CS370 Operating Systems

Colorado State University

Yashwant K Malaiya

Spring 2022 L26

Cloud Computing, Reliability, Security

Slides based on

- **Text by Silberschatz, Galvin, Gagne**
- **Various sources**

Project Presentations

- Final reports due 4/28. Please check requirements.
- All slides need to be posted on Teams channel *Project Slides and Videos* by 4/29 using the specified format.
- Research projects: also post videos there by 4/29.
- Development Demos: sign-up for a 15min slot for coming Mon-Wed. Sheet will be available soon. Demos will be recorded.
- Peer reviews will be needed.

Course Survey

- CSU Course Survey form is available on Canvas.
- Online students: special feedback survey quiz available.

Why do we need to study …..

- Why do we need to study advanced topics like HDFS, Virtualization, Containers etc.? Because …
- Why did we need to study several simple schemes with toy examples for:
  - CPU Scheduling (FCFS etc.)
  - Synchronization, deadlocks
  - Page replacement algorithms
  - Disk space allocation
- So that you can understand real and advanced schemes.

FAQ

Kernel vs OS: OS = {Kernel, UI, libraries/binaries}

- Include: header files, lib: runtime libraries, bin: executable programs.

Hypervisor Type 1 vs kernel:

- In a simple system, the kernel provides access to hardware resources (memory, secondary storage, ports/network)
- With Hypervisor Type 1, the kernel thinks it is providing access to the hardware resources. In reality, the hypervisor is managing the illusion.
- With Hypervisor Type 2, the guest kernel the kernel thinks it is providing access to the hardware resources. In reality, the hypervisor is managing the illusion, by passing the responsibility to the host kernel/

FAQ

HDFS

- Hadoop Distributed File System: Large storage, large files, distributed and replicated storage
- 64MB blocks saved as 4KB blocks of underlying FS (ext4 etc), likely on different nodes, allowing access in parallel.
- Why 3 copies? 1 replica: MTDL ≈ 1 year, 2 replicas ≈ 10 years, 3 replicas ≈ 100 years assuming independent failures

- Final: comprehensive but questions will mostly be from the second half.
- Questions of various types
- Sec 001 (& local 801)
  - Wed May 11, 2022, 6:20-8:20pm
  - SDC: schedule with them for May 11 (4-8 PM window)
- Sec 801 (non-local)
  - 2 hours,
  - window Wed May 11, 6:20PM-Thurs 8:20pm.

Updates

- Project:
  - Final report: 04/28/2022, 2-columns, citations
    - Slides/Videos also needed on Project Slides channel
    - Peer reviews needed
  - Specific requirements for Option A and Option B
  - Option B: 15 minutes demos each team
    - Sign-up sheet available on MS Team for Mon, Tu, Wed (May 4-6)
    - Demos using MS Team recorded
- Project day: next Tuesday
- Course review: next Thursday

Course Survey

- Will be available on Canvas.

CS370 Students: Future dreams and nightmares

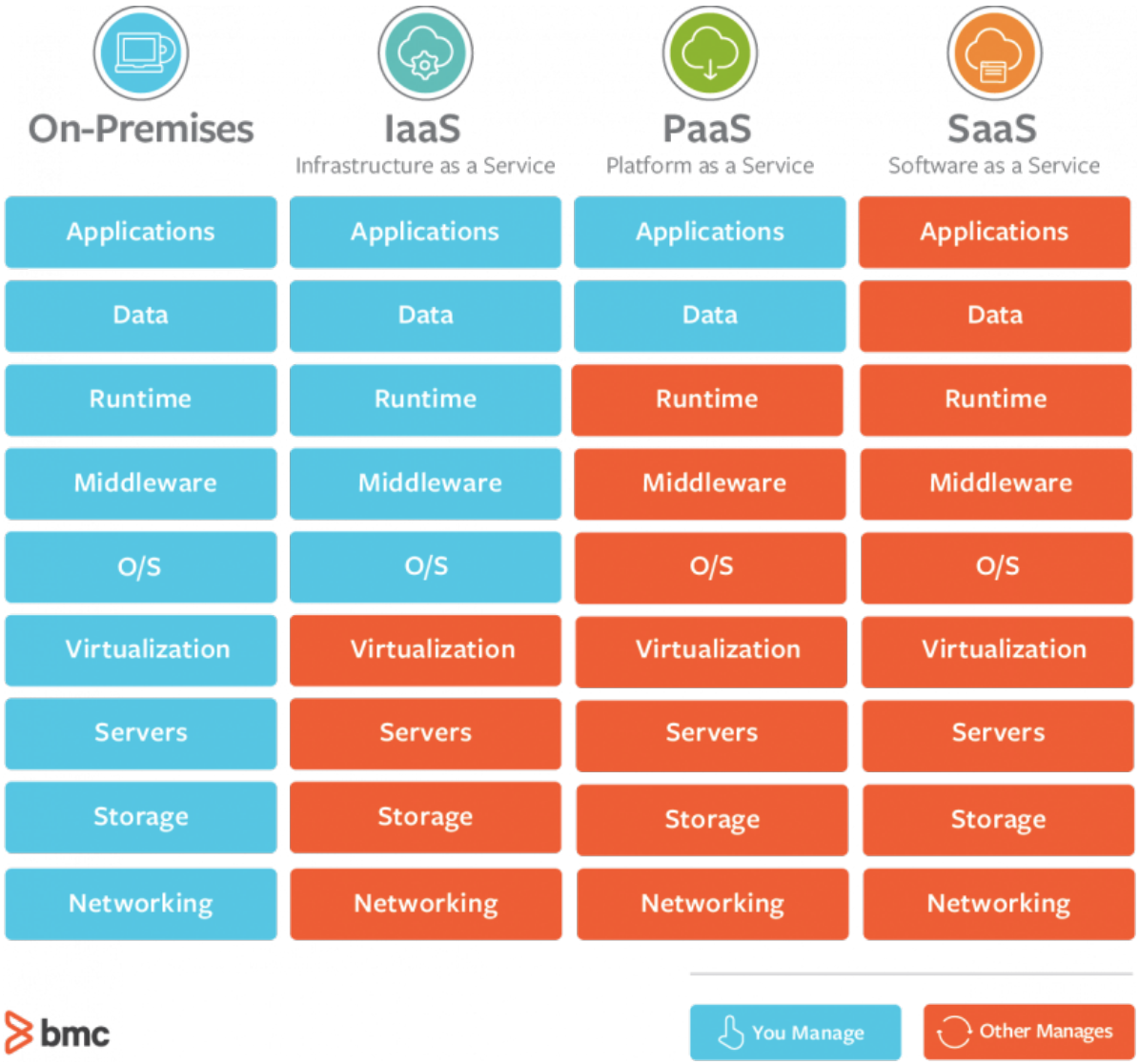
- Man on Mars
- Other dreams
  - AI with human rights
  - More powerful phones
  - natural language models being used to preserve some of the 50%-90% of human languages that will otherwise disappear by the year 2100
  - I expect to see hoverboards in the next 20 years
  - Vaccination created dynamically from AI
  - Flying cars
- Nightmares
  - Cloud gaming virtually everywhere
  - working class be taken over by robots
  - resource wars in my lifetime caused by human greed and the inability to adapt to growing problems as a species
  - future will have less progression than the past decades
  - implants of small operating systems in human brains within the next 20-40 years. (2)
  - anything that does need human intellect to operate will be done by a computer

