Making Sense of the R&E Network Topology

Petya Vasileva / University of Michigan

IRIS-HEP Topical Meeting: What is going on in networking, May 2, 2022
https://indico.cern.ch/event/1138837/

Goals

Extract knowledge about the nature of the issues on our network

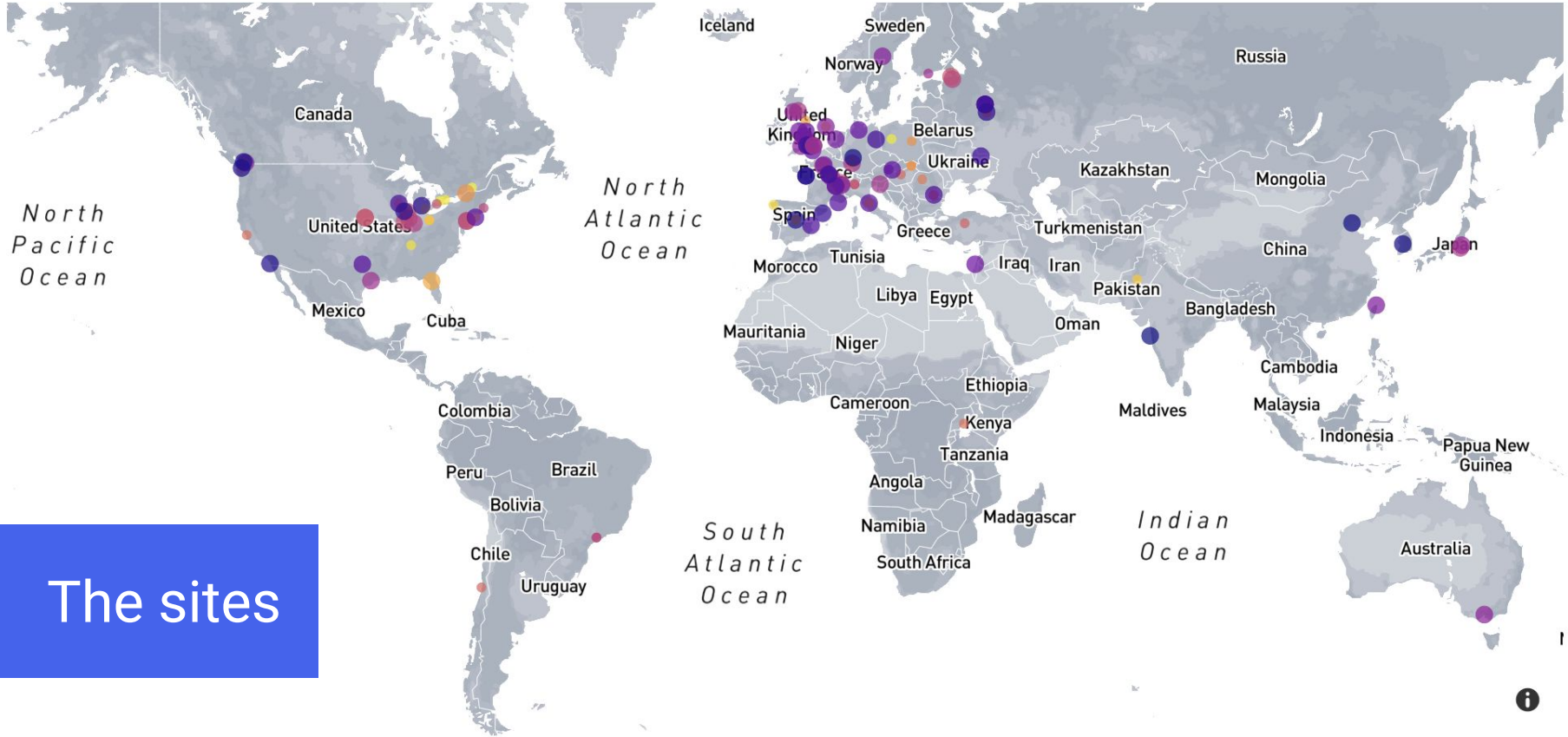
Find trivial and hard to detect problems

Determine locations of the root cause for a network problem

Automate

Notify relevant people

The sites


Iceland
Sweden
Norway
Russia
Canada
Belarus
Ukraine
Kazakhstan
Mongolia
North Pacific Ocean
North Atlantic Ocean
United States
Spain
Greece
Turkmenistan
China
Morocco
Tunisia
Iraq
Iran
Pakistan
Libya
Egypt
Bangladesh
Mexico
Cuba
Oman
Mauritania
Niger
Cambodia
Ethiopia
Colombia
Cameroon
Maldives
Malaysia
Kenya
Indonesia
Papua New Guinea
Tanzania
Peru
Brazil
Angola
Bolivia
Madagascar
Namibia
South Atlantic Ocean
Indian Ocean
Chile
South Africa
Australia
Uruguay

The types of tests

Latency

- ps_owd — Latency in ms between 2 hosts
- ps_packetloss — Percent lost packets between the 2 hosts

