.1.7 Hash Functions

Hash functions are central to key-based cryptosystems. They are relatively easy to compute, but almost impossible to decrypt. A hash function takes a variable size input and returns a fixed size string (sometimes called a Message Digest), usually 128 bits. Hash functions are used to detect modification of a message (i.e. provides a digital signature).

TYPE	DESCRIPTION
MD2	Slowest, optimized for 8-bit machines
MD4	Fastest, optimized for 32-bit machines Now broken
MD5	Most commonly used of the MD functions similar to MD4, but with added security features making it 33% slower than MD4 Provides data integrity Considered secure
SHA (Secure Hash Algorithm)	Produces 160-bit hash values from variable-sized input Proposed by NIST and adopted by the US Government as a standard Designed for use with the proposed DSS (Digital Signature Standard) and part of the US Government's Capstone project

6.1.8 Authentication Mechanisms

These mechanisms securely and reliably confirm identity or authenticity.

TYPE	DESCRIPTION
Passwords or PINs (Personal Identification Numbers)	Something a user knows and shares with the entity at the other end Typically part of a two way handshake Can be exchanged in both directions to obtain mutual authentication
One-time password	Password provided is never reused Time is often used as the constantly changing value on which the password is based
CHAP (Challenge Handshake Authentication Protocol)	One side initiates an authentication exchange, is presented with a unique and unpredictable challenge value, and based on a secretly shared value, is able to calculate and return an appropriate response Can be used to provide user authentication as well as device authentication
Callback	Dialing in over a telephone to a server which is configured to dial back to a specified number associated with the user

6.1.9 Digital Signatures and Time Stamps

A digital signature provides data integrity, but does not provide confidentiality. The digital signature is attached to the message and both can be encrypted if confidentiality is desired. The addition of a timestamp to a digital signature provides a limited form of non-repudiation.

TYPE	COMMENTS
DSA (Digital Signature Authorization)	Public key algorithm used for digital signatures but not for encryption Private hashing and public verification—only one person can produce the hash for a message, but everyone can verify that the hash is correct Based on the difficulty of taking logarithms in finite fields
RSA	Patented RSA digital signature proves the contents of a message as well as the identity of the signer The sender creates a hash of the message, and then encrypts it with the sender's private key. The receiver uses the sender's public key to decrypt the hash, hashes the message himself, and compares the two hashes.
MAC (Message Authentication Code)	Digital signature, using hashing schemes similar to MD or SHA, but the hash value is a function of both the pre-image and a private key
DTS (Digital Timestamp Service)	Issues timestamps which associate a date and time with a digital document in a cryptographically strong manner

Section References

6.0 Chandler, Janet, Cryptography 101: Technical White Paper, Signal 9 Solutions, Kanata Ontario.

7.1 What Is a Virus?

Computer viruses are programs which replicate themselves, attach themselves to other programs, and perform unsolicited and often malicious actions. Self-replication is the key trait that distinguishes viruses from other destructive programs. For instance, a Trojan Horse is a program which performs unsolicited actions, but it cannot replicate and spread on its own.

Viruses are destructive to productivity as well as data. An example of productivity damage is the Stoned virus which simply writes "Your computer is stoned" on the screen. Data damage is exemplified by the Hare virus (popularized in Summer 1996), which erases data from hard drives. In any case, viruses always cause some degradation of system resources, and some degree of wasted time for computer users. Since they are unsolicited and concealed, it does not seem accurate to call any virus "benign."

Critical to a virus' "success" is the ability to remain undetected for a long enough period to replicate and spread to new hosts. By the time the virus' presence is revealed, through unusual computer "behavior," damage to data or taunting messages, it usually will have been quite some time since the original infection took place.

This delay in time, between infection and manifestation, obviously makes it more difficult to trace the origin of the virus and/or the route it took to reach one's system. So delays are often made an inherent "feature" within a virus' design. A virus may monitor for a trigger event, which is a computer condition that, when it occurs, will cause the virus' payload to be delivered.

Examples of trigger events include dates (such as March 6 for the infamous Michelangelo virus), times, number of file saves or disk accesses, or file sizes. Specific keystroke sequences, in any predictable combination, can also be triggers.

A payload is an action performed by a virus - usually, but not always, the action that reveals the virus' presence. Examples of payloads include:

- "Amusing" or political messages (such as the Nuclear macro virus which asks for a ban on the French nuclear testing)
- Prevention of access to one's disk drives (the Monkey virus)

- A stealth boot virus overwriting data as it attempts to write pre-infected boot information to another part of the disk
- Inconspicuous activity and minute data damage spread out over a long period of time - probably the most lethal type of virus effect (the Ripper virus)

Typical Signs that a Virus May Be Present:

- Unusual messages displayed
- Files are missing or have increased in size
- System operates slower
- Sudden lack of disk space
- Cannot access disk

7.1.0 Boot vs File Viruses

Before the inception and rapid proliferation of the Macro category, most IBM-compatible and Macintosh viruses fell into two basic categories: Boot, such as "Michelangelo" and File, such as "Jerusalem."

Boot viruses activate upon system start-up and are more common. They infect a system's floppy or hard disk and then spread (by replicating and attaching) to any logical disks available. File viruses are actually programs which must be executed in order to become active, and include executable files such as .com, .exe and .dll. Once executed, file viruses replicate and attach to other executable files. Since most viruses attach at the beginning or end of processes, their execution goes unnoticed.

7.1.1 Additional Virus Classifications

Other troublesome general virus sub-classes that are active today include Stealth (active and passive), Multipartite, Encrypted, Polymorphic, and Macro.

Stealth viruses (such as "Tequila") are difficult to detect because, as their name implies, they actually disguise their actions. Passive Stealth viruses can increase a file's size, yet present the appearance of the original file size, thus evading Integrity Checking - one of the most fundamental detection tactics.
Active Stealth viruses may be written so that they actually attack installed anti-virus software (generic or brand-specific), rendering the product's detection tools useless.

Multipartite viruses, such as "Natas," have the characteristics of both boot and file viruses. "Cascade" is a well-known Encrypted virus. The challenge of Encrypted viruses is not primarily one of detection, per se. The encryption engine of this type of virus masks its viral code – making identification, as opposed to detection, more difficult.

The Polymorphic category ("SMEG" is an example) has grown considerably, presenting a particular detection challenge. Each polymorphic virus has a built-in mutation engine. This engine creates random changes to the virus' signature on given replications. Therefore, detection and prevention of recurring infections further requires frequent anti-virus component updates from a given vendor.

7.2 The New Macro Virus Threat

If you send or receive documents or spreadsheets, chances are your computer has been or will be infected at one time or another by a macro virus. Relatively new on the computing scene, these computer viruses are spreading faster than most anti-virus software makers can find ways to detect and remove them. Macro viruses are now the most prevalent computer viruses in the world, largely due to the new way in which they spread--they attach themselves to word processor and spreadsheet documents, which often are transmitted as e-mail attachments via the Internet throughout the world.

This new means of virus proliferation calls for new methods of virus detection. One such approach is based on intelligent, rule-based scanning -- a technique that searches for and removes even macro viruses never before analyzed. This approach combines the following elements:

- OLE2 technology to efficiently extract only that portion of files that can carry viruses
- Pattern matching for detection of known viruses, as well as intelligent rule-based scanning to detect unknown viruses

7.2.0 Background

Despite a significant increase in the usage of anti-virus products, the rate of computer virus infection in corporate America has nearly tripled in the past year, according to a survey released in April 1997 by the International Computer Security Association (ICSA), formerly the National Computer Security Association. Virtually all medium and large organizations in North America experienced at least one computer virus infection firsthand, and the survey indicated that about 40 percent of all computers used in the surveyed companies would experience a virus infection within a year.

Macro viruses, which unlike their predecessors, are carried in common word processing documents and spreadsheets, are the biggest problem, representing 80% of all infections. Moreover, the instances of macro virus infection doubled about every four months in 1996. This makes these viruses the fastest to spread in the history of the ICSA.

The Number One macro virus encountered in the survey, by far, was the Concept virus, also known as prank macro, wm-Concept, winword.Concept, wordmacro.Concept, ww6, and ww6macro. Within months of its discovery in the fall of 1995, the Concept virus accounted for more than three times the number of virus encounters reported for the previous leader, the "Form virus." Today, the Concept virus has infected almost one-half of all ICSA survey sites (see Figure 1).

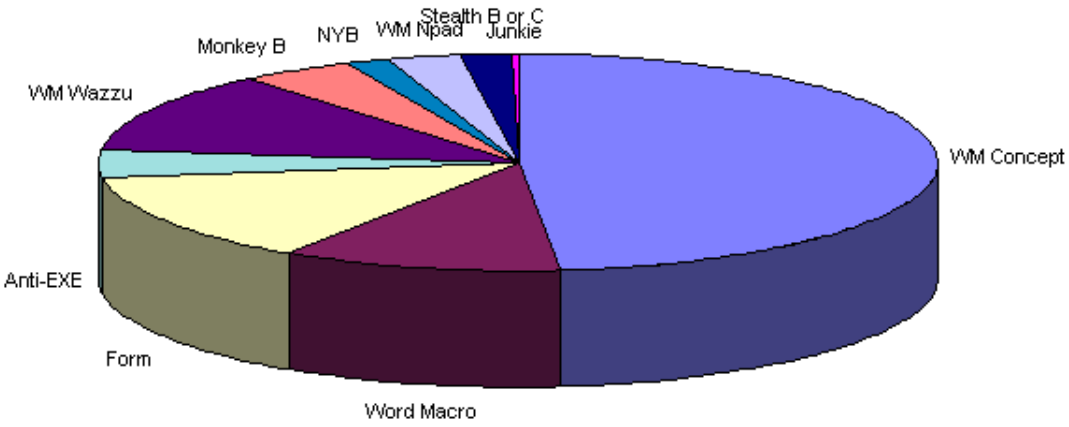


Figure 1. The Concept virus and other Word macro viruses were the dominant viruses encountered in 1997, according to a virus prevalence survey conducted by the International Computer Security Association.

Perhaps even more worrying than the meteoric rise in infections by this particular virus is what it bodes for the future. Microsoft Word™, Microsoft Excel™, and other document and spreadsheet files were once thought to be immune to

infection. Since these virus carriers are now the most prevalent types of files exchanged in the world, the threat of viruses has evolved in a big way. With the exponential growth of the Internet for e-mail and file exchange, macro viruses now represent the most widespread virus threat ever.

"Macro viruses are incredibly successful viruses," says Eva Chen, CTO of Trend Micro. "Because they hitchhike on document and spreadsheet files, they can travel both on floppy diskettes and across computer networks as attachments to electronic mail. Then they spread quickly by taking advantage of e-mail, groupware, and Internet traffic."

Adding to growing concern about these viruses is the ease of their creation. Prior to the macro virus era, creating a virus required some knowledge of assembly language or other complex programming language. Today, almost anyone can write a macro virus using Visual Basic, which uses English-like commands (see Figure 2). There is even a guided step-by-step template for creating Word macro viruses available on the Internet.

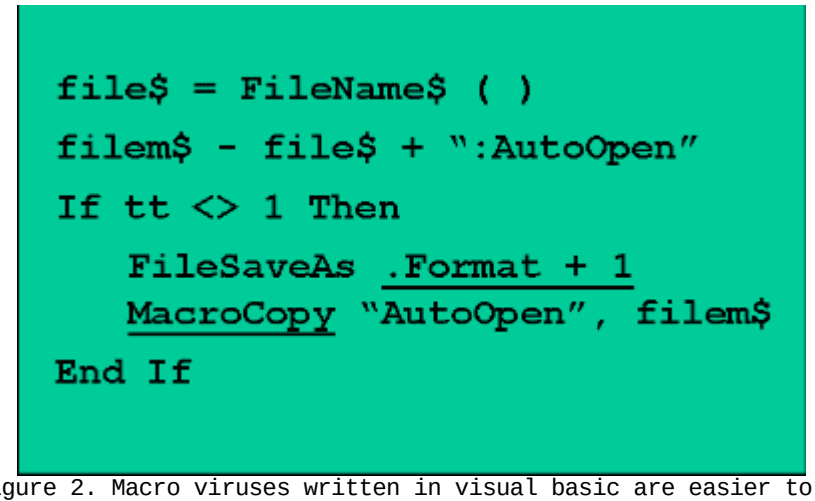


Figure 2. Macro viruses written in visual basic are easier to write than their assembly language predecessors.

While most of the more than 500 macro viruses known at the time of this writing are not destructive, many cause a considerable loss of productivity and staff time. Average financial cost per 'virus disaster,' according to the ICSA, rose to $8366 in 1997, and Figure 3 shows that virus incident costs are shifting from predominantly low levels to intermediate levels. Concept restricts file saving operations, and other macro viruses have been known to manipulate information, control data storage, and even reformat hard drives. This potential destructiveness has system administrators buzzing about how to address this new threat.

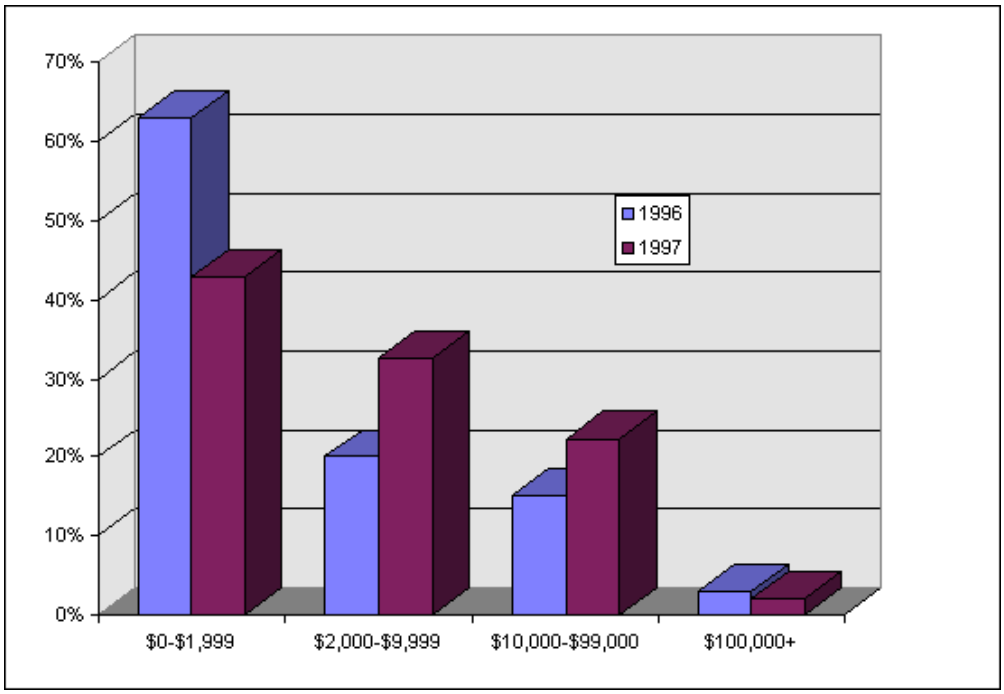


Figure 3. According to the ICSA 1997 Computer Virus Prevalence Survey, the stated costs of virus incidents tended to shift from less than $2000 to the range of $2000-$99,000 [1].

7.2.1 Macro Viruses: How They Work

Understanding how to protect against macro viruses requires some knowledge about what makes these viruses tick. Just when we thought we understood how viruses work--by attaching executable code to other executable code in software--along come viruses that attach themselves to document files and spreadsheets. How do macro viruses pull this off?

The answer is that there is more to today's word processing or spreadsheet file than meets the eye.

Traditional files like these consist solely of text. But today's increasingly sophisticated word processing and spreadsheet files carry macros with them that can provide a variety of features to your documents and spreadsheets. For example, macro commands can perform key tasks, such as saving files every few minutes, or they can prompt you to type in information, such as a name and address into a form letter. These macros, part of the document itself, travel with the file as it is transferred from user to user, either via floppy diskette, file transfer, or e-mail attachment.

Some of these macro commands have special attributes that force them to execute automatically when the user performs various standard operations. For example, Word uses five predefined macros, including the AutoOpen macro, which executes when a user opens a Word document, and AutoClose, which runs when you close the document.

Macro viruses gain access to word processing and spreadsheet files by attaching themselves to the executable portion of the document--in AutoOpen, AutoExec, AutoNew, AutoClose, AutoExit, and other file macros. For example, the Concept virus attaches itself to AutoOpen and FileSaveAs in Word
(See Figure 4).

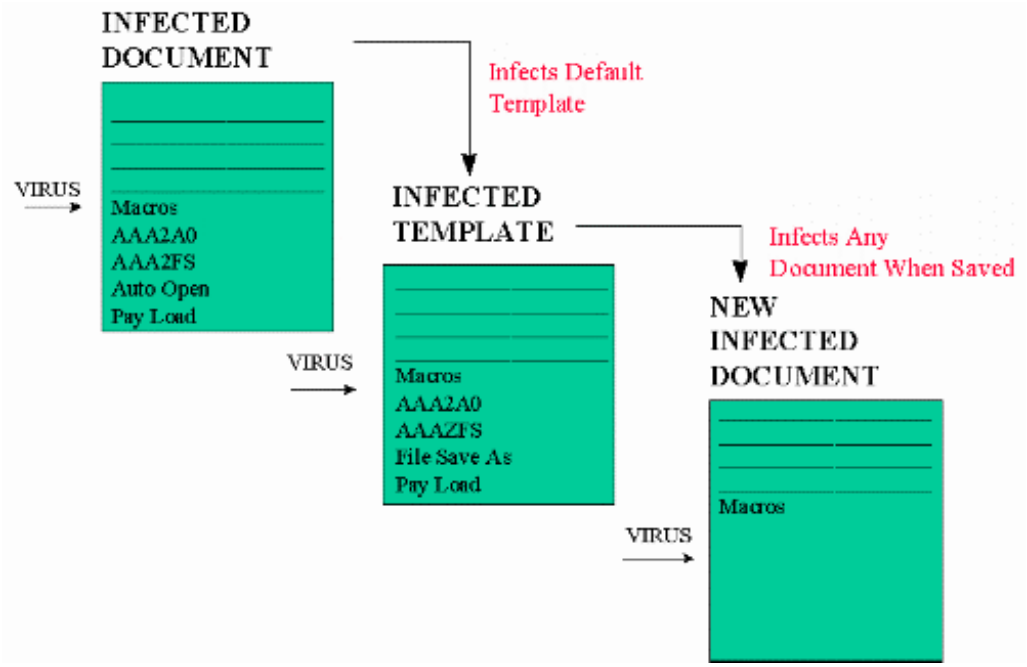


Figure 4. Concept latches onto one macro that is automatically run in Word: AutoOpen. By attaching itself to AutoOpen, the virus takes control as soon as an infected document is opened. Next, it infects the default template. Then, by attaching itself to FileSaveAs, the virus effectively spreads itself to any other document when it is saved.

Macro viruses are particularly difficult to eradicate because they can hide in attachments to old e-mail messages. For example, the administrator of a network infected by a macro virus may take pains to eliminate it. But when an employee returns from a vacation and opens an e-mail attachment with the virus and forwards it to others on the network, the virus can spread again, necessitating a second round of detection and disinfection.

This migration of viruses to word processing and spreadsheet files mirrors user computing patterns. In fact, this parallel evolution of viruses and computing media has been going on for years. When the primary means of exchanging files was the floppy diskette, the most prevalent viruses were boot sector infectors, which resided on the first sector of a diskette. Later, the wide use of internal networks built around file servers allowed viruses to spread by modifying executable files. Today, the ICSA reports that commonly exchanged word processed and spreadsheet files sent over the Internet as e-mail attachments are the most common carrier of viruses [1].

7.2.2 Detecting Macro Viruses

The increase in virus incidence despite rising anti-virus usage can lead to but one conclusion. "It is obvious that existing virus protection software isn't working," says

Chen. "Traditional methods have not been successful in combating viruses entering networks from new entry points--e-mail and the
Internet." Hence, the Concept virus seems to be aptly named, since dealing with it and viruses like it reliably and effectively requires new concepts in virus detection.

The traditional approach to virus detection has been to gather samples of suspicious code, conduct analysis, create new virus signature files, and distribute them to customers.

Assuming that users periodically download updates of anti-virus software, this approach works well for viruses that do not spread quickly and for viruses without large numbers of variants. Many anti-virus software packages that take this approach use pattern-matching algorithms to search for a string of code that signals malicious actions. When virus writers began to foil this "fingerprint analysis" by encrypting their code, anti-virus software developers responded by using the decryption routine included with the virus, emulating operation of the code in an isolated environment, and determining if the code was malicious.

Unfortunately, the Concept virus and other macro viruses often elude these techniques for several reasons. The ease with which these viruses can be developed, coupled with the vast number of word processing and spreadsheet documents exchanged throughout the world every day via the Internet, is leading to the rapid proliferation of many variants of each macro virus. Essentially, macro viruses are spreading and mutating so fast that anti-virus software designed to detect and remove them is obsolete soon after it is shipped to users.

Stopping Macro Viruses Requires New Approaches

The solution is to supplement pattern matching with a more sophisticated technique--analyzing the behavior of each macro and determining whether the macro's execution would lead to malicious acts.

This enables detection and cleaning of even those macro viruses that have not yet been captured and analyzed. But implementing this approach is not easy, requiring intelligent, rule-based scanning.

A rule-based scanning engine should complement pattern matching with algorithms to examine macro commands embedded in word processed and spreadsheet files and identify malicious code. This type of solution should also instantly detects and cleans known and unknown macro viruses, eliminating the time-consuming steps that traditional virus approaches require (see Figure 5).

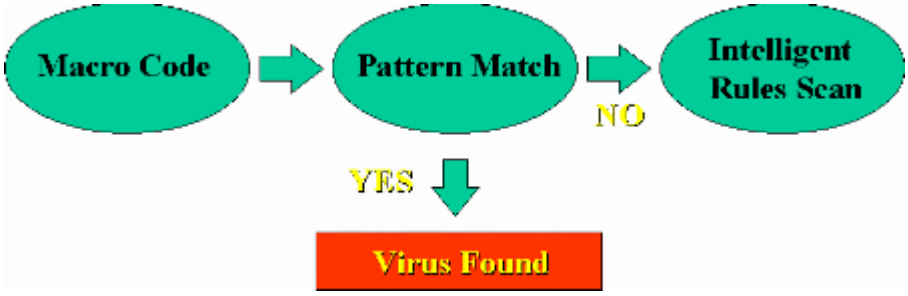


Figure 5. A new approach to stopping macro viruses detects and removes even previously unknown macro viruses from word processed and spreadsheet files.

To efficiently extract only the macro portion of each word processed or spreadsheet file it examines, this new approach is based on OLE2 (object linking and embedding) technology. Files such as those created in Word are also based on OLE2 structure, which organizes each file into discreet components (e.g., document and objects).

This new approach examines the document portion of the file only to identify key information about the macros that accompany the document, such as the locations of the macros (i.e., which "object" locations contain macros, as expressed in the macro table). The anti-virus technology does not scan the (sometimes very long) text portion of the file, since this portion cannot contain viruses. In addition to maintaining high-speed scanning performance, this approach reduces the likelihood of false positive virus indications -- possible when large text files are scanned.

After extracting the macro code, this approach compares it with patterns from known viruses. If a match is found, the user is alerted. Otherwise, the anti-virus software applies a comprehensive set of intelligent binary rules that can detect the presence of almost all macro viruses. For example, if the macro code indicates it would reformat a hard drive without prompting the user for approval to do so, the user would be alerted of the virus. This is one part of several sets of such checks that are performed. Since some macro viruses are activated when files are simply opened, virus detection is performed on files before they are even opened by any application.

Macro Virus Dependencies:

Application Popularity- The more common and "horizontal" the application, the greater the risk. More specialized or vertical market-specific programs aren't attractive enough to offer a large "breeding ground" for macro viruses.
Macro Language Depth- The extent of the application's macro language affects a virus writer's ability to create a successful macro virus. Macro Implementation- Not all programs embed macro commands into data files. For instance, AmiPro documents will not necessarily contain "invisible" macro information. The easier it is to transfer and execute the macro from within the application, the faster the spread of the virus.

Major Virus Classifications:	BOOT			FILE			MACRO	
Sub-Classes:	Encrypted	Stealth: Passive or Active		Multipartite		Polymorphic	Application	
Examples:	*Cascade*	*Tequila*	*Michelangelo*	*Natas*	*Jerusalem*	*SMEG*	*Word.Concept*	*Laroux*

Current estimates recognize as many as 8,500 distinct viruses. Rate of growth is estimated at 100 to 200 new viruses each month.

7.3 Is It a Virus?

Viruses Are Often Blamed for Non-Virus Problems

As awareness of computer viruses has grown, so has the tendency to blame "some kind of virus" for any and every type of computing problem.

In fact, more cases of "not a virus" are encountered by customer support staff at anti-virus vendors than are actual virus infections, and not only with inexperienced

users. Typical symptoms of viral infection such as unusual messages, screen color changes, missing files, slow operation, and disk access or space problems may all be attributable to non-virus problems.

Possible culprits include lost CMOS data due to a faulty system battery, another user's misuse, fragmented hard disks, reboot corruption, or even a practical joke. For instance, some PCs play the Happy Birthday song through their speakers every November 13. Sounds like a virus payload, but it happens only in computers containing BIOS chips from a certain batch that was sabotaged by a former programmer at the BIOS vendor. Switching out the BIOS chip eliminates the annual singing message.

Even deliberately written unwelcome programs are not always viruses...

As stated before, a multitude of hardware and software incompatibilities and/or bugs may cause virus-like symptoms, but there is also the in-between world of destructive, deliberately designed programs which still are not viruses. Again, it is important to remember that the key distinction of viruses is their ability to replicate and spread without further action by their perpetrators. Some non-virus programs are more destructive than many actual viruses.

Non-virus threats to user systems include Worms, Trojan Horses and Logic Bombs. In addition to the potential for damage these programs can bring by themselves, all three types can also be used as vehicles for virus program propagation.

7.3.0 Worms

Network worm programs use network connections to spread from system to system, thus network worms attack systems that are linked via communications lines. Once active within a system, a network worm can behave as a computer virus, or it could implant Trojan horse programs or perform any number of disruptive or destructive actions. In a sense, network worms are like computer viruses with the ability to infect other systems as well as other programs. Some people use the term virus to include both cases.

To replicate themselves, network worms use some sort of network vehicle, depending on the type of network and systems. Examples of network vehicles include:

- a network mail facility, in which a worm can mail a copy of itself to other systems,
- a remote execution capability, in which a worm can execute a copy of itself on another system,
- a remote login capability, whereby a worm can log into a remote system as a user and then use commands to copy itself from one system to the other.

The new copy of the network worm is then run on the remote system, where it may continue to spread to more systems in a like manner. Depending on the size of a network, a network worm can spread to many systems in a relatively short amount of time, thus the damage it can cause to one system is multiplied by the number of systems to which it can spread.

A network worm exhibits the same characteristics as a computer virus: a replication mechanism, possibly an activation mechanism, and an objective. The replication mechanism generally performs the following functions:

- searches for other systems to infect by examining host tables or similar repositories of remote system addresses
- establishes a connection with a remote system, possibly by logging in as a user or using a mail facility or remote execution capability
- copies itself to the remote system and causes the copy to be run

The network worm may also attempt to determine whether a system has previously been infected before copying itself to the system. In a multi-tasking computer, it may also disguise its presence by naming itself as a system process or using some other name that may not be noticed by a system operator.

The activation mechanism might use a time bomb or logic bomb or any number of variations to activate itself. Its objective, like all malicious software, is whatever the author has designed into it. Some network worms have been designed for a useful purpose, such as to perform general "house-cleaning" on networked systems, or to use extra machine cycles on each networked system to perform large amounts of computations not practical on one system. A network worm with a harmful objective could perform a wide range of destructive functions, such as deleting files on each affected computer, or by implanting Trojan horse programs or computer viruses.

Two examples of actual network worms are presented here. The first involved a Trojan horse program that displayed a Christmas tree and a message of good cheer (this happened during the Christmas season). When a user executed this program, it examined network information files, which listed the other personal computers that could receive mail from this user. The program then mailed itself to those systems. Users who received this message were invited to run the Christmas tree program themselves, which they did. The network worm thus continued to spread to other systems until the network was nearly saturated with traffic. The network worm did not cause any destructive action other than disrupting communications and causing a loss in productivity [BUNZEL88].

The second example concerns the incident whereby a network worm used the collection of networks known as the Internet to spread itself to several thousands of computers located throughout the United States. This worm spread itself automatically, employing somewhat sophisticated techniques for bypassing the systems' security mechanisms. The worm's replication mechanism accessed the systems by using one of three methods:

- it employed password cracking, in which it attempted to log into systems using usernames for passwords, as well as using words from an on-line dictionary
- it exploited a trap door mechanism in mail programs which permitted it to send commands to a remote system's command interpreter
- it exploited a bug in a network information program which permitted it to access a remote system's command interpreter

By using a combination of these methods, the network worm was able to copy itself to different brands of computers, which used similar versions of a widely used operating system. Many system managers were unable to detect its presence in their systems, thus it spread very quickly, affecting several thousands of computers within two days. Recovery efforts were hampered because many sites disconnected from the network to prevent further infections, thus preventing those sites from receiving network mail that explained how to correct the problems.

It was unclear what the network worm's objective was, as it did not destroy information, steal passwords, or plant viruses or Trojan horses. The potential for

destruction was very high, as the worm could have contained code to effect many forms of damage, such as to destroy all files on each system.

7.3.1 Trojan Horses

A Trojan horse program is a useful or apparently useful program or command procedure containing hidden code that, when invoked, performs some unwanted function. An author of a Trojan horse program might first create or gain access to the source code of a useful program that is attractive to other users, and then add code so that the program performs some harmful function in addition to its useful function. A simple example of a Trojan horse program might be a calculator program that performs functions similar to that of a pocket calculator. When a user invokes the program, it appears to be performing calculations and nothing more, however it may also be quietly deleting the user's files, or performing any number of harmful actions. An example of an even simpler Trojan horse program is one that performs only a harmful function, such as a program that does nothing but delete files. However, it may appear to be a useful program by having a name such as CALCULATOR or something similar to promote acceptability.

Trojan horse programs can be used to accomplish functions indirectly that an unauthorized user could not accomplish directly. For example, a user of a multi-user system who wishes to gain access to other users' files could create a Trojan horse program to circumvent the users' file security mechanisms. The Trojan horse program, when run, changes the invoking user's file permissions so that the files are readable by any user. The author could then induce users to run this program by placing it in a common directory and naming it such that users will think the program is a useful utility. After a user runs the program, the author can then access the information in the user's files, which in this example could be important work or personal information. Affected users may not notice the changes for long periods unless they are very observant.

An example of a Trojan horse program that would be very difficult to detect would be a compiler on a multi-user system that has been modified to insert additional code into certain programs as they are compiled, such as a login program. The code creates a trap door in the login program, which permits the Trojan horse's author to log onto the system using a special password. Whenever the login program is recompiled, the compiler will always insert the trap door code into the program; thus, the Trojan horse code can never be discovered by reading the login program's source code. For more information on this example, see [THOMPSON84].

Trojan horse programs are introduced into systems in two ways, they are initially planted and unsuspecting users copy and run them. They are planted in software repositories that many people can access such as on personal computer network servers, publicly accessible directories in a multi-user environment, and software bulletin boards. Users are then essentially duped into copying Trojan horse programs to their own systems or directories. If a Trojan horse program performs a useful function and causes no immediate or obvious damage, a user may continue to spread it by sharing the program with other friends and co-workers. The compiler that copies hidden code to a login program might be an example of a deliberately planted Trojan horse that could be planted by an authorized user of a system, such as a user assigned to maintain compilers and software tools.

7.3.2 Logic Bombs

Logic Bombs are a favored device for disgruntled employees who wish to harm their company after they have left its employ. Triggered by a timing device, logic bombs

can be highly destructive. The "timer" might be a specific date (i.e., the logic bomb that uses Michelangelo's birthday date to launch "his" virus embedded within). An event can also be the designed-in trigger (such as after the perpetrator's name is deleted from a company's payroll records).

7.3.3 Computer Viruses

Computer viruses, like Trojan horses, are programs that contain hidden code, which performs some usually unwanted function. Whereas the hidden code in a Trojan horse program has been deliberately placed by the program's author, the hidden code in a computer virus program has been added by another program, that program itself being a computer virus or Trojan horse. Thus, computer viruses are programs that copy their hidden code to other programs, thereby infecting them. Once infected, a program may continue to infect even more programs. In due time, a computer could be completely overrun as the viruses spread in a geometric manner.

An example illustrating how a computer virus works might be an operating system program for a personal computer, in which an infected version of the operating system exists on a diskette that contains an attractive game. For the game to operate, the diskette must be used to boot the computer, regardless of whether the computer contains a hard disk with its own copy of the (uninfected) operating system program. When the computer is booted using the diskette, the infected program is loaded into memory and begins to run. It immediately searches for other copies of the operating system program, and finds one on the hard disk. It then copies its hidden code to the program on the hard disk. This happens so quickly that the user may not notice the slight delay before his game is run. Later, when the computer is booted using the hard disk, the newly infected version of the operating system will be loaded into memory. It will in turn look for copies to infect. However, it may also perform any number of very destructive actions, such as deleting or scrambling all the files on the disk.

A computer virus exhibits three characteristics: a replication mechanism, an activation mechanism, and an objective.

The replication mechanism performs the following functions:

- searches for other programs to infect
- when it finds a program, possibly determines whether the program has been previously infected by checking a flag
- inserts the hidden instructions somewhere in the program
- modifies the execution sequence of the program's instructions such that the hidden code will be executed whenever the program is invoked
- possibly creates a flag to indicate that the program has been infected

The flag may be necessary because without it, programs could be repeatedly infected and grow noticeably large. The replication mechanism could also perform other functions to help disguise that the file has been infected, such as resetting the program file's modification date to its previous value, and storing the hidden code within the program so that the program's size remains
the same.

The activation mechanism checks for the occurrence of some event. When the event occurs, the computer virus executes its objective, which is generally some unwanted, harmful action. If the activation mechanism checks for a specific date or time before executing its objective, it is said to contain a time bomb. If it checks for a

certain action, such as if an infected program has been executed a preset number of times, it is said to contain a logic bomb. There may be any number of variations, or there may be no activation mechanism other than the initial execution of the infected program.

As mentioned, the objective is usually some unwanted, possibly destructive event. Previous examples of computer viruses have varied widely in their objectives, with some causing irritating but harmless displays to appear, whereas others have erased or modified files or caused system hardware to behave differently. Generally, the objective consists of whatever actions the author has designed into the virus.

As with Trojan horse programs, computer viruses can be introduced into systems deliberately and by unsuspecting users. For example, a Trojan horse program whose purpose is to infect other programs could be planted on a software bulletin board that permits users to upload and download programs. When a user downloads the program and then executes it, the program proceeds to infect other programs in the user's system. If the computer virus hides itself well, the user may continue to spread it by copying the infected program to other disks, by backing it up, and by sharing it with other users. Other examples of how computer viruses are introduced include situations where authorized users of systems deliberately plant viruses, often with a time bomb mechanism. The virus may then activate itself at some later point in time, perhaps when the user is not logged onto the system or perhaps after the user has left the organization.

