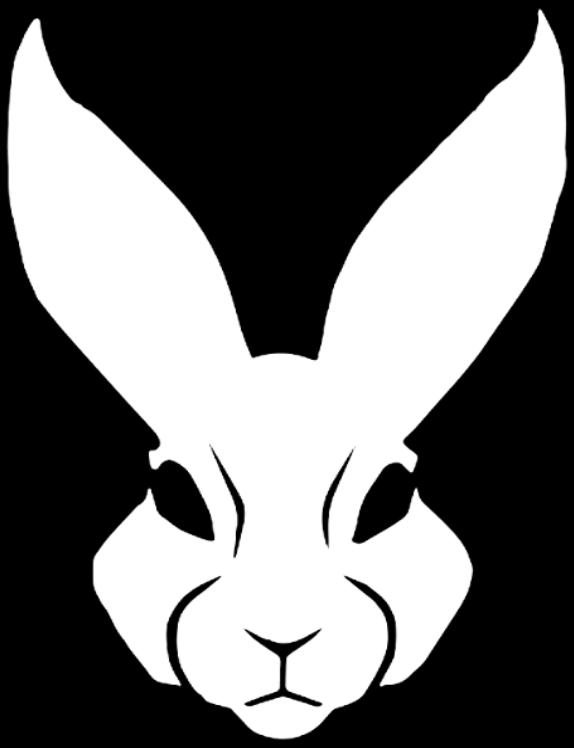


Thank you for downloading! This document is intended to be used as a learning and reference tool. In it you will find all of my compiled notes from various courses I have taken, and helpful information I've collected. I Intend to update my GitHub regularly as I gather more information, resources, and continue my efforts. Enjoy and use responsibly.

-Chocka

https://github.com/xChockax

## Table of Contents:

# Q1: What is a DMZ

A: In computer security, a DMZ Network (sometimes referred to as a “demilitarized zone”) functions as a subnetwork containing an organization's exposed, outward-facing services. It acts as the exposed point to an untrusted networks, commonly the Internet.
The goal of a DMZ is to add an extra layer of security to an organization's local area network. A protected and monitored network node that faces outside the internal network can access what is exposed in the DMZ, while the rest of the organization's network is safe behind a firewall. When implemented properly, a DMZ Network gives organizations extra protection in detecting and mitigating security breaches before they reach the internal network, where valuable assets are stored.

All services accessible to users on communicating from an external network can and should be placed in the DMZ, if one is used. The most common services are:

- **Web servers:** Web servers responsible for maintaining communication with an internal database server may need to be placed into a DMZ. This helps ensure the safety of the internal database, which is often storing sensitive information. The web servers can then interact with internal database server through an application firewall or directly, while still falling under the umbrella of the DMZ protections.
- **Mail servers:** individual email messages, as well as the user database built to store login credentials and personal messages, are usually stored on servers without direct access to the internet. Therefore, an email server will be built or placed inside the DMZ in order to interact with and access the email database without directly exposing it to potentially harmful traffic.
- **FTP servers:** These can host critical content on an organization's site, and allow direct interaction with files. Therefore, an FTP server should always be partially isolated from critical internal systems

DMZ Designs: There are numerous ways to construct a network with a DMZ. The two major methods are a single firewall (sometimes called a three-legged model), or dual firewalls. Each of these system can be expanded to create complex architectures built to satisfy network requirements:
**Single firewall**: A modest approach to network architecture involves using a single firewall, with a minimum of 3 network interfaces. The DMZ will be placed Inside of this firewall. The tier of operations is as follows: the external network device makes the connection from the ISP, the internal network is connected by the second device, and connections within the DMZ is handled by the third network device. **Dual firewall**: The more secure approach is to use two firewalls to create a DMZ. The first firewall (referred to as the “frontend” firewall) is configured to only allow traffic destined for the DMZ. The second firewall (referred to as the “backend” firewall) is only responsible for the traffic that travels from the DMZ to the internal network. An effective way of further increasing protection is to use firewalls built by separate vendors, because they

are less likely to have the same security vulnerabilities. While more effective, this scheme can be more costly to implement across a large network.

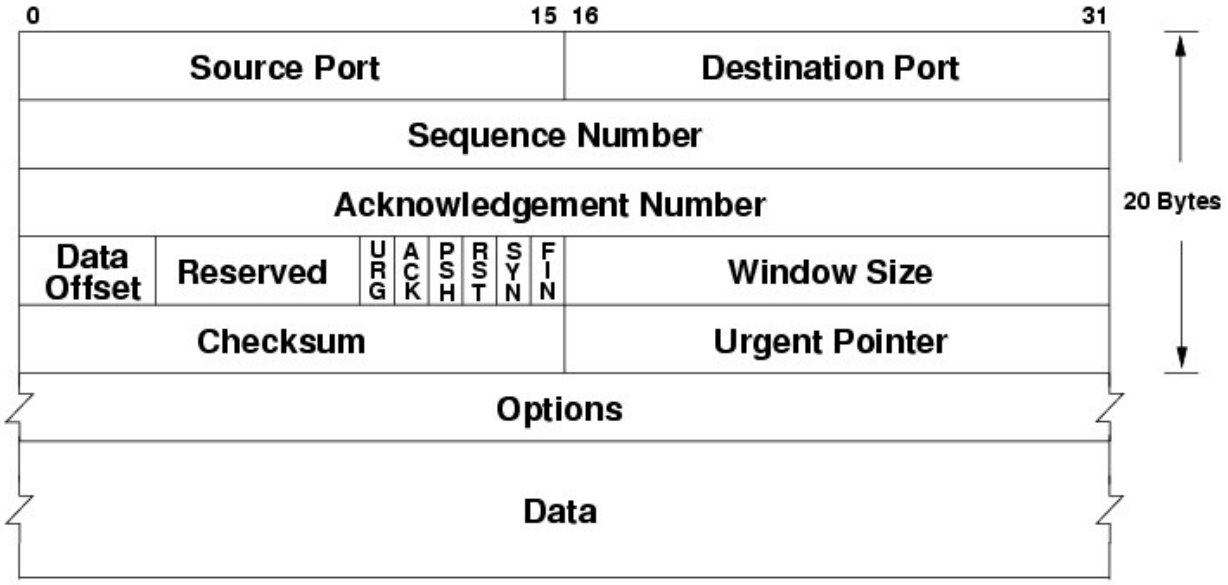
## Q2: Explain TCP / UDP and their differences

**TCP**

TCP came before UDP. It stands for Transmission Control Protocol. You'll often see it referred to as **TCP/IP**, though there's no distinction between that and TCP.

The IP protocol breaks up data into packets and sends them to a destination over the internet, but how do you put those packets back together once they arrive? That's what TCP was invented for. Once packets reach their destination, they are reassembled by the receiving device back into their original form.

TCP requires both parties to communicate in order to establish a connection and send data. TCP guarantees the recipient will receive packets in order according to **sequence numbers** included in the header. The recipient will send a message back to the sender for each packet, **acknowledging** that they've been received. Any packets not acknowledged by the recipient are sent again. Packets are checked for errors using a **checksum**, which is also included in the header.




Because of all this back-and-forth between client and server, TCP can reliably ensure the integrity of data exchanged over the internet. Put simply, it can guarantee the data arrives exactly as it was sent with no modifications or missing parts. This makes TCP useful for a huge range of applications, and it's the most commonly used protocol on the internet. Any time you click a

link, download a file in your web browser, update an application, or open an email, TCP is probably used.

