Key to reading the document:

**Week number:** purple, underlined and bold.

**Chapter name:** blue, underlined and bold.

Sub topic: blue, underlined.

**Sub - sub - topic**: black, bold, underlined

Sub - sub- sub topic :black underlined

Important words : coloured in blue

THE DOCUMENT DOES NOT REFLECT THE CHANGES, YET

**FOR ALL EDITORS, leave a comment so i can see any changes u have made.**

# Week 1

# Chapter 1: Introduction to Networks

## What is the internet?

- It is a collection of billions of devices, networks, servers, hubs and many more devices that keep the internet running.
- The devices that are located on the edge of the internet are known as *hosts* or *end systems*.
- The data that is sent over the internet is known as *packet switches.*
- There are many different types of links joining the internet, such as *fibre optic cables*, *copper cables* , and *satellites*.
- It can be described as a network of networks.
- Whenever devices communicate with each other over the internet they use a set of rules called *protocols*.eg http, smtp, ftp and many more.
- The *IEFT(internet engineering task force)* sets the standards for Internet specifications, communications protocols, procedures, and events, this document is called the *RFC(Request for comments*).

## What are protocols?

- When we meet a new person or enter a new organisation, we as humans follow some rules and etiquettes.
- We talk to our elders in a formal way and an informal way with our friends that are of our age group.
- All these manners are heavily affected by the geo-political and socio-economic factors of that person, but generally all cultures have some sort of framework to follow when it comes to manners, respect, and etiquette.
- Similarly when the devices communicate they with the help of protocols.
- These define the format, order, network entities, and actions taken on message transmission.

# Week 2

## The structure of the internet

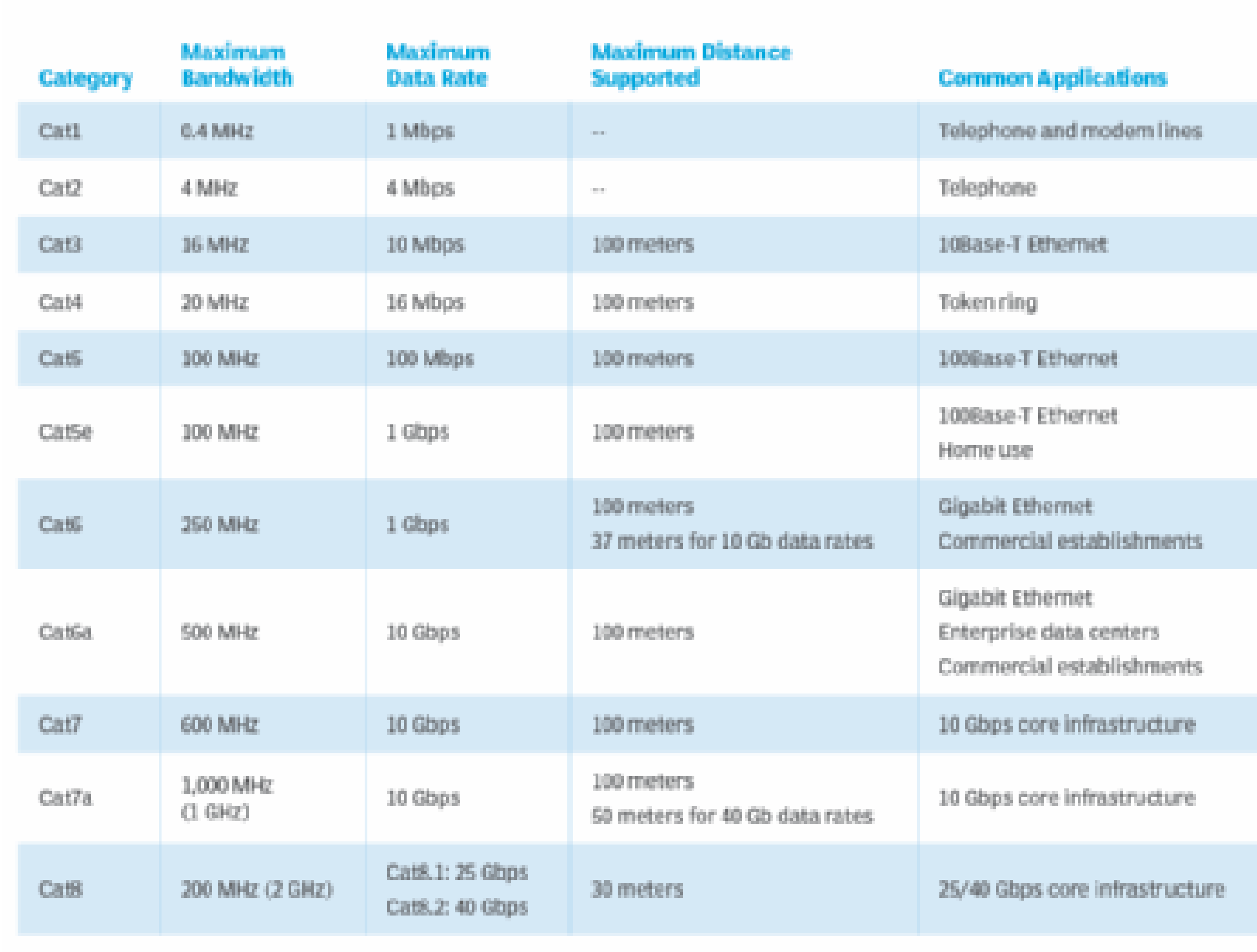
- The Networks edge often refers to the host, clients and servers located at the edge of the internet
- The edge is connected via access networks consisting of wired and wireless connections.
- The access networks are interconnected with each other and core networks, which are the backbone of the internet.
- Core networks allow for international communication via global ISPs.

## Sending and Receiving Data:

- When a host wants to send data over the network/ internet. Protocols breakdown the data into smaller chunks called packets.
- These packets are then sent over the internet.
- If we have a chunk of data that is L mbs and a transmission link with a transmission speed of R mbps then the formula for the rate of transmission or bandwidth is : $\frac{L}{R}$

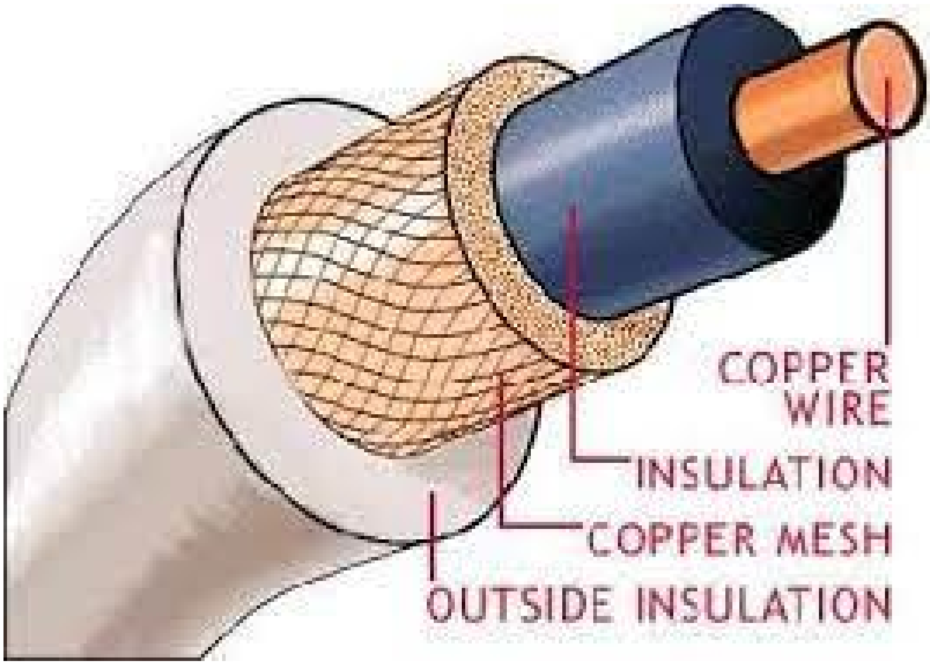
## Links: Physical Media

- The chunks of data that propagate between a transmitter and receiver are called a bit.
- When there is a physical connection between 2 devices it is called a physical link, this is also called guided media
- Twisted Pair (TP):
  - These are a number of copper cables twisted together and are insulated with the help of compounds.
  - The twisting helps to prevent the interference from electo-magnetic radiation and resists external interference from disrupting the data.
  - The below chart shows the different types of the Twisted pair cables.

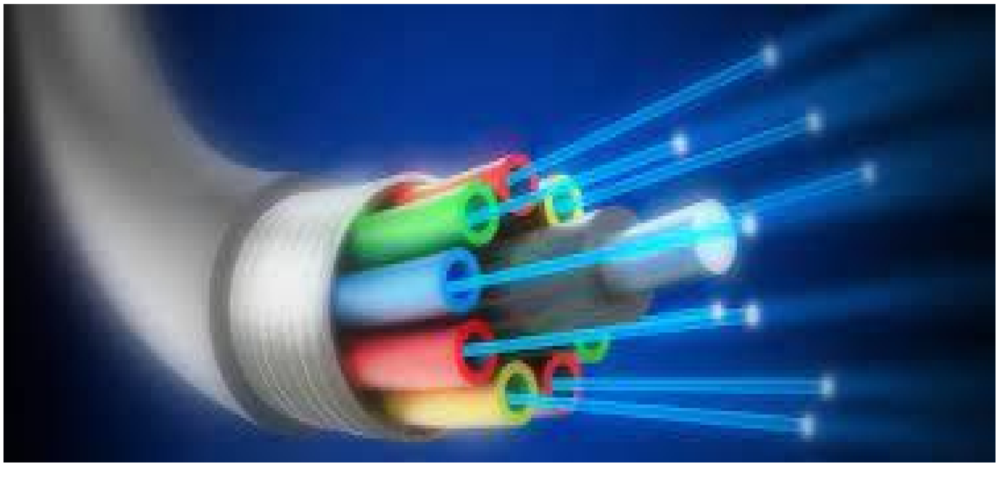


| Category | Maximum Bandwidth | Maximum Data Rate | Maximum Distance Supported | Common Applications |
|---|---|---|---|---|
| Cat1 | 0.4 MHz | 1 Mbps | -- | Telephone and modem lines |
| Cat2 | 4 MHz | 4 Mbps | -- | Telephone |
| Cat3 | 16 MHz | 10 Mbps | 100 meters | 10Base-T Ethernet |
| Cat4 | 20 MHz | 16 Mbps | 100 meters | Token ring |
| Cat5 | 100 MHz | 100 Mbps | 100 meters | 100Base-T Ethernet |
| Cat5e | 100 MHz | 1 Gbps | 100 meters | 100Base-T Ethernet<br>Home use |
| Cat6 | 250 MHz | 1 Gbps | 100 meters<br>37 meters for 10 Gb data rates | Gigabit Ethernet<br>Commercial establishments |
| Cat6a | 500 MHz | 10 Gbps | 100 meters | Gigabit Ethernet<br>Enterprise data centers<br>Commercial establishments |
| Cat7 | 600 MHz | 10 Gbps | 100 meters | 10 Gbps core infrastructure |
| Cat7a | 1,000 MHz<br>(1 GHz) | 10 Gbps | 100 meters<br>50 meters for 40 Gb data rates | 10 Gbps core infrastructure |
| Cat8 | 200 MHz (2 GHz) | Cat8.1: 25 Gbps<br>Cat8.2: 40 Gbps | 30 meters | 25/40 Gbps core infrastructure |

- Coaxial Cable:
    - It consists of a copper core  that is surrounded by a concentric copper woven shield, separated by an insulator
    - It is bidirectional meaning it can send and receive data at the same time.



- Fibre optic:
    - This cable consists of glass fibers that carry light pulses
    - This type of cable has the highest transmission speed as well as the lowest error rate.
    - This is also immune to interference and electromagnetic noise.



## Links: Non-Physical Media

- When the connection between 2 devices is through medium it is referred to as non-physical media or unguided media.
- Data bits are transmitted via electromagnetic waves
- When designing a wireless network , Architects have to take many factors into considerations such as:
  - Reflection
  - Obstructions by objects
  - Interference and noise

## Packet/circuit switching,internet structure

- Forwarding, aka switching,aka local action, is the act of moving packets of data from an input link to the appropriate destination link.
- Routing, aka global action, is the act of determining which path a packet of data takes.
  - All routers have a routing table, which has a header value and the corresponding output link.
- Eg: in a city forwarding is taking different paths, while routing is taking different junctions to change roads.

## Packet-switching:

### Store-and-forward:

- When a packet of data reaches the router, the router waits for the complete packet to reach the router.
- After it has loaded completely it is then forwarded to the next hop router.
- 1 hop = next closest router in the link

- There is a delay, known as packet transmission delay.

Queueing:

- It is defined as the accumulation of data at a network point due to the difference in transmission speeds.
- This can cause the data to be lost.

Performance: loss, delay, throughput:

- There are 4 types of delays.
    - Queuing: this is caused when there is a difference between the transmission speeds in the network.
    - Processing: this is the process of looking up the header for the data packet, this is usually ignored.
    - Transmission: the time it takes for data to cross a link. To calculate use: $\frac{L}{R}$ , where L is the size of the packet and R is the link speed.
    - Propagation: the time it takes the data to cross the whole network. To calculate use: $\frac{d}{s}$ , where d is the distance between the sender and receiver and s is the speed of propagation.
- Throughput is defined as the measure of how many units of information a system can process in a given amount of time
    - Instantaneous: throughput at a certain point
    - Average: throughput recorded between more than 2 links.
    - A bottleneck link is defined as the limiting link, i.e. a link with the slowest transfer speed.