7.3.4 Anti-Virus Technologies

Without control of the "human element" and proper implementation, anti-virus software alone cannot provide full protection.

However, it is still the critical element in the fight against viruses. As stated before, non-virus problems may appear to be virus related, even to sophisticated users. Without anti-virus software, there is no conclusive way to rule out viruses as the source of such problems and then arrive at solutions.

Effective anti-virus software must be capable of performing three main tasks: Virus Detection, Virus Removal (File Cleaning) and Preventive Protection. Of course, detection is the primary task ad the anti-virus software industry has developed a number of different detection methods, as follows.

Five Major Virus Detection Methods:

- Integrity Checking (aka Checksumming) - Based on determining, by comparison, whether virus-attacked code modified a program's file characteristics. As it is not dependent on virus signatures, this method does not require software updates at specific intervals.

- Limitations - Does require maintenance of a virus-free Checksum database; allows the possibility of registering infected files; Unable to detect passive and active stealth viruses; Cannot identify detected viruses by type or name.

- Interrupt Monitoring - Attempts to locate and prevent a virus "interrupt calls" (function requests through the system's interrupts).

- Limitations - Negative effect on system resource utilization; May flag "legal" system calls and therefore be obtrusive; Limited success facing the gamut of virus types and legal function calls.

- Memory Detection - Depends on recognition of a known virus' location and code while in memory; Generally successful.

- Limitations - As in Interrupt Monitoring, can impose impractical resource requirements; Can interfere with valid operations.

- Signature Scanning - Recognizes a virus' unique "signature," a pre-identified set of hexadecimal code, making it highly successful at virus identification.

- Limitations - Totally dependent on maintaining current signature files (as software updates from vendor) and scanning engine refinements; May make false positive detection in valid file.

- Heuristic/Rules-based Scanning - Faster than traditional scanners, method uses a set of rules to efficiently parse through files and quickly identify suspect code (aka Expert Systems, Neural Nets, etc.).

- Limitations - Can be obtrusive; May cause false alarms; Dependent on the currency of the rules set.

All five techniques can usually perform on-access or on-demand scans, for both network servers and work-stations. On-access scanning is analogous to a building" automatic sprinkler system –virus scanning is automatically initiated on file access, such as when a disk is inserted, a file is copied or a program is executed. On-demand scanning is more like a fire extinguisher - requiring user initiation (but may also be set up to continue scanning at regular intervals or at system startup).

Today, all effective products leverage a combination of detection methods because of the large number of virus types and their many tricks for invasion and disguise. Anti-virus software is a constantly evolving field, and as the knowledge base deepens, vendors can further refine these methods and develop even more effective future solutions.

7.4 Anti-Virus Policies and Considerations

The best anti-virus software in the world cannot protect you if it is not deployed systematically throughout the enterprise (even if "the enterprise" is a single home-based computer!).

Many people think they can dismiss a disk, shared or e-mailed file because it came from someone they know and trust. What they aren't considering is that their friend colleague, customer or vendor is working on another system, with its own set of vulnerabilities from different outside conditions.

Computer users must recognize that the virus threat is too pervasive today to be ignored by anyone...the number of users who never come into contact with others' files is small and becoming smaller every day, especially with the tremendous growth of online services and Internet usage.

7.4.0 Basic "Safe Computing" Tips

- Use and update anti-virus software regularly
- Scan any newly received disks and files before loading, opening, copying, etc.
- Never assume disks and/or files are virus-free
- To help avoid boot viruses, do not leave diskettes in your computer when shutting it down.
- Change your computer's CMOS boot sequence to start with the C drive first, then the A drive.

For offices or homes with one or two computers, following these basic rules faithfully is probably adequate protection. However, in organizations with multiple PCs, especially in networks, a sound anti-virus strategy will necessarily be more complex.

This is because vulnerability to viruses increases in proportion to the number of machines, the extent of their interconnection, and the number of non-technical users who may view anti-virus vigilance as "someone else's job." (In contrast, a solo entrepreneur is likely to take the virus threat seriously because he or she will have to deal with infection results personally or pay an outside consultant.)

All organizations are different in the way they operate and the industries they serve, so no one anti-virus scheme is correct for all enterprises. However, at the very least, a company's program should include ongoing user education and a system for tracking virus activity (suspect and real) in addition to using anti-virus software.

Ultimately, your goal is to provide consistent, effective protection and a "damage control and recovery" plan for virus infections that may occur despite your efforts. In addition, and perhaps most importantly, you want to achieve this while minimizing any negative impact on staff productivity and system/network resources.

Therefore, to formulate a comprehensive anti-virus plan, it is necessary to first analyze the "bit picture" of your organization along with its more detailed computing characteristics.

5 Key Factors in Anti-Virus Program Planning

1. The number and density of personal computers
   The more PCs you have, or the higher the ratio of computers to people, the more you need a formalized, thoroughly documented anti-virus program.

2. The degree of interconnection between computers
   "Interconnection" does not necessarily mean electronically networked. If data is frequently moved from one PC to another via diskettes or other media, those computers are effectively connected, whether they are separated by a few yards or many miles. Again, the frequency of data interchange may be as important as the methods of transfer.

3. How many locations are involved in the anti-virus plan
   Assuming that multiple locations are involved because they are linked via data communications, more locations will require more coordination and reporting between the various IT staffs, as well as more user training.

4. The operational pace of the enterprise
   Every organization has an inherent pace of operations, mostly dependent on the nature of its business. No matter how "busy" it is, a research laboratory's pace will not be as fast as that of a securities brokerage firm. In general, the faster the pace of operations, the greater the risk of virus infection because of the faster rate at which new data is being generated and distributed. faster pace = more frequent new data = greater risk !

5. Whether there is a high level of transaction processing
   If massive and timely data exchange is typical, the plan must yield the highest possible level of anti-virus security, along with comprehensive backup. Even weekly backups won't be adequate if vital data captured in real-time has been violated by a virus infection since the last backup.

Balance: Implementing Security by Function

Whatever the profile of your organization's computing characteristics and virus vulnerability, it is important to remember that anti-virus measures must be balanced in relation to the actual functions of various machines and their users.

Even within a specific location of the enterprise, there may be computers for which you need to sacrifice some level of anti-virus security in order to maintain necessary throughput and/or productivity. Cost is another factor that must be balanced against "ideal" protection levels, for all equipment and personnel in the organization.

7.4.1 Anti-Virus Implementation Questions

- Are there any PCs that should not be included in the anti-virus program? (For instance, computers that are isolated, diskless or used solely for manual data entry.)
- What special procedures should apply to the headquarters network, as opposed to branch offices?
- How should user reports of suspected virus activity be handled? What is a realistic (vs desired) response time?
- In response to an apparent virus infection, what procedures should users be authorized and trained to perform by themselves?
- How should suspected and/or actual virus infections, and resulting counter measures, be recorded and reported? (It is important to log routine anti-virus scans as well as suspicious situations.)
- Who is responsible for maintaining these possibly exhaustive records?
- What improvements to existing backup procedures might be necessary? (Note that the common practice of rotating backup media might cause clean data to be replaced by infected data.)
- An anti-virus policy and procedures manual will need to be created and then maintained...who will take charge?
- How will you establish a "baseline" virus-free environment for the new anti-virus program to maintain?
- How will the schedule for adoption of a new virus control program be established? How will you balance simultaneous needs for speed and low cost?
- Who will provide the funding for the anti-virus program staff, development and software? Is upper management fully behind the program?

7.4.2 More Virus Prevention Tips

- Write-protect any data source diskette before inserting it in the drive, and then use anti-virus software to scan it before doing anything else.
- Include in your policy and training that employees who work on computers at home must follow the same anti-virus procedures they use at the office (whether on personal machines or company-supplied portables.)
- Even with the above policy in place, handle disks brought back from employees' homes as foreign disks, following the write-protect and scanning procedure
- Consider any suspicious computer behavior to be possible virus-related and follow-up accordingly.
- Files that must be received from outside the organization, such as from the Internet, should be downloaded directly to quarantined scanning areas whenever possible.
- You may want to consider dedicating an isolated computer (not connected in any way to the network) to the task of testing all new files and/or diskettes. Then all files on the control machine can be systematically scanned for viruses before anyone has access to them. (Note that some compressed files may have to be decompressed before scanning.)

Take Advantage of Vendor Expertise

The larger your network, and/or the more sensitive your enterprise's data security position, the more you should seek guidance from industry peers and the anti-virus software industry before finalizing your plan.

Representatives from the leading vendors have experience in providing anti-virus solutions for many different kinds of distributed environments, in many different industries. Plus, their training programs and consulting services can be invaluable, helping to prevent both costly virus incidents and ensuring that your program is more cost-effective.

7.4.3 Evaluating Anti-Virus Vendors

Although anti-virus software companies design their products to detect and remove viruses, there is more to making a smart choice than comparing detection rates and/or product prices.

The fact that anti-virus software is necessary for everyone in the enterprise means that it must work alongside a variety of applications, and probably on multiple computing platforms within the location. Therefore, a common anti-virus product that can work "seamlessly" throughout the enterprise is desirable, for both cost-effectiveness and simpler administration.

The software must also be effective against the majority of common and damaging viruses, yet be as unobtrusive to productivity as possible. (Bear in mind that this is as important for user compliance as for the bottom line - if users feel hampered by anti-virus procedures they may "overlook" them in their haste to get work done.)

Another major factor to consider is the burgeoning number of viruses - as many as 200 new ones each month. Anti-virus software that does not include regular updates cannot provide adequate protection for long.

7.4.4 Primary Vendor Criteria

To ensure that you are providing the best possible solution, the anti-virus vendor you ultimately choose should satisfy the following primary criteria:

- Technological Strength - Demonstrably superior virus detection rates; leadership, quality assurance and timeliness in releasing new products and updates; Good grasp of technological trends that may impact your organization in the future.
- Infrastructure - Company resources in terms of financial health and strategic alliances to provide for ongoing development; Size and experience level of customer support staff; Size and scope of current user base; Ability to handle complex contracts smoothly.
- Relationships - Vendors who offer only technological strength, or excellent service with mediocre technology, will be inferior choices for an enterprise-wide anti-virus program. To get the most out of your anti-virus efforts, base them on software from a company that can sustain long-term relationships and provide excellent anti-virus technology.

While investigating anti-virus vendors and products, be sure to also assess these cost of ownership issues:

- Types of licenses available
- Variety of platforms supported
- Cost of updates for virus signatures and product releases
- Emergency services available
- Customer training (on and/or off-site)
- Consulting services available
- Maintenance agreements
- Contract terms and guarantees

In determining what is needed from the vendor, and the best contract arrangements,, evaluators should also consider their in-house support and training resources, as well as the organization's growth potential and plans for introducing any new computing platforms.

Section References

7.1 NAI White Paper. “Current Computer Virus Threats, Countermeasures and Strategic Solutions”.1997

7.2 Landry, Linda, Trapping the World's Most Prevalent Viruses. Trend Micro, Inc. 1998

"ICSA 1997 Computer Virus Prevalence Survey, ICSA.
"Roll-Your-Own Macro Virus," Virus Bulletin, September, 1996, p. 15.
Joe Wells, "Concept: Understanding the Virus and Its Impact," Trend Micro, Incorporated.
"ICSA 1997 Computer Virus Prevalence Survey, ICSA.

7.3 NAI White Paper. “Current Computer Virus Threats, Countermeasures and Strategic Solutions”.1997

7.3.0 Wack, John P and Carnahan, Lisa J. Computer Viruses and Related Threats:A Management Guide. NIST Special Publication 500-166. U.S Dept of Commerce

BUNZEL88 Bunzel, Rick; Flu Season; Connect, Summer 1988.

DENNING88 Denning, Peter J.; Computer Viruses; American Scientist, Vol 76, May-June, 1988.

DENNING89 Denning, Peter J.; The Internet Worm; American Scientist, Vol 77, March-April, 1989.

FIPS73 Federal Information Processing Standards Publication 73, Guidelines for Security of Computer Applications; National Bureau of Standards, June, 1980.

FIPS112 Federal Information Processing Standards Publication 112, Password Usage; National Bureau of Standards, May, 1985.

MACAFEE89 McAfee, John; The Virus Cure; Datamation, Feb 15, 1989.
NBS120 NBS Special Publication 500-120; Security of Personal Computer Systems: A Management Guide; National Bureau of Standards, Jan 1985.

SPAFFORD88 Spafford, Eugene H.; The Internet Worm Program: An Analysis; Purdue Technical Report CSD-TR-823, Nov 28, 1988.

THOMPSON84 Thompson, Ken; Reflections on Trusting Trust (Deliberate Software Bugs); Communications of the ACM, Vol 27, Aug 1984.

7.3.1 Wack, John P and Carnahan, Lisa J. Computer Viruses and Related Threats:A Management Guide. NIST Special Publication 500-166. U.S Dept of Commerce

BUNZEL88 Bunzel, Rick; Flu Season; Connect, Summer 1988.

DENNING88 Denning, Peter J.; Computer Viruses; American Scientist, Vol 76, May-June, 1988.

DENNING89 Denning, Peter J.; The Internet Worm; American Scientist, Vol 77, March-April, 1989.

FIPS73 Federal Information Processing Standards Publication 73, Guidelines for Security of Computer Applications; National Bureau of Standards, June, 1980.

FIPS112 Federal Information Processing Standards Publication 112, Password Usage; National Bureau of Standards, May, 1985.

MACAFEE89 McAfee, John; The Virus Cure; Datamation, Feb 15, 1989.

NBS120 NBS Special Publication 500-120; Security of Personal Computer Systems: A Management Guide; National Bureau of Standards, Jan 1985.

SPAFFORD88 Spafford, Eugene H.; The Internet Worm Program: An Analysis; Purdue Technical Report CSD-TR-823, Nov 28, 1988.

THOMPSON84 Thompson, Ken; Reflections on Trusting Trust (Deliberate Software Bugs); Communications of the ACM, Vol 27, Aug 1984.

7.3.2 NAI White Paper. “Current Computer Virus Threats, Countermeasures and Strategic Solutions”.1997

7.3.3 Wack, John P and Carnahan, Lisa J. Computer Viruses and Related Threats:A Management Guide. NIST Special Publication 500-166. U.S Dept of Commerce

BUNZEL88 Bunzel, Rick; Flu Season; Connect, Summer 1988.

DENNING88 Denning, Peter J.; Computer Viruses; American Scientist, Vol 76, May-June, 1988.

DENNING89 Denning, Peter J.; The Internet Worm; American Scientist, Vol 77, March-April, 1989.

FIPS73 Federal Information Processing Standards Publication 73, Guidelines for Security of Computer Applications; National Bureau of Standards, June, 1980.

FIPS112 Federal Information Processing Standards Publication 112, Password Usage; National Bureau of Standards, May, 1985.

MACAFEE89 McAfee, John; The Virus Cure; Datamation, Feb 15, 1989.

NBS120 NBS Special Publication 500-120; Security of Personal Computer Systems: A Management Guide; National Bureau of Standards, Jan 1985.

SPAFFORD88 Spafford, Eugene H.; The Internet Worm Program: An Analysis; Purdue Technical Report CSD-TR-823, Nov 28, 1988.

THOMPSON84 Thompson, Ken; Reflections on Trusting Trust (Deliberate Software Bugs); Communications of the ACM, Vol 27, Aug 1984.

7..3.4 NAI White Paper. “Current Computer Virus Threats, Countermeasures and Strategic Solutions”.1997

7.4 NAI White Paper. “Current Computer Virus Threats, Countermeasures and Strategic Solutions”.1997

8.1 Making Sense of Virtual Private Networks

The VPN market is on the verge of explosive growth. A virtual private network (VPN) broadly defined, is a temporary, secure connection over a public network, usually the Internet. Though the term is relatively new, everyone from the telcos, to operating system vendors, to firewall suppliers and router companies has rushed to offer some type of VPN capability. Why? Because VPNs make sense, and as a result, the market is expected to reach at least several billion dollars by the year 2001.

By leveraging the Internet, VPNs offer significant cost savings, greater flexibility, and easier management relative to traditional internetworking methods, such as leased lines and dial-up remote access.

However, choosing an appropriate solution from the recent flood of VPN offerings can be a difficult task for information technology managers who have no spare time. Each solution presents varying levels of security, performance, and usability, and each has its benefits and drawbacks.

Though a catch-all Internet security solution sounds appealing, there is currently no product that can equally address the different aspects of securing online communication. As a result, the VPN market has begun to stratify according to corporate demands for tighter security, better performance, and effortless usability and management. To select an appropriate product, IT managers should be able to define their corporation's particular business needs. For instance, does the company only need to connect a few trustworthy remote employees to corporate headquarters, or does the company hope to create a secure communications channel for its branch offices, partners, suppliers, customers, and remote employees?

At minimum, a VPN should encrypt data over a dynamic connection on a public network to protect the information from being revealed if intercepted. Beyond that basic function, VPN features customarily include tools for authentication, and a limited number provide integrated access control and authorization capabilities. In addition to enumerating the possible VPN components, this white paper outlines the predominate VPN technologies and interprets the nuances of different VPN approaches so IS professionals can better decide how to secure their corporate communication.

8.2 Defining the Different Aspects of Virtual Private Networking

Before online business can truly reach its potential, corporations must feel comfortable using the Internet as the backbone for secure communication. VPNs are the first real step toward that end. When implemented correctly, they protect networks from viruses, snoops, corporate spies, and any other known threat that results from mistakes in configuration, poorly implemented access controls, lack of system management, weak authentication, and "back-door" entry points to the network.

Sample VPN Requirements to Consider

Security	Interoperability	Ease-of-Use
• Can the VPN support Strong authentication, including token cards, smart cards, biometrics (i.e. fingerprint and iris scanning),x.509 certificates and Kerberos? • Can the VPN support strong encryption, including key sizes 40, 56, and 128 and ciphers RC4, DES, and Triple DES? • Can the VPN filter datastreams, including viruses, file types, Java and Active X, and protocols such as FTP, Telnet, etc.? • Can the VPN support role-based access control according to parameters such as type of authentication, type of encryption, user identity, time of day, source address, destination address, and type of application? • Can the VPN monitor, log, and audit all network traffic? • Does the VPN have some type of alarm to notify an administrator of specific events?	• Is the VPN based on public standards? • Can the VPN be integrated easily with perimeter security, such as a firewall or router? • Is the VPN compatible with other protocols such as IPv4, IPSec, and PPTP/L2TP? • Can the VPN support all critical authentication and encryption standards? • Can the VPN support all application types? • Can the VPN function in a cross-platform environment, including all Windows and UNIX operating systems? • Does the VPN map to standard NT, Netware, RADIUS, and ACE databases? • Does the VPN support a variety of methods of load balancing?	• Does the VPN offer a low-impact client for the desktop? Is the client transparent to the end-user? • Does the VPN permit single sign-on, or does the user have to log on each time an application is launched? • Can the VPN system scale to support hundreds of thousands of users? • Does the VPN centralize management of the security system? • Does the VPN run on standard NT and UNIX operating systems?

The three fundamental features that define virtual private networking are encryption, authentication, and access control. While strong authentication and encryption are critical components of the VPN, they are relatively simple to deploy and verify. Access control, on the other hand, is relatively complex because its deployment is tied intimately to every other security tool. Roughly speaking, the security of a VPN is a function of how tightly authentication, encryption, and access control are connected. If one component is lacking, the VPN will be lacking.

Where a company might use a guarded gate in the physical world to block all unauthorized visitors, a firewall might be used in the analogous VPN world. Until

recently, that's as far as the comparison could be drawn, because in the VPN world there hasn't been a way to provide varying levels of access. Now, with emerging VPN technologies and solutions, companies can verify someone's identity with strong authentication technologies like token cards, digital certificates, or even fingerprints. Once identified, users are granted access to resources according to very detailed profiles based on identity and often a user's role within a larger group. VPNs are also beginning to provide tools to monitor a user's activity once inside the corporate network.

Prior to even connecting to the Internet, corporations should develop a security policy that clearly identifies who can have access to what resources, leaving room for growth and change. And before implementing a VPN, corporations should evaluate their current security paradigm to determine what equipment can be leveraged for a VPN. Once the budget has been decided on, companies should consider all of the benefits they hope to derive from a VPN, such as increased profits through streamlined processes, improved customer service from providing more personalized information faster, and stronger strategic relationships from the easy and secure exchange of information.

A comprehensive solution might incorporate a firewall, router, proxy server, VPN software or hardware, or all of the above. Any one of those pieces might sufficiently secure communication, depending on a company's priorities, but it's more likely that a combination of tools will provide the most comprehensive solution.

IS professionals can effectively use VPNs to address three predominant internetworking scenarios:

1. between a corporation and its branch offices, which will be referred to as an "intranet VPN";
2. between a corporation and its remote or traveling employees, which will be referred to in this paper as a "remote access VPN";
3. and between a corporation and its business associations, such as partners, customers, suppliers, and investors, which will be referred to as an "extranet VPN."

8.2.0 Intranet VPNs

Intranets are defined here as semi-permanent WAN connections over a public network to a branch office. These types of LAN-to-LAN connections are assumed to carry the least security risk because corporations generally trust their branch offices and view them as an extension of the corporate network.
In this case, the corporation generally controls both the source and destination nodes. IS administrators should ask whether or not this assumption holds true for their company.

General Case

When the two endpoints of a data channel are relatively trusted, a company can comfortably opt for a VPN solution that focuses on performance over security, which is limited to the strength of the encryption and authentication methods between the two routers. High volumes of data are often exchanged between LANs on an intranet VPN, so the premium is wisely placed on speed and smooth interoperability. The LANs that are connected by centralized corporate databases or other enterprise-wide computing resources should appear to be part of the same corporate network. Many of the firewall, router, and frame relay vendors, as well as

some of the ISPs, are offering solutions that adequately secure intranet VPNs while transferring data quickly and reliably.

Highly Secure Case

Security threats often come from within an organization. In fact, according to a study issued jointly by the FBI and the Computer Security Institute, almost half of all computer break-ins occur within a company.

If a company is concerned about proprietary information being leaked by employees, whether intentionally or accidentally, or if a company routinely applies different levels of trust to branch offices or individuals, then it should consider investing in a VPN solution that can control the information flow on an authenticated, user-specific policy level rather than on a trusted subnet basis. IT managers should look closely at solutions that provide reasonable ways to implement and manage these advanced role-based policies.

8.2.1 Remote Access VPNs

Corporations are just now beginning to realize the advantages the Internet offers over traditional direct dial-up remote access. Many corporations, burdened by the effort of maintaining large modem pools and the expense associated with long distance charges, are finding that using the Internet as a backbone for remote access is much more affordable and easier to implement and maintain than traditional solutions.
In any remote access VPN scenarrio, usability is an importan criterion. Most security flaws are attributed to configuration errors, so the easier the system is to manage, the less likely is the chance for oversight. On the client side, simplicity is critical because many traveling employees and telecommuters either lack the technical proficiency or the access to technical resources for troubleshooting. Clients should not have to manually build a VPN tunnel, "manually" meaning having to launch VPN software each time the user wants to establish a secure communication channel. Instead, the VPN software should launch automatically at start-up and run transparently in the background. On the server side, centralized and easy management is essential because monitoring large numbers of users and adding and removing users on a regular basis can quickly become chaotic and can create a security risk.

General Case

With most remote access VPNs, it is assumed that a corporation trusts the person at the other end of the link, which is typically a traveling or remote salesperson. Rather than worrying that the employee might do damage to the network or steal proprietary information, the company is probably more concerned with the unknown element between the two end points. These companies will generally assume a "transparent access" policy, best described as: "The remote employee should have unfettered access to all resources that would be available to them if they were sitting at their desk at corporate headquarters."

The priority, therefore, becomes encrypting the data in transit so that only the intended recipient can decipher it. Most VPNs can meet this basic security requirement, so evaluators should consider additional criteria, such as the strength of the encryption cipher and the authentication method for providing additional security.

Highly Secure Case

The industries that are the most leery of any kind of security risk, such as the financial, health, and government sectors, are paradoxically the earliest adopters of VPN technologies, which have the perception of being less secure than traditional means of networking. In reality, the best VPN technologies are much more secure than most leased lines and dial-up remote access, because highly secure VPNs encrypt all data and generally provide very detailed user profiles for access control.

Highly secure remote access solutions are deployed by sophisticated IT shops with a strong understanding of the security risks inherent in any network communication. These shops generally adopt a "controlled access" policy for their remote users. This is best described by the following policy statement: "The remote employee should have tightly controlled access to specific resources on the network according to the requirements of their job function."

These companies deploy policy-driven VPNs to provide highly secure remote access over the public networks. Secure policy-driven VPNs authenticate individual users, not just IP addresses, so that a corporation knows which employee is trying to gain access to the network. This can be accomplished through common passwords digital certificates, token cards, smart cards, or biometrics, such as fingerprint or iris scanning. Once an employee has authenticated to the corporate VPN server, he or she is granted a certain level of access depending on his or her profile, which is usually set up by a network administrator to match the corporate security policy and enforced by a sophisticated system of datastream filters and access control parameters. This three-tier system is essential for companies that allow their employees to access mission-critical information, particularly when those employees are not entirely trusted.

Any time a company wants to provide varying levels of access so that different resources can be made available to different employees when appropriate, or when a company wants to prevent "back-door" holes into the network, which is common in some systems, then a more robust VPN solution is recommended. In other words, a highly secure VPN should be able to intercept network traffic destined for a particular host, add the required encryption, identify individual users, and apply restrictions and filter content accordingly.

8.2.2 Extranet VPNs

Unlike intranets that are relatively isolated, extranets are intended to reach partners, customers, and suppliers, as well as remote employees. Securing that wide area network requires diligence and the right tools. An extranet VPN needs to be able to provide a hierarchy of security, with access to the most sensitive data being nested under the tightest security control. It should secure all applications, including TCP and UDP applications, such as Real Audio, FTP, etc.; corporate vertical applications, such as SAP, BAAN, PeopleSoft, Oracle, etc.; and "homegrown" applications, such as Java, Active X, Visual Basic, etc. Because most corporate computing environments are heterogeneous with many legacy systems, a sound VPN solution should be extremely versatile and interoperable with multiple platforms, protocols, and authentication and encryption methods.

General vs. Highly Secure Case

The main objective of an extranet or business-to-business VPN is to ensure that mission-critical data arrive intact and in the proper hands without ever exposing

protected resources to potential threats, so companies should only consider implementing the most secure breed of VPNs.

The security elements of a VPN can be prioritized differently, but with an extranet VPN, all the fundamental pieces 3/4 encryption, authentication, and access control 3/4 should be integrated tightly with some type of perimeter security. Usually this means a company will place a VPN proxy server behind an impenetrable firewall that blocks all unauthenticated traffic. Any traffic that is allowed in is then funneled through a common portal directly to the VPN server, which filters traffic according to company policy. It is essential for the connection between the firewall and the VPN to be strong and reliable, and the client software should be as transparent as possible.

8.3 VPN Architecture

The most secure VPNs are built around a "directed" architecture, as opposed to a bi-directional "tunneled" method. Directed VPNs transmit encrypted information at a higher level in the networking protocol stack than tunneled VPNs, and security and control increase as functionality moves up the network hierarchy. Directed VPNs act as proxy servers, which means they do not open any direct connections into corporate networks, preventing IP addresses from being "spoofed," or mapped. Tunneling hides information in IP packets at the packet level, exposing them more easily to attack. Because all data is proxied in directed VPNs, administrators can tell at a glance who has been trying to gain access to the network and how often.

Unlike tunneled VPNs, directed VPNs protect connected networks from each other’s security flaws. Directed VPNs do not assume a two-way trusted relationship between connecting parties. If security is breached in the directed model, only the attacked network is exposed, not the linked networks. In the tunneled model, when one network is attacked, each successive network is susceptible to the same attacker. In the directed model, each company's IS managers can set their own access privileges and be confident they are not exposing their networks to unknown security problems.

Tunneled VPNs, as the name implies, open tunnels within the Internet and secure information traveling through them with basic packet filtering. This approach gives participating companies weakly secured access to each other's networks, with no way to fine-tune access control. These types of solutions often mistakenly start with the faulty assumption that there should be peer-to-peer trust among companies connected by VPNs. When trading partners or customers are involved, that is rarely the reality.

When companies conduct multi-faceted business transactions over public networks, simple encrypted tunnels will not suffice. Online business, or electronic commerce, is not restricted to credit card transactions. It involves complex negotiations and collaboration on projects. When vital, confidential information is involved, IS professionals cannot risk compromising any portion of the network. An extranet VPN should use the highest encryption available, which is currently 128 bits, except when restricted by exportation laws. In addition, the VPN should support multiple authentication and encryption methods since business partners, suppliers, and customers are likely to have varying network infrastructures and platforms.

In a true business-to-business scenario, IS managers should look for a VPN that filters access to resources based on as many parameters as possible, including source, destination, application usage, type of encryption and authentication used, and individual, group, and subnet identity. Administrators should be able to identify

individual users, not just IP addresses, either through passwords, token cards, smart cards, or any other method of authentication. Passwords are usually sufficient for casual office use, but they are not considered as secure as token or smart cards. Employees are often careless with their passwords, and they rarely change their codes, whereas token and smart cards change the passcode on a regular basis, often as frequently as every 60 seconds.

Once authenticated, administrators should be able to route authorized traffic to protected resources without jeopardizing network security. The access control is what ultimately distinguishes the level of security among VPN solutions. Without being able to control exactly who has access to each resource on a network, a VPN is virtually useless beyond the network's perimeter. Once authenticated, a user should not have carte blanche to the network. Rather, specific permissions should be granted to each user in order to retain the most control over every resource. Security should increase, not lessen, as a user moves inward toward the most sensitive data. By utilizing strong encryption, authentication, and access control methods, all working seamlessly within a VPN solution, companies can seal their corporate networks from almost any security breach.

8.4 Understanding VPN Protocols

The VPN security market is young, and standards are still evolving, but a handful of protocols have emerged as the leading choices for building VPNs. An IS manager should not have to base his or her purchasing decision on the technology used, but understanding the benefits of each protocol may help clarify the related strengths and weaknesses of different VPN end products. Although there are many possible security approaches for creating a VPN, the following protocols show the most promise for lasting in the market, whether for the quality of their design or their financial backing.

8.4.0 SOCKS v5

SOCKS v5 was originally approved by the IETF as a standard protocol for authenticated firewall traversal, and, when combined with SSL, it provides the foundation for building highly secure VPNs that are compatible with any firewall. It is most appropriately applied to VPNs that require the highest degree of security, since its strength is access control.

SOCKS v5 was developed in 1990 by David Koblas and championed through the IETF by NEC Systems Laboratory. It is currently the only IETF-approved standard being used to create VPNs. Though it is not as well known as some of the other protocols, it has received widespread support from industry leaders such as Microsoft, Netscape, and IBM. SOCKS v5 is the protocol used in Aventail’s policy-based VPN solution.

Advantages

SOCKS v5 controls the flow of data at the session, or circuit, layer, which maps approximately to layer five of the OSI networking model. Because of where it functions in the OSI model, SOCKS v5 provides far more detailed access control than protocols operating at the lower layers, which permit or reject packets based solely on source and destination IP addresses. SOCKS v5 establishes a virtual circuit between a client and a host on a session-by-session basis and provides monitoring and strong access control based on user authentication without the need to reconfigure each new application.

Because SOCKS v5 and SSL operate at the session layer, they have the unique ability to interoperate on top of IPv4, IPSec, PPTP, L2TP, or any other lower-layer VPN protocol. In addition, SOCKS v5 and SSL have more information about the applications running above them than do lower-layer protocols, so they can provide very sophisticated methods of securing traffic.

SOCKS v5 stands out as the only VPN approach to use a directed architecture, which essentially protects destination computers by proxying traffic between source and destination computers. When used in conjunction with a firewall, data packets are passed through a single port in the firewall (port 1080 by default) to the proxy server, which then filters what is sent forward to a destination computer. This prevents administrators from having to open multiple holes in their firewall for different applications. For additional security, the VPN proxy server hides the address structure of the network, making it more difficult for confidential data to be cracked. Another design advantage of SOCKS v5 is that the client is non-intrusive. It runs transparently on the user's desktop and does not interfere with networking transport components, as do lower-layer protocols, which often replace the Winsock DLL, TCP/IP stack, and low-level drivers, interfering with desktop applications.

SOCKS v5 is also highly flexible. It works easily with multiple security technologies and platforms, which is critical for IS professionals managing heterogeneous computing environments. It offers modular plug-in support for many authentication, encryption, and key management methods, providing IS managers the freedom to adopt the best technologies for their needs. Plug-and-play capabilities include access control tools, protocol filtering, content filtering, traffic monitoring, reporting, and administration applications. SOCKS v5 can filter data streams and applications, including Java applets and ActiveX controls, according to very detailed specifications.

SOCKS v5 is the only VPN protocol that can interoperate with other VPN protocols, such as PPTP, IPSec, and L2TP, and it is ready for implementation today. Because the SOCKS v5 protocol is designed specifically for highly secure environments, analysts expect that SOCKS v5 and appropriate plug-ins will be used primarily for highly secure remote access and the extension of private client networks across multiple organizational perimeters.

Disadvantages

Because SOCKS v5 adds a layer of security by proxying traffic, its performance generally is slightly slower than that of lower-layer protocols, depending on how the VPN is implemented. Though it is more secure than solutions located at the lower network or transport layers, the extra security requires more sophisticated policy management than at the lower layers. Also, client software is required to build a connection through the firewall to transmit all TCP/IP data through the proxy server.

8.4.1 PPTP/L2TP

One of the most widely known VPN security choices is Point-to-Point Tunneling Protocol (PPTP) from Microsoft. It is embedded in Microsoft’s Windows NT v4.0 operating system and is used with Microsoft’s Routing and Remote Access Service. It sits at the datalink layer, which maps approximately to layer two of the OSI model. It encapsulates PPP with IP packets and uses simple packet filters and the Microsoft Domain networking controls to provide access control. PPTP and its successor, L2TP, are seen as tools to extend the current PPP dial-up infrastructure supported by Microsoft, most ISPs, and the remote access hardware vendors.

Layer Two Transport Protocol (L2TP) has evolved from the combination of Microsoft's PPTP protocol and Cisco Systems' Layer 2 Forwarding (L2F). It supports multiple, simultaneous tunnels for a single client and is targeted at the telco and ISP markets. With L2TP, the end user dials up a local ISP POP without encryption, and the ISP, acting as an agent for the end user, creates an encrypted tunnel back into the secure destination.

PPTP and L2TP have received broad support from the current leaders in the remote access services market, which includes Cisco, Bay Networks, 3Com, Shiva, and Microsoft, because they provide an effective way for these vendors to migrate their current corporate dial-up products to Internet-based methods of building tunnels. Analysts predict that PPTP and L2TP will play a dominant role in the Internet-based remote access market when security requirements are relatively low.

Advantages

IS professionals running Microsoft-centric shops will find PPTP and L2TP ready-made to work with their systems. Because they use packet-filtering that makes use of existing network routers, they are typically less complicated to implement, and they are transparent to end users.