However, all that back-and-forth communication slows TCP down. If a packet goes missing, the entire operation is held up until it's sent again. While this only translates to milliseconds in real life, it can impact performance for applications that require a lot of bandwidth. Enter UDP.

**UDP**

UDP stands for User Datagram Protocol. Recall that a datagram and a packet are more or less the same thing. UDP, also built on top of the IP protocol, works similarly to TCP, but is **simpler and faster.**

The main difference is that **UDP doesn't require the recipient to acknowledge** that each packet has been received. Any packets that get lost in transit are not resent. This enables computers to communicate more quickly, but the data received might not exactly match the data sent.

UDP packets don't have sequence numbers, so they can arrive out of order. They do have checksums, though, so the packets that do arrive are protected against corruption or modification in transit.

For this reason, UDP is used when speed is preferred over integrity and error correction. Some common applications include streaming video and music, live broadcasts, voice and video calling (VoIP), and online gaming. In these scenarios, it doesn't really matter if you lose the occasional video frame or button press, which favors UDP. DNS traffic is usually exchanged over the UDP protocol.

Break Down of Differences

**UDP:**

- Used for streaming video, gaming, VoIP, live broadcasts
- Faster and requires fewer resources
- Packets do not necessarily arrive in order
- Allows missing packets; sender cannot know whether a packet has been received

**TCP:**

- Most widely used protocol on the internet
- TCP guarantees no packets are missing and all data sent makes it to the recipient
- TCP sends packets in order so they can be easily stitched back together
- Slower and requires more resources

## Q3: TCP 3-Way Handshake (SYN, SYN-ACK,ACK)

THREE-WAY HANDSHAKE or a TCP 3-way handshake is a process which is used in a TCP/IP network to make a connection between the server and client. It is a three-step process that requires both the client and server to exchange synchronization and acknowledgment packets before the real data communication process starts.

Three-way handshake process is designed in such a way that both ends help you to initiate, negotiate, and separate TCP socket connections at the same time. It allows you to transfer multiple TCP socket connections in both directions at the same time.

- **Step 1:** In the first step, the client establishes a connection with a server. It sends a segment with SYN and informs the server about the client should start communication, and with what should be its sequence number.
- **Step 2:** In this step **s**erver responds to the client request with SYN-ACK signal set. ACK helps you to signify the response of segment that is received and SYN signifies what sequence number it should able to start with the segments.
- **Step 3:** In this final step, the client acknowledges the response of the Server, and they both create a stable connection will begin the actual data transfer process.

**Summary**

- TCP 3-way handshake or three-way handshake or TCP 3-way handshake is a process which is used in a TCP/IP network to make a connection between server and client.
- Syn use to initiate and establish a connection
- ACK helps to confirm to the other side that it has received the SYN.
- SYN-ACK is a SYN message from local device and ACK of the earlier packet.
- FIN is used for terminating a connection.
- TCP handshake process, a client needs to initiate the conversation by requesting a communication session with the Server
- In the first step, the client establishes a connection with a server
- In this second step, the server responds to the client request with SYN-ACK signal set
- In this final step, the client acknowledges the response of the Server
- TCP automatically terminates the connection between two separate endpoints.

## Q4: What is the OSI Model?

The Open Systems Interconnection (OSI) model is a conceptual model created by the International Organization for Standardization which enables diverse communication systems to

communicate using standard protocols. In plain English, the OSI provides a standard for different computer systems to be able to communicate with each other.

The OSI model can be seen as a universal language for computer networking. It's based on the concept of splitting up a communication system into seven abstract layers, each one stacked upon the last.

Application Layer (7)
Human-Computer interaction layer, where applications can access network services

Presentation Layer (6)
Ensures that data is in a usable format and is where data encryption occurs

Session Layer (5)
Maintains connections and is responsible for controlling ports and sessions

Transport Layer (4)
Transmits data using transmission protocols including TCP and UDP

Network Layer (3)
Decides which physical path the data will take

Datalink Layer (2)
Defines the format of data on the network

Physical Layer (1)
Transmits raw bit stream over the physical medium

### *7. The Application Layer*

This is the only layer that directly interacts with data from the user. Software applications like web browsers and email clients rely on the application layer to initiate communications. But it should be made clear that client software applications are not part of the application layer; rather the application layer is responsible for the protocols and data manipulation that the software relies on to present meaningful data to the user. Application layer protocols include HTTP as well as SMTP (Simple Mail Transfer Protocol is one of the protocols that enables email communications).

### *6. The Presentation Layer*

This layer is primarily responsible for preparing data so that it can be used by the application layer; in other words, layer 6 makes the data presentable for applications to consume. The presentation layer is responsible for translation, encryption, and compression of data.

Two communicating devices communicating may be using different encoding methods, so layer 6 is responsible for translating incoming data into a syntax that the application layer of the receiving device can understand.

If the devices are communicating over an encrypted connection, layer 6 is responsible for adding the encryption on the sender's end as well as decoding the encryption on the receiver's end so that it can present the application layer with unencrypted, readable data.

Finally the presentation layer is also responsible for compressing data it receives from the application layer before delivering it to layer 5. This helps improve the speed and efficiency of communication by minimizing the amount of data that will be transferred.

### *5. The Session Layer*

This is the layer responsible for opening and closing communication between the two devices. The time between when the communication is opened and closed is known as the session. The session layer ensures that the session stays open long enough to transfer all the data being exchanged, and then promptly closes the session in order to avoid wasting resources.

The session layer also synchronizes data transfer with checkpoints. For example, if a 100 megabyte file is being transferred, the session layer could set a checkpoint every 5 megabytes. In the case of a disconnect or a crash after 52 megabytes have been transferred, the session could be resumed from the last checkpoint, meaning only 50 more megabytes of data need to be transferred. Without the checkpoints, the entire transfer would have to begin again from scratch.

### *4. The Transport Layer*

Layer 4 is responsible for end-to-end communication between the two devices. This includes taking data from the session layer and breaking it up into chunks called segments before sending it to layer 3. The transport layer on the receiving device is responsible for reassembling the segments into data the session layer can consume.

The transport layer is also responsible for flow control and error control. Flow control determines an optimal speed of transmission to ensure that a sender with a fast connection doesn't overwhelm a receiver with a slow connection. The transport layer performs error control on the receiving end by ensuring that the data received is complete, and requesting a retransmission if it isn't.

### *3. The Network Layer*

The network layer is responsible for facilitating data transfer between two different networks. If the two devices communicating are on the same network, then the network layer is unnecessary. The network layer breaks up segments from the transport layer into smaller units, called packets, on the sender's device, and reassembling these packets on the receiving device. The network layer also finds the best physical path for the data to reach its destination; this is known as routing.

### *2. The Data Link Layer*

The data link layer is very similar to the network layer, except the data link layer facilitates data transfer between two devices on the SAME network. The data link layer takes packets from the network layer and breaks them into smaller pieces called frames. Like the network layer, the data link layer is also responsible for flow control and error control in intra-network communication (The transport layer only does flow control and error control for inter-network communications).

### *1. The Physical Layer*

This layer includes the physical equipment involved in the data transfer, such as the cables and switches. This is also the layer where the data gets converted into a bit stream, which is a string of 1s and 0s. The physical layer of both devices must also agree on a signal convention so that the 1s can be distinguished from the 0s on both devices.

## Q5: Port Numbers for Common Applications

- From 0 to 1023 – well known ports assigned to common protocols and services
- From 1024 to 49151 – registered ports assigned by ICANN to a specific service
- From 49152 to 65 535 – dynamic (private, high) ports range from 49,152 to 65,535. Can be used by any service on an ad hoc basis. Ports are assigned when a session is established, and released when the session ends.

