SYNGRESS

Security+ Study Guide

Exam SY0-201

11th HOUR

The only guide you need for last-minute studying

Answers the toughest questions and highlights core topics

Can be paired with any other study guide so you are completely prepared

Ido Dubrawsky

Syngress is an imprint of Elsevier
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
Linacre House, Jordan Hill, Oxford OX2 8DP, UK

Eleventh Hour Security+ Exam SY0-201 Study Guide

Library of Congress Cataloging-in-Publication Data
Application submitted

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN: 978-1-59749-427-4

Printed in the United States of America
09 10 11 12 13 10 9 8 7 6 5 4 3 2 1

About the Authors

Author

Ido Dubrawsky (CISSP, Security+, CCNA) is the Chief Security Advisor for Microsoft's Communication Sector Americas division. His responsibilities include providing subject matter expertise on a wide range of technologies with customers as well as discussions on policy, regulatory concerns, and governance. Prior to working at Microsoft, Ido was the acting Security Consulting Practice Lead and a Senior Security Consultant at AT&T's Callisma subsidiary where he was tasked with helping to rebuild the practice. Ido has held a wide range of previous roles, including Network Security Architect for Cisco Systems, Inc. on the SAFE Architecture Team. He has worked in the systems and network administration field for almost 20 years in a variety of environments from government to academia to private enterprise and has a wide range of experience in various networks, from small to large and relatively simple to complex. Ido is the primary author of three major SAFE white papers and has written, and spoken, extensively on security topics. He has been a regular contributor to the SecurityFocus Web site on a variety of topics covering security issues. He holds a BSc and an MSc in Aerospace Engineering from the University of Texas at Austin.

Technical Editor

Michael Cross (MCSE, MCP+I, CNA, Network+) is an Internet specialist/programmer with the Niagara Regional Police Service. In addition to designing and maintaining the Niagara Regional Police's Web site (www.nrps.com) and intranet, he has also provided support and worked in the areas of programming, hardware, database administration, graphic design, and network administration. In 2007, he was awarded a Police Commendation for work he did in developing a system to track high-risk offenders and sexual offenders in the Niagara Region. As part of an information technology team that provides support to a user base of over 1,000 civilian and uniformed users, his theory is that when the users carry guns, you tend to be more motivated in solving their problems.

Michael was the first computer forensic analyst in the Niagara Regional Police Service's history, and for 5 years he performed computer forensic examinations on computers involved in criminal investigations. The computers he examined for evidence were involved in a wide range of crimes, inclusive to homicides, fraud, and possession of child pornography. In addition to this, he successfully tracked numerous individuals electronically, as in cases involving threatening

e-mail. He has consulted and assisted in numerous cases dealing with computer-related/Internet crimes and served as an expert witness on computers for criminal trials.

Michael has previously taught as an instructor for IT training courses on the Internet, Web development, programming, networking, and hardware repair. He is also seasoned in providing and assisting in presentations on Internet safety and other topics related to computers and the Internet. Despite this experience as a speaker, he still finds his wife won't listen to him.

Michael also owns KnightWare, which provides computer-related services like Web page design, and Bookworms, which provides online sales of merchandise. He has been a freelance writer for over a decade and has been published over three dozen times in numerous books and anthologies. When he isn't writing or otherwise attached to a computer, he spends as much time as possible with the joys of his life: his lovely wife, Jennifer; darling daughter Sara; adorable daughter Emily; charming son Jason; and beautiful and talented daughter Alicia

CHAPTER 1

Systems Security

Exam objectives in this chapter:

- Systems Security Threats
- Host Intrusion Detection System
- Personal Software Firewall
- Anti-Virus
- Anti-SPAM
- Pop-Up Blockers
- Hardware and Peripheral Security Risks

SYSTEMS SECURITY THREATS

There are security risks to almost any system. Any computer, network or device that can communicate with other technologies, allows software to be installed, or is accessible to groups of people faces any number of potential threats. The system may be at risk of unauthorized access, disclosure of information, destruction or modification of data, code attacks through malicious software, or any number of other risks discussed in this book.

Some of the most common threats to systems come in the form of malicious software, which is commonly referred to as *malware*. Malware is carefully crafted software written by attackers and designed to compromise security and/or do damage. These programs are written to be independent and do not always require user intervention or for the attacker to be present for their damage to be done. Among the many types of malware we will look at in this chapter are viruses, worms, Trojan horses, spyware, adware, logic bombs, and rootkits.

Privilege escalation

Privilege escalation occurs when a user acquires greater permissions and rights than he or she was intended to receive.

- Privilege escalation can be a legitimate action.
- Users can also gain elevated privileges by exploiting vulnerabilities in software (bugs or backdoors) or system misconfigurations. *Bugs* are errors in software, causing the program to function in a manner that wasn't intended.
- *Backdoors* are methods of accessing a system in a manner that bypasses normal authentication methods.
- System misconfigurations include such items as adding a user to a privileged group (such as the Administrator group in Active Directory) or leaving the root password blank or easily guessable.

Viruses and worms

Malicious software has appeared in many forms over the decades, but the problem has increased substantially as more computers and devices are able to communicate with one another.

- Before networks were commonplace, a person transferring data needed to physically transport software between machines, often using floppy diskettes or other removable media.
- To infect additional machines, the malicious software would have to write itself to the media without the user's knowledge.
- With the widespread use of networking, exploitable vulnerabilities, file sharing, and e-mail attachments made it much easier for malware to disseminate.

There are many different types of malicious code that are written with the intention of causing damage to systems, software, and data—two of the most common forms are viruses and worms.

VIRUSES

A *computer virus* is defined as a self-replicating computer program that interferes with a computer's hardware, software, or OS.

- A virus's primary purpose is to create a copy of itself.
- Viruses contain enough information to replicate and perform other damage, such as deleting or corrupting important files on your system.
- A virus must be executed to function (it must be loaded into the computer's memory) and then the computer must follow the virus's instructions.
- The instructions of the virus constitute its *payload*. The payload may disrupt or change data files, display a message, or cause the OS to malfunction.
- A virus can replicate by writing itself to removable media, hard drives, legitimate computer programs, across the local network, or even throughout the Internet.

WORMS

Worms are another common type of malicious code, and are often confused with viruses.

- A *worm* is a self-replicating program that does not alter files but resides in active memory and duplicates itself by means of computer networks.
- Worms can travel across a network from one computer to another, and in some cases different parts of a worm run on different computers.
- Some worms are not only self-replicating but also contain a malicious payload.

DIFFERENCE BETWEEN VIRUSES AND WORMS

Over time the distinction between viruses and worms has become blurred. The differences include:

- Viruses require a host application to transport itself; worms are self-contained and can replicate from system to system without requiring an external application.
- Viruses are intended to cause damage to a system and its files; worms are intended to consume the resources of a system.

DEFENDING AGAINST VIRUSES AND WORMS

Protection against viruses, worms, and other malicious code usually includes up-to-date anti-virus software, a good user education program, and diligently applying the software patches provided by vendors.

- *Anti-virus software* is an application that is designed to detect viruses, worms, and other malware on a computer system. These programs may monitor the system for suspicious activity that indicates the presence of malware, but more often will detect viruses using signature files. *Signature files* are files that contain information on known viruses, and are used by anti-virus software to identify viruses on a system.
- User education is an important factor in preventing viruses from being executed and infecting a system. As viruses require user interaction to load, it is important that users are aware that they shouldn't open attached files that have executable code (such as files with the extension .com, .exe, and .vbs), and avoid opening attachments from people they don't know.
- Updating systems and applying the latest patches and updates is another important factor in protecting against viruses and worms.
- When researchers discover a flaw or vulnerability, they report it to the software vendor, who typically works on quickly developing a fix to the flaw.

TIP

If you're really pressed for time, focus on the general characteristics of viruses and worms as they still represent some of the most challenging problems for enterprise network and security administrators.

- A *zero-day attack* is an attack where a vulnerability in a software program or operating system is exploited before a patch has been made available by the software vendor.
- You can prepare for an infection by a virus or worm by creating backups of legitimate original software and data files on a regular basis. These backups will help to restore your system, should that ever be necessary.

Trojan

A *Trojan horse* is a program in which malicious code is contained inside what appears to be harmless data or programming, and is most often disguised as something fun, such as a game or other application. The malicious program is hidden, and when called to perform its functionality, can actually ruin your hard disk.

Spyware and adware

Spyware and adware are two other types of programs that can be a nuisance or malicious software. Both of these may be used to gather information about your computer, or other information that you may not want to share with other parties.

SPYWARE

- *Spyware* is a type of program that is used to track user activities and spy on their machines.
- Spyware programs can scan systems, gather personal information (with or without the user's permission), and relay that information to other computers on the Internet.
- Spyware has become such a pervasive problem that dozens of anti-spyware programs have been created.
- Some spyware will hijack browser settings, changing your home page, or redirect your browser to sites you didn't intend to visit. Some are even used for criminal purposes, stealing passwords and credit card numbers and sending it to the spyware's creator.
- Spyware usually does not self-replicate, meaning that the program needs to be installed in each target computer.
- Some spyware programs are well behaved and even legal, with many spyware programs taking the form of browser toolbars.

ADWARE

Adware is software that displays advertising while the product is being used, allowing software developers to finance the distribution of their product as freeware (software you don't have to pay for to use). However, some types of adware can be a nuisance and display pop-up advertisements (such as through an Internet browser), or be used to install and run other programs without your permission.

- Adware can cause performance issues.

DIFFERENCE BETWEEN SPYWARE AND ADWARE

Adware and spyware are two distinctively different types of programs.

- Adware is a legitimate way for developers to make money from their programs.
- Spyware is an insidious security risk.
- Adware displays what someone wants to say; spyware monitors and shares what you do.
- Adware may incorporate some elements that track information, but this should only be with the user's permission. Spyware will send information whether the user likes it or not.

DEFENDING AGAINST SPYWARE AND ADWARE

Preventing spyware and adware from being installed on a computer can be difficult as a person will give or be tricked into giving permission for the program to install on a machine. Users need to be careful in the programs they install on a machine and should do the following:

- Read the End User License Agreement (EULA), as a trustworthy freeware program that uses advertising to make money will specifically say it's adware. If it says it is and you don't want adware, don't install it.
- Avoid installing file-sharing software as these are commonly used to disseminate adware/spyware.
- Install and/or use a pop-up blocker on your machine such as the one available with Google Toolbar, MSN Toolbar, or the pop-up blocking feature available in Internet Explorer running on Windows XP SP2 or higher. The pop-up blocker prevents browser windows from opening and displaying Web pages that display ads or may be used to push spyware to a computer.
- Be careful when using your Web browser and clicking on links. If you see a dialog box asking you to download and install an ActiveX control or another program, make sure that it's something you want to install and that it's from a reliable source. If you're unsure, do not install it.
- Use tools that scan for spyware and adware, and can remove any that's found on a machine.

Rootkits and botnets

Botnets and rootkits are tools used to exploit vulnerabilities in operating systems and other software.

- *Rootkits* are software that can be hidden on systems and can provide elevated privileges to hackers.
- A rootkit is a collection of tools used to gain high levels of access to computers (such as that of an administrator).
- Rootkits try to conceal their presence from the OS and anti-virus programs in a computer.

- Rootkits can make it easy for hackers to install remote control programs or software that can cause significant damage.
- A *bot* is a type of program that runs automatically as robots performing specific tasks without the need for user intervention.
- Bots have been developed and used by Google, Yahoo, and MSN to seek out Web pages and return information about each page for use in their search engines. This is a legitimate use for bots, and do not pose a threat to machines.
- Botnets are one of the biggest and best-hidden threats on the Internet.
- The botnet controller is referred to as the bot herder, and he or she can send commands to the bots and receive data (such as passwords or access to other resources) from them.
- Bots can be used to store files on other people's machines, instruct them to send simultaneous requests to a single site in a DoS attack, or for sending out SPAM mail.
- A Web server or IRC server is typically used as the Command and Control (C&C) server for a group of bots or a botnet.

Logic bombs

A *logic bomb* is a type of malware that can be compared to a time bomb.

- Designed to execute and do damage after a certain condition is met, such as the passing of a certain date or time, or other actions like a command being sent or a specific user account being deleted.
- Attackers will leave a logic bomb behind when they've entered a system to try to destroy any evidence that system administrators might find.

HOST INTRUSION DETECTION SYSTEM

Intrusion detection is an important piece of security in that it acts as a detective control. An *intrusion detection system* (*IDS*) is a specialized device that can read and interpret the contents of log files from sensors placed on the network as well as monitor traffic in the network and compare activity patterns against a database of known attack signatures. Upon detection of a suspected attack, the IDS can issue alarms or alerts and take a variety of automatic action to terminate the attack.

There are two types of IDSs that can be used to secure a network: host-based IDS (HIDS) and network-based IDS (NIDS). The two types are further broken down into signature-based and behavior-based IDSs. A behavior-based IDS is also known as an anomaly-based IDS.

- A *host-based IDS* is one that is installed on a single system or server and monitors the activity on that server through log analysis and server traffic analysis.
- A *network-based IDS* is a system or appliance that monitors all traffic on a network segment and compares that activity against a database of known attack signatures in an attempt to identify malicious activity.

- A *signature-based IDS* monitors access points and network segments for malicious activity, triggering on events by referencing network activity against an attack signature database.
- A *behavior-based IDS* uses rules or predefined concepts about "normal" and "abnormal" system activity (called heuristics) to distinguish malicious activity from normal system behavior and to monitor, report on, or block anomalies as they occur.

EXAM WARNING

To eliminate confusion on the Security+ exam, the simplest definition of IDS is a device that monitors and inspects all inbound and outbound network traffic, and identifies patterns that may indicate suspicious activities or attacks. Do not confuse this with a firewall, which is a device that inspects all inbound and outbound network traffic looking for disallowed types of connections.

Behavior-based vs. signature-based IDS characteristics

In this section, we'll discuss the differences between signature- and behavior-based IDS.

SIGNATURE-BASED IDSs

Here are the pros and cons of signature-based IDSs.

Pros

- Signature-based IDS examines ongoing traffic, activity, transactions, or behavior for matches with known patterns of events specific to known attacks.
- Requires access to a current database of attack signatures and some way to actively compare and match current behavior against a large collection of signatures.
- Technique works extremely well and has a good track record.

Cons

- Signature databases must be constantly updated.
- IDS must be able to compare and match activities against large collections of attack signatures.
- If signature definitions are too specific, a signature-based IDS may miss variations of known attacks.
- Signature-based IDSs can also impose noticeable performance drags on systems when current behavior matches multiple (or numerous) attack signatures, either in whole or in part.

ANOMALY-BASED IDSs

Here are the pros and cons of anomaly-based IDSs.

Pros

- An anomaly-based IDS examines ongoing traffic, activity, transactions, or behavior for anomalies on networks or systems that may indicate attack. The underlying principle is the notion that "attack behavior" differs enough from "normal user behavior" that it can be detected by cataloging and identifying the differences involved.
- By creating baselines of normal behavior, anomaly-based IDSs can observe when current behavior deviates statistically from the norm. This capability theoretically gives an anomaly-based IDS the ability to detect new attacks that are neither known nor for which signatures have been created.

Cons

- Because normal behavior can change easily and readily, anomaly-based IDSs are prone to false positives, where attacks may be reported based on changes to the norm that are "normal," rather than representing real attacks. Their intensely analytical behavior can also impose heavy processing overheads on the systems they are running on.
- Anomaly-based systems take a while to create statistically significant baselines (to separate normal behavior from anomalies); they are relatively open to attack during this period.

DID YOU KNOW?

Signatures are defined as a set of actions or events that constitute an attack pattern. They are used for comparison in real time against actual network events and conditions to determine if an active attack is taking place against the network. The drawback of using attack signatures for detection is that only those attacks for which there is a released signature will be detected. It is vitally important that the signature database be kept up to date.

Finally, advances in IDS design have led to a new type of IDS, called an *intrusion prevention system* (*IPS*), which is capable of responding to attacks when they occur. By automating a response and moving these systems from detection to prevention, they actually have the ability to block incoming traffic from one or more addresses from which an attack originates. This allows the IPS the ability to halt an attack in process and block future attacks from the same address.

IDS DEFENSES

By implementing the following techniques, IDSs can fend off expert and novice hackers alike. Although experts are more difficult to block entirely, these techniques can slow them down considerably:

- Breaking TCP connections by injecting reset packets into attacker connections causing attacks to fall apart
- Deploying automated packet filters to block routers or firewalls from forwarding attack packets to servers or hosts under attack
- Deploying automated disconnects for routers, firewalls, or servers

ANTI-SPAM

SPAM is also known as unsolicited bulk e-mail (UBE) and accounts for nearly 75–80% of all e-mail traffic on the Internet. SPAM is the digital equivalent of unsolicited postal mail sent by marketing companies on a daily basis across the United States. On a given day, a user is likely to receive 10 times more unsolicited ads or other unwanted e-mail messages than legitimate, useful messages. Anti-SPAM systems use a combination of algorithms and heuristics to identify SPAM based on context or even just word content. Many anti-SPAM systems also use lists of known IP addresses in a database that have been reported as sources of SPAM. These databases are known as real-time black hole lists, or RBLs. The anti-SPAM software checks the originating IP address of the e-mail to determine if it is listed in an RBL and, if so, rejects the e-mail. Not all anti-SPAM programs are successful, and inevitably some SPAM does tend to make it through the filters.

POP-UP BLOCKERS

Many modern Web browsers include some form of *pop-up blocker* to prevent sites from indiscriminately opening up new browser windows against the user's desire. In many cases, vendors have bundled this pop-up blocking capability with browser toolbars that have been made available. Many of the most common browser toolbars can block pop-up applications before the Web browser can process them, which helps prevent a large number of spyware-related applications from being installed. These toolbars also provide many other utilities that enhance the Web surfing experience or additional security features that are not normally found in the Web browsers. Some pop-up blockers may end up missing many forms of pop-ups and may block legitimate windows. To test the effectiveness of a particular pop-up blocker, visit the Popup Test Web site at www.popuptest.com. The Popup Test Web site simulates a variety of pop-up window techniques to validate a particular blocker utility.

HARDWARE AND PERIPHERAL SECURITY RISKS

Having physical access to a computer or other device can enable an unauthorized or uneducated user to make changes to settings that can seriously impact its security and functionality. Conversely, a system administrator can configure

hardware settings so that authentication is required, or disable features that could be used for malicious purposes.

- Peripherals are devices that are connected to a computer using cables or wireless technologies.
- Peripherals include scanners, cameras, and other devices, as well as various storage devices like removable drives, USB Flash Drives, memory cards, and other devices and media.

BIOS

BIOS is an acronym for *Basic Input/Output System* and refers to a chip that resides on the motherboard of a computer.

- This chip contains instructions on how to start the computer and load the operating system and contains low-level instructions about how the system is to handle various hardware and peripherals.
- Information used by the BIOS is set and stored through a semiconductor chip known as the CMOS (Complementary Metal Oxide Semiconductor).
- The CMOS uses a battery on the motherboard to retain power so that settings such as the date, time, and other system settings used by the BIOS aren't lost when the computer turns off.
- A user interface allows you to edit CMOS settings so that you can configure the date, time, boot sequence, video settings, hard drive configuration, and security settings.
- After going through the Power-On Self Test (POST), the BIOS will read the boot sector of the boot drive and use the information there to begin loading the operating system.
- A password may be set to prevent unauthorized persons from accessing the setup software and making changes to the computer. Setting this password also prevents malicious users from configuring Power-On and BIOS passwords, which would restrict valid users from starting the computer or making system changes.

USB devices

USB is an acronym for *Universal Serial Bus*, a standard technology that's used to allow devices to connect through a port on a computer. USB devices can be plugged into the computer and recognized by the operating system, without the need to shut down the computer.

- USB devices are also a possible infection vector for viruses, worms, and other malicious software.

EXAM WARNING

Use encryption and/or password-protected files stored on USB devices in case a device with sensitive data is lost or stolen.

- To prevent the computer from being infected by a virus or other malware, the autoplay feature in Windows should be turned off—this is the default setting in Windows 7.
- USB storage devices should be scanned with up-to-date anti-virus software before any files are opened.

FLASH MEMORY CARDS

Flash memory cards and *sticks* are a popular medium for storing and transferring varying amounts of data.

- Memory cards typically range in size from 8 to 512 MB, but new cards are capable of storing upwards of 8 GB of data.
- Commonly used for storing photos in digital cameras and for storing and transferring programs and data between handheld computers (pocket PCs and Palm OS devices).
- Flash memory cards include:
  - Secure Digital (SD) Memory Card
  - CompactFlash (CF) Memory Card
  - Memory Stick (MS) Memory Card
  - Multi Media Memory Card (MMC)
  - xD-Picture Card (xD)
  - SmartMedia (SM) Memory Card

USB FLASH DRIVES

USB Flash Drives are small portable storage devices that use a USB (Universal Serial Bus) interface to connect to a computer. Like flash memory cards, they are removable and rewritable and have become a common method of storing data.

- USB Flash Drives are constructed of a circuit board inside of a plastic or metal casing, with a USB male connector protruding from one end.
- Some USB Flash Drives come with software that can be used to provide additional features such as encryption.
- Compression may also be used, allowing more data to be stored on the device.

Cell phones

Cell phones are handheld devices that allow people to communicate over a network. Originally only used for voice communication, today's mobile phones provide additional services such as e-mail, Internet browsing, PDA (Personal Digital Assistant) functionality, digital camera, SMS (Short Message Service) for text messaging, games, and the ability to watch video or listen to music.

- Cell phones present additional risks due to their smaller form factor and greater portability than laptops.
- Cell phones used by an organization should have as much security as possible setup on the device.

- If the cell phone supports a power-on password or has a key lock, which prevents the phone from being used unless a personal identification number (PIN) is entered, these features should be activated on the phone.
- Data stored on memory cards used by cell phones should be encrypted if the phone software supports it.
- Organizations should also decide whether to limit or prohibit the use of cameras on cell phones as a cell phone camera can be used to take pictures of sensitive data displayed on a screen or other classified information that may be displayed in plain sight.
- Viruses have been written for cell phones and could be easily disseminated to cell phone users.
- The first cell phone virus, Cabir, first appeared in 2004 and spread between cell phones that used the Symbian operating system by transmitting itself using Bluetooth.
- Cell phones can be used as modems and can allow a computer to connect to the Internet without having to go through the corporate firewall. This could allow for the unauthorized transfer of data outside of the corporate network. Another method of transferring data is using Bluetooth technology.
- *Bluetooth* is a wireless protocol and service that allows Bluetooth-enabled devices to communicate and transfer data with one another. It has a discovery mode that allows devices to automatically detect and connect with other devices. Without authentication, a person could connect to a Bluetooth-enabled cell phone or other device and download data.
- *Bluesnarfing* is a term used for someone who leaves their laptop or another device in discovery mode, so that they can connect to any nearby Bluetooth device that's unprotected.

Removable storage devices

Removable storage, also referred to as *removable media*, is any device that can be attached to a system and used for storing data. Removable storage includes devices like USB Flash Drives and memory cards but also includes devices that provide the ability to store data on such media as:

- CD
- DVD
- Blu-Ray
- Floppy Disks
- Magnetic Tape

CD/DVD/BLU-RAY

CDs and *DVDs* are rigid disks of optical media a little less than 5 inches in diameter made of hard plastic with a thin layer of coating. A laser beam, along with an optoelectronic sensor, is used to write to and read the data that is "burned" into the coating material (a compound that changes from reflective to nonreflective when heated by the laser). The data is encoded in the form of

incredibly tiny pits or bumps on the surface of the disk. The different types of disks include:

- CD-R, which is short for CD-Recordable. This type of CD is a Write Once, Read Multiple (WORM) media that allows you to record data to it once, so that you can later read the data. Once data is written to a CD-R, no additional data can be written to the CD.
- CD-RW, which is short for CD-Rewritable and allows you to erase and write to the disk multiple times.
- CD-ROM is an acronym for Compact Disk—Read Only Memory; however, the term has grown to refer to the CD-ROM drive used to read this optical storage media.
- CD-ROMs are capable of holding up to 700 MB of data and remain a common method of storing data.
- CD and DVD media are unaffected by Electromagnetic Pulse (EMP) effects, X-rays, and other sources of electromagnetic radiation.
- The primary consideration with recordable CD media (and to a lesser extent, manufactured media) is energy transfer. It takes a significant amount of energy to affect the data that the writing laser transfers to the disk. Rewritable disks (discussed later) require even more energy to erase or rewrite data.
- Blu-Ray is a high-density optical storage method that was designed for recording high-definition video. The name of this technology comes from the blue-violet laser that is used to read and write to the disks. A single-layer Blu-Ray disk can store up to 25 GB of data, while a dual-layer Blu-Ray disk can store up to 50 GB of data.

MAGNETIC TAPE

In the early days of computing, *magnetic tape* was one of the few methods used to store data. Magnetic tape consists of a thin plastic strip that has magnetic coating on which data can be stored. Today magnet tape is still commonly used to back up data on network servers and individual computers, as it is a relatively inexpensive form of removable storage.

Network attached storage

Network attached storage (*NAS*) is a system that is connected to a network to provide centralized storage of data. A NAS is only used for data storage and is scaled down to provide access only to a file system in which data is stored and management tools that are accessed remotely. A NAS consists of a set of hard disks that can be configured as RAID arrays, and supports authentication, encryption, permissions, and rights with access to the data using protocols like Network File System (NFS) or Server Message Blocks (SMB).

SUMMARY OF EXAM OBJECTIVES

System security comprises a wide range of topics—from threats such as viruses, worms, bots, and Trojans to SPAM and pop-ups. In addition, system security is

not just concerned with software security but also physical, hardware security. From the BIOS to data storage to software system, security is one of the most complex topics in the security field today.

It is important to understand that while there are a multitude of threats out there, there are also many tools that are available to combat those threats. Anti-virus software has become a mainstay of the computing environment today. Similarly, personal firewalls are more ubiquitous than ever. It is the proper use of tools such as these that helps ensure the integrity and security of an end system in today's corporate environments.

TOP FIVE TOUGHEST QUESTIONS

1. You are analyzing the current security of your network and are concerned about the possibility that users will bypass authentication and gain greater permissions than they were given. What are the two major causes of privilege escalation? Choose all that apply.

   A. Bugs in software
   B. Spyware
   C. Backdoors
   D. BIOS

2. What are good ways to protect against worms? (Select all that apply.)

   A. User education programs
   B. Correct firewall configuration
   C. Timely software patches
   D. Anti-virus scans

3. Your company's Web server suddenly gets tens of thousands of simultaneous requests for a Web page. After the Web server crashes, you restart the server and then take a look at the log files. You see that some of the requests came from your own network. What kind of attack has most likely happened?

   A. Rootkit
   B. Botnet
   C. Virus
   D. Worm

4. You have purchased a used computer in an auction. When you power-on the computer, you are asked for a password before the operating system even loads. Since you don't have it, how will you clear the password so that you can start the computer and begin using it?

   A. Clear the password in the CMOS settings.
   B. Flash the BIOS.
   C. Press F10 or DEL on the keyboard.
   D. There is nothing you can do if you don't have the power-on password.

5. You have heard that upgrading the BIOS on a computer can help to fix any bugs and provide new features. You download a new BIOS version and begin the upgrade. Everything seems to go well, and you recycle the power on the computer. It doesn't start, but produces a blank screen. What most likely is the cause of the computer not starting?

   A. The wrong BIOS version was installed.
   B. There was a power outage during the upgrade.
   C. The CMOS editor needs to be reconfigured.
   D. You should never flash the BIOS as it will cause the computer to fail.

ANSWERS

1. The correct answers are A and C. Bugs in software and backdoors are two major causes for privilege escalation. Privilege escalation occurs when a user acquires greater permissions and rights than he or she was intended to receive. This can occur as a result of bugs (which are errors in code) or backdoors in software (which can bypass normal authentication). B is incorrect because spyware is used to monitor a system and send data to a third party. D is incorrect because the BIOS is low-level software on a computer that's used for recognizing and configuring hardware on a computer and starting the machine.
2. The correct answers are B and C. Firewalls can prevent ports like SQL and NetBIOS from being available and usable to worms. Most worms use known vulnerabilities, so timely patches will defend against them. A is incorrect because worms do not require user intervention, and so user education doesn't affect them. D is incorrect because a worm is not resident, and so can only be detected in memory, where it already has infected the machine.
3. The correct answer is B. Botnet. Computers have been turned into zombie machines after being infected with bots. The bot herder can then send commands to these machines to make requests from a specific Web site, preventing the server from serving legitimate requests from Web site users. When you attempt to view who caused the attack, it will only show those who have been infected with the bot. A is incorrect because a rootkit is used to acquire elevated permissions to a computer. C and D are incorrect because computers infected with a virus or worm wouldn't make tens of thousands of computers suddenly visit a Web site.
4. The correct answer is B. Flash the BIOS. By flashing the BIOS, you are erasing the existing settings by updating the BIOS software. A is incorrect because (although power-on passwords are set in the CMOS editor) you can't start the CMOS editor until you've entered the power-on password. C is incorrect because pressing keys on the computer won't help in this situation, unless of course you're entering the password. D is incorrect because you can flash the BIOS to reset all of the settings and clear the power-on password.

5. The correct answer is A. The wrong BIOS version was installed. Flashing the BIOS with a version that was meant for another motherboard can cause all sorts of problems, including the BIOS not being able to start the computer. When you are flashing the BIOS, it is important that the correct version for your computer is used. B is incorrect because (although a power outage would cause the BIOS upgrade to fail) the scenario says that everything seemed to go well during the upgrade. C is incorrect because correctly flashing the BIOS will clear any CMOS settings, restoring them to default settings. This wouldn't affect the computer not starting. D is incorrect because you can flash the BIOS to upgrade it.

CHAPTER 2

OS Hardening

Exam objectives in this chapter:

- General OS Hardening
- Server OS Hardening
- Workstation OS

GENERAL OS HARDENING

Operating system hardening involves making the operating system less vulnerable to threats. There are numerous best practices documents that can be followed in a step-by-step approach to harden an operating system. One of the first places to look at when securing a system is the structure and security settings on files and directories.

- Start with everything accessible and lock down the things to be restricted.
- Start with everything locked down and open up those files necessary to allow access to.

Of these two potential methods, the second, which is also referred to as the rule of least privilege, is the preferred method. Least privilege starts with the most secure environment and then loosens the controls as needed. This method tends to be the most restrictive, with authorizations provided to users, processes, or applications that access these resources on a needs-only basis. Accessibility and security are usually at opposite ends of the spectrum; this means that the more convenient it is for users to access data, the less secure the network.

Here are the general steps to follow for securing an OS:

1. Disable all unnecessary services.
2. Restrict permissions on files and access to the Registry.
3. Remove unnecessary programs.
4. Apply the latest patches and fixes.

Services

Like servers, many workstations also have the ability to enable and disable services. Services can be disabled through the Services administration tool on Windows platforms, by commenting the service out of *inetd.conf*, or by disabling it through the appropriate service file in *xinetd.conf* under UNIX. It is considered a best practice to disable any services on a workstation that are not required. While considering the removal of nonessential services, it is important to look at every area of the computer's application to determine what is actually occurring and running on the system.

File system

Controlling access is an important element in maintaining system security. The most secure environments follow the "least privileged" principle, as mentioned earlier, which states that users are granted the least amount of access possible that still enables them to complete their required work tasks. Expansions to that access are carefully considered before being implemented. Law enforcement officers and those in government agencies are familiar with this principle regarding noncomputerized information, where the concept is usually termed *need to know*.

In practice, maintaining the least privileged principle directly affects the level of administrative, management, and auditing overhead, increasing the levels required to implement and maintain the environment. One alternative, the use of user groups, is a great time saver. Instead of assigning individual access controls, groups of similar users are assigned the same access. In cases where all users in a group have exactly the same access needs, this method works. However, in many cases, individual users need more or less access than other group members. When security is important, the extra effort to fine-tune individual user access provides greater control over what each user can and cannot access.

Keeping individual user access as specific as possible limits some threats, such as the possibility that a single compromised user account could grant a hacker unrestricted access. It does not, however, prevent the compromise of more privileged accounts, such as those of administrators or specific service operators. It does force intruders to focus their efforts on the privileged accounts, where stronger controls and more diligent auditing should occur.

Removing unnecessary programs

The default installation of many operating systems includes programs that are unnecessary. It is therefore very important that an organization with the resources to do so create their own operating system images and remove any unnecessary programs or features. For example, the default installation of many Linux-based operating systems includes a telnet server as part of the base install. Depending on the flavor of Linux, this server may be operational when it is not needed or desired.

Hotfixes/patches

Updates are typically provided by the manufacturer of a specific component or operating system. Updates contain improvements and new or improved components that the manufacturer believes will make the product more stable, usable, secure, or otherwise attractive to end users. For example, Microsoft updates are often specifically labeled Security Updates and can be found at www.microsoft.com/protect/default.mspx. These updates address security concerns recognized by Microsoft, and should be evaluated and installed as needed.

It's a good idea to keep up with the hotfixes and patches for operating systems, with many vendors providing regular patch releases and periodic hotfixes. Many of the hotfixes and patches will address security-related features.

Vendors' Web sites contain information regarding patches and hotfixes. One good location would be the Computer Emergency Response Team's (CERT) Web site, which may be found at www.cert.org. An equally valuable resource is the SecurityFocus Web site at www.securityfocus.com, which has operating system–specific mailing lists administrators can join to receive regular updates on available patches, information on security flaws to be aware of, and discussions on current security topics and best practices.

Service packs/maintenance updates

HOTFIXES

Hotfixes are packages that can contain one or more patches for software. They are generally created by the vendor either when a number of clients indicate there is a compatibility or functional problem with a manufacturer's products used on particular hardware platforms or when a vulnerability in an operating system's software component is discovered. These are mainly fixes for known or reported problems that may be limited in scope.

SERVICE PACKS

Service packs are accumulated sets of updates or hotfixes. Service packs are usually tested over a wide range of hardware and applications in an attempt to assure compatibility with existing patches and updates, and to initiate much broader coverage than just hotfixes. The recommendations discussed previously also apply to service pack installation.

Service packs must be fully tested and verified before being installed on live systems. Although most vendors of OS software attempt to test all of the components of a service pack before distribution, it is impossible for them to test every possible system configuration that may be encountered in the field.

Patch management

PATCHES

Patches for operating systems and applications are available from the vendor supplying the product. These are available by way of the vendor's Web site or

from mirror sites around the world. They are often security-related, and may be grouped together into a cumulative patch to repair many problems at once. Except for Microsoft, most vendors issue patches at unpredictable intervals; it is therefore important to stay on top of their availability and install them after they have been tested and evaluated in a nonproduction environment. The exception to this is when preparing a new, clean install. In this case, it is considered a best practice to download and install all known patches prior to introducing the machines to the network.

SCRIPTS

Scripts are a versatile way to manage patches. They can be used to perform custom installations, automatic installations, and pretty much anything a programmer is clever enough to write a script for.

PATCH MANAGEMENT SYSTEMS

As operating systems have become more complex, the need for patch management became more critical. There are many systems out there for managing patches, including open source patch management systems, "home grown" systems, Symantec's Altiris, Microsoft's System Management Server/System Center, and Microsoft's Windows Software Update Services.

Altiris

Symantec's Altiris management software allows for the management of a wide spectrum of clients, including Windows, UNIX, Linux, and MacOS machines—all from a single management platform. Altiris has the ability to discover, catalog, and inventory software on Windows, UNIX, Linux, and Mac systems, which can help determine the patch level of the computers in your organization. In addition, the Altiris system can push patches to the end clients as well as verify their system configurations and tune them if necessary.

System Management Server (SMS)/System Center

Microsoft's SMS 2003 and System Center 2007 products are designed to aid in monitoring system health and also can be used to distribute software and settings out to different groups of computers in an organization. SMS 2003 and System Center rely heavily on Active Directory and integrate tightly with Windows group policy.

Windows Software Update Services

Windows Software Update Services (WSUS) is a freely available product that allows enterprise users to manage Microsoft updates on their computers running the Windows operating system. WSUS in its simplest form gets the latest updates from Microsoft and allows the administrators to determine whether to approve or decline individual update as well as to distribute them across their infrastructure.

Windows group policies

Group policy in Windows allows administrators to set security settings as well as install specific software (such as virus scanning) on a group of computers. System administrators use Group Policy to manage all aspects of the client desktop environment for Windows clients (Windows Servers and Workstations), including Registry settings, software installation, scripts, security settings, etc. The possibilities of what can be done with Group Policy are almost limitless. With VBScript, Jscript, or PowerShell, administrators can write entire applications to execute via Group Policy as well as install software automatically across the network and apply patches to applications.

When you are deciding on the Group Policies to enforce on the network, it is important to keep in mind that the more policies that are applied, the more network traffic generated and hence the longer it could take for users to log onto the network. Group policies are stored in Active Directory as Group Policy Objects (GPOs). These objects are the instructions for the management task to perform.

Group Policy is implemented in four ways:

- ***Local Group Policy:*** Local Group Policy is configured on the local computer.
- ***Site Group Policy:*** Site Group Policies are linked to a "site" and can generate unwanted network traffic.
- ***Domain Group Policy:*** A Domain Group Policy is linked to an Active Directory domain and applies group policy objects to all computers and users within a domain.
- ***Organizational Unit Group Policy:*** A Group Policy object that is linked to the organizational unit (OU), which is especially useful for applying a Group Policy object to a logical grouping (organizational unit) of users or computers.

Security templates

Security templates are basically a "starting point" for defining system settings in Windows. These templates contain hundreds of possible settings that can control a single computer or a whole network of computers and can be customized extensively. Some of the areas that security templates control include user rights, password policies, system policies, and user and system permissions. The base security templates provided by Microsoft are predefined settings to accomplish a specific task. For example, *compatws* in Windows is used to reduce the security level to allow older applications to run and *hisecdc* is used to apply a high security level to a domain controller. Similarly, *hisecws* is used to apply stringent security controls on a workstation. Windows security templates can be found in C:\Windows\Security\templates in XP/Server 2003. The security templates for Windows Vista are available in the Vista Security Guide available at http://www.microsoft.com.

DID YOU KNOW?

When *making a new template*, you can save a lot of time and aggravation to start with one of the windows templates that's already created.

SE LINUX

Security Enhanced (SE) Linux allows for the application of security policies through the use of *Linux Security Modules* (*LSM*) in the kernel. Some of the capabilities introduced in SE Linux include the use of Mandatory Access Controls (MAC), controls over network sockets, file systems, directories, and processes. *Bastille UNIX* is an automated security setup tool that was originally written specifically for the Linux operating system. Bastille UNIX provides a level of security on the basis of the usage of the server. The administrator answers a series of questions, and on the basis of the answers the settings are determined and then applied. Bastille UNIX is freely available at www.bastille-unix.org.

Configuration baselines

Configuration baselines are standard setups used when configuring machines in organizations. *Configuration baselines* are used to provide a starting point where machines can then be customized with respect to their specific roles in the network. For example, a Windows domain controller may not require Windows Media Services to be installed since its primary function is that of a directory service. A Web server would not necessarily require a database to be installed. Additionally, specific services would be installed, turned off, or even removed completely on the basis of the final location of the system in the network architecture.

DETERMINING CONFIGURATION BASELINES

When you are considering baselines for an organization, it is important to always keep in mind the principle of least access. The function of each system in the network defines the appropriate baseline for that system. Each of the systems listed below requires specific baseline configurations that should be developed before the systems are deployed on their network:

- Web server
- File and print server
- Database server
- Domain controller
- Normal workstation
- Developer workstation
- DNS, DHCP server

This list is by no means exhaustive. If the above systems are all Windows systems, each category may require its own security template. For example, the domain controllers may have the *hisecdc* security template applied since they

contain user account information as well as directory services for the organization as a whole. The normal workstation may only need to have the *compatws* template applied as the end workstations will only be used by the regular users. The Web servers as well as the DNS servers will most likely have tight security requirements as they could be placed outside the corporate firewall in a DMZ that is accessible from the Internet.

It is important to remember that the generic security templates provided by Microsoft or used in such hardening tools as *Bastille UNIX* will need to be further customized by an organization in order to meet their specific security requirements.

MICROSOFT BASELINE SECURITY ANALYZER

The Microsoft Baseline Security Analyzer (MBSA) is a free tool for small and medium-sized businesses that can be used to analyze the security state of a Windows network relative to Microsoft's own security recommendations. In addition to identifying security issues, the tool offers specific remediation guidance. MBSA will detect common security misconfigurations and missing security updates on Windows systems. The MBSA is an excellent tool that will provide insight into security vulnerabilities in your organization.

SERVER OS HARDENING

Server OS hardening can be a very complex and daunting task. However, by following a standard set of procedures and utilizing tools like security templates and MBSA, this task can be made significantly easier and can result in improved security across your network. One of the first tasks to focus on is deciding which services and protocols need to be enabled and which should be disabled.

Enabling and disabling services and protocols

When you are considering whether to enable and disable services and protocols in relation to network hardening, there are extra tasks that must be done to protect the network and its internal systems. As with operating systems discussed earlier, it is important to evaluate the current needs and conditions of the network and infrastructure, and then begin to eliminate unnecessary services and protocols.

Eliminating unnecessary network protocols includes eliminating those that aren't used on your network. While removal of nonessential protocols is important, it is equally important to look at every area of the network to determine what is actually occurring and running on systems. The appropriate tools are needed to do this, and the Internet contains a wealth of resources for tools and information to analyze and inspect systems.

FTP servers

FTP servers are potential security problems as they are typically open to the Internet to support anonymous access to public resources. Incorrect file system

settings in a server acting as an FTP server allows unrestricted access to all resources stored on that server and could lead to a system breach. FTP servers exposed to the Internet should be placed in a Demilitarized Zone (DMZ) and hardened with all available operating system patches. All services other than FTP should be disabled or removed and contact from the internal network to the FTP server through the firewall should be restricted and controlled through Access Control List (ACL) entries, to prevent possible traffic through the FTP server from returning to the internal network.

Some of the hardening tasks that should be performed on FTP servers include:

- Protection of the server file system
- Isolation of the FTP directories
- Positive creation of authorization and access control rules
- Regular review of logs
- Regular review of directory content to detect unauthorized files and usage

DNS servers

Hardening DNS servers consists of performing normal OS hardening and then considering the types of control that can be done with the DNS service itself. Older versions of BIND DNS were not always easy to configure, but current versions running on Linux and UNIX platforms can be secured relatively easily.

Zone transfers should only be allowed to designated servers. Additionally, those users who may successfully query the zone records with utilities such as *nslookup* should be restricted via the access control list (ACL) settings. Windows Server 2003 DNS server added controls to prevent zone transfer operations to machines that are not approved to request such information, thus better protecting the resources in the zone files from unauthorized use. Another best practice would be to not use HINFO records in the DNS server.

Other attacks administrators must harden against include denial of service attacks (DoS) as well as cache poisoning, in which a server is fed altered or spoofed records that are retained and then duplicated elsewhere.

NNTP servers

NNTP servers are also vulnerable to some types of attacks, because they are often heavily utilized from a network resource perspective. NNTP servers that are used to carry high volumes of newsgroup traffic from Internet feeds are vulnerable to DOS attacks that can be mounted when "flame wars" occur. This vulnerability also exists in the case of *listserv* applications used for mailing lists. NNTP servers also have vulnerabilities similar to e-mail servers, because they are not always configured correctly to set storage parameters, purge newsgroup records, or limit attachments.

File and print servers

The ability to share files and printers with other members of a network can make many tasks simpler and, in fact, this was the original purpose for networking computers. However, this ability also has a dark side, especially when users are unaware that they are sharing resources. If a trusted user can gain access, the possibility exists that a malicious user can also obtain access. On systems linked by broadband connections, crackers have all the time they need to connect to shared resources and exploit them.

If a user does not need to share resources with anyone on the internal (local) network, the file- and print-sharing service should be completely disabled. On most networks where security is important, this service is disabled on all clients. This action forces all shared resources to be stored on network servers, which typically have better security and access controls than end-user client systems.

DHCP servers

DHCP servers add another layer of complexity to some layers of security, but also offer the opportunity to control network addressing for client machines. This allows for a more secure environment if the client machines are configured properly. In the case of the clients, this means that administrators have to establish a strong ACL to limit the ability of users to modify network settings, regardless of platform. Nearly all operating systems offer the ability to add DHCP server applications to their server versions.

Additional security concerns arise with DHCP. Among these, it is important to control the creation of extra DHCP servers and their connections to the network. A rogue DHCP server can deliver addresses to clients, defeating the settings and control efforts for client connection.

Data repositories

NAS and SAN configurations may present special challenges to hardening. For example, some NAS configurations used in a local area network (LAN) environment may have different file system access protections in place that will not interoperate with the host network's OS and NOS. In this case, a server OS is not responsible for the permissions assigned to the data access, which may make configuration of access or integration of the access rules more complex. SAN configuration allows for intercommunication between the devices that are being used for the SAN, and thus freedom from much of the normal network traffic in the LAN, providing faster access. However, extra effort is initially required to create adequate access controls to limit unauthorized contact with the data it is processing.

DIRECTORY SERVICES

Hardening of directory services systems requires evaluation not only of the permissions to access information, but of permissions for the objects that are contained in the database. Additionally, these systems require the use of LDAP on

the network, which also requires evaluation and configuration for secure operation. This includes setting perimeter access controls to block access to LDAP directories in the internal network if they are not public information databases. Maintenance of security-based patches and updates from the vendor is absolutely imperative in keeping these systems secure.

NETWORK ACCESS CONTROL

Another way to harden the network is to use Network Access Control (NAC). There are several different incarnations of NAC available:

1. Infrastructure-based NAC requires an organization to be running the most current hardware and OSs. Operating system platforms such as Microsoft's Windows Vista have the ability to participate in NAC.
2. Endpoint-based NAC requires the installation of software agents on each network client. These devices are then managed by a centralized management console.
3. Hardware-based NAC requires the installation of a network appliance. The appliance monitors for specific behavior and can limit device connectivity should noncompliant activity be detected.