## Basic terminology used in Computer

- Hard-wired: it refers to anything that is physically connected to a server, such as workstations, computers, and printers.
- Network interface: it is a network-specific software that communicates with the network-specific device driver, this allows the computer with the installed software to access the internet.
- Network: a link between a plug, or a connector or into a port or a jack.
- Gateway: A network node used to connect 2 networks, it also serves an entry and exit point.
- Bandwidth: it is the measure of the data transfer rate or capacity of a given network

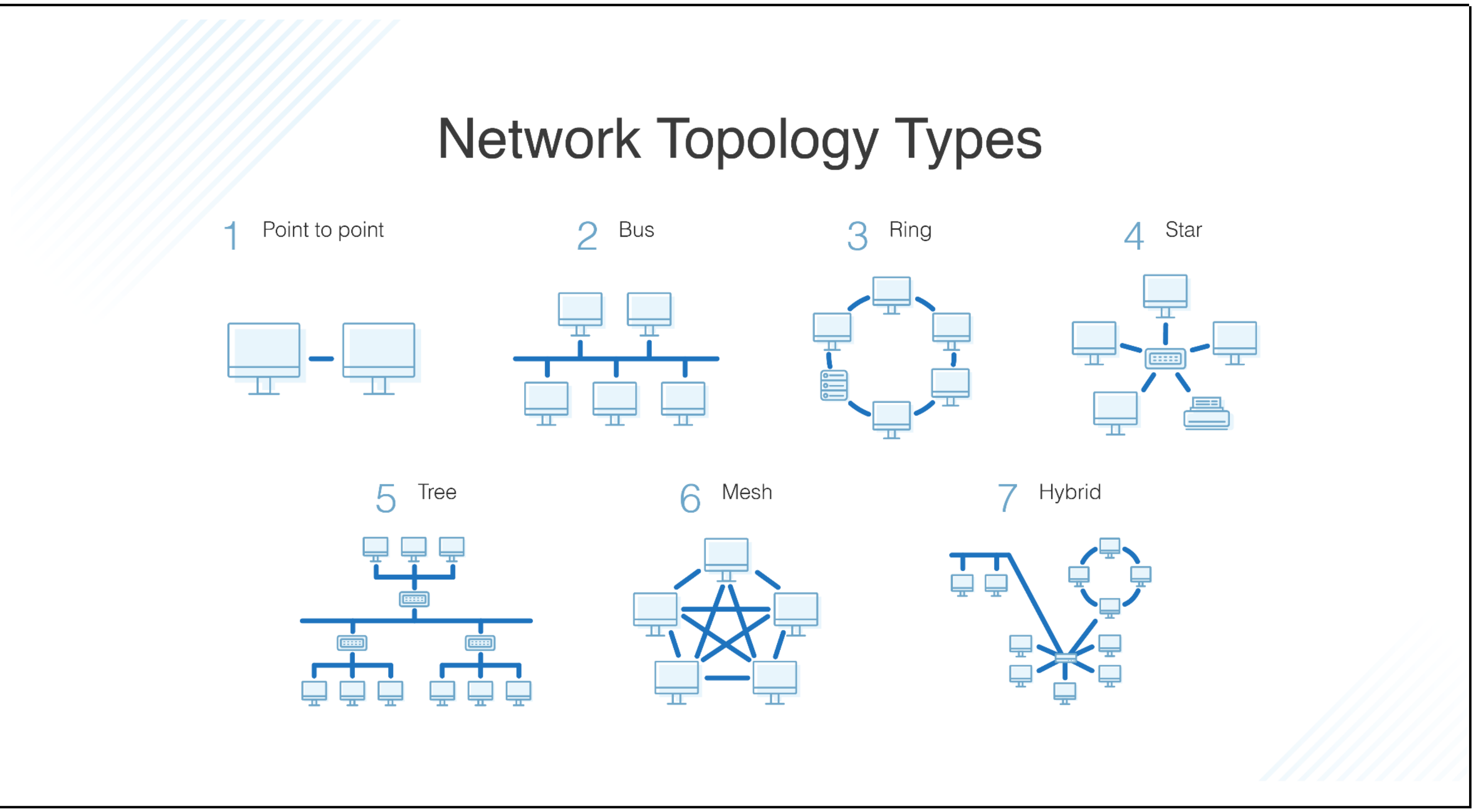
## CONNECTIONS:

- There are 2 types of connections:
    - Point to point:
        - A single path or link between 2 devices
    - Multipoint:
        - Multiple paths between 2 devices.

## PHYSICAL TOPOLOGY:

- There are 5 type of topology:
    - Mesh
        - All devices are connected to each other
    - Star
        - All devices are connected to a central node/sever
    - Bus
        - All devices are connected in serial way, or just by 1 cable

    - Ring
        - All devices connected to a central cable which loops back to the first node
    - Tree
        - Multiple nodes are connected to a serial line. Each node then branches out to form smaller sub branches.



Network Topology Types

## CATEGORIES OF NETWORKS:

- There are 4 categories of networks:
- PAN( personal area network): it usually refers to the smallest type of network, such as a home network.

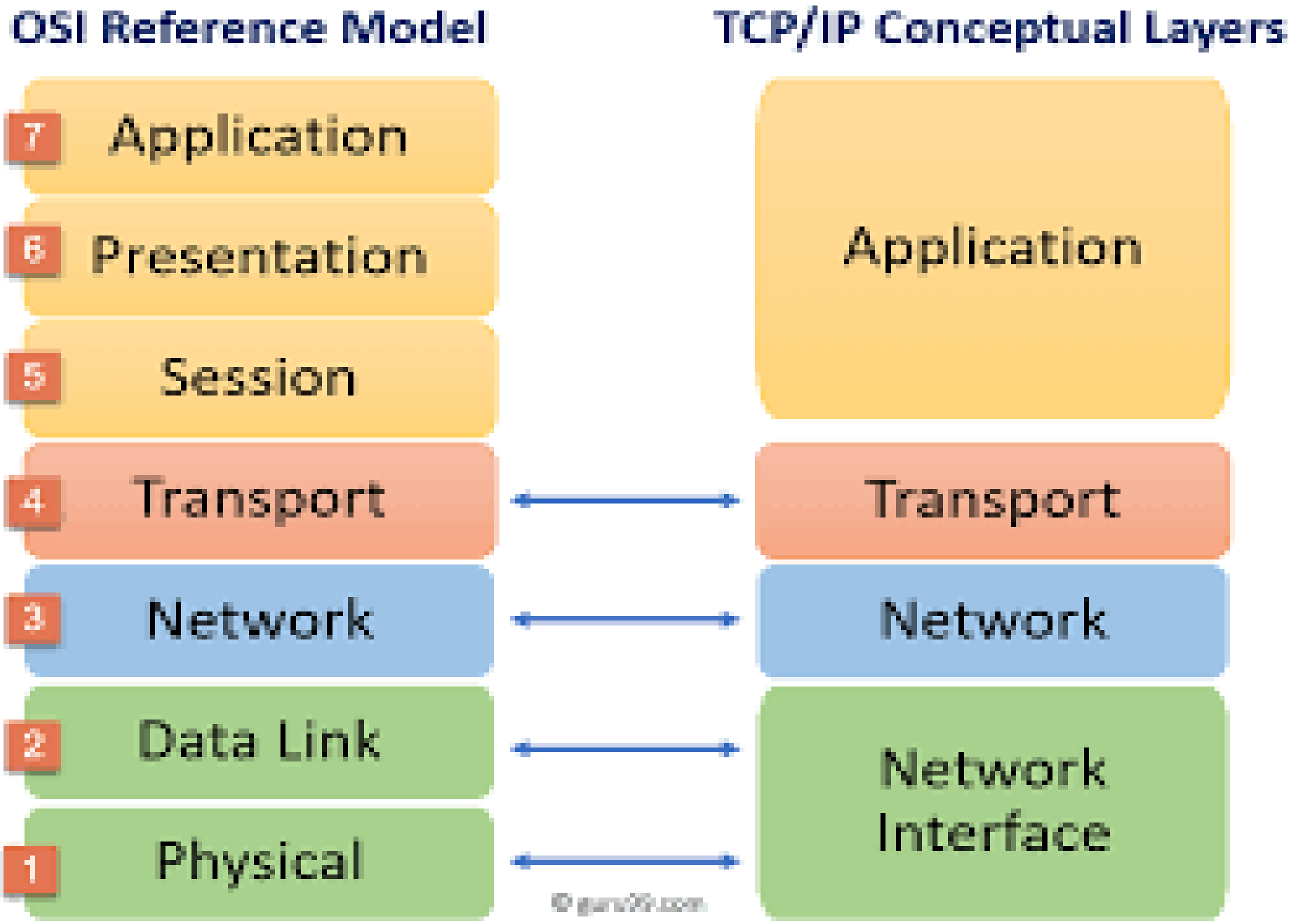
- LAN(Local area network): It refers to a network composed of multiple PANs, such as the network your entire neighbourhood is connected to.
- MAN(metropolitan area network): it refers to networks composed of multiple LANs, such as the network your entire city is connected to
- WAN (wide area network): it refers to the networks composed of multiple MANs, LANs, and PANs. It is the largest type of network. A good example of this is the nation ISP.

## LAYERING AND PROTOCOLS:

What is layering:

- It refers to the different protocols, softwares and hardwares a chunk of data goes through from sender and receiver.
- This has many benefits:
  - It allows for the modularization of the system. This means each layer's components can easily be upgraded, changed or removed.
  - It also allows us to track data more easily,  and hence makes resolving bugs easier.
  - It also breaks a complicated process into manageable chunks of data.

- There are 2 types of layering models:
    - OSI:
        - The Open Systems Interconnection (OSI) model describes seven layers that computer systems use to communicate over a network.
        - This is a theoretical model.
    - TCP/IP:
        - It stands for Transmission Control Protocol/Internet Protocol.
        - The TCP/IP model is a concise version of the OSI model. It contains four layers, unlike seven layers in the OSI model. The layers are:



## ENCAPSULATION:

- When we want to send or receive data over the internet it must pass through each of these layers.
- When the data enters a new layer, a header is attached to the message, to ensure that it doesn't get lost,it is later removed when it reaches the destination's corresponding layer, this process is known as encapsulation.
- Let's have a step by step look at this process.

| | |
|---|---|
| Fist the message is generated by the user, and is sent | M |
| The transport layer then adds the corresponding header the message is now called a segment | H M |
| Next the network layer add the its own header, the message is now called a datagram | H H M |
| And finally the link layer adds its own header, it is now called a frame. Finally it is sent over the physical network. Where the headers are removed at their corresponding layers | H H H M |

## Chapter 2: Introduction to cybersecurity:

### What is cybersecurity / information security?

- It is the field of study of computer science that studies the confidentiality, integrity, and availability of a computer system.
- It relates to the protection of a system's hardware, software, personnel, users, and the digital information present in these assets..
- It is an important part of many systems where data security is of the utmost importance such as in banks, in  armies, and the e-commerce sectors of the world.
- The goal of information security is to protect data's integrity, confidentiality and availability, wherever it may be.

### What challenges do cyber security specialists face?

- Due to the continuous evolution of technology, there are many loopholes, back doors, and inadequate procedures which have come about.
- From simple booby traps and keyloggers to large scale DDos attacks and operations such as titan rain. Attacks come in almost every form and type available, which makes it difficult to anticipate and prepare for.
- Some attacks are not even done by humans ,but by robots or very sophisticated programs.
- Similarly, a perpetrator doesn't use just one pc, many attacks involve multiple pc and systems making it harder to trace the culprit.

- By the same reason, it is has become easier to detect vulnerabilities in a system
- There is no 1 solution to this problem. As you cannot give the same medicine for cancer to patients with a fever, the same is true for cyber security.
- Nowadays, due to the internet everything is connected, this means if anyone has access to even 1 device, he/she can easily access all the devices on the network.
- Furthermore, due to the lack of security patches and updates systems remain exposed to threats for longer periods of time. Some companies even do this on purpose so that people have to upgrade their current software to get protection.
- Due to so many devices being manufactured, each device is vulnerable to an attack, thus the numbers of attacks have also drastically increased.
- Finally, it is also harder to protect systems from threats due to the lack of awareness. Many people have no idea what cyber security is or why is it important, this makes them very vulnerable to threats.

## UNDERSTANDING SECURITY:

- If you want to have a secure system then you need a number of security procedures.
- This decreases the convenience of the user. This means that the higher the security the lower the convenience.

- For example: if you have multiple doors to a house, each with a different key, it becomes harder to break in, but it takes longer and you have to carry a block of keys to enter your house

THE CIA:

- CIA stands for confidentiality, integrity, and availability.
- Confidentiality:
  - It means that no user has unnecessary access to the system's data or working tools.
  - This prevents the data from being exploited or manipulated.
  - For Example: a peon and a professor working in a university must not have the same access level to the universities records and systems.

- Integrity:
  - Integrity is defined as safeguarding the information in its original state while it is stored , in transit and while being accessed.
  - For example: if you receive your salary in your bank account. Originally it was 200,000 Rs but you only received 20,000. This means that the data was not secured and it has lost its integrity.
- Availability:
  - It means that the authorised personnel or machine has access to the data it has been cleared for.

  - For example: if a database architect has been hired to maintain a database, but he has not been given the password to access it, then that means there is no availability.

## INFORMATION SECURITY TERMINOLOGY:

- In the field you will have some trade terminologies, these are words that have meanings related to the particular field.
- Asset: something of value, such as data, hardware etc.
- Threat : anything that has the potential to get through the security measures
- Threat agent: the object that can potentially carry out the threat
- Vulnerability :Any weakness that allows the threat agent to bypass security measures.
- Exploit: To take advantage of a certain vulnerability.
- Risk: the probability that the threat agent will exploit a vulnerability.

## THREATS AND ATTACKS, A CLOSER LOOK:

- Attacks are defined as an attempt to bypass a security measure.
- There are 2 types of attacks:
  - Passive:
    - Attacks that deal in gathering information, without harming the target’s OS
    - This includes attacks like:
      - phishing attacks
      - Snooping attacks
      - Spoofing attacks

  - Active:
    - Attacks that deal with harming and exploiting the user’s information.
    - This includes attacks like:
      - DDos
      - Malware infection

## SECURITY MECHANISM:

- It is the process that is followed before, during and after a security breach.
- Each security mechanism has a:
  - Policy: it is a set of rules and procedures to follow. Each company has a different policy for a different incident
  - Mechanism: the tool or tool that is used to carry out / enforce the policy.
- Companies implement the prevention, detection, adn recovery system.
  - Prevention:
    - This aspect focuses on finding the system vulnerabilities before they are detected by threats agents.
    - This step also includes setting up the right procedures and mechanism incase of a system breach.
  - Detection:
    - This aspect focuses on detecting the threat before it has a chance to deploy.

        - This step includes direct warfare between the assailants and the cyber security specialists. This may include; tracking software, worms, and virus deployment
    - Recovery:
        - This step focuses on the aftermath of the previous 2 steps.
        - This includes repairing any broken components of the security system, checking the integrity of the data and so on.

## CHAPTER 3: MALWARE AND SOCIAL ENGINEERING ATTACKS

### WHAT IS MALWARE?

- For our course it is defined as a malicious programs that are designed to extract or waste resources from the users system
- There has been hundreds of viruses over the decades and each has been created for a specific purpose
- The first virus developed by Pakistan was done by a 19 year old, in 1986. It was considered the world's first IBM PC compatible virus.
- Worms are another type of malware, the main purpose of worms is to use up the system resources.
- Viruses are pieces of malicious software whose primary purpose is to extract information and destroy a system, from the inside out.

