Port Scanning with OS Information/Version

1. What is Port Scanning?

TCP/IP (Transmission Control Protocol/Internet Protocol) is the internet protocol suite which is the conceptual model and set of communication protocols used in the Internet and similar computer networks. In TCP/IP model, there is a notion of **IP Address** which is a **Layer-3/Network-layer Address** and used to reference **Server/Host Machines** in networks. In addition to IP address, there is also a notion of **Port** which is a **Layer-4/Transport-layer Address** and used to connect and get services from specific application running on a server/host machine in the network.

Port scanning is a systematic approach to probe a server/host machine for **Open Ports**. It, itself, is not a **Network-level Threat**. Network administrators may do port scanning to verify security policy of the network for improving overall security posture of the organizational network. But, attackers may also do port scanning to identify network services running on a server/host machine and exploit their **Vulnerabilities**. The attackers use port scanning as a vehicle for carrying out **Reconnaissance Attack**, that is, they try to get information about a particular network so that they can identify how the network may behave if **Malicious Connections** are made. If they can find an open port corresponding to a specific service and the vulnerability of that service is known, then they can attack that service via open port.

2. Ports on a Server/Host Machine

Usually, 16 bits are used to encode and identify **Port Number** on a server/host machine. Therefore, there are 2^{16} = 65,536 possible ports. But, `port 0` is left unassigned and used by **Operating System** to accept request for allocating and opening any currently unused port. Hence, there are, in total, 2^{16}-1 = 65,535 ports on a server/host machine.

Most **Internet applications** run on **Well-defined Ports**. Example of associations between Internet applications and server/host machine ports is given below:

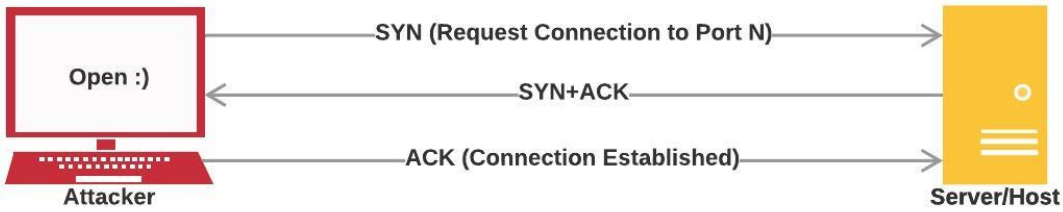
Port Number	Internet Application Running
`port 21`	FTP (File Transfer Protocol)
`port 22`	SSH (Secure Shell Protocol)
`port 23`	Telnet Protocol
`port 25`	SMTP (Simple Mail Transfer Protocol)
`port 53`	DNS (Domain Name System)
`port 80`	HTTP (Hypertext Transfer Protocol)
`port 443`	HTTPS (HTTP Secure)
`port 1433`	SQL Server

Normally, all ports on a server/host machine are not scanned during port scanning. Instead, ports with known applications running or known exploitable vulnerabilities are cherry-picked and scanned.

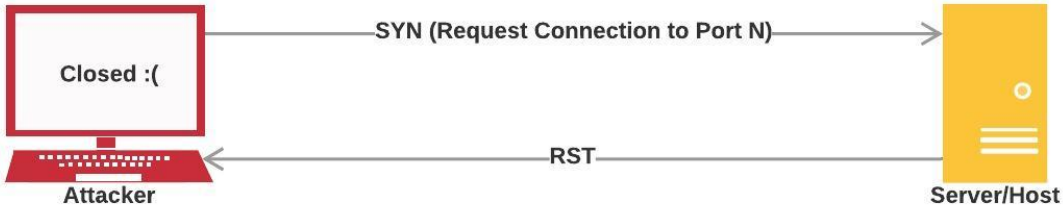
Typically, port scanning works by trying to **Systematically** establish connections to all available ports on a given server/host machine. When such an attempt to establish connections is made, one may get one of the three different types of **Responses** from the corresponding port on server/host machine. These responses are:

