ISC CISSP

Certified Information Systems Security Professional

Study Guide

Version 2.0

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3DES	Triple Data Encryption Standard (Triple DES)
AAA	Authentication, Authorization, and Accounting
ACK	Acknowledgement (Message)
ACL	Access Control List
ADSL	Asymetrical Digital Subscriber Line
AES	Advanced Encryption Standard
ALE	Annual Loss Expectancy
ALU	Arithmetic Logic Unit
AM	Active Monitor (Token Ring)
ANSI	American National Standards Institute
ARO	Annual Rate of Occurrence
ARP	Address Resolution Protocol
AS	Authentication Server (Kerberos)
ATM	Asynchronous Transfer Mode
AV	Asset Value
BCP	Business Continuity Planning
BGP	Border Gateway Protocol
BIA	Business Impact Assessment
CA	Certificate Authority
CAN	Campus Area Network
CATV	Cable Television
CCTV	Closed-Circuit Television
CDDI	Copper Distribution Data Interface
CEI	Computer Ethics Institute
CER	Crossover Error Rate
CFR	Code of Federal Regulations
CHAP	Challenge Handshake Authentication Protocol
CIA	Confidentiality, Integrity, and Availability
CISC	Complex Instruction Set Computing
CLNS	Connectionless Network Service
CMNS	Connection-Mode Network Service
COI	Conflict of Interest
COPPA	Children’s Online Privacy Protection Act
CORBA	Common Object Request Broker Architecture
CPE	Customer Premises Equipment
CPU	Central Processing Unit
CSMA/CD	Carrier Sense Multiple Access Collision Detect (Ethernet)
CSS	Content Scrambling System

CSSPAB	Computer System Security and Privacy Advisory Board
DAC	Discretionary Access Control
DBMS	Database Management System
DBPSK	Differential Binary Phase-Shift Keying (Ethernet)
DCE	Data Circuit-Terminating Equipment
DDoS	Distributed Denial of Service
DES	Data Encryption Standard
DISA	Deploy Direct Inward System Access
DLCI	Data Link Connection Identifier (Frame Relay)
DMZ	Demilitarized Zone
DNS	Domain Naming System
DoD	U.S. Department of Defense
DoS	Denial of Service
DQPSK	Differential Quaternary Phase-Shift Keying (Ethernet)
DRP	Disaster Recovery Planning
DSL	Digital Subscriber Line
DSU/CSU	Data Service Unit/Channel Service Unit
DTE	Data Terminal Equipment
EAP	Extensible Authentication Protocol
EAP-LTS	Extensible Authentication Protocol-Transport Layer Security
EF	Exposure Factor
EIA	Electronics Industry Association
EIGRP	Enhanced Interior Gateway Routing Protocol
EMI	Electromagnetic Interference (Noise)
ESP	Encapsulating Security Payload
ETR	Early Token Release (Token Ring)
FCS	Frame Check Sequence
FDDI	Fiber Distribution Data Interface
FISA	Foreign Intelligence Surveillance Act
FR	Frame Relay
FSM	Finite State Machine
FTP	File Transfer Protocol
GRE	Generic Routing Encapsulation
HDSL	High-bit-rate Digital Subscriber Line
HIDS	Host-based Intrusion Detection System
HIPAA	Health Insurance Portability and Accountability Act
HVAC	Heating, Ventilating, and Air Conditioning
IAB	Internet Activities Board
ICMP	Internet Control Message Protocol
IDC	IBM Data Connector
IDEA	International Data Encryption Algorithm

IDL	Interface Definition Language
IDS	Intrusion Detection System
IDSL	ISDN Digital Subscriber Line
IEEE	Institute of Electrical and Electronics Engineers
IGRP	Interior Gateway Routing Protocol
IKE	Internet Key Exchange
IMAP	Internet Message Access Protocol
IP	Internet Protocol
IPL	Initial Program Load
IPSec	IP Security
IR	Infrared
IRM	Information Resources Management
ISDN	Integrated Services Digital Network
IS-IS	Intermediate System-to-Intermediate System
ISP	Internet Service Provider
IT	Information Technology
KDC	Key Distribution Center (Kerberos)
L2F	Layer 2 Forwarding
L2TP	Layer 2 Tunneling Protocol
LAN	Local Area Network
LCP	Link Control Protocol
LED	Light Emitting Diode
LLA	Locally Administered Address (Token Ring)
LLC	Logical Link Control
MAA	Mutual Assistance Agreement
MAC	Mandatory Access Control
MAC	Media Access Control (Address)
MAN	Metropolitan Area Network
MBI	Manpower Buildup Index
MBR	Master Boot Record
MD2	Message Digest 2
MD5	Message Digest 5
MIME	Multipurpose Internet Mail Extensions
MMF	Multimode Fiber (Cable)
MOSS	MIME Object Security Services
MPPE	Microsoft Point-to-Point Encryption
MSAU	Multistation Access Unit (Token Ring)
MTD	Maximum Tolerable Downtime
MTTF	Mean Time To Failure
MTTR	Mean Time To Repair
MTU	Maximum Transmission Unit

NAT	Network Address Translation
NAUN	Nearest Active Upstream Neighbor (Token Ring)
NIC	Network Interface Card
NIDS	Network-based Intrusion Detection System
NIST	National Institute of Standards and Technology
NOS	Network Operating System
NSA	National Security Agency
OECD	Organization for Economic Cooperation and Development
OIRA	Office of Information and Regulatory Affairs
OLAP	On-Line Analytical Processing
OMB	Office of Management and Budget
OMG	Object Management Group
OOP	Object-Oriented Programming
ORA	Object Request Architecture
ORB	Object Request Broker
OS	Operating System
OSI	Open Systems Interconnection (Model)
OSPF	Open shortest Path First
P2P	Peer-To-Peer (Network)
P3P	Platform for Privacy Preferences
PAP	Password Authentication Protocol
PAT	Port Address Translation
PBX	Private Branch Exchange
PCMCIA	Personal Computer Memory Card International Association (Interface)
PEAP	Protected Extensible Authentication Protocol
PEM	Privacy Enhanced Mail
PF	Productivity Factor
PGP	Pretty Good Privacy
PII	Personal Indentifiable Information
PIN	Personal Identification Number
PKCS	Public Key Cryptography Standard
PKI	Public Key Infrastructure
POP3	Post Office Protocol 3
POTS	Plain Old Telephone Service
PPP	Point-to-Point Protocol
PPTP	Point-to-Point Tunneling Protocol
PSN	Packet-Switched Network
PSTN	Public Switched Telephone Network
PVC	Permanent Virtual Circuit (Frame Relay)
QPSK/CCK	Quaternary Phase-Shift Keying/Complimentary Code Keying (Ethernet)
RA	Registration Authority

RADIUS	Remote Authentication Dial-In User Service
RADSL	Rate-Adaptive Digital Subscriber Line
RAM	Random Access Memory
RARP	Reverse Address Resolution Protocol
RAS	Remote Access Service
RBAC	Role-Based Access Control
RFC	Request For Comment (Document)
RFI	Radio Frequency Interference (Noise)
RICO	Racketeer Influenced and Corrupt Organization (Act)
RIP	Routing Information Protocol
RISC	Reduced Instruction Set Computing
ROM	Read-Only Memory
RPS	Ring Parameter Server (Token Ring)
RSA	Rivest, Shamir, & Adleman (Encryption)
S/MIME	Secure Multipurpose Internet Mail Extensions
SBU	Sensitive But Unclassified
SDLC	System Development Life Cycle
SDSL	Symmetrical Digital Subscriber Line
SESAME	Secure European System and Applications in a Multivendor Environment
SFTP	Secure File Transfer Protocol
SLA	Service Level Agreement
SLCM	Software Life Cycle Model
SLE	Single Loss Expectancy
SMDS	Switched Multimegabit Data Service
SMF	Single-Mode Fiber (Cable)
SMTP	Simple Mail Transfer Protocol
SNMP	Simple Network Management Protocol
SPAP	Shiva Password Authentication Protocol
SSH	Secure Shell
STP	Shielded Twisted Pair (Cable)
SUI	Sensitive Unclassified Information
SVC	Switched Virtual Circuits (Frame Relay)
TACACS	Terminal Access Controller Access Control System
TCB	Trusted Computer Base
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
TCSEC	Trusted Computer System Evaluation Criteria
TDM	Time-Division Multiplexed
TFTP	Trivial File Transfer Protocol
TGS	Ticket-Granting Server (Kerberos)
TIA	Telecommunications Industry Association

UART	Universal Asynchronous Receiver-Transmitter
UDP	User Datagram Protocol
UPS	Uninterruptible Power Supply
UTP	Unshielded Twisted Pair (Cable)
VC	Virtual Circuit (Frame Relay)
VDSL	Very-high-bit-rate Digital Subscriber Line
VNC	Virtual Network Computing
VoIP	Voice over IP
VPN	Virtual Private Network
W3C	World Wide Web Consortium
WAN	Wide Area Network
WAP	Wireless Access Point
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network
XTACACS	Extended TACACS

Certified Information Systems Security Professional

Certifications:

Certified Information Systems Security Professional (CISSP)	**Core**
(ISC)^2 Associate for CISSP (Associate of (ISC)^2)	**Core**

Prerequisites:
At least 4 years' experience or a college degree with 3 years' experience as a practicing security professional. Candidates without the required experience can become an (ISC)^2 Associate for CISSP (Associate of (ISC)^2).

About This Study Guide
This Study Guide provides all the information required to pass the (ISC)^2 CISSP exam. It however, does not represent a complete reference work but is organized around the specific skills that are tested in the exam. Thus, the information contained Study Guide is specific to the CISSP exam and not Information Systems security. It includes the information required to answer questions related to the CISSP exam. Topics covered in this Study Guide includes: Understanding Security Management, Risk Management, and Risk Assessment; Identifying Threats and Vulnerabilities; Performing Quantitative and Qualitative Assessment of Assets; Understanding Security Policies and Procedures, including Security Policy Objectives, Security Policy Standards, Guidelines and Procedures, and the Various Types of Information Classification; Providing Security Training and Education; Understanding and Implementing Access Control and Accountability; Understanding the Various Access Control Models, including Discretionary Access Control (DAC), Mandatory Access Control (MAC) and Role-based Access Control (RBAC); Understanding Access Control Types, including Passwords, Tokens, Biometrics, Multifactor Authentication, Single Sign-On, Kerberos, Secure European System and Applications in a Multivendor Environment (SESAME), KryptoKnight and NetSP; Understanding Access Control Systems, including Centralized Access Control, Remote Authentication Dial-In User Service (RADIUS) and DIAMETER, and Terminal Access Controller Access Control System, as well as Decentralized/Distributed Access Control; Understanding Threats against Access Control, including Password Attacks, Dictionary Attacks, Brute-Force Attacks, Back Door Attacks, Spoofing, Man-in-the-Middle Attacks, Replay Attacks, Denial of Service (DoS) and Distributed Denial of Service (DDoS) Attacks, TCP Hijacking, Social Engineering, Dumpster Diving and Software Exploitation; Monitoring Information Systems for Possible Intrusion and Implementing Intrusion Detection; Understanding Intrusion Detection System (IDS) and Honeypots; Performing Penetration Testing; Understanding Telecommunications and Network Security; Understanding the OSI Reference Model; Understanding the Protocols of the Transmission Control Protocol/Internet Protocol (TCP/IP) Architecture; Understanding and Implementing Communication and Network Security; Identifying the Various Types of Networks, Network Topologies and Network Cabling; Understanding Wireless Networks, including IEEE 802.11x, Bluetooth, and IrDA; Understanding Network Technologies, including Ethernet, and Token Ring; Understanding Data Network Services, including File Transfer Services (FTP), Secure File Transfer Protocol (SFTP), Trivial File Transfer Protocol (TFTP), Secure Shell (SSH) and Secure Shell version 2 (SSH-2), Understanding Wide Area Networks, including the Internet, Intranets, and Extranets, Understanding WAN Technologies, including Dedicated Lines, WAN Switching, Circuit-Switched

Networks, and Packet-Switched Networks Understanding Network Address Translation (NAT); Understanding and Implementing Remote Access; Understanding Virtual Private Networks (VPNs) and VPN Applications; Integrating VPN in a Routed Intranet; Understanding and Implementing E-mail Security and E-mail Security Solutions; Understanding and Implementing Cryptography; Understanding Data Encryption, including Symmetric and Asymmetric Algorithms; Understanding and Implementing a Public Key Infrastructure (PKI); Understanding Certificates and Certificate Policies; Understanding System Architecture, including Computer Architecture; Understanding Security Policies and Computer Architectures; Implementing Security Mechanisms for Computer Architectures, including Process Isolation, Rings of Protection and Trusted Computer Base (TCB); Understanding Single-State and Multistate Systems; Understanding the Various Security Models, including the State Machine Model, the Bell-LaPadula Model, the Biba Integrity Model, the Clark-Wilson Integrity Model, the Information Flow Model, the Noninterference Model, the Take-Grant Model, the Access Control Matrix, and the Brewer and Nash Model; Understanding and Implementing Operational Security; Understanding the Role of Employees in Operational Security; Implementing New-Hire Orientation, Understanding the Importance of Separation of Duties and Job Rotation; Understanding Threats and Vulnerabilities to Operational Security, including Traffic Analysis, Insecurities Associated with Default and Maintenance Accounts, Data-Scavenging Attacks, Initial Program Load Vulnerabilities, Social Engineering, and Network Address Hijacking; Understanding the Importance of Auditing, Monitoring and Intrusion Detection; Understanding Audit Trails; Understanding Controls for Operational Security, including Orange Book Controls; Understanding and Implementing Operations Controls, including Resource Protection, Hardware Controls, Software Controls, Privileged Entity Controls, Media Security Controls, Media Viability Controls, and Physical Access Controls; Understanding Application and System Development; Understanding, Identifying and Protecting against Malicious Code, including Viruses, Worms, Logic Bombs, Trojan Horses, Active Content, Spyware, and SQL Injection; Understanding the System Development Life Cycle (SDLC); Understanding Software Development Models, including the Waterfall Model, the Spiral Model, and Cost Estimation Models; Understanding and Implementing Information Security and the Life Cycle Model; Understanding Object-Oriented Programming; Understanding Implementing Secure Database Management; Understanding the Importance Business Continuity Planning and Disaster Recovery Planning; Understanding and Implementing Alternate Recovery Sites, including Cold Sites, Hot Sites, Warm Sites, and Mobile Sites; Understanding Computer Crimes, including the Laws Related to Computer Crimes and the of Computer Crimes Understanding Information Privacy and Privacy Laws; Understanding Computer Forensics; Understanding Ethical Computing and the Various Codes of Ethics; Implementing Physical Security; Designing Secure Facilities; Implementing Physical Access Controls; Understanding Environment and Personnel Safety; Implementing Environmental Controls, including Heating, Ventilating, and Air Conditioning (HVAC), and Fire Detection and Suppression; and Understanding Electromagnetic Interference (EMI) and Radio Frequency Interference (RFI)

Intended Audience

This Study Guide is targeted specifically at people who wish to take the $(ISC)^2$ CISSP exam. This information in this Study Guide is specific to the exam and is not a complete reference work. Although our Study Guides are aimed at new comers to the world of IT, the concepts dealt with in the exam, and consequently in this Study Guide are rather complex. We therefore suggest that a sound knowledge of CompTIA's A+, N+ and Server+ course work material would be advantageous.

How To Use This Study Guide

To benefit from this Study Guide we recommend that you:

- Study each chapter carefully until you fully understand the information. This will require regular and disciplined work.

> **Note:** Remember to pay special attention to these note boxes as they contain important additional information that is specific to the exam.

- Perform all labs that are included in this Study Guide to gain practical experience, referring back to the text so that you understand the information better. Remember, it is easier to understand how tasks are performed by practicing those tasks rather than trying to memorize each step.

- Be sure that you have studied and understand the entire Study Guide before you take the exam.

Good luck!

Topic 1: Security Management

Security management concepts and principles are inherent elements in a security policy and solution deployment. They encompass of critical documents, such as policies, procedures, and guidelines that define the basic parameters needed for a secure an information system. These documents identify the organization's information assets and define the organization security practices.

The primary goals and objectives of security are contained within the **CIA Triad**, which is the three primary security principles: confidentiality, integrity, and availability. Security controls must address one or more of these three principles.

Confidentiality

Confidentiality is the process of ensuring that sensitive information is not disclosed to unauthorized persons. When there is an unintentional release of information, confidentiality is lost. Attacks on confidentiality include sniffing, keystroke monitoring, and shoulder surfing

Integrity

Integrity is the process of ensuring that data is consistent and that it has not modified without authorization. This applies to data in use, data in storage and data in transit.

Availability

Availability ensures that data and systems are always available and can be accessed by authorized personnel whenever needed.

Section 1.1: Risk Assessment

Risk is the possibility of experiencing some form of loss. It does not mean the risk will be realized, but that it has the potential to occur. Risk management is used to determine what risks are potential threats and to deal with these risks. By taking a proactive approach to risks, the damage that can occur from them is minimized. Risk identification is the process of ascertaining what threats pose a risk to an organization. There are many different types of risks that can affect an organization. Each business must identify the risks they may be in danger of confronting. Disasters can be naturally occurring or the result of accidents and malfunctions.

Natural disasters include storms, floods, fires, earthquakes, tornadoes, or any other environmental event. They also include situations that may cause damage to an organization, such as when a fire breaks out due to faulty wiring, a pipe bursts, or a power outage occurs. In addition to these risks, the organization is commonly at risk for equipment failures. There are a number of different risks that result from malicious persons and the programs they use and disseminate. Trojan horse attacks, viruses, hackers, and various other attacks can devastate an organization as effectively as any natural disaster. An attack on systems can result in disruption of services or the modification, damage, or destruction of data. Internal risks are risks in which consequences result from the actions of persons employed by an organization. Software and data are also targets of corporate theft. Employees may steal installation CDs or make copies of software to install at home. A single program can cost one thousand dollars or more, while copied CDs that are illegally installed could result in piracy charges and legal liability. If an employee takes sensitive data from a company and sells it to a competitor, the company could lose millions of dollars or face liability suits or even criminal charges if the stolen data breaches client confidentiality.

1.1.1: Risk Management

Risk management is the act of determining what threats the organization faces, analyzing vulnerabilities to assess the threat level, and determining how risk should be dealt with. This could involve developing a risk-management team, identifying threats and vulnerabilities, placing a value on the organization's assets, and determining the risks that are uncover will be dealt with. There are three important concepts in risk management: **threat**, which is a man-made or natural event that could have a negative impact on the organization; **vulnerability**, which is a potential weakness resulting from a flaw, loophole, oversight, or

error that could be exploited to violate system security policy; and **controls**, which can be corrective, detective, preventive, or deterrent mechanisms that an organization can use to restrain, regulate, or reduce vulnerabilities.

1.1.2: Identifying the Threats and Vulnerabilities

Identifying threats and vulnerabilities is an important part of the risk-management process. Threats can occur as a result of human or natural factors, and can be caused by internal or external events. Threats can also occur because of errors in computer code, accidental buffer overflows, or the unintentional actions of employees.

You can start to analyze the threats, vulnerabilities, and risks the organization faces by creating a table such as the one shown in Table 1.1. This helps demonstrate the relationship among threats, vulnerabilities, and risk. For example, an intruder can represent a threat that exposes the organization to theft of equipment because there is no security guard or controlled entrance.

Table 1.1: Threat, Vulnerability and Risk

Threat	Vulnerability	Risk
Intruder	No security guard or controlled entrance	Theft
Hacker	Misconfigured firewall	Stolen credit card information
Current employee	Poor accountability; no audit policy	Loss of integrity; altered data
Fire	Insufficient fire control	Damage or loss of life
Hurricane	Insufficient preparation	Damage or loss of life
Virus	Out-of-date antivirus software	Virus infection and loss of productivity
Hard drive failure	No data backup	Data loss and unrecoverable downtime

1.1.3: Assessing Asset Value

Identifying the assets that are the most crucial to the organization and that should be protected is as important as identifying threats and vulnerabilities because it would be foolish to exceed the value of the asset by spending more on the countermeasure than the asset is worth. Organizations usually have limited funds and resources, so countermeasures must be effectively deployed to protect the most critical assets. For this reason you must assess the value of assets held by the organization. This can be a quantitative assessment, in monetary value, or a qualitative assessment, in importance.

1.1.3.1: Quantitative Assessment

In a quantitative assessment you attempt to assign a monetary value to the components of a risk assessment and to the assets and threats of a risk analysis. All elements of the risk analysis process, which includes asset value, impact, threat frequency, safeguard effectiveness, safeguard costs, uncertainty, and probability, must be quantified. When performing this assessment, you should include all associated costs, such loss of

productivity; cost of repair; value of the damaged equipment or lost data, and cost to replace the equipment or reload the data. The quantitative assessment involves three steps:

- **Single loss expectancy (SLE)**, which is the estimated potential one time losses of an asset. The SLE is calculated as the asset value multiplied by its exposure factor, which is the percent of damage that a realized threat would have on the asset.
- **Annual rate of occurrence (ARO)**, which is the estimated number of times an event may occur within a year.
- **Annual loss expectancy (ALE)**, which combines the SLE and ARO to determine the magnitude of the risk. The ALE is calculated as SLE multiplied by the ARO.

1.1.3.2: Qualitative Assessment

It is difficult, if not impossible, to assign monetary values to all components of a risk analysis. Therefore some qualitative measures must be applied to quantitative elements. Qualitative assessment does not attempt to assign monetary values to components of the risk analysis. Instead it rates the severity of threats and the sensitivity of assets and categories components on the basis of this rating. The categories components can be rated as are:

- **low**, when loss of a component would be a minor inconvenience that could be tolerated for a short period of time.
- **medium**, when loss of a component could result in damage to the organization or could result in a moderate expense to repair.
- **high**, when loss of a component would result in loss of goodwill between the organization and clients or employees, and could result in a legal action or fine, or cause the organization to lose revenue or earnings.

1.1.4: Handling Risk

Risk can be dealt with in four ways: by implementing a countermeasure to alter or reduce the risk, this is referred to as **risk reduction**; by taking out insurance to transfer a portion or all of the potential cost of a loss to a third party, this is referred to as **risk transference**; by accepting the potential cost and loss if the risk occurs, this is referred to as **risk acceptance**; and by ignoring a risk, this is referred to as **risk rejection**. Obviously, you could combine these measures.

Section 1.2: Security Policies and Procedures

Security policies are official, high-level security documents that are created in accordance with the security philosophy of an organization. These documents are a general statement about the organization's assets and the level of protection each asset or set of assets should enjoy. Well-written security policies would contain a set of rules to which users in the organization should adhere to when accessing the network resources. It would include a list of acceptable and unacceptable activities and would define responsibilities in respect of security. However, security policies do not dictate how the policies should be implemented. The security policy is therefore a guide for administrators to use when planning security efforts and subsequent reactions.

Security policies can be written to meet advisory, informative, and regulatory needs, each with its own unique role and function.

- Advisory policies ensure that all employees know the consequences of their actions.

- Informative policies are designed to inform and enlighten employees about company procedures.
- Regulatory policies ensure that the organization complies with local, state, and federal laws.

The security policy could also be divided into smaller documents, with each document addressing a specific topic. These documents could include:

- **Use Policy**, which addresses suitable use of items such as email and Internet access.
- **Configuration Policy**, which spells out which applications are to be arranged on the network and should assign a particular build for each system**.** This is important in making sure that all the network systems follow a set arrangement, thus cutting down on troubleshooting time.
- **Patch Management** system, which explains the upgrade and testing of new patches before being implemented. After approval, it is added to the standard build. This guarantees that all new systems are in accordance with the approved patch.
- **Infrastructure Policy**, which spells out how the system is to be managed and maintained, and by whom. It also addresses the following:
  - Service Quality
  - Checking and Controlling the systems
  - Processing and Consolidating of Logs
  - Managing change
  - The scheme for addressing network
  - The standard for naming
- **User Account Policy**, which spells out which users should be designated what permissions. It is important to ensure adherence to the PC configuration policy. This can be achieved by limiting user permissions.
- **Other policie**s: The amount of policies varies according to each organization. Factors such as encryption, backup, the handling of data and password requirements, like time span and size, as well as remote access, could be included in other policies.

1.2.1: The Objectives of a Security Policy

- The initial objective would be to **advise the technical team on their choice of equipment**. Because of the policy not being in the nature of a technical document, it neither dictates nor stipulates which equipment or designs are to be used.
- The second objective would be to **guide** the team in arranging the equipment. It might state that the team uses their efforts to block users from viewing unacceptable websites. It does not however stipulate specific sites.
- The third objective spells out the **responsibilities** of users and administrators. This assists management and technicians in evaluating the effectiveness of the security measures.
- The fourth objective would set out the **consequences** of a policy violation.
- The last objective would be to spell out the **reactions** to network threats. If there is no plan for fending off an attack, the result would be bewilderment. It is also significant to describe escalation steps for items that are not as easily recognized on the network. Each member should know the steps to be taken in the event of a problem.

1.2.2: Standards, Guidelines and Procedures

Standards, guidelines, and procedures are documents at the level directly below the policies and constitute the three elements of policy implementation. These three documents contain the actual details of the policy, such as how it should be implemented and what standards and procedures should be used.

- **Standards** define the processes that need to be implemented to meet a policy requirement but does not define the method of implementation.
- **Guidelines** are recommendation or suggestion of how policies or procedures should be implemented. These are meant to be flexible enough to allow customization for individual situations.
- **Procedures** are the most specific and most detailed security documents. They define the actual step-by-step implementation of security mechanisms to meet the requirements of the security policies. Procedures are linked to specific technologies and devices and could change as equipment changes.

1.2.3: Roles and Responsibility

Members of an organization often fulfill different roles. However, from a security administration perspective, it is important that roles and responsibilities are clearly defined. Roles and responsibilities are also central to the **separation of duties**, a security concept that enhances security through the division of responsibilities in the production cycle.

Table 1.2: Roles and Responsibilities

Role	Description
Senior Manager	Although senior management may delegate the function of security, they have the ultimate responsibility for security of information where liability is concerned.
Information Systems Security (InfoSec) Officer	The Information Systems Security (InfoSec) Officer has the functional responsibility for implementing and maintaining security. The InfoSec Officer's duties include the design, implementation, management, and review of the organization's security policy, standards, guidelines, and procedures.
Data owner	Data owners are primarily responsible for determining the data's sensitivity or classification levels. They can also be responsible for maintaining the information's accuracy and integrity.
Data custodian	The Data Custodian has the responsibility of preserves the data's confidentiality, integrity and authenticity. They are responsible for the day-to-day management of data, controlling access, adding and removing privileges for individual users, and ensuring that the proper controls have been implemented.
User	Users are responsible for following the procedures defined by the organization's security policies during the course of their normal daily tasks.
Security auditor	Examines security and provides reports on the effectiveness of the security controls to the senior management. They

also check that security policies, standards, guidelines, and procedures comply with the company's stated security objectives.

Section 1.3: Information Classification

Not all data has the same value to an organization. Some data is more valuable than others to specific users within the organization. Some data, such as trade secrets, formulas, and new product information, are of great value and their loss could result in significant decline for enterprise. For these reasons information classification has a higher, enterprise-level benefit. The primary purpose of information classification is to enhance confidentiality, integrity, and availability of the information, and to minimize the risks to the information. It is also a means of rating an organization's informational assets. Each level of classification that is established should have specific requirements and procedures.

The two most common information classification models are military information classification and commercial information classification. Both the military and commercial information classification models have predefined labels and levels. Table 1.3 lists the levels of both military and commercial information classification models.

Table 1.3: Commercial and Military Information Classifications

Commercial Classifications	Military Classifications
	Top secret
Confidential	Secret
Private	Confidential
Sensitive	Sensitive
Public	Unclassified

Section 1.4: Security Training and Awareness

People are often the weakest point in security, often because they are not aware of the implications their actions have on the overall security of the system. Security awareness is crucial in overcoming this problem but is the most overlooked element of security management.

Security awareness programs can effectively increasing employee understanding of security and the implications of their actions on overall security. All employees need education in the basic concepts of security and its benefits to an organization. However, security awareness training must be developed differently for the different groups of employees that make up the organization. There are three primary groups that security awareness training should be targeted at. These are: senior management, data custodians, and users.

Note: Security awareness increases management's ability to hold employees accountable for their.

There are three tenets of security awareness: awareness, training and education.

- Security **awareness** refers to the general awareness amongst an organization's personnel of the importance of security and security controls. Once personnel have a clear understanding of the need for security, they are considered to be "security aware".
- **Training** programs are classroom based or one-on-one training that usually teach individual users a specific skill. They are of a short duration.
- Unlike training, **education** is usually classroom based but with broader objectives and of longer duration.

Topic 2: Access Control and Accountability

Access control and accountability are important concepts in understanding computer and network security. These concepts are used to protect property, data, and systems from deliberate or not deliberate damage. These two concepts, together with Auditing, are used to support the **Confidentiality**, **Integrity**, and **Availability (CIA)** security concept, as well as for access to networks and equipment using **Remote Authentication Dial-In User Service (RADIUS)** and **Terminal Access Controller Access Control System (TACACS/TACACS+)**.

Access control is best described as the process by which use of resources and services is granted or denied. This can be accomplished by using advanced component such as a smart card, a biometric device, or network access hardware, or simply by means of user name and passwords. It can include routers, remote access points such as **remote access service (RAS)** and **virtual private networks (VPNs)**, or the use of **wireless access points (WAPs)**. It can also be a file or shared resource permission assigned through the use of a network operating system (NOS).

Section 2.1: Access Control Models

The three common models of access control are:

- **Discretionary access control (DAC)**;
- **Mandatory access control (MAC)**; and
- **Role-Based access control (RBAC)**

2.1.1: Discretionary Access Control (DAC)

Discretionary access control (DAC) allows the owner of an object to manage access control at his or her own discretion. In other words, the holder of an object sets the access permissions at his or her judgment whereas with **mandatory access control (MAC)** and **role-based access control (RBAC)**, access to the information is synchronized by a formal set of rule. The user or application has control of the settings. This is the standard type of access control used and includes setting permissions on files, folders, and shared resources. You apply access control in every state that information is found in your enterprise. This includes computerized data as well as hard-copy files, photographs, displays, and communication packets. With DAC, an **access control list (ACL)** is maintained that lists the users with permissions to resources and the type of access they are permitted. With discretionary authentication, the **ACL** can become quite large if individual users are added. This can become difficult to manage and can impact the overall system performance as well.

There are a number of risks associated with DAC. These include:

- Software might be executed or updated by illegal personnel.
- Classified information might be by chance or intentionally compromised by users who are not intended to have access.
- Auditing of file and resource accesses might be difficult.

Note: In a large or small network, one of your objectives is to supply users with access to the resources they need. To control users in larger networks,

you should organize and manage them as groups as opposed to individuals. Organize users that require similar access to resources into groups and assign permissions to the group rather than individuals. As users leave or change positions, their access capabilities change. Using groups with intuitive names to occupy ACLs and adding users to the groups is a better, more secure management method.

2.1.2: Mandatory Access Control (MAC)

Mandatory access control (MAC) is usually built into and applied within the operating system being used. MAC components are present in UNIX, Linux, Microsoft's Windows NT-based operating systems, Open BSD, and others.

Mandatory access control is also known as multilevel security. You categorize all users and resources and assign a security label to the classification. It also compares the user's clearance level with the current classification of the information. It also allows licensed or cleared persons a suitable level of access. Mandatory controls are usually hard-coded, and set on each object or resource individually.

MAC techniques reduce the need for you to maintain ACLs because the access decision logic is built into the classification hierarchy. When set up a MAC policy, clients are not allowed to change permissions or rights associated with objects.

2.1.3: Role-based Access Control (RBAC)

Role-based access control (RBAC) takes the combination of mandatory and discretionary to the next level. Users and processes are identified for access to the information by the role they play within the enterprise. People in the budget department could access and use sensitive budget data, whereas people in other parts of the enterprise would be denied access. RBAC is an alternative to DAC and MAC, giving you the ability to specify and enforce enterprise-specific security policies in a way that maps naturally to an organization's structure. RBAC seeks to group users by common behaviors and access needs.

When creating the appropriate groupings, you have the ability to centralize the function of setting the access levels for various resources within the system. Roles are mapped to a particular resource or a particular user group. When roles are mapped to a resource, the resource name defined in the role is verified and then it is determined if access is permitted to proceed. RBACs allow for a more granular and defined access level, without the generality that exists within the group environment. This type of access control requires more development and cost, but is superior to MAC in that it is flexible and can be redefined more easily.

Section 2.2: Access Control Types

The types of access control provide different levels of protection, each of which can be configured to meet the needs of the organization. This provides the security administrator with a very granular level of control over the security mechanisms, providing the organization with true defense in depth.

The primary purpose of security control mechanisms is to prevent, detect, or recover from problems. The purpose of preventive controls is to inhibit the occurrences harmful attacks; the purpose of detective controls are to discover if harmful incidents have attacks; and the purpose of corrective controls are to restore systems that are victims of harmful attacks.

To implement these measures, controls can be administrative, technical, and physical.

- **Administrative controls** are the policies and procedures implemented by the organization. The security awareness training, strong password policies, pre-employment background checks and work habit checks, and increased supervision are **preventive administrative controls**.
- **Technical controls** are the logical controls organization put in place to protect the IT infrastructure. These include the restriction of access to systems though mechanisms such as strong authentication, network segmentation, and demilitarized zones (DMZs); and the protection of information through encryption and antivirus controls.
- **Physical controls** protect organizations against theft, loss, and unauthorized access. Physical access controls include building security though guards, gates, locks, guard dogs, closed-circuit television (CCTV), and alarms; the securing of server rooms or laptops; the protection of cables, the separation of duties; and the backing up of data and files.

Preventive and detective control types can be combined with the administrative, technical and physical methods of implementation to create the following pairings:

- **Preventive administrative** controls, which focus on the mechanisms that support the access control objectives and includes organizational policies and procedures, pre-employment background checks, employment agreements, employee termination procedures, vacation scheduling, labeling of sensitive materials, increased supervision, security awareness training, behavior awareness, and sign-up procedures to obtain access to information systems and networks.
- **Preventive technical** controls, which use technology to enforce access control policies. These technical controls can be built into the operating system, be software applications, or can be supplemental hardware/software units. Some common preventive/technical controls are protocols, encryption, smart cards, biometrics, local and remote access control software packages, call-back systems, passwords, constrained user interfaces, menus, shells, database views, limited keypads, and virus scanning software.
- **Preventive physical** controls, which are designed to restrict the physical access to areas with systems holding sensitive information or areas that are used for storage of the backup data files. Often, the restricted area is protected by a security perimeter that is under access control. These security perimeters include fences, badges, man-traps, magnetic card entry systems, biometrics, guards, dogs, and environmental control systems. Preventive/physical measures also apply to areas.

Protocols, Encryption and Smart Cards

Protocols, encryption, and smart cards are technical control mechanisms that protect information and passwords from disclosure.

- **Detective administrative** controls, which can be applied for prevention of future security policy violations or to detect existing violations. The mechanisms implemented by this control pairing are organizational policies and procedures, background checks, vacation scheduling, the labeling of sensitive materials, increased supervision, security awareness training, and behavior awareness. Additional detective/administrative controls are job rotation, the sharing of responsibilities, and reviews of audit records.

Biometrics

Biometrics use devices that scan physically unique attributes such as fingerprint, retina, and iris to authenticate individuals requesting access to resources.

Constrained User Interfaces

Constrained user interfaces limit the functions of a software program that can be selected by a user.

- **Detective technical** controls, which use technical mechanisms to detect the violations of security policy. These mechanisms include intrusion detection systems and automatically-generated violation reports from audit trail information. These reports can indicate variations from “normal” operation or detect known signatures of unauthorized access events. Due to the importance of the audit information, audit records should be protected at the highest level of sensitivity in the system.

- **Detective physical** controls, which rely on sensors or cameras to detect a violation. Some of these control types are motion detectors, thermal detectors, and video cameras. However, these devices usually require human intervention to determine if the detected violation is real...

Section 2.3: Identification and Authentication

Identification and authentication are the key elements of an access control system. Identification is performed by the user or service supplying a system with a user ID, which is usually a user name or computer name. Authentication is the process the system uses to verify the identity of the user or service that requests access to an information system. This allows a sender and receiver of information to validate each other as the entities with which they want to communicate. If entities wishing to communicate cannot properly authenticate each other, there can be no trust in the activities or information provided by either party.

Authentication can be based on three factor types:

- Something only you know, such as a PIN number or a password.
- Something only you have, such as a smart card.
- Something biometric, such as a fingerprint or an iris scan.

2.3.1: Passwords

Of the three authentication factor types, the most widely used factor is passwords. However, passwords are the easiest factor to crack, particularly as most users use passwords that are easy to remember, such as a birthday, an anniversary, or a child's name. These types of passwords can be easily guessed. Ideally, password should be used only once. This type of password, which is called a **one-time password,** provides highest possible security because a new password is required for each new log-on. However, users may have problems remembering the password required for the next log-on. They would prefer a password that remains the same for each log-on. This type of password is called a **static password**. The longer a password is used, the greater the probability of it being compromised. Therefore, security administrators require that passwords change frequently. This can be monthly, quarterly, or after a set number of log-ons, depending on the sensitivity of the information requiring protection.

2.3.2: Tokens

Tokens are used in the second authentication factor – something you have. They are regarded more secure than passwords, because it makes impersonation and falsification difficult. These are in the form of credit card-size memory cards, called smart cards, or those resembling small calculators that are used to supply static and dynamic passwords. Smart cards are often used with a personal identification number (PIN) so that only you have control over the token. They also provide additional processing power on the card. These devices generate authentication credentials that are often used as one-time passwords. They can also be used for multifactor authentication.

2.3.3: Biometrics

Biometrics are used in the third authentication factor and is based on a behavioral or physiological characteristic that is unique to an individual user. As such, it provides the most accurate means of authentication, but is also more expensive. Biometric systems work by recording physical information that is

very minute and is unique to the individual user. This can be finger prints, palm prints, voice patterns, retina patterns and iris recognition.

The different biometric systems offer different levels of accuracy, which is measured by the percentage of Type I and Type II errors it produces. **Type I errors**, which is the false rejection rate, are a measurement of the percentage of legitimate users that were denied access by the system, while **Type II errors**, which is the false acceptance rate, are a measurement of the percentage of unauthorized users that were erroneously permitted access.

The point at which the Type I and Type II errors are at the same level indicates the accuracy of the system. This point is known as the **crossover error rate (CER)**. The lower the CER, the better.

2.3.4: Multifactor Authentication

Multifactor authentication is the combination of two or more authentication factors, such as using a token together with a PIN. A user must possess both the token and the PIN to log on successfully. Using multifactor authentication greatly increases the security of the system.

2.3.5: Single Sign-On (SSO)

Most information systems consist of multiple systems, resources and data that users will require access to. Each of these require access control, however, requiring users to pass their credentials every time they require access to a different system, resource or piece of data would be cumbersome. Often users will use the same password, or write down the different user IDs and passwords that they require of each system or resource, creating a greater security risk.

Single sign-on addresses this problem by requiring users to authenticate once to a single authentication authority and then allowing them access to all other protected systems and resources without requiring them to reauthenticate.

This system has two drawbacks – it is expensive and if an attacker can gain entry to the system, that person then has access to all systems resources in the information system. Therefore, the passwords should not be stored or transmitted in the clear-text. Kerberos, SESAME, IBM's KryptoKnight, and NetSP are all authentication server systems with operational modes that can be used to implement single sign-on.

