LHCONE status report

2nd Asia Tier Centre Forum
Nakhon Ratchasima, Thailand

1st December 2016

edoardo.martelli@cern.ch

Content

- LHCOPN update
- LHCONE update
- LHCONE in Asia
- Future developments

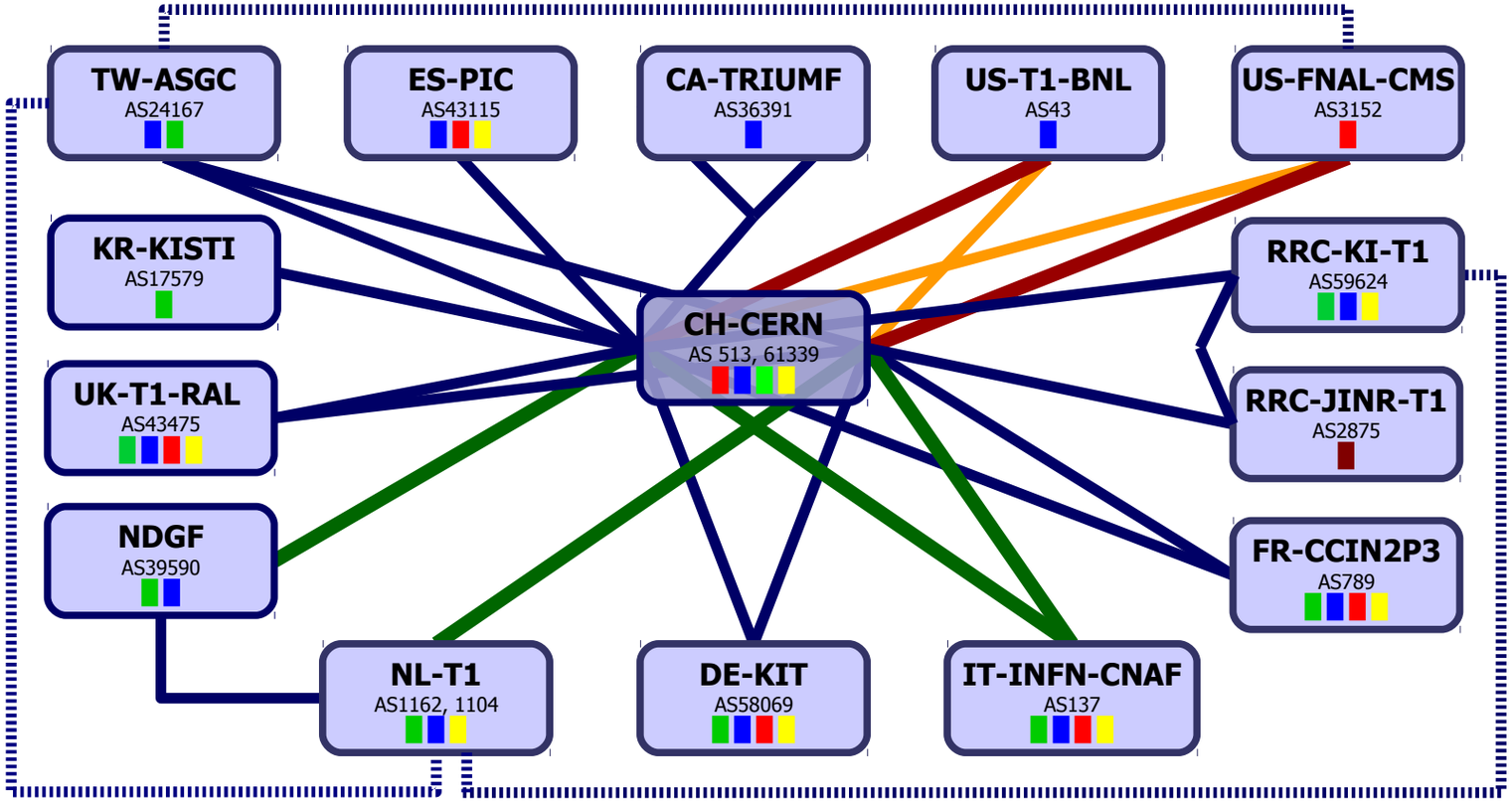
LHCOPN update

LHCOPN

Private network connecting Tier0 and Tier1s

- Reserved to LHC data transfers and analysis
- Single and bundled long distance 10G and 100G Ethernet links
- Star topology
- BGP routing: communities for traffic engineering, load balancing.
- Security: only declared IP prefixes can exchange traffic.

LHCOPN


TW-ASGC
AS24167
ES-PIC
AS43115
CA-TRIUMF
AS36391
US-T1-BNL
AS43
US-FNAL-CMS
AS3152
KR-KISTI
AS17579
CH-CERN
AS 513, 61339
RRC-KI-T1
AS59624
UK-T1-RAL
AS43475
RRC-JINR-T1
AS2875
NDGF
AS39590
FR-CCIN2P3
AS789
NL-T1
AS1162, 1104
DE-KIT
AS58069
IT-INFN-CNAF
AS137
T0-T1 and T1-T1 traffic
T1-T1 traffic only
= Alice = Atlas = CMS = LHCb
edoardo.martelli@cern.ch 20161010
10Gbps
20Gbps
40Gbps
100Gbps

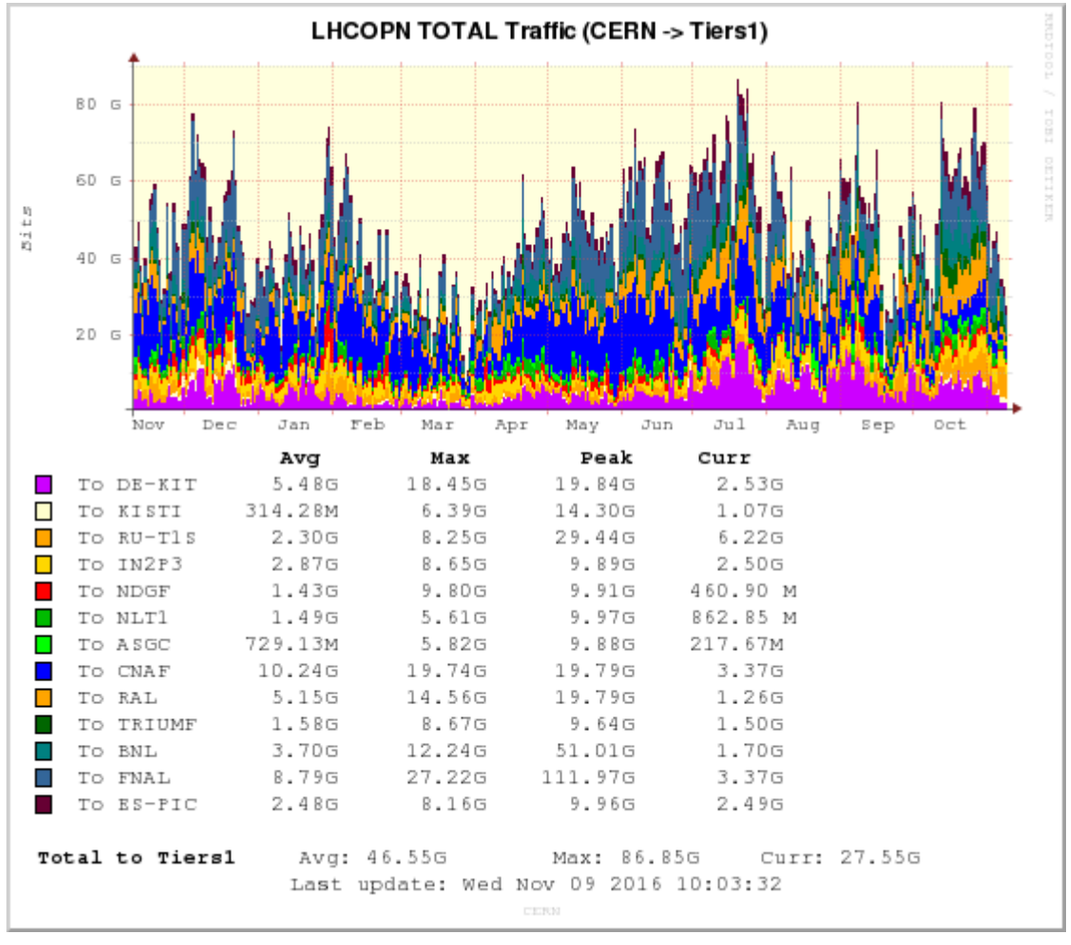
Latest developments

Traffic volume constantly growing
- already increased of 100% from the beginning of Run2

5 Tier1s have doubled link capacity in the last year:
- NL-T1 (2x10G)
- FR-IN2P3 (2x10G)
- NDGF (2x10G)
- IT-INFN-GARR (4x10G)
- UK-T1-RAL (2x10G, load balancing on existing backup link)