1. `open/accepted`: service in question is **Listening for New Connections** on specified port
2. `closed/denied`: service in question is **Not Listening for New Connections** on specified port
3. `filtered/dropped/blocked`: **No Reply** from specified port running service in question

Response: Open



Response: Closed

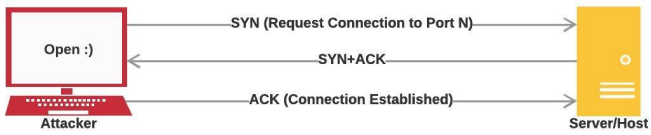


Response: Blocked

3. Implementation

There are many techniques that are typically used to implement port scanning. These techniques are discussed in the following subsections.

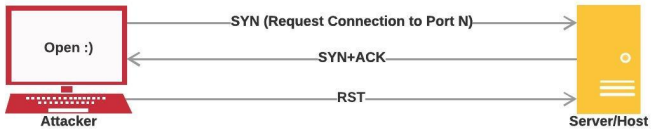
3.1 TCP `connect()` (Full-open) Port Scanning



TCP Full-open Port Scanning

This is the most common and simplest technique used in port scanning implementation. **TCP** `connect()` is a **System Command** used to establish connection with server/host machine through a port. The connection is established via a **Three-way Handshaking** scheme as depicted in the above diagram.

3.2 TCP `SYN` (Half-open) Port Scanning



TCP Half-open Port Scanning

TCP `SYN` port scanning works in a similar fashion to **TCP** `connect()` port scanning. The difference is that after receiving `SYN+ACK` from server/host machine, the attacker responds with **TCP** `RST` instead of **TCP** `ACK` to terminate the connection.

3.3 TCP `connect()` Port Scanning vs TCP `SYN` Port Scanning

The problems with TCP `connect()` port scanning are discussed below:

- During TCP three-way handshaking, server/host machine allocates some space in its memory to store details on the connection at targeted port and state of the three-way handshaking. If attacker stops responding and does not complete TCP three-way handshaking, then server/host machine will not release previously allocated memory for that connection. Eventually, server/host machine may run out of memory and stop working.
- From attacker's perspective, an attempt to scan a port on server/host machine should go unnoticed. **IPS (Intrusion Prevention System)** looks for number of ports on server/host machine with fully established connection (connection established via three-way handshaking) and decides on whether the machine is a victim of port scanning attack. Hence, completing three-way handshake risks the port scanning detection.
- As soon as the three-way handshake is completed, the application running on the targeted port takes control of the established connection. Consequently, the application as well as the entire server/host machine may crash if attacker tries to abruptly terminate the connection.

Therefore, an attacker may not prefer TCP `connect()` port scanning. TCP `SYN` port scanning tackles the issue with TCP `connect()` port scanning.

3.4 Other Port Scanning Implementation Techniques

3.4.1 TCP `FIN` Port Scanning

Often, **Firewall** and **Packet Filters** monitor specified ports so that the previously mentioned techniques of TCP port scanning can be detected. To bypass this detection, **TCP Packet with FIN Flag** may be used. Usually, this packet is used to finish an established connection with server/host machine. If TCP FIN packet is sent before establishing a connection, an open port discards the packet and a closed port replies with a **TCP Packet with RST Flag**. So, open and closed ports can be distinguished between by looking at the TCP RST reply from the target port on server/host machine. Though some systems, regardless of whether the port is open, reply to TCP FIN packet with TCP RST packet to prevent this type of scanning method, TCP FIN port scanning may come in handy for many systems.

