Security Audit of Network Time Protocol

Sudipta Halder
M.Tech CSIS
Roll: 2021202011

Most computers have clocks

- Everything from the clock widget on your desktop, meeting reminders, posts on social media to online shopping use the time.
- Many devices have a battery backed Real Time Clock (RTC), like you PC. But many devices don't, like the Raspberry Pi. Smartphones generally rely on the main battery having enough power to keep the clock going, even when the phone switches off due to lack of power.
- In all cases these devices need to check and synchronize their clocks to make sure the time is accurate.

Clocks drift

- Computers clocks tend to drift.
- Switched off synchronization on a PC which started to lose about 0.1 second per day.
  - After a couple of weeks that will be approaching 2 seconds.
- Have a PC not connected to the Internet.
  - During the course of the year the clock is wrong by several minutes.

Some computers need high accuracy

- Financial transactions (stock market etc).
- Telecommunications.
- Electric power distribution.
- Data centers.
- Distributed databases.

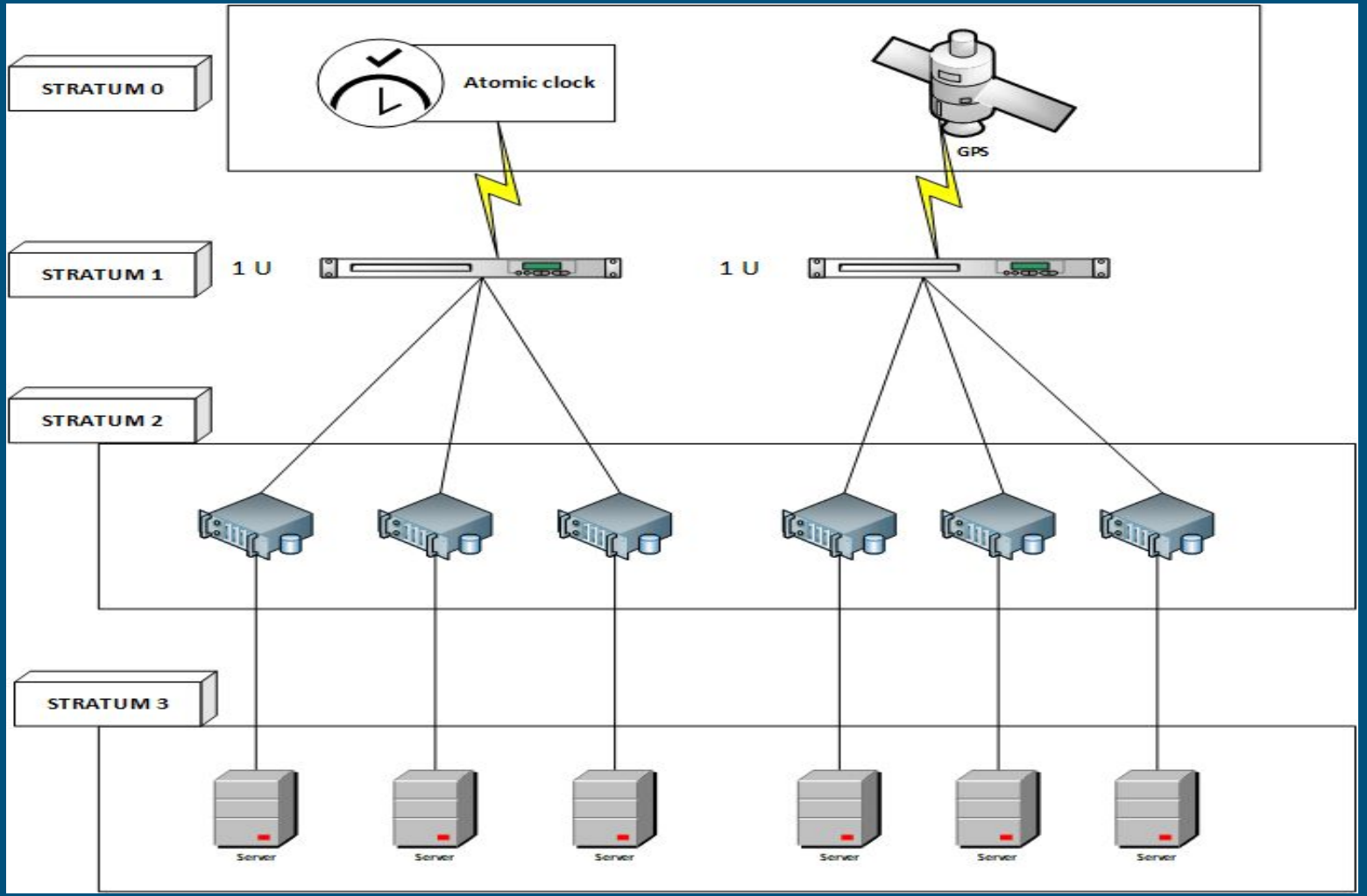
How do any of us know the time?