Rising IPv6 adoption
- 9 Tier1s and the Tier0 peering over IPv6
- dual-stack perfSONAR installed in all of them

Last year traffic statistics



Plans

Complete IPv6 deployment:
- connect all Tier1s
- make all LHCOPN perfSONAR probes dual-stack
- use IPv6 for production data transfers

Upgrade Tier0-Tier1 links when necessary
- move to 100G when cost effective

LHCONE status

LHCONE concept

- Serving any LHC sites according to their needs and allowing them to grow
- Sharing the cost and use of expensive resources
- A collaborative effort among Research & Education Network Providers
- Traffic separation: no clash with other data transfer, resource allocated for and funded by the HEP community
- Trusted peers: common security policies

LHCONE services

L3VPN (VRF): routed Virtual Private Network - *operational*

P2P: dedicated, bandwidth guaranteed, point-to-point links - *in development*

perfSONAR: monitoring infrastructure - *operational*

LHCONE L3VPN service

Layer3 (routed) Virtual Private Network

Dedicated worldwide backbone connecting **Tier1s, T2s and T3s** at high bandwidth

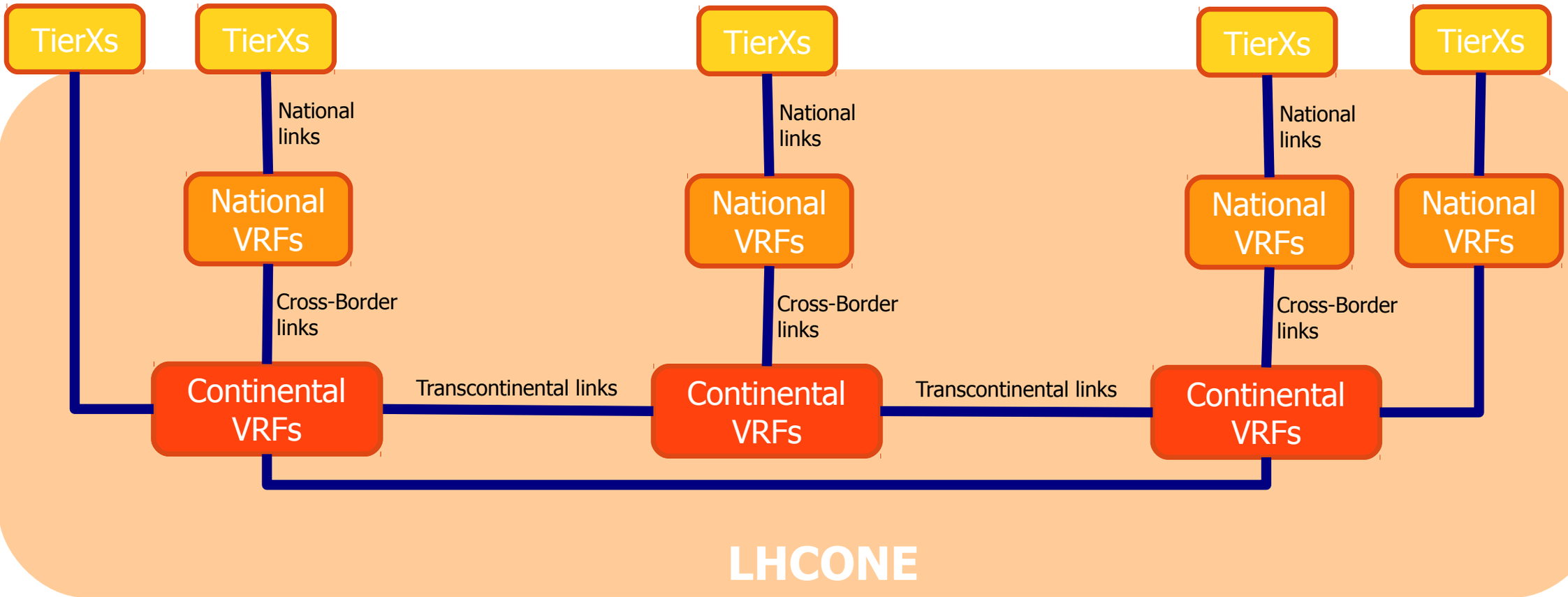
Bandwidth dedicated to LHC data transfers, no contention with other research projects

Trusted traffic that can bypass slow perimeter firewalls

LHCONE L3VPN architecture

- TierX sites connected to National-VRFs or Continental-VRFs
- National-VRFs interconnected via Continental-VRFs
- Continental-VRFs interconnected by trans-continental/trans-oceanic links

Acronyms: **VRF** = Virtual Routing Forwarding (virtual routing instance)



L3VPN: status

Over 18 national and international Research Networks

Interconnections at Open Exchange Points including NetherLight, StarLight, MANLAN, WIX, CERNlight, Hong Kong and others

Trans-Atlantic connectivity provided by ESnet, GEANT, Internet2, NORDUnet and SURFnet

Trans-pacific connectivity provided by ASGCnet, KREOnet, SINET

~74 end sites connected to LHCONE:
- 14 Tier1s
- 60 Tier2s

L3VPN: latest developments

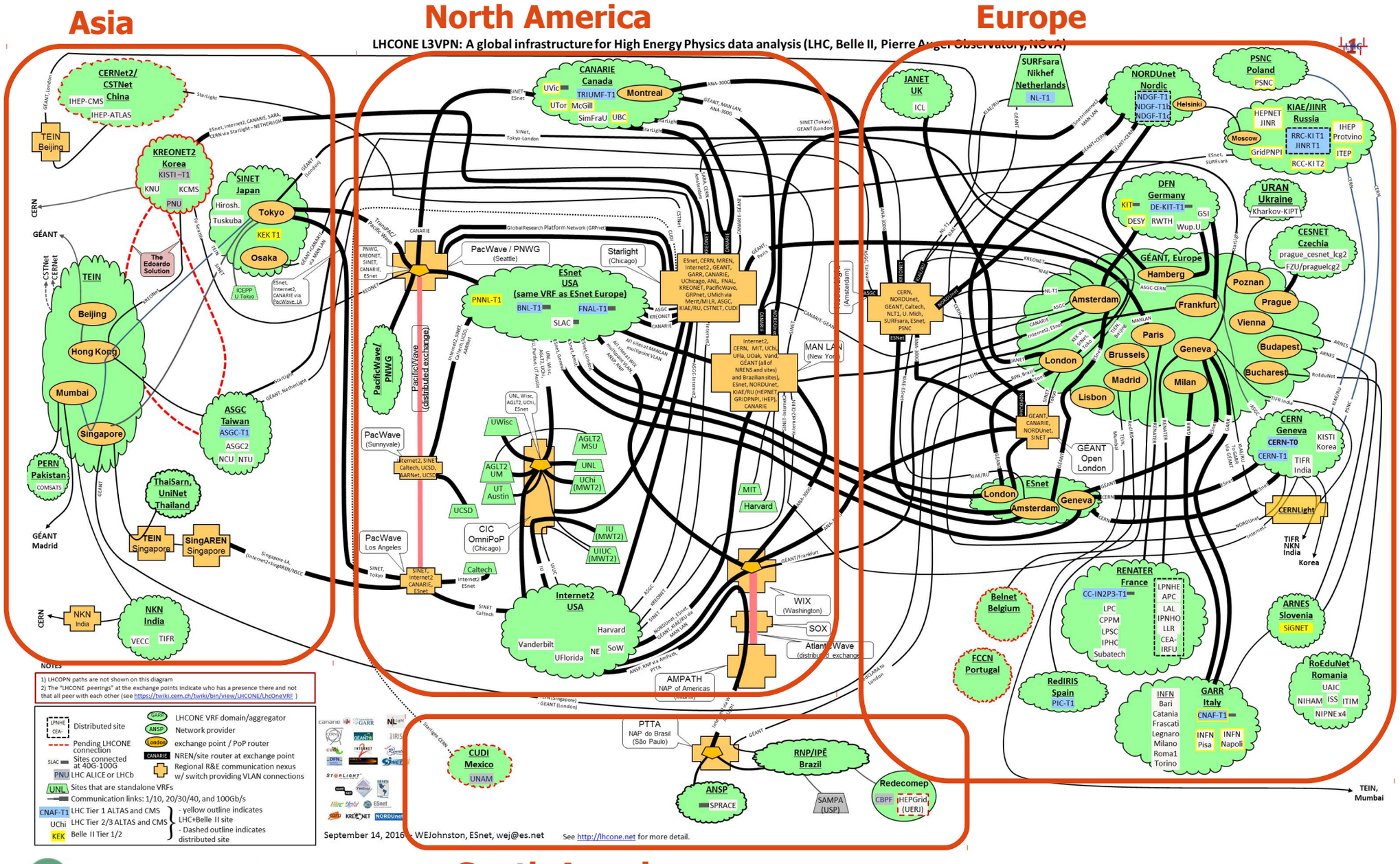
The LHCONE network is expanding
- Asia: SINET (JP), ASGC (TW), KISTI (KR), TEIN are connected
- Europe: Russia, Poland, Ukraine and Belgium have recently connected; Portugal is willing to join
- North America: bigger sites connecting at 100Gbps
- South America: Brazil is now a stable partner, Chile is interested to join
- Africa: first contacts with South Africa

