Networking for WLCG: LHCOPN and LHCONE

ISGC Taipei, 17th March 2016

Edoardo.Martelli@cern.ch

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Summary

- LHC data challange
- Networking for WLCG
  - LHCOPN
  - LHCONE

LHC data challenge

LHC Data Challenge

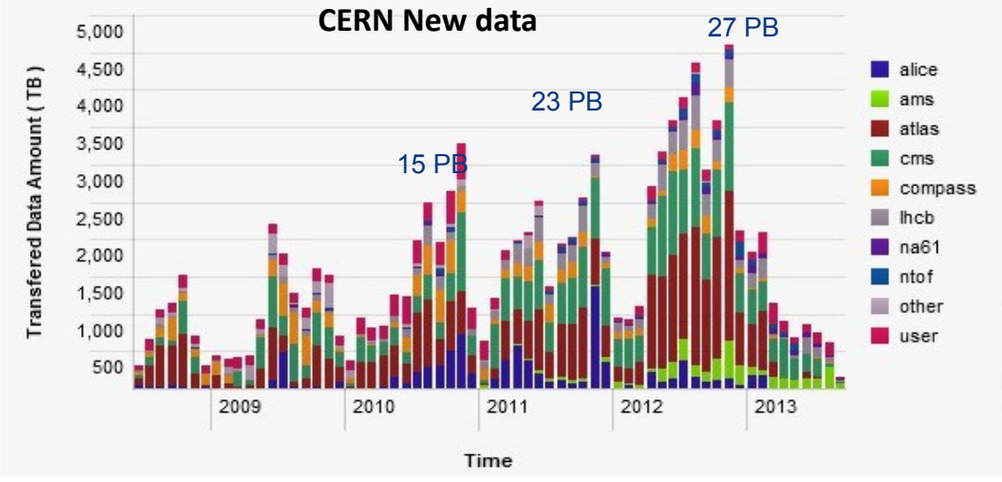
40 million collisions per second

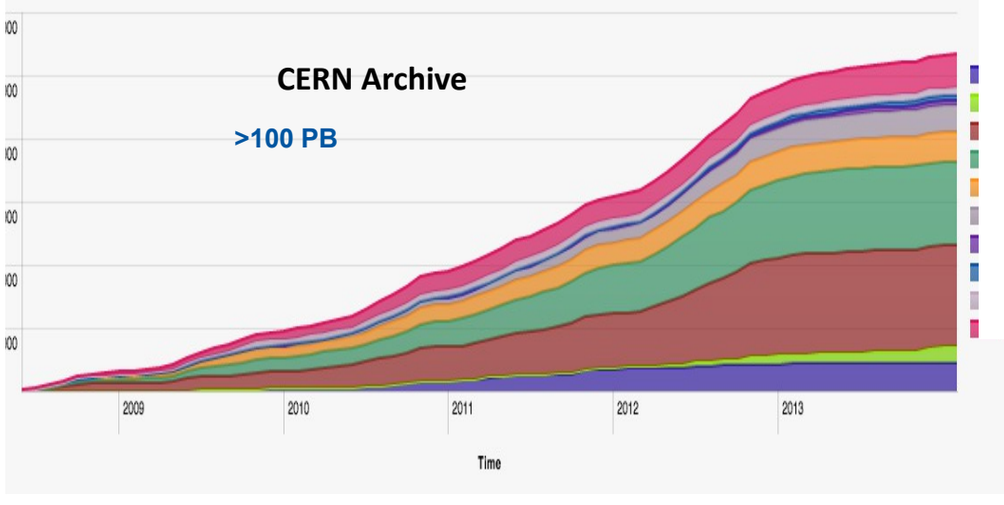
After filtering, only few hundreds of collisions of interest per second