- There are 4 types of viruses:
    - Computer virus:
        - It is malicious code which replicates itself os the same computer
    - Program Virus:
        - Malicious code which infects the exe file
    - Macro virus:
        - Viruses which are written in the native language of the software
        - They are usually a series of instructions that can be grouped together as a single command.
    - Armoured virus:
        - They are viruses that have a variety of mechanisms in places to make it difficult to detect.
- There are 4 ways of classifying viruses:
    - Circulation:
        - This the spreading ability of the virus
        - There are many ways a virus can spread; via email attachments, through usbs and many other ways.
        - Worms spread through a system, by exploiting the vulnerabilities of that system.
        - Some of the ways a virus can enter circulation are,using an appender which appends itself to the end of the file, but this can be easily detected by virus scanners
        - Some ways armoured virus avoid detection:

- Swiss cheese:The virus breaks itself into chunks and scatters itself into the program. When reassembled all the segments come together to infect the program.
- Split infection: the virus is encrypted and placed at the end of the program. It then scatters the decryption key all over the program. After all the parts of the decryption key are collected the virus is activated and the file dies (RIB: rest in bits)

- Infection:
  - This is the ability to embed itself in a system, aka when does the virus activate and for how it is active.
  - There are many examples of this:
    - Trojan virus:
      - In ancient mythology the Greek besieged the city of troy. But were on the verge of losing, so they surrendered and gifted the people of Troy a giant wooden horse. At night when everyone was asleep, the soldiers exited the giant horse and opened the gate to the rest of the awaiting Greek army.
      - Similarly to the myth the trojan virus masquerade itself as normal software, but it contains some hidden code. This allows it to harm the infected system.

      - Some trojan viruses even give remote access to the system, these are called RAT or remote access trojans
    - Another example of infected files is ransomware. Which are software or programs that when installed prevent the access of the user to the system.
    - Crypto ransomware is any software that encrypts the whole system and can only be unlocked if the hijackers are paid.
- Concealment:
  - This is the ability of the virus to hide
  - Many assailants/ viruses use Rootkits to hide themselves when they are in the system.
  - This can be done by removing entry logs.
  - They can also alter the OS with specific files which may harm the computer even more.
- Payload:
  - This is the ability of the virus to damage or collect information from a system.
  - Generally these fall into 1 of these categories:
    - Collection of data:
      - This includes the use of softwares to collect data against victims and identify the system’s vulnerabilities.

    - A software of such nature is a **keylogger**; which collects all the keystrokes of the users or the infected system and sends it to the perpetrator.
    - **Keyloggers** come in both hardware and software form
    - Adware is another such software, but instead of collecting data it spams the user with unwanted advertisements, which may lead to degradation of the users productivity or even the system crashing.
- Deletion of data:
    - Logic bombs are code that is added to a system and lies dormant, until it is activated by a certain event like a date or a specific value.
- Modify the security system
    - Some malware can even access the security system to modify it, an example of this is a backdoor, which grants remote access to the system.
    - Backdoors are dangerous because they prevent the security system from detecting the backdoor
- Launch attacks:

- It is one of the most popular ways malware infects the software.
- An infected robot is called a zombie or bot
- These can spread and become a swarm and the one who controls them is called a bot herder
- They can be controlled by using C&C (command and control) instructions.

## SOCIAL ENGINEERING ATTACKS:

- They are a type of attack which focuses on the vulnerability of a system and collecting information about a user's system.
- these often involve using psychological and physical techniques
- Some psychological tactics include:
  - Gaining the trust of the victims
  - Tailgating: which includes following the victims and spying on them physically, often referred to as reconnaissance, shoulder surfing, and piggy backing.
  - Going through their trash, which is referred to as dumpster diving. The online version of this is known as google dorking.
- Some non physical tactics includes:
  - Phishing: which includes luring ,usually by an email ,a victim by pretending to be a legitimate source and then redirecting them to a fake website and asking them to enter their credentials.

  - Some variations of this tactic include:**spear phishing**, in which specific targets are attacked. **Whaling** refers to the act of targeting large organisations. **Vishing** is phishing but instead of emails, social engineers use calls.
  - Spam is all the unwanted messages and email u receive, much of today's malware uses spam to spread virus and worms. A variation of this is **image spam** where instead of messages attackers use images to bypass filters.
  - Hoaxes are false warning messages from unauthorised sources, whose aim is to get unsuspecting users to change their security settings to allow them to gain access to their system.
  - A water hole attack is an attack where a group of individuals who frequently visit a website are attacked.

## Chapter 3: Cryptography

### What is cryptography?

- It is the process of taking readable or plain text and converting it to cipher or encrypted text.
- This makes it impossible for hackers to read the data, thus this keeps the confidentiality and integrity of the data. It also preserves the authenticity of the data as well as the availability.
- Cryptography is not the solution to all of the challenges, but it helps us secure many types of data.

- Cryptography also helps us verify the proof of origin aka non-repudiation. This is done with the help of digital signatures.

How is this process done?

- Algorithms are used to generate keys which are used to encrypt and decrypt the data.
- If the same key is used to encrypt and decrypt the data, then it is known as symmetric encryption. If 2 keys are used then it is known as asymmetric encryption.

Types of cryptography:

There are 2 main types of cryptography:

**Monoalphabetic cipher:**

- These ciphers map all letters to exactly 1 letters, we will do 2 types of cipher:
- Shift cipher:
  - also known as caesar cipher.
  - This involves moving or shifting the letters of the alphabets so they map on some other letter of the alphabet.
  - The sequence of the letters are preserved.
  - Let's say we need to encrypt the words: hello world

First we write all the alphabets and their corresponding shift:
Plain:    ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: DEFGHIJKLMNOPQRSTUVWXYZABC

Now we substitute each letter with its pair, this becomes:

| |
|---|
| LHOOR ZRUOG |
| We can decrypt it by reversing the method. |

- Substitution cipher:
    - They use a similar system to shift ciphers but instead of shifting it uses an equation to get the next letter.
    - The letters sequence is not preserved

## Polyalphabetic cipher:

- These ciphers map 1 letter onto various letters.

## The strength the of a cryptographic algorithm:

- The strength of an algorithm is determined by the pseudorandom number generator, the diffusion, and the confusion.
- A pseudorandom number generator(PRNG) is the ability of the software to develop a sequence of numbers which resemble a random number.
- Diffusion means that each character has multiple ciphers when it is put through the algorithm. The more the diffusion the harder it is to break the encryption.
- Confusion means that each binary digit (bit) of the ciphertext should depend on several parts of the key, obscuring the connections between the two. The better the confusion the harder it is to crack the algorithm.

## Types of Cryptographic algorithms:

## Hash algorithms:

### **What is a hash?**

- It is an algorithm that creates a unique digital fingerprint that is able to verify that you have sent the message.
- The hashed message is called a digest.
- Some common examples include:
  - Message digest
  - Secure hash algorithm
  - RACE Integrity Primitives Evaluation Message Digest
  - Hashed Message Authentication Code.

### Message digest 5:

- There are 3 main types:
  - **Message digest 2**, which has a processor size of 16 bit. Messages are processed as 128 bit sections.
  - **Message digest 4,** which has a processor size of 32 bit, messages are processed as 512 bit sections
  - **Message digest 5,** which has a processor size of 32, but it uses 4 rotating 32 bit variables.

### Secure hash algorithm:

- Created by NSA and NIST
- 3 subtypes:
  - **SHA 1 -** works like MD4, but creates hashes 160 bits in length

    - **SHA 2 -** has 4 subtypes: SHA-224, SHA-256, SHA-384, SHA-512. Each of these hashes produces a hash that is in the name, for example SHA-224 produces a hash to the length of 224 bits
    - **SHA 3 -** it is more secure than its siblings because it is impervious to length based attacks(Attacks where a hacker can use the length of the hash to figure out the original)

Features of that make a hashing algorithm secure:

- **Fixed size:** no matter what data we put it must always give us a digest of a fixed length
- **Unique:**All digests must be unique.
- **Original:** the algorithm must not be unique
- **Secure:** reverse engineering is impossible or at least very hard

Hashed Message Authentication Code(HMAC):

- **Step 1:** first a message is put through a hashing algorithm, usually md2, or SHA1 . The message is then encrypted using the key and it is sent over the network. Message + digest
- **Step 2:** when the message + digest is received the message is decrypted using the same key and it is put through the same hashing algorithm and then the generated hash is compared with the hash that was sent.
- **Step 3:** if the 2 digests match then the message is authentic, otherwise it is said that the message has been tampered with.

## Stream cipher:

**Definition:**

- It is a type of cipher which encrypts individual characters of the plain text.

**Working:**

1. Each letter is taking and put through an algorithm to encrypt it.

**Advantages:**

1. Faster than block cipher
2. Less prone to error propagation

**Disadvantages:**

1. Easier to decipher
2. Low diffusion
3. More susceptible to malicious insertions

## Block cipher:

**Definition:**

- It is a type of cipher which encrypts entire blocks of plain text.

**Working:**

- Breaks plain text into 8 to 16 bytes segments
- These are then encrypted independently of each other.

**Advantages:**

1. If we change a single letter in the plain or cipher text, then half of the bits in the plain or cipher text will change, ie high diffusion
2. Immunity to insertion symbol

**Disadvantages:**

1. Slower than stream
2. Padding is required
3. More prone to error propagation

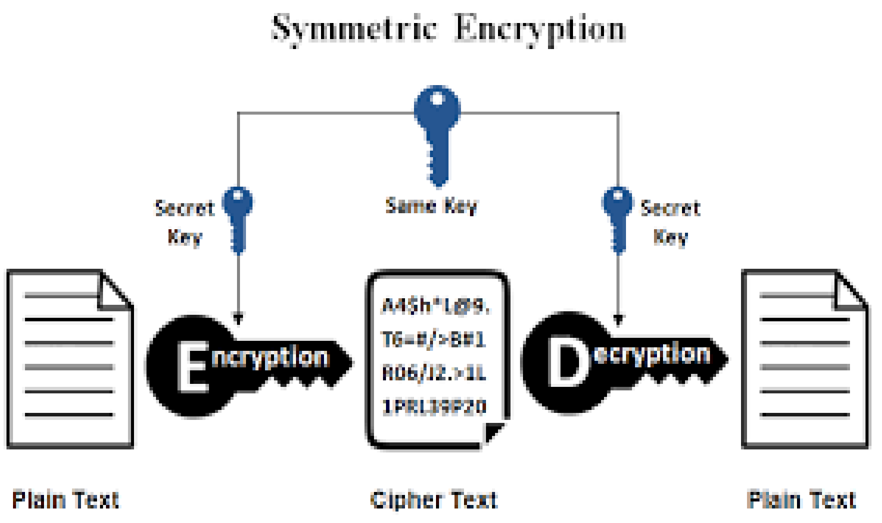
One time pad(OTP):

**Definition:**

- It is a type of a cipher which uses a key to encrypt the message. The receiver must also have the same pad to reverse the cipher.
- The uniqueness of this cipher is that it is nearly impossible to break. This is because the key is only and only used once and then it is destroyed.
- First deployed by the german diplomats to send and receive messages

Symmetric cryptographic algorithms:

**Definitions:**

| | |
|---|---|
| They are those ciphers which use a single key to encrypt and decrypt the message |  Symmetric Encryption<br>Secret Key<br>Same Key<br>Secret Key<br>Encryption<br>A4$h*L@9.<br>T6=#/>B#1<br>R06/J2.>1L<br>1PRL39P20<br>Decryption<br>Plain Text<br>Cipher Text<br>Plain Text |

**<u>Types:</u>**

<u>DES:</u>

Definition:

- Stands for Data Encryption standard.
- It is part of the symmetric encryption family, which means that it uses 1 key encrypt and decrypt a messages

Working:

1. First the plain text is broken into 64 bit chunks
2. Each segment is encrypted using a 56 bit key
3. It passes the segment through 16 rounds of encryption

<u>Double DES:</u>

- It has the same definition as DES , the key it creates has uses 2 sub keys
- It also passes the chunk of data through the encryption process 32 times, or 2 rounds.

<u>Triple DES:</u>

- It has the same definition as DES, the key it creates has 3 sub keys
- It also passes the chunk of data through the encryption process 48 times or 3 rounds.

AES:

Definition:

- Stands for Advanced encryption standard
- Has never been broken, yet…

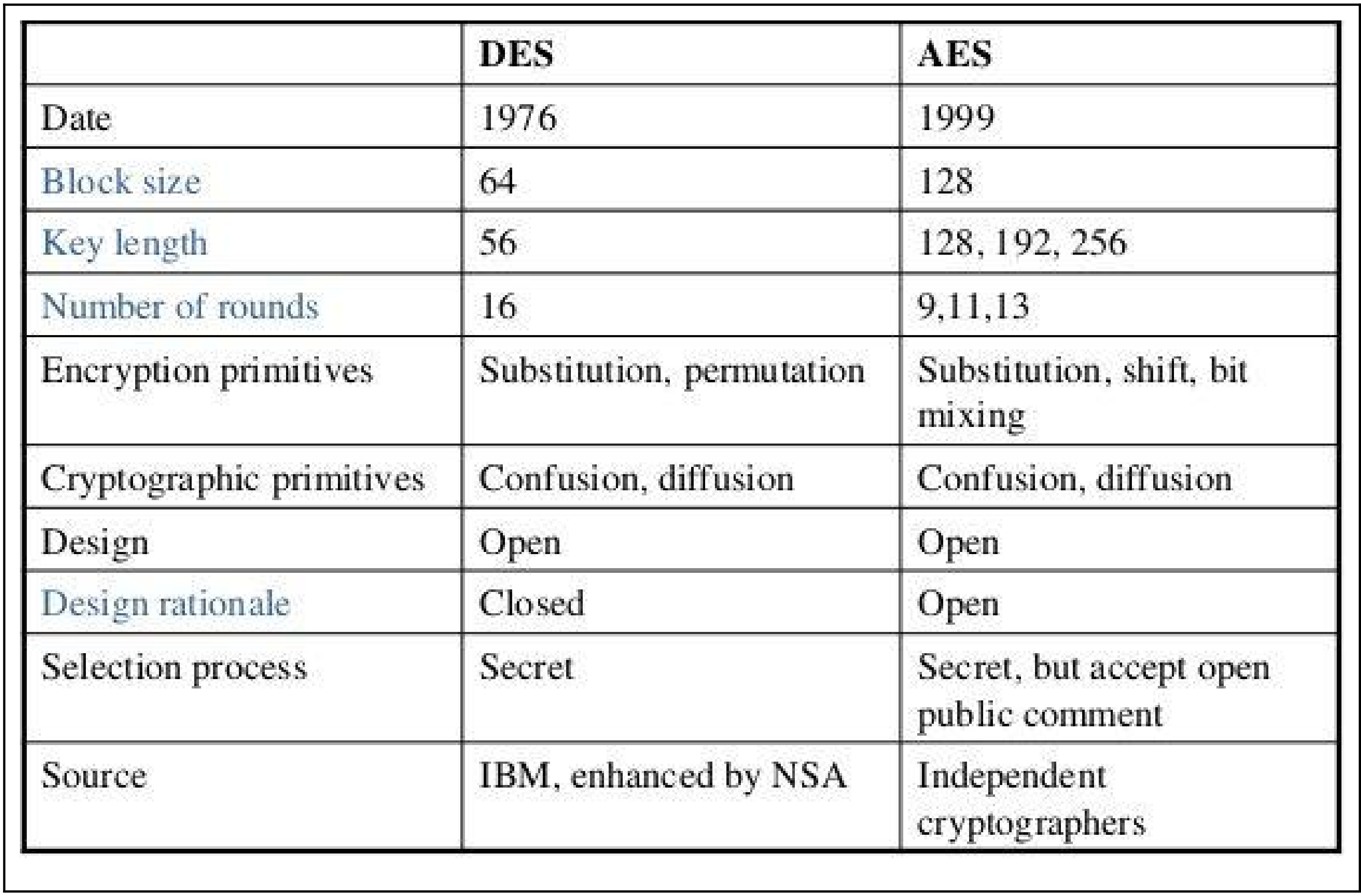
Working:

1. First a block of 128 bits is select from the plain text
2. Encrypted using the selected key and then text passes through the corresponding rounds to produce the final output.

