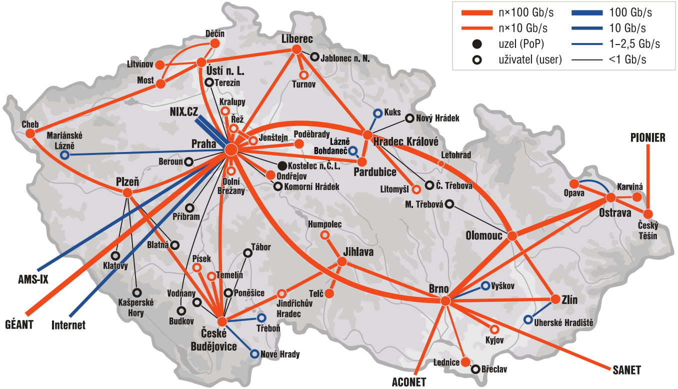
ExaFS: mitigating unwanted traffic

Ondřej Caletka

13th November 2019

The specifics of NREN backbone

- an attack can be fatal to a single customer
- the network can be dangerous to others
- formerly **no filtering by default**[1]



[1] unless required (BCP 38) or requested by client

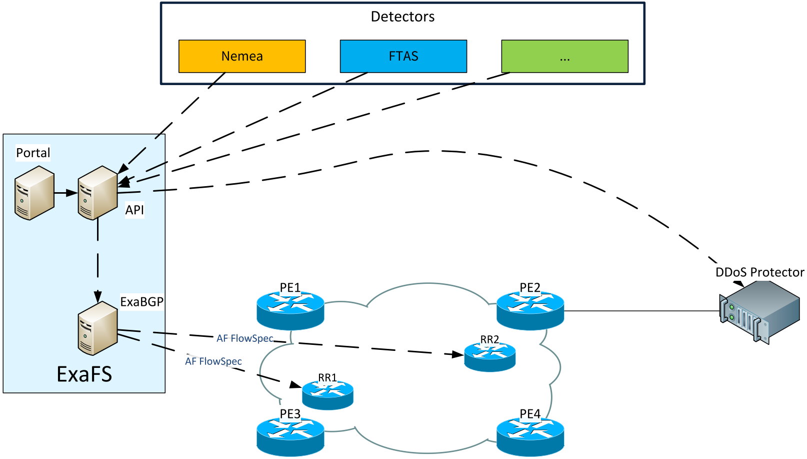
DoS mitigation strategies in CESNET

- per-protocol QoS on the network perimeter
  - for connection-less protocols like NTP, SNMP,...
  - sum of NTP flows typical ~2 Mbps
  - different packet sizes of legitimate and attack flows
- many QoS groups for DNS and fragments (cca. one per customer)
  - hard to recognize attack on the perimeter
  - crucial service for *eyeball* experience
- Remote-Triggered Black Hole filtering for BGP-connected customers
  - for attacks targetted to small number of IP addresses
  - eliminates saturation of the last mile link

BGP Flowspec

- allows fine-grained selection of flows to filter
  - but tricky to set up properly by hand
- we found no ready-made solution allowing easy access:
  - to customers’ network admins
  - to the CSIRT team
  - to automated tools for mitigation of well known attack patterns
- we decided to build our own open source solution **ExaFS**
  - Flowspec-based filtering and RTBH control
  - user accounts with permissions for IP ranges
  - automatic expiration of rules
  - API for robots

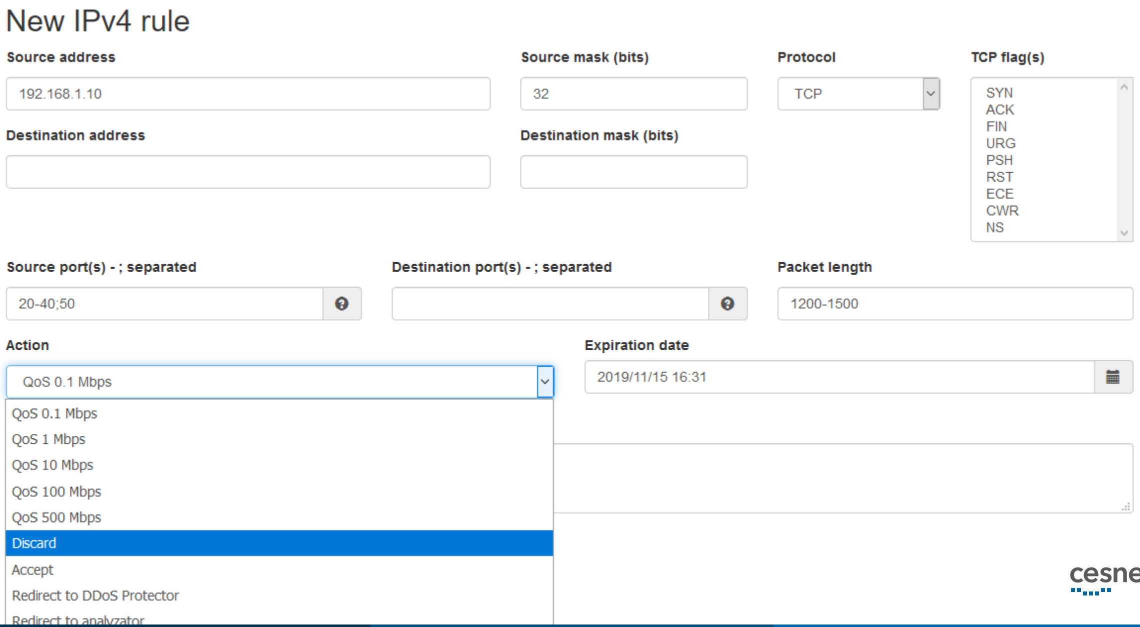
The big picture



ExaFS components

- ExaBGP 4.1.2
- Python 3.6
- MariaDB
- Flask + WTForms + SQLAlchemy

- ready for Shibboleth Single sign-on federated identity login
- sources on `https://github.com/CESNET/exafs`
- API documentation on `https://exafs.docs.apiary.io`
- open-source with MIT license