Tickets

A ticket is a block of data that allows users to prove their identity to a service. Limiting the length of time a ticket is valid shrinks the chances of a hacker obtaining a ticket and being able to use it for unauthorized access.

Tickets issued by the Kerberos server provide the credentials required to access additional network resources.

2.3.5.1: Kerberos

Kerberos is the preferred single sign-on authentication protocol in many medium and large information systems. It is a network protocol designed to centralize the authentication information for any entity requesting access to resources. Kerberos accomplishes this by using symmetric key cryptography and assigning tickets to the entity that requests access. When a user attempts to log-on to the local system, a local agent or process sends an authentication request to the Kerberos **ticket-granting server (TGS)**. The TGS returns the encrypted credentials for the user attempting to sign onto the system. The local agent decrypts the credentials using the user-supplied password. If the correct password has been supplied, the user is validated and assigned **authentication tickets**, which allow the user to access other Kerberos-authenticated services. A user is also assigned a set of **cipher keys** that can be used to encrypt all data sessions. All services and users in the realm receive tickets from the TGS and are authenticated by an **authentication server (AS)**. This provides a sole

source of authority to register and authenticate with. Realms can trust one another, providing the capability to scale Kerberos authentication.

Kerberos is implemented by a **Key Distribution Center (KDC)** that contains the information that allows Kerberos clients to authenticate. The information is contained in a database that makes single sign-on possible. The KDC's database does not work in the same manner as many other databases. The KDC and the client establish trust relationships using public key cryptography. The trust relationship allows the KDC to then determine exactly which services a host and user can access.

Kerberos addresses the confidentiality and integrity of information but not availability. All the secret keys are stored on the TGS and authentication is performed on Kerberos TGS and the authentication servers. In addition, a client's secret key and session keys are stored temporarily on the client workstation. These systems are vulnerable to both physical attacks and attacks from malicious code. Furthermore, because a client's password is used in the initiation of the Kerberos request for the service protocol, password guessing can be used to impersonate the client.

2.3.5.2: Secure European System and Applications in a Multivendor Environment (SESAME)

The Secure European System and Applications in a Multivendor Environment (SESAME) project was developed to address some of the weaknesses in Kerberos. SESAME uses public key cryptography for the distribution of secret keys as well as MD5 and CRC32 one-way hash functions and two certificates or tickets. One of the certificates is used to provide authentication, as in Kerberos, and the other is used to control the access privileges assigned to a client.

Like Kerberos, SESAME has its weaknesses. SESAME authenticates using only the first block of a message and not the complete message. It is also subject to password guessing.

2.3.5.3: KryptoKnight and NetSP

IBM's KryptoKnight system was designed to provide authentication, single sign-on, and key distribution services for computers with widely varying computational capabilities. KryptoKnight uses a trusted Key Distribution Center (KDC) that knows the secret key of each party. A key difference between Kerberos and KrytpoKnight is that there is a peer-to-peer relationship among the parties and the KDC, rather than a server/client relationship. To implement single sign-on, the KDC has a party's secret key that is a one-way hash of their password. The initial exchange from the party to the KDC is the user's name and a value, which is a function of a nonce, which is a randomly-generated one-time use authenticator, and the password. The KDC authenticates the user and assigns the user a ticket encrypted with the user's secret key. This ticket is decrypted by the user and can be used for authentication to obtain services from other servers on the system.

NetSP is based on KryptoKnight. It uses a workstation as an authentication server.

Section 2.4: Access Control Systems

Access control requirements are diverse; hence the implementation of access control systems can be just as diverse. Despite this diversity, access control systems can be divided into two categories: centralized access control and decentralized or distributed access control. Depending on the organization's environment and requirements, one system would work better than the other.

2.4.1: Centralized Access Control

Centralized access-control systems stores user IDs, rights, and permissions in a database or a file on a central server. Remote Authentication Dial-In User Service (RADIUS), TACACS, and DIAMETER are common centralized access-control systems.

2.4.1.1: Remote Authentication Dial-In User Service (RADIUS) and DIAMETER

Remote Authentication Dial-In User Service (RADIUS) is a client/server-based system that provides authentication, authorization, and accounting (AAA) services for remote dial-up access while securing the information system against unauthorized access.

RADIUS facilitates centralized user administration by storing all user profiles in one location that all remote services have access to. To authenticate to a RADIUS server, a user must enter his or her credentials, which is encrypted and sent to the RADIUS server in an **Access-Request** packet. On receiving the credentials, the RADIUS server accepts, rejects or challenges the credentials. If the RADIUS server accepts the credentials, it sends an **Access-Accept** packet and the user is authenticated successfully. If the RADIUS server rejects the credentials, it sends an **Access-Reject** packet. If the RADIUS server challenges the credentials, it sends an Access-Challenge packet, which prompts the user to provide additional information which the RADIUS server can use to authenticate the user.

For remote dial-up access, RADIUS also provides callback security, in which the RADIUS server terminates the connection and establishes a new connection to the user by dialing a predefined telephone number to which the user's modem is attached. This provides an additional layer of protection against unauthorized access over dial-up connections.

Due to the success of RADIUS, an enhanced version of RADIUS called DIAMETER was developed. DIAMETER is designed for use on all forms of remote connectivity and not just dial-up connectivity.

2.4.1.2: Terminal Access Controller Access Control System

There are three versions of Terminal Access Controller Access Control System (TACACS): TACACS, Extended TACACS (XTACACS), and TACACS+. All three versions authenticate users and deny access to users who do not have a valid username/password pairing. TACACS integrates the authentication and authorization functions. XTACACS allows the division of the authentication, authorization, and auditing functions, providing the administrator more control over its deployment. TACACS+ also allows the division of the authentication, authorization, and auditing but also provides two-factor authentication.

The TACACS authentication process is similar to that of RADIUS and it provides the same functionality as RADIUS. However, RADIUS is based on an Internet standard, whereas TACACS is a proprietary protocol. For this reason, TACACS has failed to gain the popularity of RADIUS.

2.4.2: Decentralized/Distributed Access Control

A decentralized access-control system store user IDs, rights, and permissions in different locations throughout the network. These locations are often servers on subnets that are closer to the user requesting access and make use of linked or relational databases.

Section 2.5: Threats Against Access Control

Attackers use a variety of tools and techniques to try to bypass or crack access control mechanisms, making access control one of the most targeted security mechanisms.

2.5.1: Password Attacks

Access control on most systems is accomplished by means of user name and passwords. However, most users do not practice good password security. Attackers are well aware of this and use the information to launch common password attacks. Attackers typically use one of two methods to crack passwords: a dictionary attack or a brute-force attack.

2.5.1.1: Dictionary Attacks

A dictionary attack uses a predefined dictionary file that a program will cycle through to find a match with the user's password. Usually, passwords are stored in a hashed format. Most password-cracking programs use a technique called comparative analysis in which all common permutations of each word in the dictionary file is hashed and compared to the encrypted password. If a match is obtained, the password is cracked. Thus, if a user's password is well-known, dictionary-based words, dictionary tools will crack them quickly.

2.5.1.2: Brute-Force Attacks

A brute force attack systematically attempts every possible combination of letters, numbers, and symbols in an attempt to discover passwords for user accounts. With the speed of modern computers and with distributed computing, brute force attacks are proving successful, even against strong passwords. The longer the password, the more secure against brute force attacks it becomes as a greater period of time is required for a force brute attack to crack it. However, most passwords of 14 characters or less can be cracked within 7 days.

A variation of the traditional brute-force attack uses a rainbow table. In this variation, all possible passwords are pre-computed in advance of the attack. When this time-consuming process is complete, the passwords and their corresponding encrypted values are stored in a file called the rainbow table. An encrypted password can then be compared to the values stored in the rainbow table and cracked within a few seconds.

2.5.2: Back Door Attacks

A back door attack allows an attacker access to a system from a different interface or with different credentials. Programmers often put back doors in place to allow them to debug and change code during test deployments of the software under development. Another type of back door can be put in place on a current production system by malicious code, allowing unregulated access to systems or services. Some of the common software used as a back door includes:

- Virtual Network Computing (VNC)
- Back Orifice
- NetBus
- Sub7 (or SubSeven)
- PC Anywhere
- Terminal Services

The malicious code can also be hidden in another application. When malicious code is hidden inside another application, it is called a Trojan horse.

2.5.3: Spoofing

Spoofing usually involves the modification of a packet at the TCP level. The attacker sends a packet with an IP source address of a known and trusted host to a target host, thus impersonating the trusted host. The attacker can also impersonate services such as Web, FTP, and email.

2.5.4: Man-in-the-Middle Attacks

A man-in-the-middle attack is usually used to gather information that is transmitted between two hosts. In this attack the attacker is able to gain a position between the two hosts while remaining invisible to them. This is usually accomplished by the attacker altering routing information and DNS values, steal IP addresses, or defraud ARP lookups to impersonate the two legitimate hosts.

The man-in-the-middle attack allows the attacker to intercept logon credentials or sensitive data that is being transmitted, as well as to change the data before passing it on to the intended destination host. To safeguard against this type of attack, you need to protect your DNS by restricting access to its records and name-caching system.

2.5.5: Replay Attacks

A replay attack, which is also referred to as a playback attack, is similar to the man-in-the-middle attack. In this attack, the attacker records the traffic between a client and server and then retransmits the packets to the server with slight variations of the time stamp and source IP address. This may allow the attacker to restart the previous communication link with the server, allowing the attacker the opportunity to obtain data or additional access. To safeguard against this type of attack, you need to use complex sequencing rules and time stamps to prevent retransmitted packets from being accepted as valid.

2.5.6: Denial of Service (DoS) and Distributed Denial of Service (DDoS) Attacks

The purpose for Denial-of-service (DoS) and Distributed DoS (DDoS) attacks is to consume resources to the point the use of resources or services is not possible. This attack can be executed with out the need to identify a legitimate user. Utilities that test network connectivity and host availability are generally used to perform these forms of attacks.

Types of DoS and DDoS attacks include:

- **Smurf**, which is based on ping and the Internet Control Message Protocol (ICMP) echo reply function. This attack consists of three sites - the source site, the bounce site, and the target site. An attacker (the source site) will send a modified or spoofed ping packet to the broadcast address of a network (the bounce site). The modified ping packet will contain the address of the target site. This causes the bounce site to broadcast the misinformation to all of the devices on its local network. All of these devices now respond with a reply to the target system, which will then be saturated with these replies.
- **Buffer Overflow**, in which a process receives much more data than it can handle. If that process has no routine to deal with the excessive amount of data, it acts in an unexpected way that an attacker can exploit.

- **Ping of death**, which is a type of buffer overflow attack. In this attach an attacker exploits ICMP by sending an illegal ECHO packet of more than 65K octets of data, which can cause an overflow of system variables and lead to a system crash.
- **Teardrop attack**, which targets the UDP in the operating system. The attacker modifies length and fragmentation offset fields in sequential UDP packets and sends them to a system. When the system tries to rebuilds the packets from the fragments, the fragments overwrite each other as the system receives contradictory instructions on how the fragments are offset on these packets. This causes the target system to crash.
- **SYN attack**, in which an attacker exploits the use of the buffer space during a three-way Transmission Control Protocol (TCP) session initialization handshake. A source host sends a TCP SYN request when it requires connection session with a destination host. Usually the destination host will respond with an acknowledgement (ACK) and returns a SYN response. The source host usually sends a final ACK packet, but in this attack the attacker sends a flood of SYN requests and never sends the final ACK. This causes the target system to time out while waiting for the proper response, which makes the system crash or become unusable.

2.5.7: TCP Hijacking

In a TCP hijacking attack, the attacker hijacks a session between a trusted client and network server form the trusted client by substitutes its IP address for that of the trusted client. When a session is hijacked the attacker can create a new back door account, or can view files and use services to which the legitimate host would normally have access. This attack usually occurs after the trusted client has connected to the network server.

2.5.8: Social Engineering

Social engineering is not an active attack against a system but is one of the easiest methods an attacker can use to acquire the information needed to compromise information systems. It is also the most difficult to defend against. This attack uses social skills to obtain information such as passwords or PIN numbers that can be used to gain access to a system. This can be accomplished by an attacker impersonating someone in an organization and make phone calls to employees of that organization requesting passwords for use in maintenance operations, an attacker befriending an employee and soliciting them for information or physical access to the system, etc.

There is not real defense against social engineering, except awareness training and the creation and dissemination of security policies regarding disclosure of information.

2.5.9: Dumpster Diving

Dumpster diving is also not an active attack against a system. Instead, the attacker attempts to acquire information that has been discarded by a user or by an organization. Often the information found in the discarded trash can be of great value to the attacker as the discarded information may include technical manuals, password lists, and organization charts. It may also include telephone umbers and user names, which the attacker can use in social engineering attacks.

2.5.10: Software Exploitation

Software exploitation is not a single attack but the exploitation of the weaknesses in the code of a software program. This type of attack is most often directed at the vulnerabilities in various operating systems. These

vulnerabilities can often be exploited by an attacker to gain access to information systems, resources and data.

Some examples of software exploitation are:

- An attacker can cause a DoS buffer overflow attack against a Novell Web Server by sending a large GET request to the remote administration port. This causes the data being sent to overflow the storage buffer and reside in memory as executable code.
- An attacker can discover passwords in a system running the AIX operating system by using diagnostic commands.
- An attacker to gain root access to a system running the IRIX operating system by exploiting the operating system's buffer overflow vulnerability.
- An attacker to locate system and screensaver passwords on systems running Windows 9x, thereby gaining unauthorized logon access to that system.
- An attacker can use privilege exploitation software to gain administrative access to systems running the Windows NT operating system

Section 2.6: Monitoring and Intrusion Detection

2.6.1: Monitoring

Monitoring is both the process of holding authenticated users accountable for their actions while on a system, as well as the process of detecting unauthorized or abnormal activities on a system and system failures.

Accountability is maintained by logging the activities of users and system services that maintain the operating environment and the security mechanisms. These logs can help reconstruct events, provide evidence for prosecution, and produce problem reports and analysis. The process of analyzing logs is called **auditing** and is usually native features of an operating system.

The **audit trails** created by logging system events can be used to evaluate a system's health and performance. System crashes may indicate faulty programs, corrupt drivers, or intrusion attempts. The event logs leading up to a crash can often be used to discover the reason a system failed.

2.6.2: Intrusion Detection System (IDS)

An Intrusion Detection System (IDS) is a technical detective access control system designed to continuously monitor the network activities and to detect any scanning and probing activities, or patterns that appear to be attempts at unauthorized access to the information system in real-time. It can also be configured to scan for attacks, track an attacker's movements, alert an administrator to an ongoing attack, test the system for possible vulnerabilities, and can be configured to put preventative measures into place to stop any additional unauthorized access. IDSs can also be used to detect system failures and system performance. Attacks detected by an IDS can come from external connections, viruses, malicious code, trusted internal users attempting to perform unauthorized activities, and unauthorized access attempts from trusted locations.

IDS systems can be divided into two broad categories: network-based intrusion-detection systems (NIDS) and host-based intrusion-detection systems (HIDS), based on their implementation. They also use two different methods that an IDS uses to detect malicious events. These are: knowledge-based, or signature-based detection and behavior-based detection. These two methods take different approaches to detecting intrusions.

2.6.2.1: Host-Based IDS (HIDS)
Host-Based IDS (HIDS) is installed on individual systems and is meant to protect that particular system. They are more closely related to a virus scanner in their function and design and are generally more reliable than NIDS's in detecting attacks on individual systems because HIDS can examine events in much greater detail than a network-based IDS can. HIDS typically use operating system audit trails and system logs. Operating system audit trails are quite reliable for tracking system events as monitor traffic and attempt to detect suspect activity. Suspect activity can range from attempted system file modification to unsafe activation of ActiveX commands. However, HIDS have difficulty with detecting and tracking denial of service (DoS) attacks, especially those that consume bandwidth. In addition, HIDS consumes system resources from the computer being monitored, and thus reduces the performance of that system.

2.6.2.2: Network-Based IDS (NIDS)
A network-based IDS (NIDS) captures and evaluates network traffic, inspecting each network packet as it passes through the system. A single NIDS can monitoring a large network if installed on a backbone of that network or can protect multiple hosts on a single network segment. As such, an NIDS provides an extra layer of defense between the firewall and the hosts. However, an NIDS may not be able to keep up with the flow of data on networks with extremely large volumes of traffic. This could result in the NIDS missing an attack that occurred during high traffic levels. In addition, NIDSs do not work well on switched networks, particularly if the routers do not have a monitoring port.

NIDS implementations often rely on the placement of sensors at various points on a network. The sensors are usually bastion hosts that run only the IDS sensor software. This allows the bastion host to be hardened against attack, reduces the number of vulnerabilities to the NIDS, and allows the NIDS to operate in stealth mode, which means that NIDS is invisible to the network. Thus, an attacker would need to know of the exact location and system identification of the NIDS to discover it. Furthermore, an NIDS has little impact on the overall performance of the network.

2.6.2.3: Knowledge-Based IDS
Knowledge-based IDS, which is also referred to as **signature-based IDS**, **pattern-matching IDS**, or **rule-based IDS**, rely on a signature database of known attacks and attack patterns. This system examines data and attempts to match the monitored data to a signature pattern in the signature database. If an event does match a signature, the IDS assumes that an attack is occurring or has occurred and will respond appropriately. This might include an alarm, an alert, or a change to the firewall configuration.

The main disadvantage of a knowledge-based IDS is that it is only effective against known attack methods. New attacks or slightly modified versions of known attacks are often missed by the IDS. Thus, the knowledge-based IDS is only as effective as its signature database. Hence, the signature database must be kept as current as possible.

2.6.2.4: Behavior-based IDS
A behavior-based IDS, which is also referred to as a **statistical intrusion IDS**, **profile-based IDS** (anomaly detection) , and **heuristics-based IDS**, observers normal activities and events on the system or network and detects abnormal activities or events which are deemed possible malicious activities or events. This allows behavior-based IDS to detect new and unknown vulnerabilities, attacks, and intrusion methods.

However, behavior-based IDS also produce many **false positives** or false alarms because patterns of normal activities and events can vary widely from day to day. This is the main disadvantage of behavior-based IDS. The more it produces false positives, the less likely security administrators will respond to the raised alarm.

2.6.3: Honeypots

A honeypot is similar to an IDS in that it also detects intrusion attempts. However, a honeypot is also a means of luring an intruder away from vulnerable systems by posing as a valuable system. Honeypots usually contain fake files, services, and databases to attract and entrap a intruder. This makes the honeypot an ideal system for IDS monitoring.

Section 2.7: Penetration Testing

Penetration testing, which is also referred to as **ethical hacking**, is the process to testing a systems defense against attacks, or to perform a detailed analysis of the information system's weaknesses. A penetration test can also be used to determine how a system would react to an attack and what information can be acquired from the system. There are three types of penetration tests:

- **Full knowledge test**, in which the penetration testing team has as much knowledge as possible about the information system to be tested. This simulates the type of attack that might be attempted by a knowledgeable employee of an organization.
- **Partial knowledge test**, in which the penetration testing team has knowledge that might be relevant to a specific type of attack.
- **Zero knowledge test**, in which the penetration testing team is provided with no information about the system and must gather the information on its own and as part of the test. This simulates an attack by a hacker that has no prior knowledge of the information system.

Penetration testing can also be described as open-box testing or closed-box testing. In **open-box penetration testing**, the penetration testing team has access to internal system code. This simulates possible scenarios in attacks against general-purpose operating systems such as UNIX or Linux. In **closed-box penetration testing**, the penetration testing team does not have access to internal code. This type of testing is applied to specialized systems that do not execute user code.

Other types of tests should be considered beyond basic penetration testing. These include:

- **Application security testing** for organizations that offer access to core business functionality through web-based applications. Application security testing evaluates the controls over the application and its process flow.
- **Denial-of-service (DoS) testing**, which evaluates a networks susceptibility to DoS attacks.
- **War dialing**, which is an attempt to systematically call a range of telephone numbers to identify modems, remote-access devices, and maintenance connections of computers that could exist on an organization's network.
- **Wireless network testing**, which evaluates the controls over an organization's wireless access policies and ensure that no misconfigured devices have been introduced that have caused additional security exposures.
- **Social engineering testing**, which refers to techniques using social interaction, typically with the organization's employees, suppliers, and contractors, to gather information and penetrate the organization's systems.

Topic 3: Telecommunications and Network Security

Section 3.1: OSI Reference Model

The Open Systems Interconnection (OSI) reference model was developed by the International Standards Organization (ISO) in 1984 to define network communications, and describe the flow of data on a network. The OSI reference model consists of seven layers that start with the physical connection and end with the application. As illustrated in Figure 3.1, the seven OSI layers are:

- The Physical Layer
- The Data-Link Layer
- The Network Layer
- The Transport Layer
- The Session Layer
- The Presentation Layer
- The Application Layer

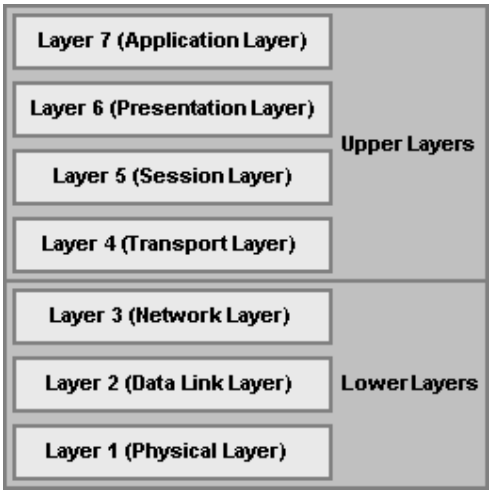

Figure 3.1: The OSI Reference Model

Each layer performs particular functions and can have one ore more sublayers. The upper layers of the OSI reference model define functions focused on the application, while the lower three layers define functions to facilitate the transport and delivery of data from the source to the destination.

The upper layers of the OSI reference model, i.e., the Application Layer, the Presentation Layer, and the Session Layer, define functions focused on the application. The lower four layers, i.e., the Transport Layer, the Network Layer, the Data-Link Layer, and the Physical Layer, define functions focused on end-to-end delivery of the data.

- The **Physical Layer** deals with the physical characteristics of the transmission medium, such as signaling specifications, cable types and interfaces. It also describes voltage levels, physical data rates, and maximum transmission distances. In short, the physical layer deals with the electrical, mechanical, functional, and procedural specifications for the physical links between networked systems, and includes connectors, pins, use of pins, electrical currents, encoding, and light modulation are all part of different physical layer specifications.
- The **Data-Link Layer** deals with the reliable transport of data across one particular link or medium. Data at the data link layer is encapsulated into frames. Data-link specifications deal with the sequencing of frames, flow control, synchronization, error notification, physical network topology, and physical addressing. At this layer, data is converted from frames into bits when it is sent across the physical media and is converted back into frames when it is received from the physical media. **Bridges** and **switches** operate at the data-link layer.

  In the IEEE 802, the data link layer is subdivided into two sublayers, the **Logical Link Control (LLC)** sublayer and the **Media Access Control (MAC)** sublayer. The upper sublayer is called the **Logical Link Control (LLC)** sublayer and manages the communications between devices. The lower layer is called the **Media Access Control (MAC)** sublayer and manages protocol access to the physical media.
- The **Network Layer** deals with the routing of data, which is called packets at this layer, and defines the methods to facilitate this, including how routing works, logical addressing, how routes are learned; as

well as how packets are fragmented into smaller packets to accommodate media with smaller maximum transmission unit (MTU) sizes.

Note: As you would expect, routers operate at this layer but they also perform Data-Link Layer functions. Thus a router is best described as a Layer 2/3 device.

Note: The functions of the IP and IPX protocols most closely match the OSI Network Layer (Layer 3) and are therefore called Layer 3 protocols.

- The **Transport Layer** performs several functions, including the choice of protocols. This layer provides reliable, transparent transport of data segments from upper layers. The most important Layer 4 functions are **error recovery** (retransmission) and **flow control** to prevent unnecessary congestion by attempting to send data at a rate that the network can accommodate, depending on the choice of protocols. Multiplexing of incoming data for different flows to applications on the same host is also performed. Messages are assigned a sequence number at the transmission end. At the receiving end the packets are reassembled, checked for errors, and acknowledged. Reordering of the incoming message when packets arrive out of order is also performed at this layer.
- The **Session Layer** defines how to start, control, and end communication sessions between applications. Communication sessions consist of service requests and responses that occur between applications on different devices. This includes the control and management of multiple bidirectional messages so that the application can be notified if only some of a series of messages are completed. This allows the presentation layer to have a seamless view of an incoming stream of data. The management of sessions also involves the synchronization of dialog control by using checkpoints in the data stream.
- The **Presentation Layer** ensures that data sent from a sending application on the source system is readable by the application layer on the destination system by providing data representation with a variety of coding and conversion functions. This includes defining the conversion of character representation formats, such as ASCII text, EBCDIC text, binary, BCD, and JPEG; data compression schemes; and encryption schemes. Voice coding schemes are specified at this layer.
- The **Application Layer** provides to network communication services to the end user or operating system. It interacts with software applications by identifying communication resources, determining network availability, and distributing information services. It also provides synchronization between the peer applications that reside on separate systems.

3.1.1: Inter-OSI Layer Interaction

When a host receives a data transmission from another host on the network, that data is processed at each of the OSI layers to the next higher layer, in order to render the data transmission useful to the end-user. To facilitate this processing, headers and trailers are created by the sending host's software or hardware, that are placed before or after the data given to the next higher layer. Thus, each layer has a header and trailer, typically in each data packet that comprises the data flow. The sequence of processing at each OSI layer, i.e., the processing between adjacent OSI layers, is as follows:

- The **Physical Layer** (Layer 1) ensures **bit synchronization** and places the received binary pattern into a buffer. It notifies the Data Link Layer (Layer 2) that a frame has been received after decoding the incoming signal into a bit stream. Thus, Layer 1 provides delivery of a stream of bits across the medium.
- The **Data-Link Layer** (Layer 2) examines the **frame check sequence (FCS)** in the trailer to determine whether errors occurred in transmission, providing **error detection**. If an error has occurred, the frame is

discarded. The current host examines data link address is examined to determine if the data is addressed to it or whether to process the data further. If the data is addressed to the host, the data between the Layer 2 header and trailer is handed over to the Network Layer (Layer 3) software. Thus, the data link layer delivers data across the link.

- The **Network Layer** (Layer 3) examines the destination address. If the address is the current host's address, processing continues and the data after the Layer 3 header is handed over to the Transport Layer (Layer 4) software. Thus, Layer 3 provides end-to-end delivery.
- If error recovery was an option chosen for the **Transport Layer** (Layer 4), the counters identifying this piece of data are encoded in the Layer 4 header along with acknowledgment information, which is called **error recovery**. After error recovery and reordering of the incoming data, the data is given to the Session Layer (Layer 5).
- The **Session Layer** (Layer 5) ensures that a series of messages is completed. The Layer 5 header includes fields signifying sequence of the packet in the data stream, indicating the position of the data packet in the flow. After the session layer ensures that all flows are completed, it passes the data after the Layer 5 header to the Presentation Layer (Layer 6) software.
- The **Presentation Layer** (Layer 6) defines and manipulates the data format of the data transmission. It converts the data to the proper format specified in the Layer 6 header. Typically, this header is included only for initialization flows, not with every data packet being transmitted. After the data formats have been converted, the data after the Layer 6 header is passed to the Application Layer (Layer 7) software.
- The **Application Layer** (Layer 7) processes the final header and examines the end-user data. This header signifies agreement to operating parameters by the applications on the two hosts. The headers are used to signal the values for all parameters; therefore, the header typically is sent and received at application initialization time only.

In addition to processing between adjacent OSI layers, the various layers must also interact with the same layer on another computer to successfully implement its functions. To interact with the same layer on another computer, each layer defines additional data bits in the header and, in some cases, trailer that is created by the sending host's software or hardware. The layer on the receiving host interprets the headers and trailers created by the corresponding layer on the sending host to determine how that layer's processing is being defined, and how to interact within that framework.

Section 3.2: Transmission Control Protocol/Internet Protocol (TCP/IP)

As illustrated in Figure 3.2, the TCP/IP architecture consists of four layers, each of which can have several sublayers. These layers correlate roughly to layers in the OSI reference model and define similar functions. Some of the TCP/IP layers correspond directly to layers in the OSI reference model while others span several OSI layers. The four TCP/IP layers are:

- The **TCP/IP Application Layer** refers to communications services to applications and is the interface between the network and the application. It is also responsible for presentation and controlling

- communication sessions. It spans the Application Layer, Presentation Layer and Session Layer of the OSI reference model. Examples include: HTTP, POP3, and SNMP.
- The **TCP/IP Transport Layer** defines several functions, including the choice of protocols, error recovery and flow control. The transport layer may provide for retransmission, i.e., error recovery, and may use flow control to prevent unnecessary congestion by attempting to send data at a rate that the network can accommodate, or it might not, depending on the choice of protocols. Multiplexing of incoming data for different flows to applications on the same host is also performed. Reordering of the incoming data stream when packets arrive out of order is included. It correlates with the Transport Layer of the OSI reference model. Examples include: TCP and UDP, which are called **Transport Layer**, or **Layer 4**, protocols. TCP provides connection oriented service while UDP provides connectionless service in the Transport Layer.

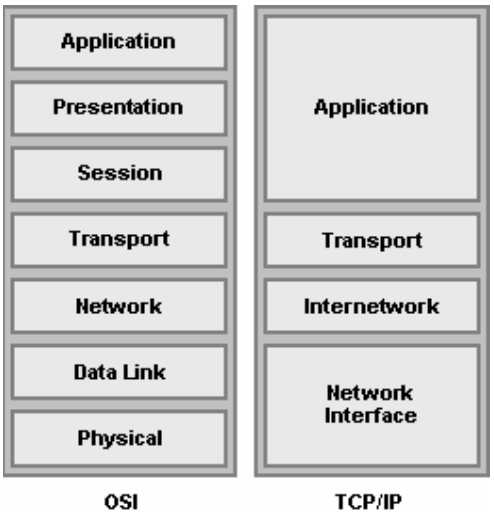

Figure 3.2: OSI and TCP/IP

- The **TCP/IP Internetwork Layer**, or **Internet Layer**, defines end-to-end delivery of packets and defines logical addressing to accomplish this. It also defines how routing works and how routes are learned; and how to fragment a packet into smaller packets to accommodate media with smaller maximum transmission unit sizes. It correlates with the Network Layer of the OSI reference model. However, while the OSI network-layer protocols provide connection-oriented (Connection-Mode Network Service (CMNS), X.25) or Connectionless Network Service (CLNS), IP provides only **connectionless** network service. The routing protocols are network layer protocols with an IP protocol number. One exception is the Border Gateway Protocol (BGP), which uses a TCP port number. Another is Intermediate System-to-Intermediate System (IS-IS), which resides over the data-link layer.
- The **TCP/IP Network Interface Layer** is concerned with the physical characteristics of the transmission medium as well as getting data across one particular link or medium. This layer defines delivery across an individual link as well as the physical layer specifications. It spans the Data Link Layer and Physical Layer of the OSI reference model. Examples include: Ethernet and Frame Relay.

Note: TCP/IP's architecture does not have a Presentation Layer and a Session Layer. Therefore, the Application Layer protocols use the Transport Layer services directly.

3.2.1: TCP/IP Protocols

Transmission Control Protocol/Internet Protocol (TCP/IP) consists of a suite of protocols that were originally developed by the U.S. Department of Defense (DoD) in the 1970s to support the construction of the Internet. The protocols included in this suite are: Transmission Control Protocol (TCP), User Datagram Protocol (UDP), Internet Protocol (IP), Address Resolution Protocol (ARP), Reverse Address Resolution Protocol (RARP) and Internet Control Message Protocol (ICMP).

- **Transmission Control Protocol (TCP)** is the most widely used protocol and accounts for the most of the traffic on a TCP/IP network. It is a connection-oriented protocol that provides a full-duplex. TCP ensures that data delivery across any IP link by using mechanisms such as connection startup, flow control, slow start, and acknowledgments. Incoming TCP packets are sequenced to match the original transmission sequence numbers. Because any lost or damaged packets are retransmitted, TCP is very costly in terms of network overhead.

- **User Datagram Protocol (UDP)** is similar to TCP is connectionless, which means that it does not create a virtual circuit and does not contact the destination before delivering the data. Furthermore, UDP does not offer error correction, does not sequence the packet segments, and does not care in which order the packet segments arrive at their destination. Consequently, it is referred to as an unreliable protocol. Because of this, UDP has much less overhead which makes it a better choice for applications, such as streaming video or audio, that are not adversely affected by occasional packet loss. Both TCP and UDP use port numbers to communicate with the upper layers.
- **Internet Protocol (IP)** is a widely used network layer protocol that uses unique IP addresses to identify or define each distinct host or end system on a network. The unique IP address allows communication between hosts on an IP network. Each IP packet contains the IP address of the sender, called the **source IP address**, and the IP address of the recipient, called the **destination IP address**. Intermediate devices between the sender and the recipient makes routing decisions based upon the packet's destination IP address
- **Address Resolution Protocol (ARP)** maps the destination IP address to the physical hardware address, called the **MAC address**, of the recipient host. An ARP request that holds the destination IP address, the sender's IP address and MAC address is transmitted to each host within a subnet when the destination MAC address is not listed in the ARP table. When a device receives the ARP request, and it has the IP address, it passes the associated MAC address to the sender of the ARP request.
- **Reverse Address Resolution Protocol (RARP)**maps the MAC address to the IP address. When the MAC address is known but not he IP address, the RARP protocol sends out a packet that includes its MAC address along with a request to be informed of which IP address should be assigned to that MAC address. A RARP server responds with the correct IP address.
- **Internet Control Message Protocol (ICMP)** is a management protocol and messaging service provider for IP. It reports errors and provides other information relevant to IP packet processing, such as informing hosts of a better route to a destination if there is trouble with an existing route, and it can help identify the problem with a route. **`PING`** is an ICMP utility that can be used to test physical connectivity between two hosts on a network.
- **Telnet** is an application layer protocol. It allows a user that is running a Telnet client program to connect to a remote Telnet system. The TCP destination port number is 23 and is commonly used to manage routers and switches. However, Telnet is an insecure protocol, because data flows in plain text and the Telnet passwords can be sniffed. SSH is more secure for remote logins.
- **File Transfer Protocol (FTP)** is a popular file transfer protocols used in TCP/IP networks. FTP is TCP-based. When an FTP client attempts to connect to an FTP server, a TCP connection is established to the FTP server's well-known **port 21**. The data is transferred over a separate FTP data connection, another TCP connection, established to well-known **port 20**. This prevents a file transfer from impacting on the control connection.
- **Trivial File Transfer Protocol (TFTP)** is a more basic file transfer protocol that use a small set of features, takes little memory to load, and little time to program. TFTP uses User Datagram Protocol (UDP), so there is no connection establishment and no error recovery by the transport layer. However, TFTP uses application layer recovery by embedding a small header between the UDP header and the data.
- **Simple Network Management Protocol (SNMP)** is an application layer protocol for the management of IP devices. SNMP allows network administrators to inspect or change parameters on a device remotely, and manage network performance over a period of time. There are three versions of SNMP: SNMP version 1 (SNMPv1); SNMP version 2 (SNMPv2); and SNMP version 3 (SNMPv3)

- **Simple Mail Transfer Protocol (SMTP)** is used to provide e-mail services to IP devices over the Internet. Typically, two mail servers use SMTP to exchange e-mail. After the e-mail is exchanged, the users can retrieve their mail from the mail server and read it. This can be done using any mail client, which use different protocols, such as **Post Office Protocol 3 (POP3)**, to connect to the server. SMTP uses well-known ports TCP port 25 and UDP port 25, although SMTP applications use only TCP port 25.
- **BOOTP** is a protocol that allows a booting host to configure itself by dynamically obtaining its IP address, IP gateway, and other information from a remote server. This allows you to use a single server to centrally manage numerous network hosts without having to configure each host independently.

Section 3.3: Communication and Network Security

A network is defined as a group of two or more computers linked together for the purpose of communicating and sharing information and other resources, such as printers and applications. Networks are constructed around a cable connection or a wireless connection that use radio wave or infrared signals that links the computers. For a network to function it must provide connections, communications, and services.

- **Connections** are defined by the hardware or physical components that are required to connect a computer to the network. This includes the **network medium**, which refers to the hardware that physically connects one computer to another, i.e., the network cable or a wireless connection; and the **network interface**, which refers to the hardware that attaches a computer to the network medium and is usually a network interface card (NIC).
- **Communications** refers to the network protocols that are used to establish the rules governing network communication between the networked computers. Network protocols allow computers running different operating systems and software to communicate with each.
- **Services** define the resources, such as files or printers, that a computer shares with the rest of the networked computers.

3.3.1: Types of Networks

These network definitions can be divided into two types of networks, based on how information is stored on the network, how network security is handled, and how the computers on the network interact. These two types are: **Peer-To-Peer (P2P) Networks** and **Server/Client Networks**. The latter is often also called Server networks.

- On a **Peer-To-Peer (P2P) Network**, there is no hierarchy of computers; instead each computer acts as either a server which shares its data or services with other computers, or as a client which uses data or services on another computer. Furthermore, each user establishes the security on their own computers and determines which of their resources are made available to other users. These networks are typically limited to between 15 and 20 computers. Microsoft Windows for Workgroups, Windows 95, Windows 98, Windows ME, Windows NT Workstation, Windows 2000, Novell's NetWare, UNIX, and Linux are some operating systems that support peer-to-peer networking.

- A **Server/Client Network** consists of one or more dedicated computers configured as servers. This server manages access to all shared files and peripherals. The server runs the network operating system (NOS) manages security and administers access to resources. The client computers or workstations connect to the network and use the available resources. Among the most common network operating systems are Microsoft's Windows NT Server 4, Windows 2000 Server, and Novell's NetWare. Before the release of Windows NT, most dedicated servers worked only as hosts. Windows NT allows these servers to operate as an individual workstation as well.

3.3.2: Network Topologies

The layout of a LAN design is called its topology. There are four basic types of topologies: the star topology, the bus topology, the ring topology, and the mesh topology. Hybrid combinations of these topologies also exist.

- In the **star topology**, all computers and devices are connected to a centrally located hub or switch. The hub or switch collects and distributes the flow of data within the network. This is the most predominant network type and is based on the Ethernet standard.

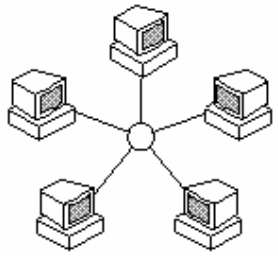

Figure 3.3: The Star Topology

- In the **bus topology**, all computers and devices are connected in series to a single linear cable called a trunk. The **trunk** is also known as a **backbone** or a segment. Both ends of the trunk must be terminated to stop the signal from bouncing back up the cable.

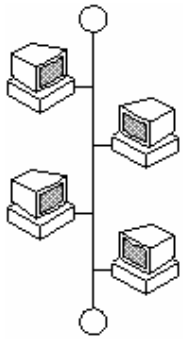

Figure 3.4: The Bus Topology

- In a **ring topology**, all computers and devices are connected to cable that forms a closed loop. On such networks there are no terminating ends; therefore, if one computer fails, the entire network will go down. Each computer on such a network acts like a repeater and boosts the signal before sending it to the next station. This type of network transmits data by passing a "**token**" around the network. If the token is free of data, a computer waiting to send data grabs it, attaches the data and the electronic address to the token, and sends it on its way. When the token reaches its destination computer, the data is removed and the token is sent on. Hence this type of network is commonly called a **token ring** network.