The cloud Service Models

Service models

- **IaaS: Infrastructure as a Service**
  - infrastructure components traditionally present in an on-premises data center, including servers, storage and networking hardware
  - e.g., Amazon EC2, Microsoft Azure, Google Compute Engine
- **PaaS: Platform as a Service**
  - supplies an environment on which users can install applications and data sets
  - e.g., Google AppEngine, Heroku, Apache Stratos
- **SaaS: Software as a Service**
  - a software distribution model with provider hosted applications
  - Microsoft Office365, Amazon DynamoDB, Gmail

The Service Models



https://www.bmc.com/blogs/saas-vs-paas-vs-iaas-whats-the-difference-and-how-to-choose/

bmc

Colorado State University

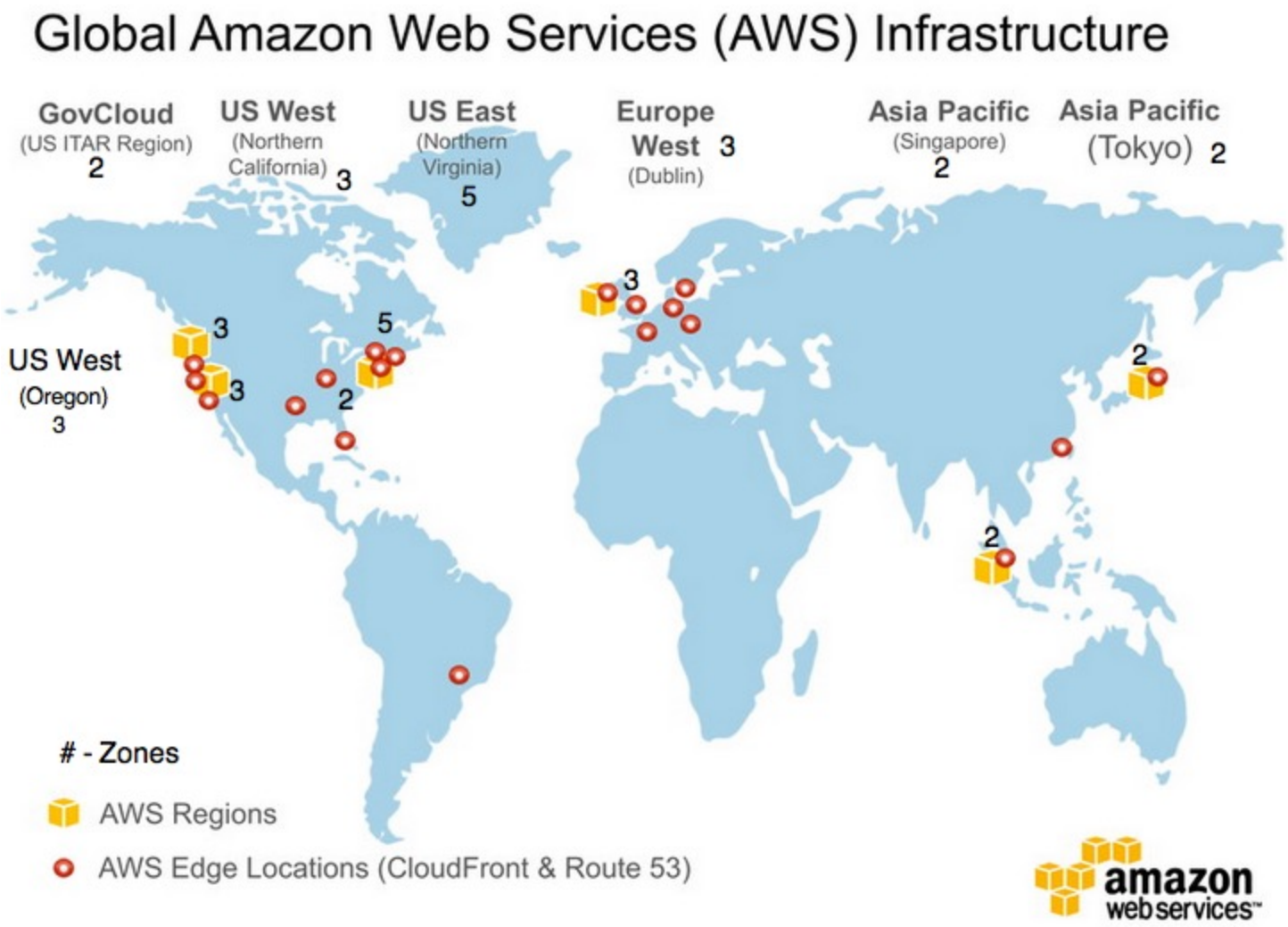
Cloud Management models

- **Public clouds**
  - Utility model
  - Shared hardware, no control of hardware,
  - Self-managed (e.g., AWS, Azure)
- **Private clouds**:
  - More isolated (secure?)
  - Federal compliance friendly
  - Customizable hardware and hardware sharing
- **Hybrid clouds**:
  - a mix of on-premises, private cloud and third-party, public cloud services.
  - Allows workloads to move between private and public clouds as computing needs and costs change.