New IPv4 rule

Source address: 192.168.1.10

Source mask (bits): 32

Protocol: TCP

TCP flag(s):

Destination address:

Destination mask (bits):

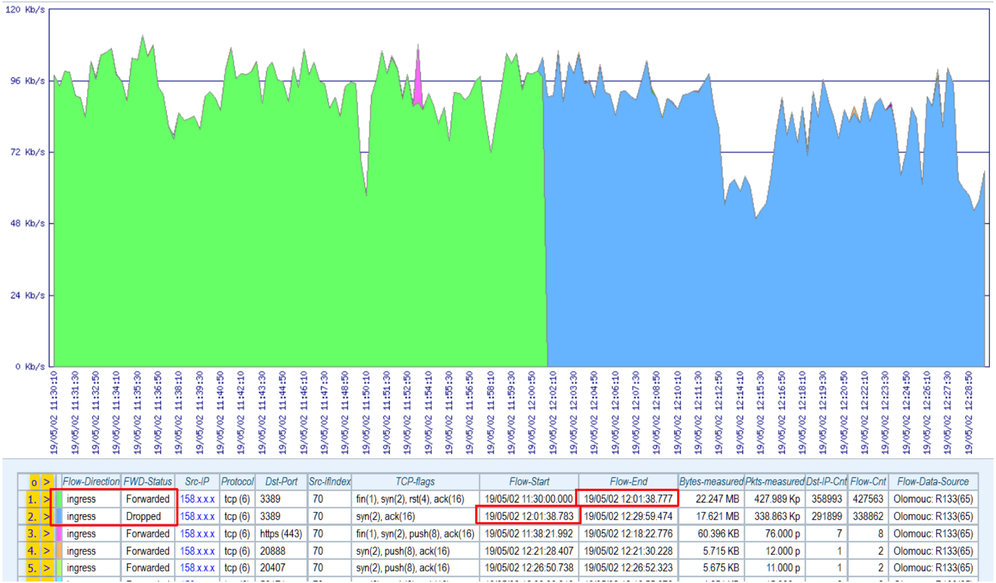
Source port(s) - ; separated: 20-40;50

Destination port(s) - ; separated:

Packet length: 1200-1500

Action: QoS 0.1 Mbps

Expiration date: 2019/11/15 16:31



o	>		Flow-Direction	FWD-Status	Src-IP	Protocol	Dst-Port	Src-ifIndex	TCP-flags	Flow-Start	Flow-End	Bytes-measured	Pkts-measured	Dst-IP-Cnt	Flow-Cnt	Flow-Data-Source
1.	>	🟩	ingress	Forwarded	158.x.x.x	tcp (6)	3389	70	fin(1), syn(2), rst(4), ack(16)	19/05/02 11:30:00.000	19/05/02 12:01:38.777	22.247 MB	427.989 Kp	358993	427563	Olomouc: R133(65)
2.	>	🟦	ingress	Dropped	158.x.x.x	tcp (6)	3389	70	syn(2), ack(16)	19/05/02 12:01:38.783	19/05/02 12:29:59.474	17.621 MB	338.863 Kp	291899	338862	Olomouc: R133(65)
3.	>	🟪	ingress	Forwarded	158.x.x.x	tcp (6)	https (443)	70	fin(1), syn(2), push(8), ack(16)	19/05/02 11:38:21.992	19/05/02 12:18:22.776	60.396 KB	76.000 p	7	8	Olomouc: R133(65)
4.	>	🟧	ingress	Forwarded	158.x.x.x	tcp (6)	20888	70	syn(2), push(8), ack(16)	19/05/02 12:21:28.407	19/05/02 12:21:30.228	5.715 KB	12.000 p	1	2	Olomouc: R133(65)
5.	>	🟩	ingress	Forwarded	158.x.x.x	tcp (6)	20407	70	syn(2), push(8), ack(16)	19/05/02 12:26:50.738	19/05/02 12:26:52.323	5.675 KB	11.000 p	1	2	Olomouc: R133(65)