NAC offers administrators a way to verify that devices meet certain health standards before they're allowed to connect to the network. Laptops, desktop computers, or any device that doesn't comply with predefined requirements can be prevented from joining the network or can even be relegated to a controlled network where access is restricted until the device is brought up to the required security standards.

DATABASES

Database servers may include servers running SQL or other databases such as Oracle. These types of databases present unique and challenging conditions when considering hardening the system. For example, in most SQL-based systems, there is both a server function and a client front end that must be considered. In most database systems, access to the database information, creation of new databases, and maintenance of the databases are controlled through accounts and permissions created by the application itself. Although some databases allow the integration of access permissions for authenticated users in the directory services system, they still depend on locally created permissions to control most access. This makes the operation and security of these types of servers more complicated than is seen in other types.

Unique challenges exist in the hardening of database servers. Most require the use of extra components on client machines and the design of forms for access to the data structure, to retrieve the information from the tables constructed by the database administrator. Permissions can be extremely complex, as rules must be defined to allow individuals to query database access to some records and no access to others. This process is much like setting access permissions, but at a much more granular and complex level.

EXAM WARNING

Spend a few minutes reviewing port and protocol numbers for standard services provided in the network environment. This will help when you are analyzing questions that require configuration of ACL lists and determination of appropriate blocks to install to secure a network.

The Security+ exam can ask specific questions about ports and what services they support. It's advisable to learn common ports before attempting the exam.

21 FTP

22 Secure Shell (SSH)

23 Telnet

25 Simple Mail Transfer Protocol (SMTP)

53 DNS

80 HTTP

110 Post Office Protocol (POP)

161 Simple Network Management Protocol (SNMP)

443 SSL

Memorizing these will help you with the Security+ exam.

WORKSTATION OS

Workstations can present special challenges. Depending on a user's knowledge and capabilities, they may modify the steps it takes to secure their workstation and violate company policy when it comes to best practices. As laptops become more commonplace, they present specific challenges to the organization when it comes to securing operating systems, including configuration of the appropriate services as well as user and group rights.

User rights and groups

Ideally, the minimum required rights for a person to perform their job should be given. Under older Windows operating systems (XP and 2000 most notably), the user of a machine was given administrative rights or was added to the "Power Users" group in order to gain full functionality from the operating system. However, if a user account is compromised, the entire machine could be compromised, which could potentially lead to the entire domain being compromised. Under Vista and Windows 7, users no longer need to have administrative privileges to their systems in order

TIP

Remember the principle of least access! In many cases, this will help you to make the correct choice.

to be able to be fully functional. This allows the system administrator to reduce the rights assigned to regular users and follows the principle of least access.

SUMMARY OF EXAM OBJECTIVES

This chapter looked at the broad concept of infrastructure security and specifically discussed the concepts and processes for hardening various sections of systems and networks. OS security and configuration protections were discussed as were file system permission procedures, access control requirements, and methods to protect the core systems from attack. Security+ exam objectives were studied in relation to OS hardening and in relation to hardening by visiting potential problem areas including configuration concerns, ACLs, and elimination of unnecessary protocols and services from the computer. We also looked at how these hardening steps might improve and work with the OS hardening and ways to obtain, install, and test various fixes and software updates.

TOP FIVE TOUGHEST QUESTIONS

1. As part of the overall operating system hardening process, you are disabling services on a Windows server machine. How do you decide which services to disable?

   A. Disable all services, and then reenable them one by one.
   B. Research the services required and their dependencies, then disable the unneeded services.
   C. Leave all services enabled, since they may be required at some point in the future.
   D. Disable all workstation services.

2. Robby is preparing to evaluate the security on his Windows XP computer and would like to harden the OS. He is concerned as there have been reports of buffer overflows. What would you suggest he do to reduce this risk?

   A. Remove sample files.
   B. Upgrade his OS.
   C. Set appropriate permissions on files.
   D. Install the latest patches.

3. Yesterday, everything seemed to be running perfectly on the network. Today, the Windows 2003 production servers keep crashing and running erratically. The only events that have taken place are a scheduled backup, a CD/DVD upgrade on several machines, and an unscheduled patch install. What do you think has gone wrong?

   A. The backup altered the archive bit on the backup systems.
   B. The CD/DVDs are not compatible with the systems in which they were installed.

C. The patches were not tested before installation.
D. The wrong patches were installed.

4. You have been asked to review the general steps used to secure an OS. You have already obtained permission to disable all unnecessary services. What should be your next step?

A. Remove unnecessary user accounts and implement password guidelines.
B. Remove unnecessary programs.
C. Apply the latest patches and fixes.
D. Restrict permissions on files and access to the Registry.

5. During a routine check of a file server, you discover a hidden share someone created that contains 100 GB of music content. You discover that the share was created on a drive that everyone has full control over. What steps should you take to ensure this doesn't happen again?

A. Define an acceptable use policy.
B. Remove full control from the "Everyone" group.
C. Remove full control from the offending user.
D. Remove the files and the directory.

ANSWERS

1. The correct Answer is B. It is important that you understand why services are needed and what their dependencies are. Answer A is wrong as you may not know which services are needed or not. Answer C is wrong as it leaves too many services running on the machine. Answer D is wrong as the workstation services are still required even on a Windows server machine.

2. The correct answer is D. It is important to keep systems updated to the latest patches in order to protect the system from known vulnerabilities and exploits. Answers A, B, and C are wrong as, while they do provide some level of protection, the best method of protecting a system against buffer overflows is to apply the latest patches for the system.

3. The correct answer is C. Answer A is incorrect as a backup would not cause a system-wide failure of all the Windows 2003 servers. Answer B is incorrect as all the Windows 2003 servers are behaving erratically—not just the ones that had a CD/DVD upgrade. Answer D is incorrect as operating system patches to the Windows operating system are system type specific and the installation process prevents patches that are not meant for a specific operating system to be installed on that system.

4. The correct answer is A. Answer B is incorrect as removing unnecessary programs would come after the removal of unnecessary user accounts

and the implementation of password guidelines. Answer C should come as the first step before disabling unnecessary services. Answer D, restricting permissions on files and Registry access, will be one of the last steps done to secure the OS.

5. The correct answers are A, B, and D. Answer C is incorrect since everyone has full control over the drive and the hidden share could have been created by someone but ownership could have been set to another account by the user to hide their connection to the music and the share.

CHAPTER 3

Application Security

Exam objectives in this chapter:

- Threats Are Moving Up the Stack
- Application Security Threats

THREATS ARE MOVING "UP THE STACK"

Data must pass through multiple layers of communication when sent from one network device to another. The *OSI model* details seven layers of communication, and when you view the model from the bottom up, each layer ultimately supports the layer above it. The OSI model consists of:

- ***Application Layer:*** Network process to application
- ***Presentation Layer:*** Data representation and encryption
- ***Session Layer:*** Interhost communication
- ***Transport Layer:*** End-to-end connections and reliability
- ***Data Link Layer:*** Physical addressing
- ***Network Layer:*** Logical addressing using IPv4 or IPv6
- ***Physical Layer:*** Media, signal, and binary transmission

Over recent years, there has been a large shift in the focus of computer-related attacks moving from lower layers of the OSI model to the application layer. This shift is due to changes in network architecture and security technologies as well as efforts by vendors of operating systems (Sun, Microsoft, etc.) to harden the underlying operating system from attack.

EXAM WARNING

In preparation for the exam, you should understand the seven layers of the OSI model as well as the specific function within each.

Rationale

The motive behind computer attacks has shifted from generating large Denial-of-Service (DoS) to covert financially motivated attacks. Financially motivated attacks involve data that is withheld, manipulated, or resold for financial benefit, including personal information such as health and financial data being prime targets of cyber crime.

Threat modeling

Threat modeling is a comprehensive process for assessing a system's security risks and can be applied to any information system. A traditional vulnerability assessment performed within the corporate world involves the following tasks:

- Running an automated vulnerability scanning tool against an infrastructure
- Generating scan results and associating findings with a generic risk rating that was developed by the vulnerability scanning tool vendor
- Qualifying scan results and sending them out to the appropriate individuals for remediation

Automated scans look primarily at common forms of vulnerabilities such as:

- Insecure coding practices
- Misconfigurations
- Missing patches

Threat modeling uses a systematic approach and takes a holistic view of security to identify the threats and vulnerabilities that threaten defined objectives. Threat modeling can be subdivided into five stages:

- ***Security Objective Definition:*** In this phase, the security objectives placed on the application are identified, thus helping to control the scope of the threat modeling process.
- ***Application Review:*** In this stage, the application solution and design documentation are reviewed to identify key functionalities, with special attention being placed on the application architecture and technologies in use, how the application is used, and the security mechanisms in use.
- ***Application Decomposition:*** This stage focuses on the in-depth review of application internals such as ingress and egress data flows and application trust boundaries. Trust boundaries mark areas within applications that require a change in trust.
- ***Threat Identification:*** Threats to the earlier defined security objectives are identified factoring in knowledge gained during *Application Decomposition* where participants in brainstorming sessions review prior collected information to identify possible areas of attack.
- ***Vulnerability Identification:*** On the basis of the earlier documented threats, the application is reviewed and specific vulnerabilities are documented.

DID YOU KNOW?

Threat modeling can be an effective process if the correct people are involved. The best threat models involve representation from many groups within an organization, including security analysts, developers, business analysts, and information architects. Combined, this group will provide a comprehensive understanding of the application, associated technologies, and vulnerabilities, thereby creating a much better threat model than one created solely by a security analyst.

APPLICATION SECURITY THREATS

Application security involves securing both custom-developed as well as Common Off-The-Shelf (COTS) applications.

Browser

The primary purpose of using a Web browser is to navigate and interact with Web-based applications. With over 248 million Internet users in North America alone, it's not difficult to see why these widely deployed applications are a target for cyber crime. Browser-based vulnerability was ranked the number one threat in 2007 by the SysAdmin, Audit, Network, Security Institute (SANS) in its report titled "SANS Top 20 2007 Security Risks (2007 Annual Update)" (see www.sans.org/top20/) and again in 2008 within its report titled "Top Ten Cyber Security Menaces for 2008" (see www.sans.org/2008menaces/).

DRIVE-BY-DOWNLOAD

Drive-by-download attacks occur when a user navigates to or is unknowingly directed to a malicious Web site and hostile content is automatically downloaded and executed on their computers. This code when executed can provide a hacker full control of the visiting user's computer, and the user normally has no idea that this attack has occurred. One of the most widely used Web technologies actively exploited by hackers to carry out drive-by-download and other forms of attacks is ActiveX.

ACTIVEX

ActiveX enables software applications to share and reuse software components, called ActiveX controls. These controls are tiny applications that can be developed using various programming languages such as C-Sharp (C#), Visual C++, Visual Basic, and Java with controls written in one language actually sharing code with controls written in another. ActiveX controls greatly enhance Web applications.

Securing ActiveX controls within the Web browser

Numerous vulnerabilities have been identified with both vendor-shipped and third-party–developed ActiveX controls. To help minimize this risk, there

are some steps users can take to safeguard their machines against ActiveX exploitation:

- Ensure that the computer is up to date with security patches.
- Don't click on suspicious links or navigate to Web sites you are not familiar with. Avoiding sites and links you are not familiar with can be an effective way to avoid the execution of malicious code.
- Utilize browser-based security zones—granular ActiveX restrictions should be implemented using zones.

A *zone* is a named collection of Web sites (from the Internet or a local intranet) that can be assigned a specific security level.

Each zone is assigned a predefined security level or a custom level can be created. These possible settings are:

- **Low**, which provides the least security and allows all ActiveX content to run.
- **Medium–Low**, the default setting for the Local intranet zone and provides the same security as the Medium level except that users aren't prompted.
- **Medium**, the default level for trusted sites and the lowest setting available for the Internet zone; unsigned ActiveX content isn't downloaded, and the user is prompted before downloading potentially unsafe content.
- **Medium–High**, which is the default setting for the Internet zone, as it is suitable for most Web sites. Unsigned ActiveX content isn't downloaded, and the user is prompted before downloading potentially unsafe content.
- **High**, which is not only the default level for restricted sites but also the only level available for that zone. It is the most restrictive setting and has a minimum number of security features enabled.

Custom security levels can be defined to fit the specific security restrictions of an environment. Within a custom security level, there are numerous individual security controls related to how ActiveX, downloads, Java, data management, data handling, scripting, and logon are handled.

Developing secure ActiveX controls

In response to vulnerabilities within ActiveX controls, Microsoft introduced Authenticode to help ensure the integrity and nonrepudiation of ActiveX controls. Authenticode is a method of code signing that allows developers to obtain a digital certificate generated by a Certificate Authority (CA) and digitally sign an ActiveX control. Developers can use the following recommendations to help minimize the number of vulnerabilities that exist within developed ActiveX controls:

- **Follow secure coding practices:** Secure coding practices including data validation can be obtained from the Microsoft Development Network (MSDN). (See http://msdn.microsoft.com/en-us/library/aa752035.aspx.)
- **Use Authenticode:** Sign controls with a certificate issued from a trusted CA to ensure that ActiveX controls are not tampered with after they are developed.

JAVA

Java is a programming language, developed by Sun Microsystems, that is used to make small applications (applets) for the Internet as well as stand-alone programs utilizing an interpreter called the *Java Runtime Environment* (*JRE*). A core component of the JRE is the *Java Virtual Machine* (*JVM*), which is a collection of programs that execute applications and scripts and supports a computer intermediate language referred to as Java bytecode. The JVM also incorporates security features such as the bytecode verifier, which verifies the code for a list of predetermined insecurities, and sandboxing, which isolates executing code in a reserved area of memory to limit the damage potentially malicious code could inflict on the user's machine.

Crunch Time

ActiveX-related vulnerabilities will be covered on the exam. In preparation, you should ensure that you are familiar with IE security zones, default permissions, and how to add or remove sites from zones.

Developing secure Java applets

Developers who write Java applets can help secure their code by implementing code signing. The JVM uses sandboxing to restrict the damage a Java applet can inflict on a user's computer; however, when a control is digitally signed, it is allowed to leave the sandbox and obtain access to client resources, possibly resulting in a security issue.

Securing the execution of Java applets

A key security component within the JVM is a built-in Security Manager that controls the level of restrictions placed on executing Java bytecode. This includes what code must run within a sandbox. Digitally signed Java applets (similar to Authenticode within ActiveX) are, however, allowed to escape the sandbox for a greater level of access to client system resources.

These restrictions are controlled by the user through security policies, which are similar to zones in Internet Explorer. To secure the execution of Java applets on local clients, the following recommendations can be followed:

- **Ensure that systems are regularly patched:** Java applets like other browser-based technologies are developed by numerous third-party organizations and require vigilance to ensure that the latest security patches have been applied to correct vulnerabilities.
- **Use Java security policies:** Local security policies can be used to restrict the level of privileges downloaded Java applets (including signed applets) have on the local computer.

- **Don't click on suspicious links or navigate to Web sites you are not familiar with:** User vigilance is an important element of Java security. Avoiding unfamiliar sites and links can be an effective way to avoid the execution of malicious code.

SCRIPTING

Unlike ActiveX and Java applets, which are developed in actual programming languages (Visual Basic/C++ and Java, respectively), lightweight scripting was released by Microsoft and Netscape to allow people with no formal programming experience to develop flexible Web pages. However, similar to ActiveX and Java applets, these scripts could be exploited, resulting in many attacks, including the drive-by-download discussed earlier. The Internet today is dominated by a handful of scripting languages: JavaScript, Active Scripting, VBScript, and Jscript.

Javascript

Javascript performs client-side Web development and the reuse functionality within other Web objects. Javascript was designed to look like Java, but it is a much simpler language to grasp and carries the same type of vulnerabilities as Java. JavaScripts are downloaded and run inside a sandbox, which prevents execution of privileged tasks such as reading and writing files on the local computer or accessing additional information.

Active Scripting

Active Scripting is a Microsoft-developed scripting language similar to ActiveX that enables software components to share information and interact with each other. It was commonly used to support animation and dynamic content within Web pages and/or e-mail clients. Active Scripting has been deprecated in favor of .NET and ASP.NET.

.NET

.NET is Microsoft's software framework running on the Windows operating system. It utilizes a Common Language Runtime (CLR) environment to execute software programs in an application virtual machine as well as provides security, memory management, and exception handling services. The most current version of the .NET framework is v3.5, which is available in Windows 7 and Windows Server 2008 R2 as well as a downloadable addition to Vista, XP, and Server 2003/2008.

VBScript and Jscript

VBScript is a scripting language developed by Microsoft to compete with Netscape's *JavaScript* and was regarded by many as even easier to use than Java. After seeing the widespread adoption and success Netscape achieved with JavaScript, Microsoft developed Jscript in 1996 as a comparable language for Microsoft systems. VBScript and Jscript scripts are tiny pieces of code that are similar to Active Scripting and allow developers to extend and reuse Web functionality. When a user connects to a Web server, the scripts are downloaded and

executed on the user's machine depending on the security zone the site resides within on recent editions of Internet Explorer.

Vulnerabilities with JavaScript, Active Scripting, VBScript, and Jscript

JavaScript, Active Scripting, VBScript, and Jscript all suffer from similar vulnerabilities as do ActiveX and Java applets; poorly written code can be exploited by attackers, as can the scripting engine that processes the downloaded scripts.

Securing the execution of client-side scripts

In order to secure browsers against Active Scripting vulnerabilities, you can follow similar recommendations to those we looked at to secure ActiveX and Java:

1. Ensure that application updates are downloaded regularly.
2. Use browser security zones.
3. Don't click on suspicious links or navigate to Web sites you are not familiar with.

EXAM WARNING

Remember that the exploitation of browser-based vulnerabilities within ActiveX, Java, and scripting only directly impact the client computers. However, it is possible for a compromise of credentials stored or typed into a client computer to result in unauthorized access to the very Web application who issued the vulnerable code.

COOKIES

Cookies are small text files that are downloaded and locally stored by a user's browser, and typically contain information about the user's session and/or preferences. Occasionally, Web sites also store authentication-related information such as usernames and passwords. Each time the user visits the Web site, the cookie is retrieved by the site's Web application and data from the cookie is processed. Storing this information within client-side cookies prevents Web sites from having to store and maintain information about all user sessions and preferences.

There are three main cookie types: Session, Persistent, and Tracking.

- ***Session Cookies:*** Session cookies are used by Web applications to store information, and when a user closes his or her Web browser session, the cookie is deleted. Session cookies can often contain authentication-related information about the user's session, such as display preferences and in some cases session identifiers or usernames and passwords.
- ***Persistent Cookies:*** Persistent cookies are also used by Web applications to store information about the user connection. Persistent cookies

are typically used to store user preferences about a Web site that are nonsensitive; therefore, there is less of a concern with it persistently stored on a user's hard drive. These cookies, however, are not deleted when the user closes his or her Web browser (session). Instead, the cookies have a timeout value set by the application and the cookies are downloaded by a user's Web browser and stored up until the expiration of the timeout value.

- ***Tracking Cookies:*** When a user connects to certain Web sites, a tracking cookie may be downloaded in the background. As their name suggests, these cookies are used to record users' Web activity, such as the type and specific sites they visit. Many sites may use the same form of tracking cookie. If the same tracking cookie is used by multiple sites, then these sites can all read and write to the contents of it.

Cookie vulnerabilities

Applications use cookies as files to store data for processing. Anytime there are data inputs into an application, there are potential security risks, and cookies are no different. As small and seamless as cookies are, they present a large security concern and are the target of many types of attacks.

Cookie hijacking

Cookies containing sensitive information such as usernames, passwords, or session identifiers can be a target for hackers. Attackers can sniff network traffic and capture a cookie downloaded from a site to a Web browser or gain access to a computer and view a cookie stored on the local hard drive. By capturing cookies, it's possible for an attacker to initiate another session to the same Web site and submit your cookie in order to bypass site authentication and perform actions within your account without your knowledge.

DID YOU KNOW?

A *common countermeasure to cookie hijacking attacks* is to send cookies over encrypted channels. This prevents cookies from being intercepted and replayed or disclosed. However, a recent tool named Cookie Monster showed how easy it really is to hijack cookies even sent over encrypted channels. Cookie Monster is an advanced cookie hijacking tool that was debuted in 2008 at DefCon 16, a popular hacking security conference. There, a demonstration was provided showing how Cookie Monster will monitor network traffic and filter for HTTPS connections. Cookie Monster stores information about the connection that it will later use to replay to the Web site the next time the user uses the Internet in order to trick the Web site to sending the authentication cookie over HTTP—a nonencyrpted channel. Cookies are harvested and can be used by an attacker to masquerade as a legitimate user and gain unauthorized access to many popular Web sites. Additional information can be found on the http://defcon.org Web site.

Cookie poisoning

Some cookies used by popular Web sites store authentication data such as session identifiers, usernames, and passwords. Cookie poisoning involves the modification of the data stored within a cookie. When the cookie containing the modified contents is used by the application, the values entered by the attacker are processed by the application.

Cookie leaking

Cookie leaking occurs when sensitive information such as usernames, passwords, and account numbers are stored within cookies and then the information is obtained by unauthorized users.

Preventing against cookie attacks

The following steps can be used to prevent or minimize the impact of cookie-related attacks:

For Web developers

- Use Secure bit to ensure that the cookie is transferred over only secure channels. Always initiate SSL connections to supported sites to help prevent the network interception of cookies.
- Web site developers should avoid using cookies to hold sensitive information. If absolutely necessary, the data values should be encrypted.

For Users

- Block third-party cookies. Third-party cookies are any cookie that does not originate from the domain you are visiting. Third-party cookies typically scan newly added cookies added to your system and on the basis of key words will generate pop-up advertisements.

CROSS-SITE SCRIPTING

Cross-site scripting (*XSS*) attacks occur when one user injects malicious code into a Web site where it is downloaded and executed by another user. These attacks are performed without an attacker needing to modify Web site files or binaries. Injected data is stored within a Web application and executed on the computers of unsuspecting victims. The typical method used to load and execute malicious code stored in a vulnerable Web site is in the form of client-side scripts. XSS attacks generally fall into one of two categories: reflected and stored.

Reflected XSS attacks

Reflected XSS attacks involve an attacker reflecting (echoing) code of a Web application to another user where it is downloaded and executed and then performs actions crafted by the attacker. Hackers usually combine reflected XSS attacks with another attack such as social engineering to trick a victim into

navigating to the site where the malicious code is echoed and executed by the user's Web browser. An example of a nonpersistent attack is as follows:

1. A malicious Web site user identifies an XSS vulnerability within a Web application.
2. The malicious user crafts a hyperlink inclusive of malicious code and sends it to a victim enticing them to click on the hyperlink.
3. When the victim clicks on the hyperlink, the Web page within the hyperlink is loaded and the malicious code developed by the hacker is input into a variable within a Web page that is downloaded by the victim and executed on their local machine.

In this example, once the code has finished executing, that is the end of the attack. If the attacker would like to launch another attack against the victim, they would need to craft another malicious hyperlink and entice the victim to open it as well.

Stored XSS attacks

Stored XSS attacks occur when the data supplied by a user is stored on the server by the Web application. This may be in the Registry, file system, or database. This data is later retrieved by the Web application and downloaded by the Web site visitor and executed on their local machine. An example of this type of attack is as follows:

1. A malicious Web site user identifies an XSS vulnerability within the bulletin board feature of a Web application.
2. The hacker crafts a message inclusive of malicious code that is uploaded through the bulletin board application and stored within the bulletin board's database.
3. When any user views the bulletin board message, the Web application retrieves the message's text (and malicious code) from the database and it is then downloaded and executed automatically by the victims' Web browser without their knowledge.

In this example, the attack is stored and every user who visits the site and views the message page will have the malicious code executed on their machines even if the user reboots his or her machine in between visits.

Preventing XSS attacks

The following recommendations can be used to help prevent XSS attacks:

- **Ensure that all application data input is properly validated:** Data input can come in many forms, including form fields, HTTP headers, cookies, and application variables.
- **Encode user-supplied data:** Encoding is the process of converting data from one format to another. Encoding data input will not prevent the

reflection of malicious code but rather will change malicious code into a format that is nonexecutable to block the attack.
- **Don't click on unknown or malicious hyperlinks.**
- **Implement restrictive security zones:** Security zones will help limit the impact of hostile code executing on the local machine.

Buffer overflows

Buffer overflows are based on the way the C or C++ programming languages work. Many function calls do not check to ensure that the buffer will be big enough to hold the data copied to it. Programmers can use calls that do this check to prevent overflows.

A *buffer* is a holding area for data. When more information is put into the buffer than it is able to handle, a buffer overflow occurs. Overflows can be caused deliberately by hackers and then exploited to run malicious code. There are two types of overflows: *stack* and *heap*.

- The stack is a location in memory where function calls are stored.
- The heap is a location in memory where dynamically allocated variables are stored.

Creating a buffer overflow attack requires that the hacker understand assembly language as well as technical details about the OS to be able to write the replacement code to the stack. However, the code for these attacks is often published so that others, who have less technical knowledge, can use it. Buffer overflows constitute one of the top flaws for exploitation on the Internet today.

INPUT VALIDATION

Input validation is a core requirement of building a secure application. Information received from a user should be treated as untrusted and validated by a trusted application component prior to processing, regardless of whether or not the user providing the input has been successfully authenticated. Input validation is a single programming practice that if implemented properly can result in code that is immune from many types of attacks, including the following:

- Cross-site scripting
- Response splitting
- Buffer overflows
- Data injection
- Directory transversals
- Denial-of-Service

Preventing input validation–related attacks

Proper data input validation needs to be implemented whenever and wherever data is received by the application and performed prior to processing, regardless

of the method used to receive it. Data input validation should evaluate data at a minimum for the following criteria:

- ***Type:*** Verifies that data received is within the specified format. For example, if an integer is expected, the application should not process alpha values such as ABC.
- ***Length:*** Ensures that the length of received data does not fall outside of an expected number of characters.
- ***Format:*** Verifies that data is received within the specified format, for example, YY/DD/MM.
- ***Range:*** Ensures that the data falls within a specified range of values, for example, a value between 1 and 1000.

Data input that does not pass all validation checks should be rejected.

PACKET SNIFFERS AND INSTANT MESSAGING

Packet sniffers are tools that can capture packets of data off a network. Packet sniffers are effective as they rely on the cleartext communication of information inside of a network. Instant messaging (IM) is a real-time communication method that allows people to communicate with one another through text messages as though they were in an actual phone conversation.

Instant messaging

More and more people and businesses rely on communicating in real time, which has allowed *Instant Messaging* (*IM*) to grown by leaps and bounds. IM involves using clients such as ICQ, AOL Instant Messenger (AIM), Yahoo! Messenger, Google Talk, and Windows Live Messenger/Windows Messenger, as well as Lotus Sametime and Microsoft's Office Communicator. Generally, each of these IM clients ties into a service that transfers messages between other users with the same client software. IM technologies pose significant security risks. Each of the messenger programs has been exploited with DoS attacks as well as remote execution attacks. The following security issues that are related to using IM technology must be acknowledged:

- IM clients are constantly exploited via buffer overflow attacks.
- IP address exposure is prominent and provides a way that an attacker can isolate a user's home machine, crack into it, and then exploit it.
- IM technology includes a file transfer capability, with some providing the ability to share folders (containing groups of files) with other users. All kinds of worms and viruses can be downloaded (circumventing the firewall), which could cause huge problems on an internal network.
- Companies' Human Resources (HR) policies need to be addressed to track IM communication out of the box. Some IM clients like Microsoft's Office Communicator allow IM conversations to be stored in the Exchange environment and subjected to retention policies.

There are software products available to better control IM traffic and log and archive IM communications. Such products add to the security of *IM*.

Peer-to-peer

Peer-to-Peer (*P2P*) networks have become a mainstream application, with two of the largest P2P networks being BitTorrent and eMule. In P2P networks, each computer communicates with other systems, and in order for this to work properly, a firewall rule would be required to allow traffic to and from all addresses that existing and future clients may use. Aside from limitations on your inability to implement restrictive network-based Access Control Lists, P2P networks are associated with the following additional risks:

- **Used as a target ingress path for Trojans and viruses:** The port(s) used between P2P clients to share data can include viruses.
- **Used as an egress vector to transfer stolen data:** An attacker can access data and then transfer the stolen information off a client\network to a location under the attacker's control.
- **Information disclosure:** Some P2P clients, such as Kazaa and Gnutella, provide backdoor file system access to other peers on the P2P network.

SECURING P2P CLIENTS

To help limit the risks associated with P2P networks, you can implement the following safeguards:

- **Virus scan all files retrieved from P2P networks:** All files downloaded from a P2P network should be scanned using an up-to-date anti-virus client before execution.
- **Implement strict restrictions on which folders are shared to other P2P clients:** This will help ensure which file folders are being shared over the P2P network and prevent sensitive information from being stored in them.

SMTP open relays

Organized crime is a significant driving force behind *SPAM* (unsolicited electronic mail), with research from Cisco showing that SPAM accounted for 90% of e-mail messages sent over the Internet during 2008. Most SPAM is sent from open SMTP relays existing on corporate networks or home computers without the knowledge or consent of the owner. An SMTP open relay is a mail server that accepts e-mail from anyone on the Internet through it. Mail servers should restrict e-mail destined to or originating from known users only, but an open relay allows anyone to send mail through it.

When an attacker relays a mail message, they bounce it off an open mail relay that in turn forwards the message to an address of the attackers choosing. Aside from serving as the source of SPAM messages that could lead authorities back

to your organization during a cyber investigation, there are additional risks associated with sending SPAM via an open mail relay:

- **DoS conditions:** Although businesses often use large Internet pipes, a single mail relay can quickly clog this pipe, causing a DoS condition where legitimate business functions cease to operate owing to lack of network resources.
- **Damage to brand**: SPAM messages may include advertising, images, or in some cases viruses. If one of these viruses were to impact another organization, especially a client, then they could see the incidents as an indication to the company's internal security challenges and may affect future business.
- **Blacklisted on SPAM sites:** Blacklisting sites track computers on the Internet that have been reported to be a major source of originating SPAM. These blacklists are maintained by SPAM-fighting organizations and are used by several anti-SPAM products as the gospel to block e-mails originating from computers on these SPAM lists.

SECURING MAIL RELAYS

To prevent the relaying of mail, mail relays and mail servers need to be properly configured to prohibit relaying of mail. In cases where mail relaying is legitimately needed, restrictions should be placed on the mail application to restrict which systems can relay mail off it.

SUMMARY OF EXAM OBJECTIVES

In this chapter, we reviewed application security–related Security+ exam objectives. We looked at how the focus of cyber crime has shifted in past years and how threat modeling, a relatively new method of risk assessment, can help secure these complex applications. Internet usage is higher than ever and browser-based threats are prevalent in the industry. We looked at some of the most serious browser-based threats and how they can be secured both by the developers who write them as well as the users in which they are downloaded to and executed. Understanding the information, threats, and countermeasures in this chapter will not only help you prepare for the Security+ exam but also provide you the insight needed to tackle application security, one of the largest threats to information security today.

TOP FIVE TOUGHEST QUESTIONS

1. A user contacts you with concerns over cookies found on his hard disk. The user visited a banking site several months ago, and when filling out a form on the site, provided some personal information that was saved to a cookie. Even though this was months ago, when the user returned to the site, it displayed his name and other information on the Web

page. This led the user to check his computer and find that the cookie created months ago is still on the hard disk of his computer. What type of cookie is this?

A. Temporary
B. Session
C. Persistent
D. Tracking

2. Proper input validation should be inclusive of which of the following checks? (Select all that apply.)

A. Data type
B. Data length
C. IP address of data transmission
D. Name of the user submitting the data
E. Range of values

3. Your developer contacts you for guidance on how to secure the ActiveX controls he plans on using within his Web application. What advice would you provide him?

A. Be sure to follow secure coding practices and sign the control before publishing.
B. Only transfer the control over SSL sessions to and from the Web browser.
C. Write the ActiveX control within Java.
D. Perform a threat model on the ActiveX control.

4. You are tasked with creating a threat model for a new application your company is developing. Whom should you include in the threat modeling process?

A. A member of the corporate security team
B. Members of the security team and upper management
C. Members of the security team and middle management
D. Members of the security team and members from all teams responsible for the design and operation of the application

5. You perform a security assessment of your company's Web server and identify a cross-site scripting vulnerability. What recommendation can you provide to your company to correct the vulnerability? (Choose the best answer.)

A. Advise Web site users to ensure that cookies are only transferred over secure connections.
B. Implement a policy mandating that Web site users disable ActiveX support within their Web browsers.

C. Implement a policy mandating that Web site users disable Java applet support within their Web browsers.
D. Advise the Web administrator to ensure that all Web application data inputs are validated prior to processing.

ANSWERS

1. The correct answer is C. Persistent. Persistent cookies are created to store for a long-term basis, so the person doesn't have to login each time they visit, or to save other settings like the language you want content to be displayed in, your first and last name, or other information. Answers A and B are incorrect because Temporary and Session cookies are created on a temporary basis and removed from the computer when the Web browser is shut down. Answer D is incorrect because the user filled out a form on a banking site and it is retrieving this information months later to display on a Web page when the user returns to the site. This is the behavior of a persistent cookie. Tracking cookies are different, because they are used to retain information on sites visited by a user.

2. Answers A, B, and E are correct and all pertain to ensuring that the entered data is in a form that the application is expecting. Answers C and D are related to authentication as opposed to data validation.

3. Answer A is correct. Following securing coding practices will help prevent the existence of vulnerabilities within the code. Signing the control will allow the developer to ensure that the control has not been tampered with after development and publication. Answer B would ensure that the control is transferred to a Web user over a secure channel but would not secure the ActiveX control created by the developer. Answer C is incorrect as the language used to create an ActiveX control is not the largest security concern. Answer D is incorrect as performing a threat model will help identify vulnerabilities within the ActiveX control and countermeasures to be applied; however, a separate action of applying the threat model results are required to effectively secure the control.

4. Answer D is correct. To create an accurate threat model, you will need participation from all teams responsible for the design and operation of the targeted application. This representation will help ensure that vulnerabilities from all operational aspects are identified. Answers A, B, and C are incorrect because they do not cover participation from all application design and operational teams.

5. Answer D is correct as the best way to address cross-site scripting vulnerabilities is to validate data input. This would fix occurrences of XSS on

ActiveX controls and Java applets downloaded to the client as well as any vulnerability located on server-side code within the application.

Incorrect answers and explanations: Answer A is incorrect because disabling cookies is not a countermeasure against XSS. Answers B and C are also incorrect, as although XSS vulnerabilities may exist within downloaded Java applets or ActiveX controls, these controls are executed on the client and would not address the server-side XSS vulnerability.

CHAPTER 4

Virtualization Technologies

Exam objectives in this chapter:

- The Purpose of Virtualization
- Benefits of Virtualization
- System Virtualization
- Application Virtualization

THE PURPOSE OF VIRTUALIZATION

Virtualization is the ability to allow one physical computer to run multiple instances of an operating system or multiple operating systems on the same physical computer. The basic concepts of virtualization are not new but come from the mainframe computing world, where they were originally designed to maximize the resource utilization of expensive hardware and software so businesses could get the best most efficient utilization of their mainframe processing capacity. This ability of the more modern servers also presents both security challenges and benefits. With more virtual machines there are more patches that need to be applied, more servers to be secured, virtual machines to be created and just as important removed, and users accessing both internal and external resources.

In addition to server virtualization there is application virtualization technology. *Virtual applications* run on servers located remote from the users. These users do not need to have the application or data loaded on their desktop devices. Application virtualization allows applications that may be sensitive or not compatible with a user's desktop to operate as if they were loaded locally. These virtual applications also do not leave a trace on the client machine, so they are safe to use from computers outside the trusted network.

BENEFITS OF VIRTUALIZATION

With the cost of servers remaining basically flat, their power and capabilities are ever increasing. This has created a situation where very little of the power

and performance of the physical computers is actually used in running the process or application that has been tasked on that server. It has been shown by several different studies that most modern servers are only running at 2–20% of their capacity. This is an inefficient use of the resources. Businesses want to get a better value for the money they spend on servers.

One of the key benefits of a virtual infrastructure is that all the virtual machines have standard virtual hardware regardless of the physical platform they are currently running on. This feature creates a utility computing environment where virtual machines simply work on whatever physical server the organization chooses. A hypervisor is simply a program that allows multiple operating systems to share a single physical host. Leveraging the advanced features of many hypervisors, an administrator can move running virtual machines to other physical servers without interruption to the users accessing the virtual server. The old physical server can be upgraded, repaired, or replaced, all without changing the virtual machine.

This utility computing feature allows for rapid recovery in case of disaster or security breach in that the virtual server configuration files and virtual disks can be copied or snapshots taken and transferred to a remote facility or separate storage and then used to restart the virtual machine in a different location without regard to drives or physical hardware differences. This feature allows for recovery of virtual machines in minutes instead of hours or days using traditional servers.

There are other side-expenses to consider when determining the value of virtualization, such as the cost of network ports, power connections, heating and cooling, space requirements, maintenance and upgrades, replacement and disposal of equipment, and the amount of manpower it takes to manage and maintain a physical infrastructure. For the organization, the benefits can be:

- Reduced cost of hardware
- Reduced space requirements
- Rapid deployment of new servers
- High availability
- Hosting multiple environments
- Separation of virtual
- Ability to maintain a Test/Development Environment in an easy fashion
- Lower costs for software testing

Types of virtualization

There are basically four types of virtualization, Hosted, Binary Translation, Paravirtualization, and Hardware Assist.

- ***Hosted:*** This type of virtualization uses a base operating system to run the physical computer and the hypervisor manages access to the physical resources through the operating system. The base operating system is normally Windows or Linux, but there are hosted virtualization versions for the Mac. VMware Virtual Server and Microsoft Virtual Server 2005 are examples of a server-based hypervisor using a hosted design.

- ***Binary Translation:*** This type of virtualization has a very thin operating system below the hypervisor. The hypervisor captures all system calls for hardware resources and translates the virtual calls to physical calls. By translating all system calls, each virtual machine is completely isolated from the underlying hardware. VMware ESX server and Microsoft's Server Core are examples of this type of hypervisor.
- ***Paravirtualization:*** This design of hypervisor allows some specific system call to be passed directly to the physical resources. The remaining system calls are still translated before passing to the physical resources. In a true paravirtualization hypervisor, small pieces of the guest operating system are modified to modify the dangerous kernel operations. These changes are picked up by the hypervisor and translated to the physical resources. Some less disruptive hardware calls are allowed to pass directly to the physical resources.
- ***Hardware Assist:*** This type of hypervisor leverages the benefits of the paravirtualization design and takes it a step further by adding specific CPU calls from the guest virtual machines. This allows for an even thinner hypervisor and increased performance of the virtual machines. Both Intel VT and AMD-V are examples of hardware assist in a paravirtualized hypervisor. Commercial versions of this type of hypervisor can be found in Citrix XenServer, Microsoft Hyper-V, and VMware ESX 3.5.

Designing a virtual environment

The differences in hypervisors have now been explained and some of the benefits of a virtual environment explored. These new tools, while very flexible and powerful, can also present challenges to the security team if the environment is not well designed and manageable.

Virtual machines are isolated both from the physical host computer and each other for the most part. It is important to remember that most of the physical resources are shared even though there is a separation between the virtual machines. You should take advantage of the physical capabilities of the hypervisor and add additional NICs, separate your storage, and use the snapshot and backup features of the hypervisor. If you properly allocate your physical resources, you can create a robust and secure environment for your virtual infrastructure.

The virtual infrastructure is very similar to a physical infrastructure in what can be done. It is possible to connect virtual machines to internal switches, physical NIC bonds or teams, VLANs, and internal and external storage. These features allow you to design and connect the different virtual machines to the necessary resources and still maintain your security design.

PROCESSORS

Most modern processors are now multi-core and have the hardware assist features for virtualization. They are mostly all x64-bit technology and will support

both 64-bit and 32-bit guests. If you do run across an older processor, you may need to use a hosted hypervisor for your virtual machines. You would then be limited by the restrictions of the host operating system. There are hosted hypervisors for both Windows and Linux.

Multi-core processors are like adding additional physical processors to your server. They appear as either a two- or four-processor system. Some hypervisors do require at least two physical processors. Once loaded, they will utilize each core as a separate processor.

There are some limitations you should be aware of when considering the processor selection for your virtual environment.

- **Total number of processor cores:** Most hypervisors will only support up to 32 processor cores. Current versions will allow up to 256 or more, but you should check the limitations of the hypervisor you select. It sounds like a lot of capacity until you do the math. A standard dual-core dual processor server is four processor cores. If the same server has quad-core processors, we have eight cores. If we have a four processor server with quad core processors, we have 16 cores. Given that eight core processors are coming out soon, it will not be difficult to hit the limits as it once was.
- **Pick a processor family:** Both Intel and AMD make multi-core processors with hardware assist virtualization built in. Everyone has their own particular favorite and each will change position based on their latest release, so we won't get into the debate on which is better. Just pick the one you prefer and stick with that processor family. Some hypervisors allow motion of running virtual machines between physical servers, but they must have processors from the same family. That means if you select one vendor as your processor of choice you should stick with that family of processors. It is possible to move virtual machines between processor families, but it normally requires you to shut down the virtual machine.

NETWORKING

It is normally possible to support between 15 and 32 physical NIC cards on a host server. Each virtual machine can have four or more virtual NIC cards. These NICs can be connected to internal switches or external port groups. As can be clearly seen there is plenty of flexibility for the virtual infrastructure.

DID YOU KNOW?

You can *use VLANs for dividing you virtual machines* and setting up different network connections. You can also use an internal switch to connect virtual machines without connecting them to an external network. Take the time and draw the connection maps to make sure you are meeting all the requirements of the questions.

Because the network cards are shared among virtual machines, it is recommended that gigabit NIC cards be used where possible.

STORAGE

Storage is where the virtual machines are kept along with their data, which can include the server's local disk drives or Direct Attached Storage Devices (DASD), a Storage Area Network (SAN) or a Network Attached Storage (NAS), or a combination of each for this purpose. Remember that virtual machines are really big files that must be managed by the physical server. While this is not normally a problem for the hypervisor, the type of storage chosen can make a big difference in the performance and availability of the virtual machines.

How storage is used can also increase the security of the virtual infrastructure. Authentication protocols and encryption can be applied to the storage used to increase the security and control access to the shared storage.

The different storage types each have benefits, and the basic advantages are listed below:

- ***Direct Attached Storage Devices:*** This type is the most common and familiar. This is the local hard drives in the physical server. These may be connected to a Redundant Array of Independent Disks (RAID) controller or just connected to the internal disk controller.
- ***Storage Area Network:*** The SAN is a stand-alone device that can share the storage among multiple physical servers. These connections are typically made using either Fiber Channel (FC) or iSCSI connections. Notice the SAN switch in between the servers and the storage. This is the component that allows for multiple physical servers to connect to the storage. This switch must match the protocol of the SAN.

EXAM WARNING

You should know the different types of storage and any security associated with them. Fiber Channel has the least storage because of the design and protocol. The FC-SP is being adopted, but because this is a fiber network that only moves disk IO traffic and is typically local only to the data center, there is less opportunity for compromise of the data. Both iSCSI and NAS storage use Ethernet and TCP/IP for communications and therefore are more vulnerable to compromise. The iSCSI uses CHAP authentication, and NAS devices use either NFS or NTFS and rely on a user name and password for access.

- ***Fiber Channel SAN:*** These connections can transfer data between 1 Gb and 8 Gb per second. They use a special interface card called a Host Bus Adapter (HBA) and are typically connected using a fiber optic cable. Fiber Channel Security Protocol (FC-SP) is designed to secure the transfer of data across the network between the storage and the server. It does not address the data stored on the SAN.
- ***iSCSI SAN:*** These connections transfer data at 1 Gb per second using normal Ethernet protocols. This disk traffic should be isolated on a

separate VLAN to improve performance and security. The iSCSI protocol can take advantage of Jumbo Frames on an Ethernet network. This feature must be supported by the network switch before it can be used. It is also recommended that an iSCSI HBA be used instead of a normal server NIC. Because of the ability to transfer data over a normal network, security is built into the protocols. Challenge Handshake Authentication Protocol (CHAP) is a protocol that is used to authenticate the connection and is based upon sharing a security key that is similar to a password.

- ***Network Attached Storage:*** This type of storage is similar to the SAN except it uses normal server NICs and a protocol called Network File System (NFS). This type of shared storage was originally developed for sharing files to individual computers by allowing the storage to be mapped to the local system as a local disk drive. The transfer of data is limited to the speed of the network.

As you can see it is very easy to start using multiple NICs in a virtualized infrastructure. Planning the implementation and leveraging the features of the hypervisor will help you maintain the security policies while still providing a robust and flexible virtual environment.

SYSTEM VIRTUALIZATION

Every hypervisor has tools for performing the virtualization of both existing systems as well as new systems. However, while these functions all look a bit different, the functionality is common across all of the major hypervisors. Each hypervisor may even have its own file format for the virtual systems. Some will read the different virtual file systems, whereas others may even use other format directly in some instances.

When a virtual machine is created there are at least two files created, a configuration file and a virtual hard drive. The format of the configuration file may vary from one hypervisor to the next, but it contains similar information—the location of the virtual hard drive, the name of the virtual machine, the amount of memory allocated to the virtual machine, the number of virtual NICs, and any other virtual hardware or connections for this specific virtual machine.

The virtual disk file is where the operating system and data files are stored for the virtual machines. Depending on the specific features of the hypervisor, this file may be created all at once, a 20 GB file, or it may be allocated for the specified size and created in 2 GB chunks. This makes the virtual hard disk much faster to create and does not use any space that it really doesn't need. This saves storage space but still makes the operating system believe it has full access to the allocated storage.

Creating virtual machines can be done in one of two ways: physical to virtual conversion and creating a new system.

- *Physical to virtual conversion:* This method is best if moving from a physical infrastructure to a virtual infrastructure. The method of conversion can be done on or off line, depending on the utility used.
- *Creating new virtual servers:* This is the same process as loading a new physical server. The difference is that since it is a virtual machine it can be done faster and in many instances a clone or template can be used of an existing virtual server.