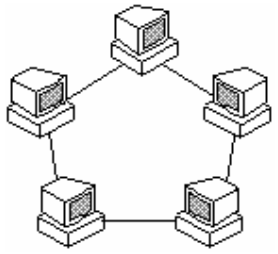

Figure 3.5: The Ring Topology

- In a **mesh topology**, all computers and devices are connected with many redundant interconnections between network nodes. There are two types of mesh topologies: **full mesh** and **partial mesh**.
  - In a **full mesh topology** every computer or device has a link connecting it to every other computer or device in a network. Full mesh is very expensive to implement but yields the greatest amount of redundancy, so in the event that one of those nodes fails, network traffic can be directed to any of the other nodes. Full mesh is usually reserved for **backbone** networks.
  - In a **partial mesh topology** some devices are organized in a full mesh scheme while others are only connected to one or two other devices in the network. Partial mesh topology is less expensive to implement and yields less redundancy than full mesh topology. Partial mesh is commonly found in peripheral networks connected to a full meshed backbone.

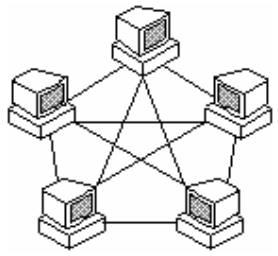

Figure 3.6: The Mesh Topology

3.3.3: Network Cabling

There are three types of cable used to build networks: coaxial, twisted pair, and fiber optic. Coaxial and twisted pair cables are copper-based, and fiber optic cables use glass or plastic conductors. There are also cable-free networks called wireless networks which use radio signals and microwave frequencies.

3.3.3.1: Coaxial Cable

Coaxial cable contains two conductors within the sheath. It has one conductor inside the other. At the center of the cable is the copper core, which actually carries the electrical signals. The core is solid copper or composed of braided strands of copper. Surrounding the core is a layer of insulation, and surrounding that is the second conductor, which is made of braided copper mesh. This second conductor functions as the cable's ground. The entire assembly is encased in an insulating sheath made of PVC or Teflon. There are two types of coaxial cable:

3.3.3.1.1: Thick Ethernet

Ethernet cable is known as Thick Ethernet cable. It is also called 10Base5 and is graded as RG-8. With Thick Ethernet, a station attaches to the main cable via a vampire tap, which clamps onto the cable. A vampire tap is named because a metal tooth sinks into the cable. The tap is connected to an external transceiver that, in turn, has a 15-pin AUI connector to which you attach a cable that connects to the station. DIX got its name from the companies that worked on this format—Digital, Intel, and Xerox. The other option that is found occasionally is the N-series connector. The N connector comes in a male/female screw-and-barrel configuration. A CB radio uses the PL-259 connector, and the N connector looks similar

3.3.3.1.2: Thin Ethernet

Thin Ethernet, also referred to, as Thinnet or 10Base2, is a thin coaxial cable. It is as thick as the coaxial cable except that the diameter of the cable is smaller. Thin Ethernet coaxial cable is RG-58. With Thinnet cable, you use BNC connectors to attach stations to the network. The BNC connector locks securely with a quarter-twist motion.

Table 3.1: Coaxial Cable Specifications

RG Rating	Popular Name	Ethernet Implementation	Cable Type
RG-58 U	None	None	Solid copper
RG-58 AU	Thinnet	10Base2	Stranded copper
RG-8	Thicknet	10Base5	Solid copper
RG-62	ARCnet	None	Solid/stranded

3.3.3.2: Twisted Pair Cable

Twisted pair cable wired in a star topology is used in LANs today. These consist of shielded twisted pair (STP) cable and unshielded twisted pair (UTP) cable. Most of the LANs use UTP cable. UTP cable contains eight separate conductors, as opposed to the two used in coaxial cable. Each conductor is a separate insulated wire, and the eight wires are arranged in four pairs of twisted conductors. The twists prevent the signals on the different wire pairs from interfering with each other, and offer resistance to outside interference. The connectors used for twisted pair cables are called RJ-45; they are the same as the connectors used on standard telephone cables, except that they have eight electrical contacts instead of four.

Twisted pair cable has been used for telephone installations, and its adaptation to LAN use is recent. Twisted pair cable has replaced coaxial cable in the data networking world. Twisted pair cable contains eight separate wires; the cable is more flexible than the more solidly constructed coaxial cable. This makes it easier to flex, which simplifies installation. An advantage is that there are thousands of qualified telephone cable installers who can easily adapt to installing LAN cables.

3.3.3.2.1: UTP Cable Grades

UTP cable comes in different grades, called "categories" by the Electronics Industry Association (EIA) and the Telecommunications Industry Association (TIA). The two most significant UTP grades for LAN use are Category 3 and Category 5. Category 3 cable was designed for voice-grade telephone networks and came to be used for Ethernet. Category 3 cable is sufficient for 10 Mbps Ethernet networks, but it is not used for Fast Ethernet, except under certain conditions. If you have an existing Category 3 cable installation, you can use it to build a standard Ethernet network, but virtually all new UTP cable installations today use at least Category 5 cable.

Table 3.2: EIA/TIA UTP Cable Grades

Cable Grade	Suitability
Category 1	Used for voice-grade telephone networks; not for data transmissions
Category 2	Used for voice-grade telephone networks, as well as IBM dumb-terminal connections to mainframe computers
Category 3	Used for voice-grade telephone networks, 10 Mbps Ethernet, 4 Mbps Token Ring, 100BaseT4 Fast Ethernet, and 100VG AnyLAN
Category 4	Used for 16 Mbps Token Ring networks
Category 5	Used for 100BaseTX Fast Ethernet, SONet, and OC-3 ATM
Category 5e	Used for Gigabit (1000 Mbps) Ethernet protocols
Category 6	Also used for Gigabit (1000 Mbps) Ethernet protocols

Category 5 UTP is best suited for 100BaseTX Fast Ethernet networks running at 100 Mbps. In addition to the officially ratified EIA/TIA categories, there are other UTP cable grades available that have not yet been standardized. A cable standard called Level 5 by a company called Anixter, Inc. is currently being marketed using names such as Enhanced Category 5. It increases the bandwidth of Category 5 from 100 to 350 MHz, making it suitable to run the latest Gigabit Ethernet protocol at 1,000 Mbps (1 Gbps).

3.3.3.2.2: STP Cable Grades

Shielded twisted pair cable is same in construction to UTP. It has two pairs of wires and it has additional foil or mesh shielding around each pair. The shielding in STP cable makes it preferable to UTP in installations where electromagnetic interference is a problem, often due to the proximity of electrical equipment. The various types of STP cable were standardized by IBM, who developed the Token Ring protocol. STP networks use Type 1A for longer cable runs and Type 6A for patch cables. Type 1A contains two pairs of 22 gauge solid wires with foil shielding, and Type 6A contains two pairs of 26 gauge stranded wires with foil or mesh shielding. Token Ring STP networks use large, connectors called IBM data connectors (IDCs). Most Token Ring LANs today use UTP cable.

3.3.3.3: Fiber Optic Cable

Fiber optic cable is a different type of network medium. Instead of carrying signals over copper conductors in the form of electrical voltages, fiber optic cables transmit pulses of light over a glass or plastic conductor. Fiber optic cable is resistant to the electromagnetic. Fiber optic cables are less subject to attenuation than are copper cables. **Attenuation** is the tendency of a signal to weaken as it travels over a cable. On copper cables, signals weaken to the point of unreadability after 100 to 500 meters. Some fiber optic cables can span distances up to 120 kilometers without excessive signal degradation. This makes fiber optic the medium of choice for installations that span long distances. Fiber optic cable is more secure than copper, because it is not possible to tap into a fiber optic link without affecting the normal communication over that link.

There are two types of fiber optic cable: single-mode and multimode. The difference between the two is in the thickness of the core and the cladding.

Single-mode fiber uses a single-wavelength laser as a light source, and as a result, it can carry signals for extremely long distances. Single-mode fiber is commonly found in outdoor installations that span long distances, such as telephone and cable television networks. This type of cable is less suited to LAN installations because it is much more expensive than multimode and has a higher bend radius. Multimode fiber uses a light emitting diode (LED) as a light source instead of a laser and carries multiple wavelengths. Multimode fiber cannot span distances as long as single-mode, but it bends around corners better and is much cheaper.

3.3.3.4: Wireless Networks

Conventional Ethernet networks require cables connected computers via hubs and switches. This has the effect of restricting the computer's mobility and requires that even portable computers be physically connected to a hub or switch to access the network. An alternative to cabled networking is wireless networking.

Wireless networks use network cards, called Wireless Network Adapters, that rely on radio signals or infrared (IR) signals to transmit and receive data via a Wireless Access Point (WAP) . The WAP uses has an RJ-45 port that can be attached to connect to a 10BASE-T or 10/100BASE-T Ethernet hub or switch and contains a radio transceiver, encryption, and communications software. It translates conventional Ethernet signals into wireless Ethernet signals it broadcasts to wireless network adapters on the network and performs the same role in reverse to transfer signals from wireless network adapters to the conventional Ethernet network. WAP devices come in many variations, with some providing the Cable Modem Router and Switch functions in addition to the wireless connectivity.

Table 3.3: Types of Wireless Network and Their Standards

Network	Description	Standard
Wireless personal area network (WPAN) also known as Bluetooth	Used for PDAs, cell phones and laptops. For short distances, it could use infrared.	802.15
Wireless local area network (WLAN).	To connect devices that is in corporate or campus environments.	802.11 802.11b 802.11a 802.11g

3.3.3.4.1 Wireless Network Standards

In the absence of an industry standard, the early forms of wireless networking were single-vendor proprietary solutions that could not communicate with wireless network products from other vendors. In 1997, the computer industry developed the IEEE 802.11 wireless Ethernet standard. Wireless network products based on this standard are capable of multivendor interoperability.

The IEEE 802.11 wireless Ethernet standard consists of the IEEE 802.11b standard, the IEEE 802.11a standard, and the newer IEEE 802.11g standard.

- **IEEE 802.11** was the original standard for wireless networks that was ratified in 1997. It operated at a maximum speed of 2 Mbps and ensured interoperability been wireless products from various vendors. However, the standard had a few ambiguities allowed for potential problems with compatibility between devices. To ensure compatibility, a group of companies formed the Wireless Ethernet Compatibility Alliance (WECA), which has come to be known as the **Wi-Fi Alliance**, to ensure that their products would work together. The term **Wi-Fi** is now used to refer to any IEEE 802.11 wireless network products that have passed the Wi-Fi Alliance certification tests.
- **IEEE 802.11b**, which is also called 11 Mbps Wi-Fi, operates at a maximum speed of 11 Mbps and is thus slightly faster than 10BASE-T Ethernet. Most IEEE 802.11b hardware is designed to operate at four speeds, using three different data-encoding methods depending on the speed range. It operates at 11 Mbps using quaternary phase-shift keying/complimentary code keying (QPSK/CCK); at 5.5 Mbps also using QPSK/CCK; at 2 Mbps using differential quaternary phase-shift keying (DQPSK); and at 1 Mbps using differential binary phase-shift keying (DBPSK). As distances change and signal strength increases or decreases, IEEE 802.11b hardware switches to the most suitable data-encoding method.

  Wireless networks running IEEE 802.11b hardware use the 2.4 GHz radio frequency band that many portable phones, wireless speakers, security devices, microwave ovens, and the Bluetooth short-range networking products use. Although the increasing use of these products is a potential source of interference, the short range of wireless networks (indoor ranges up to 300 feet and outdoor ranges up to 1,500 feet, varying by product) minimizes the practical risks. Many devices use a spread-spectrum method of connecting with other products to minimize potential interference.

  IEEE 802.11b networks can connect to wired Ethernet networks or be used as independent networks.
- **IEEE 802.11a** uses the 5 GHz frequency band, which allows for much higher speeds, reaching a maximum speed of 54 Mbps. The 5 GHz frequency band also helps avoid interference from devices that cause interference with lower-frequency IEEE 802.11b networks. IEEE 802.11a hardware maintains relatively high speeds at both short and relatively long distances.

  Because IEEE 802.11a uses the 5 GHz frequency band rather than the 2.4 GHz frequency band used by IEEE 802.11b, standard IEEE 802.11a hardware cannot communicate with 802.11b hardware. A solution to this compatibility problem is the use of dual-band hardware. Dual-band hardware can work with either IEEE 802.11a or IEEE 802.11b networks, enabling you to move from an IEEE 802.11b wireless network at home or at Starbucks to a faster IEEE 802.11a office network.
- **IEEE 802.11g** is also known as Wireless-G and combines compatibility with IEEE 802.11b with the speed of IEEE 802.11a at longer distances. This standard was ratified in mid-2003, however, many network vendors were already selling products based on the draft IEEE 802.11g standard before the final standard was approved. These early IEEE 802.11g hardware was slower and less compatible than the

specification promises. In some cases, problems with early-release IEEE 802.11g hardware can be solved through firmware upgrades.

3.3.3.4.2: Wireless Network Modes

Wireless networks work in one of two modes that are also referred to as topologies. These two modes are ad-hoc mode and infrastructure mode. The mode you implement depends on whether you want your computers to communicate directly with each other, or via a WAP.

- In **ad-hoc mode**, data is transferred to and from wireless network adapters connected to the computers. This cuts out the need to purchase a WAP. Throughput rates between two wireless network adapters are twice as fast as when you use a WAP. However, a network in ad-hoc mode cannot connect to a wired network as a WAP is required to provide connectivity to a wired network. An ad-hoc network is also called a peer-to-peer network.
- In **infrastructure mode**, data is transferred between computers via a WAP. Because a WAP is used in infrastructure mode, it provides connectivity with a wired network, allowing you to expand a wired network with wireless capability. Your wired and wirelessly networked computers can communicate with each other. In addition, a WAP can extend your wireless network's range as placing a WAP between two wireless network adapters doubles their range. Also, some WAPs have a built-in router and firewall. The router allows you to share Internet access between all your computers, and the firewall hides your network. Some of these multifunction access points include a hub with RJ-45 ports.

3.3.3.4.3: Bluetooth

Bluetooth operates in the 2.4 GHz frequency spectrum. The Bluetooth specification allows a maximum data connection speed of 723 Kbps. Bluetooth uses much lower power levels than wireless LAN technologies (802.11x). Because it uses less power, its radio wave is not as strong as the 802.11.

Bluetooth uses a peer-to-peer networking model. Bluetooth does not require line of sight between any of the connected devices. Bluetooth can also connect multiple devices together in a point-to-multipoint fashion.

Cell phone and PDAs are not the only devices that can use Bluetooth. The value of Bluetooth would be lessened if it was not for the network effect of a networked device that increases exponentially as the number of networked devices increases.

There are three classifications of Bluetooth. These are:

- **Class 1**, which has a range of up to 100m and 100mW of power.
- **Class 2**, which has a range of up to 20m and has 2.5mW of power.
- **Class 3**, which is the most widely implemented classification. It has a range of up to 10m and has 1mW of power.

3.3.3.4.4: IrDA

IrDA uses infrared (IR) signals to transmit data for IrDA 1.1. Normal distance is 3 to 6 meters, although some IR technologies have a maximum distance of 1.5 miles. Because IR signals are used to transmit data, distance limits for long-range IR depend on weather conditions (such as humidity). Additionally, IR is a line-of-sight technology that requires a clear path between the transmitter and receiver.

3.3.4: Networking Devices

Network devices can be categorized based on their function relative to the OSI model. The main network devices are:

- **Hubs and repeaters**, which operate at the physical layer of the OSI model and basically transmit the incoming data (bits) out all other ports on the device. These devices are not aware of frames or packets; they amplify the signal and send out all ports. Repeaters do not break up broadcast or collision domains and are said to be protocol transparent because they are not aware of upper-layer protocols, such as IP, IPX, DECnet, etc.
- **Bridges and Layer-2 switches**, which operate at the data-link layer of the OSI model. Bridges learn the MAC layer addresses of each node of the segments and build tables of MAC addresses and ports, identifying which interface a particular MAC address is connected to. If the destination MAC address of an incoming frame is not in the table, bridges forward the frame to all ports except the port from which the frame came. If the destination MAC address is in the table, bridges forward the frame through the port to which the destination MAC address is attached if the destination MAC address is not on the same port from which the frame came, other wise the bridge filters (drops) the frame. Bridges are store-and-forward devices. They store the entire incoming frame and verify the checksum before forwarding the frame. If a checksum error is detected, the frame is discarded.

Collision Domains

A collision domain is a set of network interface cards (NICs) for which a frame sent by one NIC could result in a collision with a frame sent by any other NIC in the same collision domain. In a collision domain all devices on the network compete for the same bandwidth.

Broadcast Domains

A broadcast domain is a set of NICs for which a broadcast frame sent by one NIC is received by all other NICs in the same broadcast domain.

Layer-2 Switching and Routing

The major difference between Layer-2 switching and routing is that switching occurs at Layer 2 of the OSI reference model and routing occurs at Layer 3. Switches forward frames based on MAC address information while Routers forward packets based on logical addresses (IP address).

Switches use fast integrated circuits that reduce the latency common to bridges. Some switches have the capability to run in **cut-through mode** where the switch does not wait for the entire frame to enter its buffer; instead, it begins to forward the frame as soon as it finishes reading the destination MAC address. This increases the probability that error frames are propagated on the network because the frame is forwarded before the entire frame is buffered and checked for errors. Each port on a bridge or switch is a separate **collision domain** but all ports in a switch are in the same **broadcast domain** because bridges and switches do not control broadcasts. Instead they flood broadcasts out all ports.

- **Routers and Layer-3 switches**, which operate in the network layer of the OSI model and make forwarding decisions based on network layer addresses, such as IP addresses. Routers define both collision and broadcast domains as each router interface is a separate broadcast domain that is defined by a separate subnetwork. Routers are protocol aware, which means that they are capable of forwarding packets of routable protocols such as IP, IPX, DECnet, and AppleTalk. Routers are configured to run routing protocols, such as Routing Information Protocol (RIP); Interior Gateway Routing Protocol (IGRP); Open shortest Path First (OSPF); Intermediate System-to-Intermediate System (IS-IS); Enhanced Interior Gateway Routing Protocol (EIGRP); and Border Gateway Protocol (BGP), to determine the best paths to a destination. Routers exchange information about destination networks and their interface status by using these routing protocols. Routers can also be configured manually with static routes.

  LAN switches that are capable of running routing protocols and can communicate with routers as peers are called Layer-3 switches. Layer-3 switches off-load local traffic from wide-area network (WAN) routers by performing network-layer forwarding within the local-area networks (LANs). Both routers

and Layer-3 switches make forwarding decisions based on IP addresses and not MAC addresses. Both participate in the exchange of route information based on the dynamic routing protocol they participate in.

- **Firewalls**, which are computers, routers, or software components implemented to control access to a protected network. They filter potentially harmful incoming or outgoing traffic and can also partition internal networks from the Internet, as well as protect individual computers. The three types of firewalls are packet filtering firewalls; proxy servers, which includes application filtering firewalls and circuit-level firewalls, and stateful inspection firewalls.
  - **Packet Filtering**: A packet filtering firewall checks each packet crossing the device. It also inspects the packet headers of all network packets going through the firewall.
    - **Source IP Address:** It identifies the host that is sending the packet. Attackers can modify this field in an attempt to conduct IP spoofing. Firewalls are configured to reject packets that arrive at the external interface, that is either an erroneous host configuration or an attempt at IP spoofing.
    - **Destination IP Address:** This is the IP address that the packet is trying to reach.
    - **IP Protocol ID:** Each IP header has a protocol ID that follows. For example, Transmission Control Protocol (TCP) is ID 6, User Datagram Protocol (UDP) is ID 17, and Internet Control Message Protocol (ICMP) is ID 1.
    - **Fragmentation Flags:** Firewalls examine and forward or reject fragmented packets. A successful fragmentation attack can allow an attacker to send packets that could compromise an internal host.
    - **IP Options Setting:** This field is used for diagnostics. The firewall is configured to drop network packets that use this field. Attackers can use this field in conjunction with IP spoofing to redirect network packets to their systems.
  - **Application Filtering:** This device will intercept connections and performs security inspections. The firewall acts as a proxy for connections between the internal and external network. The firewall enforce access control rules specific to the application. It is also use to check incoming e-mails for virus attachments. These firewalls are often called e-mail gateways.
  - **Proxy Server:** A proxy server takes on responsibility for providing services between the internal and external network. Proxy server can be used to hide the addressing scheme of the internal network. It can also be used to filter requests based on the protocol and address requested.
  - **Circuit-Level:** A circuit-level firewall controls TCP and UDP ports, but doesn't watch the data transferred over them. If a connection is established, the traffic is transferred without any further checking.
  - **Stateful Inspection:** An inspection firewall works at the Network layer. It assesses the IP header information. It also monitors the state of each connection. Connections are rejected if they attempt any actions that are not standard for the given protocol. These listed firewall features can be implemented in combination by a given firewall implementation. Placing a lot of firewalls in series is a common practice to increase security at the network perimeter.

3.3.5: Network Technologies

Various network technologies can be used to establish network connections, including Ethernet, Fiber Distribution Data Interface (FDDI), Copper Distribution Data Interface (CDDI), Token Ring, and Asynchronous Transfer Mode (ATM). Of these, Ethernet is the most popular choice in installed networks because of its low cost, availability, and scalability to higher bandwidths.

3.3.5.1: Ethernet

Ethernet is based on the Institute of Electrical and Electronics Engineers (IEEE) standard **IEEE 802.3** and offers a bandwidth of 10 Mbps between end users. Ethernet is based on the carrier sense multiple access

collision detect (CSMA/CD) technology, which requires that transmitting stations back off for a random period of time when a collision occurs.

Coaxial cable was the first media system specified in the Ethernet standard. Coaxial Ethernet cable comes in two major categories: **Thicknet** (10Base5) and **Thinnet** (10Base2). These cables differed in their size and their length limitation. Although Ethernet coaxial cable lengths can be quite long, they are susceptible to electromagnetic interference (EMI) and eavesdropping.

Table 3.4: Coaxial Cable for Ethernet

Cable	Diameter	Resistance	Bandwidth	Length
Thinnet (10Base2)	10 mm	50 ohms	10 Mbps	185 m
Thicknet (10Base5)	5 mm	50 ohms	10 Mbps	500 m

Most wired networks use twisted-pair media for connections to the desktop. Twisted-pair also comes in two major categories: **Unshielded twisted-pair (UTP)** and **Shielded twisted-pair (STP)**. One pair of insulated copper wires twisted about each other forms a twisted-pair. The pairs are twisted top reduce interference and crosstalk. Both STP and UTP suffer from high attenuation, therefore these lines are usually restricted to an end-to-end distance of 100 meters between active devices. Furthermore, these cables are sensitive to EMI and eaves dropping. Most networks use 10BaseT UTP cable.

An alternative to twisted-pair cable is fiber optic cable (10BaseFL), which transmits light signals, generated either by light emitting diodes (LEDs) or laser diodes (LDs), instead of electrical signals. These cables support higher transmission speeds and longer distances but are more expensive. Because they do not carry electrical signals, fiber optic cables are immune to EMI and eavesdropping. They also have low attenuation which means they can be used to connect active devices that are up to 2 km apart. However, fiber optic devices are not cost effective while cable installation is complex.

Table 3.5: Twisted-Pair and Fiber Optic Cable for Ethernet

Cable	Technology	Bandwidth	Cable Length
Twisted-Pair	(10BaseT)	10 Mbps	100 m
Fiber Optic	(10BaseFL)	10 Mbps	2,000 m

3.3.5.2: Fast Ethernet

Fast Ethernet operates at 100 Mbps and is based on the **IEEE 802.3u** standard. The Ethernet cabling schemes, CSMA/CD operation, and all upper-layer protocol operations have been maintained with Fast Ethernet. Fast Ethernet is also backward compatible with 10 Mbps Ethernet. Compatibility is possible because the two devices at each end of a network connection can automatically negotiate link capabilities so that they both can operate at a common level. This negotiation involves the detection and selection of the highest available bandwidth and half-duplex or full-duplex operation. For this reason, Fast Ethernet is also referred to as **10/100 Mbps Ethernet**.

Cabling for Fast Ethernet can be either UTP or fiber optic. Specifications for these cables are shown in Table 3.7.

Table 3.6: Fast Ethernet Cabling and Distance Limitations

Technology	Wiring Type	Pairs	Cable Length
100BaseTX	EIA/TIA Category 5 UTP	2	100 m
100BaseT2	EIA/TIA Category 3,4,5 UTP	2	100 m
100BaseT4	EIA/TIA Category 3,4,5 UTP	4	100 m
100BaseFX	Multimode fiber (MMF) with 62.5 micron core; 1300 nm laser	1	400 m (half-duplex) 2,000 m (full-duplex)
	Single-mode fiber (SMF) with 62.5 micron core; 1300 nm laser	1	10,000 m

3.3.5.3: Gigabit Ethernet

Gigabit Ethernet is an escalation of the Fast Ethernet standard using the same **IEEE 802.3** Ethernet frame format. Gigabit Ethernet offers a throughput of 1,000 Mbps (1 Gbps). Like Fast Ethernet, Gigabit Ethernet is compatible with earlier Ethernet standards. However, the physical layer has been modified to increase data transmission speeds: The IEEE 802.3 Ethernet standard and the American National Standards Institute (ANSI) X3T11 FibreChannel. IEEE 802.3 provided the foundation of frame format, CSMA/CD, full duplex, and other characteristics of Ethernet. FibreChannel provided a base of high-speed ASICs, optical components, and encoding/decoding and serialization mechanisms. The resulting protocol is termed IEEE 802.3z Gigabit Ethernet.

Gigabit Ethernet supports several cabling types, referred to as 1000BaseX. Table 3.8 lists the cabling specifications for each type.

Table 3.7: Gigabit Ethernet Cabling and Distance Limitations

Technology	Wiring Type	Pairs	Cable Length
1000BaseCX	Shielded Twisted Pair (STP)	1	25 m
1000BaseT	EIA/TIA Category 5 UTP	4	100 m
1000BaseSX	Multimode fiber (MMF) with 62.5 micron core; 850 nm laser	1	275 m
	Multimode fiber (MMF) with 50 micron core; 1300 nm laser	1	550 m
1000BaseLX/LH	Multimode fiber (MMF) with 62.5 micron core; 1300 nm laser	1	550 m
	Single-mode fiber (SMF) with 50 micron core; 1300 nm laser	1	550 m
	Single-mode fiber (SMF) with 9 micron core; 1300 nm laser	1	10 km
1000BaseZX	Single-mode fiber (SMF) with 9 micron core; 1550 nm laser	1	70 km
	Single-mode fiber (SMF) with 8 micron core; 1550 nm laser	1	100 km

3.3.5.4: Token Ring

The IEEE standard for Token Ring is IEEE 802.5. Token Ring was originally developed by IBM for the forwarding of data on a logical unidirectional ring. Like Ethernet, Token Ring is a LAN technology that provides shared media access to multiple connected hosts and is implemented in the Data-Link Layer. Token Ring networks pass a small frame, called a **token**, around from host to host the network. Possession of the token grants the holder the right to transmit a frame onto the ring. After a station has the token, it modifies it into a data frame, appends the data for transmission, and sends the frame to the next station. No token is on the ring until the data frame is received by the source station marked as read and copied, and releases a token back into the ring. This means that only one station can transmit at a given time, and prevents a Token Ring network from experiencing collisions.

A Token Ring network offers a bandwidth of 4 Mbps or 16 Mbps. At the higher rate, hosts are allowed to introduce a new token as soon as they finish transmitting a frame. This early token release increases efficiency by letting more than one host transmit a frame during the original token's round trip. One station is elected to be the ring monitor, to provide recovery from runaway frames or tokens. The ring monitor will remove frames that have circled the ring once, if no other station removes them.

Traditional Token Ring networks use **multistation access units (MSAUs)** to provide connectivity between hosts. MSAUs have several ports that a host can connect to, with either a **B** connector for **Type 2** cabling or an **RJ-45** connector for Category 5 UTP cabling. Internally, the MSAU provides host-to-host connections to form a ring segment. The Ring-In and Ring-Out connectors of a MSAU can be chained to other MSAUs to form a complete ring topology.

Token Ring includes an optional **priority system** that permits stations configured with a higher priority value to use the network more frequently than permitted by the default round-robin operation. Eight levels of priority are provided using a 3-bit reservation field and a 3-bit priority field. As an information frame passes, a station sets a higher priority in the reservation field, which reserves the token. The transmitting station then sends a token out with the higher priority set. After the high priority station completes sending its frame, it releases a token with the normal or previous priority.

3.3.5.4.1: Token Ring Operation

One station on the Token Ring is selected to be the **Active Monitor (AM)**. This station removes continuously circulating frames that are not removed by a failed transmitting station. As a frame passes the AM, the monitor count bit is set. If a frame passes with the monitor count bit set, the AM assumes that the original sender of the frame was unable to remove the frame from the ring. The AM purges this frame, sends a **Token Soft Error** message to the **Ring Error Monitor**, and generates a new token.

The AM also provides timing information to ring stations. It inserts a 24-bit propagation delay to prevent the end of a frame from wrapping onto the beginning of the frame, and also confirms that a data frame or token is received every 10 milliseconds.

Standby Monitors and **Ring Error Monitors** are also on Token Ring networks. **Standby Monitors** take over as AM if the primary AM is removed from the ring or no longer performs its functions. **Ring Error Monitors** can also be present on the ring to collect ring status and error information.

If a station does not receive any more frames—either a data frame or a token—from its upstream neighbor, it sends a **beacon MAC** frame, which includes the beaconing station's MAC address and the address of the station's nearest active upstream neighbor (NAUN), and indicates that the problem lies between the two stations. An adapter keeps beaconing until it begins to receive frames again.

Ring Insertion

The process for a station to insert into the token ring follows five phases:

- **Phase 0 - Lobe media test. The transmitter and receiver of the adapter and the cable between the adapter and the MSAU are tested.**
- **Phase 1 - Physical insertion. The adapter opens a relay on the MSAU. After the MSAU opens, the adapter determines an AM is present on the ring, which indicates successful completion of phase 1.**
- **Phase 2 - Address verification. This phase verifies that the MAC address is unique to the ring. This phase can detect if two Locally Administered Addresses (LAAs) are configured with the same MAC address. This phase is also called the duplicate address test.**
- **Phase 3 - Participation in ring poll. The station learns its upstream neighbor's address and informs its downstream neighbor of the inserting adapter's address and produces a station list. If the adapter successfully participates in a ring poll, it proceeds into the final phase of insertion.**
- **Phase 4 - Request initialization. The adapter sends request initialization MAC frames to the functional address of the Ring Parameter Server (RPS). The RPS responds with information such as the ring number and speed. The adapter uses its own default values and reports successful completion of the insertion process if no RPS is present.**

3.3.5.4.2: Early Token Release (ETR)

In a normal token ring operation, the station that transmitted the data frame removes it and then generates a free token.

With ETR, a token is released immediately after the sending station transmits its frame. The sending station does not wait for the data frame to circle the ring. ETR is only available on 16 Mbps rings. Stations running ETR can coexist with stations not running ETR. With ETR, a free token can circulate the ring with multiple data frames.

3.3.6: Areas of the Network

The network is classified into areas based on where traffic originates from, and where its destination is. An area can be:

- **Trusted**, which is normally a private sector of the network that needs protection against security threats and attacks. Traffic coming from the less trusted areas of the firewall is blocked. In this manner, security is implemented and computers in this area enjoy more protection and security.
- **Untrusted**, which are areas of the network like the Internet segment of the firewall that are open to the element of security threats.
- **Demilitarized Zone (DMZ)**, which is an area like a web server, that normally supports computers or services that are used by trusted authorized users and untrusted external individuals. The Demilitarized

Zone therefore resides between a trusted and untrusted area. The Demilitarized Zone is considered untrusted when classifying the area from within the private trusted network. Accordingly, traffic originating from the DMZ is blocked.

A simple firewall configuration comprises of an inside trusted interface and an outside untrusted interface. A DMZ area can be established amid two firewalls that are jointly stacked. A standard router, known as a perimeter router, is normally used to supply the Internet Service Provider (ISP) connection. Three-pronged **firewalls** refer to the more sophisticated firewall models that have no fewer than three interfaces: an inside trusted interface, an outside untrusted interface and a DMZ connecting to an area that is partially trusted.

Section 3.4 Common Data Network Services

3.4.1: File Transfer Protocol (FTP)

File Transfer Protocol (FTP) is a TCP-based application that has many options and features, including the capability to change directories, list files using wildcard characters, transfer multiple files with a single command, and use a variety of character sets or file formats. It can be configured to allow anonymous access without requiring passwords, or it can be set up to require a valid username and password. It also provides a simple interface resembling a standard UNIX file directory.

When an FTP client attempts to connect to an FTP server, a TCP connection is established to the FTP server's well-known **port 21**. The FTP client is required to enter a username and password, which the server uses to authenticate the files available to that user for read and write permissions. This security is based on the file security on the server's platform. All the commands used to control the transfer of a file are sent across this connection. At this point, the user has a variety of commands available to enable settings for transfer, change directories, list files, etc. However, whenever a **`get`** (**`mget`** for multiple files) or **`put`** (or **`mput`** for multiple files) command is entered, or the equivalent button is clicked on the user interface, a file is transferred. The data is transferred over a separate FTP data connection, another TCP connection, established to well-known **port 20**. This prevents a file transfer from impacting on the control connection.

3.4.2: Secure File Transfer Protocol (SFTP)

SFTP (Secure File Transfer Protocol) is a secure version of the File Transfer Protocol (FTP) that includes strong encryption and authentication. It provides secure file transfer functionality using SSH or SSH-2. Like FTP, SFTP can be used to transfer files between a client and a server over a network. It can also provide secure file system access to a remote server. An SFTP server can be designed to provide only file transfer access, or it can provide system command access as well. SFTP can restrict users to their home directories, is not vulnerable to the "flashfxp" transfer utility, and is much less vulnerable to remote exploitation than FTP. It can be configured to authorize users with certificates as well as passwords.

3.4.3: Secure Shell (SSH) and Secure Shell version 2 (SSH-2)

SSH is an open standard used for remote administration and file transfer over a network. It establishes an encrypted tunnel between an SSH client and an SSH server and can authenticate the client to the server. SSH is designed to replace clear-text telnet sessions that are highly insecure. SSH uses port 22 and can be used in place of both FTP and Telnet.

All SSH communications are encrypted using the International Data Encryptions Algorithm. Rivest, Shamir, & Addleman (RSA) methods are used for key exchange. Keys are destroyed and regenerated every hour. SSH protects from the following attacks:

- **IP spoofing or IP source routing:** The attacker replays the source IP address in his packets. Then it looks like it did come from a trusted IP address.
- **DNS spoofing:** It occurs when a attacker forges name server records in the DNS.
- **Real-time data modification:** This occurs when an intermediary host hijacks active communication and impersonates both parties to both parties. The attacker receives information sent by the real sender, modifies it and forwards it to the recipient on behalf of the sender.
- **Authentication replay attacks:** The attacker records the stream of data and cutoff all user replies from the stream to establish a connection. If a hacker gets into a workstation where SSH is used and gains root access privileges, he can then modify the SSH application to his liking.

Secure Shell version 2 (SSH-2) is an security enhanced version SSH and should be used in place of SSH.

3.4.4: Trivial File Transfer Protocol (TFTP)

Trivial File Transfer Protocol (TFTP) is a more basic version of FTP. It has a small set of features, takes little memory to load, and less time to program. It has no directory browsing abilities and can only send and receive files. TFTP is commonly used to capture router configuration files by logging a terminal session during a configuration session and then storing that configuration on a TFTP server. The TFTP server is then accessed during the configuration session to save or retrieve configuration information to the network device. However, unlike FTP, session authentication does not occur. Therefore, TFTP is insecure.

Section 3.5: Types of Data Networks

A **data network** consists of two or more computers that are connected for the purpose of sharing files, printers, data, etc. To communicate on the network, every workstation must have a network interface card (NIC); a transmission medium, such as copper, fiber, or wireless; a Network Operating System (NOS); and a network connectivity device, such as a hub, bridge, router, or switch.

Data networks can be classified and defined according to geographical area that the network covers. There are four geographically defined networks: a **Local Area Network (LAN)**, a **Campus Area Network (CAN)**, a **Metropolitan Area Network (MAN)**, and a **Wide Area Network (WAN)**. There are three additional network definitions based on the network infrastructure. These are: the **Internet**, an **intranet** and an **extranet**. These network definitions are described in Table 3.1.

Table 3.8: Network Definitions

Definition	Description
Local Area Network (LAN)	A LAN is defined as a network that is contained within a closed environment and does not exceed a distance of 1.25 mile (2 km). Computers and peripherals on a LAN are typically joined by a network cable or by a wireless network connection. A LAN that consists of wireless connections is referred to as a **Wireless LAN (WLAN)**.
Campus Area Network (CAN)	A CAN is limited to a single geographical area but may exceed the size of a LAN
Metropolitan Area Network (MAN)	A MAN is defined as a network that covers the geographical area of a city that is less than 100 miles.

Wide Area Network (WAN)	A WAN is defined as a network that exceeds 1.25 miles. A WAN often consists of a number of LANs that have been joined together. A CAN and a MAN is also a WAN. WANs typically connected numerous LANs through the internet via telephone lines, T1 lines, Integrated Services Digital Network (ISDN) lines, radio waves, cable or satellite links.
Internet	The Internet is a world wide web of networks that are based on the TCP/IP protocol and is not own by a single company or organization.
Intranet	An intranet uses that same technology as the Internet but is owned and managed privately by the organization. A LAN or a WAN is usually an intranet. Intranets are also referred to as a private network or an internal network.
Extranet	An extranet is essentially the same as an intranet with the exception that it is private network owned and managed by another organization.

Of these network definitions, the most common are the Internet, the LAN and the WAN.

Section 3.6: Wide Area Networks

A Wide Area Network (WAN) is a conglomerate of physically or logically interconnected subnetwork that covers a larger geographic area than LANs and crosses metropolitan, regional, or national boundaries. WANs differ from LANs in the following ways:

- WANs cover greater distances.
- WAN speeds are slower.
- WANs can be connected on demand or it can be connected permanent; LANs have permanent connections between stations.
- WANs can use public or private network transports.
- WANs use private network transports.

3.6.1: Internet

The Internet is a TCP/IP based WAN that was originally developed for the U.S. Department of Defense (DoD). The Internet is a global network of public networks that is connected by Internet Service Providers (ISPs). Both public and private networks can be connected to the Internet.

3.6.2: Intranet

An intranet is a logical network that uses an organization’s internal, physical network infrastructure that can also span large areas. Intranets use TCP/IP and HTTP standards and are commonly used to publish corporate web sites that are accessible by all employees on the intranet. This is more secure and controlled than publishing corporate web sites on the Internet.

3.6.3: Extranet

An extranet is similar to an intranet. It is a private network that uses Internet protocols and can be used to publish corporate web sites. However, an extranet is accessible users outside the organization, such as business partners.

3.6.4: WAN Technologies

3.6.4.1: Dedicated Lines

A dedicated line, such as a **leased line** or a **point-to-point link**, is a communications line that is continuously available for transmission, rather than being switched on and off as transmission is required. These lines run over physical time-division multiplexed (TDM) networks and use private synchronous circuits, which is a dedicated analog or digital point-to-point connection that can interconnect diverse networks. Synchronous circuits must have the same clock so that the receiving side knows exactly when each frame bit is received.

There are many dedicated line speeds available. These line speeds are based on the basic **digital signal level 0 (DS-0)** rate of 64 kbps. The T-carriers are the most common dedicated lines in North America.