Caveats of Flowspec filters

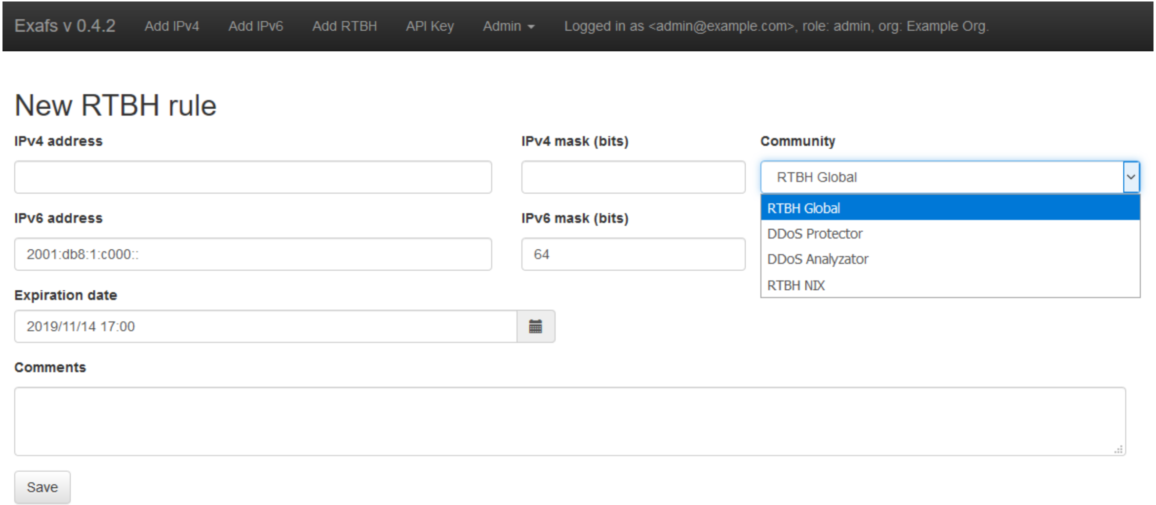
- no universal support for all features
  - our Nokia boxes cannot do QoS together with packet length matching
- fragmented traffic has port numbers set to 0
- ordering of rules is not always intuitive (RFC 5575 5.1)
  1. Destination prefix
  2. Source prefix
  3. IP protocol
  4. Port
  5. Destination port
  6. Source port

BGP Flowspec rules ordering example

```
Sequence:  1513     Flow
  :Dest:192.0.2.1/32,Source:198.51.100.128/26,
   Proto:=17,DPort:=3702
Sequence:  1572     Flow
  :Dest:192.0.2.1/32,Proto:=17,DPort:=3702
Sequence:  1575     Flow
  :Dest:192.0.2.0/31,Source:198.51.100.188/32,
   Proto:=17,DPort:=3702
Sequence:  1579     Flow
  :Dest:192.0.2.0/31,Source:198.51.100.128/26,
   Proto:=17,DPort:=3702
Sequence:  1586     Flow
  :Dest:192.0.2.0/24,Source:198.51.100.128/26,
   Proto:=17,DPort:=3702
```

RTBH support in ExaFS

- not limited only to BGP-connected clients
- particularly useful for large volumetric attacks
- RTBH rules **can be propagated** to peering partners and transit providers
- can be also used for redirection to DDoS protector
- support for standard, extended and large BGP communities



Logged in as <admin@example.com>, role: admin, org: Example Org.

New RTBH rule

IPv4 address:

IPv4 mask (bits):

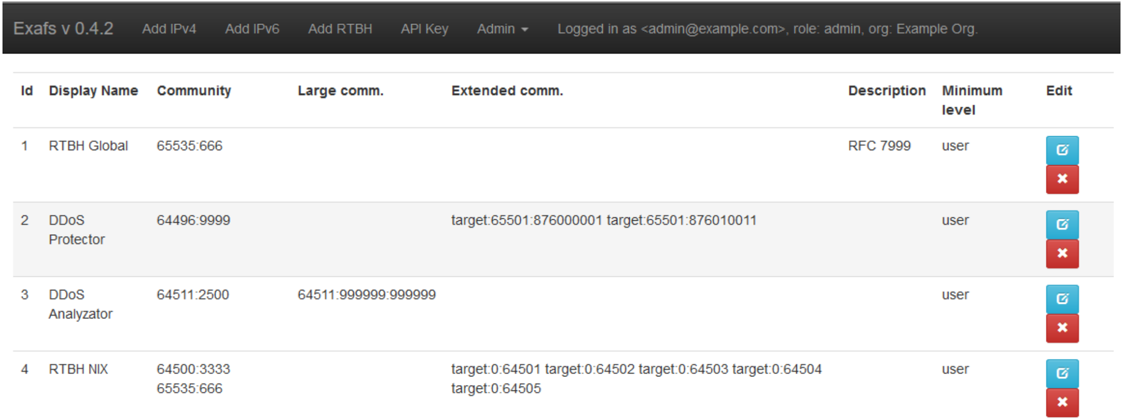
Community: RTBH Global

IPv6 address: 2001:db8:1:c000::

IPv6 mask (bits): 64

Expiration date: 2019/11/14 17:00

Comments:



Logged in as <admin@example.com>, role: admin, org: Example Org.

Id	Display Name	Community	Large comm.	Extended comm.	Description	Minimum level	Edit
1	RTBH Global	65535:666			RFC 7999	user	✎ ✖
2	DDoS Protector	64496:9999		target:65501:876000001 target:65501:876010011		user	✎ ✖
3	DDoS Analyzator	64511:2500	64511:999999:999999			user	✎ ✖
4	RTBH NIX	64500:3333 65535:666		target:0:64501 target:0:64502 target:0:64503 target:0:64504 target:0:64505		user	✎ ✖

Conclusion

- easy to use web-based front-end for BGP Flowspec and RTBH
- also a **very dangerous weapon** that can kill your network pretty easily
- automatic **expiration of rules**
- **JSON API** for automated mitigation of well known attacks
  - but we are still a little bit scared to keep humans out of the loop

Live demo:[2]

`https://exafs-demo.cesnet.cz`

[2] Available only for limited time.

Thank You!