Management of virtual servers

Most hypervisors have a management console to control the virtual environment. This is typically loaded on a separate server beside the physical host. There is also a connection client for some that will allow connections to specific virtual machines for management of that VM. This feature allows the security administrator to control administrative access to only the necessary level of access and specify the virtual machines a user can access or the administrative tasks that can be performed.

Most hypervisors provide for the use of an ISO file store. This is used to copy the installation media of operating systems and applications that are used for virtual machines. Each virtual machine can mount either the physical DVD of the host or an ISO image to the virtual DVD drive.

APPLICATION VIRTUALIZATION

Another method of virtualization is to virtualize applications. This technology allows users to be presented with either a desktop or a list of available applications for performing their tasks. The applications actually run on a server that may be located either on the local area network or across wide area links. Because the applications are executed on a server and only the display, keystrokes, and mouse movements are presented across the network, performance is seen as if the application is being executed locally. The connection can be configured to support full encryption from end to end on the connections. There are two common types of application virtualization in the Windows environment: Microsoft Terminal Services (now called Remote Desktop Services) and Citrix XenApp. There are also versions of XenApp that will run on UNIX with support for X Windows clients.

Crunch Time

Application virtualization is the easiest method of deploying and updating applications to users. You just need to update the applications on the terminal servers and all the users will receive the new or updated version the next time they log on.

Application streaming

Application streaming uses a sequencer that monitors how an application starts up. It records the different program modules as they load and monitors when the application becomes functional. The application is sequenced so the most functional blocks are loaded first and the application can be used. As the user requests additional features, those blocks are streamed to the application.

When the application is sequenced, the streaming file is stored on a file server. The application is presented to the user either through XenApp, a RemoteApp, or even from an Active Directory installation file. When the user clicks on the application, the file server is contacted and the application is streamed to the user's computer or session for processing. When the user closes the application, all files are removed from the computer executing the application.

SUMMARY OF EXAM OBJECTIVES

Virtual technologies allow for the rapid deployment of servers of different operating systems on common physical hardware. These virtual machines are fully functional and present themselves as physical servers to users and the outside world. This technology can be leveraged in order to deploy security servers alongside normal production servers. In addition to server virtualization, applications can also be virtualized so that they may be hosted on a remote server and accessed securely from outside the network and from untrusted clients by using a gateway device or software. Another form of application virtualization is application streaming, where only the necessary application modules for the user to begin working are sent to the client while waiting for a request for the remaining modules.

TOP FIVE TOUGHEST QUESTIONS

1. You are the security administrator for Versa Corp. You have been assigned the task of creating a "honey pot" server on the company's Internet DMZ. You have decided to use virtualization and a virtual machine for this purpose. One of the best reasons for using a virtual machine is:

   A. Virtual machines run Windows only and cannot have security templates applied to them.
   B. Virtual machines can be rapidly restored when breached.
   C. Virtual machines cannot join the production Active Directory.
   D. Virtual machines are not vulnerable to viruses.

2. You are the security administrator for Versa Corp. You have three virtual machines running on HP DL380 servers. There are IBM x3350 servers also running the same hypervisor and processor family with available resources. You have moved your virtual machines to the IBM servers.

What should you do to configure your virtual machines to run on the IBM servers?

A. Replace the network and RAID controller drivers on all the virtual machines immediately after powering them up.
B. Replace only the RAID controller drivers.
C. Replace only the network drivers.
D. Nothing.

3. You are the security administrator for Versa Corp. You have been tasked with designing a single server solution for the remote branch offices. You must have in your solution a Linux based firewall, a mail server in a DMZ, a domain controller, and a file server.

The system administrator has created a virtual host to run the necessary virtual machines and has asked you how you want the NICs connected to the virtual servers. You have provided him/her the diagram shown in Figure 4.1:

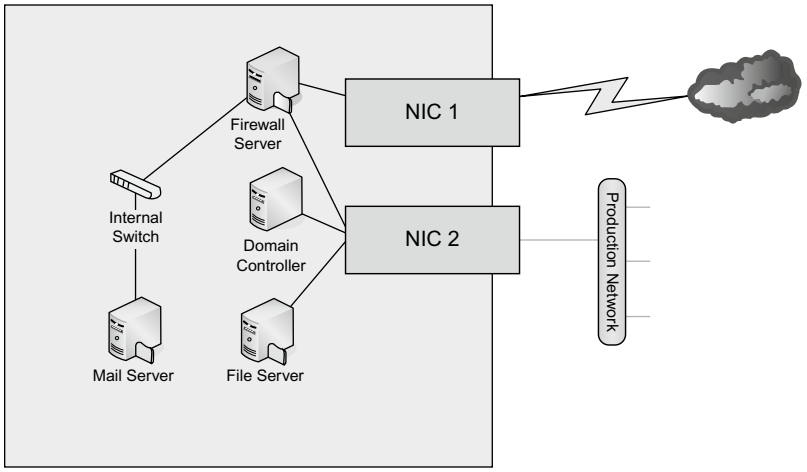

FIGURE 4.1
Virtual machine configuration.

This solution:

A. Fails to meet the requirements specified.
B. Exposes the virtual servers to the open Internet.
C. Prevents users from receiving their e-mail.
D. Meets all requirements.

4. You are the security administrator for Versa Corp. You have been asked to virtualize 10 security servers without altering their configurations. Your manager wants to retain the physical servers just in case there is

a problem later. What is your best course of action to accomplish the assigned tasks?

A. Build new virtual machines on the physical host to match the security servers. Once loaded, you copy the data files from each of the original server to the virtual servers. You leave the original servers online until the new servers are verified as working.
B. You copy the disk drives of the original servers to the SAN. Once completed, you create new virtual machines and attach the data on the SAN to the virtual machine. You shut down the original servers.
C. You use a physical to virtual migration tool to copy the disk drives of the physical servers to the new virtual machines. Once completed, you shut down the original server and power on the new virtual server.
D. You create a new virtual machine and use a bulk copy utility to copy all the data from the source servers to the new virtual machines. When complete, you leave the original servers online until the new servers are verified.

5. What is a benefit of application virtualization?

A. Applications are executed on the local clients instead of the application server.
B. Applications are all Web based.
C. Only Windows clients can access the published applications.
D. Any device that can run the client can access the applications.

ANSWERS

1. The correct answer is B. Virtual machines can be rapidly restored if they are breached. By using the snapshot features or using a template, a new virtual machine can quickly be restored at a known state without the breach that was detected. Answer A is incorrect because virtual machines are just like physical machines and can have security templates applied to them the same as any physical server. Answer C is incorrect because a virtual machine can join an Active Directory and can even be a domain controller. Answer D is incorrect because just like a physical server virus attacks are capable of compromising the server.

2. The correct answer is D. Virtual machines have the same hardware drivers when running on the same hypervisor. The hypervisor isolates the virtual machines from differences in the physical hardware. Answers A, B, and C are incorrect because the virtual machines never see the physical network or RAID controllers. The hypervisor and the virtual drivers ensure that all the virtual machines see the same device drivers, regardless of the physical resources.

3. The correct answer is D. By using both internal network connections and managing the external physical NICs, all requirements have been met and a secure environment can be deployed to the remote office. Answer A is incorrect because all requirements have been met. Answer B is incorrect because only the firewall is connected to NIC 1, which is connected to the Internet. There is no direct path to the Internet from any other server. Answer C is incorrect because the mail server is connected to the internal switch with the firewall. The firewall has its DMZ ports connected to this internal switch. Users should be able to access their e-mail through NIC 2 to the firewall then be routed to the internal switch connected to the e-mail server.

4. The correct answer is C. Using a physical to virtual migration tool copies all the configuration and settings to the virtual hard disk without altering the configuration of the source. Once complete, the original can be powered down and the new virtual server powered on. Answer A is incorrect because just coping the data files will not get all the configuration information from the original server. Even if this did work, you would have two servers with the same name and IP address online at the same time. This would cause a conflict. Answer B is incorrect because merely copying the files to the SAN will not make them usable on the virtual machines. The data must be copied to the virtual hard disk, not just a folder on the SAN. Answer D is incorrect because just copying the file to the virtual hard disk may not allow them to be used by the new virtual machines. You also left the original servers online, so both a name and IP conflict would exist if the copy did work.

5. The correct answer is D. Any device that can run the client software can access and run the published application. This is especially true of XenApp. There are clients for a wide variety of operating systems and processors. Answer A is incorrect because the application is executed on the server not the client device. Answer B is incorrect because applications do not need to be Web based to function on a terminal server. Answer C is incorrect because Windows is only one of several operating systems that can have a terminal services or XenApp client loaded.

CHAPTER 5

Network Security

Exam objectives in this chapter:

- General Network Security
- Network Ports, Services, and Threats
- Network Design Elements and Components
- Network Security Tools

GENERAL NETWORK SECURITY

In today's network infrastructures, it is critical to know the fundamentals of basic security infrastructure. Before any computer is connected to the Internet, planning must occur to make sure the network is designed in a secure manner. Many of the attacks that hackers use are successful because of an insecure network design. That is why it is so important for a security professional to use secure topologies and tools like intrusion detection and prevention. By understanding each of these items, you will see how they can be used to build a layered defense against attack.

Network services and risks associated with them

All networks contain services that provide some type of functionality. Some of the services are essential to the health of the network, or required for user functionality, but others can be disabled or removed since they are superfluous. When services that are not actively being used exist on networks, the chances of exploitation are increased. Simply by having a service enabled offers additional opportunity for hackers to attempt entrance into the infrastructure. It is important to evaluate the current needs and conditions of the network and infrastructure, and then begin to eliminate unnecessary services. This leads to a cleaner network structure, which then becomes less vulnerable to attack.

Network design elements

Not all networks are created the same; thus, not all networks should be physically laid out in the same fashion. The judicious usage of differing security topologies in a network can offer enhanced protection and performance.

Network security tools

Many tools exist today that can help you better manage and secure your network environment: specifically, intrusion detection and protection, firewalls, honeypots, content filters, and protocol analyzers. These tools will monitor, detect, and help contain malicious activity in an environment.

NETWORK PORTS, SERVICES, AND THREATS

In order to properly protect a network it is important to first identify the existing vulnerabilities, network ports, services, and potential threats. Knowing what exists in a network is the best first defense. Monitoring required services and removing all others reduces the opportunity for an attack and begins to make the environment more predictable.

Network ports and protocols

As discussed earlier in Chapter 2, "OS Hardening", unnecessary network ports and protocols in an environment should be eliminated whenever possible. Many, if not nearly all, internal networks today utilize TCP/IP as the primary protocol. So for most that means eliminating the following protocols: Internetwork Packet Exchange (IPX), Sequenced Packet Exchange (SPX), and/or NetBIOS Extended User Interface (NetBEUI). It is also important to look at the specific operational protocols used in a network, such as Internet Control Messaging Protocol (ICMP), Internet Group Management Protocol (IGMP), Service Advertising Protocol (SAP), and the Network Basic Input/Output System (NetBIOS) functionality associated with Server Message Block (SMB) transmissions in Windows-based systems.

The question as to which ports should be open is a matter of policy and risk assessment. Even for ports that are allowed and have been identified by scanning tools, decisions must be made as to which of these ports are likely to be vulnerable, and then the risks of the vulnerability weighed against the need for the particular service connected to that port. Port vulnerabilities are constantly updated by various vendors, and should be reviewed and evaluated for risk at regular intervals to reduce potential problems. It is important to remember that scans of a network should be conducted initially to develop a baseline of what services and protocols are active on the network. Once the network has been secured according to policy, these scans should be conducted on a periodic basis in order to ensure that the network is in compliance with the policy.

Network threats

Network threats exist in today's world in many forms. One of the more exciting and dynamic aspects of network security relates to the threat of attacks. A great deal of media attention and many vendor product offerings have addressed the topics of attacks and attack methodologies. While there are many different

varieties and methods of attack, they can generally all be grouped into several categories:

- General target of the attack (application, network, or mixed)
- Whether the attack is an active or passive one
- How the attack works (e.g., via password cracking, or by exploiting code and cryptographic algorithms)

TCP/IP HIJACKING

TCP/IP hijacking, or *session hijacking*, is a problem that has appeared in most TCP/IP-based applications, ranging from simple Telnet sessions to Web-based e-commerce applications. In order to hijack a TCP/IP connection, a malicious user must first have the ability to intercept a legitimate user's data, and then insert themselves into that session much like a man-in-the-middle (MITM) attack.

Two other more interesting and malicious forms of session hijacking involve Web-based applications (especially e-commerce and other applications that rely heavily on cookies to maintain session state).

- The first scenario involves hijacking a user's cookie, which is normally used to store login credentials and other sensitive information, and using that cookie to then access that user's session. The legitimate user will simply receive a "session expired" or "login failed" message and probably will not even be aware that anything suspicious happened.
- The second scenario with Web server applications that can lead to session hijacking is incorrectly configured session timeouts. A Web application is typically configured to timeout a user's session after a set period of inactivity. If this timeout is too large, it leaves a window of opportunity for an attacker to potentially use a hijacked cookie or even predict a session ID number and hijack a user's session.

The use of encrypted sessions is key to preventing these types of attacks as with other TCP/IP-based attacks.

NULL SESSIONS

Null sessions are unauthenticated connections. When someone attempts to connect to a Windows machine and does not present credentials, they can potentially successfully connect as an anonymous user, thus creating a null session. Null sessions present vulnerabilities in that once someone has connected to a machine a lot of information can be learned about the machine. The more that is exposed about the machine, the more ammunition a hacker will have to attempt to gain further access.

IP SPOOFING

Spoofing is a result of some inherent flaws in TCP/IP. TCP/IP basically assumes that all computers are telling the truth, and there is little or no checking done

to verify that a packet really comes from the address indicated in the IP header. When the TCP/IP protocols were being designed in the late 1960s, engineers didn't anticipate that anyone would or could use the protocol maliciously. In fact, one engineer at the time described the system as flawless because "computers don't lie." There are different types of IP spoofing attacks. These include *blind spoofing attacks*, in which the attacker can only send packets and has to make assumptions or guesses about replies, and *informed attacks*, in which the attacker can monitor, and therefore participate in, bidirectional communications.

Spoofing is not always malicious. Some network redundancy schemes rely on automated spoofing in order to take over the identity of a downed server. This is due to the fact that the networking technologies never accounted for the need for one server to take over for another.

Technologies and methodologies exist that can help safeguard against spoofing of these capability challenges. These include:

- Using firewalls to guard against unauthorized transmissions
- Not relying on the expectation that using undocumented protocols will provide protection (i.e., on security through obscurity)
- Using cryptographic algorithms to provide differing levels of authentication

MAN-IN-THE-MIDDLE ATTACKS

One issue that has resulted from IPv4's lack of security is the *MITM* attack. A TCP/IP connection is formed with a three-way handshake. A host (Client) that wants to send data to another host (Server) will initiate communications by sending a SYN packet. The SYN packet contains, among other things, the source and destination IP address as well as the source and destination port numbers. The server will respond with a SYN/ACK. The SYN from the server prompts the client to send another ACK and the connection is established.

If a malicious individual can place himself/herself between the client and the server, he/she can then monitor the packets moving between the two hosts. It is then possible for the malicious individual to analyze and change packets coming and going to the host. The key to this attack is being able to predict the right TCP sequence number and properly modify the data for this type of attack to actually work—all before the session times out waiting for the response.

REPLAY ATTACKS

In a *replay attack* a malicious person captures an amount of sensitive traffic and then simply replays it back to the host in an attempt to replicate the transaction. Another potential scenario for a replay attack is this: An attacker replays the captured data with all potential sequence numbers, in hopes of getting lucky and hitting the right one, thus causing the user's connection to drop, or in some cases, to insert arbitrary data into a session.

DENIAL OF SERVICE

Even with the most comprehensive filtering in place all firewalls are still vulnerable to **denial of service (DoS)** attacks. These attacks attempt to render a network inaccessible by flooding a device such as a firewall with packets to the point that it can no longer accept valid packets. This works by overloading the processor of the firewall by forcing it to attempt to process a number of packets far past its limitations.

Distributed denial of service

An alternative attack that is more difficult to defend against is the *distributed denial of service* (*DDoS*) attack. This attack is worse, because it can come from a large number of computers at the same time. This is accomplished either by the attacker having a large distributed network of systems all over the world (unlikely) or by infecting normal users' computers with a Trojan horse application, which allows the attacker to force the systems to attack specific targets without the end-user's knowledge. These end-user computers are systems that have been attacked in the past and infected with a Trojan horse by the attacker. By doing this, the attacker is able to set up a large number of systems (called zombies) to perform a DoS attack at the same time. This type of attack constitutes a DDoS attack. Performing an attack in this manner is more effective due to the number of packets being sent. In addition, it introduces another layer of systems between the attacker and the target, making the attacker more difficult to trace.

DOMAIN NAME KITING

Domain Name Kiting is when someone purchases a domain name, then soon after deletes the registration only to immediately reregister it. Since there is normally a five-day registration grace period offered by many domain name registrars, domain kiters will abuse this grace period by canceling the domain name registrations in order to avoid paying for them. This way they can use the domain names without cost.

DOMAIN NAME SERVICE POISONING

Domain Name Service (*DNS*) *poisoning*, or *DNS cache poisoning*, occurs when a server is fed with altered or spoofed records that are then retained in the DNS server cache. Once the DNS cache on a server has been "poisoned" in this fashion, since servers use their cache as the first mechanism to respond to incoming requests, all additional queries for the same record will be responded to with the falsified information. Attackers can use this method to redirect valid requests to malicious sites.

ADDRESS RESOLUTION PROTOCOL POISONING

Address Resolution Protocol (*ARP*) *poisoning* occurs when a client machine sends out an ARP request for another machine's Media Access Control (MAC) address information and is sent falsified information instead. This is done on a local

area network (LAN) where the requesting machine knows the target's IP address but needs to associate that IP address with a MAC address. A *MAC address* is a unique identifier or address that is assigned to most network adapters or network interface cards and, technically, each MAC address on a given LAN should be unique. The spoofed ARP message allows the attacker to associate a MAC address of their choosing to a particular IP address, which means any traffic meant for that IP address would be mistakenly sent to the attacker instead. This opens the door for numerous attack mechanisms to be employed including:

- IP spoofing attacks
- MITM attacks
- DoS attacks

NETWORK DESIGN ELEMENTS AND COMPONENTS

Understanding the components and elements used in network design and how they work together is a good first step to building an effective design. This section discusses following components of network design:

- Demilitarized zone (DMZ)
- Subnets
- VLANs
- Network Access Translation
- Network Access Control/Network Access Protection
- IP Telephony

While differing components can be effectively used together, in some instances they need to be used completely separately from each other. The different pieces that make up a network can be considered as discrete network segments holding systems that share common requirements. These are sometimes called security zones, and some of these common requirements can be:

- The types of information the zone handles
- Who uses the zone
- What levels of security the zone requires in order to protect its data

Firewalls

A firewall is the most common device used to protect an internal network from outside intruders. When properly configured, a *firewall* blocks access to an internal network from the outside, and blocks users of the internal network from accessing potentially dangerous external networks or ports.

There are three firewall technologies examined in the Security+ exam.

- ***Packet filtering*** works at the network layer of the Open Systems Interconnect (OSI) model and is designed to operate rapidly by either allowing or denying packets.

- ***Application-layer gateways*** operate at the application layer of the OSI model, analyzing each packet and verifying that it contains the correct type of data for the specific application it is attempting to communicate with.
- ***Stateful inspection*** checks each packet to verify that it is an expected response to a current communications session.

In the early days of business Internet connectivity, the concept of security zones was developed to separate the systems available to the public Internet from private systems available for internal use by an organization. A device called a firewall was utilized to separate the zones.

Many of these early firewalls had only basic abilities and usually functioned only as a packet filter. Packet filters rely on *Access Control Lists* (*ACLs*). ACLs allow the packet filter to be configured to block or allow traffic based on attributes such as IP address and source and destination port. Most firewalls provide the ability to:

- Block traffic based on certain rules.
- Mask the presence of networks or hosts to the outside world.
- Log and maintain audit trails of incoming and outgoing traffic.
- Provide additional authentication methods.

Some newer firewalls include more advanced features, such as integrated *virtual private networking* (*VPN*) applications that allow remote users to access local systems through a secure, encrypted tunnel, and some firewalls have integrated Intrusion Detection Systems in their product and can make firewall rule changes based on the detection of suspicious events happening at the network gateway.

It is the network security professional's job to make sure that all traffic on the intranet will be secure and safe from the prying eyes on the Internet. While a security breach of a DMZ system can be costly to a company, a breach that occurs inside an intranet could be *extraordinarily* costly and damaging. If this happens, customers and business partners might lose faith in the company's ability to safeguard sensitive information, and other attackers will likely make the network a favorite target for future attacks.

To ensure that all traffic on the intranet is secure, the following issues should be addressed:

- Make sure that the firewall is configured properly to stop attack attempts at the firewall.
- Make sure that the firewall is configured properly to prevent unauthorized network traffic from the internal network reaching the Internet.
- Make sure the firewall will watch traffic that egresses or leaves the network from trusted hosts, and ensure that it is not intercepted and altered en route; steps should also be taken to try to eliminate spoofing from attackers.

- Make sure that the anti-virus software is in use and up to date.
- Educate users on the necessity of logging out of their computers when not working on them.
- If possible, implement Internet Protocol Security (IPSec) on the intranet between all clients and servers to prevent eavesdropping.
- Conduct regular, but unannounced, security audits and inspections. Be sure to closely monitor all logs that are applicable.
- Do not allow the installation of modems or unsecured wireless access points on any intranet computers.
- Do not allow any connection to the Internet except through the firewall and proxy servers, as applicable.

Of course, there are literally hundreds of other issues that may need to be addressed, but these are some of the easiest ones to take care of and the most commonly exploited ones.

Extranets are a special implementation of the intranet topology. Creating an extranet allows for access to a network, or portions of the network, by trusted customers, partners, or other users. These users can then be allowed to access private information stored on the internal network that they would not want to place on the DMZ for general public access. The amount of access that each user or group of users is allowed to have to the intranet can be easily customized to ensure that each user or group gets what they need and nothing more. Additionally, some organizations create extranets to allow their own employees to have access to certain internal data while away from the private network.

Users attempting to gain access to an extranet require some form of authentication before they are allowed access to resources. The type of access control implemented can vary, but some of the more common include usernames/passwords, smart cards, tokens, and digital certificates. Once an extranet user has been successfully authenticated, they can gain access to the resources that are allowed for their access level.

EXAM WARNING

Be able to readily define an extranet. You must know the difference between the Internet, intranet, and extranet.

APPLICATION-LAYER FIREWALLS

The second firewall technology is called *application filtering*, or an *application-layer gateway*. This technology is more advanced than packet filtering, as it examines the entire packet and determines what should be done with the packet based on specific defined rules. One of the major benefits of application-layer gateway technology is its application-layer awareness. Since application-layer gateway technology can determine more information from a packet than

a simple packet filter can, application-layer gateway technology uses more complex rules to determine the validity of any given packet. These rules take advantage of the fact that application-layer gateways can determine whether data in a packet matches with what is expected for data going to a specific port. For example, the application-layer gateway can tell if packets containing controls for a Trojan horse application are being sent to the HTTP port (80) and, thus, can block them.

What is a DMZ?

In computer security, the *demilitarized zone* (*DMZ*) is a "neutral" network segment where systems accessible to the public Internet are housed, which offers some basic levels of protection against attacks. The term DMZ is derived from the military and is used to describe a "safe" or buffer area between two countries where, by mutual agreement, no troops or war-making activities are allowed. In the next sections we will explore this concept in more detail.

DMZ DESIGN

There are usually strict rules regarding what is allowed within a zone. When applying this term to the IT security realm, it can be used to create DMZ segments in usually one of two ways:

- Layered DMZ implementation
- Multiple interface firewall implementation

In the first method, the systems that require protection are placed between two firewall devices with different rule sets, which allow systems on the Internet to connect to the offered services on the DMZ systems, but prevent them from connecting to the computers on the internal segments of the organization's network (often called the protected network). The second method is to add a third interface to the firewall and place the DMZ systems on that network segment. In either case, the DMZ systems are offered some level of protection from the public Internet while they remain accessible for the specific services they provide to external users.

A good first step in building a strong defense is to harden the DMZ systems by removing all unnecessary services and unneeded components. The result is a *bastion host*. This scenario allows for public services while still maintaining a degree of protection against attack.

The role of the firewall in all of these scenarios is to manage the traffic between the network segments. If an Internet system attempts to connect to a service not made public, the firewall drops the traffic and logs the information about the attempt. Systems on a protected network are allowed to access the Internet as they require, and they may also access the DMZ systems for managing the computers, gathering data, or updating the content.

Creating and managing security controls such as firewall rules, IDS signatures, and user access regulations is a large task. These processes should be kept as simple as possible without compromising security or usability. It is best to start with *deny-all strategies* and permit only the services and network transactions required to make the site function, and then carefully manage the site's performance making small changes to the access controls to more easily manage the rule sets.

EXAM WARNING

The concept of a denial-all strategy will be covered on the Security+ exam. A denial-all strategy means that all services and ports are disabled by default, and then only the minimum level of service is activated as a valid business case is made for each service.

VLANs

A *VLAN* can be thought of as the equivalent to a broadcast domain. VLANs are a way to segment a network into logical LANs that use a basis other than a physical location to map the computers that belong to each separate VLAN. VLANs require the use of switches and switching technology.

DID YOU KNOW?

A broadcast domain consists of a group of nodes (computers) that receive layer 2 broadcasts sent by other members of the same group. Typically, broadcast domains are separated by creating additional network segments or by adding a router.

Do not confuse broadcast domains with collision domains. Collision domains refer specifically to Ethernet networks. The area of network cabling between layer 2 devices is known as a collision domain. Layer 2 devices typically include switches that rely on the physical address (MAC address) of computers to route traffic.

Each VLAN functions like a separate network due to the combination of hardware and software features built into the switch itself. Thus, the switch must be capable of supporting VLANs in order to use them. The following are typical characteristics of VLANs when implemented on a network:

- Each VLAN is the logical equivalent of a physically separate network as far as traffic is concerned.
- A VLAN can span multiple switches, limited only by imagination and the capabilities of the switches being used.
- Trunks carry the traffic between each switch that is part of a VLAN. A *trunk* is defined as a point-to-point link from one switch to another switch. The purpose of a trunk is to carry the traffic of multiple VLANs over a single link.
- Cisco switches, for example, use the Cisco proprietary Inter-Switch Link (ISL) and IEEE 802.1Q protocol as their trunking protocols.

EXAM WARNING

Know that VLANs implement security at the switch level. If you are not on the same VLAN as another user on your network and access is not allowed, you can secure communications from such hosts.

Network Address Translation

Network Address Translation (*NAT*) was developed because of the explosive growth of the Internet and the increase in home and business networks—the number of available IP addresses was simply not enough. A computer must have an IP address in order to communicate with other computers on the Internet. NAT allows a single device, such as a router, to act as an agent between the Internet and the local network. This device or router provides a pool of addresses to be used by your local network. Only a single, unique IP address is required to represent this entire group of computers. Common types of NAT include:

- ***Static NAT***—Used by businesses to connect Web servers to the Internet.
- ***Dynamic NAT***—Larger businesses use this type of NAT because it can operate with a pool of public addresses.
- ***Port Address Translation (PAT)***—Most home networks using Digital Subscriber Line (DSL) or cable modems use this type of NAT.

NAT has several benefits, one of which is its ability to hide the IP address and network design of the internal network. The ability to hide the internal network from the Internet reduces the risk of intruders gleaning information about the network and exploiting that information to gain access. NAT enables internal clients to use nonroutable IP addresses, such as the private IP addresses defined in RFC 1918, but still enables them to access Internet resources. The three ranges of IP addresses RFC 1918 reserved includes:

```
10.0.0.0-10.255.255.255 (10/8 prefix)
172.16.0.0-172.31.255.255 (172.16/12 prefix)
192.168.0.0-192.168.255.255 (192.168/16 prefix)
```

Network access control/network access protection

Another way to harden the network is to use *network access control* (*NAC*). As a brief aside, there's a bit of semantics that need to be dealt with. There are several different incarnations of NAC available, including:

- *Infrastructure-based NAC* requires an organization to be running the most current hardware and OSs. OSs, such as Microsoft Vista, have the ability to perform NAC.

- *Endpoint-based NAC* requires the installation of software agents on each network client. These devices are then managed by a centralized management console.
- *Hardware-based NAC* requires the installation of a network appliance. The appliance monitors for specific behavior and can limit device connectivity should noncompliant activity be detected.

NAC offers administrators a way to verify that devices meet certain health standards before they're allowed to connect to the network. Laptops, desktop computers, or any device that doesn't comply with predefined requirements can be prevented from joining the network or can even be relegated to a controlled network where access is restricted until the device is brought up to the required security standards.

Telephony

One area that is often overlooked in the IT security field is telecommunications. A company's business can be just as easily disrupted by having its telecommunications disabled as it can by having its computer network disabled. That makes this an important area to be aware of when developing an overall security plan.

A *private branch exchange* (*PBX*) is a device that handles routing of internal and external telephone lines. This allows a company to have a limited number of external lines and an unlimited number of internal lines. By limiting the number of external lines, a company is able to control the cost of telephone service while still providing for the communications needs of its employees.

PBXs are vulnerable to DoS attacks against their external phone lines. There is also the possibility of them being taken over remotely and used to make unauthorized phone calls via the company's outgoing lines. Voicemail capability can also be abused. Hackers who specialize in telephone systems, called *phreakers*, like to take control over voicemail boxes that use simple passwords, and change the passwords or the outgoing messages.

NETWORK SECURITY TOOLS

In this section we will discuss network security tools and how you can put them to work in your environment to help you to keep it safe and more resilient to malicious attack.

Intrusion detection and preventions systems

Firewalls and other simple boundary devices lack some degree of intelligence when it comes to observing, recognizing, and identifying attack signatures that may be present in the traffic they monitor and the log files they collect. A successful security strategy requires many layers and components. One of these components is the intrusion detection system (IDS).

PROXY SERVERS

A *proxy server* is a server that sits between an intranet and its Internet connection. Proxy servers provide features such as document caching (for faster browser retrieval) and access control. Proxy servers can provide security for a network by filtering and discarding requests that are deemed inappropriate by an administrator. Proxy servers also protect the internal network by masking all internal IP addresses—all connections to Internet servers appear to be coming from the IP address of the proxy servers.

Honeypots

A *honeypot* is a computer system that is deliberately exposed to public access—usually on the Internet—for the express purpose of attracting and distracting attackers. In other words, these are the technical equivalent of the familiar police "sting" operation. Although the strategy involved in luring hackers to spend time investigating attractive network devices or servers can cause its own problems, finding ways to lure intruders into a system or network improves the odds of being able to identify those intruders and pursue them more effectively.

The following characteristics are typical of honeypots:

- Systems or devices used as lures are set up with only "out of the box" default installations, so that they are deliberately made subject to all known vulnerabilities, exploits, and attacks.
- The systems or devices used as lures do not include sensitive information (e.g., passwords, data, applications, or services an organization depends on or must absolutely protect), so these lures can be compromised, or even destroyed, without causing damage, loss, or harm to the organization that presents them to be attacked.
- Systems or devices used as lures often also contain deliberately tantalizing objects or resources, such as files named *password.db*, folders named *Top Secret*, and so forth—often consisting only of encrypted garbage data or log files of no real significance or value—to attract and hold an attacker's interest long enough to give a backtrace a chance of identifying the attack's point of origin.
- Systems or devices used as lures also include or are monitored by passive applications that can detect and report on attacks or intrusions as soon as they start, so the process of backtracing and identification can begin as soon as possible.

HONEYNETS

A *honeynet* is a network that is set up for the same purpose as a honeypot: to attract potential attackers and distract them from your production network. In a honeynet, attackers will not only find vulnerable services or servers but also find vulnerable routers, firewalls, and other network boundary devices, security applications, and so forth.

The following characteristics are typical of honeynets:

- Network devices used as lures are set up with only "out of the box" default installations, so that they are deliberately made subject to all known vulnerabilities, exploits, and attacks.
- The devices used as lures do not include sensitive information (e.g., passwords, data, applications, or services an organization depends on or must absolutely protect), so these lures can be compromised, or even destroyed, without causing damage, loss, or harm to the organization that presents them to be attacked.
- Devices used as lures also include or are monitored by passive applications that can detect and report on attacks or intrusions as soon as they start, so the process of backtracing and identification can begin as soon as possible.

Content filters

Content filtering is the process used by various applications to examine content passing through and make a decision on the data based on a set of criteria. Actions are based on the analysis of the content and the result actions can result in the blocking or permitting of the traffic through the content filter.

Content filtering is commonly performed on e-mail, and is often also applied to Web page access as well.

Protocol analyzers

A *protocol analyzer* is used to examine network traffic as it travels along your Ethernet network. They are called by many names, such as packet analyzer, network analyzer, and sniffer, but all function in the same basic way. As traffic moves across the network from machine to machine, the protocol analyzer takes a capture of each packet. This capture is essentially a photocopy, and the original packet is not harmed or altered. Capturing the data allows a malicious hacker to obtain your data and potentially piece it back together in order to analyze the contents. A sniffer is typically a software installed on a machine that can then capture all the traffic on a designated network. Much of the traffic on the network will be destined for all machines, as in the case of broadcast traffic. These packets will be picked up and saved as part of the capture. Also, all traffic destined to and coming from the machine running the sniffer will be captured. In order to capture traffic addressed to/from another machine on the network the sniffer should be run in promiscuous mode.

SUMMARY OF EXAM OBJECTIVES

Network security covers a wide variety of topics, including the overall design and the types of tools that are used as well as the policies that are applied in the network. These tools include intrusion detection, firewalls, NAT, VLANs, and many others. It is critical for anyone taking the Security+ exam to be well versed in many of these areas—not just knowledge of their existence but an understanding of their operation and their impact to network security in general.

TOP FIVE TOUGHEST QUESTIONS

1. Your company is considering implementing a VLAN. As you have studied for you Security+ exam, you have learned that VLANs offer certain security benefits as they can segment network traffic. The organization would like to set up three separate VLANs in which there is one for management, one for manufacturing, and one for engineering. How would traffic move for the engineering to the management VLAN?

A. The traffic is passed directly as both VLAN's are part of the same collision domain.
B. The traffic is passed directly as both VLAN's are part of the same broadcast domain.
C. Traffic cannot move from the management to the engineering VLAN.
D. Traffic must be passed to the router and then back to the appropriate VLAN.

2. You have been asked to protect two Web servers from attack. You have also been tasked with making sure that the internal network is also secure. What type of design could be used to meet these goals while also protecting all of the organization?

A. Implement IPSec on Web servers to provide encryption.
B. Create a DMZ and place the Web server in it while placing the intranet behind the internal firewall.
C. Place a honeypot on the internal network.
D. Remove the Cat 5 cabling and replace it with fiber-optic cabling.

3. You have been asked to use an existing router and utilize it as a firewall. Management would like you to use it to perform address translation and block some known bad IP addresses that previous attacks have originated from. With this in mind, which of the following statements are accurate?

A. You have been asked to perform NAT services.
B. You have been asked to set up a proxy.
C. You have been asked to set up stateful inspection.
D. You have been asked to set up a packet filter.

4. You have installed an IDS that is being used to actively match incoming packets against known attacks. Which of the following technologies is being used?

A. Stateful inspection
B. Protocol analysis
C. Anomaly detection
D. Pattern matching

5. You must choose what type of IDS to recommend to your company. You need an IDS that can be used to look into packets to determine their composition. What signature type do you require?

   A. File-based
   B. Context-based
   C. Content-based
   D. Active

ANSWERS

1. Answer D is correct. Answers A and B are incorrect as the VLANs are not part of the same collision or broadcast domains, since good VLAN design requires unique IP address spaces for each VLAN. Answer C is incorrect as there is no reason why traffic would not be able to go from the management to the engineering VLAN so long as it didn't violate policy.

2. Answer B is correct. Answer A is incorrect as encryption does not provide security for the internal network. Answer C is incorrect as a honeypot is designed to detect an attack and to distract an attacker. Answer D is incorrect as the type of media does not affect the security of the internal network.

3. Answers A and D are correct. Answer B is incorrect as the router is not performing proxy services, and answer C is incorrect as the router is not doing stateful inspection but rather simple packet filtering.

4. Answer D is correct as the technology being used is to match patterns. Answer A is incorrect as stateful inspection is used by firewalls to ensure that connections that traverse the firewall are valid. Answer B only checks that the protocol used in the communication does not deviate from standards. Answer C is another IDS technology that uses behavioral analysis to identify an event that is outside the normal behavior found on the network.

5. Answer C is correct. Answer A is incorrect as there is no such thing as a file-based signature. Answer B is incorrect as the signature is looking at the contents of the packets and not how they are being used in context. Answer D is incorrect as there is no such thing as an "active" signature type.

CHAPTER 6

Wireless Networks

Exam objectives in this chapter:

- Wireless Network Design
- Service Set ID Broadcast
- Wireless Security Standards
- Rogue Access Points
- Data Emanation
- Bluetooth

WIRELESS NETWORK DESIGN

This section covers the basics of wireless network design and architectures. Before we delve too deeply into the design of wireless systems, it's a good idea to first review some wireless communication basics as wireless networks, like their wired counterparts, rely on the manipulation of an electrical charge to enable communication between devices. Changes or oscillations in signal strength from 0 to some maximum value (amplitude) and the rate of those oscillations (frequency) are used singularly or in combination with each other to encode and decode information.

Wireless communications

The primary difference between wired and wireless networks is that wireless networks use a special type of electric current known as *radio frequency* (*RF*), which is created by applying alternating current (AC) to an antenna to produce an *electromagnetic* (*EM*) ***field***. Devices for broadcasting and reception use the resulting RF field. In the case of wireless networks, the medium for communications is the *EM field*, the region of space that is influenced by electromagnetic radiation. (Unlike audio waves, radio waves do not require a medium such as air or water to propagate.) As with wired networks, amplitude decreases with distance, resulting in the degradation of signal strength and the ability to communicate. However, the EM field is also dispersed according to the properties

of the transmitting antenna, and not tightly bound as is the case with communication over a wire. The area over which the radio waves propagate from an electromagnetic source is known as the *fresnel zone*.

Like the waves created by throwing a rock into a pool of water, radio waves are affected by the presence of obstructions and can be reflected, refracted, diffracted, or scattered, depending on the properties of the obstruction and its interaction with the radio waves. Reflected radio waves can be a source of interference on wireless networks. The interference created by bounced radio waves is called *multipath interference*.

In planning for a wireless network, administrators should consider the presence of common sources of multipath interference. These include metal doors, metal roofs, water, metal vertical blinds, and any other source that is highly reflective to radio waves. Antennas may help to compensate for the effects of multipath interference, but must be carefully chosen. Many wireless access points (APs) have two antennas precisely for this purpose.

Another source of signal loss is the presence of obstacles. While radio waves can travel through physical objects, they are degraded according to the properties of the object they travel through. For example, a window is fairly transparent to radio waves, but may reduce the effective range of a wireless network by between 50% and 70%, depending on the presence and nature of the coatings on the glass. A solid core wall can reduce the effective range of a wireless network by up to 90% or greater.

To mitigate the effects of interference from these devices and other sources of electromagnetic interference, RF-based wireless networks employ *spread spectrum* technologies. *Spread spectrum* provides a way to "share" bandwidth with other devices that may be operating in the same frequency range. Rather than operating on a single, dedicated frequency such as is the case with radio and television broadcasts, wireless networks use a "spectrum" of frequencies for communication.

Spread spectrum technology

Conceived of by Hedy Lamarr and George Antheil in 1940 as a method of securing military communications from jamming and for eavesdropping during WWII, spread spectrum defines methods for wireless devices to use to send a number of narrowband frequencies over a range of frequencies simultaneously for communication. The narrowband frequencies used between devices change according to a random-appearing but defined pattern, allowing individual frequencies to contain parts of the transmission. Someone listening to a transmission using spread spectrum would hear only noise, unless their device understood in advance what frequencies were used for the transmission and could synchronize with them.

FREQUENCY HOPPING SPREAD SPECTRUM

As the name implies, *frequency hopping spread spectrum* (*FHSS*) works by quickly moving from one frequency to another according to a pseudorandom pattern.

The frequency range used by the frequency hop is relatively large (83.5 MHz), providing excellent protection from interference. The amount of time spent on any given frequency is known as *dwell time*, and the amount of time it takes to move from one frequency to another is known as *hop time*. FHSS devices begin their transmission on one frequency and move to other frequencies according to a predefined pseudorandom sequence and then repeat the sequence after reaching the final frequency in the pattern. The frequency hopping sequence creates a channel, allowing multiple channels to coexist in the same frequency range without interfering with each other.

DIRECT SEQUENCE SPREAD SPECTRUM

Direct sequence spread spectrum (*DSSS*) works somewhat differently. With DSSS, the data is divided and simultaneously transmitted on as many frequencies as possible within a particular frequency band (the channel). DSSS adds redundant bits of data known as *chips* to the data to represent binary 0s or 1s. The ratio of chips to data is known as the *spreading ratio*: The higher the ratio, the more immune to interference the signal is, because if part of the transmission is corrupted, the data can still be recovered from the remaining part of the chipping code. This method provides greater rates of transmission than FHSS, which uses a limited number of frequencies, but fewer channels in a given frequency range. And, DSSS also protects against data loss through the redundant, simultaneous transmission of data. However, because DSSS floods the channel it is using, it is also more vulnerable to interference from EM devices operating in the same range.

Wireless network architecture

The seven-layer open systems interconnect (OSI) networking model defines the framework for implementing network protocols. Wireless networks operate at the *physical* and *data link* layers of the OSI model. Both FHSS and DSSS are implemented at the physical layer. The data link layer is divided into two sublayers, the Media Access Control (MAC) and Logical Link Control (LLC) layers.

The MAC layer is responsible for such things as:

- Framing data
- Error control
- Synchronization
- Collision detection and avoidance

The Ethernet 802.3 standard, which defines the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) method for protecting against data loss as a result of data collisions on the cable, is defined at this layer.

CSMA/CD and CSMA/CA

In contrast to Ethernet 802.3 networks, wireless networks defined by the 802.11 standard do not use CSMA/CD as a method to protect against data loss resulting

from collisions. Instead, 802.11 networks use a method known as *Carrier Sense Multiple Access with Collision Avoidance* (*CSMA/CA*). CSMA/CD works by detecting whether a collision has occurred on the network and then retransmitting the data in the event of such an occurrence. However, this method is not practical for wireless networks because it relies on the fact that every workstation can hear all the other workstations on a cable segment to determine if there is a collision.

CSMA/CA solves the problem of potential collisions on the wireless network by taking a more active approach than CSMA/CD, which kicks in only after a collision has been detected. Using CSMA/CA, a wireless workstation first tries to detect if any other device is communicating on the network. If it senses it is clear to send, it initiates communication. The receiving device sends an acknowledgment (ACK) packet to the transmitting device indicating successful reception. If the transmitting device does not receive an ACK, it assumes a collision has occurred and retransmits the data. However, it should be noted that many collisions can occur and that these collisions can be used to compromise the confidentiality of Wired Equivalent Privacy (WEP) encrypted data.

SERVICE SET ID BROADCAST

The 802.11 standard provides for two modes for ad-hoc and infrastructure wireless clients to communicate. The ad-hoc mode is geared for a network of stations within communication range of each other. Ad-hoc networks are created spontaneously between the network participants. In infrastructure mode, APs provide more permanent structure for the network. An infrastructure consists of one or more APs as well as a distribution system (i.e., a wired network) behind the APs that tie the wireless network to the wired network.

802.11 traffic can be subdivided into three parts:

- ***Control frames*** include such information as Request to Send (RTS), Clear to Send (CTS), and ACK messages.
- ***Management frames*** include beacon frames, probe request/response, authentication frames, and association frames.
- ***Data frames*** 802.11 frames that carry data, which is typically considered network traffic, such as Internet Protocol (IP) encapsulated frames.

All this communication requires that systems have a means to distinguish different wireless networks from one another; the 802.11 standard defines the *Service Set Identifier* (*SSID*). The SSID is considered the identity element that "glues" various components of a wireless local area network (LAN) together. Traffic from wireless clients that use one SSID can be distinguished from other

wireless traffic using a different SSID. Using the SSID, an AP can determine which traffic is meant for it and which is meant for other wireless networks. Unless otherwise configured to block such activity, wireless networks will regularly broadcast their SSID. This is known as an SSID broadcast. While SSID broadcast can be disabled, the SSID is still needed to direct packets to and from the AP; this basically means that it is still discoverable to an attacker with the right tools. It is important to understand that hiding the SSID is not true security, but rather security through obscurity.

WIRELESS SECURITY STANDARDS

The IEEE 802.11 standard covers the communication between WLAN components. RF poses challenges to privacy in that it travels through and around physical objects. Because of the nature of the 802.11 wireless LANs, the IEEE working group initially implemented a mechanism to protect the privacy of the individual transmissions, known as the *WEP* protocol. WEP, however, was never intended to be the absolute authority in wireless security and quickly became an example of how *not* to design a cryptographic security protocol. The IEEE 802.11 standard stated that WEP should provide for the same amount of protection as a wired network. However, this provides only a basic level of privacy and it was quickly determined that WEP had some fatal flaws. The flaws in WEP were so severe that wireless vendors and other organizations quickly provided recommendations in how to improve the security of wireless networks. For networks that required higher degrees of security, other mechanisms should be utilized on top of the WEP encryption provided by wireless networks. The mechanisms include strong authentication, access control, password protection, and virtual private networks (VPNs). It is important to review the effect of key size on the overall security of WEP as an illustration of one of its major weaknesses.

