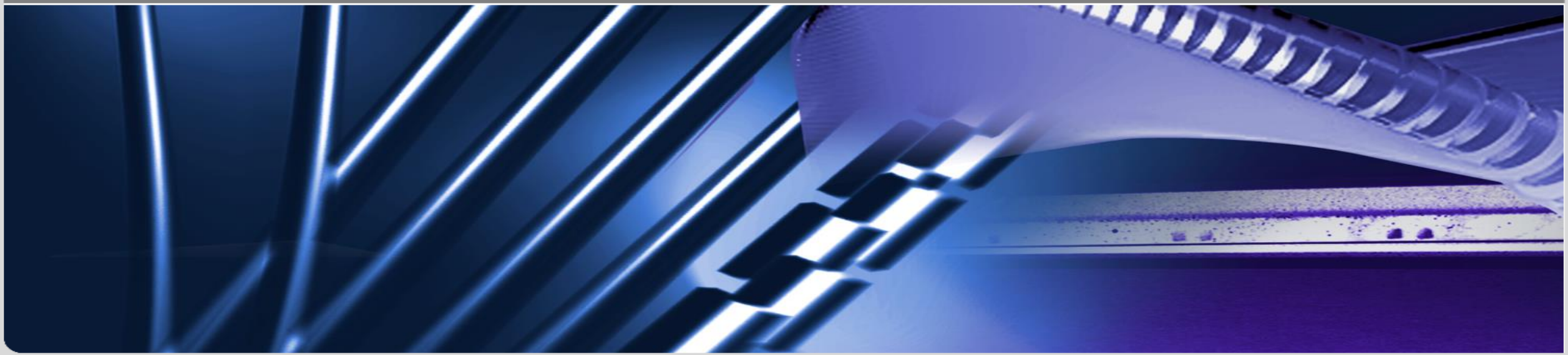
unroutable LHCONE traffic

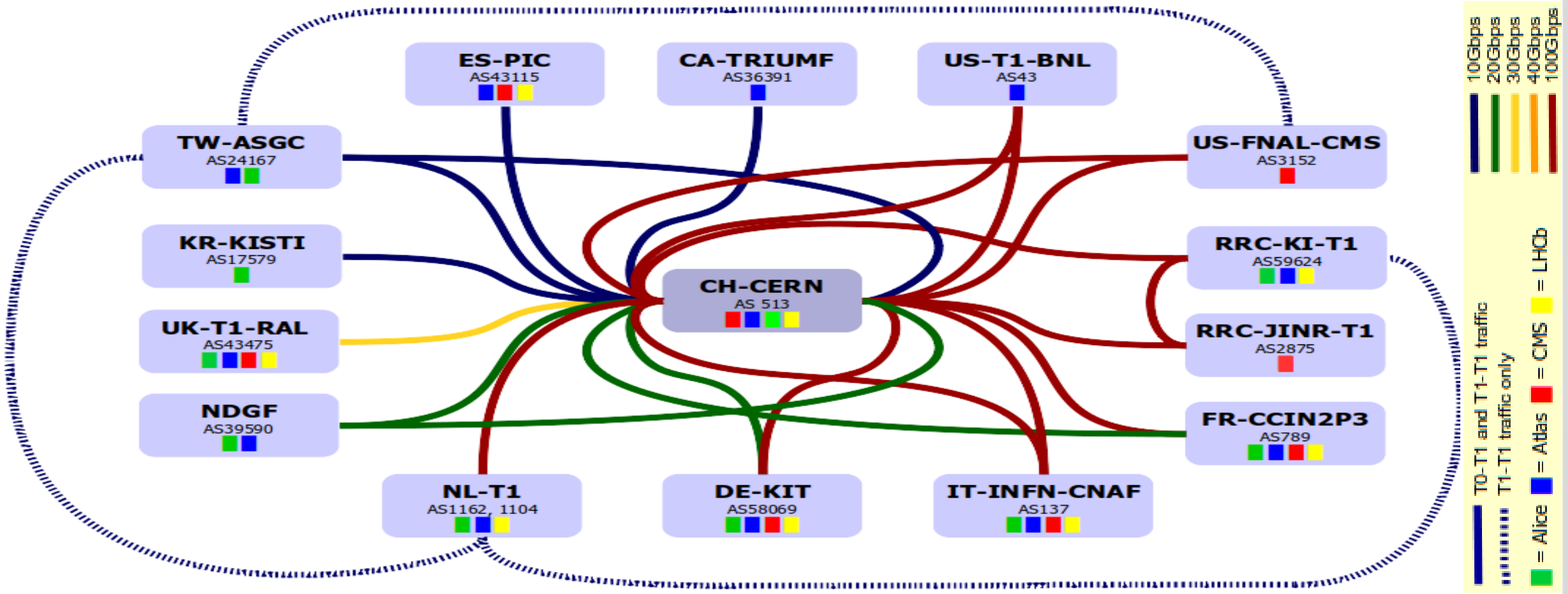
Bruno Hoeft / KIT
Michael O'Connor / ESnet
Richard Cziva / Esnet
Samuel Ambroj Perez / KIT
Magnus Bergroth / NORDUnet