Different Regions to Achieve HA

- AWS datacenters is divided into regions and zones,
  - that aid in achieving availability and disaster recovery capability.
- Provide option to create point-in-time snapshots to back up and restore data to achieve DR capabilities.
- The snapshot copy feature allows you to copy data to a different AWS region.
  - This is very helpful if your current region is unreachable or there is a need to create an instance in another region
  - You can then make your application highly available by setting the failover to another region.

Different Regions to Achieve HA



Colorado State University

Reliability -Parity Bit

Parity bit for error detection

- Bits to be flipped during transmission or storage.
- Single bit error can be detected by adding a redundant parity bit.
- Even parity:
  - The number of 1-bit must add up to an even number
- Example 1:
  - Information bits:1000000
  - Parity bit: 1
  - String sent 10000001: 2 bits are "1"
  - If received string is
    - 11000001 – **Incorrect parity detected**
    - 10000000 – **Incorrect parity detected**
    - 10000001 – **Correct parity**
- Example 2:
  - Information bits:1110010
  - Parity bit: 0
  - String sent 11100010: 2 bits are "1"
- Odd parity:
  - The number of 1-bit must add up to an odd number

Parity bit multiple errors?

- Example 3:
  - Information bits:1000000
  - Parity bit: 1
  - String sent 10000001: 2 bits are "1"
  - If received string is

    11001001 – **Correct parity!**

  Error detection capability is limited.
- If Pr{1 bit error}=0.01,
  - Pr{2 errors} = 0.01x0.01 = 0.0001 if if errors are statistically independent
  - Pr(2 errors} << Pr{1 bit error}
  - Single errors are much more likely

Coding theory

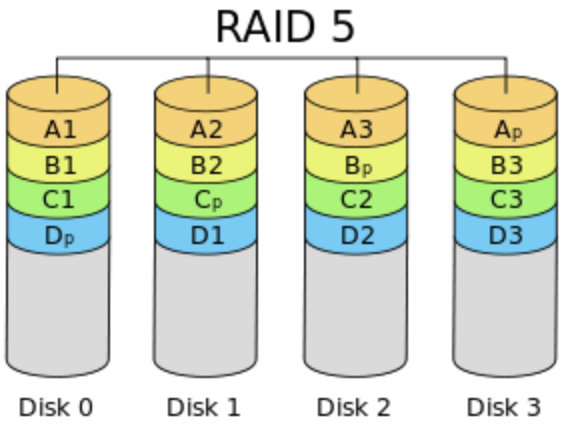
- Information bits + check bits = code word
- Several types of codes are possible.
  - Parity bit: single error detection
  - Hamming codes: single error correction capability
  - CRC (Cyclic redundancy codes): Burst-error detection capability
    - Used in magnetic storage
    - Transmission of information packets on the internet

Disks and Redundancy

- Each block on a disk is protected by a CRC, allowing any corruption to be detected.
- A few bits in a block may get flipped. A block may be restored using redundant information.
- A disk may go bad. A replacement disk can be created using redundant information.
- RAID 1: mirroring. A block/disk may be restored using its mirror.
- RAID 5, 6: A block/disk may be restored using redundant information.

RAID 5: Reconstruction of bad block

- If one disk fails, its data can be reconstructed using a spare



Parity block = Block1 ⊕ block2 ⊕ block3
10001101 block1
01101100 block2
11000110 block3

00100111 parity block (ensures even number of 1s)