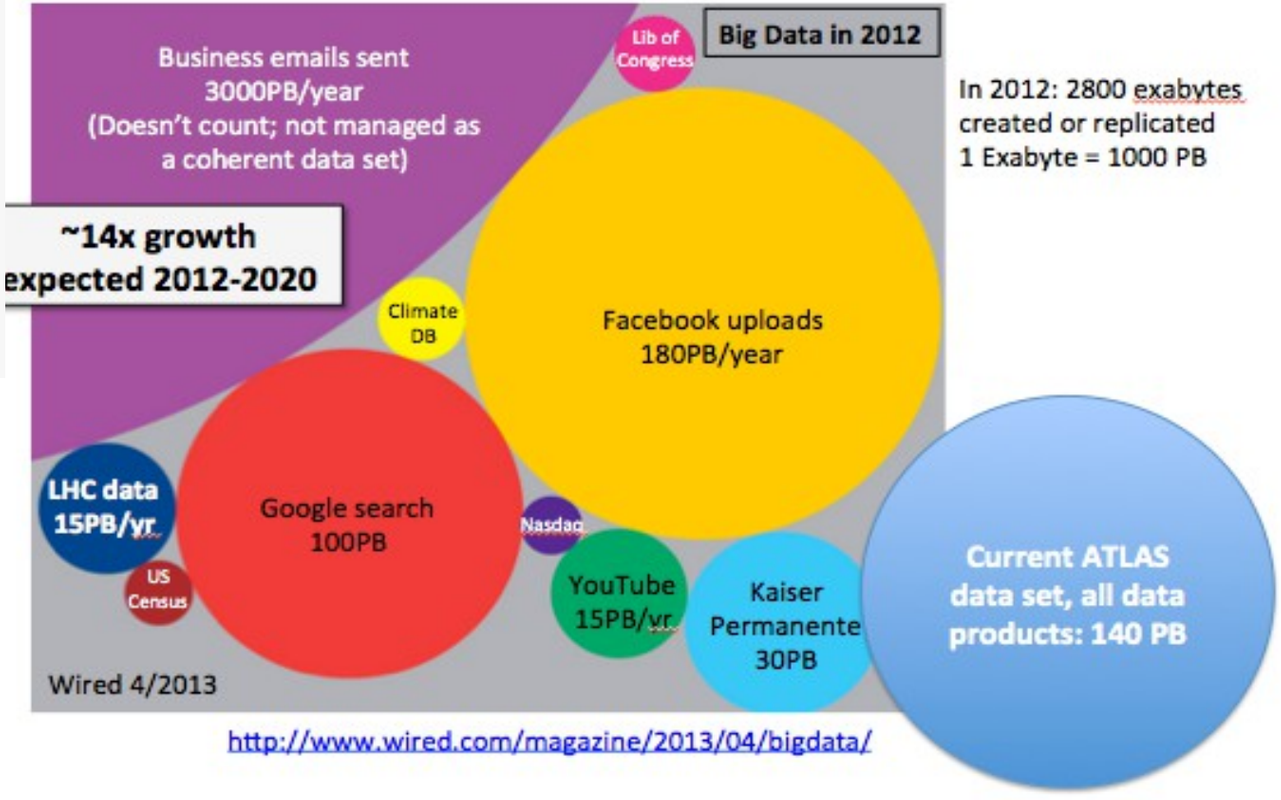
10^{10} collisions recorded each year

=> more than 25 Petabytes/year of data

LHC data production

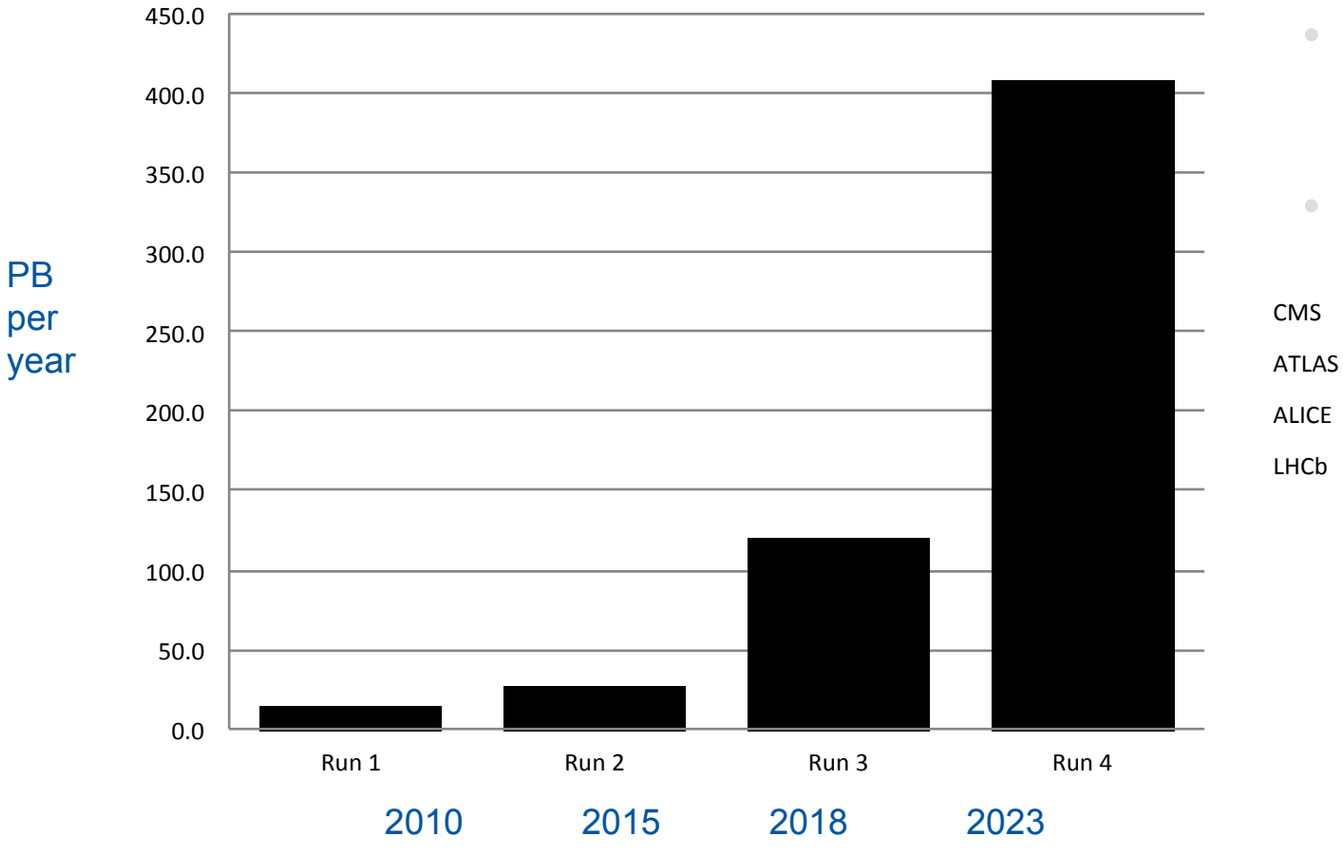






Credit: Ian Bird, CERN

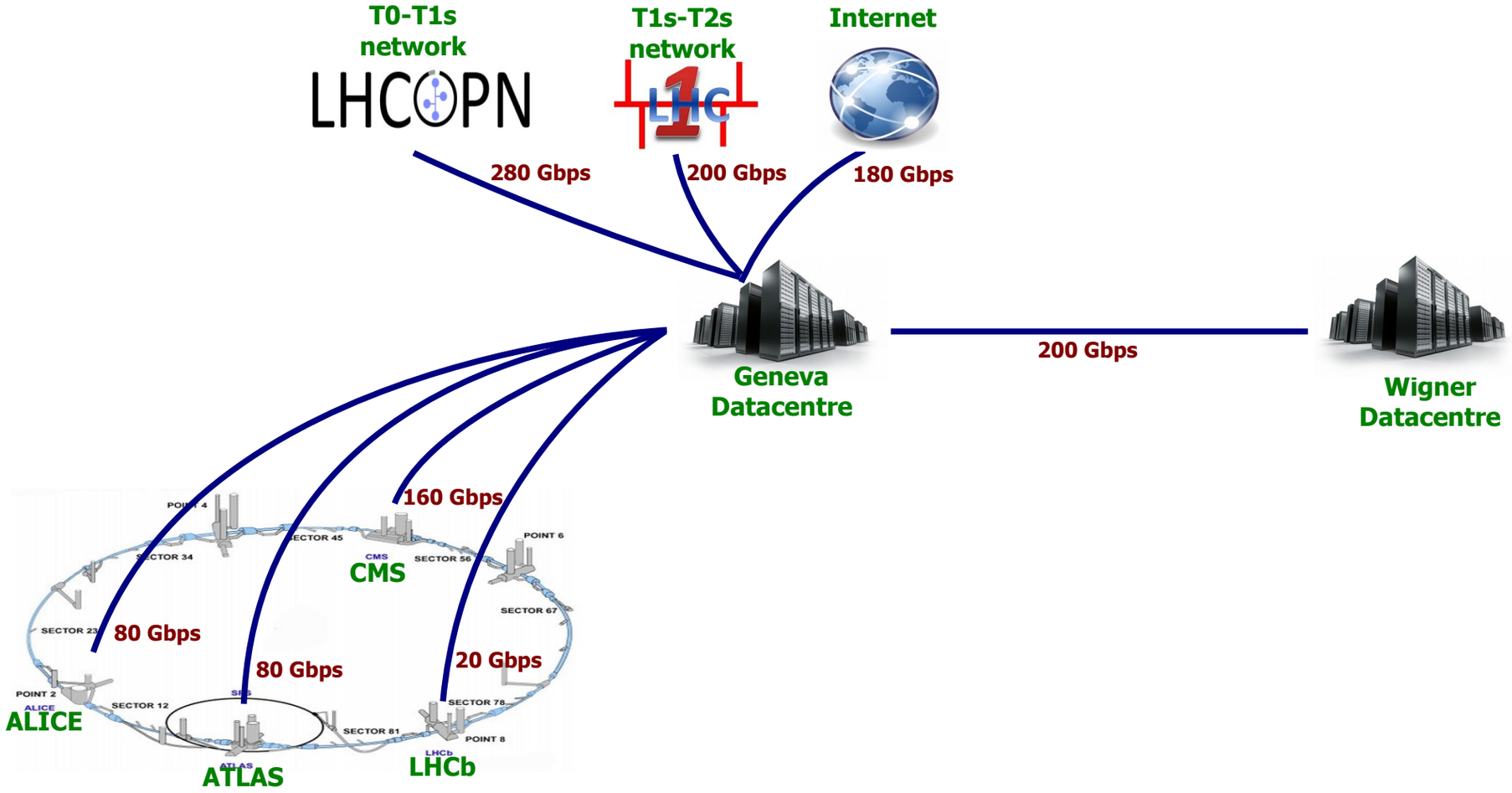
LHC data growth



- Expecting to record 400PB/year by 2023
- Compute needs expected to be around 50x current levels, if budget available

Credit: Ian Bird, CERN

Networks for LHC data



WLCG

WLCG resources:

1 Tier0 (CERN)

13 Tier1s

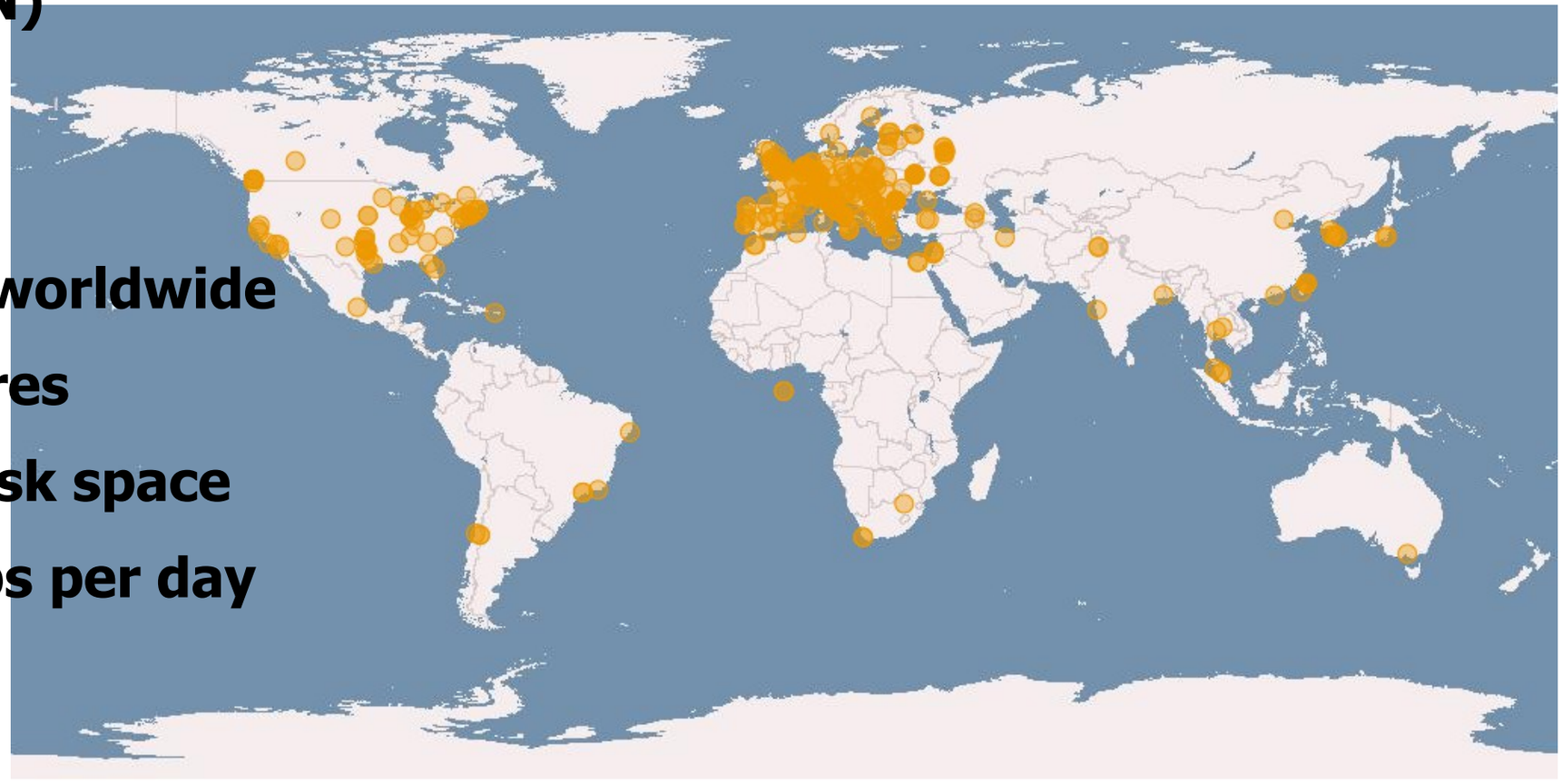
~170 Tier2s

>300 Tier3s worldwide

~350,000 cores

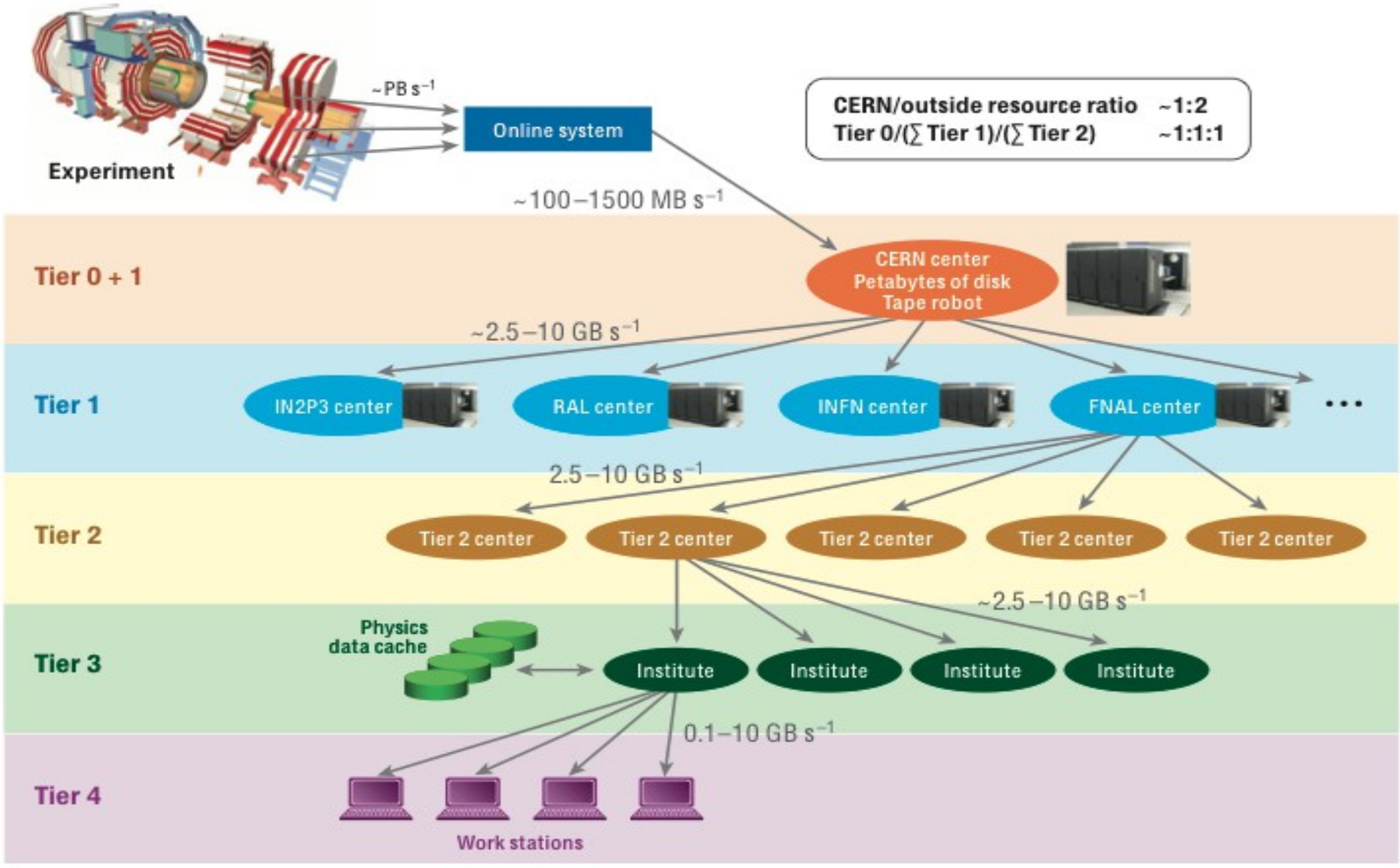
~500PB of disk space

2 millions jobs per day



LHCOPN – Tier0/1 network

Original Computing Model



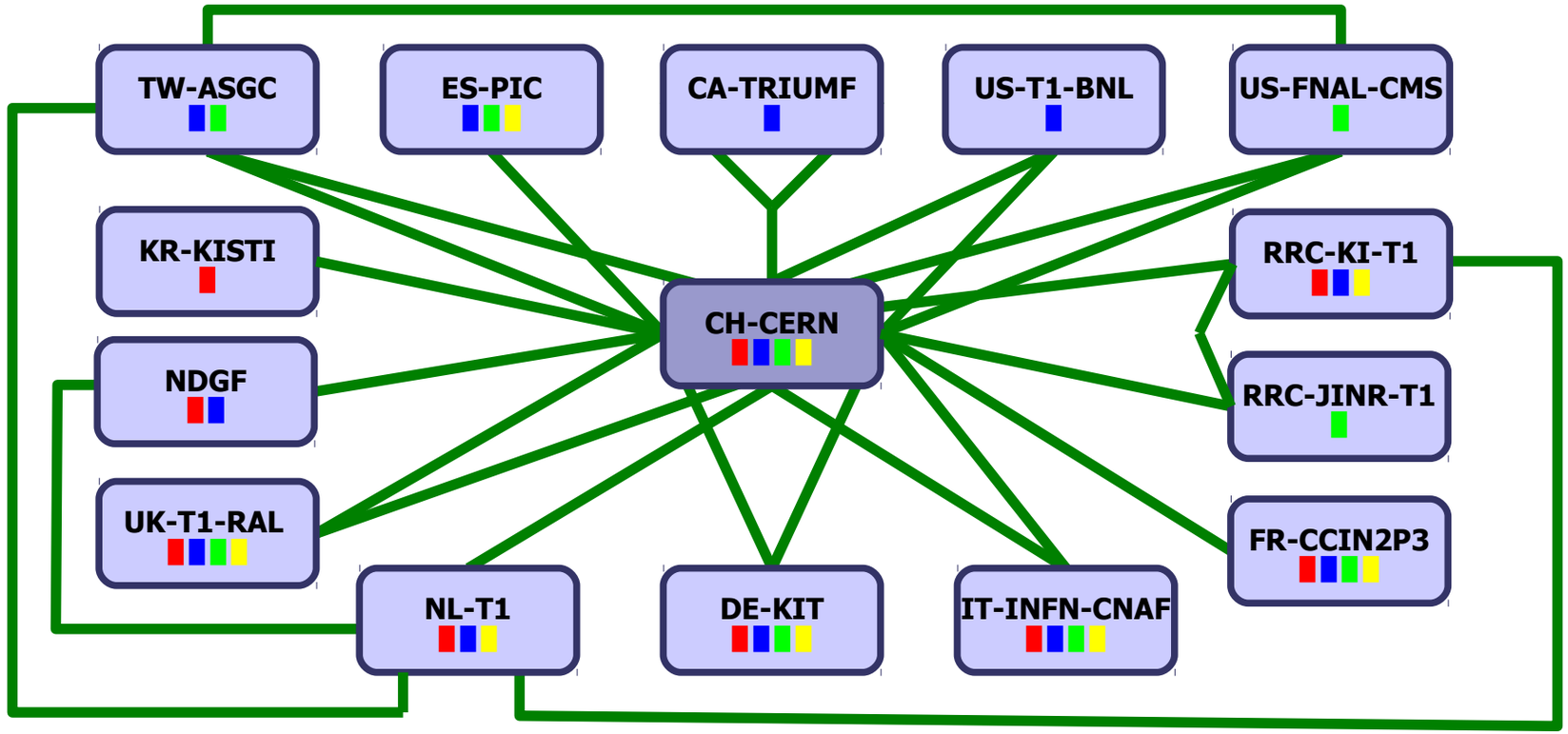
LHCOPN

Network designed to fit the Tier0-Tier1 LHC data distribution:

- private network connecting Tier0 and Tier1s
- reserved to LHC data transfers and analysis
- dedicated, large bandwidth links from each Tier1 to the Tier0

