Cybersecurity Concepts

Malware Types

Malware includes a wide range of software that has malicious intent. Malware is not software that you would knowingly purchase or download and install. Instead, it is installed onto your system through devious means. Infected systems give various symptoms, such as running slower, starting unknown processes, sending out email without user action, rebooting randomly, and more.

Malware Types

- Viruses
- Worms
- Ransomware
- Crypto-Malware
- Trojan Horse
- Rootkit
- Keylogger
- Adware, Spyware
- Bot
- Botnet
- Logic Bomb

Day in the Life of a Hacker

350,000 NEW malware is detected every day. (AV-Test Institute)

7 billion malware attacks were discovered in 2019. (SonicWall)

Malware Process

1. A virus takes advantage of a vulnerability.
2. Installs malware that includes a remote access backdoor. The bot may be installed later.
3. The target computer must run a program. Three mechanisms for execution of a malicious program are: Clicking on an email link, Web page pop-up, Drive-by download.

Viruses

A virus is malicious code that attaches itself to a host application. The host application must be executed to run, and the malicious code executes when the host application is executed. The virus tries to replicate by finding other host applications to infect with the malicious code. At some point, the virus activates and delivers its payload. Typically, the payload of a virus is damaging. It may delete files, cause random reboots, join the computer to a botnet, or enable backdoors that attackers can use to access systems remotely. Some older viruses merely displayed a

message or pop-up. Most viruses won't cause damage immediately. Instead, they give the virus time to replicate first. A user will often execute the virus (though unknowingly), but other times, an operating system will automatically execute it after user interaction. For example, when a user plugs in an infected USB drive, the system might automatically execute the virus, infecting the system.

Worms

A worm is self-propagating malware that travels throughout a network without the assistance of a host application or user interaction. A worm resides in memory and can use different transport protocols to travel over the network. One of the significant problems caused by worms is that they consume network bandwidth. Worms can replicate themselves hundreds of times and spread to all the systems in the network. Each infected system tries to locate and infect other systems.

Backdoors

A backdoor provides another way of accessing a system, similar to how a backdoor in a house provides another method of entry. Malware often installs backdoors on systems to bypass normal authentication methods. While application developers often code backdoors into applications, this practice is not recommended. For example, an application developer might create a backdoor within an application intended for maintenance purposes. However, if attackers discover the backdoor, they can use it to access the application.

Effective account management policies help prevent ex employees from creating backdoors after they are fired. For example, if an employee loses network access immediately after being fired, the employee cannot create a backdoor account. In contrast, if an administrator retains network access, he might create another administrative account. IT personnel might disable his account after they learn he has been fired, but he can still use this new backdoor account. That's exactly what a John Doe Unix engineer did after being told he was fired.

John Doe's account management policy did not revoke his elevated system privileges right away, giving him time to create a backdoor account. After going home, he accessed the system remotely and installed a logic bomb script scheduled to run at 9:00 a.m. on January 31. If another administrator hadn't discovered the logic bomb, it would have deleted data and backups for about four thousand servers, changed their passwords, and shut them down.

Ransomware

A specific type of Trojan is ransomware. Attackers encrypt the user's data or take control of the computer and lock out the user. Then, they demand that the user pay a ransom to regain access to the data or computer. Criminals often deliver ransomware via drive-by downloads or embedded in other software delivered via email. Attackers originally targeted individuals with ransomware. However, they have increasingly been targeting organizations demanding larger and larger ransoms. Ransomware types continue to evolve. In early versions, they sometimes just locked the user out of the system. However, this is rarely done anymore. Instead, attackers typically encrypt the user's data to ensure that users can't retrieve it. Ransomware that encrypts the user's data is sometimes called crypto-malware.

Some ransomware has added in a new blackmail technique called doxing. If the user doesn't pay the ransom to decrypt the files, the attacker threatens to publish the

files along with the victim's credentials. Malware that uses doxing is sometimes called doxingware.

Trojan Horse

A trojan, also called a Trojan horse, looks like something beneficial, but it’s actually something malicious. Trojan horses are named after the infamous horse from the Trojan War. In computers, a Trojan horse can come as pirated software, a useful utility, a game, or something else that users might be enticed to download and try. Attackers are increasingly using drive-by downloads to deliver Trojans. In a drive-by download, web servers include malicious code that attempts to download and install itself on user computers after the user visits. Here are the typical steps involved in a drive-by download.

1. Attackers compromise a web site to gain control of it.
2. Attackers install a Trojan embedded in the web site’s code.
3. Attackers attempt to trick users into visiting the site. Sometimes, they simply send the link to thousands of users via email hoping that some of them click the link.
4. When users visit, the web site attempts to download the Trojan onto the users’ systems.

Remote Access Trojans (RATs)

A remote access Trojan (RAT) is a type of malware that allows attackers to take control of systems from remote locations. It is often delivered via drive-by downloads. Once installed on a system, attackers can then access the infected computer at any time, and install additional malware if desired.

Some RATs automatically collect and log keystrokes, usernames and passwords, incoming and outgoing email, chat sessions, and browser history as well as take screenshots. The RAT can then automatically send the data to the attackers at predetermined times.

Additionally, attackers can explore the network using the credentials of the user or the user’s computer. Attackers often do this to discover, and exploit, additional vulnerabilities within the network. It’s common for attackers to exploit this one infected system and quickly infect the entire network with additional malware, including installing RATs on other systems.

Rootkits

A rootkit is a group of programs that hides the fact that the system has been infected or compromised by malicious code. A user might suspect something is wrong, but antivirus scans and other checks indicate everything is fine because the rootkit hides its running processes to avoid detection.

In addition to modifying the internal operating system processes, rootkits often modify system files such as the Registry. In some cases, the rootkit modifies system access, such as removing users’ administrative access.

Rootkits have system-level access to systems. This is sometimes called root-level access, or kernel-level access, indicating that they have the same level of access as the operating system. Rootkits use hooked processes, or hooking techniques, to intercept calls to the operating system. In this context, hooking refers to intercepting system-level function calls, events, or messages. The rootkit installs the hooks into memory and uses them to control the system’s behavior.

Antivirus software often makes calls to the operating system that could detect malware, but the rootkit prevents the antivirus software from making these calls. This is why antivirus software will sometimes report everything is OK, even if the system is infected with a rootkit. However, antivirus software can often detect the hooked processes by examining the contents of the system's random access memory (RAM).

Another method used to detect rootkits is to boot into safe mode, or have the system scanned before it boots, but this isn't always successful. It's important to remember that rootkits are very difficult to detect because they can hide so much of their activity. A clean bill of health by a malware scanner may not be valid.

It's important to remember that behind any type of malware, you'll likely find an attacker involved in criminal activity. Attackers who have successfully installed a rootkit on a user's system might log on to the user's computer remotely, using a backdoor installed by the rootkit. Similarly, attackers might direct the computer to connect to computers on the Internet and send data. Data can include anything collected from a keylogger, collected passwords, or specific files or file types stored on the user's computer.

Keyloggers

A keylogger attempts to capture a user's keystrokes. The keystrokes are stored in a file, and are either sent to an attacker automatically, or the attacker may manually retrieve the file.

While a keylogger is typically software, it can also be hardware. For example, you can purchase a USB keylogger, plug it into the computer, and plug the keyboard into the USB keylogger. This hardware keylogger will record all keystrokes and store them within memory on the USB device.

Adware

When adware first emerged, its intent was primarily to learn a user's habits for the purpose of targeted advertising. As the practice of gathering information on users became more malicious, more people began to call it spyware. However, some traditional adware still exists. Internet marketers have become very sophisticated and use a combination of web analytics with behavioral analytics to track user activity. They then provide targeted ads based on past user activity.

The term adware also applies to software that is free but includes advertisements. The user understands that the software will show advertisements and has the option to purchase a version of the software that does not include the ads. All of this is aboveboard without any intention of misleading the user.

Spyware

Spyware is software installed on users' systems without their awareness or consent. Its purpose is often to monitor the user's computer and the user's activity. Spyware takes some level of control over the user's computer to learn information and sends this information to a third party. If spyware can access a user's private data, it results in a loss of confidentiality.

Some examples of spyware activity are changing a user's home page, redirecting web browsers, and installing additional software within the browser. In some situations, these changes can slow a system down, resulting in poorer performance. These examples

are rather harmless compared with what more malicious spyware (called privacy-invasive software) might do.

Privacy-invasive software tries to separate users from their money using data-harvesting techniques. It attempts to gather information to impersonate users, empty bank accounts, and steal identities. For example, some spyware includes keyloggers. The spyware periodically reads the data stored by the keylogger, and sends it to the attacker. In some instances, the spyware allows the attacker to take control of the user's system remotely.

Spyware is often included with other software like a Trojan. The user installs one application but unknowingly gets some extras. Spyware can also infect a system in a drive-by download. The user simply visits a malicious web site that includes code to automatically download and install the spyware onto the user's system.

Bots & Botnets

Generically, bots are software robots. For example, Google uses bots as search engine spiders to crawl through the Internet looking for web pages. However, attackers also use bots for malicious purposes. A botnet combines the words robot and network. It includes multiple computers that act as software robots (bots) and function together in a network (such as the Internet), often for malicious purposes. The bots in a botnet are often called zombies and they will do the bidding of whoever controls the botnet.

Bot herders are criminals who manage botnets. They attempt to infect as many computers as possible and control them through one or more servers running command-and-control software. The infected computers periodically check in with the command-and-control servers, receive direction, and then go to work. The user is often unaware of the activity.

Most computers join a botnet through malware infection. For example, a user could download pirated software with a Trojan or click a malicious link, resulting in a drive-by download. The malware then joins the system to a botnet.

Bot herders have been using **Mirai** to create large botnets. Mirai infects Linux systems that are running out-of-date versions of Linux and join them to a botnet. This includes Linux software running on Internet of things (IoT) devices such as digital cameras connected to the Internet. Infected devices search for other IoT devices on the Internet and infect them. Attackers have published the source code for Mirai in public forums, making it easily accessible by many attackers.

A Mirai botnet launched an attack in October 2016 against Domain Name System (DNS) servers. It included about 100,000 simple devices such as digital cameras and printers that were connected to the Internet. The bot herders directed the devices to repeatedly query DNS servers in a protracted distributed denial-of-service (DDoS) attack. This attack overwhelmed the DNS servers and prevented users in the United States and Europe from accessing many common web sites, such as Amazon, Second Life, Twitter, CNN, BBC, Fox News, Tumblr, Reddit, and many more.

Logic Bomb

A logic bomb is a string of code embedded into an application or script that will execute in response to an event. The event might be a specific date or time, or a user action such as when a user launches a specific program.

There’s an often-repeated story about a company that decided it had to lay off an engineer due to an economic downturn. His bosses didn’t see him doing much, so they thought they could do without him. Within a couple of weeks after he left, they started having all sorts of computer problems they just couldn’t resolve.

They called him back, and within a couple of weeks, everything was fine. A few months later, they determined they had to lay him off again. You guessed it. Within a couple of weeks, things went haywire again.

The engineer had programmed a logic bomb that executed when the payroll program ran. It checked for his name on the payroll, and when it was there, things were fine, but when his name wasn’t there, ka-boom! the logic bomb exploded.

Real world logic bombs

March 19, 2013 in South Korea:

Email with malicious attachment sent to South Korean organizations.

Posed as a bank email.

Trojan installs malware.

March 20, 2013, 2 p.m. local time:

Malware logic-bomb activates.

Storage and master boot record deleted, system reboots.

Boot device not found.

Please install an operating system on your hard disk.

December 17, 2016, 11:53 p.m in Kiev, Ukraine High-voltage substation:

Logic bomb begins disabling electrical circuits.

Malware mapped out the control network

Began disabling power at a predetermined time.

Customized for Supervisory Control and Data Acquisition (SCADA) networks

Attack Types

Phishing

Phishing is the practice of sending email to users with the purpose of tricking them into revealing personal information or clicking on a link. A phishing attack often sends the user to a malicious website that appears to the user as a legitimate site.

The classic example is where a user receives an email that looks like it came from eBay, PayPal, a bank, or some other well-known company.The “phisher” doesn’t know if the recipient has an account at the company, just as a fisherman doesn’t know if any fish are in the water where he casts his line. However, if the attacker sends out enough emails, the odds are good that someone who receives the email has an account. The email may look like this,

“We have noticed suspicious activity on your account. To protect your privacy, we will suspend your account unless you are able to log in and validate your credentials. Click here to validate your account and prevent it from being locked out.”

The email often includes the same graphics that you would find on the vendor’s web site or an actual email from the vendor. Although it might look genuine, it isn’t. Legitimate companies do not ask you to revalidate your credentials via email. If you go directly to the actual site, you might be asked to provide additional information to prove your identity beyond your credentials, but legitimate companies don’t send emails asking you to follow a link and input your credentials to validate them.

Spearfishing

Spear phishing is a targeted form of phishing. Instead of sending the email out to everyone indiscriminately, a spear phishing attack attempts to target specific groups of users, or even a single user. Spear phishing attacks may target employees within a company or customers of a company.

As an example, an attacker might try to impersonate the CEO of an organization in an email. It’s relatively simple to change the header of an email so that the From field includes anything, including the CEO’s name and title. Attackers can send an email to all employees requesting that they reply with their password. Because the email looks like it’s coming from the CEO, these types of phishing emails fool uneducated users.

One solution that deters the success of these types of spear phishing attacks is to use digital signatures. The CEO and anyone else in the company can sign their emails with a digital signature. This provides a high level of certainty to personnel on who sent the email.

Whaling

Whaling is a form of spear phishing that attempts to target high-level executives. Las Vegas casinos refer to the big spenders as whales, and casino managers are willing to spend extra time and effort to bring them into their casinos. Similarly, attackers consider high-level executives the whales, and attackers are willing to put in some extra effort to catch a whale because the payoff can be so great. When successful, attackers gain confidential company information that they might not be able to get anywhere else.

As an example, attackers singled out as many as 20,000 senior corporate executives in a fine-tuned phishing attack. The emails looked like official subpoenas requiring the recipient to appear before a federal grand jury and included the executive’s full name and other details, such as their company name and phone number. The emails also included a link for more details about the subpoena. If the executives clicked the link, it took them to a web site that indicated they needed a browser add-on to read the document. If they approved this install, they actually installed a keylogger and malware. The keylogger recorded all their keystrokes to a file, and the malware gave the attackers remote access to the executives’ systems.

Real world phishing

March 19, 2016 John Podesta, Former White House Chief of Staff, as well as Former Counselor to the President of the United States and Former chairman of the 2016 Hillary Clinton United States presidential campaign. His personal Gmail account was compromised by a phishing email, exposing messages from 2007 through 2016. Podesta

used the **bit.ly link** in the email to “reset” his password. It wasn’t actually a Google reset link. Ten years of personal emails were unlocked and downloaded. Every email was made available on WikiLeaks. The good, the bad, and the ugly. Don’t underestimate the effects of phishing. It can have significant repercussions.

Vishing

Voice phishing that is done via phone call or voicemail.

Caller ID spoofing is a common component to appear as trusted organization. Often involves scammer asking victim for senstive data such as passwords, account info, PINs, and even credit card information.

Smishing

Phishing via text messages, aka SMS phishing.

A form of phone number spoofing (a phone number that disguised as a more recognizable phone number for deceptive purposes). Will often include a link or ask for personal info.

Pharming

Unlike other phishing schemes, user is not tricked by scammer acting as trustworthy source but instead the victim is redirected from legit site to bogus one. Everything appears right to users. Difficult for anti-virus software to detect.

Poisoned DNS server or client vulnerabilities allow this to happen.

Typo Sqatting

Typo sqatting, aka URL hijacking, relies on typographic errors users make. (goooogle.com) Often set up to record login information. Attackers will set up a dummy site with the common misspelling and try to trick users into entering login creds.

Influence Campaigns

Hacking public opinion, swaying people on political and social issues. Made so much easier with social media! Hybrid warfare - Non-traditional way to wage war. Not new but the cyber aspect takes it to different level.

Invoice Scams

Usually starts with Spear Phishing, knowing who pays the bills and targeting them. Fake invoice sent from spoofed address creating authority (CEO, company letterhead) Trusting the authority of the invoice the bill is payed. May also contain a link to pay, then the payment details can be captured by attacker.

Credential Harvesting

The collection of login credentials without the users knowledge. Often done via link in an email which prompts the user to in put their data.

Tailgating

Tailgating is the practice of one person following closely behind another without showing credentials. For example, if Homer uses a badge to gain access to a secure

building and Francesca follows closely behind Homer without using a badge, Francesca is tailgating.

Employees often do this as a matter of convenience and courtesy. Instead of shutting the door on the person following closely behind, they often hold the door open for the person. However, this bypasses the access control, and if employees tailgate, it’s very easy for a non-employee to slip in behind someone else. Often, all it takes is a friendly smile from someone like Francesca to encourage Homer to keep the door open for her.

Impersonation

Some social engineers impersonate others to get people to do something. For example, many have called users on the phone claiming they work for Microsoft. The Police Virus (a form of ransomware) attempts to impersonate a law enforcement agency. Other times, social engineers attempt to impersonate a person of authority, such as an executive within a company, or a technician.

Dumpster Diving

Dumpster diving is the practice of searching through trash or recycling containers to gain information from discarded documents. Many organizations either shred or burn paper instead of throwing it away.

For example, old copies of company directories can be valuable to attackers. They may identify the names, phone numbers, and titles of key people within the organization. Attackers may be able to use this information in a whaling attack against executives or social engineering attacks against anyone in the organization. An attacker can exploit any document that contains detailed employee or customer information, and can often find value in seemingly useless printouts and notes.

On a personal basis, pre-approved credit applications or blank checks issued by credit card companies can be quite valuable to someone attempting to gain money or steal identities. Documentation with any type of Personally Identifiable Information (PII) or Protected Health Information (PHI) should be shredded or burned.

Shoulder Surfing

Shoulder surfing is simply looking over the shoulder of someone to gain information. The goal is to gain unauthorized information by casual observation, and it’s likely to occur within an office environment. This can be to learn credentials, such as a username and password, or a PIN used for a smart card or debit card. Recently, attackers have been using cameras to monitor locations where users enter PINs, such as at automatic teller machines (ATMs).