- Can reconstruct any missing block from the others

Internet

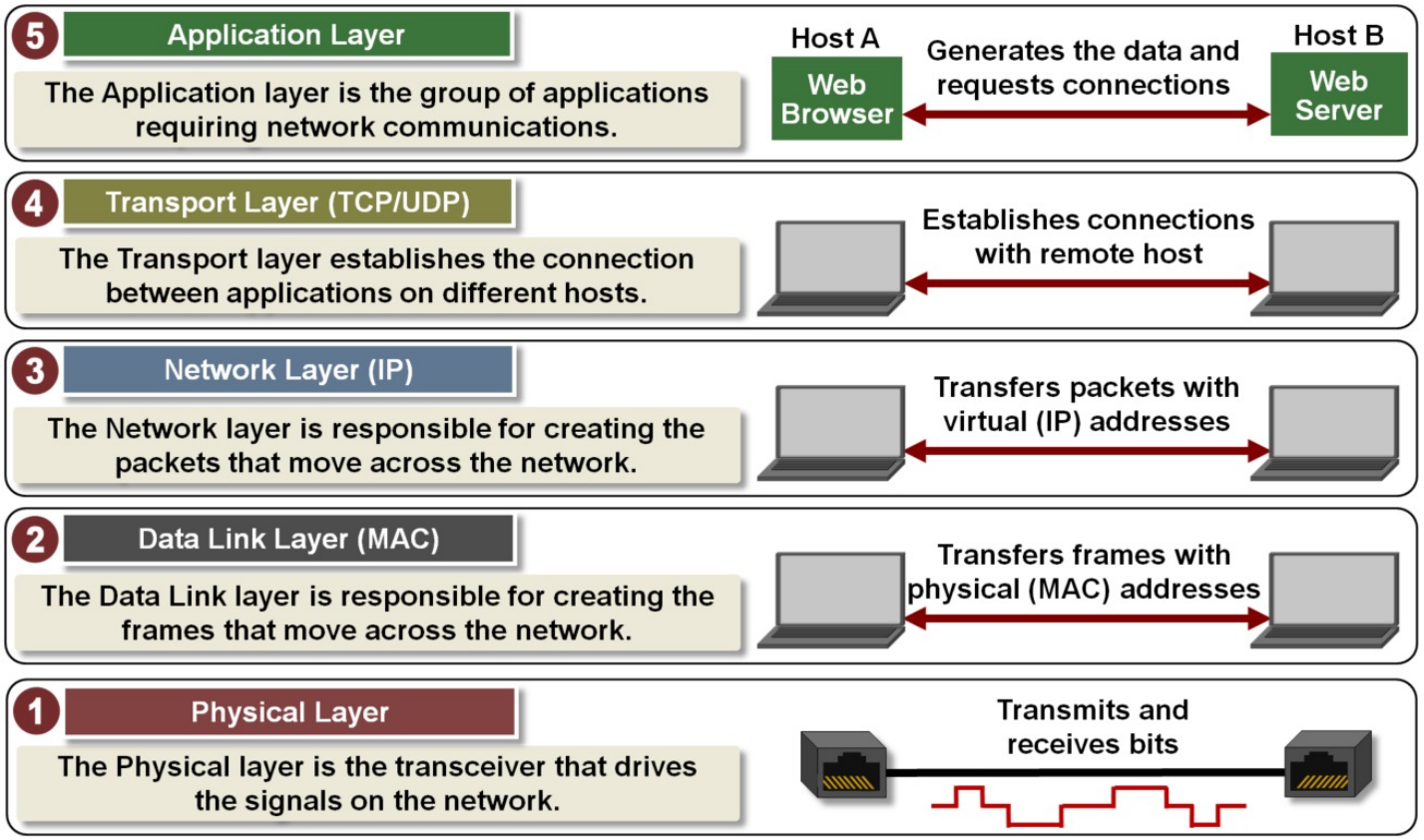
- IP: internet protocol: node to node transfer
- All notes have a unique IP address
  - IPv4: 32 bits. Example 198.51.100.222
  - IPv6: 128 bits
- Packets are routed from source IP to destination IP through routers.
  - Information is packaged into packets including IP addresses, sequence numbers and CRC.
  - Receiving node acknowledges receiving good packets. Those corrupted or lost are transmitted again.
  - TCP: end to end transfer
- In the machine at an IP address, there are a number of ports. Port number : 16 bits, identifies associated application/service.
- The port numbers 0 -1023 reserved for well known services (e.g. SSH 22, HTTP 80 etc). Higher numbers are not reserved.

Security System Architecture

- Networked systems
  - Use of firewalls
- Single computing System: OS
  - Isolation of processes, cgroups, virtual machines

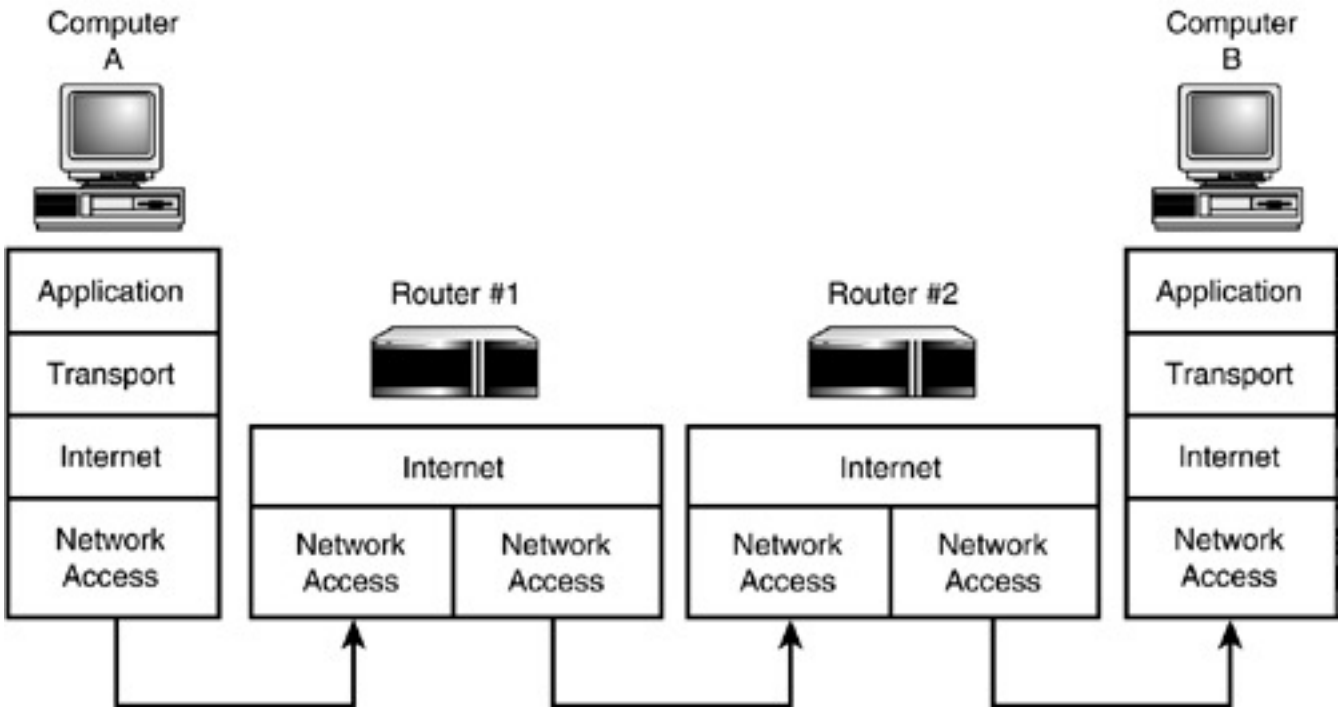
Internet: IP, TCP

- Networking protocols have multiple layers.