3.4.2 UDP Raw ICMP Port Unreachable Scanning

This port scanning implementation technique differs from the aforementioned techniques in that it uses **UDP (User Datagram Protocol)** instead of TCP. While this protocol is simpler compared to TCP, scanning with this protocol is significantly more difficult. This is because open ports are not obliged to send an acknowledgement in response to the scanning probe and close ports are not even required to send an error packet. Fortunately, most server/host machines do send an `ICMP_PORT_UNREACH` error when a packet is sent to a closed UDP port. Thus, open and closed ports can be distinguished between by looking at the receipt of `ICMP_PORT_UNREACH` error. Again, neither UDP packet nor **ICMP (Internet Control Message Protocol) Error** are guaranteed to arrive. So, UDP port scanner of this type must also implement **Packet Retransmission Mechanism**. This port scanning implementation technique is very slow. Again, this technique requires the attacker to have **Root Privileges**.

3.4.3 UDP `recvfrom()` and `write()` Port Scanning

An attacker without root privileges can not read `ICMP_PORT_UNREACH` error directly. But, **Linux** can indirectly notify the attacker when the error arrives. This port scanning implementation technique is used to check whether the port is open.

The reason behind so many implementation techniques for port scanning being developed by attackers is that any of these techniques may potentially get detected by network security systems like **IDS (Intrusion Detection System)**, **Firewall**, etc. Therefore, many implementation techniques for port scanning have been developed over the time so that an attacker can successfully bypass the detection and carry out the desired attack on server/host machine. Each of these implementation techniques has its own subtle method along with pros and cons.

4. OS Information/Version Detection

Sometimes, inside a network, it is beneficial to know the **Operating System (OS)** of a server/host machine. Accessing a machine is easier when the OS it is running on is known to the attacker because attacker can, then, specifically search for known **Security Vulnerabilities** in that OS. Though, these security vulnerabilities are usually patched quickly, but the attacker just needs to know when and where a certain security vulnerability exists.

4.1 OS Detection Database

Each OS has unique characteristics associated with its **TCP/IP Stack Implementation** which may serve to identify the OS in a network. **TCP/IP Stack Fingerprinting** is the passive collection of configuration attributes from a **Remote** server/host machine during **Standard Transport Layer (Layer 4) Network Communications**. The combination of attributes may then be used to infer the OS running on that remote machine (known as **OS Fingerprinting**). There are many techniques and methods for getting a **Good Estimate** of the OS running on a certain remote server/host machine. Two of these techniques and methods are discussed in the following subsections.

4.1.1 Active Fingerprinting

Special Probe Packets are sent to certain server/host machine and based on the response from that remote machine, certain OS is assumed to be running on that machine.

4.1.2 Passive Fingerprinting

Legitimate Traffic is analyzed and then, compared for certain key differences in the TCP/IP stack implementation on different types and versions of OS.

NMAP (Network Mapper) is a free and open-source **Network Scanner** created by *Gordon Lyon*. This tool offers a **Database** which is installed together with NMAP. This database is named `nmap-os-db` and used when carrying out OS detection through OS fingerprinting. This database does not automatically get updated. The updated version can be found on the Internet from svn.nmap.org/nmap/nmap-os-db.

4.2 OS Detection Process

There are five different **Probes** being performed during OS fingerprinting. Each of these probes may consist of one or more **Packets**. The response to each packet from the target server/host machine helps the attacker to estimate the type and version of OS. The five different probes are discussed in the following subsections.

4.2.1 Sequence Generation

The **Sequence Generation Probe** consists of six packets. These packets are all **TCP** `SYN` **Packets** and sent 100 milliseconds apart from one another. The response for each TCP SYN packet contributes to estimating the type and version of OS.

4.2.2 ICMP Echo

The **ICMP Echo Probe** consists of two packets. These packets are sent to the target server/host machine with varying settings. The responses contribute to verifying the type and version of OS.