- Nowadays the gold standard is generally Global Navigation Satellite Systems (GNSS): GPS, GLONASS, Galileo, etc.
  - But atomic clocks and radio clocks are also high-precision timekeeping devices.

- These are called Stratum 0 devices. They are broadcast devices.

How do any of us know the time?

- Servers which synchronize with Stratum 0 devices and offer time synchronization services over a network are called Stratum 1 devices or primary time servers.


STRATUM 0
Atomic clock
GPS
STRATUM 1
1 U
1 U
STRATUM 2
STRATUM 3
Server
Server
Server
Server
Server
Server

The NTP Stratum Model

- The NTP Stratum model is a representation of the hierarchy of time servers in an NTP network, where the Stratum level (0-15) indicates the device's distance to the reference clock.

The NTP Stratum Model...

- Stratum 0 means a device is directly connected to e.g., a GPS antenna. Stratum 0 devices cannot distribute time over a network directly, though, hence they must be linked to a Stratum 1 time server that will distribute time to Stratum 2 servers or clients, and so on.

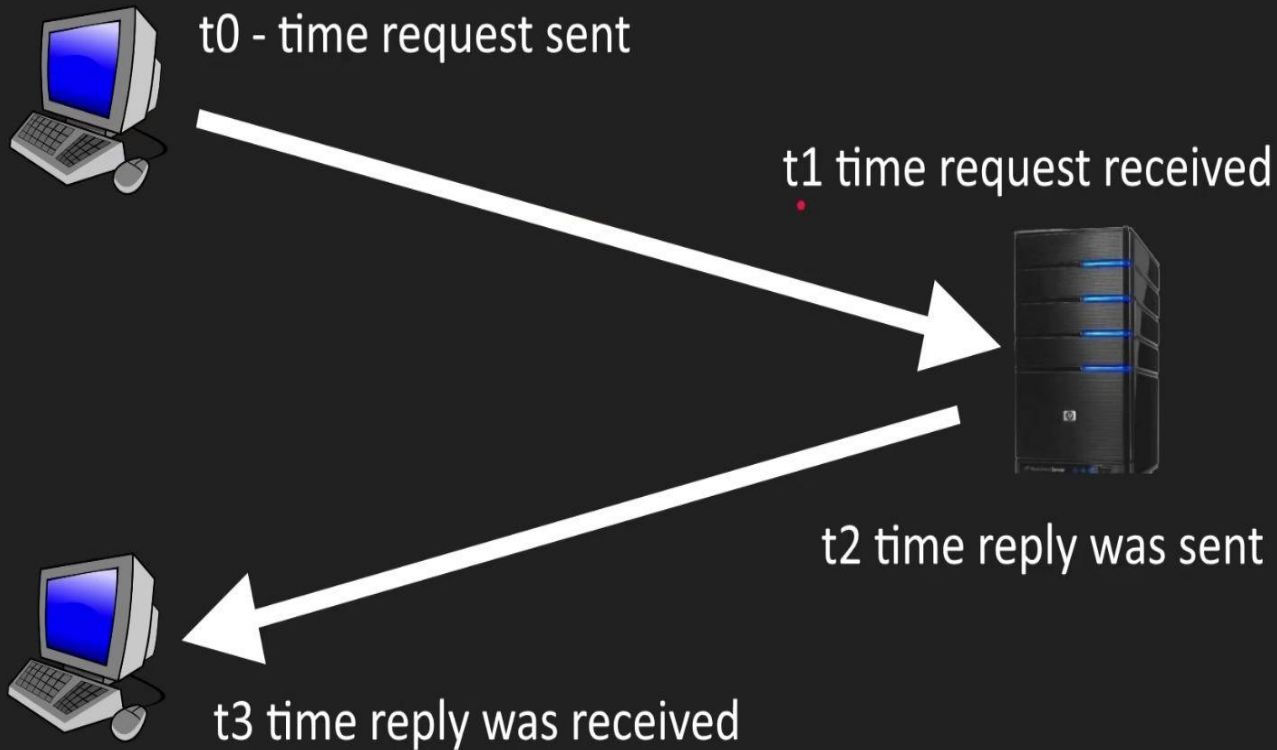
- The NTP protocol does not allow clients to accept time from a Stratum 15 device, hence Stratum 15 is the lowest NTP Stratum.