Preventing Shoulder Surfing

A simple way to prevent shoulder surfing is to position monitors and other types of screens so that unauthorized personnel cannot see them. This includes ensuring people can’t view them by looking through a window or from reception areas. Another method used to reduce shoulder surfing is to use a screen filter placed over the monitor. This restricts the visibility of the screen for anyone who isn’t looking directly at the monitor.

Computer Hoaxes

A hoax is a message, often circulated through email, that tells of impending doom from a virus or other security threat that simply doesn’t exist. Users may be encouraged to delete files or change their system configuration.

An older example is the teddy bear virus (jdbgmgr.exe), which was not a virus at all. Victims received an email saying this virus lies in a sleeping state for 14 days and then it will destroy the user’s system. It then told users that they can protect their system by deleting the file (which has an icon of a little bear), and provided instructions on how to do so. Users who deleted the file lost some system capability.

More serious virus hoaxes have the potential to be as damaging as a real virus. If users are convinced to delete important files, they may make their systems unusable. Additionally, they waste help-desk personnel’s time due to needless calls about the hoax or support calls if users damaged their systems in response to the hoax.

Watering Hole Attack

A watering hole attack attempts to discover which web sites a group of people are likely to visit and then infects those web sites with malware that can infect the visitors. The attacker’s goal is to infect a web site that users trust already, making them more likely to download infected files.

As an example, an attack discovered in late 2016 initially targeted Polish banks. The attack was discovered by a single Polish bank that discovered previously unknown malware on internal computers. Symantec reported the source of the attack was servers at the Polish Financial Supervision Authority. This is a well-trusted institution by Polish bank employees, and they are likely to visit the organization’s web sites often. This isn’t an isolated incident though. Symantec reported over 100 similar attacks located in over 30 countries.

Social Engineering

Social engineering is the practice of using social tactics to gain information. It’s often low- tech and encourages individuals to do something they wouldn’t normally do, or cause them to reveal some piece of information, such as user credentials.

Social Engineering Principles

Authority

People tend to comply when they believe the person is an authority on the matter. Titles, uniforms, badges, and expertise are all elements of Authority.

The social engineer is in charge. I’m calling from the help desk/office of the CEO/police.

Intimidation

Plays on fear that bad things will happen if victim doesn't comply. Examples: Payroll check won't be processed without providing info. You may be fired if you don't provide requested information.

Consensus/social proof

Plays on the "safety in numbers" idea. By mentioning friends or coworkers who have complied a SE can trick a victim into sharing details they normally would not.

Convince based on what’s normally expected. Your co-worker Joe did this for me last week.

Scarcity

The situation will not be this way for long. Must make the change before time expires.

Urgency

Threat of dire consequences if action isn't take immediately. Scarcity: Offer or situation has limited time of availability. Urgency: SE hopes that normal reasoning skills will fail due to the need to make a quick decision. Works alongside scarcity. Act quickly, don’t think.

Familiarity

SE tries to find common ground (mutual friends, hobbies, interests) to create connection with victim because when you like someone you're more likely to trust them.

Someone you know, we have common friends.

Trust

Bad guy is seen as someone safe. Consensus and liking are used to gain the trust of the victim.

Someone who is safe. I’m from IT, and I’m here to help.

$50,000 lost from Twitter Username

Naoki Hiroshima - @N

1. Bad guy calls PayPal and uses social engineering to get the last four digits of the credit card on file.
2. Bad guy calls GoDaddy and tells them he lost the card, so he can’t properly validate. But he has the last four, and asks if this would help.
3. GoDaddy let the bad guy guess the first two digits of the card.
4. He was allowed to keep guessing until he got it right.

How to steal $50,000 Twitter name Bad guy is now in control of every domain name. And there were some good ones.

Bad guy extorts a swap Once Naoki Hiroshima is made aware that

Twitter reviewed the case for a month Eventually restored access to @N

Denial of Service (DOS) Attack

A denial-of-service (DoS) attack is an attack from one attacker against one target. A distributed denial-of-service (DDoS) attack is an attack from two or more computers against a single target. DDoS attacks often include sustained, abnormally high network traffic on the network interface card of the attacked computer. Other system resource usage (such as the processor and memory usage) will also be abnormally high. The goal of both is to perform a service attack and prevent legitimate users from accessing services on the target computer.

Man-in-the-Middle Attack

A man-in-the-middle (MITM) attack is a form of active interception or active eavesdropping. It uses a separate computer that accepts traffic from each party in a conversation and forwards the traffic between the two. The two computers are unaware of the MITM computer, and it can interrupt the traffic at will or insert malicious code.

For example, imagine that Alice and Bob are exchanging information with their two computers over a network. If Hacker Henry can launch an MITM attack from a third computer, he will be able to intercept all traffic. Alice and Bob still receive all the information, so they are unaware of the attack. However, Hacker Henry also receives all the information. Because the MITM computer can control the entire conversation, it is easy to insert malicious code and send it to the computers. Address Resolution Protocol (ARP) poisoning is one way that an attacker can launch an MITM attack.

Kerberos helps prevent man-in-the-middle attacks with mutual authentication. It doesn't allow a malicious system to insert itself in the middle of the conversation without the knowledge of the other two systems.

Meet-in-the-Middle Attack

> *A Meet-in-the-Middle (MitM) Attack is a kind of cryptanalytic attack where the attacker uses some kind of space or time tradeoff to aid the attack. - Source: (https://www.hypr.com/meet-in-the-middle-mitm-attack/)*

> *Specifically, MitMs attempt to reduce the amount of difficulty required to carry out the assault in its original state. MitMs can take the form of dividing the target communication into two so that each piece can be addressed individually. It could mean transforming an attack requiring X amount of time into one requiring Y time and Z space. The aim is to significantly reduce the effort needed to perform a brute-force attack. - Source: (https://www.hypr.com/meet-in-the-middle-mitm-attack/)*

> *Meet-in-the-middle and man-in-the-middle (MitMs, both) are often conflated. The difference between the two is that the "man" variant is where the attacker places themselves between the two users, eavesdropping or altering the conversation to carry out an attack. The "meet" variant is not interactive, and indeed the term "meet" refers to "let's meet in the middle" or find middle ground by halving, for example, the perceived time that is required to crack encryption when the problem is first encountered. - Source: (https://www.hypr.com/meet-in-the-middle-mitm-attack/)*

Example

> *"Meet-in-the-Middle adversaries try to reconcile the difficulty involved in a large cryptanalytic attack by 'meeting in the middle', or halving the portion of what they are analyzing to make the effort feasible or reasonable in their view." - Source: (https://www.hypr.com/meet-in-the-middle-mitm-attack/)*

Buffer Overflow Attack

A buffer overflow occurs when an application receives more input, or different input, than it expects. The result is an error that exposes system memory that would otherwise be protected and inaccessible. Normally, an application will have access only to a specific area of memory, called a buffer. The buffer overflow allows access

to memory locations beyond the application's buffer, enabling an attacker to write malicious code into this area of memory.

The buffer overflow exposes a vulnerability, but it doesn't necessarily cause damage by itself. However, once attackers discover the vulnerability, they exploit it and overwrite memory locations with their own code. If the attacker uses the buffer overflow to crash the system or disrupt its services, it is a DoS attack.

More often, the attacker's goal is to insert malicious code in a memory location that the system will execute. It's not easy for an attacker to know the exact memory location where the malicious code is stored, making it difficult to get the computer to execute it. However, an attacker can make educated guesses to get close.

A popular method that makes guessing easier is with no operation (NOP, pronounced as "no-op") commands, written as a NOP slide or NOP sled. Many Intel processors use hexadecimal 90 (often written as x90) as a NOP command, so a string of x90 characters is a NOP sled. The attacker writes a long string of x90 instructions into memory, followed by malicious code. When a computer is executing code from memory and it comes to a NOP, it just goes to the next memory location. With a long string of NOPs, the computer simply slides through all of them until it gets to the last one and then executes the code in the next instruction. If the attacker can get the computer to execute code from a memory location anywhere in the NOP slide, the system will execute the attacker's malicious code.

The malicious code varies. In some instances, the attackers write code to spread a worm through the web server's network. In other cases, the code modifies the web application so that the web application tries to infect every user who visits the web site with other malware. The attack possibilities are almost endless.

A buffer overflow attack includes several different elements, but they happen all at once. The attacker sends a single string of data to the application. The first part of the string causes the buffer overflow. The next part of the string is a long string of NOPs followed by the attacker's malicious code, stored in the attacked system's memory. Last, the malicious code goes to work.

SQL Injection Attack

In a SQL injection attack, the attacker enters additional data into the web page form to generate different SQL statements. SQL query languages use a semicolon (;) to indicate the end of the SQL line and use two dashes (--) as an ignored comment. With this knowledge, the attacker could enter different information into the web form like this,

```
John Doe'; SELECT * FROM Downloaders;--
```

If the web application plugged this string of data directly into the SELECT statement surrounded by the same single quotes, it would look like this,

```
SELECT * FROM Apps WHERE Developer
='John Doe'; SELECT * FROM
Downloaders;
--'
```

The first line retrieves data from the database, just as before. However, the semicolon signals the end of the line and the database will accept another command.

The next line reads all the data in the Downloaders table, which can give the attacker access to names, credit card data, and more. The last line comments out the second single quote to prevent a SQL error.

If the application doesn't include error-handling routines, these errors provide details about the type of database the application is using, such as an Oracle, Microsoft SQL Server, or MySQL database. Different databases format SQL statements slightly differently, but once the attacker learns the database brand, it's a simple matter to format the SQL statements required by that brand. The attacker then follows with SQL statements to access the database and may allow the attacker to read, modify, delete, and/or corrupt data.

Cross-Site Scripting (XSS)

Cross-site scripting (XSS) is another web application vulnerability that can be prevented with input validation techniques. Attackers embed malicious HTML or JavaScript code into a web site's code. The code executes when the user visits the site.

You may be wondering why the acronym isn't CSS instead of XSS. The reason is that web sites use Cascading Style Sheets identified as CSS and CSS files are not malicious.

The primary protection against XSS attacks is at the web application with sophisticated input validation techniques. Developers should avoid any methods that allow the web page to display untrusted data. Additionally, OWASP strongly recommends the use of a security encoding library. When implemented, an encoding library will sanitize HTML code and prevent XSS attacks. OWASP includes more than 10 rules that developers can follow to prevent XSS attacks.

Cross-Site Request Forgery

Cross-Site Request Forgery (XSRF or CSRF) is an attack where an attacker tricks a user into performing an action on a web site. The attacker creates a specially crafted HTML link and the user performs the action without realizing it.

As an innocent example of how HTML links create action, consider this HTML link: http://www.google.com/search?q=Success. If users click this link, it works just as if the user browsed to Google and entered Success as a search term. The ?q=Success part of the query causes the action.

Many web sites use the same type of HTML queries to perform actions. For example, imagine a web site that supports user profiles. If users wanted to change profile information, they could log on to the site, make the change, and click a button.

Privilege Escalation

In many attack scenarios, the attacker first gains access to a low-level system or low-level account. For example, an attacker might gain access to Homer's computer using Homer's user account. Homer has access to the network, but doesn't have any administrative privileges. However, attackers use various techniques to gain more and more privileges on Homer's computer and his network.

Advanced persistent threats (APTs) often use remote access Trojans (RATs) to gain access to a single system. Attackers trick a user into clicking a malicious link, which gives them access to a single computer. Attackers then use various scripts to

scan the network looking for vulnerabilities. By exploiting these vulnerabilities, the attackers gain more and more privileges on the network.

DNS Poisoning

A DNS poisoning attack attempts to modify or corrupt DNS results. For example, a successful DNS poisoning attack can modify the IP address associated with google.com and replace it with the IP address of a malicious web site. Each time a user queries DNS for the IP address of google.com, the DNS server responds with the IP address of the malicious web site.

There have been several successful DNS poisoning attacks over the years. Many current DNS servers use Domain Name System Security Extensions (DNSSEC) to protect the DNS records and prevent DNS poisoning attacks.

Zero Day Attack

A zero-day vulnerability is a weakness or bug that is unknown to trusted sources, such as operating system and antivirus vendors. A zero-day attack exploits an undocumented vulnerability. Many times, the vendor isn't aware of the issue. At some point, the vendor learns of the vulnerability and begins to write and test a patch to eliminate it. However, until the vendor releases the patch, the vulnerability is still a zero-day vulnerability.

In most cases, a zero-day vulnerability is a new threat. However, there have been zero-day vulnerabilities that have existed for years. As an example, a bug existed in the virtual DOS machine (VDM) that shipped with every version of 32-bit Windows systems from 1993 to 2010. The bug allowed attackers to escalate their privileges to full system level, effectively allowing them to take over the system. Google researcher Tavis Ormandy stated that he reported the bug to Microsoft in mid-2009. At this point, Microsoft (the vendor) knew about the bug, but didn't release a workaround until January 2010 and a patch until February 2010. Because the bug wasn't known publicly until January 2010, it remained a zero-day vulnerability until then.

Replay Attack

In a replay attack, an attacker captures data sent between two entities, modifies it, and then attempts to impersonate one of the parties by replaying the data. WPA2 using CCMP and AES is not vulnerable to replay attacks. However, WPA using TKIP is vulnerable to replay attacks.

WPA uses a sequence counter to number the packets and an access point will reject packets received out of order. Additionally, TKIP uses a 64- bit Message Integrity Check (MIC) to verify the integrity of the packets. While this sounds secure, security experts identified a method to discover the MIC key. After discovering the key, an attacker can transmit and decrypt packets. Later, other security experts improved this attack allowing them to launch a replay attack. This is one of the reasons that TKIP was deprecated in 2012 and should not be used.

Client Hijacking

URL Hijacking

Typo squatting (also called URL hijacking) occurs when someone buys a domain name that is close to a legitimate domain name. People often do so for malicious purposes. As an example, Amazon hosts the amazon.com web site. If an attacker purchases the name

amazom.com with a slight misspelling at the end of amazon, some users might inadvertently go to the attacker's web site instead of the legitimate web site.

Clickjacking

Clickjacking tricks users into clicking something other than what they think they're clicking. As a simple example, imagine Bart is browsing Facebook. He sees a comment labeled Chalkboard Sayings, so he clicks it. He's taken to a page with a heading of "Human Test" and directions to "Find the blue button to continue." This looks like a Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA), but it isn't. When he clicks the blue button, he is actually clicking on a Facebook share button, causing him to share the original comment labeled Chalkboard Sayings with his Facebook friends.

Session Hijacking

Session hijacking takes advantage of session IDs stored in cookies. When a user logs on to a web site, the web site often returns a small text file (called a cookie) with a session ID. In many cases, this cookie is stored on the user's system and remains active until the user logs off. If the user closes the session and returns to the web site, the web site reads the cookie and automatically logs the user on. This is convenient for the user, but can be exploited by an attacker. In a session hijacking attack, the attacker utilizes the user's session ID to impersonate the user. The web server doesn't know the difference between the original user and the attacker because it is only identifying the user based on the session ID.

Driver Manipulation

Shimming & Refactoring

Occasionally, an application needs to support an older driver. For example, Windows 10 needed to be compatible with drivers used in Windows 8, but all the drivers weren't compatible at first. Shimming provides the solution that makes it appear that the older drivers are compatible. A driver shim is additional code that can be run instead of the original driver. When an application attempts to call an older driver, the operating system intercepts the call and redirects it to run the shim code instead. Refactoring code is the process of rewriting the internal processing of the code, without changing its external behavior. It is usually done to correct problems related to software design.

Developers have a choice when a driver is no longer compatible. They can write a shim to provide compatibility or they can completely rewrite the driver to refactor the relevant code. If the code is clunky, it's appropriate to rewrite the driver.

Attackers with strong programming skills can use their knowledge to manipulate drivers by either creating shims, or by rewriting the internal code. If the attackers can fool the operating system into using a manipulated driver, they can cause it to run malicious code contained within the manipulated driver.

Spoofing

Spoofing occurs when one person or entity impersonates or masquerades as someone or something else. Two common spoofing attacks are media access control (MAC) address spoofing and Internet Protocol (IP) address spoofing.

MAC Spoofing

Host systems on a network have a media access control (MAC) address assigned to the network interface card (NIC). These are hard-coded into the NIC. However, it’s possible to use software methods to associate a different MAC address to the NIC in a MAC spoofing attack.

IP Address Spoofing

In an IP spoofing attack, the attacker changes the source address so that it looks like the IP packet originated from a different source. This can allow an attacker to launch an attack from a single system, while it appears that the attack is coming from different IP addresses.

Rogue Access Points

A rogue access point (rogue AP) is an AP placed within a network without official authorization. It might be an employee who is bypassing security or installed by an attacker. If an employee installs a rogue AP, the chances are higher that this AP will not be managed properly, increasing vulnerabilities to the network.

Attackers may connect a rogue access point to network devices in wireless closets that lack adequate physical security. This access point acts as a sniffer to capture traffic passing through the wired network device, and then broadcasts the traffic using the wireless capability of the AP. The attacker can then capture the exfiltrated data files while sitting in the parking lot. Data exfiltration is the unauthorized transfer of data from an organization to a location controlled by an attacker. Additionally, attackers may be able to use the rogue access point to connect into the wired network. This works the same way that regular users can connect to a wired network via a wireless network. The difference is that the attacker configures all the security for the counterfeit access point and can use it for malicious purposes.

Wireless Evil Twins

An evil twin is a rogue access point with the same SSID as a legitimate access point. For example, many public places such as coffee shops, hotels, and airports include free Wi-Fi as a service. An attacker can set up an AP using the same SSID as the public Wi-Fi network, and many unsuspecting users will connect to this evil twin. Once a user connects to an evil twin, wireless traffic goes through the evil twin instead of the legitimate AP. Often, the attacker presents bogus logon pages to users to capture usernames and passwords. Other times, they simply capture traffic from the connection, such as email or text typed into web page text boxes, and analyze it to detect sensitive information they can exploit.