Types:

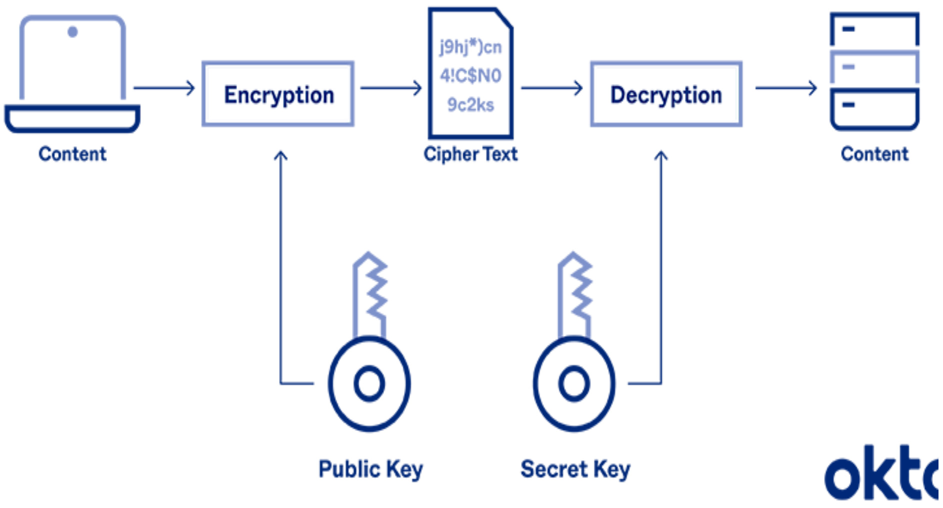
1. AES-128: has a key size of 128 bits and performs 9 rounds of encryption.
2. AES-192: has a key size of 192 bits and performs 11 rounds of encryption.
3. AES-256: has a key size of 256 bits and performa 13 rounds of encryption.

## Main difference between AES and DES:



| | DES | AES |
|---|---|---|
| Date | 1976 | 1999 |
| Block size | 64 | 128 |
| Key length | 56 | 128, 192, 256 |
| Number of rounds | 16 | 9,11,13 |
| Encryption primitives | Substitution, permutation | Substitution, shift, bit mixing |
| Cryptographic primitives | Confusion, diffusion | Confusion, diffusion |
| Design | Open | Open |
| Design rationale | Closed | Open |
| Selection process | Secret | Secret, but accept open public comment |
| Source | IBM, enhanced by NSA | Independent cryptographers |

## Asymmetric cryptographic Algorithms:

| | |
|---|---|
|  Content<br>Encryption<br>j9hj*)cn<br>4!C$N0<br>9c2ks<br>Cipher Text<br>Decryption<br>Content<br>Public Key<br>Secret Key | They are those ciphers and algorithms which use 2 keys to encrypt and decrypt. |

- One key is known by everyone on the network, aka the public key.
- One key is known by only the user or holder, aka the private key

**Digital Signature Algorithm:**

- A digital signature is a mathematical technique used to verify people on the internet.

- A digital signature can:
    1. Verify the sender
    2. The sender can never disown the message
    3. Prove the integrity of the message
- There are 4  primary condition, which validate the use of DS:
    1. If someone signs a message with a unique signature, then no one on the network must be able to sign a message with the same signature.
    2. It must be authentic; if a person receives a digitally signed message, he must be able to reverse the encryption and authenticate the sender
    3. It must not be alterable by anyone
    4. It must not be reusable

**Working:**

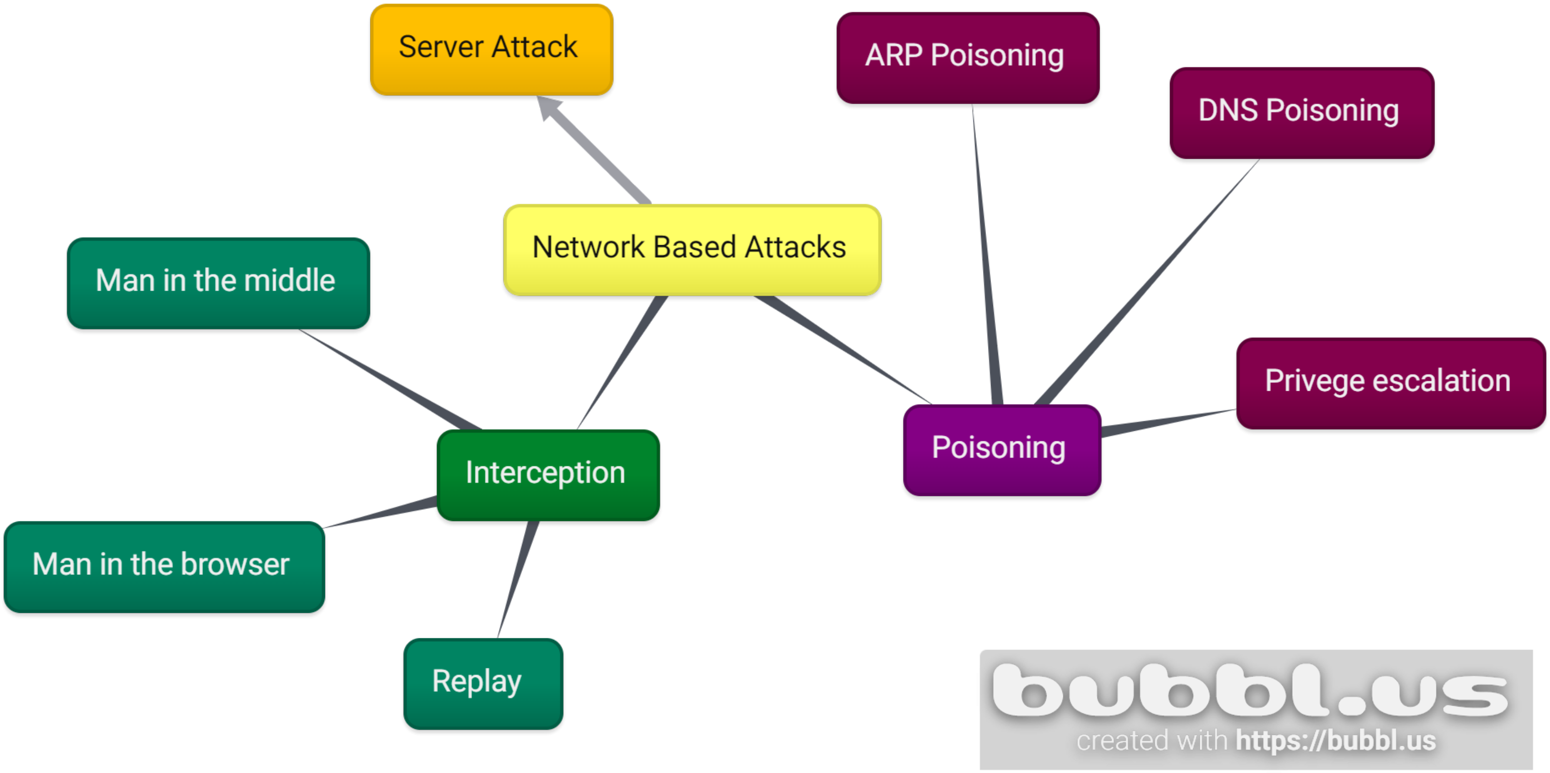
1. After a message has been written, a digest is generated
2. The digest is encrypted with the sender's private key
3. The message and the encrypted digest(now known as a digital signature) is sent to the receiver.
4. The receiver then decrypts the signature using his/her own public key, which reveals the digest.
5. If the message cannot be decrypted that means that the message was not sent by the sender, but by someone else.
6. The message is then hashed again with the same algorithm the sender used.

7. If the digest match then the message is authenticated, otherwise that means that the message has been tampered

Cryptoanalysis:

- It is the field of cryptography that deals with the reverse engineering of cryptographic algorithms.
- It involves tedious process of:
    1. Recognizing patterns
    2. Analysing unusual frequency patterns and meanings
    3. Deciphering a key to break the algorithm
    4. Searching and identifying weakness of a system
    5. Searching and identifying weakness of an algorithm.

## Chapter 4: Network based attacks:



- There are 3 main classifications of network based attacks:
    1. Interception attacks
    2. Poisoning attacks
    3. Server attacks

## Interception attacks:

**Definition:**

- Interception attacks are those attacks which focus on coming into the communication links between the sender(s) and receiver(s) in order to steal or extract information or data.

**What makes a network susceptible to Interception attacks?(NEEds to be revisited)**

1. A large network means that there are points of contact from where the attacks can launch the attack.

**TYPES:**

Man in the middle:

- It is a type of interception attack where an attacker positions himself/herself between the sending and receiving party.
- Both the parties do not know that the messages have been breached.

- The attacker is able to send forged messages to the sender, acting as the receiver. And vice versa

Man in the browser:

- It is a type of interception attack where an attacker positions an trojan between the web browser and the security mechanisms of the computer
- This is done to manipulate and steal information, mainly internet fund transactions

Working:

- A trojan infects a the systems target, which then install an extension to the browser,
- Now the virus waits, until the user visits a webpage that can be exploited to.
- After the user enters his account number, password, and anything of importance.
- As soon as the user hits submit the virus captures the data, and modifies the data(let's say if he captured your bank account detail, then he might modify it to wire some money to his own account.)
- When anything is sent back to the user, the virus modifies to make it seem like it was an ordinary interaction(for example if he wired some money to his account, the virus will modify the receipt to show you the amount that you entered and hide the actual amount.)

Replay Attacks:

- It is a variation of MITM attack, but it differs is that the attackers can save the contents of the message and reuse or the data to impersonate the sender.
- A simple replay would involve the MITM capturing logon credentials between the user's computer and the server. Once that session has ended, the MITM would attempt to log on and replay the captured user credentials.
- Reply attacks can be mitigated by using a key which has a limited lifespan and only the sender and receiver know.

Poisoning attacks:

**Definition:**

- Poisoning is the process of intentionally introducing a "substance" into a system inorder to extract information or harm the system.

**Types:**

ARP Poisoning:

- ARP stands for address resolution protocol, this is a protocol that is used to maintain a table also known as an ARP cache which contains the MAC address and its corresponding IP address.
- Whenever a new device is added, or data packets arrive at a network the gateway asks the ARP program for this table to be able to reroute the data to that particular device.

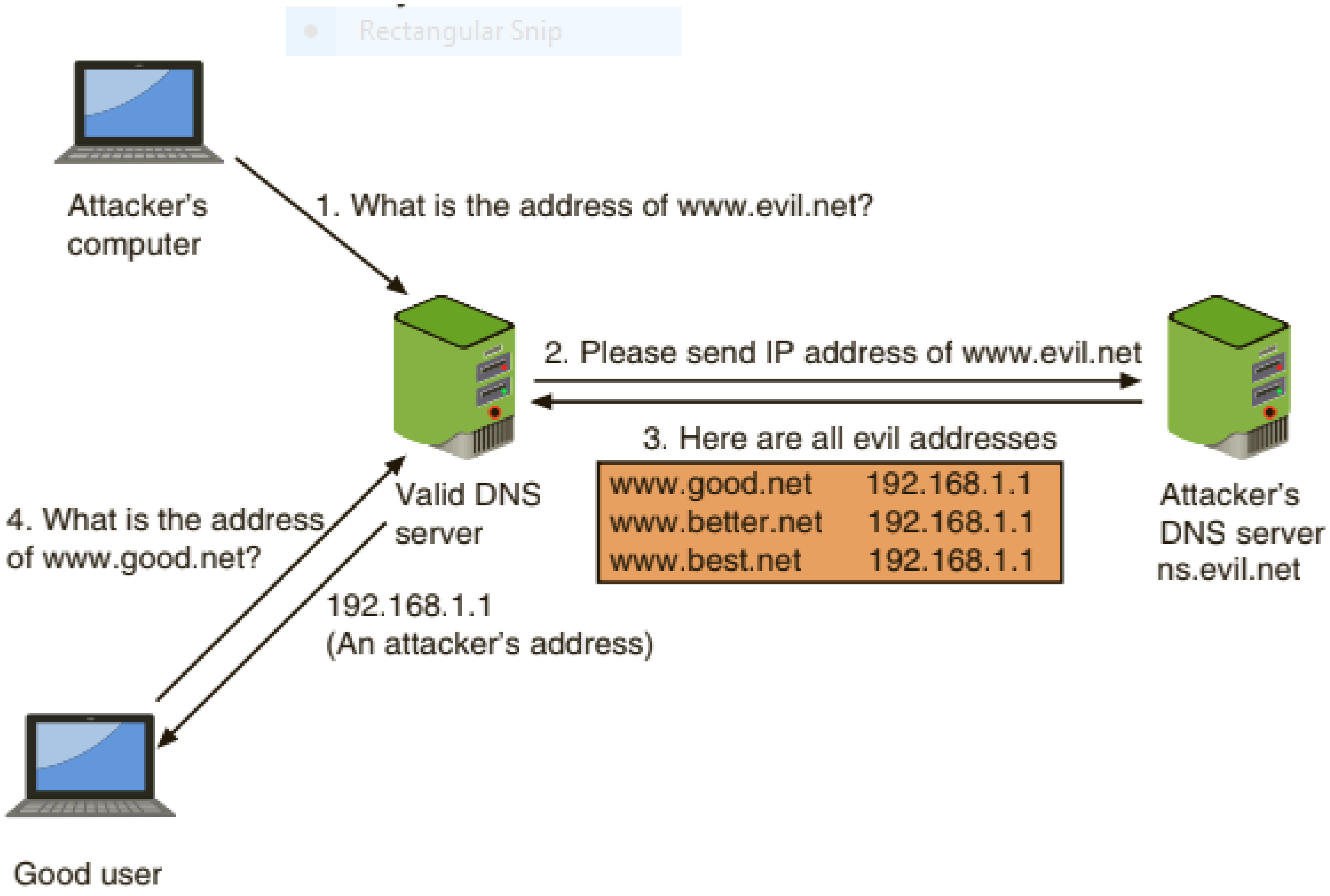
- In ARP poisoning a hacker successfully manipulates the ARP cache so that the victim's MAC address is replaced with his own, so that any messages sent to the victim will be first sent to the hacker.
- A hacker may be able to hijack the victim’s computer or deny him any service.
- He can transition this attack into a MITM.

