amsix

Fighting DDoS attacks @ AMS-IX: A story of pain and tears

Stavros Konstantaras
GRNOG 15
25-10-2023

Few intro words

- **About me**
  - Sr. Network Engineer @ NOC (~7yr)
  - MSc in System & Network Engineering from UvA
  - Main focus in big technical projects (design, implement & operate)
  - Active member of RIPE, EURO-IX, NLNOG & GRNOG

- **About AMS-IX**
  - 16 locations in NL
  - 11,6 Tbps of traffic,
  - ~ 870 ASNs
  - ~1300 MACs
  - Route Servers export
    - ~297.000 IPv4
    - ~ 88.000 IPv6
  - Our own stub network
    - AS1200

Types of DDoS attacks

- **Volume based Attacks (Gbps)**
    - UDP/ICMP/other floods
    - End goal is to saturate the bandwidth

- **Protocol Attacks (Pps)**
    - SYN floods/Ping of Death/Fragmented Packets etc.
    - End goal is to consume host resources but also resources from intermediate nodes

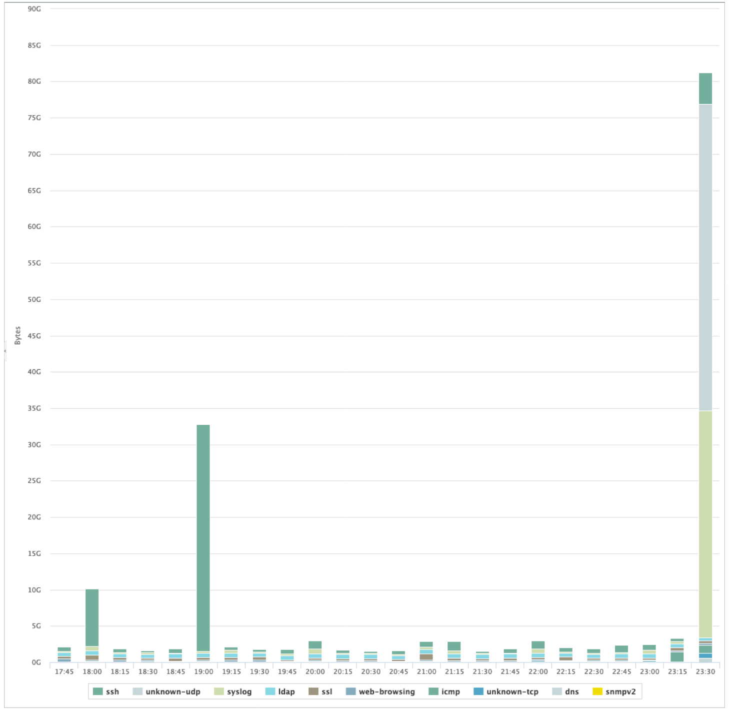
- **Application Layer Attacks (Rps)**
    - GET/POST attacks, OS vulnerabilities, etc. etc.
    - End goal is to make the software handling the requests to crash.

Our DDoS attack saga

- **It all started back in June 2020**
  - Sudden disruption of office connectivity & VPN users
  - No email/instant messaging/Nagios/DNS/transit, etc. (therefore, no access to internal and external resources)
  - We became blacked out for several minutes and then recovery was happening by itself (but very slowly)
  - AMS-IX customers and peering LAN were not affected though, transit and BGP sessions didn’t flap either.
  - And after that incident, every month the same story … ☹

The anatomy of our attack

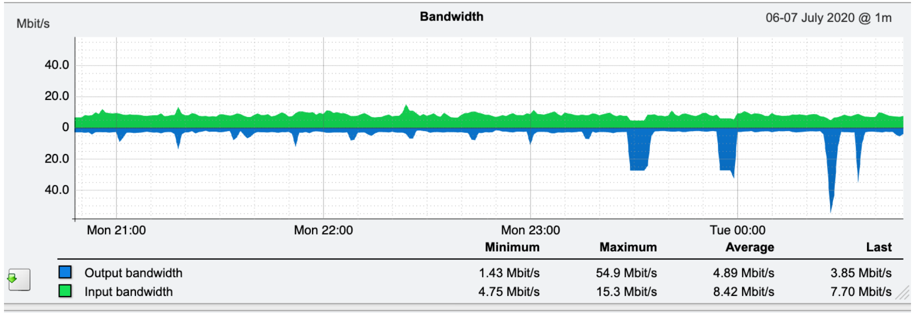
- **UPD at destination port 53 (small to medium size packets)**
- **Destination IP 185.55.136.36 (our public facing nameserver)**
- **Source IP: <*>**
- **Source port: <*>**



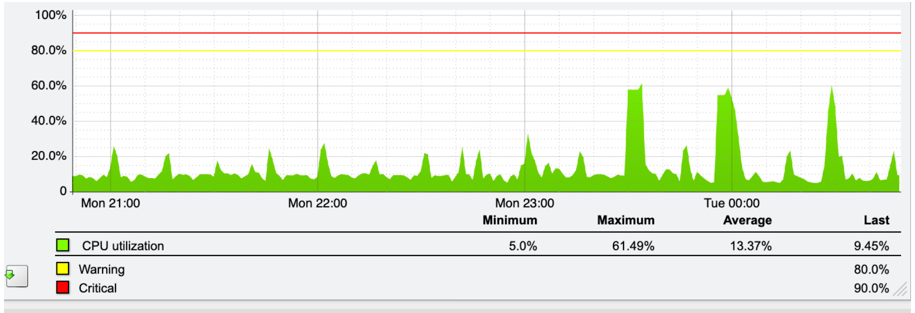
- **Overall volume of the traffic was few Mbps!!!**

Here comes the puzzle

- **If DDoS attack is only few Mbps, then how did our network collapse?**
  - Is there a bottleneck on the network?
  - Did all nameservers collapse simultaneously?