In typical Microsoft fashion, PPTP is free. Microsoft includes it as a component of its RAS and router software, formerly known as Steelhead. When affordability in a Microsoft-only environment is an issue, PPTP is a viable solution. L2TP will likely follow the same path and be included in upcoming versions of NT servers and Windows 32-bit desktop clients.

Most VPNs secure TCP/IP traffic, but PPTP and L2TP support additional networking protocols such as Novell's IPX, NetBEUI, and AppleTalk. They also support flow control, which keeps traffic from overwhelming clients and servers. They enhance network performance by minimizing dropped packets, thus cutting down on re-transmission.

Disadvantages

PPTP and L2TP are typical tunneled approaches to VPN security, which means they encapsulate non-secure IP packets within secure IP packets. They use IP frames to create an open data passageway between two computer systems. Once a tunnel is open, source and destination identification is no longer required. The tunnel is bi-directional, so while it encrypts data traveling along the Internet, it does not provide a way to monitor or control what gets passed between the two points.

One often overlooked limitation is that PPTP and L2TP are limited to 255 concurrent connections. In addition, end users are required to manually establish a tunnel prior to connecting to the intended resource, which can be a hassle. Also, the selection of authentication and encryption standards is very limited, and currently no strong encryption or authentication is supported.

Another concern is that there are currently no versions of PPTP or L2TP available for older Microsoft operating systems or UNIX. PPTP is still very narrowly targeted for Microsoft-specific networking.

PPTP and L2TP are currently only proposed standards. PPTP is presently supported by Microsoft's Windows NT 4.0 server, NT workstation, and Windows 95.

Remote access vendors, such as Ascend and Shiva, are backing L2TP, and Microsoft plans to incorporate L2TP into Windows NT server version 5.0.

8.4.2 IPSec

Internet Protocol Security (IPSec) has gained a lot of recent attention in the industry. It evolved from the IPv6 movement, and as a standard promoted by the IETF, IPSec will be a broad-based, open solution for VPN security that will facilitate interoperability between VPNs. IPSec can be configured to run in two distinct modes 3/4 tunnel mode and transport mode. In tunnel mode, IPSec encapsulates IPv4 packets within secure IP frames to secure information from one firewall to another. In transport mode, information is encapsulated in such a way that it can be secured from endpoint to endpoint. In other words, the security wrapper does not obscure the end routing information as it does in the tunnel mode. Tunnel mode is the most secure method for deploying IPSec, but it results in significant overhead on a per-packet basis.

IPSec has had a very slow adoption cycle due primarily to dissension among the various IETF committees over key management standards and other issues. Intranet VPN applications using IPSec should start to be introduced to the market sometime in 1998. Commercial implementations using IPSec are still relatively immature, but the greatest supporters of the standard are the router vendors and the VPN hardware vendors who hope to usurp the router vendors in the market for building intranet (LAN-to-LAN) VPNs. Analysts predict that IPSec will be the primary standard for this segment of the VPN market.

Advantages

IPSec defines a set of standard protocols for authentication, privacy, and data integrity that are transparent to the application and the underlying network infrastructure. Unlike PPTP, IPSec supports a wide variety of encryption algorithms, such as DES, Triple DES, and IDEA. It also checks the integrity of transmitted packets to make sure they have not been tampered with en route.

IPSec was designed to provide security between multiple firewalls and routers, which makes it an optimum solution for LAN-to-LAN VPNs. IPSec's promise is that, because it is a natural extension to IP, it could be applied very broadly to the VPN market, ensuring interoperability among VPNs running over TCP/IP.

Disadvantages

For a number of years now, IPSec has been held out to the Internet community as the way to do secure networking. While IPSec holds great promise and will be a critical standard in IP-based networking, to date many attempts to deploy IPSec have been frustrated by the IETF committee infighting, which has delayed true interoperability between IPSec implementations. IPSec will likely be very successful in the LAN-to-LAN environments, but it will be of limited utility in the client/server configuration over the next few years.

IPSec in the client-to-server configuration has a number of disadvantages that may be difficult to get around. IPSec, for all practical intents, requires a public key infrastructure. Today's PKIs, including Internet Security Association Key Management Protocol (ISAKMP), have achieved relatively limited penetration, and concerns about overall scalability still exist. In addition, IPSec implementations require a known range of IP addresses or fixed IP addresses to establish identity.

This makes them impractical to use in dynamic address environments, which are common to Internet service providers.

IPSec does not support network protocols other than TCP/IP. As a standard, it does not specify a methodology for access control other than simple packet filtering. And, because it uses IP addressing as part of its authentication algorithm, it is seen as less secure than some of the higher-layer protocols that identify individual users.

Probably the most significant drawback to IPSec today is Microsoft's lukewarm support of the standard. Microsoft has been noticeably silent about support for IPSec in its client operating systems. Since IPSec in some ways competes with PPTP and L2TP for client tunnel building, it is not clear that Microsoft will make wholesale changes to the core IPv4 stack to support IPSec on the desktop. Regardless, replacing current IP stacks or widespread deployment of new device drivers is seen as extremely expensive and labor intensive.

8.5 Matching the Right Technology to the Goal

Corporations are finding that without Internet connectivity, they cannot compete in their respective markets. The Internet offers immediate access to information, which is tremendously beneficial as long as it is not coupled with security risks. Vendors are offering a number of VPN options to provide the necessary security to make internetworking worthwhile, but no solution today can solve every corporate need for secure communications. Each has its own benefits and drawbacks. Network administrators should carefully consider their priorities and base their decision on matching criteria. While some VPNs are easy to set up, others are more secure. And those that are fast may lack interoperability. The one certainty corporations can count on is an evolving market.

Each corporation has its own business style. Smaller shops may just need to provide their traveling sales representatives with a way to remotely access the corporate network. The larger the organization, the more likely it is to use an intranet to share information among its employees and branch offices. As security is added to VPNs, companies are extending those intranets and implementing full extranets. Considering that VPNs are moving in the direction of secure extranet VPNs, which are basically supersets of remote access and intranet VPNs, network managers should carefully review the scalability and potential of VPN solutions to support future business-to-business transactions over untrusted networks.

In general, the better performing pure VPN solutions will be targeted at the intranet (LAN-to-LAN) and less secure remote access VPN environments, and the more secure policy-based VPNs will be targeted at the extranet (business-to-business) and highly secure remote access VPN environments. VPNs that are implemented at layers two and three of the OSI model should demonstrate better performance than those at higher layers, and VPNs at layer five and above should offer much greater security. With that in mind, the following recommendations reflect the best practices for the different approaches to VPN
implementation.

Best Practices for Virtual Private Networking

Type of VPN	OSI Layer	Technology
Trusted LAN-to-LAN	Layers 1 and 3	IPSec, Frame Relay, SMDS, etc.
Basic Remote Access	Layer 2	PPTP, L2TP
Secure Remote Access	Layer 5	SOCKS v5, SSL
Business-to-Business	Layer 5	SOCKS v5, SSL

VPNs based on SOCKS v5 are best used by companies that need to provide highly secure, client-to-server connectivity for comprehensive business solutions, such as building a supply-chain extranet or highly secure remote access infrastructure. Because SOCKS v5 is an open standard that sits at the session layer, it can operate apart from lower-level protocols or add value to the VPN tunneling protocols that lack security features like access controls.

IPSec contains the most appropriate functionality to support trusted LAN-to-LAN VPNs. It does not require a client piece of software, so it provides appealing solutions for companies that want to exchange large amounts of data as fast as possible, typically the intranet VPN scenario. Because many of the router and internetworking vendors are building IPSec functionality into their platforms, IPSec will probably be the most important standard for this part of the market.

As mentioned earlier, PPTP and its variant, L2TP, are most appropriately used for remote access VPNs, as long as the limited encryption and authentication seem sufficient, and as long as Windows is the platform used. PPTP and L2TP will more than adequately meet many IT shops' basic remote access requirements.

According to a November 1997 issue of The Forrester Report, the VPN market is still immature, but early adopters of VPNs, ranging from start-ups to Fortune 50 companies, have been optimistic. Forrester predicted that the appeal of VPNs will broaden as security, performance, and interoperability wrinkles are smoothed out. One Forrester respondent from an aerospace company summed up a common forecast, saying, "Our VPN usage will explode over the next two years. Any application we need to share 3/4 internal Web, database access, personnel data, and benefits 3/4 will run over the network." Whatever solution a corporation decides on, it should adopt a security framework that can utilize the best of evolving technologies, function in a heterogeneous corporate environment, and map real-world trust relationships to the network.

Section References

8.0 "Making Sense of Virtual Private Networks", Aventail Corporation

9.1 NT Security Mechanisms

Many DOE sites have been upsizing from Windows 3.11 or Windows 95 to the Windows NT operating system. In today's environment, it is important to migrate to Windows NT because it was built from its inception to incorporation networking, security and audit reporting as services within the operating system.

What is the basis for NT security? It is designed to help enforce an organization's security policy (See Appendix A for details on Security Policies). This policy specifies an organization's information protection requirements, access controls, and audit requirements. NT enables you to
configure your network to allow information to be separated by departments or users in need-to-know groups and to control access by "outsiders". It further enables you to manage network and organizational resources as a group of objects and to enforce security rules controlling access and authentication.

Since NT is built to be secure, you don't have to worry about someone breaking into your system, right? Wrong. NT provides the ability to have a highly secure system only with the correct configuration and object access controls. Operating systems don't make security problems go away. There is not an operating system available today that can provide you with a complete security solution.

Remember you must define a security plan that defines the level of security needed in your organization, and integrate Windows NT with its security features into that plan. Security plans must detail both physical and logical security measures, to build the best protection against intrusion on your systems.

9.2 NT Terminology

9.2.0 Objects in NT

Described in this section are the basic concepts in the Windows NT environment. The concept of objects is important to the overall security theme in this operating system. The difference between the different types of NT software is defined, as well as the difference between domains and workgroups. Other terminology included in this section is concepts regarding the NT Registry and C2 Security.

Most elements in the NT operating system are represented as objects. Objects can be files, directories, memory, devices, system processes, threads, or desktop windows. Objects are what provide the NT operating system with a high level of security. They hide data from the outside and provide information only as defined by the object's functions. This gives layer of protection against external processes accessing internal data directly. NT obtains its high security level by preventing programs direct access to objects. All actions on objects must be authorized and performed by the operating system.

Objects can be secured in NT by setting attributes described by a security descriptor, or access token, containing the following:

- Owner/User Security ID (SID) indicating who owns the object.
- Group SID only used by the POSIX subsystem.
- Discretionary access control list contains access permissions for users and groups, controlled by the owner of the object.

- System Access Control List (ACL) controls the creation of auditing messages.

There are two types of objects: container objects and non-container objects. Container objects hold other objects; non-container objects do not have the ability to include other objects. Directories are container objects and files are non-container objects. Child objects created within a parent container inherit permissions from the parent object.

9.2.1 NT Server vs NT Workstation

There are two different types of Windows NT software available: Windows NT Workstation and Windows NT Server. The Server version is the same as the Workstation version except that it provides additional features for networking. Only ten users can access a Windows NT Workstation at a time, and NT Server can be accessed by an unlimited number of users
dependent upon the license purchased.

There may be some confusion between a server and a Windows NT Server. Windows NT Server is a piece of software, where a server is a piece of hardware.

9.2.2 Workgroups

There are two types of networking configurations in Windows NT:

Workgroups and Domains.

A workgroup is an organizational unit of a single system, or multiple systems not belonging to a domain. Systems in a workgroup individually manage their own user and group account information and their own security and account policy databases. They do not share this information with any other systems. If a system is not part of a domain, it is automatically part of a
workgroup. The best use of the workgroup configuration is for small groups of systems with few users, or where the network is configured without an NT Server.

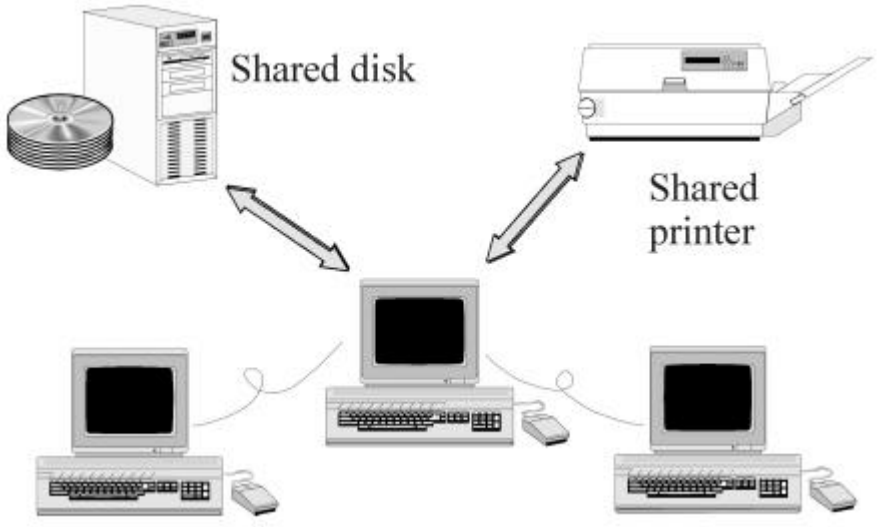


Figure 1: Workgroup Model Illustration

Warning: Security for Workgroups with systems running Windows 95, Windows 3.x, or Windows for Workgroups is virtually eliminated due to the fact that anyonecan access the computers and copy files to a diskette. There is no secure logon process or object access controls to prevent users from accessing sensitive files. Therefore, the workgroup model is not recommended unlessthe systems are all running Windows NT.

9.2.3 Domains

A domain is a collection of servers that are grouped together sharing a security policy and a user account database. Centralizing the user account database and security policy provides the system administrator with an easy and effective way to maintain the security policies across the network. Domains consist of a Primary Domain Controller (PDC), Backup Domain Controllers (BDC), servers and workstations. Domains can be set up to segregate different parts of your organization. Setting up proper domain configurations cannot guarantee a secure network, but it can give administrators a start in controlling user access on the network.

TIP: Isolate mission critical departments and services into separate domains, and limit the number of user accounts in these domains, to have more control over users actions.

Domain Controller

A PDC is a server in the domain that maintains the security and user account databases for that domain. Other servers in the domain can act as BDCs that hold a copy of the security database and user account information. The PDC, as well as the BDC can authenticate logon requests.
The BDC provides the network with a backup in case the PDC crashes important data will not be lost. Only one PDC is permitted in each domain. The master copy of the Security Account Manager (SAM) database is located on the PDC, where all account modifications are made. The BDCs are not permitted to make any modifications to the databases.

9.2.4 NT Registry

The Registry is a database that contains applications, hardware, and device driver configuration data, as well as network protocols and adapter card settings. This data is stored in the registry to provide a repository that stores and checks configuration data in one centralized location.

The functions of many files are combined in the Registry including the CONFIG.SYS, AUTOEXE.BAT, SYSTEM.INI, WIN.INI, PROTOCOL.INI, LANMAN.INI, CONTROL.INI and other .INI files. It is a fault-tolerant database that is difficult to crash. Log files provide NT with the ability to recover and fix the database if the system fails.

The Registry database structure has four subtrees:

- HKEY_LOCAL_MACHINE: Contains information about the local system including hardware and operating system data, startup control data and device drivers.

- HKEY_CLASSES_ROOT: Includes data pertaining to object linking and embedding (OLE) and file-class associations.
- HKEY_CURRENT_USERS: Contains information about users currently logged on the system, which includes the user's profile groups, environment variables, desktop settings, network connections, printers and application preferences.
- HKEY_USERS: Stores all actively loaded user profiles, including profiles of any users who have local access to the system. Remote user profiles are stored in the Registry of the remote machine.

Each of the subtrees contains value entries which are called keys, and each key can have many subkeys. The data in the four Registry subtrees is derived from sets of files called hives. Each hive consists of two files: data and log files. Each hive represents a group of keys, subkeys, and values that are rooted at the top of the Registry hierarchy.

9.2.5 C2 Security

Requirements for a C2 compliant system are defined by the National Computer Security Center (NCSC) of the United States Department of Defense, in the Trusted Computer System Evaluation Criteria document, better known as the Orange Book. Although a useful reference, the Orange
Book only applies to stand-alone systems. NCSC security ratings range from A to D, where A is the highest level of security and D is used mostly to evaluate business software. Each range is divided into classes, and in the C division there are C1 and C2 levels of security.

C2 represents the highest level of security in its class. Windows NT 3.5 Server, as a standalone system, was designed from the ground up to comply with the NCSC's C2 level requirements, and has been successfully evaluated as such. Certain processes such as identification, authentication, and the ability to separate accounts for operator and administrator functions, have met B2 requirements, an even higher level of security. These processes fulfill requirements for the B2 Trusted Path and B2 Trusted5 Facility Management.

Windows NT Server 4.0 is currently in NCSC evaluation as the networking component of a secure system. This is defined by the Red Book which is NCSC's Trusted Network Interpretation of the Trusted Computer System Evaluation Criteria, or Orange Book. The requirements are not changed in the Red Book, they just define how a networked system needs to operate in order to meet Orange Book requirements for a C2 level system.

C2 implementation on the Windows NT Server 3.5 is based solely on the software. In order to have a C2 compliant system setup, you must:

Have no network access to the system. Remove or disable floppy disk drives. Change standard file system access to be more restrictive.

> ***TIP:*** The C2 Config tool is available through the Windows NT Resource Kit, which can help you achieve a C2 level secure system.

The most important C2 level requirements featured in Windows NT 3.5 are:

- Discretionary access control (DAC): allows an administrator or user to define access to the objects they own.

- Object reuse: Memory is protected to prevent read access after it is freed from a process. When objects are deleted, users will be denied access to the object even when that object's disk space has been reallocated.
- Identification and authentication: Users must uniquely identify themselves before any access to the system is obtained. This is accomplished by entering a unique name, password, and domain combination, which will produce a users unique identity.
- Auditing: Must be able to create, maintain, and protect against modifications of an audit trail of access to objects. Access to the audit information must be restricted to a designated administrator.

9.3 NT Security Model

The Windows NT security model affects the entire Windows NT operating system. It provides a central location through which all access to objects is verified so that no application or user gets access without the correct authorization.

NT Security Subsystem

The Windows NT security model is based on the following components:

Local Security Authority (LSA)
Security Account Manager (SAM)
Security Reference Monitor (SRM)

In addition to these components, NT also includes logon processing, access control and object security services. Together these elements form the foundation of security in the Windows NT operating system, which is called the security subsystem. This subsystem is known as an integral subsystem since it affects the entire operating system.

9.3.0 LSA: Local Security Authority

The LSA is the heart of the security subsystem. It has the responsibility of validating local and remote logons to all types of accounts. It accomplishes this by verifying the logon information from the SAM database. It also provides the following services:

- Checks user access permissions to the system
- Generates access tokens during the logon process
- Manages local security policies
- Provides user validation and authentication
- Controls the auditing policy
- Logs audit messages generated by the SRM

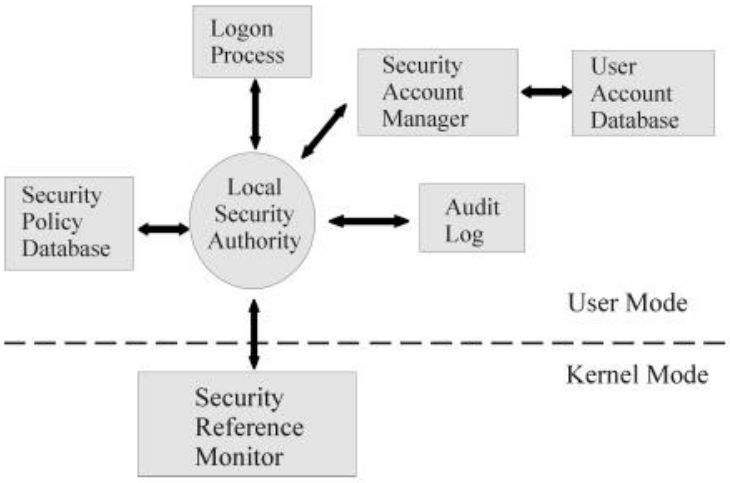


Figure 2: NT Security Model

9.3.1 SAM: Security Account Manager

The SAM manages a database which contains all user and group account information. SAM provides user validation services which are used by the LSA, and are transparent to the user. SAM is responsible for checking logon input against the SAM database and returning a secure identifier (SID) for the user, as well as a SID for each group to which the user belongs. When a user logs on, the LSA creates an access token which includes the SID information along with the user's name and associated groups.

From this point on, every process that runs under this user's account will have a copy of the access token. When a user requests access to an object, a comparison is made between the SID from the access token and the object's access permissions list to validate that the user has the correct permissions to access the object.

The SAM database supports a maximum of 10,000 accounts. SAM databases may exist on one or more NT systems, depending on the network configuration. The types of network configurations include:

- When separate user accounts are on each system, the local SAM database is accessed.
- The SAM database is located on the domain controller when a single domain with a centralized source of user accounts is the configuration.
- In the master domain configuration, where user accounts are also centralized, the SAM database is located on the Primary Domain Controller (PDC), which is copied to all Backup Domain Controllers (BDC) in the master domain.

9.3.2 SRM: Security Reference Monitor

The SRM runs in kernel mode and is a component of the Windows NT Executive. It is responsible for the enforcement of access validation and audit generation policies required by the LSA. SRM provides services for access validation to objects and access privileges to user accounts. It also protects objects from being accessed by

unauthorized users. To ensure that objects are protected regardless of their type, the SRM maintains only one copy of the access validation code on the system. Instead of accessing objects directly, users requesting access to objects must have SRM validation. The steps used to determine user access to objects are as follows:

- When access to an object is requested, a comparison is made between the file's security descriptor and the SID information stored in the user's access token. The user will obtain access to the object given sufficient rights. The security descriptor is made up of all the Access Control Entries (ACE) included in the object's Access Control List (ACL).
- When the object has an ACL, the SRM checks each ACE in the ACL to determine if access to the object is granted. If the object has no ACL associated with it, SRM automatically allows access to everyone. If the object has an ACL with no ACEs, all access requests to that object will be denied.
- After the SRM grants access to the object, continued validation checks are not needed to access the particular object. Any future access to the object is obtained by the use of a handle which was created when the access was initially validated.

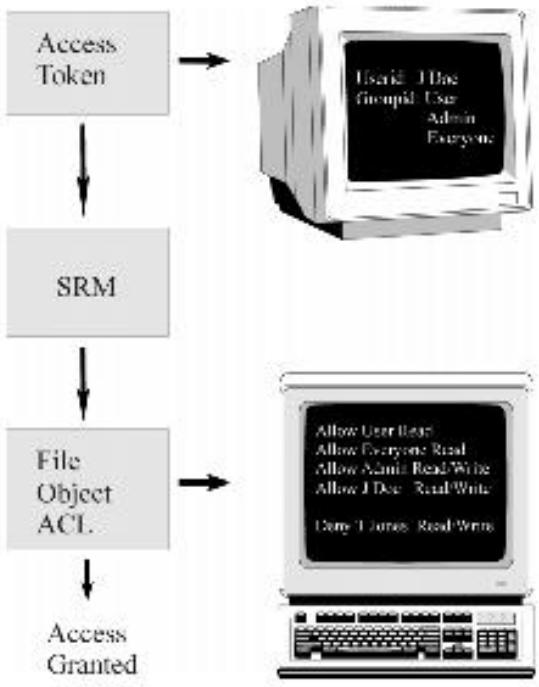


Figure 3: SRM Access Validation Process

9.4 NT Logon

Windows NT logon processes provide mandatory logon for user identification and cannot be disabled. Before accessing any resources on the system, the users go through the logon process so that the security subsystem can authenticate the user name and password.

To protect against an application running in background mode, such as a Trojan logon program, the logon process begins with a Welcome message box that requests the user to press Ctrl, Alt and Del keys before activating the actual logon screen.

> ***Note:*** The Ctrl, Alt, Del sequence guarantees that a valid Windows NT logon sequence will be initiated. This key sequence should always be used when logging on to a machine, even if it appears that the logon screen is already displayed.

Logon Banner

A logon banner, also referred to as a warning banner, should be added to warn individuals who may try gaining access to a system without authorization. If activated, this message is displayed after the Welcome message in a dialog box that must be confirmed. The text and style of the legal notice is set in the Registry Editor.

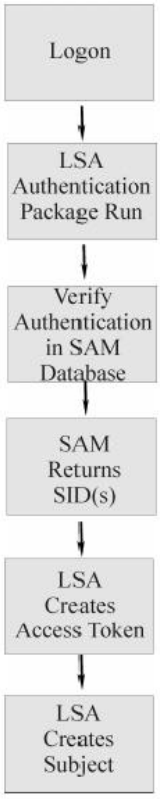


Figure 4 NT LOGIN

9.4.0 NT Logon Process

Outlined in Figure 4 is the Windows NT logon process:
A Welcome dialog is displayed which requires a user name, password and the server/domain the user would like to access. If the user information is valid, the system proceeds to authenticate the user.

User authentication is determined by passing the user input from the Welcome screen to SAM via the security subsystem.

SAM does a comparison between the user logon information and the server's SAM database. If the data matches, the server notifies the workstation of the approval. The server also stores information about the user, such as account privileges, home directory location and workstation variables.

The LSA now constructs the access token. The access token is connected with each process the user runs.
This process and token information together form a subject. When a user requests access to an object, the contents of the subject's token are compared to the object's ACL through an access validation procedure. This access validation procedure grants or denies permission to the user's request.

9.5 Designing the NT Environment

NT security components enable you to design a network configuration that separates highly sensitive data and applications from less sensitive data and applications. By designing your network according to information protection needs, you greatly simplify the application of your security policies. The NT environment uses the concept of domains as a means for grouping resources together that share common information and have common security needs. Communication between domains is then controlled by trust relationships.

For example, many areas of an organization may need access to data located within the financial domain; however, user in the financial domain probably doesn't need

access to data within the medical domain. Additional ways to protect your systems are achieved by group management, access control of objects, and file system configurations, which are all discussed in this section.

9.5.0 Trusts and Domains

Trust Relationships

Trusts are an administrative way to link together two domains allowing one domain's users access to the other domain. Trust relationships between domains are a way to centralize administrative tasks. They enable user accounts and groups to be used in a domain outside of where those accounts originated. Trusts combine two or more domains into an administrative group. There are two parts to a trust: the trusted domain and the trusting domain. The trusted domain makes accounts available for use in the trusting domain. Users only need one name and password to access multiple domains.

> ***Tip:*** The best policy in setting up trust relationships between domains is to provide the least amount of service possible. Evaluate the services you have running on domains. Do not allow trust relationships to a domain that might allow users to disrupt services providing critical information, and avoid running high security risk services in domains which are accessed by any users other than administrators.

Trust Relationship Models

Trust relationships are defined in only one direction. To obtain a two-way trust, both domains must trust each other. The trusted domain is where the accounts reside, known as the account domain. The trusting domain contains the resources, known as the resource domain.

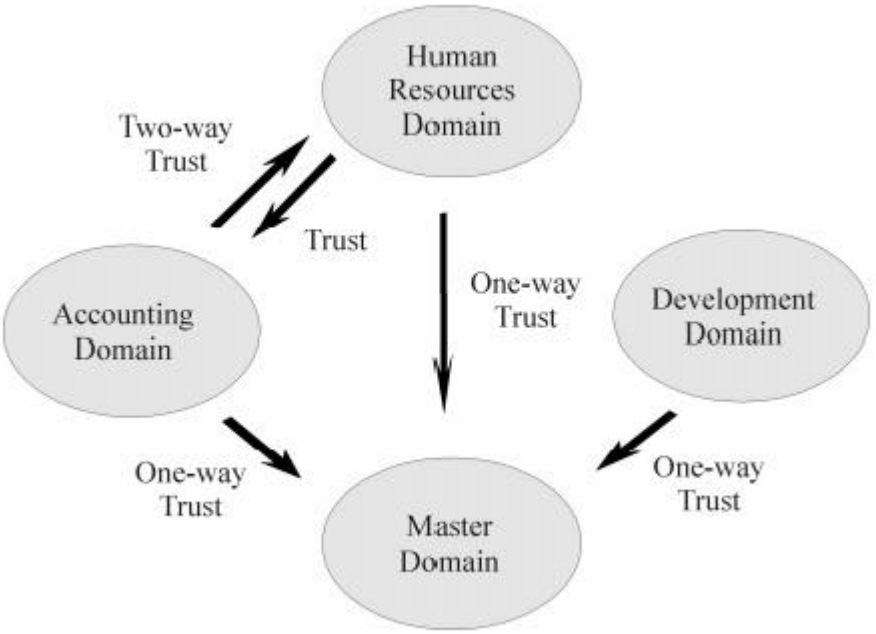


Figure 5: Trust Relationships

The following are the types of Trust Relationship Models:

- Single Domain
- Master Domain
- Multiple Master Domain

Single Domain Model

The Single Domain is the best model for organizations with fewer than 10,000 users. There is only one domain in this model; therefore there is no administration of trust relationships. Administration of user accounts is centralized, and global groups are used for accessing resources.

Master Domain Model

The Master Domain model includes multiple domains, with one being the master domain. The master domain is trusted by all other resource domains, but does not trust any of them. The resource domains do not trust each other. This model provides the benefits of centralized administration and multiple domains.

Administration of user accounts and resources are in separate domains. Resources are managed locally on the trusting domains, while user accounts are controlled on the trusted master domain. The master domain model is used in organizations with less than 10,000 users. The number of
users is limited because the accounts are all maintained on the master domain.

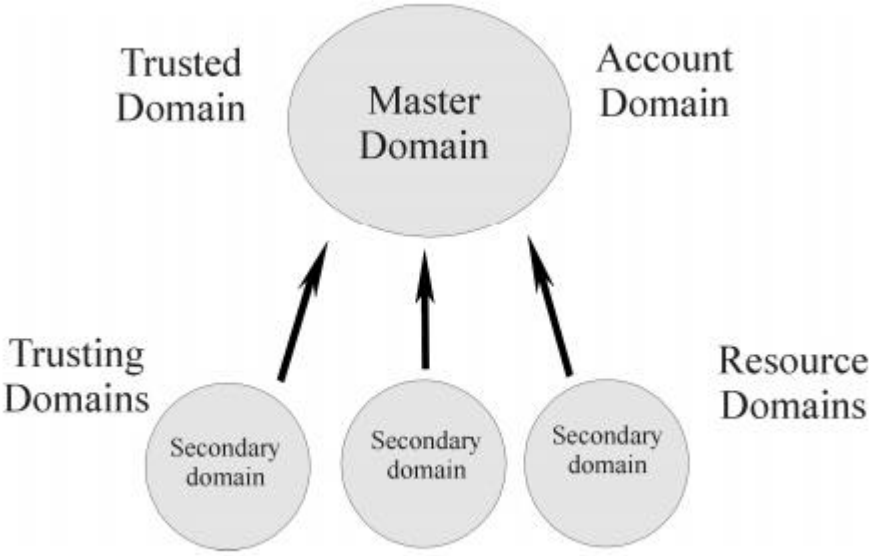


Figure 6: Master Domain Model

> ***Note:*** If done correctly, this model can provide a secure configuration because administration is managed for the entire network in one centralized location.

Multiple Master Domain Model

The Multiple Master Domain model is used for organizations with computer resources grouped into logical divisions, such as by departments or location. This model is identical to the Master Domain model except that there is more than one master domain. All master domains have a two-way trust with each other. Each resource domain trusts all master domains, but the resource domains do not trust each other. Since master domains trust each other, only one copy of the user account database is needed. This model is designed for organizations with more than 10,000 users.

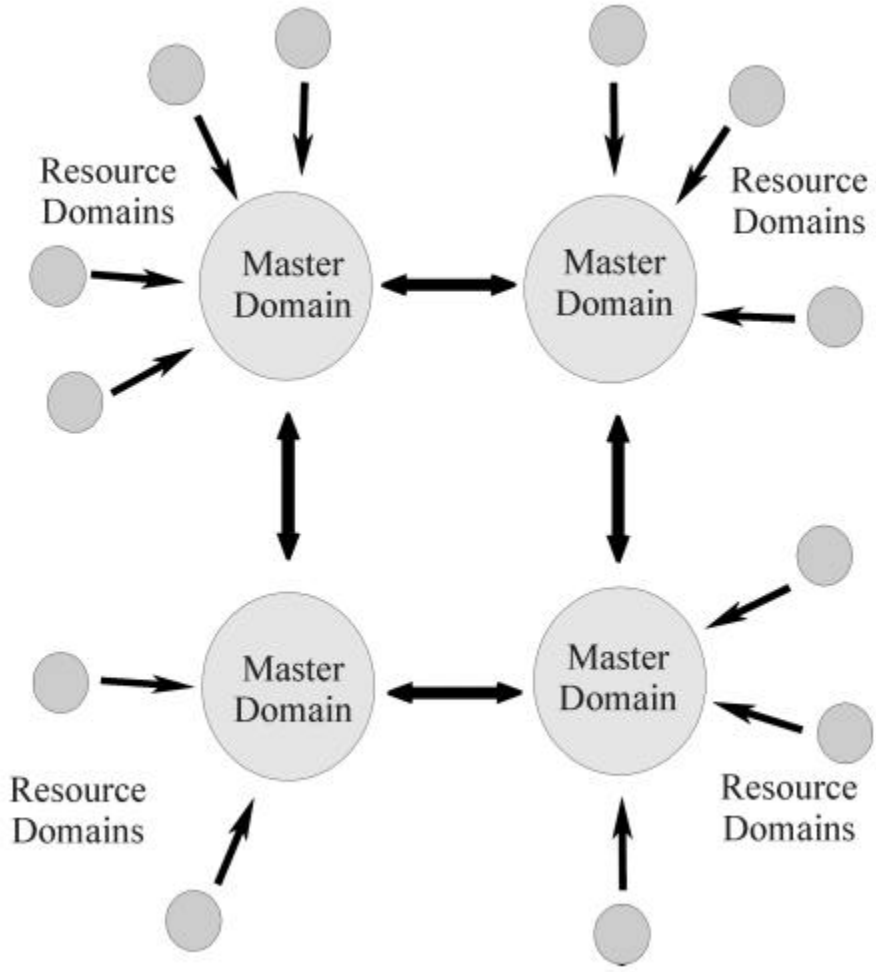


Figure 7: Multiple Master Domain Model

9.6 Group Management

Groups are an administrative tool used to provide a collection of users, with common needs, the permissions and rights they require to perform their job. As previously mentioned, a group is essentially an account containing other accounts in Windows NT. A user in a group is a member of the group and access permissions, rights, and restrictions assigned/granted to the group are assigned/granted to each of the group members.

For example, if a directory is established for the Payroll Department to hold their common files, it is much easier for a system administrator to have everyone in the Payroll Department in a group and then assign that group permissions on the directory and the files in it. Otherwise, the system administrator would have to go through and assign permissions to every user in the Payroll Department.

In addition, groups can be used to restrict the access a collection of users has to certain objects. For example, the system administrator could utilize the Payroll

group to prevent the users in the Payroll Department from printing to a printer in a remote location (because their data could be
potentially very sensitive), while allowing access for all other users, by placing a deny ACE for the Payroll group in the ACL for the printer. It is normally easier to administer rights by granting them to groups and then making the users who need the right a member of the group. For example,
if there are users who need to logon to a server locally, create a group called Local Logon. Add the users to the group, and grant the Log on Locally right to the group. This group could then be reused again should this group of users need some other common right or access permission.