DNS Poisoning:

- DNS poisoning is substituting another DNS address, so that whenever that address is sent the person is redirected to the fraudulent website.
- We will only study External.
- There are 2 ways for DNS poisoning; internal and external.
- In ARP poisoning we are able to take over the MAC address of the victim, but in DNS poisoning, hackers take advantage of a concept called zone transfer, which is where DNS servers exchange information to resolve addresses.
- The process is as follows:

1. First the attacker requests a valid DNS server to resolve(i.e. return an IP address) an address which is located on his own server.

2. To get this information, the DNS server asks the responsible server, which is the hackers DNS server, which contains all his own fraudulent IP addresses.
3. The servers perform a zone transfer, now all the fraudulent IP addresses are now in the DNS server ,and this server is now infected.
4. Any request to this server will result in spreading the hackers fraudulent websites all over the world.:)


Rectangular Snip
Attacker's computer
1. What is the address of www.evil.net?
2. Please send IP address of www.evil.net
3. Here are all evil addresses
Valid DNS server
Attacker's DNS server ns.evil.net
4. What is the address of www.good.net?
192.168.1.1 (An attacker's address)
Good user

| | |
|---|---|
| www.good.net | 192.168.1.1 |
| www.better.net | 192.168.1.1 |
| www.best.net | 192.168.1.1 |

**Figure 5-4** DNS server poisoning

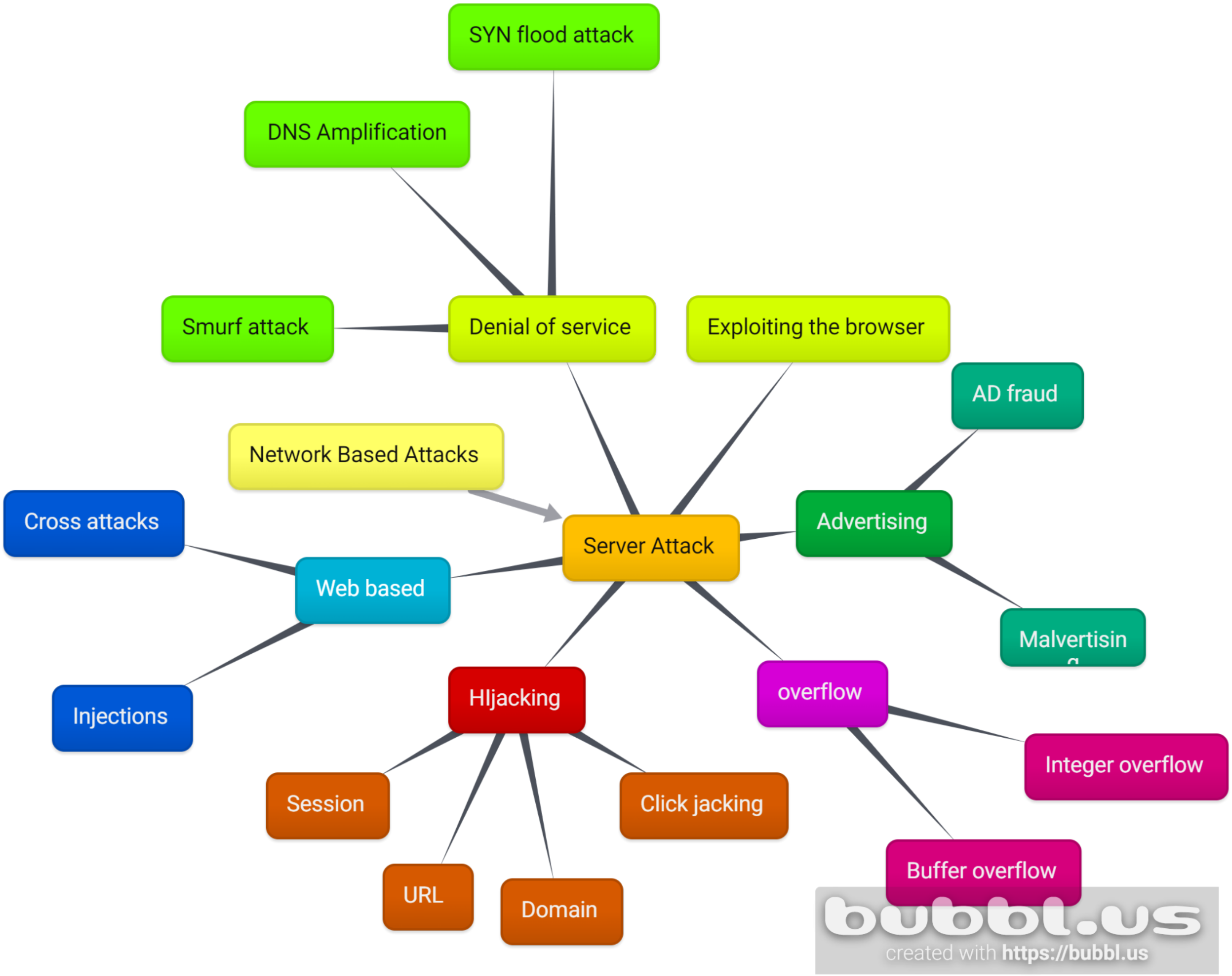
## Privilege Escalation:

- Privilege escalation is the exploitation of the software's vulnerability to access the resources that are normally restricted.
- There are 2 types: horizontal escalation and vertical escalation.
- Horizontal escalation is accessing the resources that have minimal differences, such as a senior and junior accountant. The senior account may have access to many things which the junior account may not have access to.

- Vertical escalation is accessing resources that someone higher in the hierarchy than you.
- Like a normal employee has access to the CEO’s access.

## Server attacks:

- There are many types of server attacks and there are many subtypes of these attacks.:



**Denial of service attacks:**

Definition:

- It is a type of attack where the attacker overwhelms the networks by sending a large number of requests.
- This prevents other users from accessing the site or the system.

Types:

1. Smurf attacks: in these attacks an attacker spoofs his own address with the victim's ip address(called ip spoofing) and then broadcasts a request to a network, which causes multiple computers to send hundreds of requests, which overwhelms the computer.
2. DNS Amplification Attacks: in these attacks the agent will send a DNS lookup request spoofed as the victims address. This will cause the server to send the data to the victim. This usually involves multiple servers and as and added bonus attackers can craft the lookup to send the entire table of data to the unsuspecting victim. A great example of this is:a malicious teenager calling a restaurant and saying "I'll have one of everything, please call me back and tell me my whole order." When the restaurant asks for a callback number, the number given is the targeted victim's phone number. The target then receives a call from the restaurant with a lot of information that they didn't request.

3. <u>SYN flood attack/ SYN Spoofing:</u> the hacker take advantage of the TCP/IP protocol, where to establish a connection the client first sends a SYN request, to which the server replies SYN + ACK and waits for the ACK from the device to communicate. During the SYN + ACK stage the server waits for some time to allow for slow devices to respond. But the threat agents use spoofed source IP addresses which don't exist or cannot respond, so the port remains open. In that time frame the hackers send more of these requests until the server cannot process the queries of the actual customers.

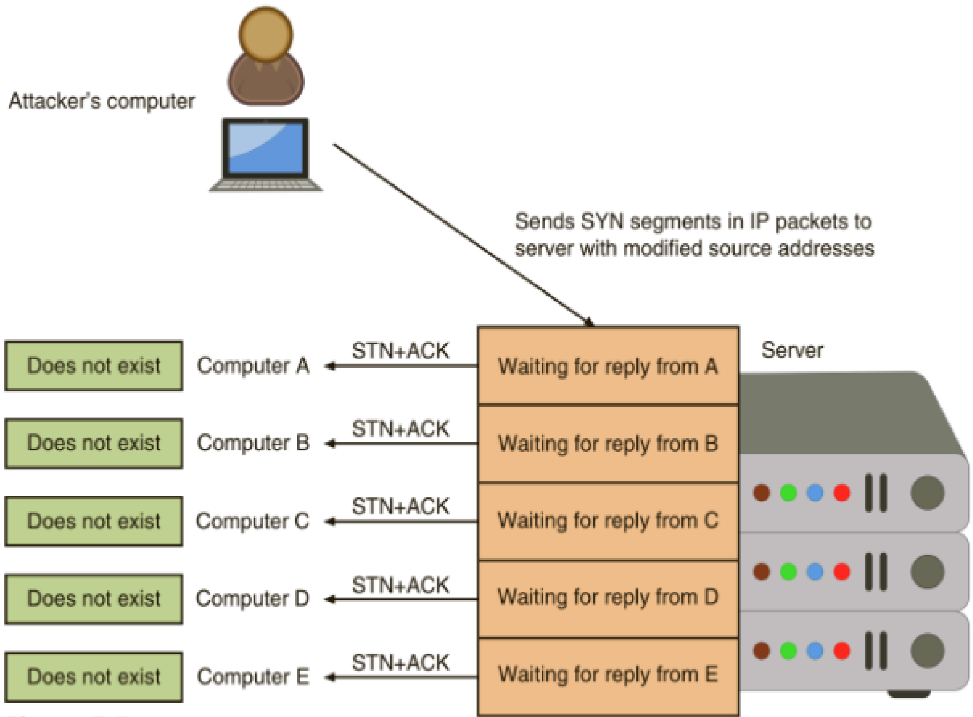


Figure 5-5 SYN flood attack

## <u>**Classic DoS:**</u>

<u>Definition:</u>

- This is the classic type of denial of service, where a single computer with a spoofed IP address attacks a server or a system.

- It has the same purpose as a DDoS, which is to limit the availability of some service or a resource offered by a system.
- There are 3 types in our course.

Types:

1. Ping flooding attack: the purpose of this attack is to overwhelm the network connection of the network to the target.The attack involves flooding the victim's network with request packets, knowing that the network will respond with an equal number of reply packets.
2. Source address spoofing:aka IP spoofing. Essentially attackers get their hands on a trusted ip address and use that ip address to fool the server into thinking that the attacker is from a trusted source. The attacker then generates large volumes of data packets, with the target system as the destination address. This causes congestion, which inturn causes the router to lower the capacity link.

**Application based bandwidth attacks:**

Definition:

- These attacks follow the principle of forcing the user's computer to perform such operations which will cause it to waste resources.
- There are 2 such types of attacks.

Types:

1. Session Initiation Protocol(SIP) flood attacks: SIP is a protocol that is used for video calling, and instant messaging over the internet. Attackers generate thousands of INVITEs which requires a lot of resources, this causes the proxy server to be overwhelmed by consuming resources to resolve the INVITE requests and by overloading the network's capacity.
2. Hyper Text Transfer Protocol(HTTP) flood attack: raiders and hijackers target servers by overwhelming servers with HTTP requests so that they are unable to respond to normal traffic.

Preventing DDoS/Dos:

1. Anti Spoofing software
2. Monitoring traffic and identifying abnormal traffic.
3. Deploy packet tracing
4. Have a contingency plan
5. Deploying CDNs, the content distribution network helps boots efficiency by using caches and finding the shortest path for the end user to get resources

**Web Server application attacks:**

Definition:

- These are the types of attacks that try to exploit web applications, which are much harder to secure as they are very dynamic.

- Attacks which exploit previously unknown vulnerabilities are known as Zero day attacks.

Types:

Coss-site attacks: There 2 main subtypes

1. Cross-site attacks(XSS):
    a. When a website takes an input from the user without authentication, it allows the attacker to store malicious instructions.
    b. XSS attacks requires a website that meets the following conditions:
        i. A website which takes an input from the user without authentication
        ii. The input is used in a response
2. Cross-site Forgery(XSRF):
    a. In this attack an attacker can use the user browser's settings to impersonate the user and use his credentials.
    b. This can be further improved by using social engineering to send legit looking email request and if the user has saved these passwords on the browser then using those credentials the attacker can take multiple actions such as:
        i. Transfers funds
        ii. Access Sensitive data and records
        iii. Impersonate a person on any website

Injection: Entering a specific or new input to exploit certain vulnerabilities in a system.

- One of the most common injection attacks include SQL injection attacks, which involve attacks using SQL to manipulate data in a database.
- Nowadays most forms have a username and password option.
- These are connected to a database.
- An assaulter ca check if the data is being properly filtered by entering an email address with a single quote at the end if an error like *unknown email address entered* is shown that means that the data is being filtered properly, but if an error like *server failure* is shown that means the data is not being filtered and it can be exploited.
- If the invader enters this command into the input:

SELECT fieldlist FROM table WHERE field w = = ‘ ’ hatever or a‘’ ‘’

- Then the attacker can see all the valid email address
- He can also look for specific people, delete the table and even look up sensitive information like phone numbers and names.

**<u>Hijacking:</u>**

<u>Definition:</u>

- An attack which uses illegally seized assets to utilise for attacks

Types:

1. Session Hijacking:
    a. Whenever a user logs in to a secure web application, the user receives a session token, it is a string of random integers and characters.
    b. If a hijacker is able to steal this token he/she will be able to impersonate the user.
    c. These tokens can be stolen with the previously mentioned techniques.
2. URL hijacking:
    a. If you ever enter an incorrect URL you would have been show a 404 error, but nowadays people have created fake and bogus website so that if a URL has an incorrect domain or a misspelling then they will be redirected to their own website to generate revenue and steal information
    b. Cyber criminals are also buying domains of the website(aka bitsquatting).
    c. Because giant or international websites have many instances of a domain in a DNS server.
    d. The best way to understand this is that, if even a single 1 is in an incorrect position, the person will be redirected to a different website. So for example if changing a single bit into the binary of g will convert it to f, the criminal will buy the domain foogle.com, as people will visit their site.