Traffic within LHCONE is steadily growing
- GÉANT has seen peaks of over 100Gbps
- Growth of over 65% from Q2 2015 to Q2 2016

Some NRENs and sites need to upgrade their connection
- GÉANT is already working with the NRENs for the upgrades

Expected to see further traffic volume increases after the upgrades

L3VPN Current topology



LHCONE Point-to-Point service

On demand point-to-point (P2P) links over a multi-domain network

Provides P2P links between any pair of TierX. The P2P links have guaranteed bandwidth (protected from any other traffic)

Accessible and configurable via software API

Still in development phase, not yet a service

Point-to-Point service status

Demonstrated routing over dynamic circuits using BGP Router Servers

Demonstrated dynamic circuit provisioning using AutoGOLE fabric and NSI *[Network Service Interface protocol]*

Demonstrated high speed data transfer (100G from a single disk server) using DTNs *[high performance Data Transfer Node]*

Looking for contact people in the LHC Experiments to progress with application interfaces

LHCONE perfSONAR service

- LHCONE Network monitoring infrastructure

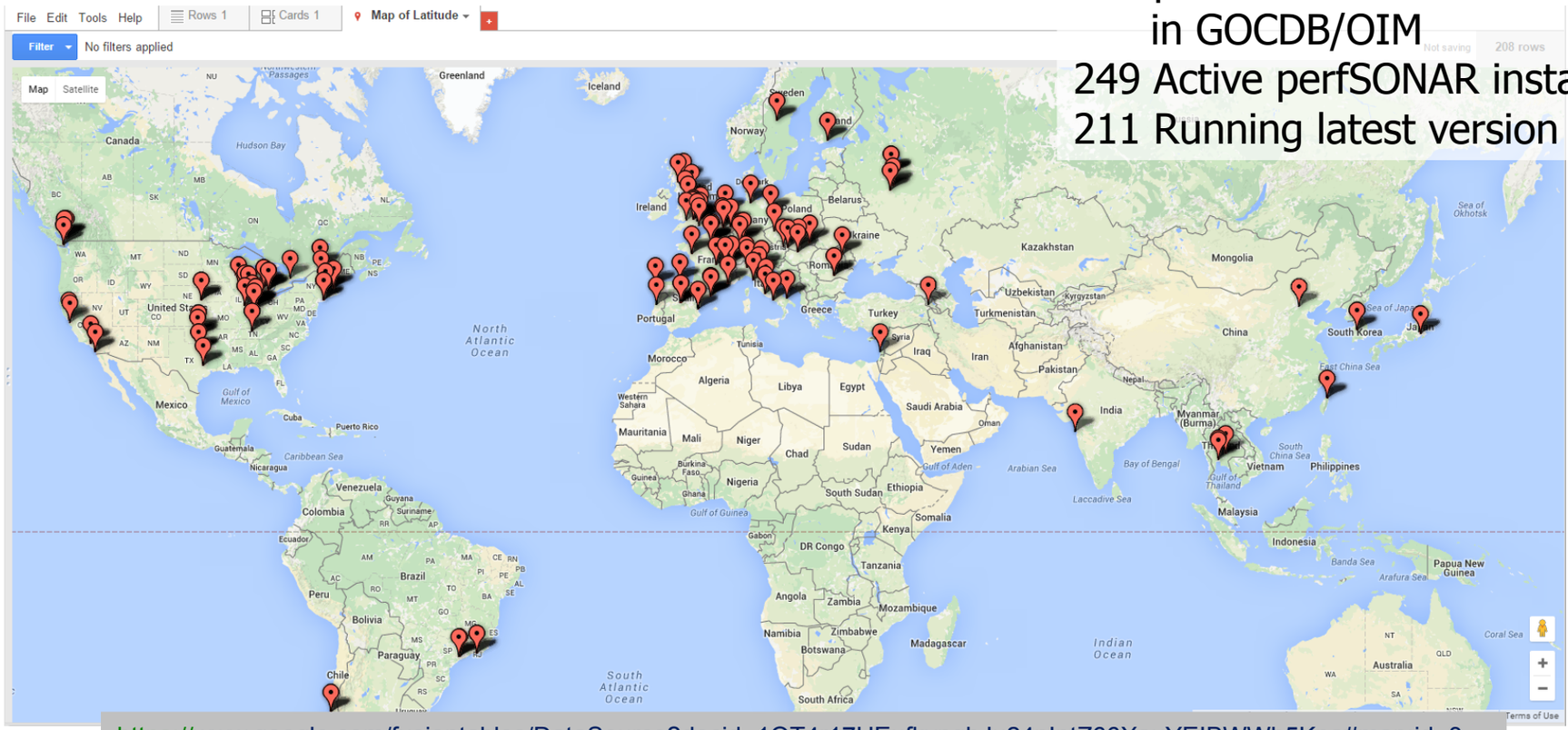
- perfSONAR probes installed at the VRFs interconnecting points and at the TierX sites

- Accessible to any TierX to check network healthiness

LHCONE perfSONAR: status

http://grid-monitoring.cern.ch/perfsonar_report.txt for stats

278 perfSONAR instances registered in GOCDB/OIM
249 Active perfSONAR instances
211 Running latest version (3.5+)



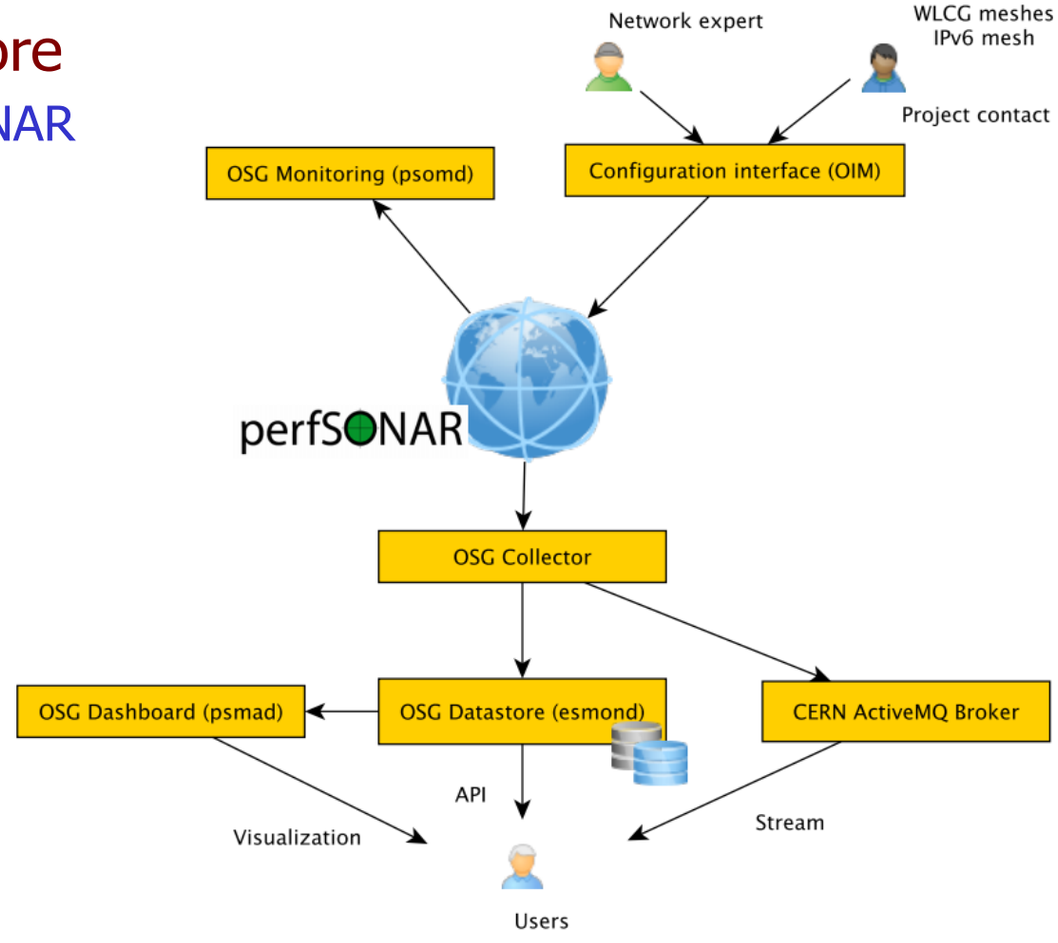
https://www.google.com/fusiontables/DataSource?docid=1QT4r17HEufkvnqhJu24nIptZ66XauYEIBWWh5Kpa#map:id=3