- The **T1** carrier can carry 24 DS-0s for a capacity of 1.544 Mbps.
- The **T3** carrier is a dedicated phone connection. It consists of 672 individual DS-0 channels and supports data rates of approximately 45 Mbps. The T3 is also commonly called DS-3 and carries 28 T1 lines.
- The **E1** carrier is the most common dedicated lines in Europe and other countries, and can carry 30 DS-0s for a capacity of 2.048 Mbps.

3.6.4.2: WAN Switching

WAN switching is required in networks that need more than a single point-to-point connection. There are two main types of WAN switching: circuit switching and packet switching.

3.6.4.2.1: Circuit-Switched Networks

In a circuit-switched network, a dedicated point to point connection or circuit must exist between the sender and receiver for the duration of the data transmission. This network type is used heavily in telephone company networks.

Integrated Services Digital Network (ISDN) is an example of a circuit-switched network. It provides permanent, always-on WAN connectivity and is the most popular choice for connectivity between routers. ISDN uses digital signals, which allows for faster speeds than analog lines. These speeds are in increments of 64 kbps. For Internet access, ISDN has been usurped by competing technologies such as Digital Subscriber Line (DSL), Asymmetric Digital Subscriber Line (ADSL) cable modems, and simply faster analog modems. ISDN, however, remains a popular option for temporary connectivity between routers and is frequently used to create a backup link when the primary leased line or Frame Relay connection is lost.

3.6.4.2.2: Packet-Switched Networks

In a packet-switched network (PSN), nodes share bandwidth with each other by sending small data units called **packets**. Unlike circuit-switched networks, the data in packet-switched networks is broken up into packets and then sent to the next destination based on the router's routing table. At that destination, the packets are reassembled based on their originally assigned sequence numbers. PSNs are far more cost

effective than dedicated circuits because they create virtual circuits, which are used as needed. PSNs include:

- **X.25** is a connection-oriented packet-switching network, in which the packets travel over virtual circuits and is defined by the International Telecommunications Union (ITU-T). The ITU-T specifications defines the point-to-point communication between Data Terminal Equipment (DTE), Data Circuit-Terminating Equipment (DCE), or a Data Service Unit/Channel Service Unit (DSU/CSU), which supports both switched virtual circuits (SVCs) and permanent virtual circuits (PVCs). Routers and other devices perform the roles of data terminal equipment (DTE) and data circuit-terminating equipment (DCE). Routers are typically DTEs that are connecting to modems or packet switches, which perform the DCE function. X.25 was designed to operate effectively regardless of the type of systems that are connected to the network. It has become an international standard and is currently much more prevalent outside the United States.
- **Link Access Procedure-Balanced (LAPB)** was created for use with X.25, LAPB specifies methods for exchanging frames, monitoring frame sequence and missing frames, and executing frame acknowledgements and retransmission when necessary.
- **Frame Relay** is a high-performance, connection-oriented WAN connection technology. It is a successor to X.25 and LAPB and operates at speeds from 56 Kbps to 45 Mbps. It is very flexible and offers a wide array of deployment options. Frame Relay operates by statistically multiplexing multiple data streams over a single physical link. Each data stream is known as a **virtual circuit (VC)**. Frame Relay VCs come in two types: **Permanent Virtual Circuits (PVCs)** and **Switched Virtual Circuits (SVCs)** . Each VC is tagged with an identifier to keep it unique. The identifier, known as a **Data Link Connection Identifier (DLCI)**, is determined on a per-leg basis during the transmission. It must be unique and agreed upon by two adjacent Frame Relay devices. As long as the two agree, the value can be any valid number, and the number does not have to be the same end to end. Valid DLCI numbers are 16-1007. For DLCI purposes, 0-15 and 1008-1023 are reserved. The DLCI also defines the logical connection between the Frame Relay (FR) switch and the customer premises equipment (CPE).
- **Switched Multimegabit Data Service (SMDS)** is a high-speed, connectionless, packet-switched public network service. It is generally delivered over a SONET ring with a maximum effective service radius of around 30 miles. It provides bandwidth to organizations that exchange large amounts of data over WANs on a bursty or non-continuous basis.
- **Asynchronous Transfer Mode (ATM)** is a connection-oriented high-bandwidth, low-delay transport technology that uses both switching and multiplexing. It was developed to meet the need to transmit voice, data, and video across enterprise and service provider networks and uses 53-byte, fixed size cells rather than frames. It can allocate bandwidth on demand, making it ideal for bursty applications. ATM requires a high speed, high-bandwidth medium like fiber optics.
- **Voice over IP (VoIP)** is a multi-service digital access technology that combines different types of data into a single IP packet, including data, voice, audio and video. This provides major benefits in the areas of cost, interoperability, and performance.

3.6.5: Network Address Translation (NAT)

The advantage of using private IP addresses is that it allows an organization to use private addressing in a network, and use the Internet at the same time, by implementing Network Address Translation (NAT).

NAT is defined in RFC 1631 and allows a host that does not have a valid registered IP address to communicate with other hosts through the Internet. Essentially, NAT allows hosts that use private addresses or addresses assigned to another organization, i.e. addresses that are not Internet-ready, to continue to be

used and still allows communication with hosts across the Internet. NAT accomplishes this by using a valid registered IP address to represent the private address to the rest of the Internet. The NAT function changes the private IP addresses to publicly registered IP addresses inside each IP packet that is transmitted to a host on the Internet.

There are several variations of NAT. These include Static NAT; Dynamic NAT; and Overloading NAT with Port Address Translation (PAT).

- In **Static NAT**, the IP addresses are statically mapped to each other. Thus, the NAT router simply configures a one-to-one mapping between the private address and the registered address that is used on its behalf. Supporting two IP hosts in the private network requires a second static one-to-one mapping using a second IP address in the public address range, depending on the number of addresses supported by the registered IP address. .
- **Dynamic NAT** is similar to static NAT in that the NAT router creates a one-to-one mapping between an inside local and inside global address and changes the IP addresses in packets as they exit and enter the inside network. However, the mapping of an inside local address to an inside global address happens dynamically. Dynamic NAT accomplishes this by setting up a pool of possible inside global addresses and defining criteria for the set of inside local IP addresses whose traffic should be translated with NAT.

  With dynamic NAT, you can configure the NAT router with more IP addresses in the inside local address list than in the inside global address pool. When the number of registered public IP addresses is defined in the inside global address pool, the router allocates addresses from the pool until all are allocated. If a new packet arrives, and it needs a NAT entry, but all the pooled IP addresses are already allocated, the router discards the packet. The user must try again until a NAT entry times out, at which point the NAT function works for the next host that sends a packet. This can be overcome through the use of Port Address Translation (PAT).
- **Overloading NAT with Port Address Translation (PAT)** is used in some networks where most, if not all, IP hosts need to reach the Internet. If that network uses private IP addresses, the NAT router needs a very large set of registered IP addresses. If you use static NAT, each private IP host that needs Internet access needs a publicly registered IP address. Dynamic NAT lessens the problem, but if a large percentage of the IP hosts in the network need Internet access throughout normal business hours, a large number of registered IP addresses would also be required. These problems can be overcome through overloading with port address translation. Overloading allows NAT to scale to support many clients with only a few public IP addresses.

  To support lots of inside local IP addresses with only a few inside global, publicly registered IP addresses, NAT overload uses Port Address Translation (PAT), translating the IP address as well as translating the port number. When NAT creates the dynamic mapping, it selects not only an inside global IP address but also a unique port number to use with that address. The NAT router keeps a NAT table entry for every unique combination of inside local IP address and port, with translation to the inside global address and a unique port number associated with the inside global address. And because the port number field has 16 bits, NAT overload can use more than 65,000 port numbers, allowing it to scale well without needing many registered IP addresses.

  NAT can also be used in organizations that do not use private addressing but use a network number registered to another company. If one organization uses a network number that is registered to another organization, and both organizations are connected to the Internet, NAT can be used to translate both the source and the destination IP addresses. However, both the source and the destination addresses must be changed as the packet passes through the NAT router..

Section 3.7: Remote Access

There are different ways to connect a system to a remote location. From the network layer up, a remote connection is the same as a direct LAN connection, but the data-link and physical layers can take different forms.

- **Public Switched Telephone Network (PSTN)** is a telephone service. You can connect a modem to the line to convey data to any location. The PSTN service uses copper-based twisted pair cable. These PSTN connections go to a central office where the calls can be routed to anywhere in the world. To convey data the system must convert its data to analog signals to enable the telephone network to carry it. The **modem** does the conversion job. The modem converts the digital signals to analog signals and transmits them over the PSTN. PSTN connections are slow. The quality of the link depends on the location and the state of the cables.

  Almost all of the modems support the Plug and Play standard. The systems of today will detect and install the correct drivers for it. With external modems, the IRQ and I/O address are assigned to the serial port that is used to connect the modem to the system. Most of the system is supported by two serial ports. It is assigned to the four: COM1 and COM2 share IRQ4, and COM2 and COM4 share IRQ3. A chip, **universal asynchronous receiver-transmitter (UART)**, is used to manage the communication.

- A **Virtual Private Network (VPN)** is a connection between a remote system and a server on a private network that uses the Internet as its backbone. A remote user can connect to the Internet by the use of a modem and connect to an ISP. The remote system and the network server then set up a secured connection that protects the data exchanged between them by using tunneling. The protocol that allows tunneling is the Point-to-Point Tunneling Protocol (PPTP). PPTP works with PPP to establish a connection between the client computer and a server. When the tunneling is active the system sends its data by encapsulating the PPP data.

- **Integrated Services Digital Network (ISDN)** was design to replace the analog telephone system. Digital Subscriber Line (DSL) and cable television (CATV) services are faster and cheaper than ISDN. ISDN is a dial-up service which means that no modem is required. Because ISDN uses the dial-up connection, it could connect various different sites. The speed of ISDN is better than PSTN. ISDN uses the same wiring as the PSTN, but additional equipment is required at the terminal locations. The telephone company offers a U-interface, which provide a four wire connection. Because of the speed, the length of the connection is limited is.

- **Digital Subscriber Line (DSL)** is a range of digital communication services that use standard telephone lines. The speed that it transfer data is faster that the PSTN and the ISDN.

  Digital Subscriber Line (DSL), which is also called xDSL, is a blanket term for a variety of digital communication services that use standard telephone lines but provide much faster data transfer speeds than the PSTN and ISDN. DSL uses a higher frequency than the normal telephone services, and support special signaling schemes. The connection of the DSL is direct and permanent. The available DSL services include:
  - **High-bit-rate Digital Subscriber Line (HDSL)**, which has a transmission rate of 1.544 Mbps in full-duplex using two wire pairs, or a 2.048 Mbps in full-duplex using three wire pairs. It has a connection length of 12,000 to 15,000 feet and is used by large networks as a replacement for T1 leased line connections, LAN and PBX interconnections, or Frame Relay traffic aggregation.
  - **Symmetrical Digital Subscriber Line (SDSL)**, which has a transmission rate of 1.544 Mbps in full-duplex or 2.048 Mbps in full-duplex using a pair of wires. It has a connection length of 10,000 feet. Like HDSL, it is use by large networks as a replacement for T1 leased line connections, LAN and PBX interconnections, or Frame Relay traffic aggregation.

- **Asymmetrical Digital Subscriber Line (ADSL)**, which has a transmission rate of 1.544 to 8.448 Mbps downstream and 16 Kbps to 640 Kbps upstream. It has a connection length of 10,000 to 18,000 feet and is used for internet/intranet access, remote LAN access, virtual private networking, video-on-demand, voice-over-IP.
- **Rate-Adaptive Digital Subscriber Line (RADSL)**, which has a transmission rate of 640 Kbps to 2.2 Mbps downstream and 272 Kbps to 1.088 Mbps upstream. It has a connection length of 10,000 to 18,000 feet. Like ADSL, it is used for internet/intranet access, remote LAN access, virtual private networking, video-on-demand, voice-over-IP, however, the transmission speed is dynamically adjusted to match the link length and signal quality.
- **ADSL Lite**, which has a transmission rate of up to 1 Mbps down-stream and up to 512 Kbps upstream. It has a connection length is 18,000 feet and is used for internet/intranet access, remote LAN access, IP telephony, and video conferencing.
- **Very-high-bit-rate Digital Subscriber Line (VDSL)**, which has a transmission rate of 12.96 to 51.84 Mbps downstream and 1.6 to 2.3 Mbps upstream. It has a connection length is 1,000 to 4,500 feet and is used for Multimedia Internet access, high- definition television delivery.
- **ISDN Digital Subscriber Line (IDSL)**, which has a transmission rate of up to 144 Kbps in full-duplex. It has a connection length is 18,000 feet and is used for Internet/intranet access, remote LAN access, IP telephony, and video conferencing.

- **Cable Television (CATV)** Internet access is inexpensive and it can work at 512 Kbps. CATV uses broadband transmission, which means that one network medium carries a lot of signals simultaneously. If some of this bandwidth should go to data transmission, it could deliver Internet data as fast as the television signals. CATV data connections are not secure connections. If you run Windows on your computer and browse the network, you will see your neighbors' system on the same network as yours. Because of this you would share bandwidth and your security is zero. There are firewall products available for protection. CATV connection is not expensive, but it cannot be used to connect your PC with your office LAN.

3.7.1: Remote Access Requirements

The elements needed to establish a remote network connection:

- **Common protocols -** the computers that are linked must share common protocols at the data-link layer and above. Both computers should use a data-link layer protocol suitable for point-to-point connections, such as PPP or SLIP. There must be network and transport layer protocols in common, such as TCP/IP, IPX, or NetBEUI.
- **TCP/IP configuration -** if your remote system will be using the TCP/IP protocols to communicate with the host network, the computer must be assigned an IP address and other configuration parameters appropriate for that network. Most remote networking solutions enable the network server to assign configuration parameters automatically using DHCP.
- **Host and remote software -** the remote system needs a client program that can use the physical layer medium to setup a connection.
- **Security -** the host computer and the other systems on the network to which it is attached must have security mechanisms in place that control access to the network resources.

3.7.2: Virtual Private Networks (VPNs)

A Virtual Private Network (VPN) allows remote users to create virtual private networks to the corporate network through the internet. This allows remote users to make local calls to a local Internet service provider (ISP) rather than having remote users make a long distance calls to connect to a corporate network, the user

can call his or her local ISP. Using the connection to the local ISP, a VPN is created between the dial-up user and the corporate VPN server across the Internet. You can either use dedicated lines or dial-up lines to connect to an ISP when creating a VPN connection.

A VPN allows secure remote connections through an otherwise insecure public network such as the Internet, enabling the **confidentiality**, **integrity** and **availability** of data, while also allowing for authentication. In other words, a VPN connection is a remote access connection that allows for the same level of security that is available on a LAN. VPNs accomplish this through the process of tunneling, which is a method of transferring data from one system to another by encapsulating the data packets in an additional header. The additional header provides routing information so that the encapsulated payload can be transmitted across the public network. For a tunnel to be established, both the tunnel client and the tunnel server must be using the same tunneling protocol. Tunneling technology can be based on either a Layer 2 or a Layer 3 tunneling protocol. However, tunneling, and the use of a VPN, is not intended as a substitute for encryption/decryption. In cases where a high level of security is necessary, the strongest possible encryption should be used within the VPN itself.

3.7.2.1: VPN Applications

There are two basic configurations for a VPN:

- A client-to-gateway VPN, which is used when a remote user connects to a private network using a VPN. This is similar to using dial-up access, but the user can connect through any dial-up provider or a separate LAN with Internet access rather than over the phone system.
- A gateway-to-gateway VPN, which is used to form a permanent link between two VPN servers on separate networks, each with its own Internet connectivity.

There are also different VPN applications based on the basic VPN configuration and the network infrastructure. These are:

- Remote Access, which is based on a client-to-gateway VPN
- Intranet Access, which is based on a gateway-to-gateway VPN
- Extranet Access, which is also based on a gateway-to-gateway VPN

3.7.2.1.1: Remote Access VPN

Many organizations supply their own VPN connections via the Internet. Remote users running VPN client software are assured private access through a publicly shared environment through the ISP(s). By using analog, ISDN, DSL, cable technology, dial and mobile IP, VPNs are implemented over extensive existing network infrastructures. Email, database and office applications can also be used over these secure remote VPN connections.

3.7.2.1.2: Intranet Access VPN

Gateway-to-gateway VPNs allow an organization to extend its internal network to remote branch offices. These VPNs enable a secure point of contact between two end devices, which are usually two routers. The users on each remote LAN that are connected to the local router can communicate with the other LAN via this link. The data available through such a VPN would depend on what information is sharable and is determined by the organization’s security policies. The organization’s data is kept secure by the use of dedicated circuits. Frame Relay, Asynchronous Transfer Mode (ATM), or point-to-point circuits are examples of infrastructures used by VPNs.

3.7.2.1.3: Extranet Access VPN
Extranet Access VPNs are similar to Intranet Access VPNs but provide remote access for a brokers, agents, business partners or any other applicable non-employees. Extranet VPNs enable these connections to the organizations private network. A combination of remote access and intranet access infrastructures are used. The distinction would be the rights that are assigned to these users. Some level of security or authentication would be necessary to access the network, protect network resources, and prevent others from accessing the information.

3.7.2.1.4: Integrating VPN in a Routed Intranet
In some organizations, data of a highly sensitive nature needs to be isolated from the rest of the organization's network. Although this protects the sensitive data, it creates information accessibility problems for those users not physically connected to the isolated LAN.

VPNs allow the isolated LAN to be physically connected to the rest of the organization's network but separated by a VPN server. In this arrangement the VPN server does not act as a router between the rest of the network and the isolated LAN; but as a access control mechanism. Users on the network must have the appropriate permissions and authority to establish a VPN connection with the VPN server and gain access to the isolated data. In addition, all communication across the VPN can be encrypted to ensure data confidentiality.

3.7.2.2: VPN and Remote Access Protocols
VPN connections use a tunneling protocol to encrypt packets of data and transmit them through a public network. The two popular VPN protocols are: Point-to-Point Tunneling Protocol (PPTP) and Layer 2 Tunneling Protocol (L2TP). VPNs also make use of IP Security (IPSec) for encryption. The other remote access protocols include: Point-to-Point protocol (PPP), RADIUS and TACACS.

3.7.2.2.1: Point-to-Point Protocol (PPP)
The Point-to-Point protocol (PPP) is used to implement TCP/IP over point-to-point connections. It has the basic function of encapsulating Network Layer (Layer 3) protocol information over point-to-point links, which can be used to establish ISDN connections, dialup connections, or serial connections. PPP has its own framing structure, which enables the encapsulation of any Layer 3 protocol. Because PPP is, by nature, point-to-point, no mapping of protocol addresses is necessary. PPP makes use of the Link Control Protocol (LCP) to communicate between PPP client and host. LCP tests the link between client and PPP host and specifies PPP client configuration.

PPP has a number of capabilities that make it flexible and versatile, including:

- Multiplexing of network layer protocols
- Link configuration
- Link quality testing
- Authentication
- Header compression
- Error detection

- Link parameter negotiation

For authentication, PPP offers a number of options, including: **Password Authentication Protocol (PAP)**, **Shiva Password Authentication Protocol (SPAP)**, **Challenge Handshake Authentication Protocol (CHAP)** and **Extensible Authentication Protocol (EAP)**. These two protocols offer differing degrees of protection. Both protocols require the definition of usernames and passwords. This can be done on the router itself or on a TACACS or RADIUS authentication server.

- **Password Authentication Protocol (PAP)** is a clear text exchange of username and password information. When a user dials in, a username request is sent. Once that is entered, a password request is sent. All communications are in clear text form and no encryption is used. PAP is a one-way authentication between the router and the host.
- **Shiva Password Authentication Protocol (SPAP)** is a reversible encryption mechanism employed by Shiva. A client uses SPAP when connecting to a Shiva LAN Rover. This form of authentication is more secure than PAP but less secure than CHAP.
- **Challenge Handshake Authentication Protocol (CHAP))** is much more secure than PAP. It implements a two-way encrypted authentication process. Usernames and passwords still must exist on the remote router, but they are not transmitted as they were with PAP. Instead, when a user dials in, the access server issues a challenge message to the remote user after the PPP link is established. The remote end responds with a one-way hash function. This hash is generally an **MD5** entity. If the value of the hash matches what the router expects to see, the authentication is acknowledged. If not, the connection terminates. CHAP repeats a challenge every two minutes for the duration of the connection. If the authentication fails at any time, the connection is terminated. The access server controls the frequency of the challenges.
- **Extensible Authentication Protocol (EAP)** is an authentication protocol that can be extended with additional authentication methods that you can install separately. It enables an arbitrary authentication mechanism to authenticate a remote access connection.
  - **EAP with MD5-Challenge** uses the same challenge handshake protocol as PPP-based CHAP, but the challenges and responses are sent as EAP messages. A typical use for MD5-Challenge is to authenticate non-Windows remote access clients. EAP with MD5-Challenge does not support encryption of connection data.
  - **Protected Extensible Authentication Protocol (PEAP)** is primarily used to authenticate wireless users with a user name and password.
  - **Extensible Authentication Protocol-Transport Layer Security (EAP-TLS)** is used to enable remote access authentication with a smart card or a public key certificate.

3.7.2.2.2: Point-to-Point Tunneling Protocol (PPTP)

Point-to-Point Tunneling Protocol (PPTP) is an extension of PPP that takes advantage of the authentication, compression, and encryption mechanisms of PPP. PPTP is the most common protocol used for dial-up remote access but is used for single client-to-server connections as it allows only a single point-to-point connection per session. It encapsulates PPP frames into IP datagrams for transmission over an IP network. PPTP tunnels the PPP frame within a **Generic Routing Encapsulation (GRE)** header using IP protocol 47 and a TCP header using port 1723. To allow PPTP traffic through a firewall, the firewall must allow TCP port 1723 and IP protocol 47. PPTP is commonly used by Windows clients for asynchronous communications. It allows IP, IPX, or NetBEUI traffic to be encrypted and then encapsulated in an IP header. PPTP uses Microsoft Point-to-Point Encryption (MPPE) and compression from PPP. PPTP tunnels must be authenticated by using the same authentication mechanisms as PPP connections (PAP, CHAP, and EAP).

3.7.2.2.3: Layer 2 Tunneling Protocol (L2TP)
Layer 2 Tunneling Protocol (L2TP) is a combination of PPTP and Layer 2 Forwarding (L2F). L2TP supports the same authentication methods as PPTP but is a more secure tunneling protocol as it relies on IPSec for encryption. Like PPTP, it allows only a single point-to-point connection per session and is thus used for single client-to-server connections. L2TP also allows IP, IPX, or NetBEUI traffic to be encrypted and then sent over any medium that supports point-to-point datagram delivery, such as IP, X.25, Frame Relay, or ATM networks. This combination of L2TP and IPSec is known as **L2TP/IPSec**. When using IP as its datagram transport, L2TP can be used as a tunneling protocol over the Internet. L2TP tunnels must be authenticated by using the same authentication mechanisms as PPP connections. PPP encryption is not used because it does not meet the security requirements of L2TP as PPP encryption can provide confidentiality but not per packet authentication, integrity, or replay protection, instead data encryption is provided by IPSec, which uses **Data Encryption Standard (DES)** or **Triple DES (3DES)** by using encryption keys generated from IPSec's **Internet Key Exchange (IKE)** negotiation process. L2TP/IPSec used the source and destination IP addresses for authentication and embedded this information inside the encrypted portion of the packet. Therefore, NAT servers were incapable of changing the source and destination IP addresses. NAT Traversal (NAT-T), a new capability of L2TP/IPSec, enables you to use L2TP to connect to an L2TP Server when the client is located behind a NAT server. However, the client, the server, and the NAT Server must all support NAT-T. New VPN server installations should use L2TP rather than PPTP.

3.7.2.2.4: IP Security Protocol (IPSec)
IPSec is a series of standards that support the secured transfer of information across an IP internetwork. IPSec **Encapsulating Security Payload (ESP)** Tunnel mode supports the encapsulation and encryption of entire IP datagrams for secure transfer across a private or public IP internetwork. When IPSec is used, the two computers involved in the communication negotiate the highest common security policy. Then the computer initiating communication uses IPSec to encrypt the data before it sends the data across the network. On receiving the data, the destination computer decrypts the data before passing it to the destination process. This encryption and decryption process is done transparently.

3.7.2.2.5: Remote Authentication Dial-In User Service (RADIUS) and DIAMETER
Remote Authentication Dial-In User Service (RADIUS) is a client/server-based system that provides authentication, authorization, and accounting (AAA) services for remote dial-up access while securing the information system against unauthorized access.

RADIUS facilitates centralized user administration by storing all user profiles in one location that all remote services have access to. To authenticate to a RADIUS server, a user must enter his or her credentials, which is encrypted and sent to the RADIUS server in an **Access-Request** packet. On receiving the credentials, the RADIUS server accepts, rejects or challenges the credentials. If the RADIUS server accepts the credentials, it sends an **Access-Accept** packet and the user is authenticated successfully. If the RADIUS server rejects the credentials, it sends an **Access-Reject** packet. If the RADIUS server challenges the credentials, it sends an Access-Challenge packet, which prompts the user to provide additional information which the RADIUS server can use to authenticate the user.

For remote dial-up access, RADIUS also provides callback security, in which the RADIUS server terminates the connection and establishes a new connection to the user by dialing a predefined telephone number to which the user's modem is attached. This provides an additional layer of protection against unauthorized access over dial-up connections.

Due to the success of RADIUS, an enhanced version of RADIUS called DIAMETER was developed. DIAMETER is designed for use on all forms of remote connectivity and not just dial-up connectivity.

3.7.2.2.6: Terminal Access Controller Access Control System

There are three versions of Terminal Access Controller Access Control System (TACACS): TACACS, Extended TACACS (XTACACS), and TACACS+. All three versions authenticate users and deny access to users who do not have a valid username/password pairing. TACACS integrates the authentication and authorization functions. XTACACS allows the division of the authentication, authorization, and auditing functions, providing the administrator more control over its deployment. TACACS+ also allows the division of the authentication, authorization, and auditing but also provides two-factor authentication.

The TACACS authentication process is similar to that of RADIUS and it provides the same functionality as RADIUS. However, RADIUS is based on an Internet standard, whereas TACACS is a proprietary protocol. For this reason, TACACS has failed to gain the popularity of RADIUS.

Section 3.8: E-Mail Security

E-mail is one of the most widely and commonly used Internet services. The e-mail infrastructure consists of e-mail servers that use the **Simple Mail Transfer Protocol (SMTP)** to accept messages from clients and to transmit those messages to other e-mail servers, and e-mail clients that use **Post Office Protocol version 3 (POP3)** or **Internet Message Access Protocol (IMAP)** to send and retrieve e-mail to and from the e-mail server. Although these protocols offer efficient delivery of e-mail messages, they are not secure and lacks controls to provide for confidentiality, integrity, and availability. Fortunately there are means for providing security to e-mail messaging.

A security policy is an important element in e-mail security. The e-mail security policy must address acceptable use policies for e-mail, which defines the activities that can and cannot be performed over the organization's e-mail infrastructure. This often allows the sending and receiving of work related e-mail but limits sending and receiving of personal e-mail. Also, illegal, immoral, or offensive e-mail can be restricted, as well as personal business e-mail. Access control over e-mail should be implemented to ensure that users have access only to their own inbox and e-mail archive databases.

The mechanisms and processes used to manage the organizations' e-mail infrastructure should be defined. End users need not know the specifics of e-mail management, but they must be informed as to whether e-mail is or is not considered private communication. End users also need to be informed as to whether e-mail is to be backed up and stored in archives for future use, and if e-mail is to be reviewed for violations by an auditor.

3.8.1: E-Mail Security Issues

POP3, IMAP and SMTP, the protocols that support e-mail, do not provide security. They do not provide encryption, source verification, or integrity checking.

- Because e-mail protocols do not provide encryption of transmitted e-mail, the interception and eavesdropping of e-mail is a real vulnerability.
- The e-mail protocols do not offer verification of the sender, or the source of the e-mail. Indeed, e-mail address spoofing is a simple process for even a novice hacker. E-mail headers can be modified at their

source and during transition. It is also possible to deliver e-mail directly to a user's inbox on an e-mail server by directly connecting to the e-mail server's SMTP port.

- The e-mail protocols also do not provide integrity checks to ensure that an e-mail message was not altered during transmission.

In addition, e-mail itself can be used as an attack mechanism. E-mail attachments are also the most common method by which malicious code, such as viruses, worms, and Trojan horses are introduced in an information system. **Mailbombing** is another attack. It is a denial of service (DoS) attack in which a large number of e-mail messages are directed to a user's inbox or through a specific STMP server, inundating a system with messages and can result in storage capacity consumption or processing capability utilization. Finally, there is unwanted **spam mail**. This is often little more than a nuisance, but it does waste system resources both locally and over the Internet. It is often difficult to stop spam because the source of the messages is usually spoofed.

3.8.2: E-Mail Security Solutions

There are several protocols, services, and solutions that can be employed to add security to an existing e-mail infrastructure. These include S/MIME, MOSS, PEM, and PGP:

- **Secure Multipurpose Internet Mail Extensions (S/MIME)** implements e-mail authentication through X.509 digital certificates and privacy through Public Key Cryptography Standard (PKCS) encryption. Two types of messages can be formed using S/MIME: signed messages and enveloped messages. A signed message provides integrity and sender authentication. An enveloped message provides integrity, sender authentication, and confidentiality. All major email vendors support S/MIME.
- **MIME Object Security Services (MOSS)** can be used to provide authenticity, confidentiality, integrity, and nonrepudiation for e-mail messages. It uses Message Digest 2 (MD2) and MD5 algorithms; Rivest, Shamir, and Addleman (RSA) public key; and Data Encryption Standard (DES) to provide authentication and encryption services.
- **Privacy Enhanced Mail (PEM)**, which is an e-mail encryption mechanism, can be used that provide authentication, integrity, confidentiality, and nonrepudiation. It uses RSA, DES, and X.509.
- **Pretty Good Privacy (PGP)** is an asymmetric public-private key system that uses the **IDEA algorithm** to encrypt, decrypt, and digitally sign files and e-mail messages. It is not a standard but is widely supported on the Internet.
  - PGP creates your key pair, which is your public and private key.
  - PGP allows you to store other users' public keys on a local key ring.
  - The sender uses the recipient's public key to encrypt messages. The recipient uses his or her own private key (or secret key) to decrypt those messages.

Section 3.9: Voice Communications

With the convergence of voice, data and video, through technologies such as **Voice over IP (VoIP)**, securing voice communication is related to network security. When voice communications occur over a network infrastructure, issues of confidentiality, integrity, and authentication become important.

Normal **private branch exchange (PBX)** or **plain old telephone service (POTS)** voice communications are vulnerable to interception, eavesdropping, and tapping. Usually, physical security is required to maintain control over voice communications within the physical locations of the organization. Security of voice communications outside of the organization is usually the responsibility of the telephone company.

Many PBX systems can be exploited by malicious attackers, known as phreakers, to avoid toll charges and hide their identity. Phreakers may be able to gain unauthorized access to personal voice mailboxes, redirect messages, block access, and redirect inbound and outbound calls. Countermeasures to phreaking include logical or technical controls, administrative controls, and physical controls.

- Replace remote access or long-distance calling through the PBX with a credit card or calling card system.
- Restrict dial-in and dial-out features to only authorized users.
- Use unpublished phone numbers that are outside of the prefix block range of your voice numbers for your dial-in modems.
- Block or disable any unassigned access codes or accounts.
- Define an acceptable use policy.
- Log and audit all activities on the PBX and review the audit trails regularly.
- Disable maintenance modems and accounts.
- Change all default configurations, especially passwords and capabilities related to administrative or privileged features.
- Block remote calling.
- Deploy Direct Inward System Access (DISA) technologies to reduce PBX fraud by external parties.

The tools used by phreakers are referred to as colored boxes. These include:

- Black boxes, which are used to manipulate line voltages to steal long-distance services. They are usually custom-built circuit boards with a battery and wire clips.
- Red boxes, which are used to simulate tones of coins being deposited into a pay phone. They are usually small tape recorders.
- Blue boxes, which are used to simulate 2600 Hz tones to interact directly with telephone network trunk systems. This could be a whistle, a tape recorder, or a digital tone generator.
- White boxes, which are used to control the phone system. A white box is a DTMF or dual-tone multifrequency generator. It can be a custom-built device or one of the pieces of equipment that most telephone repair personnel use.

Topic 4: Cryptography

Algorithms are the underlying foundation of cryptography. Encryption, a type of cryptography, refers to the process of scrambling information so that the casual observer cannot read it. An algorithm is a set of instructions for mixing and rearranging an original message, called plaintext, with a message key to create a scrambled message, referred to as ciphertext. A cryptographic key is a piece of data used to encrypt plaintext to ciphertext, or ciphertext to plaintext, or both. Crypto has its origins in the Greek word kruptos, which means hidden. The objective of cryptography is to hide information so that only the intended recipients can "unhide" it. The hiding of information is called encryption, and when the information is unhidden, it is called decryption. There are two different subclasses of algorithms, block ciphers and stream ciphers. Block ciphers work on "blocks" or chunks of text in a series.

Section 4.1: Encryption

Encryption is a form of cryptography that "scrambles" plain text into unintelligible ciphertext. Encryption is the foundation of such security measures as digital signatures, digital certificates, and the public key infrastructure (PKI). Computer-based encryption techniques use keys to encrypt and decrypt data. A key is a variable that is a large binary number. Key length is measured in bits, and the more bits in a key, the more difficult the key will be to "crack." The key is only one component in the encryption process. It must be used in conjunction with an encryption algorithm to produce the cipher text. Encryption methods are usually categorized as either symmetric or asymmetric, depending on the number of keys that are used.

4.1.1: Symmetric Algorithms

A symmetric algorithm uses the same key for encrypting and decrypting data. Symmetric algorithms provide confidentiality by encrypting data or messages. Some of the past and current symmetric key encryption algorithms include Data Encryption Standard (DES), Triple DES (3DES), Advanced Encryption Standard (AES), International Data Encryption Algorithm (IDEA), Blowfish, and RC4.

- **Speed:** The algorithms used with symmetric encryption are relatively fast, so they impact system performance less and are good for encrypting large amounts of data.
- **Strength:** Symmetric algorithms are difficult to decipher without the correct algorithm. They are not easy to break. Well-tested symmetric algorithms such as 3DES and AES are nearly impossible to decipher without the correct key. Technique can be used in which encrypted data can be encrypted a second or even third time. Some of the disadvantages of using symmetric keys are as follows:
- **Poor key distribution mechanism:** There is no easy way to securely distribute a shared secret; therefore wide-scale deployment of symmetric keys is difficult.
- **Single key:** There is a single key (single shared secret); therefore if the secret is compromised, the impact is widespread. Because there is a single key that can be shared with some or many, symmetric keys are not suited to provide integrity, authentication, or nonrepudiation.
- Some of the characteristics of specific symmetric keys are as follows:
- **DES:** 56-bit key, U.S. Government standard until 1998, but not considered strong enough for today's standards, relatively slow.
- **Triple DES:** Performs 3DES operations, equivalent of 168-bit keys, more secure than DES, widely used, relatively slow.
- **AES:** Variable key lengths, latest standard for U.S. Government use, replacing DES.

- **IDEA:** 128-bit key, requires licensing for commercial use.
- **Blowfish:** Variable key length, free algorithm, extremely fast.
- **RC4:** Variable key length, stream cipher, effectively in public domain.

4.1.2: Asymmetric Algorithms

Asymmetric algorithms use different keys to encrypt and decrypt data. One method of asymmetric encryption is called public key cryptography. Public key cryptography uses two keys that form a key pair. These two keys are known as the public key and the private key. The key that encrypts the plaintext cannot be used to decrypt the ciphertext. The public key encrypts the plaintext, and the private key decrypts the ciphertext.

- **Public key**, which provided to everyone who needs to send you encrypted data.
- **Private key**, which is the key that only you possess. When a plaintext message is encrypted using the public key, only the person with the private key can decrypt the ciphertext. When a plaintext message is encrypted using the private key, it can be decrypted by everyone who possesses the public key. The person can be certain the plaintext message originated with the person who possessed the private key. Asymmetric keys provide authentication, integrity, and nonrepudiation. They can also support confidentiality when used for key management.

There are advantages and disadvantages to using asymmetric keys.

If there is a public key and a private key, the public key can be provided to anyone that you want to send you encrypted information, but only you can decrypt that information. This helps ensure data confidentiality.

You can use a private key to create a digital signature, which can be used to verify that you are who you claim to be. This helps provide an authentication method and nonrepudiation. Some of the disadvantages of using asymmetric keys are that asymmetric algorithms are generally slower than symmetric algorithms due to the increased computational complexity required encrypting and decrypting data; therefore it is not suited to provide confidentiality for large amounts of data.

Some characteristics of specific asymmetric keys are:

- **RSA:** Variable-length key, de facto standard for public key encryption.
- **Diffie-Hellman:** Variable-length key, used to securely establish a shared secret.
- **Elliptic curve cryptography:** Variable-length key, currently too slow for widespread implementation. The purpose of using cryptography is to ensure confidentiality, integrity, identification and authentication, and nonrepudiation. Given enough time, though, a hacker can decrypt the information. The strength of symmetric and asymmetric keys comes from the length of the key and the algorithm that is used to encrypt the data.
- **DES and Triple DES:** DES was based on the Lucifer algorithm invented by Horst Feistel, which never saw widespread use. Essentially, DES uses a single 64-bit key—56 bits of data and 8 bits of parity—and operates on data in 64-bit chunks. Each round consists of a substitution phase, wherein the data is substituted with pieces of the key, and a permutation phase. Substitution operations occur within S-boxes. Similarly, permutation operations, sometimes called diffusion operations, are said to occur in P-boxes. Both of these operations occur in the “F Module” of the diagram. The security of DES lies in the fact that since the substitution operations are non-linear. The permutation operations add another layer of security by scrambling the already partially encrypted message. Triple DES (3-DES) is a method that attempts to use the DES cipher in a way that increases its security. In 3-DES, two to three 56-bit DES

subkeys are connected to form a single 112- or 168-bit 3-DES key. The resulting ciphertext can withstand attacks from all currently known brute-force attacks and techniques. For 112-bit security, two different DES keys are used and one is repeated, and for 168-bit security, three different DES keys are connected together.

Section 4.2: Advanced Encryption Standard (Rijndael)

Because of its small key size of 56 bits, DES is no longer able to withstand coordinated brute-force attacks using modern cryptanalysis. Consequently, The National Institute of Standards and Technology has selected the Advanced Encryption Standard to be the authorized Federal Information Processing Standard for all non-secret communications by the U.S. government. NIST is also realizing widespread use in the private sector.

Rijndael was selected by NIST from a group that included four other finalists: MARS, RC6, Serpent, and Twofish. NIST seems resistant to side-channel attacks such as power- and timing based attacks Power- and timing-based attacks measure the time it takes to encrypt a message or the minute changes in power consumption during the encryption and decryption processes. These attacks are sufficient enough to allow hackers to recover keys used by the device. Rijndael uses iterative rounds like International Data Encryption Algorithm.

A hashing algorithm is used to provide data integrity. A hash is a one-way mathematical function (OWF) that creates a fixed-sized value. Some common hash algorithms currently in use include the MD4, MD5, and SHA-1 algorithms.

- **MD4:** Produces a 128 bit message digest very fast, appropriate for medium security usage.
- **MD5:** Produces a 128 bit message digest, fast more secure than MD4, and widely used.
- **SHA-1:** Produces a 160 bit message digest, standard for the U.S. government, but slower than MD5.

Section 4.3: Public Key Infrastructure (PKI)

Using asymmetric key pairs is simple enough to implement, but when scaled beyond a small community using a few applications. When a private key is compromised, it is difficult to locate and remove that key.