There are three types of groups in Windows NT:

- Local Groups
- Global Groups
- Special Groups

Local Groups

Local groups are maintained on a local system or domain and may have user accounts or global groups as members. At the local system level, local groups would be used to administer permissions and rights for the system on which they reside. At the domain level, local groups would be used to administer permissions and rights on Windows NT Servers within the domain where the groups reside. To summarize, local groups are only utilized in the user account database for the local system or domain where they are created.

Windows NT provides some built-in local groups each with established permissions and rights. At the local system level they are:

- Administrators - can fully administer the system.
- Power Users - can share directories and printers.
- Users - normal users.
- Guests - granted guest access.
- Backup Operators - can bypass file security in order to complete backups.

At the domain level, the built-in groups are:

- All listed above except Power Users.
- Server Operators - can manage domain servers.
- Account Operators - can manage user accounts and groups.
- Print Operators - can manage printers.
- Replicator - supports file replication.

Global Groups

Global groups maintained on a Windows NT domain may have domain user accounts as members, and are used to administer domain users. System administrators can effectively use global groups to sort users based on their needs. This can be accomplished by placing the global group in the appropriate local groups, assigning the users permissions and granting them the rights they need to perform their jobs. As mentioned, global groups can only have domain user accounts as members. No other groups can be members of a global group. This is due to the fact that the system administrator assigns permissions and grant rights to the local groups (because the local system or domain server holds the resources)

and then makes the global groups members of the local groups. Windows NT provides two built-in global groups each with established permissions and rights. They are:

- Domain Admins - contains the domain administrator account by default and is a member of the domain level Administrators local group and the system level Administrators local group for Workstations in the domain.
- Domain Users - contains all the domain users.

Special Groups

Special groups are created by Windows NT for unique or specific purposes and can not be viewed, changed, or have members added to them in the User Manager. A user's membership to a special group is determined by how they access resources on the system. Special groups may be assigned access permissions in some cases and may be seen when a system administrator is assigning permissions on Windows NT objects.

The following is a list special groups and a description of their membership:

- Network - any user connected to a system via the network.
- Interactive - any user logged on interactively at a local system
- Everyone - any user logged on to the system (both the Network and
- Interactive groups).
- Creator Owner - the user that created or took ownership of an object.
- System - the Windows NT operating system.

> ***Note:*** If the user were the system administrator or other user that is a member of the Administrators group, the Administrator group would be a member of the Creator Owner group.

The special group that system administrators must pay close attention to is the Everyone group. As stated above, all users logged on are members of this group. Therefore, any access permissions assigned to the Everyone group allowing or denying access to objects is by default assigned to all users.

For example, if a file should only be accessed by a certain group, the system administrator could not assign permissions to that group allowing file access and then assign permissions to the Everyone group denying file access. Since Windows NT acts on all deny ACEs before allow ACEs, it would stop when it found the deny ACE for the Everyone group and no one would be allowed access including the group with permissions assigned to allow access to the file.

9.7 Access Control

Each file and directory object has an Access Control List (ACL) that contains a list of Access Control Entries (ACEs). ACEs provide information regarding access or auditing permissions to the object for a user or group of users. Along with the file system, they protect objects from unauthorized access. There are three different types of ACEs:

- System Audit
- Access Allowed
- Access Denied

System Audit is a system ACE used for logging security events and audit messages. Access Allowed and Access Denied are known as discretionary ACEs. They are prioritized by the type of access: Denied and Granted.

Deny always overrides grant access. If a user belongs to a group with Access Denied privileges to an object, the user will be denied access regardless of any granted access he possesses from his own user account, or in other groups to which he is included.

Discretionary ACLs allow owners to control the access of their objects. Controls over objects can be applied to individual users, multiple users, and groups. They can be set by the object's owner, a user who has an administrator account, or any user with correct permissions to control resources on the system. If a discretionary ACL is not specified for an object, a default ACL is created. Default ACL file objects inherit access controls from their parent directories.

> ***Warning:*** Be sure to evaluate your object's ACLs after installing Windows NT. Most versions are shipped with file ACLs set to give Everyone Full Control access.

User Rights

User authorization to perform specified actions on a system is called rights. Rights apply to the entire system. They are usually assigned to groups or users by the system administrator. Rights give users access to services such as backing up files and directories, shutting down the computer, logging on interactively or changing system times, that normal discretionary access controls do not provide.

9.8 Managing NT File Systems

Due to NT's modular approach of file system management, multiple file systems are supported. NT uses low-level drivers as a part of the NT Executive to support each file system. This provides the ability to expand to additional file systems as they are introduced by simply installing a new
driver.

NT 4.0 supports two file systems: NTFS and FAT.

9.8.0 FAT File System

The File Allocation Table (FAT) file system is named after it's organizational method. The FAT file system was originally designed for small disks and simple directory structures. Its design has since evolved to support larger disks and more powerful systems. It is most widely used for systems that run the DOS operating system.

The FAT file system doesn't support the security features or the automatic disk restoration utilities that NT provides. Using the FAT file system is not recommended for volumes shared across the network. The following configurations do require the FAT file system structure:

- Dual-boot system configurations with DOS or OS/2 volumes.

- FAT is the only file system available for formatting diskettes on Windows NT.
- RISC-based systems must provide a FAT partition to boot system files.
- NT provides a tool to secure the FAT system partition on this type of system.

> ***Tip:*** If there is no need to boot DOS, and the system is not an RISC architecture, using FAT file systems are not recommended.

9.8.1 NTFS File System

NTFS was developed to support the Windows NT file and directory security features. It is the only file system available on NT that provides the capability to assign permissions to individual files. The NTFS driver that allows access to an NTFS volume is loaded in NT so unauthorized users cannot access NTFS volumes by booting the system from a DOS diskette.

NTFS also prevents users from undeleting files or directories that have been removed from NTFS volumes. Since NT doesn't give undeleted programs access to work on an NTFS volume, even files that still exist on the disk are not available.
NTFS provides file system recovery where disk activities can be logged to enabling activities to be restored in the case of a system crash.
Chances of corrupting data, due to power or hardware failures, are small with NTFS.

Physical Security and NTFS

NTFS file system security is only valid if the ability to access the system from DOS, or another operating system is eliminated. The following precautions for physical security should be examined:

- Remove or lock floppy drives.
- Require boot passwords on servers and set the BIOS to disable booting from a floppy drive. In most cases, removing the battery disables the BIOS lock.
- Do not create any DOS partition on the server.
- Lock the system in a secure location.
- Set alarms alerting you to when a server is shut down, so an intruder can be caught during a potential attack.

> ***Warning:*** A program called ntfsdos.exe is available to read files protected by Windows NTFS. The program is run after booting a system with a DOS diskette. This is not a security risk if the proper physical security measures are taken or floppy drives are not available on the system.

NTFS vs FAT

NTFS provides extended security features not available with the FAT file system.
NTFS is built for speed. It uses a binary tree structure for directories to reduce the access time needed to locate files.

NTFS minimizes file fragmentation in large disk volumes.
NTFS uses small cluster sizes (512 bytes) to prevent wasted disk space.
NTFS provides the ability to selectively compress individual files and directories or actual volumes on disks.

Shares

The Shared Directory feature in the File Manager allows sharing of files and directories over the network. Shared object permissions can be established for FAT or NTFS file structures. The user must be a member of the Administrator group or Server Operator group to work with shared directory permissions. Users are unable to access files on a system through the network until there is a shared directory available.

Once a directory has been shared on the system, users can log on to that system and be able to access the shared directory. To use the directory, the user must assign the share to an unassigned drive letter. When the directory is assigned a drive letter, the share can be accessed just like another hard disk on the system. Directory sharing can be viewed and stopped by an Administrator or Server Operator.

9.9 Object Permissions

File and directory permissions are the foundation of most user-controlled security in Windows NT. Permissions are the rules associated with a particular object, which describe which users can access what objects, and how they have access to the objects. Object permissions for files are only available for files stored on NTFS volumes. File and directory permissions are cumulative, but the No Access permission overrides all other permissions.

The types of file access permissions are:

- No Access
- Read
- Change
- Full Control
- Special Access

For directory access the following permissions are added:

- List
- Add
- Read

Object Ownership

Object ownership allows the user to change permissions on the owned object. The user who is the creator of a file or directory is usually the owner. Users can't give away ownership of their objects, but they can give other users permission to take ownership. This prevents users from creating objects and making them appear to be owned by another user. Ownership of a file or directory can be taken by an Administrator without the owner's consent, but the Administrator can't transfer ownership to others. Administrators cannot access private files without leaving some trails behind, because after claiming ownership, Administrators cannot return ownership to the original owner.

9.10 Monitoring System Activities

Monitoring is a continuous evaluation of system-level attributes that could reveal system compromise. Monitoring also provides reporting and follow-up mechanisms on attempted violations to the system. Auditing systems validates compliance when using monitoring procedures. In addition, auditing is used in follow-up actions.

There are two types of security monitoring: status and event monitoring. Status monitoring involves current states or processes of the system. Event monitoring evaluates audit trails, which occurs after processes have finished running. Auditing is provided to evaluate the control structure, assess risk, determine compliance, report on exceptions and make improvements to the system. Systems should be evaluated against the organization's security policies and compliant technical platforms to the security implementation standards.

The monitoring section of a site security plan should include:

Systems and subsystems to audit
Tools and configuration settings
Schedules for periodic auditing tasks
Review and testing of audit coverage and functionality

Section References:

9.0 Kelley, Marcey and Mayson, Wendall. "*Windows NT Network Security A Manager's Guide CIAC-2317*". CIAC *Department of Energy* Lawrence Livermore National Laboratory. December 1997

If you suspect or have been notified that your computer system has been or is under attack, you must determine:

- if there really is or was an attack
- if the attack was successful
- and, to what degree the attack compromised the system

This can be routine, quite challenging, or extremely difficult. Modern operating systems are large, complex, and imperfect dynamic systems, with many places for attackers to hide and many opportunities for them to cover their tracks.

CIAC has collected and developed techniques to discover traces of an attack. Almost all attacks leave detectable remnants that may be uncovered and used in an investigation.

This section contains step-by-step instructions to follow if you are investigating an actual security incident. It can also be used as a tutorial in general techniques for use if an attack occurs.

This guide helps you with these security scenarios...	**By providing you with detailed information on these topics...**
A person's system is linked to the Internet; there is "a feeling" that something is wrong. A security problem might exist, but you can't be sure. You are notified by CIAC that someone from another site that had an intruder found your site's name in an intruder's log file. You know that an intruder has at least "touched" your system. The extent of the contact is unknown. An incident response team informs you that an intruder was located, and the team's log files indicate the intruder came from *your* site. finding the footprints left by an intruder You get a call that someone is performing an illegal action (either breaking into another system, or breaking into that particular system) right NOW. Action must be swift in order to minimize damage. You suspect you have a sniffer on your system, but don't have the slightest idea where to start looking for it.	displaying the users logged in to your system displaying active processes detecting a sniffer finding files and other intrusions left by an intruder

10.1 Displaying the Users Logged in to Your System

If you suspect that there is an active intruder on your system, first determine where they are and what they are doing. This section shows you how to use these commands to find out who is on your system:

- the “w” command
- the “finger” command
- the “who” command
- The “w” Command

> ☛ **These commands are only useful when a suspected intruder is logged in to your system.**

10.1.0 The “W” Command

The “**w**” command gives you a general overview of all users and their active programs on the system. A sample output is shown here.

```
Prompt % w

  3:47pm  up 18 days,  3:02,  7 users,  load average: 0.02, 0.00, 0.00
User     tty      login@   idle     JCPU     PCPU     what
user1    ttyp0    25Mar94  2:08     39:15    4        -tcsh
user2    ttyp1    5Apr94   8        5:51     5:28     emacs
user2    ttyp2    3:46pm                     w
user3    ttyp3    Mon 2pm           2:04     1        -csh
user3    ttyp4    Mon 3pm  41       21                -csh
user2    ttyp6    5Apr94   3        1:38     6        -tcsh
user2    ttyp7    Wed 2pm  5:31     17       1        -tcsh
Prompt %
```

The first line displayed, the status line, gives general information: the present time, how long the system has been running, and the load on the system for various periods of time. The rest of the output from the “w” command shows you who is currently logged in to the system, which TTY they are using, and what each user is currently doing.

What to Look For

Verify that:

- all users are valid
- users have not been logged in for an abnormal length of time
- users are not running suspicious software

Vulnerabilities
The output listing from the “w” command can be easily modified to hide a skilled intruder’s existence on the system.

10.1.1 The "finger" Command

Another command that displays who is on the system is the **"finger"** command. A sample output is shown here. The "finger" command shows you who is currently logged in to the system, which TTY they are using, the time they logged in, and where they are logged in from.

```
Prompt %  finger
Login    Name         TTY      Idle       When      Where
user1    user name     p0        26    Fri 11:46    host1.sub.domain
user2    user name     p1        34    Tue 10:42    host2.sub.domain
user4    user name     p2              Mon 14:04    host3.sub.domain
user3    user name     p3        44    Mon 14:06    host5.sub.domain
user2    user name     p4              Mon 16:43    host4.sub.domain
user2    user name     p6      3:45    Tue 11:06    host2.sub.domain
user2    user name     p7         1    Wed 14:47    host2.sub.domain
user3    user name     p8      3:04    Thu 11:04    host5.sub.domain
user3    user name     p9      1:02    Fri 13:52    host5.sub.domain
Prompt %
```

What to Look For

Verify that:

- all users are valid
- users have not been logged in for an abnormal length of time
- the login location of each user is valid

Vulnerabilities
The output from the "finger" command can easily be modified to hide a skilled intruder's existence.

10.1.2 The "who" Command

The **"who"** command lists information about the users currently on the system. This information is retrieved from the /etc/utmp file. A sample output is shown here. This command lists who is currently logged in to the system, which TTY they are using, login time, and where they are logged in from.

```
Prompt %  who
user1    ttyp0    Mar    25    11:46    (host1.sub.domain)
user2    ttyp1    Apr    5     10:42    (host2.sub.domain)
user4    ttyp2    Apr    18    14:04    (host3.sub.domain)
user3    ttyp3    Apr    11    14:06    (host5.sub.domain)
user2    ttyp4    Apr    18    16:43    (host4.sub.domain)
user2    ttyp6    Apr    5     11:06    (host2.sub.domain)
user2    ttyp7    Apr    6     14:47    (host2.sub.domain)
user3    ttyp8    Apr    14    11:04    (host5.sub.domain)
user3    ttyp9    Apr    15    13:52    (host5.sub.domain)
Prompt %
```

What to Look For

Verify that:

- all users are valid
- users have not been logged in for an abnormal length of time
- the login location of each user is valid

Vulnerabilities
The output from the "who" command can easily be modified to hide a skilled intruder's existence, as the command gets its information from the /etc/utmp file.

10.2 Displaying Active Processes

Even if an intruder is no longer logged in to a (potentially) penetrated system, a process may have been left running by the intruder to continue performing tasks. This section shows you how to use these commands to display the active processes running on your system:

the “ps” command
the “crash” command

10.2.0 The “ps” Command

The **“ps -agux”** command lists the processes that are executing on your system.

The command’s “a” parameter displays all processes running on the system, not just those owned by you. The command’s “g” parameter displays all processes, as opposed to those which “ps” decides are simply “interesting” (refer to the “ps” man page for the definition of “interesting”).
The “u” parameter displays user-oriented output. The “x” parameter includes processes without control terminals.

The “ps” command is a reliable way to see what programs are being executed on the system. A shortened sample output is shown here.

```
Prompt % ps -agux
USER       PID  %CPU %MEM    SZ   RSS  TT   STAT START   TIME COMMAND
user5    28206   8.1  0.4    48   280  p4   S    13:55   0:00 man inetd.conf
user5    28208   3.9  0.5    56   312  p4   S    13:55   0:00 more -s /usr/man/cat5/in
root         2   0.0  0.0     0     0  ?    D    Mar 25  0:02 pagedaemon
root        87   0.0  0.0   176     0  ?    IW   Mar 25  0:16 sendmail: accepting conn
root         1   0.0  0.0    56     0  ?    IW   Mar 25  0:04 /sbin/init -
user3    15547   0.0  0.0    88     0  ?    IW   Apr 5   0:00 selection_svc
user1      184   0.0  0.0   192     0  p0   IW   Mar 25  0:06 -tcsh (tcsh)
user2    28209   0.0  0.8   208   520  p5   R    13:55   0:00 ps -agux
user2    21674   0.0  0.4   112   248  p5   S    16:24   0:00 -tcsh (tcsh)
user3    16834   0.0  0.0    88     0  ?    IW   Apr 5   0:00 selection_svc
user3    27350   0.0  0.0   112     0  p3   IW   Apr 11  0:01 -csh (csh)
user4    23846   0.0  0.0    80     0  pa   IW   11:12   0:00 -csh (csh)
user3    23801   0.0  0.0    80     0  p8   IW   11:04   0:00 -csh (csh)
user2    18590   0.0  0.0   120     0  p7   IW   Apr 6   0:01 -tcsh (tcsh)
user2    15591   0.0  0.0   120     0  p6   IW   Apr 5   0:06 -tcsh (tcsh)
user2    15588   0.0  0.1  9288    72  p1   I    Apr 5   7:08 emacs
user2    15496   0.0  0.0   112     0  p1   IW   Apr 5   0:00 -tcsh (tcsh)
Prompt %
```

What to Look For

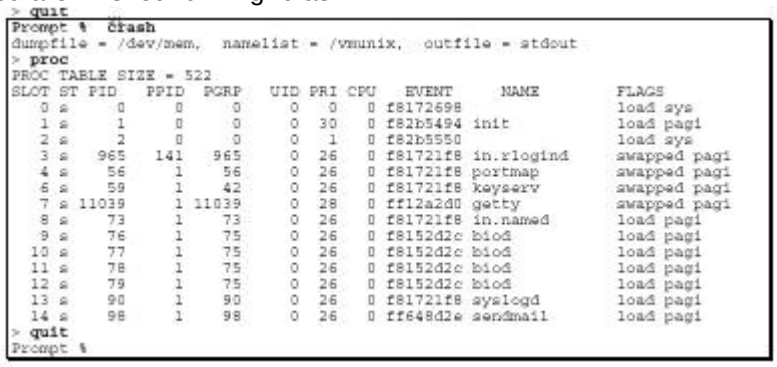
The following may indicate undesired activity:

- processes that take a long time
- unusual start times (such as 3:00 a.m.)
- unusual names
- a process that has an extremely high percentage of CPU (this may indicate a sniffer process)
- processes without a controlling terminal (a “?” in the TT column) that are executing unusual programs

Vulnerabilities
In some cases, compromised systems have been found to contain a Trojaned version of “ps” which does not display intruder processes. Also, if an invalid process is running but has a valid process name, it may be difficult to distinguish the suspicious process. For example, intruders
often run sniffer processes under names such as “sendmail” or “inetd”.

10.2.1 The "crash" Command

You can use the "**crash**" command to list all processes. This functions as a cross-check against the "ps" command. That is, finding a process with "crash" output that does not appear in "ps" output (matching pids). Once you execute "crash," you will receive a ">" prompt. Type **proc** in response and **quit** when you are finished running "crash".



What to Look For
The following may indicate undesired activity:

- processes that do not appear in the ps list (use the PID column to identify)
- a high value in the CPU column
- unusual commands in the NAME column

Vulnerabilities
Names can be faked. Like any command, "crash" can be Trojaned.

10.3 Finding the Footprints Left by an Intruder

If you suspect that an intruder has been on your system but is gone, use the commands and files described in this section to find the "footprints" the intruder may have left behind. This section shows you how to use these commands and files:

the "last" command
the "lastcomm" command
the "/var/log/syslog" file
the "netstat" command

10.3.0 The "last" Command

The **"last"** command displays information about logins and logouts on the system from the /var/adm/wtmp file. If you can determine the username the intruder used to log in, this command can show you how long the intruder was logged in and where they logged in from. The command's "-*n*" parameter is used to display the last *n* entries in the /var/adm/wtmp file.

A sample output is shown here.

```
Prompt % last -20
user1    ftp      host1.sub.domain    Fri Apr15 15:09 - 15:10 (00:00)
user3    ttyp9    host5.sub.domain    Fri Apr 15 13:52 still logged in
user6    ttyp2    host7.sub.domain    Fri Apr 15 13:45 - 14:1 (00:26)
user6    ttyp2    host7.sub.domain    Fri Apr 15 10:34 - 10:34 (00:00)
user6    ftp      host7.sub.domain    Fri Apr 15 10:32 - 10:33  (00:01)
user4    ttyp4    host3.sub.domain    Fri Apr 15 10:17 still logged in
user5    ttyp2    host6.sub.domain    Fri Apr 15 09:20 - 10:29 (01:09)
user1    ttyh1                        Thu Apr 14 20:33 - 22:00 (01:26)
user4    ftp      host3.sub.domain    Thu Apr 14 14:21 - 14:22 (00:01)
user4    ttyp2    host3.sub.domain    Thu Apr 14 14:01 - 16:36 (02:35)
user4    ftp      host3.sub.domain    Thu Apr 14 13:43 - 13:44 (00:00)
user5    ttyp4    host6.sub.domain    Thu Apr 14 13:38 - 14:56 (01:18)
user4    ttyp2    host3.sub.domain    Thu Apr 14 13:37 - 13:47 (00:10)
user4    ftp      host3.sub.domain    Thu Apr 14 13:16 - 13:18 (00:01)
user4    ttyp2    host3.sub.domain    Thu Apr 14 13:12 - 13:18 (00:05)
user4    ttypa    host3.sub.domain    Thu Apr 14 11:13 - 15:05 (03:52)
user4    ttyp9    host3.sub.domain    Thu Apr 14 11:12 - 13:08 (01:55)
user3    ttyp8    host5.sub.domain    Thu Apr 14 11:04  still logged in
user1    ftp      host1.sub.domain    Thu Apr 14 11:01 - 11:02 (00:00)
Prompt %
```

The first column contains the username, followed by the terminal device the user is connected to. If the connection used a network device, the name of a remote system is displayed in the next column. For serial devices such as dial-up modems, the column will be blank. This is followed by the login and logout time and an indication of the length of the session.

What to Look For

- examine the log entries made around the time of the suspected attack for ones that appear to be out of the ordinary, including logins to accounts that had previously been dormant, logins from unexpected locations, logins at unusual times, and short login times a missing /var/adm/wtmp file or one with gaps in the output (this may indicate that an intruder attempted to hide their existence)

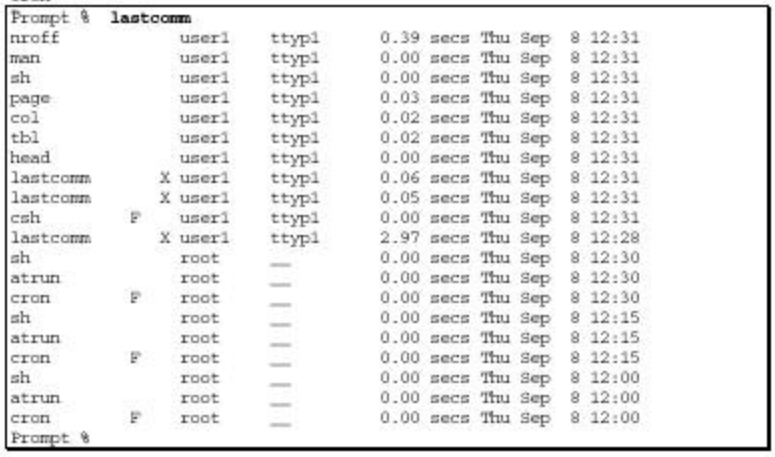
As a general rule, many system administrators never delete this file. Therefore, it can be quite large and include activity from when the system was first loaded.

Vulnerabilities

An intruder who breaks into a system can hide their tracks by deleting or modifying the /var/adm/wtmp file.

10.3.1 The "lastcomm" Command

The **"lastcomm"** command displays the last commands executed. This command is only available if you have process accounting turned on. With this command, you can see every command issued by anyone on the system. A sample output is shown here.



```
Prompt % lastcomm
nroff            user1    ttyp1      0.39 secs Thu Sep  8 12:31
man              user1    ttyp1      0.00 secs Thu Sep  8 12:31
sh               user1    ttyp1      0.00 secs Thu Sep  8 12:31
page             user1    ttyp1      0.03 secs Thu Sep  8 12:31
col              user1    ttyp1      0.02 secs Thu Sep  8 12:31
tbl              user1    ttyp1      0.02 secs Thu Sep  8 12:31
head             user1    ttyp1      0.00 secs Thu Sep  8 12:31
lastcomm       X user1    ttyp1      0.06 secs Thu Sep  8 12:31
lastcomm       X user1    ttyp1      0.05 secs Thu Sep  8 12:31
csh          F   user1    ttyp1      0.00 secs Thu Sep  8 12:31
lastcomm       X user1    ttyp1      2.97 secs Thu Sep  8 12:28
sh               root     __         0.00 secs Thu Sep  8 12:30
atrun            root     __         0.00 secs Thu Sep  8 12:30
cron         F   root     __         0.00 secs Thu Sep  8 12:30
sh               root     __         0.00 secs Thu Sep  8 12:15
atrun            root     __         0.00 secs Thu Sep  8 12:15
cron         F   root     __         0.00 secs Thu Sep  8 12:15
sh               root     __         0.00 secs Thu Sep  8 12:00
atrun            root     __         0.00 secs Thu Sep  8 12:00
cron         F   root     __         0.00 secs Thu Sep  8 12:00
Prompt %
```

What to Look For

This command is an excellent way of seeing what a user did while on your system because it lists all commands executed by all users.

Vulnerabilities
This command produces a file that tends to get quite large very quickly as it saves the data needed to track the commands issued by every user. You should periodically rename it so that you can manage smaller files.

The "lastcomm" command only tracks the command that ran a program, but not what actions were taken after the program started (for example, it may show the editor being run, but not which files were opened after the initialization of the editor).

Many times, attacks are not discovered until days after the actual event. And in these cases, the accounting logs may have been purged by the time the attack is discovered. The biggest potential intruder-style vulnerability is that the data is kept in the file /var/adm/pacct, which the intruder can
easily delete and perhaps modify if the proper privileges are obtained.

10.3.2 The /var/log/ syslog File

The **/var/log/syslog** file is a file that contains messages relating to various types of connections to your system. The content of this file is defined by the /etc/syslog.conf file. The results of this command contain extremely long lines; a shortened sample of this file is shown here.

```
Prompt %  more /var/log/syslog
Apr 20 13:04:22 host8 sendmail[15026]:
NAA15025:to=user8@sub.domain,
user7@sub.domain,user3@sub.domain, delay=00:00:02, mailer=smtp,
relay=computer.sub.domain. [128.xxx.xx.xx], stat=Sent (Mail
accepted)

Apr 20 13:04:23 host8 sendmail[15026]: NAA15025:
to=user5,user2, delay=00:00:03, mailer=local, stat=Sent

Apr 20 13:04:23 host8 sendmail[15026]: NAA15025:
to=user1@host1.sub.domain, delay=00:00:03, mailer=smtp,
relay=host1.sub.domain. [198.128.36.1], stat=Sent (Ok)

Apr 20 13:06:20 host8 in.telnetd[15032]: connect from
computer.sub.domain (198.xxx.xx.xx)
Prompt %
```

Most messages are from the sendmail program, and display the status of messages sent and received by your system. This file may also contain in.telnetd connection messages and other previously defined messages.

What to Look For
Since this file saves data on incoming as well as outgoing information, especially sendmail information, one of the things to look for is outbound E-mail to suspicious hosts. This may indicate that an intruder sent out information from your system to a remote system.

Telnet connections, both incoming and outgoing, should be examined. A short file may be suspicious, as it may indicate that this file has been edited or deleted. A ‘hole’ in the file
(a large chunk of time when no messages occur) may indicate that an intruder deleted the messages related to their time on the system. Note that this ‘hole’ may be useful in tracking down when the intruder used the system. In general, look for things that may appear out of the ordinary.

Vulnerabilities
In many cases, the /var/log/syslog file is world writable and must remain so for operational reasons. Therefore, its data may be suspect and untrustworthy.

This file tends to be very long. Investigating all connections, especially sendmail messages, can be difficult. This is because at least one line is written to the /var/log/syslog file for each mail message. In addition, users tend to delete messages and forget exactly who sent them the messages, when they were received, and what they were about.

10.3.3 The /var/adm/ messages File

The **/var/adm/messages** file usually contains a listing of all messages that are sent to the console. The actual content of this file is defined in the /etc/syslog.conf file. A sample of this file is shown here.

```
Prompt % more /var/adm/messages
Mar 21 10:36:04 host8 su: 'su root' failed for user1 on
/dev/ttyp2
Mar 21 10:36:08 host8 su: 'su aaa' succeeded for user1 on
/dev/ttyp2
Mar 21 16:00:59 host8 xntpd[121]: Previous time adjustment
didn't complete
Mar 24 15:01:44 host8 login: REPEATED LOGIN FAILURES ON
console, user3
Mar 25 11:42:51 host8 shutdown: reboot by user1
Mar 25 11:42:53 host8 syslogd: going down on signal 15
Mar 25 11:48:04 host8 su: 'su aaa' succeeded for user1 on
/dev/ttyp0
Mar 28 15:47:19 host8 login: ROOT LOGIN REFUSED ON ttyp3 FROM
machine.sub.domain
Mar 28 16:12:12 host8 login: ROOT LOGIN console
Apr 13 15:58:35 host8 su: 'su aaa' failed for user1 on
/dev/ttyp0
Apr 13 15:58:55 host8 su: 'su aaa' succeeded for user1 on
/dev/ttyp0
Apr 15 08:48:22 host8 named[2682]: starting. named 4.9.2 Wed
Nov 17 13:17:49 PST 1993
Apr 15 08:48:22 host8 named[2683]: Ready to answer queries.
Prompt %
```

What to Look For

The following may indicate undesired activity:

- an unauthorized user logging into the root directory
- attempts to “su” to root or a privileged account
- failed login attempts may be from a valid user making mistakes or from an intruder

In the sample file above, you would make sure that “user1” is a valid user logging into the aaa root privileged account.

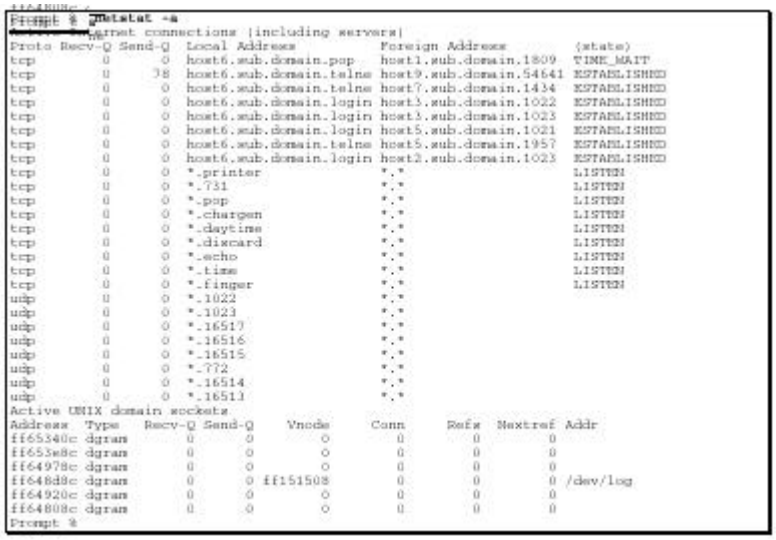
Vulnerabilities
Once an intruder obtains root access, this file can be modified or deleted quite easily. Also, if the syslog.conf file is compromised, logging to this file may be discontinued.

10.3.4 The "netstat" Command

The **"netstat"** command displays listening and connected processes. You should compare the output from this command with the output from the "last *-n*" command.

The command's "-a" parameter is used to display the status of all sockets.

A sample output is shown here.



```
Prompt % netstat -a
Internet connections (including servers)
Proto Recv-Q Send-Q  Local Address          Foreign Address        (state)
tcp        0      0  host6.sub.domain.pop   host1.sub.domain.1809  TIME_WAIT
tcp        0     38  host6.sub.domain.telne host9.sub.domain.54641 ESTABLISHED
tcp        0      0  host6.sub.domain.telne host7.sub.domain.1434  ESTABLISHED
tcp        0      0  host6.sub.domain.login host3.sub.domain.1022  ESTABLISHED
tcp        0      0  host6.sub.domain.login host3.sub.domain.1023  ESTABLISHED
tcp        0      0  host6.sub.domain.login host5.sub.domain.1021  ESTABLISHED
tcp        0      0  host6.sub.domain.telne host5.sub.domain.1957  ESTABLISHED
tcp        0      0  host6.sub.domain.login host2.sub.domain.1023  ESTABLISHED
tcp        0      0  *.printer              *.*                    LISTEN
tcp        0      0  *.731                  *.*                    LISTEN
tcp        0      0  *.pop                  *.*                    LISTEN
tcp        0      0  *.chargen              *.*                    LISTEN
tcp        0      0  *.daytime              *.*                    LISTEN
tcp        0      0  *.discard              *.*                    LISTEN
tcp        0      0  *.echo                 *.*                    LISTEN
tcp        0      0  *.time                 *.*                    LISTEN
tcp        0      0  *.finger               *.*                    LISTEN
udp        0      0  *.1022                 *.*
udp        0      0  *.1023                 *.*
udp        0      0  *.16517                *.*
udp        0      0  *.16516                *.*
udp        0      0  *.16515                *.*
udp        0      0  *.772                  *.*
udp        0      0  *.16514                *.*
udp        0      0  *.16513                *.*
Active UNIX domain sockets
Address  Type   Recv-Q Send-Q    Vnode     Conn     Refs  Nextref Addr
ff655340c dgram       0      0         0        0        0        0
ff6553e8c dgram       0      0         0        0        0        0
ff64978c dgram        0      0         0        0        0        0
ff648d8c dgram        0      0 ff151508         0        0        0 /dev/log
ff64920c dgram        0      0         0        0        0        0
ff64808c dgram        0      0         0        0        0        0
Prompt %
```

What to Look For

The following may indicate undesired activity:

- you have a telnet connection that does not correlate with the output from the "who" or "w" commands other network connections

Vulnerabilities
In some cases, compromised systems have been found to contain a Trojaned version of "netstat" that does not show connections to or from the source of the intrusion.

10.4 Detecting a Sniffer

Sniffers are a major source of contemporary attacks. This section shows you how to use the "ifconfig" command to determine if a sniffer has been installed.

10.4.1 The "ifconfig" Command

The **"ifconfig"** command displays the current configuration of your network interface. Most Ethernet adaptors are (and should be) configured to accept only messages intended for themselves. An attacker must set a computer's adaptor to "promiscuous mode," in order to listen to (and record) everything on its segment of the Ethernet.

