

Networking for WLCG: LHCOPN and LHCONE

ISGC Taipei, 17th March 2016
Edoardo.Martelli@cern.ch

Summary

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

LHC data challenge

LHC Data Challenge

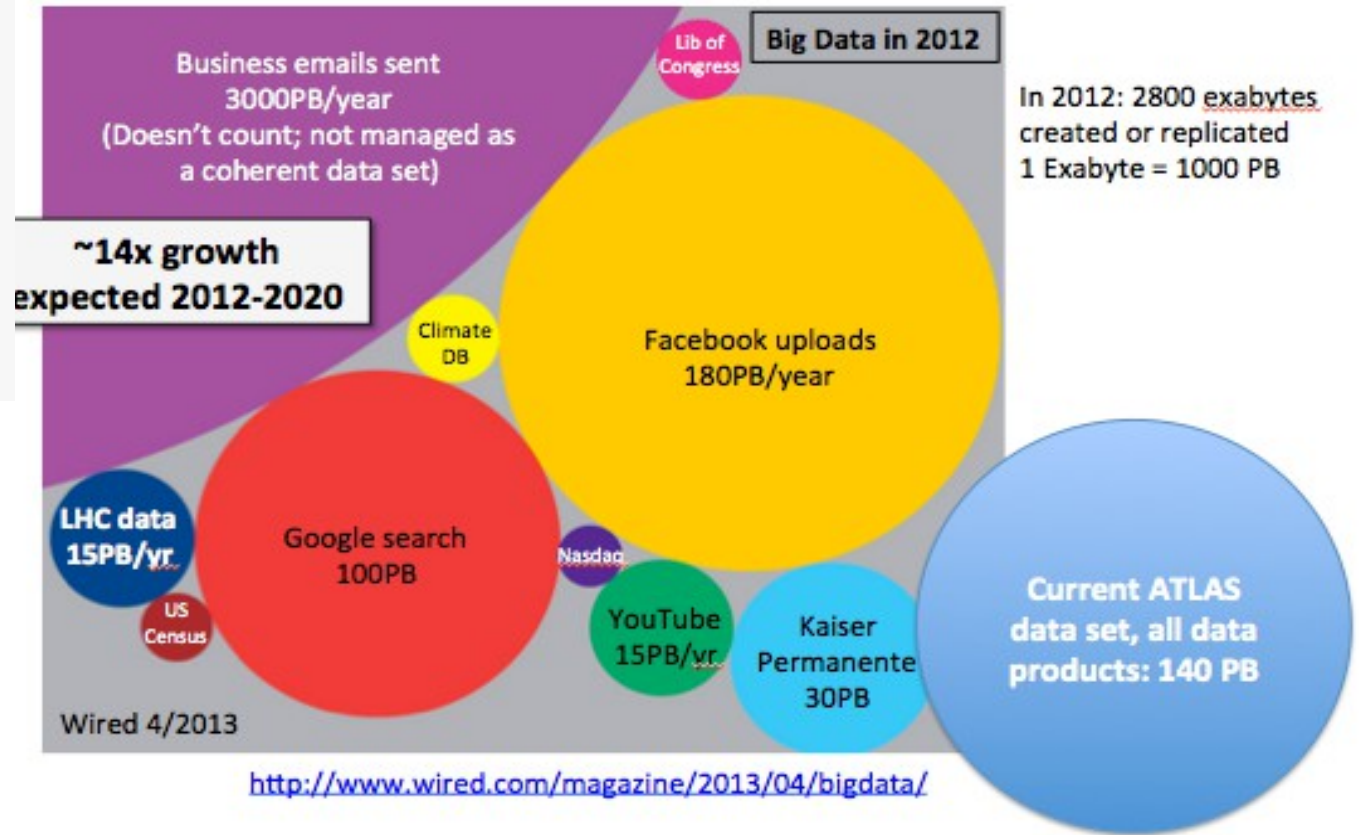
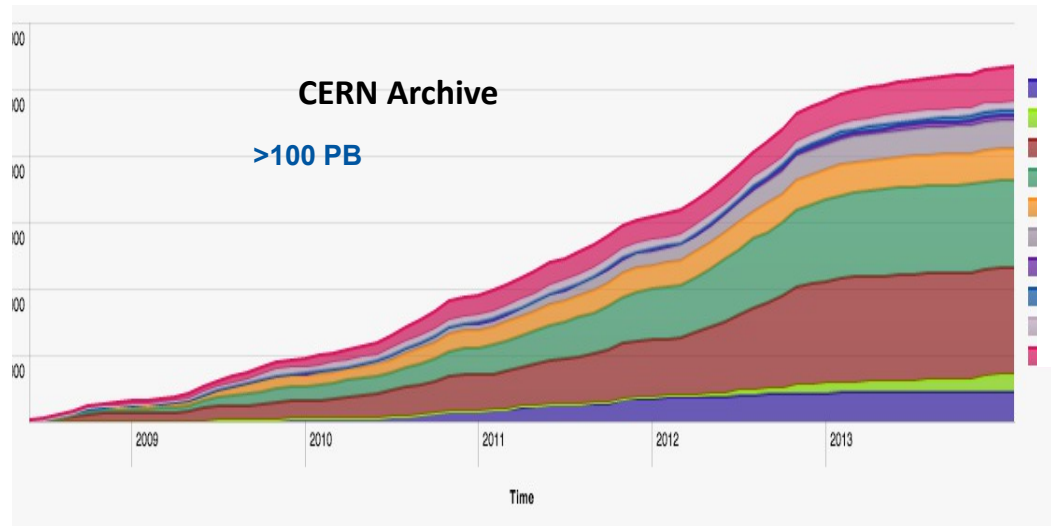
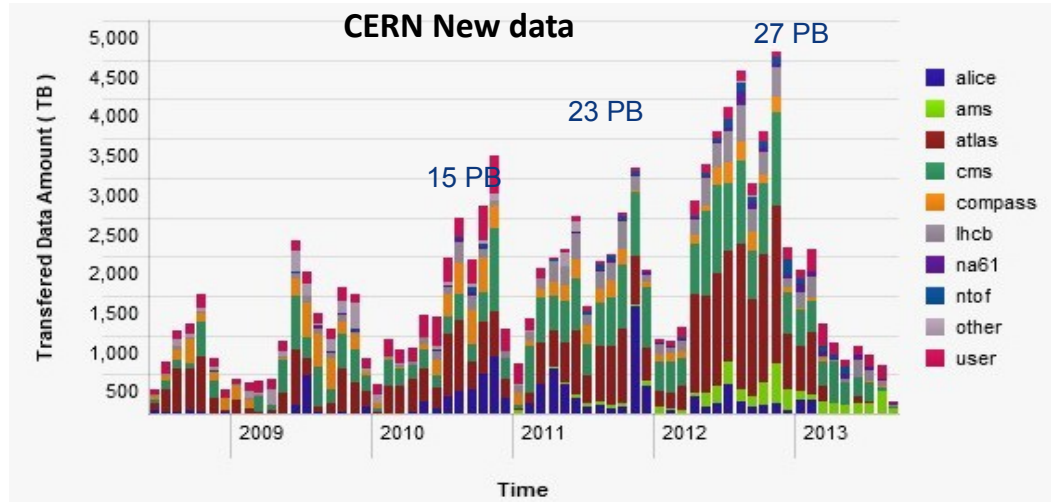
40 million collisions per second