STEINBUCH CENTRE FOR COMPUTING - SCC



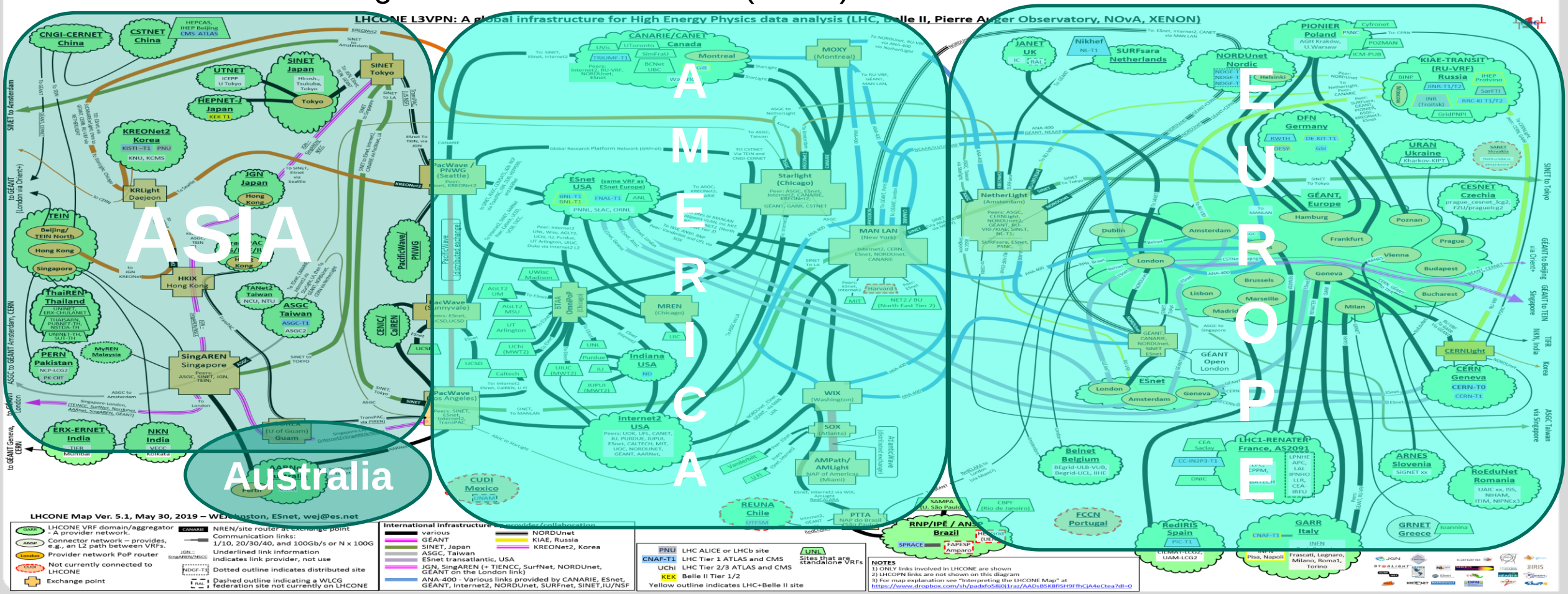
LHC → VPNs

- LHCOpticalPrivatNetwork → a CERN centric star VirtualPrivatNetwork
  - VPN with 15 participants only, mutual agreements between connected partner, no formal rules



LHCOpenNetworkEnvironment – MAP

- LHCONE → a still growing international distributed VPN
  - VPN with 104 endsites connected through 26 VirtualRouting andForwarding implenentations at the connecting Network Service Providers (NSPs)



NSP Packet Filtering Requirements

All LHCONE Traffic is subject to the following conditions:

- Traffic injected into the LHCONE must only be originated from addresses within an LHCONE routable prefix
- Only address ranges present in the LHCONE routing table should be transported on the network

Objective: In order to maintain route symmetry and access control, each NSP will implement policy and packet filters to manage their connected customer address prefix ranges.

- Ensures that a return route exists in the LHCONE network
- Blocks spoofed packets (Similar to BCP 38)

https://twiki.cern.ch/twiki/pub/LHCONE/LhcOneVRF/LHCONEconnectionguide-1.2.pdf

NSP Border Gateway Protocol (BGP) Import Policy

Prefix Lists will be negotiated between connecting institutions and their NSP within the constraints imposed by the LHCONE Acceptable Use Policy (AUP).

LHCONE NSPs have agreed to to configure:

1. BGP import filters
2. Source address packet filters

End sites are encouraged to implement source address filters at their edge in order to eleminate their own unroutable LHCONE packets. NSPs will generally discard none compliant packets without informing the site.

Connecting institutions/sites will not add prefixes to the LHCONE routing table without direct cooperation with their NSP.

AUP: https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneAup

The Investigation

ESnet

- Three months of ESnet netflow IPv4 & IPv6 sampling from Feb. 2019 - April 2019 for the following connected sites and peers

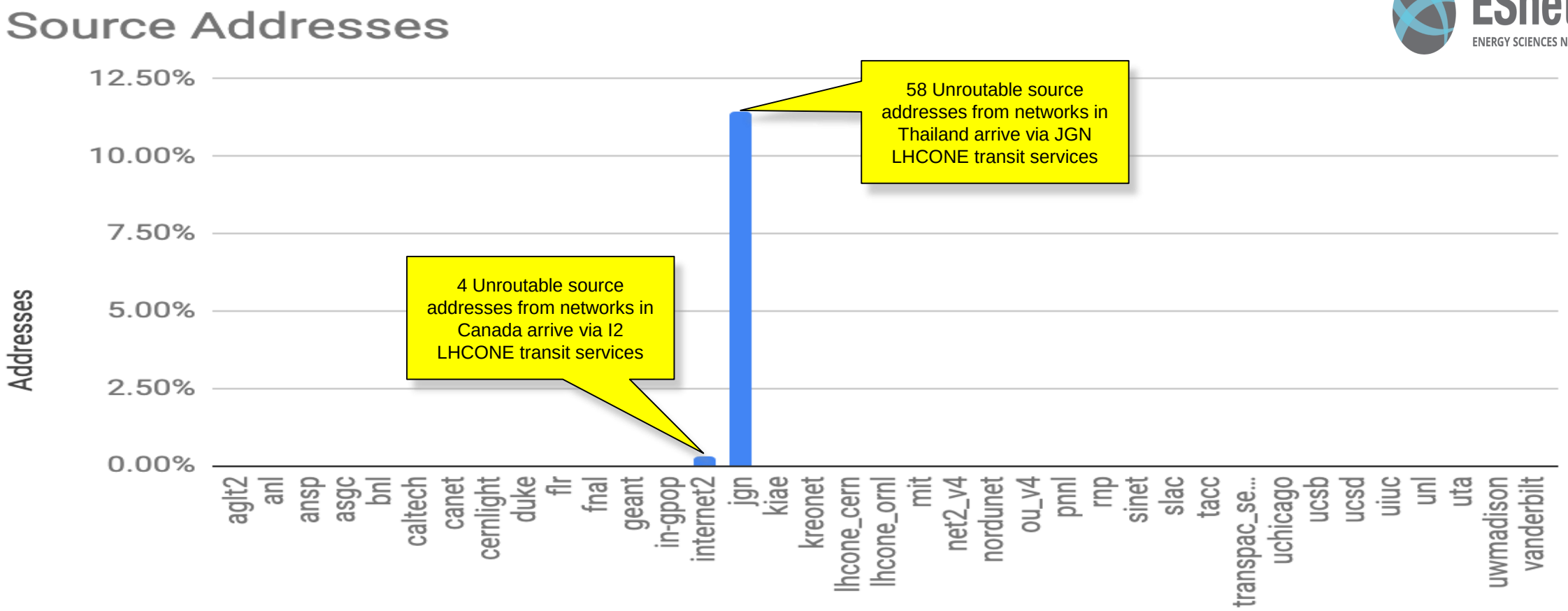
aarnet	duke	ornl	transpac
aglt2	flr	mit	uchicago
anl	fnal	net2	ucsb
ansp	geant	nordunet	ucsd
asgc	ind-gpop	ou	uiuc
bnl	internet2	pnnl	unl
caltech	JGN	rnp	uta
canet	kiae	sinet	uwmadison
cern	kreonet	slac	vanderbilt
cernlight	cern	tacc	

ESnet counted:

- All LHCONE ingress packets
- Unroutable source packets
- Packets with non-lhcone/missing origin ASN