Throughput

- ps_throughput — Bandwidth
- ps_retransmits — Number of retransmitted packets

Network path

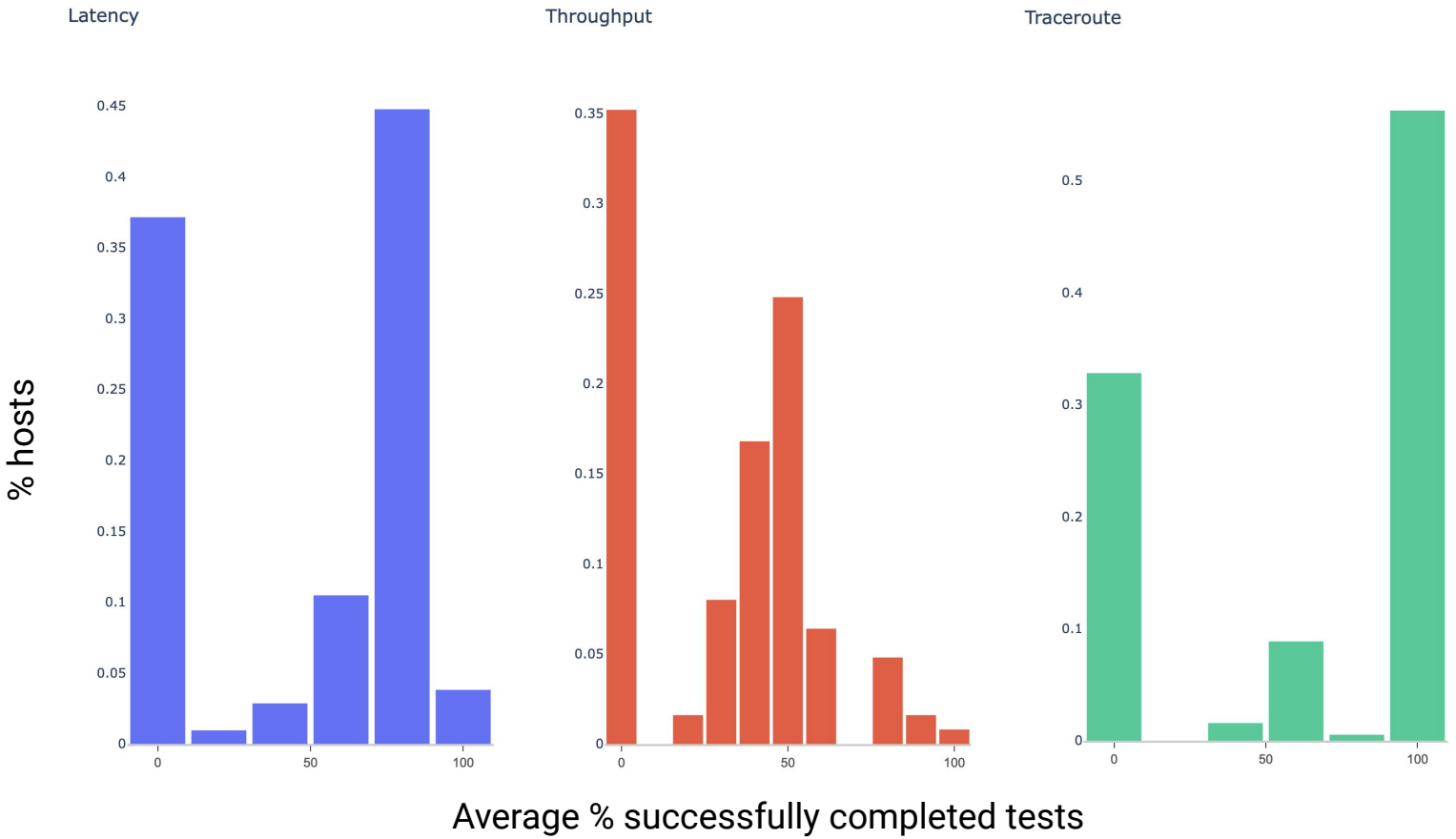
- ps_trace — The path between a source and destination
  - List of hops
  - List of ASNs
  - List of time to live values
  - List of round trip times
  - isPathComplete
  - isDestinationReached
  - hasLooping

The data volume

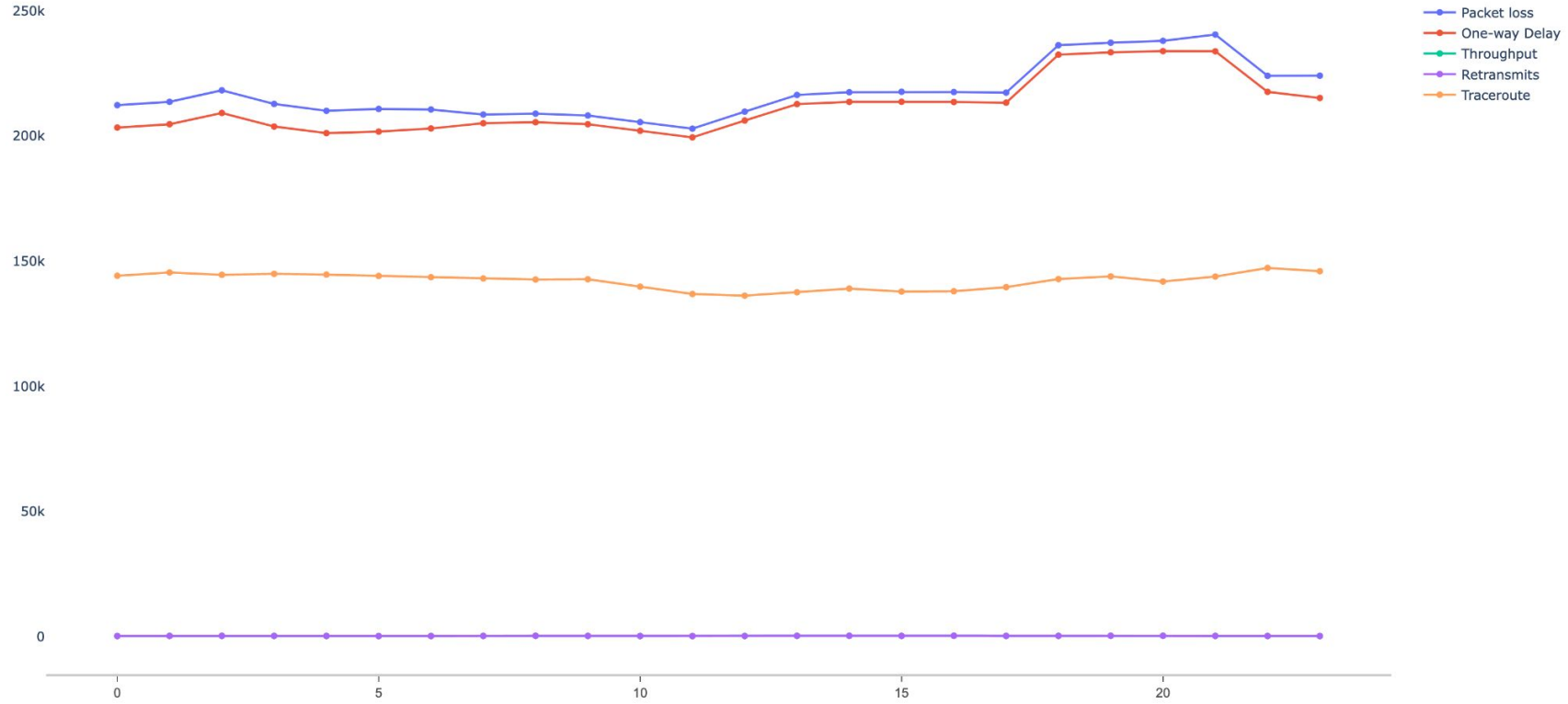
Type	Tests	Tests/day	Storage
Latency	6.91B	7.95M	3.1TB
Packet loss	7.00B	8.08M	2.4TB
Retransmits	14.7M	18.8K	6.3GB
Throughput	15.6M	19.2K	7.0GB
Network path	1.28B	2.14M	1.5TB