A sample output of a system in promiscuous mode is shown here.

```
Prompt %  ifconfig -a
ie0: flags=63<UP,BROADCAST,NOTRAILERS,RUNNING,PROMISC>
        inet 987.654.32.1 netmask ffffff00 broadcast 987.654.32.255
lo0: flags=49<UP,LOOPBACK,RUNNING>
        inet 127.0.0.1 netmask ff000000
Prompt %
```

Note "PROMISC" is the last parameter of the flag's description.

What to Look For
In conjunction with positive results from the above command, the following may indicate undesired activity:

- newly created files
- a process that has an extremely high percentage of CPU

Vulnerabilities
Like any command, "ifconfig" can be Trojaned. If you suspect that a sniffer has been installed, obtain "cpm" from CIAC or CERT and run it. The cpm tool will test the network interface directly and report if it is in promiscuous mode.

10.5 Finding Files and Other Evidence Left by an Intruder

When an intruder breaks into a system, information related to the attack is occasionally left behind. This information includes, but is not limited to directories, shell scripts, programs, and text files.

This section describes various files that have been found on compromised systems. Because file names can be easily changed, the actual name of the file may be different than the file names listed in this section. Many times, intruders try to hide files; methods for achieving and detecting this will be also be described.

What to Look For

When you look for files left behind by an intruder, you should:

- obtain a baseline of what your normal operating system looks like
- find files and file and directory names commonly used by intruders
- examine system logs
- inspect log files
- inspect processes
- inspect targeted files
- look for X windows intrusions

Each of these tasks is described on the following pages.

Obtaining a Baseline of What Your Normal Operating System Looks Like To obtain a baseline of your normal operating system, you should periodically run the

commands described in this document. Record and become familiar with the output from these commands. Also, obtain and periodically use SPI and Tripwire.

Finding Files and File and Directory Names Commonly Used by Intruders The file names given in this section are commonly used by intruders. Start by looking for these file names, but realize that, as intruders learn that their bogus file names are discovered, they will change them. You must ultimately look for a name or names that do not belong.

Suspicious Files
Often, the best indication of whether or not a system has been compromised comes from a thorough examination of its file systems. The creation or modification of files is often a strong indication of intruder activity on a system. Occasionally, the intruder will modify ("Trojan") system
programs to hide the intrusion. Some system administrators have discovered that a command such as "ps" will be Trojaned to ignore the intruder's processes. Keep this in mind when running **any** command, because if a command has been Trojaned, the results of the command will be questionable.

```
Prompt #  find / -mtime -ndays -ls
```

The **"find"** command, run preferably as root, will list all files that have been modified in the previous *n* days: Note that many intruders routinely change file modification times to hide changes made to the system. Many of these modifications may still be detected by examining a file's inode change time, which is more difficult for an intruder to forge. The following command will locate all files with inode change times that have changed in the last *n* days:

While examining the results generated by the above commands, consider the hidden files and directories often used by attackers described in the next section,

```
Prompt #  find / -ctime -ndays -ls
```

"Hidden Files and Directories."

Hidden Files and Directories
Intruders often attempt to conceal their presence on a system by using hidden files or directories; that is, those with names that begin with a "." (period). They are not displayed by the "ls" command, unless the "-a" parameter is used. The following names are commonly used by intruders:

- "..." (period period period)
- ".. " (period period space)
- ".. " (period period space space)
- ".hushlogin"
- ".sh"
- ".xx"
- ".test"

Password Files and Crack
In many cases, intruders use compromised hosts to store and crack password files from other systems. Finding files that contain password entries from other systems or finding password cracking software (such as Crack) probably indicates intruder activity on your system.

Setuid Files
Unix systems allow users to temporarily elevate their privileges through a mechanism called setuid. When a file with the setuid attribute is executed by a user, the program is given the effective access rights of the owner of the file. For example, the "login" program is typically a setuid file owned by root. When a user invokes "login", the program is able to access the system with super-user privileges instead of the user's normal privileges. Intruders often create setuid files that enable them to quickly gain root access during later visits to the system. Often, the file is placed in a hidden directory or has a hidden filename (e.g., ".sh"). Setuid files appear in directory lists with an "s" in place of the "x" in the execute bit position. For example, the output of the **"ls -l .sh"** command would display output similar to the following:

```
-r-sr-xr-x   1 root     other      86012 Jun  2 01:09 .sh
```

Note that a typical Unix system contains dozens of legitimate setuid programs necessary for normal operation of the system. Setuid files that should be suspected include:

- files that appear to have been modified recently unfamiliar files
- files stored in user or temporary directories

To list all setuid files on your system, use the following command:

```
Prompt #  find / -user root -perm -4000 -print
```

10.6 Examining System Logs

All Unix systems provide some level of accounting, recording the actions of both users and system processes. The amount of information recorded can vary significantly depending on both the version of Unix and its configuration. The default for many systems is to record little more than login/logout times for users. At the other end of the spectrum, systems running at an Orange Book C2 level of assurance can easily generate several megabytes of log information per hour.

To detect an intrusion, begin by examining whatever logs are available on your system. Bear in mind, however, that if an intruder gained access to your system, the information stored in the logs may have been modified to hide the intruder's tracks. Use the "last" and/or "lastcomm" commands discussed in the next two sections (and previously described above) to help you examine the logs.

The "last" Command
The **"last"** command, available on almost every version of Unix, displays login and logout activity for the system. This can be a useful place to begin an investigation. Check the login times and locations for all users and compare them to expected norms. Refer to the previous discussion of the "last" command for more information and a sample output.

The "lastcomm" Command (Accounting)
On systems with process level accounting enabled, the **"lastcomm"** command will generate a detailed list of all commands executed by each user on the system.

Unusual or inappropriate system activity can often be discovered in the results from this command. For example, “lastcomm” output indicating repeated executions of the “tftp” program might indicate attempts to steal password files using TFTP. For information on enabling process accounting on a specific Unix system, refer to the man page for “acct”. Refer to the previous discussion of the “lastcomm” command for more information and a sample output.

10.7 Inspecting Log Files

Many system process events generate messages. For example, the “su” utility often makes a log entry when a user attempts to become the “super-user.” These messages may prove useful in discovering unusual activity possibly caused by an intruder.

These messages are often archived in log files for later examination. Commonly used files include /var/log/syslog and /var/adm/messages; however, the file names may vary from system to system. Refer to the sections about these files in this guide or to the man page for “syslog” for more information.

~/.history

Some shells, tcsh for example, keep a record of the most recently executed commands for each user. This information is usually stored in a file in the user’s home directory and is often called “.history”. Examining this file may allow the reconstruction of the recent activities of a specific user.

Inspecting Processes Look for:

- process names that are unfamiliar
- processes using an unusual amount of CPU time
- processes with names such as Crack or ypsnarf
- an unusually large number of processes

Keep in mind that process names can be changed.

Inspecting Targeted Files

/etc/passwd

Look for:

- new accounts
- changed uid
- no password
- a “+::” entry

~/.forward

The ~/.forward file is used to manipulate E-mail forwarding. When examining this file, look for any
suspicious entries (that is, would it make sense for a legitimate user to manipulate his or her E-mail in that manner?).

~/.rhosts and hosts.equiv

The ~/.rhosts file can be used to allow remote access to a system and is sometimes used by intruders to create easy backdoors into a system. If this file has recently been modified, examine it for evidence of tampering. Initially and periodically verify that the remote host and user names
in the files are consistent with local user access requirements. View with extreme caution a "+" entry; this allows users from any host to access the local system.

An older vulnerability is systems set up with a single "+" in the /etc/hosts.equiv file. This allows any other system to log in to your system. The "+" should be replaced with specific system names. Note, however, that an intruder cannot gain **root** access through /etc/rhosts entries.

~/ftp Files

Directories which can be written to by anonymous FTP users are commonly used for storing and exchanging intruder files. Do not allow the user "ftp" to own any directories or files.

System Executables in User Directories
Copies of what may appear to be system executables in user directories may actually be an attempt to conceal malicious software. For example, recent attacks have made use of binaries called "vi" and "sed", two commonly used Unix utilities. However, these particular binaries were actually renamed intrusion software files, designed to scan systems for weaknesses.

System binaries found in unusual locations may be compared to the actual executable using the **"cmp"** command:

```
Prompt %  cmp /home/jdoe/sed /usr/bin/sed
```

Determining if System Executables Have Been Trojaned SPI or Tripwire must be set up before an exposure in order to determine if your system executables have been Trojaned.

Use your CD-ROM to make sure you have a good copy of all your system executables, then run the above mentioned products according to the instructions that accompany them to create a basis for later comparison. Periodically, run SPI or Tripwire to detect any modification of the system executables.

/etc/inetd.conf
Print a baseline listing of this file for comparison. Look for new services.

/etc/aliases
Look for unusual aliases and those that redirect E-mail to unlikely places. Look for suspicious commands.

cron
Look for new entries in cron tab, especially root's. Look at each user's table.

/etc/rc*
Look for additions to install or reinstall backdoors or sniffer programs. Use SPI or Tripwire to detect changes to files.

NFS Exports
Use the **"showmount -a"** command to find users that have file systems mounted.

Check the /etc/exports (or equivalent) file for modifications. Run SPI or Tripwire to detect changes.

Changes to Critical Binaries
Run SPI or Tripwire initially and then periodically. Use the **“ls -lc”** command to determine if there have been inappropriate changes to these files.

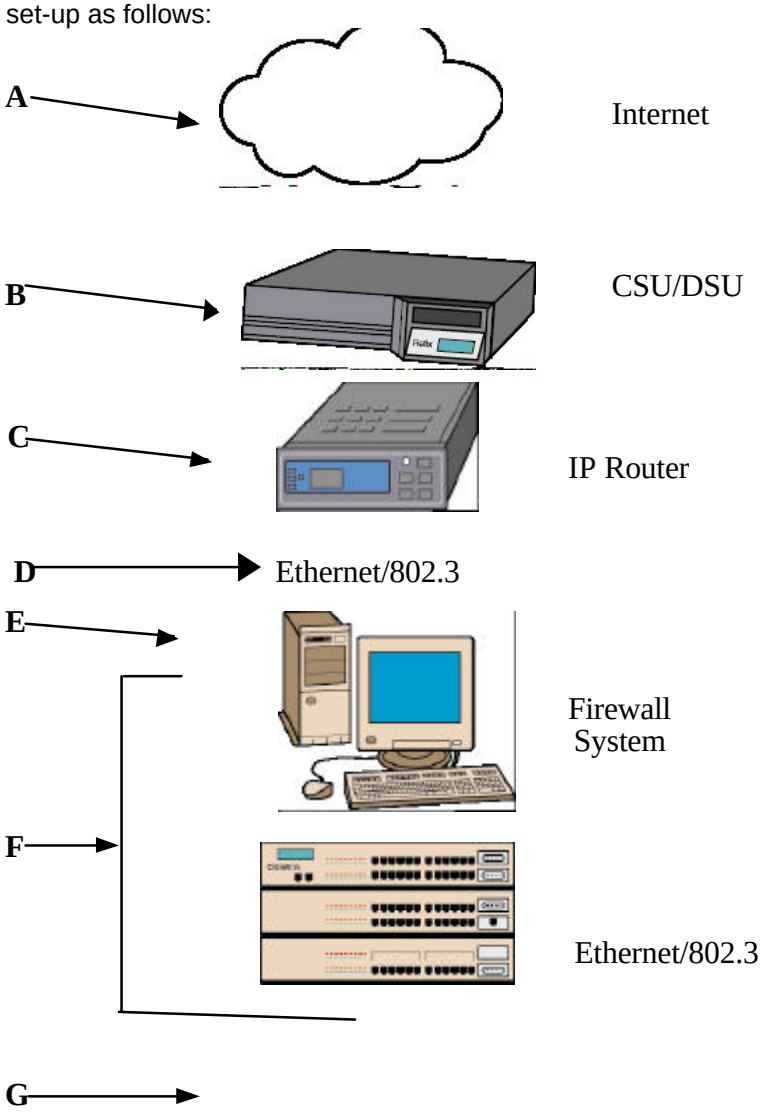
Note that the change time displayed by the “ls -lc” command can be changed and the command itself can be Trojaned.

Section References:

Pichnarczyk, Karen, Weeber, Steve & Feingold, Richard. "*Unix Incident Guide: How to Detect an Intrusion CIAC-2305 R.1*". C I A C Department of Energy. December, 1994.

All firewalls from any vendor that will be providing Internet firewall facilities require a routed connection to the Internet to provide traffic flow between the Internet and in-house network facilities. There are usually more than one router involved in such connections. With some effort, connections are successful but usually difficult to monitor and manage.

A typical set-up with an Internet Service Provider where a firewall is configured in the network is set-up as follows:



In the above diagram, the network and firewall connection parts are as follows:

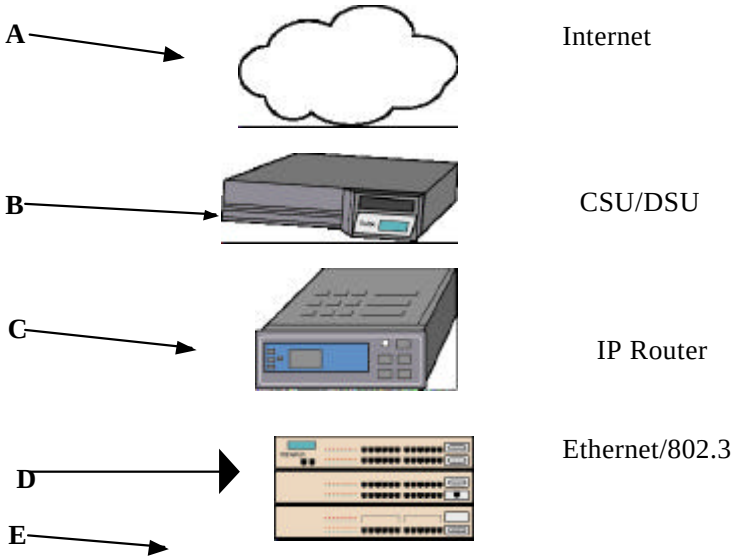
a) Internet connection provided by an Internet Service Provider (ISP)
b) A CSU/DSU interface to the telephone drop from the local equipment company (LEC)

c) A router system to connect to the ISP's router connection to the Internet
d) An Ethernet/802.3 or Token Ring/802.5 UTP connection from the router to the firewall
e) A "dual-homed gateway" firewall system with two LAN controllers (in this diagram, two Ethernet/802.3 connections are provided)
f) An Ethernet/802.3 UTP connection from the firewall to the internal network
g) An internal network configuration. In this case, a simple stacked hub architecture (e.g. Cabletron Mini-MAC)

The above is an illustration of a typical, but simple, network configuration between a customer network and the Internet where information provision (e.g. a Web Site) will not be used.

Using a Router as a "Screen"

One of the more popular configurations of a "firewall" is to use an external router as the singular security facility between an untrusted network (e.g. Internet) and the internal, trusted network. This configuration is called a "screening router" set-up. A typical configuration is as follows:



Trusted Network Hub

The network configuration for a "screening router" is as follows:

a) Internet connection provided by an Internet Service Provider (ISP)
b) A CSU/DSU interface to the telephone drop from the local equipment company (LEC)
c) A router system to connect to the ISP's router connection to the Internet. On this router, there are a variety of "filter" rules, which provide some level of security between the trusted internal network and the untrusted Internet connection.
d) An Ethernet/802.3 or Token Ring/802.5 UTP connection from the router to the internal network

e) An internal network configuration. In this case, a simple stacked hub architecture (e.g. Cabletron Mini-MAC)

While the router is a required part of the network connection, there are some definitive problems with using screening routers as the *only* network security interface to an untrusted network, including:

- Configuration of filters and security facilities in the router may be difficult to accomplish and knowledge about the intricacies of routing is required to do it correctly
- There usually is little or no auditing or logging of traffic and security information as most routers are diskless in nature and have no easy way to get information to secondary (disk) storage. Further, routers are built to route and not necessarily to handle logging of network traffic.
- It can be quite difficult for the network and security managers to get information out of the router on the paths and security rule base that was implemented
- Adding authentication is difficult, time consuming and expensive even if the router vendor supports such functions
- Sessions from other parts of the network may be "tunneled" on top of each other and, therefore, non-filterable by the router itself
- There is usually a user demand to open up features in a router that are not screenable by the router and therefore put the network (trusted side) at risk
- Any bug in the router's operating environment may not be detected and can compromise the network's security (there are numerous CERT and CIAC alerts about router bugs and security issues over the years)
- Routers can be "spoofed" with some types of IP header options that would cause the router to believe that an external packet "looks" like an internal packet to the router tables
- Over time, multiple connections on the router usually do not get the same security screening rules. This means that one path through the router may not have the same security facilities as another and this may allow alternate paths to compromise the security of the router.
- Routers are configured to route. Enabling any filtering facility in a router will degrade the router's performance. As more filters are added, the router's performance may degrade to a totally unacceptable performance level for traffic. As a result, many sites opt to remove necessary filtering for security to gain performance and end up compromising trusted network security and integrity.

Using a router on a network connection is a normal, essential function. Relying on the router as the only screen for security facilities is dangerous.

The following 20 base factors comprise the basic costing issues in the ownership of firewall products:

1. **Firewall requirements analysis prior to vendor selection.** This phase involves the technology assessment issues a company must go through to determine the threat to the corporate information structures, the risk of loss that would be associated with a connection that is unprotected, the risk of loss that could happen if the connection is breached, the known corporate information resources that must be protected and their relative priorities of protection categories, corporate security policies and procedures as related to any external network connection, corporate audit measurement and adherence requirements, technical details on what facilities are on-line and are threatened, etc...
2. **Corporate decisions on exactly what security policies need to be in-place** in any firewall to satisfy the corporate security requirements as defined in the initial needs analysis. This step is crucial to properly identifying to the firewall vendor WHAT the firewall will be programmed to protect. The vendors will need this list to identify if their product can provide the levels of protection required by the corporate need.
3. **Vendor product evaluation to determine a list of finalist vendors**. Typically, a corporate committee will be appointed to evaluate vendor offerings vis-a-vis the corporate firewall requirements list. In this stage of costing, the meeting with vendors and selection of, typically, no more than five finalists for the firewall product set is completed.
4. **Evaluation of finalist vendors.** This costing factor involves the testing and technical evaluation of the firewall vendor finalists to ensure that the selected vendor products can really provide the required corporate security services in the firewall product, that the product meets quality and management standards as defined in the requirement definition phase, that the firewall product(s) function as advertised by discussing the product with existing customers, that the firewall product performs technically as expected and provides required throughput to solve the firewall connectivity requirements and that the vendors meet corporate requirements of technical support, maintenance and other requirements that may have been defined.
5. **Selection of a vendor's product.** This phase involves the selection of a vendor and the political jostling that always takes place just prior to a decision in a corporate culture.
6. **Acquisition of hardware/software and basic set-up effort**. In this costing phase, the basic hardware, system software, firewall software and layered/additional products are acquired, configured and set-up so that security policies may be later added. Items would also include basic system management (backup/restore, system tuning, system and network management tool set-up, system/network management account set-up, etc.), network hardware interconnection and set-up (router installation, service acquisition from the Internet feed provider, cabinet and cable installation, power hook-up, basic hardware configuration and activation, etc.), etc...
7. **Training on the creation/definition/management of security policies** for the selected firewall. If the company intends to properly manage and maintain the firewall product set, training must be supplied to the technical staff which will be installing and maintaining the firewall facilities. If the staff is not familiar with technical aspects of firewall technologies, then additional training on firewall concepts, network security concepts, advanced network security technologies and security management must be undertaken. Failure to provide adequate

training on the firewall product will result in a much higher manpower costing factor for in-house personnel as well as a higher consultation costing factor due to the recurring need to secure outside help to make modifications to the firewall facilities to satisfy corporate needs as time goes on.

8. **Definition and installation of security policies for the firewall**. Using the requirements definitions, security filters are created that mirror the security requirements for use of the network connection that is provided via the firewall facilities. How long this phase takes depends heavily on the training provided to in-house personnel or the expertise in the system and firewall product set for the consultant(s) hired to implement the security policy filter baseline. There can be a very wide variance in manpower requirement from product to product.
9. **Testing of the firewall with the security policies installed**. This phase of costing is critical to reduce corporate risk factors and to ensure that the firewall is functioning properly. Typically, the filters are fully tested by in-house or consulting personnel and then a third party is contracted to provide a penetration study to verify integrity of the firewall and proper implementation of security policies implemented as filters in the firewall product set. How much testing is required is a function of corporate risk factors, estimated usage metrics, importance of reliability and many other issues.
10. **Release of the firewall connection to the user population.** For a period of time, there is a requirement to provide modifications and changes to satisfy a shake-down period of user access. This is usually a higher manpower requirement than the day-to-day management function that eventually settles into corporate use.
11. **Day-to-day technical management effort.** This costing factor involves the typical day-to-day functions required to keep the firewall functioning properly (checking of logs, events, backup/restore, disk maintenance, etc.) as well as the modifications and additions to the security policy rule base to accommodate new users, changes of service to existing users, moves of users, readdressing issues of systems on the network, added service facilities, etc. There may also be report-writing requirements to the company to show management and maintenance of the firewall as well as disposition of serious events and problems that need to be addressed as the product is used.
12. **Periodic major maintenance and upgrades.** As time goes on, there will be required down-time network activities that are required to satisfy hardware and software operational needs. The hardware will need to be periodically updated with additional disk space or memory, faster processing may be required via a new processing system, additional network controllers or faster network controllers may be added to the configuration and so on. Software-wise, the operating system may require upgrades to patch or fix problems, bug fixes and updates to the firewall software will be required, new security threats may be identified by vendors and updates to the security filters are required, etc. Further major maintenance may be required in the form of major system upgrades to support higher-speed Internet connectivity or to support multiple network feeds from Internet, customers, sister companies, etc.
13. **Remedial training for technical personnel.** As the systems and software are upgraded over time, the firewall software and operating environment will undergo extensive transformations to take into account new security facilities as well as new user facilities. This will require remedial training and updates to technical personnel to allow them to properly take advantage of the new facilities as well as to properly identify potential security risks and isolate them before they become problems for the company. Remedial training may also include attendance at national and international security conferences and outside training events for firewall and security efforts.
14. **Investigation of infiltration attempts.** As the firewall product set is used and connected to a publicly available network, chances are extremely likely that

unauthorized connections will be attempted by hackers and other disreputable individuals on the network. When these infiltration attempts occur, someone within the company will be required to investigate the whys and hows of the penetration attempt, report on the attempt and help management make decisions on what to do to defeat such infiltrations in the future as well as modify existing policies, filtering rules and other firewall functions to ensure security integrity in the firewall set-up. This effort, depending upon the visibility of the company, can be time consuming and expensive. It is labor intensive as tools on firewalls are only one component of the investigator's repertoir of facilities required to accomplish their mission.

15. **Corporate audits.** Needless to say, corporate EDP audit functionaries will require someone who understands the firewall set-up to work with them to ensure that corporate security requirements are properly implemented in the firewall facilities. For those companies without proper corporate audit expertise, an outside consultancy may be hired to evaluate the firewall set-up and operations from time to time to ensure integrity and reliability. In either case, someone familiar with the technical operations of the firewall set-up must be made available to the audit functionary and this takes time.
16. **Application additions to the network firewall connection.** As the network connection via the firewall increases in popularity and criticality to corporate business, the need to add application facilities and access to remote network facilities will increase. This leads to multiple meetings between firewall management team personnel and users/application implementers who wish to add applications over the firewall facilities. This will eventually result in new security policy filters, additional firewall packet loading and other performance and labor-related functions which affect overall cost of ownership. It may also require hardware and software upgrades faster than expected due to packet or application loading increases.
17. **Major outage troubleshooting.** From time-to-time, all technological components break and a firewall is no exception. When such outages occur, someone has to spend time defining the problem(s), finding solutions, implementing solutions and restoring the status quo ante. How much time this will take varies, but it usually is significant and intense as the firewall becomes a locus of activity during an outage of any kind.
18. **Miscellaneous firewall and network security meeting time (technical and political).** This factor is a catch-all for time spent explaining the firewall facilities to interested corporate groups or management as well as functioning as a "go-between" for information on facilities available to users. This factor can be extremely time consuming and does not generate any measurable progression as a general rule. It is manpower time required to keep things running smoothly and is, therefore, a cost factor.
19. **New firewall and network security technology assessment (ongoing).** As the firewall lifetime progresses, the need to evaluate new threats and new technologies that defeat new threats is important. Further, additional vendor features for a particular firewall product may need to be evaluated for inclusion into the existing facilities. For instance, if a new standard for remote authentication via firewalls is added to most products, this facility will need to be evaluated for use with the existing facilities. This takes time and technical effort.
20. **Application changes and network re-engineering.** All applications and network components change with time on any network. Prudent engineering requires that firewall facilities be re-evaluated for any changes in application set-up or network hardware changes that could affect the integrity of the firewall facility. Again, a time-consuming effort is involved.

As can be seen, properly (and improperly) defined and installed firewalls consume a great deal of time and resources. This makes them fairly expensive resources as

well as a strategic corporate resource - not a tactical one. The cost of a firewall is not the firewall itself - it is all the ancilliary functions and time involved. The more the extra costs are eliminated, the better the costing solution for the customer.

1. **Abuse of Privilege**: When a user performs an action that they should not have, according to organizational policy or law.
2. **Application-Level Firewall**: A firewall system in which service is provided by processes that maintain complete TCP connection state and sequencing. Application level firewalls often re-address traffic so that outgoing traffic appears to have originated from the firewall, rather than the internal host.
3. **Authentication**: The process of determining the identity of a user that is attempting to access a system.
4. **Authentication Token**: A portable device used for authenticating a user. Authentication tokens operate by challenge/response, time-based code sequences, or other techniques. This may include paper-based lists of one-time passwords.
5. **Authorization**: The process of determining what types of activities are permitted. Usually, authorization is in the context of authentication: once you have authenticated a user, they may be authorized different types of access or activity.
6. **Bastion Host**: A system that has been hardened to resist attack, and which is installed on a network in such a way that it is expected to potentially come under attack. Bastion hosts are often components of firewalls, or may be "outside" Web servers or public access systems. Generally, a bastion host is running some form of general purpose operating system (e.g., UNIX, VMS, WNT, etc.) rather than a ROM-based or firmware operating system.
7. **Challenge/Response**: An authentication technique whereby a server sends an unpredictable challenge to the user, who computes a response using some form of authentication token.
8. **Chroot**: A technique under UNIX whereby a process is permanently restricted to an isolated subset of the filesystem.
9. **Cryptographic Checksum**: A one-way function applied to a file to produce a unique "fingerprint" of the file for later reference. Checksum systems are a primary means of detecting filesystem tampering on UNIX.
10. **Data Driven Attack**: A form of attack in which the attack is encoded in innocuous-seeming data which is executed by a user or other software to implement an attack. In the case of firewalls, a data driven attack is a concern since it may get through the firewall in data form and launch an attack against a system behind the firewall.
11. **Defense in Depth**: The security approach whereby each system on the network is secured to the greatest possible degree. May be used in conjunction with firewalls.
12. **DNS spoofing**: Assuming the DNS name of another system by either corrupting the name service cache of a victim system, or by compromising a domain name server for a valid domain.
13. **Dual Homed Gateway**: A dual homed gateway is a system that has two or more network interfaces, each of which is connected to a different network. In firewall configurations, a dual homed gateway usually acts to block or filter some or all of the traffic trying to pass between the networks.
14. **Encrypting Router**: *see Tunneling Router and Virtual Network Perimeter.*
15. **Firewall**: A system or combination of systems that enforces a boundary between two or more networks.
16. **Host-based Security**: The technique of securing an individual system from attack. Host based security is operating system and version dependent.
17. **Insider Attack**: An attack originating from inside a protected network.

18. **Intrusion Detection**: Detection of break-ins or break-in attempts either manually or via software expert systems that operate on logs or other information available on the network.
19. **IP Spoofing**: An attack whereby a system attempts to illicitly impersonate another system by using its IP network address.
20. **IP Splicing / Hijacking**: An attack whereby an active, established, session is intercepted and co-opted by the attacker. IP Splicing attacks may occur after an authentication has been made, permitting the attacker to assume the role of an already authorized user. Primary protections against IP Splicing rely on encryption at the session or network layer.
21. **Least Privilege**: Designing operational aspects of a system to operate with a minimum amount of system privilege. This reduces the authorization level at which various actions are performed and decreases the chance that a process or user with high privileges may be caused to perform unauthorized activity resulting in a security breach.
22. **Logging**: The process of storing information about events that occurred on the firewall or network.
23. **Log Retention**: How long audit logs are retained and maintained.
24. **Log Processing**: How audit logs are processed, searched for key events, or summarized.
25. **Network-Level Firewall**: A firewall in which traffic is examined at the network protocol packet level.
26. **Perimeter-based Security**: The technique of securing a network by controlling access to all entry and exit points of the network.
27. **Policy**: Organization-level rules governing acceptable use of computing resources, security practices, and operational procedures.
28. **Proxy**: A software agent that acts on behalf of a user. Typical proxies accept a connection from a user, make a decision as to whether or not the user or client IP address is permitted to use the proxy, perhaps does additional authentication, and then completes a connection on behalf of the user to a remote destination.
29. **Screened Host**: A host on a network behind a screening router. The degree to which a screened host may be accessed depends on the screening rules in the router.
30. **Screened Subnet**: A subnet behind a screening router. The degree to which the subnet may be accessed depends on the screening rules in the router.
31. **Screening Router**: A router configured to permit or deny traffic based on a set of permission rules installed by the administrator.
32. **Session Stealing**: *See IP Splicing.*
33. **Trojan Horse**: A software entity that appears to do something normal but which, in fact, contains a trapdoor or attack program.
34. **Tunneling Router**: A router or system capable of routing traffic by encrypting it and encapsulating it for transmission across an untrusted network, for eventual de-encapsulation and decryption.
35. **Social Engineering**: An attack based on deceiving users or administrators at the target site. Social engineering attacks are typically carried out by telephoning users or operators and pretending to be an authorized user, to attempt to gain illicit access to systems.
36. **Virtual Network Perimeter**: A network that appears to be a single protected network behind firewalls, which actually encompasses encrypted virtual links over untrusted networks.
37. **Virus**: A self-replicating code segment. Viruses may or may not contain attack programs or trapdoors.

1. Firewall and System Probing

Hackers are using sophisticated, automated tools to scan for vulnerabilities of a company's corporate firewall and systems behind the firewall. These hacker tools have proved to be quite effective, with the average computer scan taking less than three minutes to identify and compromise security.

Companies can prevent this by ensuring that their systems sit behind a network firewall and any services available through this firewall are carefully monitored for potential security exposures.

2. Network File Systems (NFS) Application Attacks

Hackers attempt to exploit well-known vulnerabilities in the Network File System application, which is used to share files between systems. These attacks, usually through network firewalls, can result in compromised administrator access.

To combat this, ensure systems do not allow NFS through the firewall, and enable NFS protections to restrict who can access files.

3. Electronic Mail Attacks

Hackers can compromise network systems by simply sending an e-mail to it. Companies who accept e-mail from the Internet and who have exposed versions of the sendmail program are potential targets from this attack. Last year more than 20,000 systems were compromised due to this exposure.

To prevent this from occurring, check with vendors to ensure systems are running a correct version of sendmail or some more secure mail product.

4. Vendor Default Password Attacks

Systems of all types come with vendor-installed usernames and passwords. Hackers are well educated on these default usernames and passwords and use these accounts to gain unauthorized administrative access to systems.

Protect systems by ensuring that all vendor passwords have been changed.

5. Spoofing, Sniffing, Fragmentation and Splicing Attacks

Recently computer hackers have been using sophisticated techniques and tools at their disposal to identify and expose vulnerabilities on Internet networks. These tools and techniques can be used to capture names and passwords, as well as compromise-trusted systems through the firewall.

To protect systems from this type of attack, check with computer and firewall vendors to identify possible security precautions.

6. Social Engineering Attacks

Hackers will attempt to gain sensitive or confidential information from companies by placing calls to employees and pretending to be another employee. These types of attacks can be effective in gaining usernames and passwords as well as other sensitive information.

Train employees to use a "call-back" procedure to verify the distribution of any sensitive information over the telephone.

7. Easy-To-Guess Password Compromise

Most passwords that are easy to remember are also easy to guess. These include words in the dictionary, common names, slang words, song titles, etc. Computer hackers will attempt to gain access to systems using these easy-to-guess passwords usually via automated attacks.

Protect systems by ensuring that passwords are not easy to guess, that they are at least eight characters long, contain special characters and utilize both uppercase and lowercase characters.

8. Destructive Computer Viruses

Computer viruses can infect systems on a widespread basis in a very short period. These viruses can be responsible for erasing system data.

Protect systems from computer viruses by using anti-virus software to detect and remove computer viruses.

9. Prefix Scanning

Computer hackers will be scanning company telephone numbers looking for modem lines, which they can use to gain access to internal systems. These modem lines bypass network firewalls and usually bypass most security policies. These "backdoors" can easily be used to compromise internal systems.

Protect against this intrusion by ensuring modems are protected from brute force attacks. Place these modems behind firewalls; make use of one-time passwords; or have these modems disabled.

10. Trojan Horses

Hackers will install "backdoor" or "Trojan Horse" programs on businesses computer systems, allowing for unrestricted access into internal systems, which will bypass security monitoring and auditing policies.
Conduct regular security analysis audits to identify potential security vulnerabilities and to identify security exposures.