The failure of WEP

WEP had many weaknesses associated with it. WEP required the use of a 24-bit Initialization Vector (IV) along with either a 40-bit or a 104-bit secret key that was set by the administrator. The secret key was concatenated onto the end of the IV, and the combination was then input into the RC4 stream cipher to generate the ciphertext for each packet. The combination of the IV and the WEP secret key was determined to be problematic due to a flaw in the Key Scheduling Algorithm of the RC4 cipher, which was quickly identified by Scott Fluhrer, Itsik Mantin, and Adi Shamir in their paper "Weaknesses in the Key Scheduling Algorithm of RC4" [1]. Using this information, Adam Stubblefield, John Ionnadis, and Aviel Rubin showed in their paper "Using the Fluhrer, Mantin and Shamir Attack to Break WEP" [2] how weak the WEP algorithm was and how easily it can be broken. It didn't matter whether you used 40-bit or 104-bit WEP keys (the extra 24-bits were provided by the IV), WEP was a seriously flawed security algorithm.

WPA and WPA2

The issues with WEP were cause enough for concern that the WiFi Alliance created a certification program for its replacements, *Wi-Fi Protected Access* (*WPA*) and WPA2. These improvements were needed to address the serious weaknesses in the way in which WEP was implemented. WPA was designed to meet these short-term needs of wireless security as a stopgap measure. One big change between WEP and WPA was the advancement of *Temporal Key Integrity Protocol* (*TKIP*). TKIP increases the Initialization Vector (IV) from 24-bits to 48-bits. WPA was designed to also use a different secret key for each packet and also featured Message Integrity Code (MIC), which was designed to detect invalid packets. WPA was effective in that it was designed as a stopgap measure until a completely new replacement could be approved and released. This replacement was WPA2 (802.11i).

WPA2 implemented all the elements that were requirements of the Wi-Fi Alliance and as specified in 802.11i. The standard took so long to be released that it was branded WPA-2 even though it uses a completely different method of security. WPA2 includes Robust Security Network (RSN) support, which includes added protection for ad-hoc networks, key caching, and pre-roaming authentication. WPA2 is built around the following:

- 802.1x used for authentication
- RSN used for tracking client association
- Advanced Encryption Standard (AES) with mandatory support for Counter Mode with Cipher Block Chaining Message Authentication Code Protocol (CCMP) used for confidentiality, integrity, and availability

WAP

The *Wireless Application Protocol* (*WAP*) is an open specification designed to enable mobile wireless users to easily access and interact with information and services. WAP is designed for hand-held digital wireless devices such as mobile phones, pagers, two-way radios, smartphones, and other communicators. It works over most wireless networks and can be built on many operating systems (OSs) including PalmOS, Windows CE, JavaOS, and others. The WAP specification added two significant enhancements to the above programming model: *push* and *telephony support* (Wireless Telephony Application, WTA). A WAP push is a Short Message Service (SMS) message containing a link to a WAP page. When the receiving handset gets the WAP push, the user can access the content in the page. WAP telephony support allows for telephone services to be available in a data environment so that WAP phones can be operated as both a telephone and a Web device.

WTLS

Wireless Transport Layer Security (*WTLS*) is an attempt by the WAP Forum to introduce a measure of security into WAP.

The WTLS protocol is based on the Transport Layer Security (TLS) protocol that is itself a derivative of the Secure Sockets Layer (SSL) protocol. However, several changes were made to these protocols to adapt them to work within WAP. These changes include:

- Support for both datagram- and connection-oriented protocols
- Support for long round-trip times
- Low-bandwidth, limited memory, and processor capabilities

WTLS is designed to provide privacy as well as reliability for both the client and the server over an unsecured network and is specific to applications that utilize WAP. These applications tend to be limited by memory, processor capabilities, and low-bandwidth environments.

Authentication

There are two authentication methods in the 802.11 standard:

- ***Open authentication*** also known as device-oriented authentication and can be considered a null authentication. All requests are granted.
- ***Shared-key authentication*** where the client and the AP are each preprogrammed with a common secret key that is used to prove that the client is authorized to communicate with the AP.

Crunch Time

The shared-key authentication process is a four-step process that begins when the AP receives the validated request for association. After the AP receives the request, a series of management frames are transmitted between the stations to produce the authentication. This includes the use of the cryptographic mechanisms employed by WEP as a validation and break down in the following manner:

1. The requestor (the client) sends a request for association.
2. The authenticator (the AP) receives the request, and responds by producing a random challenge text and transmitting it back to the requestor.
3. The requestor receives the transmission, encrypts the challenge with the secret key, and transmits the encrypted challenge back to the authenticator.
4. The authenticator decrypts the challenge text and compares the values against the original. If they match, the requestor is authenticated. On the other hand, if the requestor does not have the shared key, the cipher stream cannot be reproduced, therefore the plaintext cannot be discovered, and theoretically the transmission is secured.

EXAM WARNING

While the Security+ exam does not cover the authentication process in great detail, it is important to remember the two authentication mechanisms in the 802.11 standard: open and shared-key.

802.1X AUTHENTICATION

To address the weaknesses in WEP, many vendors (including Cisco and Microsoft) adopted the IEEE 802.1x authentication mechanism for wireless networks. The IEEE 802.1x standard was created for the purpose of providing a security framework for port-based access control that resides in the upper layers of the protocol stack. The most common method for port-based access control is to enable new authentication and key management methods without changing current network devices. The 802.1x authentication method is covered in more detail in Chapter 8, "Network Authentication".

EXAM WARNING

802.1x typically is covered in the access control, authentication, and auditing sections of the Security+ exam, but is relevant to wireless networks because of the fact that it is quickly becoming the standard method of securely authenticating on a wireless network. Also, do not confuse 802.1x with 802.11x.

ROGUE ACCESS POINTS

Another clever attack can be accomplished using *rogue APs*. If an attacker can put together an AP with enough strength, end users may not be able to tell which AP is the authorized one that they should be using. In fact, most will not even know that another is available. Using this technique, an attacker is able to receive authentication requests and information from the end workstation regarding the secret key and where they are attempting to connect.

Rogue APs can also be used to attempt to break into more tightly configured wireless APs. A hacker sitting in a car in front of a house or office is noticeable, and thus will generally not have enough time to finish acquiring enough information to break the key. However, if an attacker installs a tiny, easily hidden machine in an inconspicuous location, it could sit there and provide connectivity to users while capturing important information.

DATA EMANATION

Wireless systems are more vulnerable to attacks than wired systems, and data emanation is one such vulnerability. *Emanation* is simply something that is emitted or radiated. Data emanation is not just a problem with 802.11 wireless

networks but also with all types of wired and wireless equipment. Almost all activities dealing with computers or across a network involve data emanation. Consider the Cathode Ray Tube (CRT), a wireless keyboard, a Bluetooth headset, a cordless mouse. Each of these devices is at risk of some type of data emanation.

Research on this problem began back in the 1950s under the TEMPEST project. This project was designed to look at hardening devices to prevent emanations form items such as keyboards and CRTs. These early studies focused on ways to prevent interception of signals from systems that could be transmitting or holding sensitive information. One early technique was the Faraday cage. A *Faraday cage* is an enclosure made out of specific type of copper wire, which can be fashioned into an enclosure to block radio waves. When a Faraday cage is used, no electromagnetic radiation can enter or leave the item or equipment enclosed. Other techniques used to prevent data emanation include jamming or noise generators and control zones. Jamming is nothing more than the deliberate radiation of electromagnetic energy to disrupt the enemy's ability to intercept or send radio signals. Noise generators transmit broadcasting their own interference.

Finally, there are control zones. A *control zone* is designed to block radio signals; as such, it is really nothing more than a rather large Faraday cage used to block electromagnetic radiation. As an example, you may have secure equipment in one area of the building but have this area enclosed with a Faraday cage and for added protection place several noise generators outside the control zone.

BLUETOOTH

Bluetooth uses the same 2.4 GHz frequency that the IEEE 802.11b wireless networks use, but unlike those networks, Bluetooth can select from up to 79 different frequencies within a radio band. *Bluetooth* is a short-range protocol that includes three classes: 1 meter, 10 meters, and 100 meters. Unlike 802.11b networks, where the wireless client can only be associated with one network at a time, Bluetooth networks allow clients to be connected to seven networks at the same time. However, one of the main reasons that Bluetooth never succeeded like the 802.11b standard did is because of its low-bandwidth capabilities and a lack of range.

Bluetooth, by its very design, is not intended for the long ranges or high data throughput rates that 802.11 wireless networks have. This is largely due to the fact that the hop rate of Bluetooth devices is about 1600 hops per second with an average of a 625μs dwell time, thus producing exceptionally more management overhead than 802.11. While this exceptionally high hop rate does tend to make Bluetooth resistant to narrowband interference, it has the undesirable side effect of causing disruption of other 2.4 GHz-based network technologies, such as 802.11b and 802.11g. This high hop rate causes all-band interference on these 802.11 networks and can, in some cases, completely prevent an 802.11 wireless network from functioning.

Bluetooth has been shown to be vulnerable to the following attacks:

- ***Bluejacking***, which allows an individual the ability to send unsolicited messages over Bluetooth to other Bluetooth devices
- ***Bluesnarfing***, the theft of data, calendar information, or phone book entries
- ***Bluebugging***, which uses the Bluetooth protocol to establish a serial connection to the device, allowing access to the full control over the phone

SUMMARY OF EXAM OBJECTIVES

Wireless LANs are attractive to many companies and home users because of the increased mobility provided by the technology as well as its ease and low cost of deployment. Enterprise wireless networks utilize the IEEE 802.11 specification for wireless LANs. Initially the security of 802.11 networks was seriously flawed with the use of the WEP protocol but has evolved to the much stronger and robust WPA and WPA-2 standards specified in 802.11i.

TOP FIVE TOUGHEST QUESTIONS

1. The medium for communications in a wireless system is:

A. Cabling
B. Access point
C. Antenna
D. EM field

2. The area over which the radio waves propagate from an electromagnetic source is known as the:

A. Control point
B. Fresnel zone
C. Footprint
D. Wavelength

3. Which of the following is not a valid class for Bluetooth?

A. Class 0
B. Class 1
C. Class 2
D. Class 3

4. TEMPEST is best defined as:

   A. A method used to attack wired networks.
   B. A means to attack wireless networks.
   C. An investigation and study of compromising data emanation.
   D. A tool used to setup a rogue access point.

5. James is worried about the security of the wireless network and therefore has disabled SSID broadcasts. James has now made the statement that his wireless network cannot be hacked. How should you respond?

   A. Sniffing the SSID is not possible once the SSID broadcast has been disabled.
   B. Once broadcast has been disabled, sniffing the SSID is only possible with specialized expensive equipment.
   C. James is correct only if 128-bit WEP has been enabled.
   D. Even with SSID turned off, someone can still sniff the network.

ANSWERS

1. The correct answer is D. The medium for communications in a wireless system is EM field, the region of space that is influenced by electromagnetic radiation. (Unlike audio waves, radio waves do not require a medium such as air or water to propagate.) Cabling is used for a wired network. An access point is the termination point of the signal. An antenna is used by an access point as the device to emanate the EM field.

2. The correct answer is B. The area over which the radio waves propagate from an electromagnetic source is known as the Fresnel zone. Answer A is incorrect as a control point is the physical location from which a radio source's functions are controlled. Answer C is incorrect as the footprint is the surface space occupied by a structure or device. Answer D is incorrect as the wavelength is the distance (measured in the direction of propagation) between two points in the same phase in consecutive cycles of a wave.

3. The correct answer is A. The class rating of Bluetooth refers to the power class of the radio transmitter in the device. This includes Class 1 devices that have a range of 100 M. Class 2 devices that have a range of 10 M. Class 3 devices that have a range of less than 10 M.

4. The correct answer is C. TEMPEST was designed to look at hardening devices to prevent emanations form items such as keyboards and CRTs. It is not used to attack wired networks, wireless networks, or to setup a rogue access point.

5. The correct answer is D. It is possible to turn off SSID on some APs. Disabling SSID broadcasts creates a "closed network." If possible, SSID broadcasts should be disabled, although this will interfere with the ability of Windows XP to automatically discover wireless networks and associate with them. However, even if SSID broadcasts are turned off, it is still possible to sniff the network traffic and see the SSID in the frames. Answers A, B, and C are not correct.

REFERENCES

[1] Fluhrer, S., Mantin, I., and Shamir, A. "Weaknesses in the Key Scheduling Algorithm of RC4." Cisco Systems/Weizmann Institute; 2001. Available at <http://www.drizzle.com/~aboba/IEEE/rc4_ksaproc.pdf/>

[2] Stubblefield, A., Ionnadis, J., and Rubin, A. "Using the Fluhrer, Mantin, and Shamir Attack to Break WEP." ATT Labs; 2001. Available at <http://www.simovits.com/archive/break_wep.pdf/>

CHAPTER 7

Network Access

Exam objectives in this chapter:

- General Network Access
- Access Control Methods and Models
- Access Control Organization
- Logical Access Control Methods
- Physical Access Security Methods

GENERAL NETWORK ACCESS

When you are working with information security, the heart and soul of security is controlling access to objects. All other security measures and techniques are pointless if the objects they are meant to protect have no access controls.

Gaining access to network resources is based on identification through authentication, proving that you are you, requesting access, and being granted the requested access.

Access control

Access control encompasses the security controls, processes, or procedures whereby access to specific objects is either granted or denied based on preestablished policies or rules. Access control is made up of many different parts, but at its roots it is a very simple concept. The goal of access control is to allow objects to be accessed by those authorized to access it (and limit the manner in which it is accessed) while denying access to those who are not authorized.

Access control has several individual parts:

- Access Control Objects—The objects that need to be accessed. This includes hardware, networks, buildings, and data.

- Access Control Subjects—The users, programs, and processes that are requesting permission to access control objects.
- Access Control Systems—The procedures, processes, and controls in place to verify the authenticity of the request, the identity of the access control subject, and determine the levels of access that the subject should be granted to the object.

Access control can be implemented in many different ways, all of which have the end result of controlling access to data, systems, or hardware.

- Physical (i.e., biometric device to secure a door)
- Hardware (i.e., a dedicated firewall)
- Software (i.e., built-in application security)
- Policy (i.e., a workplace security policy)
- Network (i.e., secure networking protocols)

Access control models

Most access control systems are based off of several basic access control models. These models define the operating parameters for the access control system and define the manner in which they operate. The access control model also defines the way that permissions are set on access control objects and how authorization is handled in the access control system.

One of the first serious efforts to define the effectiveness of security controls in computing was the U.S. Department of Defense Trusted Computing System Evaluation Criteria (TCSEC). The centerpiece of TCSEC was the "Orange" book (named that way for the color of its cover) as well as another one, "Trusted Network Interpretation" (also known as the "Red" book). The Orange and Red books were superseded in 2005 by the Common Criteria for Information Technology Security Evaluation, also known as the Common Criteria or CC. Using the CC, software and hardware can be certified at a variety of evaluation assurance levels (EAL) similar to those available with the Orange and Red books. These levels range from EAL1 through EAL7. Using these criteria, Microsoft Windows Vista and Windows Server 2008 were graded as EAL1, the first level of certification.

The *formal models* of access control are theoretical applications of access control methods. These do not specify specific methods of controlling access, but rather specific guidelines that should be followed. They work best with static environments and are difficult to implement within dynamic systems that are constantly changing, such as those in most enterprising environments. The documentation on how these models are supposed to be implemented is very limited and does not give any specific examples.

- ***Biba***: The Biba formal model was written by K. J. Biba in 1977 and is unique as it was the first formal model to address integrity. The Biba model

bases its access control on levels of integrity. The Biba policy consists of three primary rules:

1. A subject cannot access objects that have a lower level of integrity than the access control subject has.
2. Access control subjects cannot modify objects that have a higher level of integrity than their current integrity levels.
3. Access control subject may not request services from subjects that have a higher integrity level.

- ***Clark–Wilson Model:*** Written in 1987 and updated in 1989 by David D. Clark and David R. Wilson.
  1. Similar to Biba as it addresses integrity.
  2. Addresses access to objects.
  3. Ensures integrity by specifying guidelines for processes that occur using the access control object.
  4. One of the most important guidelines to come out of Clark–Wilson is that of *segregation of duties* or *separation of duties*—no single person should perform a task from beginning to end, but that the task should be divided among two or more people to prevent fraud by one person acting alone.
- ***Bell–LaPadula:*** Written by David E. Bell and Len J. LaPadula in 1973 for use in government and military applications.
  1. Specifies that all access control objects have a minimum security level assigned to it so that access control subjects with a security level lower than the security level of the objects are unable to access the object
  2. Forms the basis for mandatory access control (MAC)

Authentication models and components

Authentication is basically the transfer of some form of information that provides proof of identity. This can be in many different forms, but there are three basic types under which all of the different forms of authentication fall.

- **Something you know** typically relies on the access control subject to memorize and know specific facts that can be used to prove who they are. This usually results in a password or a memorized Personal Identification Number (PIN).
- **Something you have** relies on some form of authentication that the access control subject physically has. This could be anything from a driver's license to authenticate them as a valid driver of vehicles to an ATM card used to authenticate them to their bank to a smart card or identification token to provide access to a network or a system.
- **Something you are** relies on *biometrics* and is based on the science of identifying people based on their physical characteristics. Common measurable physical traits include fingerprints, signatures, voiceprints, DNA, and others.

AUTHENTICATION TYPE COMBINATIONS

These three basic types cover the main types of authentication. In addition, the three types of authentication can be combined to provide even greater security. These combinations are called *factors of authentication*. A two-factor authentication method would make use of two of the three types of authentication. Three-factor authentication uses all three types of authentication and is considered the strongest form of authentication.

Some examples of authentication type combinations include:

- Requiring that a PIN be entered in combination with a 6-digit code displayed
- An authentication token
- Requiring a password, smart card, and fingerprint scan in order to enter a secure area

Identity

Identification is basically the concept of saying that you are a specific access control subject. This process can be as simple as saying, "Hi, I'm Peaches Perry," or as complex as presenting a sample of your DNA for a biometric scan. Note that identification is just the act of presenting yourself as the access control subject, nothing more. It does not involve *proving* that you are who you say you are. That is where authentication comes in.

The difference between identification and authentication is just proof.

ACCESS CONTROL METHODS AND MODELS

This section covers some of the key concepts associated with access control methods and models. Some of these directly relate to the formal models, so expect some repetition. However, the concepts themselves stand alone and can be implemented without implementing a full formal model.

Separation of duties

Separation of duties is a very important security concept, especially as it relates to fraud. *Separation of duties* is the concept that no one person should handle a transaction from beginning to end. Instead, some parts of a transaction should be executed by one person and other parts of the transaction should be executed by someone else entirely. Separation of duties allows for a basic level of fraud prevention by providing some checks and balances organizationally. The same principle applies in many areas and is actually one of the basic tenets of the anti-trust lawsuits, which are common for large organizations. When one entity controls too many portions of a transaction, there is increased risk that the transaction can be manipulated in the favor of the entity that has control over it. Separation of duties can help prevent this by establishing lines of segregation between elements of the transaction.

Least privilege

The principle of *least privilege* is very simple, but incredibly important. The concept here is to provide the lowest amount of access to an access control object necessary for the access control subject to perform their task. This has the unfortunate side effect of causing the access control subject to have to go through the steps of requesting access to what they need every time their roles or the tasks that they need to perform change. This is made more difficult by the fact that a typical end user may not know what they need access to in order to do their job. For example, they may request access to a specific file, but not the directory containing other files that the first file is linked to. In general, it is a best practice to attempt to maintain the principle of least privilege by granting access control subjects the most *appropriate* access to the requested access control object possible.

Job rotation

Job rotation is the concept of personnel changing job roles at a scheduled time into a different role. There are several reasons behind this, such as increasing education across the personnel, allowing for people to experience different parts of a company and how it functions, increased job satisfaction through change, and scheduled changes due to changes in leadership (terms in public offices). All of these reasons lead to changing the job role for personnel and raise new security concerns because of it. Going back to separation of duties, it might initially seem that job rotation violates this principle since personnel could be in roles, which would allow them to perform a transaction end to end. What prevents this from being a concern is that each role in job rotation would only be held by a single individual at one time, not multiple roles at the same time.

Mandatory access control

MAC is based off of sensitivity levels rather than access control lists (ACLs) on objects and is frequently used by government systems. In MAC, the security administrator gives every access control object and access control subjects a sensitivity level and the object owner or system user cannot change this sensitivity level. Based on the sensitivity levels of the access control objects, the access control system decides how all data will be shared and the data is restricted to the access control subjects with the required matching sensitivity label. For example, if an object has a sensitivity label of *top-secret*, an access control subject with a label of *secret* will be unable to access the object.

MAC is considered to be a more secure access control model than discretionary access control (DAC—described below) as every subject and object must have a label assigned to it. This model ensures that if a subject is not authorized to access data with a specific sensitivity label, they will not be able to access it. This works well in a strictly defined hierarchy such as the military where subjects are simply not authorized to access any information that is above their level in the hierarchy.

MAC has several major disadvantages. These include:

- Extremely difficult to implement.
- High administration overhead.
- Difficult to program applications to work with MAC due to the way objects are created and used.
- The total cost of ownership for MAC is not justified for most business purposes.

To review briefly, MAC is:

- **Non-discretionary**. The control settings are hard-coded and not modifiable by the user or owner.
- **Multi-level**. Control of access privileges is definable at multiple access levels.
- **Label-based**. May be used to control access to objects in a database.
- **Universally applied**. Applied to all objects.

Discretionary access control

DAC model is the most common access control model in use. This model bases security off of the identity of the access control subject. Every access control subject has specific permissions applied to it and based on these permissions has some level of authority.

This access control model is called discretionary because individual users or applications have the option of specifying access control requirements on specific access control objects that they own. In addition, the permission to change these access control requirements can also be delegated as a permission. Basically, the owner of the access control object is allowed to decide how they want their data protected or shared.

DAC allows for a distributed access control system to be used as the owner of the access control object has the ability to change the access control permission on objects without regard to a central authority. Also, centralized access control systems can be used with this as a single authoritative point of authorization with the permissions still being applied at the object level. The ability to use different types of access control systems with this model gives it a great deal of flexibility.

As previously mentioned, this is a very common access control model. It is used in UNIX, Windows, Novell Netware, Linux, and many other network operating systems. These systems use an *ACL* to set permissions on access control objects. These ACLs are basically a list of user IDs or groups with an associated permission level. Every access control object has an ACL, even if it is left at the default after

ACL Permissions	
Permission	**Definition**
Read	Allows the access control subject to read the data contained in the object
Write	Allows the access control subject to write data to the object
Create	Allows the access control subject to create new objects
Execute	Allows the access control subject to execute the code within the object
Modify	Combination of Read and Write, may also include Create and Execute
Delete	Allows the access control subject to delete the object
Rename	Allows the access control subject to rename the object
List	Allows the access control subject to list the contents of a directory—only applicable to directories
No Access	Explicitly denies the access control subject access to the object

the object is created. The operating systems vary in the way the permissions are defined in the ACL, but the Security+ exam is not vendor specific and does not require you to know how each operating system uses these. However, you are required to know the basic types of permissions that are defined. These are detailed in Table 7.1 along with a definition of what they mean.

EXAM WARNING

While the Security+ exam is not vendor specific and takes a general perspective of information security, you do need to know how DAC works and that many common network operating systems use DAC with ACLs as part of their access control security.

It is important to understand that DAC is assigned or controlled by the owner, rather than being hard-coded into the system. DAC does not allow the fine level of control available with MAC, but requires less coding and administration of individual files and resources.

To summarize, DAC is:

- **Discretionary**. Not hard-coded and not automatically applied by the OS/NOS or application.
- **Controllable**. Controlled by the owner of the object (file, folder, or other types).
- **Transferable**. The owner may give control away.

Role- and rule-based access control

Role- and rule-based access control (*RBAC*) can be described in different ways. The most familiar process is a comparison or illustration utilizing the "groups" concept.

In Windows, UNIX/Linux, and NetWare systems, the concept of groups is used to simplify the administration of access control permissions and settings. When creating the appropriate groupings, the administrator has the ability to centralize the function of setting the access levels for various resources within the system.

Rule-based access control is a method commonly used in applying MAC. In rule-based access control, access is granted or denied based on matching an object's sensitivity level as well as a subject's sensitivity level. Role-based access control is determined by the system rather than the owner of the asset and is typically used in multi-level access systems both commercially as well as in the military.

Although the concept of RBAC is similar, it is not the exact same structure. With the use of groups, a general level of access based on a user or machine object grouping is created for the convenience of the administrator. However, when the group model is used, it does not allow for the true level of access that should be defined, and the entire membership of the group gets the same access. This can lead to unnecessary access being granted to some members of the group.

RBAC allows for a more granular and defined access level, without the generality that exists within the group environment. A role definition is developed and defined for each job in an organization, and access controls are based on that role. This allows for centralization of the access control function, with individuals or processes being classified into a role that is then allowed access to the network and to defined resources. This type of access control requires more development and cost, but is superior to MAC in that it is flexible and able to be redefined more easily.

TIP

The best way to think of RBAC is to look at it like an organizational chart. Every person has a specific position and job function and the access control model mimics this organizational structure.

With RBAC, there is less administrative work than MAC as any objects created by a subject can be accessed by other subjects with the same role in the organization. This behavior can also be overridden in most access control systems using RBAC to increase security.

In summary, RBAC is:

- **Job-based**. The role is based on the functions performed by the user.
- **Highly configurable**. Roles can be created and assigned as needed or as job functions change.
- **More flexible than MAC**. MAC is based off of very specific information, whereas RBAC is based off of a user's role in the company, which can vary greatly.

- **More precise than groups**. RBAC allows the application of the principle of least privilege, granting the precise level of access required to perform a function.

EXAM WARNING

Be careful! RBAC has two different definitions in the Security+ exam. The first is defined as *role-based access control*. A second definition of RBAC that applies to control of (and access to) network devices is defined as *rule-based access control*. This consists of creating ACLs for those devices, and configuring the rules for access to them.

ACCESS CONTROL ORGANIZATION

When you are working with access control, it's typically easiest to control access by groups of access control subjects instead of applying security on an individual level. These access control subjects would, of course, have to have some common factor which would allow them to be grouped together such as sharing a job role, working in the same department, or even being located in the same building. By applying access controls on groups of access control subjects, an administrator can make their job a little easier while still applying good security practices.

Security groups

Because a company may have hundreds or thousands of users on a network or system, it would be an administrative nightmare to maintain access control over every single account. To make management easier, *groups* can be used to assemble user accounts together and define access control as a batch. For example, let's say a network administrator wanted branch office managers to have the ability to backup data on servers and workstations in their individual locations. The administrator could modify the account of every manager, or add each of these accounts to a Backup Operators group, which has the necessary permissions to backup data. By modifying the access control of one group, the access of each account that is a member of that group would also be affected.

User accounts and groups may be local to a computer or server, or have the ability to connect to servers on a network. This allows administrators to control what a user or group can do on a specific machine, or on the network as a whole. This is particularly useful when they want users to have different levels of access on individual machines and the network.

Some network operating systems also have the ability to control access through roles. *Roles* are similar to groups, as they can be used to control the access of numerous users as a batch. If a number of users have a similar role in an organization, the administrator can associate them with a role created on the network OS. The role would have specific access to resources such as drive mappings or other privileges unique to this role.

Security controls

Security controls refer to the access control mechanisms that we put into place to mitigate security risks. There are three levels of security controls that are typically put into place.

- **Preventative**. Prevent a security breach.
- **Detective**. Intended to detect when a security breach happens and get details on the breach.
- **Corrective**. Fix a problem after is has occurred and has been detected.

FILE RESOURCES

Despite the emphasis on group-based access permissions, a much higher level of security can be attained in all operating platforms by individually assigning access permissions. Administratively, however, it is difficult to justify the expense and time involved in tracking, creating, and verifying individual access permissions for thousands of users trying to access thousands of individual resources. Role-based access control is a method that can be used to accomplish the goal of achieving the status of least privileged access. It requires more design and effort to start the implementation, but develops a much higher level of control than does the use of groups.

Good practice indicates that the default permissions allowed in most OS environments are designed for convenience, not security. For this reason, it is important to be diligent in removing and restructuring these permissions. Applying important security concepts such as the principle of least privilege to file resources can help to prevent users from gaining access to data that they should not have as well as ensuring the integrity of the data.

PRINT RESOURCES

The organization of print resources is often forgotten from a security perspective. Print devices may not be considered critical to most organizations, but they are important, and incorrectly applying access controls to print resources can cause problems. Some of the most common methods are to organize based on physical location, functional area, or job role. Sometimes a combination of all three is used to provide for more precise access controls.

LOGICAL ACCESS CONTROL METHODS

Access control can be broken up into two primary methods, logical and physical. *Logical access control* involves applying access controls to logical entities such as data or the ability to perform a certain action on a computer. This differs from *physical access control* methods in that physical access control deals with limiting the ability to physically interact with an entity.

Access control lists

ACLs are lists of permissions associated with access control objects. Access control subjects are listed in the ACL as well as the level of permission that the

subject is granted to the object. This is all part of the DAC model as ACLs are enforced on a discretionary basis rather than being mandated by the operating system. As previously mentioned, MAC requires the definition of sensitivity levels on the access control subjects and objects to define the access level to be granted. ACLs make no use of sensitivity levels of the access control subjects and basically provide access in the exact manner defined through the rules laid out in the ACL. Another feature of ACLs is that it is possible to grant an ACL rule, which allows the access control subject to modify the ACL for the object. ACLs can typically be associated with single files, full directories, disks, ports, or any other type of access control object based on the capabilities of the access control system. There is also typically an implicit deny associated with ACLs to prevent access from being granted to access control subjects when they fall outside the defined criteria.

Group policies

Group policies are a feature of the Microsoft Active Directory technology, which allows for logical access control based on groupings of access control subjects. The access control subjects are organized by site, organizational unit (OU), or domain. After grouping the access control subjects, a *Group Policy Object* (*GPO*) can be defined, which controls many elements of a user's system access, including registry settings, auditing, software installation, and Internet Explorer settings. This GPO is created as a template, which can then be applied to the groups previously defined.

When a user is authenticated and granted access to a Windows workstation, the GPOs applicable to that user are pulled from the authenticating server and applied. In addition, the workstation will poll for new GPOs assigned to the user on a regular basis using a random delay between the ranges of 90 and 120 minutes. Any changes found between the previously applied GPOs and the new GPOs will be applied to the workstation at that time.

Domain policies

Domain policies are GPOs that are defined at the domain level within Microsoft Active Directory. GPOs, once defined, can be linked to sites, domains, or OUs. The GPOs defined at each of these levels as well as locally are always applied in the following order:

1. Local
2. Site
3. Domain
4. OU

GPOs at each level are applied with each having the ability to overwrite the previous GPO unless forced otherwise. This allows for setting specific security policies such as password expiration and required password length at the domain level overriding conflicting settings at the local level. Inheritance plays a very important role in the use of GPOs within Windows environments. While GPOs

are applied in the order listed above, there are options available to make exceptions. These are the "Block Policy Inheritance" and "No Override/ Enforced" options. "Block Policy Inheritance" prevents the GPOs at higher levels (such as domains) from being inherited by lower levels. The "No Override" option has recently been renamed to "Enforce" and prevents lower levels from overriding higher-level settings.

In the order of precedence, "Enforce" wins over "Block Inheritance." This means that a GPO at the domain level that requires a 14-character password and has the "Enforce" parameter set cannot be overridden at the OU level with an 8-character password requirement.

Time of day restrictions

Another form of logical access control is use of *time of day restrictions*. This is often used in the home environment to ensure that children are unable to use a computer outside of their allowed hours, but it actually began in corporate environments. Similar to the way a time-lock safe works, time of day restrictions prevent specific applications or systems from being used outside of specific hours. Some situations where this can be useful would be to restrict the size of a print job allowed to be processed during working hours or to prevent a bank teller workstation from being accessed after the bank is closed.

Account expiration

When you are dealing with access control, the most common task is granting access to access control objects. The second most common is removing access. This can be required due to a variety of reasons, including account revocation, changes in roles, and account expiration.

Often it makes sense for an account to only be valid for a specific duration. For example, an account might only be active for a contractor for the duration of their contract. Or an account allowing remote system access might only be valid during the term of a support contract. In cases such as this, it is logical to specify a specific duration of validity to access that has been granted and automatically revoke that access when the duration expires. Of course the option is always there to extend the expiration date if needed.

Logical tokens

Logical tokens are strings of values that can be used in lieu of a password to gain access to access control subjects. This concept allows for systems external to the client and server to be responsible for authenticating the user and then to pass along a token indicating that the user has been authenticated to the system which then authorizes access to the access control object. This mechanism is familiar to most security professionals as it is the method used by Kerberos to authenticate users.

Most frequently, logical tokens are used to prevent passwords from being sent directly from an access control subject to the access control system. Instead, the password is sent to a third party and a token issued in return that takes the place of a password when the access control subject then uses the access control system. When working with logical tokens, there are a few requirements that must be met in order for this mechanism to be effective.

- There must be a trusted third party who authenticates the access control subject and provides the token.
- The third party must be trusted by both the access control subject and the access control system responsible for the access control object(s) in question.
- There must be a method defined by the third party for creating tokens that allow for the access control system being issued the token to validate its authenticity. This can be done by decrypting the token using a previously shared key, revalidating the token with the third-party system, or some other similar technique.
- There must be a set of policies in place that determine important things such as how long a token is valid and how it can be revoked.

With these policies in place, there are rules defining the use of the tokens, and they can then be applied to improve the security of the processes associated with the logical tokens.

PHYSICAL ACCESS SECURITY METHODS

Physical security involves protecting systems from bodily contact and requires controlling access to hardware and software so that people are unable to damage devices and the data they contain. If people are unable to have physical access to systems, they will not be able to steal or damage equipment. Physical security also limits or prevents their ability to access data directly from a machine, or create additional security threats by changing account or configuration settings.

Physical security also requires protecting systems from the environmental conditions within a business. Environmental conditions such as floods, fires, electrical storms, and other natural disasters can result in serious losses to a business. These conditions can also leave a business exposed to situations such as power outages, leakage of data due to poor shielding, and other potential threats. Without strong physical security in place, unauthorized persons can access information in a variety of ways. When designing physical security, the first step is to identify what needs to be protected and what it needs to be protected from. Inventories should be made of servers, workstations, network connectivity devices, and other equipment within an organization.

When you are designing security, it is important to strike a balance between the cost of security and the potential loss. Servers are costly and may contain

valuable data, so a higher level of security is needed to protect them. On the other hand, an old computer in the Human Resources department that is used for keyboarding tests given to prospective employees needs little or no protection. When determining value, it is important to not only consider the actual cost of something, but also how difficult it is to replace or what the cost to the organization's credibility would be. While certain data may be of relatively low cost value, it may still be important to a company and difficult to replace.

When you are creating measures to protect systems, it is important to note that threats are not limited to people outside the company. One of the greatest challenges to physical security is protecting systems from people within an organization. Corporate theft is a major problem for businesses, because employees have easy access to equipment, data, and other assets. Because an employee's job may require working with computers and other devices, there is also the possibility that equipment may be damaged accidentally or intentionally. Physical security must not only protect equipment and data from outside parties, but also those within a company.

A good way to protect servers and critical systems is to place them in a centralized location. Rather than keeping servers in closets throughout a building, it is common for organizations to keep servers, network connectivity devices, and critical systems in a single room. Equipment that cannot be stored in a centralized location should still be kept in secure locations. Servers, secondary routers, switches, and other equipment should be stored in cabinets, closets, or rooms that are locked, have limited access, are air-conditioned, and have other protective measures in place to safeguard equipment.

DID YOU KNOW?

Even if the *physical security* of a location is suitable when a server was installed, it may not be at a later date. In an office environment, people will move to different offices, renovations will be made to facilities, and equipment will be moved. Even though a server was initially placed in a secure location, the server could be moved or the location could become insecure as changes are made.

Unfortunately, many of the decision makers in a company may be unaware of the importance of physical security for network equipment, and make changes without considering implications. In a large organization where much of the network administration is done remotely, IT staff may be unaware that such changes have even occurred.

Access lists and logs

Access lists are basically pre-authorized list of people who are allowed to enter an area. Think of it as similar to the list of celebrities that a bouncer will let in to a private party. Only the specific people on the list will be allowed entry. Just like most elements of access control, the visitor must first prove their identity using some form of identification in order to be allowed in assuming that

they are not recognized by sight. This is not a very secure manner of controlling access to buildings as it is not a very complicated process to fabricate a false identification card. All a potential intruder would need to know is the name of someone who is likely to be on the list and fabricate an ID to match.

Access logs require anyone entering a secure area to sign in before entering. When visitors require entry, such as when consultants or vendor support staff need to perform work in a secure room, an employee of the firm must sign the person in. In doing so, the employee vouches for the credibility of the visitor, and takes responsibility for this person's actions. The access log also serves as a record of who entered certain areas of a building. Entries in the log can show the name of a visitor, the time this person entered and left a location, who signed them in, and the stated purpose of the visit.

Hardware locks

One of the easiest methods of securing equipment is the simplest: Keep it behind a locked door. There are a variety of different locks that can be used. Some locks require metal keys to open them, much like those used to unlock the front door of a home. Other types may be programmed and have keypads requiring a PIN number or card key to open them. With these more advanced locks, features may be available that allow logging of anyone who enters the area, which is useful for monitoring who entered a secure area at a particular time.

Whether equipment is stored in a centralized server room or a locked closet, it is important that all sides of the room or closet are secure. Windows should be locked and alarmed, if possible, so that anyone climbing through will be stopped or detected. Air vents can also provide a route into a room, so any large vents should be bolted shut with grates. Even seemingly obscure routes should be blocked or secured. Intruders may be able to crawl into a room through the area between a false ceiling and the real ceiling, or the space between a raised floor and the concrete beneath.

TIP

Remember that physical security includes all sides of a room; the walls, ceiling, and floor. Even if most of these are secure, leaving one side of the room insecure can provide an avenue of penetration. Looking at the room this way will also help to identify where security lapses exist, and what security measures should be implemented.

ID badges

ID badges are identification cards issued to individuals who need access to a specific location. These badges will often include a photograph of the individual

as well as other identifying information such as their name or a badge number. They may also include a magnetic strip or an RFID tag that allows for storage of additional data or identification that the badge is genuine.

There may be policies in place at a company that requires ID badges to be worn at all times when on the premises. This is one manner of ensuring that only authorized individuals are on site at the facility. Anyone without an ID badge could potentially be an intruder and quickly identified due to the lack of the identifying badge.

There are, of course, some challenges with relying solely on ID badges for security at a site. If the ID badges do not include a photo, they can easily be stolen and used by unauthorized personnel. In addition, there are logistics that must be considered when implementing an ID badge system such as ensuring that a process exists for visitors with no badge, handling of stolen or lost badges, and temporary badges for personnel who forget their badge. Badges with no magnetic strip or imbedded chip for identification can be easily forged, and even those with these features can be duplicated with a little effort.

Door access systems

Door access systems have increased in complexity from simple locks to complex systems that perform elements of access control such as authentication and logging of entry/exit. With the increased complexity comes increased security as well as increased management needs. It is no longer a matter of just duplicating keys and handing them out. Management of door access systems now includes having to detail out how a person (access control subject) is going to identify themselves, how to authenticate them, and how to authorize their access to the room or building in question (access control object).

These door access systems come in two major types, standalone or centrally managed. *Standalone door access systems* typically have a small imbedded electronic system built into them that allows the administrator to set up all of the rules of access control, such as who has access to the door during which time periods. A standalone door access system is only concerned with the one door that it has to protect and does not share data with any other door access systems.

With a *centralized door access system* a central computer is responsible for performing the access control necessary across multiple door access systems. This allows for simpler centralized administration of the door access systems but has its own drawbacks. In the event of a failure of the centralized system, it is possible that multiple door access systems will stop working correctly. Some systems work around this by maintaining a copy of the centralized server's rules in the memory of each local door access system so that if the central system fails, they can continue functioning using the last update they received until the problem is corrected. This works well from the perspective of keeping things running, but can be a security problem as revoked IDs will still work on systems with cached data until they receive a new rule set from the central server.

Even with the most stringent physical security in place, there are ways of bypassing these methods and gaining access. One of the simplest methods is *tailgating*, or *piggybacking*, in which an unauthorized person follows an authorized person into a secure area.

Man-trap

A *man-trap* in a security context is basically a method used for physically trapping a person if they fail to be properly authenticated. It often takes the form of a small room with two doors. After gaining entry through the first door, there are specific criteria that must be met for the second door to open. The first of these is typically that the first door is closed. This prevents the person entering the man-trap from leaving if they fail to pass the other criteria necessary to enter the second door.

Video surveillance

Chaperoning someone who has been given clearance to an area is not always possible or desirable. For example, if you have hired an outside party to install equipment that is needed for Internet access, you may not want to stand beside the installer for an extended period of time. However, workers can be monitored in high-security locations using video cameras to provide electronic surveillance. This provides a constant eye, and allows for review of their actions if an incident occurs.

SUMMARY OF EXAM OBJECTIVES

This chapter examined network access in general—both from physical and logical perspective. The three primary access security models are Biba, Bell–LaPadula, and the Clark–Wilson model. The Biba and the Clark–Wilson model attempt to address integrity, while the Bell–LaPadula model forms the basis for MAC. The primary access control used in modern operating systems is the DAC model. In addition to these concepts, physical security, ACLs, and group policies were reviewed in the context that they play in network security.

TOP FIVE TOUGHEST QUESTIONS

1. When using DAC systems with ACLs, what permission or privilege gives users the ability to read and write to an access control object?

   A. Write
   B. Create
   C. Execute
   D. Modify

2. How does role-based access control differ from DAC?

   A. Role-based access control requires that permissions be configured on every object and DAC does not.

B. Role-based access control uses the ID of the user to help determine permissions to objects and DAC does not.
C. Role-based access control uses the position of the user in the organization structure to determine permissions for objects and DAC does not.
D. Role-based access control requires that every object have a sensitivity label and DAC requires that every object have an ACL.

3. The Clark–Wilson formal access control model specifies a very important guideline related to account administration. What is this guideline and what does it mean?

A. ***Principle of least privilege***. Grant all the rights and permissions necessary to an account, but no more than what is needed.
B. ***Account administration***. Work hand-in-hand with the human resources or personnel office of the company to ensure that accounts can be authorized and created when employees are hired and immediately destroyed when they are dismissed.
C. ***Segregation of duties***. No single person should perform a task from beginning to end, but the task should be divided among two or more people to prevent fraud by one person acting alone.
D. ***Access control***. Provide access control subjects the ability to work with access control objects in a controlled manner.

4. When you are administering access control objects in a MAC system, what is an important part of your duty?

A. Declassifying data when necessary
B. Removing ACLs when necessary
C. Deleting inactive accounts regularly
D. Replacing expired access control tokens when necessary

5. You are designing the access control methodology for a company implementing an entirely new IT infrastructure. This company has several hundred employees, each with a specific job function. The company wants their access control methodology to be as secure as possible due to recent compromises within their previous infrastructure. Which access control methodology would you use and why?

A. Role-based access control because it is job-based and more flexible than MAC.
B. Rule-based access control because it is user-based and easier to administer.
C. Groups because they are job-based and very precise.
D. Groups because they are highly configurable and more flexible than MAC.

ANSWERS

1. The correct answer is D. Answers A, B, and C are incorrect as none of them give the ability to both read and write to an access control object.

2. The correct answer is C. Answer A is incorrect as DAC also requires permissions to be configured on every object. Answer B is incorrect as DAC uses the identity of the access control subject as the basis of security. Answer D is incorrect as DAC is based off the identity of the access control subject.

3. Answer C is correct. Answers A, B, and D are not specified in the formal Clark–Wilson model.

4. The correct answer is A. Data classification in a MAC system determines access rights to that data. Answer B is incorrect as the ACLs are applied based on the classification of the data. Answer C is incorrect as MAC is not concerned with active and inactive accounts. Answer D is incorrect as access control tokens are not the administrator's responsibility in MAC.

5. The correct answer is A. Answer B is incorrect as rule-based access control is based on object and subject sensitivity level. Answers C and D are incorrect as groups are neither precise nor are they more flexible than MAC.

CHAPTER 8

Network Authentication

Exam objectives in this chapter:

- Authentication Methods
- Authentication Systems

AUTHENTICATION METHODS

Authentication, authorization, and auditing (AAA) is a set of primary concepts that aid in understanding computer and network security as well as access control and is used daily to protect property, data, and systems from intentional or even unintentional damage. AAA is used to support the confidentiality, integrity, and availability (CIA) security concept, in addition to providing the framework for access to networks and equipment using Remote Authentication Dial-In User Service (RADIUS) and Terminal Access Controller Access Control System (TACACS/TACACS+).

AAA is a group of processes used to protect the data, equipment, and confidentiality of property and information. One of the goals of AAA is to provide Confidentiality, Integrity, and Availability (CIA). CIA can be briefly described as follows:

- **Confidentiality:** The contents or data are not revealed.
- **Integrity:** The contents or data are intact and have not been modified.
- **Availability:** The contents or data are accessible if allowed.

AAA consists of three separate areas that work together. These areas provide a level of basic security in controlling access to resources and equipment in networks. This control allows users to provide services that assist in the CIA process for further protection of systems and assets.

DID YOU KNOW?

The *difference between access control and authentication* is a very important distinction, which you must understand in order to pass the Security+ exam. Access control is used to control the access to a resource through some means. This could be thought of as a lock on a door or a guard in a building. Authentication, on the other hand, is the process of verifying that the person trying to access whatever resource is being controlled is authorized to access the resource. In our analogy, this would be the equivalent of trying the key or having the guard check your name against a list of authorized people. So in summary, access control is the lock and authentication is the key.

Access control

Access control is a policy, software component, or hardware component that is used to grant or deny access to a resource. Control can be enforced through a variety of ways:

- Smart Card
- Biometric device
- Network access hardware (routers, remote access points)
- Virtual private networks (VPNs)
- Wireless access points (WAPs)

Access control can also refer to file or shared-resource permissions assigned through network operating systems including:

- Microsoft Windows with Active Directory
- UNIX systems using Lightweight Directory Access Protocol (LDAP)
- Kerberos
- Sun Microsystems' Network Information System (NIS) and Network Information System Plus (NIS+)

Authentication

Authentication is the process used to verify that a machine or user attempting access to the networks or resources is, in fact, the entity being presented. *Nonrepudiation* is the method used (time stamps, particular protocols, or authentication methods) to ensure that the presenter of the authentication request cannot later deny they were the originator of the request. Authentication methods include presentation of credentials (such as a username and password, Smart Card, or personal identification number [PIN]) to a NOS (logging on to a machine or network), remote access authentication, and a discussion of certificate services and digital certificates. The authentication process uses the information presented to the NOS (such as username and password) to allow the NOS to verify the identity based on those credentials.