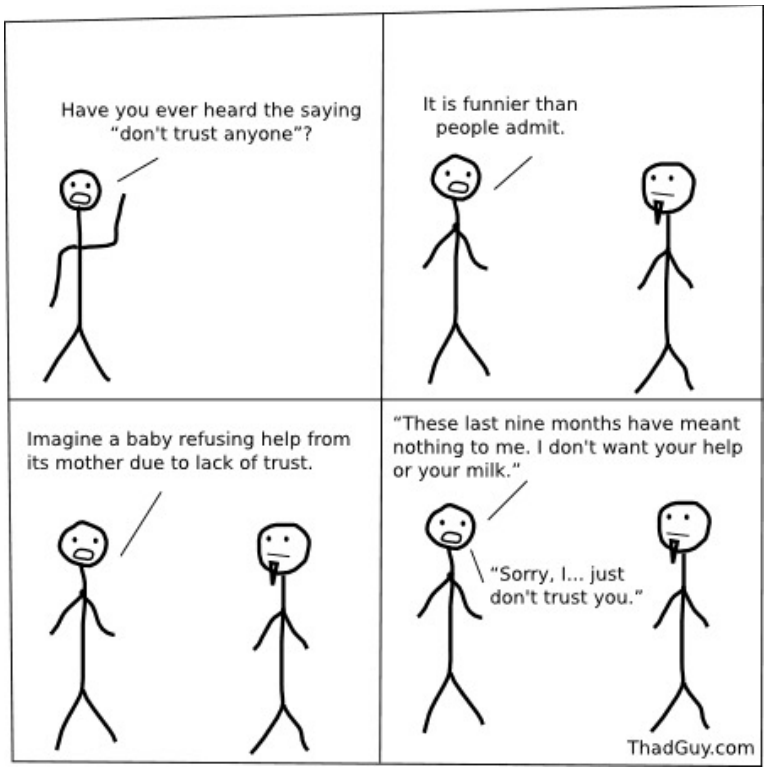
https://microchipdeveloper.com/tcpip:tcp-ip-five-layer-model

Internet architecture



https://www.yaldex.com/tcp_ip/FILES/06fig07.gif

Trusted and Untrusted Actors



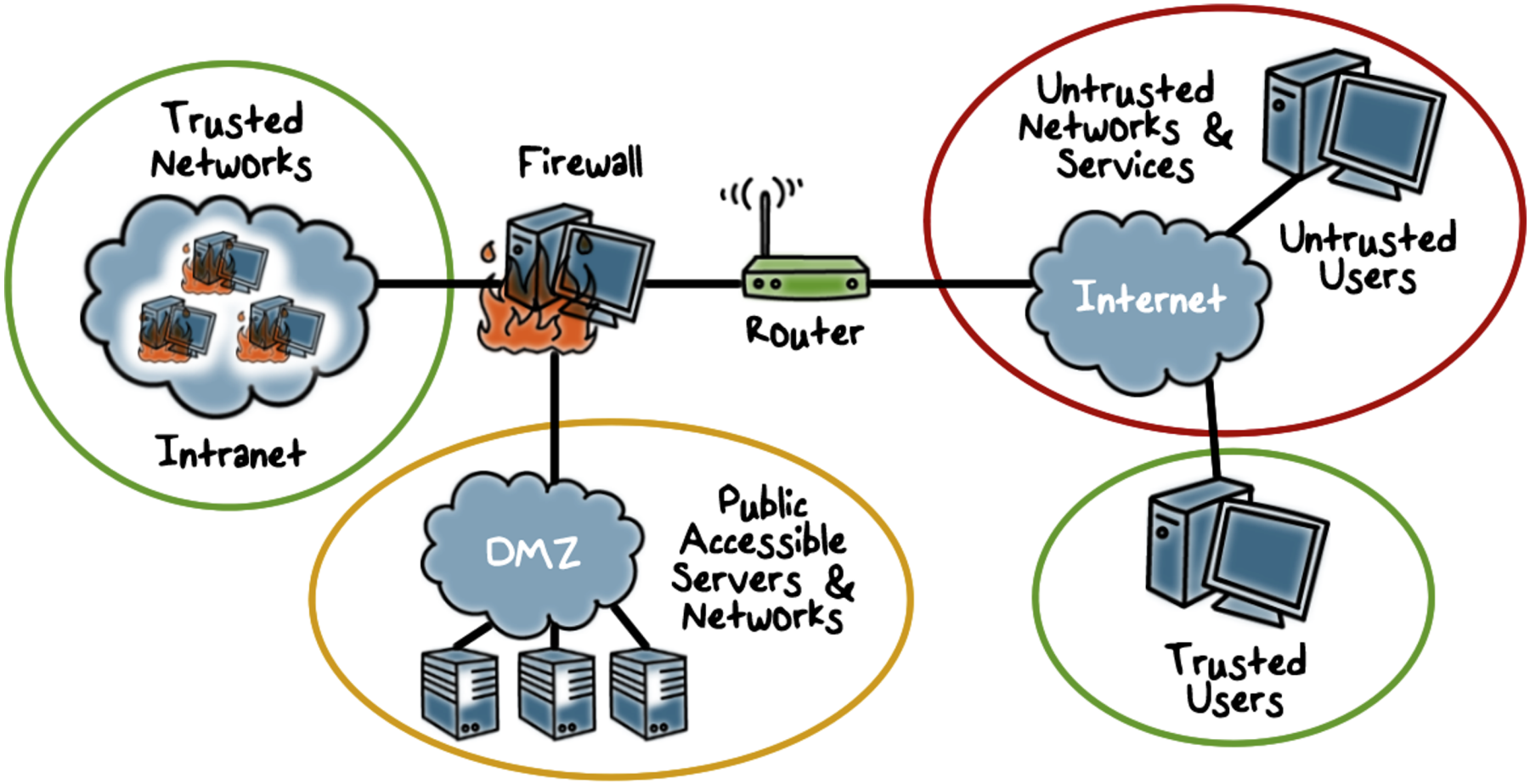
F

forbes.com/cartoons

A binary trusted/untrusted classification is an approximation.