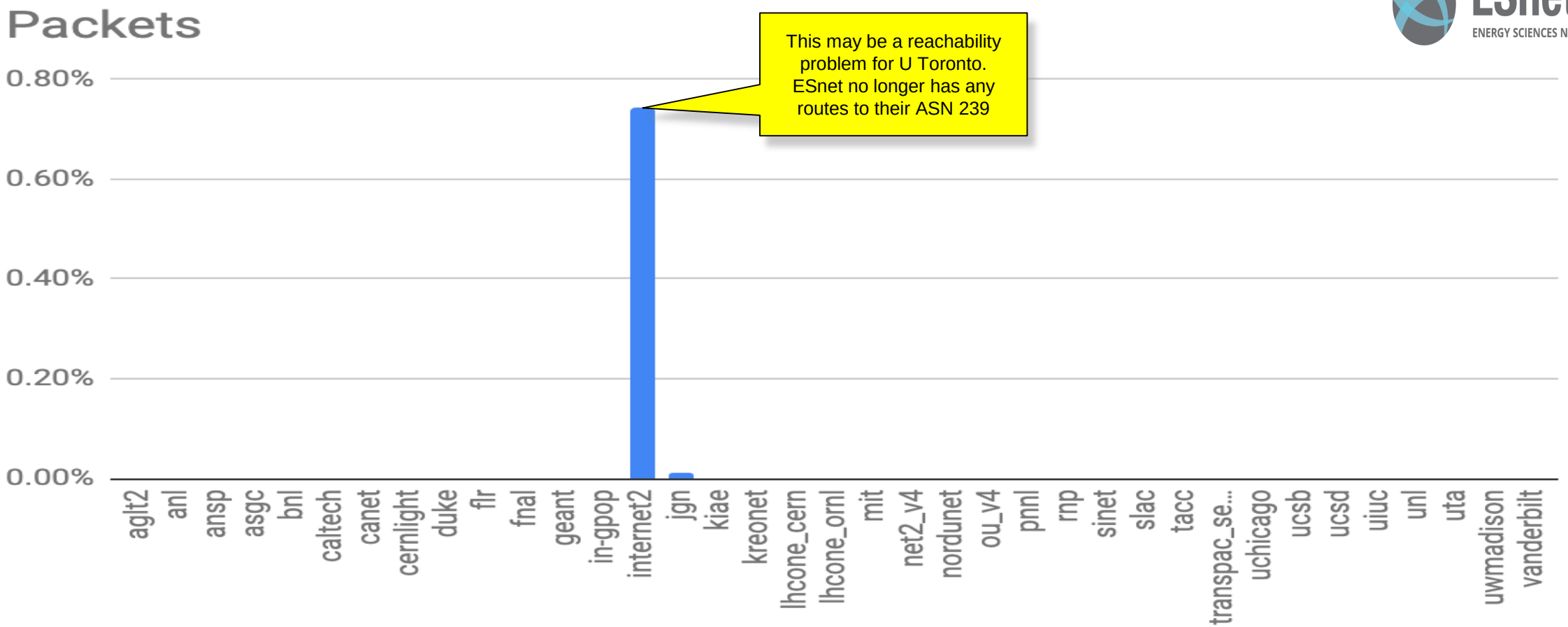
* corrected for netflow sampling rate (1000)

ESnet monitoring Unroutable Source Addresses by percentage



Very good results!!!

ESnet monitoring



A small number of U Toronto hosts are transmitting large amounts of data on LHCONE. ESnet had routes for U Toronto in the past.

DE-KIT : unroutable IPv4/6 LHCONE packet statistics

unsampled **Ingress Packet Filters** based on LHCONE routing table

	IPv4	IPv6
Total packets (during 4 weeks)	**1.044.471**	**1.376.338**
Packet/day	**37.302**	**35.290**
Privat IP destination	10.0.0.0/16, 172.16.0.0/12, 192.168.0.0/16	fe00::/16
Number of sources	44 + private	33 + private

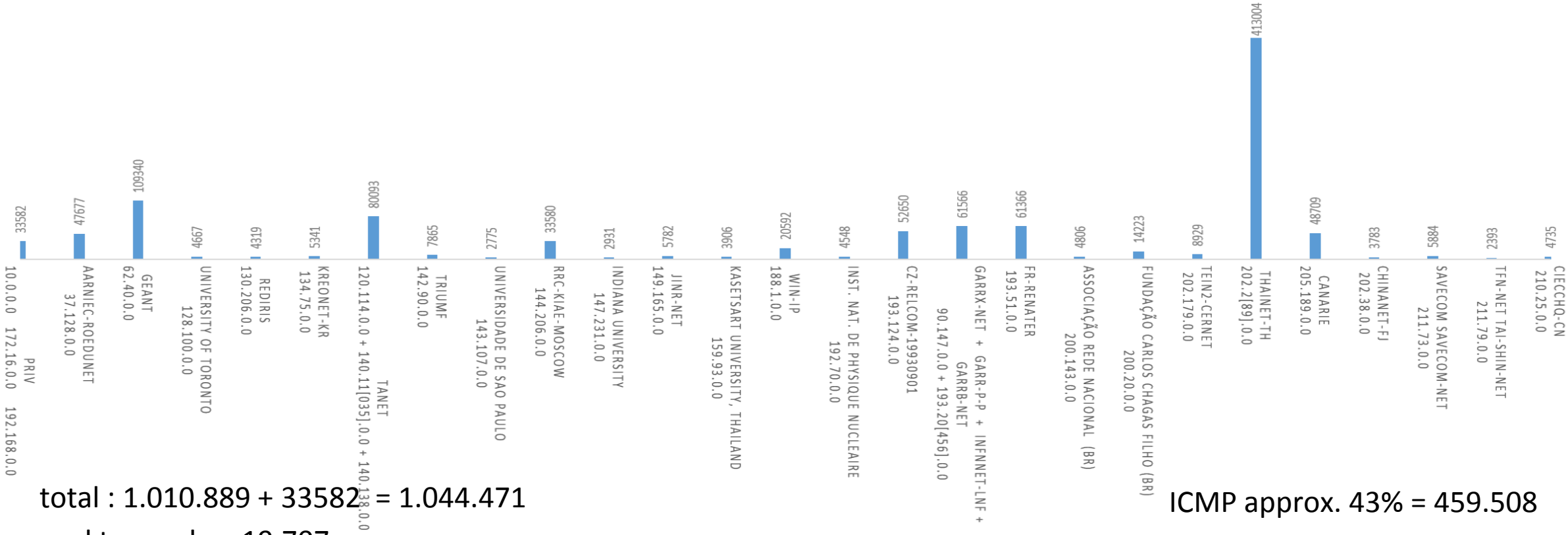
Total LHCONE traffic → 0,2% / none ICMP → approx. 57%