+ Metadata
+ Service status

Service status information for a random 24 hour window



Rate of tests per hour for a random 24 hour window



The data & the issues

Inconsistent
- metadata varies across indices

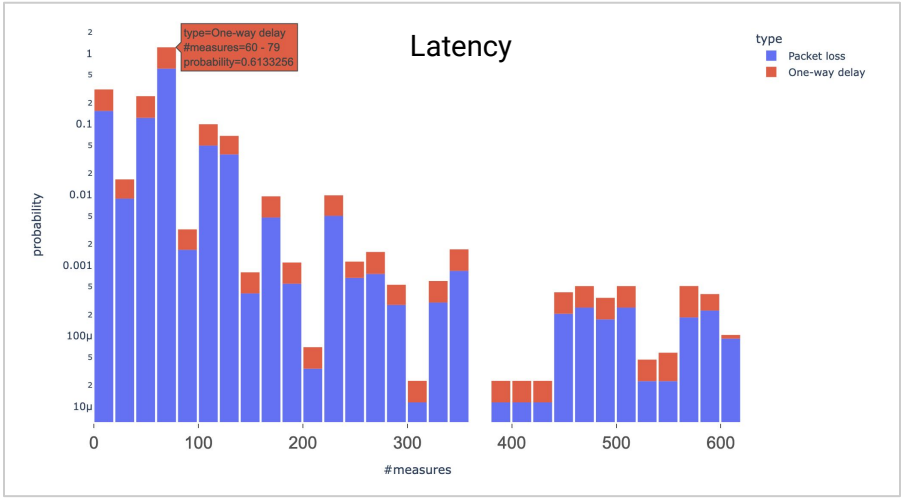
Redundant data

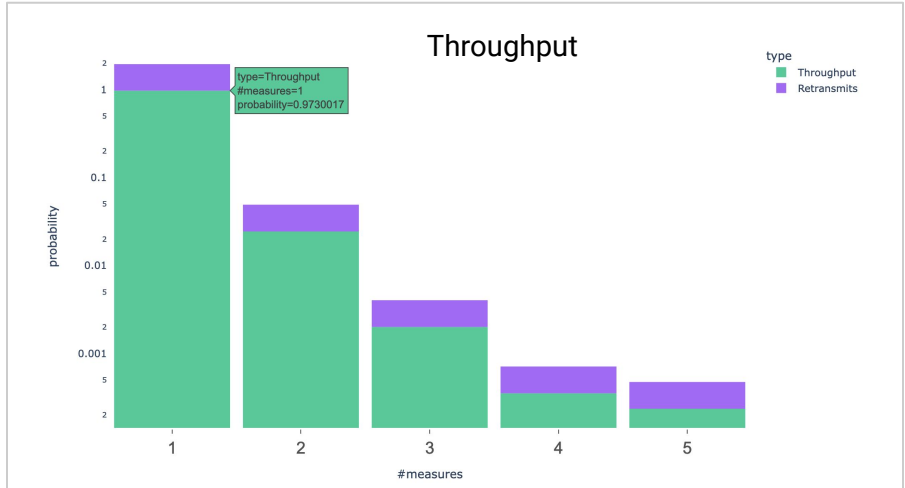
Incomplete
- test results
- metadata

Unknown/Undiscovered data

idx	src	src_site
ps_retransmits	129.107.255.30	UTA_SWT2
ps_retransmits	129.107.255.30	SWT2_CPB
ps_throughput	129.107.255.30	UTA_SWT2
ps_throughput	129.107.255.30	SWT2_CPB
ps_trace	129.107.255.30	SWT2_CPB
ps_trace	129.107.255.30	UTA_SWT2

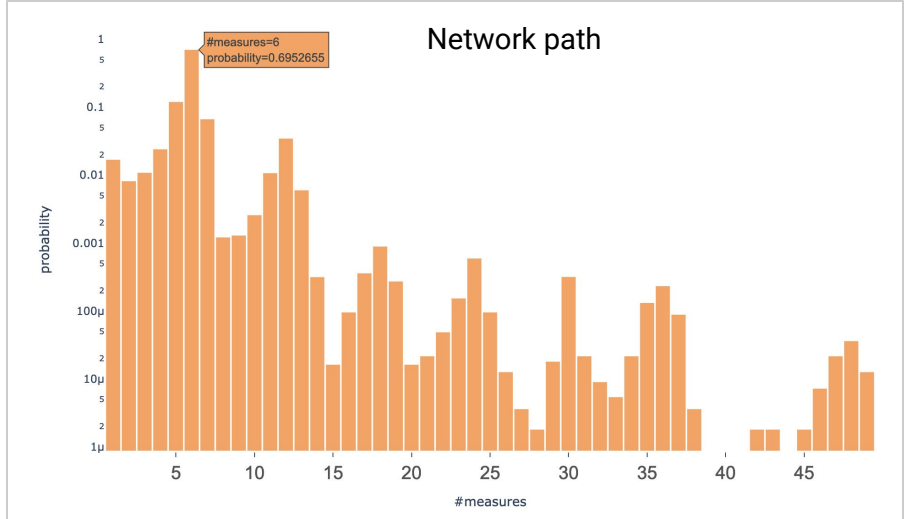
Example
(inconsistency)





Example
(redundancy)

The number of measures each tested pair makes for an hour



Example

(incompleteness)

