



SEZG566/SSZG566

Secure Software Engineering

Intrusion Detection/Prevention

BITS Pilani

T V Rao

Pilani | Dubai | Goa | Hyderabad



- The slides presented here are obtained from the authors of the books, product documentations, and from various other contributors. I hereby acknowledge all the contributors for their material and inputs.
- I have added and modified slides to suit the requirements of the course.

Why Intrusion Detection/Prevention

How much does a data breach cost?

https://www.ibm.com/reports/data-breach



BITS Pilani
Pilani | Dubai | Goa | Hyderabad

Intrusion Detection



Security Intrusion & Detection

RFC 2828 - Internet Security Glossary

- Security intrusion: a security event, or combination of multiple security events, that constitutes a security incident in which an intruder gains, or attempts to gain, access to a system (or system resource) without having authorization to do so.
- Intrusion detection: a security service that monitors and analyzes system events for the purpose of finding, and providing real-time or near real-time warning of attempts to access system resources in an unauthorized manner.



Intrusion Techniques

- Objective to gain access or increase privileges
- Initial attacks often exploit system or software vulnerabilities to execute code to get backdoor
 - e.g. buffer overflow
- Or to gain protected information
 - Password guessing or acquisition (or via social engineering)



Why Intrusion Detection?

- Applications do not detect attacks, but instead try their best to fulfill the attackers' requests
- Lack of intrusion detection allows an attacker to attempt attacks until a successful one is identified
- There are three types of requests that an application might receive:
 - Almost certainly an attack
 - Not sure whether it an attack or not
 - Almost certainly legitimate input
- In terms of the accuracy of an IDS, there are four possible states for each activity observed.
 - A true positive state is when the IDS identifies an activity as an attack and the activity is actually an attack.
 - A true negative state is similar.
 - A false positive state is when the IDS identifies an activity as an attack but the activity is acceptable behavior. A false positive is a false alarm
 - A false negative state is the most serious and dangerous state. This is when the IDS identifies an activity as acceptable when the activity is actually an attack



Intrusion detection systems

- Host-based IDS: monitor single host activity
- Network-based IDS: monitor network traffic
- Distributed or hybrid:
 Combines information from a number of sensors, often both host and network based, in a central analyzer that is able to better identify and respond to intrusion activity

Comprises three logical components:

- Sensors: collect data
- Analyzers: determine if intrusion has occurred
- User interface: view output or control system behavior

Base Rate Fallacy

• Bayes theorem:

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)}$$

More generally:

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{\sum_{i=1}^{n} P(A_i) \cdot P(B|A_i)}$$



Base Rate Fallacy (Axelsson, 1999)

The base-rate fallacy is best described through example:

Suppose a doctor performs on a patient a diagnostic test that is 99% accurate symmetrically (i.e. both for presence and absence of disease).

The doctor may report the test result as bad news + good news

- Bad News the patient is tested positive
- Good News Out of entire population, rate of incidence is only 1 in 10000.

Now what is the probability that the patient has the disease?

Base Rate Fallacy

$$P(S|P) = \frac{P(S) \cdot P(P|S)}{P(S) \cdot P(P|S) + P(\neg S) \cdot P(P|\neg S)}$$

$$P(S|P) = \frac{1/10000 \cdot 0.99}{1/10000 \cdot 0.99 + (1 - 1/10000) \cdot 0.01} = 0.00980 \dots \approx 1\%$$

• Even though the test is 99% certain, your chance of having the disease is 1/100, because the population of healthy people is much larger than sick people

Base Rate Fallacy in Intrusion Detection

- I: intrusive behavior,
- ¬I: non-intrusive behavior
- A: alarm
- \neg A: no alarm

Detection rate (true positive rate): P(A|I)

False alarm rate: $P(A | \neg I)$

Goal is to maximize both

- Bayesian detection rate, P(I|A)
- P(¬I | ¬A)

Detection Rate vs False Alarm Rate

$$P(I|A) = \frac{P(I) \cdot P(A|I)}{P(I) \cdot P(A|I) + P(\neg I) \cdot P(A|\neg I)}$$