Ondřej Caletka
Ondrej.Caletka@cesnet.cz
https://**Ondřej.Caletka.cz**

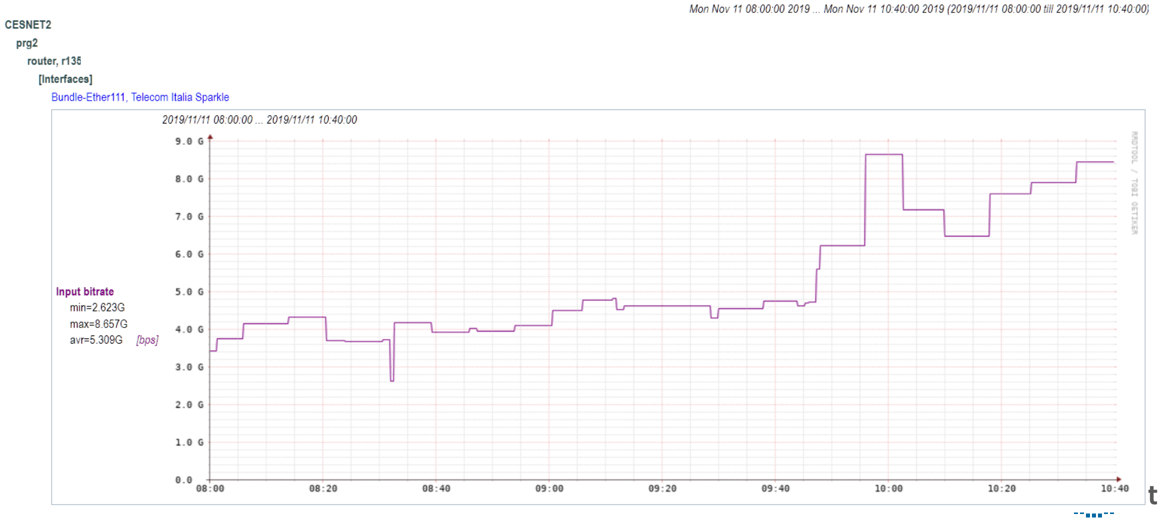
cesnet

Bonus slides

Recent attack case study

- transit connectivity link saturated with NTP replies to one IP address
- blocked at the upstream using RTBH
- smaller part of the NTP flood arived from other links, filled global policers
- we used a BGP Flowspec rule to block NTP flood towards the IP address under attack
- the global NTP policers returned to empty state, allowing normal NTP operation for the other parts of the network

The anomaly is detected



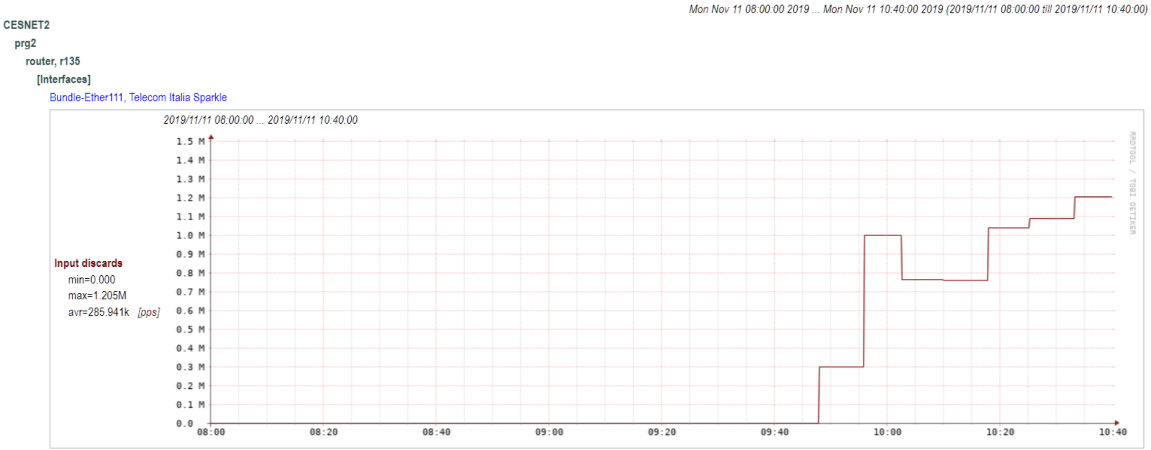
Mon Nov 11 08:00:00 2019 ... Mon Nov 11 10:40:00 2019 (2019/11/11 08:00:00 till 2019/11/11 10:40:00)

CESNET2
prg2
router, r135
[Interfaces]
Bundle-Ether111, Telecom Italia Sparkle

2019/11/11 08:00:00 ... 2019/11/11 10:40:00

Input bitrate
min=2.623G
max=8.657G
avr=5.309G *[bps]*

Global NTP QoS is being utilized



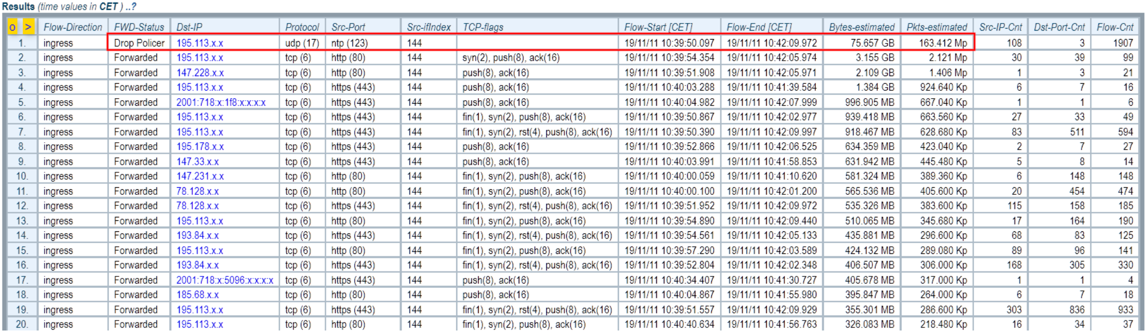
Mon Nov 11 08:00:00 2019 ... Mon Nov 11 10:40:00 2019 (2019/11/11 08:00:00 till 2019/11/11 10:40:00)