Well known ones are:

| Port | Service name | Transport protocol |
|---|---|---|
| 20, 21 | File Transfer Protocol (FTP) | TCP |
| 22 | Secure Shell (SSH) | TCP and UDP |

| | | |
|---|---|---|
| 23 | Telnet | TCP |
| 25 | Simple Mail Transfer Protocol (SMTP) | TCP |
| 50, 51 | IPSec | |
| 53 | Domain Name System (DNS) | TCP and UDP |
| 67, 68 | Dynamic Host Configuration Protocol (DHCP) | UDP |
| 69 | Trivial File Transfer Protocol (TFTP) | UDP |
| 80 | HyperText Transfer Protocol (HTTP) | TCP |
| 110 | Post Office Protocol (POP3) | TCP |
| 119 | Network News Transport Protocol (NNTP) | TCP |
| 123 | Network Time Protocol (NTP) | UDP |
| 135-139 | NetBIOS | TCP and UDP |
| 143 | Internet Message Access Protocol (IMAP4) | TCP and UDP |
| 161, 162 | Simple Network Management Protocol (SNMP) | TCP and UDP |
| 389 | Lightweight Directory Access Protocol | TCP and UDP |
| 443 | HTTP with Secure Sockets Layer (SSL) | TCP and UDP |
| 989, 990 | FTP over SSL/TLS (implicit mode) | TCP |
| 3389 | Remote Desktop Protocol | TCP and UDP |

# Q6: What is DNS and how does it work

**What is DNS?**

The Domain Name System (DNS) is the phonebook of the Internet. Humans access information online through domain names, like nytimes.com or espn.com. Web browsers interact through Internet Protocol (IP) addresses. DNS translates domain names to IP addresses so browsers can load Internet resources.

Each device connected to the Internet has a unique IP address which other machines use to find the device. DNS servers eliminate the need for humans to memorize IP addresses such as 192.168.1.1 (in IPv4), or more complex newer alphanumeric IP addresses such as 2400:cb00:2048:1::c629:d7a2 (in IPv6).

**How does DNS work?**

The process of DNS resolution involves converting a hostname (such as www.example.com) into a computer-friendly IP address (such as 192.168.1.1). An IP address is given to each device on the Internet, and that address is necessary to find the appropriate Internet device - like a street address is used to find a particular home.

When a user wants to load a webpage, a translation must occur between what a user types into their web browser (example.com) and the machine-friendly address necessary to locate the example.com webpage.

In order to understand the process behind the DNS resolution, it's important to learn about the different hardware components a DNS query must pass between. For the web browser, the DNS lookup occurs " behind the scenes" and requires no interaction from the user's computer apart from the initial request.

**There are 4 DNS servers involved in loading a webpage:**

- **DNS recursor** - The recursor can be thought of as a librarian who is asked to go find a particular book somewhere in a library. The DNS recursor is a server designed to receive queries from client machines through applications such as web browsers. Typically the recursor is then responsible for making additional requests in order to satisfy the client's DNS query.
- **Root nameserver** - The root server is the first step in translating (resolving) human readable host names into IP addresses. It can be thought of like an index in a library that points to different racks of books - typically it serves as a reference to other more specific locations.
- **TLD nameserver** - The top level domain server (TLD) can be thought of as a specific rack of books in a library. This nameserver is the next step in the search for a specific IP address, and it hosts the last portion of a hostname (In example.com, the TLD server is "com").
- **Authoritative nameserver** - This final nameserver can be thought of as a dictionary on a rack of books, in which a specific name can be translated into its definition. The authoritative nameserver is the last stop in the nameserver query. If the authoritative name server has access to the requested record, it will return the IP address for the requested hostname back to the DNS Recursor (the librarian) that made the initial request.

**What's the difference between an authoritative DNS server and a recursive DNS resolver?**

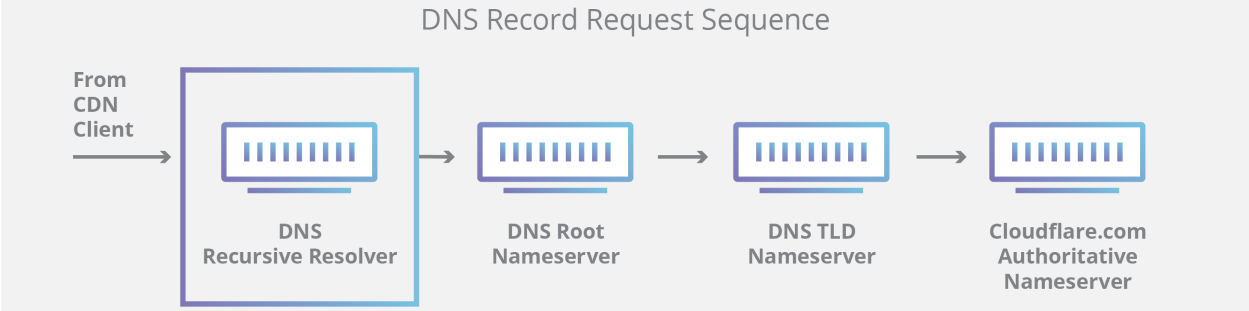
Both concepts refer to servers (groups of servers) that are integral to the DNS infrastructure, but each performs a different role and lives in different locations inside the pipeline of a DNS query.

One way to think about the difference is the recursive resolver is at the beginning of the DNS query and the authoritative nameserver is at the end.

### *Recursive DNS resolver*

The recursive resolver is the computer that responds to a recursive request from a client and takes the time to track down the DNS record. It does this by making a series of requests until it reaches the authoritative DNS nameserver for the requested record (or times out or returns an error if no record is found).

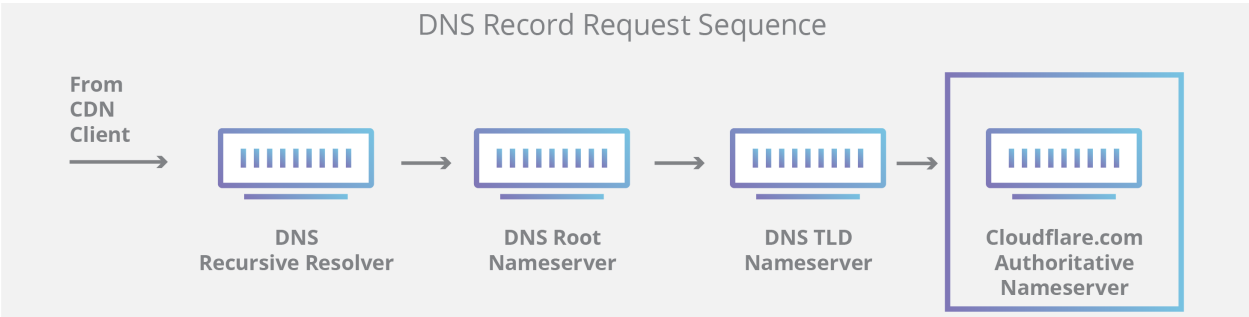
Luckily, recursive DNS resolvers do not always need to make multiple requests in order to track down the records needed to respond to a client; caching is a data persistence process that helps short-circuit the necessary requests by serving the requested resource record earlier in the DNS lookup.