Auditing

Auditing is the process of tracking and reviewing events, errors, access, and authentication attempts on a system.

AUTHENTICATION METHODS

Authentication, when looked at in its most basic form, is simply the process used to prove the identity of someone or something that wants access. This can range from a very simple process such as one-factor authentication to more complex systems involving two- and three-factor authentication.

One-factor

One-factor authentication has been used for authenticating uses for many years and includes simple authentication methods such as username and password combinations. Most operating systems have had some form of local authentication that could be used if it was designed to be used by multiple users. From a security standpoint, it is important to understand that the first line of defense of a system is the creation and maintenance of a password policy that is enforced and workable.

Password policies define the appropriate length and strength of a password as well as how often they are changed and the penalties for sharing passwords.

Good password policies should require that an acceptable password contain a combination of the following:

- Uppercase and lowercase alphabetic characters
- Numbers
- Special characters
- No dictionary words
- No portion of the username in the password
- No personal identifiers should be used including birthdays, social security number, pet's name, and so on

Password strength based solely on the length of the password is generally considered according to the following guidelines:

- **Low:**6 characters long
- **Medium**:between 8 and 13 characters
- **High:**14 or more characters

Two-factor

Two-factor authentication is typically implemented as a combination of *something you have* (e.g., Automatic Teller Machine [ATM] cards) and *something you know* (a PIN). In order to misuse a victim's authentication credentials in a two-factor authentication scheme like an ATM, both the ATM card and the PIN number must be acquired. *Token authentication* is a form of two-factor authentication and can be provided by way of either hardware- or software-based tokens:

- A hardware device that is coded to generate token values at specific intervals
- A software or server-based component that tracks and verifies that these codes are valid

The token code is entered into the server/software monitoring system during setup of the system. A user wishing to be authenticated visits the machine or resource they wish to access and enters a PIN number in place of the usual user logon password. They are then asked for the randomly generated number currently present on their token. When entered, this value is checked against the server/software system's calculation of the token value. If they are the same, the authentication is complete and the user can access the machine or resource.

Three-factor

Three-factor authentication, commonly known as *Multifactor authentication*, is the process in which another item for authentication in addition to or in place of the traditional password is used to complete the authentication process. The implementation should utilize three independent authentication mechanisms available.

Here are four possible types of factors that can be used for multifactor authentication.

- A password or a PIN can be defined as a *something you know* factor.
- A token or Smart Card can be defined as a *something you have* factor.
- A thumbprint, retina, hand, or other biometrically identifiable item can be defined as a *something you are* factor.
- Voice or handwriting analysis can be used as a *something you do* factor.

Single sign-on

Single Sign-On (*SSO*) is a process in which access to different systems is accomplished by authenticating the user once. SSO should be implemented with stringent policies for access control and authorization mechanisms and group policies to ensure simplification do not result in compromise in security. SSO can be implemented through various network operating systems.

AUTHENTICATION SYSTEMS

Authentication systems are used to provide user authentication in a wide variety of applications from local domain service to online banking systems. This section covers Remote Authentication Dial In User Service (RADIUS), Kerberos, and LDAP authentication services and TACACS+ as well as authentication protocols including PAP, CHAP, and 802.1x.

Remote access policies and authentication

Remote users may connect to a network through dial-in services using an analog line and a modem by dialing in to an organization's modem pool connected to a dial-in server, or through a VPN client software configured on their laptops or remote desktops to connect to the corporate VPN server. Even wireless clients connecting through the WAPs can be defined as remote users and restrictions can be applied on them. A remote access policy defines the conditions and remote access permissions, and creates a profile for every remote connection made to the corporate network.

Remote access policies can define:

- Grant or deny dial-in based on connection parameters such as type and time of the day
- Authentication protocols (PAP, CHAP, EAP, MS-CHAP)
- Validation of the caller ID
- Call back
- Apply connection restrictions upon successful authorization
- Create remote user/connection profile
- Assign a static IP or dynamic IP from the address pool defined for remote users
- Assign the user to a group to apply group policies
- Configure remote access permission parameters
- Define encryption parameters (for a remote access VPN client)
- Control the duration of the session including maximum time allowed and idle time before the connection is reset

Biometrics

Biometric devices can provide a higher level of authentication than, for example, a username and password combination. However, although they tend to be relatively secure, they are not impervious to attack. For instance, in the case of fingerprint usage for biometric identification, the device must be able to interpret

the actual presence of the print. Early devices that employed optical scans of fingerprints were fooled by fogging of the device lenses, which provided a raised impression of the previous user's print as it highlighted the oils left by a human finger. Some devices are also subject to silicon impressions or fingerprinting powders that raise the image. Current devices may require a temperature or pulse sense as well as the fingerprint to verify the presence of the user, or another sensor that is used in conjunction with the print scanner, such as a scale. Biometrics used in conjunction with Smart Cards or other authentication methods lead to the highest level of security.

RADIUS

Remote Authentication Dial-In User Service (*RADIUS*) was initially created by Livingston Enterprises to handle dial-in authentication and its usage broadened into wireless authentication and VPN authentication. An authentication service must be able to authenticate a user, authorize the authenticated user to perform specified functions, and log (i.e., account for) the actions of users for the duration of the connection.

When users dial into a network, RADIUS is used to authenticate usernames and passwords. One of the reasons that RADIUS is so popular is that it supports a number of protocols, including:

- Password Authentication Protocol (PAP)
- Challenge Handshake Authentication Protocol (CHAP)

Authentication process

RADIUS authentication consists of five steps:

1. Users initiate a connection with an ISP Remote Access Server (RAS) or corporate RAS. Once a connection is established, users are prompted for a username and password.
2. The RAS encrypts the username and password using a *shared secret* and passes the encrypted packet to the RADIUS server. The *shared secret* is set by the administrator between the RADIUS client and the RADIUS server.
3. The RADIUS server attempts to verify the user's credentials against a centralized database.
4. If the credentials match those found in the database, the server responds with an *access-accept* message. If the username does not exist or the password is incorrect, the server responds with an *access-reject* message.
5. The RAS then accepts or rejects the message and grants the appropriate rights.

KERBEROS

Kerberos is used as the preferred network authentication protocol in many medium and large environments, to authenticate users and services requesting access to resources. Kerberos is a network protocol designed to centralize

the authentication information for the user or service requesting a resource. Kerberos allows for the authentication of an entity requesting access (user, machine, service, or process) by a host of the resource being accessed through the use of secure and encrypted keys and tickets (*authentication tokens*) from a Key Distribution Center (KDC).

DID YOU KNOW?

The overall structure of control in Kerberos is called a *realm*. The process of authentication for access to a resource or utilizes the following sequence:

1. The user authenticates to the KDC using a username/password combination.
2. The KDC, upon validation of the authentication credentials, returns a Ticket Granting Ticket (TGT) to the client.
3. The client caches the TGT locally for future use.

A process to access a resource or service in the Kerberos realm is an extension of the aforementioned procedure. Here is a simplified version of this process.

1. The client presents a TGT (cached locally on the client) to the KDC.
2. The KDC validates the TGT and returns a session ticket to the client.
3. The session ticket is presented to the target server or resource.
4. The session is established and communication is allowed for the valid duration of the session ticket.

Kerberos uses a time stamp, and we need to understand where and when the time stamp is used. The time stamp is used to limit the possibility of *replay* or *spoofing* of credentials. Replay is the capture of information, modification of the captured information, and retransmission of the modified information to the entity waiting to receive the communication. If clocks are not synchronized between the systems, the credentials (tickets) will not be granted if the time differential exceeds the established limits.

LDAP

Directory services are used to store and retrieve information about objects, which are managed by the service. On a network, these objects can include user accounts, computer accounts, mail accounts, and information on resources available on the network. Directory services shouldn't be confused with the directory itself. The *directory* is a database that stores data on the objects managed through directory services.

Directory services are used by many different operating systems to organize and manage the users, computers, printers, and other objects making up the network. Some of the directory services that are produced by vendors include:

- Active Directory, which was developed by Microsoft for networks running Windows 2000 Server, Windows 2003 Server, or Windows 2008.

- eDirectory, which was developed by Novell for Novell NetWare networks. Previous versions for Novell NetWare 4.x and 5.x were called Novell Directory Services (NDS).
- Open Directory, which was developed by Apple for networks running Mac OS X Servers.

LDAP is a protocol that enables clients to access information within a directory service, allowing the directory to be searched and objects to be added, modified, and deleted. LDAP services are used to access a wide variety of information that's stored in a directory.

LDAP directories

LDAP directories follow a hierarchy, much in the same way that the directories on a hard drive are organized in a hierarchy—each uses a tree-like structure, branching off of a root with containers (called organizational units in LDAP) and objects (also called entries in LDAP's directory). Each of the objects has attributes or properties that provide additional information.

Because LDAP directories are organized as tree structures (sometimes called the Directory Information Tree [DIT]), the top of the hierarchy is called the *root*. The *root server* is used to create the structure of the directory, with organizational units and objects branching out from the root. Because the directory is a distributed database, parts of the directory structure may exist on different servers. Segmenting the tree based on organization or division and storing each branch on separate directory servers increases the security of the LDAP information. By following this structure, even if one directory server is compromised, only a branch of the tree (rather than the entire tree) is compromised.

Organizational units

The hierarchy of an LDAP directory is possible because of the various objects that make up its structure. These objects represent elements of the network, which are organized using containers called *organizational units* (*OUs*). Each OU can be nested in other OUs, similar to having subfolders nested in folders on your hard disk. In the same way the placement of folders on your hard disk makes a directory structure, the same occurs with OUs and objects in an LDAP directory. The topmost level of the hierarchy generally uses the domain name system (DNS) to identify the tree.

The structure of the LDAP directory is not without its own security risks, as it can be a great source of information for intruders. Viewing the placement of OUs can provide a great deal of information about the network structure, showing which resources are located in which areas of the organization.

Objects, attributes, and the schema

As mentioned, entries in the directory are used to represent user accounts, computers, printers, services, shared resources, and other elements of the network.

These objects are named, and as we discussed with organizational units, each object must have a name that's unique to its place in the namespace of the hierarchy. Just as you can't have two files with the same name in a folder on your hard disk, you can't have two objects with the same name in an OU. The name given to each of these objects is referred to as a *common name*, which identifies the object but doesn't show where it resides in the hierarchy.

The common name is part of the LDAP naming convention. Just as a filename identifies a file, and a full pathname identifies its place in a directory structure, the same can be seen in the LDAP naming scheme. The common name identifies the object, but a *distinguished name* can be used to identify the object's place in the hierarchy. An example of a distinguished name is the following, which identifies a computer named DellDude that resides in an organizational unit called Marketing in the tacteam.net domain:

DN: CN = DellDude,OU = Marketing,DC = tacteam,DC = net

The distinguished name is a unique identifier for the object, and is made up of several attributes of the object. It consists of the *relative distinguished name*, which is constructed from some attribute(s) of the object, followed by the distinguished name of the parent object. Each of the attributes associated with an object are defined in the schema. The *schema* defines the object classes and attribute types, and allows administrators to create new attributes and object classes specific to the needs of their network or company.

Securing LDAP

LDAP is vulnerable to various security threats, including spoofing of directory services, attacks against the databases that provide the directory services, and many of the other attack types discussed in this book (e.g., viruses, OS and protocol exploits, excessive use of resources and denial of service, and so forth). LDAP clients must authenticate to the server before being allowed access to the directory. Clients (users, computers, or applications) connect to the LDAP server using a distinguished name and authentication credentials (usually a password). LDAP allows for anonymous clients to send LDAP requests to the server without first performing the bind operation. LDAP can also be used over SSL, which extends security into the Internet. LDAPS is Secure LDAP, which encrypts LDAP connections by using SSL or Transport Layer Security (TLS). Some of these types of services integrate as objects, such as PKI certificates, in the authentication process using Smart Card technologies, and in the extended properties of account objects so that they can support extra security requirements. To use SSL with LDAP, the LDAP server must have an X.509 server certificate. Additionally, SSL/TLS must be enabled on the server.

PAP

Password Authentication Protocol (*PAP*) is the simplest form of authentication for a remote access. This method was used earlier to authenticate users

using username and passwords. PAP transmits the username and password in ASCII without any encryption. PAP was replaced with CHAP to provide more security.

CHAP

Challenge Handshake Authentication Protocol (*CHAP*) is a remote access authentication protocol used in conjunction with Point-to-Point Protocol (PPP) to provide security and authentication to users of remote resources. CHAP is used to periodically verify the identity of the peer using a three-way handshake. This is done upon initial link establishment, and may be repeated anytime after the link has been established.

1. After the link establishment phase is complete, the authenticator sends a "challenge" message to the peer.
2. The peer responds with a value calculated based on an ID value, a random value, and the password using a "one-way hash" function such as MD5.
3. The authenticator checks the response against its own calculation of the expected hash value. If the values match, the authentication is acknowledged; otherwise the connection should be terminated.
4. At random intervals, the authenticator sends a new challenge to the peer, and repeats steps 1 to 3.

CHAP's shared secrets may be stored on both ends as a cleartext item, making the secret vulnerable to compromise or detection.

TACACS/TACACS+

RADIUS is not the only centralized remote authentication service. *TACACS* is another remote authentication service that has gone through three major "generations": TACACS, XTACACS, and TACACS+.

Crunch Time

For the Security+ exam you need to know only about TACACS and TACACS+.

TACACS

TACACS was first developed during the days of ARPANET and is detailed in RFC 1492. Although TACACS offers authentication and authorization, it does not offer any accounting tools.

TACACS+

TACACS+ is a proprietary version of TACACS developed by Cisco Systems. TACACS+ separates the AAA functions and uses individual databases for each

AAA function. TACACS+ offers secure communications between the TACACS+ client and server utilizing TCP as its transport protocol.

EXAM WARNING

Make sure you understand the difference between TACACS and TACACS+. The most important thing to remember is that TACACS uses UDP as its transport protocol while TACACS+ uses TCP. Also, TACACS+ is a proprietary version owned by Cisco.

Vulnerabilities

The largest vulnerability in TACACS+ is the comparative weakness of the encryption mechanism. In addition, one of the biggest complaints regarding TACACS+ is that it does not offer protection against *replay attacks*. Other common weaknesses of TACACS+ include:

- **Birthday attacks**: The pool of TACACS+ session IDs is not very large; therefore, it is reasonable that two users could have the same session ID.
- **Buffer overflow**: Like RADIUS, TACACS+ can fall victim to buffer-overflow attacks.
- **Packet sniffing**: The length of passwords can be easily determined by "sniffing" a network.
- **Lack of integrity checking**: An attacker can alter accounting records during transmission because the accounting data is not encrypted during transport.

MUTUAL AUTHENTICATION

Mutual authentication is a process where both the requestor and the target entity must fully identify themselves before communication or access is allowed. This can be accomplished through a share secret or a Diffie–Hellman key exchange that protects the secret being used for the verification and authentication process. Certificates represent another method that can be used for mutual authentication. To verify the identities, the Certificate Authority (CA) must be known to both parties, and the public keys for both must be available from the trusted CA.

Mutual authentication provides for confidence that communication is not being intercepted by a man-in-the-middle (MITM) attacker or being redirected in any way. Mutual authentication provides more secure communications by positively identifying both sides of a communication channel.

802.1x and EAP provide for a mutual authentication capability. This makes the clients and the authentication servers mutually authenticating end points, and

assists in the mitigation of attacks from MITM types of devices. Any of the following EAP methods provide for mutual authentication:

- **TLS**: Requires that the server supply a certificate and establish that it has possession of the private key
- **IKE**: Requires that the server show possession of a preshared key or private key (this can be considered certificate authentication)
- **GSS_API (Kerberos)**: Requires that the server can demonstrate knowledge of the session key

802.1X METHODS

The *IEEE 802.1x* standard was created for the purpose of providing a security framework for port-based access control that resides in the upper layers of the protocol stack. The most common method for port-based access control is to enable new authentication and key management methods without changing current network devices. The benefits that are the end result of this work include the following:

- Significant decrease in hardware cost and complexity.
- More options, allowing administrators to pick and choose their security solutions.
- The latest security technology can be installed and should still work with the existing infrastructure.
- Quick response to security issues as they arise.

When a client device connects to a port on an 802.1x-capable AP, the AP port determines the authenticity of the devices. Here is a typical 802.1x authentication process.

1. Authenticator places client in an unauthorized state.
2. Authenticator sends request for user credentials.
3. User responds with their username and password.
4. Authenticator forwards credentials to authentication server.
5. Authentication server verifies credentials. If they are valid, then the client is allowed onto the WLAN.

User identification and strong authentication

With the addition of the 802.1x standard, clients are identified by username, not by the MAC addresses of the devices. This design not only enhances security, but also streamlines the process of authentication, authorization, and accountability (AAA) for the network. 802.1x was designed to support extended forms of authentication using password methods (such as one-time passwords or GSS_API mechanisms like Kerberos) and nonpassword methods (such as biometrics, Internet Key Exchange [IKE], and Smart Cards).

Dynamic key derivation

The IEEE 802.1x standard allows for the creation of per-user session keys. Wired Equivalency Privacy (WEP) keys, which are also called *network keys*, do not have

to be kept at the client device or at the WAP when using 802.1x. These WEP keys are dynamically created at the client for every session, thus making it more secure. The Global key, like a broadcast WEP key, can be encrypted using a Unicast session key, and then sent from the AP to the client in a much more secure manner.

EAP

Extensible Authentication Protocol (*EAP*) is an authentication protocol designed to support several different authentication mechanisms. It runs directly over the data link layer and does not require the use of Internet Protocol (IP).

Per-packet authentication

Per-packet authentication and integrity protection work for the following (packet is encrypted unless otherwise noted):

- TLS and IKE derive session key
- TLS ciphersuite negotiations (not encrypted)
- IKE ciphersuite negotiations
- Kerberos tickets
- Success and failure messages that use a derived session key (through WEP)

TIP

It is helpful to write out a table showing the various authentication methods used in 802.11 networks (e.g., open authentication, shared-key authentication, and 802.1x authentication) with the various properties each of these authentication methods require. This will help keep them straight in your mind when taking the test.

PEAP

Protected Extensible Authentication Protocol (*PEAP*) is a member of the family of EAP. PEAP uses TLS to create an encrypted channel between the client supplicant and the RADIUS server. PEAP provides additional security for the client-side EAP authentication protocols, such as EAP-MS-CHAPV2, that can operate through the TLS encrypted channel.

The advantages of PEAP are:

- Support in Windows Server 2008, Windows Server 2003, Windows 2000, Windows XP, and Pocket PC 2002.
- PEAP uses a TLS channel to protect the user credentials.

- Using the TLS channel from the client to the authentication server, PEAP offers end-to-end protection, not just over the wireless data link.
- PEAP supports any EAP compatible methods. PEAP is also defined as an extensible authentication method that can embrace new EAP authentication schemes as they become ratified.
- Within the TLS channel, PEAP hides the EAP type that is negotiated for mutual client and server authentication. Also, because each packet sent in the TLS channel is encrypted, the integrity of the authentication data can be trusted by the PEAP client and server.
- PEAP offers strong protection against the deployment of unauthorized WAPs because the client verifies the RADIUS server's identity before proceeding ahead with further authentication or connectivity. The WAP is unable to decrypt the authentication messages protected by PEAP.
- PEAP offers highly secure keys that are used to encrypt the data communications between the clients and WAP. New encryption keys are derived for each connection and are shared with authorized WAPs accepting the connection.
- PEAP does not require the deployment of certificates to wireless clients. Only the PEAP server (authentication server) needs to be assigned a certificate.
- Microsoft offers native support for PEAP so that a user can use the same logon credentials for all network connections and applications. PEAP integrates seamlessly with Microsoft Windows domain policy, Group Policy, and logon scripts.
- PEAP is an open standard supported under the security framework of the IEEE 802.1x specification.
- PEAP offers security and efficiency when used with roaming wireless devices through quick re-authentication. PEAP supports this capability through the TLS session resumption facility.
- PEAP provides support for EAP authentication methods such as EAP-TLS and EAP-MS-CHAPV2 that can perform computer authentication.
- The PEAP specifies an option of hiding a user's name known as **identity privacy**.

SUMMARY OF EXAM OBJECTIVES

Authentication has changed significantly since the early days of computing when username and passwords were the only method available and passwords were statically stored. At the root of authentication systems this is still the case; however, the methods of authentication have changed due to new requirements

of technology as well as system architecture. The need for stronger authentication has resulted in the proliferation of multifactor authentication technologies such as smart cards, tokens, and even biometric systems.

In addition to stronger authentication methods there are also stronger authentication systems. RADIUS, TACACS+, LDAP, and Kerberos all provide greater protection than the straightforward username and password combination although, at the heart of each one, is a username and password that provides the final validation of user identity.

Early experiences with 802.11 wireless technologies and the flaws in the authentication and privacy systems in the original 802.11b specification has led to an explosion and widespread use of 802.1x authentication systems based on EAP and PEAP. These systems provide not just client but also network authentication and thus provide stronger security for network and security administrators.

TOP FIVE TOUGHEST QUESTIONS

1. You are a security consultant for a large company that wants to make its intranet available to its employees via the Internet. They want to ensure that the site is as secure as possible. To do this, they want to use multifactor authentication. The site uses an ID and password already but they want to add security features that ensure that the site is indeed their site, not a spoofed site, and that the user is an authorized user. Which authentication technology supports this?

   A. Certificates
   B. CHAP
   C. Kerberos
   D. Tokens

2. You are attempting to query an object in an LDAP directory using the distinguished name of the object. The object has the following attributes:
   cn: 4321
   givenName: John
   sn: Doe
   telephoneNumber: 905 555 1212
   employeeID: 4321
   mail: mailto:jdoe@nonexist.com jdoe@nonexist.com
   objectClass: organizationalPerson
   Based on this information, which of the following would be the distinguished name of the object?
   A. dc = nonexist, dc = com
   B. cn = 4321
   C. dn: cn = 4321, dc = nonexist, dc = com
   D. mailto:jdoe@nonexist.com jdoe@nonexist.com

3. You are creating a new LDAP directory in which you will need to develop a hierarchy of organizational units and objects. To perform

these tasks, on which of the following servers will you create the directory structure?

A. DIT
B. Tree server
C. Root server
D. Branch server

4. Choose the correct set of terms: When a wireless user, also known as the _____ wants to access a wireless network, 802.1x forces them to authenticate to a centralized authority called the _____.

A. Authenticator; supplicant
B. Supplicant; authenticator
C. Supplicant; negotiator
D. Contact; authenticator

5. EAP is available in various forms, including:

A. EAPoIP, EAP-TLS, EAP-TTLS, RADIUS, EAP-FAST
B. EAPoIP, EAP-TLS, EAP-MPLS, RADIUS, EAP-FAST
C. EAPoIP, EAP-TLS, EAP-TTLS, RADIUS, Cisco PEAP
D. EAPoIP, EAP-TLS, EAP-TTLS, Kerberos, EAP-FAST

ANSWERS

1. The correct answer is A. Certificates can be used not only to ensure that the site is the company's Web site, but also that the user is an authorized user. The Web server can be configured to require client-side certificates. Answer B is incorrect because CHAP does not support two-way authentication in this manner. Answer C is incorrect because Kerberos can authenticate the user in a method similar to this, but could not serve to authenticate the server. Answer D is incorrect because tokens are used for one-way authentication.

2. The correct answer is C. The distinguished name is a unique identifier for the object, and is made up of several attributes of the object. It consists of the relative distinguished name, which is constructed from some attribute(s) of the object, followed by the distinguished name of the parent object. Answer A is incorrect because this identifies the root of the tree. Answer B is incorrect because this identifies the common name of the object. Answer D is incorrect because this is the user account's e-mail address.

3. The correct answer is C. The root server is used to create the structure of the directory, with organizational units and objects branching out from the root. Because LDAP directories are organized as tree structures, the

top of the hierarchy is called the root. Answer A is incorrect because the DIT is the name given to the tree structure. Answers B and D are incorrect because there is no such thing as a Branch server or Tree server in LDAP.

4. The correct answer is B. Supplicant is the client that wants to access a wireless network and authenticator performs the authentication. Answer A is incorrect in order; Answer C is incorrect as there is not negotiator in the process. Answer D is incorrect as Contact is not the right term used while defining authentication process.

5. The correct answer is A. EAP comes in several forms: EAP over IP (EAPoIP), Message Digest Algorithm/Challenge-Handshake Authentication Protocol (EAP-MD5-CHAP), EAP-TLS, EAP-TTLS, RADIUS and Cisco LEAP. The incorrect answers are B, C, and D. EAP-MPLS, Cisco PEAP, and Kerberos are not the EAP forms.

CHAPTER 9

Risk Assessment and Risk Mitigation

Exam objectives in this chapter:

- Conduct Risk Assessments and Implement Risk Mitigation
- Carry Out Vulnerability Assessments Using Common Tools
- Use Monitoring Tools on Systems and Networks

CONDUCT RISK ASSESSMENTS AND IMPLEMENT RISK MITIGATION

Risk assessments are a critical tool in ensuring clients and your internal needs are being met in regards to security. Often enough, laws and regulation will mandate you to have periodic risk assessments performed. The Health Insurance Portability and Accountability Act (HIPAA) is often a driving need, along with the Gramm–Leach–Bliley Act (GLBA). Conducting these risk assessments uncovers weaknesses within the IT infrastructure, procedures, policies, or business applications.

In many cases, risk assessments are carried out using vulnerability assessment tools. The concept is that vulnerabilities are directly related to overall risk. These tools do not consider the impact to risk of poor policies, standards, and procedures and are beyond the scope of this discussion.

Vulnerability assessment tools

The most common vulnerability assessment tool utilized is a *port scanner*, which is used to search a network for open ports. This is often the most basic of tools utilized by IT, security staff, and third-party organizations to review the security of their networks both internally and externally. Common port scanning software includes Nmap, Scanmetender, Supercan, and NHS Nohack Scanner.

VULNERABILITY SCANNERS

Vulnerability scanners are designed to map systems for weaknesses. They can often perform port scanning as well as checking for any applications that may

be running. Vulnerability scanners also run reports to show what information they can determine about the system, such as the operating system, service pack level, or applications installed.

PROTOCOL ANALYZERS

Protocol analyzers are a vital part of a network administrator's and security administrator's tool kit. Protocol analyzers can monitor the traffic on a network and expose data and protocols that are being passed along the wire.

OVAL

The *Open Vulnerability and Assessment Language*, or *OVAL*, is a language to determine the presence of vulnerabilities and confirmation of problems on a computer system. Prior to OVAL, there was no common means for system administrators to determine if software vulnerabilities existed or if patches were installed on local systems. The language standardizes the three main steps of assessments:

1. Representing configuration information systems for testing
2. Analyzing the system for the presence of the specified machine state (vulnerability, configuration, patch state, etc.)
3. Reporting the results of the assessment

Some of the benefits of using OVAL include:

- OVAL is restricted to publicly known configuration issues and vulnerabilities.
- OVAL definitions help users determine the presence of vulnerabilities or configuration issues on systems before they can be exploited.
- You must have root-level or system administrator access to employ the vulnerability information in an OVAL definition.
- The availability of the detailed technical information about vulnerabilities or configuration issues in OVAL definitions reduces the need for releasing exploit code to the public.

For more information see the main OVAL Web site at http://oval.mitre.org/

Password crackers

One aspect of a vulnerability assessment is to run a *password cracker* to test the strength of user and administrator passwords in a network environment. Password crackers range from the old L0phtCrack (originally a freely available tool that later became a commercial product) to the open source John the Ripper to Elcomsoft's Password Auditor to rainbow tables. For brevity, only rainbow tables will be discussed in this section.

RAINBOW TABLES

A *Rainbow table* is a lookup table to recover plaintext passwords from hash function output. In essence a rainbow table is a big one-to-one mapping of a password to the hash output that is stored in a password file or the Windows

security accounts manager (SAM) database. All possible password combinations have their hashes calculated and then stored in either a database or a simple text file, making lookups easy and eliminating the need for brute force calculations of passwords.

Network mapping tools

Network mapping tools are used by administrators to discover and ensure what devices are on their network. These tools enable an administrator to map out the network, discovering any new devices as well as any potential security holes due to unauthorized applications and services.

System and network scanning, when viewed from the context of a security system specialist or security administrator, is the use of appropriate technologies to detect and repair potential areas of vulnerability within a network or a system. This effort involves:

- Checking the strength of and compliance with password policies
- Measuring the ability to access networks from an outside or foreign network
- Providing analysis of known security vulnerabilities in network operating systems or hardware devices
- Testing a system's responses to various scenarios that could lead to Denial of Service (DoS) or other problems such as a system crash

There are a number of security scanning options available and a list of some of these can be found at http://sectools.org.

USE MONITORING TOOLS ON SYSTEMS AND NETWORKS

Many large networks employ some form of ongoing monitoring or diagnostic routine to continually keep administrators aware of the status of the network and allow for proactive corrective actions to potential problems. This can be done with monitoring software such as a network sniffer or with dedicated devices located on the network.

Crunch Time

Remember that sniffing a network is a passive attack but can provide a huge amount of information that can later be used for active attacks.

Workstations

The term *workstation* basically refers to any computer system that the end users of a network work on, assuming that the end users do not use servers for their

normal day-to-day work. The largest security concern in relation to workstations is the end user. End users always have local access (the ability to work at the local console) to their workstation, which can cause some big security problems, ranging from changing a password to something a hacker can easily guess, to inadvertently opening e-mails with viruses or Trojan horse applications running.

Another critical aspect of workstation security is to make sure that the operating systems or software applications always have the latest security patches in place. Often a vendor will release security patches that address individual vulnerabilities so technicians will be able to apply them faster, rather than having to wait for a full service pack.

Crunch Time

It is important to understand the differences between workstations and servers. You should know that workstations are typically used by a single local user and are designed to support fast front-end processing. Servers are designed to support a large number of remote users and provide fast back-end processing and file sharing.

INTRUSION DETECTION SYSTEMS

An *intrusion detection system* (*IDS*) is designed to monitor access points, hostile, and activities. These systems typically trigger on events by referencing network activity against an attack signature database or by monitoring network behavior. If an attack is detected (or believed to be detected), an alert takes place and the event is logged for future reference. Creating and maintaining the attack signature database is the most difficult part of working with IDS technology. It is important to always keep the IDS up to date with the latest signature database provided by the vendor as well as updating the database with the signatures found in testing.

EXAM WARNING

The Security+ exam expects you to understand the different types of IDSs, what they are used for, and how they can help protect your network.

LOGGING AND AUDITING

Auditing provides methods for tracking and logging activities on networks and systems, and links these activities to specific user accounts or sources of activity. In the case of simple mistakes or software failures, audit trails can be extremely useful in restoring data integrity. They are also a requirement for trusted systems to ensure that the activity of authorized individuals can be traced to their specific actions, and that those actions comply with defined policy. Audits also allow for a method of collecting evidence to support any investigation into improper or illegal activities.

Auditing systems

Auditing of systems must occur with a thorough understanding of the benefits of the process. To assist in catching mistakes and reducing the likelihood of fraudulent activities, the activities of a process should be split among several people. This process is much like the role based access control (RBAC) concepts discussed in Chapter 7. This segmentation of duties allows the next person in line to possibly correct problems simply because they are being viewed with fresh eyes.

From a security point of view, segmentation of duties requires the collusion of at least two people to perform any unauthorized activities. The following guidelines assist in assuring that the duties are split so as to offer no way other than collusion to perform invalid activities.

- **No access to sensitive combinations of capabilities.**
- **Prohibit conversion and concealment.** Segregation ensures that there is supervision for people who have access to assets.
- **The same person cannot both originate and approve transactions.**

These principles, whether manual or electronic, form the basis for why audit logs are retained. They also identify why people other than those performing the activities reported in the log should be the ones who analyze the data in the log file.

DID YOU KNOW?

One of the major problems with auditing is the simple fact that many network administrators do not have time to read and analyze the log files on a regular basis. Auditing provides us with the ability not only to provide a chronological path of access or attack, but also to spot trends of unauthorized activity so that they can be blocked before they do any damage. Unfortunately, many organizations do not devote the time to examine audit logs until after an attack. Good maintenance and procedures regarding the analysis of the log files will benefit your security efforts.

This may seem a daunting task when a large amount of log data is concerned. Tools have been developed that can help with this, such as Microsoft Log Parser or other free tools geared toward this purpose. By analyzing the log files for patterns or specific data, you can reduce the time required to review the log files. The difference between looking through logs line by line vs. scanning the logs for suspicious activity can be hours of time savings.

System Logs

System Logs are critical, as they will provide information on what is occurring on a system. One key concern is the amount of space the system logs take up. Log rotation prevents the situation of running out of space on the disk or partition that the system logs are written to. Old logs should be archived and retained in accordance with the logging and audit policy that is in place.

Performance Logs

Performance Logs provide insight as to how a system is performing. The performance of a system can be affected by operating system components, applications, as well as by other executables running on the system. Performance logs capture point-in-time statistics providing specific variables at specific points in time. This enables a system administrator to correlate system performance with specific events as they occur.

Access Logs

Access Logs require anyone entering a secure area to sign in before entering. When visitors require entry, such as when consultants or vendor support staff needs to perform work in a secure room, an employee of the firm must sign the person in. In doing so, the employee vouches for the credibility of the visitor, and takes responsibility for this person's actions. The access log also serves as a record of who entered certain areas of a building. Entries in the log can show the name of a visitor, the time this person entered and left a location, who signed them in, and the stated purpose of the visit.

AUDITS

The need for periodic audits is critical to the security of any organization. Periodic security audits are critical to a company's security. Audits reveal any security flaws or the need for the company to update any standards and technical configurations. Third-party audits are also recommended on a periodic basis depending on many factors, including internal company policy as well as external regulatory requirements.

SUMMARY OF EXAM OBJECTIVES

Risk assessment and mitigation are important parts of the overall security process in an organization or enterprise. Risk assessments are typically carried out using a variety of tools such as vulnerability scanners, password crackers, and protocol analyzers. In addition, the results of the risk assessments should be reviewed in order to be able to identify critical risks as well as mitigation techniques that could reduce the overall exposure of the organization.

Logging and auditing play critical roles in securing an organization's environment. System, performance, and access logs all provide key pieces of information regarding the day-to-day security of a network. Periodic auditing of logs such as these is critical to identify any flaws or the need to update standards and technical configurations.

TOP FIVE TOUGHEST QUESTIONS

1. You are the security officer of a company, and you have been asked to implement an employee security program. Where would you start?

   A. Security scan
   B. Security policy

C. Security audit
D. By locking down access for everyone

2. Vulnerability scanners are designed to:

A. Map systems for weaknesses.
B. Monitor the traffic on a network and expose data and protocols that are being passed along the wire.
C. Never attempt to exploit a known vulnerability.
D. Packet sniffing.

3. You have identified a number of risks to which your company's assets are exposed, and you want to implement policies, procedures, and various security measures. In doing so, what will be your objective?

A. Eliminate every threat that may affect the business.
B. Manage the risks so that the problems resulting from them will be minimized.
C. Implement as many security measures as possible to address every risk that an asset may be exposed to.
D. Ignore as many risks as possible to keep costs down.

4. You have decided that you are going to have an audit performed within your organization. What are the things not to consider?

A. External regulatory requirements
B. Your last external audit
C. Internal policies
D. Change control procedures

5. What is the goal of a risk assessment?

A. To test the basic strength of your systems and create a report for your executive team.
B. To test everything possible and create a report for your executive team.
C. To test everything possible and create a report that will be read by your management and customers, showing what was performed, what was discovered, and how issues were addressed.
D. To test everything possible and create a report that shows you have no issues and will be read by your management and customers.

ANSWERS

1. The correct answer is B. Incorrect answers: A, C, and D. The first step is to develop a security policy from which a security scan can be conducted (Answer A) and then an audit of the findings of the scan (Answer C). From this audit, the need for an access lock down can be identified (Answer D).

2. The correct answer is A. Incorrect answers: B, C, and D. Vulnerability scanners are not designed to monitor traffic on a network (Answer B) as that is the role of a packet sniffer. Vulnerability scanners, by definition, do not sniff packets on a network wire (Answer D), and they typically do attempt to exploit a known vulnerability (Answer C) in order to validate the response from the system and verify that a vulnerability exists.

3. The correct answer is B. Incorrect answers: A, C, and D. The objective of any security administrator is to manage the risks to the business and to minimize their impact. Eliminating every threat (Answer A) is impossible and implementing as many security measures as possible (Answer C) will be cost and administratively prohibitive. Ignoring risks (Answer D) is the surest way for a disaster to occur.

4. The correct answer is B. Incorrect answers: A, C, and D. When you are conducting an external audit, items such as external regulatory requirements (Answer A), internal policies (Answer C), and change control procedures (Answer D) are all items to consider. The last external audit should have no bearing on the current audit but provides a baseline with which to compare the current results.

5. The correct answer is C. The goal of a risk assessment is not to test the strength of your systems (Answer A), and the report is not meant for the executive team (Answer B) but rather management and, secondarily, customers. A risk assessment is not meant to show that there are no issues (Answer D) but rather what was found and any issues that were identified if there were any.

CHAPTER 10

General Cryptographic Concepts

Exam objectives in this chapter:

- General Cryptography
- Encryption Algorithms
- Protocols
- Cryptography in Operating Systems

GENERAL CRYPTOGRAPHY

Locks and keys have been used for centuries to keep items—and communications—hidden. In many locks in the physical world, it is possible to close the lock without having the key (or the code, for a combination lock). Other locks require a key to lock them, that key also being required to unlock them.

There are analogous mechanisms in the online world, except that the terminology is different. Locking and unlocking are not common terms in cryptography—instead, "*encrypting*" data turns it from readable into unreadable, and "*decrypting*" data changes it from unreadable to readable. Locks are now "algorithms" or methods by which encryption is done, and keys, thankfully are still represented by the term "keys." The two kinds of encryption (same key, different key) are referred to as symmetric and asymmetric, respectively.

Symmetric key cryptography

"Symmetric" means that two sides are in balance, or equal. *Symmetric key cryptography* is the use of one key to both encrypt and decrypt, and the encryption algorithm is sometimes the same as the decryption algorithm.

A simple example is the "Caesar" cipher, in which letters are shifted by a number of characters. The number of the shift is the key, and the operation is a shift to the right to encrypt, and a shift to the left to decrypt. Many children used code wheels to create their own Caesar cipher texts, and challenged others to break their codes.

For instance, if the message we wanted to secretly communicate (the "plain text") was "THERE ARE THREE PEOPLE TO SEE CAESAR," and the shift was 8, the letter T would be replaced by B (when a shift takes us to Z, we wrap around to A), H becomes P, etc., to give the encrypted message, known as "cipher text," as follows:

THERE ARE THREE PEOPLE TO SEE CAESAR

BPMZM IZM BPZMM XMWXTM BW AMM KIMBIZ

The big problem that faces any symmetric key cipher is that their simplest use requires that the key be shared—a "shared secret"—between the two parties in the cipher. In some senses, that is to be expected, because the encrypted text is itself a secret that is shared between the sender and recipient. However, if the sender has multiple recipients, he/she will need to generate multiple keys to be sure that each recipient sees only those messages meant for him.

Asymmetric key cryptography

Clever as the solutions to key exchange may be, there are still some problems to overcome with symmetric key cryptography—the most notable is that the sender has no way to verify the identity of the recipient. Asymmetric key cryptography helps to solve that issue.

A British invention of the 1970s, "non-secret encryption" provided for asymmetric key cryptography, in which a pair of unequal (hence, "asymmetric") keys is created. (The British government held the discovery of "non-secret encryption" as a state secret until 1997, so it is fortunate that American researchers later independently discovered and published their own public key algorithms.) For each pair of keys in asymmetric cryptography, one key is held privately, the other is published. They are referred to as the "*private key*" and the "*public key*," respectively. The public key can be given out freely without compromising the private key at all. A sender can encrypt a message using the receiver's public key, and be sure that it can only be decrypted using the related private key—which means that the encrypted message can then only be read by the holder of the private key. As long as the public key is trusted to be associated with the intended recipient, and the private key has not been exposed, only the intended recipient will be able to decrypt the message.

Asymmetric key cryptography is computationally expensive. This is mostly because of the size of the keys involved, which have to be much larger than keys providing similar protection strength for symmetric cryptography.

Hashes and applications

In cryptography, a *hash* is when a piece of plain text is sliced, diced, and otherwise rendered into a small "digest" or "hash" that bears no obvious resemblance to the original text, except that the same text will always produce the same hash.

Hash algorithms are based on mathematical "one-way functions"—these are functions that are relatively easy to calculate going forward, but the inverse of the function is such a complex procedure that it is significantly harder to reverse the function than it would be to simply try every single possible input against the function to try and match its result.

There are a couple of reasons to generate hashes:

- A hash can serve as a check that a document is untampered.
- A cryptographic hash may serve as a placeholder for a document.
- A cryptographic hash can be used to verify a piece of information that is too sensitive to store.

Cryptographic hashes are required to possess a few specific properties in order to be useful:

- It must be practically impossible to reconstruct the original data from the hash.
- Two similar documents must produce vastly different hashes.
- It must be practically impossible to construct two pieces of data that generate the same hash.
- Computing the hash must be a quick process.

Care should be taken when using a hash for calculating a digest of small pieces of data—for passwords or credit cards, for instance. An attacker with access to the hash codes may be able to mount a bulk attack against the entire database at once, unless appropriate care is taken. Typically, the addition of a random component, known as a "salt," to each piece of data being hashed will protect against this kind of bulk attack—that way, two users with the same password will not have the same hash. Storing the salt with the hashed value is necessary to ensure that the hash can be regenerated when checking the hash.

Here are some definitions that you should know:

- ***SHA1***: SHA is an abbreviation for "Secure Hash Algorithm" and is the name for a series of algorithms that each were selected by the NIST—National Institute of Standards and Technology—to provide standardized cryptographic hash functions for widespread public use.
  - SHA-0 and SHA-1 each produce a 160-bit "digest" from any input message up to $2^{64}-1$ bits in length.
  - SHA-2 is a family of hash functions that provide output digests in a number of different lengths—224, 256, 384, and 512.
  - SHA-3 has not yet been selected at the time of writing, but will be chosen from a public competition between entries submitted by a number of cryptographers. The winner is scheduled to be announced in 2012.

- ***MD5***: "MD5" means "Message Digest 5"—there are others from its stable, particularly MD2 and MD4. All of these Message Digest algorithms are known to have collision-resistance flaws that could result in the creation of two documents that have the same hash. This has been demonstrated in a number of interesting ways, including the creation of two X.509 certificates, one benign, and the other quite definitely malign in nature. It is strongly recommended that instead of MD5, a different hashing algorithm such as any of the SHA-2 family is used for new cryptographic designs.
- ***LANMAN***: The LAN Manager (LANMAN, or LM), Hash is frequently used as an example of a hash that has long outlived its ability to protect against attack, but has survived in use simply because there are so many old systems that use it, and it is therefore considered by some to be too risky to turn off, because applications that still use it may break. Such applications include older versions of Windows, and non-Windows implementations of the Server Message Block protocol used for network file sharing.
- ***NTLM***: The replacement of the original LanManager Hash function was achieved with the *NTLM* protocol, which provides the ability to authenticate a user against either the LM Hash or the new "NT Hash." Where possible, disabling the LM Hash at the Domain Controller means that the NTLM protocol will only use the NT Hash to verify authentication attempts.

  The NT Hash is an MD4 hash of the user's password—it is a true hash, and it uses the full Unicode character set available to Windows, whereas the LM Hash used only a limited subset of the US-ASCII character set. As such, it is less susceptible to brute-force guessing, or bulk cracking, if the passwords are strong. The NT Hash is still not salted, and as a result it is possible to tell if two users have the same password, because they will have the same hash.

Digital signatures

Digital signatures represent a combination of cryptographic hashes and asymmetric encryption.

One of the most frequently asked questions by novice developers of cryptographic applications is how to encrypt using the private key, or decrypt using the public key—both of these operations are often forbidden by cryptographic frameworks. The reason is to prevent applications from accidentally using keys inappropriately and thereby posting freely decryptable text across a network.

Encrypting data with a private key does not protect it against interception—a public key is supposed to be public, and so it must be assumed that your attacker has a copy of it. So any data encrypted with a private key must be either public already, or protected in some other way. By encrypting data with a private key, however, it is clear that the file was encrypted by owner of the private key.

A digital signature is created by creating a cryptographic hash of the document to be signed and then encrypting the hash with the private key of the signer. This has several benefits—a single document can be signed by multiple parties; signature of any document is fast; a signature can be sent or held separately from a document or along with it.

Certificates

A *digital certificate* lists a set of claimed qualities about the person, organization, or computer identified in the certificate, and confirms those claims by the presence of a digital signature of a reputable body. It ties those claims to a public/private key pair, so that the user of the private key can be verified as the subject of the claims in the certificate. The subject of a certificate, the person, organization, or computer about which the certificate holds details, is known as the Subject. The reputable body that signs the certificate is known as the Issuer.

The standard for digital certificates is ITU-T X.509. There are a number of different versions of this certificate standard—at the time, of writing version 1 up to version 3, each adding features over the last. X.509 is a standard for Public Key Infrastructure in general, and covers other topics such as certificate revocation lists, and certificate path validation rules.

SINGLE

Most certificates in current use are what the CompTIA Security+ Objective refers to as *single certificates*—single in the sense that they are independent of any certificates other than their Issuer's.

The word "single" does not imply single use—these certificates may be used for a single purpose, or multiple purposes, indicated by values known as "Key Usage" and "Enhanced Key Usage" values. The "Key Usage" value is a set of bits that can be on or off.