CESNET2
prg2
router, r135
[Interfaces]
Bundle-Ether111, Telecom Italia Sparkle

2019/11/11 08:00:00 ... 2019/11/11 10:40:00

Input discards
min=0.000
max=1.205M
avr=285.941k *[pps]*

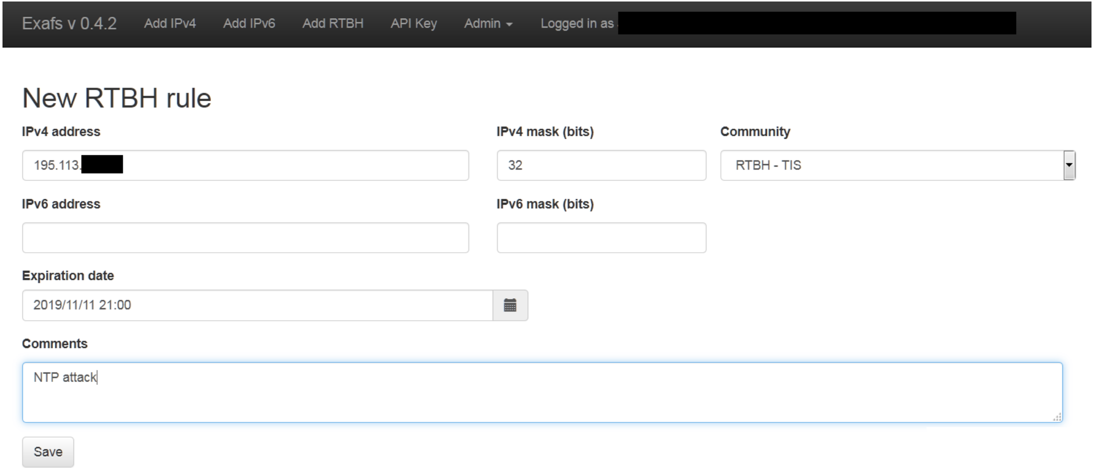
The attack target is located



Results *(time values in* ***CET*** *) ..?*

o >	Flow-Direction	FWD-Status	Dst-IP	Protocol	Src-Port	Src-ifIndex	TCP-flags	Flow-Start [CET]	Flow-End [CET]	Bytes-estimated	Pkts-estimated	Src-IP-Cnt	Dst-Port-Cnt	Flow-Cnt
1.	ingress	Drop Policer	195.113.x.x	udp (17)	ntp (123)	144		19/11/11 10:39:50.097	19/11/11 10:42:09.972	75.657 GB	163.412 Mp	108	3	1907
2.	ingress	Forwarded	195.113.x.x	tcp (6)	http (80)	144	syn(2), push(8), ack(16)	19/11/11 10:39:54.354	19/11/11 10:42:05.974	3.155 GB	2.121 Mp	30	39	99
3.	ingress	Forwarded	147.228.x.x	tcp (6)	http (80)	144	push(8), ack(16)	19/11/11 10:39:51.908	19/11/11 10:42:05.971	2.109 GB	1.406 Mp	1	3	21
4.	ingress	Forwarded	195.113.x.x	tcp (6)	https (443)	144	push(8), ack(16)	19/11/11 10:40:03.288	19/11/11 10:41:39.584	1.384 GB	924.640 Kp	6	7	16
5.	ingress	Forwarded	2001:718:x:1f8:x:x:x:x	tcp (6)	https (443)	144	push(8), ack(16)	19/11/11 10:40:04.982	19/11/11 10:42:07.999	996.905 MB	667.040 Kp	1	1	6
6.	ingress	Forwarded	195.113.x.x	tcp (6)	https (443)	144	fin(1), syn(2), push(8), ack(16)	19/11/11 10:39:50.867	19/11/11 10:42:02.977	939.418 MB	663.560 Kp	27	33	49
7.	ingress	Forwarded	195.113.x.x	tcp (6)	https (443)	144	fin(1), syn(2), rst(4), push(8), ack(16)	19/11/11 10:39:50.390	19/11/11 10:42:09.997	918.467 MB	628.680 Kp	83	511	594
8.	ingress	Forwarded	195.178.x.x	tcp (6)	https (443)	144	push(8), ack(16)	19/11/11 10:39:52.866	19/11/11 10:42:06.525	634.359 MB	423.040 Kp	2	7	27
9.	ingress	Forwarded	147.33.x.x	tcp (6)	https (443)	144	push(8), ack(16)	19/11/11 10:40:03.991	19/11/11 10:41:58.853	631.942 MB	445.480 Kp	5	8	14
10.	ingress	Forwarded	147.231.x.x	tcp (6)	http (80)	144	fin(1), syn(2), push(8), ack(16)	19/11/11 10:40:00.059	19/11/11 10:41:10.620	581.324 MB	389.360 Kp	6	148	148
11.	ingress	Forwarded	78.128.x.x	tcp (6)	http (80)	144	fin(1), syn(2), push(8), ack(16)	19/11/11 10:40:00.100	19/11/11 10:42:01.200	565.536 MB	405.600 Kp	20	454	474
12.	ingress	Forwarded	78.128.x.x	tcp (6)	https (443)	144	fin(1), syn(2), rst(4), push(8), ack(16)	19/11/11 10:39:51.952	19/11/11 10:42:09.972	535.326 MB	383.600 Kp	115	158	185
13.	ingress	Forwarded	195.113.x.x	tcp (6)	http (80)	144	fin(1), syn(2), push(8), ack(16)	19/11/11 10:39:54.890	19/11/11 10:42:09.440	510.065 MB	345.680 Kp	17	164	190
14.	ingress	Forwarded	193.84.x.x	tcp (6)	https (443)	144	fin(1), syn(2), rst(4), push(8), ack(16)	19/11/11 10:39:54.561	19/11/11 10:42:05.133	435.881 MB	296.600 Kp	68	83	125
15.	ingress	Forwarded	195.113.x.x	tcp (6)	http (80)	144	fin(1), syn(2), push(8), ack(16)	19/11/11 10:39:57.290	19/11/11 10:42:03.589	424.132 MB	289.080 Kp	89	96	141
16.	ingress	Forwarded	193.84.x.x	tcp (6)	https (443)	144	fin(1), syn(2), rst(4), push(8), ack(16)	19/11/11 10:39:52.804	19/11/11 10:42:02.348	406.507 MB	306.000 Kp	168	305	330
17.	ingress	Forwarded	2001:718:x:5096:x:x:x:x	tcp (6)	https (443)	144	push(8), ack(16)	19/11/11 10:40:34.407	19/11/11 10:41:30.727	405.678 MB	317.000 Kp	1	1	4
18.	ingress	Forwarded	185.68.x.x	tcp (6)	http (80)	144	push(8), ack(16)	19/11/11 10:40:04.867	19/11/11 10:41:55.980	395.847 MB	264.000 Kp	6	7	18
19.	ingress	Forwarded	195.113.x.x	tcp (6)	https (443)	144	fin(1), syn(2), rst(4), push(8), ack(16)	19/11/11 10:39:51.557	19/11/11 10:42:09.929	355.301 MB	286.600 Kp	303	836	933
20.	ingress	Forwarded	195.113.x.x	tcp (6)	http (80)	144	fin(1), syn(2), push(8), ack(16)	19/11/11 10:40:40.634	19/11/11 10:41:56.763	326.083 MB	218.480 Kp	16	34	37