Network Time Protocol

- The most common system for synchronizing clocks on servers, PCs, laptops, smartphones, tablets, smart TVs, etc is NTP.
- NTP is one of the oldest Internet protocols in current use. NTP was designed by David L. Mills of the University of Delaware in the 1980s.
- Typically a NTP client polls one or more NTP servers and requests the correct time.

NTP in action!!

Next slide


t0 - time request sent
t1 time request received
t2 time reply was sent
t3 time reply was received

Maths

- t0 is the client's timestamp of the request packet transmission,
- t1 is the server's timestamp of the request packet reception,
- t2 is the server's timestamp of the response packet transmission and
- t3 is the client's timestamp of the response packet reception

- `Offset = ((t1 - t0) + (t2-t3)) / 2`

- `Round Trip Delay = (t3 - t0) - (t2 -t1)`

Example

Offset = ((t1 - t0) + (t2-t3)) / 2
Round Trip Delay = (t3 - t0) - (t2 -t1)

- It is 17:01:00 on client. It is 17:01:30 on server (correct time)

- *t0* is 17:01:00 (17:01:30)
  - 2s
- ***t1*** is 17:01:32 (17:01:32)
  - 1s
- *t2* is 17:01:33 (17:01:33)
  - 2s
- *t3* is 17:01:05 (17:01:35)

- Offset = (32 + 28) / 2 = 30s
- Delay = 5 - 1 = 4s round trip (2s one-way trip)
- So at *t3* client can set correct time to:
  - t2 + one-way delay = 17:01:35
  - t3 + offset: 17:01:35