ATTACK NAME	SYMPTOMS	DESCRIPTION	NOTES
Boink (similar to Bonk, Teardrop and New Tear/Tear2), a hack	System seizure	Bad fragment attack	Sends bad packet fragments that cannot be correctly reassembled, causing the system to fail
DoS (Denial of Service)	Lack of access to resources and services	Denial of Service attacks tie up system resources doing things you do not want so you cannot get service	Examples include floods (which soak up bandwidth and CPU) and disconnects (which prevent you from reaching hosts or networks)
Floods (Nukes), a DoS attack	n/a	Large amounts of ICMP (usually) or UDP useless packets	Ties up system by making it respond to floods of useless garbage
ICMP flooding (flood ping), a DoS attack	Loss of bandwidth (slow responses from the Internet) and poor response time on the desktop	A flood of ICMP (ping) requests that tie your system in knots responding to garbage traffic. This is analogous to wasting your time answering the door to never-ending doorbells that do nothing.	Ties up CPU time and wastes your bandwidth with the garbage traffic. For example, "Pingexploit" typically attacks Unix systems with oversized ICMP packet fragments.
Identification flooding (Identd), a DoS attack	Loss of bandwidth (slow responses from the Internet) and poor response time on the desktop	Similar to an ICMP flood, but requests information from your system (TCP port 113)	Very often slows the CPU down (even more than an ICMP flood) since identification responses take more time than ICMP responses to generate
Jolt (SSping, IceNuke), a hack	System seizure	Oversized, fragmented packet which causes the system to seize up	System stops working and must be rebooted
Land, a hack	System seizure forcing cold reboot	Spoofing attempt which establishes TCP/IP connection to you from you. This SYN request forces the system to connect to itself, thereby locking itself up.	The attacked system attempts to connect to itself and seizes up

Hack	N/A	An application or a packet that exploits a weakness in operating system, application or protocol	Varied results. Examples include smurf, teardrop, land, newtear, puke, ssping, jolt, etc.
Pong, a hack	Loss of bandwidth (slow responses from the Internet) and poor response time on the desktop	Flood of spoofed ICMP packets, usually changing the spoofed source address with every packet	Reboot to solve
Puke, a hack	Disconnection from a server (usually IRC)	Spoofs an ICMP unreachable error to a target. This forces a disconnect from a server.	Usually preceded by an ICMP port scan where "pings" are sent to a system to find a vulnerable port being used to connect to a server
Scan, a generic technique and a DoS attack	System slows	A progressive, systematic testing of ports for an "opening." This attack can chew into system resources since its target is usually changing. It often requires a proper firewall or large, multi-port block to prevent.	Usually used prior to a hack to find a vulnerable attack spot. This is considered a brutish form of attack and is not as effective as other floods for tying up resources. It usually precedes a more "elegant" attack form.
Smurf, a hack	A very effective CPU crushing flood-like attack. Apparent system seizure.	Spoofs ICMP packets requesting a response and triggering multiple responses	A form of flood that is very dangerous since it can get a "many-for-one" effect, tying up lots of CPU cycles for relatively few packets sent
Spoofing (IPspoof)	N/A	An attack masking style that makes traffic appear to come from a legitimate target or that attempts to frame innocent bystanders for attacks for which they are not responsible	Particularly nasty attack because hacks, floods and nukes are illegal in most countries and subject to prosecution

unreachable (dest_unreach)- a DoS attack	"Destination Unreachable" messages and disconnection from a server	There are 2 forms of this—client unreachable and server unreachable. The server unreachable attack sends an ICMP message to the system fooling it into thinking its traffic can no longer reach the server, so it gives up. The client unreachable form does the same thing to the server with respect to your system.	
WinNuke, a hack and a DoS attack, but not a flood	Loss of networking resources	Sends OOB (Out-of-Band) data to port 139 and exploits Win 3.11, Win95, Win NT 3.51 and Win NT 4.0 systems	Does not crash the system, but it causes a fatal exception requiring a reboot to regain TCP/IP (Internet) connectivity

1. Firewall Sensitive Systems

Ensure corporate systems are protected from Internet attacks. Deploy a firewall between these systems and the Internet to guard against network scans and intrusions.

2. Obtain Security Alert Information

Subscribe to security alert mailing lists to identify potential security exposures before they become problems. CERT (Computer Emergency Response Team at Carnegie Mellon University) is a good place to start. The URL for CERT's Web site is cert-advisory-request@cert.org. The e-mail address is cert@cert.org.

3. Review System Audit Trails Regularly

Regularly check logging data and audit trails to look for unusual or suspicious activity.

4. Backup Data

Don't be a victim of accidental or malicious data erasure. Backup all sensitive data on a regular basis.

5. Purchase and Deploy Anti-Virus Software

Computer viruses can spread throughout a system in minutes. Check systems for viruses on a regular basis.

6. Change Passwords On A Regular Rotational Basis

Don't pick easy to remember passwords and change them often. Consider the use of one-time password tokens to avoid password compromise threats.

7. Deploy Vendor Security Patches

Consult with vendors and obtain any system security patches that can be used to add additional layers of protection.

8. Establish and Enforce A Security Policy

Develop and enforce a company-wide computer and physical security policy.

9. Employee Awareness

Ensure all employees and management are briefed regularly on security threats, policies, corrective measures and incident reporting procedures.

10. Make Use Of Public Domain Security Tools

A variety of public domain security tools exist on the Internet, many of which can be used to assist in the protection of computer systems.

Back Door: An entry to a program, or system created by its designer to allow special access; often without proper security checks. A classic back door was used by a teen-age hacker in the movie "War Games".

Bacterium: A program which spreads to other users or systems by copying itself as a by product of execution. It doesn't infect other programs, but acts independently.

Bogus Programs: Programs which do not do what they have been advertised to do. A example is XTRATANK, which claims to double your hard drive space. It merely diddles the file allocation to double the reported size of the disk.

Boot Sector Virus: A virus secreted in the boot sector or replacing the boot sector on a floppy disk. Also a virus on the master boot block of a hard disk, or in the partition table of a hard disk. N.B. even non-systems floppy disks still have a boot sector; they just lack the boot program on that block ! Examples are Stoned and Michelangelo viruses.

Bug: An error in the design or implementation of a program, that causes the program to do something unintended. Remember even viruses have bugs. The original "bug" was a moth stuck in a relay of ENIAC.

Checksum: a number that uniquely defines a file, block or other bit of computer code. A checksum is calculated by applying an algorithm to each byte of the code and rotating it, logically ANDing or ORing it to some standard, or otherwise encoding it. The result is a single number which is a numeric finger-print. See cyclic redundancy check (CRC).

Cracks: Programs with the anti-copying protection removed, disabled or by-passed. Both hard-ware and software anti-pirating techniques can be broken with the appropriate knowledge and software.

Cyclic Redundancy Check (CRC) - A unique numeric finger-print of a file, block or other bit of computer code. This is usually calculated using a look-up table. It is common in error checking protocols. See checksum.

Device Bomb: A program which executes based on the presence of a particular device, such as a com port, hard-drive D:, etc., usually with malicious actions.

Droppers: Programs which have a legitimate use, but contain viruses which are secretly planted in system. Droppers may actually be commercial software hacked to drop viruses.

FAT: File Allocation Tables. These areas of the formatted floppy or hard disk contain information used by the system to locate and maintain the file structure.

File Viruses: These viruses infect files with *.COM or *.EXE extensions. Friday the 13th is an example. Also included in this category are viruses which use the "corresponding files" technique. These viruses search for directories with files with .EXE extensions and then creates a file of the same name with a .COM extension. Since DOS executes files with the *.COM extension before those with the .EXE extension, the virus is executed and then passes control to the .EXE file.

Hacks: Software which has been illegally modified by a system expert. See cracks, pirates, droppers, etc.. This may be as simple as modifying parts of the code with a debugger; to patching the system to snatch interrupts.

Hoaxes: Programs which claim to do the impossible; and don't. An example is a file 2496 which claims to provide instructions on running a 2400 bps modem at 9600 or even 14400 bps. If you follow the instructions, you get a modem which runs at 0 bps.

Immunization: An anti-virus strategy to prevent virus infection. This may involve putting a virus signature into software to be immunized in hopes of fooling a virus into believing the code is already infected. It may also involve creating checksums for each file which can be compared during later anti-virus examinations to guard against virus infection.

Interrupt: A hardware or software signal which indicates to the OS some event such as a keystroke has happened. It is typically taken care of by an interrupt handler which services the event.

Jokes: Programs which do something intended to be amusing, without causing serious harm, or replicating. BUGS, which cause little bugs to run across the screen when executed is an example.

Logic bomb: A program which executes on the occurrence, or lack of occurrence of a set of system conditions. Classic examples are programs which cease functioning if the programmer's name is removed from the company's payroll list.

Multi-partite Viruses: These viruses infect both boot sectors and files. Tequila is an example.

Pirates: Any illegally obtained software. Also software which has had the copy-right notices, or other identification altered or removed.

Polymorphic Viruses: These viruses change their characteristics as they replicate. Many of these utilize the Bulgarian Dark Avenger's mutating engine. The Whale virus is an example.

Rabbit: A program designed to exhaust a system resource (e.g. CPU time, disk space, terminal I/O, etc.) by replicating itself without limit. It differs from a bacterium in that it is specifically targeted at a system resource; and from a virus in that it is a self contained program.

Rogue Program: A program that is no longer under the control of its owner, the system or its executing terminal; a.k.a. zombie. A virus is the ultimate rogue program!

Stealth Viruses: These viruses conceal the results of infection; keeping file length unchanged for example, or modifying the file in such a way that the checksum is not changed. They may simply alter the system so that the file length is reported unchanged although it is actually increased. Hundred years is an example.

Systemic Viruses: These viruses infect parts of the system other than the boot block. The file allocation table (FAT), device tables, directories, device drivers and COMMAND.COM are typical targets. Number of the Beast is an example.

Time Bomb: A logic bomb activated after a certain amount of time, or on a certain date. The classic example is a program that ceases functioning on a given date, as a control for leasing it. Such a program is often re-activated by an appropriate password.

Trojan Horse Programs: A program which has a hidden aspect which causes malicious damage. The classic is AIDS, which purports to be an AIDS data base, but actually destroys the hard disk when executed. False logon screens which snatch the users logon ID and password are another example.

Virus (pl. viruses): a program that can "infect" other software by modifying them to include a copy of itself. A program need not cause malicious damage to be a virus; the act of "infecting" other programs is central to the definition.

Worm: A program that spreads copies of itself through-out a network. The first use of the term was applied to a program that copied itself benignly around a network, to use otherwise unused resources for distributed computation. A worm becomes a security problem when it spreads against the wishes of the system owners, and disrupts the network by overloading it.

AAL An acronym for *ATM adaptation layer*, which interprets the type and format of user data messages, and then translates these messages into ATM format by packaging them into the 48-byte payload portion of an ATM cell. The AAL's interpretation of data type and format is based on the specific class of service assigned to the data by the application. The AAL provides support for four different service classes and provides five different AAL types to accommodate a particular service class. *AAL1* is used for data that require connection-oriented, constant-bit rate transmissions (e.g., voice transmissions); *AAL2* is used for data that require connection-oriented variable-bit rate transmissions (e.g., a videoconferencing application); *AAL3* and *AAL4* are used for connection-oriented or connectionless variable-bit rate transmissions (e.g., bursty data typical of LAN applications such as those found on frame relay and SMDS networks); and AAL5, which is an improvement to AAL3, is used for transmissions in which higher layer protocols provide error recovery.

AAUI Apple Computer Corporation's proprietary attachment unit interface (AUI). "AAUI" stands for "Apple Attachment Unit Interface."

Access Line A term used in *frame relay* to denote the *local loop*. Also called *port connection*.

Active Monitor A station on a token ring network that oversee the ring and ensure that it is functioning properly. Also called a *monitor station*.

Address A unique number assigned to a device to identify its location within a network. An address also can uniquely identify a network application process.

Addressing A network concept that describes the process of assigning unique identification numbers (called *addresses*) to a networked device.

ADSL An acronym for *asynchronous digital subscriber line*, which is a *DSL* variant in which traffic is transmitted at different rates in different directions. Downstream rates range from 1.5 Mbps to 9 Mbps; upstream rates range from 16 kbps to 1 Mbps. Rates depend on line quality and local loop distance. Suitable for Internet or intranet access, video-on-demand, database access, remote LAN access.

ADSL Lite A slower *ADSL*; also called G.lite. Downstream rates equal 1 Mbps; upstream rates equal 128 kbps. Intended primarily for homes.

Alignment Error An Ethernet/802.3 frame that does not end on a "byte-boundary."

Always On/Dynamic ISDN (AO/DI) An initiative from the Vendor's ISDN Association (VIA) in which a portion of the D channel, which is always active and constantly connected to the provider's switch, is used to transmit user packet data.

Ambient Noise Electrical *noise* that is always present and is generated primarily by transmission equipment like transmitters, receivers, and repeaters. Ambient noise also can be induced by external sources such as fluorescent light transformers, electrical facilities, and heat. Ambient noise makes it difficult for receiving equipment to distinguish between incoming signals. Also called *thermal noise*.

Analog Refers to any physical device or signal that varies continuously in strength or quantity over an infinite range of voltages or currents. An example is voltage in a circuit.

Analog Communication Refers to any communication method based on analog principles. In analog communications, signals flow across a wire in the form of electromagnetic waves. These waves resemble a sine curve and have the following three characteristics: *amplitude*, which is the level of voltage on a wire (or the intensity of a light beam when dealing with fiber-optic cable); *frequency*, which is the number of oscillations, or cycles, of a wave in a specified length of time; and *phase*, which is the point a wave has advanced within its cycle. Typically associated with voice transmission rather than data transmission because voice transmission facilities, such as the telephone, were initially analog-based.

Application Gateway Firewall See *proxy server*.

Application Program Software that performs a specific function such as e-mail.

Application Protocol Defines how an application is to be implemented on a network. Also includes specific user programs for interacting with an application.

ARP An acronym for *address resolution protocol*, which is an Internet protocol that binds a node's *IP address* to its corresponding *MAC* sublayer (hardware) address.

Asynchronous Communication A data transmission method that requires the sending node to encapsulate special start and stop bits within each unit of data being transmitted. Thus, data can be transferred at any time by the sending node without the receiving node having any advance notification of the transfer.

ATM An acronym for *asynchronous transfer mode*, which is a connection-oriented, full- duplex, and point-to-point high-speed cell-switched network architecture that was created in the late 1980s/early 1990s to apply circuit switching concepts to data networks. Designed to carry data in 53-octet cells, ATM can be used to transmit data, voice and video—separately or simultaneously—over the same network path. Although not based on any specific physical layer protocol, ATM is generally carried over *SONET*. Also known as *cell relay* to distinguish it from *frame relay*.

Attenuation The decrease in signal strength, which occurs as the signal travels through a circuit or along a cable. The longer the cable, the greater the attenuation. Also, the higher the frequency of the signal, the greater the attenuation.

AUI A 15-pin “universal” connector that allows a device to be connected to UTP, thick or thin coax, or fiber-optic cable via an external transceiver. “AUI” stands for “attachment unit interface.

Autonomous System (AS) A collection of networks controlled by a single administrative authority, and which share a common routing strategy. Routers connecting networks within an AS trust each other and exchange routing information using a mutually agreed upon routing protocol. Also known as a routing domain or protocol area.

Auto-wrapping A term used to describe the “self healing” of a token or FDDI ring that has been cut in a single spot. The break in the active ring is corrected by establishing a loopback connection to the inactive ring. This creates a single virtual ring and allows the network to continue to function at full speed.

B Channel A 64 kbps ISDN clear channel (no signaling information is sent on the channel) used to transmit computer data (text and graphics), digitized voice, and digitized video. Most basic ISDN services are based on multiple B channels.Also called a *bearer channel*.

Backbone Switch A term used to describe one application of an Ethernet switch in which the switch serves as the backbone for the entire LAN. In this application, the network topology is called a “collapsed backbone.”

Bandwidth In *analog communications*, bandwidth is the total capacity of a communications channel measured in Hertz (Hz). It is the difference between the highest and lowest frequencies capable of being carried over a channel. The greater the bandwidth, the more signals that can be carried over a given frequency range. In *digital communications* and networking, bandwidth is the theoretical capacity of a communications channel expressed in bits per second (bps), which is called *data rate*.

Baseband Cable Uses the entire bandwidth of the cable to carry a single signal.

Baud A unit of signaling speed, named after the French engineer Jean Maurice Emile Baudot (1845-1903). It is another term used to express the capacity of a channel, but is different from bits per second.

Baud Rate A measure of the number of times line conditions (i.e., frequency, amplitude, voltage, or phase) change each second. At low speeds (under 300 bps) data rate (measured in bps) and baud rate are the same because signaling methods are relatively simple. As speed increases, signaling methods become more complex. Baud rate then differs from data rate because several bits are typically encoded per baud. That is, each signal can represent more than one bit of information.

Bearer Channel See *B channel*.

BECN An acronym for *backward explicit congestion notification*, which is a one-bit field in a *frame relay* frame that is set to 1 by a frame relay switch to denote that a frame transmitted toward the sending node experienced congestion.

Bend Radius The radius in which cable (copper or fiber) can be curved or "bent" without breaking. Fiber is much more flexible than copper cable and can be bent in much smaller radii than equivalent copper.

B-ISDN An acronym for *broadband ISDN*, which is an extension of ISDN that provides full-duplex data transmission at OC-12 rates (622.08 Mbps) and is designed for delivery of interactive services (e.g., videoconferencing and video surveillance), and distribution services (e.g., cable TV and high definition TV). B-ISDN is also the basis for *ATM*.

Bit-Time A unit of measure equal to 0.1 µs. Thus, a one bit transmission requires 0.1 µs. Transmitting a 64-byte Ethernet/802.3 frame requires 512 bit-times or 51.2 µs.

BNC Connector A type of connector used with thin coaxial cable. There are several interpretations of BNC, including Bayonet Neill-Concelman (named after its developers), Bayonet Nut Connector, Barrel Nut Connector., and British National Connector.

BONDING An acronym for *bandwidth on demand interoperability network group*, which is a protocol that aggregates two *ISDN B* channels into a single 128 Mbps circuit.

BRI An acronym for *basic rate interface*, which is an *ISDN* basic access channel that comprises two 64 kbps B channels, one 16 kbps D channel, and 48 bits of overhead used for framing and other functions. Commonly written as 2*B* + *D*.

Bridge A layer 2 device that interconnects two or more individual LANs or LAN segments. A transparent bridge is used in Ethernet/802.3 and 802.5 (Token Ring) networks; a source routing bridge (introduced by IBM) is used exclusively in token ring networks. Bridges keep local traffic local, but forward traffic destined for a remote network. Forwarding/filtering decisions are based on MAC sublayer (i.e., hardware) addresses. Bridges partition Ethernet/802.3 networks into multiple collision domains.

Broadband Cable Shares the bandwidth of a coaxial cable among multiple signals.

Broadcast A data transmission that s destined to all hosts connected to a network. A broadcast message is a special *multicast* message.

Broadcast Design A network configuration that consists of nodes sharing a single communications channel. Every node connected to this shared medium "hears" each other's transmissions.

Broadcast Storm A network phenomenon that occurs when several broadcast messages are transmitted at the same time. Broadcast storms can use up a substantial amount of network bandwidth, and in many cases, can cause a network to crash or shut down.

Brouter A combination bridge-router; a bridge with routing capabilities.

Bus Design A specific design based on a broadcast topology. All nodes are directly connected to the same communications channel.

Cable See *wire*.

Cable Modem A modem that uses cable television lines for data communications. These lines use broadband coaxial cable, which has a multitude of frequencies available and significantly higher bandwidth than the UTP cable used by the telcos. Cable modems provide an Ethernet/802.3 network interface that enables a computer to connect to the cable. Once connected, it is as if the PC were connected to an Ethernet/802.3 LAN. The connection is always "up," and multimegabit data rates are possible. Depending on the cable operator and service, current upstream rates fro cable modems are somewhere between 500 Kbps to 3 Mbps; downstream rates range between 10 Mbps to 30 Mbps.

Capacitance The property of a circuit that permits it to store an electrical charge. The capacitance of a cable determines its ability to carry a signal without distortion. The lower the capacitance, the longer the distance a signal can travel before signal distortion becomes unacceptable.

Carrier Sense Protocol A network protocol that requires nodes to first listen ("sense") for the "sound" of another node's transmission prior to accessing a shared channel.

CCITT An acronym for *Consultative Committee for International Telephony and Telegraphy*, which was formerly an international standards organization. CCITT is now part of *ITU*.

CDDI An acronym for *copper distributed data interface*, which is an interface that provides a 100 Mbps data transmission rate over copper. A CDDI network is similar to an FDDI network. CDDI also is restricted to connections between concentrators on the ring and single attachment devices, not for the ring itself.

Cell A unit of data that is transmitted across a network. Similar to a data *frame*. When used in the context of *ATM*, a contains exactly 53-bytes—48 bytes for user data and 5 bytes for overhead.

Cells in Frames (CIF) Defines a method for transporting ATM protocols over Ethernet and token ring LANs. CIF is a LAN technology that provides LANs with ATM features including QoS and the seamless integration of data, voice, and video.

Centralized System A single computer that provides all the computing resources for all offices and departments within an organization via computer *terminals* that are connected to the centralized system.

Check Bits See *redundancy bits*.

Checksum A parameter used to detect errors. Checksums are calculated using a predetermined *generator polynomial* and assigned to a specific checksum field of a data frame.

CIDR An acronym for *classless inter-domain routing*, which allows sites to advertise multiple *IPv4* Class C networks by using a single prefix.

Ciphertext A coded message. See *encryption*.

CIR An acronym for *committed interface rate.*

Circuit Gateway Firewall A device or product that involves monitoring the session set- up between a system and the user security options relative to that system for a particular user. For instance, a circuit gateway might check user IDs and passwords, or it might implement proxy connection authorization or other types of authentication services. A circuit firewall is also responsible for logging who came from where and went to what.

Circuit-switched Network A network design in which a dedicated physical circuit is established between the source and destination nodes before any data transmission can take place. Furthermore, this circuit must remain in place for the duration of a transmission.

CIX An acronym for *commercial Internet exchange*, a subscription organization consisting of a consortium of commercial and nonprofit regional network providers that began offering Internet service independent of the NSFNET backbone and without NSF's restriction on traffic type. Today, CIX serves as an Internet interconnect site similar to a *NAP*.

Class I Repeater A type of repeater used in Fast Ethernet LANs. Class I repeaters support both of Fast Ethernet's signaling schemes—100BASE-T4 and 100BASE TX/FX.

Class II Repeater A type of repeater used in Fast Ethernet LANs. Class II repeaters support only one of Fast Ethernet's signaling scheme—100BASE-T4 or 100BASE TX/FX.

Class of Service (CoS) A data prioritization scheme that tags data with a specific priority level. Higher priority data get delivered before lower priority data.

CLEC An acronym for *competitive local exchange carrier*, which refers to a new telecommunication service provider formed after the Telecommunications Act of 1996 in the United States.

Client A networked device that requests resources from a *server*.

Client-Server. A model or paradigm that describes network services and the programs used by end users to access these services. The client side (or front end) provides a user with an interface for requesting services from the network, and the server side (or back end) is responsible for accepting user requests for services and providing these services transparent to the user.

Coaxial Cable A type of cable that consists of a single-wire conductor, surrounded by a dielectric material and two types of shielding, a foil shield and a braided shield, arranged concentrically and encased in a PVC or Teflon outer jacket.

Collapsed Backbone A term used to describe a network topology in which all LAN segments are interconnected via a bridge or switch, which serves as the network backbone.

Collision The term used to describe what happens when two or more nodes attempt to transmit data simultaneously on an Ethernet/802.3 network: Their signals collide resulting in a collision.

Collision Domain A "field" within a single Ethernet/802.3 network where two nodes can cause a collision. In the case of a single-segmented Ethernet/802.3 LAN, the independent segment represents the collision domain; in a multisegmented Ethernet/802.3 LAN, the collective segments comprise the collision domain.

Committed Burst *(B_c)* A term used in *frame relay* to denote the maximum amount of data a provider guarantees to deliver within a specified time period, *T*. $CIR = B_c/T$. Most providers use a one-second time interval to calculate the average amount of bandwidth utilization. Thus, CIR is usually equal to B_c. The difference between these two parameters is their units. CIR is measured in bps; B_c is measured in bits. See also *excessive burst.*

Committed Information Rate (CIR) The amount of throughput a *frame relay* provider guarantees to support under normal network loads. A CIR, which is assigned to a PVC when the network is initially configured, can range from 16 kbps to T3 (44.8 Mbps) and is the minimum guaranteed throughput of a PVC. If a PVC's assigned CIR is greater than or equal to the average amount of traffic transmitted across a PVC over a specified period of time (e.g., one second), then data transmissions are guaranteed. If the assigned CIR is less than this average, then data transmissions are not guaranteed.

Compression A process that codes repetitive patterns within a data set. Compressed files can be sent at a faster rate than uncompressed files.

Computer Emergency Response Team (CERT) A formal organization operated by the Software Engineering Institute at Carnegie Mellon University and dedicated to addressing computer and network security issues. CERT also serves as a clearinghouse for identifying and resolving security "holes" in network-related software or operating systems.

Computer Network A collection of computers and other devices that use a common network protocol to share resources with each other over a network medium.

Conductor That part of a wire which serves as the medium for the physical signal. It is composed of either copper wire, glass, or plastic fiber. In the case of copper, the wire can be stranded (composed of several thin wires) or solid (a single, “thick” strand). Furthermore, the thickness of a wire is given in terms of gauge, which represents the conductor’s diameter. The lower the gauge, the thicker the wire. Most often, wire gauges are expressed in terms of AWG—American Wire Gauge—which is a classification system for copper wire based on a wire’s cross-section diameter.

Congestion A term used to describe a situation when a network is consumed with excessive network traffic (i.e., lots of packets) resulting in performance degradation. Congestion occurs when routers are too slow, causing queues to lengthen, or when routers are too fast, causing queues to build up whenever input traffic is greater than the capacity of output lines. The ultimate level of congestion is known as deadlock, which occurs when one router cannot proceed until a second router does something, and the second router cannot

proceed because it is waiting for the first router to do something. Congestion control is provided by layer 3 of the *OSI* model.

Connectionless Service A type of service in which messages are partitioned into *packets* and routed through the network. Each packet is independent of the other packets that carry parts of the message, and each packet carries a destination address. Unlike connection-oriented service, no physical link is established between sending and receiving nodes prior to data transmission.

Connection-oriented Service A type of service in which prior to the transfer of data a physical (and virtual) link is established between the sending and receiving nodes. This link remains in effect for the duration of the session. After the session is completed, the link is removed. Characteristics of a connection-oriented service include: wasted bandwidth (link must remain established even during idle periods of a transmission); a high potential for a hung network (there is always a possibility that a link will not be terminated); and guaranteed sequential arrival of packets at the destination node.

Connector A layer 1 device that attaches network components together.

Consortia Standards Network standards that are designed and agreed upon by a group of vendors who have formed a consortium for the express purpose of achieving a common goal. These vendors pledge their support for the standards being developed by the consortium and also develop and market products based on these mutually agreed upon set of standards.

Contention A phenomenon in which more than one node competes to access a shared medium simultaneously.

Contention Protocol A network protocol that specifies procedures nodes are to follow when competing for access to the same communications channel at the same time. Also called *random access protocol.*

CRC An acronym for *cyclic redundancy check*.

CRC Checksum The result of a polynomial division that uses a predetermined *generator polynomial* as the divisor.

CRC Error An invalid CRC checksum.

Crosstalk Electrical interference (i.e., *noise*) that occurs when energy radiated from one wire-pair of a twisted pair wire "spills over" into another pair. In one type of crosstalk, called *near-end crosstalk* (abbreviated NEXT), a signal on the transmit pair is so strong that it radiates to the receive pair. A direct consequence of this spilled-over radiation is that the receiving device cannot decipher the real signal.

Cryptology The practice or art of encoding messages.

CSMA An acronym for *carrier sense multiple access*, which serves as the basis for various *random access protocols*. CSMA-based protocols include *one-persistent CSMA, nonpersistent CSMA, CSMA with collision Detection CSMA/CD),* and *CSMA with collision avoidance (CSMA/CD)*.

CSMA with Collision Avoidance (CSMA/CA) A variant of *CSMA/CD* except that it specifies a implementation scheme for *collision avoidance* instead of *collision detection*.

CSMA with Collision Detection (CSMA/CD) A variant of either *1-persistent* or *nonpersistent CSMA* that specifies what a node is to do upon detecting a collision. One-persistent CSMA/CD is the MAC sublayer protocol used in Ethernet/802.3 LANs.

CSU An acronym for *channel service unit*, which is a device used for terminating T*x* circuits. A CSU regenerates the signal, monitors the line for electrical anomalies, provides proper electrical termination, performs framing, and provides remote loopback testing for diagnosing line problems. Usually combined with a *DSU* to form a single unit called a *CSU/DSU* or DSU/CSU.

CSU/DSU An acronym for *channel service unit/data* (or d*igital*) *service unit*, which is a device that combines the functions of a *CSU* and a *DSU*. A CSU/DSU works exclusively with digital signals; it provides an interface between a digital computing device and a digital transmission medium.

Cut-Through A term used to describe a network switch architecture. Cut-through switches begin forwarding frames from one switch port to another as soon as the frame's destination address is read.

Cyclic Redundancy Check (CRC) An *error detection* method that constructs a polynomial whose terms' coefficients are the values of each of the bits of a data

frame. This polynomial is divided by a predetermined *generator polynomial*. The remainder of this division, called the CRC *checksum*, is then assigned to a frame's checksum field. The most common CRC used in most LAN protocols is CRC-32, a 32-bit checksum.

D Channel A 16 kbps or 64 kbps ISDN circuit that is used to carry signal and control information for circuit-switched user data. The D channel transmits call initiation (call- setup) and termination (call tear-down) information between an ISDN device and the telco's central office for each *B channel*. The D channel also can be used to transmit packet-switched user data (provided that no signal or control information is needed), data from security alarm signals of remote sensing devices that detect fire or intruders, and low speed information acquired from telemetry services such as meter reading. The "D" stands for "delta."

Datagram A grouping of bits organized as a logical unit of data at the network layer. *IP* datagrams serve as the Internet's primary unit of information. In the OSI model, a datagram is generically referred to as a *packet*.

Data Link Layer The second layer (layer 2) of the OSI Reference Model. The data link layer regulates and formats transmission of information from software on a node to the network cabling facilities. This layer is partitioned into two sublayers: The *logical link control sublayer* (LLC), which provides framing, flow control, and error control; and the *media access control sublayer* (MAC), which media access management protocols for accessing a shared medium.

Data Rate A measure of the amount of data that can be transferred over a communications medium in a given period. Data rate is measured in bits per second (bps) and can vary considerably from one type of channel to another.

DB Connector Layer 1 device that serves as an interface between a computer and a peripheral device such as a printer or external modem; "DB" stands for "data bus.")

DCE An acronym for *data communications equipment*. Generally used as a synonymous term for *modem.* A DCE device is placed between *DTEs* and is responsible for establishing, maintaining, and terminating the link connecting the two DTEs.

DCE-to-DCE Rate The speed at which two modems "talk" to each other. This rate is fixed and is a function of a modem's speed. Typical rates are 14,400 bps (V.32), 28,800 bps (V.34), and 57,600 bps (V.90).

Decentralized System Computer systems that are independent of each other, and maintain separate databases germane to specific activities.

Decryption The process of taking an encrypted (coded) message and translating it into its original, meaningful form.

***De Facto* Standards** Network standards, placed in the public domain, that have been met with widespread industry acceptance instead of formal approval from a standards organizations ("De facto" is Latin for "from the fact.")

***De Jure* Standards** Network standards approved by a formal, accredited standards organization such as ANSI or ITU. ("De jure" is Latin for "by right, according to law.")

Demand Access Multiplexing (DAM) A multiplexing technique in which a pool of frequencies is managed by a "traffic cop." Pairs of communications frequencies are assigned to a requesting station—one pair for transmission, a second pair for reception ("demand"). These two pairs of frequencies are connected to another set of frequencies ("access"). When one or both stations are finished communicating, the allocated frequencies are de- allocated and returned to the frequency pool, where they are made available for other incoming requests ("multiplexing").

Demand Priority A MAC sublayer protocol used in *100VG-AnyLAN* networks. Demand priority specifies the manner in which repeater hubs poll their ports to identify which nodes have data to transmit and the order of these transmissions.

DES An acronym for *data encryption standard*, which is a specific coding technique developed by the National Institute of Standards and Technology (formerly the National Bureau of Standards) and IBM for protecting sensitive data during transmission.

Desktop Another name for a networked device. See *workstation*.

Device Any entity that is connected to a network. Examples include terminals, printers, computers, or special network-related hardware units such as communication servers, repeaters, bridges, switches, and routers. Local or sending devices originate communications; remote or receiving devices are the recipient of such communications.

Differential Manchester Encoding A data transmission encoding scheme similar to Manchester—each bit-period is partitioned into two intervals and a transition between "high" and "low" occurs during each bit-period. In differential Manchester coding, though, the interpretation of these low-to-high and high-to-low transitions is a function of the previous bit-period. The presence of a transition at the beginning of a bit period is coded 0, and the absence of a transition at the beginning of a bit period is coded 1. Differential Manchester encoding is used for clocking purposes only.

Diffused IR A "broadcast" infrared transmission method in which a transmitter "floods" a specific area with a strong infrared signal that is spread over a wide angle. The IR signal is transmitted by reflecting off of ceilings, walls, and other surfaces.

Digital Refers to any device or signal that varies discreetly in strength or quantity between two values, usually zero and one. Zero implies "off"; one implies "on." Digital signals are represented as binary digits called "bits," and are *discrete*.

Digital Certificate An electronic passport that consists of a numerical pattern, value, or key and used for personal identification. Creating a digital certificate involves a user identifying a specific personal trait to a trusted third party, which issues the certificate.

Digital Communication Refers to any type of communication in which data are represented in the form of binary digits.

Digital Signature A security authorization method in which a user "signs" a document so that the document's authenticity can be confirmed by checking the signature. A digital signature proves a message was not modified.

Digital Subscriber Loop The formal term used to denote the *local loop*, which is the circuit between a *customer's premise equipment* (CPE) and the telco's equipment.

DIN Connector Similar to a DB connector, but is circular instead of rectangular and typically used to connect a keyboard to a computer; "DIN" stands for "Deutsche Industrie Norm," a German industrial standard.

Directed IR A "point-to-point" infrared transmission method that requires an unobstructed line-of-sight connection between transmitter and receiver. It is basically a "point and beam" medium.

Discard Eligibility The name of a field in a *frame relay* frame, which, if set to 1 by an end node, denotes that the frame can be discarded in the presence of congestion Discarded frames will then be retransmitted at a later time when congestion has subsided.

Distance-vector Algorithm A routing algorithm that determines the distance between source and destination nodes by calculating the number of router hops a packet traverses en route from the source network to the destination network. An example of a distance- vector algorithm is the Bellman-Ford algorithm.

Distributed System Computers that are linked together to provide, in a transparent manner, the required computing resources and information processing needs of an entire organization. Distributed systems bear the greatest resemblance to computer networks.

DLCI An acronym for *data link connection identifier*, which is a term used in *frame relay* to denote virtual circuit addresses assigned to *PVC*s or *SVC*s.

Domain Name A logical name assigned to an *IP address* and used as another type of addressing construct for identifying Internet nodes. The translation between logical name and IP address is called name resolution, which is provided by a *domain name service*.

Domain Name Service (DNS) An Internet translation service that resolves *domain names* to *IP addresses* and vice versa. DNS is provided by DNS servers.

DQDB An acronym for *distributed queue dual bus*, which is a data link layer protocol (IEEE 802.6) that specifies the medium access method for MANs. Used in *SMDS*.

DS-0 A single, digital voice channel rated at 64 kbps. The notation *DS-0* stands for *digital signal at level 0*, which refers to a voice channel multiplexed into a digital signal.

DS-1 A digital signal that carries 24 *DS-0* channels plus one 8 kbps channel reserved for framing for an aggregate bandwidth of 1.544 Mbps. A *T1* circuit carries a DS-1 signal.

DS-2 A digital signal that carries 4 *DS-1* channels for an aggregate bandwidth of 6.312 Mbps. A *T2* circuit carries a DS-2 signal.

DS-3 A digital signal that carries 28 *DS-1* channels for an aggregate bandwidth of 44.736 Mbps. A *T3* circuit carries a DS-3 signal.

DS-4 A digital signal that carries 168 *DS-1* channels for an aggregate bandwidth of 274.176 Mbps. A *T4* circuit carries a DS-4 signal.

DSL An acronym for *digital subscriber line*, which is a technology that enables data, voice, and video to be mixed and carried over standard analog, (copper) telephone lines. This is accomplished by using the unused frequencies that are available on a telephone line. Thus, DSL can deliver data services without interfering with voice transmissions.