RTBH rule is created to free up transit



Logged in as [redacted]

New RTBH rule

IPv4 address: 195.113.[redacted]

IPv4 mask (bits): 32

Community: RTBH - TIS

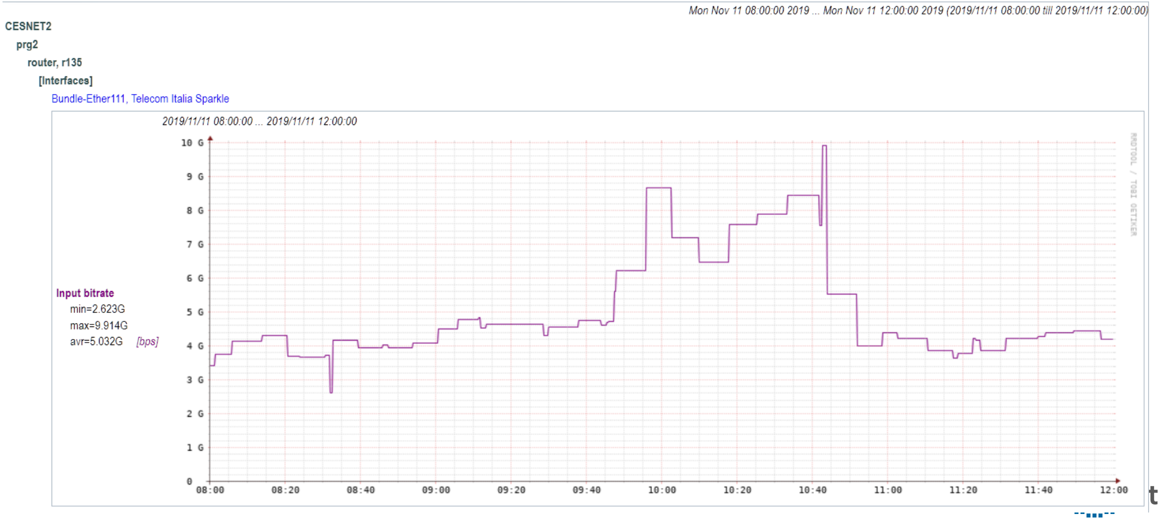
IPv6 address:

IPv6 mask (bits):