Suppose there are about 20 intrusions Per million $P(I)=1\left/\frac{1\cdot 10^6}{2\cdot 10}\right.=2\cdot 10^{-5};$ $P(\neg I)=1-P(I)=0.99998$

$$P(I|A) = \frac{2 \cdot 10^{-5} \cdot P(A|I)}{2 \cdot 10^{-5} \cdot P(A|I) + 0.99998 \cdot P(A|\neg I)}$$

• False alarm rate becomes more dominant if P(I) is very low

Detection Rate vs False Alarm Rate

$$P(I|A) = \frac{P(I) \cdot P(A|I)}{P(I) \cdot P(A|I) + P(\neg I) \cdot P(A|\neg I)}$$

If there are about 20 intrusions Per million

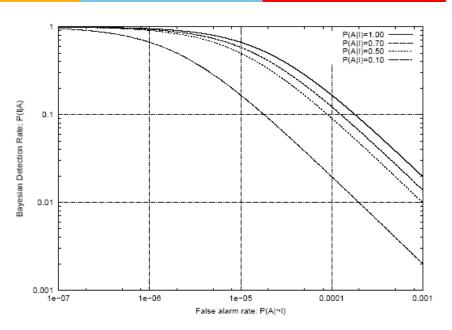
$$P(I) = 1 / \frac{1 \cdot 10^6}{2 \cdot 10} = 2 \cdot 10^{-5};$$

 $P(\neg I) = 1 - P(I) = 0.99998$

$$P(I|A) = \frac{2 \cdot 10^{-5} \cdot P(A|I)}{2 \cdot 10^{-5} \cdot P(A|I) + 0.99998 \cdot P(A|\neg I)}$$

Say, we got a system with 99% accuracy, i.e. P(A|I)=0.99 and $P(A|\neg I)=0.01$, then $P(I|A) \sim 0.002$, i.e. there are more than 99% false alarms

Detection Rate vs False Alarm Rate



 Axelsson: We need an extremely low false alarm rate to achieve a reasonable Bayesian detection rate

Intrusion Detection Evaluation Metrics: Confusion Matrix



Confusion Matrix:

Given m classes, an entry, $CM_{i,j}$ in a confusion matrix indicates # of tuples in class i that were labeled by the classifier as class j

May have extra rows/columns to provide totals

Predicted class ->	C ₁	¬ C ₁
Actual class↓		
C ₁	True Positives (TP)	False Negatives (FN)
¬ C ₁	False Positives (FP)	True Negatives (TN)



Intruder Behavior

Target acquisition and information gathering

Initial access

Privilege escalation

Information gathering or system exploit

Maintaining access

Covering tracks

Also see Adversary tactics & techniques at https://attack.mitre.org/

(a) Target Acquisition and Information Gathering

- Explore corporate website for information on corporate structure, personnel, key systems, as well as details of specific web server and OS used. Gather information on target network using DNS lookup tools such as dig, host, and
- others; and query WHOIS database.
- Map network for accessible services using tools such as NMAP.
- Send query email to customer service contact, review response for information on mail client, server, and OS used, and also details of person responding.
 - Identify potentially vulnerable services, eg vulnerable web CMS.

(b) Initial Access

- Brute force (guess) a user's web content management system (CMS) password.
- Exploit vulnerability in web CMS plugin to gain system access.
 - Send spear-phishing email with link to web browser exploit to key people.

(c) Privilege Escalation

- Scan system for applications with local exploit.
- Exploit any vulnerable application to gain elevated privileges.
- Install sniffers to capture administrator passwords.
- Use captured administrator password to access privileged information.

(d) Information Gathering or System Exploit

- Scan files for desired information.
- Transfer large numbers of documents to external repository.
- Use guessed or captured passwords to access other servers on network.

(e) Maintaining Access

- Install remote administration tool or rootkit with backdoor for later access.
- Use administrator password to later access network.
- Modify or disable anti-virus or IDS programs running on system.

(f) Covering Tracks

- Use rootkit to hide files installed on system.
- Edit logfiles to remove entries generated during the intrusion.

Table 8.1

Examples of Intruder Behavior

(Table can be found on pages 255-256 in the textbook.)