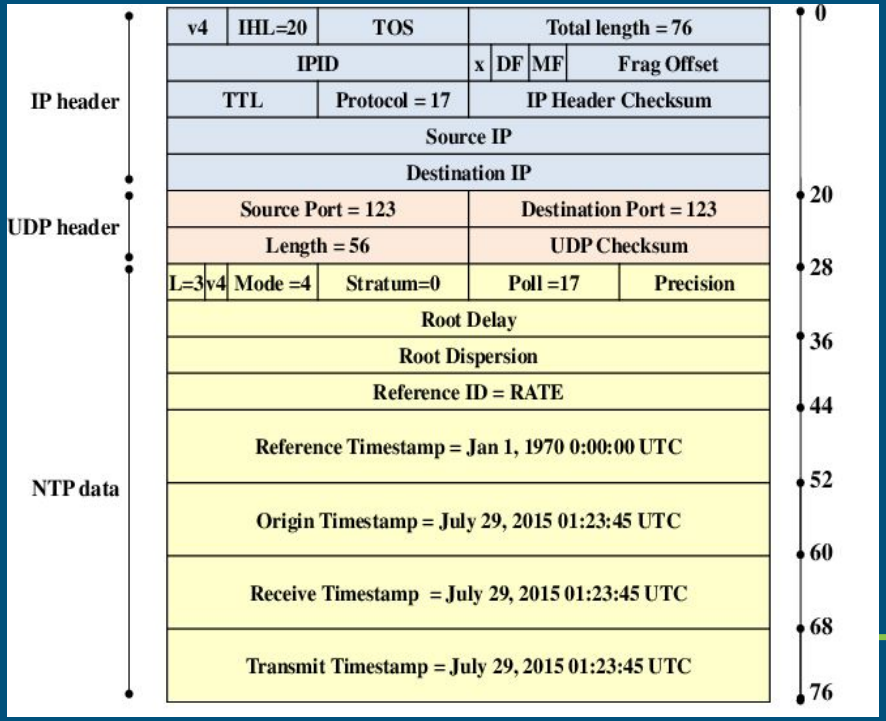
Security Attacks on NTP

- Denial of Service by spoofed Kiss-of-Death Packet.
- Denial of Service by priming the pump.
- NTP Amplification Attack.
- Time Shifting by Reboot.
- Delay attacks.

Denial of Service by spoofed Kiss-of-Death Packet

- Kiss-o-Death packets are used by NTP servers to rate-limit ntp client requests that query too frequently. By sending a KoD packet with a larger poll value and a spoofed source IP address for any pre-configured NTP servers, a remote, unauthenticated attacker could disable NTP on a targeted system.