3. Domain Hijacking:
    a. It occurs when a hacker is able to change the domain pointer, which links a domain name to a specific web server, is changed
    b. So whenever a person visits the website it will automatically redirect him to a different physical web server.
4. Clickjacking:
    a. Hijacking the click of a mouse is called clickjacking.
    b. For example a felon might have a big button serving some function, but when clicked it has a hidden function(for example to transfer some funds to the felon bank account or something like that).
    c. This is done with help of zero pixel IFrame, which allows 2 html links to be placed inside one another, they are invisible to the naked eye.

**Overflow:**

Definition:

- They are the attacks they focus on storing values greater than the intended storage capacity to have an overflow occur.
- There are 2 main types: buffer overflow and integer overflow.

Types:

1. Buffer overflow:
    a. A buffer overflow attack occurs when a process attempts to store data in RAM beyond the boundaries of a fixed-length storage buffer.
    b. Usually these buffers contain the return address memory location, these can be overridden with the attackers malware code, so that when the program tries to run it instead runs the malware of the hacker :(.

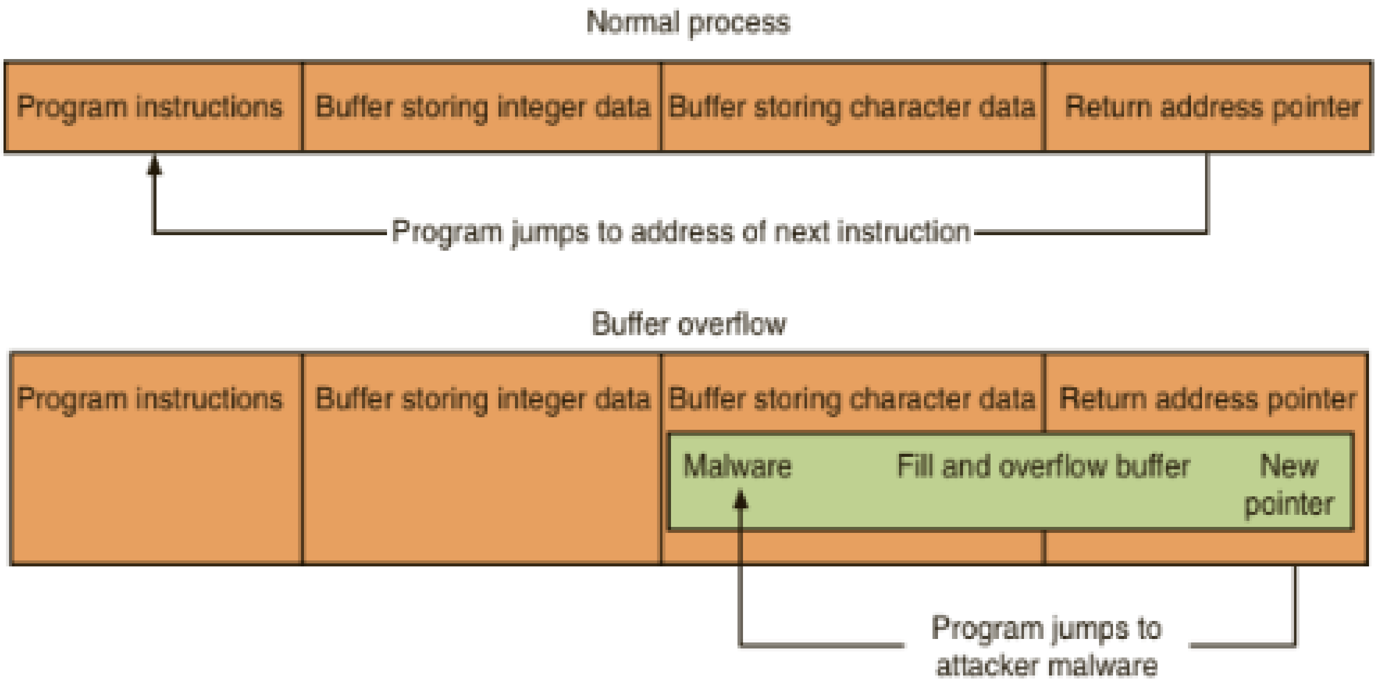

Normal process
Program instructions
Buffer storing integer data
Buffer storing character data
Return address pointer
Program jumps to address of next instruction
Buffer overflow
Program instructions
Buffer storing integer data
Buffer storing character data
Return address pointer
Malware
Fill and overflow buffer
New pointer
Program jumps to attacker malware

Figure 5-12 Buffer overflow attack

2. Integer overflow:

    a. an integer overflow attack, an attacker changes the value of a variable to something outside the range that the programmer had intended by using an integer overflow.
    b. This can have many problems for example if a in supermarket an integer overflow would actually show a refund to the customer instead of calculating the correct amount of money.
    c. An extreme example of an integer overflow attack would be withdrawing $1 from an account that has a balance of 0, which could cause a new balance of $4,294,967,295!

**<u>Advertising Attacks:</u>**

<u>Definition:</u>

- These are attacks that rely on advertising to trick users into seeing and or clicking their malware infected websites.
- There 2 main types:Malvertising and ad fraud.

<u>Types:</u>

1. Malvertising:
    a. This involves using reputable companies to infect them with malware infested ads and when the user clicks on anything they will show hundreds of them.
    b. These will usually to downloading ransomware and trojans
    c. Preventing malvertising is difficult because:
        i. Website operators are unaware of the types of ads that are being displayed.

ii. Users have a false sense of security going to a “mainstream” website.

iii. Turning off ads that support plug-ins such as Adobe Flash often disrupts the user’s web experience

2. Ad fraud:

a. When we watch pre rolls(10-15 second video used for promotion) the advertiser who will display their ad bid against each other, the one who offers the highest bid is offered the slot,

b. The ad is then directly sent from the advertiser and is received by the browser.

c. The browser is also responsible for verifying that the user has actually watched the ad.

d. Thieves have created a bot that is able to spoof all the necessary interactions needed to initiate, carry out, and complete the ad auction. This bot is called Methbot

e. Methbot contacts the advertiser, acting like the website, asking them to display their ads.

f. The company adds them to the auction list, after the auction process Methbot acts like the browser who verifies that the user has actually watched the ads, but in reality they are played by spoofed IP addresses and they trick the advertisers into giving the creators of methbot fake views.

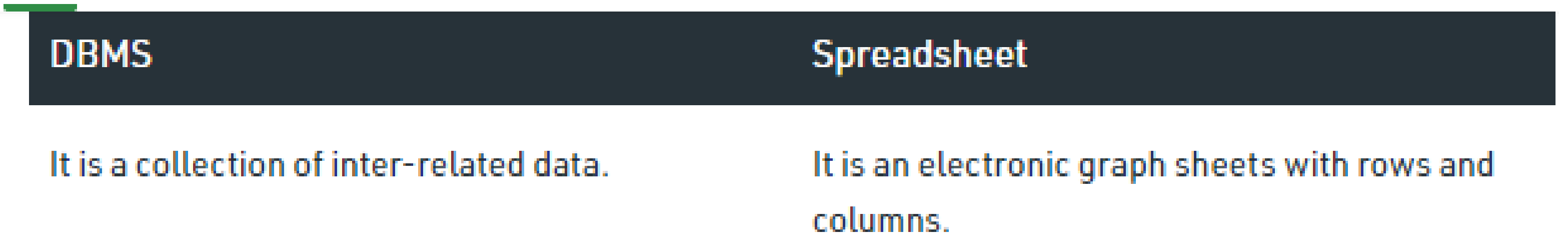
g. This way nearly $5 million dollars are stolen a day.

# Chapter 5: Database security

## Introduction Databases:

### Terminology and Definitions:

- A database is a collection of data organised in rows and columns, based on certain rules.
- Horizontal rows are referred to as a record, and vertical columns are referred to as fields or elements.
- The name of each column is referred to as an attribute, a set of columns is a relation.
- These rules allow the database administrator to control the type and format of the data that is entered.
- If a user wants to interact with the database then he/she does it through the help of a DBMS or a database management system.
- The logical structure of the database is called Schema, users only have access to a small portion of the schema called a sub schema.
- To perform any operation in a database we use a query, this allows us to modify, delete, add, or retrieve a field or a record from the database.
- We can also add conditions to our query to make mass changes, the most common database query language is SQL.
- Here is the difference between a DBMS and spreadsheet



| DBMS | Spreadsheet |
|---|---|
| It is a collection of inter-related data. | It is an electronic graph sheets with rows and columns. |

Advantages of databases over spreadsheet systems:

1. Shared access: Many users can access the system.
2. Controlled access: Only authentic users can access the system to modify it, others can only access it.
3. Minimal redundancy: There is no repetition of data because all the users do not need to maintain their own tables.

4. Data consistency: Any changes made in the central database will be available for all users, as they all share the same access.
5. Data integrity: A DBMS keeps the data integrity by preventing any accidental deletion ,malicious and undesirable changes.

**Security Requirements of databases:**

Why do we need to secure Databases:

- Databases are used to store an array of sensitive data that includes storing the name, location and power of an ICBM, intercontinental ballistic missile, the most wanted criminals in the country and region and so on.
- To protect the Confidentiality of such data it is important to give the right people the right amount of access in a database.
- Databases also need to be able to maintain their integrity and the integrity of the information stored, as even a single error can have consequences.
    - Databases also need to be able to maintain the integrity of relationships in databases as they never operate by themselves, which means that the DB can be subjected to inference attacks
- DBs also need to be available to the system's users to maximise usage and efficiency, but this comes at a cost: too many users can become a liability and a very small amount of DBm team can create bottlenecks in the working of the system.

- To Maintain the CIA of a DB a database management system or DBMS is deployed. Traditionally it has 2 parts: the front end, which everyone can access and the back end which only the Database administrators can access.
- A good DBMS has many security requirement, these include:

1. <u>Physical Database integrity:</u>
    - This means that the physical components of the DBMS are also safe from any intention and unintentional damage.
    - These damages range from power failures, human errors, SQL injection attacks, and many others.
    - These include all the servers, hubs, computers and any other physical systems the DBMS is connected to.
2. <u>Logical Database integrity:</u>
    - Any changes made to the database must be approved by the DBM or from the authentic users.
    - If an element is unreadable or the data is damaged in some way the DBMS will lose its integrity.
    - The integrity is the responsibility of the DBMS and its manager.
    - One way to maintain the integrity is backing up the file up to 3 times.
3. <u>Element integrity:</u>
    - An element maintains its integrity by maintaining accurate data, which must be done by authentic users.

- A DBMS can facilitate this operation by:
    1. Maintaining validation checks for each field.
    2. Maintaining access control
    3. Maintaining a changelog for the database

4. <u>Auditability:</u>
    - It is the name given to the process by which a report is generated which contains all the changes that have been made to the database.
    - It also contains who made the changes as well as when it was made.
    - Another way a DBMS maintains its auditability is by displaying sensitive or protected information incrementally.
    - This means that to access protected information the user has to take many steps to access the data.
    - This audit trail allows administrators to identify the users who have done these steps and stop them.

5. <u>Access control:</u>
    - A good DBMS must have a good DBm who grants access to only authentic users
    - This means that not everyone will be able to access all the data in DB.

    - It is a DBMS job to only display what is appropriate to each appropriate individual who has the specified level of access.

6. User authentication:
    - A good DBMS must be able to differentiate between a normal user and a malicious one.
    - This means rigorous user identification may be needed.
    - Because a DBMS runs an application on top of the OS, the system must verify each individual user. As there are no trusted paths from the application to DBMS.
7. Availability:
    - A DBMS is a hybrid system where it is run as a system and a program.
    - This means that  a majority of the users run the application of the system, which means that a DBMS must be able to cater to its users' requests.
    - This means that the data must also be displayed according to the user's authentication level.
    - There may also be cases where due to the maintenance of the DB a user may not be able to access the database or see certain information.

**Reliability and OS protection and CIA controls:**

1. A reliable DB preserves the integrity of the elements and the database as a whole, and the accuracy of the elements.

2. Some basic protection offered by the OS for DBs:
    - File backups
    - Access controls
    - Integrity checks
3. Some CIA controls for a DB include:
    - Two-phase commit protocols for updates:
        - This is a system where to update anything in the DB you need to first make a change in the local DB, which then updates the commit flag. This phase is called the First phase or the intent phase. After this there is no turning back
        - After a certain time period the DBMS updates the new changes, this is called the permanent change phase.
    - Redundancy/internal consistency controls:
        - DBMS use many error detection techniques such as: Hamming codes, parity bits and cyclic redundancy.
        - Shadow fields are also used, which are copies of the fields, and are used temporarily to display data when an error occurs.
    - DB recovery: DBMS maintain the log of user access and when they were made. Many DBMS also maintain backup safes in isolated spaces and are updated periodically.
    - Concurrency/consistency control:

**Database Disclosure:**

Definition:

- Data disclosure is the voluntary sharing of any and all information that is considered relevant to a given situation.
- This means that information can either be so sensitive that only a handful people know about it or it can be public knowledge.
- If the data is partially sensitive then the DBMS must have varying levels of access for each and everyone of its users.
- If the complete data is sensitive then everyone of the users must be verified to be able to access the Database.

Factors that decide if the data is sensitive or not:

1. The value itself or the nature of the data in sensitive such as the ICBM type, the rarest animal on the continent and so on.
2. The source where the data is from is sensitive such as a working ISI field agent, or files left by a criminal.
3. Sometimes it becomes the case where previous declared data now becomes sensitive. Such as uncovering the identity of a civilian who was murder,so his entire will be declared sensitive informatin by the DMA.
4. We mentioned relational databases, and it is sometimes the case that due to data being sensitive in 1 table, any data related to it also becomes sensitive. For example the serial killer, after he is discovered his entire family data will become sensitive and the authorities will also go after them as well.

Types of data that is classified as sensitive for disclosure:

1. Exact data: Names, dates, years…
2. Range value: float values, mass, temperature…
3. Negative predictive value: values that exist in the table that do not satisfy some conditions
4. Existence: for example checking if a name exists on a hit list.
5. Probability: using formulas to check if a record exists on the basis of probability.

Types of disclosure attacks:

- There are many ways people can manipulate multiple parts of the system to get sensitive information
- Other than the techniques we learned in previous there is a an that is used to derive sensitive information from publicly available information, this is called inference attacks.
- These attacks can be done on human components as well as non human components.