Sites

CIT_CMS_T2 → RAL-LCG2

Hosts

perfsonar2.ultralight.org → lcgps01.gridpp.rl.ac.uk

IP addresses

198.32.44.4 → 130.246.176.109

198.32.44.4 → None

2605:d9c0:1:5::1:2 → 2001:630:58:1820::82f6:b06d

2605:d9c0:1:5::1:2 → None

Metadata lookup table

	ip	timestamp	host	site	administrator	email	lat	lon	site_meta	site_index	last_update
0	109.105.125.232	1.620854e+12	pship01.csc.fi	FI_HIP_T2	Johan Guldmyr	johan.guldmyr@csc.fi	60.1915	24.8993	None	FI_HIP_T2	2021-05-12 21:18
1	117.103.105.191	1.620855e+12	lhc-latency.twgrid.org	Taiwan-LCG2	Wenshui Chen	chenws@twgrid.org	25.0411	121.6166	ASGC	Taiwan-LCG2	2021-05-12 21:35
2	128.142.208.134	1.620855e+12	psl01-gva.cern.ch	CERN-PROD	CERN perfSONAR support	perfsonar-cern@cern.ch	46.2324045	6.0457234	CERN	CERN-PROD	2021-05-12 21:35
3	128.40.37.73	1.620855e+12	lcg-sonar01.hep.ucl.ac.uk	UKI-LT2-UCL-HEP	Ben Waugh	lcg-admin@hep.ucl.ac.uk	51.525269	0.134406	None	UKI-LT2-UCL-HEP	2021-05-12 21:28
4	129.215.213.70	1.620855e+12	gridpp-ps-lat.ecdf.ed.ac.uk	UKI-SCOTGRID-ECDF	UKI-SCOTGRID-ECDF	wlcg-support-ecdf@mlist.is.ed.ac.uk	55.9500	-3.2000	UKI-SCOTGRID-ECDF	UKI-SCOTGRID-ECDF	2021-05-12 21:28
...	...	...	...	...	...	...	...	...	...	...	...
215	2001:b30:1002:612::1119	1.646132e+12	perfsonar-grid.uaic.ro	RO-16-UAIC	admin	ciprian.pinzaru@uaic.ro	None	None	perfsonar-grid	RO-16-UAIC	2022-03-01 10:57
237	85.122.31.119	1.646132e+12	perfsonar-grid.uaic.ro	RO-16-UAIC	admin	ciprian.pinzaru@uaic.ro	None	None	perfsonar-grid	RO-16-UAIC	2022-03-01 10:57
55	194.36.11.38	1.649278e+12	perfsonar-latency.esc.qmul.ac.uk	UKI-LT2-QMUL	edg site admin	edg-site-admin@qmul.ac.uk	51.523	-0.041	UKI-LT2-QMUL	UKI-LT2-QMUL	2022-04-06 20:48
190	2605:9a00:10:200a:be97:e1ff:fee8:a1b0	1.649321e+12	uct2-net2.mwt2.org	MWT2	MWT2	mwt2-support@lists.uchicago.edu	41.789889	-87.600453	MWT2	MWT2	2022-04-07 08:48
64	2001:2f8:3e:cc21::4b	NaN	perfsonar2.cc.kek.jp	JP-KEK-CRC-02	None	None	36.148	36.148	None	JP-KEK-CRC-02	None

393 rows × 11 columns

PS-Dash App

perfSONAR Toolkit Information | Kibana: Packet Loss in OSG/WLCG | Kibana: Packet Loss Tracking | MEPHi Tracer: Traceroute explorer

ps dash | SITES | LINKS | PLOTS

Sites' ranking based on their measures

The darker the color, the worse their performance. Smaller points indicate missing set of measures. Click on a site in the map to see an overview over the past days

IN2P3-SUBATECH

Latency hosts		Throughput hosts	
IPv4	IPv6	IPv4	IPv6
2	2	2	2

PACKETLOSS (packets)

TODAY IN	TODAY OUT
0	0

Change over the past 3 days (%)

	27/04	28/04	29/04
IN	-0.94	-0.05	-0.77
OUT	-0.1	-0.36	-0.84

THROUGHPUT (MBps)

TODAY IN	TODAY OUT
3007.26	2034.13

Change over the past 3 days (%)

	27/04	28/04	29/04
IN	-0.04	+0.06	-0.02
OUT	+0.07	+0.27	0

OWD (ms)

TODAY IN	TODAY OUT
-34.05	39.54

Change over the past 3 days (%)

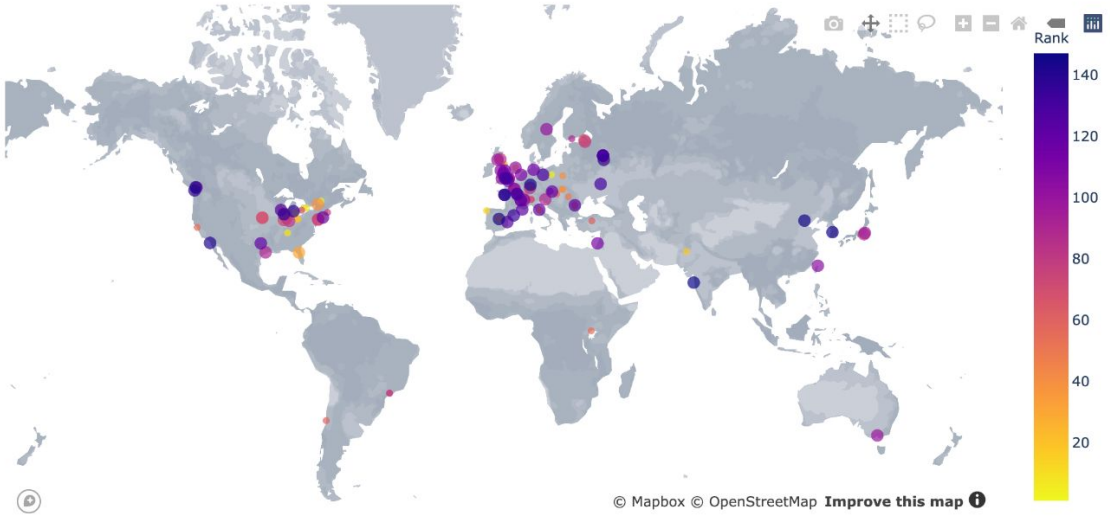
	27/04	28/04	29/04
IN	+0.26	-0.21	+0.06
OUT	+0.19	-0.23	+0.18

RETRANSMITS (packets)