Wireless Jamming

Attackers can transmit noise or another radio signal on the same frequency used by a wireless network. This interferes with the wireless transmissions and can seriously degrade performance. This type of denial-of-service attack is commonly called jamming and it usually prevents all users from connecting to a wireless network. In some cases, users have intermittent connectivity because the interference causes them to lose their association with the AP and forces them to try to reconnect.

In some cases, you can increase the power levels of the AP to overcome the attack. Another method of overcoming the attack is to use different wireless channels. Each wireless standard has several channels you can use, and if one channel is too noisy, you can use another one. Although this is useful to overcome interference in home

networks, it won’t be as effective to combat an interference attack. If you switch channels, the attacker can also switch channels.

WPS Attack

Wi-Fi Protected Setup (WPS) allows users to configure wireless devices without typing in the passphrase. Instead, users can configure devices by pressing buttons or by entering a short eight-digit personal identification number (PIN).

For example, a user can configure a new wireless device by pressing a button on the AP and on the wireless device. It will automatically configure the device within about 30 seconds with no other actions needed. These buttons can be physical buttons on the devices, or virtual buttons that the user clicks via an application or web page. When using the PIN method, users first identify the eight-digit PIN on the AP and then enter the PIN on the new wireless device.

Unfortunately, WPS is susceptible to brute force attacks. A WPS attack keeps trying different PINs until it succeeds. As an example, Reaver is an open source tool freely available that allows attackers to discover the PIN within 10 hours, and often much quicker. Once it discovers the PIN, it can then discover the passphrase in both WPA and WPA2 wireless networks.

Bluejacking

Bluejacking is the practice of sending unsolicited messages to nearby Bluetooth devices. Bluejacking messages are typically text, but can also be images or sounds. Bluejacking is relatively harmless, but does cause some confusion when users start receiving messages.

Bluesnarfing

Bluesnarfing refers to the unauthorized access to, or theft of information from, a Bluetooth device. A bluesnarfing attack can access information, such as email, contact lists, calendars, and text messages. Attackers use tools such as `hcitool` and `obexftp`.

RFID (Radio-Frequency Identification)

Radio-frequency identification (RFID) systems include an RFID reader and RFID tags placed on objects. They are used to track and manage inventory, and any type of valuable assets, including objects and animals.

There’s an almost endless assortment of tags available for multiple purposes. This includes tags implanted into animals, packaging for any type of product (such as computers), pharmaceuticals, transportation systems (such as shipping containers, railcars, and busses), and controlled substances (such as pharmaceutical containers). Some tags are only slightly larger than a grain of rice.

RFID Attacks

Sniffing or Eavesdropping

Because RFID transmits data over the air, it’s possible for an attacker to collect it by listening. A key requirement is to know the frequency used by the RFID system and have a receiver that can be tuned to that frequency. The attacker also needs to know the protocols used by the RFID system to interpret the data.

Replay

Successful eavesdropping attacks allow the attacker to perform a replay attack. For example, an attacker can configure a bogus tag to mimic the tag attached to a valuable object. The attacker can then steal the valuable object without the theft being easily detected.

DoS

A denial-of-service (DoS) attack attempts to disrupt services. If an attacker knows the frequency used by the RFID system, it’s possible to launch a jamming or interference attack, flooding the frequency with noise. This prevents the RFID system from operating normally.

Wireless Disassociation Attack

A disassociation attack effectively removes a wireless client from a wireless network. To understand the attack, it’s valuable to first understand the normal operation.

After a wireless client authenticates with a wireless AP, the two devices exchange frames, causing the client to be associated with the AP. At any point, a wireless device can send a disassociation frame to the AP to terminate the connection. This frame includes the wireless client’s MAC address. When the AP receives the disassociation frame, it deallocates all the memory it was using for the connection.

In a disassociation attack, attackers send a disassociation frame to the AP with a spoofed MAC address of the victim. The AP receives the frame and shuts down the connection. The victim is now disconnected from the AP and must go through the authentication process again to reconnect.

Cryptographic Attacks

Birthday Attack

A birthday attack is named after the birthday paradox in mathematical probability theory. The birthday paradox states that for any random group of 23 people, there is a 50 percent chance that 2 of them have the same birthday. This is not the same year, but instead one of the 365 days in any year. In a birthday attack, an attacker is able to create a password that produces the same hash as the user’s actual password. This is also known as a hash collision.

A hash collision occurs when the hashing algorithm creates the same hash from different passwords. This is not desirable. As an example, imagine a simple hashing algorithm creates three-digit hashes. The password “success” might create a hash of 123 and the password “passed” might create the same hash of 123. In this scenario, an attacker could use either “success” or “passed” as the password and both would work.

Known Plaintext Attack (KPA)

Many cryptographic attacks attempt to decrypt encrypted data. Plaintext is human-readable data. An encryption algorithm scrambles the data, creating ciphertext.

An attacker can launch a known plaintext attack if he has samples of both the plaintext and the ciphertext. As an example, if an attacker captures an encrypted message (the ciphertext) and knows the plaintext of the message, he can use both sets of data to discover the encryption and decryption method. If successful, he can use the same decryption method on other ciphertext.

Rainbow Tables

Rainbow table attacks are a type of attack that attempts to discover the password from the hash. A rainbow table is a huge database of precomputed hashes. It helps to look at the process of how some password cracker applications discover passwords without a rainbow table. Assume that an attacker has the hash of a password.

Dictionary Attacks

A dictionary attack is one of the original password attacks. It uses a dictionary of words and attempts every word in the dictionary to see if it works. A dictionary in this context is simply a list of words and character combinations. Dictionaries used in these attacks have evolved over time to reflect user behavior. Today, they include many of the common passwords that uneducated users configure for their accounts. For example, even though 12345 isn't a dictionary word, many people use it as a password, so character sets such as these have been added to many dictionaries used by dictionary attack tools. These attacks are thwarted by using complex passwords. A complex password will not include words in a dictionary.

Brute Force Attack

A brute force attack attempts to guess all possible character combinations. The two types of brute force attacks are online and offline. An online password attack attempts to discover a password from an online system. For example, an attacker can try to log on to an account by repeatedly guessing the username and password. Many tools are available that attackers can use to automate the process. For example, ncrack is a free tool that can be used to run online brute force password attacks.

Offline password attacks attempt to discover passwords from a captured database or captured packet scan. For example, when attackers hack into a system or network causing a data breach, they can download entire databases. They then perform offline attacks to discover the passwords contained within the databases.

Threat Actors

When considering attacks, it's important to realize that there are several different types of threat actors, and they each have different attributes. Don't let the phrase threat actors confuse you. It's just a fancier name given to attackers anyone who launches a cyberattack on others. One common method that attackers often use before launching an attack is to gather information from open-source intelligence. This includes any information that is available via web sites and social media. For example, if attackers want to get the name of the chief executive officer (CEO) of a company, they can probably find it on the company's web site. Similarly, many organizations post information on social media sites such as Facebook and Twitter.

Script Kiddies

Script kiddie is an attacker who uses existing computer scripts or code to launch attacks. Script kiddies typically have very little expertise or sophistication, and very little funding. Many people joke about the bored teenager as the script kiddie, attacking sites or organizations for the fun of it. However, there isn't any age limit for a script kiddie. More important, they can still get their hands on powerful scripts and launch dangerous attacks. Their motivations vary, but they are typically launching attacks out of boredom, or just to see what they can do.

Hacktivist

Hacktivist launches attacks as part of an activist movement or to further a cause. Hacktivists typically aren’t launching these attacks for their own benefit, but instead to increase awareness about a cause.

Organized Crime

Organized crime is an enterprise that employs a group of individuals working together in criminal activities. This group is organized with a hierarchy with a leader and workers, like a normal business. Depending on how large the enterprise is, it can have several layers of management. However, unlike a legitimate business, the enterprise is focused on criminal activity.

Nation States / APT

Advanced Persistent Threat (APT) is a targeted attack against a network. The attacks are typically launched by a group that has both the capability and intent to launch sophisticated and targeted attacks. They often have a significant amount of resources and funding. Additionally, individuals within an APT group typically have very specific targets, such as a specific company, organization, or government agency. Successful attacks often allow unauthorized access for long periods of time, allowing attacks to exfiltrate a significant amount of data.

Insiders

Insider is anyone who has legitimate access to an organization’s internal resources. Common security issues caused by insider threats include loss of confidentiality, integrity, and availability of the organization’s assets. The extent of the threat depends on how much access the insider has. For example, an administrator would have access to many more IT systems than a regular user.

Penetration Testing

Penetration testing actively assesses deployed security controls within a system or network. It starts with passive reconnaissance, such as a vulnerability scan, but takes it a step further and tries to exploit vulnerabilities by simulating or performing an attack.

Security testers typically perform a penetration test to demonstrate the actual security vulnerabilities within a system. This can help the organization determine the impact of a threat against a system. In other words, it helps an organization determine the extent of damage that an attacker could inflict by exploiting a vulnerability.

Although it’s not as common, it’s also possible to perform a penetration test to determine how an organization will respond to a compromised system. This allows an organization to demonstrate security vulnerabilities and flaws in policy implementation. For example, many organizations may have perfect policies on paper. However, if employees aren’t consistently following the policies, a penetration test can accurately demonstrate the flaws.

Because a penetration test can exploit vulnerabilities, it has the potential to disrupt actual operations and cause system instability. Because of this, it’s important to strictly define boundaries for a test. Ideally, the penetration test will stop right before performing an exploit that can cause damage or result in an outage. However, some tests cause unexpected results.

Passive Reconnaissance

Collects information about a targeted system, network, or organization using open-source intelligence. This includes viewing social media sources about the target, news reports, and even the organization's web site. If the organization has wireless networks, it could include passively collecting information from the network such as network SSIDs. Note that because passive reconnaissance doesn't engage a target, it isn't illegal.

Active Reconnaissance

Includes using tools to send data to systems and analyzing the responses. It typically starts by using various scanning tools such as network scanners and vulnerability scanners. It's important to realize that active reconnaissance does engage targets and is almost always illegal. It should never be started without first getting explicit authorization to do so.

Vulnerability Scanning

A key part of a vulnerability assessment is a vulnerability scan. Security administrators often use a vulnerability scanner to identify which systems are susceptible to attacks. Vulnerability scanners identify a wide range of weaknesses and known security issues that attackers can exploit. Most vulnerability scanners combine multiple features into a single package.

Identify Vulnerability

Open ports Open ports can signal a vulnerability, especially if administrators aren't actively managing the services associated with these ports. For example, all web servers do not use File Transfer Protocol (FTP), so if TCP ports 20 and 21 are open, it indicates a potential vulnerability related to FTP. Similarly, Telnet uses port 23, so if this port is open, an attacker can try to connect to the server using Telnet.

Weak passwords Many scanners include a password cracker that can discover weak passwords or verify that users are creating strong passwords in compliance with an organization's policy. It is more efficient to use a technical password policy to require and enforce the use of strong passwords. However, if this isn't possible, administrators use a separate password cracker to discover weak passwords.

Default accounts and passwords Operating systems and applications can have default usernames and passwords. Basic operating system and application hardening steps should remove the defaults, and a scan can discover the weaknesses if operating systems and applications aren't secured properly. F o r example, some SQL database systems allow the sa (system administrator) account to be enabled with a blank password. Scanners such as Nessus will detect this.

Sensitive data Some scanners include data loss prevention (DLP) techniques to detect sensitive data sent over the network. For example, a DLP system can scan data looking for patterns such as Social Security numbers or key words that identify classified or proprietary data.

Security and configuration errors Vulnerability scans can also check the system against a configuration or security baseline to identify unauthorized changes.

Vulnerability Types

There are many types of vulnerabilities that exist today.

Race Condition Improper Certificate and Key Management End-of-life Vulnerabilities Weak Cipher Suites Embedded System Vulnerabilities Lack of Vendor Support Improper Input Handling Improper Error Handling Improperly Configured Accounts System Sprawl/Undocumented Assets Vulnerable Business Processes Misconfiguration/Default Configuration Untrained Users

Memory/Buffer Vulnerabilities

Memory Leak Integer Overflow Buffer Overflow NULL Pointer Dereference DLL Injection

Architecture/Design Weaknesses

Examine Every Part of the Network

Ingress VPN Third-Party Access Internal Controls Account Access Front Door Access Conference Room Access

New Threats = Zero Day

Security Components

Firewalls

Firewall filters incoming and outgoing traffic for a single host or between networks. In other words, a firewall can ensure only specific types of traffic are allowed into a network or host, and only specific types of traffic are allowed out of a network or host.

Firewalls start with a basic routing capability for packet filtering, including the use of an implicit deny rule. More advanced firewalls go beyond simple packet filtering and include advanced content filtering.

An application-based firewall is typically software running on a system. For example, host- based firewalls are commonly application-based. A network-based firewall is usually a dedicated system with additional software installed to monitor, filter, and log traffic. For example, Cisco makes a variety of different network-based firewalls. Many of them are dedicated servers with proprietary firewall software installed.

A network-based firewall would have two or more network interface cards (NICs) and all traffic passes through the firewall. The firewall controls traffic going in and out of a network. It does this by filtering traffic based on firewall rules and allows only authorized traffic to pass through it. Most organizations include at least one network-based firewall at the border, between their intranet (or internal network) and the Internet.

Stateless firewall

Stateless firewalls use rules implemented as ACLs to identify allowed and blocked traffic. This is similar to how a router uses rules. Firewalls use an implicit deny strategy to block all traffic that is not explicitly allowed.

Stateful firewall

A stateful firewall inspects traffic and makes decisions based on the context, or state, of the traffic. It keeps track of established sessions and inspects traffic

based on its state within a session. It blocks traffic that isn’t part of an established session. As an example, a TCP session starts with a three-way handshake. If a stateful firewall detects TCP traffic without a corresponding three-way handshake, it recognizes this as suspicious traffic and can block it. A common security issue with stateless firewalls is misconfigured ACLs. For example, if the ACL doesn’t include an implicit deny rule, it can allow almost all traffic into the network.

Host-Based Firewall

A host-based firewall monitors traffic going in and out of a single host, such as a server or a workstation. It monitors traffic passing through the NIC and can prevent intrusions into the computer via the NIC. Many operating systems include software-based firewalls used as host- based firewalls. For example, Microsoft has included a host-based firewall on operating systems since Windows XP. Additionally, many third-party host-based firewalls are available.

Web Application Firewall

A web application firewall (WAF) is a firewall specifically designed to protect a web application, which is commonly hosted on a web server. In other words, it’s placed between a server hosting a web application and a client. It can be a stand-alone appliance, or software added to another device.

VPN Concentrators

VPN Concentrator

A virtual private network (VPN) is another method used for remote access. VPNs allow users to access private networks via a public network.Larger organizations often use a VPN concentrator, which is a dedicated device used for VPNs. A VPN concentrator includes all the services needed to create a VPN, including strong encryption and authentication techniques, and it supports many clients.

When using a VPN concentrator, you would typically place it in the DMZ. The firewall between the Internet and the DMZ would forward VPN traffic to the VPN concentrator. The concentrator would route all private VPN traffic to the firewall between the DMZ and the intranet.

Remote Access VPN

The VPN client first connects to the Internet using a broadband connection to an Internet Service Provider (ISP). After connecting to the Internet, the VPN client can then initiate the VPN connection. The VPN server is in the DMZ and reachable through a public IP address. This makes it accessible from any other host on the Internet. A VPN server needs to authenticate clients. A common method is to use an internal Remote Authentication Dial-in User Service (RADIUS) server. When a user logs on, the VPN server sends the user’s credentials to the RADIUS server.

Site-to-Site VPN

A site-to-site VPN includes two VPN servers that act as gateways for two networks separated geographically. For example, an organization can have two locations. One is its headquarters and the other is a remote office.A benefit of the site-to-site model is that it connects both networks without requiring additional steps on the part of

the user. Users in the remote office can connect to servers in the headquarters location as easily as if the servers were in the remote office.

IP Sec (Internet Protocol Security)

Tunnel mode encrypts the entire IP packet used in the internal network, and is the mode used with VPNs transmitted over the Internet. The benefit is that the IP addressing used within the internal network is encrypted and not visible to anyone who intercepts the traffic. If someone does intercept the traffic, he can see the source IP address from the client and the destination address to the VPN server, but the internal IP address information remains hidden.

Transport mode only encrypts the payload and is commonly used in private networks, but not with VPNs. If traffic is transmitted and used only within a private network, there isn’t any need to hide the IP addresses by encrypting them.

Authentication Header (AH)

Authentication IPsec includes an Authentication Header (AH) to allow each of the hosts in the IPsec conversation to authenticate with each other before exchanging data. AH provides authentication and integrity. AH uses protocol number 51.

Encapsulation Security Payload (ESP)

Encryption IPsec includes Encapsulating Security Payload (ESP) to encrypt the data and provide confidentiality. ESP includes AH so it provides confidentiality, authentication, and integrity. ESP uses protocol number 50.

Intrusion Detection & Prevention

Intrusion Detection System (IDS)

HIDS A host-based intrusion detection system (HIDS) is additional software installed on a system such as a workstation or server. It provides protection to the individual host and can detect potential attacks and protect critical operating system files. The primary goal of any IDS is to monitor traffic. For a HIDS, this traffic passes through the network interface card (NIC).

Many host-based IDSs have expanded to monitor application activity on the system. As one example, you can install a HIDS on different Internet facing servers, such as web servers, mail servers, and database servers. In addition to monitoring the network traffic reaching the servers, the HIDS can also monitor the server applications.

NIDS A network-based intrusion detection system (NIDS) monitors activity on the network. An administrator installs NIDS sensors or collectors on network devices such as routers and firewalls. These sensors gather information and report to a central monitoring server hosting a NIDS console.

A NIDS is not able to detect anomalies on individual systems or workstations unless the anomaly causes a significant difference in network traffic. Additionally, a NIDS is unable to decrypt encrypted traffic. In other words, it can only monitor and assess threats on the network from traffic sent in plaintext or non-encrypted traffic.