- priv
- AARNIEC-RoEduNet
- AIT-TH
- ASGC-NET
- Associação Rede Nacional (BR)
- CANARIE
- Bcnet Vancouver
- CAS-PRG-6TCZ
- CAS-TCZ
- CERNET
- CHINANET-FJ
- CIECCHQ-CN
- CIEMAT
- CZ-RELCOM-19930901
- DFN WIN-IPV6
- ERNET-IN
- ES-REDIRIS-20010521
- Fundação Carlos Chagas Filho (BR)
- FR-RENATER
- FR-IN2P3-LAL-ORSAY
- FR-IN2P3-LLR-PALAISEAU
- FR-CEA-SACLAY-GRILLES
- FR-IN2P3-LPNHE-PARIS
- FR-IN2P3-LPC-CLERMONT-AUBIERE
- FR-IN2P3-LAPP-ANNECY
- FR-IN2P3-CPPM-MARSEILLE
- GEANT
- GARR-P-P
- GARRB-NET
- GARRX-NET
- GR-GRNET-19991208
- GZIN
- IANA – reserved
- IHEP-IPv6
- Imperial College London
- Indiana University
- INFNNET-LNF
- Inst. Nat. de Physique Nucleaire
- IT-GARR-20011004
- JINR-NET
- Kasetsart University, Thailand
- KREONet-KR
- NL-GEANT-20020131
- PNPI
- REDIRIS
- RoEduNet-IPv6-NET-1
- RRC-KIAE-Moscow
- RU-ROSNIIROS-20180806
- RWTH Aachen
- SAVECOM SAVECOM-NET
- SINET-JPNIC
- SUT-TH
- T-NCU.EDU.TW-NET
- T-NSYSU.EDU.TW-NET
- T-NTHU.EDU.TW-NET
- TANET
- TANET-B T-HCRC.EDU.TW-NET
- TANET-BNETA
- TANET-BNETS Taiwan
- TANET-NET
- TANET-NET Taiwan
- TANET Taiwan
- TEIN2-CERNET
- THAINET-TH
- TFN-NET TAI-SHIN-NET
- TRIUMF
- UAM
- UK-GEANT-20020131
- UNI Michigan
- UNIVERSIDADE DE SAO PAULO
- UNI of Nebraska-Lincoln
- University of Toronto
- IJS-IPv6-NET - Ljubljana
- VANDERBILT
- WIN-IP

Color legend : IPv6 and IPv4 / IPV4 / IPv6

unroutable IPv4 packets

- Reduced to 30 different sites only by
  - Combining private address areas
  - Removed sites with less than 1000 packets (per month)
  - Pull different subnets of one source/site together