TODAY IN	TODAY OUT
1544.34	182.24

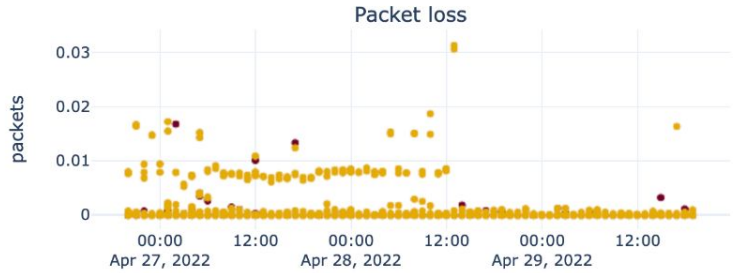
Change over the past 3 days (%)

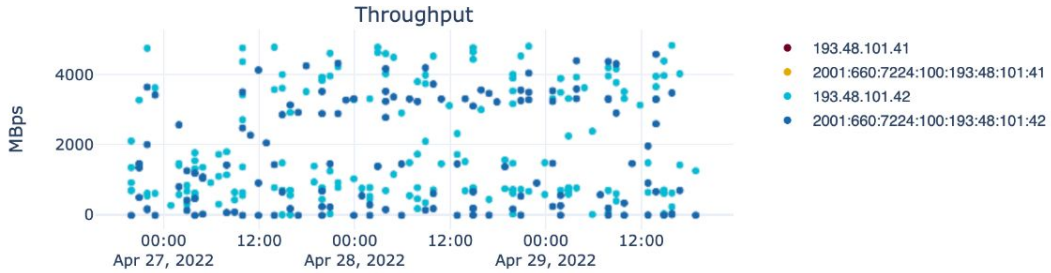
	27/04	28/04	29/04
IN	-0.06	+0.77	-0.5
OUT	+1.15	-0.46	-0.52

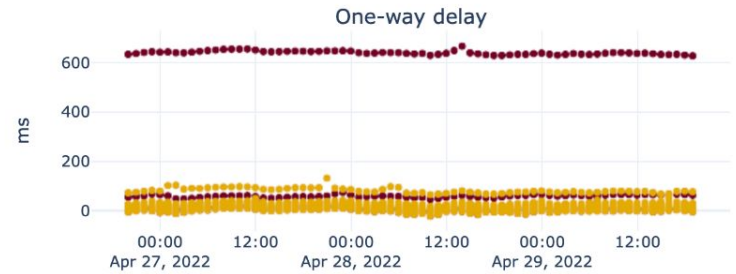


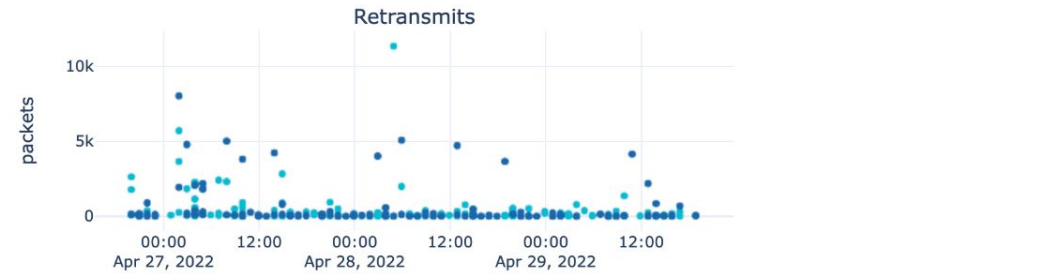
IN2P3-SUBATECH as destination of measures