The Enhanced Key Usage values are stored as hierarchical numeric values—Object Identifiers, or OIDs—in a format called ASN.1.

DUAL-SIDED

Dual-sided certificate is a term encountered only in the CompTIA Security+ Objectives and documents derived from them. The more usual term for this is a *dual key pair*, or *dual key certificates*. Two key pairs, and two certificates, are generated. One certificate and its related key pair is used for encryption, the other is used for data signing (and non-repudiation) purposes. The key pair used for signing is generated and held by the user, and is not stored in any kind of key management system outside of the user's control. The key pair used for encryption may be backed up in a key management system for later recovery.

CIA—For all your security needs

Practitioners of information security rely on a number of different taxonomies in order to ensure that they cover their system's security needs entirely. One of the simplest such groupings available is "CIA"—confidentiality, integrity, and availability. A security solution should generally be assessed as to how well it fits each, and all, of these three categories.

CONFIDENTIALITY

Confidentiality addresses a system's ability to keep information hidden from those people, systems, and processes that are not meant to see it. Encryption is the process of taking a readable document, and producing from it a document that is unreadable unless you possess a proper key to decrypt it.

INTEGRITY

Integrity addresses a system's ability to ensure—and to prove—that information being processed is the result of the application of approved processes to the original data. What this means is that data cannot be modified without approval, or where it is modified, that such modifications are detected.

AVAILABILITY

Availability addresses a system's ability to be present and to provide data to its approved users.

It is often considered to be the reverse side of confidentiality—if a document is encrypted and then the key is thrown away, the document is very definitely confidential, but it is not available anymore.

Non-repudiation

Non-repudiation is a component of information security that tries to remove the possibility that someone could perform an action and then later claim that was not truly them, but someone using their name without permission. A digital signature can be used as proof that the owner of the key associated with the signature approved of, or at least was aware of, a document or an action.

Key management

To maintain availability it is important to consider how to handle the inadvertent loss of a key, or how to prevent it, and how to manage a key's life cycle. *Key management* includes determining whether it is acceptable for anyone other than the key holder to be able to have access to the key. While the initial reaction is to say "no," this ignores the possibility that the key holder may deliberately or accidentally destroy the key. If a key's purpose is to identify the user, pointing to use of the key is proof of the user's involvement, then the answer is quite definitely that no one other than the user should have access to that key. Such a key should be stored only in the users' private certificate

store, ideally on a smart card or other hardware device subject to anti-tampering protection. The life cycle of such a key follows from creation through use, renewal, and finally to either revocation or expiration. It is the simplest of key life cycles.

ENCRYPTION ALGORITHMS

A number of encryption algorithms are available for use, and more are created over time—this list is by no means complete, but represents some of the more important algorithms at the time of writing—and covers the requirement for the CompTIA Security+ Objectives.

DES

Standing for *Data Encryption Standard*, this is another of those generic names that indicates it comes out of NIST. The DES algorithm uses a 56-bit key, and as expected from something with such a small key size, it is a symmetric key encryption algorithm—asymmetric keys are usually more than a thousand bits in length. It is also a "block" encryption algorithm, meaning that it encrypts in blocks—in DES's case, a block is 64 bits—one block at a time. While block ciphers are not in themselves designed to encrypt streams of data, it is possible to use what is known as a "mode of operation" to encrypt a stream using the block cipher.

3DES

3DES is an abbreviation that is generally read as "triple DES"—as the name implies, it's an algorithm built from three applications of the DES algorithm. Rather than the obvious process of running the DES encryption three times, 3DES first encrypts using DES and the first key, then decrypts using the second key, and finally encrypts using the third key. This method was chosen in large part so that a hardware implementation of 3DES could be used to also implement DES by setting all three keys to the same 56-bit value.

RSA

RSA, named after Rivest, Shamir, and Adleman, its inventors, is the name of a company whose focus is on public key infrastructure and cryptography, as well as the name of an asymmetric cryptography algorithm, which is why we encounter this abbreviation in this section on cryptographic algorithms.

Since the RSA algorithm relies on mathematical operations—particularly exponentiation—it is possible to apply it to any size of input. The RSA algorithm can be quite slow and, as such, the RSA algorithm—and most asymmetric key cryptographic algorithms—will generally be used only to encrypt an exchange of a suitable symmetric key for a stream or block cipher to be used for bulk encryption.

AES

The *Advanced Encryption Standard* (*AES*) was the name NIST gave to the winner of a competition to develop a good encryption algorithm to replace DES. It was previously called Rijndael, an amalgam of the names of its two designers, Joan Daemen and Vincent Rijmen. The AES cipher is actually a specialization of the Rijndael cipher, as the AES cipher has a block size of 128 bits, whereas the Rijndael cipher can have any block or key size from the selection of 128, 160, 192, 224, and 256 bits. The AES cipher supports key sizes of 128, 192, or 256 bits. The cipher is often known by its name and the number of bits of key—for instance, AES-128 and AES-256 are often supported ciphers, and refer to AES with 128-bit and 256-bit keys, respectively.

Elliptic curve cryptography

Elliptic curve cryptography is another mathematically based technique for cryptographic operations, rather than being based in bitwise logic, and like RSA, can essentially be used with any size of key. The bit-strength of Elliptic Curve encryption is theorized to be roughly half the size of the key—so a 256-bit ECC key has strength of about 128 bits, compared to an RSA key for the same strength, which would need to be 3,072 bits in length.

One-time pads

One-time pads are the perfect encryption method—and have been mathematically proven to be so. They are, unfortunately, very impractical. The way a one-time pad works is that a stream of random characters (the one-time pad) is generated, and distributed securely between the sender and recipient. The stream must be of at least the same size as the stream to be encrypted. When it is time to send the encrypted traffic, the plain text is combined with the stream of random characters, usually using a simple XOR combination, to generate cipher text. At the recipient's end, the cipher text is decrypted by reversing the process.

Transmission encryption

General-purpose encryption algorithms are often adapted to fit specific uses. Two such protocols that the CompTIA Security+ Objectives specifically call out are WEP and TKIP, both invented to protect IEEE 802.11 wireless network traffic (aka "WiFi").

EXAM WARNING

WEP has long since been discredited as a valid cryptographic security system. It is not going to be covered in great detail in the Security+ exam. It is provided here for reference and quick review.

WEP

WEP is particularly worth knowing about, if for no other reason than it is an object lesson in the old maxim "don't write your own cryptography unless you are a cryptography expert," along with its corollary, "you are not a cryptography expert."

TKIP

While WEP was being broken by attackers, the WiFi Alliance approved a subsequent protocol, *TKIP*—the Temporal Key Integrity Protocol. TKIP was approved as a part of the Wi-Fi Protected Access (WPA) protocol.

TKIP uses RC4 as well, but has several advantages over WEP—most notably, each data packet is encrypted using a different key, and instead of merely concatenating the IV and the key, TKIP combines them using a key mixing function. TKIP also uses a sequence counter, so that replay attacks fail, as the sequence counter is different when the replay attack is attempted.

The final result of the WEP debacle was the adoption of WPA and the follow-up, WPA2.

PROTOCOLS

Cryptographic algorithms are ephemeral—their useful life span is measured in years, rarely decades. Because of the constantly changing list of reliable cryptographic algorithms and also because it makes good engineering sense, the first thing that needs to be done when deciding to use cryptography is to agree between sender and recipient on a choice of cryptographic algorithm, and to exchange keying information where necessary. This process of agreement prior to exchange is defined in a "protocol" specification—just as in real life, a protocol is the set of agreed upon rules by which work is done, rather than the doing of the work itself.

SSL/TLS

Secure Socket Layer (*SSL*) and *Transport Layer Security* (*TLS*) refer to essentially the same protocol, but in different versions. SSL was originally invented by Netscape, the makers of the Web browser Netscape Navigator, as a means to allow credit card transactions to be carried out securely over the World Wide Web. SSLv2 had some flaws that needed correcting, and while Netscape produced SSL 3 to correct these flaws, Microsoft developed their own Private Communication Technology (PCT) standard, which also corrected the flaws in SSLv3.

In an effort to achieve a harmony of standards, as well as to address the flaws in SSLv3, Netscape allowed for automatic and royalty-free licensing of the SSLv3 protocol, and worked with members of the Internet Engineering Task Force (IETF) to produce a new unified standard based on SSLv3, and with lessons learned from SSL and PCT.

This new standard is called *Transport Layer Security*, or *TLS*, and although it is a new name, it is functionally a logical development from SSL. In fact, the version number embedded in a TLS 1.0 stream is "3.1," essentially declaring that a TLS-capable client is actually an SSL 3.1 client.

TLS was designed as an extra layer on top of TCP/IP, but underneath an application, so as to make it easy to add TLS to an existing application. This means that TLS is divided into three sections—negotiation, communication, and closure. Four sections, if you count the inclusion of error information.

TLS is not completely a "black box" addition—there are some subtleties to developing an SSL/TLS compliant program.

HTTP vs. HTTPS vs. SHTTP

The usual use case for TLS still remains the one for which it was originally designed—that of protecting World Wide Web transactions over HTTP. HTTP itself is a text-based protocol, which makes debugging and analysis by humans easy, but also makes theft of data in transit by humans and machines alike even easier.

HTTPS is the most naïve possible implementation of TLS with HTTP, and served for many years as the model for how TLS should be used. A new port was reserved—443 for HTTPS, vs. port 80 for HTTP—and all connections to that port would be considered to be using TLS from the moment they connected (and therefore, starting with a Server Hello message) until such time as the HTTPS connection was terminated, leading to the implicit closure of the TLS connection.

S-HTTP, by contrast, does not use TLS at all, but instead treats each HTTP request and response as an individual message to be protected using the Cryptographic Messaging Syntax (CMS) encapsulation protocol. This allows each request and response to be signed or encrypted, according to options specified on each resource available from the Web server.

Other protocols with TLS

Many other TCP-based protocols have added options to use TLS to protect their content over the years—SMTP, LDAP, and NNTP use a STARTTLS command to indicate a request from the client to use TLS, and FTP uses an "AUTH TLS" command with options to indicate whether commands, data, or both are clear or encrypted.

S/MIME

S/MIME offers signing and encryption of e-mail messages in a similar way to the S-HTTP protocol—in S/MIME, the components being signed or encrypted are the MIME parts of the message. Again, CMS is used with some minor modifications to fit the CMS protocol in with MIME. By using detached signatures, where one MIME part containing clear text is signed to provide a separate MIME

part containing its signature, S/MIME signed messages can be sent without worrying about whether or not the recipient is capable of displaying or interpreting S/MIME messages. A recipient who is unable to process the S/MIME part will simply see the text message, without the ability to verify the signature on the message.

SSH

SSH stands for *Secure Shell*, which was invented in 1995 in an attempt to prevent password-sniffing attacks on the rlogin and telnet protocols commonly used for shell (console) access to UNIX systems.

Since then, SSH has developed into more than just a shell access tool, and it operates as a secured network layer, equivalent in many ways to SSL and TLS. However, there are some significant differences that are worth commenting on:

- SSH essentially provides its own transport layer on top of TCP.
- SSH authenticates using public and private keys.
- SSH is generally provided along with a small suite of applications.
- SSH is popular among Linux/UNIX users and mostly spreads out from the community of Open Source admirers. Microsoft Windows does not natively support SSH, nor does Microsoft offer any SSH implementations or toolkits, which can make it difficult to ask for your partners to use SSH-based protocols if they are Microsoft-heavy environments.
- While SSH has recently become an IETF documented standard with RFC documents, some aspects are still not standardized or fully documented—particularly SFTP.

SSH and SSL are often compared, but while they both achieve the aim of secure communications protected by public key authentication, they do so in different ways. When assessing the need to protect communications, both options should be considered, depending on who you are connecting to and what the communications will focus on.

IPSec

In contrast to SSH and SSL, *IPSec* works below the transport layer and above the IP packet layer. By adding IPSec rules to connections, it is possible to allow any application that uses an IP-based protocol to take advantage of the security offered by IPSec. IPSec provides two security mechanisms for general data—Authentication Header (AH) and Encapsulating Security Payload (ESP). The other part of IPSec, Internet Key Exchange (IKE), operates before communication protected by AH or ESP to establish a Security Association (SA) between two hosts, negotiating between those hosts to authenticate, as well as to agree on encryption methods and keys.

IKE operates over UDP, on port 500, and negotiates using a number of different methods. IKE can authenticate parties and establish encryption keys by using public keys or by using a pre-shared key (PSK).

Each IKE negotiation results in two SAs, one inbound and one outbound, at each host. Obviously, one host's inbound SA will match the other host's outbound SA, and vice versa. The SA consists of an IP address, a Security Parameters Index (SPI), and the key associated with the SA. The SPI is simply a random number generated by the host that created its associated key, and along with the IP address of that host, can be viewed as an index into the database of SAs.

AH operates as IP protocol number 51 (thus, it is neither UDP nor TCP, and does not have an associated port number), and inserts a header into each protected data packet containing the SPI of the negotiated SA to which the packet is associated, a Sequence Number to prevent replay attacks, and an Integrity Check Value (ICV), which is generally a keyed MAC of the AH header (excluding the ICV) and any data following it. This allows each packet to be verified independently of any other packets (other than the key exchange performed by IKE).

ESP operates as IP protocol number 50, and its IP header contains the SPI of the connection, a Sequence Number to prevent replay attacks, encrypted payload data (the IP packet that has been encrypted), encrypted padding to align the payload data with block sizes for block ciphers, an encrypted Next Header value, and an ICV just as in AH protocol. The Next Header value refers to the header inside the payload data, rather than a header following the ESP header—there usually is no such following header, but the decrypted payload data can be considered to be the logically following data.

Comparing IPSec to the protocols previously discusses it is clear that there are some advantages and some disadvantages.

- IPSec authenticates hosts to one another, and cannot authenticate users.
- IPSec protects any application, without that application being aware of its being protected.
- IPSec requires that routers accept and pass protocols 50 and 51.
- IPSec builds its ICV, and creates its SA, over values that include the IP address of each host—this makes it difficult to use across a NAT (Network Access Translation) router. To deal with this, an encapsulation known as NAT-T (NAT Tunneling) was developed.
- IPSec ESP can operate either in Transport mode, in which it uses addresses on the local network, or in Tunnel mode, in which the source and destination IP addresses inside are from different physical—and logical—networks than those that are carrying the outer packets.

PPTP

PPTP is the least fully featured or secure by itself. PPTP stands for the Point-to-Point Tunneling Protocol. Described in RFC 2637, it is a relatively simple encapsulation of PPP (the Point-to-Point Protocol) over an existing TCP/IP

connection. It consists of two connections (perhaps more in multilink environments, although this is less common today):

- The control connection is a TCP connection to port 1723.
- The IP tunnel connection is carried over the GRE (Generic Routing Encapsulation) protocol, carrying the user's data itself.

Because PPTP is so simple, it is frequently implemented, even in non-Microsoft operating systems such as Mac OS X and Linux. As a simple protocol, it is ideal for small low-power devices, such as mobile phones and PDAs. PPTP's biggest hindrance is that it uses a protocol (GRE) other than TCP or UDP, which may be blocked at firewalls, NATs, and routers.

L2TP

L2TP is the *Layer 2 Tunneling Protocol*, and was defined originally in RFC 2661, with the current version, L2TPv3, defined in RFC 3931. The name refers to the fact that Layer 2 (the same layer as Ethernet) traffic is tunneled over UDP, a layer 4 protocol. Unlike PPTP, L2TP uses one data stream only, on UDP port 1701. L2TP packets are divided between control and data by a flag in the header. Because L2TP operates over UDP, it has to implement its own acknowledgement and retransmission mechanisms for the control messages it uses.

L2TP's main usability benefit comes in its use of a single pseudo-connection over a protocol that is forwarded by most routers, UDP. L2TP's biggest security benefit also comes from the use of a well-defined protocol, IPSec.

CRYPTOGRAPHY IN OPERATING SYSTEMS

With increased use of laptops, third-party data centers, and hard drive backup storage, it has become clear that protecting data "in flight" should include protecting that data in storage on the operating system.

Over the years, this has seen a number of developments in encryption technologies, leading to the widespread deployment of cryptography within most enterprises.

File and folder encryption

The first application of encryption to stored data was that of file and folder encryption. In many cases, this encryption is carried out in a rather manual and ad hoc manner—running an application to encrypt a file, then running the companion decryption application when the file was needed to be edited. For folders, often an archiving utility was used that would compress and store a folder (or a set of files) into the archive file, which was encrypted. The *Zip* format is one example of an archive format that supported encryption of its compressed archives, initially using RC4 encryption, and in current incarnations using the AES algorithm with 128 or 256 bits.

With Windows 2000, Microsoft added a more automated approach to file and folder encryption, by adding its *Encrypting File System* (*EFS*) to the file system used in Windows—NTFS (New Technology File System). EFS has continually been added to in subsequent releases of the operating system, with new encryption algorithms such as AES replacing the original DES and 3DES, and new options for encryption and recovery of encrypted files.

E-mail

E-mail encryption is generally performed in one of the two ways—either by encrypting the network connection or by encrypting the message itself using a protocol such as S/MIME.

Encrypting the connection between mail client and mail server—or more properly, the MUA (Mail User Agent) and the MTA (Mail Transport Agent)—is particularly useful when the authentication method chosen is that of a simple username and password. By encrypting the connection itself, a user makes it impossible for anyone listening in on the network stream to read his or her password.

Encrypting and/or signing the message itself allows for that message to be encrypted and signed while it sits in storage, allowing the message to remain protected for the foreseeable life of the message. Again, as with file and folder encryption, it is possible to copy data out of an encrypted message, and save it to an unencrypted format. Preventing this copying of encrypted content to a plain text version would require the use of a Digital Rights Management (DRM) technology.

Whole disk encryption

Laptops are an easy target for thieves, as they are quick to steal and often have significant monetary value not only in themselves as a free computer but also in the data they contain, since personal details and bank information can raise a high price on the black market. To reduce the cost of this threat to their organization's data, most enterprises have taken to requiring encryption of their laptops' hard drives. Rather than rely on EFS and a user's individual preferences as to which documents are secret and which should be publicly available, these enterprises choose to encrypt the entire hard drive as a complete unit.

Once a disk has been encrypted, the overhead of encrypting and decrypting data is minimal—typically much less than 10%. This is because the encryption or decryption of one block can be done while waiting for the next block to be read from, or written to, the disk.

Trusted platform module

The *Trusted Platform Module* (*TPM*) is a relatively recent addition to a computer's arsenal, and is a specialized, tamper-resistant, hardware device designed to

engage in a few simple cryptographic operations. As with most cryptographic systems, the TPM has a single root key, called the *Storage Root Key* (*SRK*), which is strongly protected inside of the TPM, and is used to protect all the other keys the TPM device handles. In addition to the SRK, the operating system can request a number of other key pairs to be generated and encrypted (or "wrapped") with the SRK, such that the private key is only available inside the TPM for decryption or signing operations.

In addition to wrapping a key, the TPM can "seal" the key, such that it can only be used in the event that a number of system measurements (selected at the time of sealing) are the same that they were when the key was sealed. These system measurements include the BIOS code and settings stored in the computer's firmware, as well as the boot sector of the disk.

SUMMARY OF EXAM OBJECTIVES

Cryptography is a very complex and involved topic that plays an exceedingly important role in modern computer and network security. Symmetric and asymmetric cryptography helps provide the basis of confidentiality in network communications and data, while hashes provide integrity verification. The complex protocols used in network communications ranges from SSL/TLS to PPTP/L2TP to IPSec. Each has its place in the overall scheme of network and data protection and each continues to evolve to meet the new threats that are constantly arising.

TOP FIVE TOUGHEST QUESTIONS

1. What cryptographic properties should a strong symmetric cipher have?

   A. The number of bits in the key should be large, so as to discourage brute-force cracking.
   B. Encryption should be slow, so as to discourage brute-force cracking.
   C. Bits in the cipher text should never be the same value as the corresponding bit in the plain text.
   D. The same plain text should always generate the same cipher text.
   E. The cipher should prevent the use of keys chosen by poor random number generators.

2. What technique improves the protection given by a cryptographic hash of small data?

   A. Signing the hash with a private key.
   B. Padding the data with null bytes to match the block size of the hash algorithm.
   C. Prefixing the small data with a random value prior to hashing it.
   D. Repeating the data two or more times.

3. Which encoding should be used for exporting X.509 digital certificate information in the most portable way possible?

   A. DES encryption, followed by base 64 encoding
   B. PKCS 12/PFX
   C. DER text encoding
   D. Dumping the contents of an SSL key exchange

4. What IPsec protocol offers authentication verification and data integrity protection? (Choose one or more answers.)

   A. AH
   B. ESP
   C. L2TP /IPSec
   D. IKE

5. What is the difference between TPM "wrap" and "seal"?

   A. The "wrap" operation will allow a key to be used at any time; the "seal" operation will allow a key to be used only when system measurements match those present at the time of sealing the key.
   B. The "wrap" operation allows a key to be revealed if the system measurements match those at its creation; the "seal" operation never allows a key to be revealed, but may allow it to be used.
   C. The "wrap" operation uses symmetric cryptography keys; the "seal" operation uses asymmetric keys.
   D. The "seal" operation is designed not to leak the key, the "wrap" operation may leak the key under some attacks.

ANSWERS

1. Answers A and D are correct. Answer B is incorrect as symmetric ciphers tend to be relatively fast. Answer C is incorrect as that is not a cryptographic property determining whether a cipher is strong. Answer E is incorrect as the cipher has no control over which keys are used.

2. Answer C is correct and is known as adding a salt to the data. Answer A will not improve the protection of the cryptographic hash but does create a digital signature. Answer B is incorrect as null byte padding to match the algorithm block size is already done in order to create the hash of the data. Answer D is incorrect, does nothing to the output hash.

3. Answer C is correct. Answer A is incorrect as DES encryption does not specify an encoding format. Answer B is incorrect as PKCS 12 is used when providing the private key as well and is not the most portable way possible.

4. Answer B is correct. Answer A is incorrect because AH only provides authentication verification. Answer C is not an IPSec protocol, and

Answer D is the key exchange that takes place to set up an IPSec security association and does not provide data integrity protection.

5. Answer A is correct as that is the actual operation of the wrap and seal operations. Answer B is incorrect as it is the exact opposite of the "wrap" and "seal" operations. Answers C and D are incorrect as the "wrap" and "seal" operations are not concerned with whether symmetric or asymmetric keys and both operations are designed not to leak the key.

CHAPTER 11

Public Key Infrastructure

Exam objectives in this chapter:

- PKI Overview
- Components of PKI
- Registration
- Recovery Agents
- Implementation
- Certificate Management

PKI OVERVIEW

Cryptographic technologies such as *public key infrastructure* (*PKI*) provide a way to identify both users and servers during network use. The primary function of the PKI is to address the need for privacy throughout a network. For the administrator, there are many areas that need to be secured. A few examples include:

- Internal and external authentication
- Encryption of stored and transmitted files
- E-mail privacy

PKI is the underlying cryptography system that enables users or computers that have never been in trusted communication before to validate themselves by referencing an association to a trusted third party (TTP). Once this verification is complete, the users and computers can securely send messages, receive messages, and engage in transactions that include the interchange of data. PKI is used in both private networks (intranets) and on the World Wide Web (the Internet).

DID YOU KNOW?

Cryptography refers to the process of encrypting data; *cryptanalysis* is the process of decrypting, or "cracking" cryptographic code. Together, the two make up the science of cryptology.

PKI encryption

Ciphering text has always played an important role in wars and politics. As modern times provided new communication methods, scrambling information became increasingly more important. World War II brought about the first use of the computer in the cracking of Germany's Enigma code. In 1952, President Truman created the National Security Agency at Fort Meade, Maryland. This agency, which is the center of US cryptographic activity, fulfills two important national functions: It protects all military and executive communication from being intercepted and it intercepts and unscrambles messages sent by other countries.

There are three types of cryptographic functions:

- The hash function
- The secret key method of encryption
- Public/private key cryptography

There are basically two types of symmetric algorithms:

- **Block symmetric algorithms** work by taking a given length of bits known as blocks.
- **Stream symmetric algorithms** operate on a single bit at a time.

PKI standards

The *Public-Key Cryptography Standards* (*PKCS*) are a set of standard protocols issued for securing the exchange of information through PKI. The list of these standards was actually established by RSA laboratories—the same organization that developed the original RSA encryption standard—along with a group of participating technology leaders that included Microsoft, Sun, and Apple.

Here is a list of active PKCS standards:

- **PKCS #1: RSA Cryptography Standard**
- **PKCS #3: Diffie–Hellman Key Agreement Standard**
- **PKCS #5: Password-based Cryptography Standard**
- **PKCS #6: Extended-Certificate Syntax Standard**
- **PKCS #7: Cryptographic Message Syntax Standard**
- **PKCS #8: Private-Key Information Syntax Standard**
- **PKCS #9: Selected Attribute Types**
- **PKCS #10: Certification Request Syntax Standard**
- **PKCS #11: Cryptographic Token Interface Standard**
- **PKCS #12: Personal Information Exchange Syntax Standard**

TIP

On the day of the test, do not concern yourself too much with what the different standard numbers are. It is important to understand why they are in place and what PKCS stands for.

PKI solutions

Specifically, PKI seeks to provide solutions for the following:

- **Authentication:** The ability to verify a claim of identity by an individual or site
- **Trust:** The degree of confidence that an individual or a site is who they actually claim to be
- **Confidentiality:** The ability to ensure that data cannot be viewed and understood by a third party
- **Integrity:** The confidence that data has arrived in its original form without change or edit
- **Nonrepudiation:** Proof that an individual or party actually sent a particular piece of data

COMPONENTS OF PKI

Technologies such as virtual private networks (VPNs), digital signatures, access control (SSH), secure e-mail (PGP and S/MIME), and secure Web access (Secure Sockets Layer, or SSL) each includes an implementation of PKI for managing trusted communications between a host and a client. While PKI exists at some level within the innards of several types of communications technologies, its form can change from implementation to implementation. As such, the components necessary for a successful implementation can vary depending on the requirements, but in public key cryptography there is always:

- A *private key:* A key that is kept secret by an entity and used to decrypt messages encrypted with the mathematically related public key
- A *public key:* A key whose value is mathematically related to a private key that can be shared *publicly* to encrypt messages and validate digital signatures
- A *TTP:* An entity that helps facilitate transactions or communications between two parties as it is trusted by both parties

Since a public key must be associated with the name of its owner, a data structure known as a *public key certificate* is used. The certificate typically contains the owner's name, their public key and e-mail address, validity dates for the certificate, the location of revocation information, the location of the issuer's policies, and possibly other affiliate information that identifies the certificate issuer with an organization such as an employer or other institution.

Whether to use centralized or decentralized key management depends on the size of the organization. With *decentralized key management*, the private key can be assumed to belong only to its intended owner; with *centralized key management*, there is a possibility for abuse of other users' private keys by the administrators of the central key store.

- Digital certificates
- Certification authorities (CA)
- Certificate revocation lists (CRL)
- Recovery agents

Digital certificates

In PKI, a *digital certificate* is a tool used for binding a public key with a particular owner. A digital certificate in PKI serves the same function as a driver's license. Various systems and checkpoints may require verification of the owner's identity and status and will reference the TTP for validation. It is the certificate that enables this quick handoff of key information between the parties involved.

Certification authority

Certificates are created by a TTP called a *Certification Authority* (CA), which may also be called a *Certificate Authority*. CAs are systems that create, distribute, store, and validate digitally created signature and identity verification information about machines, individuals, and services. This CA may be a commercially available service point, such as VeriSign or Thawte. A CA can also be created within an enterprise to manage and create certificates that are used only within an organization or with trusted partners.

Certificate revocation list

A *certificate revocation list* is just as the name implies—a list of certificates issued by a specific CA that are no longer valid or have been revoked. CRLs are discussed in greater detail later in the chapter.

Recovery agents

A key *recovery agent* is an individual or entity that is responsible for recovering lost or damaged certificates for individuals. Recovery agents are discussed in greater detail later in the chapter.

X.509

The information contained in the certificate is actually part of the *X.509* certificate standard. X.509 is actually an evolution of the X.500 directory standard. Initially intended to provide a means of developing easy-to-use electronic directories of people that would be available to all Internet users, it became a directory and mail standard for a very commonly known mail application: Microsoft Exchange 5.5.

X.509 is the standard used to define what makes up a digital certificate. Within this standard, a description is given for a certificate as allowing an association between a user's distinguished name (DN) and the user's public key. The DN is specified by a naming authority (NA) and used as a unique name by the CA who will create the certificate.

PUBLIC KEYS

Since the PKI identification process is based on the use of unique identifiers known as keys, each person using PKI creates two different keys, a public key and

a private key. These keys are mathematically related, whereas things encrypted with one key can then be decrypted with the other—they are commonly referred to as a key pair.

Public keys are generally transported and stored in digital certificates. The public key is openly available to the public, while only the person the keys were created for has the private key. The utilization of these key pairs for public key cryptography brings major security technologies to the desktop. The network now is provided with the ability to allow users to safely:

- Transmit over insecure channels.
- Store sensitive information on any commonly used media.
- Verify a person's identity for authentication.
- Prove that a message was generated by a particular person.
- Prove that the received message was not tampered with in transit.

Algorithms based on public keys can be used for all these purposes. The most popular public key algorithm is the standard RSA, which is named after its three inventors: Rivest, Shamir, and Adleman.

PRIVATE KEYS

A *private key* is a secret key that is shared between two parties in symmetric cryptography and is only kept by one party in asymmetric cryptography. A private key can be used to decrypt information encrypted with the corresponding public key as well as used to create the digital signature of a file or certificate. A private key is meant to be kept secret, but it is common for individuals to make the mistake of sending their private keys to others to decrypt files.

EXAM WARNING

In a Windows Server 2008 PKI, a user's public and private keys are stored under the user's profile. For the administrator, the public keys would be under *Documents and Settings\Administrator\System Certificates\My\Certificates* and the private keys would be under *Documents and Settings\Administrator\Crypto\RSA* (where they are double encrypted by Microsoft's Data Protection API, or DPAPI). Although a copy of the public keys is kept in the registry, and can even be kept in Active Directory, the private keys are vulnerable to deletion. If you delete a user profile, the private keys will be lost!

Taking this a step further, imagine what could happen if a root CA key was not stored in a secure place; all of the keys that used the CA as their root certificate would have to be invalidated and regenerated. So, how to store private keys in a manner that guarantees their security? Not storing them in a publicly accessible FTP folder is just a start. There are also several options for key storage, most falling under either the software storage category or the hardware storage category.

DID YOU KNOW?

Keeping a root CA's private keys secure should be priority number one in PKI security. The work that goes into revoking and replacing a compromised root CA key is tremendous. Not only does the root CA have to be revoked and recreated, but so do any certificates created by a subordinate CA now suspect of being compromised. Also, the revocation of the root CA's key must be communicated to anyone who has ever trusted the root CA.

It is generally accepted that software storage is not a reliable means of storing high-security private keys. To overcome the issues of software storage, *Hardware Storage Modules* (*HSMs*) were created. HSMs, such as Smart Cards, Personal Computer Memory Card International Association (PCMCIA) cards, and other hardware devices, store private keys and handle all encryption and decryption of messages so that the key does not have to be transmitted to the computer. (Using magnetic media is really the equivalent of software key storage with an off-line file store, and should not be thought of as hardware storage of keys.) Keeping the keys off the computer prevents information about the keys from being discovered in computer memory. *Smart Cards* are the most flexible method of storing personal private keys using the hardware storage method. Since Smart Cards are normally about the size of a credit card, they are easily stored and can resist a high level of physical stress. Smart Cards are also not very expensive. Unlike a credit card that has a magnetic strip, Smart Cards store information using microprocessors, memory, and contact pads for passing information.

EXAM WARNING

Make sure that you understand what an HSM is and why a Smart Card is the most popular form of these modules.

Certificate authority

By definition, a *certificate authority* is an entity (computer or system) that issues digital certificates of authenticity for use by other parties. With the ever-increasing demand for effective and efficient methods to verify and secure communications, our technology market has seen the rise of many TTPs into the market. In a third-party or external PKI, it is up to the third-party CA to positively verify the identity of anyone requesting a certificate from it.

Within an organization, several options exist for building this trust relationship. Each of these begins with the decisions made around selecting and implementing CAs. There are two major roles or types of CAs to be aware of:

- **Root CA:** A root CA functions as a top-level authority over one or more levels of CAs beneath it. The CAs below the root CA are called subordinate CAs. Root CAs serve as a *trust anchor* to all the CAs beneath it and to the

users who trust the root CA. A trust anchor is an entity known to be trusted without requiring that it be trusted by going to another party, and therefore can be used as a base for trusting other parties.

- **Subordinate CA:** Any CA that is established after the root CA is a subordinate CA. Subordinate CAs gain their authority by requesting a certificate from either the root CA or a higher-level subordinate CA. Once the subordinate CA receives the certificate, it can control CA policies and/or issue certificates itself, depending on the PKI structure and policies.

There are two ways to view PKI trust models: single CA and hierarchical. In a *single CA model*, PKIs are very simplistic; only one CA is used within the infrastructure. Anyone who needs to trust parties vouched for by the CA is given the public key for the CA. That single CA is responsible for the interactions that ensue when parties request and seek to verify the information for a given certificate.

Since there is nothing above the root CA, no one can vouch for its identity; it must create a self-signed certificate to vouch for itself. With a *self-signed certificate*, both the certificate issuer and the certificate subject are exactly the same. Being the trust anchor, the root CA must make its own certificate available to all of the users (including subordinate CAs) that will ultimately be using that particular root CA.

Sometimes, subordinate CAs also issue certificates to other CAs below them on the tree. These CAs are called *intermediate CAs*. In most hierarchies, there is more than one intermediate CA. Subordinate CAs that issue certificates to end users, server, and other entities but do not issue certificates to other CAs are called *leaf CAs*.

Certificate revocation list

It is sometimes necessary to revoke a person's (or company's) certificate before the expiration date. Usually, revocation occurs when:

- A company changes ISPs, if its certificate was based on its ISP's Domain Name Server (DNS) name or its IP address, rather than the company's own DNS name, or if the ISP had access to the private key.
- A company moves to a new physical address, so that the address information in the certificate becomes incorrect.
- The contact listed on a certificate has left the company.
- A private key has been compromised or is lost.

Do not get tripped up by a question about a certificate being revoked. The thing to remember is that crucial information in the certificate has changed or the key has been compromised.

Another method of verifying the state of a certificate is called the *Online Certificate Status Protocol* (*OCSP*). OCSP was defined to help PKI certificate revocation get past the limitations of using CRL schemes. OCSP returns information relating only to certain certificates that have been revoked. With OCSP, there is no need for the large files used in a CRL to be transmitted.

With OCSP, a query is sent to a CA regarding a particular certificate over transport protocols such as Hypertext Transfer Protocol (HTTP). Once the query is received and processed by the CA, an OCSP responder replies to the originator with the status of the certificate, as well as information regarding the response. An OCSP response consists of:

- The status of the certificate ("good," "revoked," or "unknown")
- The last update on the status of the certificate
- The next time the status will be updated
- The time that the response was sent back to the requestor

One of the most glaring weaknesses of OCSP is that it can only return information on a single certificate, and it does not attempt to validate the certificate for the CA that issued it.

Key escrow

Key escrow works in the same way as a financial escrow. When a company uses key escrow, they keep copies of their private key in one or more secured locations where only authorized persons are allowed to access them. A simple key escrow scheme would involve handing a copy of a party's keys to an escrow company, who would only divulge the keys back to the party (or its legal successor) upon presentation of sufficient credentials. In a more advanced key escrow scheme, there may be two or more escrow agencies. The keys are split up and one half is sent to each of the two different escrow companies. Using two different escrow companies is a separation of duties, preventing one single escrow company from being able to compromise encrypted messages by using a client's key set.

Key escrow is a sore spot with many people and companies, because many proposed key escrow schemes are designed to allow a government or law enforcement authority to have access to keys.

REGISTRATION

Some PKI implementations use one or more *Registration Authorities* (*RA*). An RA is used to take some of the burden off the CA by handling verification of credentials prior to certificates being issued. In a single CA model, an RA can be used for verifying the identity of a subscriber, as well as setting up the preliminary trust relationship between the CA and the end user.

An RA is generally an out-of-band service provider, whose task is usually to verify identity documentation before confirming that a CA may issue a certificate. The RA is usually a physical outlet, at which a party will present itself, its

documentation, and its certificate request. The RA verifies the physical documentation, ensures that it matches the information in the certificate request, and that the documentation is sufficient to prove the identity claimed by the desired certificate. The RA typically also takes payment on behalf of itself and the CA, and on the basis of complete identification and payment, will request the CA to issue the requested certificate. RAs are found in stand-alone or hierarchical models where the workload of the CA may need to be offloaded to other servers.

EXAM WARNING

Make sure you understand the difference between a CA and an RA. You will need to know when an RA would be used within a PKI.

RECOVERY AGENTS

Sometimes it is necessary to recover a lost key. One of the problems that often arise regarding PKI is the fear that documents will become lost forever—irrecoverable because someone loses or forgets their private key. Let's say that employees use Smart Cards to hold their private keys. If a user were to leave his smart card in his wallet that was left in the pants that he accidentally threw into the washing machine, then that user might be without his private key and therefore incapable of accessing any documents or e-mails that used his existing private key.

Many corporate environments implement a key recovery server solely for the purpose of backing up and recovering keys. Within an organization, there typically is at least one *key recovery agent*. A key recovery agent has the authority and capability to restore a user's lost private key. Some key recovery servers require that two key recovery agents retrieve private user keys together for added security. This is similar to certain bank accounts that require two signatures on a check for added security. Some key recovery servers also have the ability to function as a key escrow server, thereby adding the ability to split the keys onto two separate recovery servers, further increasing security.

CAs and recovery servers also require certain information before they allow a key to be recovered. This is known as *Key Recovery Information* (*KRI*). KRI usually consists of:

- The name of the key owner
- Information verifying that the person requesting key recovery is authorized to recover the key on behalf of that key owner
- The time that the key was created
- The issuing CA server

Once the CA (or the key recovery agent) verifies the KRI, the key recovery process can begin.

IMPLEMENTATION

Certificates are used more frequently since the development and expansion of Internet-based transactions have grown. X.509 is an ITU-T standard for PKI, and X.509 certificates are now used for Web-based authentication for access to remote systems and for encryption of information on local machines. They are also used for directory services access in various operating systems, Smart Cards, digital signatures for e-mail, and encrypting e-mail. Additionally, they may be used for authentication when implementing a secure network protocol such as IPSec to protect data transmission within systems.

EXAM WARNING

Remember that certificates must be issued from a verifiable and identifiable CA. This can be a commercial entity, such as VeriSign or Thawte, or a stand-alone or enterprise CA within your organization. The path to the CA must be unbroken, or the certificate may be viewed as invalid. A compromised or physically unsecured CA will require recreation of your entire PKI infrastructure.

Certificates can be installed via the Web browser on client machines to identify and authenticate users. In some operating systems, such as Windows 2003 and Windows 2008, certificates can be mapped to user accounts in Active Directory, and then associated with the access tokens generated by the operating system when the user logs on, making the local installation of the certificate optional on the workstation being used. Web servers must have a Web server certificate installed in order to participate in SSL.

Multiple aspects of the certificate may be verified including the certificate expiry date, the domain associated with the certificate, and the validity of the CA. It is important to note that if the software verifying the certificate is not configured to trust the CA, the certificate will be considered invalid.

CERTIFICATE MANAGEMENT

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

Mechanisms that play a part in the life cycle of a certificate are:

- Centralized vs. decentralized key management
- Storage of private keys

- Key escrow
- Certificate expiration
- Certificate revocation
- Certificate suspension
- Key recovery
- Certificate renewal

SUMMARY OF EXAM OBJECTIVES

PKI has become an indispensable part of network operations and security operations in modern networks. PKI is predominantly based on public key cryptography and provides for multiple services that include authentication, data encryption, and identity validation. A PKI can be either hierarchical in nature, where a root CA is used to generate an *anchor of trust* and intermediate or subordinate CAs provide the certificate distribution to the end systems and users, or a stand-alone implementation, where a single CA provides both root and subordinate services to end devices and users. A stand-alone PKI implementation involves a single CA providing both root and subordinate services to end devices and users. Certificates are confirmed against a CA to ensure that the certificate is still valid and not revoked through the publication in a CRL or through the OCSP. Finally, key management is critical to a successful PKI implementation because of the need for key renewal and recovery.

TOP FIVE TOUGHEST QUESTIONS

1. You are applying for a certificate for the Web server for your company. Which of these parties would you not expect to be contacted in the process?

   A. A registration authority (RA)
   B. A leaf CA
   C. A key escrow agent
   D. A root CA

2. What statement best describes the transitive trust in a simple CA model?

   A. Users trust certificate holders, because the users and the certificate holders each trust the CA.
   B. Users trust certificate holders, because the users trust the CA and the CA trusts the certificate holders.
   C. Certificate holders trust users, because the certificate holders trust the CA and the CA trusts its users.
   D. Users trust certificate holders, because the certificate holders have been introduced to the users by the CA.

3. In a hierarchical CA model, which servers will use self-signed certificates to identify themselves?

   A. Root CAs
   B. Intermediate CAs

C. Leaf CAs
D. Subordinate CAs
E. All CAs

4. Which of the following certificate life-cycle events is best handled without revoking the certificate?

A. The contact e-mail address for the certificate changes to a different person.
B. The certificate reaches its expiry date.
C. The company represented by the certificate moves to a new town in the same state.
D. The certificate's private key is accidentally posted in a public area of the Web site.

5. Which statement is true about a CRL?

A. A CRL may contain all revoked certificates or only those revoked since the last CRL.
B. A CRL is published as soon as a revocation is called for.
C. A CRL only applies to one certificate.
D. A CRL lists certificates that can never be trusted again.

ANSWERS

1. The correct answer is D. A root CA. You will most likely contact an RA (Answer A) to prove your identity as a representative of your company, and you will be receiving your issued certificate from the leaf CA (Answer B). You will also want to escrow your private key with a key escrow agent (Answer C) so that it can be recovered in the event of your departure from the company or your losing the key. However, you will never want to contact the root CA, because the root CA is only used to form the trust anchor at the root of the certificate chain.

2. The correct answer is B. Users trust the CA, the CA trusts the certificate owners, and therefore the users trust the certificate owners. Answer A is wrong because there is no trust from the certificate holders up to the CA. Answer C is wrong for the same reason, and also because there is no trust from the CA to its users. Answer B is wrong because it does not involve the PKI model in any way.

3. The correct answer is A. Any CA other than the root must chain up to the root; only the trust anchor is able to vouch for itself with no other authority to support its claim. Intermediate CAs (Answer B) are signed by another CA; leaf CAs (Answer C) are signed by the intermediate or root CA above them; subordinate CAs (Answer D) are signed by the CA above them. Answer E—all CAs—cannot be true unless all A–D are true.

4. The correct answer is B. When the certificate reaches its expiry date, it naturally expires everywhere, and you should already have requested a renewal certificate with a later expiry date. The other answers are all reasons to revoke the certificate as soon as possible. Answer A, a change of contact e-mail address, requires revoking the certificate to prevent the old e-mail contact from being able to submit a request for a changed certificate; a change of address (Answer C) voids information in the certificate, so that it is no longer a true statement of identity; accidental (or deliberate) exposure of the private key to unauthorized parties results in the certificate being unreliable as a uniquely identifying piece of information.

5. The correct answer is A. A CRL may be simple, containing all certificates that have been revoked, or delta, containing all certificates that have been revoked since the last CRL was published. Answer B is not true. CRLs are published to a schedule. Answer C is not true of CRLs but is true of OCSP. Answer D is not true because some of the certificates on the CRL may be merely "suspended," and will be trustable later.

CHAPTER 12

Redundancy Planning

Exam objectives in this chapter:

- Alternate Sites (Hot, Warm, and Cold)
- Redundant Systems
- RAID
- Spare Parts
- Backup Generator
- UPS

ALTERNATE SITES

Alternate sites are important to certain companies, so they can experience minimal downtime or almost no downtime at all. In a disaster, it's possible that the facilities, servers, or other network devices are damaged or destroyed. In such a case, the company would require a temporary location in which data can be restored to servers and business functions can resume. Without such a facility, the company would need to find a new business location, purchase new equipment, set it up, and then go live. Alternate sites get the business up and running quicker, allowing the business to continue running until their existing facilities are repaired or a new permanent site is established.

There are different types of alternate sites that can be used, with each having their own benefits and drawbacks. They are:

- **Hot site**, which has everything needed and is ready to go live
- **Warm site**, in which some equipment and services need to be setup, and data needs to be restored from backups before going live
- **Cold site**, which is the least expensive to maintain, but requires the most amount of preparation before going live

Creating alternate or backup sites can take considerable planning. Companies need to identify what equipment needs to be available, and how fast they need backup systems to go live after a disaster. When deciding on appropriate locations for such sites, it is important that they be in different geographical locations.

DID YOU KNOW?

The terrorist activities of September 11, 2001, which resulted in the destruction of the World Trade Center in New York City, caused many companies to seriously consider their *business continuity plans*. Companies may have planned for a localized disaster (such as a fire) affecting their business, but the decimation caused by airliners slamming into buildings was something no one had accounted for. A wide-scale disaster resulting in the loss and inaccessibility of employees, facilities, and other assets wasn't something that many considered.

Hot site

A *hot site* is a facility that has the necessary hardware, software, phone lines, and network connectivity to allow a business to resume normal functions almost immediately. This can be a branch office or data center, but must be online and connected to the production network. A copy of data is held on a server at that location, so little or no data is lost. Replication of data from production servers may occur in real time, so that an exact duplicate of the system is ready when needed. In other instances, the bulk of data is stored on servers, so only a minimal amount of data needs to be restored. This allows business functions to resume very quickly, with almost zero downtime.