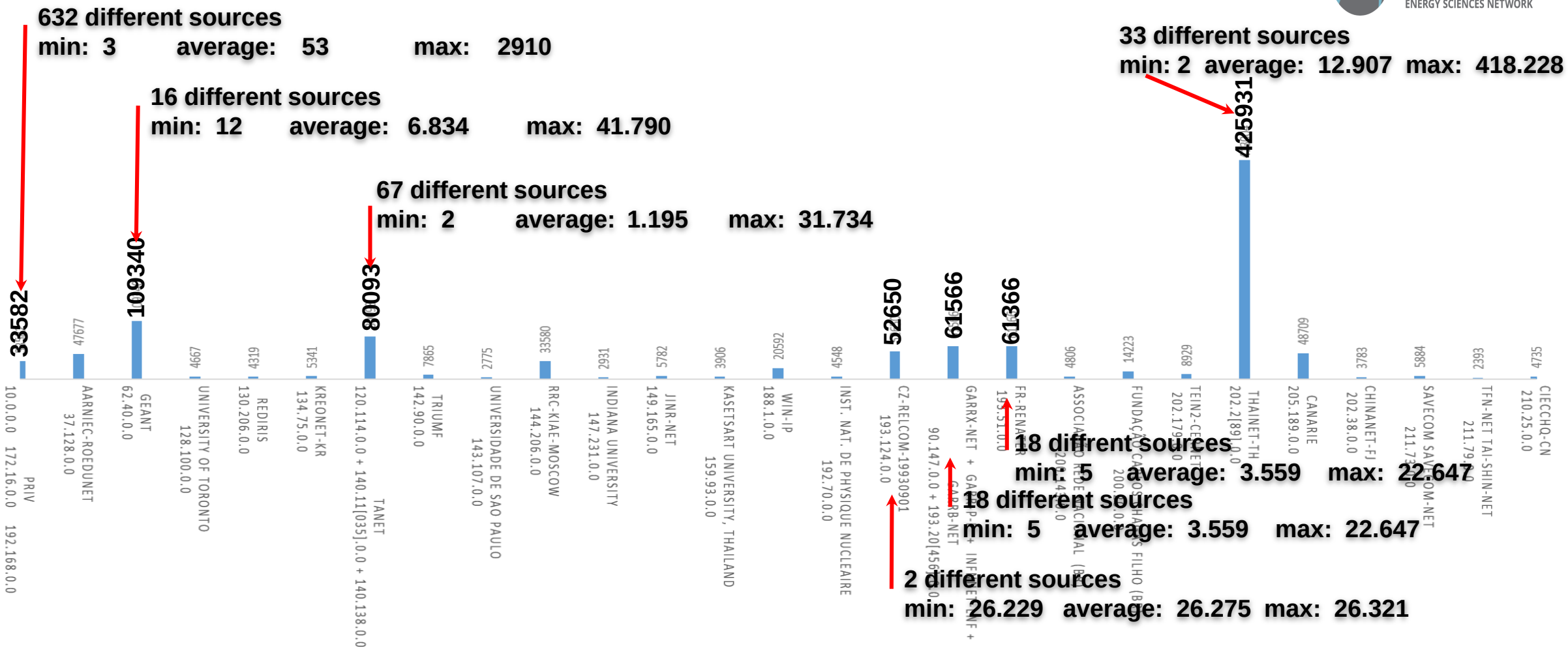
total : 1.010.889 + 33582 = 1.044.471

packts per day: 19.707

ICMP approx. 43% = 459.508

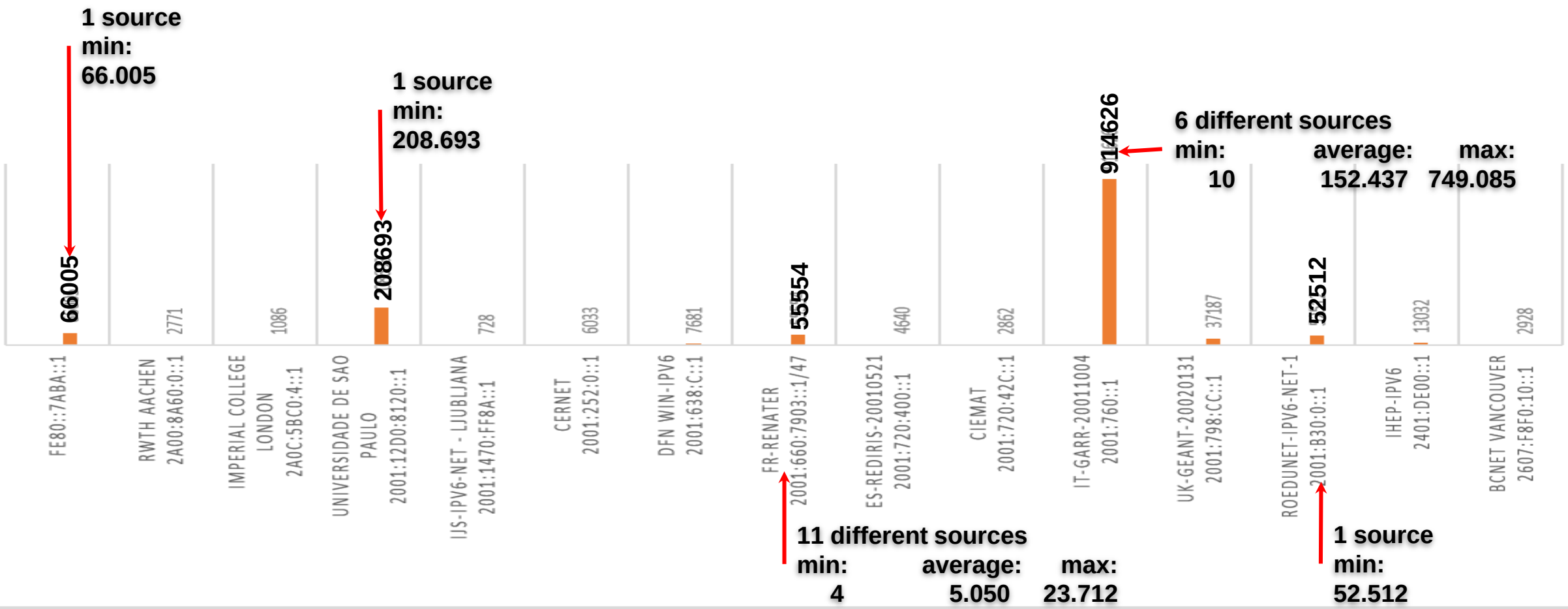
none ICMP = 584.963

Reduced by 1 Gbps packets FR-Renater and GARR (ICMP only)



IPv6 Filter : unroutable LHCONE packet

- filter Privat address space (link local)
- FR-Renater → ICMP packts of network devices
- IT-GARR → ICMP packet of network devices and the main injector (2 Perfsonar Server) are became part of LHCONE



Project @ DE-KIT : unroutable packet -- web portal

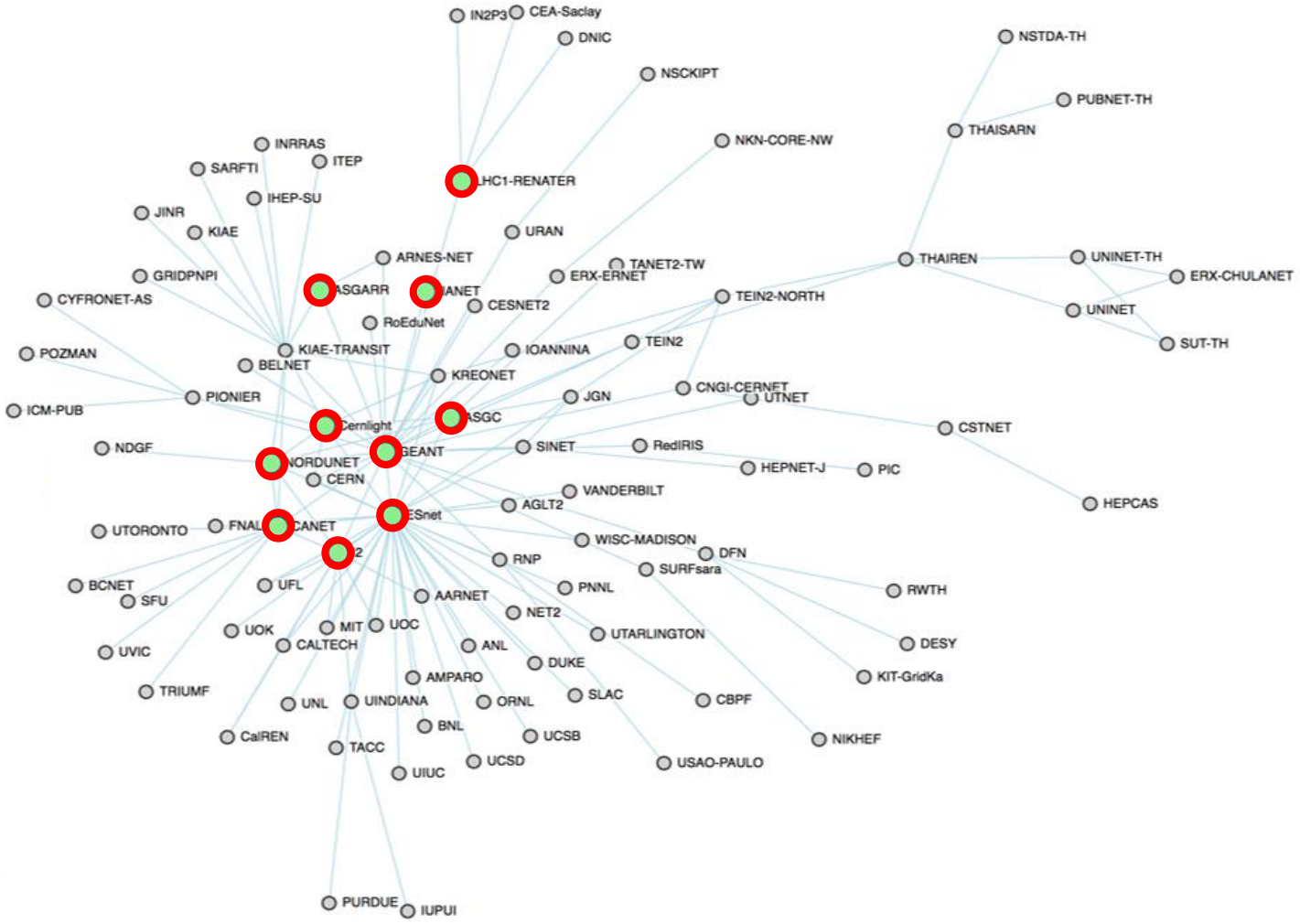
- automate unroutable packet information gathering
  - → store into a organized and structured database (kibana)

  - visualise the data (packet beat / elastic search / Kibana / grafana) with different levels
    - abstract overview
    - and zoomable into detailed view (up to source/dest. of a single packet)
      this data shall be available for the LHCONE connected sites (but not for the world),
      one idea:
    - community securing the data
    - restrict access --> personal authentication enabled (via eduGAIN)

project just started,
working on first results by the end of this year (2019)