The security infrastructure created to solve these problems is known as a public key infrastructure (PKI). PKI uses asymmetric key pairs and combines software, encryption technologies, and services to provide a means of protecting the security of communications RFC 2459 defines the X.509 PKI, which is the PKI defined for use on the Internet. It is comprised of certificates, certification authorities (CAs), certificate management tools, and certificate-enabled applications.

4.3.1: Components of a PKI

A PKI uses public key encryption technologies to bind public keys to their owners and to help with reliable distribution of keys across networks. PKI provides a framework of services, technologies, protocols, and standards that enable you to deploy and manage a strong and scalable information security system. With the PKI in place, companies can conduct business electronically and be assured of the following:

- The person or process sending a transaction is the actual originator.
- The person or process receiving a transaction is the actual receiver.
- The integrity of the data has not been compromised.

4.3.2: Digital Certificates

X.509 was developed from the X.500 standard. X.500 is a directory service standard that was ratified by the International Telecommunications Union. It was intended to provide a means of developing an easy-to-use electronic directory of people that would be available to all Internet users.

The X.500 directory standard specifies a common root of a hierarchical tree. The root of the tree is depicted at the top level, and all other containers are below it. A CN= before a username represents it is a "common name," a C= precedes a "country," and an O= precedes "organization." Each X.500 local directory is considered a directory system agent. The DSA can represent either single or multiple organizations. Each DSA connects to the others through a directory information tree, which is a hierarchical naming scheme that provides the naming context for objects within a directory. X.509 is the standard used to define what makes up a digital certificate.

Digital Certificates

An electronic credential used to authenticate users.

Certification Authority (CA)

A computer that issues digital certificates, maintains a list of invalid certificates, and maintains a list of invalid CAs.

Registration Authority (RA)

An entity that is designed to verify certificate contents for a CA.

Certificate Publication Point

A location where certificates are stored and published.

4.3.2.1: Certificate Policies

A CA can issue a certificate for a number of different reasons, but must indicate exactly what the certificate will be used for. The set of rules that indicates exactly how a certificate may be used is called a certificate policy. The X.509 standard defines certificate policies as "a named set of rules that indicates the applicability of a certificate. Different entities have different security requirements. Users want a digital certificate for securing e-mail. TestKing wants a digital certificate for their online store. The Department of Defense wants a digital certificate they can use to protect secret information regarding nuclear submarines. The certificate policy is a plaintext document that is assigned a unique object.

4.3.2.2: Certificate Practice Statements

It is important to have a policy in place to state what is going to be done. A CPS describes how the CA plans to manage the certificates it issues. If a CA does not have a CPS available, users should consider finding another CA.

4.3.2.3: Revocation

Certificates are revoked when the information contained in the certificate is no longer considered valid or trusted. It happens when a company changes Internet Service Providers, moves to a new physical address, or the contact listed on the certificate has changed. Some policies are of a technical nature, some refer to the procedures used to create and manage certificates. When dealing with security systems, it is important to make sure the CA has a policy covering each item required. The most important reason to revoke a certificate is if the private key has been compromised in any way. If a key has been compromised, it should be revoked immediately. It is important to notify all certificate users of the date that the certificate will no longer be valid. After notifying users and the CA, the CA is responsible for changing the status of the certificate and notifying users that it has been revoked. If a certificate is revoked because of key compromise, you must publish the date the certificate was revoked, as well as the last date that communications were considered trustworthy. When a certificate revocation request is sent to a CA, the CA

must be able to authenticate the request with the certificate owner. Once the CA has authenticated the request, the certificate is revoked and notification is sent out.

A certificate issued by a CA includes an expiration date that defines how long the certificate is valid. If a certificate needs to be revoked before that date, the CA can be instructed to add the certificate to its CRL. When a certificate is revoked, the CA administrator must supply a reason code. PKI-enabled applications are configured to check CAs for their current CRL, and do not operate if they cannot verify that the certificate has not been added to the CRL.

4.3.3: Standards and Protocols

Without standards and protocols, a juggernaut like PKI would become unmanageable. The Public-Key Cryptography Standards (PKCS) are standard protocols used for securing the exchange of information through PKI. The list of PKCS standards was created by RSA laboratories.

- **PKCS #1: RSA Cryptography Standard** outlines the encryption of data using the RSA algorithm. The purpose of the RSA Cryptography Standard is in the development of digital signatures and digital envelopes. PKCS #1 also describes syntax for RSA public keys and private keys. The public-key syntax is used for certificates, while the private-key syntax is used for encrypting private keys.
- **PKCS #3: Diffie-Hellman Key Agreement Standard** outlines the use of the Diffie-Hellman Key Agreement, a method of sharing a secret key between two parties. The secret key is used to encrypt ongoing data transfer between the two parties. Whitfield Diffie and Martin Hellman developed the Diffie-Hellman algorithm in the 1970s as the first asymmetric cryptographic system. Diffie-Hellman overcomes the issues of symmetric key systems because management of the keys is less difficult.
- **PKCS #5: Password-Based Cryptography Standard** defines a method for encrypting a string with a secret key that is derived from a password. The result of the method is an octet string (8-character string).
- **PKCS #6: Extended-Certificate Syntax Standard** deals with extended certificates. Extended certificates are made up of the X.509 certificate plus additional attributes. The additional attributes and the X.509 certificate can be verified using a single public-key operation. The issuer that signs the extended certificate is the same as the one that signs the X.509 certificate.
- **PKCS #7: Cryptographic Message Syntax Standard** is the foundation for Secure/Multipurpose Internet Mail Extensions (S/MIME) standard. Is also compatible with Privacy-Enhanced Mail and can be used in several different architectures of key management.
- **PKCS #8: Private-Key Information Syntax Standard** describes a method of communication for private-key information that includes the use of public-key algorithms and additional attributes. PKCS #8 is primarily used for encrypting private keys when they are being transmitted between computers.
- **PKCS #9: Selected Attribute Types** defines the types of attributes for use in extended certificates (PKCS#6), digitally signed messages (PKCS#7), and private-key information (PKCS#8).
- **PKCS #10: Certification Request Syntax Standard** describes syntax for certification requests. A certification request consists of a distinguished name, a public key, and additional attributes. Certification requests are sent to a CA, which then issues the certificate.
- **PKCS #11: Cryptographic Token Interface Standard** specifies an application program interface for token devices that hold encrypted information and perform cryptographic functions, such as Smart Cards and USB pigtails.
- **PKCS #12: Personal Information Exchange Syntax Standard** specifies a portable format for storing or transporting a user's private keys and certificates. Ties into both PKCS #8 (communication of private

key information) and PKCS #11 (Cryptographic Token Interface Standard). Portable formats include diskettes, Smart Cards, and Personal Computer Memory Card International Association (PCMCIA) cards.

PKI standards and protocols are living documents, meaning they are always changing and evolving. Additional standards are proposed every day, but before they are accepted as standards they are put through rigorous testing and scrutiny.

4.3.4: Key Management Life Cycle

Certificates and keys have a life cycle. Different factors play into the lifecycle of a particular key. Many things can happen to affect the life span of a key. They may become compromised or they may be revoked or destroyed. Keys also have an expiration date. Just like a license or credit card, a key is considered valid for a certain period of time. Once the end of the usable time for the key has expired, the key must be renewed or replaced

4.3.4.1: Centralized versus Decentralized Keys

Different PKI implementations use different types of key management. The hierarchical model uses centralized key management. The key management in the hierarchical model is centralized, because all of the public keys are held within one central location. Older implementations of PGP used decentralized key management, since the keys are contained in a PGP user's key ring and no one entity is superior over another. Whether to use centralized or decentralized key management depends on the size of the organization. Under older versions of PGP, you could only hold the keys of those PGP users that you trust. For a large organization of 10,000 that requires all of their employees to use digital signatures when communicating, managing PGP keys would be impossible.

Whether using centralized management or decentralized management for keys, a secure method of storing those keys must be designed.

4.3.4.1.1: Storage

If a person left a wallet on a counter in a department store and someone took it. They would have to call their credit card companies to close out their accounts, they would have to go to the DMV to get a duplicate license, they would have to change their bank account numbers, etc. Imagine what would happen if Company X put all of their private keys into a publicly accessible File Transfer Protocol (FTP) site. Once hackers discovered that they could obtain the private keys, they could very easily listen to communications between the company and clients and decrypt and encrypt messages being passed.

4.3.4.1.2: Software Storage

With software storage of an archived key, the key can be stored on a floppy disk or other type of removable media. When you need to provide a user with a key, you can copy the key to a floppy disk and use the copy of the key to perform the operation. When the key is in use, it is loaded into active memory on the computer. To ensure the integrity of the key, it can be stored in an approved cryptographic module. When you are finished using the copy of the private key, you must destroy the media it was copied to in a secure manner. With this type of storage, distribution is relatively simple and cost effective, but it is also somewhat easier to compromise than a hardware solution.

4.3.4.1.3: Hardware Storage

With hardware storage of a key, the key is placed on a hardware storage medium, such as a smart card or hardware security module. HSMs also generate the keys on the hardware device to prevent the need to transmit the private key over a network connection or other medium. When you need to provide a user with a key, you program the smart card with the key and give the key to the user. This is a secure way to store keys and is difficult to compromise. It requires specialized equipment and is more costly than the software storage solution.

4.3.4.2: Centralized Key Management

4.3.4.2.1: Private Key Protection

Keeping private keys stored in secure location must be priority one when dealing with PKI. Many people take private keys for corporate CAs completely offline, store them in a secure place, and use them only when they need to generate a new key.

4.3.4.2.2: Key Escrow

Private key escrow is probably one of the most sensitive topics in the PKI community. Private key escrow occurs when a CA maintains a copy of the private key associated with the public key that has been signed by the CA. This allows the CA or entity to have access to all information encrypted using the public key from a user's certificate. A corporate PKI solution usually includes a key escrow element. An employee is bound by information security policies that allow a corporation to have access to all intellectual property generated by a user for the company as part of that person's terms of employment. A corporation must have the ability to access data an employee generates to maintain the operations of the business. Key escrow also helps an organization overcome the problem of lost or forgotten passwords.

4.3.4.2.3: Certificate Expiration

When a certificate is created, it is stamped with **Valid From** and **Valid To** dates. The period in between these dates is the duration of time that the certificate and key pairs are valid. Once a certificate has reached the end of its validity period, it must be either renewed or destroyed.

4.3.4.2.4: Certification Revocation List

The X.509 standard requires that CAs publish CRLs. The list in its simplest form is a published form listing the revocation status of certificates that the CA manages. There are several forms that the revocation list may take.

- A simple CRL is a container that holds the list of revoked certificates.
- A simple CRL contains the name of the CA, the time and date the CRL was published, and when the next CRL will be published.
- A simple CRL is a single file that continues to grow over time.
- The fact that only information about the certificate is included and not the certificate itself control the size of a simple CRL container.
- Delta CRLs were created to handle the issue that simple CRLs cannot—size and distribution.
- Although a simple CRL only contains certain information about the revoked certificate, it can still become a large file.

- In a Delta CRL configuration, a base CRL is sent out to all end parties to initialize their copies of the CRL. After the base CRL is sent out, updates known as deltas are sent out on a periodic basis to inform the end parties of changes.

4.3.5: M of N Control

This is the concept of backing up the public and private key material across multiple systems or devices. This method is designed to ensure that no one individual can re-create her private and public key material from the backup. The key materials are backed up and then mathematically distributed across several systems or devices. Usually three people are defined with separate job responsibilities and from separate portions of the organization. This is intended to prevent collusion between the individuals so that they do not recovering the keys without proper permission. The mathematical equation supports any number of users up to 255 for the splitting activity.

Assuming your key makes it through the entire period of time it is valid without the need for revocation, you will need to renew it. You do not have to prove your identity again to get a new certificate. If the certificate is in good standing, and you are renewing it with the same CA, you can use the old key to sign the request for the new key. The reason behind this is that since the CA trusts you based on your current credentials, there is no reason why they should not trust your request for a new key. There is a second method of renewal, called key update, where a new key is created by modifying the existing key. The key renewal process that is used will depend on the user and the requirements of the CA. The renewal process is also true of a CA's key pair. Eventually, a CA will need to renew its own set of keys. Again, a CA can use its old key to sign the new key. The PKI renewal process is performed by creating three new keys:

- The CA creates another self-signed certificate. This time, the CA signs the new public key using the old private key that is about to retire.
- Next, the CA server signs the old public keys with the new private key. This is done so that there is an overlap between when the new key comes online and when the old key expires.
- Finally, the new public key is signed with the new private key. This will be the new key that will be used after the old key expires.

The reason for this process is two-fold. Since a CA verifies the credentials of other parties, there has to be a degree of difficulty to renewing the CA's own certificate. Second, creating all of these keys makes the changeover from old keys to new keys transparent to the end user.

When a key pair and certificate are no longer valid and must be destroyed. If the key pair is used for digital signature purposes, the private key portion should be destroyed to prevent future signing activities with the key. If the key pair is used only for privacy purposes, you might need to archive a copy of the private key. This is because the private key might need to be used to decrypt archived data that was encrypted using it. The digital certificate must be added to the CRL as soon as the certificate is no longer valid. This activity takes place regardless of the archive or nonarchive status of the private key for future use. Depending on the sensitivity of the key in question, it might also be necessary to contact the individuals who use this certificate and trust the credentials it represents to inform them to no longer trust this certificate.

4.3.6: Key Usage

Key pairs are used in a variety of different functions. In most PKI implementations, only single key pairs are used.

It is sometimes necessary for a CA to generate multiple key pairs. This situation arises when there is a need to back up private keys, but the fear of a forged digital signature exists. For example, if someone is the backup operator, he is responsible for the backup of all data, including user's private keys. If he comes in after a long weekend and decides that he deserves a raise. Since he has access to all of the private keys, he can recover the CIO's private key, send a message to the Human Resources department requesting a raise, and sign in using the CIO's certificate. Since the CIO's digital signature provides non-repudiation, the Human Resources manager would have no reason to question the e-mail.

To avoid this problem, many public key infrastructures support the use of dual keys. In the example above, the CIO has two separate key pairs. The first key pair is used for authentication or encryption, while the second key pair is used for digital signatures. The private key used for authentication and encryption can still be backed up for his safekeeping. The second private key would never be backed up and would not provide the security loophole that using single keys creates. The CIO could continue using his second private key for signing e-mails without fear of the key being misused.

Topic 5: System Architecture and Models

Section 5.1: Computer Architecture

Computer architecture is the design and construction of computing systems at a logical level. It also refers to the view a programmer has of the computing system when viewed through its instruction set. The main hardware components of a computer are the **central processing unit (CPU)**, **memory**, and **input/output** devices.

5.1.1: The Central Processing Unit (CPU)

The central processing unit (CPU) is the computer component that controls the operation of the computer. It consists of an **arithmetic logic unit (ALU)**, which performs all arithmetic and logical operations, and a control unit, which performs the decoding of instructions and execution of the requested instructions. Early computers used several chips to handle the task. Some functions are still handled by support chips, which are often referred to collectively as a chip set.

The CPU requires two inputs to perform its operations: instructions and data. The data is passed to the CPU for manipulation, where it is typically processed in either supervisor or problem state. In problem state, the CPU works on the data with nonprivileged instructions. In supervisor state, the CPU executes privileged instructions.

There are two basic designs of CPUs manufactured for modern computer systems. These are:

- **Reduced Instruction Set Computing (RISC)**, which uses simple instructions that require a reduced number of clock cycles.
- **Complex Instruction Set Computing (CISC)**, which performs multiple operations for a single instruction.

> **Note:** A superscalar processor can execute multiple instructions at the same time, while a scalar processor can execute only one instruction at a time.

The CPU can also be classified based on its functionality. However, both the hardware and software must be able to support these features. These categories include: **multitasking**, which is when the CPU handles more than one task at a time; **multiprogramming**, which is similar to multitasking but the CPU alternates between the processes rather than perform them simultaneously; **multithreading**, which allows multiple concurrent tasks to be performed within a single process; and **multiprocessing**, which is a system that has more than one CPU.

Multiprocessing

There are two types of multiprocessing: symmetric multiprocessing (SMP), where the processors share a common operating system, a common data bus and memory resources; and massively parallel processing (MPP), where each processor has its own operating system, its own data bus and its own memory resources. SMP systems process simple tasks at extremely high rates, while MPP systems process very large, complex, computationally intensive tasks that can be processed as a number of subordinate parts.

5.1.2: Memory

Computer memory is used to hold data that is to be manipulated by the CPU. Computer memory can be divided into two major classes: nonvolatile memory, which is retained even when the computer is powered off; and volatile memory, which is lost when the computer loses power.

- **Read-only memory (ROM)** is an example of nonvolatile memory and is retained when the computer is powered off. ROM is usually a chip that has its contents "burned in" during the manufacturing process and cannot be altered. These ROM chips often contain "bootstrap" information and the power-on self-test (POST) diagnostics that computers use to start up prior to loading the operating system.
- **Random access memory (RAM)** is an example of volatile memory. If power to the system is lost, all data held in RAM is lost. RAM is readable and writeable memory that holds data the CPU needs when processing a task. RAM retains its contents only when power is continuously supplied to it. The term random is applied because the CPU can access or move data to and from any addressable RAM on the system.
- **Cache memory** retains recently or frequently accessed data in special, high-speed memory chip located on or close to the CPU. Cache memory is much faster than RAM, but is also much more expensive. Because the CPU is constantly requesting and using data and executing code, it requires quick access to that data. The closer the required data is to the CPU, the faster the system can locate it and execute the operation. Caches are organized into layers. The highest layer is known as the level 1 (L1) cache and is closest to the CPU. Modern computers can also have level 2 (L2) and even level 3 (L3) cache memory.
- **Virtual memory** is an area on the hard disk drive that the operating system uses as part of its memory management functions. The Windows pagefile or swap file is an example of virtual memory. The pagefile is a specially formatted file that contains data previously stored in memory but that has not been accessed recently. It can also be used when the system is low on RAM. However, accessing virtual memory is relatively slow and requires significant computer overhead. This slows down the entire system.

Data that is held in memory is directly accessible to the CPU, and is called active memory. Data is located outside the computer system's active memory, is said to be held in storage or in secondary memory.

5.1.3: Data Storage

Data storage devices are magnetic and optical devices or media that are used to store data. Storage devices include hard drives, floppy disks, magnetic tapes, compact discs (CDs), digital video disks (DVDs), flash memory cards and flash drives.

There are three main concerns when it comes to the security of data storage devices:

- Data remnants may remain on data storage devices even after it has been erased. This data is retrievable. It is also possible to retrieve data from a disk that has been reformatted.
- Data storage devices are also prone to theft.
- Data storage devices are also prone to unauthorized access, particularly removable storage devices, such as floppy disks, flash memory cards and flash drives.

5.1.4: Input and Output Devices

Input and output devices can present security risks to a system. A technology known as **TEMPEST**, for example, can compromise the security of data displayed on a monitor. TEMPEST allows the electronic emanations from the monitor to be read from another location. Print outs from a printer can also represent a security risk as they may be intercepted or read by unauthorized users.

Section 5.2: Security Policy and Computer Architecture

A security policy is important to the design and implementation of information systems. The security policy is a document that defines a set of rules, practices, and procedures that describe how the system should manage, protect, and distribute sensitive information. As such, it informs and guides the design, development, implementation, testing, and maintenance of the information system.

- The **principle of least privilege** is very important to the design of computers and operating systems. When designing operating system processes, you should ensure that they run in user mode whenever possible. The greater the number of processes that execute in privileged mode, the higher the number of potential vulnerabilities that a malicious individual could exploit to gain supervisory access to the system.
- The principle of **separation of privilege** requires the use of granular access permissions; that is, different permissions for each type of privileged operation. This allows designers to assign some processes rights to perform certain functions without granting them unrestricted access to the system.
- **Accountability** is important in the security design.

Modern computing is based on a client/server model where users operate independent fully functional desktop computers but also access services and resources on networked servers. This means that clients have computing and storage capabilities. Therefore, all computers must be properly secured and protected, as well as the network links between clients and servers.

5.2.1: Vulnerabilities

Desktop systems can contain sensitive information that may be at some risk of being exposed and must therefore be protected. Individual users may lack general security awareness which the underlying architecture has to compensate for. Client computers can provide access to critical information systems elsewhere in a distributed network environment because users require access to networked servers and services.

Communications equipment can also provide vulnerable points of entry into a distributed environment. A modem that is attached to a desktop computer that is also attached to an organization's network can make that network vulnerable to dial-in attacks. Likewise, users who download data from the Internet increase the risk of infecting the system with viruses, Trojan horses, and back doors. Client computers and their storage devices may not be secure from physical intrusion or theft. While data on client computers may not be protected with a proper backup.

5.2.2: Safeguards

Distributed environments require numerous safeguards to implement appropriate security and to ensure that vulnerabilities are eliminated, mitigated, or remedied. Clients must be subjected to policies that impose safeguards on their contents and their user's activities. These safeguards include:

- The screening of e-mail to prevent it from introducing malicious software to the system and the implementation of e-mail policies that defines appropriate use and limits potential liability.
- Download/upload policies must be implemented to allow for incoming and outgoing data to be screened and suspect materials to be blocked.
- Robust access controls, which could include multifactor authentication and/or biometrics, should be used to restrict access to client computers and to prevent unauthorized access to servers and services.

- Graphic user interface mechanisms and database management systems should be installed, and their use required, to restrict and manage access to critical information.
- The encryption of sensitive files and data stored on client computers.
- The isolation of processes that run in user and supervisory mode so that unauthorized access to high-privilege processes and capabilities is prevented.
- Protection domains should be created so that compromise of a client computer will not compromise the entire network.
- Disks and other sensitive materials should be protected from unauthorized access.
- Client computers should be backed up regularly.
- Security awareness training should be provided for client computer users.
- Client computers and their storage devices should also be protection against environmental hazards.
- Client computers should be included in disaster recovery and business continuity planning.

Section 5.3: Security Mechanisms

To ensure security, you have should have mechanisms in place to control processes and applications. These mechanisms could include process isolation, protection rings, and trusted computer base (TCB).

5.3.1: Process Isolation

Process isolation, which is performed by the operating system, maintains a high level of system trust by enforcing memory boundaries. Without process isolation, it would not be possible to prevent one process from spilling over into another process's memory space, corrupting data or possibly rendering the system unstable.

The operating system must also prevent unauthorized users from accessing areas of the system to which they should not have access. This is often achieved by means of a virtual machine, which allows the user to believe that they have the use of the entire system, but in reality, processes are completely isolated. Some systems also implement hardware isolation for even greater security. With hardware isolation, the processes are segmented both logically and physically.

5.3.2: Single-State and Multistate Systems

Single-state and multistate systems were developed to meet the requirements of handling sensitive government information with categories such as sensitive, secret, and top secret. These systems have a bearing on whether the sensitive information that is processed and stored on a system is managed by the system itself, or by an administrator.

- **Single-state systems**, which are also referred to as **dedicated systems**, are designed to handle only one category of information and are dedicated to one mode of operation. Here the responsibility of management falls upon the administrator who must develop the policy and procedures to manage this system. The administrator must also determine who has access to the system and what type of access they should have.
- **Multistate systems** allow more than one user to log in to the system and access various categories of information, depending upon the users' levels of clearance. Multistate systems can operate as a compartmentalized system and can segment data on a need-to-know basis.

5.3.4: Rings of Protection

The operating system relies on rings of protection to determine a user's level of clearance. It provides the operating system with various levels at which to execute code or restrict its access. The rings are organized into domains, with the most privileged domain located in the center and the least-privileged domain in the outermost ring. This is illustrated in Figure 5.1.

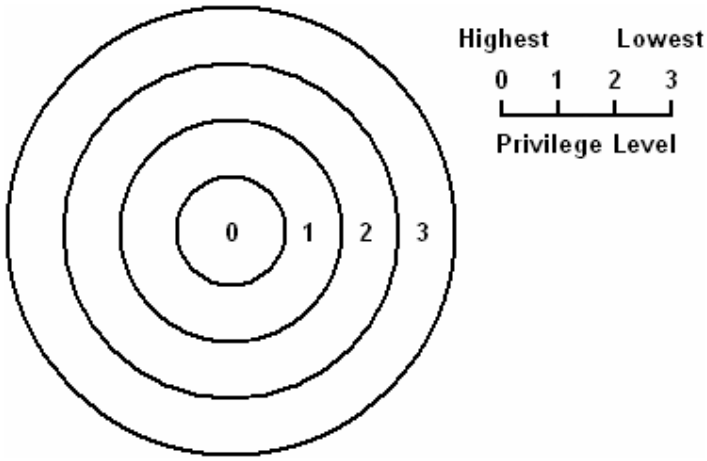

Figure 5.1: Rings of Protection

- Layer 0 is the most trusted level. The operating system kernel resides at this level. Any process running at layer 0 is said to be operating in privileged mode.
- Layer 1 contains non-privileged portions of the operating system.
- Layer 2 is where I/O drivers, low level operations, and utilities reside.
- Layer 3 is where applications and processes operate. This is the level at which individuals usually interact with the operating system. Applications operating here are said to be working in user mode.

Thus, access rights decrease as the ring number increases, thus, the most trusted processes reside in the center rings and system components are placed in the appropriate ring according to the principle of least privilege. This ensures that the processes have only the minimum privileges necessary to perform their functions.

5.3.5: Trusted Computer Base (TCB)

The trusted computer base (TCB), defined in the U.S. Department of Defense standard known as "the Orange Book" (DoD Standard 5200.28), is the sum of all the protection mechanisms within a computer, including hardware, software, controls, and processes, that work together to form a trusted base to enforce an organization's security policy. It is the part of an information system that can be trusted to adhere to and enforce the security policy. The TCB is responsible for confidentiality and integrity and monitors four basic functions:

- **Input/output operations**, which are a security concern because operations from the outermost rings might need to interface with rings of greater protection.
- **Execution domain switching**, which are a security concern because applications running in one domain often invoke applications or services in other domains.
- **Memory protection**, which must be monitored to verify confidentiality and integrity in storage.
- **Process activation**, which are a security concern because registers, process status information, and file access lists are vulnerable to loss of confidentiality in a multiprogramming environment.

The **reference monitor** is an important component of the TCB and is an abstract machine that is used to validate access to objects by authorized subjects. The reference monitor is implemented by the **security**

kernel, which handles all user/application requests for access to system resources. The security kernel must control all access, be protected from modification or change, and be verified and tested to be correct.

Section 5.4: Security Models

Information security models are means of formalizing security policies as they are intended to provide an explicit set of rules that a computer can follow to implement the fundamental security concepts, processes, and procedures that make up a security policy. These models can be abstract or intuitive.

5.4.1: State Machine Model

The **state machine model** describes a system that is always secure regardless of the state it is in. According to the state machine model, a **state** is a snapshot of a system at a specific moment in time. The state machine model is based on the computer science definition of a **finite state machine (FSM)**, which combines an external input with an internal machine state to model all types of systems, including parsers, decoders, and interpreters. Given an input and a state, an FSM transitions to another state and may create an output. A transition occurs when accepting input or producing output and always results in a new state. All state transitions must be evaluated and if all aspects of the state meet the requirements of the security policy, then the state is considered secure. When each possible state transitions results in another secure state, the system can be called a secure state machine. Many other security models are based on the secure state concept.

5.4.2: Bell-LaPadula Model

The Bell-LaPadula Model was developed to formalize the U.S. Department of Defense (DoD) multi-level security policy. The DoD classifies resources into four different levels. These are, in order form least sensitive to most sensitive: **Unclassified**, **Confidential**, **Secret**, and **Top Secret**. According to the Bell-LaPadula model, a subject with any level of clearance can access resources at or below its clearance level. However, only those resources that a subject requires access to are made available. Thus, a subject cleared for Confidential can access only the Confidential-labeled documents that are necessary for that subject to perform an assigned job function. With these restrictions, the Bell-LaPadula model maintains the confidentiality of objects. It does not address integrity or availability of objects.

Subjects

A subject is an active entity that is seeking rights to a resource or object. A subject can be a person, a program, or a process.

Objects

An object is a passive entity, such as a file or a storage resource. In some cases, an item can be a subject in one context and an object in another.

The Bell-LaPadula model is based on the **state machine model**. It also employs **mandatory access controls** and the **lattice model**. The lattice tiers are the **classification levels** used by the security policy of the organization. In this model, secure states are delimited by two rules, which are called properties:

- The **Simple Security Property (SS Property)**, which states that a subject at a specific classification level cannot read data with a higher classification level.
- The *** Security Property (* Property)**, which states that a subject at a specific classification level cannot write data to a lower classification level.

The Bell-LaPadula does not address integrity or availability, access control management, and file sharing. It also does not prevent covert channels, which are means by which data can be communicated outside of normal, expected, or detectable methods.

5.4.3: Biba Integrity Model

The **Biba model** was developed as a direct analogue to the Bell-LaPadula model and is also a state machine model based on a classification lattice with mandatory access controls. It was developed to address three integrity issues:

- The prevention of object modification by unauthorized subjects.
- The prevention of unauthorized object modification by authorized subjects.
- The protection of internal and external object consistency.

In this model there are three axioms:

- The **Simple Integrity Axiom (SI Axiom)**, which states that a subject at a specific classification level cannot read data with a lower classification level.
- The *** Integrity Axiom (* Axiom)**, which states that a subject at a specific classification level cannot write data to a higher classification level.
- A subject at one level of integrity cannot invoke a subject at a higher level of integrity.

The Biba model only addresses integrity, not confidentiality or availability. It focuses on protecting objects only from external threats and assumes that internal threats are handled programmatically. It also does not address access control management, nor does it provide a way to assign or change an object's or subject's classification level. In addition, it does not prevent covert channels.

5.4.4: Clark-Wilson Integrity Model

The **Clark-Wilson model** is an integrity model that was developed after the Biba model. It approaches integrity protection from a different perspective. Rather than employing a lattice structure, it uses a three-part relationship of **subject**/**program**/**object** known as a **triple**. In this relationship, subjects can only access objects through programs and do not have direct access to objects. The Clark-Wilson model provides integrity through the use of two principles: well-formed transactions and separation of duties.

- **Well-formed transactions** take the form of programs. A subject is able to access objects only by using a program. Each program has specific limitations on what it can and cannot do to an object, effectively limiting the subject's capabilities. If the programs are properly designed, then the triple relationship provides a means to protect the integrity of the object.
- **Separation of duties** is the practice of dividing critical functions into two or more parts. Each part must be completed by a different subject. This prevents authorized subjects from making unauthorized modifications to objects, further protecting the integrity of the object.
- The Clark-Wilson model requires **auditing** in addition to these two principles. Auditing tracks changes and access to objects as well as inputs from outside the system.

5.4.5: Information Flow Model

The **information flow model** is based on a state machine model, and consists of objects, state transitions, and lattice states. Information flow models are designed to prevent unauthorized, insecure, or restricted information flow, which can be between subjects and objects at the same classification level, or between subjects and objects at different classification levels. It allows all authorized information flows, either within the same classification level or between classification levels, while preventing all unauthorized information flows, either within the same classification level or between classification levels.

The Bell-LaPadula model and the Biba model are both information flow models. Bell-LaPadula is concerned with preventing information from flowing from a high security level to a low security level. Biba is concerned with preventing information from flowing from a low security level to a high security level.

5.4.6: Noninterference Model

The **noninterference model** is based on the information flow model but is concerned with how the actions of a subject at a higher security level affect the system state or actions of a subject at a lower security level.

In this model, the actions of a subject at a higher security level should not affect the actions of a subject at a lower security level and should not even be noticed by a subject at a lower security level.

5.4.7: Take-Grant Model

The **Take-Grant model** is a confidentiality-based model that uses a directed graph to specify the rights that can be passed from one subject to another or from a subject to an object. This model allows subjects with the take right to take rights from other subjects. Subjects with the grant right can grant rights to other subjects.

5.4.8: Access Control Matrix

An access control matrix is a table that lists a subject's access rights on an object. A subject's access rights can be of the type read, write, and execute. Each column of the access control matrix is called an **Access Control** List (ACL) while each row is called a **capability list**. An ACL is linked to the object and lists actions each subject is allowed to perform on a specific object. A capability list is linked to the subject and lists the actions that a specific subject is allowed to perform on each object. The access matrix model supports discretionary access control because the entries in the matrix are at the discretion of the individual(s) who have the authorization authority over the table.

5.4.9: Brewer and Nash Model

The Brewer and Nash model is similar to the Bell-LaPadula model and is also referred to as the **Chinese Wall model**. This model was created to permit access controls to change dynamically based on a user's previous activity. This model applies to a single integrated database; it seeks to create security domains that are sensitive to the notion of **conflict of interest (COI)**. This model creates a class of data that defines which security domains are potentially in conflict and prevents any subject with access to one domain that belongs to a specific conflict class from accessing any other domain that belongs to the same conflict class. Thus, this model uses the principle of data isolation within each conflict class to keep users out of potential conflict of interest situations.

Topic 6: Operational Security

The operations security addresses the day-to-day activities that are required to maintain the confidentiality, integrity and availability of the system after it has been designed and deployed. This involves using hardware controls, media controls, and subject controls that are designed to protect against asset threats, as well as daily activities such as responding to attacks and violations, good administrative management and control, and establishing a threshold to determine notable violations.

> **Note:** Violations to operational security are not always of a malicious nature. These violations could be accidental or system failures. Therefore, operational security must also be prepared to deal with occurrences of this nature.

Section 6.1: Employees and Operational Security

Employees can affect operational security. Therefore, when hiring an individual, you should validate the individual's CV. Particularly in relation to of level of education and skills. You should also perform a background check, if needed. You should hire the individual under a probationary period, specify whether individual has to obtain special qualifications or security clearances for the job, and have the individual sign a noncompete agreement, a nondisclosure agreement, and possibly a nonsolicitation agreement.

Once an individual has been hired, there are additional operational security controls that can be put in place. This includes instituting a new hire orientation, separation of duties, job rotation, least privilege, mandatory vacations, audit controls, and effective termination practices.

6.1.1: New-Hire Orientation

A **new-hire orientation** training program can be used to ensure that new employees become aware of and familiar with and the organization's policies to perform. The aim of the new-hire orientation training program should be to teach the new employees about the organization's established security policies and procedures, and to inform them of acceptable use policies.

Thereafter you can keep security awareness up by sending periodic security-awareness emails or newsletters that reinforce the practices of good security. You may also hold regular policy reviews so that employees can review current policies and receive a signed copy that they have agreed to.

6.1.2: Separation of Duties

The **separation of duties** is the practice of dividing a job task into smaller components so that more than one person is required to complete a task. This concept closely related to the principle of least privilege and prevents authorized subjects from making unauthorized modifications to objects, further protecting the integrity of the object

6.1.3: Job Rotation

In addition to providing job redundancy and backup, job rotation also allows an organization to identify fraudulent activities more easily

6.1.4: Least Privilege

The principle of least privilege demands that employees have access only to the resources they require to accomplish their required job tasks. This reduces the opportunity for resource misuse. However, over time least privilege can result in authorization creep, in which employees move from job to job and keep picking up more rights and access. The rights and access they no longer need should be removed.

6.1.5: Mandatory Vacations

Employees that never take leave are not always good workers. They may not have taken vacation because they are performing fraudulent activities. By remaining at work, these employees are able and available to provide cover for their scheme. Such fraudulent schemes could be uncover when the employees are required to take their vacations. A week provides plenty of time for illicit activities to be discovered.

6.1.6: Termination

Employee termination sometimes is necessary; however, standardized termination procedures should be used so as to protect the organization and its resources. Standardized termination procedures ensure that all employees are treated equally and that employees do not have the opportunity to destroy or damage company property. Steps that should be incorporate into this procedure include:

- Disabling computer access at the time of notification
- Monitoring the employee while he or she packs belongings
- Ensuring that at no time the employee is left alone after the termination process
- Verifying that the employee returns company identification and any company property, including access tokens and laptops
- Escorting the employee from the building

Section 6.2: Threats, Vulnerabilities and Attacks

A threat is any event that can cause damage to a system and can create a loss of confidentiality, availability, or integrity. Threats can be malicious or accidental. A vulnerability, on the other hand, is an inherent weakness in a system that can be exploited by a threat. Reducing the vulnerable of a system can reduce risk and can also reduce the impact of threats on the system.

6.2.1: Threats

Threats can be grouped into several categories, including malicious activities, accidental loss and inappropriate actions.

6.2.1.1: Malicious Activities

Malicious activities are intentional threats usually for personal financial gain or for destruction. Malicious activities include **eavesdropping, which includes data** scavenging, traffic or trend analysis, social engineering, economic or political espionage, sniffing, dumpster diving, keystroke monitoring, and shoulder surfing, and are used to gain information or to create a foundation for a later attack; **fraud**, which includes collusion, falsified transactions, data manipulation, and other altering of data integrity for gain; **theft**, which includes the theft of information or trade secrets for profit or unauthorized disclosure, and physical theft of

hardware or software; **sabotage**, which includes denial of service (DoS), production delays, and data integrity sabotage; as well as **external attack**, which includes malicious cracking, scanning, and probing to gain infrastructure information, demon dialing to locate an unsecured modem line, and the insertion of a malicious code or virus.

6.2.1.2: Accidental Loss

Accidental loss is a loss that is incurred unintentionally. Accidental loss can include **input errors and omissions** by an operator, or **transaction processing errors** that are introduced into the data through faulty application programming or processing procedures.

6.2.1.3: Inappropriate Activities

Inappropriate activities may not be of a criminal nature but might be grounds for dismissal. These include using organizational systems to store **inappropriate content** such as pornography, entertainment, political, or violent content; **waste of organizational resources**; using email or other resources for the **inappropriate distribution** of material such as pornography, entertainment, political, violent content, material that may constitute sexual or racial harassment; and the **abuse of privileges**, which includes using unauthorized access levels to violate the confidentiality of sensitive company information.

6.2.2: Vulnerabilities and Attacks

6.2.2.1: Traffic Analysis

Traffic analysis, which is also referred to as trend analysis, involves analyzing data characteristics such as message length and message frequency, and the patterns of transmissions. It is used by an attacker to infer information that might be useful to that attacker. Countermeasures to traffic analysis include:

- **Padding messages**, which is the filling empty space in the data to ensure that all messages are of a uniform data size.
- **Sending noise**, which is the transmission of non-informational data elements together with real information to disguise the real message.
- **Covert channel analysis**, which is the analysis of covert channels. Covert Channels are information paths that are not normally used for communication within a system and are therefore not protected by the system's normal security mechanisms. Covert Channel Analysis is discussed in more detail in Section 6.5.1.

6.2.2.2: Default and Maintenance Accounts

Default and maintenance accounts are vulnerabilities that can be used to break into information systems, especially default and maintenance accounts that still have factory-set or easily guessed passwords. Physical access to the hardware by maintenance personnel can also constitute a security violation.

6.2.2.3: Data-Scavenging Attacks

Data scavenging is the process of gathering bits of data over time and piecing them together. There are two common types of data-scavenging attacks:

- **Keyboard Attacks**, which uses normal utilities and tools to glean information that is available to normal system users who are sitting at the keyboard.
- **Laboratory Attacks**, which uses very complex and precise electronic equipment.

6.2.2.4: Initial Program Load Vulnerabilities

The initial installation of a system is referred to as the initial program load (IPL) and presents very specific system vulnerabilities. During the IPL, the system administrator brings up the facility's system and has the ability to put a system into a single-user mode, without full security features. In this state, the system administrator could load unauthorized programs or data, reset passwords, rename various resources, reset the system's time and date, and reassign the data ports or communications lines. In a local area network (LAN), a system administrator could also bypass the operating system's security settings by starting the system from a tape, CD-ROM, or floppy disk.