Topology

All Tier1s directly connected to CERN-Tier0



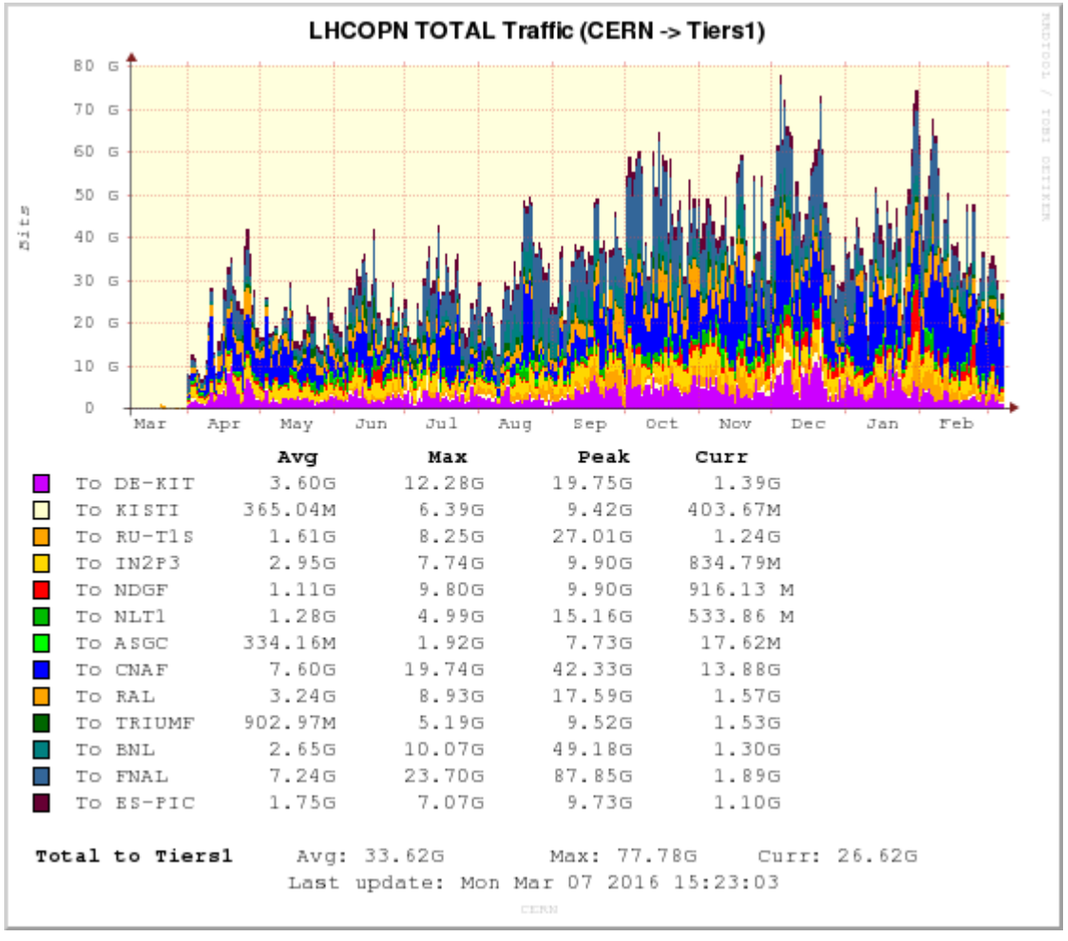
= Alice = Atlas = CMS = LHCb

edoardo.martelli@cern.ch 20150821

Main characteristics

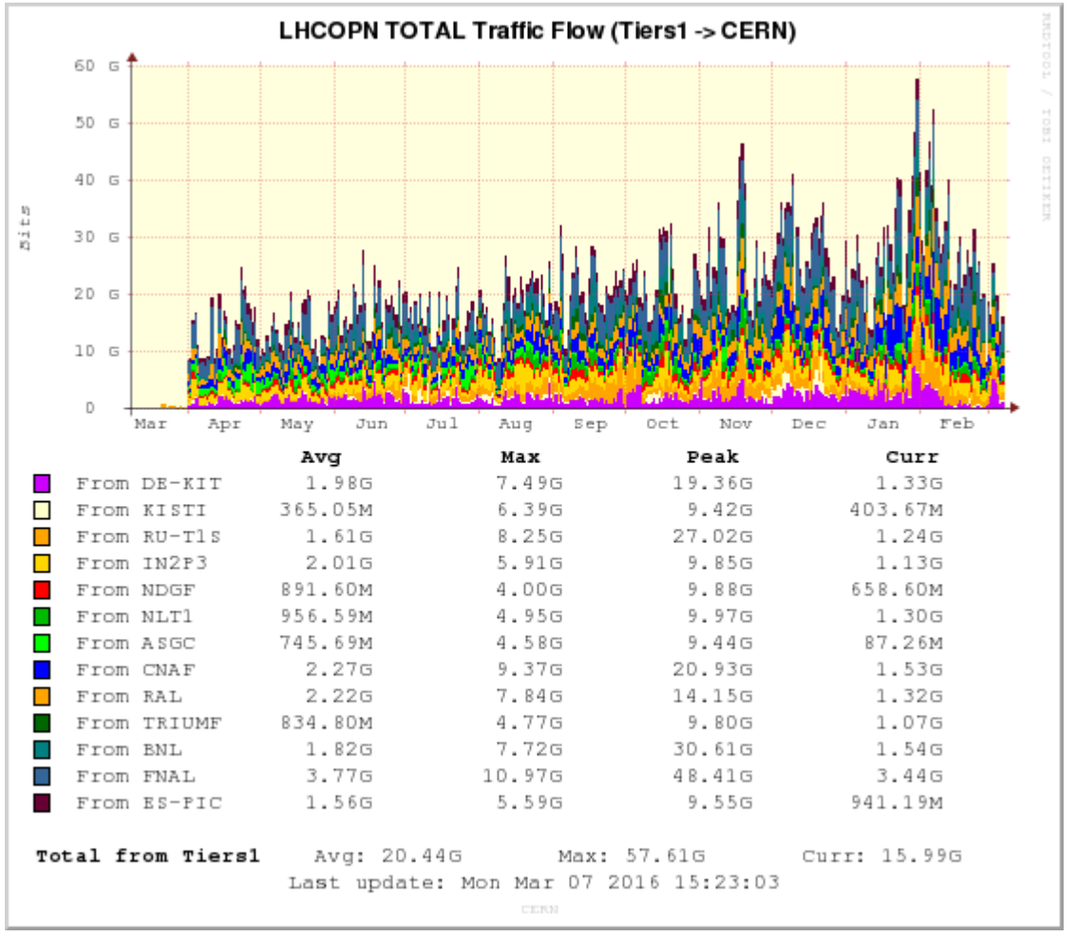
- **High bandwidth**: single and bundled long distance 10G and 100G Ethernet links
- **L3 Routing**: BGP peerings managed by the connected sites
- **Secured**: only packet to and from declared IP prefixes can cross the network

LHCOPN Traffic



	Avg	Max	Peak	Curr
To DE-KIT	3.60G	12.28G	19.75G	1.39G
To KISTI	365.04M	6.39G	9.42G	403.67M
To RU-T1S	1.61G	8.25G	27.01G	1.24G
To IN2P3	2.95G	7.74G	9.90G	834.79M
To NDGF	1.11G	9.80G	9.90G	916.13 M
To NLT1	1.28G	4.99G	15.16G	533.86 M
To ASGC	334.16M	1.92G	7.73G	17.62M
To CNAF	7.60G	19.74G	42.33G	13.88G
To RAL	3.24G	8.93G	17.59G	1.57G
To TRIUMF	902.97M	5.19G	9.52G	1.53G
To BNL	2.65G	10.07G	49.18G	1.30G
To FNAL	7.24G	23.70G	87.85G	1.89G
To ES-PIC	1.75G	7.07G	9.73G	1.10G

Total to Tiers1 Avg: 33.62G Max: 77.78G Curr: 26.62G
Last update: Mon Mar 07 2016 15:23:03



	Avg	Max	Peak	Curr
From DE-KIT	1.98G	7.49G	19.36G	1.33G
From KISTI	365.05M	6.39G	9.42G	403.67M
From RU-T1S	1.61G	8.25G	27.02G	1.24G
From IN2P3	2.01G	5.91G	9.85G	1.13G
From NDGF	891.60M	4.00G	9.88G	658.60M
From NLT1	956.59M	4.95G	9.97G	1.30G
From ASGC	745.69M	4.58G	9.44G	87.26M
From CNAF	2.27G	9.37G	20.93G	1.53G
From RAL	2.22G	7.84G	14.15G	1.32G
From TRIUMF	834.80M	4.77G	9.80G	1.07G
From BNL	1.82G	7.72G	30.61G	1.54G
From FNAL	3.77G	10.97G	48.41G	3.44G
From ES-PIC	1.56G	5.59G	9.55G	941.19M

Total from Tiers1 Avg: 20.44G Max: 57.61G Curr: 15.99G
Last update: Mon Mar 07 2016 15:23:03

Source: https://netstat.cern.ch/monitoring/network-statistics/ext/?q=LHCOPN&p=LHCOPN&mn=00-Total-Traffic&t=all

LHCOPN evolution

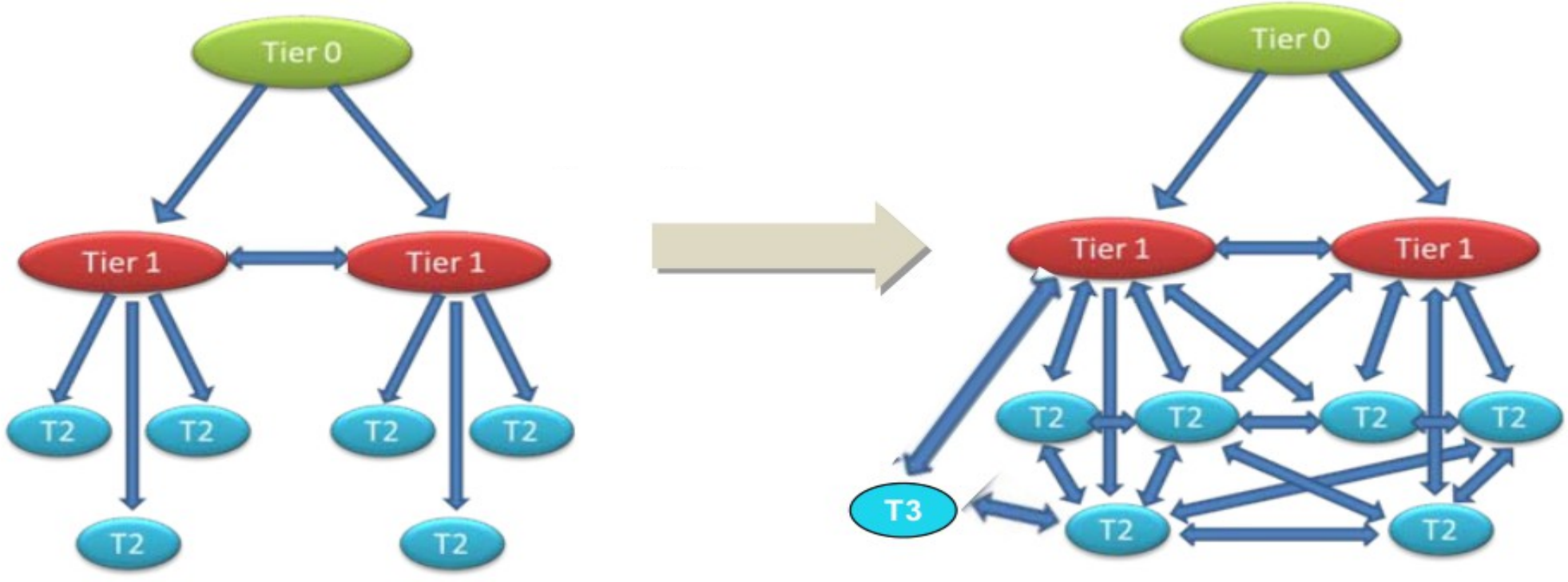
- The LHCOPN is kept as the main network to exchange data among Tier0 and Tier1s

- Tier1-Tier1 traffic and backup functionalities progressively moving to LHCONE (the Tier1/2/3 network)

- Links being upgraded to multiple 10Gbps and 100G

LHCONE - Tier1/2/3 network

Computing model evolution



Original MONARCH model

Model evolution

LHCONE

LHCONE is the network that serves the WLCG and HEP sites around the world

It is the result of the effort and resources provided by several major Research and Education network providers

LHCONE services

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

perfSONAR: monitoring infrastructure - *operational*

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

LHCONE L3VPN service

LHCONE L3VPN service

Layer3 (i.e. routed) Virtual Private Network

Dedicated worldwide backbone connecting Tier1s, Tier2s and Tier3s at high bandwidth