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

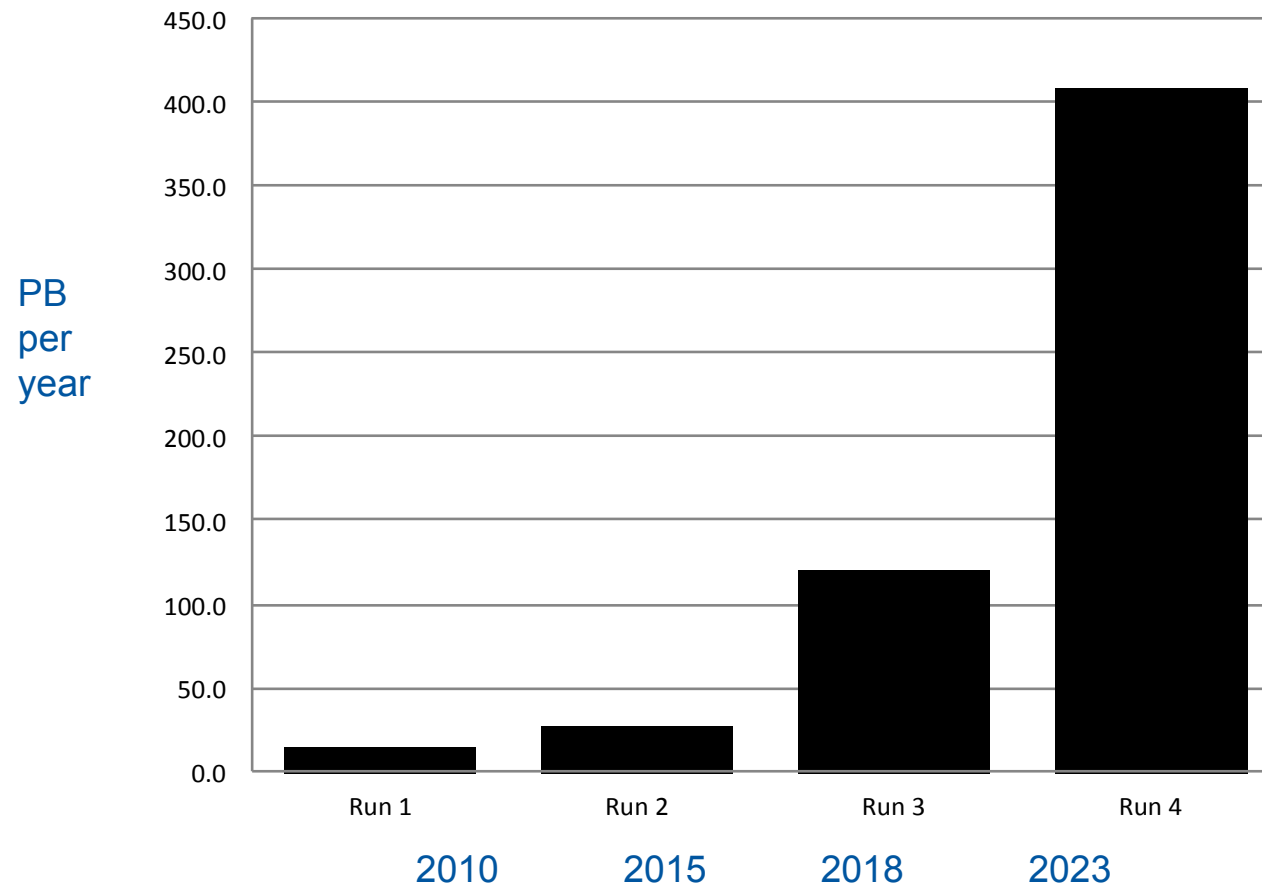
10^{10} collisions recorded each year

=> more than 25 Petabytes/year of data

LHC data production



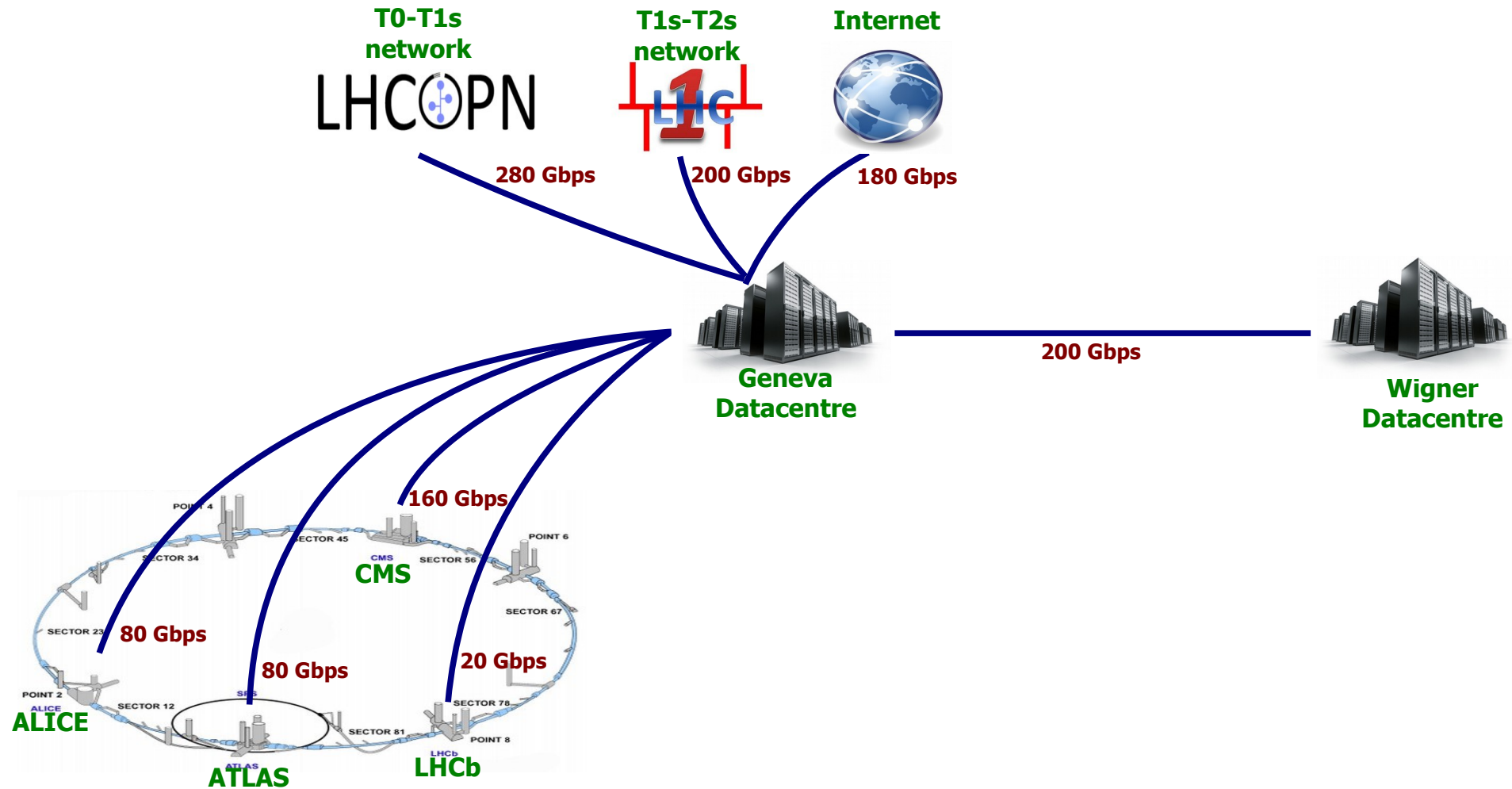
LHC data growth

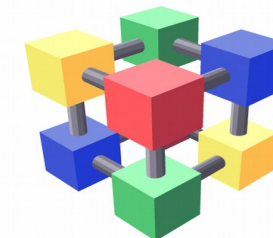


- Expecting to record 400PB/year by 2023
- Compute needs expected to be around 50x current levels, if budget available

CMS
ATLAS
ALICE
LHCb

Networks for LHC data





WLCG
Worldwide LHC Computing Grid

WLCG resources:

1 Tier0 (CERN)

13 Tier1s

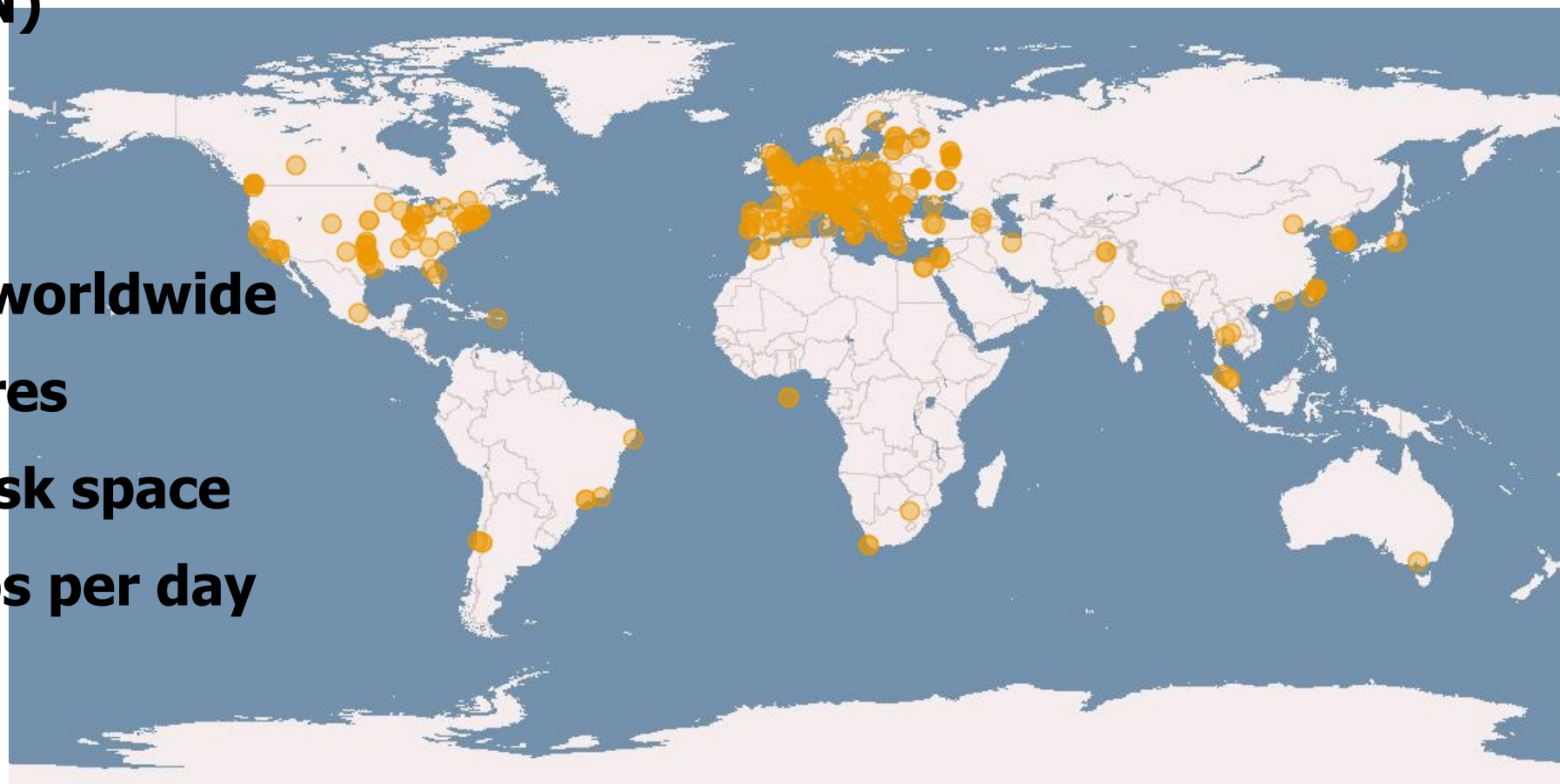
~170 Tier2s

>300 Tier3s worldwide

~350,000 cores