Expiration date: 2019/11/11 21:00

Comments: NTP attack

The transit link is not saturated anymore



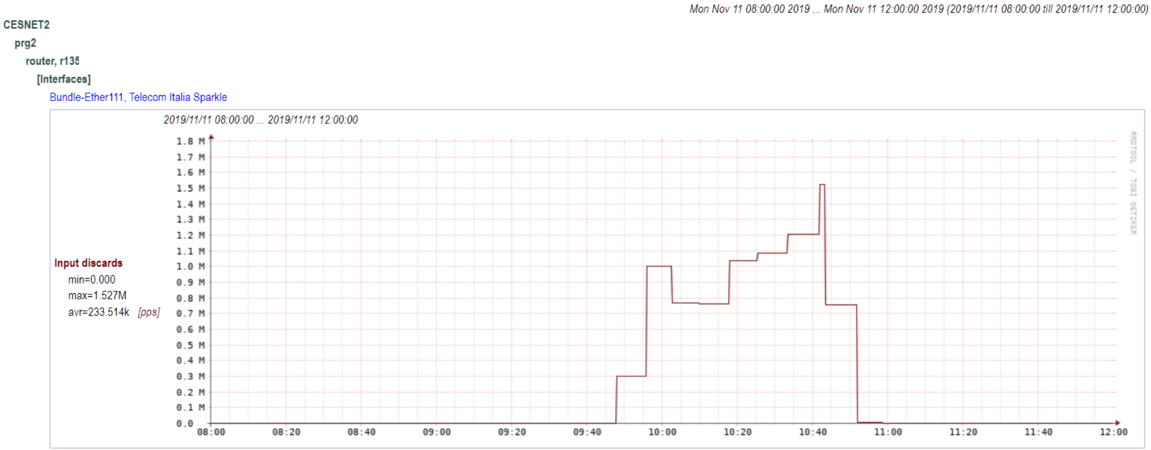
Mon Nov 11 08:00:00 2019 ... Mon Nov 11 12:00:00 2019 (2019/11/11 08:00:00 till 2019/11/11 12:00:00)

CESNET2
prg2
router, r135
[Interfaces]
Bundle-Ether111, Telecom Italia Sparkle

2019/11/11 08:00:00 ... 2019/11/11 12:00:00

Input bitrate
min=2.623G
max=9.914G
avr=5.032G *[bps]*

No more discards on the trasit link



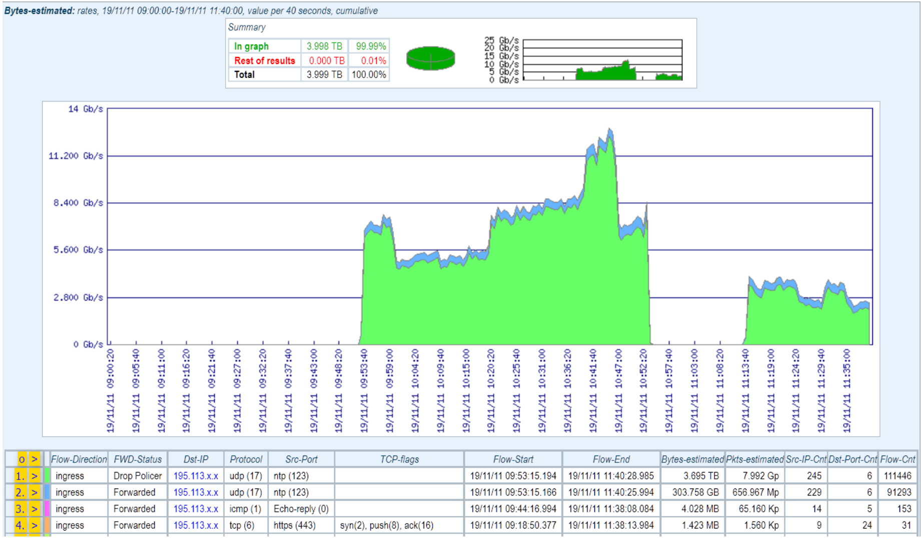
Mon Nov 11 08:00:00 2019 ... Mon Nov 11 12:00:00 2019 (2019/11/11 08:00:00 till 2019/11/11 12:00:00)

CESNET2
prg2
router, r135
[Interfaces]
Bundle-Ether111, Telecom Italia Sparkle

2019/11/11 08:00:00 ... 2019/11/11 12:00:00

Input discards
min=0.000
max=1.527M
avr=233.514k [pps]

The attack is coming from other sources as well



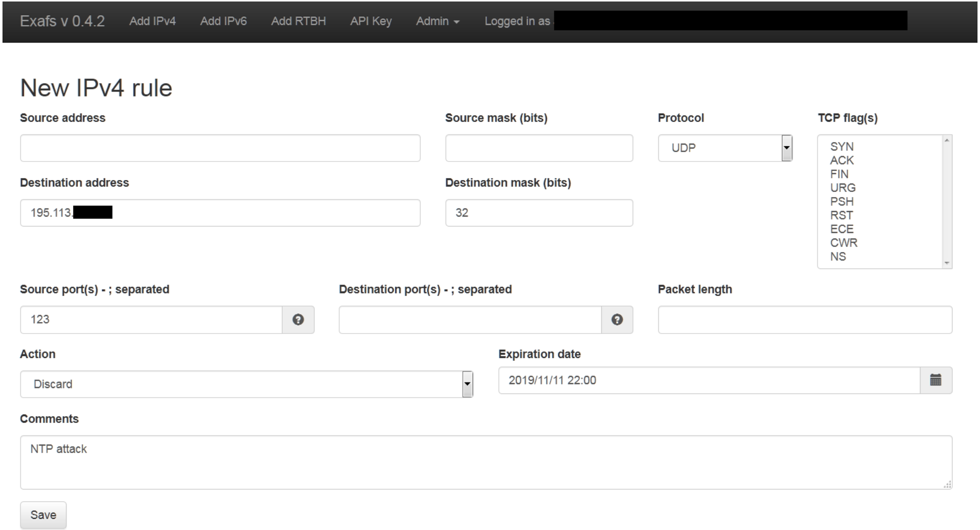
Bytes-estimated: *rates, 19/11/11 09:00:00-19/11/11 11:40:00, value per 40 seconds, cumulative*