6.2.2.5: Social Engineering

In social engineering, an attacker uses social skills to obtain information needed to compromise information systems from an unsuspecting user. This information could include passwords or PIN numbers that can be used to gain access to a system. Social engineering can be accomplished by:

- **Impersonation**, in which the attacker impersonates someone in authority and uses that person's position to solicit information or to convince the unsuspecting user to alter system settings.
- **Intimidation**, which includes browbeating the user with harsh language or threatening behavior to permit access or release information.
- **Flattery**, which is positive reinforcement used to coerce the user into giving access or information for system access

6.2.2.6: Network Address Hijacking

An attacker may be able to reroute data traffic from a server or network device to his or her personal system, by device address modification or by network address hijacking. This diversion enables the intruder to capture traffic to and from the devices for data analysis or modification or to obtain the password file from the server and gain access to user accounts. By rerouting the data output, the intruder can obtain supervisory terminal functions and bypass the system logs.

Section 6.3: Auditing, Monitoring and Intrusion Detection

Operational security requires the regular reviewing an operational system to ensure that security controls are functioning correctly and effectively. This can be accomplished by means of regular auditing and monitoring. Both auditing and monitoring rely on accountability.

6.3.1: Auditing and Audit Trails

Effective auditing is impossible without proper accountability, which is maintained by logging the activities of users and system services that maintain the operating environment and the security mechanisms. If you cannot verify that individual users perform specific actions, and hold them accountable for those actions, auditing becomes useless and security policies cannot be enforced. Logging can help reconstruct events, provide evidence for prosecution, and produce problem reports and analysis. The process of analyzing logs is called **auditing** and is usually native features of an operating system.

The **audit trails** are created by logging security events and is a set of records that provide documentary evidence of user actions. These audit trails may be limited to specific events, or they may encompass all of the activities on a system and can be used to identify whether a user has violated security policies. They

allow a security practitioner to trace user activity over time, and include information about additions, deletions, or modifications to the data within a system. However, audit trails are not preventative controls as they are usually examined after the event.

> **Note:** Audit controls are detective controls that are usually implemented to detect illegal activities. Accountability, on the other hand, is the ability to track specific actions, transactions, changes, and resource usage to a specific user within the system.

6.3.2: Monitoring

System monitoring is integral to almost all of the domains of information security. The main purpose of monitoring is the identification of problems such as unauthorized or abnormal computer usage. Indeed, most organizations use network utilities, such as Snort and TCPdump, to monitor network traffic for suspicious activity and anomalies. Failure recognition and response, which includes reporting mechanisms, is an important part of monitoring. An intrusion-detection system (IDS) is another monitoring mechanism. It is a technical detective access control system designed to continuously monitor the network activities and to detect any scanning and probing activities, or patterns that appear to be attempts at unauthorized access to the information system in real-time. It can also be configured to scan for attacks, track an attacker's movements, alert an administrator to an ongoing attack, test the system for possible vulnerabilities, and can be configured to put preventative measures into place to stop any additional unauthorized access. IDSs can also be used to detect system failures and system performance. Attacks detected by an IDS can come from external connections, viruses, malicious code, trusted internal users attempting to perform unauthorized activities, and unauthorized access attempts from trusted locations. Intrusion Detection Systems were discussed in more detail in Section 2.6.2.

Clipping levels play an important role in system monitoring. It allows users to make the occasional mistake before investigation or notification begins. It acts as a violations threshold before violations are recorded or some other type of response occurs. Once that threshold is exceeded, investigation or notification begins.

Section 6.4: Controls for Operational Security

Operational security is implemented through different types of controls. These controls provide different levels of protection and can be placed into six broad categories:

- **Preventive controls**, which are designed to reduce the impact and frequency of accidental errors and to prevent unauthorized access to the system. Data validation mechanisms review procedures are examples of preventative operational security controls.
- **Detective controls**, which are designed to detect errors once they have occurred. These controls operate after the fact and can be used to track an unauthorized transaction for prosecution, or to reduce the impact of an error on the system by identifying it early. An audit trail is an example of a detective operational security control.
- **Corrective or recovery controls**, which are designed to mitigate against the impact of a loss event through data recovery procedures. Redundant systems, RAID and tape backup are examples of corrective operational security controls.
- **Deterrent or directive controls**, which are designed to encourage compliance with external controls and to discourage violations. These controls are meant to complement other types of controls. An administrative policy stating that those who place unauthorized modems or wireless devices on the network could be fired is an example of a deterrent operational security control.

- **Application controls**, which are the controls that are designed into a software application to minimize and detect the software's operational irregularities.
- **Transaction controls**, which are designed to provide control over the various stages of a data transaction. There are several types of transaction controls:
  - **Input controls**, which are designed to ensure that transactions are properly inputted and are inputted only once into the system.
  - **Processing controls**, which are designed to guarantee that transactions are valid and accurate and that wrong entries are reprocessed correctly and promptly.
  - **Output controls**, which are designed to protect the confidentiality of an output and verifies the integrity of an output by comparing the input transaction with the output data.
  - **Change controls**, which are designed to preserve data integrity in a system while changes are made to the configuration.
  - **Test controls**, which are implemented during the testing of a system to prevent violations of confidentiality and to ensure a transaction's integrity.

Section 6.5: Orange Book Controls

The Orange Book is one of the National Security Agency's Rainbow Series of books on evaluating "Trusted Computer Systems". This is the main book in the Rainbow Series and defines the **Trusted Computer System Evaluation Criteria (TCSEC)**. As illustrated by Table 6.1, the TCSEC defines hierarchical classes of security by the letters D for the least secure through A for the most secure.

Table 6.1: TCSEC Hierarchical Classes of Security

Class	Level of Protection
D	Minimal protection
C (C1, C2)	Discretionary protection
B (B1, B2, B3)	Mandatory protection
A (A1)	Verified protection

The Orange Book also defines assurance requirements for secure computer operations that are meant to ensure that a trusted computing base's security policy has been correctly implemented and that the system's security features have accurately implemented that policy. Two types of assurances are defined in the Orange Book. These are:

- **Operational assurance**, which focuses on the basic features and architecture of a system. These include **system architecture**, **system integrity**, **covert channel analysis**, **trusted facility management**, and **trusted recovery**.
- **Life cycle assurance**, which focuses on the controls and standards that are necessary for building and maintaining a system. These include **security testing**, **design specification and testing**, **configuration management**, and **trusted distribution**.

Channels

A channel is a communication channel within a system that allows a process to transfer information, but it is also used to refer to the mechanism by which the

6.5.1: Covert Channel Analysis

communication channel is effected.

A covert channel is a communication channel within a system that is not normally used for communication and is therefore not protected by the system's security mechanisms. It is therefore a vulnerability that could be exploited to violate a system's security policy. There are two common types of covert channels: covert storage channels and covert timing channels.

- **Covert storage channels** transfer information by changing data on a resource, such as a hard disk drive, that is shared by two subjects at different security levels. This can be accomplished by a program transferring information to a less secure program by changing the amount or the patterns of free hard disk space, or by changing the characteristics of a file.
- **Covert timing channels** are covert channels in which one process signals information to another by manipulating one of its observable system resources in such a way that it affects the real response time observed by the second process. This usually takes advantage of some kind of system clock or timing device in a system. Information is then transferred by using timing elements such as the elapsed time required to perform an operation, the amount of CPU time expended, or the time occurring between two events.

6.5.2: Trusted Facility Management

Trusted facility management is the appointment of a specific user to administer the security-related functions of a system. This must serve requirements for B2 systems and B3 systems. The B2 systems require that the trusted computing base support separate operator and administrator functions, while the B3 systems require that the functions performed in the role of a security administrator be clearly identified. This requires that the security administrator should only be able to perform security administrator functions after taking a distinct auditable action to assume the security administrator role on the system. Other functions that can be performed in the security administrator should be limited strictly to those functions that are essential to the security role.

In addition, trusted facility management relies on the concept of **least privilege** and is also related to the **separation of duties** and **need to know** concepts.

6.5.3: Trusted Recovery

A system failure represents a serious security risk because the security controls might be bypassed when the system is not functioning normally. However, trusted recovery is designed to ensure that the security of a system cannot be violated in the event of such a system failure. It required for B3-level and A1-level systems and ensures that the system can be restarted without compromising its required protection levels and that it can recover and roll back without being compromised after the failure. Two activities are involved in trusted recovery: preparing for system failure and recovering from system failure.

6.5.3.1: Failure Preparation

Preparing for system failure involves performing regular backups of all critical data. This preparation must enable the data recovery in a protected and orderly manner while ensuring the continued security of the system. These procedures might also be required if a system problem, such as a missing resource, an inconsistent database, or any kind of compromise, is detected, or if the system needs to be stopped and restarted.

6.5.3.2: System Recovery

Secure system recovery procedures include rebooting the system into a single-user mode so that no other user access is enabled at this time; recovering all file systems that were active at the time of the system failure; restoring any missing or damaged files from the most recent backups; recovering the required security; and verifying the integrity of security-critical files, such as the system password file. Once these steps have been performed and the system's data is secure, users can be allowed to access the system.

Section 6.6: Operations Controls

Operations controls are the procedures used to ensure operational security. These include resource protection, hardware controls, software controls, privileged-entity controls, media controls, and physical access controls.

6.6.1: Resource Protection

Resource protection is the protection, from both loss and compromise, of an organization's computing resources, such as hardware, software, and data that is owned and used by the organization. Resource protection is designed to reduce the possibility of damage that can result from the unauthorized access and/or alteration of data by limiting the opportunities for its misuse.

- **Hardware resources** that require protection include communications devices such as routers, firewalls, gateways, switches, modems, and access servers; storage devices such as floppy disks, removable drives, external hard drives, tapes, and cartridges; processing systems such as file servers, mail servers, Internet servers, backup servers, and tape drives; standalone computers; and printers and fax machines.
- **Software resources** that require protection include program libraries and source code; software application packages; and operating system software and systems utilities.
- **Data resources** that require protection include backup data; user data files; password files; operating data directories; and system logs and audit trails.

6.6.2: Hardware Controls

Hardware controls includes controls over hardware maintenance; hardware accounts; diagnostic ports and physical hardware.

- **Hardware maintenance** usually requires that by support and operations staff, vendors, or service providers have physical or logical access to a system. Security controls are required during this access and could include background investigations of the service personnel and supervising and escorting the maintenance personnel.
- Most computer systems have built-in **maintenance accounts** that are usually supervisor-level accounts created at the factory with default passwords that are widely known. These passwords, and where possible, the account name, should be changed. Alternatively, the account should be disabled until it is needed. If an account is used remotely, authentication of the maintenance provider can be performed by using callback or encryption.
- Most systems have **diagnostic ports** through which troubleshooting can be performed. This usually provides direct access the hardware. These ports should be used only by authorized personnel and should not enable either internal or external unauthorized access.
- **Physical controls**, including locks and alarms are required for data processing hardware components, including operator terminals and keyboards, media storage cabinets, server equipment, data centers, modem pools, and telecommunication circuit rooms

6.6.3: Software Controls

Software controls involves software support, and controlling the software that is and can be used in a system. Elements of software controls include anti-virus management, software testing, software utilities, software storage, and backup controls.

- **Anti-virus management** involves controlling which applications and utilities can be loaded or executed on a system so as to limit the vulnerability to viruses, unexpected software interactions, and the subversion of security controls.
- Vigorous **software testing** is required to determine compatibility of custom software applications with the system and to identify any unforeseen or unexpected software interactions. Software testing should also be applied to software upgrades.
- System utilities and **software utilities** can compromise the integrity of the operations system as well as logical access controls. Therefore, it is important that the use of system utilities as well as software utilities be controlled by a security policy.
- Secure **software storage** requires the implementation of both logical and physical access controls to ensure that the software and copies of backups have not been altered without proper authorization.
- **Backup controls** are required to ensure that backup data is stored securely and to test the restore accuracy of a backup system.

6.6.4: Privileged Entity Controls

Privileged entity access, which is also referred to as privileged operations functions, is the extended or special access to computing resources given to operators and system administrators. This extended or special access can often be divided into distinct classes, making it possible to assign users to a class based on their job title or job requirements. Examples of privileged entity operator functions include special access to system commands; access to special parameters; and access to the system control program.

6.6.5: Media Resource Protection

Media resource protection can be classified as media security controls, which are implemented to prevent any threat to confidentiality, integrity and authenticity, and media viability controls, which are implemented to preserve the proper working condition of the media.

6.6.5.1: Media Security Controls

Media security controls are designed to prevent the loss of sensitive information when the media is stored outside the system. This can be accomplished through **logging** the use of data media, which provides accountability and assists in physical inventory control; physical **access control**, which is used to prevent unauthorized personnel from accessing the media; and **sanitization** of the media to prevent data remnants and to ensure safe and proper **disposal** of the data media.

Three techniques can be used for media sanitization, namely overwriting, degaussing, and destruction.

- **Overwriting** is the process of copying new data to the media. This, however, simply copying new data to the media is not recommended because the new data may not completely overwrite the old data. In addition, bad sectors on the media may not permit the proper overwriting of the old data. The DoD

requires that the media be overwritten with a pattern, then its complement, and finally with another pattern. The DoD also requires that a character be written to all data locations on the disk.

- **Degaussing** is the best method for sanitizing magnetic media. It is the process of erasing the magnetic media and returning it to its initial virgin state. Degaussing may be accomplished by AC erasure or DC erasure:
  - In **AC erasure**, the media is degaussed by applying an alternating field that is reduced in amplitude over time from an initial high value.
  - In **DC erasure**, the media is saturated by applying a unidirectional field.
- The physical **destruction** of paper reports, diskettes and optical media are required before disposal. Disposal techniques can include shredding or burning documentation, physically breaking CD-ROMS and diskettes, or destroying them with acid. However, to retain security, paper reports should be destroyed by personnel with the proper level of security clearance.

6.6.5.2: Media Viability Controls

The viability of the data media can be protected by numerous physical controls. The aim of these controls is to protect the media from damage during handling and transportation as well as during short-term or long-term storage. Proper marking and labeling of the media is also required in the event of a system recovery process. The labels can be used to identify media with special handling instructions or to log serial numbers or bar codes for retrieval during a system recovery. Proper handling of the media is also important and involves handling the media with care, cleanliness of the media and the protection from physical damage to the media during transportation to the archive sites. In addition, the media should be stored in a clean storage environment of tolerable heat and humidity levels as data media is sensitive to heat, moisture, magnetism, smoke, and dust.

6.6.6: Physical Access Controls

The control of physical access to the hardware and software resources is also required. All personnel require some sort of control and accountability when accessing physical resources. However, some personnel will require special physical access to perform their job functions. These include IT department personnel, cleaning and maintenance personnel, service contract personnel, as well as consultants, contractors, and temporary staff.

Topic 7: Application and System Development

Because of the possible exploitation of vulnerabilities in software applications, the applications need to be well written to ensure good security. This requires that applications be developed with a strong emphasis on security.

Section 7.1: Malicious Code

Malicious code include a range of programmed computer security threats that exploit various network, operating system, software, and physical security vulnerabilities to spread malicious payloads to computer systems. It can be defined as any programmed code that is written specifically to damage, penetrate, or break a system, and includes viruses, worms, Trojans horses, denial-of-service tools, logic bombs, and back doors.

Some malicious code, such as viruses and Trojan horses, and rely on unsuspecting users to spread from system to system, while others, such as worms, spread rapidly among vulnerable systems without requiring any user interaction.

7.1.1: Viruses

Viruses are one of the earliest forms of malicious code to attack information systems and are highly prevalent with major outbreaks occurring regularly. Many viruses carry malicious payloads, some of which can cause the complete destruction of data stored on the local hard drive. It is estimated that there were approximately 65,000 strains of viruses on the Internet in early 2004.

Viruses have two main functions: propagation and destruction. The propagation function defines how the viruses spread from system to system, while the virus's payload implements the malicious and often destructive activity designed by the author of virus.

There are three broad categories of viruses, based on their propagation methods. These are boot sector viruses, file viruses and macro viruses.

- **Master Boot Record (MBR) viruses** are the oldest form of viruses. It attacks the MBR of floppy disks or the hard disk drive, which the computer uses to load the operating system during the boot process. This type of virus is spread between systems through the use of infected floppy disks and was highly effective when floppy disks were the main source of file sharing between systems.
- **File viruses** are executable files having *.exe*, *.com* or *.bat* file extensions and rely on unsuspecting users to execute the file. Social engineering is usually employed to persuade the user to execute the file. Alternatively, the virus may replace an operating system file and would be triggered when the operating system attempts to execute that file.
- **Macro viruses** exploit the scripting functionality implemented by common software applications, such as the applications belonging to the Microsoft Office suite, that may be installed on the computer. This type of virus represents the most modern form of virus and first appeared in the mid-1990s.

> **Note:** Most anti-virus programs use file signature and virus definitions to detect possible viruses by examining boot sectors, files, and sections of program code. However, these programs need to be updated regularly so that newer viruses with different file signatures can be detected.

7.1.2: Worms

Worms contain the same destructive potential as viruses, but that they do not require user interaction in order to replicate and spread. The Internet Worm was the first major security incident to occur on the Internet. Since then hundreds of new worms have been released on the Internet. The destructive power of these worms poses a high risk to the modern Internet. This presents a strong case for system administrators to ensure that the latest security patches are applied to their Internet-connected systems.

7.1.3: Logic Bombs

Logic bombs are malicious code that infect a system but remains dormant until they are triggered by the occurrence of one or more logical conditions. At that time, they spring into action, delivering their malicious payload to unsuspecting computer users. Simple logic bombs may be triggered based upon the system date or time while others may use more complex criteria such as the deletion of a file or user account, or the changing of permissions and access controls. Many viruses and Trojan horses, such as the famous Michelangelo virus, contain a logic bomb component.

7.1.4: Trojan Horses

A Trojan horse is a useful looking application in which a piece of malicious code is concealed. Some Trojan horses are fairly innocuous while others could destroy all the data on a system and cause a large amount of damage in a short period of time. Back Orifice is a well-known Trojan horse that the Windows operating system. To deploy Back Orifice on a system, a malicious attacker place Back Orifice within the installation package of useful application or utility. When an unsuspecting user installs the useful application or utility, they also install Back Orifice, which then runs in the background and allows the malicious attacker to gain administrative access remotely to the target computer.

7.1.5: Active Content

Active content on the web sites that users may visit represents another vector of attack. The delivery of active content is dependant on web applications that are downloaded to users' computers for execution. These applications are based on technologies like Java applets and ActiveX controls. This reduces the load on the web server and improves response time. However, unsuspecting users may download malicious active content, known as hostile applets, from an untrusted source and allow it to execute on their systems, resulting in a significant vulnerability. These hostile applets can cause a variety of damage, including a denial of service attack that consumes system resources or the theft and/or destruction of data. Fortunately, most web browsers require the user to choose to have the active content automatically downloaded, installed, and executed from trusted sites. However, a policy should be put in place to ensure the proper user control of active content.

7.1.6: Spyware

Spyware applications are usually similar in deployment to Trojan horses. They are installed when an unsuspecting user downloads and installs a free application from the Internet. However, more advanced spyware applications could be installed onto a user's computer when the user uses a browser that is vulnerable to visit an untrusted website. Defenses against spyware include not downloading or installing adware-supported applications, and applying the latest security patches for the operating system as well as the browser, or switching to a more secure browser.

7.1.7: SQL Injection

Developers of application that require or allow user input should be aware of malicious users who will seek out and exploit possible vulnerabilities in the protocol or application. An example of this is malformed input or SQL injection which is directed at database applications. An attacker can attempt to insert database or SQL commands to disrupt the normal operation of the database. This could cause the database to become unstable and leak information. In this attack, the attacker searches for web applications in which to insert SQL commands. They use logic such as `1 = 1--` or a single quote, such as `'` to test the database for vulnerabilities. Feedback from the database application indicates that the database is vulnerable to attack. These types of attacks can be prevented by implementing pre-validation; post-validation; and client-side validation.

Section 7.2: System Development Life Cycle (SDLC)

The System Development Life Cycle (SDLC) is a framework for system development. Its purpose is to control the development process and add security at each stage of the development process. The principal elements of the SDLC are listed in **Generally Accepted Principles and Practices for Securing Information Technology Systems** (SP 800-14, National Institute of Standards and Technology, September 1996) and **Security Considerations in the Information System Development Life Cycle** (SP 800-64, National Institute of Standards and Technology, September, October 2003). The five stages of the SDLC are listed in NIST SP 800-14. These are:

- **Initiation**, which involves identifying the need for the system and documenting its purpose and includes an evaluation of the sensitivity of the system and the information to be processed. This is called a **sensitivity assessment**.
- **Development/Acquisition**, which involves the design, development, programming and acquisition of the system. In this stage programmers develop the application code while focusing on security to ensure that input and output controls, audit mechanisms, and file-protection schemes are used.
- **Implementation**, which involves the testing, security testing, accreditation, and installation of the system. This occurs once the application coding has been completed. The testing should not be performed by the programmers but by auditors or quality assurance engineers. If the code is written and verified by the same individuals, errors can be missed and security functions can be bypassed. Therefore, separation of duties is important here.
- **Operation/Maintenance**, which involves the use of the system to perform its designed functions. This stage includes security operations, modification/addition of hardware and/or software, administration, operational assurance, monitoring, and audits.
- **Disposal**, which involves the disposition of the system or system components and products, such as hardware, software, and information; disk sanitization; archiving files; and moving equipment. This stage is usually reached when the system is no longer required.

Section 7.3: Application Development

The development of quality software applications is impossible without a development process model. A process model guides the order of project activities and represents the life cycle of a project. It ensures that the application meets the customer's requirements and that its development remains within the budget and time schedule. A number of process models have emerged over the last few decades. Historically, some process models are static and others do not allow checkpoints. Two such process models are the **waterfall model** and the **spiral model**. These two models provide different approaches to the project life cycle. Other models include:

7.3.1: The Waterfall Model

The waterfall model is one of the most well-known software-development process models. This model works on the assumption that a set of tasks can be encompassed in a single stage and that each stage flows logically to the next stage. Thus, one stage must be completed before developers can move on to the next stage. Project milestones are used as transition and assessment points of each stage. Each subsequent stage might require modifications to only the preceding stage in the model. Therefore, developers are limited to going back only one stage. If modification is required, the stage is not officially pronounced as ending. The modifications must be accomplished and the project milestone met before the stage is officially recognized as completed.

The possible stages of the waterfall model include:

- System feasibility
- Software plans and requirements
- Product design
- Detailed design
- Code
- Integration
- Implementation
- Operations and maintenance

This model is ideal for projects in which the project requirements can be clearly defined and are not likely to be changed in the future. An advantage of this method is that it provides a sense of order and is easily documented. However, it is not suited to and complex projects.

7.3.2: The Spiral Model

The spiral model is based on the continual need to refine the requirements and estimates for a project. This model is effective when used for rapid application development of very small projects. Each stage in the spiral model begins with a design goal and ends with a client review. Thus, this model can generate great synergy between the development team and the client because the client is involved in all stages by providing feedback and approval. In addition, each stage requires its own a risk assessment at which point the estimated costs to complete the project, as well as the schedules are revised. Decision can then be made as to whether the project should be continued or canceled. However, the spiral model does not incorporate clear checkpoints. Consequently, the development process might become chaotic.

7.3.3: Cost Estimation Models

Cost estimation models are not really software development models, but are used to estimate the cost of software development projects. An early example is the **Basic COCOMO Model**, which estimates software development effort and cost as a function of the size of the software product in source instructions. This model uses two equations:

- "The number of man-months (MM) required to develop the most common type of software product, in terms of the number of thousands of delivered source instructions (KDSI) in the software product"
  $MM = 2.4\ (KDSI)^{1.05}$

- "The development schedule (TDEV) in months"
  $TDEV = 2.5(MM)^{0.38}$

A more advanced model, the **Intermediate COCOMO Model**, takes into account hardware constraints, personnel quality, use of modern tools, and other attributes and their aggregate impact on overall project costs. An even more advance model is the **Detailed COCOMO Model**, which takes into account the effects of the additional factors used in the intermediate model on the costs of individual project phases.

The **Function Point Measurement Model** is another cost estimation model but it does not require the user to estimate the number of delivered source instructions. Instead, it focuses on user functions, including external input types; external output types; logical internal file types; external interface file types; and external inquiry types. These functions are calculated and weighted according to complexity and used to determine the software development effort.

A third type of model applies the Rayleigh curve to software development cost and effort estimation. A prominent model using this approach is the **Software Life Cycle Model (SLIM)** estimating method. SLIM makes its estimates by focusing on the number of lines of source code that are modified by the **manpower buildup index (MBI)**, which estimates the rate of buildup of staff on the project; and a **productivity factor (PF)**, which is based on the technology used.

Section 7.4: Information Security and the Life Cycle Model

Information security should be an integral component in the software development process. Not only will this result in secure applications but will reduce development costs and code reworking.

7.4.1: Testing

Testing of the software modules and unit testing should be performed as the modules are being designed. However, testing should not be performed by the programmers because errors can be missed and security functions can be bypassed if the code is written and verified by the same individuals. Therefore, separation of duties is important. The test data should be part of the specifications. Testing should not only check the modules using normal and valid input data, but it should also check for incorrect types, out-of-range values, and other bounds and/or conditions.

7.4.2: The Software Maintenance and Change Control

Change management is a formalized process designed to control changes made to systems and programs, and to analyze the request, examine its feasibility and impact, and develop a timeline to implement approved changes. The change-management process offers all stakeholders opportunity to make an input before changes are made. These are the six steps in change management:

- Define change-management processes and practices
- Receive change requests
- Plan and document the implementation of changes
- Implement and monitor the changes
- Evaluate and report on implemented changes
- Modify the change-management plan, if necessary

One way of approaching the maintenance stage is to divide it into the following three substages: request control; change control; and release control.

- The **request control** manages the users' requests for changes to the software product and gathers information that can be used for managing this process. Te steps are included in this process are:
  - Establishing the priorities of requests
  - Estimating the cost of the changes requested
  - Determining the interface that is presented to the user
- The **change control** process is the principal step in the maintenance stage and addresses the following issues:
  - Recreating and analyzing the problem
  - Developing the changes and corresponding tests
  - Performing quality control
  - The tool types to be used in implementing the changes
  - The documentation of the changes
  - The restriction of the changes' effects on other parts of the code
  - Recertification and accreditation, if necessary
- **Release control** is the process of issuing the latest release of the software. This involves deciding which requests will be included in the new release, archiving of the release, configuration management, quality control, distribution, and acceptance testing.

Section 7.5: Object-Oriented Programming

Object-oriented programming (OOP) is a modular form of programming that allows pieces of software to be reusable and interchangeable between programs. The method of reusing tested and reliable objects is more an efficient method of programming and results in lower programming costs. Because it makes use of modules, a programmer can easily modify an existing program. Code from one class can be passed down to another through the process of inheritance. Thus, new modules that inherit features from existing objects can be inserted into the program. These objects can be controlled through an object program library that controls and manages the deposit and issuance of tested objects to users. To provide protection from disclosure and violations of the integrity of objects, security controls must be implemented for the program library. In addition, objects can be made available to users through **Object Request Brokers (ORBs)**, which is intended to support the interaction of objects in heterogeneous, distributed environments.

Objects

In object-oriented programming (OOP), objects can perform certain operations or exhibit specific behaviors upon request. The objects have an identity and can be created during the execution of the program. Furthermore, objects are encapsulated so they can be accessed directly to request performance of their defined operations.

Classes

In object-oriented programming (OOP), objects that share a particular structure and behavior are said to belong to a particular class and each object that belongs to a particular class is an instance of that class.

Thus, ORBs locate and distribute objects across networks and can be considered **middleware** because they reside between two other entities. An ORB is a component of the **Object Request Architecture (ORA)**, which is a high level framework for a distributed environment and was developed by the **Object Management Group (OMG)**. The other components of the ORA are object services, application objects, and common facilities.

The OMG has also developed a **Common Object Request Broker Architecture (CORBA)**, which defines an industry standard that enables programs written in different languages and using different platforms and operating systems to interface and communicate. To implement this compatible interchange, a user develops

a small amount of initial code and an **Interface Definition Language (IDL)** file. The IDL file then identifies the methods, classes, and objects that are the interface targets.

Section 7.6: Database Management

A database system can be used to store and manipulate data in one or more tables without the need to write specific programs to perform these functions. A **Database Management System (DBMS)** provides high-level commands to operate on the data in the database. The DBMS allows the database administrator to control all aspects of the database, including its design, functionality, and security. There are different types of databases, including:

- Hierarchical databases, which links structures in a tree structure. Each record can have only one owner and because of this, a restriction hierarchical database often cannot be used to relate to structures in the real world.
- Mesh databases, which are more flexible than the hierarchical database and use a lattice structure in which each record can have multiple parent and child records.
- Relational databases, which consists of a collection of tables that are linked by their primary keys. Many organizations use software based on the relational database design. Most relational databases use SQL as their query language.
- Object-oriented database, which were designed to overcome some of the limitations of large relational databases. Object-oriented databases do not use a high-level language such as SQL but support modeling and the creation of data as objects.

7.6.1: Transaction Processing

In database management, transaction management is required to ensure that only one user at a time can alter data and that transactions are valid and complete. When more than one database user attempt to modify data in a database at the same time, which is referred to as concurrency, a system of controls must be implemented so that modifications effected by one user does not adversely affect the modifications effected by another user. This can be accomplished by data locking.

Locking solves the problems associated with concurrency and ensures the successful completion of all transactions. Locking also ensures the isolation of the transaction, allowing all transactions to run in complete isolation from one another, even though more than one transaction can be running at any time. This is called serializability and is achieved by running a set of concurrent transactions equivalent to the database state that would be achieved if the set of transactions were executed serially, i.e., one after the other. Although serialization is important to transactions to ensure that the data in the database is correct at all times, many transactions do not always require full isolation. The level at which a transaction is prepared to accept inconsistent data is called the isolation level and is the degree to which one transaction must be isolated from other transactions. A lower isolation level increases concurrency at the expense of data correctness. Conversely, a higher isolation level ensures that data is correct but can negatively affect concurrency.

If the transaction does not generate any errors when it is being executed, all of the modifications in the transaction become a permanent part of the database. All transactions that include data modifications will either reach a new point of consistency and become committed or will be rolled back to the original state of consistency. Transactions are not left in an intermediate state if the database is not consistent. To qualify as a transaction, a logical unit of work must exhibit four properties, called the **ACID** properties (**A**tomicity, **C**onsistency, **I**solation and **D**urability). Theses properties dictate that:

- A transaction must be an **atomic** unit of work, i.e., either all of its data modifications are performed, or none of them is performed.
- When completed, a transaction must leave all data in a **consistent** state. In a relational database, all rules must be applied to the transaction's modifications in order to maintain all data integrity and all internal data structures must be correct at the end of the transaction.
- Modifications made by concurrent transactions must be **isolated** from the modifications made by any other concurrent transactions. A transaction either sees data in the state it was in before another concurrent transaction modified it or it sees the data after the second transaction has completed, but it does not see an intermediate state. This situation is referred to as **serializability**, because it results in the capability to reload the starting data and replay a series of transactions in order to end up with the data in the same state it was in after the original transactions were performed.
- After a transaction has completed, its effects are permanently in place in the system. The modifications must persist even in the event of a system failure, i.e., they must be **durable**.

When a transaction is started, a DBMS must hold many resources to the end of the transaction to protect the ACID properties of the transaction. If data is modified, the modified rows must be protected with exclusive locks that prevent any other transaction from reading the rows, and exclusive locks must be held until the transaction is committed or rolled back.

Table 7.1: Database Terminology

Term	Explanation
Aggregation	The process of combining several low-sensitivity items, with the result being that these items produce a higher-sensitivity data item.
Attribute	A component of a database, such as a column.
Field	The smallest unit of data within a database.
Foreign key	An attribute in one table whose value matches the primary key in another table.
Granularity	Term that refers to the control one has over the view someone has of the database. Highly granular databases have the capability to restrict certain fields or rows from unauthorized individuals.
Relation	Data that is represented by a collection of tables.
Tuple	Represents a relationship among a set of values. In an RDBMS, it is synonymous with record.
Schema	The structure of the entire database. It defines how it is structured.
Primary key	Uniquely identifies each row and assists with indexing the table.
View	Addresses what the end user can see and access.

7.6.2: Data Warehousing

A data warehouse is a repository of data from numerous different databases that is available to users for making queries. These warehouses have been combined, integrated, and structured so that they can be used to provide trend analysis and make business decisions. Data warehousing is used to get a strategic view.

To create a data warehouse, data is taken from an operational database, redundancies are removed, and the data is **normalized**. Then the data is placed into a relational database and can be analyzed by using **On-Line Analytical Processing (OLAP)** and statistical modeling tools. The data in the data warehouse must be maintained to ensure that it is timely and valid.

Normalization

Normalization is the process of eliminating data redundancy by dividing data into related tables. Normalizing a database involves using formal methods to separate the data into multiple, related tables. Hence, a normalized database is said to conform to the rules of relational structure.

7.6.3: Data Mining

Data mining is the process of analyzing data to find and understand patterns and relationships about the data. The result of data mining is metadata, or data about data. The patterns discovered in this data can help companies understand their competitors and understand usage patterns of their customers to carry out targeted marketing. Therefore, information obtained from the metadata should, however, be sent back for incorporation into the data warehouse to be available for future queries and metadata analyses. Data mining can be applied to information system security as an intrusion detection tool to discover abnormal system characteristics in order to determine whether there are aggregation or inference problems and for analyzing audit information.

Data Scrubbing

Data scrubbing is the process of deleting data that is unreliable or no longer relevant from the data warehouse.

7.6.4: Data Dictionaries

A data dictionary is a database for system developers. It records all the data structures used by an application. Advanced data dictionaries incorporate application generators that use the data stored in the dictionary to automate some of the program production tasks. The data dictionary interacts with the DBMS, the program library, applications, and the information security system. In some instances, the data dictionary system is organized into a primary data dictionary and one or more secondary data dictionaries. The primary data dictionary provides a baseline of data definitions and central control, while the secondary data dictionaries are assigned to separate development projects to provide backup to the primary dictionary and to serve as a partition between the development and test databases.

7.6.5: Knowledge Management

Knowledge management seeks to make use of all the knowledge of the organization. It attempts to tie together databases, document management, business processes, and information systems. It is used to interpret the data derived from these systems and automate the knowledge extraction. This knowledge-discovery process takes the form of data mining. These are the three main approaches to knowledge extraction:

- Classification approach, which is used for pattern discovery and in situations when large databases must be reduced to only a few individual records.
- Probabilistic approach, which is used in planning and control systems, and in applications that involve uncertainty.

- Statistical approach, which is used to construct rules and generalize patterns in the data.

Topic 8: Business Continuity Planning and Disaster Recovery Planning

Every organization is exposed to natural disasters, such as hurricanes or earthquakes, or manmade disasters such as a riot or explosion, that threatens the very existence of the organization. Therefore, every organization requires business continuity and disaster recovery planning to mitigate against the effects such disasters. Business Continuity and Disaster Recovery Planning involve the preparation, testing, and updating of the processes, policies and procedures required to protect critical business processes from major disruptions to normal business operations. **Business Continuity Planning (BCP)** involves the assessment of risks to organizational processes and the creation of policies, plans, and procedures to mitigate against those risks, while **Disaster Recovery Planning (DRP)** describes the actions required for the organization resume normal operations after a disaster.

Section 8.1: Business Continuity Planning (BCP)

The BCP process, as defined by $(ISC)^2$, has five main stages:

- Project Scope and Planning
- Business Impact Assessment (BIA)
- Continuity Planning
- Plan Approval and Implementation
- Documentation

8.1.1: Project Scope and Planning

The Scope and Plan Initiation is the first stage in the creation of a business continuity plan (BCP). It entails creating the scope for the plan and the other elements needed to define the parameters of the plan. This step requires a structured analysis of the organization's operations and support services from a crisis planning point of view. Scope activities could include: creating a detailed account of the work required, listing the resources to be used, and defining the management practices to be employed.

8.1.1.1: Business Organization Analysis

An analysis of the business organization is one of the first responsibilities of the individuals responsible for business continuity planning. A business organization analysis is used to identify all departments and individuals who have a stake in the BCP process. This could include:

- **Operational departments** that are responsible for business core services
- **Critical support services**, such as the IT department, plant maintenance department, and other groups responsible for the maintenance of systems that support the operational departments
- **Senior executives** that are responsible for the continued viability of the organization

The business organization analysis is usually performed by the individuals spearheading the BCP effort. However, a thorough review of the analysis should be one of the first tasks performed by the BCP team when it is convened.

8.1.1.2: BCP Team Selection

The BCP team should be inclusive and should not be limited to the IT and/or security departments. Instead the BCP team should include, as a minimum, representatives from each of the organization's departments responsible for the core services performed by the organization; representatives from the key support departments identified by the organizational analysis; IT representatives with technical expertise in areas covered by the BCP; security representatives with knowledge of the BCP process; legal representatives familiar with corporate legal, regulatory, and contractual responsibilities; and representatives from senior management. This will ensure that the plan takes into account knowledge possessed by the individuals responsible for the day-to-day operation of the business and ensures that these individuals are informed about plan specifics before implementation.

8.1.1.3: Resource Requirements

After the team validates the business organization analysis, they should turn to an assessment of the resources required by the BCP effort. This involves the resources required by three distinct BCP phases:

- **BCP development**, which will require some resources as the BCP team performs the four elements of the BCP process. The major resource consumed by this BCP phase will probably be manpower expended by members of the BCP team and the support staff they call upon to assist in the development of the plan.
- **BCP testing, training, and maintenance**, which will require some hardware and software commitments. However, manpower on the part of the employees involved in those activities will be the major resource requirement.
- **BCP implementation**, which occurs when a disaster strikes and the BCP team deems it necessary to conduct a full-scale implementation of the business continuity plan. During this phase significant resources will be required, including a large amount of and the utilization of "hard" resources.

8.1.2: Business Impact Assessment (BIA)

The purpose of a **Business Impact Assessment (BIA)** is to create a document that identifies the resources that are critical to the continued via viability of the organization, the threats posed to those resources, the likelihood that each threat will occur, and the impact of the occurrence those threats on continued via viability of the organization. The impact might be financial or operational. A vulnerability assessment is often part of the BIA process. The BIA has three primary goals:

- **Criticality Prioritization**, which involves the identification and prioritization of critical business unit process and the evaluation of the impact a disruptive to these processes.
- **Downtime Estimation**, which assists in the estimation of the **Maximum Tolerable Downtime (MTD)** that the organization can tolerate and still remain viable. It is often found during the BIA process that this time period is much shorter than expected; that is, the company can tolerate only a much briefer period of interruption than was previously thought.
- **Resource Requirements**, which involves the identification of the resource requirements for the critical processes. Here the most time-sensitive processes should receive the most resource allocation.

8.1.2.1: Priority Identification

The first BIA task is the identification of business priorities that are most essential to the day-to-day operations of the organization. The priority identification task involves creating a comprehensive list of business processes and ranking them in order of importance. One way of performing this task would be to

divide the workload of this process among the team members is to assign each participant responsibility for drawing up a prioritized list that covers the business functions that their department is responsible for. Then a master prioritized list for the entire organization can be created using the departmental prioritized lists.

Priority identification helps identify business priorities from a qualitative point of view. As part of this process, the BCP team should draw up a list of organization **assets** and then assign an **asset value (AV)** in monetary terms to each asset. These numbers will be used in the remaining BIA steps to develop a financially based BIA. The BPC team must also develop is the **maximum tolerable downtime (MTD)** for each business function.