Colorado State University

Firewalls



DMZ: “Demilitarized zone”, distributed firewalls, From Georgia Tech
Note multiple levels of trust.

Firewalls

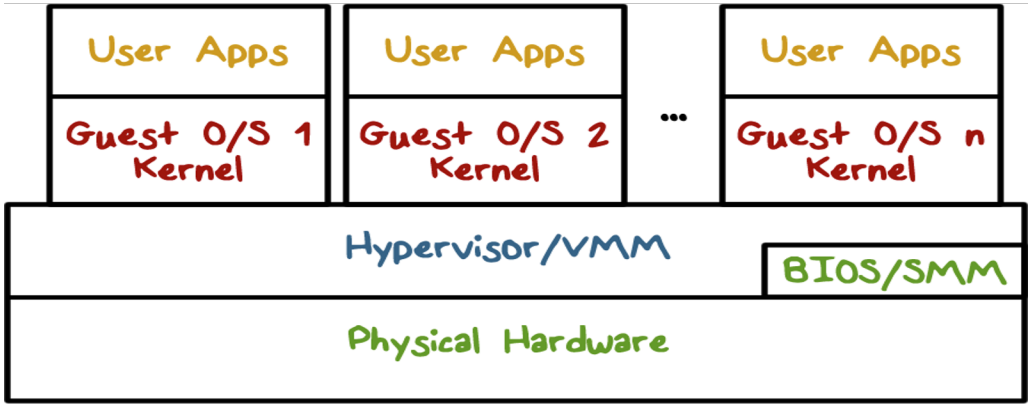
- A firewall checks traffic (packets or sessions) passing through it
- Can be programmed to check address ranges (IP addresses, ports), protocols, applications and content types.
- Can provide address translation, encryption

OS - trusted computing base

- The operating system serves as as trusted computing base (TCB) that controls access to protected resources.
  - Must establish the source of a request for a resource (authentication is how we do it)
  - Authorization or access control
  - Mechanisms that allow various policies to be supported
- How
  - Hardware support for memory protection
  - Processor execution modes (system and user modes)
  - Privileged instructions - can only be executed in system mode
  - System calls - transfer control between user and system code

Isolation in a system

- OS isolates address spaces of different processes using address translation. Also data vs code isolation.
  - Page tables governed by OS.
- In virtualization, hypervisor isolates virtual machines.
- Containers (Docker): Linux cgroups isolate process groups.



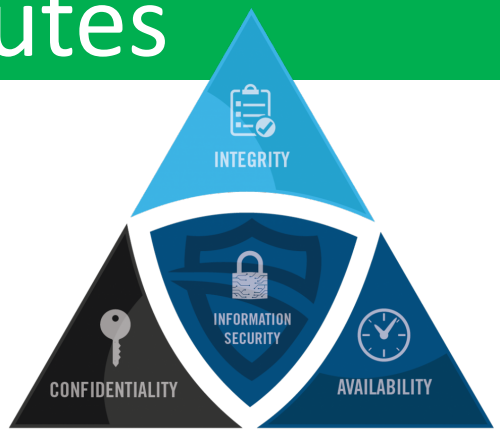
Quantitative Security

Colorado State University
Yashwant K Malaiya

Security Terminology

CSU Cybersecurity Center
Computer Science Dept

Key Security Attributes



Confidentiality: Preserving authorized restrictions on information access and disclosure,

- including means for protecting personal privacy and proprietary information.

Integrity: Guarding against improper information modification or destruction. Can be considered to include

- **Authenticity**: The property of being genuine and being able to be verified and trusted; confidence in the validity of a transmission, a message, or message originator.
- **Non-repudiability/Accountability**: requirement for actions of an entity to be traced uniquely to that entity. This supports nonrepudiation, deterrence, fault isolation, intrusion detection and prevention, and after-action. recovery and legal action.

Availability

Ensuring timely and reliable access to and use of information.

Computer security : Principles and Practice, William Stallings, Lawrie Brown

Adversary, Attack, Countermesure

Adversary (threat agent): Individual, group, organization, or government that conducts or has the intent to conduct detrimental activities.

Attack: Any kind of malicious activity that attempts to collect, disrupt, deny, degrade, or destroy information system resources or the information itself. Attack types

- Passive – attempt to learn or make use of information from the system that does not affect system resources
- Active – attempt to alter system resources or affect their operation
- Insider – initiated by an entity inside the security parameter
- Outsider – initiated from outside the perimeter

Countermeasure: A device or techniques that has as its objective the impairment of operational effectiveness of undesirable or adversarial activity, or prevention of espionage, sabotage, theft, or unauthorized access to or use of sensitive information or information systems.

RFC 2828, Internet Security Glossary

Assets, Risk, Threat, Vulnerability

System Resource (Asset): A major application, general support system, high impact program, physical plant, mission critical system, personnel, equipment, or a logically related group of systems.

Risk: A measure of the extent to which an entity is threatened by a potential circumstance or event, and typically a function of 1) the adverse impacts that would arise if the circumstance or event occurs; and 2) the likelihood of occurrence.

- To be studiesd in detail.

Threat: Any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, or the nation through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service.

Vulnerability: Weakness in an information system, system security procedures, internal controls, or implementation that could be exploited or triggered by a threat source.

RFC 2828, Internet Security Glossary

Assets, Risk, Threat, Vulnerability

System Resource (Asset): what needs protection by the defenders.

Risk: A measure of the adverse impacts and the likelihood of occurrence.

Threat: potential attempts by an adversary.

Vulnerability: Weakness in an information system that may be exploited.

Note of caution: In pre-cyber-security days, classical risk literature used the term vulnerability with a different meaning.