Summary

In graph	3.998 TB	99.99%
Rest of results	0.000 TB	0.01%
Total	3.999 TB	100.00%

o	>		*Flow-Direction*	*FWD-Status*	*Dst-IP*	*Protocol*	*Src-Port*	*TCP-flags*	*Flow-Start*	*Flow-End*	*Bytes-estimated*	*Pkts-estimated*	*Src-IP-Cnt*	*Dst-Port-Cnt*	*Flow-Cnt*
1.	>	🟩	ingress	Drop Policer	195.113.x.x	udp (17)	ntp (123)		19/11/11 09:53:15.194	19/11/11 11:40:28.985	3.695 TB	7.992 Gp	245	6	111446
2.	>	🟦	ingress	Forwarded	195.113.x.x	udp (17)	ntp (123)		19/11/11 09:53:15.166	19/11/11 11:40:25.994	303.758 GB	656.967 Mp	229	6	91293
3.	>	🟪	ingress	Forwarded	195.113.x.x	icmp (1)	Echo-reply (0)		19/11/11 09:44:16.994	19/11/11 11:38:08.084	4.028 MB	65.160 Kp	14	5	153
4.	>	🟧	ingress	Forwarded	195.113.x.x	tcp (6)	https (443)	syn(2), push(8), ack(16)	19/11/11 09:18:50.377	19/11/11 11:38:13.984	1.423 MB	1.560 Kp	9	24	31

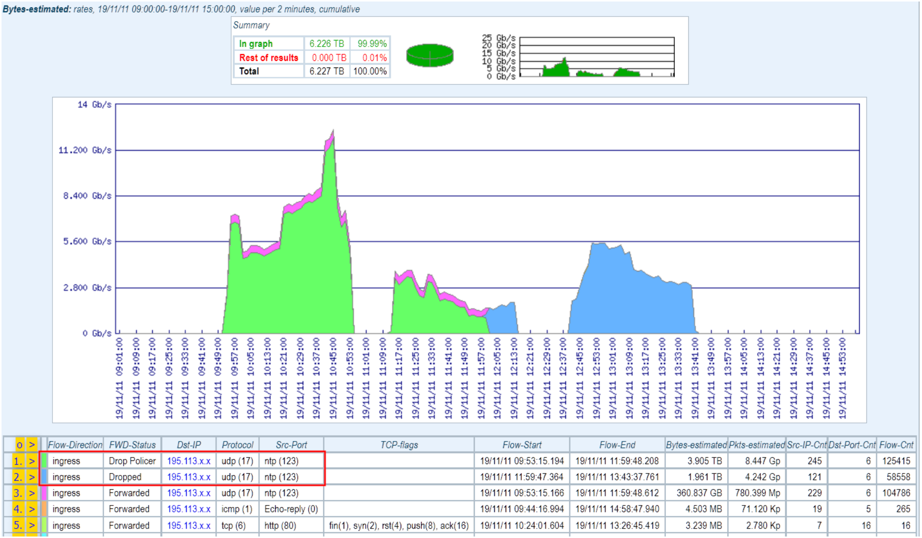
Let’s discard it using BGP Flowspec



New IPv4 rule

Source address	
Source mask (bits)	
Protocol	UDP
TCP flag(s)	SYN ACK FIN URG PSH RST ECE CWR NS
Destination address	195.113.[redacted]
Destination mask (bits)	32
Source port(s) - ; separated	123
Destination port(s) - ; separated	
Packet length	
Action	Discard
Expiration date	2019/11/11 22:00
Comments	NTP attack

The global NTP QoS is not in use anymore



Bytes-estimated: *rates, 19/11/11 09:00:00-19/11/11 15:00:00, value per 2 minutes, cumulative*

Summary

In graph	6.226 TB	99.99%
Rest of results	0.000 TB	0.01%
Total	6.227 TB	100.00%

o	>		Flow-Direction	FWD-Status	Dst-IP	Protocol	Src-Port	TCP-flags	Flow-Start	Flow-End	Bytes-estimated	Pkts-estimated	Src-IP-Cnt	Dst-Port-Cnt	Flow-Cnt
1.	>	🟩	ingress	Drop Policer	195.113.x.x	udp (17)	ntp (123)		19/11/11 09:53:15.194	19/11/11 11:59:48.208	3.905 TB	8.447 Gp	245	6	125415
2.	>	🟦	ingress	Dropped	195.113.x.x	udp (17)	ntp (123)		19/11/11 11:59:47.364	19/11/11 13:43:37.761	1.961 TB	4.242 Gp	121	6	58558
3.	>	🟪	ingress	Forwarded	195.113.x.x	udp (17)	ntp (123)		19/11/11 09:53:15.166	19/11/11 11:59:48.612	360.837 GB	780.399 Mp	229	6	104786
4.	>	🟧	ingress	Forwarded	195.113.x.x	icmp (1)	Echo-reply (0)		19/11/11 09:44:16.994	19/11/11 14:58:47.940	4.503 MB	71.120 Kp	19	5	265
5.	>	🟩	ingress	Forwarded	195.113.x.x	tcp (6)	http (80)	fin(1), syn(2), rst(4), push(8), ack(16)	19/11/11 10:24:01.604	19/11/11 13:26:45.419	3.239 MB	2.780 Kp	7	16	16