	Minimum	Maximum	Average	Last
🟦 Output bandwidth	1.43 Mbit/s	54.9 Mbit/s	4.89 Mbit/s	3.85 Mbit/s
🟩 Input bandwidth	4.75 Mbit/s	15.3 Mbit/s	8.42 Mbit/s	7.70 Mbit/s

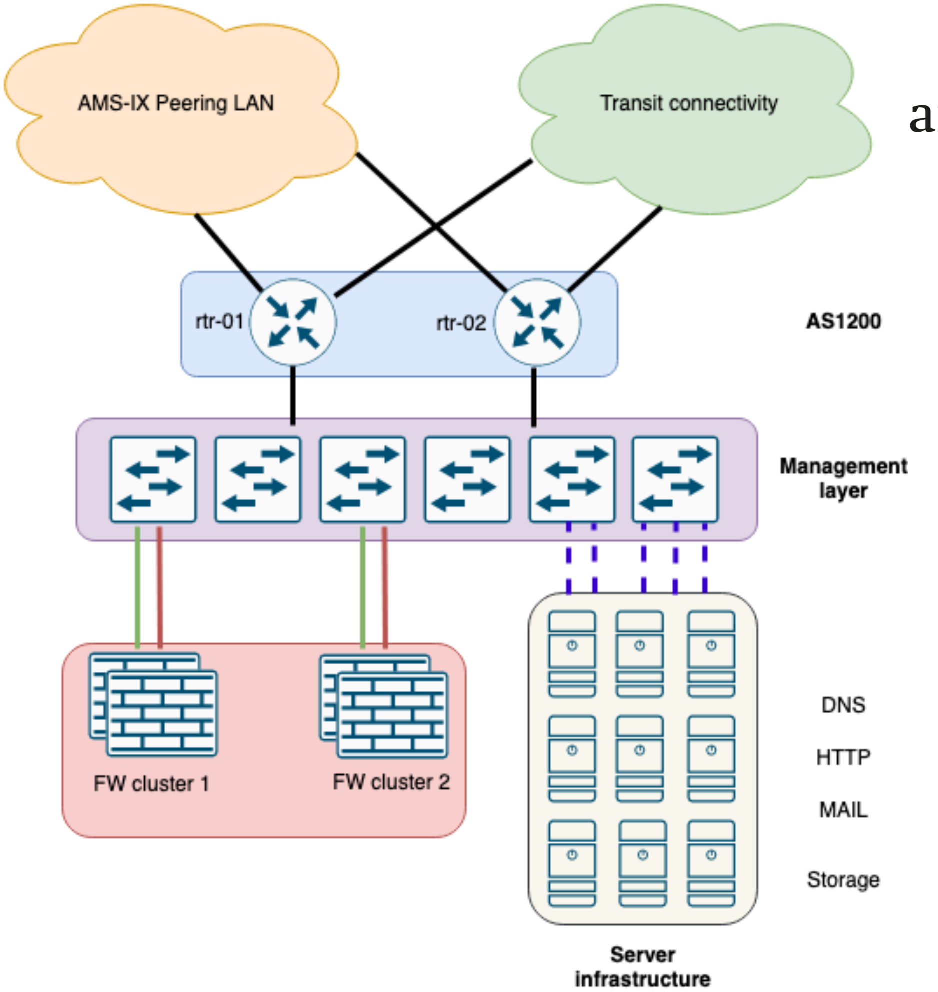


	Minimum	Maximum	Average	Last
🟩 CPU utilization	5.0%	61.49%	13.37%	9.45%
🟨 Warning				80.0%
🟥 Critical				90.0%

Overview of our Admin Network

Key components are:

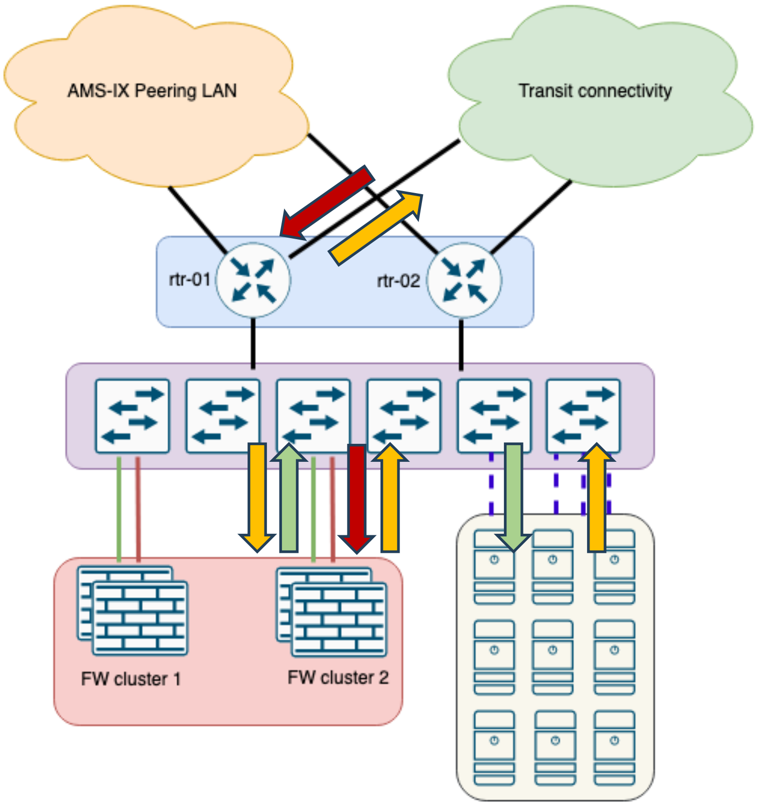
- 2 Cisco ASR 1001 Routers
- 2 firewall clusters of 2 PA 3050 (act/pass)
- A management layer of several Dell switches running Pluribus OS in a spine/leaf topology utilizing fabric technology
- Redundant Nameservers running on PowerDNS



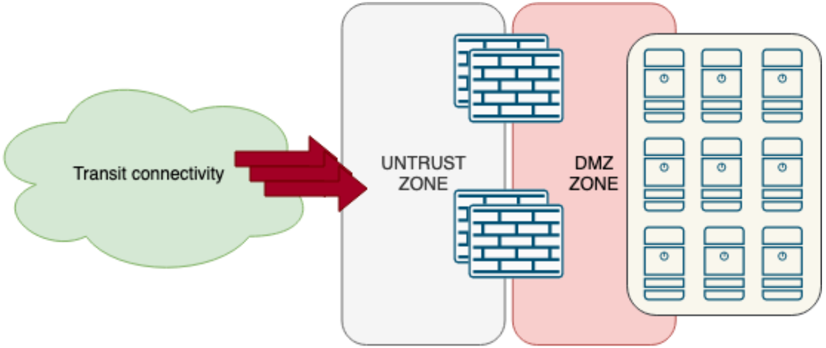
Handling incoming requests

Let's take a DNS query for example

1. Query arrives at border router.
2. Border router performs initial check, forwards the packet to the firewall.
3. Firewall performs in-depth check of the query packet.
4. If valid, query packet is forwarded to DNS server.
5. DNS server crafts a reply and sends it to default gateway.
6. Firewall receives the response, registers it and forwards it to border router.
7. Border router sends it to next hop.