Identification Technologies

Signature-based IDSs (also called definition-based) use a database of known vulnerabilities or known attack patterns. For example, tools are available for an attacker to launch a SYN flood attack on a server by simply entering the IP address of the system to attack. The attack tool then floods the target system with synchronize (SYN) packets, but never completes the three-way Transmission Control Protocol (TCP) handshake with the final acknowledge (ACK) packet. If the attack isn't blocked, it can consume resources on a system and ultimately cause it to crash. Heuristic/behavioral-based detection (also called anomaly-based detection) starts by identifying normal operation or normal behavior of the network. It does this by creating a performance baseline under normal operating conditions. The IDS provides continuous monitoring by constantly comparing current network behavior against the baseline. When the IDS detects abnormal activity (outside normal boundaries as identified in the baseline), it gives an alert indicating a potential attack.

Intrusion Prevention System (IPS)

Intrusion prevention systems (IPSs) are an extension of IDSs. Just as you can have both a HIDS and a NIDS, you can also have a HIPS and a NIPS, but a network-based IPS (NIPS) is more common. There are some primary distinctions of an IPS when compared with an IDS,

An IPS can detect, react, and prevent attacks. In contrast, an IDS monitors and will respond after detecting an attack, but it doesn't prevent them. An IPS is inline with the traffic. In other words, all traffic passes through the IPS and the IPS can block malicious traffic. This is sometimes referred to as in-band. In contrast, an IDS is out-of-band. It monitors the network traffic, but the traffic doesn't go through the IDS. This is sometimes referred to as passive.

False Positives & Negatives

While IDSs use advanced analytics to examine traffic, they are susceptible to both false positives and false negatives. A false positive is an alert or alarm on an event that is nonthreatening, benign, or harmless. A false negative is when an attacker is actively attacking the network, but the system does not detect it. Neither is desirable, but it's impossible to eliminate both. Most IDSs trigger an alert or alarm when an event exceeds a threshold.

Router & Switch Security

Routers

A router connects multiple network segments together into a single network and routes traffic between the segments. As an example, the Internet is effectively a single network hosting billions of computers. Routers route the traffic from segment to segment.

Because routers don't pass broadcasts, they effectively reduce traffic on any single segment. Segments separated by routers are sometimes referred to as broadcast domains. If a network has too many computers on a single segment, broadcasts can result in excessive collisions and reduce network performance. Moving computers to a different segment separated by a router can significantly improve overall performance. Similarly, subnetting networks creates separate broadcast domains.

Access Control Lists (ACLs)

Access control lists (ACLs) are rules implemented on a router (and on firewalls) to identify what traffic is allowed and what traffic is denied. Rules within an ACL provide rule-based management for the router and control inbound and outbound traffic. Router ACLs provide basic packet filtering. They filter packets based on IP addresses, ports, and some protocols, such as ICMP or IPsec, based on the protocol identifiers:

IP addresses and networks You can add a rule in the ACL to block access from any single computer based on the IP address. If you want to block traffic from one subnet to another, you can use a rule to block traffic using the subnet IDs.

Ports You can filter traffic based on logical ports. For example, if you want to block HTTP traffic, you can create a rule to block traffic on port 80. Note that you can choose to block incoming traffic, outgoing traffic, or both. In other words, i t ’s possible to allow outgoing HTTP traffic while blocking incoming HTTP traffic.

Protocol numbers Many protocols are identified by their protocol numbers. For example, ICMP uses a protocol number of 1 and many DoS attacks use ICMP. You can block all ICMP traffic (and the attacks that use it) by blocking traffic using this protocol number.

Anti-Spoofing

Attackers often use spoofing to impersonate or masquerade as someone or something else. In the context of routers, an attacker will spoof the source IP address by replacing the actual source IP address with a different one. This is often done to hide the actual source of the packet. You can implement anti-spoofing on a router by modifying the access list to allow or block IP addresses. As an example, private IP addresses should only be used in private networks. Any traffic coming from the Internet using a private IP address as the source IP address is obviously an attempt to spoof the source IP address.

Switches

A switch can learn which computers are attached to each of its physical ports. It then uses this knowledge to create internal switched connections when two computers communicate with each other.

Switch Port Security

Port security limits the computers that can connect to physical ports on a switch. At the most basic level, administrators disable unused ports. For example, individual RJ-45 wall jacks in an office lead to specific physical ports on a switch. If the wall jack is not being used, administrators can disable the switch port. This prevents someone from plugging in a laptop or other computer into the wall jack and connecting to the network.

Network Access Control (NAC)

Network access control (NAC) methods provide continuous security monitoring by inspecting computers and preventing them from accessing the network if they don’t pass the inspection. Most administrators have complete control over computers in their network. For example, they can ensure the clients have up-to-date antivirus software installed, operating systems have current patches applied, and their firewalls are enabled. However, administrators don’t have complete control of computers employees use at home or on the road. NAC provides a measure of control for these other computers. It ensures that clients meet predetermined characteristics prior to

accessing a network. NAC systems often use health as a metaphor, indicating that a client meets these predetermined characteristics. Just as doctors can quarantine patients with certain illnesses, NAC can quarantine or isolate unhealthy clients that don’t meet the predefined NAC conditions.

Loop Prevention

In some situations, a network can develop a switching loop or bridge loop problem. The effect is similar to a broadcast storm and it can effectively disable a switch. For example, if a user connects two ports of a switch together with a cable, it creates a switching loop where teh switch continuously sends and resends unicast transmissions through the switch. In addition to disabling the switch, it also degrades performance of the overall network. This is trivial for many network administrators, because most current switches have Spanning Tree Protocol (STP) or the newer Rapid STP (RSTP) installed and enabled for loop prevention. However, if these protocols are disabled, the switch is susceptible to loop problems. The simple solution is to ensure that switches include loop protection such as STP or RSTP.

Flood Guard

Many switches include a flood guard to protect against MAC flood attacks. When enabled, the switch will limit the amount of memory used to store MAC addresses for each port. For example, the switch might limit the number of entries for any port to 132 entries. This is much more than you need for normal operation. If the switch detects an attempt to store more than 132 entries, it raises an alert. The flood guard typically sends a Simple Network Management Protocol (SNMP) trap or error message in response to the alert. Additionally, it can either disable the port or restrict updates for the port. By disabling the port, it effectively blocks all traffic through the port until an administrator intervenes. If it restricts updates, the switch will use currently logged entries for the port, but ignore attempts to update it. All other ports will continue to operate normally.

Proxies

Many networks use proxy servers (or forward proxy servers) to forward requests for services (such as HTTP or HTTPS) from clients. They can improve performance by caching content and some proxy servers can restrict users’ access to inappropriate web sites by filtering content. A proxy server is located on the edge of the network bordering the Internet and the intranet.

Administrators configure internal clients to use the proxy server for specific protocols. The proxy accepts their requests, retrieves the content from the Internet, and then returns the data to the client. Most proxy servers only act as a proxy for HTTP and HTTPS. However, proxy servers can also proxy other Internet protocols, such as FTP.

Application Proxies

An application proxy is used for specific applications. It accepts requests, forwards the requests to the appropriate server, and then sends the response to the original requestor. A forward proxy used for HTTP is a basic application proxy. However, most application proxies are multipurpose proxy servers supporting multiple protocols such as HTTP and HTTPS.

A reverse proxy accepts requests from the Internet, typically for a single web server. It appears to clients as a web server, but is forwarding the requests to the web server and serving the pages returned by the web server.

Load Balancers

A load balancer can optimize and distribute data loads across multiple computers or multiple networks. For example, if an organization hosts a popular web site, it can use multiple servers hosting the same web site in a web farm. Load-balancing software distributes traffic equally among all the servers in the web farm, typically located in a DMZ.

The term load balancer makes it sound like it’s a piece of hardware, but a load balancer can be hardware or software. A hardware-based load balancer accepts traffic and directs it to servers based on factors such as processor utilization and the number of current connections to the server. A software based load balancer uses software running on each of the servers in the load balanced cluster to balance the load.

Load balancing primarily provides scalability, but it also contributes to high availability. Scalability refers to the ability of a service to serve more clients without any decrease in performance. Availability ensures that systems are up and operational when needed. By spreading the load among multiple systems, it ensures that individual systems are not overloaded, increasing overall availability.

Scheduling

A load balancer uses a scheduling technique to determine where to send new requests. Some load balancers simply send new requests to the servers in a round-robin fashion. The load balancer sends the first request to Server 1, the second request to Server 2, and so on.

Affinity

Some load balancers use source address affinity to direct the requests. Source affinity sends requests to the same server based on the requestor’s IP address.

Active/Passive Load Balancing

Some servers are active Others are on standby If an active server fails, the passive server takes its place

Access Points

Wireless Access Point (WAP)

A wireless access point (AP) connects wireless clients to a wired network. However, many APs also have routing capabilities. Vendors commonly market APs with routing capabilities as wireless routers so that’s how you’ll typically see them advertised. Two distinctions are,

All wireless routers are APs These are APs with an extra capability routing.

Not all APs are wireless routers Many APs do not have any additional capabilities. They provide connectivity for wireless clients to a wired network, but do not have

routing capabilities.

SSID Management

Wireless networks are identified by a service set identifier (SSID), which is simply the name of the wireless network. Some APs still come with default SSIDs, though most vendors have moved away from this practice. For example, the default SSID of some older Linksys APs is “Linksys.” Some newer APs force you to enter a name for the SSID when you first install it and do not include a default. From a defense-in-depth perspective, it’s a good idea to change the name of the SSID if a default is used. It simply gives attackers less information.

MAC Filtering

Enabling media access control (MAC) filtering provides a small measure of security to a wireless network. MAC filtering is a form of network access control. It’s used with port security on switches and you can use it to restrict access to wireless networks.

Band Selection and Bandwidth

Wireless networks use two primary radio bands: 2.4 GHz and 5 GHz. However, wireless devices don’t transmit exactly on 2.4 GHz or 5 GHz. Instead, the two bands have multiple channels starting at about 2.4 GHz and 5 GHz. There isn’t a single standard that applies to every country, so you’ll find that the number of channels within each band varies from country to country.

Antennas

The most commonly used wireless antenna on both APs and wireless devices is an omnidirectional (or omni) antenna. Omnidirectional antennas transmit and receive signals in all directions at the same time. This allows wireless devices to connect to an AP from any direction. Another type of antenna is a directional antenna. A directional antenna transmits in a single direction and receives signals back from the same direction. Because the power of the antenna is focused in a single direction, the directional antenna has greater gain than an omni antenna, and it can transmit and receive signals over greater distances. The directional antenna also has a very narrow radiation pattern, focusing the signal in a specific area.

Managing Wireless Configurations

A fat AP, also known as a stand-alone, intelligent, or autonomous AP, includes everything needed to connect wireless clients to a wireless network. It typically includes features such as a routing component, NAT, DHCP, wireless security options, access control lists (ACLs), and more.

A thin AP is a controller-based AP, meaning that it isn’t a stand-alone AP, but rather an AP managed by a controller. Administrators use a wireless controller to configure and manage thin-based APs. This streamlines the administration by consolidating it in one place.

Security Information and Event Management (SIEM)

A security information and event management (SIEM) system provides a centralized solution for collecting, analyzing, and managing data from multiple sources. They combine the services of security event management (SEM) and security information

management (SIM) solutions. A SEM provides real-time monitoring, analysis, and notification of security events, such as suspected security incidents. A SIM provides long-term storage of data, along with methods of analyzing the data looking for trends, or creating reports needed to verify compliance of laws or regulations.

SIEMs are very useful in large enterprises that have massive amounts of data and activity to monitor. Consider an organization with over 1,000 servers. When an incident occurs on just one of those servers, administrators need to know about it as quickly as possible. The SIEM provides continuous monitoring and provides real-time reporting. For example, in a large network operations center (NOC), the SIEM might display alerts on a large heads-up display. A benefit is that the monitoring and reporting is automated with scripts with the SIEM.

Time Synchronization

All servers sending data to the SIEM should be synchronized with the same time. This becomes especially important when investigating an incident so that security investigators know when events occurred. Additionally, large organizations can have locations in different time zones. Each of these locations might have servers sending data to a single centralized SIEM. If the server logs use their local time, the SIEM needs to ensure that it compensates for the time offset. One method is to convert all times to Greenwich Mean Time (GMT), which is the time at the Royal Observatory in Greenwich, London.

Syslog

A SIEM typically includes methods to prevent anyone from modifying log entries. This is sometimes referred to as write once read many (WORM). As logs are received, the SIEM will aggregate and correlate the log entries. After processing the logs, it can archive the source logs with write protection.

Event De-Duplication

Deduplication is the process of removing duplicate entries. As an example, imagine 10 users receive the same email and choose to save it. An email server using deduplication processing will keep only one copy of this email, but make it accessible to all 10 users. Imagine a NIDS collects data from a firewall and a SIEM collects data from the NIDS and the firewall. The SIEM will store only a single copy of any duplicate log entries, but also ensure that the entries are associated with both devices.

Automated Alerting and Triggers

A SIEM typically comes with predefined alerts, which provide notifications of suspicious events. For example, if it detects a port scan on a server, it might send an email to an administrator group or display the alert on a heads-up display. SIEMs also include the ability to create new alerts. Automated triggers. Triggers cause an action in response to a predefined number of repeated events. As an example, imagine a trigger for failed logons is set at five. If an attacker repeatedly tries to log on to a server using Secure Shell (SSH), the server’s log will show the failed logon attempts. When the SIEM detects more than five failed SSH logons, it can change the environment and stop the attack. It might modify a firewall to block these SSH logon attempts or send a script to the server to temporarily disable SSH. A SIEM includes the ability to modify predefined triggers and create new ones.

Aggregation

Aggregation refers to combining several dissimilar items into a single item. A SIEM can collect data from multiple sources, such as firewalls, intrusion detection systems, proxy servers, and more. Each of these devices formats the logs differently. However, the SIEM can aggregate the data and store it in such a way that it is easy to analyze and search.

Correlation Engine

A correlation engine is a software component used to collect and analyze event log data from various systems within the network. It typically aggregates the data looking for common attributes. It then uses advanced analytic tools to detect patterns of potential security events and raises alerts. System administrators can then investigate the alert.

Data Loss Prevention (DLP)

Organizations often use data loss prevention (DLP) techniques and technologies to prevent data loss. They can block the use of USB flash drives and control the use of removable media. They can also examine outgoing data and detect many types of unauthorized data transfers.

Data Exfiltration

Data exfiltration is the unauthorized transfer of data outside an organization and is a significant concern. In some cases, attackers take control of systems and transfer data outside an organization using malware. It’s also possible for malicious insiders to transfer data.

USB Blocking

It’s common for an organization to include security policy statements to prohibit the use of USB flash drives and other removable media. Some technical policies block use of USB drives completely.

A DLP solution is more selective and it can prevent a user from copying or printing files with specific content. For example, it’s possible to configure a DLP solution to prevent users from copying or printing any classified documents marked with a label of Confidential. The DLP software scans all documents sent to the printer, and if it contains the label, the DLP software blocks it from reaching the printer.

In addition to blocking the transfer, a DLP solution will typically log these events. Some DLP solutions will also alert security administrators of the event. Depending on the organization’s policy, personnel may be disciplined for unauthorized attempts to copy or print files.

Cloud Based DLP

It’s common for personnel within organizations to store data in the cloud. This makes it easier to access the data from any location and from almost any device. Cloud-based DLP solutions allow an organization to implement policies for data stored in the cloud. As an example, an organization can implement policies to detect Personally Identifiable Information (PII) or Protected Health Information (PHI) stored in the cloud. After detecting the data, a DLP policy can be configured to take one or more

actions such as sending an alert to a security administrator, blocking any attempts to save the data in the cloud, and quarantining the data.

Network Access Control (NAC)

Allowing remote access to your private network can expose your network to a significant number of risks from the clients. If a user logs on to a VPN with a malware-infected computer, this computer can then infect other computers on the internal network. Network access control (NAC) methods provide continuous security monitoring by inspecting computers and preventing them from accessing the network if they don’t pass the inspection.

Most administrators have complete control over computers in their network. For example, they can ensure the clients have up-to-date antivirus software installed, operating systems have current patches applied, and their firewalls are enabled. However, administrators don’t have complete control of computers employees use at home or on the road.

NAC provides a measure of control for these other computers. It ensures that clients meet predetermined characteristics prior to accessing a network. NAC systems often use health as a metaphor, indicating that a client meets these predetermined characteristics. Just as doctors can quarantine patients with certain illnesses, NAC can quarantine or isolate unhealthy clients that don’t meet the predefined NAC conditions.

Posture Assessment

Administrators set predefined conditions for healthy clients and those that meet these preset conditions can access the network. The NAC system isolates computers that don’t meet the conditions. Common health conditions checked by a NAC are:

Up-to-date antivirus software, including updated signature definitions Up-to-date operating system, including current patches and fixes Firewall enabled on the client

NAC systems use authentication agents (sometimes called health agents) to inspect NAC clients. These agents are applications or services that check different conditions on the computer and document the status in a statement of health. When a client connects to a NAC-controlled network, the agent reports the health status of the NAC client.

Health Checks/Posture Assessment

Persistent agents Permanently installed onto a system Periodic updates may be required

Dissolvable agents No installation is required Runs during the posture assessment Terminates when no longer required

Agentless NAC Integrated with Active Directory Checks are made during login and logoff Can’t be scheduled

Agents

Agents on clients can be either dissolvable or permanent.

A permanent agent (sometimes called a persistent NAC agent) is installed on the client and stays on the client. NAC uses the agent when the client attempts to log on

remotely. This is the most common implementation for corporate-owned devices, and for approved laptops and PCs that employees use to connect remotely.

A dissolvable agent is downloaded and run on the client when the client logs on remotely. It collects the information it needs, identifies the client as healthy or not healthy, and reports the status back to the NAC system. Some dissolvable NAC agents remove themselves immediately after they report back to the NAC system. Others remove themselves after the remote session ends.