IDS requirements

monitor itself from Run continually with minimal recover from crashes human supervision change by intruder Run continually Be fault tolerant Resist subversion Impose a Configured Adapt to minimal according to changes in overhead on system security systems and system policies users Scale to monitor Provide graceful Allow dynamic large numbers degradation of reconfiguration of systems service

if one component fails, others should continue to work



What are the current Threats?

IBM X-Force Threat Intelligence Index 2022

https://www.ibm.com/reports/threat-intelligence/



BITS Pilani
Pilani | Dubai | Goa | Hyderabad

Detection techniques



Detection techniques

Anomaly detection

Involves the collection of data relating to the behavior of legitimate users over a period of time

Current observed behavior is analyzed to determine whether this behavior is that of a legitimate user or that of an intruder

Signature/Heuristic detection

Uses a set of known malicious data patterns or attack rules that are compared with current behavior

Also known as misuse detection Can only identify known attacks for which it has patterns or rules



Anomaly Detection

A variety of classification approaches are used:

Statistical

 Analysis of the observed behavior using univariate, multivariate, or time-series models of observed metrics

Knowledge based

 Approaches use an expert system that classifies observed behavior according to a set of rules that model legitimate behavior

Machine-learning

Approaches
 automatically
 determine a
 suitable
 classification
 model from the
 training data
 using data mining
 techniques



Example of metrics

- Counters: e.g., number of logins during an hour, number of times a cmd executed
- Gauge: e.g., the number of outgoing messages [pkts]
- Interval time: the length of time between two events, e.g., two successive logins
- Resource utilization: quantity of resources used (e.g., number of pages printed)
- Mean and standard deviations



Signature/heuristic detection

Signature approaches



Match a large collection of known patterns of malicious data against data stored on a system or in transit over a network



The signatures need to be large enough to minimize the false alarm rate, while still detecting a sufficiently large fraction of malicious data



Widely used in anti-virus products, network traffic scanning proxies, and in NIDS

Rule-based heuristic identification



Involves the use of rules for identifying known penetrations or penetrations that would exploit known weaknesses



Rules can also be defined that identify suspicious behavior, even when the behavior is within the bounds of established patterns of usage



Typically rules used are specific



SNORT is an example of a rule-based NIDS



Example rules in a signature detection IDS

- Users should not be logged in more than one session
- Users do not make copies of system, password files
- Users should not read in other users' directories
- Users must not write other users' files
- Users who log after hours often access the same files they used earlier
- Users do not generally open disk devices but rely on high-level OS utilities



BITS Pilani

Host-based IDS



Host-based IDS

- Specialized software to monitor system activity to detect suspicious behavior
 - primary purpose is to detect intrusions, log suspicious events, and send alerts
 - can detect both external and internal intrusions
 - used on sensitive systems e.g. database servers, administrative systems
- Two approaches, often used in combination:
 - Anomaly detection: consider normal/expected behavior over a period of time; apply statistical tests to detect intruder
 - threshold detection: for various events (#/volume of copying)
 - profile based (time/duration of login)
 - Signature detection: defines proper (or bad) behavior (rules)



Common data sources

- Fundamental component of intrusion detection is the sensor that collects data. Common data sources include
 - System call traces
 - Audit (log file) records
 - File integrity checksums
 - Registry access