RFC 2828, Internet Security Glossary

Assets and threats

	Availability	Confidentiality	Integrity
Hardware	Equipment is stolen or disabled, thus denying service.	An unencrypted CD-ROM or DVD is stolen.	
Software	Programs are deleted, denying access to users.	An unauthorized copy of software is made.	A working program is modified, either to cause it to fail during execution or to cause it to do some unintended task.
Data	Files are deleted, denying access to users.	An unauthorized read of data is performed. An analysis of statistical data reveals underlying data.	Existing files are modified or new files are fabricated.
Communication Lines and Networks	Messages are destroyed or deleted. Communication lines or networks are rendered unavailable.	Messages are read. The traffic pattern of messages is observed.	Messages are modified, delayed, reordered, or duplicated. False messages are fabricated.

Question: where does ransomwere fit? Viruses?

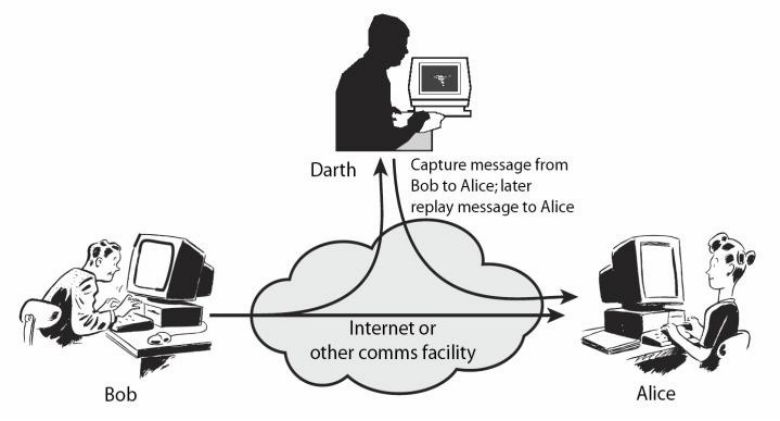
Computer security : Principles and Practice, William Stallings, Lawrie Brown

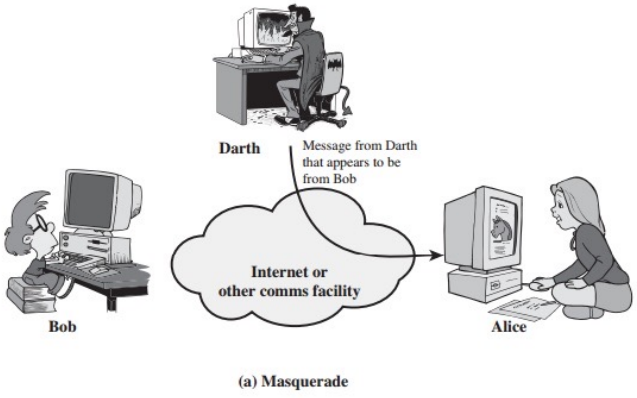
Attacks

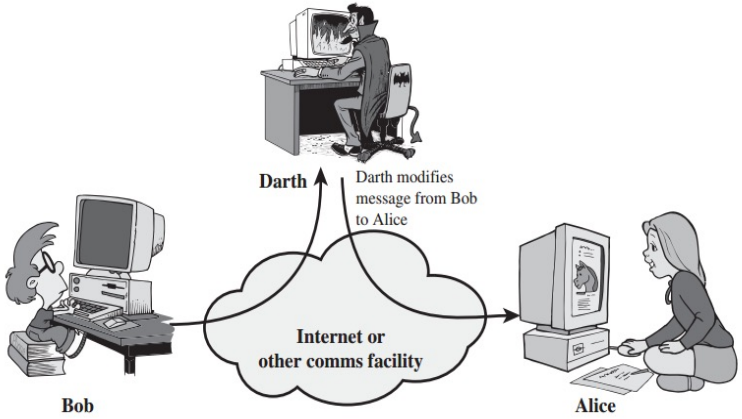
Passive Attack	Active Attack

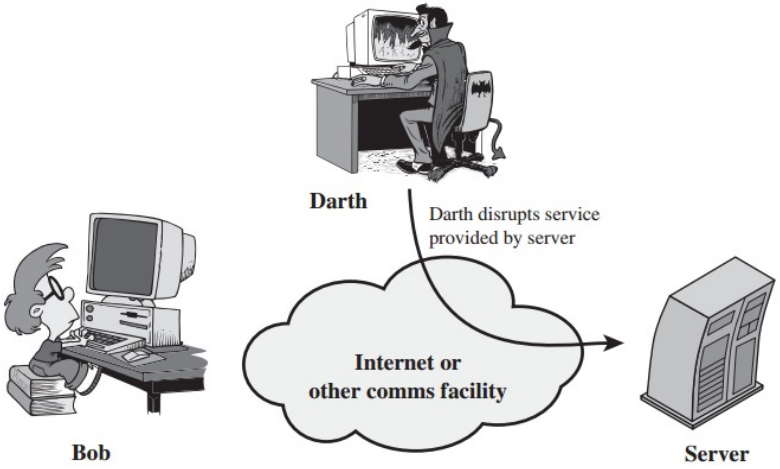
• Attempts to learn or make use of information from the system but does not affect system resources
• Eavesdropping on, or monitoring of, transmissions to obtain information that is being transmitted
• Two types:
  • Release of message contents
  • Traffic analysis
	• Attempts to alter system resources or affect their operation
	• Involve some modification of the data stream or the creation of a false stream
	• Four categories:
	  • Replay
	  • Masquerade
	  • Modification of messages
	  • Denial of service