Warm site

A *warm site* is not as equipped as a hot site, but has part of the necessary hardware, software, and other office needs to restore normal business functions. Such a site may have most of the equipment necessary, but will still need work to bring it online and support the needs of the business. With such a site, the bulk of data will need to be restored to servers, and additional work (such as activating phone lines or other services) will need to be done. No data is replicated to the server, so backup tapes must be restored so that data on the servers is recent. Warm sites cost less than hot sites, which makes them an attractive alternative.

Cold site

A *cold site* requires the most work to set up, as it is neither online nor part of the production network. It may have all or part of the necessary equipment and resources needed to resume business activities, but installation is required and data needs to be restored to servers. Additional work (such as activating phone lines and other services) will also need to be done. The major difference between a cold site and hot site is that a hot site can be used immediately when a disaster occurs, while a cold site must be built from scratch. A cold site is the least expensive type of alternate site, but isn't an option for companies that can't afford to wait for servers and equipment to be set up.

Crunch Time

The exam will expect you to know the difference between cold, warm, and hot sites. Don't get too stressed out trying to remember all of the features. A quick and dirty way of keeping them straight is to remember that a hot site is active and functional, a cold site is offline and nonfunctional, and a warm site is somewhere in between.

REDUNDANT SYSTEMS

A single point of failure can be the Achilles' heel that brings down a system. *High availability* is a network's ability to keep systems operating and services available in the event of an outage. How this is provided is through redundant systems and fault tolerance. *Redundancy* is a duplication of services and systems. If the primary method used to store data, transfer information, or other operations fails, then a secondary method is used to continue providing services. *Fault tolerance* refers to a systems ability to continue working in the event of such a failure. If one component stops working, it will fail over to another component.

Crunch Time

Don't get confused between the terms high availability, redundancy, and fault tolerance. High availability means that things are up and running most of the time, regardless of a problem. Redundancy means that services and systems are duplicated, so if one goes down, the other can still be used. Fault tolerance is that system's ability to continue working if a component or service fails. These terms will probably appear on your exam, so you should be familiar with each of them.

Servers

There are many ways of providing fault tolerance and redundancy in servers, which involves duplicate components or duplicate servers. For example, a server may have multiple network cards installed on it, so that if one of the cards fails, the data on that server can still be accessed through the second card. A server that provides important services or runs critical programs may use a failover server. The *failover server* duplicates the services and data of the primary server, and checks at regular intervals that the primary server is running.

If the failover server doesn't receive a response during one of these checks, it will then takeover the role of the primary server and provides services to users.

Many operating systems provide the ability to cluster servers together. *Server clusters* are groups of independent servers that are connected together, so that if one fails the other will continue to provide services. Each server handles its own local resources and has a copy of the services and applications that run on other servers in the cluster. In many clusters, the servers share a single disk system, and appear on the network as a single entity. When a user makes a request for a resource, it is sent to the cluster. If one of these servers failed, the others in the cluster would still function and be able to take over processing requests from the network and providing services.

There are two forms of server clusters that can be used on a network: active/active and active/passive.

- An *active/active cluster* has all of the servers actively responding to requests, so that if one server fails, all of the other servers in the cluster can continue processing requests.
- An *active/passive cluster* has servers that are only used if the active server fails. In this type of cluster, the active server process requests while the other only becomes active if the first one fails.

EXAM WARNING

Remember that server clusters provide fault tolerance and redundancy, allowing users to continue making requests to servers even if one of the servers in the cluster fails. An active/active cluster provides high availability because all of the servers are regularly responding to requests. This isn't true in an active/passive cluster, where the passive server only becomes active when the primary server fails.

Connections

Redundancy is often found in networks, such as when multiple links are used to connect sites on a wide area network (WAN). Network lines may be used to connect two sites, with a separate network line set up in case the first goes down. If this first link fails, the network can be switched over to use the second link. In other instances, additional lines may be set up in other ways to provide redundancy.

ISP

Many companies depend on Internet connectivity almost as much as network connectivity. In some cases, such as e-commerce businesses, they depend on it more. A redundant Internet Service Provider (ISP) can be used to provide connectivity when an organization's primary ISP's service becomes unavailable. The

link to the secondary ISP can be configured as a low-priority route, while the primary ISP is advertised as high priority. Such a configuration will have users using the primary ISP for normal usage, but automatically switching over to the low-priority connection when the first one fails. If a secondary ISP is not desired, the administrator should ensure that the ISP uses two different points of presence. A *point of presence* is an access point to the Internet, therefore having multiple points of presence will allow access to the Internet if one goes down.

RAID

Data is a commodity of any business, so it's important to ensure that it is always available to those who need it. *Redundant Arrays of Inexpensive Disks* (*RAID*) was developed to prevent the loss of data and/or improve performance. RAID provides several methods of writing data across multiple disks, and writing to several disks at once. Rather than losing a single disk and all the information, administrators can replace the damaged disk and regenerate the data quickly. When determining which level of RAID to use, it is important to remember that some RAID levels only increase performance, some only prevent loss of data, but not all will do both. The different levels of RAID available include:

- **RAID 0 (Disk Striping)**: Data is written (striped) across two or more disks, but no copies of the data are made. This improves performance because data is read from multiple disks, but there is no fault tolerance if a disk fails.
- **RAID 0 + 1 (Disk Striping with Mirroring)**: Combines features of RAID 0 and RAID 1. Allows four or more disks to be used as a set, but provides full redundancy and the same fault tolerance as RAID 5.
- **RAID 1 (Mirroring or Duplexing)**: Data that is written to one disk is also written to another, so that each drive has an exact copy of the data.
- **RAID 2**: Similar to RAID 0, except that error correction codes are used for drives that do not have built-in error detection.
- **RAID 3**: Data is striped across three or more drives, but one drive is used to store the parity bits for each byte that is written to the other disks.
- **RAID 4**: Similar to RAID 3, but stripes data in larger blocks.
- **RAID 5 (Disk Striping with Parity)**: Data is striped across three or more disks, but parity information is stored across multiple drives.
- **RAID 10**: Allows four or more drives to be used in an array, and has data striped across them with the same fault tolerance as RAID 1.
- **RAID 53**: Allows a minimum of five disks to be used in an array, but provides the same fault tolerance as RAID 3.

RAID is available through hardware or software. Hardware RAID generally supports more levels of RAID, and provides higher performance. This type of RAID can also support *hot swapping* (discussed in the next section), in which a disk can be removed from the server without having to take the server down.

Software RAID is provided through OSs, such as Windows. When RAID is provided through the software, the levels of RAID supported may be limited.

Crunch Time

RAID 0, 1, 3, and 5 are the most commonly used levels of RAID. While there are other levels of RAID that could possibly be used on a network, these four RAID levels are the ones most likely to appear on your exam. Focus studying on RAID 0, 1, 3, and 5.

SPARE PARTS

Spare parts refer to additional hardware components that are necessary for servers or other network devices to operate. If a network card or power supply on a server failed, having an extra component on hand makes it possible to replace the part and get the server up and running. Hardware components may provide features that improve the ease and speed of replacing faulty hardware. The following allow you to replace a faulty component without having to completely shut down the system:

- **Hot swap:** Refers to the ability to replace hardware components without having to shutdown the computer.
- **Warm swap:** Similar to hot swapping but the computer need not be shut down but rather put into a suspended state (such as hibernation) while the hardware is being replaced.
- **Hot spare:** Installed on the system but isn't used until the primary component fails. When the component fails, the system might be configured to detect this and automatically switch over to the hot spare.

EXAM WARNING

Don't confuse a hot spare with some of the other "hot" topics we've discussed in this chapter. A hot spare is installed in the computer, and is only used when the primary component fails.

BACKUP GENERATOR

Even if an administrator is comfortable with the internal measures they have taken to protect data and other assets, outside sources may still have an impact on systems. Utility companies supply essential services, such as electricity and communication services. In some disasters, such as major storms or earthquakes, these services may become unavailable. Without them, servers and

other vital systems are left without power and unable to phone for assistance to bring them back online when power is restored. To continue doing normal business functions, administrators need to implement equipment that will provide these services when the utility companies cannot.

When power is out for lengthy periods of time, additional measures may be necessary to supply electricity to equipment. *Power generators* can run on gasoline, kerosene, or other fuels for an extended time, and provide energy to a building. Certain power outlets may be connected to the generator, so that any systems plugged into these outlets will receive power when normal power is lost.

DID YOU KNOW?

In August 2003, *a major power outage* affected parts of the United States and Ontario, Canada. An estimated 45 million Americans and 10 million Canadians were left without power for a day. Because of preparation for Y2K a few years before and other factors, a number of homes and businesses had various kinds of power generators. Unfortunately, people who owned gas generators and didn't have a supply of gasoline on hand were faced with a surprising fact: Gas pumps were electrically powered. Gas stations affected by the blackout had no way of powering the pumps to get the gas out of the ground, and had to close (even though there were plenty of potential customers driving around looking for fuel).

UPS

Uninterruptible power supplies (*UPS*) are power supplies that can switch over to a battery backup when power outages occur. Multiple devices can be plugged into a UPS similar to a power bar, and the UPS generally provides such functions as surge protection and noise filtering. When a drop in voltage occurs, the UPS detects it and switches over to battery backup. Components plugged into the UPS can then receive power for a limited amount of time (often ranging from 10 to 45 minutes), until normal power is restored or the system can shut down properly. This does not allow you to continue normal business functions, but will protect data from corruption caused by sudden losses of power and improper shutdowns.

EXAM WARNING

UPS are used for short-term power, while backup generators are designed for providing power for longer periods of time.

SUMMARY OF EXAM OBJECTIVES

A number of methods are available for providing redundancy of systems and preparing for potential threats that could impact an organization's ability to

function. In cases where the business's facility or networking capabilities are damaged or destroyed, alternate sites can be used. These sites can take various amounts of preparation to get up and running. Hot sites take little work to get online and have a copy of data on servers; warm sites require restoring backed-up data to servers and may require some equipment, while cold sites must be made from scratch. Redundant systems can also be used to reduce the impact of potential threats, by having duplicate components or systems available in case one fails. This can include having servers clustered on a network, having spare components available to install or bring online when a failure occurs, or implementing RAID. These ensure that the servers have high availability, can fail over, or allow data to be restored if a disaster occurs. Because power is so important to a business, methods of providing power during an outage must also be available on a network. UPS can be used to provide power for short periods of time, allowing a computer to be shut down gracefully. For longer periods of power outages, backup generators can be used to provide power for hours or days at a time. Together, redundancy in systems protects a business from a wide variety of threats. They allow systems to continue functioning throughout a disaster, and allow companies to continue doing business.

TOP FIVE TOUGHEST QUESTIONS

1. You are deciding on appropriate locations for a cold site that will be used in case of a disaster. You decide to set up the cold site in a nearby facility, which is used by the company to store equipment and office supplies. The building has an old Halon system for fire suppression in key areas, has air conditioning in all areas, and is dry. Should a disaster occur, members of the organization will simply move down the street and set up operations at this location? Based on the features and location of the site, is it suitable to set up a cold site?

   A. The facility is a perfect location for a cold site.
   B. The fire suppression system, air conditioning, and other environmental conditions make it unsuitable for a cold site.
   C. The physical proximity to the company makes it unsuitable for a cold site.
   D. The fact that it is not part of the production network makes it unsuitable for a cold site.

2. A service runs on a network server that users access with an application on their workstations. The application is used to process requests and access data in a database. If the server or service fails, you want users still be able to access this data. What method of fault tolerance will you use so that network users can still continue to work?

   A. Install two network cards on the server, so that if one card fails, users can still access the data through the second card.
   B. Use server clustering to provide fault tolerance.

C. Implement RAID.
D. Connect the server to a UPS.

3. You have decided to set up server clustering on your network, so that there is no loss of availability to data. Which of the following will you use?

A. Active/active clustering, so that all of the servers are able to become active if one of them fails.
B. Active/active clustering, so that all of the servers are actively processing requests.
C. Active/passive clustering, so that if the active server fails, the passive server will become active and begin processing requests.
D. Active/passive clustering, so that all of the servers are actively processing requests.

4. You have decided to implement a RAID for fault tolerance, and want data to be striped across multiple disks with parity information stored on multiple drives. Which of the following levels of RAID will you use?

A. RAID 0
B. RAID 1
C. RAID 3
D. RAID 5

5. You have decided to purchase spare hardware components that you can replace on a server without having to shutdown the computer. Which of the following is being used?

A. Hot swapping
B. Warm swapping
C. Hot spare
D. Hot site

ANSWERS

1. The correct answer is **C**. The physical proximity to the company makes it unsuitable for a cold site. When deciding on appropriate locations for such sites, it is important that they be in a different geographical location. If it is not a significant distance from the primary site, it can fall victim to the same disaster. Both sites would experience the same disaster, so there would be no alternate site available to resume business. Answer A is incorrect because the location makes it unsuitable for a cold site. Answer B is incorrect because the site uses an old but functional fire suppression system, and has other environmental conditions that make it suitable for a cold site. Answer D is incorrect because a cold site is not part of the production network.

2. The correct answer is B. Use server clustering to provide fault tolerance. A is incorrect because if the service failed, it wouldn't matter that there were two network cards installed. Users still wouldn't be able to access the data because the service wouldn't be available. C is incorrect because the RAID array would become unavailable when the server failed. D is incorrect because if the service failed, the data would still be unavailable even though the server still had power.

3. The correct answer is B. Use active/active clustering. A is incorrect because an active/active cluster already has all of the servers actively responding to requests. C is incorrect because there is a loss of availability during the time when the passive server identifies that the active server is no longer active. D is incorrect because this describes an active/active cluster, and not an active/passive cluster.

4. The correct answer is D. RAID 5 is disk striping with parity. Data is striped across multiple disks, but parity information is stored across multiple drives. It provides fault tolerance because a single disk that fails in the set can be restored from the parity information on the other disks. Answer A is incorrect because RAID 0 provides no fault tolerance. Data is written (striped) across multiple disks, but no copies of the data are made. This improves performance because data is read from multiple disks, but the data on the entire set will be lost if one disk fails. Answer B is incorrect because RAID 1 is disk mirroring or duplexing. Data that is written to one disk is also written to another, so that one disk's data is a mirror image of the other's. Parity information is not stored on multiple drives with this method. Answer C is incorrect because RAID 3 has data striped across several drives, but one drive is used to store the parity bits for each byte that is written to the other disks. When a disk fails, it can be replaced and data can be restored to it from the parity information. If two or more disks in the set fail, then data cannot be recovered.

5. The correct answer is A. Hot swapping. Hot swapping refers to the ability to replace hardware components without having to shutdown the computer. If a component fails, you don't need to power off the machine. B is incorrect because warm swapping requires the computer to be put into a suspected state (such as hibernate) while the hardware is being inserted or removed. C is incorrect because a hot spare is a spare component that is installed on the system, but isn't used until the primary component fails. When the component fails, the system might be configured to detect this and automatically switch over to the hot spare. D is incorrect because a hot site has a copy of data stored on the servers, and doesn't need additional equipment brought to the site in the case of a disaster.

CHAPTER 13

Controls and Procedures

Exam objectives in this chapter:

- Environmental Controls
- Implementing Disaster Recovery and Incident Response Procedures

ENVIRONMENTAL CONTROLS

Even with educated users and with all critical systems locked behind closed doors, equipment and data are still at risk if the environment beyond those locked doors is insecure. *Environment* refers to the surroundings in which the computers and other equipment reside. If an environment is insecure, data and equipment can be damaged. To prevent the environment from affecting a system's safety and ability to function, the following elements should be considered:

- Fire suppression
- Temperature
- Humidity
- Airflow
- Electrical and other types of interference
- Electrostatic discharge (ESD)

Fire suppression

Fire is a major risk in any environment that contains a lot of electrical equipment, so *fire suppression systems* must be put in place to protect servers and other equipment. Because problems with moisture and flooding can damage or destroy equipment, water sprinklers are not an option in server rooms or other areas storing devices. Other problems may occur if the fire suppression system releases foam that damages equipment, creates significant smoke when putting out a fire, or causes other potential problems that can result in collateral damage.

Halon is a fire suppressant often found in older facilities. When a fire occurred, this chemical would be dumped into the room at high pressure, removing necessary elements needed to work with the oxygen and fuel the fire. Although it

worked, it was found to be damaging to the ozone and was banned from new installations of fire suppression systems. There are many different alternatives to Halon that can be used safely without negative impacts on the environment. These include:

- Inergen (IG-541)
- Heptafluoropropane (HFC-227ea)
- Trifluromethane (FE-13)
- Carbon dioxide systems

When you are deciding on a fire suppression system, it is important to examine whether it will damage equipment or is toxic to people when the fire suppression system is deployed.

EXAM WARNING

Remember that Halon isn't manufactured anymore, so fire suppressants used in new systems use other chemicals to put out a fire without damaging equipment.

DETECTION SYSTEMS

Before a fire suppressant becomes activated, the signs of a fire must be detected. There are several types of devices that will detect the signs of fire, which in turn should then activate the fire suppressant system in a server room or other location where equipment resides. These are:

- *Smoke detection*, which is the most common method of detecting a fire
- *Heat detection*, which is used to monitor the temperature levels of a room
- *Flame detection*, which is used to detect the movement of flames or certain types of energy (i.e., ultraviolet and infrared) that indicates a fire has occurred

EXAM WARNING

The three indicators of fire used by detection systems are smoke, heat, and flame. Once these systems detect a fire, the fire suppression system is activated.

HVAC

HVAC is an acronym for Heating, Ventilation, and Air Conditioning. It is the control systems used to control humidity, temperature, and airflow. The environment in server rooms and other areas where sensitive equipment resides needs to have controlled conditions to operate properly. If temperature or humidity is too high or too low, it can damage the equipment and result in the loss of data. If a

computer overheats, the components inside it can be permanently damaged. While the temperature of the server room may feel comfortable to you, the inside of a computer can be as much as 40° warmer than the air outside the case. The hardware inside the case generates heat, raising the interior temperature. Computers are equipped with fans to cool the power supply, processor, and other hardware, so that temperatures do not rise above 110°. If these fans fail, the heat can rise to a level that destroys the hardware. The machine may not fail immediately, rather it may experience reboots, "blue screens of death," memory dumps, and other problems that occur randomly.

Chip creep (also known as *socket creep*) can occur due to the expansion and contraction that occur in motherboards and other circuit boards. As the circuit boards expand and contract, it causes the computer chips on these boards to move until they begin to lose contact with the sockets in which they are inserted.

Electrostatic discharge (*ESD*) is another threat to equipment, as static electricity can damage hardware components so they cease to function. To damage a computer chip, it requires only a discharge of 20 or 30 volts. Humidity levels can increase ESD. If the humidity in a room is below 50%, the dry conditions create an atmosphere that allows static electricity to build up. This creates the same situation as mentioned in the previous paragraph. A humidity level that is too high can also cause ESD, as water particles that conduct electricity can condense and stick to hardware components.

Shielding

Shielding refers to materials that are used to prevent data signals from being affected by external sources. This not only applies to wireless data escaping outside of an office, but also pertains to external signals or interference affecting data being carried along cables. Shielding can be used to prevent wireless transmissions from escaping a building or office area. Shielding blocks signals from escaping, but may also have the unwanted effect of blocking cellular communications.

Shielding is also necessary to prevent data from being damaged in transmission from *radio frequency interference* (*RFI*) and *electromagnetic interference* (*EMI*).

- ***RFI*** is caused by radio frequencies emanating from microwaves, furnaces, appliances, radio transmissions, and radio frequency–operated touch lamps and dimmers. Network cabling can pick up these frequencies much as an antenna would, corrupting data traveling along the cabling.
- ***EMI*** is caused by electromagnetism generated by heavy machinery such as elevators, industrial equipment, and lights.

When cabling travels past sources of EMI and RFI, a higher grade of cabling should be used, which has better shielding and can protect the wiring inside from interference. *Shielded twisted pair* (*STP*) is a type of cabling that uses a series of individually wrapped copper wires encased in a plastic sheath.

DID YOU KNOW?

An alternative to copper cabling and wireless technologies is using *fiber-optic cabling*, in which data is transmitted by light. Fiber-optic cable has a core made of light-conducting glass or plastic surrounded by a reflective material called cladding. A plastic sheath surrounds all of this for added protection. Because the signal is transmitted via light, data that travels along fiber-optic cable is not affected by interference from electromagnetism or radio frequencies. This makes it an excellent choice for use in areas where there are sources of EMI or RFI.

IMPLEMENTING DISASTER RECOVERY AND INCIDENT RESPONSE PROCEDURES

After the events of September 11, 2001, the widespread effects of a disaster became evident. Equipment, data, and personnel were destroyed, staggering amounts of money were lost by individual businesses, and the economic ripples were felt internationally. While some companies experienced varying levels of downtime, some never recovered and were put out of business. To deal with the various incidents and disasters that can affect an organization, procedures need to be in place so that professionals within the company can deal with them.

Disaster recovery

Backups of data need to be performed daily to ensure data can be recovered, plans need to be created that outline what tasks need to be performed by whom, and other issues need to be addressed as well. While it is hoped that such preparation is never needed, it is vital that a strategy is in place to deal with incidents. The disaster recovery plan should identify as many potential threats as possible and include easy-to-follow procedures. When discussing disaster recovery plans in greater detail, a plan should provide countermeasures that address each threat effectively.

DISASTER RECOVERY PLAN

Disaster recovery plans are documents that are used to identify potential threats and outline the procedures necessary to deal with different types of threats. When creating a disaster recovery plan, administrators should try to identify all the different types of threats that may affect their companies. Disasters are not limited to acts of nature, but can be caused through electronic methods also. Risk analysis should be performed to determine what is at risk when a disaster occurs. This should include such elements of a system as:

- Loss of data
- Loss of software and hardware
- Loss of personnel

Recovery methods discussed in the disaster recovery plan should focus on restoring the most business-critical requirements first. Another important factor

in creating a disaster recovery plan is cost. As discussed in Chapter 12, hot, warm, and cold sites require additional cost such as rent, purchasing hardware that may not be used until a disaster occurs (if one ever does), stock office supplies, and other elements that allow a business to run properly.

BACKUP TECHNIQUES AND PRACTICES

Backing up data is a fundamental part of any disaster recovery plan. When data is backed up, it is copied to a type of media that can be stored in a separate location. The type of media will vary depending on the amount of data being copied, but can include:

- Digital audio tape (DAT)
- Digital linear tape (DLT)
- Compact discs (CDR/CD-RW) and DVDs
- A folder location on a separate server

When making backups, the administrator needs to decide what data will be copied to alternative media. Critical data, such as trade secrets that a business relies on to function and other important data crucial to a business' needs must be backed up. Other data such as temporary files, applications, and others may not be backed up as they can easily be reinstalled or missed in a backup. Such decisions, however, will vary from company to company.

Once the administrator has decided on what information needs to be backed up, they can determine the type of backup that will be performed. Common backup types include:

- ***Full backup,*** which backs up all data in a single backup job
- ***Incremental backup,*** which backs up all data that was changed since the last backup and changes the archive bit
- ***Differential backup,*** which backs up all data that has changed since the last full backup but does not change the archive bit
- ***Copy backup,*** which makes a full backup but does not change the archive bit

TIP

Make sure you know the difference between the different types of backups you can perform. The backup types are full, incremental, differential, and copy. Each of these may be used for different purposes, and can affect whether the archive bit is reset on a file.

ROTATION SCHEMES

A popular rotation scheme is the *Grandfather-Father-Son* (*GFS*) rotation, which organizes rotation into a daily, weekly, and monthly set of tapes. With a GFS backup schedule, at least one full backup is performed per week, with differential or incremental backups performed on other days of the week. At the end of the week, the daily and weekly backups are stored off-site, and another set is used through the next week. Because it is too expensive to continually use new

tapes, old tapes are reused for backups. A tape set for each week in a month is rotated back into service and reused.

In the GFS rotation scheme, the full backup is considered the "Father" and the daily backup is considered the "Son." The "Grandfather" segment of the GFS rotation is an additional full backup that is performed monthly and stored off-site. The Grandfather tape is not reused, but is permanently stored off-site. Each of the Grandfather tapes can be kept for a specific amount of time (such as a year), so that data can be restored from previous backups, even after the Father and Son tapes have been rotated back into service. If someone needs data restored from several months ago, the Grandfather tape enables a network administrator to retrieve the required files.

A backup is only as good as its ability to be restored. To ensure that data is being backed up properly and can be restored correctly, administrators should perform test restores of data to the server.

OFF-SITE STORAGE

Off-site storage can be achieved in a number of ways. If a company has multiple buildings, such as in different cities, the backups from other sites can be stored in one of those buildings and the backups for servers in that building can be stored in another building. If this is not possible, there are firms that provide off-site storage facilities. The key is to keep the backups away from the physical location of the original data.

Exam Warning

Backups are an important part of disaster recovery, so it is possible there will be a question or two dealing with this topic.

When deciding on an off-site storage facility, administrators should ensure that it is secure and has the environmental conditions necessary to keep the backups safe. They should also ensure that the site has air conditioning and heating, as temperature changes may affect the integrity of data. The facility should also be protected from moisture and flooding and have fire protection. The backups need to be locked up, and policies in place of who can pick up the data when needed.

SECURE RECOVERY

Dealing with damaged hardware will vary in complexity, depending on the availability of replacement equipment and the steps required when restoring data to the network. Some companies may have additional servers with identical configurations to damaged ones, for use as replacements when incidents occur. Other companies may not be able to afford such measures or do not have enough additional servers to replace damaged ones. In such cases, the

administrator may have to put data on other servers, and then configure applications and drive mappings so the data can be accessed from the new location.

SECURE RECOVERY RESTORATION

Administrators also need to determine how data will be restored from backups. There are different types of backups that can be performed. Each of these takes differing lengths of time to restore, and may require additional work.

- ***Full backups*** will back up everything, so additional tapes are not needed.
- ***Incremental backups*** take the longest to restore. Incremental backups contain all data that was backed up since the last backup, thus many tapes may be used since the last full backup was performed. When this type of backup is used, administrators need to restore the last full backup and each incremental backup that was made since.
- ***Differential backups*** take less time and fewer tapes to restore than incremental backups. Because differential backups back up all data that was changed since the last full backup, only two tapes are needed to restore a system. The administrator needs to restore the tape containing the last full backup and the last tape containing a differential backup.

Incident response

Incidents can be any number of adverse events affecting a network or computer system or violations of existing policy. They can include, but are not limited to, unauthorized access, denial or disruptions of service, viruses, unauthorized changes to systems or data, critical system failures, or attempts to breach the policies and/or security of an organization.

A good *incident response policy* outlines who is responsible for specific tasks when a crisis occurs. It will include such information as:

- Who will investigate or analyze incidents to determine how they occurred and what problems are faced because of it
- Which individuals or departments are to fix particular problems and restore the system to a secure state
- How certain incidents are to be handled and references to other documentation

Including such information in the incident response policy ensures that the right person is assigned to a particular task. Determining who should respond and deal with specific incidents allows the restoration of a system to a secure state more quickly and effectively.

Incident response policies should also provide information on how to deal with problems when they occur, or provide references to procedures. Procedures should be clearly defined so that there is no confusion as to how to deal with an incident. Once an incident has been dealt with, the Incident Response (IR) Team should determine ways to ensure the same incident will not happen again. Simply resolving the crisis but not changing security methods increases the likelihood

that the same incident may occur again in the exact same manner. Taking a proactive approach to future incidents decreases the chance of recurring problems.

INCIDENT RESPONSE TEAMS

Incident Response Teams are IT professionals used to handle incidents that occur in a company, and may be formed in a number of ways. Some organizations use the people who have on-call duties, and are used to respond to any problems that users may encounter or with the network after hours. Because these people are trained and have the experience to troubleshoot and handle situations after hours, and generally are selected from a group of IT staff with diverse duties, many companies select them as the obvious choice for responding to incidents as a group. Other companies may form a formal team of selected individuals, while others create them as needed based on the type of incident being encountered.

In responding to an incident, the team should be trained in best practices and proper procedures. In an incident, they will go through a process of the following steps:

1. **Identification:** The Incident Response Team identifies the type of incident occurring.
2. **Investigation:** The team analyzes what has occurred and is impacted by an intrusion or other incident.
3. **Repair:** Once a system has been compromised, it must be repaired.
4. **Documentation:** All of the information that was gathered in the previous steps is written as a permanent record of the incident.

FORENSICS

Computer forensics is the application of computer skills and investigation techniques for the purpose of acquiring evidence. It involves collecting, examining, preserving, and presenting evidence that is stored or transmitted in an electronic format. Because the purpose of computer forensics is its possible use in court, strict procedures must be followed for evidence to be admissible. Even when an incident isn't criminal in nature, forensic procedures are important to follow.

Awareness

Management and employees need to be aware of the need to support computer forensic examinations. Management needs to provide funding for tools and ongoing training in examination procedures or to hire outside parties to perform the investigation. If law enforcement is called in whenever there is an incident, then there are no direct costs but there is still the need of cooperation with investigators.

Because digital evidence may be damaged or destroyed by improper handling or examination, management must also be aware that considerable time may be involved to effectively investigate an incident. Vital systems or facilities might be unavailable while evidence is being gathered, and it might be necessary

for equipment to be removed from service to be examined and stored as evidence until a criminal case has reached its conclusion.

Conceptual knowledge

Computer forensics is a relatively new field that emerged in law enforcement in the 1980s. Since then, it has become an important investigative practice for both police and corporations. It uses scientific methods to retrieve and document evidence located on computers and other electronic devices. By retrieving this information, it may result in the only evidence available to convict a culprit, or enhance more traditional evidence obtained through other investigative techniques.

Crunch Time

Forensics has four basic components: evidence must be collected, examined, preserved, and presented. The tasks involved in forensics will either fall into one of these groups or be performed across most or all of them. A constant element is the need for documentation, so that every action in the investigation is recorded. When taking the test, remember the four basic components and that everything *must* be documented.

Understanding

Because any evidence may be used in possible criminal proceedings, thorough documentation cannot be stressed enough. Documentation provides a clear understanding of what occurred to obtain the evidence, and what the evidence represents. Information should include the date, time, conversations pertinent to the investigation, tasks that were performed to obtain evidence, names of those present or who assisted, and anything else that was relevant to the forensic procedures that took place.

When conducting the investigation, a person must be designated as being in charge of the scene. This person should be knowledgeable in forensics and directly involved in the investigation. The person in charge should have authority to make final decisions on how the scene is secured, and how evidence is searched, handled, and processed. There are three major roles that people may perform when conducting an investigation. These roles are:

- First responder
- Investigator
- Crime scene technician

FIRST RESPONDERS

The *first responder* is the first person to arrive at a crime scene, who has the knowledge and skill to deal with the incident. The first responder may be an

officer, security personnel, a member of the IT staff or Incident Response Team, or any number of other individuals. The first responder is responsible for:

- **Identifying the scope of a crime scene:** What is affected and where could evidence exist?
- **Establishing a perimeter:** Protecting the crime scene requires cordoning off the area where evidence resides.
- **Preserving volatile evidence:** If a source of evidence is on the monitor screen, they should take steps to preserve and document it so it isn't lost.

The first responder shouldn't touch anything that is within the crime scene. Depending on how the crime was committed, 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. In terms of digital evidence, it is important for the first responder not to touch anything or attempt to do anything on the computer(s) as it may alter, damage, or destroy data or other identifying factors. When investigators arrive on the scene, it is important that the first responder provide 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 a report. Any observations should be mentioned, as this may provide insight into resolving the incident.

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. The technician is responsible for preserving evidence, and will make great efforts to do so. Evidence is further packaged to reduce the risk of damage, such as from ESD or jostling during transport. Once transported, the evidence is then stored under lock and key to prevent tampering, until such time that it can be properly examined and analyzed.

CHAIN OF CUSTODY

A *chain of custody* must be established to show how evidence made it from the crime scene to the courtroom. It proves where a piece of evidence was at any given time and who was responsible for it. By documenting this, you can establish that the integrity of evidence wasn't compromised. If the chain of custody is broken, it could be argued that the evidence fell into the wrong hands and may have been tampered with, or that other evidence was substituted. To prevent this from happening, policies and procedures dealing with the management of evidence must be adhered to.

DAMAGE AND LOSS CONTROL

Damage and loss control are steps in the process of attempting to reduce or minimize the impact of an incident. When an incident occurs, it is vital for members of the Incident Response Team to know what they should and shouldn't do to prevent a problem from spreading (as in the case of a virus), or stop an attacker from causing further damage. Because systems can be complex, these

procedures need to be documented before a problem occurs. Providing information on what not to do is just as important as knowing what to do during an incident.

REPORTING/DISCLOSURE

Procedures on disclosing and reporting information about an incident should also be outlined in an incident response policy. When an incident occurs, it may be up to a public relations person within the company to decide whether a media release is issued about the incident or if it will be kept quiet. Beyond disclosing an incident to the public, there are also other organizations that may be contacted, including:

- Operating system, application, or equipment manufacturers. If you believe the incident occurred due to vulnerabilities in a particular system, notifying the manufacturer of that software or hardware could help in having a security patch created to prevent the incident occurring again (to your company and others who use it).
- CERT (Computer Emergency Response Team), which is located at Carnegie Mellon University and coordinates communication during computer security emergencies. By notifying CERT (www.cert.org), others can be made aware of the attack so that it doesn't become widespread.
- Legal authorities. Contacting local police about an attack can begin the process of having an attacker arrested when he or she is found.

Defending against social engineering

Hacking may be done through expert computer skills, programs that acquire information, or through an understanding of human behavior. This last method is called *social engineering*. When social engineering is used, hackers misrepresent themselves or trick a person into revealing information. Using this method, a hacker may ask a user for their password or force the user to reveal other sensitive information.

Hackers using social engineering to acquire information will often misrepresent themselves as authority figures or someone in a position to help their victim. Social engineering often involves more subtle methods of acquiring information than simply asking for a password. In many cases, the hacker will get into a conversation with the user and slowly get the person to reveal tidbits of information. Social engineering is not confined to computer hacking.

PHISHING

A variation of social engineering is *phishing*, or *phising*, in which a hacker uses e-mail to acquire information from the recipient. A hacker will send e-mail to groups of people, posing as some authoritative source, and request the recipient to provide specific information. This may be a single department, the entire company, or (most often) sent as spam across the Internet. The e-mail asks for personal and credit card information or it links to a compromised

Web site where the hacker set up a form to capture such information, which can then be used to steal the person's identity. To prevent problems resulting from phishing, it is important to educate users and implement policies to specify how such information is to be collected.

HOAXES

E-mail hoaxes are those e-mails sent around the Internet about concerned parents desperately searching for their lost children, gift certificates being offered from retail stores for distributing e-mails for them, and dangerous viruses that have probably already infected the user's computer. There are a lot of different ways to separate hoaxes from real information. Most of the time, it comes down to common sense. The best rule of thumb is timeless—*if something seems too good to be true, it probably is.*

Virus hoaxes are warnings about viruses that do not exist. In these cases, the hoax itself becomes the virus because well-meaning people forward it to everyone they know. Some virus hoaxes are dangerous, advising users to delete certain files from their computer to "remove the virus," when those files are actually very important OS files. In other cases, users are told to e-mail information such as their password (or password file) to a specified address so the sender can "clean" the system of the virus. Instead, the sender will use the information to hack into the user's system and may "clean" it of its valuable data.

SHOULDER SURFING

Shoulder surfing is a method of obtaining passwords by watching what the person types on a keyboard, PIN (Personal Identification Number) pad, or other device that's used to enter a password. Users should be aware to protect what they are entering on a keypad.

DUMPSTER DIVING

Dumpster diving is the process of physically digging through a victim's trash in an attempt to gain information. Often it is easy to find client or product information, internal memos, and even password information that have been placed in wastebaskets. It is important to make sure that your organization has a method of securely disposing of the hard copies of confidential information.

There are many solutions to resolving dumpster diving as a security issue, including locking dumpsters with a padlock to limit access or keeping them in locked garages or sheds until they're ready for pickup. Companies can also implement a shredding policy, so that any sensitive information is shredded and rendered unusable by anyone who finds it. If documents aren't shredded, the recycling containers make it even easier to find information.

USER EDUCATION AND AWARENESS TRAINING

The best way to protect an organization from social engineering is through education. People reveal information to social engineers because they are

unaware they are doing anything wrong. Often they do not realize they have been victimized, even after the hacker uses the information for illicit purposes. There are many ways of disseminating educational material, inclusive of posting information on a corporate intranet site, e-mailed newsletters with tips on securing information, and having the information taught in formal training classes. Teaching users how social engineering works and stressing the importance of keeping information confidential will make them less likely to fall victim to social engineering.

SUMMARY OF EXAM OBJECTIVES

While many risks can negatively impact an organization's security, there are also many methods of prevention. HVAC systems are used to control temperature, humidity, and airflow, thereby preventing sensitive equipment from being damaged. Fire detection and prevention systems can also be implemented to warn and extinguish fires without harming the equipment.

Planning is the key to taking a proactive approach to possible threats. Disaster recovery plans provide procedures for recovering after a disaster occurs, and provides insight into methods for preparing for the recovery should the need arise. Incident response plans are similarly used to provide insight as to how Incident Response Teams should handle incidents.

Social engineering is another risk that organizations face, as it relies on taking advantage of human behavior rather than technology. A user may be asked questions that reveal seemingly innocuous information that can be pieced together to obtain a person's password, or phishing may be used to get the user to unwittingly reveal personal information and passwords. Such information can also be obtained through other methods, such as observing the user entering names, credit card numbers, or passwords on the computer. Another low-tech method is to simply look in the company's trash or recycling bins. This technique is called dumpster diving. The success of each of these methods relies on users being unaware that they're doing anything wrong, or how they can protect themselves.

TOP FIVE TOUGHEST QUESTIONS

1. Your organization is planning on installing a new fire suppression system in a server room. The system must be able to successfully extinguish the fire without causing damage to the servers and other equipment in the room. Which of the following will you use?

   A. Water sprinkler system
   B. A system that releases a fine mist of water to extinguish the fire
   C. A system that uses Halon to extinguish the fire
   D. A system that uses Inergen to extinguish the fire

2. You receive a complaint from the network administrator of another company regarding an attempted hacking of their Web site. Their firewall

logs show that the attempt came from an IP address from your company. Upon hearing the IP address, you find that this is the IP address of the proxy server belonging to your company. Further investigation on your part will be needed to identify who actually performed the attempted intrusion on the other company's Web site. Who will you notify of this problem before starting the investigation?

A. Media outlets to publicize the incident
B. The Incident Response Team
C. Users of the network to ensure they are aware that private information dealing with employees may need to be shared with the other company
D. No one

3. You are designing a backup regime that will allow you to recover data to servers in the event of a disaster. Should a disaster occur, you want to use a backup routine that will take minimal time to restore. Which of the following types of backups will you perform?

A. Daily full backups
B. A full backup combined with daily incremental backups
C. A full backup combined with daily differential backups
D. A combination of incremental and differential backups

4. You are the administrator of a network that is spread across a main building and a remote site several miles away. You make regular backups of the data on your servers, which are centrally located in the main building. Where should you store the backup tapes so they are available when needed in the case of a disaster?

A. Keep the backup tapes in the server room within the main building, so they are readily at hand. If a disaster occurs, you will be able to obtain these tapes quickly and restore the data to servers.
B. Keep the backup tapes in another section of the main building.
C. Keep the backup tapes in the remote site.
D. Keep the backup tapes in the tape drives of the servers so that a rotation scheme can be maintained.

5. You have created a backup regime as part of a disaster recovery plan. Each day, data on a server is backed up. After implementing it, you decide you want to make a separate backup of all data on the server but do not want it to interfere with the current backup jobs. Which of the following types of backups would you perform?

A. Full backup
B. Incremental backup
C. Differential backup
D. Copy backup

ANSWERS

1. The correct answer is D. Inergen 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. Answers A and B are incorrect because water-based systems can cause significant damage to equipment. C is incorrect because Halon isn't manufactured anymore or used in new systems because of the damage it causes to the ozone.

2. The correct answer is B. The Incident Response Team would deal with incidents such as hacking, and would be the appropriate people to notify. The Incident Response Team could assist or take over the investigation, and provide insight dealing with issues related to the intrusion attempt. Answers A, C, and D are incorrect because they would have no possible reason to be notified before conducting the investigation, if they ever are notified. Information that needs to be kept on a "need to know" basis means that only the people who need information are given it. At the beginning of an investigation, the Incident Response Team or a designated member of the company should only be notified.

3. The correct answer is A, Daily full backups. A full backup backs up all data in a single backup job. Because the data is backed up on a single tape or tape set, it will take the least amount of time to restore. While this may not be the most efficient method of performing backups, as combining full backups with incremental or differential backups takes less time to backup each day, it is the fastest to restore and uses fewer backup tapes. Answer B is incorrect because a combination of a full backup and daily incremental backups would take the least amount of time to backup each day, but the most amount of time to restore. When restoring the data, the full backup must be restored first, followed by each incremental backup that was taken since. Answer C is incorrect because a combination of a full backup with daily differential backups would require you to restore the last full backup and the last differential backup. This is still one more tape than if daily full backups were performed. Answer D is incorrect because incremental and differential backups cannot be combined together. Each would be part of a different backup regime and both would require a full backup to be restored.

4. The correct answer is C. Keep the backup tapes in the remote site. Since the company has a remote location that is miles from the main building, the tapes can be kept there for safekeeping. A firm can also be hired to keep the tapes in a storage facility. When a disaster occurs, you can then retrieve these tapes and restore the data. Answers A, B, and D are incorrect because a disaster that affects the server room or main building could also destroy the backup tapes if they were stored in these locations.

5. The correct answer is D. Copy backup. This type of backup is the same as a full backup, but does not change the archive bit. The archive bit is used to indicate that a file was backed up. Because the archive bit is not marked, any other backup jobs will not detect that a backup of data took place. As far as any incremental or differential setup on the server is concerned, it will appear as if the backup never took place. Answer A is incorrect because a full backup will change the archive bit, affecting any other backup jobs. Answer B is incorrect because an incremental backup will also change the archive bit, but is also wrong because a full backup would need to be used in conjunction with the incremental backup to acquire all data on the server. Answer D is incorrect because although a differential backup will not affect the archive bit, it will need to be used in conjunction with a full backup to acquire all data on the server. If the full backup were performed as part of this backup, the archive bit would be changed on files, affecting other backup jobs.

CHAPTER 14

Legislation and Organizational Policies

Exam objectives in this chapter:

- Secure Disposal of Systems
- Acceptable Use Policies
- Password Complexity
- Change Management
- Information Classification
- Vacations
- Personally Identifiable Information
- Due Care
- Due Process
- Due Diligence
- Service Level Agreements (SLAs)
- User Education and Awareness Training
- Security-Related HR Policies

SECURE DISPOSAL OF SYSTEMS

The first step regarding disposal and destruction is deciding what needs to be disposed of and destroyed. Because data can become obsolete or is legally required to be removed after a period of time, certain data needs to be removed from a system. Organizations often incorporate a data retention policy, which outlines the period of time when data and printed records become obsolete.

When files, records, or paperwork are destroyed, a policy dealing with disposal and destruction of data should be used. Such a policy can also be referred to when determining what to do with data that is destroyed daily, such as forms that are incorrectly filled out or corporate memos that are read but no longer needed. This policy provides clear guidelines of how an organization expects this material to be discarded.

Disk erasing software wipes the disk clean by erasing all of the files and overwriting the disk space with a series of ones and zeros. In doing so, every sector of the disk is overwritten, making the data unrecoverable. If anyone attempted

to recover data on the disk, that person wouldn't be able to retrieve anything because the data is completely destroyed. Shredder utilities like Active@ Kill Disk (www.killdisk.com) are widely used to wipe the disks before they are disposed.

A *degausser* or *bulk demagnetizer* is hardware that can be used to destroy data stored on magnetic media such as floppy disks and backup tapes. A degausser is a powerful magnet that erases all data from magnetic media so that no one can retrieve information from it. Hard disks can also have data erased with a degausser, performing a low level format that erases all data from the disk.

Retention/storage

A policy regarding the *retention* of data decides how long a company will retain data before destroying it. The length of time data is stored can be dictated by legal requirements or corporate decision making. Using this policy, certain data will be kept for a specified length of time, so that it can be referred to if needed. Retention and storage documentation is necessary to keep track of data, so that it can be determined what data should be removed and/or destroyed once a specific date is reached. Such documentation can be as simple as backup logs, which list what was backed up and when. By referring to the date the data was backed up, administrators can determine if the necessary period of time has elapsed to require destruction of this data.

EXAM WARNING

An organization should have clear policies on how long data and documentation are to be retained, and how this is to be stored. These policies ensure that data isn't destroyed too soon, and that it's stored in a safe and secure manner.

Destruction

When a retention period is reached, data needs to be destroyed. Legal requirements or policy may dictate how data is destroyed. When destroying data, it is important to follow procedures that dictate how information is to be destroyed. Even if data is destroyed on magnetic media, additional actions may be needed to destroy the media itself. Destroying the hard disks, floppy drives, backup tapes, and other media on which data is stored ensures that unauthorized persons are unable to recover data. Standard methods of physically destroying magnetic media include acid, pulverization, and incineration. When destroying data or equipment that is outdated, it is important that a log is kept of what items have been destroyed, and when and how the destruction was accomplished.

Crunch Time

Remember that how data is destroyed is as essential to maintaining privacy as storing it securely. Procedures need to be established on how to properly dispose of equipment, destroy data, and consistently purge systems of information. It's vital that outside individuals can't access data after equipment that is sold for auction or media that is thrown away.

ACCEPTABLE USE POLICIES

An *acceptable use policy* establishes guidelines on the appropriate use of technology. It is used to outline what types of activities are permissible when using a computer or network, and what an organization considers proper behavior. Acceptable use policies not only protect an organization from liability, but also provide employees with an understanding of what they can and cannot do using company resources.

Acceptable use policies also restrict the types of Web sites or e-mail an employee is allowed to access on the Internet. Acceptable use policies routinely include sections that restrict users from using equipment for their own personal use, home businesses, or other methods of financial gain.