Linux System Calls and Windows DLLs Monitored

Ubuntu Linux System Calls

accept, access, acct, adjtime, aiocancel, aioread, aiowait, aiowrite, alarm, async daemon, auditsys, bind, chdir, chmod, chown, chroot, close, connect, creat, dup, dup2, execv, execve, exit, exportfs, fchdir, fchmod, fchown, fchroot, fcntl, flock, fork, fpathconf, fstat, fstat, fstatfs, fsvnc, ftime, ftruncate, getdents, getdirentries, getdomainname, getdopt, getdtablesize, getfh, getgid, getgroups, gethostid, gethostname, getitimer, getmsg, getpagesize, getpeername, getpgrp, getpid, getpriority, getrlimit, getrusage, getsockname, getsockopt, gettimeofday, getuid, gtty, ioctl, kill, killpg, link, listen, lseek, lstat, madvise, mctl, mincore, mkdir, mknod, mmap, mount, mount, mprotect, mpxchan, msgsys, msync, munmap, nfs_mount, nfssvc, nice, open, pathconf, pause, pcfs_mount, phys, pipe, poll, profil, ptrace, putmsg, quota, quotactl, read, readlink, readv, reboot, recv, recvfrom, recvmsg, rename, resuba, rfssys, rmdir, sbreak, sbrk, select, semsys, send, sendmsg, sendto, setdomainname, setdopt, setgid, setgroups, sethostid, sethostname, setitimer, setpgid, setpgrp, setpgrp, setpriority, setquota, setregid, setreuid, setrlimit, setsid, setsockopt, settimeofday, setuid, shmsys, shutdown, sigblock, sigpause, sigpending, sigsetmask, sigstack, sigsys, sigvec, socket, socketaddr, socketpair, sstk, stat, statfs, stime, stty, swapon, symlink, sync, sysconf, time, times, truncate, umask, umount, uname, unlink, unmount, ustat, utime, utimes, vadvise, vfork, vhangup, vlimit, vpixsys, vread, vtimes, vtrace, vwrite, wait, wait3, wait4, write, writev

Key Windows DLLs and Executables

comctl32 kernel32 msvcpp msvcrt mswsock ntdll ntoskrnl user32 ws2 32

innovate achieve lead

Distributed host-based IDS

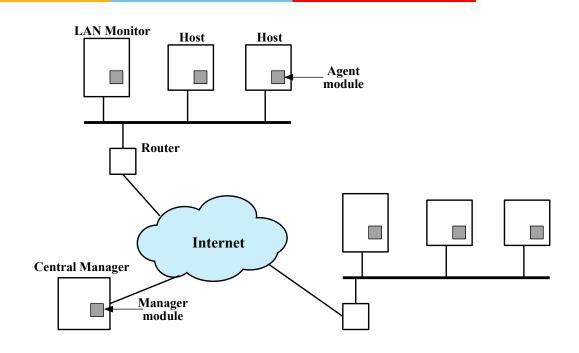


Figure 8.2 Architecture for Distributed Intrusion Detection

Agent Architecture



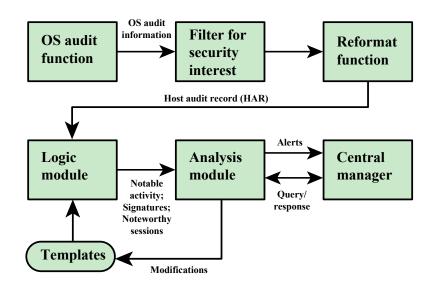


Figure 8.3 Agent Architecture



BITS Pilani
Pilani | Dubai | Goa | Hyderabad

Network-Based IDS



Network-Based IDS

Monitors traffic at selected points on a network

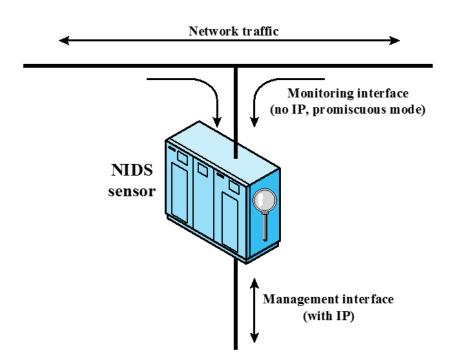
Examines traffic packet by packet in real or close to real time

May examine network, transport, and/or applicationlevel protocol activity

Comprised of a number of sensors, one or more servers for NIDS management functions, and one or more management consoles for the human interface

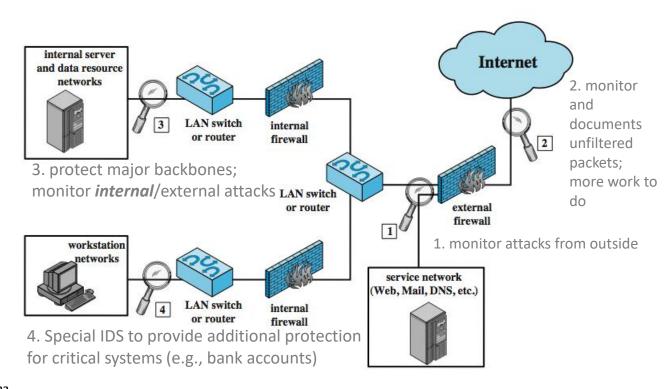
Analysis of traffic patterns may be done at the sensor, the management server or a combination of the two