8.1.2.2: Risk Identification

The next phase of the Business Impact Assessment is the identification of both the natural risks and the man-made risks the organization is exposed to. These natural risks include: violent weather conditions, such as hurricanes and tornadoes; earthquakes; mudslides and avalanches; and volcanic eruptions. The man-made risks include: terrorist acts, wars, and civil unrest; theft and vandalism; fires and explosions; power outages; building collapses; transportation failures; and labor unrest.

8.1.2.3: Likelihood Assessment

The next phase of the Business Impact Assessment is to identify the likelihood of each risk occurring. This assessment is usually expressed in terms of an **annualized rate of occurrence (ARO)** that reflects the number of times the organization expects to experience a given disaster each year. The BCP team should determine an ARO for each risk it identified. These numbers should be based upon corporate history, experience, and advice from experts, such as meteorologists, seismologists, fire prevention professionals, and other consultants.

8.1.2.4: Impact Assessment

In the impact assessment, the BPC team must analyze the data gathered during risk identification and likelihood assessment and attempt to determine what impact each of the identified risks would have upon the well being of the organization if it were to occur. There are three specific metrics the BPC team would need to examine: the exposure factor, the single loss expectancy, and the annualized loss expectancy.

- The **exposure factor (EF)**, which is the amount of damage that the risk poses to the asset, expressed as a percentage of the asset's value.
- The **single loss expectancy (SLE)**, which is the monetary loss that is expected each time the risk materializes.
- The **annualized loss expectancy (ALE)**, which is the monetary loss that the business expects to occur as a result of the risk harming the asset over the course of a year.

In addition, the BPC team should consider the non-monetary impact that interruptions might have on the organization. This could include: loss of goodwill among the organization's client base; loss of employees after prolonged downtime; the social and ethical responsibilities to the community; and negative publicity.

8.1.3: Continuity Planning

Continuity Planning focuses on the development and implementation of a continuity strategy to minimize the impact a risks might have should it occur. The first part of the Continuity Planning is to develop a

strategy that would bridges the gap between the Business Impact Assessment and the Continuity Planning stage.

8.1.3.1: Strategy Development

During the strategy development stage, the BCP team must determine which risks will be addressed by the business continuity plan based on the prioritized list created in the previous phases. The BCP team must address all the contingencies for the implementation of provisions and processes that are required for zero-downtime in the event of each and every possible risk. This would require revisiting the **maximum tolerable downtime (MTD)** estimates created during the BIA and determine which risks are deemed acceptable and which must be mitigated by BCP continuity provisions.

Once the BCP team has determined which risks require mitigation and the level of resources that will be committed to each mitigation task, the next stage of Continuity Planning, namely the provisions and processes stage may begin.

8.1.3.2: Provisions and Processes

The provisions and processes stage is the core of Continuity Planning. In this stage, the BCP team designs the specific procedures and mechanisms that will mitigate the risks deemed unacceptable during the Strategy Development stage. There are three categories of assets that must be protected through BCP provisions and processes: personnel; buildings and facilities; and infrastructure.

- The BCP must ensure that the **personnel** within the organization are safe before, during, and after an emergency. Once this has been achieved, provisions can be made to allow employees to conduct both their BCP and operational tasks in as normal a manner as possible given the circumstances. To ensure the success of BCP tasks, employees should be provided with all of the resources they would need to complete their assigned tasks, including shelter and food where required.
- Organizations that require specialized **buildings and facilities** for their critical operations, such as operations centers, warehouses, distribution/logistics centers, and repair/maintenance depots; as well as standard office facilities, manufacturing plants, etc. would require that these facilities be available for the organization's continued viability. Therefore, the BCP should outline mechanisms and procedures that can be put into place to harden existing facilities against the risks defined in the strategy development stage. In the event that it is not possible to harden a facility against a risk, the BCP should identify alternate sites where business activities can resume immediately or within a period of time that is shorter than the maximum tolerable downtime.
- Every organization depends upon an **infrastructure** for its critical processes. For most organizations, a critical part of this infrastructure is an IT network that comprises a number of servers, workstations, and critical communications links between sites. The BCP must address how these systems will be protected against risks identified during the strategy development stage. The BCP should outline mechanisms and procedures to protect systems against the risks by introducing protective measures such as computer-safe fire suppression systems and uninterruptible power supplies. In addition, either redundant components or completely redundant systems/communications links can be implemented to protect business functions.

8.1.4: Plan Approval and Implementation

Once the BCP team completes the design stage, the BCP document should be forwarded to the organization's senior management, including the chief executive officer (CEO), chairman, and board of Directors, for approval. If senior management was involvement throughout the development stages of the

plan, this should be a relatively straightforward process. However, if senior management was not involved, the BCP team should be able to provide a detailed explanation of the plan's purpose and specific provisions.

Once the BCP has been approval by senior management, the BCP team can start to implement the plan. The BCP team develop an implementation schedule that utilizes the resources dedicated to the program to achieve the stated process and provision goals in as prompt a manner as possible given the scope of the modifications and the organizational climate. After all of the resources are fully deployed, the BCP team should supervise the conduct of an appropriate BCP maintenance program to ensure that the plan remains responsive to evolving business needs.

Training and education are essential to the implementation of the BCP. All personnel who will be involved in the plan should receive training on the overall plan and their individual responsibilities. Everyone in the organization should receive at least a plan overview briefing. Personnel with direct BCP responsibilities should be trained and evaluated on their specific BCP tasks to ensure that they are able to complete them efficiently when disaster strikes. Furthermore, at least one backup person should be trained for every BCP task to ensure redundancy in the event personnel are injured or cannot reach the workplace during an emergency.

8.1.5: BCP Documentation

Documentation is a critical stage in the BCP process and provides several important benefits. These benefits are:

- Documentation ensures that BCP personnel have a written continuity document to reference in the event of an emergency, even if senior BCP team members are not present to guide the effort.
- Documentation provides an historical record of the BCP process that will be useful to future personnel seeking to both understand the reasoning behind various procedures and implement necessary changes in the plan.
- Documentation facilitates the identification of flaws in the plan. Having the plan on paper also allows draft documents to be distributed to individuals not on the BCP team for a "sanity check."

8.1.5.1: Continuity Planning Goals

The BCP document should describe the goals of continuity planning as set forth by the BCP team and senior management. The most common goal of the BCP is to ensure the continuous operation of the organization in the face of an emergency situation. Other goals may also be inserted in this section of the document to meet organizational needs.

8.1.5.2: Statement of Importance

The statement of importance reflects the importance of the BCP to the organization's continued viability. This document commonly takes the form of a letter to the organization's employees stating the reason that the organization devoted significant resources to the BCP development process and requesting the co-operation of all personnel in the BCP implementation stage.

8.1.5.3: Statement of Priorities

The statement of priorities is based on the identify priorities stage of the BIA. It involves listing the functions considered critical to the continued operation of the organization in order of importance. The

listing these priorities should also reflect the importance of the functions to continued business operations in the event of an emergency.

8.1.5.4: Statement of Organizational Responsibility

The statement of organizational responsibility also comes from a senior management and can be incorporated into the statement of importance. The statement of organizational responsibility restates the organization's commitment to Business Continuity Planning and informs the organization's employees, vendors, and affiliates that they are individually expected to do everything they can to assist with the BCP process.

8.1.5.5: Statement of Urgency and Timing

The statement of urgency and timing expresses the importance of implementing the BCP and outlines the implementation timetable decided upon by the BCP team and agreed to by senior management. This statement will depend upon the urgency assigned to the BCP process by the organization's senior management. If the statement itself is included in the statement of priorities and statement of organizational responsibility, the timetable should be included as a separate document. Otherwise, the timetable and this statement can be put into the same document.

8.1.5.6: Risk Assessment

The risk assessment of the BCP documentation essentially recaps the decision-making process undertaken during the Business Impact Assessment (BIA). It should include a discussion of all of the risks considered during the BIA as well as the quantitative and qualitative analyses performed to assess these risks. For the quantitative analysis, the actual AV, EF, ARO, SLE, and ALE figures should be included. For the qualitative analysis, the thought process behind the risk analysis should be provided to the reader.

8.1.5.7: Risk Acceptance/Mitigation

The risk acceptance/mitigation contains the outcome of the strategy development stage of the BCP process. It should cover each risk identified in the risk analysis portion of the document and outline one of two thought processes:

- For risks that were deemed acceptable, it should outline the reasons the risk was considered acceptable as well as potential future events that might warrant reconsideration of this determination.
- For risks that were deemed unacceptable, it should outline the risk mitigation provisions and processes put into place to reduce the risk to the organization's continued viability.

8.1.5.8: Vital Records Program

The BCP documentation should also outline a vital records program for the organization. This document states where critical business records will be stored and the procedures for making and storing backup copies of those records.

8.1.5.9: Emergency Response Guidelines

The emergency response guidelines outline the organizational and individual responsibilities for immediate response to an emergency situation. This document provides the first employees to detect an emergency with the steps that should be taken to activate provisions of the BCP that do not automatically activate. These

guidelines should include: immediate response procedures; whom to notify; and secondary response procedures to take while waiting for the BCP team to assemble.

Section 8.2: Disaster Recovery Planning (DRP)

Disaster recovery planning (DRP) is the identification of all the potential disasters the organization might encounter, and development of procedures required to address the realization of those disasters. The disaster recovery plan should be designed so that it would run almost automatically. Essential personnel should be well trained in their duties and responsibilities in the event of a disaster. Furthermore, the first employees on the scene should be able to immediately begin the recovery effort in an organized fashion.

8.2.1 Potential Disasters

The types of disasters an organization may face are both natural, which results from anomalous occurrences in nature and includes storms, earthquakes, and folds; and man-made disasters, such as power outages and explosions.

8.2.1.1: Natural Disasters

Natural disasters are violent occurrences that take place due to changes in the earth's surface or atmosphere that are beyond human control. These occurrences range from hurricanes, for which scientists have developed prediction techniques that provide ample warning before an occurrence, to earthquakes, which can cause massive destruction without notice. A disaster recovery plan should provide mechanisms for responding to both predictable and unpredictable of disasters.

- **Earthquakes** are likely to occur along the known fault lines, such as the San Andreas fault, that exist in many areas of the world. If you organization is located in a region along a fault line where earthquakes are likely, your DRP should include procedures that will be implemented in the event that an earthquake occurs.
- **Floods** can occur near rivers and water masses, such as lakes, and is usually the result of the gradual accumulation of rainwater. As such they are most prevalent during the rain season when rivers and water masses overflow their banks. However, flash floods can strike when a sudden and severe storm results in intense or prolonged rainfall that cannot be absorbed fast enough by the soil or the eco system. In addition, floods can occur when dams are breached.
- **Storms** can take on many forms. Sudden and severe storms can bring about intense rainfall and the risk of flash floods. Hurricanes and tornadoes, which have severe wind speeds in excess of 100 miles per hour, can threaten the structural integrity of buildings, as well as damage due to fallen trees. Thunder storms bring the risk of lightning, which can cause severe damage to sensitive electronic components. Furthermore, storms can cause damage to power lines and carry the risk of power outages.
- **Fires** can occur naturally, as the result of lightning or wildfires during the dry season, and can be devastating.

8.2.1.2: Man-Made Disasters

Intentional and unintentional man-made disasters present a number of potential risks to an organization. Some of the more common man-made disasters that should be taken into account when preparing a business continuity plan and disaster recovery plan include: fires, explosions, terrorist attacks, labor unrest, theft and vandalism.

- Man-made **fires** tend to be smaller than wildfires. Many of these smaller-scale man-made fires can result from carelessness, faulty electrical wiring, or improper fire protection practices. These fires could affect buildings, facilities, or server rooms.
- **Explosions** can result from a number of man-made occurrences, and can be unintentional, such as gas leaks, or intentional, such as a bomb blast. From a disaster planning point of view, the effects of bombings and explosions are similar to those caused by a large-scale fire.
- **Terrorist acts** pose a major challenge to disaster recovery planners due to their unpredictable nature. However, disaster recovery planners must ensure that resources are not over allocated to a terrorist threat at the expense of threats that are more likely to occur.
- You need to consider the possibility of a strike or other **labor unrest** in your disaster recovery plan. A strike could result in large segment of employees downing tools at the same time and having a detrimental impact on the organization. The BCP and DRP teams should address possible labor unrest and should consider alternative plans if labor unrest occurs.
- **Theft and vandalism** represent the same kind of threat as a terrorist attack, but on a much smaller scale. However, there is a far greater chance that an organization will be affected by theft or vandalism than by a terrorist attack. A business continuity and disaster recovery plan should include adequate preventative measures to control the frequency of these occurrences as well as contingency plans to mitigate the effects theft and vandalism have on an organization's ongoing operations.

8.2.2: Recovery Strategies

There are a number of tasks involved in designing an effective disaster recovery plan that will guide the quick restoration of normal business processes and the resumption of activity at the primary business location. These tasks include prioritizing business units, crisis management, emergency communications, and the actual recovery process. The later could include features such as cold sites, warm sites or hot sites.

8.2.2.1: Emergency Response

The disaster recovery plan should outline the procedures essential personnel should follow immediately upon the discovery that a disaster is in progress or is imminent. These procedures will be dependant upon the nature of the disaster, the type of personnel responding to the incident, and the time available before facilities need to be evacuated and/or equipment shut down. In all likelihood, these procedures will be performed in the midst of a crisis, thus not all procedure be complete. Therefore, the procedures should consist of a checklists of tasks arranges order of priority with the most essential tasks first on the checklist.

8.2.2.2: Personnel Notification

The disaster recovery plan should also contain a list of personnel that should be contacted in the event of a disaster. Normally, this will include key members of the DRP team as well as those personnel who execute critical disaster recovery tasks throughout the organization. This response checklist should include alternate means of contact as well as backup contacts for each role in the event the primary contact can not be reached or can not reach the recovery site for one reason or another. The notification checklist should be provided to all personnel who might respond to a disaster as this will enable prompt notification of key personnel.

8.2.2.3: Business Unit Priorities

In order to recover business operations with the greatest possible efficiency, the disaster recovery plan should identify the business units with the highest priority. These units should be recovered first. Thus, the

DRP team must first identify those business units and agree on an order of prioritization. This is very similar to the prioritization task the BCP team performed during the Business Impact Assessment (BIA) and could be based on the resulting documentation of the BIA prioritization task.

However, a simple listing of business units in prioritized order would not be sufficient. Instead the priority list should be broken down into a more detailed business processes list for each business unit, also in order of priority. This would be more useful as not every process performed by your highest-priority business unit will be of the highest priority. Thus, it might be possible, and better, to restore the highest-priority unit to 50 percent capacity and then move on to lower-priority units to achieve some minimum operating capacity across the organization before striving for complete recovery.

8.2.2.4: Crisis Management

An organized disaster recovery plan should counter the panic will set in when a disaster strikes an organization. The employees in the organizations who are most likely to first encounter the disaster, such as security guards, technical personnel, etc., should be trained in the disaster recovery procedures and know the proper notification procedures and immediate response mechanisms.

In addition, continuous training on disaster recovery responsibilities should be performed. If the training budget permits, crisis training should be provided for key employees. This will ensure that at least some of the employees know the proper procedures to handle emergency situations and can provide the leadership to other employees in the event of a disaster.

8.2.2.5: Emergency Communications

Communication is important to the disaster recovery process. It is important that the organization be able to communicate internally as well as externally when a disaster strikes. A disaster of any significance would be noticed, and if the organization is unable to keep the outside world informed of its recovery status, the public could assume that the organization is unable to recover. It is also essential that the organization be able to communicate internally during a disaster so that employees know what is expected of them. If a disaster, such as a storm or an earthquake, destroyed communication lines it is important to find other means of communicating both internally and externally.

8.2.3: Alternate Recovery Sites

Alternate recovery sites are important to the disaster recovery plan as they allow organizations to experience minimal downtime or almost no downtime at all in the event of a disaster. When a disaster occurs, an organization may require temporary facilities at which data can be restored to servers, and business functions can resume. Without such a facility, the organization would need to find a new business location, purchase and set up new equipment before they can resume normal business functions. These activities could require a large amount of resources, including labor, time and finance, and could result in the organization no longer being an economically viable entity.

However, an organization with alternate recovery sites can use those sites to restart business operations when the primary site is stuck by disaster. There are numerous alternative recovery sites, but the four most common alternative recovery sites used in disaster recovery planning are: cold sites, warm sites, hot sites, and mobile sites. When deciding on appropriate locations for such sites, it is important that they be in different geographical locations. If an alternate site is not a significant distance from the primary site, it can fall victim to the same disaster.

8.2.3.1: Cold Sites

A cold site is an alternative facility that is large enough to handle the processing load of an organization and has the appropriate electrical and environmental support systems. However, a cold site has no online computing facilities, does not have active broadband communications links, and has no part of the production network. It may have all or part of the necessary equipment and resources needed to resume business activities, but these would require installation and data would need to be restored to servers.

Because it has no functional computing base and not communication links, a cold site is relatively inexpensive to maintain – it does not require maintenance of workstations and servers, and there is no monthly telecommunications bill when the site is not in use. However, a cold site also requires large amounts of work to set up and there can thus be a significant expenditure of time between the time the decision is made to activate the site and the time the site is fully operational.

8.2.3.2: Hot Sites

A hot site is the exact opposite of the cold site. It is a backup facility that is fully or near operational and has all the necessary hardware, software, telecommunication lines, and network connectivity to allow an organization to resume normal business functions almost immediately. This facility has a full complement of servers, workstations, and communications links and can be a branch office or data center that is online and is connected to the production network. Furthermore, a copy of the data from the systems at the primary site is held on servers at the hot site. This can be a replication copy of the data from production servers, which may be replicated to the hot site in real time, so that an exact duplicate of the systems are ready if and when required. Alternatively, the bulk of the data can be stored on the servers, with the latest data available from backup copies. This allows an organization to resume business functions with almost zero downtime.

The cost of maintaining a fully fictional hot site is extremely high and essentially doubles the organization's budget for hardware, software, and services and requires the use of additional manpower to maintain the site.

8.2.3.3: Warm Sites

A warm site is a compromise between a hot site and a cold site. It is not as well equipped as a hot site, but has part of the necessary hardware, software, data circuits, and other resources that are required to rapidly restore normal business operations. This equipment is usually preconfigured and ready to run appropriate applications to support the organization's operations. However, no data is replicated to the servers and a backup copy is not held at the site. Therefore, the bulk of the data must be transported to the site and restored to the standby servers.

Activation of a warm site usually takes at least 12 hours from the time a disaster is declared. However, warm sites avoid the significant telecommunications and personnel costs inherent in maintaining a near-real-time copy of the operational data environment.

8.2.3.4: Mobile Sites

A mobile site usually consists of one or more self-contained trailers that have all of the environmental control systems necessary to maintain a safe computing environment. Larger corporations sometimes maintain these sites on a "fly-away" basis, ready to deploy them to any operating location around the world via air, rail, sea, or surface transportation. Smaller firms might contract with a mobile site vendor in the local area to provide these services on an as-needed basis.

8.2.3.5: Mutual Assistance Agreements

Mutual Assistance Agreements (MAAs) provide an alternate processing option. Under an MAA, two organizations pledge to assist each other in the event of a disaster by sharing computing facilities or other technological resources. They reduce the necessity for either organization to maintain expensive alternate processing sites, such as the hot sites, warm sites, cold sites, and mobile processing sites. However, there are many drawbacks to Mutual Assistance Agreements that prevent their widespread use:

- MAAs are difficult to enforce. The parties are placing trust in each other that the support will materialize in the event of a disaster. However, the unaffected organization might renege on the agreement.
- Cooperating organizations should be located in relatively close proximity to each other to facilitate the transportation of employees between sites. However, this proximity means that both organizations may be vulnerable to the same disaster.
- Confidentiality concerns often prevent organizations from trusting each other with their data.

Despite these concerns, a MAAs may be a good disaster recovery solution for organizations, especially where the cost is an overriding factor.

8.2.4: Database Recovery

For organizations that rely upon databases as part of the business process, the DRP team should include database recovery planning in the disaster recovery plans. There are various database techniques that can be implemented as part of the disaster recovery plan. These include: electronic vaulting, remote journaling, and remote mirroring. Each techniques has specific advantages and disadvantages, therefore the DRP team should analyze the organization's computing requirements and available resources in order to select the option best suited to the organization.

- **Electronic vaulting** is the process of backing up the data in a database and transferring it to a remote site by means of bulk transfers. The remote site may be a dedicated alternative recovery site, such as a warm site, or an offsite location used for the purpose of maintaining backup data. Because the data from the database are stored off-site in a backup device, there may be a significant time delay between the time you declare a disaster and the time your database is ready for operation with most current data.
- **Remote journaling** involves the backing up the data in a database and transferring it to a remote site on a more frequent basis, usually once every hour. Remote journaling also requires the transfer of copies of the database transaction logs that contain the database transactions that occurred since the previous bulk transfer. Remote journaling is similar to electronic vaulting in that the transaction logs transferred to the remote site are not applied to a live database server but are maintained in a backup device. When a disaster is declared, technicians retrieve the appropriate transaction logs and apply them to the production database.
- **Remote mirroring** is the most advanced and most expensive database backup solution. With remote mirroring, a live database server is maintained at the remote site. The remote server receives copies of the database modifications at the same time they are applied to the production server at the primary site. Therefore, the mirrored server is ready to take over an operational role at a moment's notice. Remote mirroring is a popular database backup strategy for organizations seeking to implement a hot site. However, it requires high infrastructure and personnel costs required to support the mirrored server.

8.2.5: Training and Documentation

As with the business continuity plan, the disaster recovery plan should also be fully documented and adequate training should be provided for all personnel who will be involved in the disaster recovery effort. When designing a training plan, the DRP team should consider orientation training for all new employees; initial training for employees taking on a new disaster recovery role; detailed refresher training for disaster recovery team members; and brief refresher training for all other employees.

8.2.6: Testing and Maintenance

The disaster recovery plan should also be tested on a periodic basis to ensure that the plan's provisions are viable and that it meets the changing needs of the organization. The types of tests that can be conducted will depend upon the types of recovery facilities available to the organization. The five main tests that should be conducted are: checklist tests, structured walk-throughs, simulation tests, parallel tests, and full-interruption tests.

- The **checklist test** is simply the distribution of copies of the disaster recovery checklists to the members of the DRP team, as well as key personnel for review. This ensures that key personnel are aware of their responsibilities and have that knowledge refreshed on a periodic basis. It also provides individuals with an opportunity to review the checklists for obsolete information and update any items that require modification due to changes within the organization and it allows for the identification of situations in which key personnel have left the organization and have not been replaced. In the case of the latter, the disaster recovery responsibilities assigned to those employees should be reassigned.
- The **structured walk-through** involves the role-play of a disaster scenario. Normally, this is performed by the DRP team with the scenario known only to the test moderator, who presents the details to the DRP team at the actual test. The DRP team members then refer to their copies of the disaster recovery plan and discuss the appropriate responses to that particular type of disaster.
- The s**imulation test** is similar to the structured walk-through. In simulation tests, the DRP team members are presented with a scenario and asked to develop an appropriate response. These response measures are then tested. This may involve the interruption of non-critical business activities and the use of some operational personnel.
- The **parallel test** involves the actual relocation of key personnel to the alternate recovery site and the implementation of site activation procedures. During this test, the operations at the main facility are not interrupted.
- The **full-interruption test** is similar to parallel tests, but involves the shutting down operations at the primary site and shifting them to the recovery site.

Topic 9: Law, Investigation and Ethics

The laws that apply to information systems security are complex and differ from region to region. A security professional is expected to know and understand the laws that apply to computer crimes. In addition, the security professional must be able to determine whether a crime has occurred, must be able to preserve evidence of the crime, and must understand the liabilities under the law. However, the investigation of computer crime is complicated by the ethical responsibilities the security professional has to the organizations, and to the profession as a whole.

Section 9.1: Computer Crimes

Computer crimes consist of either of crimes in which computers are used to plan or commit the crime; or of crimes in which a computer or a network is the victim of the crime. The most prominent types of computer crimes include: Denial of Service (DoS) and Distributed Denial of Service (DDoS); theft of passwords; network intrusions; emanation eavesdropping; social engineering; illegal content, such as child pornography; fraud; software piracy; dumpster diving; malicious code; spoofing of IP addresses; information warfare, which is an attacking on the information infrastructure of a nation and could include attacks on military or government networks, communication systems, power grids, and the financial; espionage; destruction or the alteration of information; masquerading; embezzlement; and the use of computers in the planning of terrorism.

Section 9.2: Common Law

There are three main categories of laws in the legal system, which is referred to as the Common Law system. These categories of laws under the Common Law system are criminal law, civil law, and administrative law, each of which is used to cover different circumstances, and impose different penalties upon the perpetrator.

- **Criminal law** serves to preserve social peace and safety and consists of the laws that the police and other law enforcement agencies are concern with. It includes laws against acts such as murder, assault, robbery, and arson. There are also a number of criminal laws that serve to protect society against computer crime, including the **Computer Fraud and Abuse Act**, the **Electronic Communications Privacy Act**, and the **Identity Theft and Assumption Deterrence Act**. These laws are established by the elected representatives that server in the legislative branch of government, as they must comply with the country's constitution. In addition, all laws are subject to judicial review by the courts. If a court finds that a law is unconstitutional, it has the power to render it invalid.
- **Civil law** serves to preserve social order and consists of the laws that govern matters that require impartial arbitration, such as contract disputes and real estate transactions. Civil laws also create the framework for the executive branch of government to carry out its responsibilities. Like criminal laws, civil laws are established by the elected representatives that server in the legislative branch of government are subject to the same constitutional constraints and judicial review procedures.
- **Administrative law** serves to facilitate good governance by outlining procedures to be followed within a federal agency. Administrative law is not established by the legislative branch of government but is published in the **Code of Federal Regulations (CFR)**. Although administrative law does not require an act of the legislative branch to gain the force of law, it must comply with all existing civil and criminal law. Thus government agencies cannot implement regulations that contradict existing laws passed by the legislature. Furthermore, administrative law must also comply with the country's constitution and are subject to judicial review.

Other categories of law under the common law system that relate to information systems are intellectual property and privacy laws.

9.2.1: Intellectual Property Law

Intellectual property law includes a number of categories designed to protect the intellectual property of the author. These categories include: patents, copyrights, trade secrets, trademarks, and warranties.

- The **patent** law protects inventions and processes, ornamental designs, and new varieties of plants. It provides the owner of the patent with a legally enforceable right to prevent others from practicing or reproducing the object covered by the patent for a specified period of time. Where a patent obtained by an individual builds on other patents, the individual must obtain permission from the owner(s) of the earlier patent(s) to exploit the new patent. In the United States, patent law that protects inventions and processes are granted for a period of 20 years from the date the application was made, patent law that protects ornamental designs are granted for a period of 14 years, and patent law that protects the creation of new varieties of plants are granted for a period of 17 years. Once the patent on an invention or design has expired, anyone is free to make, use, or sell the invention or design.
- A **copyright** protects original works of authorship and protects the right of the author to control the reproduction, adaptation, public distribution, and performance of these original works. Copyrights can also be applied to software and databases. The copyright law has two provisions that address uses of copyrighted material by educators, researchers, and librarians.
- **Trade Secret** law secures and maintains the confidentiality of proprietary technical or business-related information. However, the owner of the information has to have invested resources to develop the information; the information has to be of value to the business of the owner; and the information cannot be obvious.
- A **trademark** establishes the identity of an owner, vendor or manufacturer. This can be a name, a word, a symbol, a color, a sound, a product shape, a device, or any combination of these and distinguish them from those manufactured or sold by competitors.
- A **warranty** is a contract that commits an organization to its product. There are two types of warranties: implied warranties and express warranties.
  - An **implied warranty** is an unspoken, unwritten promise created by state law that is passed from a manufacturer or a merchant and to the customer. Implied warranties have two categories: the implied warranty of fitness for a particular purpose and the implied warranty of merchantability.
    - The **implied warranty of fitness** for a particular purpose is a commitment made by the merchant when the consumer relies on the advice of the merchant that the product is suited for a specific purpose.
    - The **implied warranty of merchantability** is the merchant ‘s or manufacturer’s promise that the product sold to the consumer is fit to be sold and will perform the functions that it is intended to perform.
  - An **express warranty** is explicitly offered by the manufacturer or merchant to the customer at the time of the sales transaction. This type of warranty contains voluntary commitments to remedy defects and malfunctions that some customers may encounter in using the product.

9.2.2: Information Privacy and Privacy Laws

Privacy is the right of an individual to protection from unauthorized disclosure of the individual’s personally identifiable information (PII). This right to privacy is embodied in the following fundamental principles of privacy:

- **Notice** — regarding collection, use and disclosure of PII
- **Choice** — to opt out or opt in regarding disclosure of PII to third parties
- **Access** — by consumers to their PII to permit review and correction of information
- **Security** — to protect PII from unauthorized disclosure
- **Enforcement** — of applicable privacy policies and obligations

9.2.2.1: Privacy Policy

Organizations develop and publish privacy policies that describe their approach to handling PII. These policies usually include:

- Statement of the organization's commitment to privacy.
- The type of information the organization would collect. This could include names, addresses, credit card numbers, phone numbers, etc.
- Retaining and using e-mail correspondence.
- Information gathered through cookies and Web server logs and how that information is used.
- How information is shared with affiliates and strategic partners.
- Mechanisms to secure information transmissions, such as encryption and digital signatures.
- Mechanisms to protect PII stored by the organization.
- Procedures for review of the organization's compliance with the privacy policy.
- Evaluation of information protection practices.
- Means for the user to access and correct PII held by the organization.
- Rules for disclosing PII to outside parties.
- Providing PII that is legally required.

9.2.2.2: Privacy-Related Legislation and Guidelines

The important legislation and recommended guidelines for privacy include:

- **The Cable Communications Policy Act**, which provides for discretionary use of PII by cable operators internally but imposes restrictions on disclosures to third parties.
- **The Children's Online Privacy Protection Act (COPPA)**, which is aimed at providing protection to children under the age of 13.
- **Customer Proprietary Network Information Rules**, which apply to telephone companies and restrict their use of customer information both internally and to third parties.
- **The Financial Services Modernization Act (Gramm-Leach-Bliley)**, which requires financial institutions to provide customers with clear descriptions of the institution's polices and procedures for protecting the PII of customers.
- **Telephone Consumer Protection Act**, which restricts communications between companies and consumers, such as in telemarketing.
- **The 1973 U.S. Code of Fair Information Practices**, which states that:
  - There must not be personal data record–keeping systems whose very existence is secret.

  - There must be a way for a person to find out what information about them is in a record and how it is used.
  - There must be a way for a person to prevent information about them, which was obtained for one purpose, from being used or made available for another purposes without their consent.
  - Any organization creating, maintaining, using, or disseminating records of identifiable personal data must ensure the reliability of the data for their intended use and must take precautions to prevent misuses of that data.
- **The Health Insurance Portability and Accountability Act (HIPAA)**, which includes Privacy and Security Rules and standards for electronic transactions and code sets.

9.2.2.3: The Platform for Privacy Preferences (P3P)

The Platform for Privacy Preferences is another privacy policy guideline. It was developed by the World Wide Web Consortium (W3C) and is meant to be implemented Web sites. This guideline can be found at www.w3.org/TR. With P3P, an organization can post its privacy policy in machine-readable form (XML) on its Web site. This policy statement should include:

- Who has access to collected information
- The type of information collected
- How the information is used
- The legal entity making the privacy statement

In addition, the P3P guidelines also define:

- A standard vocabulary for describing a Web site's data practices
- A set of data elements that Web sites can refer to in their P3P privacy policies
- A standard schema for data a Web site may wish to collect, known as the "P3P base data schema"
- A standard set of uses, recipients, data categories, and other privacy disclosures
- An XML format for expressing a privacy policy
- A means of associating privacy policies with Web pages or sites and cookies
- A mechanism for transporting P3P policies over HTTP

9.2.2.4: Electronic Monitoring

Another area that is concerned with privacy practices is keystroke monitoring, e-mail monitoring, and the use of surveillance cameras, badges and magnetic entry cards. Key issues in electronic monitoring are that the monitoring is conducted in a lawful manner and that it is applied in a consistent fashion. An organization monitoring employee e-mail should inform their employees that e-mail is being monitored. This can be accomplished by means of a prominent logon banner or some other frequent notification. The organizations should also ensure that e-mail monitoring is applied uniformly to all employees, clearly define acceptable usage of the e-mail system, clearly define back up procedures for e-mail archiving, and should not provide a guarantee of e-mail privacy. These should be contained in the organization's e-mail usage policy.

9.2.3: Computer Security, Privacy, and Crime Laws

The laws, regulations, and directives that pertaining to the protection of computer-related information include:

- The **U.S. Fair Credit Reporting Act of 1970**, which addresses consumer reporting agencies.
- The **U.S. Racketeer Influenced and Corrupt Organization (RICO) Act of 1970**, which addresses both criminal and civil crimes involving racketeers influencing the operation of legitimate businesses; crimes cited in this act include mail fraud, securities fraud, and the use of a computer to perpetrate fraud.
- The **U.S. Code of Fair Information Practices of 1973**, which addresses personal record keeping.
- The **U.S. Privacy Act of 1974**, which applies to federal agencies; provides for the protection of information about private individuals that is held in federal databases, and grants access by the individual to these databases. This Act assigns the U.S. Treasury Department the responsibilities of implementing physical security practices, information management practices, and computer and network controls.
- The **Foreign Intelligence Surveillance Act of 1978 (FISA)**, which addresses electronic surveillance and physical searches. It allows for electronic surveillance and physical searches without the requirement of a search warrant in cases of international terrorism, spying, or sabotage activities that are conducted by a foreign power or its agent and is not intended for use in prosecuting U.S. citizens.
- The **Organization for Economic Cooperation and Development (OECD) Guidelines of 1980**, which provides for data collection limitations, the quality of the data, specifications of the purpose for data collection, limitations on data use, information security safeguards, openness, participation by the individual on whom the data is being collected, and accountability of the data controller.
- The **Medical Computer Crime Act of 1984**, which addresses illegal access or alteration of computerized medical records through phone or data networks.
- The **Federal Computer Crime Law of 1984**, which was the first computer crime law passed in the U.S. and was strengthened in 1986 and amended in 1994. This law addresses classified defense or foreign relations information, records of financial institutions or credit reporting agencies, and government computers. Unauthorized access or access in excess of authorization became a felony for classified information and a misdemeanor for financial information. This law made it a misdemeanor to knowingly access a U.S. Government computer without or beyond authorization if the U.S. government's use of the computer would be affected.
- The **Computer Fraud and Abuse Act of 1986**, which was amended in 1996 and strengthened Federal Computer Crime Law of 1984 by adding three new crimes:
  - When use of a federal interest computer furthers an intended fraud
  - When altering, damaging, or destroying information in a federal interest computer or preventing the use of the computer or information that causes a loss of $1,000 or more or could impair medical treatment
  - Trafficking in computer passwords if it affects interstate or foreign commerce or permits unauthorized access to government computers
- The **Electronic Communications Privacy Act of 1986**, which addresses eavesdropping and the interception of message contents without distinguishing between private or public systems. This law updated the Federal privacy clause in the Omnibus Crime Control and Safe Streets Act of 1968 to include digitized voice, data, or video, whether transmitted over wire, microwave, or fiber optics. Court warrants are required to intercept wire or oral communications, except for phone companies, the FCC, and police officers that are party to a call with the consent of one of the parties.
- The **Computer Security Act of 1987**, which places requirements on federal government agencies to conduct security-related training, to identify sensitive systems, and to develop a security plan for those sensitive systems. A category of sensitive information called **Sensitive But Unclassified (SBU)** has to be considered. This category, formerly called **Sensitive Unclassified Information (SUI)**, pertains to information below the government's classified level that is important enough to protect, such as medical

information, financial information, and research and development knowledge. This act also partitioned the government's responsibility for security between the **National Institute of Standards and Technology (NIST)** and the **National Security Agency (NSA)**. NIST was given responsibility for information security in general, primarily for the commercial and SBU arenas, and NSA retained the responsibility for cryptography for classified government and military applications.

The Computer Security Act established the national **Computer System Security and Privacy Advisory Board (CSSPAB)**, which is a twelve-member advisory group of experts in computer and telecommunications systems security.

- The British **Computer Misuse Act of 1990**, which addresses computer-related criminal offenses.
- The **Federal Sentencing Guidelines of 1991**, which provides punishment guidelines for those found guilty of breaking federal law.
- The **OECD Guidelines to Serve as a Total Security Framework of 1992**, which includes laws, policies, technical and administrative measures, and education.
- The **Communications Assistance for Law Enforcement Act of 1994**, which requires all communications carriers to make wiretaps possible.
- The **Computer Abuse Amendments Act of 1994**, which accomplished the following:
  - Changed the federal interest computer to a computer used in interstate commerce or communications
  - Covers viruses and worms
  - Included intentional damage as well as damage done with "reckless disregard of substantial and unjustifiable risk"
  - Limited imprisonment for the unintentional damage to one year
  - Provides for civil action to obtain compensatory damages or other relief
- The **Paperwork Reduction Acts of 1980**, which was mended in 1995 and provides **Information Resources Management (IRM)** directives for the U.S. Government. This law established the **Office of Information and Regulatory Affairs (OIRA)** in the **Office of Management and Budget (OMB)**. Under the Paperwork Reduction Act, agencies must:
  - Manage information resources to improve integrity, quality, and utility of information to all users
  - Manage information resources to protect privacy and security
  - Designate a senior official, reporting directly to the Secretary of the Treasury, to ensure that the responsibilities assigned by the Act are accomplished
  - Identify and afford security protections in conformance with the Computer Security Act of 1987 commensurate with the magnitude of harm and risk that might result from the misuse, loss, or unauthorized access relative to information collected by an agency or maintained on behalf of an agency
  - Implement and enforce applicable policies, procedures, standards, and guidelines on privacy, confidentiality, security, disclosures, and sharing of information collected or maintained by or for the agency
- The **Council Directive (Law) on Data Protection for the European Union (EU) of 1995**, which declares that each EU nation is to enact protections similar to those of the OECD Guidelines.
- The **Economic and Protection of Proprietary Information Act of 1996**, which addresses industrial and corporate espionage and extends the definition of property to include proprietary economic information in order to cover the theft of this information
- The **Kennedy-Kassebaum Health Insurance and Portability Accountability Act of 1996 (HIPAA)**, which addresses the issues of personal health care information privacy, security, transactions and code sets, unique identifiers, and health plan portability in the United States.