- There are 2 types of inference attacks:
  - Direct attack: which means directly asking or generating a query which will reveal the sensitive data. This means sql injection and interrogation are extensively used
  - In direct Attacks: indirect attacks are the opposite of direct attacks where the raider uses formulas, tracers, and deduction to infer the sensitive data. For example a person might tailgate the target or a generate multiple queries so

that the Database reveals some sort of sensitive information.

Ways of preventing Data disclosure:

1. Suppression of information: suppress access of all users, all information, access time and put more rules restricting access to the database in general, this is the easiest out of the 3.
2. Tracking user disclosure: Essentially monitoring the users who have access to sensitive information. This is the most costly method, but it also has the highest rate of preventing disclosure.
3. Disguising or manipulating the data: this means that the data is stored in such a way that is not traditional and randomly permuting the data in such a way that the attacker cannot get the data or get inconsistent information.

## Chapter 6: Intrusion detection system

### Introduction:

- We have studied many ways of infiltrating a system and many easy ways of exploiting the system. We will now learn one way of protecting ourselves from such attacks.
- If you ever go to a back you will observe many security measures: the security guard, the metal detector, and sometimes even a full body scanner.

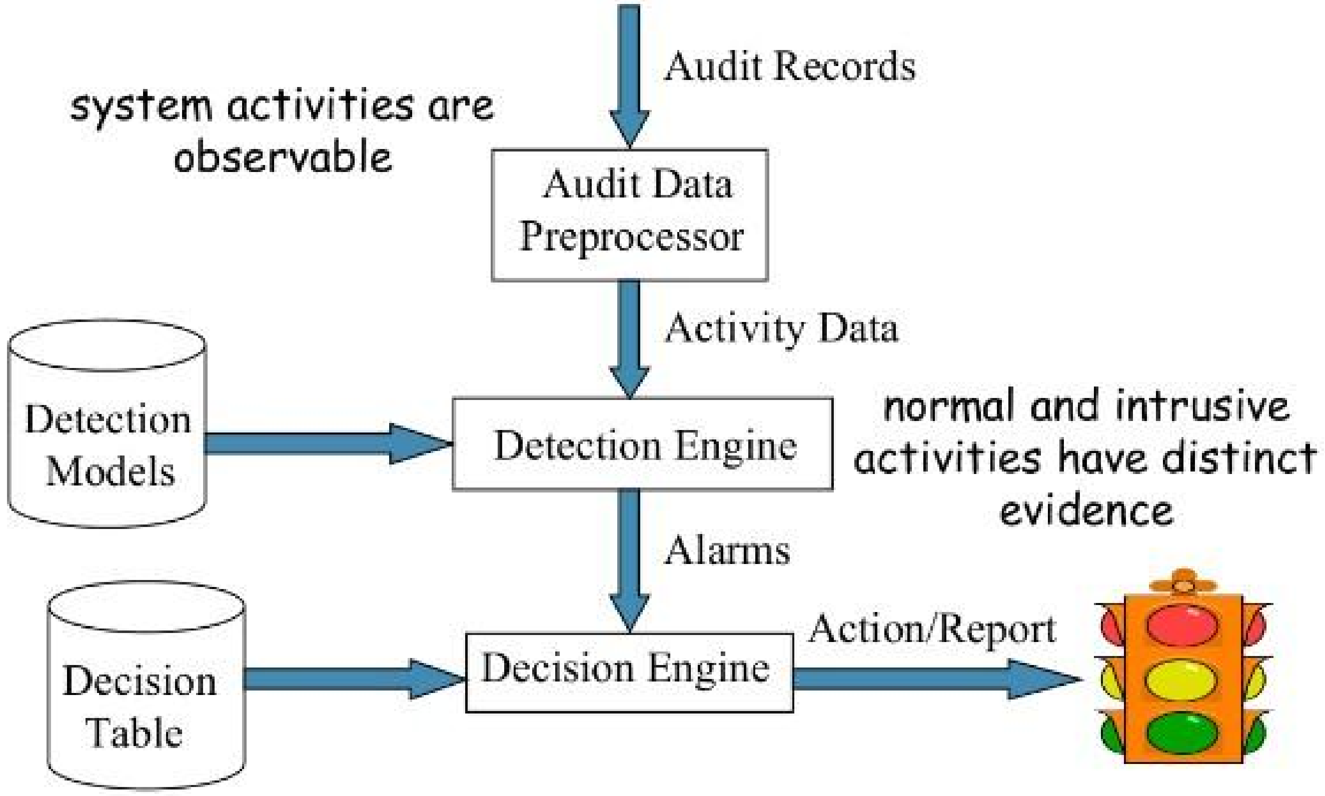
- These measures exist to catch the intruder before he is able to attack the bank or becomes a threat to the bank.
- These types of systems are called intruder detection systems or IDS.
- Intrusion is defined as a set of actions aimed to compromise or exploit resources for a system.
- Intrusion detection is the process of identifying and to intrusion activities.
- intrusion prevention is the process of ID but with exercises to control and protect computers from exploitation.
- The main reason people break into a computer system is:
    - Access information
    - Manipulate information
    - Exploit and render the resources of a system useless

## Functions and the working of IDS

- IDS can be hardware or software in nature, from a security guard or dog to a firewall.
- They are a monitoring type system where they detect and report but do not take protective measures, usually referred to as a reactive system not a pro active.
- They utilise many tools, methods and resources to identify , assess and report suspicious activity.

- Functions of the IDS include:
  - Monitor the functioning of routers, firewalls, server and files
  - Support users
  - Arranging logs and audit trail
  - Generate alarm when breaches are detected.
  - Blocking servers with suspicious activity
- The most popular way to detect intrusions is with the help of audit data generated by the system.
- When this data is arranged in a chronological order it is referred to as an audit trail.
- These help identify guilt attackers.
- There are 2 assumption we make to install an IDS:
  - The system is observable
  - There is a clear distinction between normal and intrusive activities.

- There are 3 main components of IDS:
  - Features: theses gather information and evidence, imagine a camera
  - Models : these join or link the information together. Imagine an AI which matches the images.
  - Various components: Audit processor, knowledge, decision engine, alarm generation, responses.



## Types of IDS:

- There are 2 broad categories: modelling and deployment,
- Modelling has 2 types: Signature and anomaly based
- Deployment has 2 types:
  - Network based: which monitors the network traffic
  - Host based: Which monitors Computer process.
- There are 2 types of responses;
  - Passive: these are the most common way to deal with intrusion, for example logging, Notifications and shunning.
  - Active: these involve taking steps that would actively prevent the user from doing any harm.

## **Modelling: Signature based IDS(pattern based)**

- This type uses pattern matching to identify intruders.
- They also use rules to help them identify attackers.
- They use CERTs or by analysing attack scripts from the internet.
- It can only identify those attacks which it has been trained for.
- Thus it cannot detect new or advanced attacks.
- It also requires frequent updates.

## **Modelling: Anomaly based IDS(threshold or value based)**

- This type of modelling uses thresholds to determine whether or not the person is an intruder.
- The system maintains a table of normal data from which it derives the “normal” data and anything that falls out of the range is considered an anomaly.
- It can easily detect many new and old attacks
- It struggles with false positives from abnormal normal use activity.
- A Lot of the training data does not contain realistic attack patterns.
- Attackers can incrementally increase the default values to such a level such that anomalies are registered as normal activity.
- If attackers learn about the ranges they can attack within those ranges to avoid detection, what is essentially camouflage.

- They can also be triggered by element faults or non user related components.
- These can be very costly as administrators have to locate and examine them.
- It uses counters, gauges, and resource utilisation as indicators to see if the person is an intruder.

**Living example Snot IDS:**

- Snort is a well-known and currently industry-leading tool used for packet sniffing, logging, and intrusion detection
- It was created by Cisco and can be installed on Windows as well as a few Linux distributions
- It is an anomaly Based IDS.
- It has 4 major parts:

Packet decoder: when data is sent over the network many headers are attached to it so that it is able to reach the correct location. The decoder removes these headers to access the packet.

Detection engine: the heart of the SNORT it uses predefined rules to evaluate each packet of data.

Logger: whenever an anomaly is detected or the logging option is selected the logger records the information and stores it for future use.

Alerter: if any anomaly is detected the alerter then issus an alert to the respective administrators.

## **Chapter 7: Firewalls**

### Introduction:

- We mentioned the use of firewalls in the previous chapter for an IDS, but what is a firewall?
- It is essentially a security component that is used to protect networks from external threats.
- Traditionally a firewall is run on a dedicated device to maximise efficiency.

### Design and implementation:

- A firewall implements the rules and regulation it has been configured with.
- All firewalls come with some built in protection, but they need to be configured.

- One of 3 action can take place when a packet reaches a firewall:
  - The packet can be accepted and may pass through the firewall.
  - The packet may be dropped and not allowed through the firewall.
  - The packet may be Rejected and the source of the packet will be informed why it was rejected.

- Some of the policies that the firewall use are:
  - TCP
  - UDP
  - IP address of source and the destination
  - Ports of source and the destination
  - Application payload

## Types of firewalls:

- There are 4 main types of firewalls:

### Packet filtering:

- The most basic type of firewall which either accepts the packet into the network or denies entry on the basis of rules set by the administrator.
- This also includes traffic going from the network.
- These set can be based on:
  1. SOurce IP address
  2. Destination IP address
  3. Source and destination transport level address
  4. IP protocol
- Advantages:
  1. Cheap
  2. Low resources usage
  3. Best for small network
- Disadvantages
  1. Vulnerable to spoofing

    2. Can only work on the network layer
- The firewall can only perform 2 actions on a data packet: accept it or reject it.
- These firewalls cannot “see” the message and do not maintain log of such packets
- Also known as stateless firewalls.

**Stateful inspection:**

- They perform all the functions but they also maintain tables of the contents of the packets.
- These tables maintain information related but not limited to: active connection, including the IP addresses, ports,and sequence numbers of packets.
- They also make their judgement on the basis of multiple data packets
- Advantages:
    1. Very effective in filtering.
    2. More intelligent systems can make more complex decision
    3. They have a wider logging capacity and can mitigate stronger attacks
    4. Needs less ports for processing.

**Application proxy:**

- A proxy firewall is the most secure form of firewall, which filters messages at the application layer to protect network resources. A proxy firewall, also known as an application firewall or a gateway firewall, limits the applications that a network can support, which increases security levels but can affect functionality and speed.

- The proxy firewall can easily monitor and audit the traffic.
- The user contacts the gateway using a TCP/IP application, such as Telnet or FTP, and the gateway asks the user for the name of the remote host to be accessed.
- When the user responds and provides a valid user ID and authentication information, the gateway contacts the application on the remote host and relays TCP segments containing the application data between
- Advantages:
    1. Application gateways tend to be more secure than packet filters
    2. Easy to log and audit incoming traffic

- Disadvantages:
    1. Transmission speed gets reduced due to additional processing
    2. The applications that support these features are few.

**Circuit level gateway:**

- Just like application proxy firewalls which monitor the data between application layers, circuit level gateways monitor the data between the transport layer.

- Another name for these type of firewalls is proxy firewall or servers
- They also provide security for connections like TCP , UDP and the three way handshake protocol.
- They can also monitor the connections and also maintain a log of users.
- Advantages:
    1. Hides private network data, does not require a separate proxy server for each application
    2. Easy to install
- Disadvantages:
    1. Does Not filter individual packets, which it is like a stateless firewall, where it doesn't know the contents of the packets

### Packet filtering vs Proxy server

Advantages:

| Packet filtering | Proxy server/ firewalls |
|---|---|
| simple | User authentication is possible |
| | Application protocol control can be implemented |

|  |  |
|---|---|
| Relatively cheap to implement | Logging |
|  | Detailed audit trails can be tracked |

Disadvantages:

| Packet filtering | Proxy server/ firewalls |
|---|---|
| Correctly identifying packet is difficult and is prone to error | It is expensive as each application needs a proxy |
| Reordering packets is a difficult process and makes the specifying the rules very difficult | Circuit level proxies are cheaper but do not scan application level data |

**Honeypots:**

- Honey pots are not firewalls but they are fake networks setup so that they misdirect the attackers attention from the actual server.
- Honeypots do not hold any production value to the network.
- Whenever an attacker takes the bait the server administrators can observe and study these patterns and can record them as material to train and teach IDSs.
- There are many places where honepots can be places some of them include:
    1. A honeypot outside the external firewall is useful for tracking attempts to connect to unused IP addresses within

the scope of the network. A honeypot at this location does not increase the risk for the internal network. The danger of having a compromised system behind the firewall is avoided.

2. The network of externally available services, such as Web and mail, often called the DMZ (demilitarised zone), is another candidate for locating a honeypot The security administrator must assure that the other systems in the DMZ are secure against any activity generated by the honeypot.
3. A fully internal honeypot (location 3) has several advantages. Its most important advantage is that it can catch internal attacks. A honeypot at this location can also detect a misconfigured firewall that forwards impermissible traffic from the Internet to the internal network.