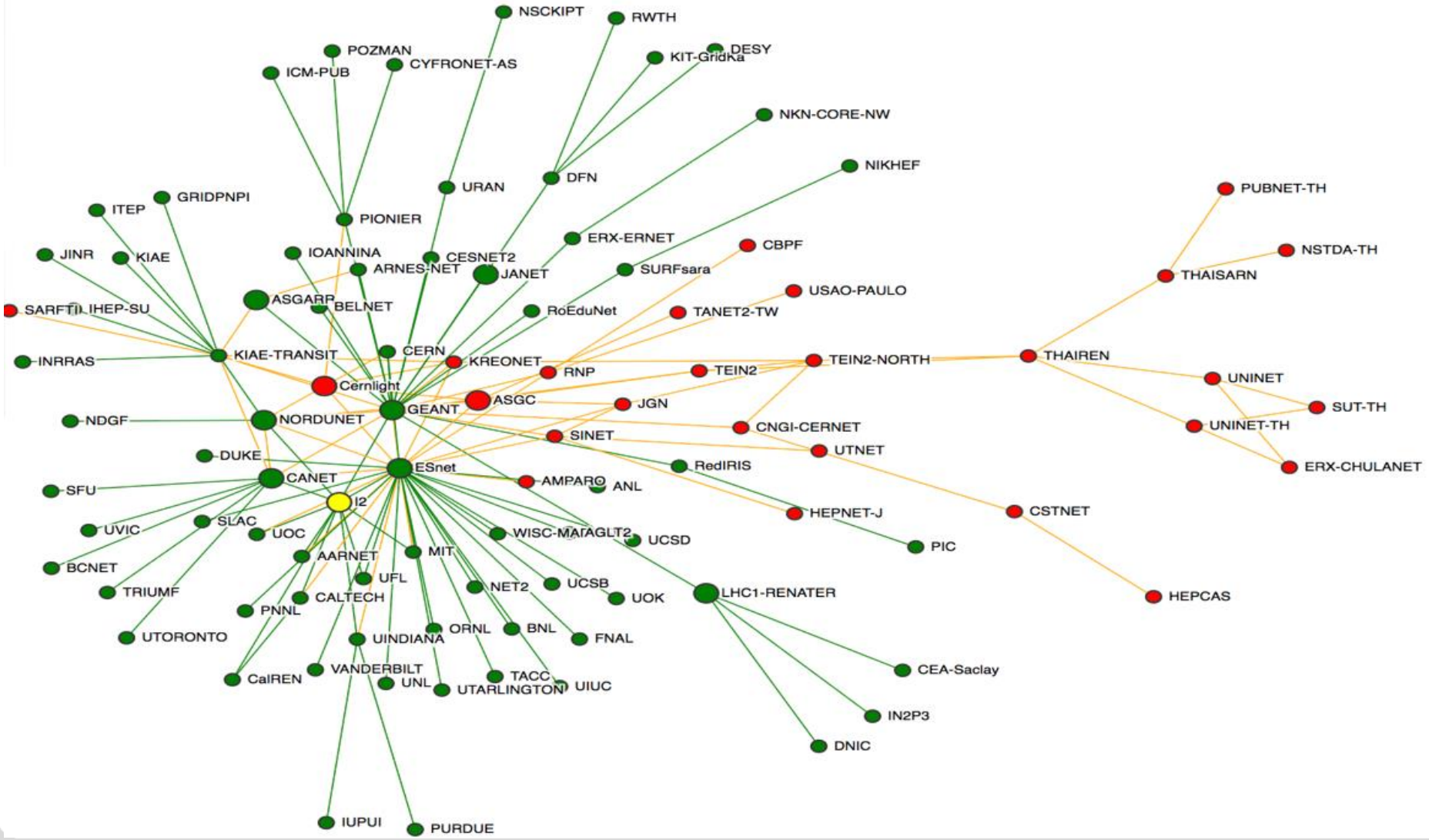
Worldmap of misrouted packets captured during last 30 days