Passive sensors





NIDS Sensor Deployment





NIDS intrusion detection techniques

- Signature detection
 - at application (FTP), transport (port scans), network layers (ICMP); unexpected application services (host running unexpected app), policy violations (inappropriate website use)
- Anomaly detection
 - of denial of service attacks, scanning, worms (significant traffic increase)
- When potential violation detected, sensor sends an alert and logs information
 - Used by analysis module to refine intrusion detection parameters and algorithms
 - by security admin to improve protection

Distributed hybrid intrusion detection

nnovate achieve lead

(host-based, NIDS, distributed host-based)

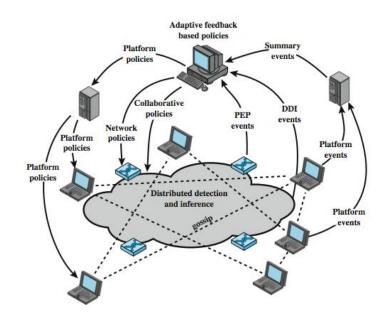
Issues:

- 1. Tools may not recognize new threats
- 2. Difficult to deal with rapidly spreading attacks

Solution:

Distributed Adaptive IDS thru Peer-to-peer gossip and cooperation

One developed by Intel, called Autonomic Enterprise Security



PEP = policy enforcement point
DDI = distributed detection and inference



Logging of alerts (for all types)

- Typical information logged by a NIDS sensor includes:
 - Timestamp
 - Connection or session ID
 - Event or alert type
 - Rating
 - Network, transport, and application layer protocols
 - Source and destination IP addresses
 - Source and destination TCP or UDP ports, or ICMP types and codes
 - Number of bytes transmitted over the connection
 - Decoded payload data, such as application requests and responses
 - State-related information



Intrusion detection exchange format

A data format used to exchange information between software enabling intrusion detection, intrusion prevention, security information collection and management systems that may need to interact with them along with personnel. Includes XML DTD

To facilitate development of a distributed IDS

Not a product, but a proposed IETF standard RFC 4765

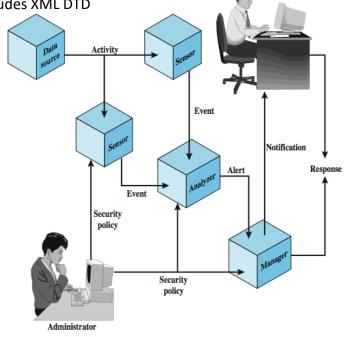
Key elements

Data source: raw data from an IDS **Sensor**: collect and forward events

Analyzer: process data

Administrator defines sec policy **Manager**: a process for operator to manage the IDS system

Operator: the user of the Manager







Honeypots



Honeypots

- Decoy systems
 - Filled with fabricated info and instrumented with monitors/event loggers
 - Lure a potential attacker away from critical systems
 - Collect information about the attacker's activity
 - Encourage the attacker to stay on the system long enough for administrators to respond
 - Divert and hold attacker to collect activity info without exposing production systems
- Initially were single systems
- More recently are/emulate entire networks



Honeypot classification

- Low interaction honeypot
 - Consists of a software package that emulates particular IT services or systems well enough to provide a realistic initial interaction, but does not execute a full version of those services or systems
 - Provides a less realistic target
 - Often sufficient for use as a component of a distributed IDS to warn of imminent attack
- High interaction honeypot
 - A real system, with a full operating system, services and applications, which are instrumented and deployed where they can be accessed by attackers