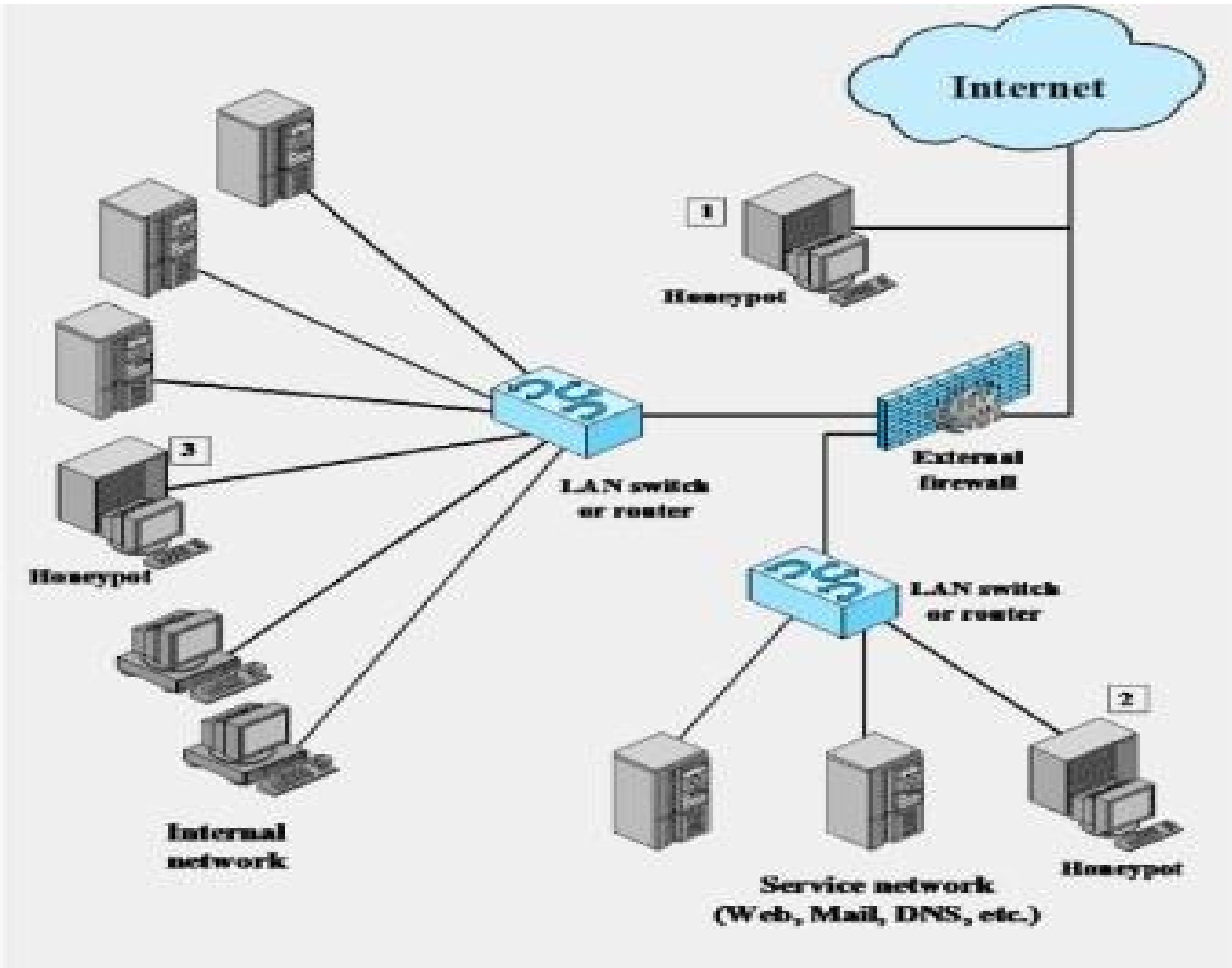


Figure 8.8 Example of Honeypot Deployment

- There are 2 broad classifications of Honey pots:

Low interaction:

- These are ghost network with just enough components and features to fool an attacker that this is a real network

High interaction:

- These are fully functional networks with real resources and working components
- These are very realistic

- If these are compromised, this may lead to other parts of the network to be exposed.

**<u>Virtual private networks(VPNs):</u>**

<u>Definition:</u>

- They are essentially connections that allow private networks to to safely extend over long physical distances via public networks as a means of transmission.
- VPN provides guarantees of data confidentiality, integrity, and authentication, despite the use of an untrusted network for transmission.
- There are 2 broad classification of VPNs:
  - Remote access VPN
  - Site to site VPN

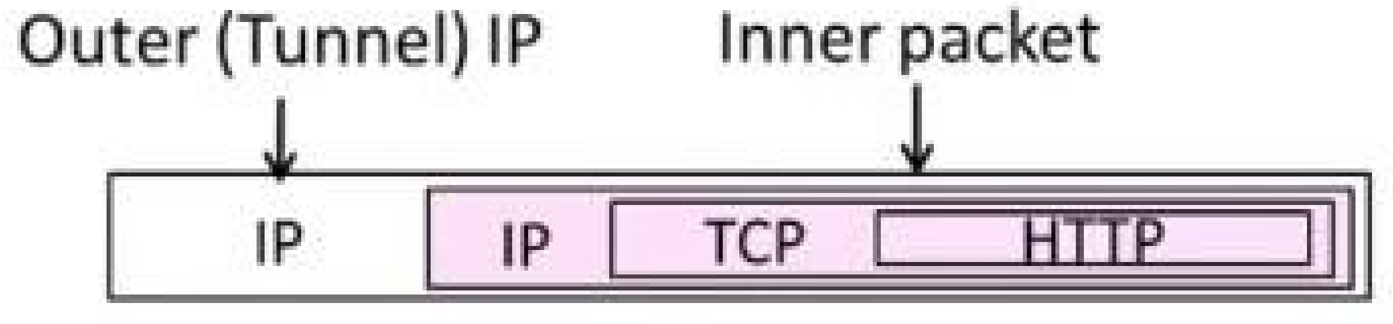
<u>Reasons why we need VPNs</u>

1. Often desirable to separate network from the Internet, e.g., a company
2. Use the public Internet instead of leased lines – cheaper!
3. Prevent eavesdropping and tampering with messages

<u>Tunnelling</u>

- Tunnelling is a process by which we are able to create virtual links across the internet.
- This allows us to set up VPNs to send our network traffic.
- These tunnels also mask out identities from the ISPs,governments, authorities, hackers, and other third parties.

- Virtual tunnels are set up by encapsulation of the IP packets in the IP address of the tunnel, by modifying the IP header for delivery to remote endpoints. This is the result:



- But these alone are not enough. Next, we need cyptography, we will see IPSEC or IP security to secure VPN tunnels.

## Chapter 8: Internet security protocols and standards

- A large amount of data is transmitted over the internet, it is very important that the data is secured.
- To do this we have 3 main Protocols:
    - MIME and S/MIME
    - Secure security Layer (SSL)
    - Transport Layer security (TLS)

### MIME and S/MIME

**Introduction:**

- Stands for Multipurpose Internet Mail extension and Secure Multipurpose Internet Mail extension.
- These protocols are used to secure mainly email messages

over the internet.

- S/MIME functionality is built into the majority of modern email software and interoperates between them.
- S/MIME is an extension of MIME, it allows the user to sign, encrypt and decrypt messages.
- There are 4 function that are supported:
    1. Enveloped data: Consists of encrypted content of any type and encrypted-content encryption keys for one or more recipients.
    2. Signed data: A digital signature is formed by taking the message digest of the content to be signed, then encrypting that with the private key of the signer. The content plus signature are then encoded using base64 encoding. A signed data message can only be viewed by a recipient with S/MIME capability.
    3. Clear-signed data: As with signed data, a digital signature of the content is formed. However, in this case, only the digital signature is encoded using base64. As a result, recipients without S/MIME capability can view the message content, although they cannot verify the signature
    4. Signed and enveloped data: Signed-only and encrypted-only entities may be nested, so encrypted data may be signed, and signed data or clear-signed data may be encrypted.

**<u>Working:</u>**

- S/MIME provides 2 security features:
    1. Digital signatures
    2. Allows for message encryption.

<u>Sending messages:</u>

- To send encrypted messages over the internet these step are followed:

1. generates random symmetric private key, KS
2. She encrypts the message with KS (for efficiency), she also encrypts KS with Bob's public key
3. She then sends both KS(m) and KB(KS) to Bob
4. Bob uses his private key to decrypt and recover KS, and uses KS to decrypt KS(m) to recover m

<u>Check sender authentication or signing and clear singed data:</u>

1. Take the message you want to send and map it into a fixed-length code of 256 bits using SHA-256.The 256-bit message digest is unique for this message making it virtually impossible for someone to alter this message or substitute another message and still come up with the same digest

2. S/MIME encrypts the digest using RSA and the sender's private RSA key
3. The result is the digital signature, which is attached to the

message. Now, anyone who gets the message can recompute the message digest then decrypt the signature using RSA and the sender’s public RSA key

4. Since this operation only involves encrypting and decrypting a 256-bit block, it takes up little time.

Enveloped Data:

- 2 algorithms are used to encrypt and decrypt S/MIME: RSA and AES.

1. First a random key is generated by S/MIME, each message will have its one unique key and it is sent with the message.
2. This key is used as a public key for the encryption algorithm, RSA,which encrypts the key with the recipient’s public RSA key
3. On the receiving end, S/MIME uses the receiver's private RSA key to recover the secret key, then uses the secret key and AES to recover the plaintext message
4. If no encryption is used then Radix-64 is used to convert the cipher text into ASCII format.

- Note Radix-64 is an algorithm that is similar to base 64 encoding, which means that it converts the characters into 64 bit format.

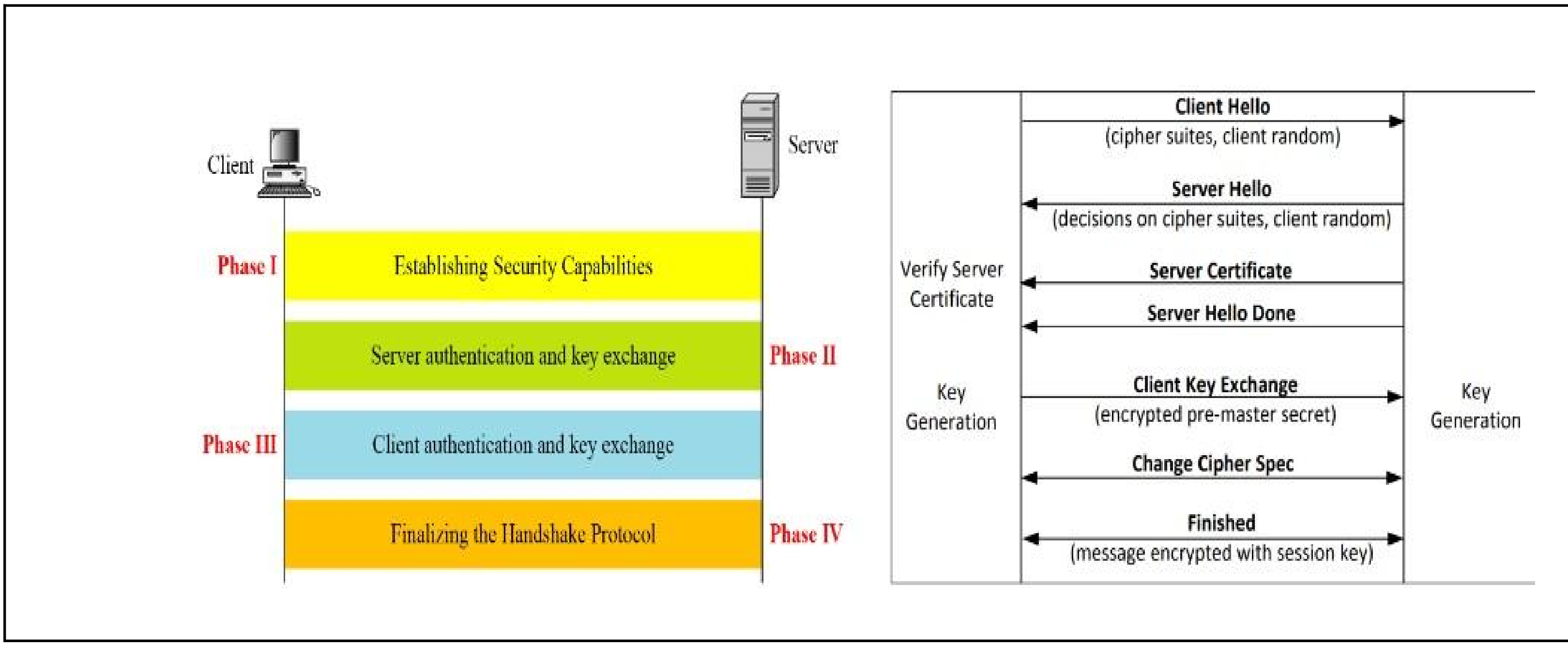
## Secure Socket Layer(SSL):

## **Introduction:**

- It is used to provide security between the web browser and the server.
- It provides privacy(Confidentiality) and Data integrity which is a part of Digital signing
- It encrypts the link between the server and browser a
- There are 4 main protocols concerning our course: handshake, record, alert, change cipher.

## **Handshake Protocol**

- Before the communication between sender and browser can begin a process called handshaking takes place.
- SSL is a protocol that utilises asymmetric encryption, which allows the browser to verify the web server and get the public key to establish a secure connection.
- There are 4 steps to this process:
    1. First the browser sends a message to the server to connect.(Phase 1)
    2. The server replies to the message, by sending the SSL certificate.(Phase 2)
    3. The browser now sends its public key to the server.(Phase 3)
    4. The connection is now finalised.(Phase 4)


Client
Server
Phase I
Establishing Security Capabilities
Server authentication and key exchange
Phase II
Phase III
Client authentication and key exchange
Finalizing the Handshake Protocol
Phase IV
Client Hello
(cipher suites, client random)
Server Hello
(decisions on cipher suites, client random)
Verify Server Certificate
Server Certificate
Server Hello Done
Key Generation
Client Key Exchange
(encrypted pre-master secret)
Key Generation
Change Cipher Spec
Finished
(message encrypted with session key)

## Transport Layer Security(TLS):

**Introduction:**

- It is a security protocol which provides data integrity and privacy communications.
- It is an advanced version of SSL protocols
- The TLS provides a secure channel between the communication of 2 applications .
- There are 3 main features of theses channels:
  1. Confidentiality: No one other than the sender and receiver can see the actual data.
  2. Integrity: the channel can detect any changes made to the data
  3. Authentication: atleast channel needs to be authenticated.

- There are many security flaws with SSL, these were patched by the first update of TLS
- TLS also adjusts to address vulnerabilities and to improve implementation and integration capabilities.
- It Also ensures that no third party can tamper with the communication in the channel.

## **TLS Connection and Session:**

### Session:

- It is the name given to the link that is created between server and browser due to the handshake protocol.
- A session can have or share multiple connections.
- It is defined by a set of security parameters.
- The main benefit of this is that you can avoid time consuming security measures, by avoid i mean that u bypass the security parameters, but u only need to perform them once and then use them to re login every time, think of when google saves your password so u dont have to remember them.

### Connection:

- A TLS connection is a temporary connection.
- Each connection can only be connected to 1 session.

### Difference between Session and connection

- The difference between connection and session is that connection is a live communication channel, and session is a set of negotiated cryptography parameters.

- You can close connection, but keep the session, even store it to disk, and subsequently resume it using another connection, may be in completely different process, or even after system reboot (of course, stored session should be kept both on the client and on the server).
- For example whenever you open a new website with a login page. You will enter the password and the username, chrome will save it, this will become the session, and u can login from any of your devices if they have chrome, this can be thought of as a connection.

IPSEC

- It is a group of protocols that are used to secure connections over the internet.
- It is related to cross protocol layers.
- It has security features for IPv6 and IPv4.

**Benefits:**

1. When implemented in a firewall or router, it provides strong
2. security to all traffic crossing the perimeter
3. In a firewall it is resistant to bypass
4. Below transport layer, hence transparent to applications
5. Can be transparent to end users
6. Can provide security for individual users
7. Secures routing architecture