LHCONE regional diffrent routing tables

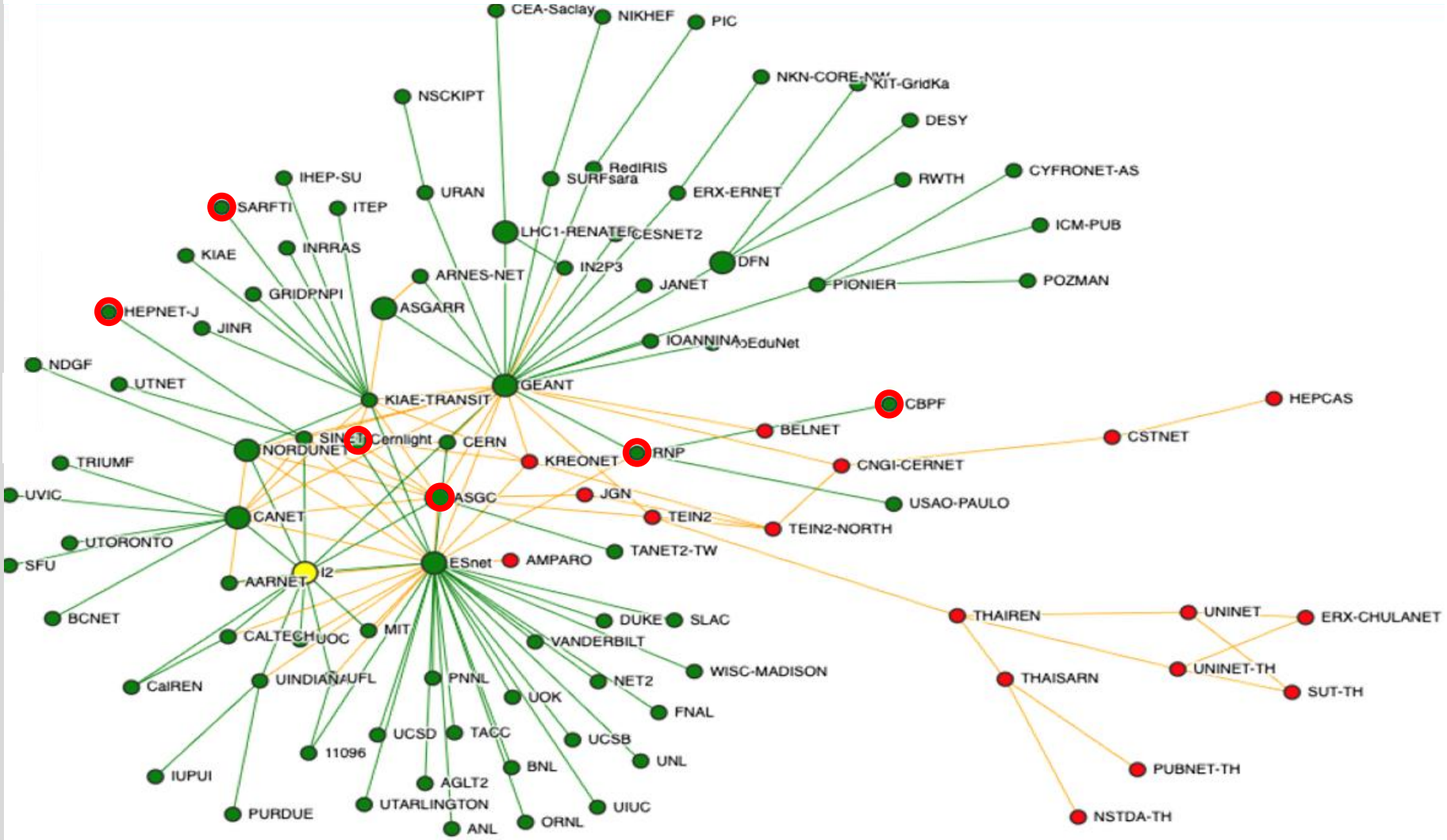


Routing-tables of collected and analysed NSPs

Internet2 IPv4 (Oct. 2018)



Internet2 IPv4 (2019)



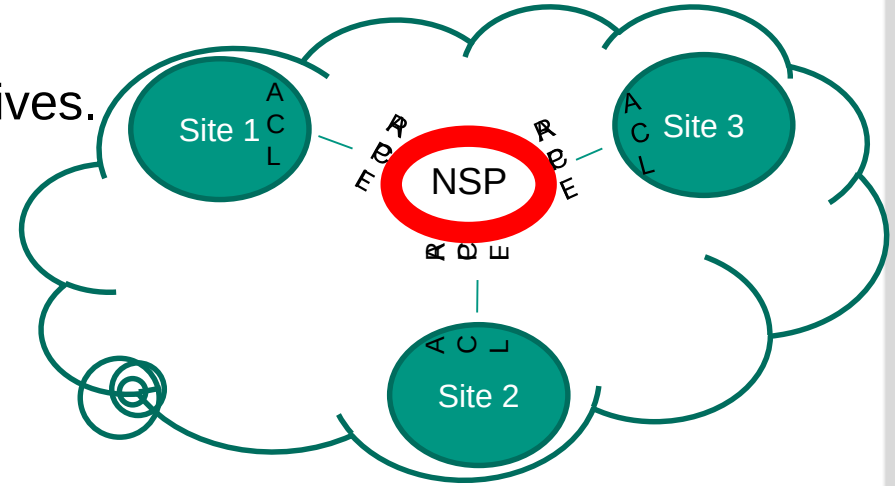
Conclusion / actions

- Number of unroutable LHCONE packets reduced from 2.4million to 1million
  found 1.4 million packtes marked as „false positive“ or could be sorted out
- LHCONE ingress filtering/control has improved dramatically since measurement began in Q1 2018!

  But still **Detection** will be beneficial:
  - Regularly scheduled monitoring?
  - Periodic NSP self run audits?

- Growth in Asia has likely contributed to a small number of exceptions.
  - Work towards **Prevention**
    - NSPs initiative at Edge Site filter configuration
      - RPF → too strict? → Rather opt for ACLs
      - Templated policy & filter configuration

- Routing table inconsistency may also be a source of false positives.



Questions

Suggestions

Discussion

Backup Slids

Recent presentations

- **unroutable LHCONE packets:**
  - **LHCOPN/ONE Meeting March 06, 2018**
    https://indico.cern.ch/event/681168/contributions/2848474/attachments/1611723/2559528/LHCONE-Filter-Policy-Practice.pdf
  - **LHCOPN/ONE Meeting Oct. 30, 2018**
    https://indico.cern.ch/event/725706/contributions/3120030/attachments/1743507/2821722/LHCONE-MTU-recommendation.pdf
  - **LHCOPN/ONE Meeting June 04, 2019**
    https://indico.cern.ch/event/772031/contributions/3360612/attachments/1855532/3047503/LHCONE_Edge_Filtering_Policy_and_Practice_Umea_1.pdf

- **regional LHCONE routing table diffrences:**
  - **LHCOPN/ONE Meeting Oct. 30, 2018**
    https://indico.cern.ch/event/725706/contributions/3149436/attachments/1744301/2823417/LHCone_routing_digging.pdf
  - **LHCOPN/ONE Meeting June 04, 2019**
    https://indico.cern.ch/event/772031/contributions/3428968/attachments/1855890/3048260/LHCone_routing_digging_2019.pdf

Edge Filtering Special Case

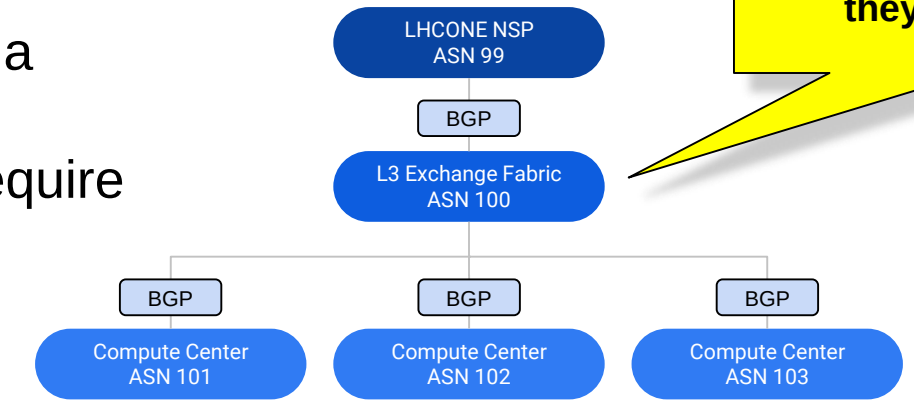
L3 Network Exchange Fabrics

An exchange is like an NSP:

- BGP import filtering
- Packet filtering
- Community based BGP filtering

An exchange is like a site:

- Require the full LHCONE table via a transit NSP
- Packet filters are configured and require maintenance