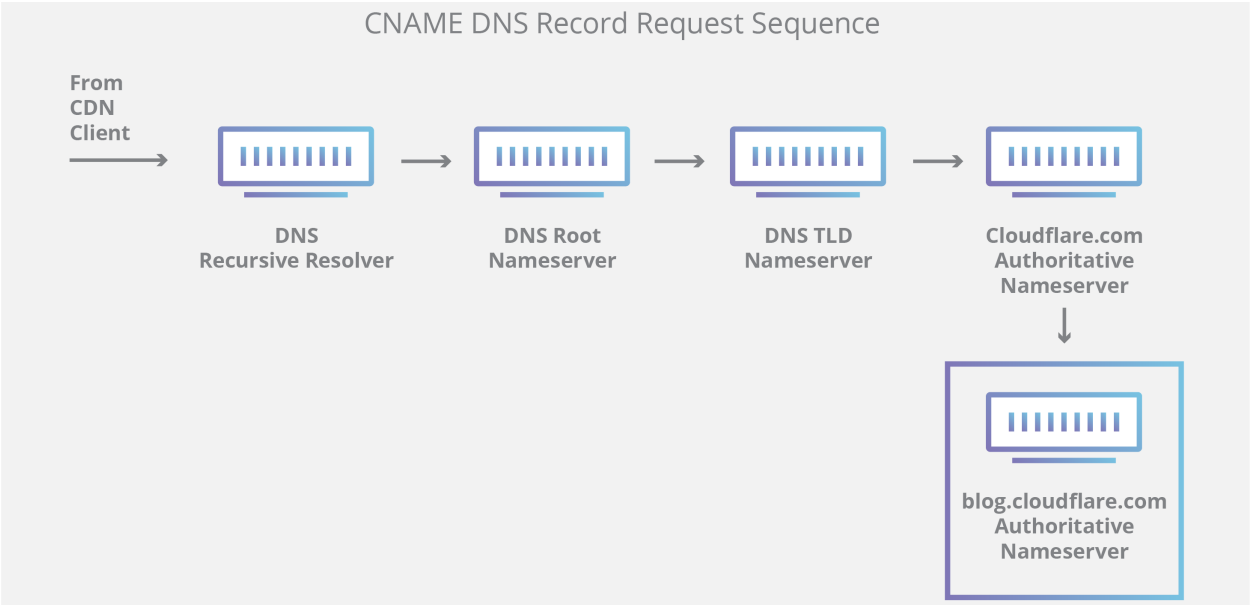

*Authoritative DNS server*

Put simply, an authoritative DNS server is a server that actually holds, and is responsible for, DNS resource records. This is the server at the bottom of the DNS lookup chain that will respond with the queried resource record, ultimately allowing the web browser making the request to reach the IP address needed to access a website or other web resources.

An authoritative nameserver can satisfy queries from its own data without needing to query another source, as it is the final source of truth for certain DNS records.




It's worth mentioning that in instances where the query is for a subdomain such as foo.example.com or blog.cloudflare.com, an additional nameserver will be added to the sequence after the authoritative nameserver, which is responsible for storing the subdomain's CNAME record.




There is a key difference between many DNS services and the one that Cloudflare provides.

Different DNS recursive resolvers such as Google DNS, OpenDNS, and providers like Comcast all maintain data center installations of DNS recursive resolvers. These resolvers allow for quick and easy queries through optimized clusters of DNS-optimized computer systems, but they are fundamentally different than the nameservers hosted by Cloudflare.

Cloudflare maintains infrastructure-level nameservers that are integral to the functioning of the Internet. One key example is the f-root server network which Cloudflare is partially responsible for hosting. The F-root is one of the root level DNS nameserver infrastructure components responsible for the billions of Internet requests per day.

Our Anycast network puts us in a unique position to handle large volumes of DNS traffic without service interruption.

**What are the steps in a DNS lookup?**

For most situations, DNS is concerned with a domain name being translated into the appropriate IP address. To learn how this process works, it helps to follow the path of a DNS lookup as it travels from a web browser, through the DNS lookup process, and back again. Let's take a look at the steps.

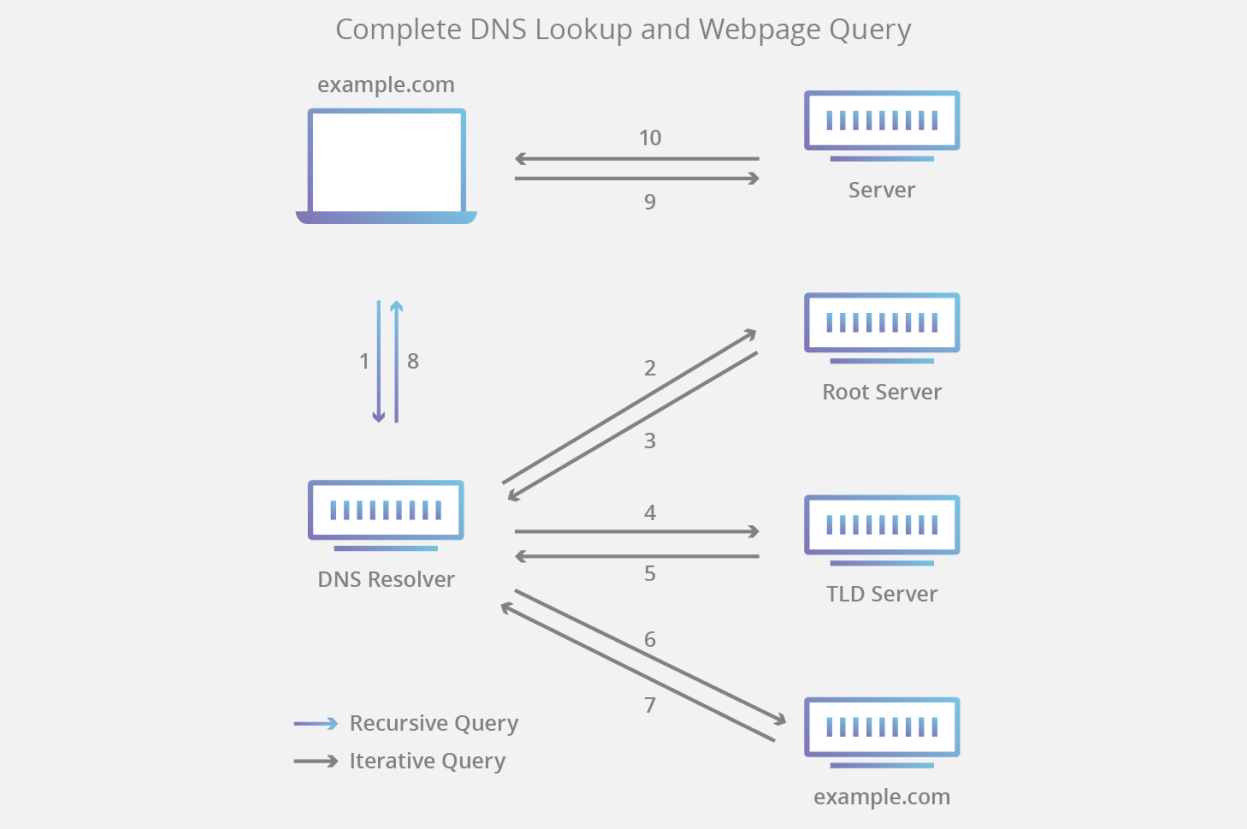
Note: Often DNS lookup information will be cached either locally inside the querying computer or remotely in the DNS infrastructure. There are typically 8 steps in a DNS lookup. When DNS information is cached, steps are skipped from the DNS lookup process which makes it quicker. The example below outlines all 8 steps when nothing is cached.