Kiss-of-Death Packet Structure

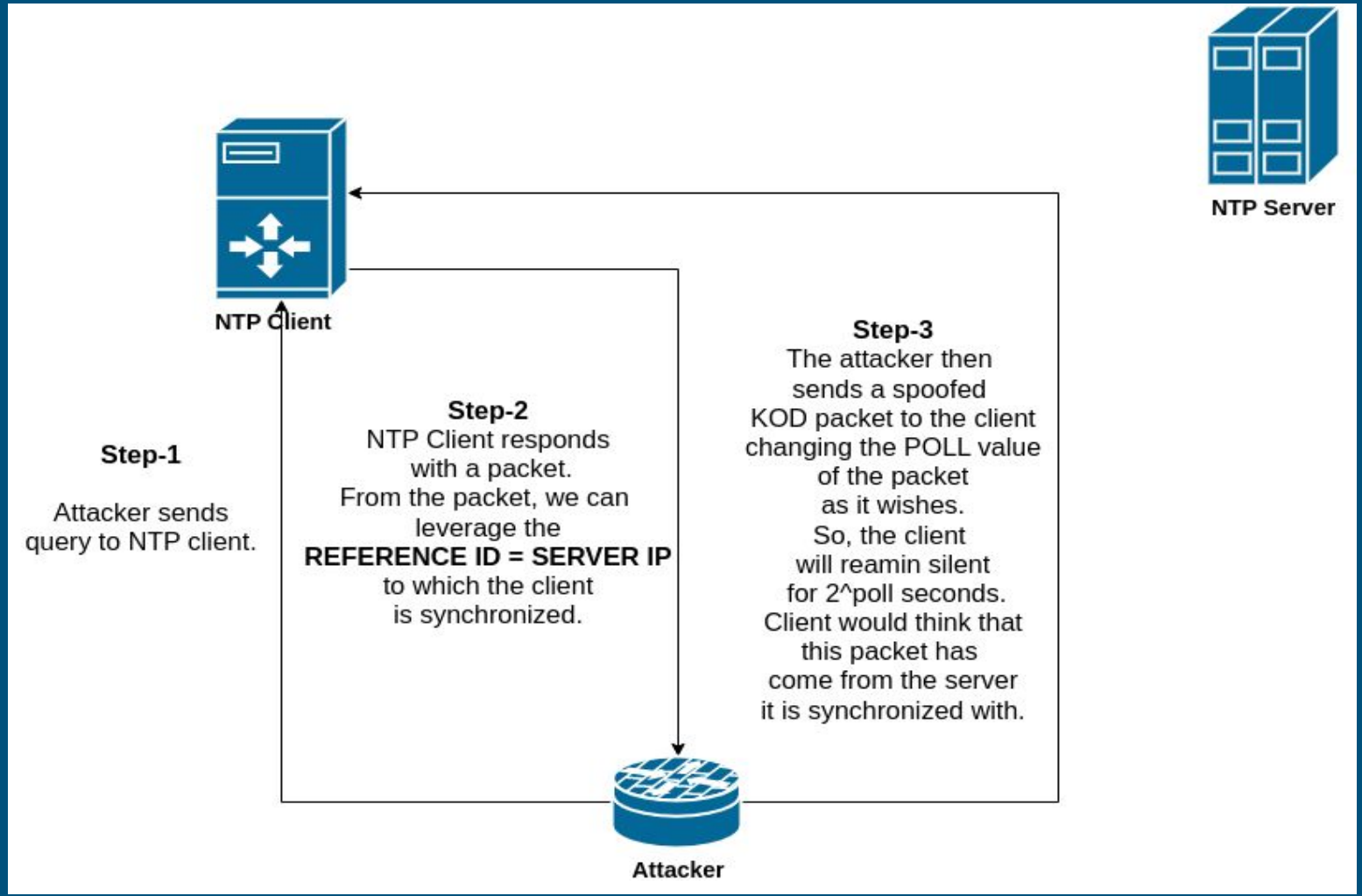


Poll Field

- We can control the the field Poll.
- 8 bits field. Highest possible value = 255.
- 2^255 seconds = some years.
- This much amount of time a NTP client can be kept silent when it receives KOD packet.
- According to RFC, highest permissible value = 17.
  2^17 seconds = 36 hours.

Challenge to perform this attack!!

- We need to know the IP of the NTP Server to which the NTP Client is connected to send the KOD packet.
- Because the NTP client should feel that the packet is coming from the server it is connected to.


NTP Server
NTP Client
Step-1
Attacker sends query to NTP client.
Step-2
NTP Client responds with a packet.
From the packet, we can leverage the
REFERENCE ID = SERVER IP
to which the client is synchronized.
Step-3
The attacker then sends a spoofed KOD packet to the client changing the POLL value of the packet as it wishes.
So, the client will reamin silent for 2^poll seconds.
Client would think that this packet has come from the server it is synchronized with.
Attacker

Mitigation

- Upgrade to ntpd v4.2.8p4.
- To see what ntpd version you are running, log into to your NTP server and type ntpq and then rv.

Denial of Service by priming the pump

- Bother NTP Server too much by faking as a client.
- So, when legitimate query comes from NTP client, server says go away with KOD packet and client will be quiet again.

Denial of Service by priming the pump...

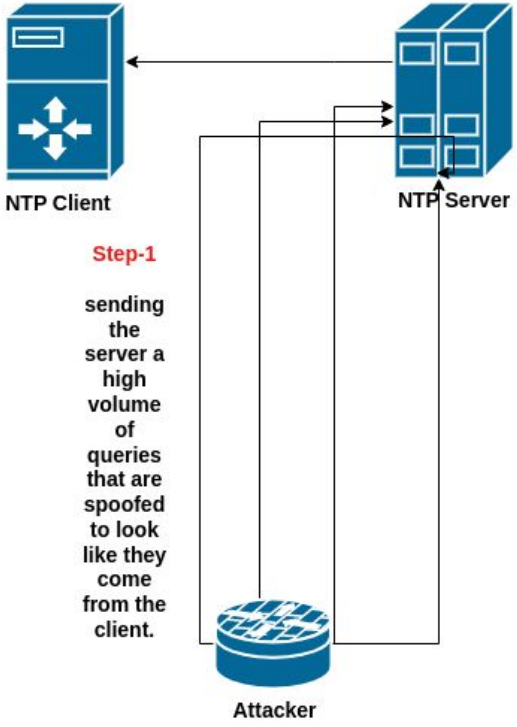
- Here, we can't control the polling time. NTP Server will decide how much time it would take to make the client silent.
- Generally, 15 mins. So, we need to do this every 15 minutes.