There are at least nine DSL variants: *ADSL, ADSL lite, HDSL, HDSL 2,IDSL, RADSL, SDSL, UDSL,* and *VDSL*.

DSLAM An acronym for *DSL access multiplexer*, which is a device that aggregates DSL signals so they can be transferred directly into a data switch for transmission across the telco's data network backbone.

DSSS An acronym for *direct sequence spread spectrum*, which is a physical layer technology used in wireless LANs (IEEE 802.11). DSSS operates by spreading a signal over a wide range of the 2.4 GHz band.

DSU An acronym for *data* (or *digital*) *service unit*, which is a device used for terminating a T*x* circuit. A DSU provides the interface (usually V.35, a type of serial interface) for connecting a remote bridge, router, or switch to a T*x* circuit. The DSU also provides flow control between the network and the CSU. Usually combined with a CSU to form a single unit called a *CSU/DSU* or DSU/CSU.

DTE An acronym for *data terminal equipment*. Computers (PCs, workstations) are data terminal equipment. DTEs are the end points of a link and communicate through their serial ports or expansion buses. See also *data communications equipment (DCE)*.

DTE-to-DCE Rate The speed at which a computer “talks” to its modem. Typical rates include a 4:1 compression ratio between DTE and DCE speeds. Thus, for a V.34 modem (28,800 bps), the DTE-DCE rate is 115,200 bps. This rate is user configurable.

Dual-attachment Station (DAS) An FDDI node that is connected to two full, dual-fiber rings and have the ability to reconfigure the network to form a valid network from components of the two rings in case of a failure. A DAS is also called *Class A* node.

E.164 An ITU-T standard network addressing format that resemble telephone numbers. E.164 addresses are 15 decimal digits long and include a country code, area or city code, and a local number. Country codes are two or three digits long and consist of a zone code followed by a one- or two-digit national identifier. Area or city codes are up to four digits long. If an address contains less than 15 digits, then it is padded with hexadecimal Fs. Australia does not use city codes, and the United States and Canada use the zone code 1 followed by a three-digit area code and a seven digit local number in lieu of county codes.

E-1 Describes the multiplexing of 30 separate 64 kbps voice channels, plus one 64 kbps control channel, into a single, wideband digital signal rated at 2.048 Mbps. E-1 is the basic telecommunications service used in Europe.

E-2 A multiplexed circuit that combines four *E-1* circuits and has an aggregate bandwidth of 8.448 Mbps.

E-3 A multiplexed circuit that combines 16 *E-1* circuits and has an aggregate bandwidth of 34.368 Mbps.

E-4 A multiplexed circuit that combines 64 *E-1* circuits and has an aggregate bandwidth of 139.264 Mbps.

E-5 A multiplexed circuit that combines 256 *E-1* circuits and has an aggregate bandwidth of 565.148 Mbps.

E-commerce Short for *electronic commerce*, which involves using the Internet for credit card purchases of items such as automobiles, airline tickets, computer hardware and software, and books.

EGP An acronym for *exterior gateway protocol*, which refers to any Internet interdomain routing protocol used to exchange routing information with other autonomous systems. Also refers to Exterior Gateway Protocol, which is a specific EGP defined in RFC 904. Another EGP is the Border Gateway Protocol (BGP), defined in RFC 1105 and RFC 1771. Both EGP and BGP are part of the TCP/IP protocol suite. Of the two, however, BGP has evolved into a robust Internet routing protocol and the term “Border Gateway Protocol” is used in favor of the term “Exterior Gateway Protocol.”

EIGRP An acronym for *enhanced IGRP*, which is routing protocol designed by Cisco that combines the best features of distance-vector and link-state routing protocols.

Encapsulation A process in which a *packet* or *frame* is enclosed or "wrapped" in a specific protocol header. For example, *routers* typically perform protocol encapsulation in which packets from one network protocol are wrapped into the header of another network protocol so the packet can be transmitted to a different network. Also called *tunneling*.

Encryption The process of coding a message so that it is incomprehensible to unauthorized users. When retrieved by authorized users, encrypted messages are then reconverted (i.e., decoded) into meaningful text. Encrypted output is called *ciphertext*.

Error Control The process of guaranteeing reliable delivery of data. Error control can be provided through *error detection* or *error correction*.

Error Correction The process in which a destination node, upon detecting a data transmission error, has sufficient information to correct the error autonomously. Error correction implies *error detection*.

Error Detection The process in which a destination node detects a data transmission error and requests a retransmission from the sending node. Error detection is also called *error correction through retransmission*.

Ethernet A local area network protocol developed jointly by Xerox, Intel, and Digital Equipment Corporation (DEC) at the Xerox Palo Alto Research Center (PARC) in the mid-1970s. The name "Ethernet" was derived from the old electromagnetic theoretical substance called *luminiferous ether*, which was formerly believed to be the invisible universal element that bound together the entire universe and all its associated parts. Thus, an "ether" net is a network that connects all components attached to the "net."

Excessive Burst (B_e) A term used in *frame relay* to denote the maximum amount of uncommitted data a provider will attempt to deliver within a specified time period. A provider will guarantee a *committed burst* of B_c bits and will attempt to deliver (but not guarantee) a maximum of $B_c + B_e$ bits.

Exchange Access SMDS (XA-SMDS) A special *SMDS* service through which *LECs* offered SMDS to *IECs* for delivery across *LATAs*.

Extranet A popular networking term that describes an interconnection from an internal intranet to a customer or noncompany network that is not the Internet connection.

4B/5B A data encoding method, which stands for *four bits in five baud*, or *four-bit to five-bit*, used in FDDI networks.

5-4-3 Repeater Rule A general rule of thumb to follow when configuring an Ethernet/ 802.3 LAN to ensure that it follows IEEE specifications. The 5-4-3 rule requires: no more than 5 segments of up to 500 m each; no more than 4 repeaters; and no more than 3 segments can have end nodes connected to them. This rule is also known as the 4-repeater

rule, or the 5-4-3-2-1 rule. In the latter, the "2" implies that two of the five segments are used as interrepeater links, and the "1" implies that a configuration using the maximum parameters permitted results into one collision domain.

Fast Ethernet 100 Mbps Ethernet (IEEE 802.3u). Three different media specifications are defined: 100BASE-TX, 100BASE-T4, and 100BASE-FX.

FDDI Fiber Distributed Data Interface. FDDI networks are described by ANSI standard X3T9.5 and created in 1986 for interconnecting computer systems and network devices typically via a fiber ring topology at 100 Mbps.

FECN An acronym for *forward explicit congestion notification*, which is a one-bit field in a *frame relay* frame that is set to 1 by a frame relay switch to denote that a frame transmitted toward the receiving node experienced congestion.

FDDI-II A now defunct second generation FDDI technology that was intended to handle traditional FDDI network traffic as well as synchronous, circuit-switched PCM data for voice or ISDN systems.

FHSS An acronym for *frequency hopping spread spectrum*, which is a physical layer technology used in wireless LANs (IEEE 802.11). FHSS operates by transmitting short bursts of data on different frequencies. One burst is transmitted on one frequency, a second burst is transmitted on a second and different frequency, and so forth.

Fiber-optic Cable A type of cable that carries data signals in the form of modulated light beams. The cable's conductor can be either glass or plastic. Fiber-optic cable is immune to electromagnetic interference (EMI) and other types of externally induced noise, including lightning, it is unaffected by most physical factors such as vibration, its size is smaller and its weight lighter than copper, it has much lower attenuation per unit of length than copper, and it can support very high bandwidth. Two general types are available: *single-mode fiber* and *multimode fiber*.

Fibre Channel A family of ANSI standards that defines a specific communications interface for high-speed data transfers between different hardware systems. Applications include the medical profession, where large images (e.g., 100 MB+ X-rays) are transferred from a scanner to a computer to a screen, and the electronic publishing industry, where large files are transferred from an designer/creator's machine to a publisher's computer. It has also become the "backbone" of high-speed data storage systems.

Firewall A device or product that allows systems or network manager to restrict access to components on a network. Five generally accepted types of firewalls are

used on Internet connections are *frame-filtering, packet-filtering, circuit gateways, stateful* and *application gateways,* and *proxy servers*.

FIX An acronym for *federal Internet exchange*, which is an Internet interconnect site similar to a *NAP*.

Flow Control A process that controls the rate at which data messages are exchanged between two nodes. Flow control provides a mechanism to ensure that a sending node does not overwhelm a receiving node during data transmission.

Fractional T1 *T1* service that is sold in 64 kbps increments.

FRAD An acronym for *frame relay access device*, which is a term used to denote any *frame relay* end node.

Fragmenting A process in which a *packet* is broken into smaller units to accommodate the maximum transmission unit a physical network is capable of supporting. Fragmented packets are sent to the destination separately and then reassembled at the destination node before it is passed to the higher levels. In *IP*, reassembly of a *datagram* occurs at the destination node and not at any of the intermediary nodes the packet traverses.

Frame A specially formatted sequence of bits that incorporates both data and control information.

Frame-filtering Firewall A firewall device or product that filters (permits or denies access) at the data link layer by examining frames for both layout and content.

Framing A data link layer process that partitions a bit stream into discrete units or blocks of data called *frames*.

Frame Relay A public WAN packet-switching protocol that provides LAN-to-LAN connectivity. Its name implies what it does, namely, relays frames across a network between two sites. Frame relay was originally part of the *ISDN* standard.

Frequency Division Multiplexing (FDM) A multiplexing technique that partitions the available transmission frequency range into narrower bands (subfrequencies), each of which is a separate channel. FDM-based transmissions are parallel in nature.

Full-duplex Transmission A data transmission method that involves the simultaneous sending and receiving of data in both directions.

GAN An acronym for *global area network*, which refers to a collection of WANs that span the globe.

Gateway A software application that converts between different application protocols. The host on which this software resides is called a *gateway machine*. Historically, this term also refers to a *router* in the *IP* community.

Gigabit Ethernet 1000 Mbps Ethernet (IEEE 802.3z).

Geostationary Earth Orbit (GEO) Satellite A satellite placed into orbit at an altitude of 22,000 miles (36,000 kilometers) above the equator. GEO satellites traverse their orbits at approximately the same rate as the Earth rotates. Thus, the satellite appears stationary with respect to the Earth's rotation. Also call *Geosynchronous Earth Orbit*. Only eight GEO satellites are needed to provide global communications coverage.

GOSIP An acronym for *Government OSI Profile*, which mandated all government organizations purchase OSI-compliant networking products beginning in 1992. In 1995, however, GOSIP was modified to include TCP/IP as an acceptable protocol suite for GOSIP compliance.

Graded-index Multimode Fiber A type of multimode fiber in which variations in the density of the core medium change its *index of refraction* such that light is refracted (i.e., bends) toward the center of the fiber.

H Channel An *ISDN* channel used for transmitting user data (not signal or control information) at higher transmission rates than a *B channel* provides. Four H channels are defined: *H0* (six B channels; 384 kbps); *H10* (United States-specific; aggregates 23 B channels; 1.472 Mbps); *H11*(equivalent of North American DS-1; 24 B channels; 1.536 Mbps); and *H12* (European-specific; comprises 30 B channels; 1.920 Mbps).

Half-duplex Transmission A data transmission method in which may travel in either direction—from sender to receiver or receiver to sender—but only one unit can send at any one time. While one node is in send mode, the other is in receive mode.

Harmonic Motion The basic model for vibratory or oscillatory motion. Examples include mechanical oscillators such as mass-spring systems and pendulums; periodic motion found in the earth sciences such as water waves, tides, and climatic cycles; and electromagnetic waves such as alternating electric currents, sound waves, light waves, radio waves, and television waves.

HDSL An acronym for *high bit-rate digital subscriber line*, which is a *DSL* variant that provides symmetrical service at T1 rates over 2 pairs of UTP, and E1 rates over 3 pairs of UTP. Telephone service not supported. Applications include connecting PBXs, serving as an alternative to T1/E1; suitable for campus networks and ISPs.

HDSL 2 A modified *HDSL* designed and packaged for corporate clients.

Hertz A measure of frequency in cycles per second. A frequency rate of one cycle per second is defined as one hertz (abbreviated Hz). Named in honor of Heinrich Rudolf Hertz (1857-1894), a German physicist who in the late 1880s was the first to produce radio waves artificially.

HFC An acronym for *hybrid fiber cable*, which describes a cable TV cable plant that has fiber-optic cable between the head end and neighborhood distribution sites,

but coaxial cable between the neighborhood distribution and residential homes and businesses.

Hold-down A strategy used by RIP that requires routers to not update their routing tables with any new information they receive for a prescribed period of time, called the hold-down time. Designed to prevent routing loops. Hold-down is not standardized.

Hop A term used to describe the passage of a *packet* through an intermediate gateway (*router*) en route to another network. For example, if a packet transverses through two routers in reaching its final destination, then we say the destination is two hops away.

Host A networked computer system (see *workstation*). Also used to describe a computer system that provides service to users (see *server*).

Hub Generically, any device that connects two or more network segments or supports several different media. Examples include repeaters, switches, and concentrators.

Hybrid Switching A data transmission method that combines the principles of circuit and packet-switching. This technique first partitions a message into packets (packet- switching) and transmits each packet via a dedicated circuit (circuit-switching). As soon as a packet is ready for transmission, a circuit meeting appropriate bandwidth requirements is established between the sending and receiving nodes. When the packet reaches its destination, the circuit is broken down so that it can be used again.

IBM Cable System (ICS) A copper wire classification system established by IBM that specifies nine cable “types” (1 through 9). Of the nine “types” defined, specifications are available for only seven; types 4 and 7 are not defined.

ICMP An acronym for *Internet control message protocol*, which uses an *IP datagram* to carry messages about the communications environment of the Internet.

IDSL An acronym for *ISDN-like digital subscriber line*, which is a *DSL* variant that provides symmetrical service at a maximum of 144 kbps each way. Uses ISDN hardware.

IEC See *IXC*.

IEEE An acronym for *Institute of Electrical and Electronics Engineers*, which is a professional society of engineers, scientists, and students. One of its many activities is to act as a coordinating body for computing and communication standards.

IEEE 802 The primary *IEEE* standard for the 802.*x* series for *LANs* and *MANs*.

IEEE 802.1 *IEEE* standard that defines an architectural overview of *LANs*.

IEEE 802.2 *IEEE* standard that defines the Logical Link Control, which describes services for the transmission of data between two nodes.

IEEE 802.3 *IEEE* standard that defines the *Carrier Sense Multiple Access/Collision Detection (CSMA/CD)* access method commonly referred to as *Ethernet*. Supplements include **802.3c** (10 Mbps Ethernet); **802.3u** (100 Mbps Ethernet known as *Fast Ethernet*), and **802.3z** and **802.3ab** (1000 Mbps Ethernet known as *Gigabit Ethernet*).

IEEE 802.4 *IEEE* standard that defines the token bus network access method.

IEEE 802.5 *IEEE* standard that defines the logical ring *LAN* that uses a token-passing access method; known also as *Token Ring*.

IEEE 802.6 *IEEE* standard that defines metropolitan area networks (*MANs*).

IEEE 802.7 *IEEE* standard that defines broadband *LANs* (capable of delivering video, data, and voice traffic).

IEEE 802.9 *IEEE* standard that defines integrated digital and video networking—Integrated Services LANs (ISLANs).

IEEE 802.10 *IEEE* standard that defines standards for interoperable *LAN/MAN* security services.

IEEE 802.11 *IEEE* standard that defines standards for wireless media access control and physical layer specifications.

IEEE 802.12 *IEEE* standard that defines the “demand priority” access method for 100Mbps LANs; known also as 100 Base-VG or *100VG-AnyLAN*.

IEEE 802.13 (Defines nothing—IEEE was concerned about the superstitious overtones associated with “13.”)

IEEE 802.14 *IEEE* standard that defines a standard for Cable-TV based broadband communication.

IGP An acronym for *interior gateway protocol*, which is any intradomain Internet protocol used to exchange routing information within an *autonomous system*. Examples include *RIP, RIP-2, OSPF, IGRP,* and *Enhanced IGRP (EIGRP)*.

IGRP An acronym for *interior gateway routing protocol*, which was developed by Cisco to address some of the problems associated with routing in large, heterogeneous networks.

ILEC An acronym for *incumbent local exchange carrier*, which is the contemporary name given to the *RBOCs* relative to the United States Telecommunications Act of 1996.

Impedance A measure of the opposition to the flow of electric current in an alternating current circuit. Measured in ohms (abbreviated by the Greek symbol,

omega, ¾), impedance is a function of capacitance, resistance, and inductance. *Impedance mismatches*, caused by mixing cables of different types with different characteristic impedances, can result in signal distortion.

Impulse Noise Electrical *noise* that consists of intermittent, undesirable signals induced by external sources such as lightning, switching equipment, and heavy electrically operated machinery such as elevator motors and copying machines. Impulse noise increases or decreases a circuit's signal level, which causes the receiving equipment to misinterpret the signal.

Infrared (IR) A line-of-sight transmission method that uses electromagnetic radiation of wavelengths between radio waves and visible light, operating between 100 GHz and 100 THz (Terahertz). IR transmission can occur in one of two ways: *directed* and *diffused*.

Insulation Material surrounding the *conductor* of a wire. The insulation serves as a protective "barrier" to the conductor by preventing the signal from "escaping" and preventing electrical interference from "entering."

Intermodulation Noise Electrical *noise* that occurs when two frequencies interact to produce a phantom signal at a different frequency. Occurs in *frequency-division multiplexed* channels.

Internet When used as a noun and spelled with a lowercase *i*, "internet" is an abbreviation for *internetwork*, which refers to a collection of interconnected networks that functions as a single network. When used as a proper noun and spelled with a capital *I*, "Internet" refers to the world's largest internetwork, which consists of hundreds of thousands of interconnected networks worldwide and based on a specific set of network standards (TCP/IP).

Internet Architecture Board (IAB) An organization that is part of the *Internet Society* responsible for the overall planning and designing of the Internet. Responsibilities include setting Internet standards, managing the publication of RFC documents, and resolving technical issues. Assigned to the IAB are the *Internet Engineering Task Force* and the *Internet Research Task Force*. Formerly known as the Internet Activities Board.

Internet Assigned Numbers Authority (IANA) An organization that has authority over all number spaces used in the Internet including *IP addresses*. IANA control will soon be transferred to the *Internet Corporation for Assigned Names and Numbers (ICANN)*.

Internet Corporation for Assigned Names and Numbers (ICANN) A private, non- profit corporation with international representation expressly formed to assume the responsibilities currently being performed by IANA and other government organizations that provide domain name service.

Internet Engineering Task Force (IETF) An organization that is part of the *Internet Architecture Board* and primarily concerned with addressing short- or

medium-term Internet engineering issues. Relies on the Internet Engineering Steering Group (IESG) to prioritize and coordinate activities.

Internet Registry (IR) A formal hierarchical system used for assigning *IP addresses*. From top to bottom, this hierarchy consists of *IANA*, Regional Internet Registries (RIR), and Local Internet Registries (LIR), and works as follows: IANA allocates blocks of IP address space to RIRs; RIRs allocate blocks of IP address space to their LIRs; LIRs then assign addresses to either end users or ISPs.

Internet Research Task Force (IRTF) An organization that is part of the *Internet Architecture Board* and primarily concerned with addressing long-term research projects. Relies on the *Internet Research Steering Group (IRSG)* to prioritize and coordinate activities.

Internet Society (ISOC) An international organization comprised of volunteers who promote the Internet as a medium for global communication and collaboration. ISOC is considered the ultimate authoritative organization of the Internet.

Internet2 A collaborative project of the University Corporation for Advanced Internet Development (UCAID), which comprises over 100 U.S. universities, government organizations, and private sector firms. Internet2's mission is to develop advanced Internet tech

nologies and applications that support the research endeavors of colleges and universities. Internet2 members use the *vBNS* to test and advance their research.

Interoperability The degree in which products (software and hardware) developed by different vendors are able to communicate successfully (i.e., *interoperate*) with each other over a network.

Intranet An internal network implementation of traditional Internet applications within a company or an institution.

Inverse Multiplexing The reverse of multiplexing. Instead of partitioning a single communication medium into several channels, an inverse multiplexer combines several "smaller" channels (i.e., low-speed circuits) into a single high-speed circuit. This technique is also sometimes generically called *line aggregation*.

IP An acronym for *Internet protocol*, a layer 3 connectionless protocol. IP receives data bits from the lower layer, assembles these bits into packets, called IP *datagrams*, and selects the "best" route based on some metric to route the packets between nodes. IP is the "IP" of *TCP/IP*.

IP Address A network address assigned to a node's network interface and used to uniquely identify (locate) the node within the Internet. Two versions are currently implemented: *IPv4* and *IPv6*.

IPSec An acronym for *IP security,* which is a suite of network security protocols that operates at layer 3 and provides address authentication, data encryption, and automated key exchanges between sender and receiver nodes.

IPv4 An acronym for *Internet protocol version 4*.

IPv4 Address An *IP address* based on *IPv4*. These addresses consist of 32 bits (0 through 31) partitioned into four groups of eight bits each (called *octets*), and organized into five classes (A through E) based on the values of bits 0 through 3.

IPv6 An acronym for *Internet protocol version 6*, which is an evolutionary replacement to IPv4. IPv6 maintains most IPv4 functions, relegates certain functions that either were not working or were rarely used in IPv4 as optional, and adds new functionality that is missing from IPv4. Sometimes called IPng (for next generation).

IPv6 Address An *IP address* based on *IPv6*. An IPv6 address consists of 128 bits and is 4 billion 4 billion times the size of the IPv4 address space (2^{96} vs. 2^{32}). Unlike IPv4 addresses, IPv6 addresses use a colon as their delimiter (instead of a "dot" notation), and they are written as eight 16-bit integers expressed in hexadecimal form.

ISDN An acronym for *integrated services digital network*, which is a carrier service that is offered by telephone companies (telcos) and designed to transmit voice and non-voice (e.g., computer data, fax, video) communications on the same network. Also known as, *I Still Don't Need it, Innovative Services users Don't Need, I Still Don't kNow,* and *It's Still Doing Nothing*, response to ISDN's long period of dormancy.

IS-IS An acronym for *intermediate system to intermediate system*, which is an intradomain routing protocol designed by *OSI* to run within an *AS* (called a "routing domain" in the OSI world). IS-IS uses a link-state routing algorithm to calculate least-cost paths, and is similar in operation *OSPF*. The formal title of this protocol is "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol."

ISO An acronym for *International Organization for Standardization*, which develops and promotes networking standards worldwide.

Isochronous A term used to describe the delivery of time sensitive data such as voice or video transmissions. Networks that are capable of delivering isochronous service (e.g., *ATM*) preallocate a specific amount of bandwidth over a regular intervals to ensure that the transmission is not interrupted.

IsoEthernet Short for *Isochronous Ethernet*, which is an IEEE standard—IEEE 802.9a, designed to support time-sensitive applications such as videoconferencing and telephony. IsoEthernet runs both conventional 10 Mbps Ethernet and ISDN B channels over the same network. The Ethernet channel is used for normal data networking needs; the ISDN B channels are used for time-sensitive applications.

ISP An acronym for *Internet Service Provider*, which is an organization that provides its customers with access to the Internet.

ITU An acronym for *International Telecommunications Union*, which is a global standards organization. ITU is the former *CCITT*.

IXC An acronym for *inter-exchange carrier*, (alternatively, IEC), which is any company that provides long distance telephone and telecommunications services. Examples include AT&T, Sprint, British Telecom (BT), and MCI Worldcom.

Jabber An *oversized* Ethernet/802.3 frame and an invalid CRC checksum.

Kerberos A client-server network security authentication system, developed at MIT, and based on DES encryption. It is an Internet standard that uses a three-pronged approach for authentication: a database that contains users' rights, an authentication server, and a ticket- granting server. Kerberos is named after *Cerberus*, the three-headed dog in Greek mythology that guarded the gates to Hades.

LAN An acronym for *local area network*, which is a network that generally interconnects computing resources within a moderately sized geographical area. This can include a room, several rooms within a building, or several buildings of a campus. A LAN's range is usually is no more than 10 km in radius).

LANE An acronym for *LAN emulation*, which is an ATM protocol that specifies a technology that enables ATM to emulate Ethernet/802.3 or token ring networks. In ATM's protocol hierarchy, LANE is above *AAL5* in the *ATM adaptation layer*. The LANE protocol defines a service interface for the network layer that functions identical to the one used by Ethernet/802.3 and token ring LANs. Data that cross this interface are encapsulated in the appropriate MAC sublayer format.

LAP-D An acronym for *link access protocol–D channel*, which is an ITU-T standard on which the *ISDN D* channel is based.

LAPM An acronym for link access procedure for modems, which uses *CRC* and *ARQ* for error control. CRC is used for error detection; ARQ prevents the modem from accepting any more data until the defective frame has been retransmitted successfully. *V.42*'s default is LAPM. Thus, if a connection is being initialized between two V.42 compliant modems, they will use LAPM for error control. If one of the modems is not V.42 compliant, then the modems will negotiate to use *MNP* 1–4.

LATA An acronym for *local access and transport area*, which is a specific geographical region in which a *local exchange carrier (LEC)* provides local telephone and telecommunications services in the United States. There are 195 LATAs. Services that cross LATA boundaries are provided by *inter-exchange carriers (IECs)*.

Latency The amount of delay a network device introduces when data frames pass through it. It is the amount of time a frame spends "inside" a network device. For example, switch latency is usually measured from the instant the first bit of a frame enters the device to the time this bit leaves the outbound (i.e., destination) port.

Layer 3 Switch A layer 2 switch that is capable of examining layer 3 header information, which is then used to filter network protocols or broadcasts. Also refers to a router that is capable of performing router table lookups and packet forwarding at hardware speeds via application specific integrated circuit (ASIC) chips.

Layer 4 Switch A router that is capable of examining upper layer (layers 4 through 7) information to make routing decisions. It is more appropriate to refer to layer 4 switches as either layer 2 or layer 3 application switches because application information from upper layers is being used for routing decisions.

Lightwave Wireless A line-of-sight laser-based connection facility that allows long-distance light-based wireless networking without the need to install cable.

Line-of-Sight A type of wireless transmission that requires the transmitter and receiver be able to "see" each other, that is, they must be in each other's "line-of-sight."

Line Set A term used by the National ISDN Users' Forum to describe the number of multiplexed *B* and *D* channels, and the type of *ISDN* service supported.

Link-state Algorithm A routing algorithm in which routers send each other information about the links they have established to other routers via a link state advertisement (LSA), which contains the names and various cost-metrics of a router's neighbors. LSAs are flooded throughout an entire router's domain. Thus, rather than storing actual paths (which is the case with *distance-vector algorithms*), link-state algorithms store the information needed to generate such paths. An example of a link-state algorithm is Dijkstra's shortest path algorithm, which iterates on length of path to determine a shortest route.

Lobe The name of a token ring node, as defined in the IBM world.

Lobe Length A term used to identify the cable length between token ring nodes.

Local Loop Refers to the circuit that connects the telephone central office or exchange (sometimes called point of presence) with a customer's location. In frame relay, this circuit is called the *port connection* or *access line*. Formally called *digital subscriber loop*.

Logical Link Control (LLC) Sublayer The top sublayer of the data link layer that provides framing, flow control, and error control Defined in IEEE 802.2.

Loop A network configuration in which nodes are connected via dedicated wiring instead of through a centralized hub (as is the case of a *star* design). Loops can be either *simple* (only one connection between any two nodes), *partial* (some nodes are interconnected by more than one link), and *complete* (every node has a connection to every other node). A loop is also referred to as a *meshed* design.

Low-Earth Orbit (LEO) Satellite A satellite placed in orbit at an altitude of 300 miles to 1,200 miles above the Earth. Depending on their orbit, a constellation of up to 48 LEO satellites are needed for global coverage.

L2F An acronym for *layer 2 forward* protocol, which provides tunneling between an ISP's dial-up server and the network.

L2TP An acronym for *layer 2 tunneling protocol*, which defines a method for tunneling PPP sessions across a network. It combines *PPTP* and *L2F*.

Manchester Encoding A data transmission encoding scheme that differs from standard digital transmission schemes. Instead of "high" equaling "1" and "low" equaling "0," a timing interval is used to measure high-to-low transitions. Furthermore, instead of a timed transmission period being "all high" or "all low" for either 1 or 0, a 1 is sent as a half-time- period low followed by a half-time-period high, and a 0 is sent as a half-time-period high followed by a half-time-period low. Consequently, the end of the last bit transmitted is easily determined immediately following the transmission of the last bit.

MAE An acronym for *metropolitan-area exchange*, which is an Internet interconnect site similar to a *NAP*. The difference between the two is a NAP is funded by the National Science Foundation and MAE is not. There are currently two MAE points, one each on the east and west coasts of the United States and known as MAE East and MAE West.

MAN An acronym for *metropolitan area network*, which interconnects computing resources that span a metropolitan area such as buildings located throughout a local county or city. MANs generally refer to networks that span a larger geographical area than LANs but a smaller geographical area than WANs.

MAU Another term for a transceiver; "MAU" stands for "Media Attachment Unit." Also, Multistation Access Unit, which is a token ring hub.

Media Access Control (MAC) Sublayer The bottom half of the data link layer that provides media access management protocols for accessing a shared medium. Example MAC sublayer protocols include IEEE 802.3 (Ethernet) and IEEE 802.5 (token ring).

Medium The physical environment used to connect networked devices.

Medium-Earth Orbit (MEO) Satellite A satellite placed in orbit at an altitude of 6,000 miles to 12,000 miles above the Earth. A constellation of 20 MEO satellites are needed for global coverage.

Media The plural of *medium*.

Media Converter A layer 1 device that enables different network media to be connected to one another.

Meshed Design A term used to describe interconnectivity among multiple nodes or sites. In a *fully-meshed* design, every node or site is connected with every other node or site. In a *partially-meshed* design, only some nodes or sites are interconnected.

Metric A generic term used in *routing* to represent different quantities such as distance, number of router *hops*, and bandwidth.

Metro-Area Satellites A proposed satellite that consists of a specially equipped jets that fly 50,000 feet above cities.

Micron One micrometer (one millionth of a meter) and abbreviated by the symbol *µm*. Used in specifying the size of fiber-optic cable.

Microwave An RF transmission method that uses high frequency waves and operates at a higher frequency in the electromagnetic spectrum (usually above 900 MHz). Microwave transmissions are considered a *line-of-sight* medium.

MNP An acronym for *Microcom Networking Protocol*, which defines various levels of error correction and compression for modems.

MNP 1-4 The first four *MNP* levels used for hardware error control. All four levels are incorporated into *V.42*.

MNP 5 The fifth level of *MNP* that incorporates the *MNP 1-4*. Also uses a data compression algorithm that compresses data by a factor of 2 to 1.

MNP 6 The sixth level of *MNP* that supports V.22 *bis* and V.29.

MNP 7 The seventh level of *MNP* that improves *MNP 5*'s data compression algorithm to a 3 to 1 compression factor.

MNP 8 The eighth level of *MNP* that extends *MNP 7*; enables half-duplex devices to operate in full-duplex mode.

MNP 9 The ninth level of *MNP* that is used in a variety of circuits.

MNP 10 The tenth level of *MNP* that is used in cellular modems and in those situations where line quality is poor.

Modem An acronym *modulator/demodulator*. A modem transforms (modulates) a computer's digital signal into analog form at the sending side so the signal can be carried across a standard telephone line. On the receiving side, a modem demodulates the signal— it reconverts the transmitted analog signal from the phone line to digital form before it is passed to the computer.

Multicast A data transmission that is destined to a group of recipients.

Multidrop Design A network configuration in which each system node is connected to a common cable plant and assigned a specific number that is used to communicate with the system and also to establish priority of when a system will be communicated with from a master control system. Primarily used in factories.

Multilink PPP (MP) An IP protocol that combines multiple physical links (i.e., telephone lines) into a single, high capacity channel. Unlike *BONDING*, which is implemented in hardware, MP is achieved via software. MP is also applicable to analog dialup connections.

Multimode Fiber A type of fiber-optic cable with a core diameter ranging from 50 µm to 100 µm. In multimode fiber, different rays of light bounce along the fiber at different angles as they travel through the core. This results in some degree of signal distortion at the receiving end. Multimode fiber can be of two types: *graded-index* or *step-index*.

Multiplexer A device that does multiplexing. Also called a *mux* for short.

Multiplexing A technique used to place multiple signals on a single communications channel. Multiplexing partitions a channel into many separate channels, each capable of transmitting its own independent signal, thereby enabling many different transmissions over a single medium.

NADH See *North American Digital Hierarchy*.

NAP An acronym for *network access point*, which is an Internet traffic exchange point that provides centralized Internet access to Internet service providers. A NAP serves as a critical, regional "switching station" where all different network backbone providers meet and exchange traffic on each other's backbone.

NSAP An acronym for *network service access point*, which is an *OSI* addressing mechanism used by private *ATM* networks. NSAPs are 20-byte addresses and include a 13-byte prefix that can be used to identify a specific location including a country, region, or end system.

National Information Infrastructure (NII) A Federal policy initiative to facilitate and accelerate the development and utilization of the nation's information infrastructure. The

perception of the NII is one of a "seamless web" of telecommunications networks consisting of computers, specialized databases, radios, telephones, televisions, and satellites. The NII is expected to provide consumers with convenient and instantaneous access to nearly any kind of information ranging from research results, to medical and educational material, to entertainment.

netstat A UNIX program that generates a local host's routing table. Similar output can be generated on a Windows NT system using the command *route print*.

Network Architecture A formal, logical structure that defines how network devices and software interact and function; defines communication protocols, message formats, and standards required for interoperability.

Network Computer (NC) An inexpensive ($500 or less) network access device with functionality that allows some applications to be run, but not as complete as what would typically be found on a PC or a workstation of some sort. NCs are stripped-down systems that use the network to access their applications dynamically.

Network Diameter The overall length between a network's two most remote nodes.

Network Ethics Refers to specific standards of moral conduct by network users for the responsible use of network devices and resources.

Network Interface Card A layer 2 device that performs standard data link layer functions, including organizing data into frames, transferring frames between the ends of a communication channel, and managing the link by providing error control, initialization, control termination, and flow control. A NIC" is also known as a LAN adapter, network adapter, network card, and network board. When used in Ethernet/802.e networks, a NIC is called an Ethernet card or adapter.

Network Operating System (NOS) Software that is installed on a system to make it network-capable. Examples include IBM's LAN Server, Banyan's VINES, and Novell's NetWare (also known as IntranetWare). A NOS is independent of a computer's native operating system—it is loaded "on top" of the computer's operating system and provides the computer with networking capability based on a particular protocol. If an operating system provides built-in network support (e.g., Microsoft's Windows NT and Sun's Solaris), then the OS is called a *networkable* operating system.

Network Protocol Suite A set of related and interoperating network protocols. For example, the TCP/IP protocol suite consists of protocols for e-mail, web service, file transfers, and routing.

Network Security Refers to the proper safeguarding of everything associated with a network, including data, media, and equipment. It involves administrative functions, such as *threat assessment*, and technical tools and facilities such as cryptographic products, and network access control products such as *firewalls*. It also involves making certain that network resources are used in accordance with a prescribed policy and only by people who are authorized to use these resources.