- Initial deployment coordinated by WLCG perfSONAR TF
- Commissioning of the network followed by WLCG Network and Transfer Metrics WG

Slide credit: Shawn McKee, University of Michigan

perfSONAR: gathering and storing metrics

- OSG is providing network metric data for its members and WLCG via the Network Datastore
  - The data is gathered from all WLCG/OSG perfSONAR instances
  - Stored indefinitely on OSG hardware
  - Data available via Esmond API
  - In production since September 14th 2015
- The primary use-cases
  - Network problem identification and localization
  - Network-related decision support
  - Network baseline: set expectations and identify weak points for upgrading



Slide credit: Shawn McKee, University of Michigan

perfSONAR update

Mesh have been reconfigured/optimized. There are now experiment specific meshes: ATLAS, CMS and LHCb which allow daily bandwidth tests

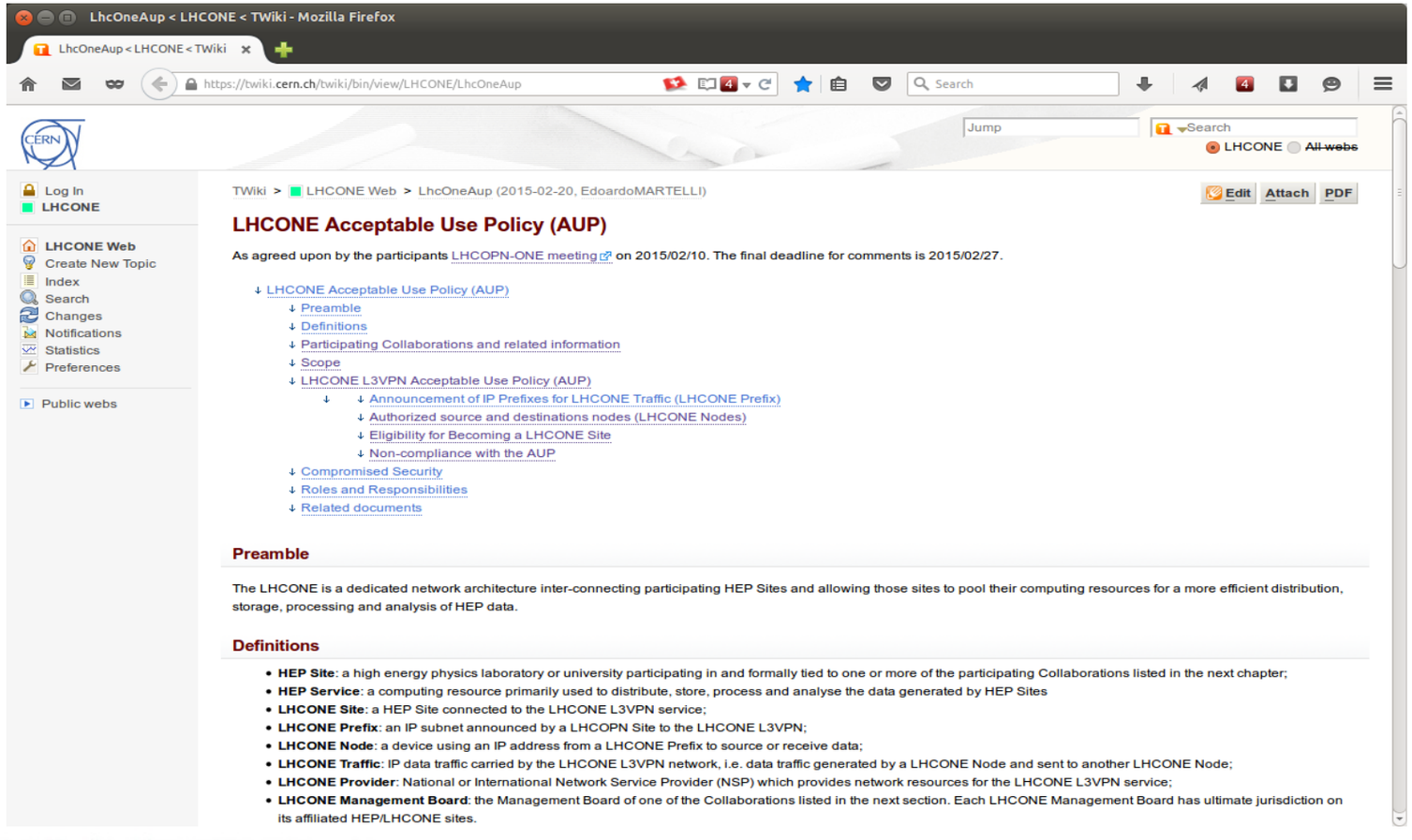
All the LHCOPN and LHCONE data is available in ElasticSearch via OSG

Generally slight improvements noted in the network quality in LHCOPN and LHCONE

New release of perfSONAR coming soon (v4.0). Focus is on control and stability. New MaDDash is part of this release.

LHCONE Acceptable Use Policy

The LHCONE AUP has been defined to regulate the utilization of the L3VPN service (https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneAup)



AUP: Procedure to join LHCONE

The WLCG Management Board has recommended that admissions to LHCONE are decided by the LHCONE community

A procedure to evaluate the admissions has been agreed:

The LHCONE community accepts or rejects based on the impact on the LHCONE. Among criteria to be used in the evaluation:

a) the collaboration must be related to Particle Physics
b) a major fraction of the sites and collaboration's resources (CPUs and storage) must be already connected to LHCONE
c) commitment to meet the technical and security requirements listed in the AUP
d) the bandwidth demand shouldn't have a significant impact on existing LHCONE data transfers
e) commitment to participating and contributing to LHCONE meetings

Current version: https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneAup#Procedure_to_include_additional

Open to other HEP collaborations

The L3VPN is now used also by:

- **Belle II experiment**

- **NOvA neutrino experiment**

- **Pierre Auger observatory**

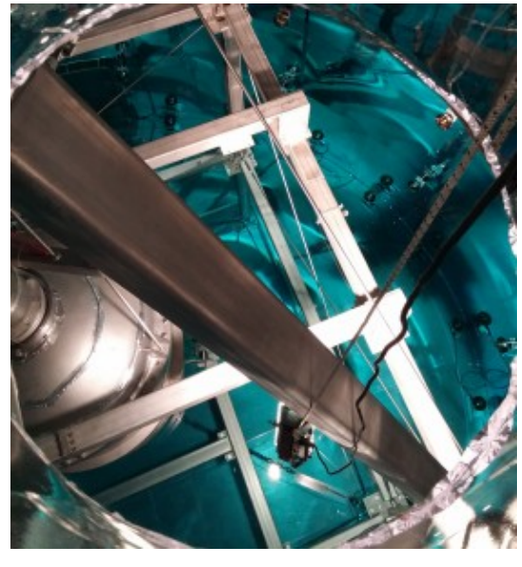
Last: XENON dark matter project

The XENON dark matter project is the last collaboration that has joined LHCONE

XENON1T: Dark matter search experiment at LNGS (Gran Sasso National Laboratory)
http://xenon.astro.columbia.edu/index3.html