Reserved to HEP data distribution

Benefits

Bandwidth dedicated to HEP, no contention with other research projects

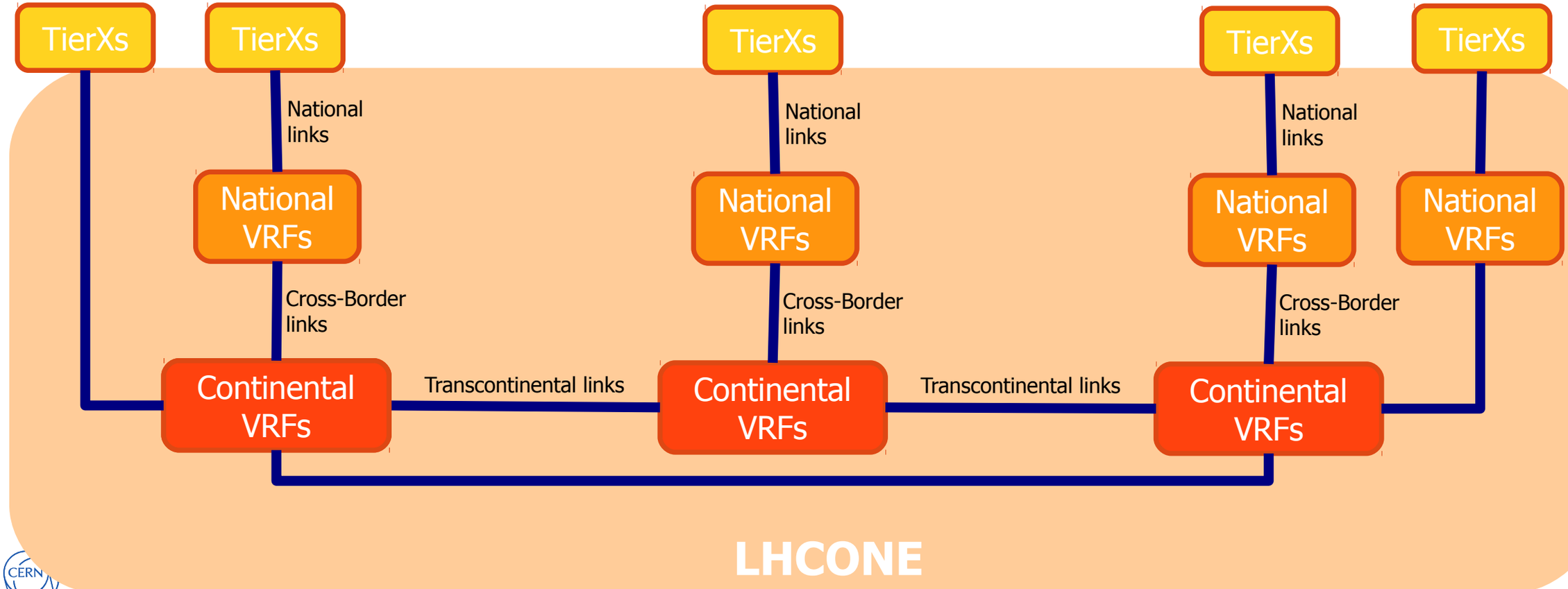
Closed network, only trusted traffic that can bypass perimeter firewalls

Well defined cost tag for HEP networking

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 (i.e. virtual routing instance)



Status

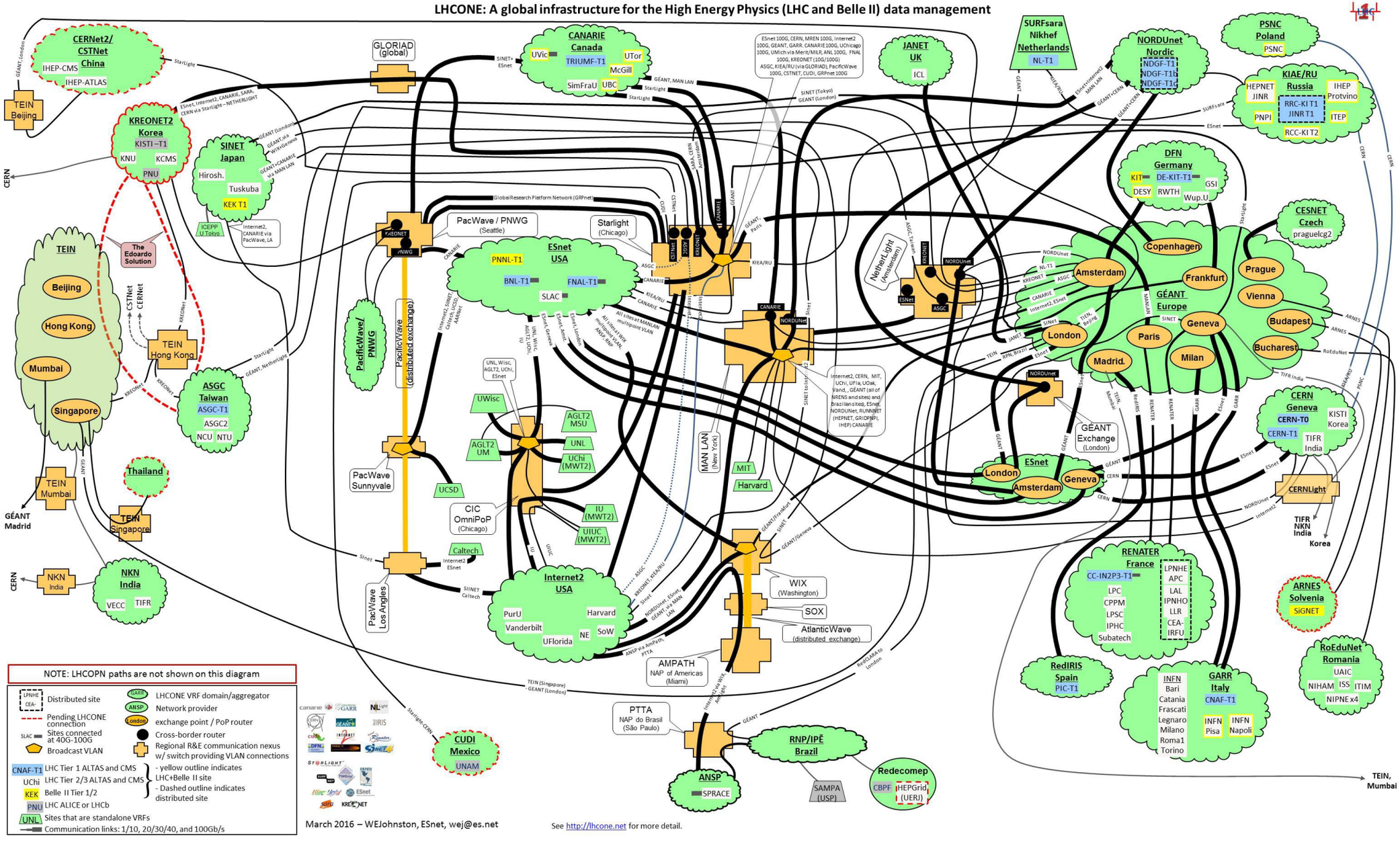
VRFs provided by over 15 national and international Research Networks

VRFs interconnections at Open Exchange Points, including NetherLight, StarLight, MANLAN, WIX, CERNlight and others

55 end sites connected to LHCONE:
- 10 Tier1s
- 45 Tier2s

More Information:
- https://indico.cern.ch/event/461511/contribution/5/attachments/1242595/1828290/2016-03-13_LHCONE_L3VPN_Update_Capone.pdf
- https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneVRF

Current L3VPN topology

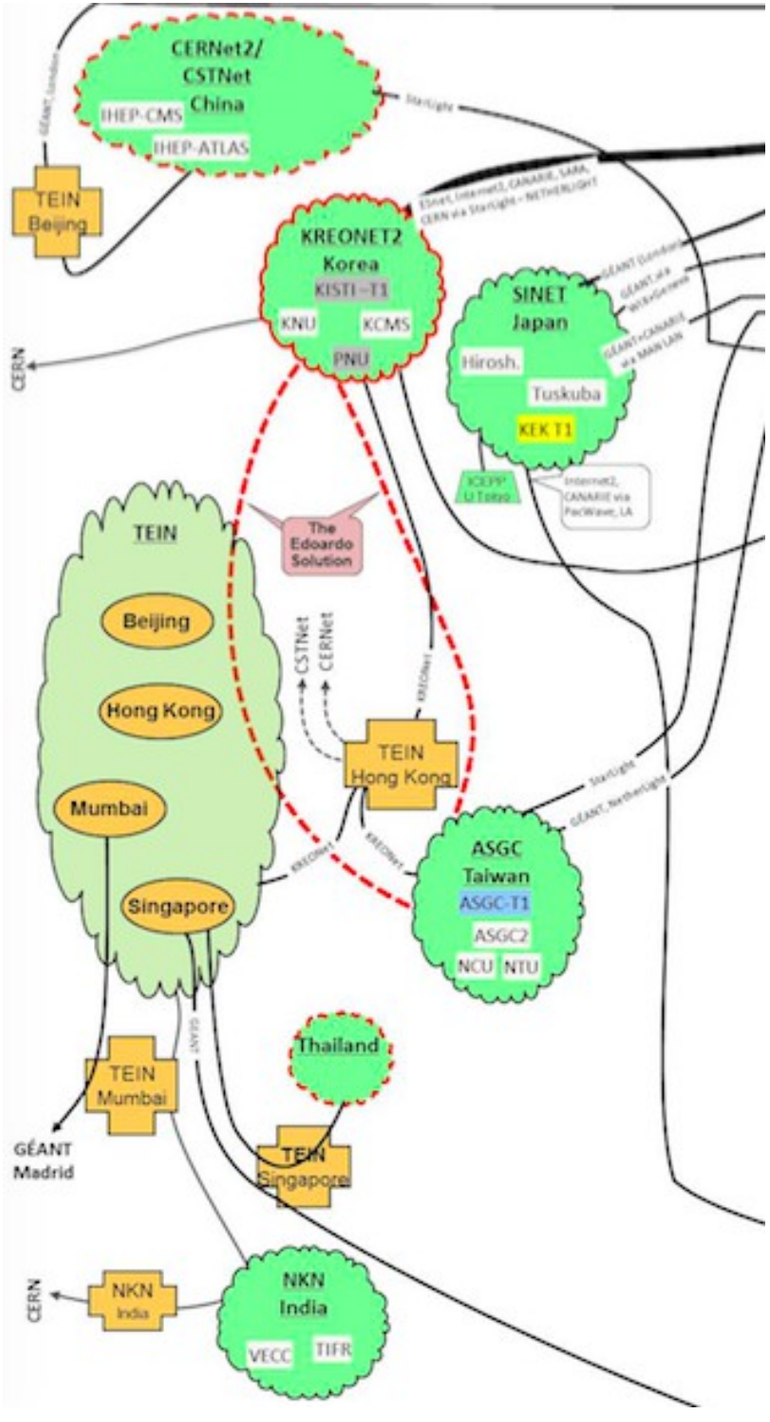


Focusing on Asia-Pacific

ASGC (TW), KISTI (KR), TEIN (Asia) are completing their VRFs

They will interconnect with most of existing VRFs and among themselves

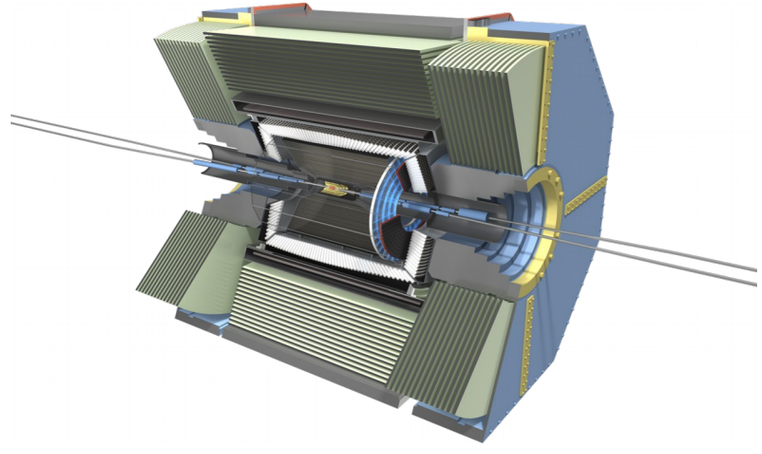
SINET (JP), CERnet/CSTnet (CN) and others are planning their own VRFs



Open to other HEP collaborations

The L3VPN is now used also by:

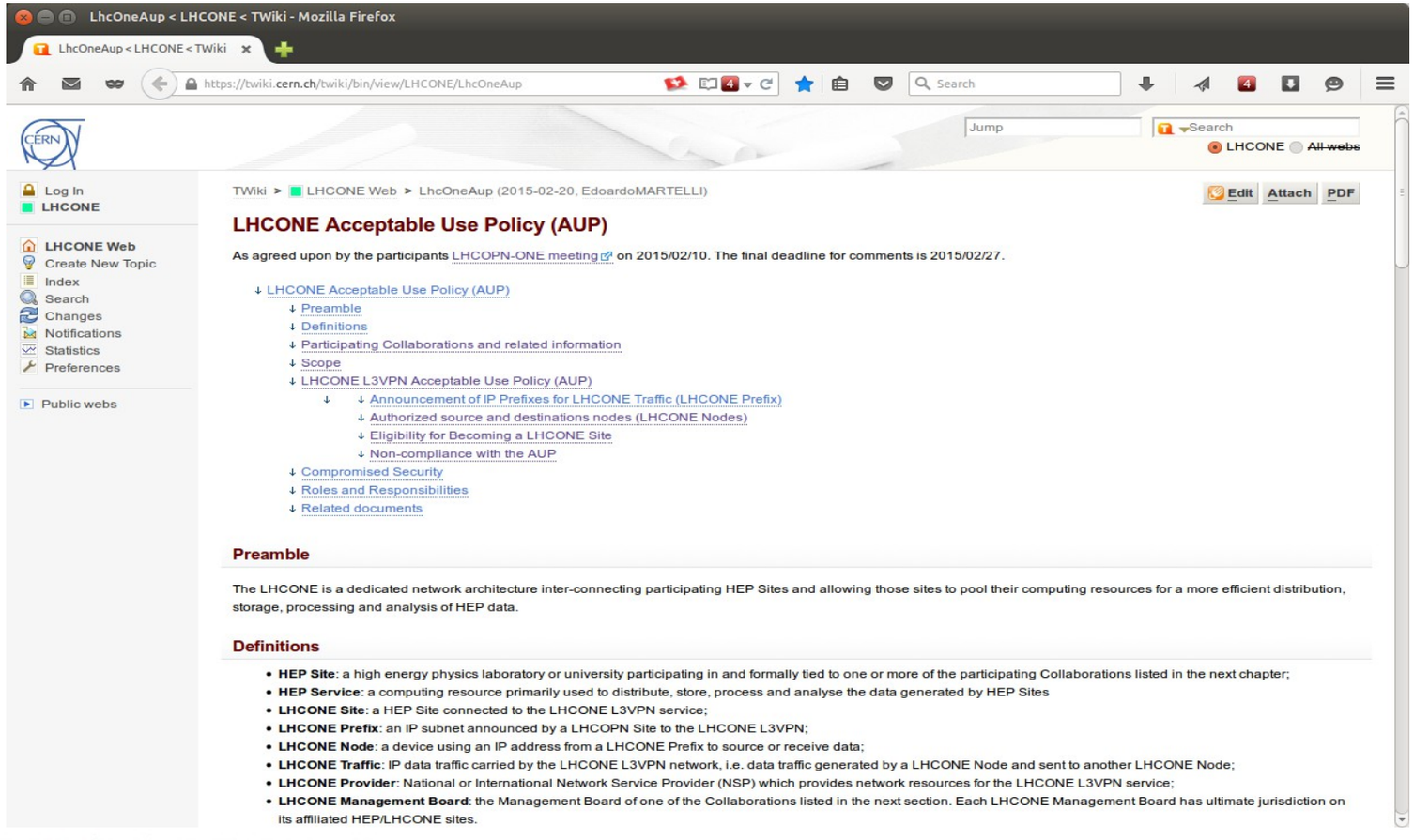
- Belle II experiment



- Pierre Auger Observatory

Acceptable Use Policy

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



LHCONE P2P service

LHCONE P2P service

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

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

Accessible and configurable via software API

Status

Work in progress: still in the prototyping phase

Challenges:

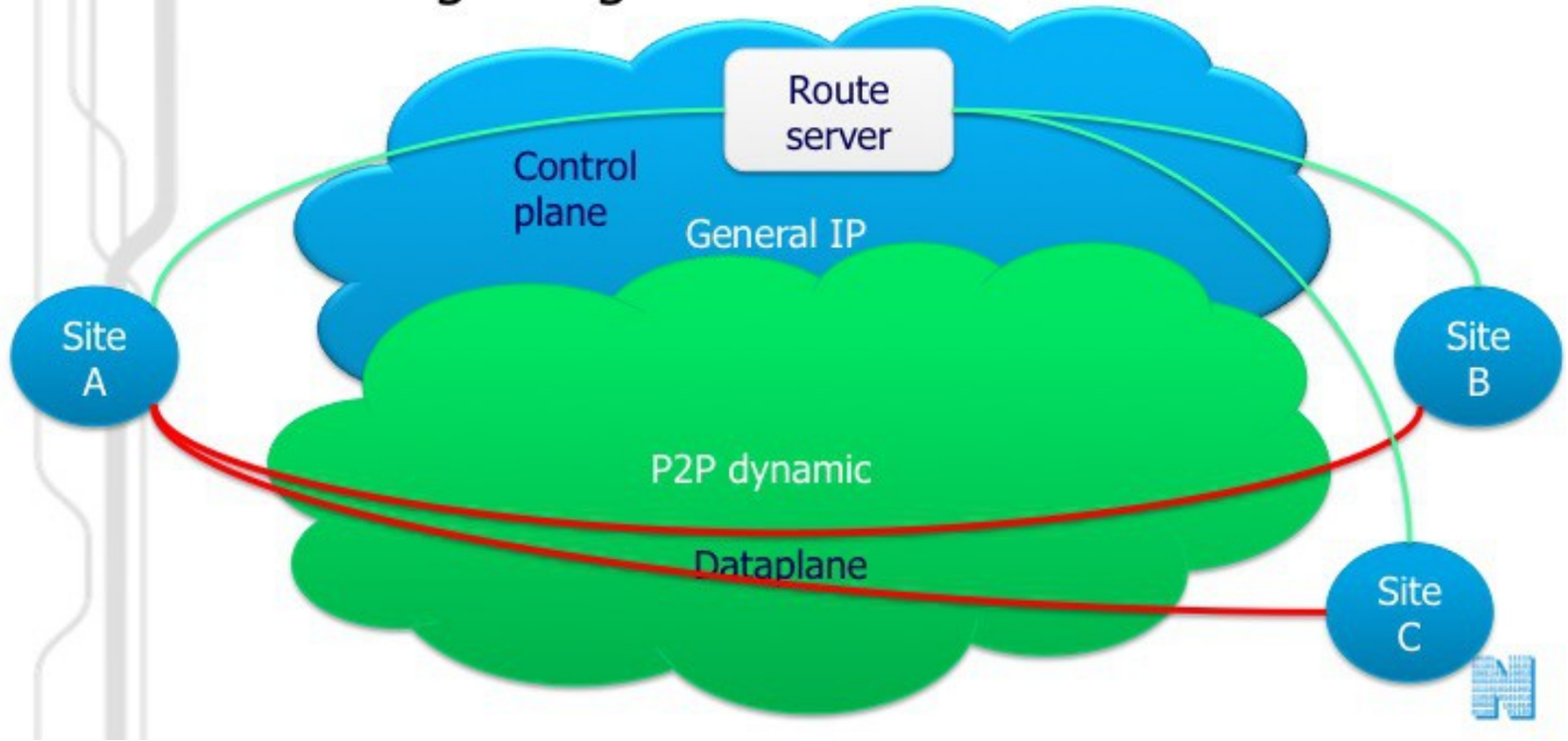
- multi-domain provisioning system
- intra-TierX connectivity
- TierX-TierY routing
- integration with WLCG software
- bandwidth allocation and protection