Network Standards A formal set of rules, developed by and agreed upon by various organizations, defining hardware interfaces, communication protocols, and network architectures. Several standards exist, including *de jure, de facto, proprietary,* and *consortia*.

Network Termination Unit (NTU) A device that terminates *E-1* circuits. An NTU provides broadly similar *CSU/DSU* functionality.

Network Topology The basic design of a computer network that details how key network components such as nodes and links are interconnected.

Next Generation Internet (NGI) An initiative to forge collaborative partnerships between the private and public sectors. Presumably, the vBNS will serve as the medium for NGI. Funding ($100 million for three years) has not been approved as of this writing.

Node Another name for a device. Usually used to identify computers that are network hosts, workstations, or servers.

Noise Any undesirable, extraneous signal in a transmission medium. There are generally two forms of noise—*ambient* and *impulse*. Noise degrades the quality and performance of a communications channel and is one of the most common causes of transmission errors in computer networks.

North American Digital Hierarchy (NADH) Describes a multiplexed *T1* structure used in North America that combines multiple T1 lines into higher rated T*x* circuits. For example, a *T2* circuit consists of four multiplexed T1 circuits and has an aggregate bandwidth of 6.312 Mbps; a *T3* link consists of 28 multiplexed T1 circuits with an aggregate bandwidth of 44.736 Mbps; and a *T4* channel consists of 168 multiplexed T1 circuits and is rated at 274.176 Mbps.

nslookup A UNIX and Microsoft NT program used to acquire the *IP address* of a *domain name*. This program can also be used for IP address resolution, which translates a numerical IP address to its corresponding domain name.

1-persistent CSMA A CSMA-based protocol in which a node continually waits a random period of time whenever it detects a busy channel. Once it senses an idle channel, it may then transmit data.

1-persistent CSMA A CSMA-based protocol in which a node continuously monitors a shared channel until it is idle and then seizes the channel and begins transmitting data. The “one” in 1-persistent represents the probability that a single waiting node will be able to transmit data once it detects an idle channel ($p = 1$).

OC An acronym for *optical carrier*, which is a fiber-optic digital transmission hierarchy used for *SONET*. OC rates range from OC-1, which is the equivalent of 28 DS-1 channels (51.84 Mbps) to OC-192, which is the equivalent of 5,376 DS-1 channels (9.953 Gbps). OC rates are the optical equivalent of *STS* rates.

OSI An acronym for *open systems interconnection*.

OSI Reference Model A network architecture for developing network protocol standards. The OSI Model formally defines and codifies the concept of *layered* network architecture. It uses well-defined operationally descriptive layers that describe what happens at each stage in the processing of data for transmission. The OSI Model consists of the following seven layers, which are numbered in descending order: Application (7), Presentation (6), Session (5), Transport (4), Network (3), Data Link (2), and Physical (1).

OSPF An acronym for *open shortest path first*, which is an *interior gateway protocol* based on a *link-state algorithm*. Designed for large, heterogeneous IP networks.

Oversized Frame An Ethernet/802.3 frame with more than 1,518 bytes but a valid CRC checksum.

Oversubscription A term used in *frame relay* to denote when the capacity of a frame relay connection into the frame relay network is less than the total bandwidth

guaranteed by the provider. More specifically, the *port speed* is less than the aggregate *CIR*.

Packet The smallest unit of information that is transferred across a packet-switched network. In TCP/IP a packet is called a *datagram*.

Packet-filter Firewall A router or a dedicated device that filters network access at the network layer by examining packet addresses (source and destination), or specific network transport protocol type.

Packet-switched Network A network design that enables nodes to share a communications channel via a *virtual circuit*. Messages are partitioned into smaller messages called *packets*, which may contain only a few hundred bytes of data, accompanied by addressing information. Packets are sent to the destination node one at a time, at any time, and not necessarily in a specific order. The network hardware delivers the packets through the virtual circuit to the specified destination node, which is responsible for reassembling them in the correct order.

PAN An acronym for *personal area network*, which refers to residential computer networks being established in private homes. Sometimes called *TANs* for *tiny area networks*.

Parallel Communication A data transmission method in which the bits representing a character of data are transmitted simultaneously on separate channels. (Also called *parallel transmission.*)

Parity Refers to the use of an extra bit (called a *parity bit* or a *redundant bit*) to detect single-bit errors in data transmissions. Parity can be specified as even, odd, or none. Even parity means that there must be an even number of 1-bits in each bit string; odd parity means that there must be an odd number of 1-bits in each bit string; and no parity means that parity is ignored. The extra bit (i.e., the parity bit) is forced to either 0 or 1 to make the total number of bits either even or odd.

Partitioning A network configuration strategy that involves dividing a LAN into several separate (but still interconnected) network segments. Also called *segmentation*.

PBX An acronym for *private branch exchange*, a telephone exchange used within an organization to provide internal telephone extensions and access to the public telephone network; it is the modern day equivalent of what used to be called a switchboard.

PC Card A layer 2 plug-in adapter used in portable or laptop computers. Three different “types” are available. Type 1 cards are 3.3 millimeters thick and enhance the memory capabilities of a device; Type II cards are 5 mm thick and used for modems and network adapters for both Ethernet and token ring; Type III cards are 10.5 mm thick and generally either miniature hard disks or wireless NICs; and Type IV cards, when produced, will be approximately 16 mm thick and support hard disk drives that have a capacity greater than what is currently available from Type III cards. PC cards were formerly known as PCMCIA Cards.

PCMCIA Card A layer 2 device that was originally designed to serve as memory cards for microcomputers. These cards are now known as PC Cards. “PCMCIA” stands for Personal Computer Memory Card International Association.

Peer-to-Peer A model or paradigm on which some network communications and applications are based. In a peer-to-peer environment, each networked host runs both the client and server parts of an application.

Period The reciprocal of the frequency. It is the amount of time it take to complete a single cycle, that is, seconds per cycle.

PGP An acronym for *pretty good privacy*, which is a *public key* application developed by Phil Zimmerman for e-mail security.

Physical Layer The lowest layer (layer 1) of the OSI Reference Model. The physical layer translates *frames* received from the *data link layer* (layer 2) into electrical, optical, or electromagnetic signals representing 0 and 1 values, or bits. Abbreviated PHY in the documentation.

ping A UNIX and Microsoft NT program used to test the communication path between source and destination nodes. Ping is an *ICMP*-based application and is an acronym for *packet Internet groper*.

Pinout The electrical signals associated with each pin and connector. Also called *pin assignment*.

Plaintext An uncoded message; a message in its original, meaningful (uncoded) form.

Plastic Fiber A type of fiber-optic cable in which the fibers (i.e., conductors) are constructed of plastic instead of glass.

Plenum Cable Any type of cable that contains an outer sheath or “jacket” that is composed of a Teflon coating. Plenum cable is used for cable “runs” through a return air system. The Teflon coating provides a low-flame spread and does not release toxic fumes as quickly as PVC does in the case the cable burns during a fire. Both PVC and Teflon give off nasty toxic gases when burning. Teflon, however, is fire retardant and takes much longer to get to a burning point.

Point-to-Point Network A network design in which only adjacent nodes (nodes that are next to each other and only one hop away) can communicate with one another.

POP An acronym for *point of presence*, which usually refers to a telco’s central office or switching station.

Port Connection A term used in *frame relay* to denote the *local loop*. Also called *access line*.

Port Speed A term commonly used in *frame relay* to denote the data transmission rate in bits per second of the *local loop*.

POTS An acronym for *plain old telephone system*.

PPTP An acronym for *point-to-point tunneling protocol*, which provides encryption and authentication for remote dial-up and LAN-to-LAN connections. PPTP establishes two types of connections: A control session for establishing and maintaining a secure tunnel from sender to receiver, and a data session for the actual data transmission.

PRI An acronym for *primary rate interface*, which is an *ISDN* primary access channel that comprises either 23 (United States) or 30 (Europe) 64 Mbps B channels and one 64 kbps D channel. Commonly written as 23*B* + *D*, or 30*B* + *D*.

Private Link A term used to describe a communications channel that provides a private, dedicated link between two sites. Also commonly referred to as *standard leased line*.

Private Switch A term used to describe one application of an Ethernet switch. A private switch supports only one MAC address per port, which provides each node with its own dedicated 10 Mbps segment. This eliminates contention for the cable, thereby liberating the end nodes from performing collision detection.

Promiscuous Mode A state in which an Ethernet interface can be placed so that it can capture every frame that is transmitted on the network. For example, an Ethernet NIC set

in promiscuous mode collects all messages placed on the medium regardless of their destination address.

Propagation Delay The amount of time a signal takes getting from one point in a circuit to another.

Proprietary Standards Network standards that are developed in a manufacturer-specific manner. Their specifications are not in the public domain and are only used and accepted by a specific vendor.

Protocol An accepted or established set of procedures, rules, or formal specifications governing specific behavior or language. When applied to networks, a *network protocol* is a formal specification that defines the vocabulary and rules of data communication.

Proxy Server A device or product that provides network protection at the application level by using custom programs for each protected application. These custom-written application programs act as both a client and server and effectively serve as proxies to the actual applications. Also called *application gateway firewall* are or *proxy gateway.*

PSTN An acronym for *public switched telephone network*, which is the traditional analog-based telephone system used in the United States that was originally designed for voice transmissions.

Public Key A special code, available in the public domain, that can be used to code and decode messages.

Pulse Code Modulation (PCM) A coding technique used to convert analog signals to digital signals and vice versa.

PVC An acronym for *permanent virtual circuit*, which is a communications channel that provides a logical connection between two sites instead of a physical one. In a *connection- oriented* protocol such as *frame relay*, PVCs appear as *private links* because a circuit must first be established between end nodes prior to data communications. The difference is PVCs are virtual circuits, not dedicated ones, and hence bandwidth is shared among multiple sites by *multiplexing* techniques. Thus, PVCs provide nondedicated connections through a shared medium, which enables data from multiple sites to be transmitted over the same link concurrently.

PVC Cable Any type of cable that contains an outer sheath or "jacket" that is composed of polyvinyl chloride (PVC). Also called *non-plenum cable*.

Quality of Service (QoS) Parameters associated with data prioritization that specify such things as the amount of bandwidth a priority data transmission requires as well as the maximum amount of latency the transmission can tolerate in order for the transmission to be meaningful. QoS is needed for transmitting real-time voice and video traffic.

Radio Frequencies (RF) A generic term used to describe a transmission method that uses electromagnetic waveforms.

Radio Transmission Refers to any wireless technique that uses *radio frequencies (RF)* to transmit information.

RADSL An acronym for *rate-adaptive digital subscriber line*, which is a *DSL* variant that provides transmission rates similar to *ADSL*. Transmission rates can be adjusted based on distance and line quality. Up to 7 Mbps downstream rate.

Random Access Protocol A network protocol that governs how nodes are to act in those instances where accessing a shared medium at will, on a first-come, first-served basis is permitted. Also called *contention protocol*.

RBOC An acronym for *regional bell operating company*, which refers to a regional telephone company in the United States formed after the AT&T breakup in 1984.

Redundancy Bits Extra bits incorporated into a data frame that provide error correction information. A data set composed of both user data and redundancy bits is called a *codeword*. Also called *check bits*.

Reliable Service A type of service that requires a sending node to acknowledge receipt of data. This is called an *acknowledged datagram service*.

Repeater A layer 1 device that provides both physical and electrical connections. Their function is to regenerate and propagate signals—they receive signals from one cable segment, regenerate, re-time, and amplify them, and then transmit these "revitalized" signals to another cable segment. Repeaters extend the diameter of Ethernet/802.3 networks but are considered to be part of the same collision domain.

RFC An acronym for *request for comments*, which are the working notes of the Internet research and development community. RFCs provide network researchers and designers a medium for documenting and sharing new ideas, network protocol concepts, and other technically-related information. They contain meeting notes from Internet organizations, describe various Internet protocols and experiments, and detail standards specifications. All Internet standards are published as RFCs (not all RFCs are Internet standards, though).

Ring Design A network design that is based on a broadcast topology in which nodes are connected to a physical ring, and data messages are transferred around the ring in either a clockwise or counterclockwise (or both) manner.

RIP An acronym for *routing Internet protocol*, a distance-vector algorithm that determines the best route by using a hops metric. RIP was at one time the *de facto* standard for IP routing.

RIP-2 An updated version of *RIP*, formally known as RIP version 2. New features include authentication, interpretation of IGP and BGP routes, subnet mask support, and multicasting support.

Risk Analysis The assessment of how much a loss is going to cost a company.

RJ A designation that refers to a specific series of connectors defined in the Universal Service Order Code (USOC) definitions of telephone circuits. "RJ" is telephone lingo for "registered jack."

RJ-11 A four-wire modular connector used for telephones.

RJ-45 An eight-wire modular connector used in 10BASE-T LANs.

Router A layer 3 device that is responsible for determining the appropriate path a packet takes to reach its destination. Commonly referred to as *gateway*.

Routing A layer 3 function that directs data packets from source to destination.

Routing Arbiter (RA) A project that facilitates the exchange of network traffic among various independent Internet backbones. Special servers that contain routing information databases of network routes are maintained so that the transfer of traffic among the various backbone providers meeting at a *NAP* is facilitated.

Routing Protocol A specific *protocol* that determines the route a packet should take from source to destination. Routing protocols are a function of network protocols. For example, if your network protocol is *TCP/IP*, then several routing protocol options are available including *RIP*, *RIP-2*, and *OSPF*. If your network protocol is OSI's CNLP, then

your routing protocol is IS-IS. Routing protocols determine the "best" path a packet should take when it travels through a network from source to destination, and maintain routing tables that contain information about the network's topology. Routing protocols rely on routing algorithms to calculate the least-cost path from source to destination.

Routing Table A data structure that contains, among others, the destination address of a node or network, known router addresses, and the network interface associated with a particular router address. When a router receives a packet it looks at the packet's destination address to identify the destination network, searches its routing table for an entry corresponding to this destination, and then forwards the packet to the next router via the appropriate interface.

RSA An acronym for Rivest, Shamir, and Adleman, which are the last names of the three individuals who designed the RSA *public-key* encryption algorithm.

RSVP An acronym for *resource reservation protocol*, which is an layer 3 protocol developed by *IETF* to provide a mechanism to control network latency for specific applications. This is done by prioritizing data and allocating sufficient bandwidth for data transmission. RSVP can be thought of as an IP-based Quality of Service (*QoS*) protocol.

Runt Frame An Ethernet/802.3 frame that has at least 8 bytes but less than 64 bytes long and have a valid CRC checksum.

SAN An acronym for *storage area network*, which is a network dedicated exclusively for storing data.

Satellite Communication System An RF-based broadcast network design involving Earth ground stations and orbiting communication satellites. Data transmissions from a land-based antenna to the satellite (called the *uplink*) are generally point-to-point, but all nodes that are part of the network are able to receive the satellite's *downlink* transmissions.

SC Connector A TIA/EIA-568A standard connector for fiber-optic cable; also called a 568SC connector.

SDH An acronym for *synchronous digital hierarchy*, which is an ITU-T physical layer standard that provides an international specification for high-speed digital transmission via optical fiber. SDH incorporates *SONET* and uses the *STM* signal hierarchy as its basic building block. SDH is essentially the same as SONET, and at OC-3 rates and higher, the two are virtually identical.

SDSL An acronym for *symmetric digital subscriber line*, which is a *DSL* variant in which traffic is transmitted at same rate in each direction. Maximum transmission rate is 768 kbps. Uses single-wire pair. Telephone service not supported. Suitable for videoconferencing.

Segmentation See *partitioning*.

Serial Communication A data transmission method in which the bits representing a character of data are transmitted in sequence, one bit at a time, over a single communications channel. (Also referred to as *serial transmission*.)

Server A networked device that provides resources to *client* machines. Examples include print servers, mail servers, file servers, and web servers. Servers are shared by more than user; clients have only a single user.

Shannon's Limit A mathematical theorem, named for the mathematician who derived it, Claude Shannon, that describes a model for determining the maximum data rate of a

noisy, analog communications channel. Shannon's Limit is given by the following formula, Maximum Data Rate (*MDR*) = $H \log_2(1 +)$, where *MDR* is given in bits per second, H = bandwidth in Hertz, and = a measure of the *signal-to-noise ratio*.

Shielded Twisted Pair (STP) Twisted pair cable in which individual wire pairs are shielded (i.e., protected from *noise*).

Signal-to-Noise Ratio (SNR) A measure of signal quality expressed in decibels (dB). It is the ratio of signal strength to background noise on a cable. More specifically, SNR is the ratio between the desired signal and the unwanted noise in a communications medium. In plain, late twentieth century English, it is a measure of how badly a line sucks.

Signal Quality Error (SQE) A signal generated by a transceiver and read by the controller of the host to which the transceiver is connected. In V2.0 Ethernet, SQE is called *heartbeat* and is generated periodically to inform the host's controller that the transceiver is "alive." In IEEE 802,3, SQE is only generated when a real signal quality error occurs.

Simplex Communication A data transmission method in which data may flow in only one direction; one device assumes the role of sender and the other assumes the role of receiver. These roles are fixed and cannot be reversed. An example of a simplex communication is a television transmission.

Single-attachment Station (SAS) An FDDI node that is connected to only the primary pair of fibers and can be isolated from the network in the case of some types of failure A SAS is also called *Class B* node.

Single Mode Fiber A type of fiber-optic cable with a core diameter ranging from 7 µm to 9 µm. In single mode fiber, only a single ray of light, called the *axial ray*, can

pass. Thus, a light wave entering the fiber exits with very little distortion, even at very long distances and very high data rates.

SIP An acronym for *SMDS interface protocol*, which consists of three protocol levels: SIP Level 3, SIP Level 2, and SIP Level 1. These three protocol levels are similar in function to the first three layers of the *OSI* model but represent SMDS's MAC sublayer and hence operate at the data link layer.

SMA Connector A fiber-optic cable connector that meets military specifications.

Smart Card A type of "credit card" with embedded integrated circuits that store information in electronic form and used for authentication. Similar to a *digital certificate*.

SMDS An acronym for *switched multimegabit data service*, a cell-based, connectionless, high-speed, public, packet-switched, broadband, metropolitan area data network.

SOHO An acronym for *small office/home office*.

SONET An acronym for *synchronous optical network*, which is an ANSI physical layer standard that provides an international specification for high-speed digital transmission via optical fiber. At the source interface, signals are converted from electrical to optical form. They are then converted back to electrical form at the destination interface. The basic building block of the SONET signal hierarchy is *STS*-1 (51.84 Mbps).

Spanning Tree A single path between source and destination nodes that does not include any loops. It is a loop-free subset of a network's topology. The spanning tree algorithm,

specified in IEEE 802.1d, describes how bridges (and switches) can communicate to avoid network loops.

SPID An acronym for *service profile identification*, which are numbers assigned by the telcos and used to identify the various processes of an ISDN device. (Used only in North America.)

Split-horizon A strategy employed by *RIP* to insure that a router never sends routing information back in the direction from which it came. Used to prevent routing loops.

Split-horizon With Poisoned Reverse A modified *split-horizon* strategy in which routing information provided by a neighbor is included in updates sent back to that neighbor. Such routes are assigned a cost factor of infinity, which makes the network unreachable.

Spread Spectrum A radio technology that refers to a security technique. Spread spectrum transmission camouflages data by mixing signals with a pseudonoise (PN)

pattern and transmitting the real signal with the PN pattern. The transmission signal is spread over a range of the frequencies in radio spectrum.

Statistical Multiplexing A multiplexing technique that allocates part of a channel's capacity only to those nodes that require it (i.e., have data to transmit). Based on the premise that, statistically, not all devices necessarily require a portion of the channel at exactly the same time.

Subnet Mask A special network address used to identify a specific subnetwork. Using a unique bit combination, a mask partitions an address into a network ID and a host ID.

Subnetting Refers to the partitioning of a network address space into separate, autonomous *subnetworks*. Key to subnetting is a network's *subnet mask*.

Subnetwork Refers to a network segment. Commonly abbreviated as *subnet*.

SVC An acronym for *switched virtual circuit*, which is a circuit between source and destination nodes that is established on the fly and then removed after data communications have ended. SVCs are logical, dynamic connections instead of logical permanent connections as with *PVCs*. Thus, SVCs provide switched, on-demand connectivity.

Synchronous Communication A data communication method that requires sending and receiving nodes to monitor each other's transmissions so that the receiving node always knows when a new character is being sent. In this instance, the sending and receiving nodes are "in synch" with each other.

Stackable Repeater Hub Individual repeater units "stacked" one on top of another. Instead of using a common shared backplane, stackable hubs use a "pseudo-backplane" based on a common connector interface. An external cable interconnects the individual hubs in a daisy-chained manner. Once interconnected, the entire chain of hubs becomes a single logical unit that functions as a single repeater.

Stacking Height The maximum number of stackable repeater hubs permitted.

Standby Monitor A station (i.e., node) on a token ring network that oversees the *active monitor*. Except for the active monitor, all token ring nodes are standby monitors.

Star A network configuration characterized by the presence of a central processing hub, which serves as a wire center for connecting nodes. All data must pass through the hub in order for nodes to communicate with each other.

Stateful Firewall A device or product that monitors all transactions between two systems and is capable of (1) identifying a specific condition in the transaction between two

applications, (2) predicting what should transpire next in the transaction, and (3) detecting when normal operational “states” of the connection are being violated.

Static Route A fixed route that is entered into a router’s *routing table* either manually or via a software configuration program.

ST Connector Similar to a BNC connector but used with fiber-optic cable.

Step-index Multimode Fiber A type of multimode fiber in which light pulses are guided along the cable from source to destination by reflecting off the cladding.

STM An acronym for *synchronous transport module*, which represents a digital transmission carrier system used for *Synchronous Digital Hierarchy (SDH)*. STM rates range from STM-1, which is equivalent to OC-3 (155.52 Mbps) to STM-64, which is equivalent to OC-192 (9.953 Gbps).

Store-and-Forward A method used by bridges and switches in which the contents of an entire frame is captured by the device before a decision is made to filter or forward the frame. A store-and-forward network switch is also called a buffering switch. A network that based on this principle is called a store-and-forward network.

STS An acronym for *synchronous transport signal*, which is a digital transmission hierarchy used for *SONET*. STS rates range from STS-1, which is the equivalent of 28 DS-1 channels (51.84 Mbps) to STS-192, which is the equivalent of 5,376 DS-1 channels (9.953 Gbps). STS rates are the electrical equivalent of *OC* rates.

Switch A network device that filters or forwards data based on specific information. A layer 2 switch (e.g., an Ethernet switch), filters or forwards frames from one node to another using Mac-level (i.e., hardware) addresses; a layer 3 switch filters or forwards packets based on network addresses; and layer 4 (or higher) switches filter or forward messages based on specific application protocols. Forwarding rates are usually done at wire speed and via “private” connections, i.e., no other node “sees” the traffic. Switches partition Ethernet/802.3 networks into multiple collision domains.

Switched Ethernet An Ethernet/802.3 LAN that is based on network switches instead of repeaters or bridges. A switched Ethernet LAN isolates network traffic between sending and receiving nodes from all other connected nodes. It also transforms traditional Ethernet/802.3 from a broadcast technology to a point-to-point technology.

T1 Describes the multiplexing of 24 separate voice channels, each rated at 64 kbps, plus one 8 kbps framing channel, into a single, wideband digital signal rated at 1.544 Mbps.

T2 A multiplexed circuit that combines four *T1* circuits and has an aggregate bandwidth of 6.312 Mbps.

T3 A multiplexed circuit that combines 28 *T1* circuits and has an aggregate bandwidth of 44.736 Mbps.

T4 A multiplexed circuit that combines 168 *T1* circuits and has an aggregate bandwidth of 274.176 Mbps.

TCP An acronym for *transmission control protocol*, which is a layer 4 connection-oriented protocol that performs several functions, including: providing for reliable transmission of data by furnishing end-to-end error detection and correction; guaranteeing that data are transferred across a network accurately and in the proper sequence; retransmitting any data not received by the destination node; and guaranteeing against data duplication between sending and receiving nodes. It is the "TCP" of *TCP/IP*.

TCP/IP An acronym for *transmission control protocol/Internet protocol*. Refers to a formal network protocol suite based on its two namesake sub-protocols, *TCP* and *IP*.

TE An acronym for *terminal equipment*, which represents a specific communication device that connects to an *ISDN* network. Two TEs are referenced in the specification: *TE1* refers to an ISDN-compatible device (e.g., digital telephone or a computer with a built-in ISDN port), and *TE2* refers to a non-compatible ISDN device (e.g., an analog telephone or a computer without a built-in ISDN port).

Telco An acronym for *telephone company*.

Terminal Adapter (TA) A device that connects non-compatible ISDN devices to an ISDN network. If a TA is used for an ISDN dialup connection, then it can be thought of as a modem. If a TA is used to connect a device to a LAN, then it can be thought of as a network interface card. It should be noted that although a TA is frequently referred to as an ISDN modem or digital modem in the context of an ISDN dialup connection, this reference is incorrect. By definition, a modem performs analog-to-digital and digital-to-analog conversions. Since ISDN is completely digital, no such conversions are necessary, which makes the expressions, ISDN modem or digital modem, incongruous.

Terminator Layer 1 device that prevents signal reflections by providing electrical resistance at the end of a cable to "absorb" signals to keep them from bouncing back and being heard again by the devices connected to the cable.

Thick Ethernet Describes IEEE 802.3 10BASE5, which uses "thick" coaxial cable (outer diameter between 0.375-inch and 0.405-inch) as its physical medium.

Thin Ethernet Describes IEEE 802.3 10BASE2, which uses "thin" coaxial cable (outer diameter between 0.175-inch and 0.195-inch) as its physical medium.

Threat Assessment An activity that involves determining how much security is necessary for proper control of system and network assets. Threat assessment is guided by answering the overriding question, "What assets are critical to the operation of my network and who do I think would want access to them?"

Throughput A realistic measure of the amount of data transmitted between two nodes in a given time period. It is a function of hardware/software speed, CPU

power, overhead, and many other items. Compared to *bandwidth*, throughput is what the channel really achieves, where bandwidth is what is theoretically possible.

Time Division Multiplexing (TDM) A multiplexing technique that assigns to each node connected to a channel an identification number and a small amount of time in which to transmit. TDM-based transmissions are serially sequenced.

Token A special frame on a token ring or token bus network. Possession of the token permits a node to transmit data.

Token Bus A local area network technology based on a token-passing protocol for media access. Defined in IEEE 802.4. A token bus network is characterized as a logical ring on a physical bus—physically, the network resembles a bus topology, but logically, the network is arranged as a ring with respect to passing the token from node to node.

Token Passing Protocol A network protocol that requires nodes to first possess a special frame, called a *token*, prior to transmitting data. Token-passing schemes are both contention-free and collision-free.

Token Ring A local area network technology based on a token-passing protocol for media access control. Defined by IEEE 802.5. A token ring LAN is implemented either as

a logical ring using a physical ring topology, or as a logical ring structure arranged in a physical star configuration.

traceroute A UNIX program that depicts the gateways a packet transverses. A corresponding Microsoft NT command is called *tracert*.

Transceiver A service used in Ethernet/802.3 networks to connect nodes to the physical medium. Transceivers serve as both the physical connection and the electrical interface between a node and the physical medium, enabling the node to communicate with the medium. Transceivers transmit and receive signals simultaneously.

Tree A network configuration in which nodes are connected to one another in a hierarchical fashion. A root node or hub is connected to second level nodes or hubs; second- level devices are connected to third-level devices, which in turn are connected to fourth- level devices, and so forth.

Triple DES A variant of *DES* that uses three DES operations instead of one.

Tunneling See *encapsulation*.

Twisted Pair Cable A type of copper cable that uses at least two insulated copper wires that have been twisted together. There are two basic type: *unshielded twisted pair (UTP)* and *shielded twisted pair (STP)*.

UDP An acronym for *user datagram protocol*, which is a connectionless protocol providing an unreliable datagram service. UDP does not furnish any end-to-end error detection or correction, and it does not retransmit any data it did not receive.

UDSL An acronym for *universal digital subscriber line*, which is a *DSL* variant that provides symmetrical service at 2 Mbps each way.

UNI An acronym for *user-to-network interface*, which is an end node's port where the *local loop* terminates at a customer's site.

Unicast A data transmission that is destined to a single recipient.

Unreliable Service A network service type that requires no acknowledgment of receipt of data from the receiving node to the sending node. This is called a *datagram service*.

Unshielded Twisted Pair (UTP) Twisted pair cable in which individual wire pairs are not shielded (i.e., protected from *noise*).

Utilization A network performance measure that specifies the amount of time a LAN spends successfully transmitting data. *Average utilization* means that over some period of time (e.g., a 10-hour period), on average, a certain percent of the LAN's capacity is used for successfully transmitting data. *Peak utilization* means that at a specific moment in time, a certain percent of the LAN's capacity was utilized.

V.22 *bis* ITU-T standard for 2400 bps full-duplex modems; cycles to 1200 bps/600 bps.

V.29 ITU-T standard for 9600 bps facsimile service.

V.32 ITU-T standard for 9600 bps modems; cycles to 4800 bps when line quality degrades, and cycles forward when line quality improves.

V.32 *bis* ITU-T standard that extends *V.32* to 7200, 12,000, and 14,400 bps; cycles to lower rate when line quality degrades; cycles forward when line quality improves.

V.32 *ter* Pseudo-standard that extends *V.32 bis* to 19,200 bps and 21,600 bps.

V.34 ITU-T standard for 28,800 bps modems. (*Note:* V.34 modems upgraded with special software can achieve data rates of 31,200 bps or 33,600 bps.)

V.FAST Proprietary, pseudo-standard from Hayes and Rockwell for modems transmitting at data rates up to 28,800 bps; served as a migration path for *V.34*

V.42 ITU-T standard for modem *error correction*. Uses *LAPM* as the primary error-correcting protocol, with *MNP* classes 1 through 4 as an alternative.

V.42 *bis* ITU-T standard that enhances *V.42* by incorporating the British Telecom Lempel Ziv data *compression* technique to V.42 *error correction*. Most *V.32, V.32 bis,* and *V.34* compliant modems come with V.42 or V.42 bis or MNP.

V.90 ITU-T standard for 57,600 bps modems (commonly called "56K modems") in which asymmetric data rates apply (i.e., the send and receive rates are different). Depending on telephone line conditions, upstream rates (send) are restricted to 33,600 bps, and downstream rates (receive) are restricted to 57,600 bps. V.90 modems are designed for connections that are digital at one end and have involve only two analog-digital conversions each way.

vBNS An acronym for *very high speed backbone network service*, which is another National Science Foundation-funded research and educational network. The vBNS is a nationwide backbone network that currently operates at 622 Mbps (OC-12) and is accessible to only those involved in high-bandwidth research activities. The backbone is expected to be upgraded to OC-48 (2.488 Gbps) in 1999.

VDSL An acronym for *very high-speed digital subscriber line*, which is a *DSL* variant that provide asymmetric service over fiber. Downstream rates range from 13 Mbps to 52 Mbps; upstream rates range from 1.5 Mbps to 2.3 Mbps. Suitable for Internet/intranet access, video-on-demand, database access, remote LAN access, and high-definition TV.

Virtual Channel Connection (VCC) A virtual circuit that provides a logical connection between an *ATM* source and destination. Data can only be transmitted in one direction via a VCC. A VCC is denoted by a *virtual channel identifier (VCI)*, which is included as part of the ATM cell header. Multiple virtual channels that share the same connection can be packaged into a single *virtual path*.

Virtual Channel Identifier (VCI) A parameter used to identify *ATM virtual channels*. VCI information is carried within an ATM cell header.

Virtual Circuit A nondedicated connection through a shared medium that gives the high- level user the appearance of a dedicated, direct connection from the source node to the destination node.

Virtual Path Connection (VPC) A semi-permanent connection that provides a logical collection of *ATM virtual channels* that have the same end points. More specifically, a VPC carries a group of virtual channels all of which have the same end points. Virtual paths enable any connection that uses the same network path from source to destination to be bundled into a single unit. A *virtual path identifier (VPI)* is used denote a virtual path and is included in a cell's header. A virtual path can also provide a form of traffic control by logically (not physically) partitioning network traffic based on the type of data being carried and associated *quality of service*.

Virtual Path Identifier (VPI) A parameter used to identify *ATM virtual path*. VPI information is carried within an ATM cell header.

VLAN An acronym for “virtual local area network.” Nodes comprising a VLAN are not physically connected to the same medium. Instead, they are connected in a virtual sense using specially designed software that groups several ports in a switch into a single work

group. Nodes connected to these ports are considered to be part of a workgroup, and network traffic from any node/port is (usually) limited to only those nodes or ports assigned to the workgroup.

VOFR An acronym for *voice over frame relay*, which refers to transmitting voice signals over a *frame relay* network.

Voice Over IP (VOIP) A technology that enables users to place telephone calls across the Internet.

VPN An acronym for *virtual private network*, which refers to an IP connection between two sites over a public IP network that has its payload traffic encrypted so that only source and destination nodes can decrypt the traffic packets. A VPN enables a publicly accessible network to be used for highly confidential, dynamic, and secure data transmissions.

WAN An acronym for *wide are network*, which interconnects computing resources that are widely separated geographically (usually over 100 km). This includes towns, cities, states, and countries. A WAN generally spans an area greater than five miles (eight kilometers). A WAN can be thought of as consisting of a collection of LANs.

Wavelength A measure of the length of a wave. It is the distance an electrical or light signal travels in one complete cycle.

Wavelength Division Multiplexing (WDM) A *multiplexing* method used with fiber-optic cables. Involves the simultaneous transmission of light sources over a single fiber- optic channel. Light sources of different wavelengths are combined by a WDM multiplexer and transmitted over a single line. When the signals arrive, a WDM demultiplexer separates them and transmits them to their respective destination receivers.

Wire A general term used to describe the physical layer of a network. The three main physical attributes of wire are *conductor, insulation,* and *outer jacket*. Wire also has three important electrical characteristics that can directly affect the quality of the signal transmitted across it: *capacitance, impedance,* and *attenuation*. Signal quality is affected most by the combination of attenuation and capacitance. The two primary forms of wire are copper and fiber. Also called *cable*.

Wireless Communications A type of communications in which signals travel through space instead of through a physical cable. There are two general types of wireless communication: *radio transmission* and *infrared transmission*.

Wire Speed A unit of measure used to describe a device’s maximum (i.e., fastest) filtering and forwarding rates. In Ethernet/802.3, wire speed is equal to 14,880

frames per second. This is frequently reported as 14,880 packets per second. (See Box 8-3.)

WLAN An acronym for *wireless LAN*.

Workgroup Switch A term used to describe one application of an Ethernet switch. A workgroup switch partitions a single, shared medium into multiple, shared media and supports more than MAC address per port. Also called *segment switches*.

Workstation A computer system that has its own operating system and is connected to a network. A workstation can be a personal computer such as a Macintosh or Intel-based PC, a graphics workstation such as those manufactured by Sun Microsystems, a super- minicomputer such as IBM's AS/400, a super-microcomputer such as DEC's Alpha, or a mainframe such as an IBM ES-9000. Also called *host, server, desktop,* or *client.*