IN2P3-SUBATECH as source of measures







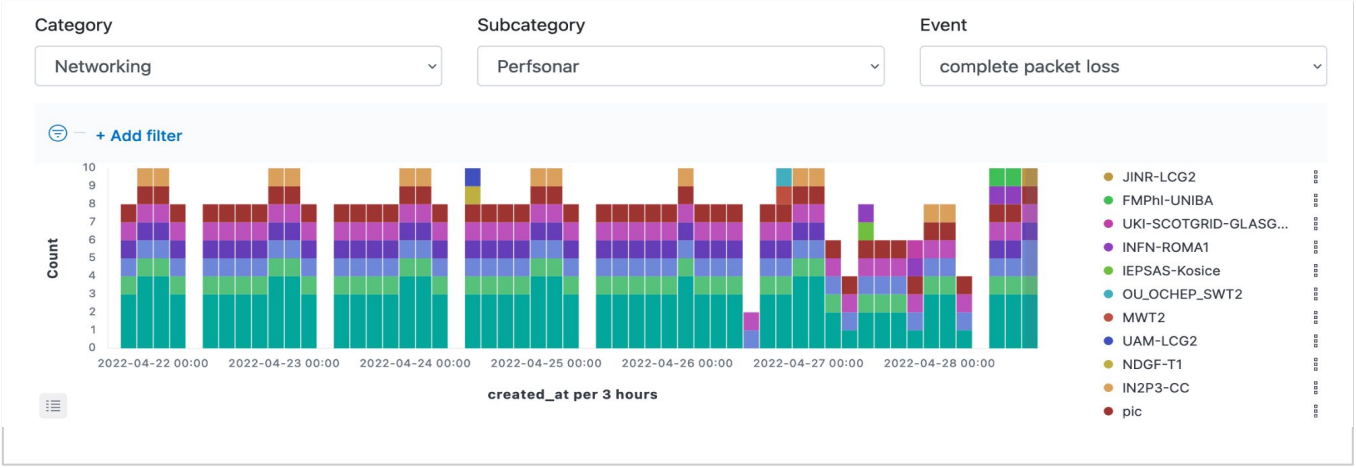


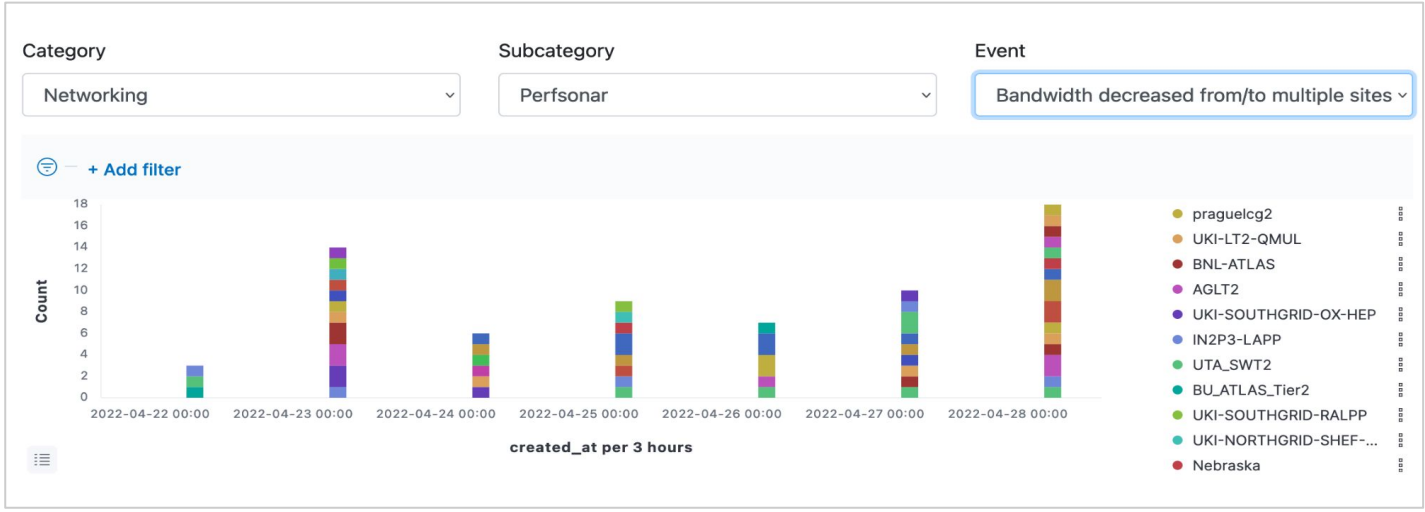
Alarms & Alerts

Current alarm types:

- Bad owd measurements
- Complete packet loss
- Firewall issue
- Large clock correction
- High packet loss on multiple links
- Bandwidth increased from/to multiple sites
- Bandwidth decreased from/to multiple sites
- Bandwidth increased
- Bandwidth decreased
- Destination cannot be reached from multiple
- Destination cannot be reached from any
- Source cannot reach any