Commissioning started in Spring 2016. Now taking data

21 sites involved. Several of them already in LHCONE



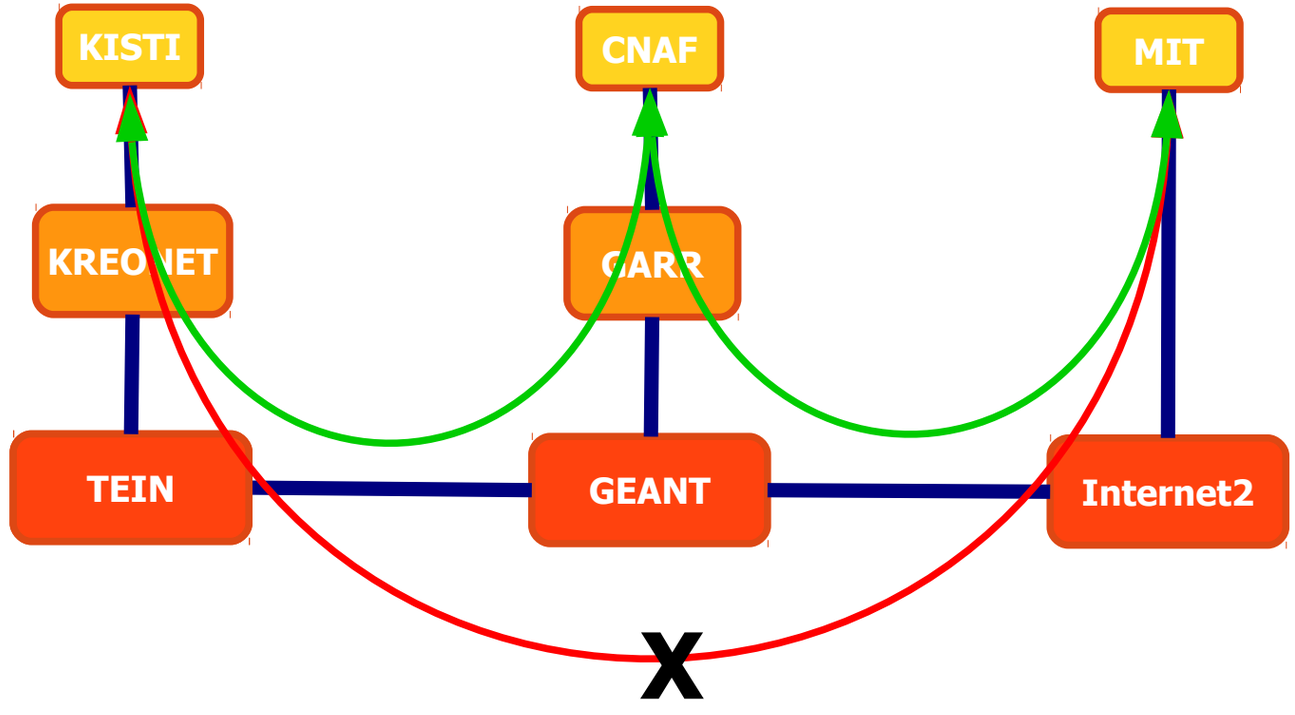
LHCONE in Asia

Challenge 1: keep delay (RTT) low

Traffic between Asian countries must be routed within Asia (avoid transit to another continent or around the world)

Challenge 2: achieve full reachability

Guarantee reachability to all LHCONE sites



Possible solutions

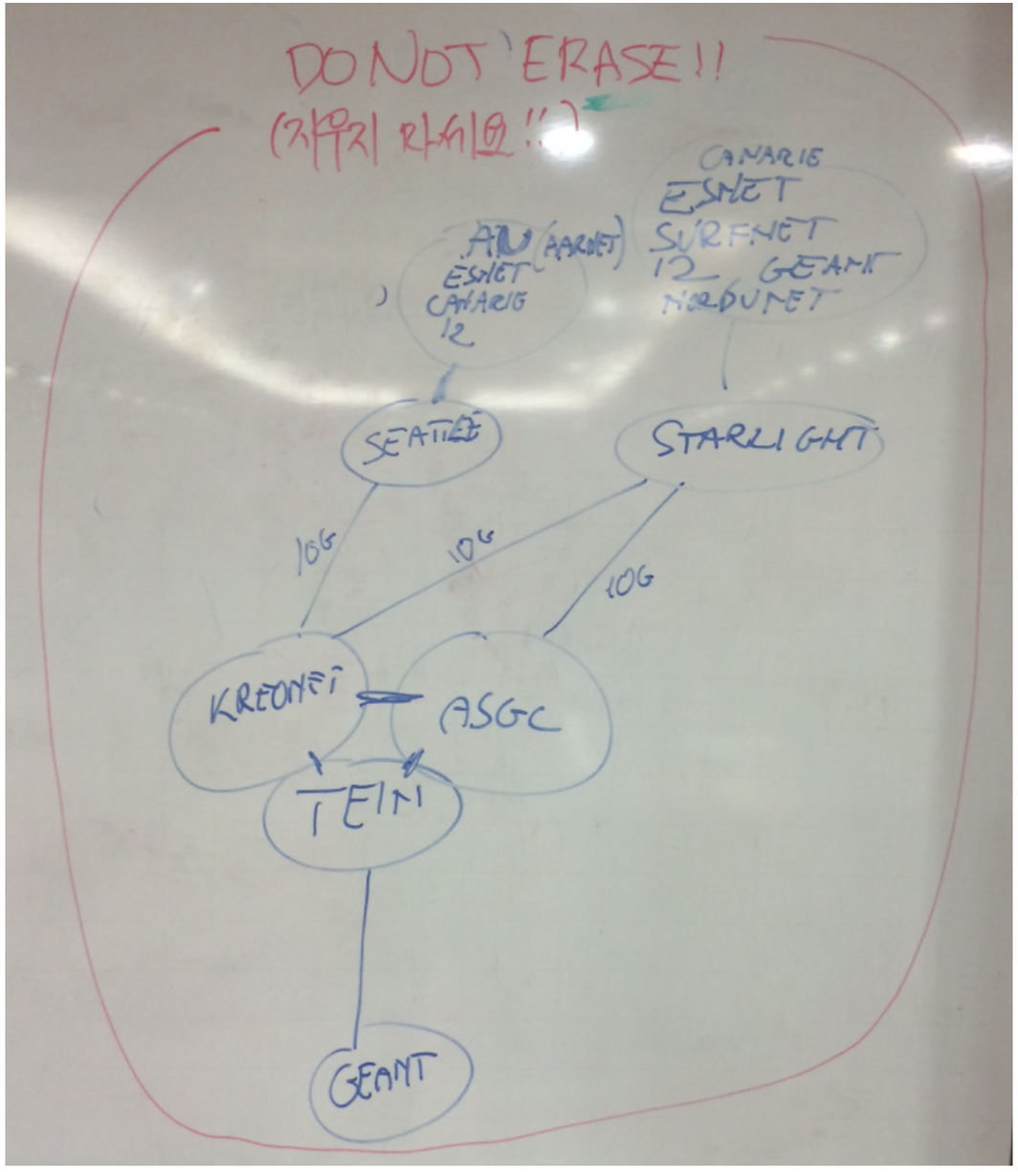
Low delay:

- Asian countries mesh networks together
- Asian countries get interconnected by common regional backbone

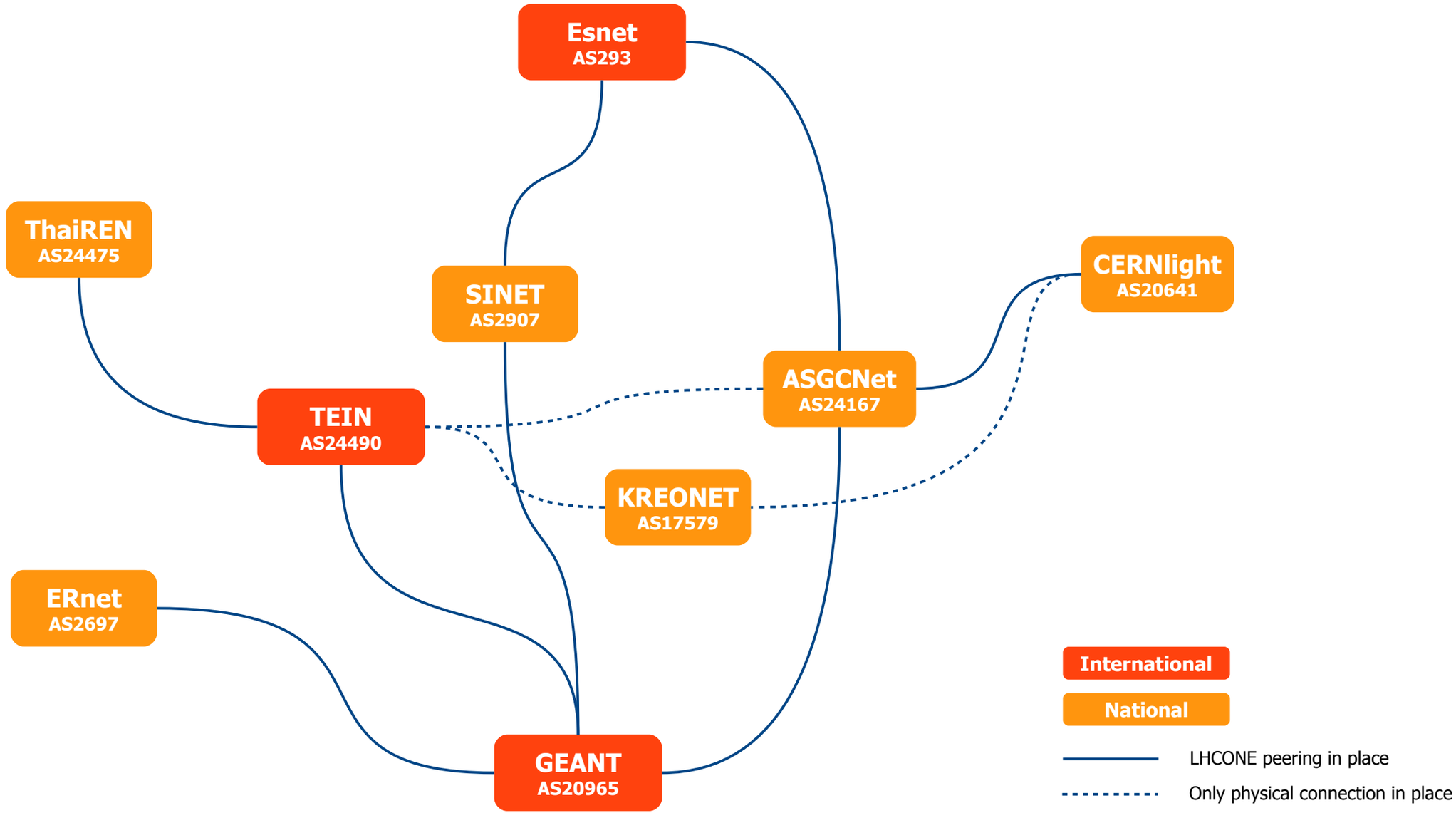
Full reachability:

- full mesh
- small VRFs get full transit from major operators

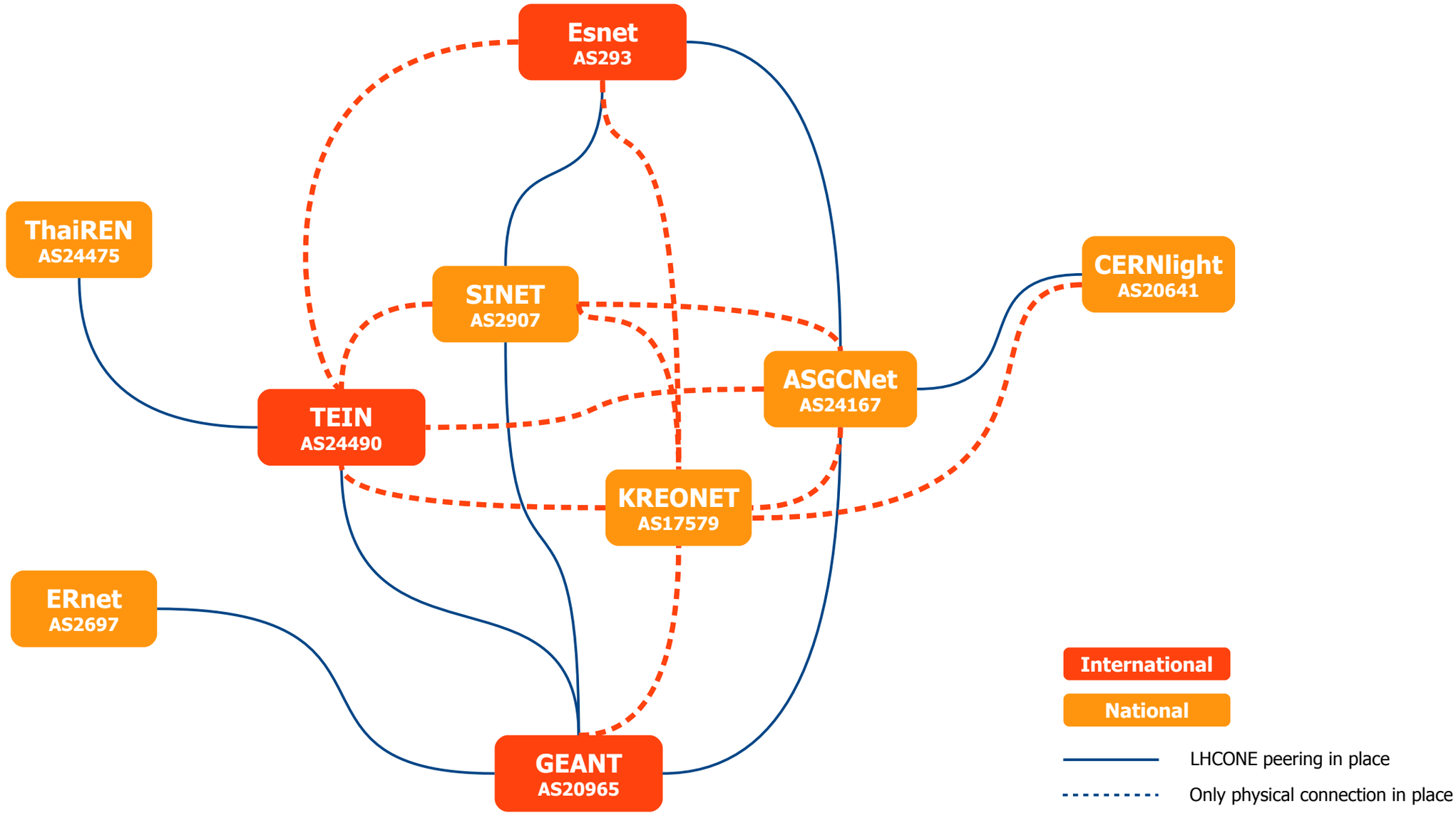
Last ATCF outcome



Status of LHCONE VRFs in Asia



Missing connections



Missing

TEIN VRF is not visible outside the GEANT domain (no transit through GEANT VRF)

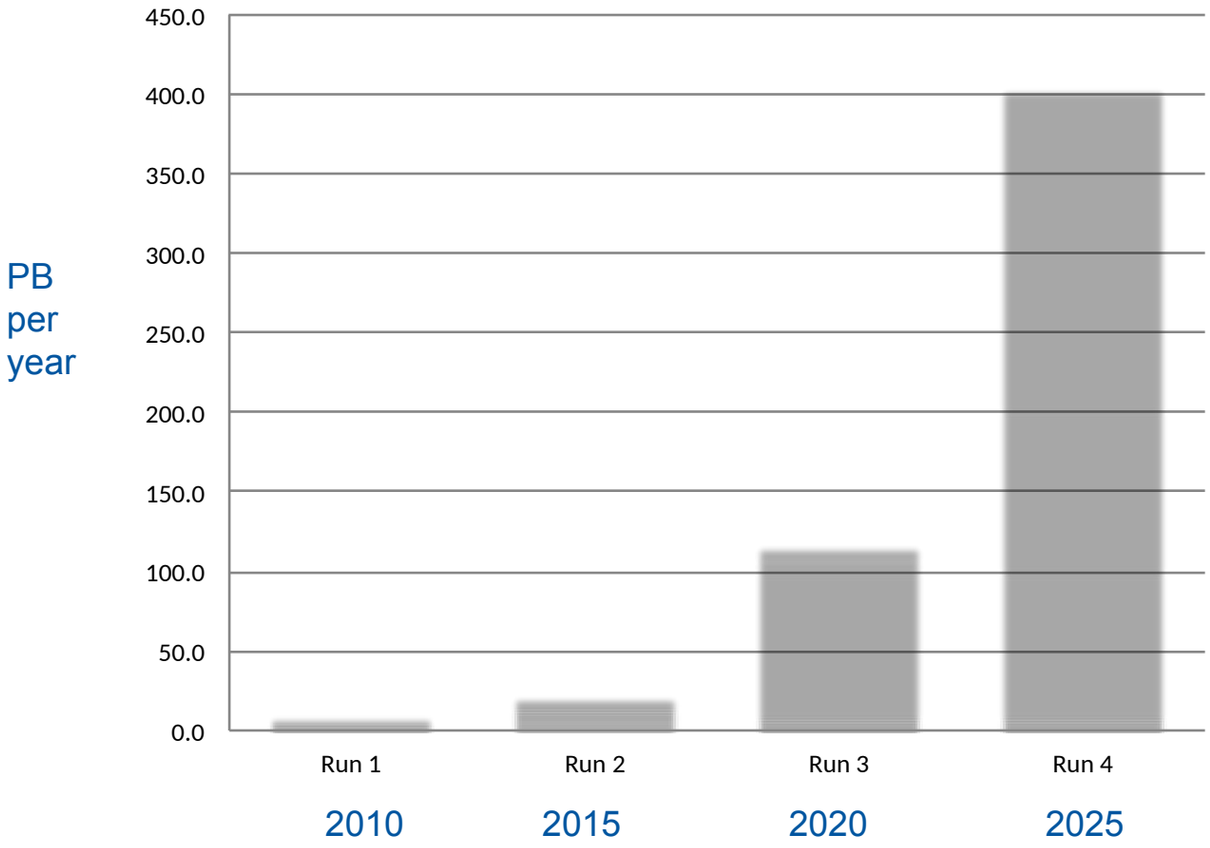
KREONET VRF not peering with anything yet

ASGCnet VRF not connected to TEIN, KREONET, SINET yet. Planning to connect to APAN-JP