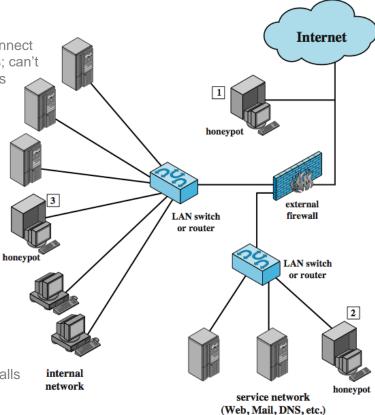
Honeypot deployment

innovate achieve lead

1. Tracks attempts to connect to an unused IP address; can't help with inside attackers

3. Full internal honeypot; can detect internal attacks

2. In DMZ; must make sure the other systems in the DMZ are secure; firewalls may block traffic to the honeypot





BITS Pilani

Pilani | Dubai | Goa | Hyderabad

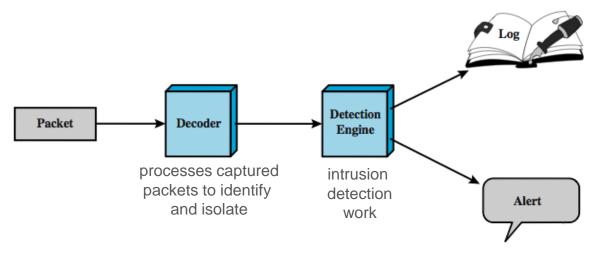
Snort IDS



Snort IDS

Lightweight IDS

- Open source (rule-based)
- Real-time packet capture and rule analysis
- Passive or inline
- Components: decoder, detector, logger, alerter





BITS Pilani
Pilani | Dubai | Goa | Hyderabad

Firewalls and Intrusion Prevention Systems



Firewalls and Intrusion Prevention Systems

- Effective means of protecting LANs
- Internet connectivity essential
 - For organization and individuals
 - But creates a threat
- Could secure workstations and servers
- Also use firewall as perimeter defence
 - Single choke point to impose security



Firewall Access Policy

- A critical component in the planning and implementation of a firewall is specifying a suitable access policy
 - Types of traffic authorized to pass through the firewall
 - Includes address ranges, protocols, applications and content types
- The policy should be developed from the organization's security risk assessment and policy
- Should be developed from a broad specification of which traffic types the organization needs to support
 - Then refined to detail the filter elements which can then be implemented within an appropriate firewall topology



Firewall Capabilities & Limits

Capabilities

- Defines a single choke point
- Provides a location for monitoring security events
- Convenient platform for some Internet functions such as NAT, usage monitoring, IPSEC, VPNs

Limitations

- Cannot protect against attacks bypassing firewall (internal systems can dial-out to an ISP)
- May not protect fully against internal threats
- Improperly secured wireless LAN
- Laptop, PDA, portable storage device infected outside then used inside



Firewall Filter Characteristics

Characteristics that a firewall access policy could use to filter traffic include:

IP address and protocol values

This type of filtering is used by packet filter and stateful inspection firewalls

Typically used to limit access to specific services

Application protocol

This type of filtering is used by an application-level gateway that relays and monitors the exchange of information for specific application protocols

User identity

Typically for inside users who identify themselves using some form of secure authentication technology

Network activity

Controls access
based on
considerations
such as the time or
request, rate of
requests, or other
activity patterns



Packet Filtering Firewall

- Applies rules to each incoming and outgoing IP packet
 - Typically a list of rules based on matches in the IP or TCP header
 - Forwards or discards the packet based on rules match

Filtering rules are based on information contained in a network packet

- Source IP address
- Destination IP address
- Source and destination transport-level address
- IP protocol field
- Interface
- Two default policies:
 - Discard prohibit unless expressly permitted
 - More conservative, controlled, visible to users
 - Forward permit unless expressly prohibited
 - Easier to manage and use but less secure



Packet-Filtering Examples

Rule	Direction	Src address	Dest addresss	Protocol	Dest port	Action
1	In	External	Internal	TCP	25	Permit
2	Out	Internal	External	TCP	>1023	Permit
3	Out	Internal	External	TCP	25	Permit
4	In	External	Internal	ТСР	>1023	Permit
5	Either	Any	Any	Any	Any	Deny



Stateful Inspection Firewall

Tightens rules for TCP traffic by creating a directory of outbound TCP connections