Mail Gateways

A mail gateway is a server that examines all incoming and outgoing email and attempts to reduce risks associated with email. Many vendors sell appliances that perform all the desired services of a mail gateway. Administrators locate it between the email server and the Internet and configure it for their purposes. All mail goes to the gateway before it goes to the email server. Additionally, many vendors include a mail gateway within a UTM appliance. The mail gateway is just another security feature within the UTM appliance.

Identifying Spam

Spam is unsolicited email and attackers commonly use spam to launch attacks. For example, spam can include malware as an attachment or it might include a link to a malicious web site. A spam filter within a mail gateway filters out spam from incoming email. By filtering out spam, it helps block attacks.

Email Encryption

Many mail gateways also support encryption. They can encrypt all outgoing email to ensure confidentiality for the data-in-transit, or only encrypt certain data based on policies. For example, if an organization is working on a project with another organization, administrators can configure the gateway to encrypt all traffic sent to the other organization.□□□DLP□Mail gateways often include data loss prevention (DLP) capabilities. They examine outgoing email looking for confidential or sensitive information and block them. As an example, imagine an organization is working on a secret project with a codeword of “DOH.” All documents associated with this project have the keyword within them. The mail gateway includes this keyword in its searches and when it detects the keyword within an email or an attachment, it blocks the email. Administrators have the choice of configuring the gateway to notify security personnel, the user who sent the email, or both when it blocks an email.

Other Security Devices

SSL Accelerators

SSL/TLS accelerators refer to hardware devices focused on handling Transport Layer Security (TLS) traffic. When using an SSL accelerator, it’s best to place it as close as possible to related devices. For example, if you’re using an SSL accelerator to off-load HTTPS sessions for a web server, place the SSL accelerator close to the web server.

SSL/TLS Decryption

Some organizations use SSL decryptors to combat many threats. For example, attackers are often using encryption to prevent inspection methods from detecting malware coming into a network.

SSL decryptors are often used with a NIPS. The NIPS is inline but malicious traffic can get through if it’s encrypted. The SSL decryptor allows the NIPS to inspect unencrypted traffic and prevent attacks.

Hardware Security Module (HSM)

A hardware security module (HSM) is a security device you can add to a system to manage, generate, and securely store cryptographic keys. High performance HSMs are external devices connected to a network using TCP/IP. Smaller HSMs come as expansion cards you install within a server, or as devices you plug into computer ports. HSMs support the security methods as a TPM. They provide a hardware root of trust, secure boot, and can be configured for remote attestation.

Security Software

Passive vs Active Tools

In the context of tools used to discover security threats and vulnerabilities, it’s important to understand the difference between passive tools and active tools. A passive tool tests systems in a non-intrusive manner and has little possibility of compromising a system. An active tool uses intrusive and invasive methods and can potentially affect the operations of a system.

Protocol Analyzers

A protocol analyzer can capture and analyze packets on a network. The process of using a protocol analyzer is sometimes referred to as sniffing or using a sniffer. Both administrators and attackers can use a protocol analyzer to view IP headers and examine packets. For example, administrators can use a protocol analyzer to troubleshoot communication issues between network systems, or identify potential attacks using manipulated or fragmented packets.

Network Scanners

A network scanner uses various techniques to gather information about hosts within a network. As an example, Nmap (covered in more depth later in this chapter) is a popular network scanning tool that can give you a lot of information about hosts within a network. Other popular network scanning tools are Netcat and Nessus. Network scanners typically use the following methods,

Ping scan A ping scan sends an Internet Control Message Protocol (ICMP) ping to a range of IP addresses in a network. If the host responds, the network scanner knows there is a host operational with that IP address.

Arp ping scan Any host that receives an ARP packet with its IP address responds with its MAC address. If the host responds, the network scanner knows that a host is operational with that IP address.

Syn stealth scan A syn stealth scan sends a single SYN packet to each IP address in the scan range. If a host responds, the scanner knows that a host is operational with that IP address.

Port scan A port scan checks for open ports on a system. Each open port indicates the underlying protocol is running on the system. For example, if port 80 is open, it indicates the host is running HTTP and it is likely running a web server.

Service scan A service scan is like a port scan, but it goes a step further. A port scan identifies open ports and gives hints about what protocols or services might be running. The service scan verifies the protocol or service.

OS detection Operating system (OS) detection techniques analyze packets from an IP address to identify the OS. This is often referred to as TCP/IP fingerprinting.

Wireless Scanners and Crackers

Wireless scanners can typically use both passive and active scans. When using a passive scan, a scanner just listens to all the traffic being broadcast on known channels within the 2.4 GHz and 5 GHz frequency ranges.

Password Crackers

A password cracker attempts to discover a password. Passwords are typically encrypted or hashed so that they aren’t easily readable. Some methods are stronger than others. If passwords are protected with weak methods, a password cracker can discover the password.

Vulnerability Scanners

A key part of a vulnerability assessment is a vulnerability scan. Security administrators often use a vulnerability scanner to identify which systems are susceptible to attacks. Vulnerability scanners identify a wide range of weaknesses and known security issues that attackers can exploit. Most vulnerability scanners combine multiple features into a single package. A vulnerability scan often includes the following actions,

Identify vulnerabilities Identify misconfigurations Passively test security controls Identify lack of security controls

Some of the vulnerabilities and common misconfigurations discovered by a vulnerability scanner include,

Open ports Open ports can signal a vulnerability, especially if administrators aren’t actively managing the services associated with these ports. For example, all web servers do not use File Transfer Protocol (FTP), so if TCP ports 20 and 21 are open, it indicates a potential vulnerability related to FTP. Similarly, Telnet uses port 23, so if this port is open, an attacker can try to connect to the server using Telnet.

Weak passwords Many scanners include a password cracker that can discover weak passwords or verify that users are creating strong passwords in compliance with an organization’s policy. It is more efficient to use a technical password policy to require and enforce the use of strong passwords. However, if this isn’t possible, administrators use a separate password cracker to discover weak passwords.

Default accounts and passwords Operating systems and applications can have default usernames and passwords. Basic operating system and application hardening steps should remove the defaults, and a scan can discover the weaknesses if operating systems and applications aren’t secured properly. F o r example, some SQL database systems allow the sa (system administrator) account to be enabled with a blank password. Sensitive

data. Some scanners include data loss prevention (DLP) techniques to detect sensitive data sent over the network. For example, a DLP system can scan data looking for patterns such as Social Security numbers or key words that identify classified or proprietary data.

Security and configuration errors Vulnerability scans can also check the system against a configuration or security baseline to identify unauthorized changes.

Banner Grabbing

Banner grabbing is a technique used to gain information about remote systems and many network scanners use it. It is often used to identify the operating system along with information about some applications. If successful, the server returns a Hypertext Markup Language (HTML) banner providing information on the server.

Command Line Security Tools

ping

Ping is a basic command used to test connectivity for remote systems. You can also use it to verify a system can resolve valid host names to IP addresses, test the NIC, and check the security posture of a network.

netstat

The netstat command allows you to view statistics for TCP/IP protocols on a system. It also gives you the ability to view active TCP/IP network connections.

traceroute

The traceroute/tracert command lists the routers between two systems. In this context, each router is referred to as a hop. Tracert identifies the IP address and sometimes the host name of each hop in addition to the round-trip times (RTTs) for each hop.

Address Resolution Protocol

Determine a MAC address based on an IP address You need the hardware address to communicate

arp

Arp is a command-line tool that is related to the Address Resolution Protocol (ARP); however, arp (the command) and ARP (the protocol) are not the same thing. ARP resolves IP addresses to MAC addresses and stores the result in the ARP cache.

ipconfig and ifconfig

The ipconfig command (short for Internet Protocol configuration) shows the Transmission Control Protocol/Internet Protocol (TCP/IP) configuration information for a system. This includes items such as the computer’s IP address, subnet mask, default gateway, MAC address, and the address of a Domain Name System (DNS) server. The command shows the configuration information for all network interface cards (NICs) on a system, including both wired and wireless NICs.

tcpdump

Tcpdump is a command-line packet analyzer (or protocol analyzer). It allows you to capture packets like you can with Wireshark. The difference is that Wireshark is a Windows-based tool and tcpdump is executed from the command line. Many administrators use tcpdump to capture the packets and later use Wireshark to analyze the packet capture.

nmap

Nmap is a network scanner. It includes many capabilities, including identifying all the active hosts and their IP addresses in a network, the protocols and services running on each of these hosts, and the operating system of the host.

netcat

You can use netcat to run a port scan against a single IP address. It allows you to specify the range of ports, such as 10 through 1024 and randomize the ports scanned to evade detection. It also supports waiting longer periods of time between port checks, again, to evade detection.

Common Security Issues

Unencrypted Credentials

Unencrypted credentials are usernames and passwords sent across a network in cleartext. One of the ways attackers can view this data is by connecting an unauthorized switch within a network to capture traffic and forward it to a system running a protocol analyzer.

Logs and Event Anomalies

Many times, security administrators are searching the logs looking for event anomalies. As a simple example, attackers sometimes try to log on to accounts by guessing passwords. Security logs will record these attempts as failed logons, which is an event anomaly. After investigating the failed logons, administrators can determine if the failed logons were part of normal operation, or a security incident.

Permission Issues

A common security issue with permissions is giving users more permissions than they need. The principle of least privilege is a core security principle and mentioned several times in this book. In short, it means that users are given only the rights and permissions they need to do their job, and no more. When users have more permissions than they need, they can accidentally, or maliciously, cause problems.

Access Violations

An access violation occurs if users access materials that they shouldn’t. As an example, imagine that Bart is a help-desk technician. During a review of logs, security administrators discover that Bart has accessed payroll data though he has no business looking at this data. This is an access violation and should be investigated. A primary objective of security investigators is to discover how Bart accessed the materials.

Certificate Issues

Before clients use a certificate, they first verify it is valid with some checks. There are many different certificate issues that can result in an invalid certificate. Browsers typically display an error describing the issue and encouraging users not to use the certificate. Applications that detect a certificate issue might display an error using a certificate, but they are typically coded to not use it.

Data Exfiltration

Data exfiltration is the unauthorized transfer of data outside an organization and is a significant concern. In some cases, attackers take control of systems and transfer data outside an organization using malware. It’s also possible for malicious insiders to transfer data.

Unauthorized Software

A third-party app store is something other than Apple’s App Store or Google Play. Apps obtained from these third-party app stores don’t undergo the same level of scrutiny as apps on the App Store or Google Play and represent a higher risk. Apple makes it very difficult to obtain apps from a third-party app store, but it is relatively easy to obtain apps from third-party stores for Android devices.

Jailbreaking refers to removing all software restrictions from an Apple device. After jailbreaking a device, users can install software from any third party source. Rooting is the process of modifying an Android device to give the user root-level (or full administrator) access to the device. Both rooting and jailbreaking introduce risks and vulnerabilities to the device, so it’s common for an MDM to block all access to a network if it detects a device has either been rooted or jailbroken.

Analyzing Security Output

Host-Based IDS/IPS

A host-based intrusion detection system (HIDS) is additional software installed on a system such as a workstation or server. It provides protection to the individual host and can detect potential attacks and protect critical operating system files. The primary goal of any IDS is to monitor traffic. For a HIDS, this traffic passes through the network interface card (NIC).

Many host-based IDSs have expanded to monitor application activity on the system. As one example, you can install a HIDS on different Internet facing servers, such as web servers, mail servers, and database servers. In addition to monitoring the network traffic reaching the servers, the HIDS can also monitor the server applications.

It’s worth stressing that a HIDS can help detect malicious software (malware) that traditional antivirus software might miss. Because of this, many organizations install a HIDS on every workstation as an extra layer of protection in addition to traditional antivirus software.

File Integrity Check

Some antivirus scanners use file integrity checkers to detect modified system files. A file integrity checker calculates hashes on system files as a baseline. It then periodically recalculates the hashes on these files and compares them with the hashes in the baseline. If the hashes are ever different, it indicates the system files have

been modified. When an antivirus scanner detects a modified file, it sends an alert. Many times, these alerts can detect rootkit infections.

Application Whitelisting

Whitelisting and blacklisting are two additional methods used to protect hosts, including workstations, servers, and mobile devices. An application whitelist is a list of applications authorized to run on a system. An application blacklist is a list of applications the system blocks.

Advanced Malware Tools

Many vendors have begun developing advanced malware tools. These go beyond just examining files to determine if they are malware. As an example, Cisco's Advanced Malware Protection (AMP) combines multiple technologies to protect a network before an attack, during an attack, and after an attack.

AMP analyzes a network to prevent attacks using threat intelligence and analytics. It collects worldwide threat intelligence from Cisco's Security Intelligence organization, Talos Security Intelligence and Research Group, and Threat Grid intelligence feeds. This information helps it detect and alert on malware similar to any antivirus software.

UTM/All-in-one Security Appliance

Unified threat management (UTM) is a single solution that combines multiple security controls. The overall goal of UTMs is to provide better security, while also simplifying management requirements. In many cases, a UTM device will reduce the workload of administrators without sacrificing security.

As IT-based threats first began appearing, security experts created various solutions to deal with each of them. When attackers began releasing malware to infect computers, vendors created antivirus software. Attackers started attacking networks, and in response, security experts developed and steadily improved firewalls. When organizations recognized a need to control what sites users can visit, organizations implemented proxies with URL filters.

UTM security appliances include multiple capabilities, including,

URL filtering URL filters within a UTM security appliance perform the same job as a proxy server. They block access to sites based on the URL. It's common to subscribe to a service and select categories to block access to groups of sites. Administrators can also configure URL filters manually to allow or block access to specific web sites. As an example, if an administrator realizes that users are routinely connecting to a peer-to- peer (P2P) file sharing site, the administrator can add the URL to the filter, and block access to that site.

Malware inspection Malware often comes into a network via spam, or malicious web pages. The malware inspection component of a UTM appliance screens incoming data for known malware and blocks it. Organizations often scan for malware at email servers and at individual systems as part of a layered security or defense-in-depth solution.

Content inspection Content inspection includes a combination of different content filters. It monitors incoming data streams and attempts to block any malicious content. It can include a spam filter to inspect incoming email and reject spam. It

can also block specific types of transmissions, such as streaming audio and video, and specific types of files such as Zip files.

DDoS mitigator A DDoS mitigator attempts to detect DDoS attacks and block them. This is similar to how intrusion prevention systems (IPSs) block attacks.

Data Execution Prevention (DEP)

Data execution prevention (DEP) is a security feature that prevents code from executing in memory regions marked as non-executable. It helps prevent an application or service from executing code from a non-executable memory region. The primary purpose of DEP is to protect a system from malware.

Mobile Connection Methods

Cellular Networks

Smartphones (and many tablets) include the ability to connect to a cellular network, such as a third generation (3G), longterm- evolution (LTE), fourth generation (4G), or 4G LTE network. The type of network you connect with is dependent on your cellular provider. Newer generations typically provide increased speed for digital transfers and improved voice communications.

Wi-Fi

Mobile devices almost always have a wireless network interface that you can configure to connect to a wireless network. Typical wireless networks require you to enter or select the service set identifier (SSID) and enter the pre-shared key or password to access the network. More secure networks use Enterprise mode with an 802.1x server.

Satellite Communications - SATCOM

Some mobile devices support connections to networks using satellite communications (SATCOM). The most common usage of SATCOM is in mobile phones rather than tablets. However, you can purchase satellite hot spots. You can connect mobile devices to the hot spot, and the hot spot provides Internet and voice access via a satellite connection. Additionally, some vehicles include satellite communication technologies that can be used for phone calls and sometimes for shared Internet access.

Near Field Communication (NFC)

NFC is most commonly used as a payment gateway allowing you to make payments simply by waving your phone in front of an NFC reader at a retailer. You can also create a peer-to-peer network between two devices with NFC. For example, Android Beam allows two users with Android devices to share data displayed on the screen by placing two devices back to back. Some applications use NFC to enable Bluetooth on the two devices, send the shared data via Bluetooth, and then disable Bluetooth.

ANT/ANT+

ANT and ANT+ are proprietary wireless protocols used by some mobile devices. Many sports and fitness sensors (such as Fitbit) collect data on users (such as heart rate, steps taken, and so on) and use ANT to send the data to a mobile device application.

IR (Infrared)

Infrared is a line-of-sight wireless technology used by some mobile devices. This is the same technology used by most remote controls for TVs and other audiovisual equipment. Many people add apps to their smartphones and use them as a universal remote for their equipment. It’s also possible to transfer files between smartphones using infrared, as long as both smartphones support infrared.

USB (Universal Serial Bus)

Mobile devices can typically be connected to a desktop PC or laptop via a USB cable. Most Apple devices have a Lightning port and can connect to PCs via a Lightning to USB cable. Many Android devices have a mini-USB cable and can connect to PCs via a mini-USB to standard USB cable.

Mobile Device Management (MDM)

Mobile device management (MDM) includes the technologies to manage mobile devices. The goal is to ensure these devices have security controls in place to keep them secure. MDM applications help administrators manage mobile devices. The following bullets describe many of the MDM concepts that apply to mobile devices.

Application Management

MDM tools can restrict what applications can run on mobile devices. They often use application whitelists to control the applications and prevent unapproved applications from being installed.

Content Management

After creating segmented storage spaces, it’s important to ensure that appropriate content is stored there. An MDM system can ensure that all content retrieved from an organization source (such as a server) is stored in an encrypted segment. Also, content management can force the user to authenticate again when accessing data within this encrypted segment.

Remote Wipe

Remote wipe capabilities are useful if the phone is lost. It sends a remote signal to the device to wipe or erase all the data. The owner can send a remote wipe signal to the phone to delete all the data on the phone. This also deletes any cached data, such as cached online banking passwords, and provides a complete sanitization of the device by removing all valuable data.

Geolocation

Mobile devices commonly include Global Positioning System (GPS) capabilities that can be used for geolocation. Applications commonly use GPS to identify the location of the device. This can also be used to locate a lost device.

Geofencing

Organizations sometimes use GPS to create a virtual fence or geographic boundary using geofencing technologies. Apps can respond when the device is within the virtual fence. As an example, an organization can configure mobile apps so that they will only run when the device is within the virtual fence. Similarly, an organization can configure a wireless network to only operate for mobile devices within the defined boundary.