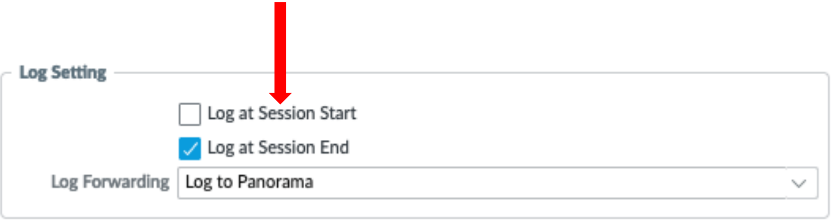


A look in the security zones



- DNS requests coming from public internet are placed in the untrust zone
- They are forwarded to DMZ zone.
- DMZ zone contains all public facing services (DNS, email, etc).

Hold on, who clicked the following at the U-D-210-DNS rule???



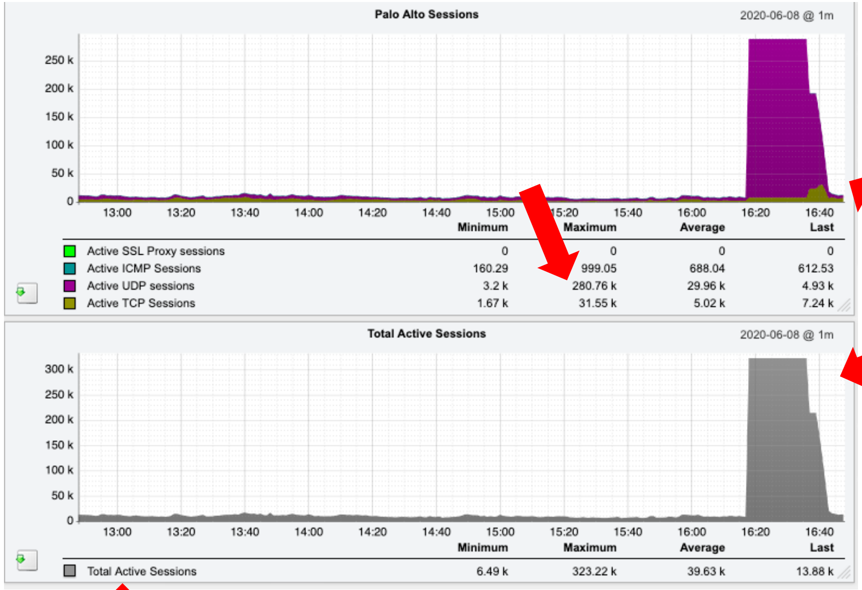
Log Setting

□ Log at Session Start

☑ Log at Session End

Log Forwarding: Log to Panorama

And again, and again, and again



Palo Alto Sessions

2020-06-08 @ 1m

	Minimum	Maximum	Average	Last
🟩 Active SSL Proxy sessions	0	0	0	0
🟦 Active ICMP Sessions	160.29	999.05	688.04	612.53
🟪 Active UDP sessions	3.2 k	280.76 k	29.96 k	4.93 k
🟫 Active TCP Sessions	1.67 k	31.55 k	5.02 k	7.24 k

Total Active Sessions

2020-06-08 @ 1m

	Minimum	Maximum	Average	Last
▨ Total Active Sessions	6.49 k	323.22 k	39.63 k	13.88 k

But now we knew where to look

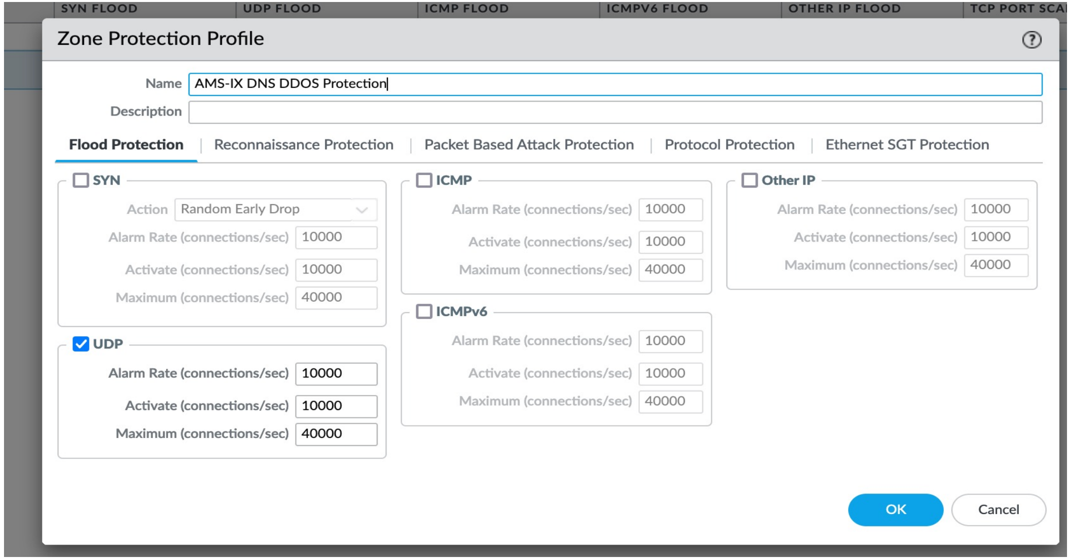
It was a chain reaction

1. Valid DNS queries arrive in our domain
2. Firewalls register the session in the session table
3. They are forwarded to our nameserver
4. Before old sessions expire, new sessions are being created
5. Session table on Firewalls gets full and firewall freaks out
6. LACP connections between FWs and Management switches drop.
7. OSPF sessions between Firewalls and RTRs drops
8. Internal infrastructure loses default gateway (firewalls)
9. Huge amount of syslog messages is being created.
10. Netflow discovered to be enabled as well!!!