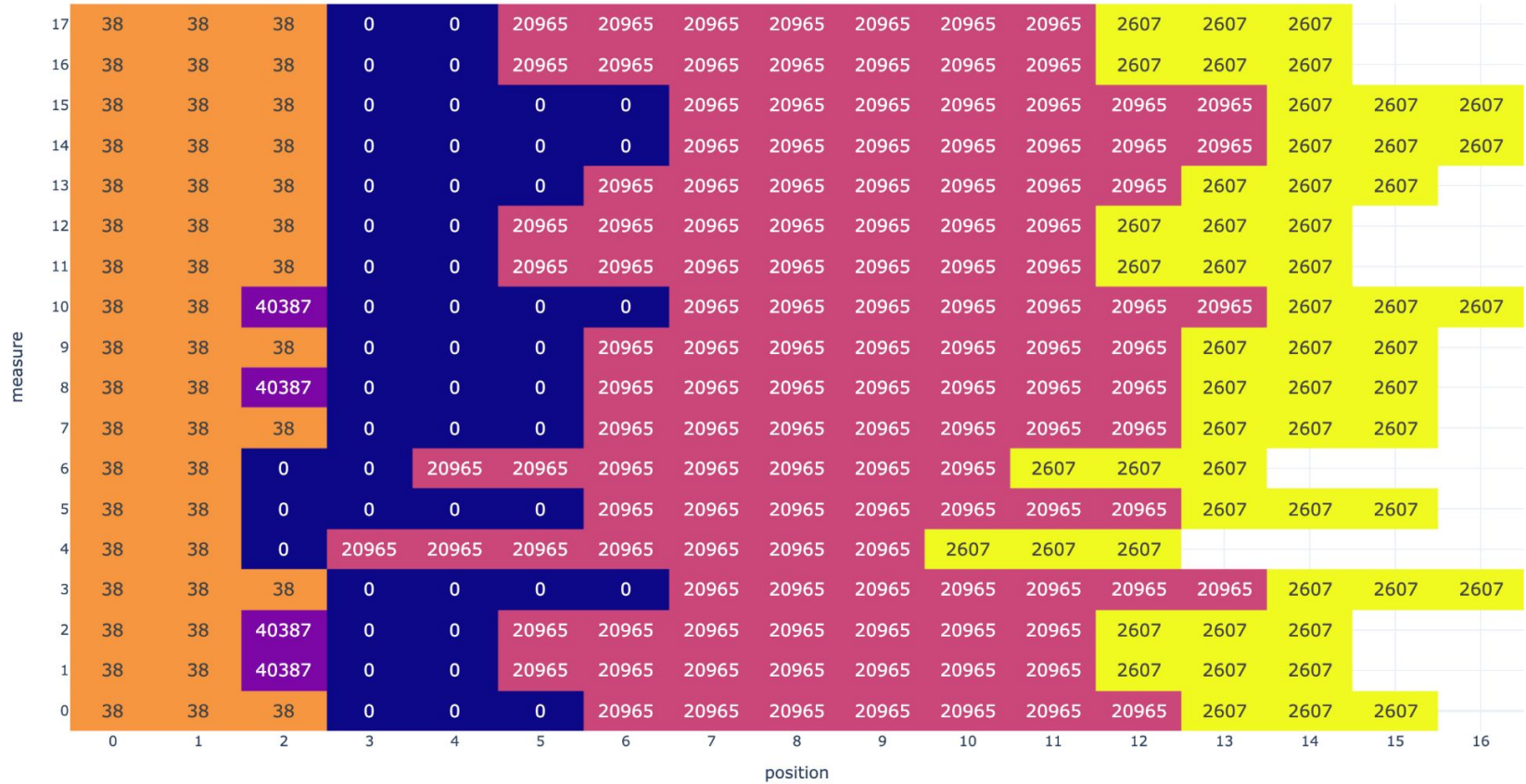
Alarms & Alerts

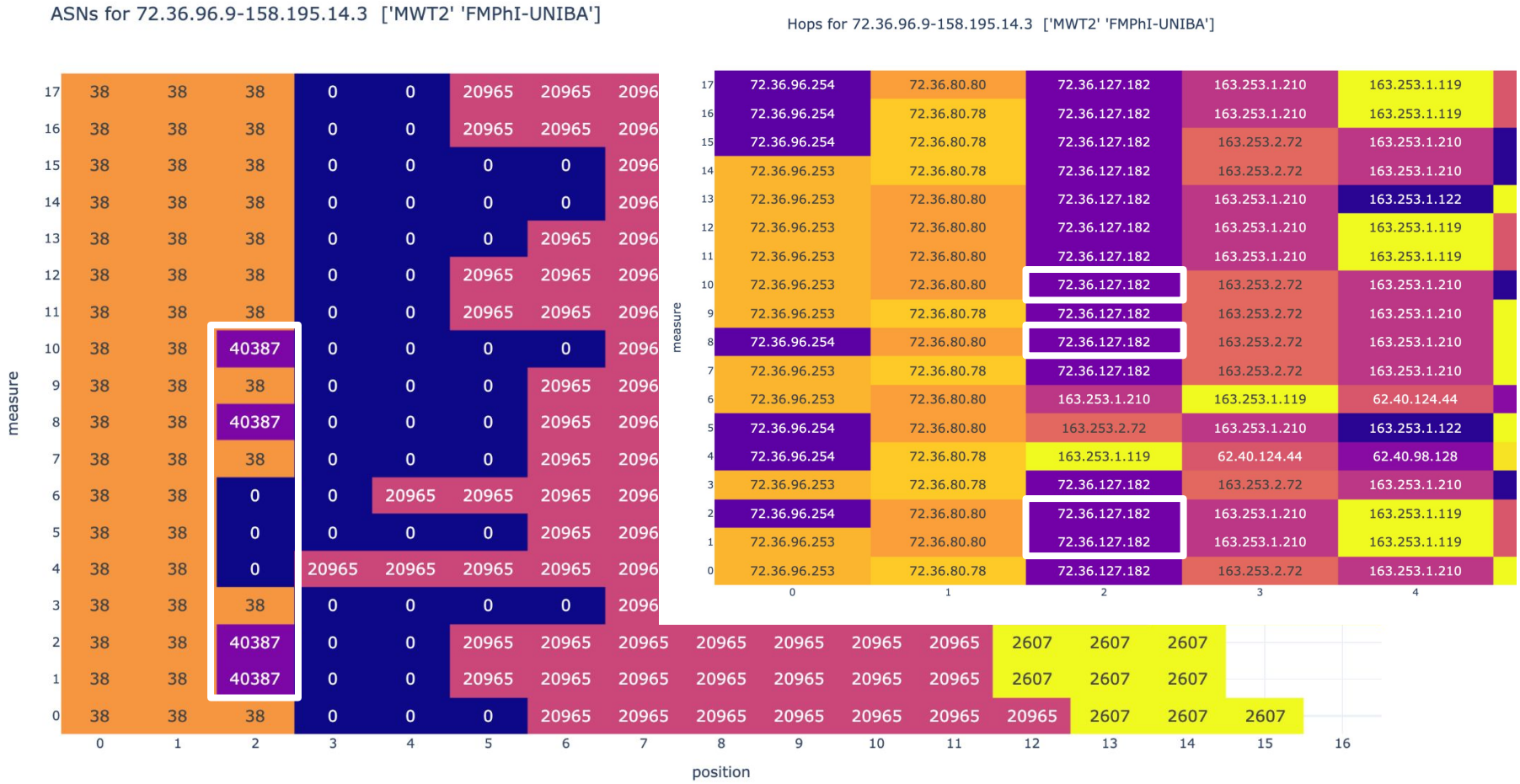




The network path analysis

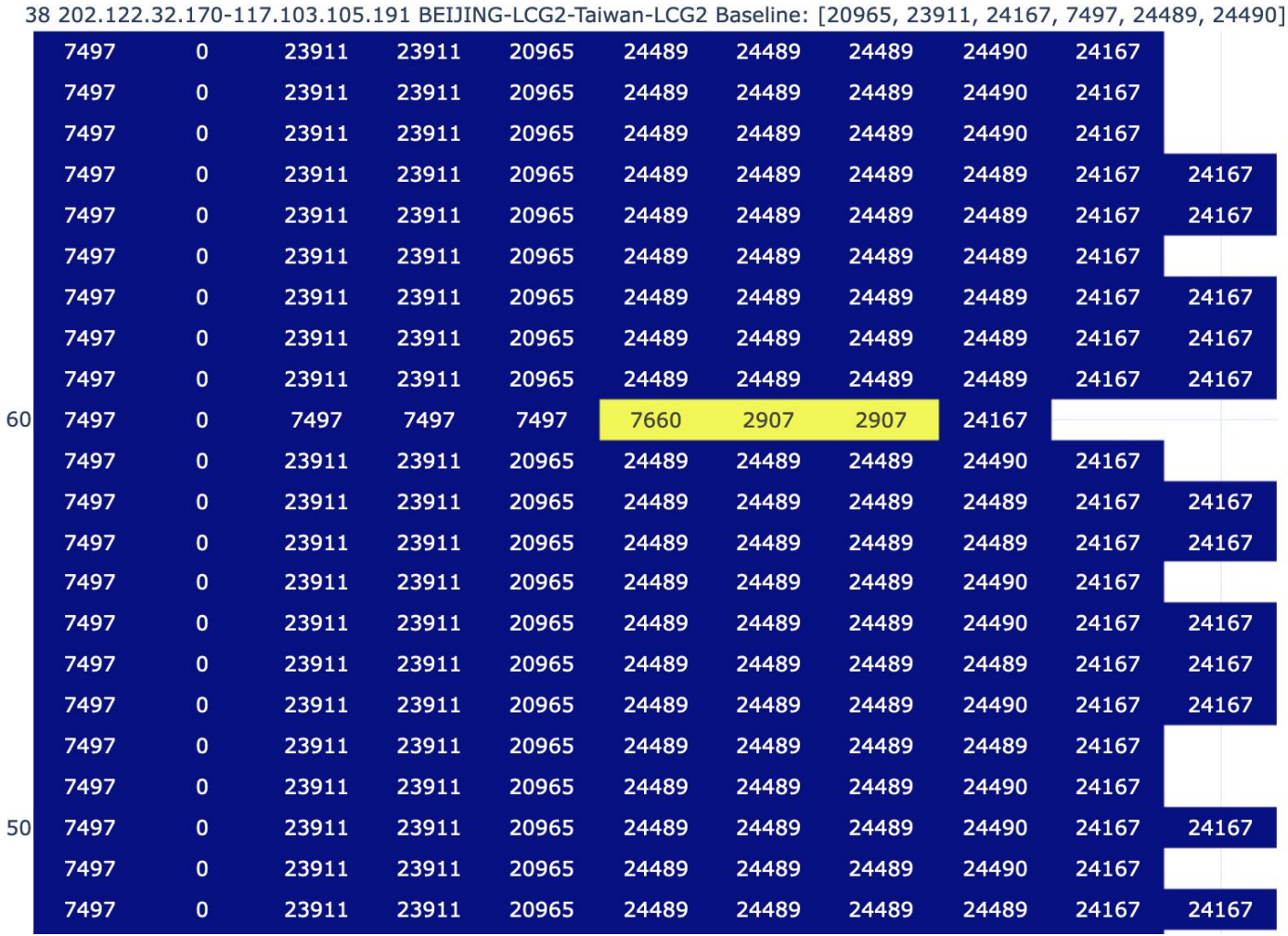
An **autonomous system** (AS) is a group of IP prefixes with a clearly defined external routing policy. In order for multiple autonomous systems to interact, each needs to have a unique identifier - an AS number.