Screen Lock

Most devices support the use of a passcode or password to lock the device. This is like a password-protected screen saver on desktop systems that automatically locks the device after a period of time. It prevents someone from easily accessing the device and the data it contains. This is often combined with an erase function. For example, if someone steals the phone and enters the incorrect passcode ten times, the smartphone will automatically erase all data on the phone.

Push Notification Services

Push notification services send messages to mobile devices from apps. As an example, if Lisa installs the Facebook app on her smartphone and enables notifications, the Facebook app will send her notifications. Software developers can configure the notifications to appear even if the device is in screen lock mode and even if the app is not running. MDM apps can send notifications to remind users of security settings, or to let them know if their device is complying with security policy requirements.

Biometrics

Many mobile devices now support biometrics for authentication. For example, you can teach the device your fingerprint and then use your fingerprint to authenticate instead of entering a password or PIN.

Context-Aware Authentication

Context-aware authentication uses multiple elements to authenticate a user and a mobile device. It can include the user’s identity, geolocation, verification that the device is within a geofence, time of day, and type of device. Combined, these elements help prevent unauthorized users from accessing apps or data.

Containerization

Containerization can also be implemented in mobile devices. By running an application in a container, it isolates and protects the application, including any of its data. This is very useful when an organization allows employees to use their own devices. It’s possible to encrypt the container to protect it without encrypting the entire device.

Full Device Encryption

Encryption protects against loss of confidentiality on multiple platforms, including workstations, servers, mobile devices, and data transmissions. Encryption methods such as full device encryption provide device security, application security, and data security.

Mobile Device Enforcement

Third-Party App Stores

A third-party app store is something other than Apple’s App Store or Google Play. Apps obtained from these third-party app stores don’t undergo the same level of scrutiny as apps on the App Store or Google Play and represent a higher risk. Apple makes it very difficult to obtain apps from a third-party app store, but it is relatively easy to obtain apps from third-party stores for Android devices.

Rooting/Jailbreaking

Jailbreaking refers to removing all software restrictions from an Apple device. After jailbreaking a device, users can install software from any third party source. Rooting is the process of modifying an Android device to give the user root-level (or full administrator) access to the device. Both rooting and jailbreaking introduce risks and vulnerabilities to the device, so it’s common for an MDM to block all access to a network if it detects a device has either been rooted or jailbroken.

Firmware OTA Updates

Mobile devices typically have the operating system stored in onboard memory such as flash memory, which retains data even without power. Because the operating system is the software and the memory is hardware, this is commonly called firmware. Updates to the operating system overwrite the firmware using over-the-air (OTA) techniques. Firmware OTA updates keep the device up to date.

SMS/MMS

Many people use text messaging services such as Short Message Service (SMS) and Multimedia Messaging Service (MMS). SMS is a basic text messaging service supported on many telephone and mobile devices. MMS is an extension of SMS that allows users to include multimedia content such as a picture, a short video, audio, or even a slideshow of multiple images.

USB OTG

MDM tools can also prevent the use of external media and Universal Serial Bus On-The-Go (USB OTG) cables. Mobile devices commonly have one or more ports where you can plug in a cable. Apple devices have a Lightning port and Android devices typically have a micro-USB or mini-USB. In some cases, it’s possible to connect external media (such as an external drive) to the device.

Hotspot/Tethering

Most smartphones support tethering, which allows you to share one device’s Internet connection with other devices. As an example, you can connect your smartphone to the Internet and then use this Internet connection with a laptop, a tablet, or any device that has a wireless connection.

Mobile Deployment Models

BYOD

(Bring Your Own Device) Some organizations allow employees to bring their own mobile devices to work and attach them to the network. Employees are responsible for selecting and supporting the device and they typically must comply with a BYOD policy when connecting their device to the network. While this is simple for the employees, it is sometimes referred to as bring your own disaster among IT professionals. Because employees can have any possible device, the IT department is now responsible for supporting, monitoring, and managing any possible device owned by employees.

CYOD (Choose Your Own Device)

To avoid some of the challenges related to supporting any possible mobile devices, some organizations create a list of acceptable devices along with a CYOD policy. Employees can purchase devices on the list and bring them to work. This gives the IT department a specific list of devices that they need to support, monitor, and manage.

COPE

(Corporate-Owned, Personally Enabled) COPE is similar to the traditional corporate-owned model, but the primary difference is that the employees are free to use the device as if it was their personally owned device. This allows employees to use the devices for personal activities in addition to connecting them to the organization's network. Because the organization owns the devices, it makes it easier to manage them.

Corporate-Owned

In this traditional deployment model, the organization purchases devices and issues them to employees.

VDI/VMI

(Virtual Desktop Infrastructure) While these are typically accessed by traditional computers within a network, it's also possible to deploy a VDI that users can access with their mobile device. This allows users to access any applications installed on their desktop. When the organization hosts a remote access solution such as a virtual private network (VPN), users can access the mobile VDI from anywhere if they have Internet access.

Secure Protocols

Voice and Video

SRTP The Real-time Transport Protocol (RTP) delivers audio and video over IP networks. This includes Voice over Internet Protocol (VoIP) communications, streaming media, video teleconferencing applications, and devices using web-based push-to-talk features. However, organizations often want to secure these transmissions. The Secure Real-time Transport Protocol (SRTP) provides encryption, message authentication, and integrity for RTP.

SRTP helps protect the confidentiality of data from these attacks while also ensuring the integrity of the data transmissions. This provides protection against replay attacks. In a replay attack, an attacker captures data sent between two entities, modifies it, and then attempts to impersonate one of the parties by replaying the data. SRTP can be used for both unicast transmissions (such as one person calling another) and multicast transmissions where one person sends traffic to multiple recipients.

Time Synchronization

NTP Within a Microsoft domain, one domain controller periodically uses the Windows Time service to locate a reliable Internet server running the Network Time Protocol (NTP). NTP is the most commonly used protocol for time synchronization, allowing systems to synchronize their time to within tens of milliseconds. Other domain controllers within the network periodically synchronize their time with the first domain controller. Last, all computers in the domain synchronize their time with one

of these domain controllers. This process ensures all the computers have the accurate time.

Email

TLS The Transport Layer Security (TLS) protocol is the designated replacement for SSL and should be used instead of SSL. Additionally, many protocols that support TLS use STARTTLS. STARTTLS looks like an acronym, but it isn't. Instead, it is a command used to upgrade an unencrypted connection to an encrypted connection on the same port.

Web

HTTPS Hypertext Transfer Protocol Secure (HTTPS) encrypts web traffic to ensure it is secure while in transit. Web browsers commonly indicate that a secure session is using HTTPS by displaying a lock icon and by including HTTPS in the Uniform Resource Locator (URL) field. HTTPS is encrypted with either SSL or TLS and it uses TCP port 443.

File Transfer

FTP File Transfer Protocol (FTP) uploads and downloads large files to and from an FTP server. By default, FTP transmits data in cleartext, making it easy for an attacker to capture and read FTP data with a protocol analyzer. FTP active mode uses TCP port 21 for control signals and TCP port 20 for data. FTP passive mode (also known as PASV) uses TCP port 21 for control signals, but it uses a random TCP port for data. If FTP traffic is going through a firewall, this random port is often blocked, so it is best to disable PASV in FTP clients.

LDAP (Lightweight Directory Access Protocol)

LDAP Lightweight Directory Access Protocol(LDAP) is the protocol used to communicate with directories such as AD DS. LDAP provides a clear syntax for object identification and management. LDAP uses TCP port 389. LDAP Secure (LDAPS) encrypts data with TLS using TCP port 636.

Remote Access

SSH Secure Shell (SSH) encrypts traffic in transit and can be used to encrypt other protocols such as FTP. Linux administrators often used Telnet when remotely administering systems, but this is not recommended because Telnet sends traffic over the network in cleartext. Instead, administrators commonly use SSH to remotely administer systems. Secure Copy (SCP) is based on SSH and is used to copy encrypted files over a network.

Domain Name Resolution

DNSSEC One risk with DNS is DNS poisoning, also known as DNS cache poisoning. When successful, attackers modify the DNS cache with a bogus IP address. For example, imagine an attacker wants to send users to a malicious web site each time they want to go to msn.com. One way is to modify the A or AAAA record in the DNS cache for msn.com. Instead of sending users to the IP address used by msn.com, it will send users to the IP address of the malicious web site.

One of the primary methods of preventing DNS cache poisoning is with Domain Name System Security Extensions (DNSSEC). DNSSEC is a suite of extensions to DNS that provides validation for DNS responses. It adds a digital signature to each record that

provides data integrity. If a DNS server receives a DNSSEC-enabled response with digitally signed records, the DNS server knows that the response is valid.

Compliance & Frameworks

Regulatory

Sarbanes-Oxley Act (SOX) The Public Company Accounting Reform and Investor Protection Act of 2002

The Health Insurance Portability and Accountability Act (HIPAA) Extensive healthcare standards for storage, use, and transmission of health care information

The Gramm-Leach-Bliley Act of 1999 (GLBA) Disclosure of privacy information from financial institutions

Non-Regulatory

A non-regulatory framework is not required by any law. Instead, it typically identifies common standards and best practices that organizations can follow. As an example, COBIT (Control Objectives for Information and Related Technologies) is a framework that many organizations use to ensure that business goals and IT security goals are linked together.

Frameworks

A framework is a structure used to provide a foundation. Cybersecurity frameworks typically use a structure of basic concepts and they provide guidance to professionals on how to implement security in various systems.

Industry-Specific Frameworks

Some frameworks only apply to certain industries. As an example, organizations that handle credit cards typically comply with the Payment Card Industry Data Security Standard (PCI DSS).

Secure Configuration Guides

Secure Configurations

A trusted operating system meets a set of predetermined requirements with a heavy emphasis on authentication and authorization. The overall goal of a trusted operating system is to ensure that only authorized personnel can access data based on their permissions. Additionally, a trusted operating system prevents any modifications or movement of data by unauthorized entities. A trusted OS helps prevent malicious software (malware) infections because it prevents malicious or suspicious code from executing. A trusted OS meets a high level of security requirements imposed by a third party.

Hardening

Hardening is the practice of making an operating system (OS) or application more secure from its default installation. It helps eliminate vulnerabilities from default configurations, misconfigurations, and weak configurations.

A core principle associated with secure systems design is least functionality. Systems should be deployed with only the applications, services, and protocols they need to meet their purpose. If a service or protocol is not running on a system, attackers cannot attack it. As a simple example, a system is not vulnerable to any File Transfer Protocol (FTP) attacks if FTP is not running and available on the system.

In addition to disabling unnecessary services to reduce vulnerabilities, it’s important to uninstall unneeded software. Software frequently has bugs and vulnerabilities. Although patching software frequently closes these vulnerabilities, you can eliminate these vulnerabilities by simply eliminating unneeded applications.

Defense-in-Depth

Layering The Defense

Control diversity is the use of different security control types, such as technical controls, administrative controls, and physical controls. For example, technical security controls such as firewalls, intrusion detection systems (IDSs), and proxy servers help protect a network. Physical security controls can provide extra protection for the server room or other areas where these devices are located. Administrative controls such as vulnerability assessments and penetration tests can help verify that these controls are working as expected.

Vendor diversity is the practice of implementing security controls from different vendors to increase security.

Defense in Depth

Defense in depth refers to the security practice of implementing several layers of protection. You can’t simply take a single action, such as implementing a firewall or installing antivirus software, and consider yourself protected. You must implement security at several different layers. This way, if one layer fails, you still have additional layers to protect you.

Firewall DMZ Hashing passwords Authentication Intrusion detection system VPN access Card/badge access Anti-virus and anti-malware software Security guard

Secure Network Topologies

Zones and Topologies

Most networks have Internet connectivity, but it’s rare to connect a network directly to the Internet. Instead, it’s common to divide the network into different zones, using different topologies.

DMZ

The demilitarized zone (DMZ) is a buffered zone between a private network and the Internet. Attackers seek out servers on the Internet, so any server placed directly on the Internet has the highest amount of risk. However, the DMZ provides a layer of protection for these Internet-facing servers, while also allowing clients to connect to them.

Extranet

An extranet is part of a network that can be accessed by authorized entities from outside of the network. For example, it’s common for organizations to provide access to authorized business partners, customers, vendors, or others.

Intranet

An intranet is an internal network. People use the intranet to communicate and share content with each other. While it’s common for an intranet to include web servers, this isn’t a requirement.

Honeypots and Honeynets

A honeypot is a sweet-looking server—at least it’s intended to look sweet to the attacker, similar to how honey looks sweet to a bear. It’s a server that is left open or appears to have been sloppily locked down, allowing an attacker relatively easy access. The intent is for the server to look like an easy target so that the attacker spends his time in the honeypot instead of in a live network. In short, the honeypot diverts the attacker away from the live network.

Security personnel often use honeypots as a tool to gather intelligence on the attacker. Attackers are constantly modifying their methods to take advantage of different types of attacks. Some sophisticated attackers discover vulnerabilities before a patch is released (also known as a zero-day exploit, or zero-day vulnerability). In some cases, security professionals observe attackers launching zero-day vulnerability attacks against a honeypot.

Honeypots never hold any data that is valuable to the organization. The data may appear to be valuable to an attacker, but its disclosure is harmless. Honeypots have two primary goals:

Divert attackers from the live network. If an attacker is spending time in the honeypot, he is not attacking live resources.

Allow observation of an attacker. While an attacker is in the honeypot, security professionals can observe the attack and learn from the attacker’s methodologies. Honeypots can also help security professionals learn about zero-day exploits, or previously unknown attacks. A honeynet is a group of honeypots within a separate network or zone, but accessible from an organization’s primary network. Security professionals often create honeynets using multiple virtual servers contained within a single physical server. The servers within this network are honeypots and the honeynet mimics the functionality of a live network.

NAT - Network Address Translation

Network Address Translation (NAT) is a protocol that translates public IP addresses to private IP addresses and private addresses back to public. You’ll often see NAT enabled on an Internet-facing firewall. A commonly used form of NAT is network address and port translation, commonly called Port Address Translation (PAT). Some of the benefits of NAT include,

Public IP addresses don’t need to be purchased for all clients A home or company network can include multiple computers that can access the Internet through one router running N AT. Larger companies requiring more bandwidth may use more than one public IP address.

NAT hides internal computers from the Internet Computers with private IP addresses are isolated and hidden from the Internet. NAT provides a layer of protection to these private computers because they aren’t as easy to attack and exploit from the Internet.

Network Segmentation

Air Gaps

Physical isolation ensures that a network isn’t connected to any other network. As an example, consider supervisory control and data acquisition (SCADA) systems. These are typically industrial control systems within large facilities such as power plants or water treatment facilities. While SCADA systems operate within their own network, it’s common to ensure that they are isolated from any other network.

This physical isolation significantly reduces risks to the SCADA system. If an attacker can’t reach it from the Internet, it is much more difficult to attack it. However, if the system is connected to the internal network, it’s possible for an attacker to gain access to internal computers, and then access any resource on the internal network.

An air gap is a metaphor for physical isolation, indicating that there is a gap of air between an isolated system and other systems. When considered literally, an air-gapped system is not connected to any other systems. As an example, many organizations use both classified (red) and unclassified (black) networks. Strict rules ensure that these two systems are not connected to each other. Some rules require that any cable from a red network must be physically separated from black network cables.

Layer 2 vs Layer 3 Switch

A traditional switch operates on Layer 2 of the Open Systems Interconnection (OSI) model. As discussed previously, a traditional switch (a Layer 2 switch) uses the destination MAC address within packets to determine the destination port. Additionally, a Layer 2 switch forwards broadcast traffic to all ports on the switch.

Routers operate on Layer 3 of the OSI model. They forward traffic based on the destination IP address within a packet, and they block broadcast traffic. A Layer 3 switch mimics the behavior of a router and allows network administrators to create virtual local area networks (VLANs). Because a Layer 3 switch forwards traffic based on the destination IP address instead of the MAC address, it is not susceptible to ARP-based attacks.

VLAN

A virtual local area network (VLAN) uses a switch to group several different computers into a virtual network. You can group the computers together based on departments, job function, or any other administrative need. This provides security because you’re able to isolate the traffic between the computers in the VLAN. Normally, a router would group different computers onto different subnets, based on physical locations. All the computers in a routed segment are typically located in the same physical location, such as on a specific floor or wing of a building. However, a single Layer 3 switch can create multiple VLANs to separate the computers based on logical needs rather than physical location. Additionally, administrators can easily reconfigure the switch to add or subtract computers from any VLAN if the need arises.

VPN Technologies

Site-to-Site VPNs

Encrypt traffic between sites Through the public Internet Use existing Internet connection No additional circuits or costs

Host-to-Site VPNs

Also called “remote access VPN” Requires software on the user device May be built-in to existing operating system

Host-to-Host VPNs

User to user encryption Software-based No hardware needed

Security Technology Placement

Sensors and Collectors

Gather information from network devices. Built-in sensors, separate devices. Integrated into switches, routers, servers, firewalls, etc.

Sensors Intrusion prevention systems, firewall logs, authentication logs, web server access logs, database transaction logs, email logs.

Collectors Proprietary consoles (IPS, firewall), SIEM consoles, syslog servers. Many SIEMs include a correlation engine to compare diverse sensor data.

Filters and Firewalls

Packet filters Simple data blocks - ignores state. Linux iptables - filter packets in the kernel. Usually placed on a device or server.

Firewalls State-based Advanced filtering by IP address, port, application, content. Usually located on the ingress/egress of a network. Some organizations place them between internal networks.

Proxy Servers

An intermediate server Client makes the request to the proxy. The proxy performs the actual request. The proxy provides results back to the client.

Useful features Access control, caching, URL filtering, content scanning.

Forward proxy Protect users from the Internet.

VPN Concentrators

VPN appliances are usually located on the edge of the network Internet-facing. Sites connect from one site to another across the Internet.

Load Balancers

Manage the load across multiple devices. The user has no idea. Placed between the users and the service.