Alice and Bob Diagrams





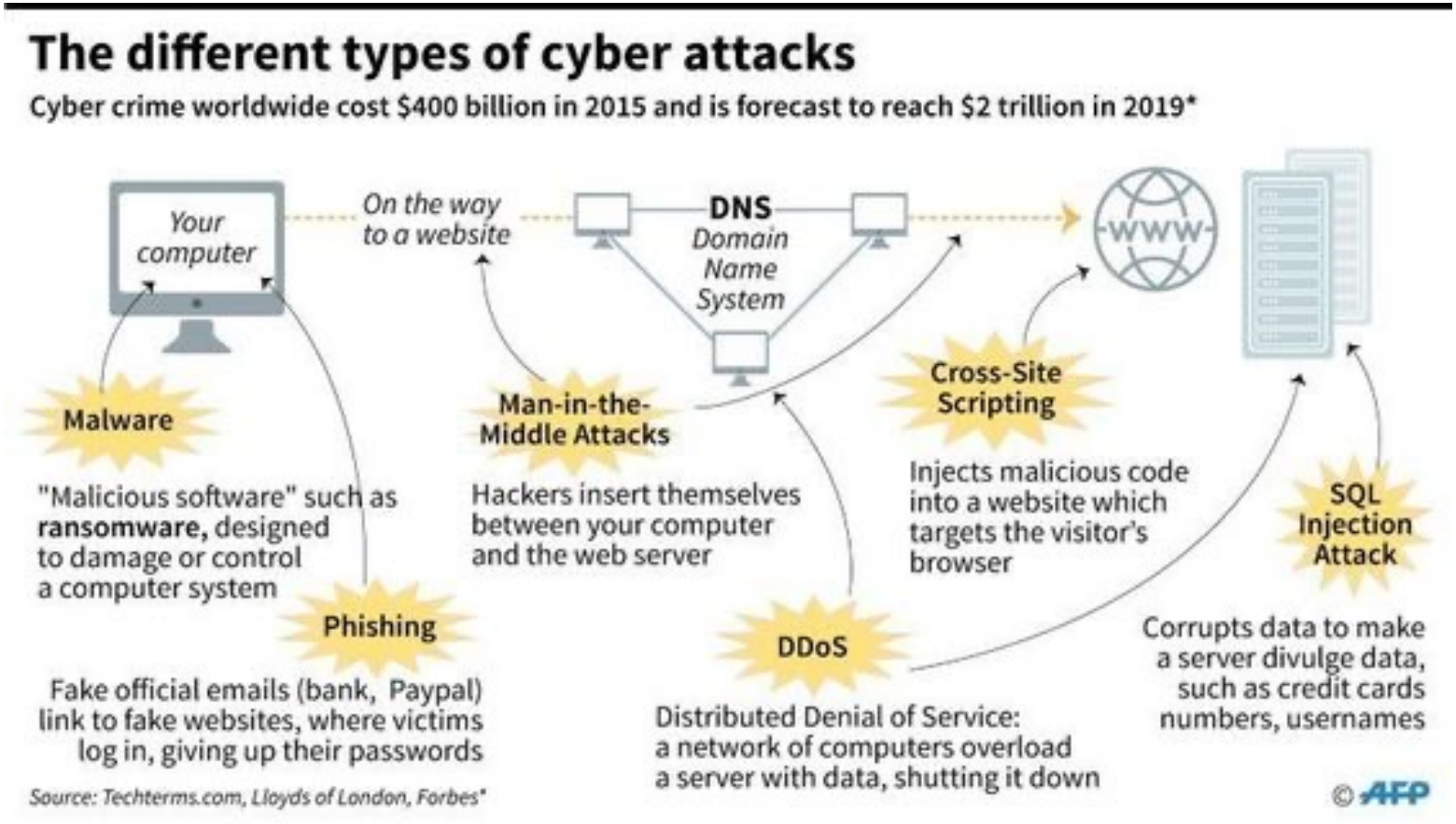




History: Rivest, Shamir, and Adleman's 1978 article "A method for obtaining digital signatures and public-key cryptosystems". a. Masquerade b. man-in-the-middle c. denial-of-service

Colorado State University

Cyber attack types



Colorado State University

Attack Surfaces

Surfaces: where the “holes” might be.

Network attack surface: vulnerabilities over an enterprise network, wide-area network, or the Internet.

- Including network protocol vulnerabilities, such as those used for a denial-of-service attack, disruption of communications links, and various forms of intruder attacks.

Software attack surface: vulnerabilities in application, utility, or operating system code.

- Web server, browser, Operating System.

Human attack surface: vulnerabilities in the personnel behavior

- social engineering, human error, and trusted insiders

Malware

- Malware ("malicious software"):
  - a catch-all term for any type of malicious software,
  - regardless of how it works, its intent, or how it's distributed.
- Virus
  - a specific type of malware that self-replicates by inserting its code into other programs. Types:
  - The **file infector** can burrow into executable files and spread through a network. A file infector can overwrite a computer's operating system or even reformat its drive.
  - The **macro virus** takes advantage of programs that support macros. Macro viruses usually arrive as Word or Excel documents attached to a spam email, or as a zipped attachment.
  - **Polymorphic viruses** modify their own code. The virus replicates and encrypts itself, changing its code just enough to evade detection by antivirus programs.

https://www.mcafee.com/enterprise/en-in/security-awareness/ransomware/malware-vs-viruses.html

Malware: Functional types

- **Worm:** a standalone program that can self-replicate and spread over a network. Unlike a virus, a worm spreads by exploiting a vulnerability in the infected system or through email as an attachment masquerading as a legitimate file.
- **Ransomware**: demands that users pay a ransom—usually in bitcoin or other cryptocurrency—to regain access to their computer.
- **Scareware**: attempts to frighten the victim into buying unnecessary software or providing their financial data.
- **Adware and spyware**: Adware pushes unwanted advertisements at users and spyware secretly collects information about the user. Spyware may record the websites the user visits, information about the user's computer system and vulnerabilities for a future attack, or the user's keystrokes.
  - Spyware that records keystrokes is called a keylogger.
- **Fileless malware**: Unlike traditional malware, fileless malware does not download code onto a computer, so there is no malware signature for a virus scanner to detect. Instead, fileless malware operates in the computer's memory and may evade detection by hiding in a trusted utility, productivity tool, or security application.

https://www.mcafee.com/enterprise/en-in/security-awareness/ransomware/malware-vs-viruses.html