SINET VRF not connected to any Asian VRF

Future developments

Networks have to follow LHC data growth



CMS
ATLAS
ALICE
LHCb

Expecting to record 400PB/year in Run4

Computing needs expected to be around 50x current levels, if budget available

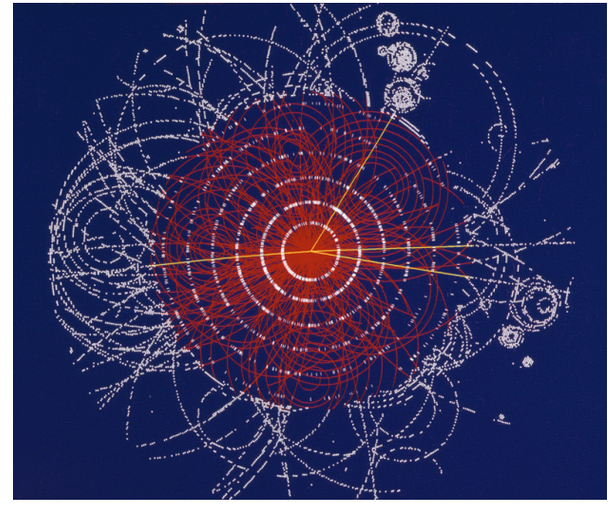
Networks must grow accordingly

Use of Commercial Cloud Services

CERN and other HEP institutes are evaluating the use of Commercial Cloud Services for HEP computing

Research and Education Networks are designing solutions to better connect Cloud Service Providers to their customers

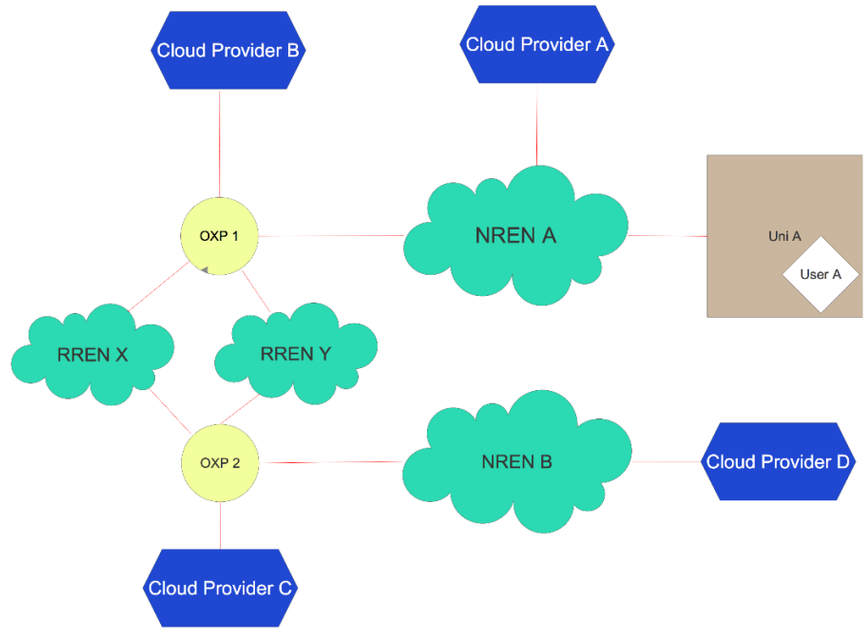
Main issues:
- deliver traffic from cloud datacentres to users in different continents
- avoid or not cloud-to-cloud traffic
- not all the research networks allow commercial traffic



Best practices for CSP connectivity

NORDUnet presented a document which suggest a way of connecting CSPs and their customers by using Open Exchange Points and connectivity provided by the REN

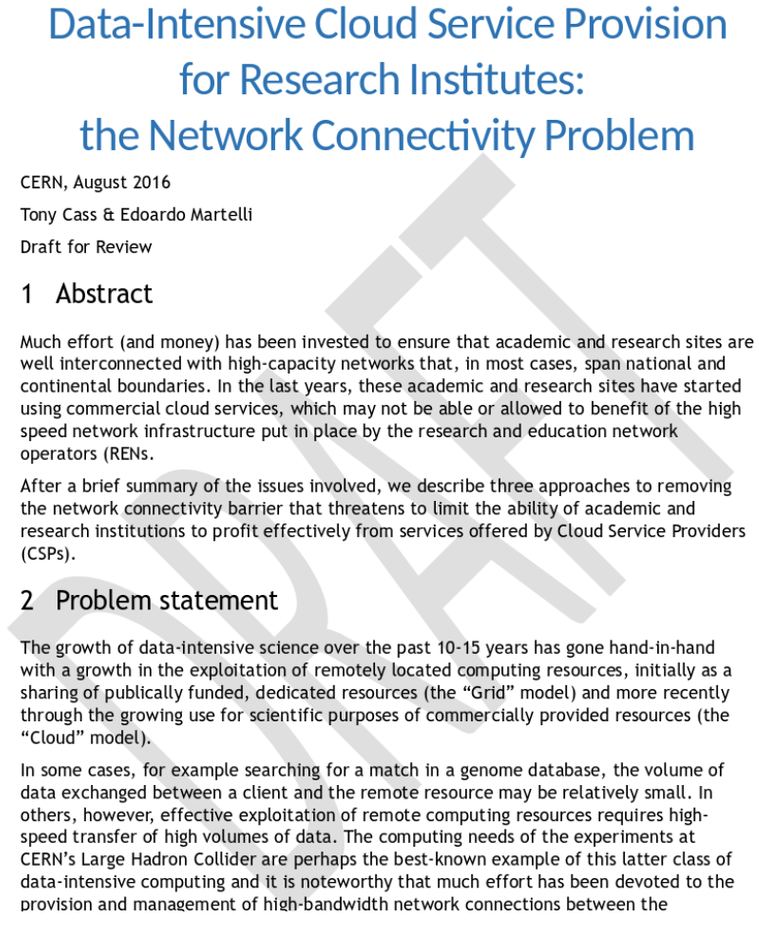
The document aims to simplify the procurement of the connectivity to CSPs



Full presentation:
https://indico.cern.ch/event/527372/contributions/2236895/subcontributions/206270/attachments/1334884/2015982/Cloud_Connectivity_Best_Practice_19-09-2016.pptx
Document:
https://indico.cern.ch/event/527372/contributions/2236895/subcontributions/206270/attachments/1334884/2007495/Best_Practices_Cloud_Providers_and_OXPs_v2.0.pdf

Network connectivity options for CSPs

CERN IT-CS is writing a document which compare the different connectivity options proposed by the RENs



Data-Intensive Cloud Service Provision for Research Institutes: the Network Connectivity Problem

CERN, August 2016

Tony Cass & Edoardo Martelli

Draft for Review

1 Abstract

Much effort (and money) has been invested to ensure that academic and research sites are well interconnected with high-capacity networks that, in most cases, span national and continental boundaries. In the last years, these academic and research sites have started using commercial cloud services, which may not be able or allowed to benefit of the high speed network infrastructure put in place by the research and education network operators (RENs.

After a brief summary of the issues involved, we describe three approaches to removing the network connectivity barrier that threatens to limit the ability of academic and research institutions to profit effectively from services offered by Cloud Service Providers (CSPs).