- The **National Information Infrastructure Protection Act of 1996**, which amended the Computer Fraud and Abuse Act of 1986 and is patterned after the OECD Guidelines for the Security of Information Systems. It addresses the protection of the confidentiality, integrity, and availability of data and systems. This path is intended to encourage other countries to adopt a similar framework, thus creating a more uniform approach to addressing computer crime in the existing global information infrastructure.
- The **Information Technology Management Reform Act (ITMRA) of 1996**, which is also known as the Clinger-Cohen Act, and relieves the General Services Administration of responsibility for procurement of automated systems and contract appeals. OMB is charged with providing guidance, policy, and control for information technology procurement. With the Paperwork Reduction Act, as amended, this Act delineates OMB's responsibilities for overseeing agency practices regarding information privacy and security.
- The **Title I, Economic Espionage Act of 1996**, which addresses the numerous acts concerned with economic espionage and the national security aspects of the crime. The theft of trade secrets is also defined in the Act as a federal crime.
- The **Digital Millennium Copyright Act (DMCA) of 1998**, which prohibits trading, manufacturing, or selling in any way that is intended to bypass copyright protection mechanisms. It also addresses ISPs that unknowingly support the posting of copyrighted material by subscribers. If the ISP is notified that the material is copyrighted, the ISP must remove the material. Additionally, if the posting party proves that the removed material was of "lawful use," the ISP must restore the material and notify the copyright owner within 14 business days. Two important rulings regarding the DMCA were made in 2001. The rulings involved DeCSS, which is a program that bypasses the **Content Scrambling System (CSS)** software used to prevent the viewing of DVD movie disks on unlicensed platforms.
- The **Uniform Computers Information Transactions Act of 1999 (UCITA)**, which addresses libraries' access to and use of software packages, as well as the licensing practices of software vendors.
- The **Electronic Signatures in Global and National Commerce Act of 2000 ("ESIGN")**, which addresses the use of electronic records and signatures in interstate and foreign commerce. It ensures the validity and legal effect of contracts entered into electronically. An important provision of the act requires that businesses obtain electronic consent or confirmation from consumers to receive information electronically that a law normally requires to be in writing. The legislation is intent on preserving the consumers' rights under consumer protection laws and went to extraordinary measures to meet this goal. Thus, a business must receive confirmation from the consumer in electronic format that the consumer consents to receiving information electronically that used to be in written form. This provision ensures that the consumer has access to the Internet and is familiar with the basics of electronic communications.
- The **Provide Appropriate Tools Required to Intercept and Obstruct Terrorism (PATRIOT) Act of 2001**, which allows for the subpoena of electronic records, the monitoring of Internet communications and the search and seizure of information on live systems, backups, and archives
- The **Generally Accepted Systems Security Principles (GASSP)**, which are not laws but are accepted principles that have a foundation in the OECD Guidelines, and dictates that:
  - Computer security supports the mission of the organization.
  - Computer security is an integral element of sound management.
  - Computer security should be cost-effective.
  - Systems owners have security responsibilities outside their organizations.
  - Computer security responsibilities and accountability should be made explicit.
  - Computer security requires a comprehensive and integrated approach.
  - Computer security should be periodically reassessed.
  - Computer security is constrained by societal factors.

- The **E-Government Act. Title III, the Federal Information Security Management Act of 2002 (FISMA)**, which addresses information security controls over information resources that support Federal operations and assets.

Section 9.3: Computer Forensics

Computer forensics is the investigation of computer crimes for the purpose of identifying and prosecuting the perpetrator. It entails the collection, examination and preservation of information from and about computer systems that can be used to identify and prosecute the perpetrator. For the gathered evidence to be admissible in a court of law, standard computer forensics techniques must be used to protect the integrity of that evidence.

Because of the information that is stored on the computer is in a digital, intangible format, there are a number of unique constraints involved in the investigation of computer crimes. Investigators and prosecutors have a limited time frame for the investigation and might interfere with the normal conduct of the business of an organization. There might also be difficulty in gathering the evidence as the data associated with the criminal investigation might be located on the same computer as data needed for the normal conduct of business.

9.3.1: Evidence

To be admissible in a court of law, the evidence must be **relevant**, **legally permissible**, **reliable**, **properly identified**, and **properly preserved**. The gathering, control, and preservation of evidence are thus critical. The evidence gathered at a computer crime scene is usually intangible and subject to easy modification without a trace. Therefore, evidence must be handled carefully and must be properly controlled throughout the **evidence life cycle**, which covers the evidence gathering and application process. This life cycle includes discovery and recognition, protection, recording, collection, identification, preservation, transportation, presentation in a court of law, and the return of evidence to owner.

The collection of evidence could include collecting all relevant storage media, making an image of the hard disk before removing power, taking and printing a screen shot, and avoiding degaussing equipment.

Preservation of evidence includes archiving and retaining all information related to the computer crime until the investigation and any legal proceedings are completed; protecting magnetic media from erasure, storing evidence in a proper environment both onsite and offsite, and defining, documenting, and following a strict procedure for securing and accessing evidence both onsite and offsite.

9.3.1.1: Categories of Evidence

Evidence gathered for a court of law can be classified into different categories. These include:

- **Best evidence**, which is the original or primary evidence rather than a copy or duplicate of the evidence
- **Secondary evidence**, which is a copy of evidence or oral description of its contents and is not as reliable as best evidence
- **Direct evidence**, which proves or disproves a specific act through oral testimony based on information gathered firsthand by a witness
- **Conclusive evidence**, which is incontrovertible evidence that overrides all other categories of evidence
- **Opinions**, which can be divided into two types:
  - **Expert opinions**, which can offer an opinion based on personal expertise and facts

  - **Nonexpert opinions**, which can testify only as to facts
- **Circumstantial evidence**, which are inferences of information from other, intermediate, relevant facts
- **Hearsay evidence**, which is evidence that is not based on personal, firsthand knowledge of the witness but was obtained from another source. Hearsay evidence is generally not admissible in court. Computer-generated records and other business records fall under the category of hearsay evidence because these records cannot be proven accurate and reliable. However, there are certain exceptions for records that are:
  - Made during the regular conduct of business and authenticated by witnesses familiar with their use
  - Relied upon in the regular course of business
  - Made by a person with knowledge of the records
  - Made by a person with information transmitted by a person with knowledge
  - Made at or near the time of occurrence of the act being investigated
  - In the custody of the witness on a regular basis

9.3.1.2 Chain of Custody

Because of the importance of evidence, it is essential that its continuity be maintained and documented. A **chain of custody**, also referred to as a **chain of evidence**, must be established to show how evidence went from the crime scene to the courtroom. Policies and procedures dealing with the management of evidence must be adhered to. Evidence management begins at the crime scene. When a crime scene is being processed, each piece of evidence must be sealed inside an evidence bag that has two-sided tape that allows it to be sealed shut. The evidence bag should then be marked or a tagged. The tag should identify the evidence, and should contain a case number, the date and time the evidence was obtained, the name of the person who discovered the evidence, and the name or badge number of the person who secured the evidence. In addition an evidence log should be established. Information on the log should include a description of each piece of evidence, serial numbers, identifying marks or numbers, and other information that is required by policy or local law. The evidence log also details the chain of custody, describing who had possession of the evidence after it was initially tagged, transported, and locked in storage, and who had access to the evidence while it was held in storage.

9.3.2: Investigation

Due to the ongoing business needs of an organization, an investigation of a suspected computer crime is complicated by numerous issues. The act of investigating a suspected computer crime may affect critical operations. It is thus important to have a plan of action for handling reports of suspected computer crimes, and a committee of appropriate personnel should be created beforehand. This committee should establish a prior liaison with law enforcement; deciding when and whether to bring in law enforcement; set up procedures for reporting computer crimes and for handling and processing reports of computer crime; planning for and conducting investigations, and ensure the proper collection of evidence.

It is important not to alert the suspect when a suspected computer crime is reported. A preliminary investigation should be conducted to determine whether a crime has been committed. This preliminary investigation could consist of examining audit records and system logs, interviewing witnesses, and assessing the damage. It is critical to determine whether and when disclosure to legal authorities is required by law or regulation. The timing of disclosure to legal authorities is also important. Law enforcement agencies in the United States are bound by the Fourth Amendment to the U.S. Constitution, which states that a warrant must be obtained prior to a search for evidence. Private citizens are not bound by the Fourth Amendment and can conduct a search for possible evidence without a warrant. However, if a private individual were asked by a law enforcement officer to search for evidence, a warrant would be required

because the private individual would be acting as an **agent** of law enforcement**.** An exception to the search warrant requirement for law enforcement officers is the **Exigent Circumstances Doctrine**. Under this doctrine, if probable cause is present and destruction of the evidence is deemed imminent, the search can be conducted without the delay of having the warrant in-hand.

9.3.2.1: The First Responder

The first responder is the first person to arrive at a crime scene. A first responder is someone who has the knowledge and skill to deal with the incident. The first responder may be an officer, security personnel, or a member of the IT staff or incident response team. The first responder is accountable for identifying the scope of the crime scene, securing it, and preserving unstable evidence. To securing a scene is important to both criminal investigations and internal incidents. Both of them use computer forensics to obtain evidence. The procedures for investigating internal policy violations and criminal law violations are basically the same. In some cases the internal investigations may not require the involvement of law enforcement. Once the crime scene has been identified, the first responder must then set up a perimeter and protect it. Protecting the crime scene is to cordoning off the area where evidence resides. Everything in an area should be considered a possible source of evidence. This includes functioning and nonfunctioning workstations, laptops, servers, handheld PDAs, manuals, and anything else in the area of the crime. Until the scene has been processed, no one should be allowed to enter the area, and people who were in the area at the time of the crime should be documented. The first responder must not touch anything that is within the crime scene. Traditional forensics may also be used to determine the identity of the person behind the crime. In the course of the investigation, police may collect DNA, fingerprints, hair, fibers, or other physical evidence. It is important for the first responder not to touch anything or attempt to do anything on the computer. Preserving volatile evidence is another important duty of the first responder.

9.3.2.2 The Investigator

When the investigator arrives on the scene, it is important that the first responder gives as much information to them as possible. If the first responder touched anything, it is important that the investigator be notified so that it can be added to the report. Any observations should be mentioned, as this may provide insight into resolving the incident. If a member of the incident response team arrives first and collects some evidence, it is important that the person in charge of the team give all evidence and information dealing with the incident to the police. If more than one member of the team was involved in the collection of evidence, documentation needs to be provided to the investigator dealing with what each person saw and did.

The investigator should make it clear that they are in charge, so that important decisions are made or presented to them. A chain of custody should also be established. It must be documented who handled or possessed evidence during the course of the investigation.

If the first responder has conducted an initial search for evidence, the investigator will need to establish what constitutes evidence and where it resides. If extra evidence is discovered, the perimeter securing the crime scene may be changed. The investigator will either have crime scene technicians begin to process the scene once its boundaries are established, or the investigator will perform the duties of the technician. The investigator or a designated person remains at the scene until all evidence has been properly collected and transported.

9.3.2.3 The Crime Scene Technician

Crime scene technicians are individuals who have been trained in computer forensics, and have the knowledge, skills, and tools necessary to process a crime scene. Technicians are responsible for preserving

evidence, and make great effort to do so. The technician may acquire data from a system's memory. He can also make images of hard disks before shutting them down. All physical evidence is sealed in a bag. It is also tagged to identify it as a particular piece of evidence. The information identifying the evidence is added to a log so that a proper inventory of each piece exists. Evidence is packaged to reduce the risk of damage such as that from electrostatic discharge or jostling during transport. Once transported, the evidence is stored under lock and key to prevent tampering; until such time that it can be properly examined and analyzed. The roles involved in an investigation have varying responsibilities, and the people in each role require special knowledge to perform it properly.

Section 9.4: Liability

The senior management of an organization has the obligation to protect the organization from losses due to natural disasters, malicious code, compromise of proprietary information, damage to reputation, violation of the law, employee privacy suits, and stockholder suits. Senior management must follow the **prudent man rule**, which requires them to perform their duties with the same diligence and due care that ordinary, prudent people would under similar circumstances. In exercising due care senior management must institute mechanisms to prevent the organization's IT infrastructure from being used as a source of attack on another organization's IT system. Failure to follow the prudent man rule would render the individual liable under the Federal Sentencing Guidelines of 1997.

Section 9.5: Ethics

In order to instill proper computing behavior, ethics should be incorporated into an organizational policy and further developed into an organizational ethical computing policy. A number of organizations have addressed the issue of ethical computing and have generated guidelines for ethical behavior.

9.5.1: (ISC)^2 Code of Ethics

The (ISC)^2 Code of Ethics dictates that a Certified Information Systems Security Professionals (CISSPs) shall:

- Conduct themselves in accordance with the highest standards of moral, ethical, and legal behavior.
- Not commit or be a party to any unlawful or unethical act that may negatively affect their professional reputation or the reputation of their profession.
- Appropriately report activity related to the profession that they believe to be unlawful and shall cooperate with resulting investigations.
- Support efforts to promote understanding and acceptance of prudent information security measures throughout the public, private, and academic sectors of our global information society.
- Provide competent service to their employers and clients, and shall avoid any conflicts of interest.
- Execute responsibilities in a manner consistent with the highest standards of their profession.
- Not misuse the information with which they come into contact during the course of their duties, and they shall maintain the confidentiality of all information in their possession that is so identified.

9.5.2: The Computer Ethics Institute's Ten Commandments of Computer Ethics

The **Coalition for Computer Ethics**, which is incorporated as the Computer Ethics Institute (CEI), focuses on the interface of advances in information technologies, ethics, and corporate and public policy. The CEI addresses industrial, academic, and public policy organizations and is concerned with the ethical issues

associated with the advancement of information technologies in society. It has established the following ten commandments of computer ethics:

- Thou shalt not use a computer to harm other people.
- Thou shalt not interfere with other people's computer work.
- Thou shalt not snoop around in other people's computer files.
- Thou shalt not use a computer to steal.
- Thou shalt not use a computer to bear false witness.
- Thou shalt not copy or use proprietary software for which you have not paid.
- Thou shalt not use other people's computer resources without authorization or the proper compensation.
- Thou shalt not appropriate other people's intellectual output.
- Thou shalt think about the social consequences of the program you are writing for the system you are designing.
- Thou shalt use a computer in ways that ensure consideration and respect for your fellow humans.

9.5.3: The Internet Activities Board (IAB) Ethics and the Internet

Under the Internet Activities Board (IAB) Ethics and the Internet, which is defined in RFC 1087, any activity is defined as unacceptable and unethical that purposely:

- Seeks to gain unauthorized access to the resources of the Internet
- Destroys the integrity of computer-based information
- Disrupts the intended use of the Internet
- Wastes resources such as people, capacity, and computers through such actions
- Compromises the privacy of users
- Involves negligence in the conduct of Internet-wide experiments

9.5.4: The U.S. Department of Health, Education, and Welfare Code of Fair Information Practices

The United States Department of Health, Education, and Welfare has developed a list of fair information practices that focuses on the privacy of individually identifiable personal information. These dictate that:

- There must not be personal data record-keeping systems whose very existence is secret.
- There must be a way for a person to find out what information about them is in a record and how it is used.
- There must be a way for a person to prevent information about them, which was obtained for one purpose, from being used or made available for other purposes without their consent.
- Any organization creating, maintaining, using, or disseminating records of identifiable personal data must ensure the reliability of the data for their intended use and must take precautions to prevent misuses of that data.

9.5.5: The Organization for Economic Cooperation and Development (OECD)

The Organization for Economic Cooperation and Development (OECD) also issued guidelines for ethical computing. These include:

- **Collection Limitation Principle**, which states that there should be limits to the collection of personal data, and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.
- **Data Quality Principle**, which states that personal data should be relevant to the purposes for which they are to be used, and to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.
- **Purpose Specification Principle**, which states that the purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfillment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.
- **Use Limitation Principle**, which states that personal data should not be disclosed, made available, or otherwise used except with the consent of the data subject, or by the authority of the law.
- **Security Safeguards Principle**, which states that personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification, or disclosure of data.
- **Openness Principle**, which states that there should be a general policy of openness about developments, practices, and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data and the main purposes of their use, as well as the identity and usual residence of the data controller.
- **Individual Participation Principle**, which states that an individual should have the right:
  - To obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him.
  - To have communicated to him data relating to him within a reasonable time at a charge, if any, that is not excessive.
    - In a reasonable manner.
    - In a form that is readily intelligible to him.
  - To be given reasons if a request is denied, and to be able to challenge such denial.
  - To challenge data relating to him and, if the challenge is successful to have the data erased, rectified, completed or amended.
- **Accountability Principle**, which states that a data controller should be accountable for complying with measures that give effect to the principles stated above.
- **Transborder Issues**, which states that a member country should refrain from restricting transborder flows of personal data between itself and another Member country except where the latter does not yet substantially observe these Guidelines or where the re-export of such data would circumvent its domestic privacy legislation. A Member country can also impose restrictions in respect of certain categories of personal data for which its domestic privacy legislation includes specific regulations in view of the nature of those data and for which the other Member country provides no equivalent protection.

Topic 10: Physical Security

Physical security concerned with facility construction and location, facility security, which includes physical access control and technical controls, and the maintenance of security. Its purpose is to protect against physical threats such as fire and smoke; water; earthquakes, landslides, and volcanoes; storms; sabotage and vandalism; explosion and other forms of destruction; building collapse; toxic materials; power outages; equipment failure; and personnel loss. In most cases, a disaster recovery plan or a business continuity plan will be needed in the event that a serious physical threat occurs.

The security controls that can be implemented to manage physical security can be administrative controls, technical controls, and physical controls.

- **Administrative physical security controls** include facility construction and selection, site management, personnel controls, awareness training, and emergency response and procedures.
- **Technical physical security controls** include access controls; intrusion detection; alarms; closed circuit television (CCTV); monitoring; heating, ventilating, and air conditioning (HVAC); power supplies; and fire detection and suppression.
- **Physical controls for physical security** include fencing, lighting, locks, construction materials, mantraps, security guards and guard dogs.

Section 10.1: Administrative Physical Security Controls

Administrative physical security controls are related to the implementation of proper administrative steps. These steps include facility requirements planning for proper emergency procedures, personnel control, and proper facility security management.

10.1.1: Facility Requirements Planning

Without proper control over the physical environment, no amount of administrative, technical, or logical access controls can provide adequate security to an organization. Control over the physical environment begins with planning the security requirements for a facility. A secure facility plan outlines the security needs of the organization and emphasizes methods or mechanisms to employ to provide security. Such a plan is developed through a process known as **critical path analysis**, which is a systematic effort to identify relationships between mission-critical applications, processes, and operations and all of the necessary supporting elements. When critical path analysis is performed properly, a complete picture of the interdependencies and interactions necessary to sustain the organization is produced. Once the analysis is complete, the results serve as a list of items to be physically secure. This needs to be performed in the early stages of the construction of a facility. One of the core physical security elements involved at the construction stage include selecting and designing a secure site that will house the organization's IT infrastructure and operations.

The security needs of the organization should be the overriding concern when selecting a site. The site should have low visibility but should be accessible to both employees and external services. It should not be located near possible hazards or in an area with a high crime rate. Another concern is the effect of natural disasters in the area. The area should not be prone to earthquakes, mud slides, sink holes, fires, floods, hurricanes, tornadoes, falling rocks, snow, rainfall, ice, humidity, heat, extreme cold, etc. finally, it is advisable that the site be located near emergency services, such as police, fire, and hospitals or medical facilities.

10.1.2: Secure Facility Design

The proper level of security for a facility must be planned and designed before the construction of the facility begins. There are a number of important security issues to consider during the design of the facility, including the combustibility of construction materials, load rating, placement, and control of items such as walls, doors, ceilings, flooring, HVAC, power, water, sewage, gas, etc. Entire walls, from the floor to the ceiling, must have an acceptable fire rating. Closets or rooms that store media must have a high fire rating. The same goes for ceilings, but ceilings as well as the floor must have a good weight-bearing rating. In addition, the floor bust be grounded against static buildup and must employ a non-conducting surface material in the data center. Electrical cables must be enclosed in metal conduit, and data cables must be enclosed in raceways. The data center should normally not have windows, but if they do, the windows must be translucent and shatterproof. Doors should resist forcible entry and have a fire rating at least equal to the walls. In addition, emergency exits must be clearly marked and monitored or alarmed. Here personnel safety is the overriding security concern. The facility should also have alternate power sources.

There should not be equal access to all locations within a facility. Areas that contain assets of higher value or importance should have restricted access. Valuable and confidential assets should be located in the heart or center of protection provided by a facility. The work areas and visitor areas should also be planned and designed. Walls or partitions can be used to separate similar but distinct work areas. Such divisions deter casual eavesdropping or shoulder surfing, which is the act of gathering information from a system by observing the monitor or the use of the keyboard by the operator. Floor-to-ceiling walls should be used to separate areas with differing levels of sensitivity and confidentiality.

The server or computer rooms should be designed to support the operation of the IT infrastructure and to prevent unauthorized physical access.

10.1.3: Facility Security Management

Audit trails and emergency procedures fall under facility security management. These are elements of the Administrative Security Controls that are not related to the initial planning of the facility but are required to ensure security on an ongoing basis.

- In information systems, an **audit trail** is a record of events that focuses on a particular type of activity, such as detecting security violations, performance problems, and design and programming flaws in applications. In physical security, audit trails are access control logs and are vital in detecting where access attempts occurred and identifying the perpetrator who attempted them. These are detective rather than preventative controls. To be effective an access logs must record the date, time and location of the access attempt, the success of failure of the access attempt, the identity of the person who attempted the access as well as the identity of the person, if any, who modified the access privileges at the supervisor level, and must be audited regularly. Some audit trail systems can also send alarms or alerts to a security officer when multiple access failure attempts occur.
- The implementation of **emergency procedures** and the proper training of employees in the knowledge of these procedures is another important part of administrative physical controls. These procedures should be clearly documented, readily accessible, and updated periodically. They should include emergency system shutdown procedures, evacuation procedures, employee training, awareness programs, and periodic drills, and periodic equipment and systems tests
- Facilities that employ **restricted areas** to control physical security need to address facility visitors. An escort can be assigned to visitors. Furthermore, a visitor's access and activities should be monitored closely.

10.1.4: Administrative Personnel Controls

Administrative personnel controls are administrative processes that are usually implemented by the Human Resources (HR) department during employee hiring and firing. These often include pre-employment screening, which entails employment, references, or educational history checks and background investigation or credit rating checks for sensitive positions; ongoing employee checks, including security clearance checks for employees that have access to sensitive information, and ongoing employee ratings or reviews by their supervisor; and post-employment procedures, including exit interviews, removal of network access and the changing of passwords, and the return of computer inventory or laptops.

Section 10.2: Physical Access Controls

There are many types of physical access control mechanisms that can be implemented to control, monitor, and manage access to a facility. These physical access control mechanisms range from deterrents to detection mechanisms. If facilities that have various sections, divisions, or areas that are designated as public, private, or restricted should have unique and focused physical access controls, monitoring, and prevention mechanisms for each of the designated areas. These mechanisms can be used to separate, isolate, and control access to the areas of the facility and include fences, gates, turnstiles and mantraps; security guards and guard dogs; badges, keys and locks; motion detectors and alarms; as well as adequate lighting.

- A **fence** can be used to differentiate between different areas and can include a wide range of components, materials, and construction methods. It can consist of stripes painted on the ground, chain link fences, barbed wire or concrete walls. Various types of fences are effective against different types of intruders: fences that are 3 to 4 feet high deter casual trespassers; fences that are 6 to 7 feet high are too hard to climb easily; and fences that are 8 feet high with three strands of barbed wire deter determined intruders.
- A **gate** can be used to control entry and exit points in a fence. The deterrent level of a gate must be equivalent to the deterrent level of the fence to sustain the effectiveness of the fence as a whole. Furthermore, the hinges and locking mechanisms of the gate should be hardened against tampering, destruction, or removal. Gates should be kept to a minimum and could be protected by security guards, guard dogs or CCTV.
- A **turnstile** is a special type of gate that permits only one person from gaining entry to a facility or room at a time and often permits entry but not exit or vice versa.
- A **mantrap** is a double set of doors that is often protected by a security guard. Its purpose is to contain an individual until their identity and authentication is verified. If they are proven to be authorized for entry, the inner door opens, allowing them to enter the facility. If they are not authorized, both doors remain locked until an escort arrives to escort them off the property or arrest them for trespassing.
- **Locks and keys** are a crude form of an identification and authorization mechanism. A user requires the correct key or combination to unlock a door. Such users are considered authorized and are permitted entry. Key-based locks are the most common and inexpensive forms of physical access control devices. Programmable or combination locks offer a broader range of control than preset locks as they can be configured with multiple valid access combinations.
- **Security guards** may be posted around a perimeter or inside to monitor access points or watch detection and surveillance monitors. They are able to adapt and react to different conditions or situations and are able to learn and recognize attack and intrusion activities and patterns. Security guards are often an appropriate security control when immediate, onsite, situation handling and decision making is necessary. However, there are a number of disadvantages to deploying, maintaining, and relying upon

security guards. Not all environments and not all facilities are designed to support security guards. Furthermore, not all security guards are reliable. In situation in which their lives may be at risk, a security guard may be more concerned about self-preservation than the preservation of the security of the facility.

- **Guard dogs** can be an alternative to security guards. They can be deployed as a perimeter security control as they are extremely effective detection and deterrent mechanisms. However, guard dogs require a high level of maintenance, and impose serious insurance and liability requirements.
- A **badge**, or an identification card is a physical identification and/or of electronic access control device. It can be a simple name tag or a smart card or token device that employs multifactor authentication to provide authentication and authorization to access a facility, specific rooms, or secured workstations. Badges often include pictures, magnetic strips with encoded data, and personal details to help a security guard verify identity. Badges may also be used in environments where physical access is controlled by security guards. In such cases, the badge serves as a visual identification mechanism for the security guards. Badges can also serve in environments guarded by scanning devices rather than security guards. In such conditions, the badge can be used either for identification or for authentication.
- Effective **lighting** is a common mechanism of perimeter security control. The main purpose of lighting is to discourage casual intruders, trespassers, and prowlers who would rather perform their misdeeds in the dark. However, lighting is not a strong deterrent and should not be the primary security control mechanism. Furthermore, lighting should not illuminate the positions of security guards, guard dogs, and patrol posts and should be combined with security guards, guard dogs, CCTV, or other forms of intrusion detection.
- A **motion detector** is a device that senses significant movement in a specific area.
- When a motion detector senses significant movement in the environment it triggers an **alarm**, which can be a deterrent, a repellant, a notification mechanism or any combination of these. Alarms that trigger deterrents may shut doors and engage locks, making further intrusion more difficult. Alarms that trigger repellants usually sound an audio siren and turn on lights. These kinds of alarms are used to discourage the intruder from continuing their malicious activities and encourage them to leave the premises. Alarms that trigger notification are often silent, but they record data about the incident and notify the security administrators, security guards, and law enforcement.
- When motion detectors and alarms are used, **secondary verification mechanisms** should be implemented to prevent false triggers. False triggers can occur when animals, birds, bugs, and authorized personnel trigger the alarm. Deploying two or more detection systems and requiring two or more triggers in quick succession to occur before an alarm is triggered may significantly reduce false alarms and increase the certainty of sensing actual intrusions or attacks.
- A **closed-circuit television (CCTV)** system is a security mechanism related to motion detectors, and alarms but is not an automated detection-and-response system. CCTV requires personnel to watch the captured video to detect suspicious and malicious activities and to trigger alarms. In most cases, CCTV is not used as a primary detection mechanism. Instead, it is used as a secondary mechanism that is reviewed after a trigger by an automated system occurs.

Section 10.3: Technical Physical Security Controls

The technical physical security controls used to control physical access include smart cards, dumb cards, proximity readers, and biometrics. Other technical physical security controls include audit trails, access logs, and intrusion detection systems (IDSs).

- A **smart card** usually resembles a credit-card and has an embedded magnetic strip, bar code, or integrated circuit chip. They can contain machine-readable identification information about the authorized user for authentication purposes.
- A **dumb card** is an ID card that usually has a photo and printed information about the authorized user. Dumb cards are for use in environments in which security guards are employed.
- A **proximity reader** can be a passive device, a field-powered device, or a transponder. The proximity device is worn or held by the authorized user. When they pass a proximity reader, the reader is able to identify the user and determine whether they have authorized access. A passive device reflects or otherwise alters the electromagnetic field generated by the reader. This alteration is detected by the reader. A field-powered device has electronics that are activated when it enters the electromagnetic field generated by the reader. A transponder device is self-powered and transmits a signal received by the reader.
- **Intrusion detection systems** used to monitor physical activity are designed to detect the attempted intrusion, breach, or attack by an unauthorized user. These systems may include security guards, automated access controls, and motion detectors, as well as burglar alarms. Physical intrusion detection systems can monitor for vibrations, movement, temperature changes, sound, changes in electromagnetic fields, etc.

Section 10.4: Environment and Personnel Safety

In all circumstances and under all conditions, the most important aspect of physical security is the protection of human life. This is the most important goal of all security solutions. Flooding, fires, release of toxic materials, and natural disasters all threaten human life as well as the stability of a facility. Thus, maintaining the environment of a facility is part of maintaining safety for personnel. However, natural disasters cannot be prevented but they can be mitigated by designing facilities that can withstand those disasters, have high fire ratings, and proper mechanisms such as fire extinguishers, etc. Furthermore, other basic elements of the environment, such as power, noise, and temperature fluctuations can be addressed.

10.4.1: Electrical Power Supply

Electronic equipment, including computer systems, is affected by the quality of the electrical power supplied by the local electric company. The power supplied by electric companies is not consistent and clean. However, most electronic equipment requires clean power to function properly and equipment damage due to power fluctuations in the quality of the power supply is common. Common power supply problems such as spikes, surges, sags, brownouts, and blackouts affect the stability and operation of the electronic equipment. These power supply problems are discussed in Table 10.1.

Table 10.1: Common Power Supply Problems

Problem	Description
Fault	A fault is a brief loss of power.
Surge	This is a massive but brief increase in the voltage source that often originates due to lightning strikes but can originate from the power source.
Inrush	This is the initial surge of power that is usually occurs when connecting equipment to a power source.
Spike	These are similar to surges but are of a very short duration being

	measured in nanoseconds, whereas a surge is measured in milliseconds.
Sag	These are brief decreases of voltage at the power source.
Brownout	If a sag lasts longer than 1 second, it is called a brownout. The overloading of a primary power source can cause brownouts. Some brownouts are "scheduled" by power companies to prevent overloading of circuits.
Blackout	A blackout is a complete power failure. When the power returns after a blackout, there is a power spike and the danger of a power surge.

Organizations can employ devices that can be used to protect electronic equipment from power supply problems. These devices include surge suppressors and uninterruptible power supplies (UPS)

- A **surge suppressor** can be used to filter out the effects of voltage spikes and surges that are present in commercial power sources and smooth out power variations. However, almost nothing will shield electronic equipment from a very close lightning strike. Surge suppressors are available from local computer dealers and superstores. Most power strips within surge protection have a red indicator light. If the light goes out, it means that the unit is no longer providing protection and needs to be replaced. If the indicator light starts flashing, it means the power strip is failing and should be replaced immediately. Surge suppressors smooth out power variations and protect the computer from power fluctuations up to a point; however, for complete protection from power fluctuations and outages, an uninterruptible power supply (UPS) is recommended.
- An **uninterruptible power supply (UPS)** is an inline battery backup that is installed between the electronic equipment and the wall outlet. A UPS protects the electronic equipment from surges and acts as a battery when the power dips or fails. It also provides a warning when the power source is above or below acceptable levels. Many UPS models can interact with computer systems and initiate a safe shutdown in the event of a complete power failure using software that runs in the background and sends a signal through one of the computer's COM ports when power is down. The amount of time that a UPS device can keep a system running is determined by battery capacity and the power demands of the equipment connected to it. A more powerful UPS device will need its own line and circuit breaker. One of the principal power drains is the monitor. To keep a system online as long as possible during a power failure, turn off the monitor immediately after the failure commences. When considering a UPS, you must take into account the amount of protection that is needed. The watt rating of the UPS must be sufficient to supply the electronic equipment and all its peripherals with power for enough time to safely shut down the system. This can be calculated by adding the power rating of all pieces of equipment that are to be connected to the UPS. You should however beware of plugging a laser printer into a UPS unless the UPS is specifically rated to support that type of device. Laser printers often require more power than a UPS is able to provide.

10.4.2: Electromagnetic Interference (EMI) and Radio Frequency Interference (RFI)

Electromagnetic interference (EMI) can cause problems in the functioning of electronic equipment and can interfere with the quality of communications, transmissions, and playback. It can affect data transmission that relies on electromagnetic transport mechanisms, such as telephone, cellular, television, audio, radio, and network mechanisms. There are two types of EMI: common mode EMI, which is generated by the difference in power between the live and ground wires of a power source or operating electronic equipment; and traverse mode EMI, which is generated by the difference in power between the live and neutral wires of a power source or operating electronic equipment.

Radio frequency interference (RFI) is similar to EMI and can affect the same systems as EMI. RFI is generated by a wide number of common electrical appliances, including fluorescent lights, electrical cables, electric space heaters, computers, elevators, motors, and electric magnets.

10.4.3: Heating, Ventilating, and Air Conditioning (HVAC)

Maintaining the environment involves control over the heating, ventilating, and air conditioning (HVAC) mechanisms. This is particularly important in computer and server rooms, which be kept should to a temperature of 60 - 75 degrees Fahrenheit or 15 - 23 degrees Celsius and a humidity level should be maintained between 40 and 60 percent. The humidity level is important in these rooms because too much humidity can cause corrosion while too little humidity can cause static electricity.

Table 10.2: Possible Damage from Static Electricity

Static Voltage	Possible Damage
40	Can cause the destruction of sensitive circuits and other electronic components.
1,000	Can cause scrambling of monitor displays.
1,500	Can cause the destruction of data stored on hard drives as well as the hard drives themselves.
2,000	Can cause an abrupt system shutdown.
4,000	Can cause printers to jam as well as damage components.
17,000	Can cause permanent circuit damage.

10.4.4: Water

Physical security policies should address water leakage even if they are not everyday occurrence. Water and electricity are not compatible, thus, a water leak can cause significant damage to electronic equipment, especially while they are operating. In addition, water and electricity represents a serious risk of electrocution to personnel. Whenever possible, you should locate server rooms and critical computer equipment away from any water source or transport pipes. You may also want to install water detection circuits on the floor around mission-critical systems. Water detection circuits sound an alarm if water is encroaches upon the equipment.

In addition to monitoring for water leaks, the facility's capability of handling severe rain or flooding should also be evaluated.

10.4.5: Fire Detection and Fire Suppression

Fire is a major risk in any environment that contains a lot of electronic equipment. Therefore fire detection and fire suppression systems must be put in place to protect the safety of personnel as well as the electronic equipment. In addition to protecting people, fire detection and suppression is designed to keep damage caused by fire, smoke, heat, and suppression materials to a minimum, especially in regard to the IT infrastructure. One of the basics of fire management is proper personnel awareness training. All personnel should be familiar with the use and location fire suppression mechanisms in the facility, as well as the evacuation routes from the facility.. Other items that can be included in fire or general emergency response training are cardiopulmonary resuscitation (CPR) training, emergency shutdown procedures, and a preestablished rendezvous location or safety verification mechanism.

Addressing fire detection and suppression includes dealing with the possible contamination and damage caused by a fire. The destructive elements of a fire include smoke and heat, but they also include the suppression medium, such as water or soda acid. Smoke is damaging to most storage devices. Heat can damage any electronic or computer component. Suppression mediums can cause short circuits, initiate corrosion, or otherwise render equipment useless. All of these issues must be addressed when designing a fire response system.

10.4.5.1: Fire Detection Systems

The implementation of an automated detection and suppression system is important to properly protect a facility from fire. There are numerous types of fire detection systems. These include:

- **Fixed temperature detection systems**, which trigger suppression when a specific temperature is reached.
- **Rate of rise temperature detection systems**, which trigger suppression when the speed at which the temperature changes reaches a specific level.
- **Flame actuated systems**, which trigger suppression based on the infrared energy of flames.
- **Smoke actuated systems**, which trigger suppression based on photoelectric or radioactive ionization sensors.

Most of these fire detection systems can be linked to fire response service notification mechanisms. When suppression is triggered, such linked systems will contact the local fire response team and request aid using an automated message or alarm.

10.4.5.2: Fire Suppression Systems

There are different types of fire extinguishers that can be used for the suppression of different types of fires. If a fire extinguisher is used incorrectly or the wrong type of fire extinguisher is used, the fire could spread and intensify instead of being suppressed. Furthermore, fire extinguishers are to be used only when a fire is still in the incipient stage. Table 10.3 lists the three common types of fire extinguishers.

Table 10.3: Fire Extinguisher Classes

Class	Type	Suppression Material
A	Common combustibles	Water, soda acid (dry powder)
B	Liquids	CO_2, Halon, soda acid
C	Electrical	CO_2, Halon

Problems may occur if the fire suppression material used in a fire extinguisher damages equipment or creates significant smoke when suppressing a fire, or causes other potential problems that can result in collateral damage. When choosing a fire suppression system, it is important to choose one that will suppress the fire without destroying the equipment in the process. The different types of fire suppression systems that you can use include water discharge systems and gas discharge systems.

- There are four main types of **water discharge systems**. These are:
  - A **wet pipe system**, which is also known as a closed head system, and is always full of water.
  - A **dry pipe system**, which contains compressed air that is released when the system is triggered, and opens a water valve that causes the pipes to fill and discharge water.

  - A **deluge system**, which is another dry pipe system that uses larger pipes and therefore a significantly larger volume of water. Deluge systems are not appropriate for environments that include electronic equipment.
  - A **preaction system**, which is a combination of dry pipe and wet pipe systems. The system exists as a dry pipe until the initial stages of a fire are detected and then the pipes are filled with water. The water is released only after the sprinkler head activation triggers are melted by sufficient heat. If the fire is quenched before the sprinklers are triggered, the pipes can be manually emptied and reset. Preaction systems are the most appropriate water-based system for environments that include both electronic equipment and personnel in the same locations.

- **Gas discharge systems** are usually more effective than water discharge systems. However, gas discharge systems are not appropriate for environments in which personnel are located as they usually remove the oxygen from the air, thus making them hazardous to personnel. Gas discharge systems employ a pressurized gaseous suppression medium, such as CO_2 or Halon.

  Halon is a very effective fire suppression compound, but it converts to toxic gases at 900 degrees Fahrenheit and is damaging to the ozone. Therefore, it is usually replaced by a more ecological and less toxic medium. The replacements for Halon include:
  - **Heptafluoropropane (HFC-227ea)**, which is also known as FM-200.
  - **Trifluromethane (HCFC-23)**, which is also known as FE-13. FE-13 molecules absorb heat, making it impossible for the air in the room to support combustion. It is considered to be one of the safest clean agents.
  - **Inergen (IG541)**, which is a combination of three different gases; nitrogen, argon, and carbon dioxide. When released, it lowers the oxygen content in a room to the point that the fire cannot be sustained.
  - **CEA-410** or **CEA 308**
  - **NAF-S-III (HCFC Blend A)**
  - **Aragon (IG55)** or **Argonite (IG01)**

Section 10.5: Equipment Failure

Failure of electronic equipment is inevitable, regardless of the quality of the equipment an organization chooses to purchase and install. Therefore, an organization has to be prepared for equipment failure so as to ensure the ongoing availability of its IT infrastructure and help to protect the integrity and availability of its resources.

Preparing for equipment failure can take many forms. In some non-mission-critical situations, simply knowing where to purchase replacement parts for a 48-hour replacement timeline is sufficient. In other situations, maintaining onsite replacement parts is mandatory. Here the response time in returning a system back to a fully functioning state is directly proportional to the cost involved in maintaining such a solution. Costs include storage, transportation, prepurchasing, and maintaining onsite installation and restoration expertise. In some cases, maintaining onsite replacements is not feasible. For those cases, establishing a **service level agreement (SLA)**, which clearly defines the response time a vendor will provide in the event of an equipment failure emergency with the hardware vendor, is essential.

Aging hardware should also be scheduled for replacement and/or repair. The schedule for such operations should be based on the **mean time to failure (MTTF)** and **mean time to repair (MTTR)** estimates for each device. MTTF is the expected typical functional lifetime of the device given a specific operating environment while the MTTR is the average length of time required to perform a repair on the device. A device can often undergo numerous repairs before a catastrophic failure is expected. All devices should be

replaced before their MTTF expires. When a device is under repair, an alternate solution or a backup device should be implemented to fill in for the duration of the repair.

INDEX

A

B

C

D

E

F

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

X