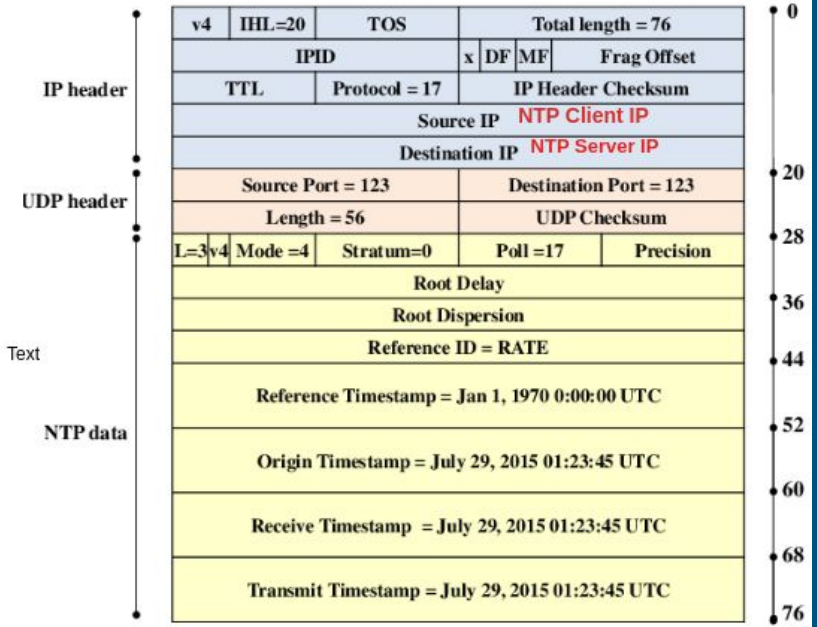
Denial of Service by priming the pump...

- This attack has the same effect as Attack 1, but requires the attacker expend more resources by sending more packets to the client's servers.

Step-2

The servers then start rate-limiting the client, responding to each of its subsequent queries with a valid KoD packet. Upon receipt of the KoD, the client stops querying its servers, and can no longer update its local clock.





Mitigation

- Upgrade to ntpd v4.2.8p4.
- To see what ntpd version you are running, log into to your NTP server and type ntpq and then rv.
- Also, monitor the system log for error messages of the form "receive: Unexpected origin timestamp from %s", which could indicate that you are subject to a priming-the-pump attack.