*The 8 steps in a DNS lookup:*

1. A user types 'example.com' into a web browser and the query travels into the Internet and is received by a DNS recursive resolver.
2. The resolver then queries a DNS root nameserver (.).
3. The root server then responds to the resolver with the address of a Top Level Domain (TLD) DNS server (such as .com or .net), which stores the information for its domains. When searching for example.com, our request is pointed toward the .com TLD.
4. The resolver then makes a request to the .com TLD.
5. The TLD server then responds with the IP address of the domain's nameserver, example.com.
6. Lastly, the recursive resolver sends a query to the domain's nameserver.
7. The IP address for example.com is then returned to the resolver from the nameserver.
8. The DNS resolver then responds to the web browser with the IP address of the domain requested initially.

   Once the 8 steps of the DNS lookup have returned the IP address for example.com, the browser is able to make the request for the web page:

9. The browser makes a HTTP request to the IP address.
10. The server at that IP returns the webpage to be rendered in the browser (step 10).

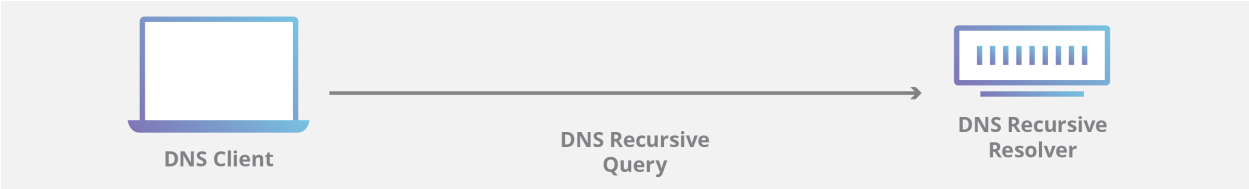



**What is a DNS resolver?**

The DNS resolver is the first stop in the DNS lookup, and it is responsible for dealing with the client that made the initial request. The resolver starts the sequence of queries that ultimately leads to a URL being translated into the necessary IP address.

Note: A typical uncached DNS lookup will involve both recursive and iterative queries.

It's important to differentiate between a recursive DNS query and a recursive DNS resolver. The query refers to the request made to a DNS resolver requiring the resolution of the query. A DNS recursive resolver is the computer that accepts a recursive query and processes the response by making the necessary requests.




**What are the types of DNS Queries?**

In a typical DNS lookup three types of queries occur. By using a combination of these queries, an optimized process for DNS resolution can result in a reduction of distance traveled. In an ideal situation cached record data will be available, allowing a DNS name server to return a non-recursive query.

*3 types of DNS queries:*

1. **Recursive query** - In a recursive query, a DNS client requires that a DNS server (typically a DNS recursive resolver) will respond to the client with either the requested resource record or an error message if the resolver can't find the record.
2. **Iterative query** - in this situation the DNS client will allow a DNS server to return the best answer it can. If the queried DNS server does not have a match for the query name, it will return a referral to a DNS server authoritative for a lower level of the domain namespace. The DNS client will then make a query to the referral address. This process continues with additional DNS servers down the query chain until either an error or timeout occurs.
3. **Non-recursive query** - typically this will occur when a DNS resolver client queries a DNS server for a record that it has access to either because it's authoritative for the record or the record exists inside of its cache. Typically, a DNS server will cache DNS records to prevent additional bandwidth consumption and load on upstream servers.

**What is DNS caching? Where does DNS caching occur?**

The purpose of caching is to temporarily stored data in a location that results in improvements in performance and reliability for data requests. DNS caching involves storing data closer to the requesting client so that the DNS query can be resolved earlier and additional queries further

down the DNS lookup chain can be avoided, thereby improving load times and reducing bandwidth/CPU consumption.

DNS data can be cached in a variety of locations, each of which will store DNS records for a set amount of time determined by a time-to-live (TTL).

### *Browser DNS caching*

Modern web browsers are designed by default to cache DNS records for a set amount of time. the purpose here is obvious; the closer the DNS caching occurs to the web browser, the fewer processing steps must be taken in order to check the cache and make the correct requests to an IP address. When a request is made for a DNS record, the browser cache is the first location checked for the requested record.

In chrome, you can see the status of your DNS cache by going to chrome://net-internals/#dns.

### *Operating system (OS) level DNS caching*

The operating system level DNS resolver is the second and last local stop before a DNS query leaves your machine. The process inside your operating system that is designed to handle this query is commonly called a “stub resolver” or DNS client.

hen a stub resolver gets a request from an application, it first checks its own cache to see if it has the record. If it does not, it then sends a DNS query (with a recursive flag set), outside the local network to a DNS recursive resolver inside the Internet service provider (ISP).

### *Recursive resolver DNS caching*

When the recursive resolver inside the ISP receives a DNS query, like all previous steps, it will also check to see if the requested host-to-IP-address translation is already stored inside its local persistence layer.

The recursive resolver also has additional functionality depending on the types of records it has in its cache:

1. If the resolver does not have the A records, but does have the NS records for the authoritative nameservers, it will query those name servers directly, bypassing several steps in the DNS query. This shortcut prevents lookups from the root and .com nameservers (in our search for example.com) and helps the resolution of the DNS query occur more quickly.
2. If the resolver does not have the NS records, it will send a query to the TLD servers (.com in our case), skipping the root server.
3. In the unlikely event that the resolver does not have records pointing to the TLD servers, it will then query the root servers. This event typically occurs after a DNS cache has been purged.

# Q7: What is a VPN, different types and how do they work?

VPN is a Virtual Private Network that allows a user to connect to a private network over the Internet securely and privately. VPN creates an encrypted connection, known as VPN tunnel, and all Internet traffic and communication is passed through this secure tunnel. Thus, keeping the user data secure and private.

There are two basic VPN types which are explained below.

### *1. Remote Access VPN*

Remote access VPN allows a user to connect to a private network and access its services and resources remotely. The connection between the user and the private network happens through the Internet and the connection is secure and private.

Remote Access VPN is useful for business users as well as home users.

A corporate employee, while traveling, uses a VPN to connect to his/her company's private network and remotely access files and resources on the private network.

Home users, or private users of VPN, primarily use VPN services to bypass regional restrictions on the Internet and access blocked websites. Users conscious of Internet security also use VPN services to enhance their Internet security and privacy.

### *2. Site – to – Site VPN*

A Site-to-Site VPN is also called as Router-to-Router VPN and is mostly used in the corporates. Companies, with offices in different geographical locations, use Site-to-site VPN to connect the network of one office location to the network at another office location. When multiple offices of the same company are connected using Site-to-Site VPN type, it is called as Intranet based VPN. When companies use Site-to-site VPN type to connect to the office of another company, it is called as Extranet based VPN. Basically, Site-to-site VPN create a virtual bridge between the networks at geographically distant offices and connect them through the Internet and maintain a secure and private communication between the networks.

Since Site-to-site VPN is based on Router-to-Router communication, in this VPN type one router acts as a VPN Client and another router as a VPN Server. The communication between the two routers starts only after an authentication is validated between the two.

**Types of VPN protocols**

The above two VPN types are based on different VPN security protocols. Each of these VPN protocols offer different features and levels of security, and are explained below:

*1. Internet Protocol Security or IPSec:*

Internet Protocol Security or IPSec is used to secure Internet communication across an IP network. IPSec secures Internet Protocol communication by authenticating the session and encrypts each data packet during the connection.