Acceptable use policies should also specify methods of how information can be distributed to the public to avoid sensitive information from being "leaked." Imposing rules on the dissemination of information may include:

- Specifications that prohibit classified information from being transmitted via the Internet (e.g., e-mail, Short Message Service [SMS], or FTP)
- Provisions on how content for the Web site is approved
- Rules on printing confidential materials
- Restricting who can create media releases, and so on

PASSWORD COMPLEXITY

Passwords are used to prevent unauthorized access to computers, networks, and other technologies by forcing anyone who wants access to provide specific information. *Password management* involves enacting policies that control how passwords are used and administered. Without good password management, security could be compromised by passwords that are easy to guess, repeatedly used, or have characteristics that make them insecure. Because of the importance of password protection, a policy should state that users are responsible for their accounts and anything that is done with them.

Strong passwords

Authentication is used to prevent unauthorized access to computers, networks, and other technologies by forcing anyone who wants access to provide specific information. *Password management* involves enacting policies that control how passwords are used and administered. Without good password management, security could be compromised by passwords that are easy to guess, repeatedly used, or have characteristics that make them insecure. Because of the importance of password protection, a policy should state that users are responsible for their accounts and anything that is done with them.

Crunch Time

Remember that password complexity makes it more difficult for a password to be cracked. It should consist of a combination of uppercase letters, lowercase letters, numbers, and/or special characters. Just in case someone has your password, the password should be changed at intervals (such as every 90 days) and not be reused for a period of time.

Password changes and restrictions

Passwords should be changed after a set period of time, so that anyone who has a particular password will be unable to use it indefinitely and others will have more difficulty guessing it. A common recommendation is forcing users to change passwords every 45 or 90 days, at the most. While changing it often is more secure, it will make it more difficult for users to remember their passwords. As with any security measure, you want authorized users to easily access the system and unauthorized users to find it difficult. For this reason, the time limit set should allow users to memorize their new passwords before forcing them to change.

Administrator accounts

Administrator passwords are another important issue that should be covered in a password policy, as anyone using an administrative account is able to make changes and access all data on a system. Because of the importance of this account, there should be limits on who knows the password to this account.

CHANGE MANAGEMENT

Change management is the process of planning and implementing changes in systems. As an Information Technology department plans, upgrades, replaces servers, deploys new software, and makes other proactive changes, documentation is created to control how these changes take place. *Change control documentation* provides information on changes that have been made to a system, and

often provides back out steps that show how to restore the system to its previous state. Without this, changes made to a system could go unrecorded causing issues in the future.

INFORMATION CLASSIFICATION

A system of classification should be explained through a corporate policy, which defines the terms used and what they mean. When you are creating these classifications, the following levels should be included:

- ***Public*** or ***unclassified***, meaning that it can be viewed by people outside of the organization.
- ***Classified***, meaning that it is only for internal use, not for distribution to outside parties.
- ***Management only***, meaning that only managers and supervisors may view the information. This can be further broken down so that only certain levels of management can view it. For example, certain information may be suitable for top management but not for supervisors of individual departments.
- ***Department specific***, so that people outside of a particular department do not view the information.
- ***Private or confidential***, denotes that the information is only for the person to whom it was specifically sent.
- ***High security levels***, such as top secret or other classifications that stress the importance of the information.
- ***Not to be copied***, denoting that hard copies are not photocopied, and data files are not printed or copied to other media (such as floppy disk).

By providing a scheme of classification, members of an organization are able to understand the importance of information and less likely to leak sensitive information. Incorporating such a scheme will also make other policies more understandable, as they can describe what information is being discussed.

EXAM WARNING

Document management systems are increasingly used in organizations that need to maintain and track large stores of documents. Classification of these documents are important to ensuring that documents are disseminated to unauthorized individuals, their importance is quickly understood by readers, and that information isn't leaked by people who don't understand whether the document contains classified information.

VACATIONS

Vacation policies dictate how and when an employee may take a vacation. Components of a mandatory vacation policy include:

- How much time a person may have based on the number of years they've worked.

- Whether an employee can only take vacations at certain times of the year.
- If employees must take all of their vacation time at once, or can split it up throughout the year.

Mandatory vacation policies exist for a number of reasons. Contracts may require specific amounts of time off from work. By having employees take time off of work, they tend to be able to do their jobs better when they get back. Another reason is to prevent employees from carrying their vacation time over to subsequent years. Before having individuals take time off of work, it is important to ensure that the job can still be performed without their presence. This means having multiple people trained in different tasks.

EXAM WARNING

Mandatory vacation policies are covered in the exam, so don't skim over the information provided here believing it won't appear on the test. Vacations are important as they have implications to the business, can be legislated or contractually agreed on, and have security requirements for insuring that individuals are available to cover the duties of employees who are unavailable.

Separation of duties

Separation of duties ensures that tasks are assigned to personnel in a manner that no single employee can control a process from the beginning to end. Separation of duties is a common occurrence in secure environments and involves each person having a different job, thus allowing each to specialize in a specific area. This provides a number of benefits to the security of an organization.

In an organization that uses a separation of duties model, there is less chance of people leaking information because of the isolated duties that each employee performs in contribution to the whole. Another benefit of separating duties is that each person (or group of people) can become an expert in his or her job.

PERSONALLY IDENTIFIABLE INFORMATION

Personally identifiable information (*PII*) is private information that identifies you, members of your organization, and your clients. PII can be found in numerous places. It can exist in databases used by your company, directory services used in your network, and various other sources that contain names, phone numbers, addresses, credit card numbers, and so on. If such information became available to unauthorized users, it could result in embarrassment, liability, and possibly even criminal charges.

EXAM WARNING

PII goes hand in hand with privacy policies. Policies within the company should adhere to legislation that ensures personal data is secure.

Privacy

Privacy policies spell out the level of privacy that employees and clients can expect, and an organization's perspective of what is considered private information. Areas typically covered in a privacy policy are:

- Unauthorized software
- E-mail
- Web site data

Privacy policies have several components:

1. They commonly state that an organization has the right to inspect the data stored on company equipment. This allows an organization to perform audits on the data stored on hard disks of workstations, laptops, network servers, and other storage media.
2. They may also authorize such audits on the basis of searching for installations of pirated or unauthorized software. *Pirated software* is software that is not licensed for use by the person or company, and can cause liability issues resulting in fines or prosecution.
3. They often state that e-mail sent or received through business e-mail addresses belongs to the organization and should not be considered private. The organization can then examine the e-mail messages, ensuring that the business e-mail account is being used properly.
4. They can state that since the Internet access is provided through the company and is therefore their property, the company has the right to investigate how employees are using this resource.

DID YOU KNOW?

Once a policy is written, you need to ensure that leaders in the company will support it. Authorization needs to be acquired from management before the policy becomes active, so it is established that the company backs the policy and will enforce it if necessary. Having senior management sign off on a policy ensures that users will not be confused as to whether the policy is part of the company's vision and will result in disciplinary actions if violated.

The policy also needs to be reviewed by legal council to ensure it does not violate any laws, and that its content and wording is not misleading or unenforceable in any way. For example, many countries have legislation dealing with privacy, so it is important that whatever privacy policy you create adheres to those laws if your business operates in those countries. As with other policies mentioned here, you should have legal counsel review your policy before publishing it to the Internet or internally.

DUE CARE

Due care is the level of care that a reasonable person would exercise in a given situation, and is used to address problems of negligence. Due care may appear as a policy or concept mentioned in other policies of an organization. Put simply, an organization and its employees must be careful with equipment, data, and other elements making up the electronic infrastructure. Irresponsible use can cause liability risks for an organization, or result in termination of a careless employee.

DUE PROCESS

Due process is the act of notifying an employee that he or she has violated existing policies of legislation, and also refers to the employee's right into a fair and impartial inquiry into the incident. For example, if a person were accused of a violation of an acceptable use policy, he or she might be notified verbally and/or in writing. Due process ensures that the employee's rights have not been violated. If his or her rights were violated, it is possible that the company itself would then face litigation.

DUE DILIGENCE

Due diligence refers to the practices of an organization in identifying risks and implementing strategies to protect the assets of a company. Assets can include data, equipment, employees, and other elements that are of value to the company. By practicing due diligence, the company proves that it has taken reasonable steps to prevent an incident.

Organizations need to show they are diligent in upholding their policies by sharing them with employees (so they are aware of the rules), keeping them up to date, and enforcing them when necessary. A company can be seen as negligent if they don't take steps to ensure that policies addressing incidents are legally binding, topical, and are enforced when necessary.

Crunch Time

Don't get confused between *due care*, *due process*, and *due diligence*. Due care is used to show whether a reasonable level of care was given to protect data and equipment by an individual or a company. Due diligence shows that the company has consistently maintained and enforced their policies. In cases where policy violations occur, a fair and impartial inquiry into the incident and a person's misconduct is held. This protects the rights of the accused, and protects the company from litigation.

SLAs

Service level agreements (*SLAs*) are agreements between clients and service providers that outline what services will be supplied, what is expected from the service, and who will fix the service if it does not meet an expected level of performance. In short, it is a contract between the parties who will use a particular service and the people who create or maintain it. Through an SLA, the expectations and needs of all parties are clearly defined so that no misunderstandings about the system will occur at a later time.

EXAM WARNING

The Security+ exam expects that you understand that an SLA is used to establish an agreement between customers and the service provider as to the services available, and the requirements and conditions in providing them. Remember that SLAs are not only used between companies and third parties, but also as a commitment between internal IT staff and the organization's user base.

SLAs can also be used internally, specifying what users of the network can expect from IT staff and procedures relating to the network.

- The SLA may specify that all equipment (such as printers, new computers, and so forth) must be purchased through the IT department. If this is not done, the IT staff is under no obligation to fix the equipment that is purchased improperly.
- An SLA may also be used to specify the services the organization expects IT staff to provide, to support applications that are developed internally, or to address other issues related to the computers and network making up the organization's electronic infrastructure.

An SLA often includes information on the amount of downtime that can be expected from systems, during which customers will be unable to use a Web site, server, or other software and equipment. This information usually provides the expected availability of the system in a percentage format, which is commonly called the "Number of Nines."

USER EDUCATION AND AWARENESS TRAINING

Education and documentation are a vital part of any secure system. Knowledgeable users can be an important line of defense, as they will be better able to avoid making mistakes that jeopardize security, identify problems, and report them to the necessary persons.

Communication

The first step to creating good methods of communication is determining what methods are available. This differs from business to business, but multiple

avenues of contacting people are always available. These may include:

- Internal or Internet e-mail
- Internal phone extensions, home phone numbers, and cell phone numbers
- Pagers
- Corporate intranets and public Web sites
- Internal mail (memoranda) and snail mail (public postal services)
- Public folders and directories containing documents that can be viewed by users across the network
- Instant messaging, text messaging, SMS, and live chat

Once all of the methods available to communicate with users are identified, the administrator can decide which ones will be used and how. Providing contact information for IT staff ensures that incidents will not remain unattended and possibly grow worse before the next scheduled workday.

In addition to having people provide notification, administrators can configure systems to automatically contact them. Users should have multiple methods of contacting IT staff so that they can acquire help and notify them of problems they are experiencing. This allows users to inform administrators of a seemingly minor problem that could grow into a major one. There are many possible methods for users to contact IT staff. Help desks are commonplace in companies, providing a single phone extension that users can call when they are experiencing problems. Signatures on e-mails can be used to provide alternative methods of contacting individual users.

User awareness

Users cannot be expected to follow rules if they are not aware of them. Organizations sometimes make the mistake of imposing policies and procedures while failing to provide effective methods of sharing that information. This has the same effect as if the policies and procedures were never created. User awareness involves taking steps to make users conscious of and responsive to security issues, rules, and practices. To make users aware, administrators can use a number of the communications methods previously mentioned.

Education

Educating users is the primary method of promoting user awareness and improving the skills and abilities of employees. When users are taught how and why certain activities need to be performed, they are generally more willing and better able to perform those tasks. In addition to enhancing work performance, education also provides the added benefit of lowering support costs, as users who are able to fix simple problems will not be as likely to call the help desk for assistance.

Online resources

With the resources available on a local network, it would be remiss not to include them in the scheme of providing education and access to documentation.

Policies, procedures, and other documentation should be available through the network, as it will provide an easy, accessible, and controllable method of disseminating information. For example, administrators can make a directory on a server accessible to everyone through a mapped drive, allowing members of an organization to view documents at their leisure. A directory that is only accessible to IT staff can also be used to provide easy access to procedures, which may be referred to when problems arise. By using network resources this way, members of an organization are not left searching for information or left unaware of its existence.

SECURITY-RELATED HR POLICIES

Human resources (HR) departments deal with a large variety of issues, and need to work closely with IT departments to ensure security needs are met. HR performs such tasks as hiring, firing, retirement, and transferring employees to different locations. HR also maintains personnel files of employees, and may be responsible for assisting in the distribution of identification cards, key cards, and other items relating to security. Because of the tasks they each perform, it is important that good communication exists between HR and IT staff.

Adding or revoking passwords, privileges, and changes in a person's employment status can affect the person's security needs dramatically. A person may need to have a network account added, disabled, or removed, and other privileges (such as access to secure areas) may need to be modified. Adding or revoking passwords, privileges, and other elements of security may need to occur under such circumstances as:

- Resignation
- Termination
- New hires
- Changes in duties or position within the company
- Investigation
- Leave of absence

Disabling accounts and passwords should also occur when a person is away from a job for extended periods of time. When people are away from the job on parental leave, sabbaticals, and other instances of prolonged absence, they do not need their accounts to remain active. To prevent others from using the person's account while they are away, the account and password should be disabled immediately after the person leaves.

Code of Ethics

Many companies have a *code of ethics*, or a statement of mission and values, which outlines the organization's perspective on principles and beliefs that employees are expected to follow. Such codes generally inform employees that they are expected to adhere to the law, the policies of the company, and other professional ethics related to their jobs. As is the case with acceptable

use policies, many companies require employees to sign a code of ethics as an agreement. Anyone failing to adhere to this code could face dismissal, disciplinary actions, or prosecution.

SUMMARY OF EXAM OBJECTIVES

Policies provide information on the standards and rules of an organization, and are used to address concerns and identify risks. They are used to provide a reference for members of an organization, and are enforced to ensure they are followed properly. Procedures provide instructions on how policies are to be carried out, and may also be used to inform users on how to perform certain tasks and deal with problems. When used in an organization, policies provide a clear understanding of what they expect from employees and how issues are to be handled.

TOP FIVE TOUGHEST QUESTIONS

1. An organization has just installed a new T1 Internet connection, which employees may use to research issues related to their jobs and send e-mail. Upon reviewing firewall logs, you see that several users have visited inappropriate sites and downloaded illegal software. Finding this information, you contact senior management to have the policy relating to this problem enforced. Which of the following policies would you recommend as applicable to this situation?

   A. Privacy policy
   B. Acceptable use policy
   C. HR Policy
   D. SLAs

2. You are configuring operating systems used in your organization. Part of this configuration involves updating several programs, modifying areas of the Registry, and modifying the background wallpaper to show the company's new logo. In performing these tasks, you want to create documentation on the steps taken, so that if there is a problem, you can reverse the steps and restore systems to their original state. What kind of documentation will you create?

   A. Change control documentation
   B. Inventory
   C. Classification
   D. Retention and storage documentation

3. An organization has decided to implement a policy dealing with the disposal and destruction of data and other materials that may contain sensitive information. They have consulted you to determine what elements

should be included in the policy. Which of the following will you tell them?

A. Data on hard disks should be deleted before hard disks are disposed of.
B. Hard disks should be shredded before being disposed of.
C. Non-classified materials, such as media releases, should be shredded before being disposed of.
D. Classified documents should be shredded before being disposed of.

4. You are concerned about the possibility of sensitive information developed by your company being distributed to the public, and decide to implement a system of classification. In creating this system, which of the following levels of classification would you apply to sensitive information that is not to be disseminated outside of the organization?

A. Unclassified
B. Classified
C. Public
D. External

5. Changes in the law now require your organization to store data on clients for 3 years, at which point the data are to be destroyed. When the expiration date on the stored data is reached, any printed documents are to be shredded and media that contains data on the client is to be destroyed. What type of documentation would you use to specify when data is to be destroyed?

A. Disaster recovery documentation
B. Retention policies and logs
C. Change documentation
D. Destruction logs

ANSWERS

1. The correct answer is B. An acceptable use policy establishes guidelines on the appropriate use of technology. It is used to outline what activities are permissible when using a computer or network, and what an organization considers proper behavior. Acceptable use policies not only protect an organization from liability, but also provide employees with an understanding of what they can and cannot do when using technology. Answer A is incorrect because a privacy policy will outline the level of privacy an employee and/or customer can expect from the company. Privacy policies generally include sections that stipulate corporate e-mail as being the property of the company, and that Internet browsing may be audited. Answer **C** is incorrect because HR policies deal with the hiring, termination, and changes of an employee within a company. They do

not provide information on the acceptable use of technology. Answer D is incorrect because SLAs are agreements between clients and service providers that outline what services will be supplied, what is expected from the service, and who will fix the service if it does not meet an expected level of performance.

2. The correct answer is A. Change control documentation provides information of changes that have been made to a system, and often provides back out steps that show how to restore the system to its previous state. Answer B is incorrect because inventories provide a record of devices and software making up a network, not changes made to the configuration of those devices. Answer C is incorrect because classification is a scheme of categorizing information, so that members of an organization are able to understand the importance of information and less likely to leak sensitive information. Answer D is incorrect because retention and storage documentation is necessary to keep track of data, so that it can be determined what data should be removed and/or destroyed once a specific date is reached.

3. The correct answer is D. Classified documents should be shredded before being disposed of. Printed materials can still be accessed after they have been disposed of. Classified documents may contain sensitive information about the company, its clients, or employees. To prevent printed materials from getting into the wrong hands, the policy should specify that these types of documents should be shredded. Answer A is incorrect because even if data is deleted from a hard disk it may still be recovered. Answer B is incorrect because it is not a standard method of physically destroying magnetic media. Answer C is incorrect because non-classified materials such as media releases are not sensitive, and are cleared for public release. There is no problem with someone outside of the organization seeing this type of material.

4. The correct answer is B. When information is designated as classified, it means that it is for internal use only and not for distribution to parties outside of the organization. Answers A and C are incorrect because when information is classified as public or unclassified, then it can be viewed by parties outside of an organization. Answer D is incorrect because external documents are those generated outside of the organization.

5. The correct answer is B. Policy regarding the retention of data will decide how long the company will retain data before destroying it. Retention and storage documentation is necessary to keep track of this data, so that it can be determined what data should be removed and/or destroyed once a specific date is reached. Answer A is incorrect because disaster recovery documentation is used to provide information on how the company can recover from an incident. Answer C is incorrect because change documentation provides information on changes that have occurred in a system. Answer D is incorrect because destruction logs are used to chronicle what data and equipment have been destroyed after the retention date has expired.

Index

0-9, and Symbols
3DES (Triple DES), 141
802.1x authentication, 84
802.1x methods, 120–1
802.11 traffic, 80

A

Acceptable use policy, 195
Access control, 89–90, 110
  methods and models, 92
    discretionary access control, 94–5
    job rotation, 93
    least privilege, 93
    mandatory access control, 93–4
    role- and rule-based access control, 96
    separation of duties, 92
  models, 90–1
  organization, 97
    security controls, 98
    security groups, 97
Access control (SSH), 155
Access control, authentication, and auditing (AAA), 109
Access control lists (ACLs), 67, 93, 98–9
Access logs, 103, 132
Active@ Kill Disk, 194
Active/active cluster, 170
Active/passive cluster, 170
Active Scripting, 36, 37
ActiveX, 33–4
Ad-hoc networks, 80
Address Resolution Protocol (ARP) poisoning, 65–6
Adleman, 157
Administrator passwords, 196
Advanced Encryption Standard (AES), 142
Adware, 4–5
  protection against, 5
  and spyware, difference between, 5
Alternate sites, 167
  cold site, 167, 168
  hot site, 167, 168
  warm site, 167, 168
Altiris management software, 20
AMD-V, 51
Anomaly-based IDS, *see* Behavior-based IDS
Anti-SPAM, 9
Antivirus software, 3
Application filtering, 68
Application security, 31
  OSI model, 31
    rationale, 31–2
    threat modeling, 32
  packet sniffers and instant messaging, 42–4
  threats
    browser, 33–41
    buffer overflows, 41–2
Application-layer firewalls, 68–9
Application-layer gateways, 67, 68
ASN.1, 139
Asymmetric key cryptography, 136
Auditing, 111, 130, 132
Auditing systems, 131
Authentication, 110
Authentication Header (AH), 145, 146
Authentication methods
  in 802.11 standard, 83–4
  one-factor, 111
  single sign-on, 112
  three-factor, 112
  two-factor, 112
Authentication models and components, 91–2
Authentication tokens, 115

B

Backdoors, 2
Backup data, 181
Backup generator, 172–3
Basic Input/Output System, *see* BIOS
Bastille UNIX, 22, 23
Bastion host, 69
Behavior-based IDS, 6, 7
  vs. signature-based IDS characteristics, 7–9
Bell–LaPadula, 91
Biba formal model, 90–1, 105
Binary Translation, 51
Biometric devices, 113–14
  802.1x methods, 120–1
  CHAP, 118
  EAP, 121
  Kerberos, 114–15
  LDAP, 115–17
  mutual authentication, 119–20
  PAP, 117–18
  PEAP, 121–2
  RADIUS, 114
  TACACS/TACACS+, 118–19
BIOS, 10
BitTorrent, 43
Blind spoofing attacks, 64
Block symmetric algorithms, 154
Blu-Ray, 13
Bluebugging, 86
Bluejacking, 86
Bluesnarfing, 12, 86
Bluetooth, 12, 85–6
Botnets, 5, 6
Browser-based vulnerability, 33
Buffer overflows, 41–2
Bugs, 2
Bulk demagnetizer, *see* Degausser

C

Caesar cipher, 135
Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA), 80
Carrier Sense Multiple Access with Collision Detection (CSMA/CD) method, 79–80
CDs, 12–13
Cell phones, 11–12
Centralized door access system, 104
Centralized key management, 155
Certificate authority, 158–9
Certificate management, 162–3
Certificate revocation list (CRL), 156, 159–60

Certification authority (CA), 156, 158–9
  root CA, 158–9
  subordinate CA, 159
Chain of custody, 186
Challenge Handshake Authentication Protocol (CHAP), 54, 118
Change control documentation, 196
Change management, 196
Chip creep, 179
Chips, 79
CIA (Confidentiality, Integrity, and Availability), 140
Citrix XenApp, 55
Citrix XenServer, 51
Clark–Wilson Model, 91
CMOS (Complementary Metal Oxide Semiconductor), 10
Code of ethics, 203
Cold site, 167, 168
Common name, 117
Communication, 155, 201–2
Compatws template, 21
CompTIA Security + Objectives, 142
Computer Emergency Response Team's (CERT) Web site, 19
Computer forensics
  awareness, 184
  conceptual knowledge, 185
  understanding, 185
Computer virus, *see* Viruses
Confidentiality, Integrity, and Availability (CIA), 109, 140
Configuration baselines, 22–3
Content filters, 74
Control zone, 85
Cookie Monster, 38
Cookies, 37
  attacks, preventing against, 39
  hijacking, 38
  leaking, 39
  persistent cookies, 38
  poisoning, 39
  session cookies, 37
  tracking cookies, 38
  vulnerabilities, 38
Copy backup, 181
Corporate theft, 102
Crime scene technicians, 186
Cross-site scripting (XSS) attacks, 38–41
Cryptanalysis, 153
Cryptographic functions, types of, 154
Cryptographic Messaging Syntax (CMS), 144
Cryptography, 153
Cryptography concepts, 135
  asymmetric key cryptography, 136
  CIA (Confidentiality, Integrity, and Availability), 140
  digital certificate, 139
  digital signatures, 138–9
  dual-sided certificate, 139
  encryption algorithms, 141
    3DES (Triple DES), 141
    Advanced Encryption Standard (AES), 142
    DES (Data Encryption Standard), 141
    Elliptic Curve Cryptography, 142
    One-Time Pads, 142
    RSA (Rivest, Shamir, and Adleman), 141
    TKIP (Temporal Key Integrity Protocol.), 143
    transmission encryption, 142
    WEP, 143
  hashes and applications, 136–8
  key management, 140–1
  non-repudiation, 140
  in operating systems, 147
    E-mail, 148
    file and folder encryption, 147–8
    Trusted Platform Module (TPM), 148–9
    whole disk encryption, 148
  protocols, 143
    HTTP vs. HTTPS vs. SHTTP, 144
    IPSec, 145–6
    L2TP (Layer 2 Tunneling Protocol), 147
    PPTP (Point-to-Point Tunneling Protocol), 146–7
    S/MIME, 144–5
    Secure Socket Layer (SSL), 143
    SSH (Secure Shell), 145
    Transport Layer Security (TLS), 144
    with TLS, 144
  single certificates, 139
  symmetric key cryptography, 135–6

D

Damage and loss control, in incident response, 186–7
Data emanation, 85
Data repositories, 25–6
Database servers, 26
Decentralized key management, 155
Decrypting, 135
Degausser, 194
Demilitarized zone (DMZ), 67, 69–70
Denial of service (DoS) attacks, 65
DES (Data Encryption Standard), 141
DHCP servers, 25
Differential backup, 181, 183
Digital certificates, 139, 156
Digital Rights Management (DRM) technology, 148
Digital signatures, 138–9
Direct Attached Storage Devices (DASD), 53
Direct sequence spread spectrum (DSSS), 79
Directory, definition of, 115
Directory services, 25–6, 115
Disaster recovery, 180
  backup techniques and practices, 181
  disaster recovery plan, 180–1
  off-site storage, 182
  rotation schemes, 181–2
  secure recovery, 182
  secure recovery restoration, 183
Disclosing and reporting information, procedures on, 187
Discretionary access control (DAC), 94–5
Disk erasing software, 193–4
Distinguished name, 114
Distributed denial of service (DDoS) attack, 65
DNS cache poisoning, 65
DNS poisoning, 65
DNS servers, 24
Domain Group Policy, 21
Domain Name Kiting, 65
Domain policies, 99–100
Door access systems, 104–5
Drive-by-download attacks, 33
Dual-sided certificate, 139
Dual key certificates, 139
Dual key pair, *see* Dual-sided certificate
Due care, 199–200
Due diligence, 200
Due process, 200
Dumpster diving, 188, 189

DVDs, 12–13
Dwell time, 79

E

E-mail, 9, 148, 199
E-mail hoaxes, 188
Educating users, 202
Electromagnetic field (EM), 77
Electromagnetic interference (EMI), 179
Electrostatic discharge (ESD), 179
Elliptic Curve Cryptography, 142
Emanation, 85
EMule, 43
Encapsulating Security Payload (ESP), 145, 146
Encrypting, 135
Encrypting File System (EFS), 148
Encryption algorithms, 141
  3DES (Triple DES), 141
  Advanced Encryption Standard (AES), 142
  DES (Data Encryption Standard), 141
  Elliptic Curve Cryptography, 142
  One-Time Pads, 142
  RSA (Rivest, Shamir, and Adleman), 141
  TKIP (Temporal Key Integrity Protocol.), 143
  transmission encryption, 142
  WEP, 143
End User License Agreement (EULA), 5
Enhanced Key Usage values, 139
Environment, meaning of, 177
Environmental controls, 177
  fire suppression, 177
    detection systems, 178
  HVAC, 178–9
  shielding, 179–80
Extensible Authentication Protocol (EAP), 121
Extranets, 68

F

Factors of authentication, 92
Failover server, 169–70
Faraday cage, 85
Fault tolerance, 169
Fiber Channel SAN, 53
Fiber Channel Security Protocol (FC-SP), 54
Fiber-optic cable, 180
File and folder encryption, 147–8
File and print servers, 25
Fire suppression systems, 177
  detection systems, 178
Firewalls, 66–9
First responders, 185–6
Flame detection, 178
Flash memory cards, 11
Frequency hopping spread spectrum (FHSS), 78–9
Fresnel zone, 78
FTP servers, 23–4
Full backup, 181, 183

G

General OS hardening, 17
  configuration baselines, 22–3
  file system, 18
  hotfixes/patches, 19
  patch management, 19–20
  security templates, 21–2
  service packs/maintenance updates, 19
  services, 18
  unnecessary programs, removing, 18
  Windows Group Policies, 21
Gramm–Leach–Bliley Act (GLBA), 127
Grandfather-Father-Son (GFS) rotation, 181–2
Group policies, 99
  in Windows, 21
Group Policy Object (GPO), 99
Groups, 97
Halon, 177

H

Hardware and peripheral security risks, 9–10
  BIOS, 10
  cell phones, 11–12
  network attached storage, 13
  removable storage devices, 12–13
  USB devices, 10–11
Hardware Assist, 51
Hardware locks, 103
Hardware Storage Modules (HSMs), 158
Hash, 136, 137
Hashes and applications, 136–8
Health Insurance Portability and Accountability Act (HIPAA), 127
Heap overflows, 41
Heat detection, 178
High availability, 169
Hisecws template, 21
Honeynets, 73–4
Honeypots, 73–4
Hop time, 79
Host-based IDS (HIDS), 6
Host Bus Adapter (HBA), 53
Host intrusion detection system, 6–7
  behavior-based vs. signature-based IDS characteristics, 7–9
Hosted virtualization, 50
Hot site, 167, 168
Hot spare, 172
Hot swap, 172
Hotfixes, 19
Hotfixes/patches, 19
HTTP vs. HTTPS vs. SHTTP, 144
HVAC (Heating, Ventilation, and Air Conditioning), 178–9
Hypertext Transfer Protocol (HTTP), 160
Hypervisor, 50, 51

I

IBM x3350 server, 56
ID badges, 103–4
Incident response, 183
  chain of custody, 186
  damage and loss control, 186–7
  first responders, 185–6
  forensics, 184–5
  Incident Response Teams, 183, 184
  reporting/disclosure, 187
Incident response policy, 183
Incidents, 183
Incremental backup, 181, 183
Information classification, 196–7
Informed attacks, 64
Input validation, 41–2
Instant messaging (IM), 42–3
Integrity Check Value (ICV), 146
Intel VT, 51
Intermediate CAs, 159
Internet Engineering Task Force (IETF), 143
Internet Key Exchange (IKE), 145, 146
Internet Protocol Security (IPSec), 145–6
Internet Service Provider (ISP), 170–1

Internetwork Packet Exchange (IPX), 62
Intrusion detection system (IDS), 6–9, 72, 130
Intrusion prevention system (IPS), 8
IP spoofing, 63–4
iSCSI SAN, 54
ITU-T X.509, 139

J

Jamming, 85
Java, 35, 36
Java Runtime Environment (JRE), 35
Java Virtual Machine (JVM), 35
Javascript, 36, 37
Job rotation, 93
Jscript, 36, 37

K

Kerberos, 114–15
Key escrow, 160
Key management, 140–1
Key recovery agents, 161
Key Recovery Information (KRI), 161
Key Usage value, 139

L

L2TP (Layer 2 Tunneling Protocol), 147
LANMAN (LAN Manager), 138
LDAP, 115–16
  directories, 115
  objects, attributes, and the schema, 116–17
  organizational units, 116
  securing, 117
Least privilege, 93
"Least privileged" principle, 18
Legislation and organizational policies, 193
  acceptable use policies, 195
  change management, 196
  due care, 199–200
  due diligence, 200
  due process, 200
  information classification, 196–7
  password complexity, 195
    administrator accounts, 196
    password changes and restrictions, 196
    strong passwords, 195
  personally identifiable information (PII), 198–9
  secure disposal of systems, 193
    destruction, 194
    retention/storage, 194
  security-related HR policies, 203
    code of ethics, 203
  service level agreements (SLAs), 200–1
  user education and awareness training, 201
    communication, 201–2
    education, 202
    online resources, 202–3
    user awareness, 202
  vacations, 197
Linux Security Modules (LSM), 22
LM Hash, *see* LANMAN
Local area network (LAN), 66
Local Group Policy, 21
Logging and auditing, 130–2
Logic bomb, 6
Logical access control methods, 98
  access control lists, 98–9
  account expiration, 100
  domain policies, 99–100
  group policies, 99
  logical tokens, 100–1
  time of day restrictions, 100
Logical Link Control (LLC) layer, 79
Logical tokens, 100–1

M

Magnetic tapes, 13
Malware, 1
Man-trap, 105
Mandatory access control (MAC), 93–4
Mandatory vacation policies, 197–8
Man-in-the-middle (MITM) attack, 63, 64
MD5 (Message Digest 5), 138
Media Access Control (MAC) layer, 65, 66, 79
Microsoft Active Directory technology, 99
Microsoft Baseline Security Analyzer (MBSA), 23
Microsoft Hyper-V, 51
Microsoft Terminal Services, 55
Microsoft updates, 19
Microsoft Virtual Server 2005; 50
MTA (Mail Transport Agent), 148
MUA (Mail User Agent), 148
Multi-core processors, 52
Multifactor authentication, *see* Three-factor authentication
Multipath interference, 78
Mutual authentication, 119–20

N

National Security Agency, 154
.NET, 36
NetBIOS Extended User Interface (NetBEUI), 25
Netscape, 143
Network access
  access control, 89–90
    models, 90–1
    security controls, 98
    security groups, 97
  authentication models and components, 91–2
  identity, 92
  logical access control methods, 98
    access control lists, 98–9
    account expiration, 100
    domain policies, 99–100
    group policies, 99
    logical tokens, 100–1
    time of day restrictions, 100
  methods and models, 92
    discretionary access control, 94–5
    job rotation, 93
    least privilege, 93
    mandatory access control, 93–4
    role-and rule-based access control, 96
    separation of duties, 92
  physical access security methods, 101–2
    access lists and logs, 102–3
    door access systems, 104–5
    hardware locks, 103
    ID badges, 103–4
    man-trap, 105
    video surveillance, 105
Network Access Control (NAC), 26, 71–2
Network access protection, 71–2
Network address translation (NAT), 71
Network Attached Storage, 54
Network attached storage (NAS), 13
Network authentication
  authentication methods
    one-factor, 111
    single sign-on, 112
    three-factor, 112
    two-factor, 112

authentication systems, 113
biometrics, 113–22
remote access policies and authentication, 113
methods, 109
access control, 110
auditing, 111
authentication, 110
Network design elements, 61
Demilitarized zone (DMZ), 69–70
firewalls, 66–9
network access control (NAC), 71–2
network access protection, 71–2
network address translation (NAT), 71
telephony, 72
VLANs, 70–1
Network File System (NFS), 54
Network keys, 120–1
Network mapping tools, 129
Network ports and protocols, 62
Network security, 61
network design elements, 61
Demilitarized zone (DMZ), 69–70
firewalls, 66–9
network access control (NAC)/ network access protection, 71–2
network address translation (NAT), 71
telephony, 72
VLANs, 70–1
network ports and protocols, 62
network services and risks, 62
network threats, 62
Address Resolution Protocol (ARP) poisoning, 65–6
denial of service (DoS) attacks, 65
distributed denial of service (DDoS) attack, 65
DNS poisoning, 65
Domain Name Kiting, 65
IP spoofing, 63–4
man-in-the-middle (MITM) attack, 64
null sessions, 63
replay attack, 64
TCP/IP hijacking, 63
network tools, 62
content filters, 74
honeypots, 73–4
Intrusion detection and preventions systems, 72–3
protocol analyzer, 74
Network services and risks, 62
Network-based IDS (NIDS), 6
NIST (National Institute of Standards and Technology), 137
NNTP servers, 24
Non-repudiation, 110, 140
Non-secret encryption, 136
NT Hash, 138
NTFS (New Technology File System), 148
NTLM, 138
Null sessions, 63

O

Off-site storage, 182
One -Time Pads, 142
One-factor authentication, 111
Online Certificate Status Protocol (OCSP), 159–60
Online resources, 202–3
Open authentication, 83
Open Systems Interconnect (OSI) model, 31, 66
rationale, 31–2
threat modeling, 32
Open Vulnerability and Assessment Language (OVAL), 128
Organizational Unit Group Policy, 21
Organizational units (OUs), 116
OS hardening, 17
general OS hardening, 17
configuration baselines, 22–3
file system, 18
hotfixes/patches, 19
patch management, 19–20
security templates, 21–2
service packs/maintenance updates, 19
services, 18
unnecessary programs, removing, 18
Windows group policies, 21
server OS hardening, 23
data repositories, 25–6
DHCP servers, 25
DNS servers, 24
file and print servers, 25
FTP servers, 23–4
NNTP servers, 24
services and protocols, enabling and disabling, 23
workstation OS, 27
user rights and groups, 27–8

P

Packet filtering, 66
Packet sniffers, 42
Paravirtualization, 51
Password Authentication Protocol (PAP), 117–18
Password changes and restrictions, 196
Password complexity, 195
administrator accounts, 196
password changes and restrictions, 196
strong passwords, 195
Password crackers, 128–9
Password management, 195
Password policies, 111
Patches, 19–20
Payload, 2
Peer-to-Peer (P2P) networks, 43
Performance Logs, 132
Persistent cookies, 38
Personal Computer Memory Card International Association (PCMCIA) cards, 158
Personally identifiable information (PII), 198–9
Phishing, 187–8
Phreakers, 72
Physical access security methods, 101–2
access lists and logs, 102–3
door access systems, 104–5
hardware locks, 103
ID badges, 103–4
man-trap, 105
video surveillance, 105
Piggybacking, 105
Pirated software, 199
PKI encryption, 154
PKI solutions, 155
PKI standards, 154
Pop-up blocker, 5, 9
Popup Test Web site, 9
Port scanner, 127
Power generators, 173
"Power Users" group, 27
PPP (Point-to-Point Protocol), 146
PPTP (Point to-Point Tunneling Protocol), 146–7

Pre-shared key (PSK), 145
Privacy policies, 199
Private Communication Technology (PCT), 143
Private key, 136, 155, 157–8
Privilege escalation, 1–2
Protected Extensible Authentication Protocol (PEAP), 121–2
Protocol analyzers, 74, 128
Proxy server, 73
Public key, 136, 155
Public key certificate, 155
Public key infrastructure (PKI), 153
  certificate management, 162–3
  components of, 155
    certificate authority, 158–9
    certificate revocation list, 156
    certificate revocation list, 159–60
    certification authority, 156
    digital certificates, 156
    key escrow, 160
    recovery agents, 156–8
  implementation, 161–2
  overview, 153
    PKI encryption, 154
    PKI solutions, 155
    PKI standards, 154
  recovery agents, 161
  registration, 160–1
Public keys, 156–7
Public-Key Cryptography Standards (PKCS), 154

R

Radio frequency (RF), 77
Radio frequency interference (RFI), 179
Rainbow table, 128–9
Realm, 115
Recovery agents, 156–8
Redundancy, 169
Redundancy planning, 167
  alternate sites, 167
    cold site, 167, 168
    hot site, 167, 168
    warm site, 167, 168
  backup generator, 172–3
  RAID, 171–2
  redundant systems, 169
    connections, 170
    ISP, 170–1
    servers, 169–70
  spare parts, 172
  UPS, 173
Redundant Arrays of Inexpensive Disks (RAID), 53, 171–2
Redundant systems, 169
  connections, 170
  ISP, 170–1
  servers, 169–70
Reflected XSS attacks, 39–40
Registration authorities (RA), 160–1
Relative distinguished name, 117
Remote Authentication Dial-In User Service (RADIUS), 114
Remote Desktop Services, 55
Removable media, *see* Removable storage devices
Removable storage devices, 12–13
Replay attack, 64
Rijndael, 142
Risk assessment and risk mitigation, 127
  audits, 132
  intrusion detection system, 130
  logging and auditing, 130–2
  monitoring tools usage, 129–30
  network mapping tools, 129
  password crackers, 128–9
  vulnerability assessment tools, 127–8
Rivest, 157
Robust Security Network (RSN), 82
Rogue access points, 84
Role-and rule-based access control (RBAC), 96
Roles, 97
Root CA, 158–9
Rootkits, 5–6
Rotation schemes, 181–2
RSA (Rivest, Shamir, and Adleman), 141

S

S/MIME, 144–5
Salt, 137
Schema, 117
Scripting languages, 36–7
Scripts, 20
Secure disposal of systems, 193
  destruction, 194
  retention/storage, 194
Secure e-mail, 155
Secure recovery, 182
Secure recovery restoration, 183
Secure Socket Layer (SSL), 143
Secure web access, 155
Security Association (SA), 145, 146
Security controls, 98
Security Enhanced (SE) Linux, 22
Security groups, 97
Security Parameters Index (SPI), 146
Security-related HR policies, 203
  code of ethics, 203
Security templates, 21–2
SecurityFocus Web site, 19
Self-signed certificate, 159
Separation of duties, 92
Sequenced Packet Exchange (SPX), 62
Server clusters, 170
Server Core, 51
Server OS hardening, 23
  data repositories, 25–6
  DHCP servers, 25
  DNS servers, 24
  file and print servers, 25
  FTP servers, 23–4
  NNTP servers, 24
  services and protocols, enabling and disabling, 23
Service level agreements (SLAs), 200–1
Service packs, 19
Service packs/maintenance updates, 19
Service Set Identifier (SSID), 80–1
Session cookies, 37
Session hijacking, *see* TCP/IP hijacking
SHA (Secure Hash Algorithm), 137
Shamir, 157
Shared-key authentication, 83
Shielded twisted pair (STP), 180
Shielding, 179–80
Shoulder surfing, 188
Signature-based IDS, 7
  vs.behavior-based IDS, 7–9
Signature files, 3
Simple Mail Transfer Protocol (SMTP) open relays, 43–4
Single CA model, 159
Single certificates, 139
Single Sign-On (SSO), 112
Site Group Policy, 21
Smart Cards, 158
Smoke detection, 178
SMS 2003 and System Center, 20
Social engineering, defending against, 187
  dumpster diving, 188
  hoaxes, 188
  phishing, 187–8

shoulder surfing, 188
user education and awareness training, 188–9
Socket creep, *see* Chip creep
SPAM, 9, 43–4
Spare parts, 172
Spoofing, 115
Spread spectrum technology, 78
direct sequence spread spectrum, 79
frequency hopping spread spectrum, 78–9
Spreading ratio, 79
Spyware, 4
and adware, difference between, 5
protection against, 5
SSH (Secure Shell), 145
SSLv3, 143
Stack overflows, 41
Standalone door access systems, 104
Sticks, 11
Storage Area Network (SAN), 53
Storage Root Key (SRK), 149
Stored XSS attacks, 40
Stream symmetric algorithms, 154
Strong passwords, 195
Subordinate CA, 159
Symmetric algorithms, types of, 154
Symmetric key cryptography, 135–6
SYN packet, 64
System Logs, 131
System Management Server (SMS)/ System Center, 20
Systems security, 1
anti-SPAM, 9
hardware and peripheral security risks, 9–13
host intrusion detection system, 6–9
pop-up blocker, 9
threats
logic bomb, 6
privilege escalation, 1–2
rootkits and botnets, 5–6
spyware and adware, 4–5
Trojan, 4
viruses and worms, 2–4

T

TACACS, 118
TACACS+, 118–19
TACACS/TACACS+, 118
Tailgating, 105
TCP/IP hijacking, 63
Telephony, 72
TEMPEST project, 85
Temporal Key Integrity Protocol (TKIP), 82
Threat modeling, 32
Three-factor authentication, 112
Time of day restrictions, 100
TKIP (Temporal Key Integrity Protocol.), 143
Token authentication, 112
Tracking cookies, 38
Transmission encryption, 142
Transport Layer Security (TLS), 144
Trojan horse, 4
Truman, 154
Trunk exception, 70
Trusted Platform Module (TPM), 148–9
Trusted third party (TTP), 153, 155
Two-factor authentication, 112

U

Uninterruptible power supplies (UPS), 173
Universal Serial Bus (USB), *see* USB devices
Unnecessary programs, removing, 18
Unsolicited bulk e-mail (UBE), *see* Anti-SPAM
U.S. Department of Defense Trusted Computing System Evaluation Criteria (TCSEC), 90
USB devices, 10–11
USB Flash Drives, 11
User awareness, 202
User education and awareness training, 201–3
User rights and groups, 27–8

V

Vacation policies, 197
duties, separation of, 198
VBScript, 36
Versa Corp., 56, 57
Video surveillance, 105
Virtual applications, 49
Virtual environment, designing, 51
networking, 52
processors, 51–2
storage, 53–4
Virtual private networks (VPNs), 67, 155
Virtualization technologies, 49
application, 55–6
benefits, 49–50
purpose, 49
system virtualization, 54
virtual servers, management of, 55
types, 50–1
virtual environment, designing, 51
networking, 52
processors, 51–2
storage, 53–4
Virus hoaxes, 188
Viruses, 2
protection against, 3–4
and worms, difference between, 3
Vista Security Guide, 21
VLANs, 70–1
VMware ESX 3.5, 51
VMware ESX server, 51
VMware Virtual Server, 50
Vulnerability assessment tools, 127–8
Vulnerability scanners, 127–8

W

Warm site, 167, 168
Warm swap, 172
Web browser, 33
Whole disk encryption, 148
Wi-Fi Protected Access (WPA), 81–2
Wide area network (WAN), 170
Windows Group Policies, 21
Windows Software Update Services (WSUS), 20
Windows Vista, security templates for, 21
Wired Equivalency Privacy (WEP) keys, 120
Wired Equivalent Privacy (WEP), 81, 143
Wireless Application Protocol (WAP), 82
Wireless networks, 77
Bluetooth, 85–6
CSMA/CD and CSMA/CA, 79–80
data emanation, 85
design, 77
architecture, 79
communications, 77–8
spread spectrum technology, 78–9
rogue access points, 84
security standards, 81
authentication, 83–4
WAP, 82

Wireless networks (*Continued*)
WEP, failure of, 81
WPA and WPA2, 81–2
WTLS, 82–3
Service Set ID Broadcast, 80–1
Wireless Transport Layer Security (WTLS), 82–3
Workstation OS, 27
user rights and groups, 27–8
Workstations, 129–30
Worms, 2, 3
protection against, 3–4
and viruses, difference between, 3
WPA2, 82

X

X.509, 139, 156, 161–2

Z

Zero-day attack, 4
Zombies, 65
Zone, definition of, 34