Can Firewalls help themselves?

PA's Zone Protection to the rescue?

A Zone Protection profile with flood protection configured defends an entire ingress zone against SYN, ICMP, ICMPv6, UDP, and other IP flood attacks.



Zone Protection Profile

Name: AMS-IX DNS DDOS Protection

Description:

Flood Protection | Reconnaissance Protection | Packet Based Attack Protection | Protocol Protection | Ethernet SGT Protection

□ SYN
- Action: Random Early Drop
- Alarm Rate (connections/sec): 10000
- Activate (connections/sec): 10000
- Maximum (connections/sec): 40000

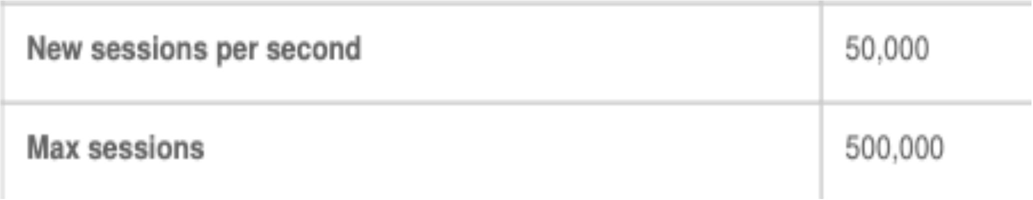
☑ UDP
- Alarm Rate (connections/sec): 10000
- Activate (connections/sec): 10000
- Maximum (connections/sec): 40000

□ ICMP
- Alarm Rate (connections/sec): 10000
- Activate (connections/sec): 10000
- Maximum (connections/sec): 40000

□ ICMPv6
- Alarm Rate (connections/sec): 10000
- Activate (connections/sec): 10000
- Maximum (connections/sec): 40000

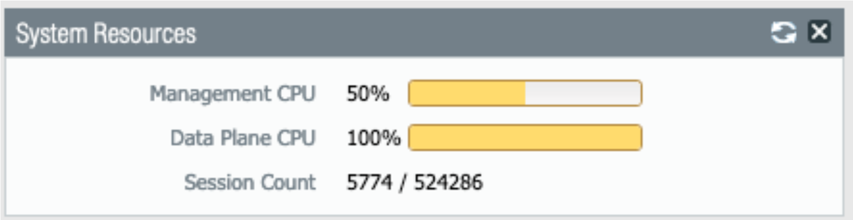
□ Other IP
- Alarm Rate (connections/sec): 10000
- Activate (connections/sec): 10000
- Maximum (connections/sec): 40000

According to Datasheet



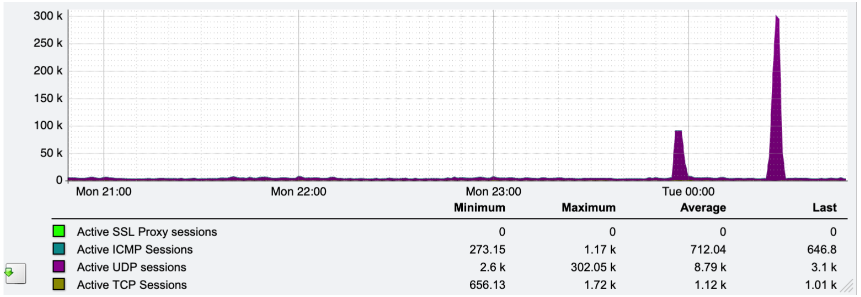
New sessions per second	50,000
Max sessions	500,000

Unfortunately, not ☹



System Resources

Management CPU	50%	
Data Plane CPU	100%	
Session Count	5774 / 524286	



		Minimum	Maximum	Average	Last
🟩	Active SSL Proxy sessions	0	0	0	0
🟦	Active ICMP Sessions	273.15	1.17 k	712.04	646.8
🟪	Active UDP sessions	2.6 k	302.05 k	8.79 k	3.1 k
🟫	Active TCP Sessions	656.13	1.72 k	1.12 k	1.01 k

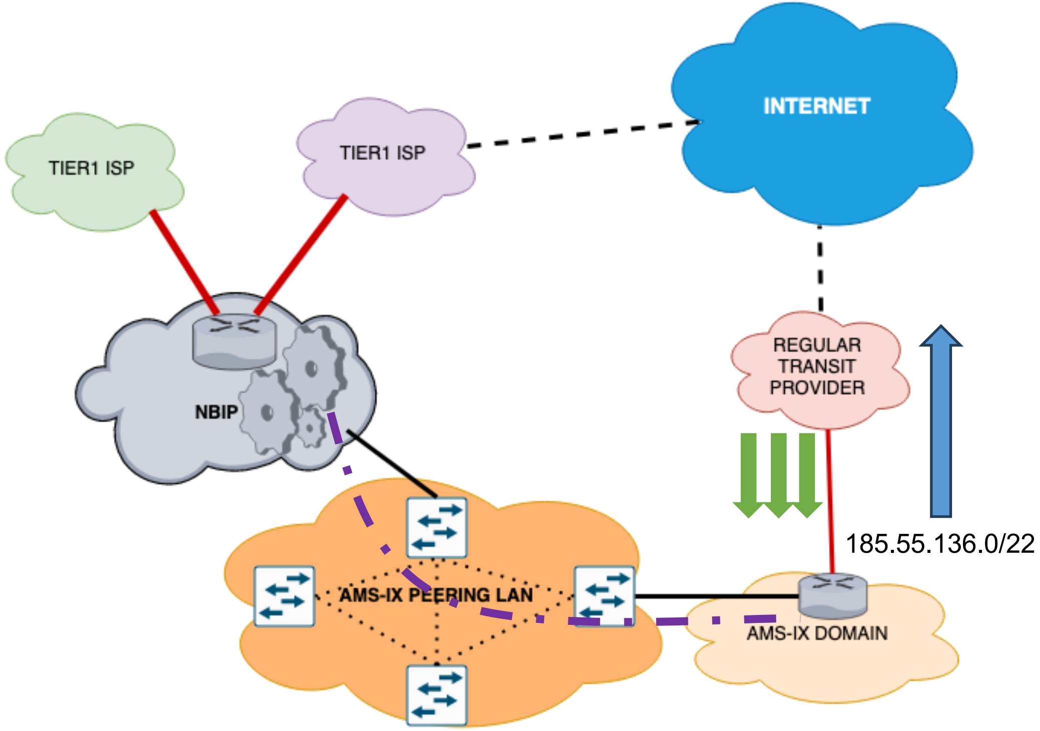
- **During the next attack we discovered the truth:**
  - The rate of new flows per second (aka new sessions) was much faster compared to what the firewall can handle.