Servers can be added and removed. Real-time response to load.

Load balancer performs constant health checks. If a server disappears, it is removed from the rotation.

SSL Accelerators

The SSL handshake requires some cryptographic overhead. A lot of CPU cycles Offload the SSL process to a hardware accelerator. Often integrated into a load balancer.

DDos Mitigation

Resist a distributed denial of service attack. Minimize the impact

Cloud-based Internet provider or reverse proxy service.

On-site tools DDoS filtering in a firewall or IPS. Positioned between you and the Internet. Literally you against the world.

Taps and Port Mirrors

Intercept network traffic Send a copy to a packet capture device.

Physical taps Disconnect the link, put a tap in the middle. Can be an active or passive tap.

Port mirror Port redirection, SPAN (Switched Port ANalyzer) Software-based tap. Limited functionality, but can work well in a pinch.

Securing SDN

SDN (Software Defined Networking)

A software defined network (SDN) uses virtualization technologies to route traffic instead of using hardware routers and switches. More specifically, an SDN separates the data planes and control planes within a network. Another way of thinking of this is that an SDN separates the logic used to forward or block traffic (the data plane) and the logic used to identify the path to take (the control plane).

Hardware routers use rules within an ACL to identify whether a router will forward or block traffic on the data plane. This is always proprietary because it’s implemented on specific hardware routers. However, an SDN implements the data plane with software and virtualization technologies, allowing an organization to move away from proprietary hardware.

Routing protocols such as Open Shortest Path First (OSPF) and Border Gateway Protocol (BGP) help routers determine the best path to route traffic on the control plane. Routers use these protocols to share information with each other, creating a map of the known network. An SDN can still use these routing protocols, but without the hardware routers.

Hardware Security

Full Disk Encryption (FDE) / Self-Encrypting Drive (SED)

Full disk encryption (FDE) encrypts an entire disk. Several applications are available to do this. For example, VeraCrypt is an open source utility that can encrypt partitions or the entire storage device.

Many hardware vendors now manufacture hardware-based FDE drives. These are sometimes referred to as self-encrypting drives (SEDs). An SED includes the hardware and software to encrypt all data on the drive and securely store the encryption keys. These typically allow users to enter credentials when they set up the drive. When users power up the system, they enter their credentials again to decrypt the drive and boot the system.

Trusted Platform Module (TPM)

A Trusted Platform Module (TPM) is a hardware chip on the computer's motherboard that stores cryptographic keys used for encryption. Many laptop computers include a TPM and you may see them on many mobile devices, too. However, if the system doesn't include a TPM, it is not feasible to add one. Once enabled, the TPM provides full disk encryption capabilities. It keeps hard drives locked, or sealed, until the system completes a system verification and authentication process.

A TPM supports secure boot and attestation processes. When the TPM is configured, it captures signatures of key files used to boot the computer and stores a report of the signatures securely within the TPM. When the system boots, the secure boot process checks the files against the stored signatures to ensure they haven't changed. If it detects that the files have been modified, such as from malware, it blocks the boot process to protect the data on the drive. A remote attestation process works like the secure boot process. However, instead of checking the boot files against the report stored in the TPM, it uses a separate system. Again, when the TPM is configured, it captures the signatures of key files, but sends this report to a remote system. When the system boots, it checks the files and sends a current report to the remote system. The remote system verifies the files are the same and attests, or confirms, that the system is safe.

The TPM ships with a unique Rivest, Shamir, Adleman (RSA) private key burned into it, which is used for asymmetric encryption. This private key is matched with a public key and provides a hardware root of trust, or a known secure starting point. The private key remains private and is matched with a public key. Additionally, the TPM can generate, store, and protect other keys used for encrypting and decrypting disks.

Hardware Security Module (HSM)

A hardware security module (HSM) is a security device you can add to a system to manage, generate, and securely store cryptographic keys. High performance HSMs are external devices connected to a network using TCP/IP. Smaller HSMs come as expansion cards you install within a server, or as devices you plug into computer ports. HSMs support the security methods as a TPM. They provide a hardware root of trust, secure boot, and can be configured for remote attestation.

UEFI BIOS

The Basic Input/Output System (BIOS) includes software that provides a computer with basic instructions on how to start. It runs some basic checks, locates the operating system, and starts. The BIOS is often referred to as firmware. It is a hardware chip that you can physically see and touch and it includes software that executes code on the computer. The combination of hardware and software is firmware.

Newer systems use Unified Extensible Firmware Interface (UEFI) instead of BIOS. UEFI performs many of the same functions as BIOS, but provides some enhancements. As an example, it can boot from larger disks and it is designed to be CPU-independent. Both BIOS and UEFI can be upgraded using a process called flashing. Flashing overwrites the software within the chip with newer software.

EMI/EMP

When designing systems, it's important to consider electromagnetic interference (EMI) a n d electromagnetic pulse (EMP). EMI comes from sources such as motors, power lines, and fluorescent lights and it can interfere with signals transmitted over wires. EMP is a short burst of electromagnetic energy. EMP can come from a wide assortment of sources and some sources can cause damage to computing equipment.

Operating System Security

Operating System Types

There are three primary types of computer operating systems (OSs): Windows, Apple's operating systems, and Linux- or Unix-based systems. While you primarily see OSs operating on desktops, laptops, and servers, they are also operating in other locations, including:

Kiosks A kiosk is a small structure in an open area used to sell something, provide information, or display advertisements. For example, an organization can create a touch-screen application installed on a computer and place it in a kiosk. This could be in a mall or store (designed to advertise something), in a medical center (designed to share information), or anywhere an organization thinks it might be useful.

Network Many network devices such as switches, routers, and firewalls include an operating system used to manage the device. These are often a version of Linux. Some Cisco network devices use the Cisco IOS (originally called the Internetwork Operating System).

Appliance A network appliance is a dedicated hardware device that bundles several features within it.

Patch Management

Patch management ensures that systems and applications stay up to date with current patches. This is one of the most efficient ways to reduce operating system and application vulnerabilities because it protects systems from known vulnerabilities. Patch management includes a group of methodologies and includes the process of identifying, downloading, testing, deploying, and verifying patches.

After testing the patches, administrators deploy them. They don't deploy the patches manually though. Instead, they use systems management tools to deploy the patches in a controlled manner.

Disabling Unnecessary Services

A core principle associated with secure systems design is least functionality. Systems should be deployed with only the applications, services, and protocols they need to meet their purpose. If a service or protocol is not running on a system, attackers cannot attack it.

In addition to disabling unnecessary services to reduce vulnerabilities, it’s important to uninstall unneeded software. Software frequently has bugs and vulnerabilities. Although patching software frequently closes these vulnerabilities, you can eliminate these vulnerabilities by simply eliminating unneeded applications.

Application Whitelisting/Blacklisting

Whitelisting and blacklisting are two additional methods used to protect hosts, including workstations, servers, and mobile devices. An application whitelist is a list of applications authorized to run on a system. An application blacklist is a list of applications the system blocks.

Peripheral Security

Wireless keyboards and mice

Wireless transmissions can sometimes be intercepted. If these devices are used with systems processing sensitive data, it might be prudent to use wired devices instead.

Displays

If displays show sensitive or private data, their view should be limited. F o r example, they shouldn’t be viewable from windows. Additionally, privacy screens can be placed over displays to limit the view of the information unless someone is looking straight at the display.

WiFi-enabled microSD

Traditional Micro Secure Digital (SD) cards need to be plugged into a port to read the data. They are typically used in digital cameras. However, newer MicroSD cards include wireless capabilities. As with any wireless devices, the risk is that wireless transmissions can be intercepted, so if these are necessary, they should be configured with strong wireless security.

External storage devices

External storage devices include any external device that has memory capabilities. It typically refers to external USB drives, but also includes other devices such as smartphones, tablets, MP3 players, and digital cameras. Users can plug them into a system and easily copy data to and from a system. They can transport malware without the user’s knowledge and can be a source of data leakage. Malicious users can copy and steal a significant amount of information using an easily concealable thumb drive.

Digital cameras

Digital cameras typically include built-in storage and support additional storage by plugging in a memory card. These include the same risks as any external storage device.

Secure Deployment

Sandboxing

Sandboxing is the use of an isolated area on a system and it is often used for testing. Administrators and security professionals also use sandboxing to test various

security controls before deploying them to a live production network. Virtualization provides a high level of flexibility when testing security controls because the environments are easy to re-create. For example, they can test the effectiveness of antivirus software to detect malware released within a sandbox. If the antivirus software doesn't detect the malware and the malware causes problems, it is easy to revert the system to a previous state. Also, the isolation within the sandbox prevents the malware from spreading.

Working Environments

A secure staging environment includes multiple environments, and typically includes different systems used for each stage. As an example, imagine a software development team is creating an application that will be used to sell products via the Internet. The different environments are:

Development Software developers use a development environment to create the application. This typically includes version control and change management controls to track the application development.

Test Testers put the application through its paces and attempt to discover any bugs or errors. The testing environment typically doesn't simulate a full production environment, but instead includes enough hardware and software to test software modules.

Staging The staging environment simulates the production environment and is used for late stage testing. It provides a complete but independent copy of the production environment.

Production The production environment is the final product. It includes everything needed to support the application and allow customers and others to use it. In this example, it would include the live web server, possibly a back-end database server, and Internet access.

Secure Baselines

A baseline is a known starting point and organizations commonly use secure baselines to provide known starting points for systems. One of the primary benefits of secure baselines is that they improve the overall security posture of systems. Weak security configuration is a common security issue, but secure baselines help eliminate this. The use of baselines works in three steps,

Initial baseline configuration Administrators use various tools to deploy systems consistently in a secure state.

Integrity measurements for baseline deviation Automated tools monitor the systems for any baseline changes, which is a common security issue. Some tools such as vulnerability scanners monitor the systems and report any changes they detect. Other tools such as Group Policy automatically reconfigure the systems to the baseline settings when they detect changes.

Remediation NAC methods can detect some changes to baseline settings and automatically isolate or quarantine systems in a remediation network. Typically, administrators need to correct the problems in these systems manually.

Embedded Systems

SCADA/ICS

An industrial control system (ICS) typically refers to systems within large facilities such as power plants or water treatment facilities. An ICS is controlled by a supervisory control and data acquisition (SCADA) system. Ideally, these systems are contained within isolated networks, such as within a virtual local area network (VLAN), that do not have access to the Internet. If they are connected to the corporate network, they are often protected by a network intrusion prevention system (NIPS) to block unwanted traffic.

Smart Devices/IoT (Internet of Things)

Wearable technology has exploded in recent years. It includes any device you can wear or have implanted. These devices can then be used to interact with other devices, such as a smartphone.

Home automation includes Internet-connected devices, such as wireless thermostats, lighting, coffee makers, and more. These devices typically connect to the home's network, which gives them Internet access.

HVAC

Heating, ventilation, and air conditioning (HVAC) systems keep computing systems at the proper temperature and with the proper humidity. Most have embedded systems to control them. If attackers can access these systems, they may be able to remotely turn off the HVAC system or trick it into keeping the temperature at 95 degrees within a data center. The resulting damage to systems within this data center could be catastrophic.

SoC (System on a Chip)

A system on a chip (SoC) is an integrated circuit that includes all the functionality of a computing system within the hardware. It typically includes an application contained within onboard memory, such as read-only memory (ROM), electrically erasable programmable ROM (EEPROM), or flash memory. Many mobile computing devices include an SoC.

RTOS (Real-TIme Operating System)

A real-time operating system (RTOS) is an operating system that reacts to input within a specific time. If it can't respond within the specified time, it doesn't process the data and typically reports an error.

Development Life Cycle Models

Development Life-Cycle Models

Software development life cycle (SDLC) models attempt to give structure to software development projects. Two popular models are waterfall and agile.

Waterfall

Requirements The developers work with the customer to understand the requirements. The output of this stage is a requirements document, which provides clear guidance on what the application will do.

Design Developers begin to design the software architecture in this stage. This is similar to creating the blueprints for a building. The design stage doesn’t include any detailed coding, but instead focuses on the overall structure of the project.

Implementation Developers write the code at this stage, based on the requirements and design.

Verification The verification stage ensures the code meets the requirements.

Maintenance The maintenance stage implements changes and updates as desired.

A challenge with the waterfall model is that it lacks flexibility. It is difficult to revise anything from previous stages. For example, if a customer realizes a change in the requirements is needed, it isn’t possible to implement this change until the maintenance stage.

Agile

The agile model uses a set of principles shared by cross-functional teams. These principles stress interaction, creating a working application, collaborating with the customer, and responding to change. Instead of strict phases, the agile model uses iterative cycles. Each cycle creates a working, if not complete, product. Testers verify the product works with the current features and then developers move on to the next cycle. The next cycle adds additional features, often adding small, incremental changes from the previous cycle.

A key difference of the agile model (compared with the waterfall model) is that it emphasizes interaction between customers, developers, and testers during each cycle. In contrast, the waterfall model encourages interaction with customers during the requirements stage, but not during the design and implementation stages. The agile model can be very effective if the customer has a clear idea of the requirements. If not, the customer might ask for changes during each cycle, extending the project’s timeline.

Secure DevOps

DevOps

DevOps combines the words development and operations and it is an agile-aligned software development methodology. Secure DevOps is a software development process that includes extensive communication between software developers and operations personnel. It also includes security considerations throughout the project. When applied to a software development project, it can allow developers to push out multiple updates a day in response to changing business needs.

Security Automation

Security automation uses automated tests to check code. When modifying code, it’s important to test it and ensure that the code doesn’t introduce software bugs or security flaws. It’s common to include a mirror image of the production environment and run automated tests on each update to ensure it is error free.

Continuous Integration

Continuous integration refers to the process of merging code changes into a central repository. Software is then built and tested from this central repository. The

central repository includes a version control system, and the version control system typically supports rolling back code changes when they cause a problem.

Immutable Systems

Immutable systems cannot be changed. Within the context of secure DevOps, it’s possible to create and test systems in a controlled environment. Once they are created, they can be deployed into a production environment. As an example, i t ’s possible to create a secure image of a server for a specific purpose. This image can be deployed as an immutable system to ensure it stays secure.

Baselining

Baselining refers to applying changes to the baseline code every day and building the code from these changes. For example, imagine five developers are working on different elements of the same project. Each of them have modified and verified some code on their computers. At the end of the day, each of these five developers uploads and commits their changes. Someone then builds the code with these changes and then automation techniques check the code. The benefit is that bugs are identified and corrected quicker. In contrast, if all the developers applied their changes once a week, the bugs can multiply and be harder to correct.

Infrastructure as Code (IAC)

Infrastructure as code refers to managing and provisioning data centers with code that defines virtual machines (VMs). Chapter 1 introduces virtualization concepts and many VMs are created with scripts. Once the script is created, new VMs can be created just by running the script.

Version Control Management

The Constant of Change

The primary purpose is to ensure that changes to systems do not cause unintended outages. Secure coding practices use version control and change management practices for the same reason—to prevent unintended outages.

Version Control

Version control tracks the versions of software as it is updated, including who made the update and when. Many advanced software development tools include sophisticated version control systems. Developers check out the code to work on it and check it back into the system when they’re done. The version control system can then document every single change made by the developer. Even better, this version control process typically allows developers to roll back changes to a previous version when necessary.

Change Management

Change management helps ensure that developers do not make unauthorized changes. As an example, if a customer wants a change or addition to the application, a developer doesn’t just implement it, no matter how easy it might be to do so. Instead, any changes to the application go through a specific, predefined process. The change management process allows several people to examine the change to ensure it won’t cause unintended consequences. Also, any change to the application becomes an added responsibility. If the customer discovers a bug due to this change after it ’s

delivered, the developer may be responsible for fixing it, even if it wasn’t authorized. In addition to preventing unauthorized changes and related problems, a change management process also provides an accounting structure to document the changes. Once a change is authorized and implemented, the change is documented in a version control document.

Provisioning and Deprovisioning

Provisioning & Deprovisioning

Provisioning and deprovisioning typically refers to user accounts. For example, when an employee starts working at an organization, administrators create the account and give the account appropriate privileges. This way, the user can use the account as authorization to access various resources. Deprovisioning an account refers to removing access to these resources and can be as simple as disabling the account.

Within the context of secure application development and deployment concepts, these terms apply to an application. Provisioning an application refers to preparing and configuring the application to launch on different devices and to use different application services.

As an example, developers who create iOS apps (running on Apple devices) provision the apps based on the devices they’ll run on. Apps can run on iPhones, iPads, and Macs. Additionally, these apps can use different services, such as an accelerometer and gyroscope to detect movement. The app needs to be properly provisioned with the appropriate code on the target device to use these services.

Deprovisioning an app refers to removing it from a device. For example, if a user decides to delete the app, the app should be able to remove it completely. Leaving remnants of the app consumes resources on the device.

Secure Coding Techniques

Secure Coding Concepts

Secure application development and deployment concepts are important for application developers to understand. Additionally, IT security managers who manage development projects should understand these concepts, too, even if they aren’t writing the code. Applications often provide an avenue for attackers to generate attacks unless developers create them using secure coding concepts.

Error and Exception Handling

Error-handling and exception-handling routines ensure that an application can handle an error gracefully. They catch errors and provide user-friendly feedback to the user. When an application doesn’t catch an error, it can cause the application to fail. In the worst-case scenario, improper error handling techniques within an application can cause the operating system to crash. Using effective error- and exception-handling routines protects the integrity of the underlying operating system.

Improper error handling can often give attackers information about an application. When an application doesn’t catch an error, it often provides debugging information that attackers can use against the application. In contrast, when an application catches the error, it can control what information it shows to the user.

Errors to users should be general Detailed errors provide information that attackers can use against the system, so the errors should be general. Attackers can analyze the errors to determine details about the system. For example, if an application is unable to connect with a database, a detailed error can let the attacker know exactly what type of database the system is using. This indirectly lets the attacker know what types of commands the system will accept. Also, detailed errors confuse most users.

Detailed information should be logged Detailed information on the errors typically includes debugging information. By logging this information, it makes it easier for developers to identify what caused the error and how to resolve it.