IPSec operates in two modes, Transport mode and Tunneling mode, to protect data transfer between two different networks. The transport mode encrypts the message in the data packet and the tunneling mode encrypts the entire data packet. IPSec can also be used with other security protocols to enhance the security system.

*2. Layer 2 Tunneling Protocol (L2TP):*

L2TP or Layer 2 Tunneling Protocol is a tunneling protocol that is usually combined with another VPN security protocol like IPSec to create a highly secure VPN connection. L2TP creates a tunnel between two L2TP connection points and IPSec protocol encrypts the data and handles secure communication between the tunnel.

*3. Point – to – Point Tunneling Protocol (PPTP):*

PPTP or Point-to-Point Tunneling Protocol creates a tunnel and encapsulates the data packet. It uses a Point-to-Point Protocol (PPP) to encrypt the data between the connection. PPTP is one of the most widely used VPN protocol and has been in use since the time of Windows 95. Apart from Windows, PPTP is also supported on Mac and Linux.

*4. Secure Sockets Layer (SSL) and Transport Layer Security (TLS):*

SSL (Secure Sockets Layer) and TLS (Transport Layer Security) create a VPN connection where the web browser acts as the client and user access is restricted to specific applications instead of entire network. SSL and TLS protocol is most commonly used by online shopping websites and service providers. Web browsers switch to SSL with ease and with almost no action required from the user, since web browsers come integrated with SSL and TLS. SSL connections have https in the beginning of the URL instead of http.

*5. OpenVPN:*

OpenVPN is an open source VPN that is useful for creating Point-to-Point and Site-to-Site connections. It uses a custom security protocol based on SSL and TLS protocol.

*6. Secure Shell (SSH):*

Secure Shell or SSH creates the VPN tunnel through which the data transfer happens and also ensures that the tunnel is encrypted. SSH connections are created by a SSH client and data is transferred from a local port on to the remote server through the encrypted tunnel.

# Q8: The OWASP Top 10

1. **Injection.** Injection flaws, such as SQL, NoSQL, OS, and LDAP injection, occur when untrusted data is sent to an interpreter as part of a command or query. The attacker's hostile data can trick the interpreter into executing unintended commands or accessing data without proper authorization.
2. **Broken Authentication.** Application functions related to authentication and session management are often implemented incorrectly, allowing attackers to compromise passwords, keys, or session tokens, or to exploit other implementation flaws to assume other users' identities temporarily or permanently.
3. **Sensitive Data Exposure.** Many web applications and APIs do not properly protect sensitive data, such as financial, healthcare, and PII. Attackers may steal or modify such weakly protected data to conduct credit card fraud, identity theft, or other crimes. Sensitive data may be compromised without extra protection, such as encryption at rest or in transit, and requires special precautions when exchanged with the browser.
4. **XML External Entities (XXE).** Many older or poorly configured XML processors evaluate external entity references within XML documents. External entities can be used to disclose internal files using the file URI handler, internal file shares, internal port scanning, remote code execution, and denial of service attacks.
5. **Broken Access Control.** Restrictions on what authenticated users are allowed to do are often not properly enforced. Attackers can exploit these flaws to access unauthorized functionality and/or data, such as access other users' accounts, view sensitive files, modify other users' data, change access rights, etc.
6. **Security Misconfiguration**. Security misconfiguration is the most commonly seen issue. This is commonly a result of insecure default configurations, incomplete or ad hoc configurations, open cloud storage, misconfigured HTTP headers, and verbose error messages containing sensitive information. Not only must all operating systems, frameworks, libraries, and applications be securely configured, but they must be patched/upgraded in a timely fashion.
7. **Cross-Site Scripting (XSS).** XSS flaws occur whenever an application includes untrusted data in a new web page without proper validation or escaping, or updates an existing web page with user-supplied data using a browser API that can create HTML or JavaScript. XSS allows attackers to execute scripts in the victim's browser which can hijack user sessions, deface web sites, or redirect the user to malicious sites.
8. **Insecure Deserialization.** Insecure deserialization often leads to remote code execution. Even if deserialization flaws do not result in remote code execution, they can be used to perform attacks, including replay attacks, injection attacks, and privilege escalation attacks.
9. **Using Components with Known Vulnerabilities.** Components, such as libraries, frameworks, and other software modules, run with the same privileges as the application. If a vulnerable component is exploited, such an attack can facilitate serious data loss or server takeover. Applications and APIs using components with known vulnerabilities may undermine application defenses and enable various attacks and impacts.
10. **Insufficient Logging & Monitoring.** Insufficient logging and monitoring, coupled with missing or ineffective integration with incident response, allows attackers to further attack systems, maintain persistence, pivot to more systems, and tamper, extract, or

destroy data. Most breach studies show time to detect a breach is over 200 days, typically detected by external parties rather than internal processes or monitoring.

## Q9: The MITRE Attack Framework

MITRE ATT&CK® is a globally-accessible knowledge base of adversary tactics and techniques based on real-world observations. The ATT&CK knowledge base is used as a foundation for the development of specific threat models and methodologies in the private sector, in government, and in the cybersecurity product and service community.

With the creation of ATT&CK, MITRE is fulfilling its mission to solve problems for a safer world — by bringing communities together to develop more effective cybersecurity. ATT&CK is open and available to any person or organization for use at no charge.

**https://attack.mitre.org/**

## Q10: Most Common Cyber Attacks

Denial-of-service (DoS) and distributed denial-of-service (DDoS) attacks
Man-in-the-middle (MitM) attack/Session Hijacking
Phishing and spear phishing attacks
Drive-by attack
Password attack
SQL injection attack
Cross-site scripting (XSS) attack
Eavesdropping attack
Birthday attack
Malware attack

### Zero-day exploit

A zero-day exploit hits after a network vulnerability is announced but before a patch or solution is implemented. Attackers target the disclosed vulnerability during this window of time. Zero-day vulnerability threat detection requires constant awareness.

### DNS Tunneling

DNS tunneling utilizes the DNS protocol to communicate non-DNS traffic over port 53. It sends HTTP and other protocol traffic over DNS. There are various, legitimate reasons to utilize DNS tunneling. However, there are also malicious reasons to use DNS Tunneling VPN services. They can be used to disguise outbound traffic as DNS, concealing data that is typically shared through an internet connection. For malicious use, DNS requests are manipulated to exfiltrate data from a compromised system to the attacker's infrastructure. It can also be used for command and control callbacks from the attacker's infrastructure to a compromised system.

### Credential Reuse

Users today have so many logins and passwords to remember that it's tempting to reuse credentials here or there to make life a little easier. Even though security best practices universally recommend that you have unique passwords for all your applications and websites, many people still reuse their passwords—a fact attackers rely on.
Once attackers have a collection of usernames and passwords from a breached website or service (easily acquired on any number of black market websites on the internet), they know that if they use these same credentials on other websites there's a chance they'll be able to log in. No matter how tempting it may be to reuse credentials for your email, bank account, and your favorite sports forum, it's possible that one day the forum will get hacked, giving an attacker easy access to your email and bank account. When it comes to credentials, variety is essential. Password managers are available and can be helpful when it comes to managing the various credentials you use.
This is just a selection of common attack types and techniques (follow this link to learn more about web application vulnerabilitites specificially). It is not intended to be exhaustive, and attackers do evolve and develop new methods as needed; however, being aware of, and mitigating these types of attacks will significantly improve your security posture.