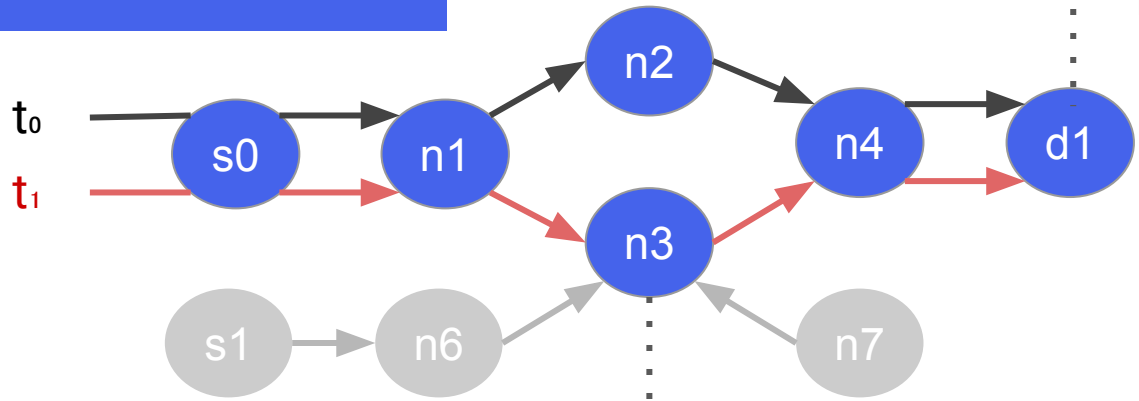

ASNs for 72.36.96.9-158.195.14.3 ['MWT2' 'FMPhI-UNIBA']
Hops for 72.36.96.9-158.195.14.3 ['MWT2' 'FMPhI-UNIBA']
measure
position
17 38 38 38 0 0 20965 20965 2096
16 38 38 38 0 0 20965 20965 2096
15 38 38 38 0 0 0 0 2096
14 38 38 38 0 0 0 0 2096
13 38 38 38 0 0 0 20965 2096
12 38 38 38 0 0 20965 20965 2096
11 38 38 38 0 0 20965 20965 2096
10 38 38 40387 0 0 0 0 2096
9 38 38 38 0 0 0 20965 2096
8 38 38 40387 0 0 0 20965 2096
7 38 38 38 0 0 0 20965 2096
6 38 38 0 0 20965 20965 20965 2096
5 38 38 0 0 0 0 20965 2096
4 38 38 0 20965 20965 20965 20965 2096
3 38 38 38 0 0 0 0 2096
2 38 38 40387 0 0 20965 20965 20965 20965 20965 20965 20965 20965 2607 2607 2607
1 38 38 40387 0 0 20965 20965 20965 20965 20965 20965 20965 20965 2607 2607 2607
0 38 38 38 0 0 0 20965 20965 20965 20965 20965 20965 2607 2607 2607
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
17 72.36.96.254 72.36.80.80 72.36.127.182 163.253.1.210 163.253.1.119
16 72.36.96.254 72.36.80.78 72.36.127.182 163.253.1.210 163.253.1.119
15 72.36.96.254 72.36.80.78 72.36.127.182 163.253.2.72 163.253.1.210
14 72.36.96.253 72.36.80.78 72.36.127.182 163.253.2.72 163.253.1.210
13 72.36.96.253 72.36.80.80 72.36.127.182 163.253.1.210 163.253.1.122
12 72.36.96.253 72.36.80.80 72.36.127.182 163.253.1.210 163.253.1.119
11 72.36.96.253 72.36.80.80 72.36.127.182 163.253.1.210 163.253.1.119
10 72.36.96.253 72.36.80.80 72.36.127.182 163.253.2.72 163.253.1.210
9 72.36.96.253 72.36.80.78 72.36.127.182 163.253.2.72 163.253.1.210
8 72.36.96.254 72.36.80.80 72.36.127.182 163.253.2.72 163.253.1.210
7 72.36.96.253 72.36.80.78 72.36.127.182 163.253.2.72 163.253.1.210
6 72.36.96.253 72.36.80.80 163.253.1.210 163.253.1.119 62.40.124.44
5 72.36.96.254 72.36.80.80 163.253.2.72 163.253.1.210 163.253.1.122
4 72.36.96.254 72.36.80.78 163.253.1.119 62.40.124.44 62.40.98.128
3 72.36.96.253 72.36.80.78 72.36.127.182 163.253.2.72 163.253.1.210
2 72.36.96.254 72.36.80.80 72.36.127.182 163.253.1.210 163.253.1.119
1 72.36.96.253 72.36.80.80 72.36.127.182 163.253.1.210 163.253.1.119
0 72.36.96.253 72.36.80.78 72.36.127.182 163.253.2.72 163.253.1.210
0 1 2 3 4

The path has changed.

Send an alarm.



Next steps: GNN



Is the destination host healthy?
Are there other tests showing the same issues to this destination?

Path1: no lost packets

Path2: dropped all packets

n3 to blame?
What's the level of certainty?

What's the status of the local network?
What's the AS number? Does it participate in other flagged paths?

Is this node important?
Are there many in/out links?

Does this happen in one direction only?
What are the previous hops? Is it an issue with the link?

Thank you!

Questions?