What else was left to help us?

- **PA’s DoS protection didn’t work**
- **No other system to protect us**
- **Contract for NBIP’s NaWas DDoS protection, but:**
  - Never tested
  - No router configuration for it
  - (Almost) no documentation

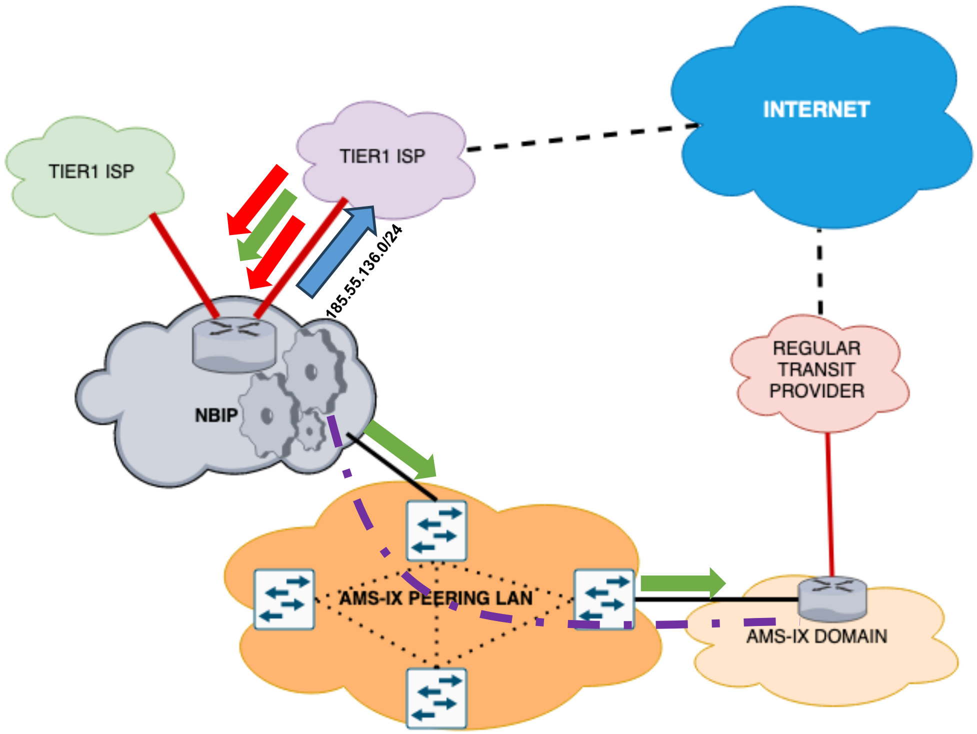
*Nationale
Beheersorganisatie
Internet
Providers

NAWAS is our shield #1



- NBIP operates a scrubbing center, connected to different Tier 1 providers.
- Under normal conditions:
  - Advertise your prefixes via transit/public peering
  - Maintain a hot standby BGP session with NBIP

NAWAS is our shield #2



- Attack scenario:
  - Advertise more specific to NBIP
  - NBIP propagates quickly to the rest of the Internet.
- NBIP will attract traffic for the specific prefix:
  - Scrubs the “dirt” packets
  - Sends the clean traffic over the dedicated BGP session.

Sleeves up and time to work

1. **SysOps actions**
   - Move public-facing nameservers in the cloud.
   - Protect them with the built-in DDoS solution.

2. **NOC actions**
   - Design and build a solution that puts an end to that.

First round of improvements

- **Review and fix the Cisco configuration**
  - Make it as simple as possible for every NOC engineer to execute it during an attack
- **Document it properly**
- **Correctly test it and fine tune it**

But how do you test it properly?

- **Shall I order a "DDoS as a Service" from Dark Web?**
  - But they don't accept my AMEX :P
- **NBIP had a testing machine**
  - But it was out-of-service that period !!!
- **Buy a VM from 3rd party hosting company and execute some tools (e.g., hping3)**
  - Unforeseen problem: all known hosting providers are AMS-IX customers (hence 1-hop away) !!!

P.S: Not to mention uRPF

DIY DDoS attack

- **Got a VM from a small Spanish hosting Provider**
  - ~ 5-6 hops away
  - uRPF disabled
  - Lots of resources (CPU & RAM)

- **Python & Scapy at hand**
  - 2 scripts (300 lines in total):
    1. a traffic generator* that produces and stores **DNS** queries in pcap files
    2. An attack script that loads the pcaps and sends the packets over the uplink as fast as possible.

P.S: Don’t ask what I used for source $source_ip

SUCCESS #1

Handmade DDoS attack was:

SUCCESS

Manual mitigation was:

Can we automate this success?

- **NOC still needs to wake up in the middle of the night to mitigate an attack of few Mbps**
- **By the time you try to mitigate, it's already too late**
  - Firewalls have already collapsed; thus, VPN concentrators are unreachable.

What are we missing?