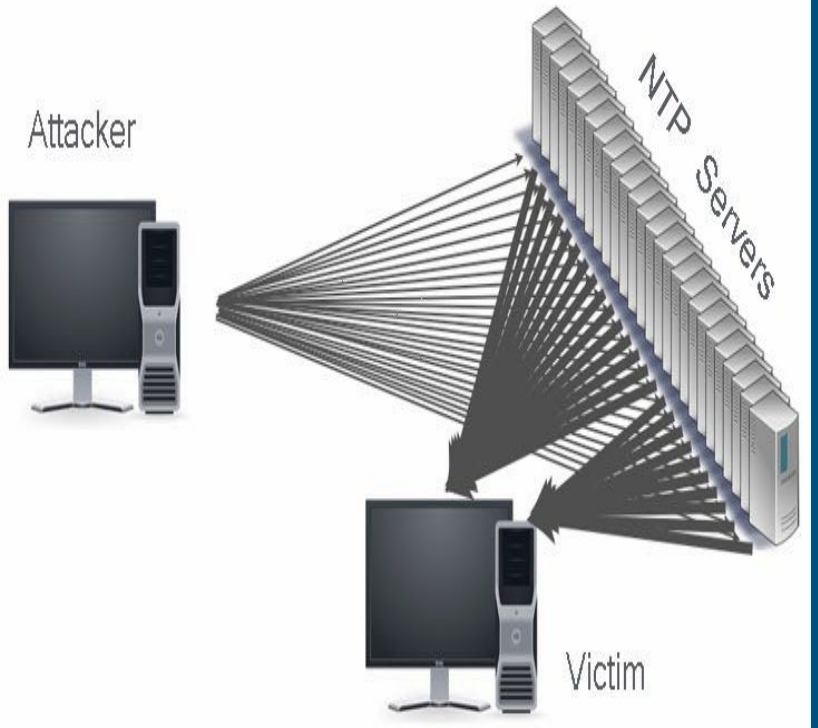
NTP Amplification Attack

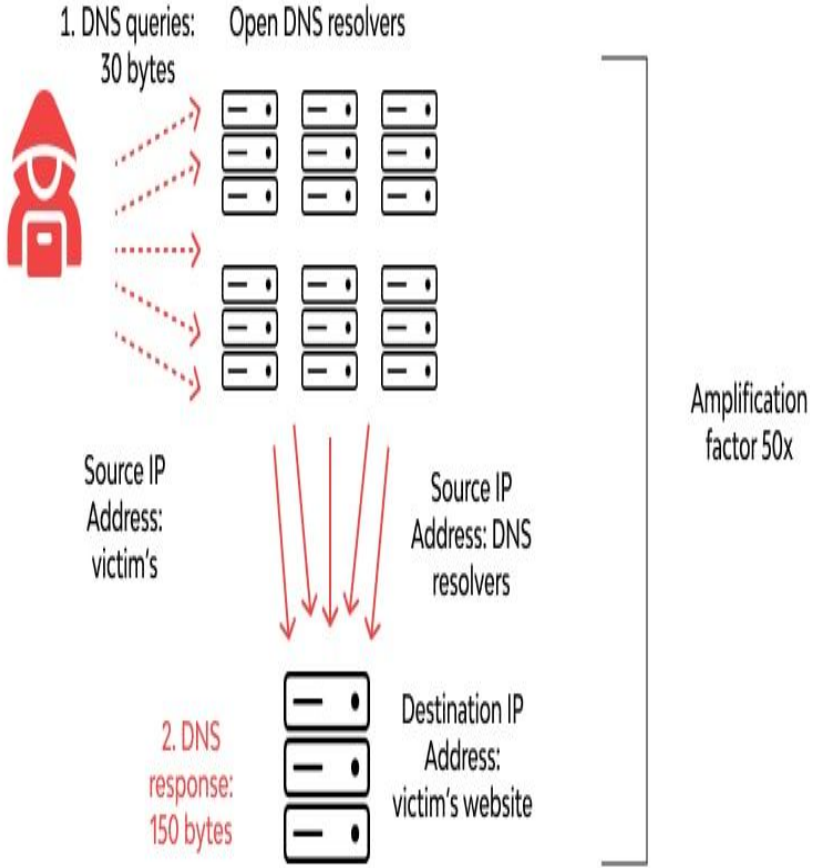
- NTP includes a monlist function, also known as MON_GETLIST, which is mainly used to monitor NTP servers.
- After the NTP server responds to the monlist, it returns the IPs of the last 600 clients that have performed time synchronization with the NTP server; the response packages are split every 6 IPs, and there will be up to 100 response packets eventually.

Key Idea

- Sending 1 packet to server with monlist command as an attacker, server sends 100 packets to the victim client.
- Now, if we send this to 10 servers, the victim client will get 1000 packets against 1 packet of ours!!!
- Huge amplification!!!