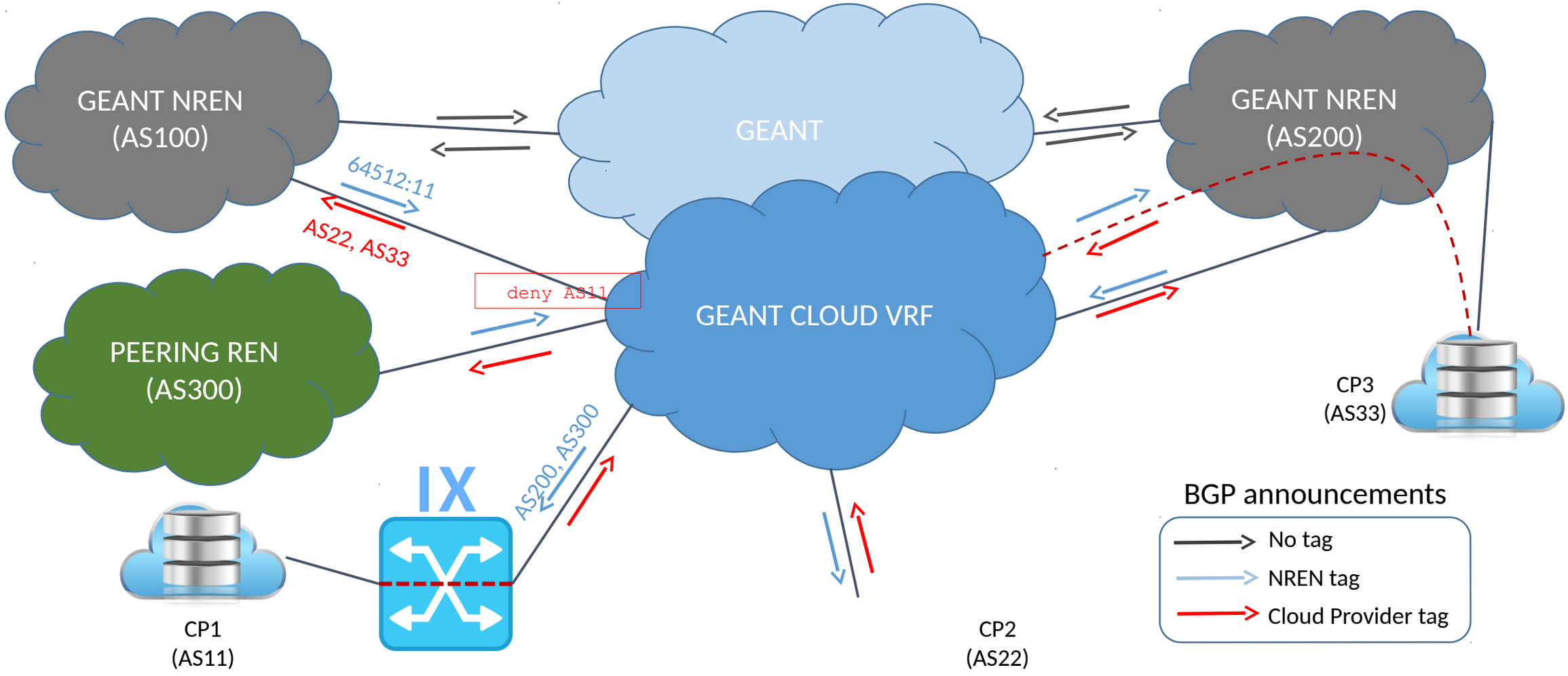
2 Problem statement

The growth of data-intensive science over the past 10-15 years has gone hand-in-hand with a growth in the exploitation of remotely located computing resources, initially as a sharing of publically funded, dedicated resources (the “Grid” model) and more recently through the growing use for scientific purposes of commercially provided resources (the “Cloud” model).

In some cases, for example searching for a match in a genome database, the volume of data exchanged between a client and the remote resource may be relatively small. In others, however, effective exploitation of remote computing resources requires high-speed transfer of high volumes of data. The computing needs of the experiments at CERN’s Large Hadron Collider are perhaps the best-known example of this latter class of data-intensive computing and it is noteworthy that much effort has been devoted to the provision and management of high-bandwidth network connections between the

Document https://indico.cern.ch/event/527372/contributions/2236895/subcontributions/208050/attachments/1338702/2015050/connectivity-options-for-clouds-draft.pdf

GEANT: dedicated VRF

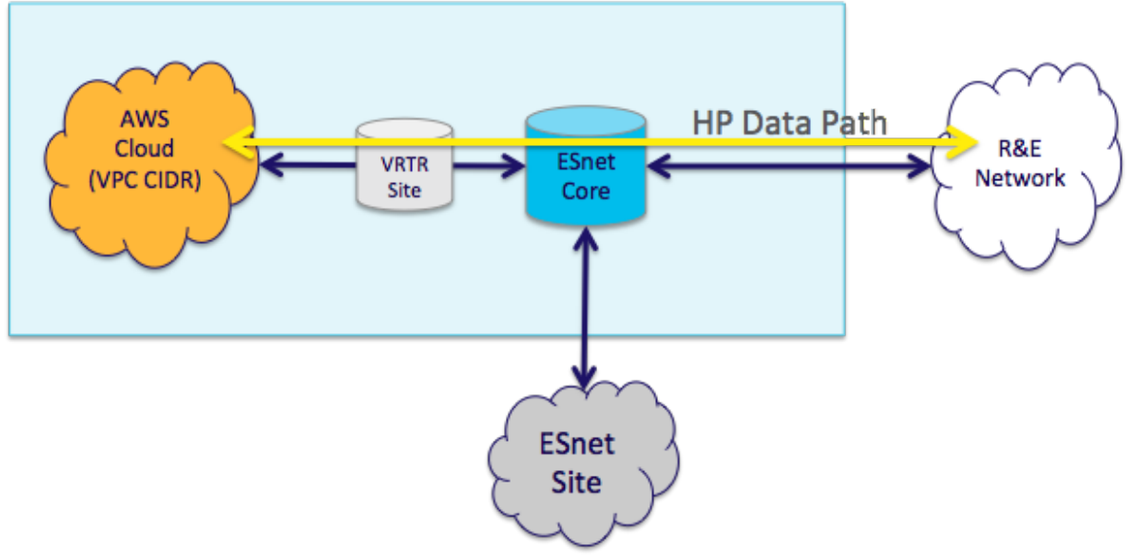


ESnet: on-net VPN termination

Virtual “Site Router” (VRTR) Service

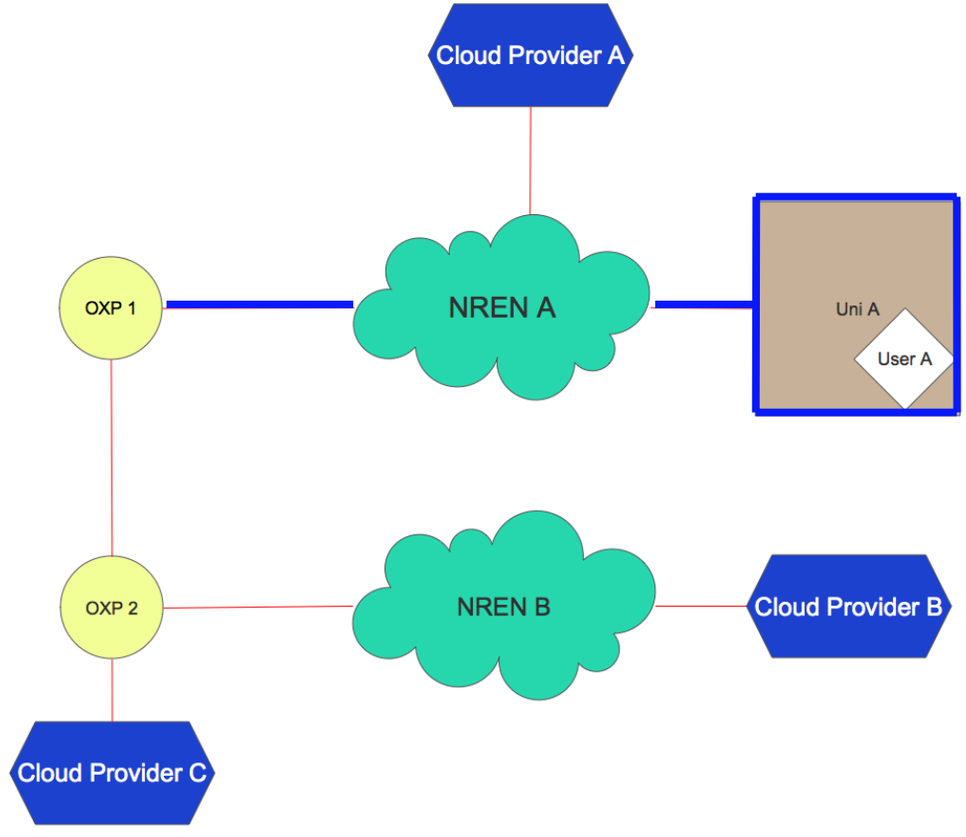
At the edge of the cloud

Virtual Site Router at AWS Exchange Point



Virtual “Site Router” improves path efficiency and takes pressure off of the site local-loop.

NORDUnet: transport to eXchange Points



Conclusion

Summary

LHCOPN: increasing capacity

LHCONE: expanding in capacity and extension

LHCONE in Asia: VRFs implemented but interconnections and transit are still missing

LHC Run4 will produce unprecedented amount of data: networks will have to grow accordingly. Commercial cloud resources are being evaluated

Upcoming meetings

WLCG pre-GDB meeting on networking

Date: 10th of January 2017
Location: CERN, Geneva - Switzerland
https://indico.cern.ch/event/571501/

Next LHCONE-LHCOPN meeting

Date: 4-5 of April 2017 (TBC)
Location: BNL, New York – U.S.
https://indico.cern.ch/event/581520/

Following LHCONE-LHCOPN meeting

Date: October 2017 (TBC)
Location: KEK, Tsukuba – Japan (TBC)

More information on LHCOPN and LHCONE

Latest LHCOPN/ONE meetings:
Taipei, March 2016: https://indico.cern.ch/event/461511/
Helsinki, September 2016: https://indico.cern.ch/event/527372/

Websites:
LHCOPN: https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome
LHCONE: https://twiki.cern.ch/twiki/bin/view/LHCONE/WebHome

Questions?

edoardo.martelli@cern.ch