- There is an entry for each currently established connection
- Packet filter allows incoming traffic to high numbered ports only for those packets that fit the profile of one of the entries in this directory

Reviews packet information but also records information about TCP connections

- Keeps track of TCP sequence numbers to prevent attacks that depend on the sequence number
- Inspects data for protocols like FTP, IM and SIPS commands



Application-Level (Proxy) Gateway

- Acts as a relay of application-level traffic
 - User contacts gateway with remote host name
 - Authenticates themselves
 - Gateway contacts application on remote host and relays TCP segments between server and user
- Must have proxy code for each application
 - May restrict application features supported
 - Some services may not be available
- More secure than packet filters
- But have higher overheads

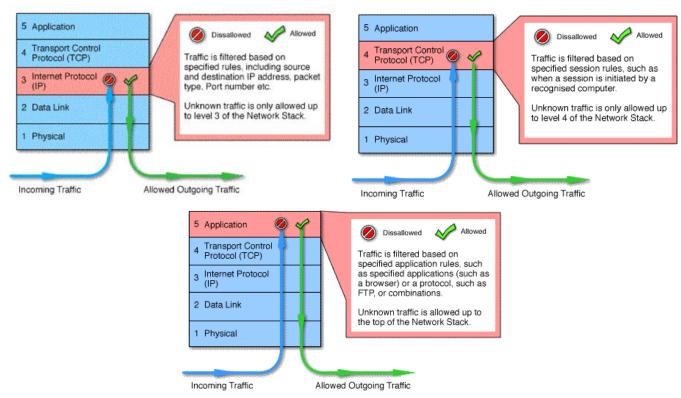


Circuit-Level Gateway

- Sets up two TCP connections, to an inside user and to an outside host
- Once connection is established, relays TCP segments from one connection to the other without examining contents
 - Hence independent of application logic
 - Just determines whether relay is permitted
- Typically used when inside users trusted
 - May use application-level gateway inbound and circuit-level gateway outbound
 - Hence lower overheads



Packet Filtering vs Gateway vs Application-Level Firewall





Firewall Basing

- Several options for locating firewall:
 - -Bastion host
 - —Individual host-based firewall
 - —Personal firewall



Bastion Hosts

- Critical strongpoint in network
- Hosts application/circuit-level gateways
- Common characteristics:
 - Runs secure O/S, only essential services
 - May require user auth to access proxy or host
 - There may be many proxy services
 - Each proxy can restrict features, hosts accessed
 - Each proxy small, simple, checked for security
 - Each proxy is independent, can be uninstalled



Host-Based Firewalls

- Used to secure individual host
- Available in/add-on for many O/S
- Filter packet flows
- Often used on servers
- Advantages:
 - Tailored filter rules for specific host needs
 - Protection from both internal/external attacks
 - Additional layer of protection to org firewall when used with a standalone firewall



Personal Firewall

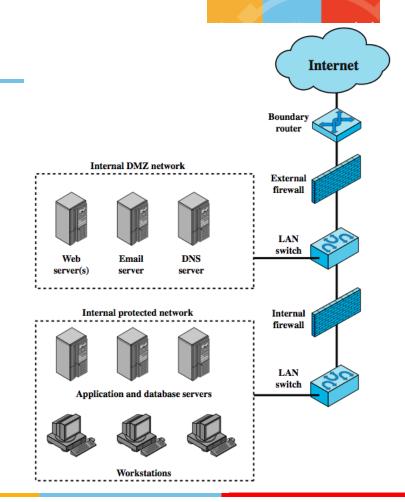
- Controls traffic flow to/from PC/workstation
- For both home or corporate use
- May be software module on PC
- Or in home cable/DSL router/gateway
- Typically much less complex
- Primary role to deny unauthorized access
- May also monitor outgoing traffic to detect/block worm/malware activity

Firewall Locations

External firewall: protection for the DMZ consistent with their need for external connectivity