## Q11: What is SQL Injection (SQLi)

SQL injection is a web security vulnerability that allows an attacker to interfere with the queries that an application makes to its database. It generally allows an attacker to view data that they are not normally able to retrieve. This might include data belonging to other users, or any other data that the application itself is able to access. In many cases, an attacker can modify or delete this data, causing persistent changes to the application's content or behavior.

In some situations, an attacker can escalate an SQL injection attack to compromise the underlying server or other back-end infrastructure, or perform a denial-of-service attack.

https://portswigger.net/web-security/sql-injection

## Q12: What is Phishing

Phishing is a cybercrime in which a target or targets are contacted by email, telephone or text message by someone posing as a legitimate institution to lure individuals into providing sensitive data such as personally identifiable information, banking and credit card details, and passwords.

Common Features of Phishing Emails

**Too Good To Be True** - Lucrative offers and eye-catching or attention-grabbing statements are designed to attract people's attention immediately. For instance, many claim that you have won an iPhone, a lottery, or some other lavish prize. Just don't click on any suspicious emails. Remember that if it seems to good to be true, it probably is!

**Sense of Urgency** - A favorite tactic amongst cybercriminals is to ask you to act fast because the super deals are only for a limited time. Some of them will even tell you that you have only a few minutes to respond. When you come across these kinds of emails, it's best to just ignore them. Sometimes, they will tell you that your account will be suspended unless you update your personal details immediately. Most reliable organizations give ample time before they terminate an account and they never ask patrons to update personal details over the Internet. When in doubt, visit the source directly rather than clicking a link in an email.

**Hyperlinks -** A link may not be all it appears to be. Hovering over a link shows you the actual URL where you will be directed upon clicking on it. It could be completely different or it could be a popular website with a misspelling, for instance www.bankofarnerica.com - the 'm' is actually an 'r' and an 'n', so look carefully.

**Attachments -** If you see an attachment in an email you weren't expecting or that doesn't make sense, don't open it! They often contain payloads like ransomware or other viruses. The only file type that is always safe to click on is a .txt file.

**Unusual Sender -** Whether it looks like it's from someone you don't know or someone you do know, if anything seems out of the ordinary, unexpected, out of character or just suspicious in general don't click on it!

Prevent Phishing Attacks:

Though hackers are constantly coming up with new techniques, there are some things that you can do to protect yourself and your organization:

-To protect against spam mails, spam filters can be used. Generally, the filters assess the origin of the message, the software used to send the message, and the appearance of the message to determine if it's spam. Occasionally, spam filters may even block emails from legitimate sources, so it isn't always 100% accurate.

-The browser settings should be changed to prevent fraudulent websites from opening. Browsers keep a list of fake websites and when you try to access the website, the address is blocked or an alert message is shown. The settings of the browser should only allow reliable websites to open

up.

-Many websites require users to enter login information while the user image is displayed. This type of system may be open to security attacks. One way to ensure security is to change passwords on a regular basis, and never use the same password for multiple accounts. It's also a good idea for websites to use a CAPTCHA system for added security.

-Banks and financial organizations use monitoring systems to prevent phishing. Individuals can report phishing to industry groups where legal actions can be taken against these fraudulent websites. Organizations should provide security awareness training to employees to recognize the risks.

-Changes in browsing habits are required to prevent phishing. If verification is required, always contact the company personally before entering any details online.

-If there is a link in an email, hover over the URL first. Secure websites with a valid Secure Socket Layer (SSL) certificate begin with "https". Eventually all sites will be required to have a valid SSL.

https://www.phishing.org/what-is-phishing

## Q13: What are the Different DDoS Attacks

- **Ping of Death –** During a Ping of Death (POD) attack the attacker sends multiple pings to one computer. POD attacks use manipulated packets to send packets to the network which have IP packets that are larger than the maximum packet length. These illegitimate packets are sent as fragments. Once the victim's network attempts to reassemble these packets network resources are used up, they are unavailable to legitimate packets. This grinds the target network to a halt and takes it out of action completely.
- **UDP Floods –** A UDP flood is a DDoS attack that floods the victim network with User Datagram Protocol (UDP) packets. The attack works by flooding ports on a remote host so that the host keeps looking for an application listening at the port. When the host discovers that there is no application it replies with a packet that says the destination wasn't reachable. This consumes network resources and means that other devices can't connect properly.
- **Ping Flood –** Much like a UDP flood attack, a ping flood attack uses ICMP Echo Request or ping packets to derail a network's service. The attacker sends these packets rapidly without waiting for a reply in an attempt to make the target network unreachable through brute force. These attacks are particularly concerning because bandwidth is consumed both ways with attacked servers trying to reply with their own ICMP Echo Reply packets. The end result is a decline in speed across the entire network.
- **SYN Flood –** SYN Flood attacks are another type of DoS attack where the attacker uses the TCP connection sequence to make the victim's network unavailable. The attacker

sends SYN requests to the victim's network which then responds with a SYN-ACK response. The sender is then supposed to respond with an ACK response but instead, the attacker doesn't respond (or uses a spoofed source IP address to send SYN requests instead). Every request that goes unanswered takes up network resources until no devices can make a connection.

- **Slowloris –** Slowloris is a type of DDoS attack software that was originally developed by Robert Hansen or RSnake to take down web servers. A Slowloris attack occurs when the attacker sends partial HTTP requests with no intention of completing them. To keep the attack going, Slowloris periodically sends HTTP headers for each request to keep the computer network's resources tied up. This continues until the server can't make any more connections. This form of attack is used by attackers because it doesn't require any bandwidth.
- **HTTP Flood –** In a HTTP Flood attack the attacker users HTTP GET or POST requests to launch an assault on an individual web server or application. HTTP floods are a Layer 7 attack and don't use malformed or spoofed packets. Attackers use this type of attack because they require less bandwidth than other attacks to take the victim's network out of operation.
- **Zero-Day Attacks –** Zero-Day attacks are attacks that exploit vulnerabilities that have yet to be discovered. This is a blanket term for attacks that could be faced in the future. These types of attacks can be particularly devastating because the victim has no specific way to prepare for them before experiencing a live attack.

More Information on DDoS Attacks

# Q14: What is a Port Scan

### Network Scanning for Host Discovery

The process for determining what systems are up and running and listening on a network is called Host Discovery. This is often the first step used by hackers in a hostile attack. There are two primary protocols used for host discovery: Address Resolution Protocol (ARP) scans, and various forms of Internet Control Message Protocol (ICMP) scans.

An ARP scan is the process of mapping IP addresses to MAC addresses on a local subnet. ARP requests can be sent out to all of the IP addresses on a Local Area Network (LAN) to determine which hosts are up based on the ones that respond with an ARP reply. Because ARP requests only work within a LAN, this requires the potential attacker to be connected to your internal network.

To conduct a network scan outside of the LAN, there are a number of different ICMP packets that can be used instead, such as echo, timestamp, and address mask requests. Echo or ping requests are used to detect if a host can be reached, while timestamp packets determine the latency between two hosts. You can use address mask requests to find out the subnet mask used on the network.

Discovering hosts on a network via ICMP messages all depends on receiving a corresponding reply from the targeted hosts. If no response is received, it could mean that there is no host at the target address or that the ICMP message type isn't supported by the target host. It could also mean that the original request was blocked by a firewall or packet filter. Generally, ICMP echo (ping) requests that do not originate from inside the network are blocked by firewalls, but timestamp and address mask requests are less likely to be blocked.