Solution 1 based on draft-ietf-idr-rs-bfd-01

NORDUnet
Nordic Gateway for Research & Education

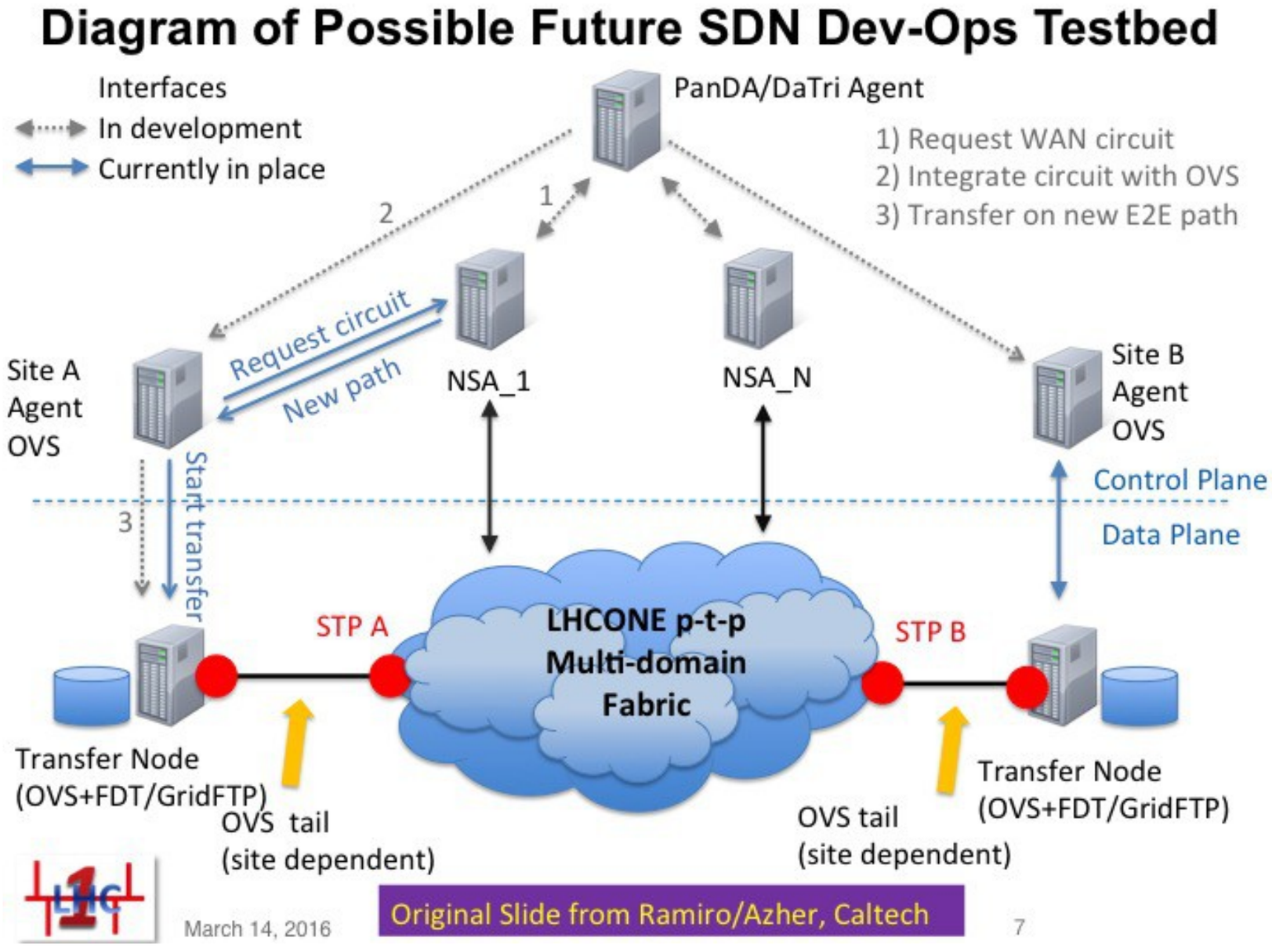
Route server

- Simplify the BGP setup
- Only one BGP session per site
- Route server with one outgoing RIB per site, steering using communities



Credit: Magnus Bergroth, NORDUnet

Solution 2 – based on SDN Openflow



Credit: Azher Mugal and Ramiro Voicu, Caltech – Shawn McKee, Univ. of Michigan

LHCONE perfSONAR service

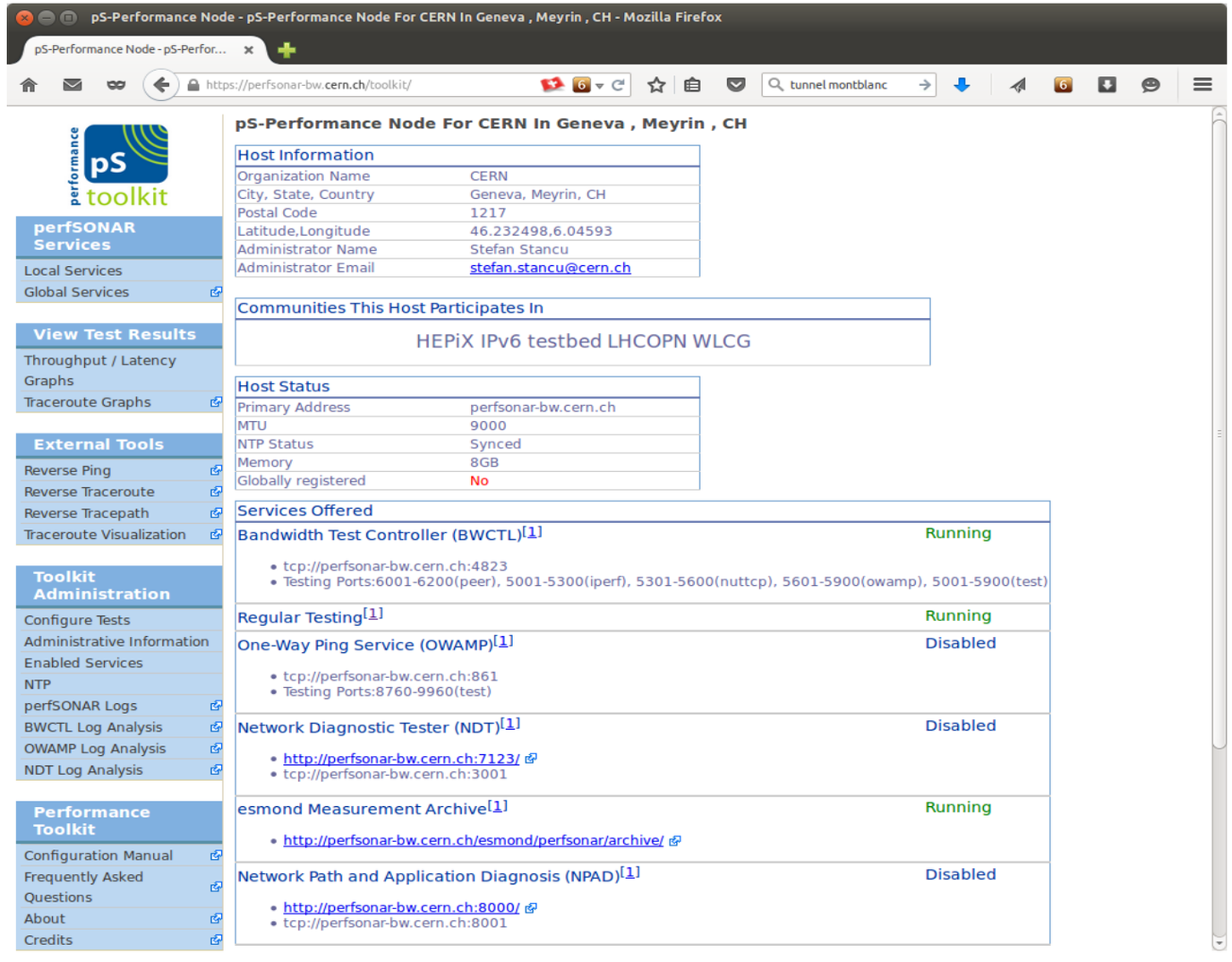
perfSONAR

- software framework for active and passive network probing

- developed by Internet2, ESnet, GEANT and others

perfSONAR

http://www.perfsonar.net/



LHCONE perfSONAR service

LHCONE Network monitoring infrastructure

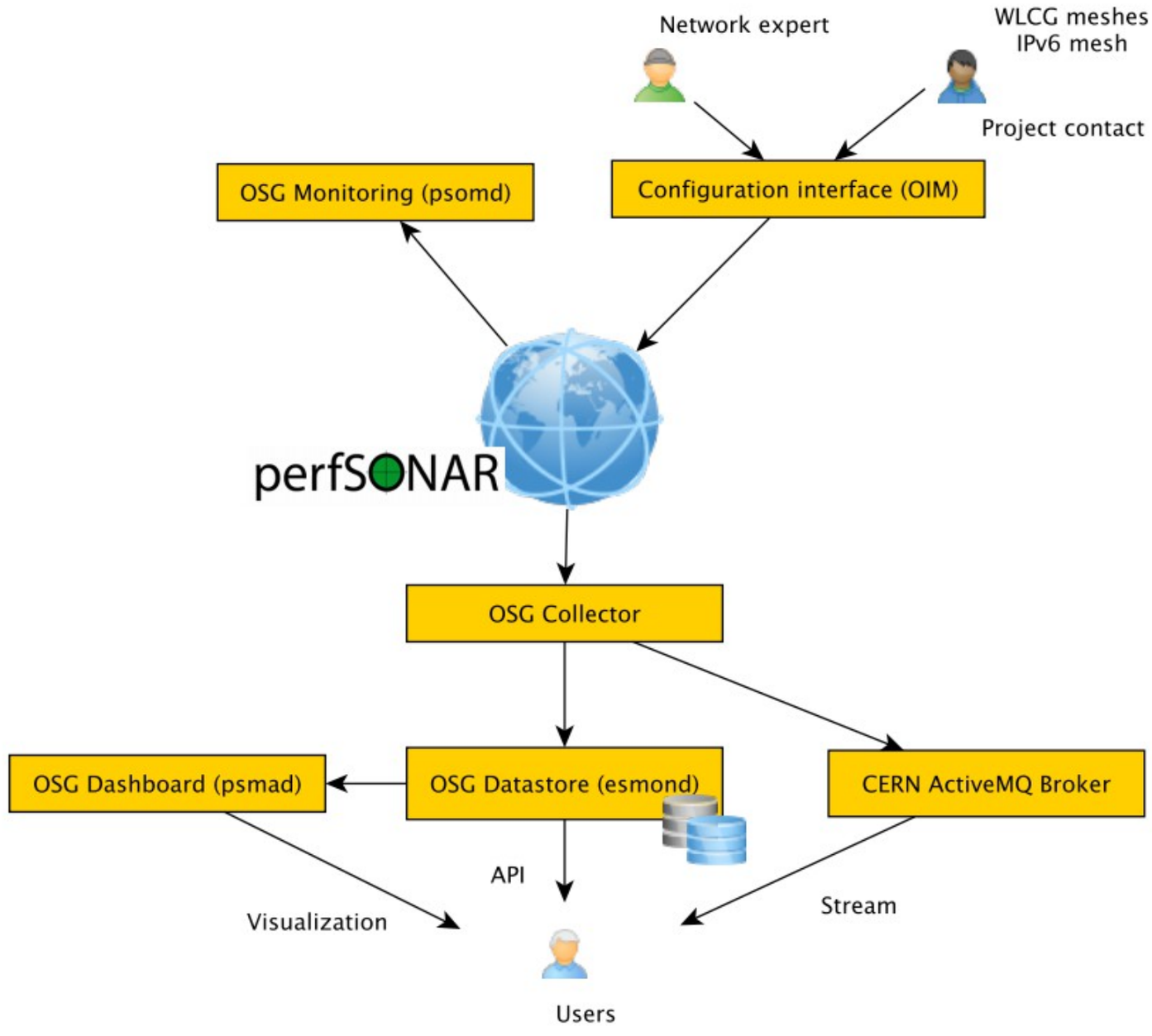
Probes installed at:
- VRFs interconnecting points
- Sites