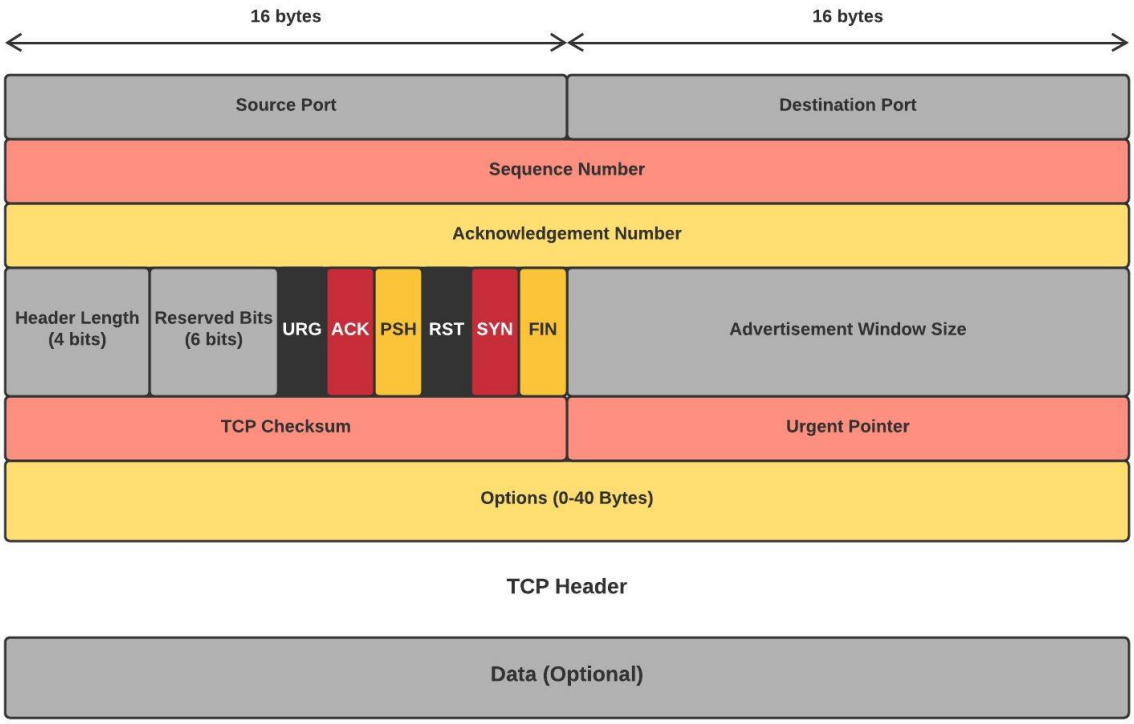
4.2.3 TCP Explicit Congestion Notification

When many packets are being generated and passing via certain router, the router becomes burdened and this situation is known as **Congestion**. Consequently, server/host machine slows down to reduce congestion so the router can stop dropping packets. Remote machine responds only for delivered packets. Since each OS handles these packets in different ways, specific values returned contribute to estimating the type and version of OS running on remote machine.

4.2.4 TCP

The **TCP Probe** consists of six packets. Each packet is sent to either open or closed port on remote server/host machine with specific settings. Responds from remote machine vary depending on the OS running on it. The **TCP Packets** are all sent with varying **Flags** as follows:

1. No Flags
2. `URG`, `PSH`, `SYN`, and `FIN`
3. `ACK`
4. `SYN`
5. `ACK`
6. `URG`, `PSH`, and `FIN`



4.2.5 UDP

The **UDP Probe** consists of a single packet sent to a closed port on remote server/host machine. If that port on the target machine is closed and an **ICMP Port Unreachable Message** is returned, then there is no firewall deployed on server/host side.

5. Why These Attacks Work

Port Scanning is a common technique used by the attackers to go and discover open ports on a remote server/host machine so that they can exploit the services with known vulnerabilities running on those ports. This attack can also reveal whether active network security measures like firewalls are being deployed in an organizational network.

On the other hand, **OS Fingerprinting** is the process used by the attackers to infer the type and version of OS running on a target server/host machine so that they can easily exploit the vulnerabilities associated with that OS.

These attacks are carried out by sending trivial networking packets from attackers to victim server/host machine. As a result, it becomes quite difficult for the active network security measures to distinguish between normal network traffic and malicious network traffic. Also, many implementation techniques have been developed over the years for carrying out these attacks. As a result, attackers have an arsenal of attacking methods at their disposal to bypass detection by these active network security measures.