## Moving On to Port Scanning

Now that the network scan has been completed and a list of available hosts has been compiled, a port scan can be used to identify the in use on specific ports by the available hosts. Port scanning will typically classify ports into one of three categories:

**Open:** The target host responds with a packet indicating it is listening on that port. It also indicates that the service that was used for the scan (typically TCP or UDP) is in use as well.

**Closed:** The target host received the request packet but responds back with a reply indicating that there is no service listening on that port.

**Filtered**: A port scan will categorize a port as filtered when a request packet is sent but no reply is received. This typically indicates that the request packet has been filtered out and dropped by a firewall.

## Port Scan Methods

TCP and UDP are generally the protocols used in port scanning, as previously mentioned and there are several methods of actually performing a port scan with these protocols.

The most commonly used method of TCP scanning is SYN scans. This involves creating a partial connection to the host on the target port by sending a SYN packet and then evaluating the response from the host. If the request packet is not filtered or blocked by a firewall, then the host will reply by sending a SYN/ACK packet if the port is open or a RST packet if the port is closed.

Another method of TCP scanning is the TCP connect scan. This involves the scanner trying to connect to a port on the target host using the TCP connect system call and initiating the full TCP handshake process. This process creates a lot of overhead in terms of packets and is a lot easier to detect, therefore making it a less utilized method of port scanning.

Other types of TCP port scans include NULL, FIN and Xmas. These three types of scans involve manipulating the TCP header flags. NULL scans send packets with no flags set in their headers, while FIN scans only have the FIN bit set. Xmas scan packets have the FIN, PSH and URG flag bits turned on, making them appear to be "lit up like a Christmas tree". Hence the name Xmas scan.

UDP scans, like TCP scans, send a UDP packet to various ports on the target host and evaluate the response packets to determine the availability of the service on the host. As with TCP scans, receiving a response packet indicates that the port is open.

## Q15: What is Malware, what is a Virus and explain their differences

There are many different types of viruses. These are the three most common examples:

- 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. Fake file names tempt the recipients to open the files, activating the viruses. An old but still prominent type of malware, macro viruses, remain popular with hackers.
- Polymorphic viruses modify their own code. The virus replicates and encrypts itself, changing its code just enough to evade detection by antivirus programs.

Malware encompasses all types of malicious software, including viruses, and may have a variety of goals. A few of the common objectives of malware are:

- Trick a victim into providing personal data for identity theft
- Steal consumer credit card data or other financial data
- Assume control of multiple computers to launch denial-of-service attacks against other networks
- Infect computers and use them to mine bitcoin or other cryptocurrencies

**The five types of malware**

Besides viruses, multiple other types of malware can infect not only desktops, laptops, and servers, but also smartphones. Malware categories include the following:

- **Worms.** A worm is 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. A graduate student created the first worm (the Morris worm) in 1988 as an intellectual exercise. Unfortunately, it replicated itself quickly and soon spread across the internet.
- **Ransomware.** As the name implies, ransomware demands that users pay a ransom—usually in bitcoin or other cryptocurrency—to regain access to their computer. The most recent category of malware is ransomware, which garnered headlines in 2016 and 2017 when ransomware infections encrypted the computer systems of major organizations and thousands of individual users around the globe.
- **Scareware.** Many desktop users have encountered scareware, which attempts to frighten the victim into buying unnecessary software or providing their financial data. Scareware

pops up on a user's desktop with flashing images or loud alarms, announcing that the computer has been infected. It usually urges the victim to quickly enter their credit card data and download a fake antivirus program.

- **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. Keyloggers steal credit card numbers, passwords, account numbers, and other sensitive data simply by logging what the user types.
- **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. An example is Operation RogueRobin, which was uncovered in July 2018. Rogue-Robin is spread through Microsoft Excel Web Query files that are attached to an email. It causes the computer to run PowerShell command scripts, providing an attacker access to the system. As PowerShell is a trusted part of the Microsoft platform, this attack typically does not trigger a security alert. Some fileless malware is also clickless, so a victim does not need to click on the file to activate it.

**Antimalware and antivirus solutions**

Because so many types of malware and viruses are in the wild—and cybercriminals are creating more every day—most antimalware and antivirus solutions rely on multiple methods to detect and block suspicious files. The four main types of malware detection are:

- **Signature-based scanning.** This is a basic approach that all antimalware programs use, including free ones. Signature-based scanners rely on a database of known virus signatures. The success of the scanner depends on the freshness of the signatures in the database.
- **Heuristic analysis.** This detects viruses by their similarity to related viruses. It examines samples of core code in the malware rather than the entire signature. Heuristic scanning can detect a virus even if it is hidden under additional junk code.
- **Real-time behavioral monitoring solutions.** These seek unexpected actions, such as an application sending gigabytes of data over the network. It blocks the activity and hunts the malware behind it. This approach is helpful in detecting fileless malware.
- **Sandbox analysis.** This moves suspect files to a sandbox or secured environment in order to activate and analyze the file without exposing the rest of the network to potential risk.

IT security professionals can augment their organization's malware and virus defenses by updating and patching applications and platforms. Patches and updates are especially critical for preventing fileless malware, which targets application vulnerabilities and cannot be easily detected with antimalware solutions.

Likewise, implementing and encouraging data security best practices can be valuable in preventing data breaches. Basic best practices for password management and role-based access to data and applications, for example, can minimize the odds of a hacker gaining access to a system and limit a hacker's ability to do damage if they gain access. Regular security updates for employees can also help them spot potential threats and remind employees to practice good security hygiene.

## Q16: Other

**OS –** It is important to have a good understanding of different operating systems. There are so many resources out there to learn this but keep sure you understand both Windows and Linux. Windows is commonly used on most user endpoints whereas a lot of security tools work on Linux.

**Cyber Risk** – Understanding risk within cybersecurity is vital, and you will likely get asked about it. This post from Upguard explains it well.

**CIA Triad** – The CIA triad (confidentiality, integrity, and availability) is a model which is designed to guide policies for information security within an organization. This post from techtarget explains it well.

**Cybersecurity Tools –** There are many tools used by a security analyst. The most common of these are SIEM, IDS/IPS, Vulnerability Scanners, PAM and Anti-Virus. Get to know these, what they are and how they work. I also have some posts on this blog which will help with this so take a look.

**Incident Response –** It is important that any analysts understands how incidents are responded to, even if they are not directly involved in incident response. You may be asked what is incident response and different ways it can be performed. My personal favourite resource on this is https://www.incidentresponse.com/playbooks/

**Threat Hunting –** Threat hunting put simply is looking for cyber threats within your network which have not triggered any security alerts. This post from crowdstrike sums it up well – https://www.crowdstrike.com/epp-101/threat-hunting/. I also have various threat hunting guides throughout my blog, so check them out.

**Threat Intelligence –** Threat intelligence is data collected and analyzed by an organization in order to understand a threat actor's motives, targets, and attack behaviors. This post from crowdstrike sums it up well – https://www.crowdstrike.com/epp-101/threat-intelligence/. Interviewers may ask what the difference is between Threat Hunting / Intelligence so watch out for that.

**Classification –** Keep sure you understand True vs. False and Positive vs. Negative. This post from google explains it well – https://developers.google.com/machine-learning/crash-course/classification/true-false-positive-negative

## Q17: Non-Technical

There are also a number of common non – technical questions which appear to come up quite a lot:

- Most common by far is being asked about how you keep up to date with cyber security. This can include things such as twitter, blogs, podcasts, CTFs and reading books. Also keep sure you know some recent news to show them you keep up to date.
- Can you explain technical points to non-technical people?
- Research the company, they will likely ask what you know about them.