Accessible by any Site for network healthiness checks and faster problem resolution

perfSONAR pipeline

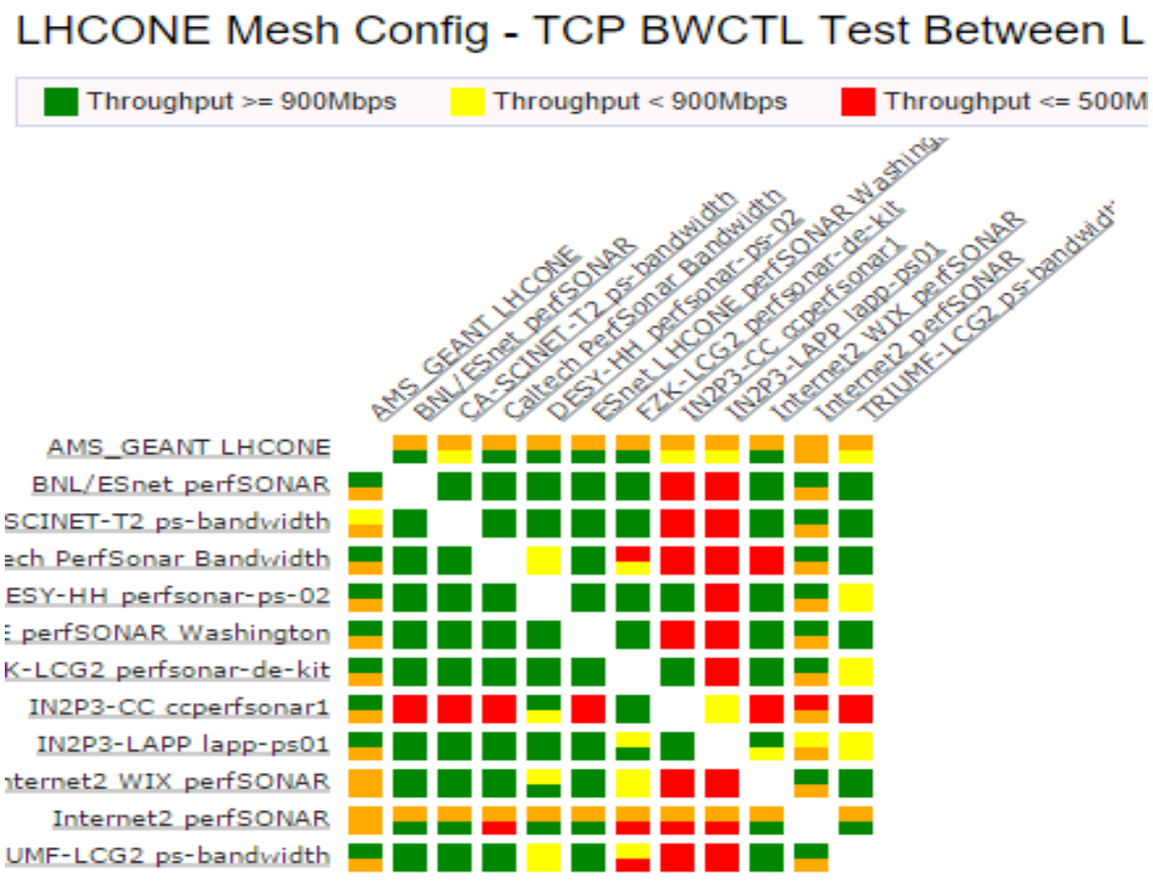
The diagram on the right provides a high-level view of how WLCG/OSG is managing the perfSONAR deployments, gathering metrics and making them available for use.

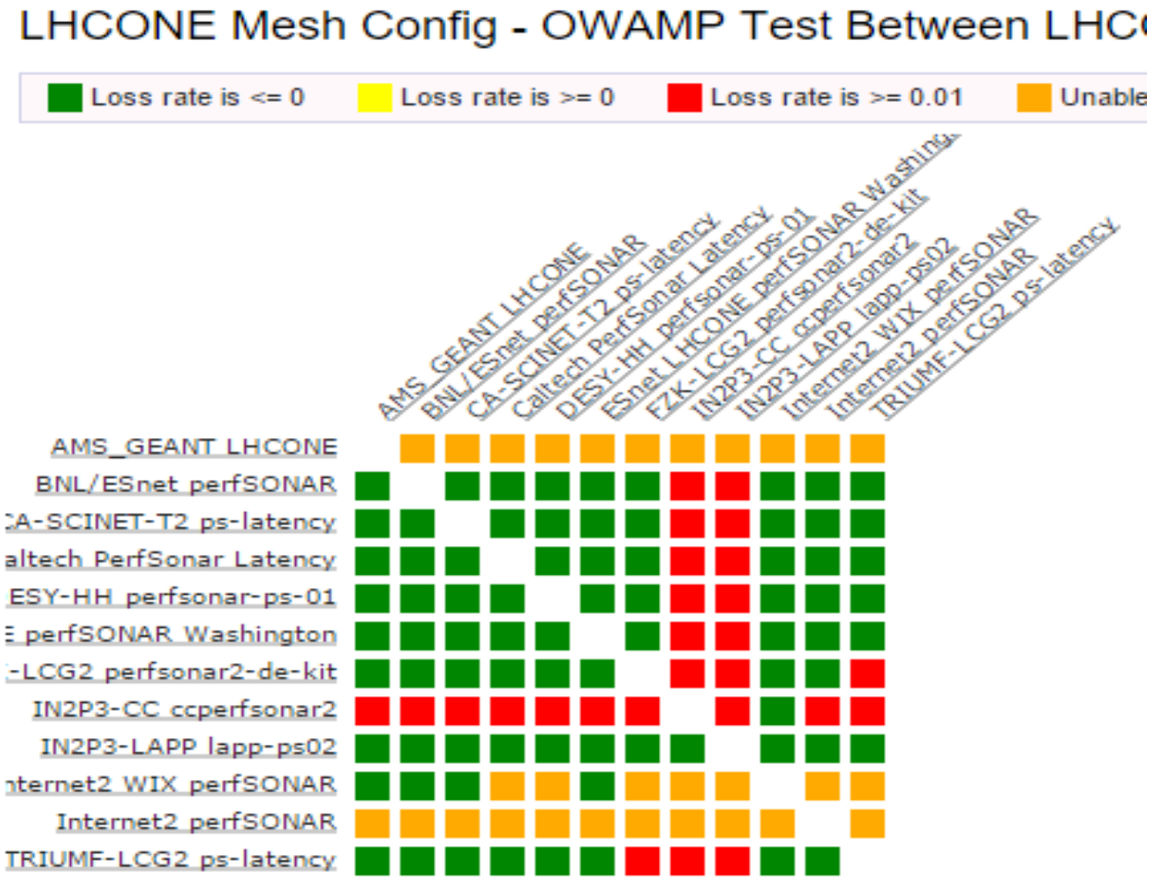
End users can get monitor the data via the OSG MaDDash instance, grab the data directly from the OSG datastore or subscribe to the ActiveMQ bus at CERN



Credit: Shawn McKee, Univ. of Michigan

perfSONAR LHCONE MaDDash





More information on perfSONAR

Deployment

https://twiki.opensciencegrid.org/bin/view/Documentation/DeployperfSONAR

MaDDash

https://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=LHCONE%20Mesh%20Config

perfSONAR

http://www.perfsonar.net/

Latest update

https://indico.cern.ch/event/461511/contribution/0/attachments/1242565/1828189/LHCONE_perfSONAR_update-Taiwan-2016.pptx

Conclusion

Summary

More and more data is generated by the LHC experiments

LHCOPN, Tier0/1 Network:
- increasing capacity to cope with growing amount LHC data

LHCONE, Tier1/2/3 Network:
- increasing capacity to serve data greedy sites
- extending worldwide coverage
- now fast growing in Asia-Pacific region
- growing monitoring infrastructure to better support operation

More information on LHCOPN and LHCONE

LHCOPN and LHCONE meetings:
https://indico.cern.ch/category/5793/

LHCOPN website:
https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome

LHCONE websites:
http://lhcone.net
https://twiki.cern.ch/twiki/bin/view/LHCONE/WebHome

Questions?

CERN