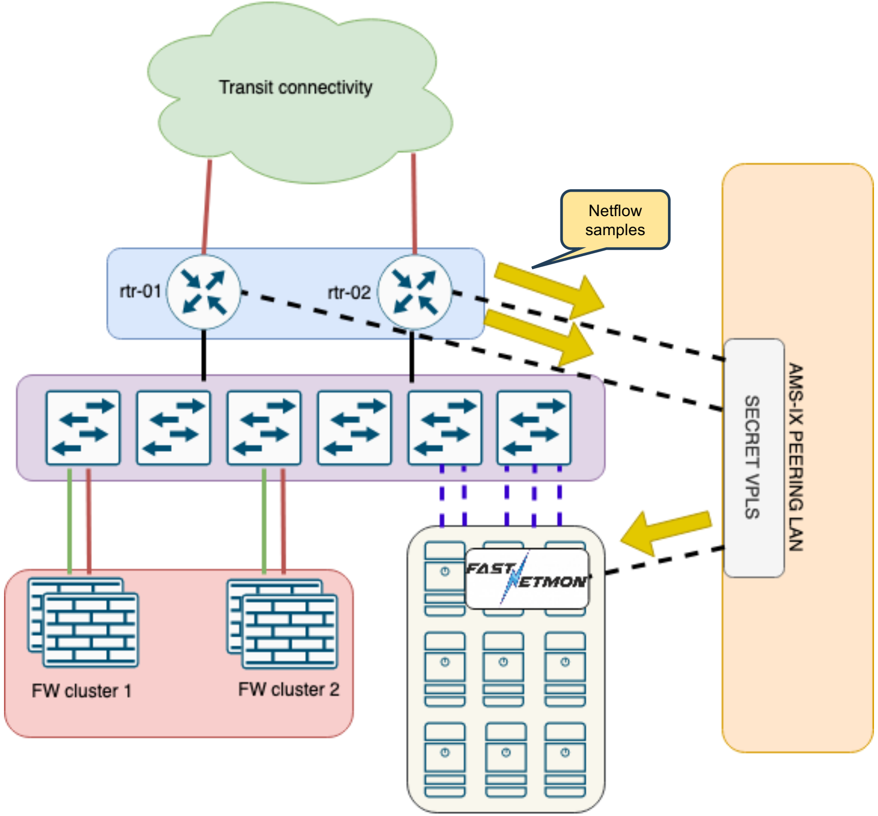
- **We have the *“shield”*, but we need the *brain* to engage it**
  - We need a “system” that:
    - Can recognize (multiple) DDoS attacks
    - Will handle the AS1200 BGP advertisements
    - Will stay up and running regardless of firewall or management network status.
    - Reliable, future proof and cheap.

- **And we need to ”glue” the brain with the weapon**

We found the brain !!!

- **FastNetmon to the rescue**
  - *Fast, Reliable and Automated DDoS detection with quick installation.*
  - Can also detect flow-based attacks (v4 only)
  - Community (free) and Advanced edition
  - Multiple sampling technologies are supported
  - Automation ready/friendly
  - Can mitigate attacks using GoBGP/ExaBGP
- **But how do you glue those parts together?**

Peering LAN is the magic glue!



To protect the traffic samples, we use the power of the peering LAN.

- Reliable, stable, with huge capacity
- We bypass the management network and the firewalls
- IXP prefix is not advertised and is not routable
- All components are NOC 24/7 monitored

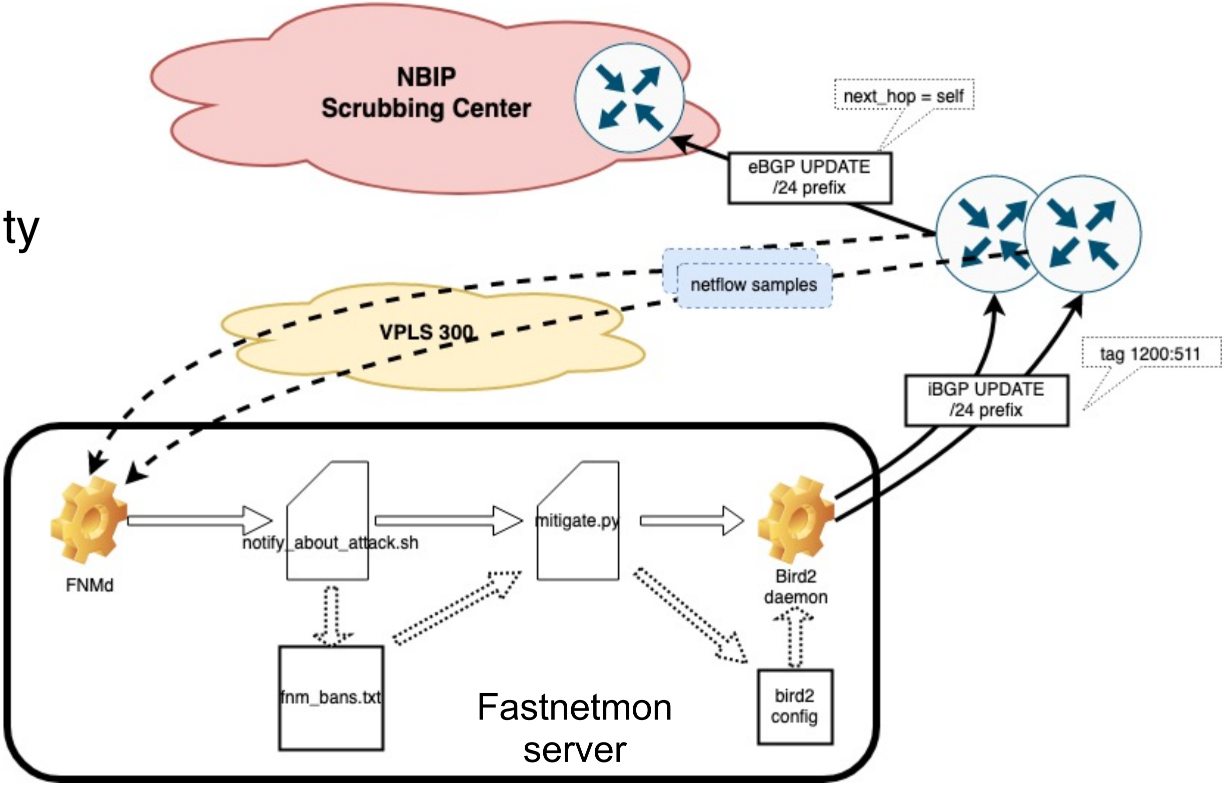
Selecting a signaling method

- **To handle the router advertisements of Cisco’s**
  - Multiple approaches were considered:
    - SSH, HTTP API, BGP
  - We opted for BGP over Bird
    - NOC team has good experience with Bird (and plenty of internal documentation)
    - BGP session can be monitored 24x7
    - Signaling over established BGP session is fast
    - We can use BGP communities for fine tuning.