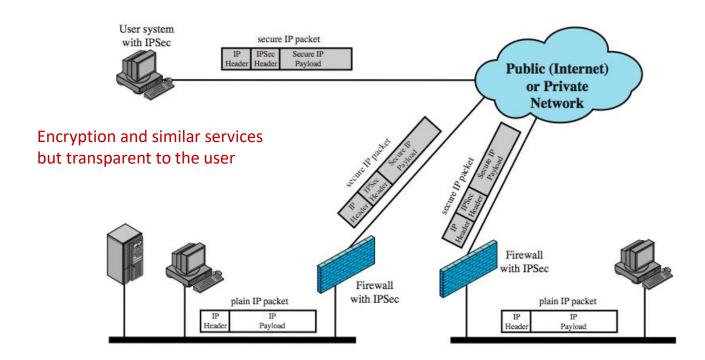
Internal firewall:

- (a) more stringent filtering capability to provide protection from external attacks
- (b) provides two way protection wrt the DMZ network





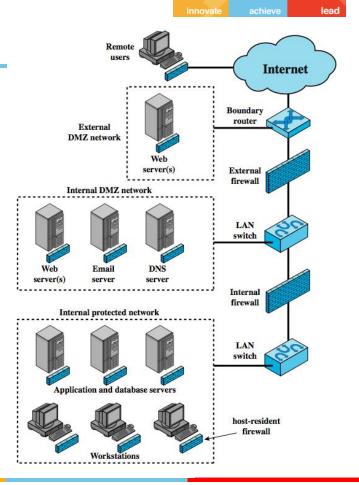
Virtual Private Networks



Distributed Firewalls

A combination of earlier firewalls

Host-resident firewall on 100s of PCs plus standalone firewalls under a central administration





Firewall vs IPS

- A firewall will block traffic based on network information such as IP address, network port and network protocol.
 - It can make some decisions based on the state of the network connectiony
- An IPS will inspect content of the request and be able to drop, alert, or potentially clean a malicious network request based on that content.
 - The determination of what is malicious is based either on behavior analysis or through the use of signatures



Intrusion Prevention Systems (IPS)

- Also known as Intrusion Detection and Prevention System (IDPS)
- Is an extension of an IDS that includes the capability to attempt to block or prevent detected malicious activity
- Can be host-based, network-based, or distributed/hybrid
- Can use anomaly detection to identify behavior that is not that of legitimate users, or signature/heuristic detection to identify known malicious behavior can block traffic as a firewall does, but makes use of the types of algorithms developed for IDSs to determine when to do so



Host-Based IPS

- Identifies attacks using both:
 - Signature techniques
 - malicious application packets
 - Anomaly detection techniques
 - · behavior patterns that indicate malware
 - Example of malicious behavior: buffer overflow, access to email contacts, directory traversal
- Can be tailored to the specific platform
 - e.g. general purpose, web/database server specific
- Can also sandbox applets to monitor behavior
- May give desktop file, registry, I/O protection



Network-Based IPS

- inline NIDS that can discard packets or terminate TCP connections
- uses signature and anomaly detection
- may provide flow data protection
 - monitoring full application flow content
- can identify malicious packets using:
 - pattern matching (for specific byte seq)
 - stateful matching (to stop attack streams rather than a single pkts)
 - protocol anomaly (deviations from stds)
 - traffic anomaly (unusual traffic like a UDP floods

Unified Threat Management Products

Reduce admin burden by replacing network products (firewall, IDS, IPS, ...) With a single device

Cisco ASA Content Security and Control (CSC) Security Services Module providing comprehensive

antivirus,

anti-spyware,

file blocking,

anti-spam,

anti-phishing,

URL blocking and filtering, and

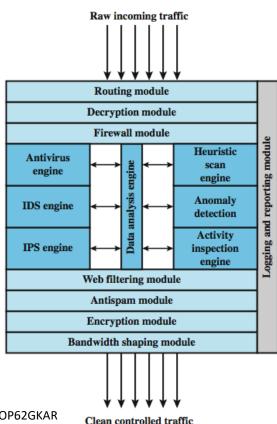
content filtering

-all available in a comprehensive easy-to-manage

solution

Also, IBM Qradar documentation at https://www.ibm.com/downloads/cas/OP62GKAR





Computer Security: Principles and Practice by William Stallings, and Lawrie Brown Pearson, 2020.

www.cigital.com

sei.cmu.edu/cert

www.owasp.com

Thank You!