Abusing Network Time Protocol (NTP) to perform massive Reflection DDoS attack





NTP Amplification Attack Steps

1. The attacker constructs the monlist command and sets the source address as the IP address of the attacked target.

2. The attacker sends the monlist command constructed in step 1 to the NTP servers open in the network.

NTP Amplification Attack Steps

- The NTP servers receive the monlist command, perform the operations, and send responses to the source address (i.e. the attacked IP) in the monlist command package.
- The target address receives a large number of responses from the NTP servers for no reason, the network bandwidth is blocked, and the normal processing processes are affected.

NTP Amplification Attack Steps

- In an NTP amplification attack, the query-to-response ratio is anywhere between 20:1 and 200:1 or more.
- This means that any attacker that obtains a list of open NTP servers (e.g., by using tool like Metasploit or data from the Open NTP Project) can easily generate a devastating high-bandwidth, high-volume DDoS attack.

Mitigation

- **Upgrade version:** For versions after the ntpd-4.2.7p26 version, the "monlist" feature has been disabled, replaced by the "mrulist" feature, which uses mode6 to control packets and implements a handshaking process to block the magnified attack from the third-party machine.

- **Modify configuration:** For versions prior to 4.2.7, the disable monitor option may be added to the ntp.conf file to disable the monlist function. restrict ... noquery or restrict ... ignore can also be used to limit the source address of the ntpd service response.

Delay Attacks

- Can’t be solved even with the best cryptography.
- NTP packets delayed by an attacker are still valid and difficult to detect.
- This leads to asymmetric packet runtimes, which results in a systematic time offset of the client.
- Hence, an attacker can always delay time of client.

Mitigation

- An effective countermeasure is a limitation of the packet round-trip time (RTT).
- The maximum possible time offset of the clients corresponds to about half of the RTT.
- If the maximum permitted runtime is set to 300 milliseconds, the possible time offset of the client is limited to approximately +/-150 milliseconds.

Time shifting by Reboot

- NTP clients do a good job at detecting and ignoring packets that indicate a large time shift.
- RFC 5905 calls this a panic threshold, which is set by default to 1000 s (16 minutes).
- However, the RFC also states that the NTP client should '**quit**' when it sees a large time shift like this.

Time shifting by Reboot...

- The problem with this is that most modern operating systems will **'restart'** a vital service like NTP when it quits after a significant time shift.
- When the NTP daemon restarts, it may be configured to ignore this threshold (Many systems invoke ntpd with the -g flag).
- This option allows the time to be set to any value without restriction.

Time shifting by Reboot...

- All anyone wishing to spoof a client needs to do is continually send data packets to it with time information indicating a large time shift.
- The NTP service will shut down and be restarted by the OS.
- It will then **accept the time supplied** in that forged data packet because it now temporarily ignores the panic threshold (just after reboot).

Mitigation

- Actively monitor system logs. Several NTP clients restarting at the same time may be an indication that a server is not being honest.
- Don’t run ntpd with the -g option enabled.
(NOTE: most operating systems run ntpd with -g as the default) (So, it’ll check panic threshold even after reboot).
- If using multiple NTP servers, increase the minimum number of servers required before the NTP clients adjust the clocks.

References

- https://www.youtube.com/watch?v=WX5E8x3pYqg
- https://www.cs.bu.edu/~goldbe/papers/NTPattacks.html
- https://weberblog.net/network-time-security-strengths-weaknesses/
- https://support.radware.com/app/answers/answer_view/a_id/17744/~/security-advisory%3A-ntp-vulnerabilities-october-2015

References

- https://safran-navigation-timing.com/how-to-protect-your-ntp-server-from-cyberattacks/
- http://events17.linuxfoundation.org/sites/events/files/slides/vangundy-ntp-security.pdf
- https://www.bodet-time.com/time-servers/articles-and-resources/1755-maintaining-network-security-when-using-ntp-time-synchronization.html