Building an automation pipeline

Components used:

- Fastnetmon Community
- Netflow
- Python + Jinja2
- Bird2
- iBGP + BGP communities
- Cisco route maps



Different strategies per AFI

- **If a prefix arrives to border router from Bird**
  - **IPv4**: if prefix is tagged with 1200:511
    - Block the propagation to transit and peers
    - Allow it to NBIP
  - **IPv6**: if prefix is tagged with 1200:511
    - Withdraw the announcement from transit and peers
    - Allow it to NBIP

Does it work?

SUCCESS

- New DNS-based exercise attack:
  - Did a combination of DNS and ICMP
  - Executed it 2 times
  - 2M packets with IPv4 destination
  - 2M packets with IPv6 destination
- ~45 seconds from time we launch the attack until the time it is completely mitigated
- **In both cases, NOC didn't perform any manual action or intervention !**
- IPv6 Mitigation didn't work ☹

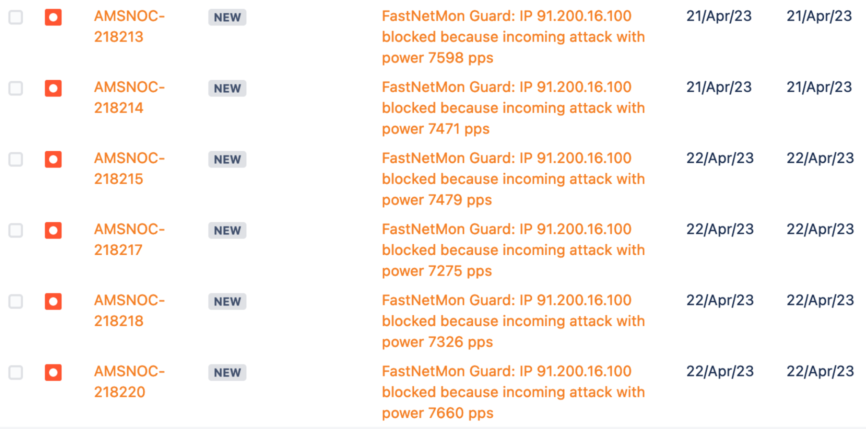
```
RTR-DR1-01_cisco#show ip bgp neighbors 194.62.128.2 advertised-
routes
Load for five secs: 16%/6%; one minute: 38%; five minutes: 30%
Time source is NTP, 17:22:52.312 CET Fri Mar 18 2022


BGP table version is 312561314, local router ID is 91.200.16.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found


     Network          Next Hop            Metric LocPrf Weight Path
V*>i 185.55.136.0/24  91.200.16.2              0    100      0 i
V*>  185.55.137.0/24  91.200.16.11            11         32768 i


Total number of prefixes 2
```

But does it really work?

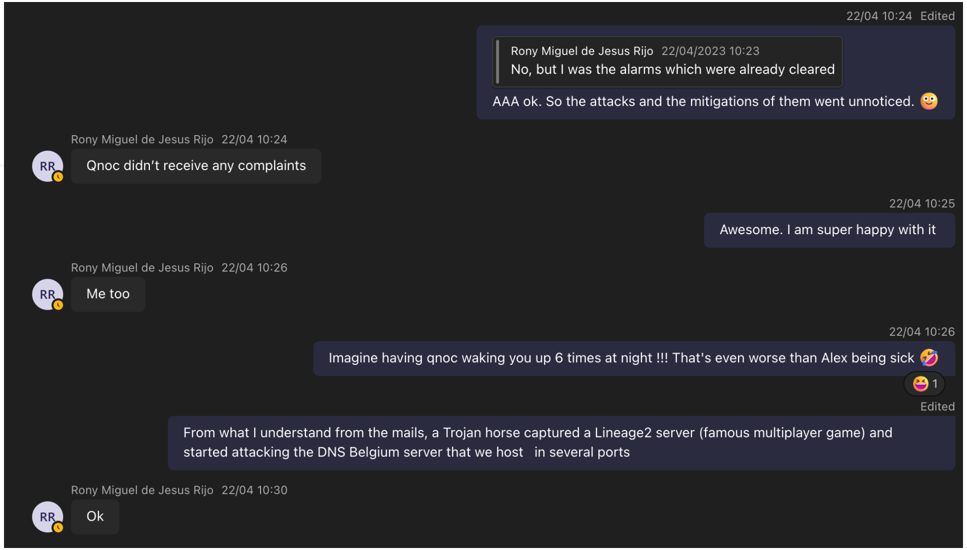


□	🟥	AMSNOC-218213	NEW	FastNetMon Guard: IP 91.200.16.100 blocked because incoming attack with power 7598 pps	21/Apr/23	21/Apr/23	*Unassigned*
□	🟥	AMSNOC-218214	NEW	FastNetMon Guard: IP 91.200.16.100 blocked because incoming attack with power 7471 pps	21/Apr/23	21/Apr/23	*Unassigned*
□	🟥	AMSNOC-218215	NEW	FastNetMon Guard: IP 91.200.16.100 blocked because incoming attack with power 7479 pps	22/Apr/23	22/Apr/23	*Unassigned*
□	🟥	AMSNOC-218217	NEW	FastNetMon Guard: IP 91.200.16.100 blocked because incoming attack with power 7275 pps	22/Apr/23	22/Apr/23	*Unassigned*
□	🟥	AMSNOC-218218	NEW	FastNetMon Guard: IP 91.200.16.100 blocked because incoming attack with power 7326 pps	22/Apr/23	22/Apr/23	
□	🟥	AMSNOC-218220	NEW	FastNetMon Guard: IP 91.200.16.100 blocked because incoming attack with power 7660 pps	22/Apr/23	22/Apr/23	

We had a flow-based attack
6 times at the same night!!!