Stored Procedures

As mentioned previously, input validation provides strong protection against SQL injection attacks. Before using the data entered into a web form, the web application verifies that the data is valid.

Additionally, database developers often use stored procedures with dynamic web pages. A stored procedure is a group of SQL statements that execute as a whole, similar to a mini- program. A parameterized stored procedure accepts data as an input called a parameter. Instead of copying the user's input directly into a SELECT statement, the input is passed to the stored procedure as a parameter. The stored procedure performs data validation, but it also handles the parameter (the inputted data) differently and prevents a SQL injection attack.

User Input Validation

One of the most important security steps that developers should take is to include input validation. Input validation is the practice of checking data for validity before using it. Input validation prevents an attacker from sending malicious code that an application will use by either sanitizing the input to remove malicious code or rejecting the input.

Improper input handling (or the lack of input validation) is one of the most common security issues on web-based applications. It allows many different types of attacks, such as buffer overflow attacks, SQL injection, command injection, and cross-site scripting attacks.

Encryption

In general, sensitive data is often encrypted to prevent the unauthorized disclosure of data. If an application is accessing any sensitive data, developers need to ensure that this access doesn't result in inadvertent data exposure. For example, if an application accesses encrypted data on a different server, the application needs to ensure that the data is encrypted while in transit.

Applications need to decrypt data before processing it. When done processing the data, applications need to encrypt the data before storing it. Additionally, applications need to ensure that all remnants of the data are flushed from memory.

Certificates are used for various purposes such as authenticating users and computers. They can also be used to authenticate and validate software code. As an example, developers can purchase a certificate and associate it with an application or code. This code signing process provides a digital signature for the code and the certificate includes a hash of the code. This provides two benefits. First, the

certificate identifies the author. Second, the hash verifies the code has not been modified. If malware changes the code, the hash no longer matches, alerting the user that the code has been modified.

Obfuscation /Camouflage

Obfuscation attempts to make something unclear or difficult to understand. Code obfuscation (or code camouflage) attempts to make the code unreadable. It does things like rename variables, replace numbers with expressions, replace strings of characters with hexadecimal codes, and remove comments.

Code Reuse/Dead Code

However, when reusing code, developers should ensure that they are using all the code that they copy into another application. As an example, imagine a developer has created a module that has three purposes: create users, modify users, and authenticate users. While working on a new application, he realizes he needs a module that will authenticate users. If he simply copies the entire module into the new application, it creates dead code. Dead code is code that is never executed or used. In this example, the copied code to create and modify users isn't used in the new application, so it is dead code.

Logic errors can also create dead code. For example, imagine a function tests the value of a variable called Donuts. If Donuts has a value (such as 12), it squares it. If Donuts is null (a value of nothing), it returns an error and exits the function.

Validation Points

It's possible to perform input validation at the client and the server. Client-side execution indicates that the code runs on the client's system, such as a user's web browser. Server-side execution indicates that the code runs on the server, such as on a web server.

Client-side input validation is quicker, but is vulnerable to attacks. Server-side input validation takes longer, but is secure because it ensures the application doesn't receive invalid data. Many applications use both. Imagine Homer is using a web browser to purchase the newest version of Scrabbleships through the Duff web site. Customers cannot purchase more than three at a time.

In client-side input validation, the validation code is included in the HTML page sent to Homer. If he enters a quantity of four or more, the HTML code gives him an error message, and doesn't submit the page to the server until Homer enters the correct data.

Unfortunately, it's possible to bypass client-side validation techniques. Many web browsers allow users to disable JavaScript in the web browser, which bypasses client-side validation.

It's also possible to use a web proxy to capture the data sent from the client in the HTTP POST command and modify it before forwarding to the server.

Server-side input validation checks the inputted values when it reaches the server. This ensures that the user hasn't bypassed the client-side checks. Using both client-side and server-side validation provides speed and security. The client- side validation checks prevent round-trips to the server until the user has entered the

correct data. The server-side validation is a final check before the server uses the data.

Code Quality and Testing

Static Code Analyzers

Static code analyzers. Static code analysis examines the code without executing it. Automated tools can analyze code and mark potential defects. Some tools work as the developer creates the code, similar to a spell checker. Other tools can examine the code once it is semi-finalized.

Dynamic Analysis (fuzzing)

Dynamic analysis checks the code as it is running. A common method is to use fuzzing. Fuzzing uses a computer program to send random data to an application. In some cases, the random data can crash the program or create unexpected results, indicating a vulnerability. Problems discovered during a dynamic analysis can be fixed before releasing the application.

Stress Testing

Stress testing methods attempt to simulate a live environment and determine how effective or efficient an application operates with a load. As an example, a web application is susceptible to a DDoS attack. A stress test can simulate a DDoS attack and determine its impact on the web application.

Sandboxing

A sandbox is an isolated area used for testing programs. The term comes from a sandbox in a playground. Children can play in the sandbox where they are relatively safe (and parents can easily keep their eyes on them). Similarly, application developers can test applications in a sandbox, knowing that any changes they make will not affect anything outside the sandbox. Virtual machines (VMs) are often used for sandboxing. For example, Java virtual machines include a sandbox to restrict untrusted applications.

Model Verification

Testing helps identify and remove bugs. However, it’s also important that the software does what it’s meant to do. Model verification is the process of ensuring that software meets specifications and fulfills its intended purpose.

Virtualization Overview

Virtualization

Virtualization is a popular technology used within large data centers and can also be used on a regular personal computer (PC). It allows you to host one or more virtual systems, or virtual machines (VMs), on a single physical system. With today’s technologies, you can host an entire virtual network within a single physical system and organizations are increasingly using virtualization to reduce costs.

Virtualization typically provides the best return on investment (ROI) when an organization has many underutilized servers. For example, imagine an organization has

nine servers with each using only about 20 percent processing power, memory, and disk space. You could convert three physical servers to virtual hosts and run three guest servers on each physical server. Assuming all the servers are similar, this wouldn’t cost any more money for the physical servers. Additionally, three physical servers consume less electricity and require less heating and ventilation to maintain.

In contrast, imagine the organization has nine servers with each using about 80 percent of their processing power, memory, and disk space. Although it is possible to convert them all to virtual servers, it requires the purchase of additional hardware. The savings from less electricity and less heating and ventilation is offset by the cost of the new servers.

Hypervisors

Type 1 Type 1 hypervisors run directly on the system hardware. They are often called bare-metal hypervisors because they don’t need to run within an operating system. For example, VMware has a family of ESX/ESXi products that are Type I hypervisors.

Type 2 Type 2 hypervisors run as software within a host operating system. For example, the Microsoft Hyper-V hypervisor runs within a Microsoft operating system.

Virtualization Security

VM Sprawl

VM sprawl occurs when an organization has many VMs that aren’t managed properly. Most organizations have specific policies in place to ensure physical servers are kept up to date and personnel only make changes to these servers after going through a change management process. These same policies should also apply to virtual servers.

Consider this scenario. Bart creates a VM running a Microsoft Windows Server version to test a software application. After testing the application, he leaves the VM running. Later, Microsoft releases security patches for the server. The IT department tests these patches and applies them to all of the known servers that need them. However, because Bart didn’t tell anyone he was creating the VM, it remains unpatched and vulnerable to attack.

Another challenge with VM sprawl is that each VM adds additional load onto a server. If unauthorized VMs are added to physical servers, they can consume system resources. The servers might become slower and potentially crash.

VM Escape

VM escape is an attack that allows an attacker to access the host system from within the virtual system. As previously mentioned, the host system runs an application or process called a hypervisor to manage the virtual systems. In some situations, the attacker can run code on the virtual system and interact with the hypervisor.

Most virtual systems run on a physical server with elevated privileges, similar to administrator privileges. A successful VM escape attack often gives the attacker unlimited control over the host system and each virtual system within the host.

When vendors discover VM escape vulnerabilities, they write and release patches. Just as with any patches, it is important to test and install these patches as soon as possible. This includes keeping both the physical and the virtual servers patched.

Real world VM Escaping

March 2017 - Pwn2Own hacking contest You pwn it, you own it - along with some cash.

JavaScript engine bug in Microsoft Edge. Code execution in the Edge sandbox.

Windows 10 kernel bug. Compromise the guest operating system.

Hardware simulation bug in VMware. Escape to the host. Patches were released soon afterwards.

Cloud Deployment Models

Platform as a Service (PaaS)

Platform as a Service (PaaS) provides customers with a preconfigured computing platform they can use as needed. It provides the customer with an easy-to-configure operating system, combined with appropriate applications and on-demand computing. Many cloud providers refer to this as a managed hardware solution.

Software as a Service (SaaS)

Software as a Service (SaaS) includes any software or application provided to users over a network such as the Internet. Internet users access the SaaS applications with a web browser. It usually doesn’t matter which web browser or operating system a SaaS customer uses. They could be using Microsoft Edge, Chrome, Firefox, or just about any web browser.

As mentioned previously, web-based email is an example of SaaS. This includes Gmail, Yahoo! Mail, and others. The service provides all the components of email to users via a simple web browser.

If you have a Gmail account, you can also use Google Docs, another example of SaaS. Google Docs provides access to several SaaS applications, allowing users to open text documents, spreadsheets, presentations, drawings, and PDF files through a web browser.

Infrastructure as a Service (IaaS)

Infrastructure as a Service (IaaS) allows an organization to outsource its equipment requirements, including the hardware and all support operations. The IaaS service provider owns the equipment, houses it in its data center, and performs all the required hardware maintenance. The customer essentially rents access to the equipment and often pays on a per-use basis.

Many cloud providers refer to this as a self-managed solution. They provide access to a server with a default operating system installation, but customers must configure it and install additional software based on their needs. Additionally, customers are responsible for all operating system updates and patches.

IaaS can also be useful if an organization is finding it difficult to manage and maintain servers in its own data center. By outsourcing its requirements, the company limits its hardware footprint. It can do this instead of, or in addition to, virtualizing some of its servers. With IaaS, it needs fewer servers in its data center and fewer resources, such as power, HVAC, and personnel to manage the servers.

Security as a Service

Another entry into cloud computing is Security as a Service. It includes any services provided via the cloud that provide security services, and is commonly viewed as a subset of the Software as a Service (SaaS) model.

A key benefit of Security as a Service is that it outsources the administrative tasks associated with implementing the service. Additionally, professionals are focused on the specific security services offered, eliminating the need for employees to be experts on everything.

Organizations that use cloud resources often add a cloud access security broker (CASB) for additional security. This is a software tool or service deployed between an organization’s network and the cloud provider. It monitors all network traffic and can enforce security policies. As an example, it can ensure that all data stored in the cloud is encrypted.

Cloud Deployment Models

There are four categories of cloud deployment models: public, private, community, and hybrid. These identify who has access to the cloud infrastructure.

Public cloud services are available from third-party companies, such as Amazon, Google, Microsoft, and Apple. They provide similar services to anyone willing to pay for them.

A private cloud is set up for specific organizations. For example, the Shelbyville Nuclear Power Plant might decide it wants to store data in the cloud, but does not want to use a third- party vendor. Instead, the plant chooses to host its own servers and make these servers available to internal employees through the Internet.

Communities with shared concerns (such as goals, security requirements, or compliance considerations) can share cloud resources within a community cloud. As an example, imagine that the Shelbyville Nuclear Power Plant and several schools within Springfield decided to share educational resources within a cloud. They could each provide resources for the cloud and only organizations within the community would have access to the resources.

Not all cloud implementations fit exactly into these definitions, though. A hybrid cloud is a combination of two or more clouds. They can be private, public, community, or a combination. These retain separate identities to help protect resources in private clouds. However, they are bridged together, often in such a way that it is transparent to the users.

Security in the Cloud

On-Premise & Hosted

On-premise All resources are owned, operated, and maintained within the organization’s building or buildings.

Hosted Organizations can rent access to resources from a specific organization. Note that the line is blurred between hosted and cloud services. In some cases, you know exactly where the services are hosted. However, in most cases, hosted services are somewhere within the cloud.

VDI (Virtual Desktop Infrastructure)

Virtualize the user’s In addition to virtualization servers, it’s also possible to virtualize desktops. In a virtual desktop infrastructure (VDI) or virtual desktop environment (VDE), a user’s desktop operating system runs as a VM on a server.

One benefit of using a VDI/VDE is that user PCs can have limited hardware resources. If the PC can connect to a server over a network, it can run a full-featured desktop operating system from the server.

A primary consideration when running virtual desktops is whether they will support persistence or non-persistence. In a persistent virtual desktop, each user has a custom desktop image. Users can customize them and save their data within the desktop. A drawback is the amount of disk space required on the server to support unique desktop images for all users.

Virtual desktops that support non-persistence serve the same desktop for all users. When a user accesses the remote server, it provides a desktop operating system from a preconfigured snapshot. Although users can make changes to the desktop as they’re using it, it reverts to a known state (the original snapshot) when they log off. Another way of viewing this is that it rolls back to a known configuration.

Cloud Access Security Broker (CASB)

Organizations that use cloud resources often add a cloud access security broker (CASB) for additional security. This is a software tool or service deployed between an organization’s network and the cloud provider. It monitors all network traffic and can enforce security policies. As an example, it can ensure that all data stored in the cloud is encrypted.

Security as Service (SECaaS)

Another entry into cloud computing is Security as a Service. It includes any services provided via the cloud that provide security services, and is commonly viewed as a subset of the Software as a Service (SaaS) model.

A key benefit of Security as a Service is that it outsources the administrative tasks associated with implementing the service. Additionally, professionals are focused on the specific security services offered, eliminating the need for employees to be experts on everything.

Resiliency & Automation

Automation and Scripting

Resiliency and automation strategies include automation, scripting, and templates and they can help deploy systems securely, and keep them in a secure state. As an example, administrators often use Group Policy in Microsoft domains to automatically check and configure systems. It provides automated courses of action by applying security and other settings.

Additionally, Microsoft has created several security templates with various levels of security. Administrators can modify these templates to fit their needs, import them into a Group Policy Object (GPO), and then apply them to systems within the domain. Some organizations deploy a master image to all systems, and then use the security templates to automatically apply different security settings to different groups of systems based on their security needs.

Master Image

One of the most common methods of deploying systems is with images starting with a master image. An image is a snapshot of a single system that administrators deploy to multiple other systems. Imaging has become an important practice for many organizations because it streamlines deployments while also ensuring they are deployed in a secure manner.

Elasticity and Scalability

Elasticity and scalability refer to the ability to resize computing capacity based on the load. For example, imagine one VM has increased traffic. You can increase the amount of processing power and memory used by this server relatively easily. Similarly, it’s relatively easy to decrease the resources when the load decreases.

Redundancy & Fault Tolerance

Distributive Allocation

Distributive allocation is another option to provide both high availability and scalability, though it is typically used primarily in scientific applications. In a distributed application model, multiple computers (often called nodes) are configured to work together to solve complex problems. These computers are configured within a local network. A central processor divides the complex problem into smaller tasks. It then coordinates tasking of the individual nodes and collecting the results. If any single nodes fail, the central processor doesn’t task it anymore, but overall processing continues, providing high availability. This also provides high scalability because it is relatively easy to add additional nodes and task them when they come online.

Cluster and Load Balancer

The primary purpose of a failover cluster is to provide high availability for a service offered by a server. Failover clusters use two or more servers in a cluster configuration, and the servers are referred to as nodes. At least one server or node is active and at least one is inactive. If an active node fails, the inactive node can take over the load without interruption to clients.

A load balancer can optimize and distribute data loads across multiple computers or multiple networks. For example, if an organization hosts a popular web site, it can use multiple servers hosting the same web site in a web farm. Load-balancing software distributes traffic equally among all the servers in the web farm, typically located in a DMZ.

The term load balancer makes it sound like it’s a piece of hardware, but a load balancer can be hardware or software. A hardware-based load balancer accepts traffic and directs it to servers based on factors such as processor utilization and the number of current connections to the server. A software based load balancer uses software running on each of the servers in the load balanced cluster to balance the load.

Load balancing primarily provides scalability, but it also contributes to high availability. Scalability refers to the ability of a service to serve more clients without any decrease in performance. Availability ensures that systems are up and

operational when needed. By spreading the load among multiple systems, it ensures that individual systems are not overloaded, increasing overall availability.

High Availability

High availability refers to a system or service that needs to remain operational with almost zero downtime. Utilizing different redundancy and fault-tolerance methods, it’s possible to achieve 99.999 percent uptime, commonly called five nines.

Physical Security Controls

Physical Security Controls

A physical security control is something you can physically touch, such as a hardware lock, a fence, an identification badge, and a security camera. Physical security access controls attempt to control entry and exits, and organizations commonly implement different controls at different boundaries, such as the following:

Perimeter Military bases and many other organizations erect a fence around the entire perimeter of their land. They often post security guards at gates to control access. In some cases, organizations install barricades to block vehicles.

Buildings Buildings commonly have additional controls for both safety and security. For example, guards and locked doors restrict entry so only authorized personnel enter. Many buildings include lighting and video cameras to monitor the entrances and exits.

Secure work areas Some companies restrict access to specific work areas when employees perform classified or restricted access tasks. In some cases, an organization restricts access to all internal work areas. In other words, visitors can enter the lobby of a building, but they are not able to enter internal work areas without an escort.

Server and network rooms Servers and network devices such as routers and switches are normally stored in areas where only the appropriate IT personnel can access them. These spaces may be designated as server rooms or wiring closets. It’s common for an organization to provide additional physical security for these rooms to prevent attackers from accessing the equipment. For example, locking a wiring closet prevents an attacker from installing illicit monitoring hardware, such as a protocol analyzer, to capture network traffic.

Hardware Additional physical security controls protect individual systems. For example, server rooms often have locking cabinets to protect servers and other equipment installed in the equipment bays. Cable locks protect laptop computers, and smaller devices can be stored in safes.

Airgap An airgap is a physical security control that ensures that a computer or network is physically isolated from another computer or network. As an example, you can isolate a computer from a network by ensuring that it is not connected to any other system in the network. This lack of connectivity provides an airgap. This is often done to separate classified networks from unclassified networks.