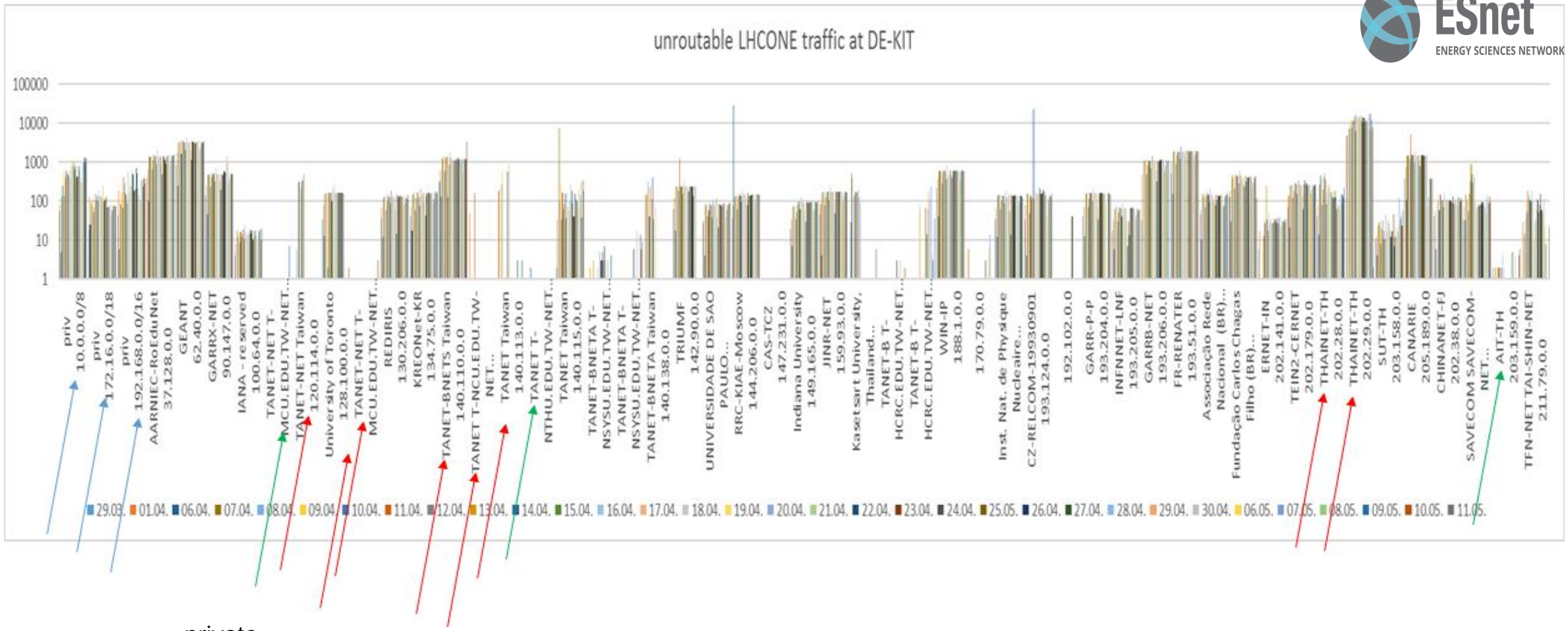


Is an L3 Exchange an edge site or an NSP?
What process defines how they add new sites?

Indiana GigaPOP is a current ESnet example.
SOX is planned to be the second and will connect UFL, FSU and others.

- Will L3 Exchange Fabrics implement and maintain LHCONE specific services?
- Should there be an LHCONE defined role for these network organizations?
- Are they permitted to attach new sites?

All counted packets (44 Sites)



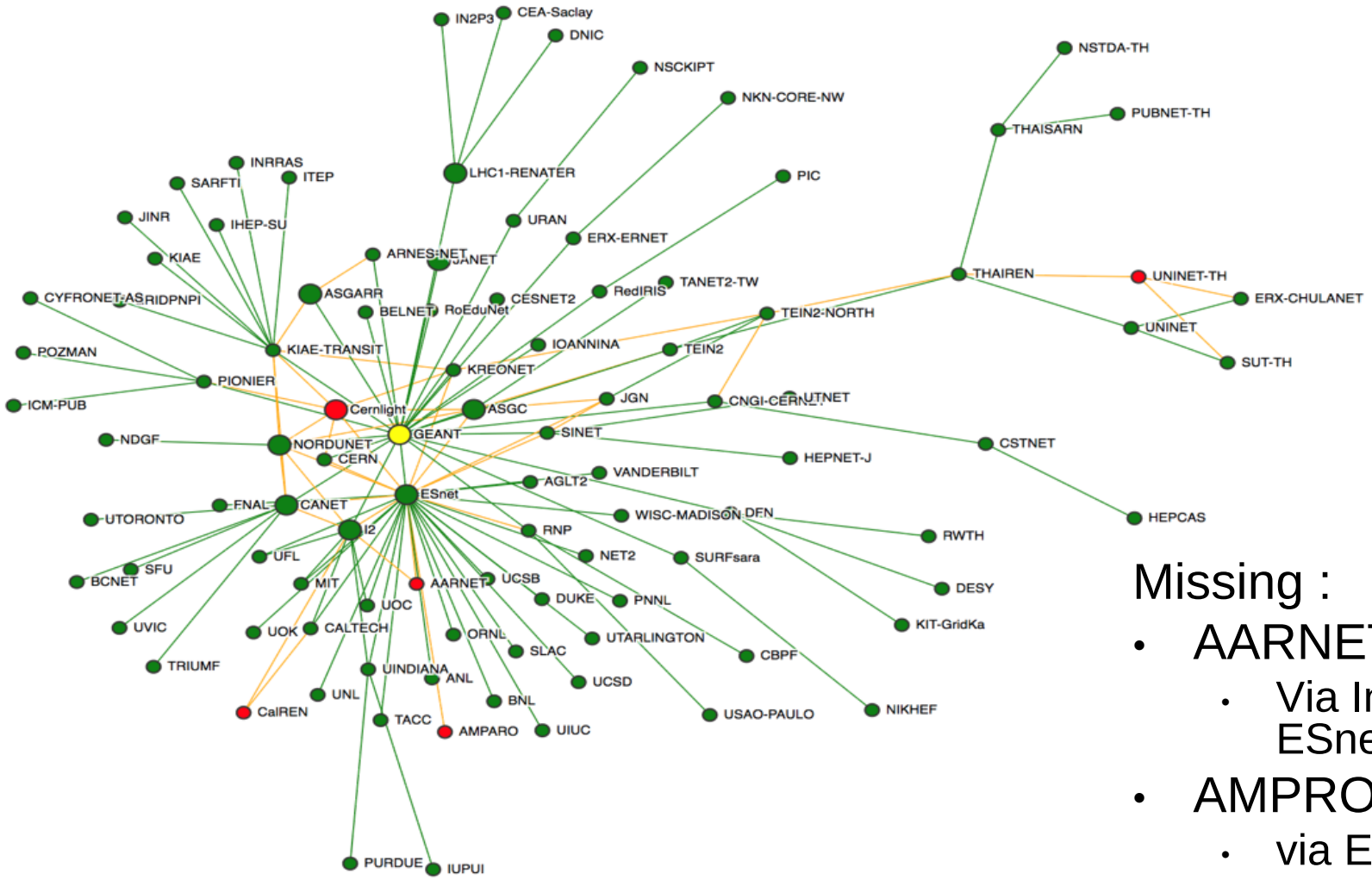
private

grouping different CIDR of one organization

removing sites with less than 1000 unroutable packets over four weeks

Routing tabel diffrenz e.g: Géant (IPv6)

No change between Oct. 2018 and June 2019



Missing :

- AARNET
  - Via Internet2 ESnet
- AMPRO
  - via ESnet