Attacks registered successfully at the ticketing system but:

- Standby engineer was **not** called
- Attacks mitigated **successfully**
- **No** complains received to FLS



22/04 10:24 Edited

> Rony Miguel de Jesus Rijo 22/04/2023 10:23
> No, but I was the alarms which were already cleared

AAA ok. So the attacks and the mitigations of them went unnoticed. 😳

Rony Miguel de Jesus Rijo 22/04 10:24

RR Qnoc didn't receive any complaints

22/04 10:25

Awesome. I am super happy with it

Rony Miguel de Jesus Rijo 22/04 10:26

RR Me too

22/04 10:26

Imagine having qnoc waking you up 6 times at night !!! That's even worse than Alex being sick 🤣

😝 1

Edited

From what I understand from the mails, a Trojan horse captured a Lineage2 server (famous multiplayer game) and started attacking the DNS Belgium server that we host in several ports

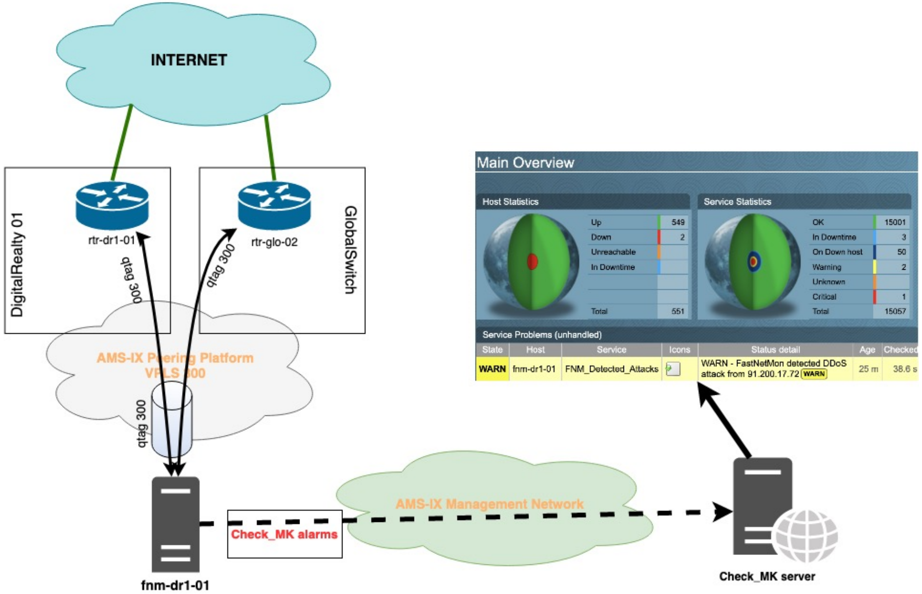
Rony Miguel de Jesus Rijo 22/04 10:30

RR Ok

From zero to hero !

amsix

Some final automation touch



Main Overview

Host Statistics

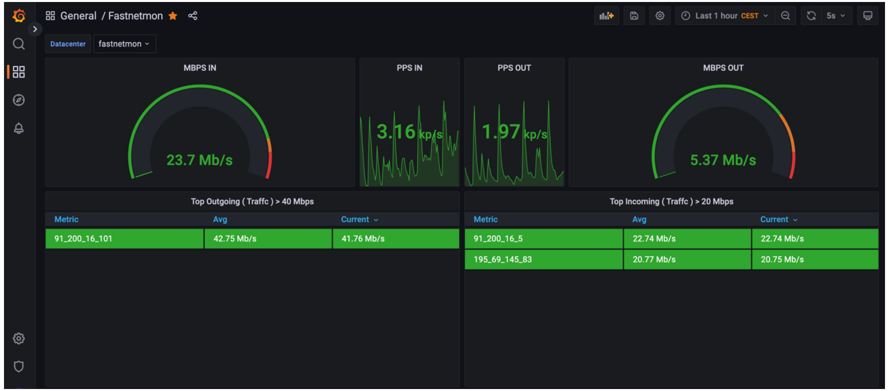
Up	549
Down	2
Unreachable	
In Downtime	
Total	551

Service Statistics

OK	15001
In Downtime	3
On Down host	50
Warning	2
Unknown	
Critical	1
Total	15057

Service Problems (unhandled)

State	Host	Service	Icons	Status detail	Age	Checked
WARN	fnm-dr1-01	FNM_Detected_Attacks		WARN - FastNetMon detected DDoS attack from 91.200.17.72 WARN	25 m	38.6 s



Top Outgoing (Traffc) > 40 Mbps

Metric	Avg	Current
91_200_16_101	42.75 Mb/s	41.76 Mb/s

Top Incoming (Traffc) > 20 Mbps

Metric	Avg	Current
91_200_16_5	22.74 Mb/s	22.74 Mb/s
195_69_145_83	20.77 Mb/s	20.75 Mb/s

Link it to our NMS

Grafana Dashboard

Lessons learned

- **It was a bumpy ride**
  - We had to build everything from scratch
  - We had to tweak NBIP’s thresholds
  - Lots of Netflow tuning (please don’t use it)
  - ROAs had to be adjusted (no max length)
  - IPv6 still needs work (at FNM side)
- **We had to train ourselves on these situations**
- **Sometimes management needs to understand the impact**

Future steps

- **Fastnetmon Community → Advanced**
  - Almost done

- **Border routers replacement**
  - Cisco ASR 1001 → Juniper MX204 (WIP)

- **Netflow → IPFIX**
  - Improve reaction time

- **Improve mitigation algorithm**
  - Use RTBH for specific cases

- **Adopt Flowspec**

Key take-aways

- **If you are a small (stub) network:**
    1. Consider adopting a DDoS protection solution **now**
    2. You can have a complete & reliable implementation with open-source tools and small budget
    3. Keep your router’s OS & documentation up-to-date
    4. Consider thresholds for traffic redirection
    5. Implement for IPv6 attacks as well
    6. Re-think your ROAs

YOU GET A DDOS ATTACK!
2004
AND YOU GET A DDOS ATTACK!
AND YOU GET A DDOS ATTACK!
EVERYONE GETS A DDOS ATTACK!
quickmeme.com

30% discount to all GRNOG members