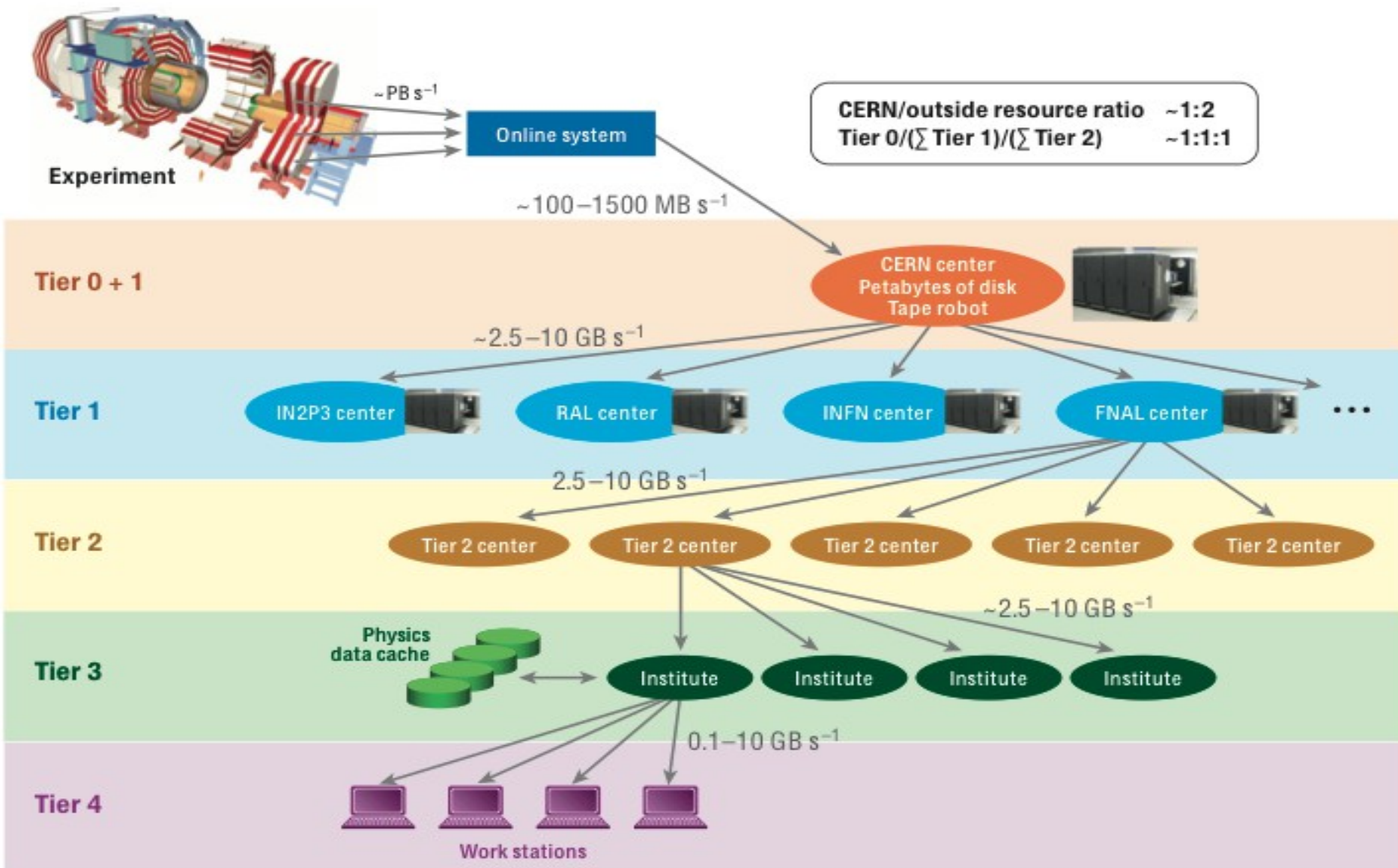
~500PB of disk space

2 millions jobs per day



LHCOPN – Tier0/1 network

Original Computing Model

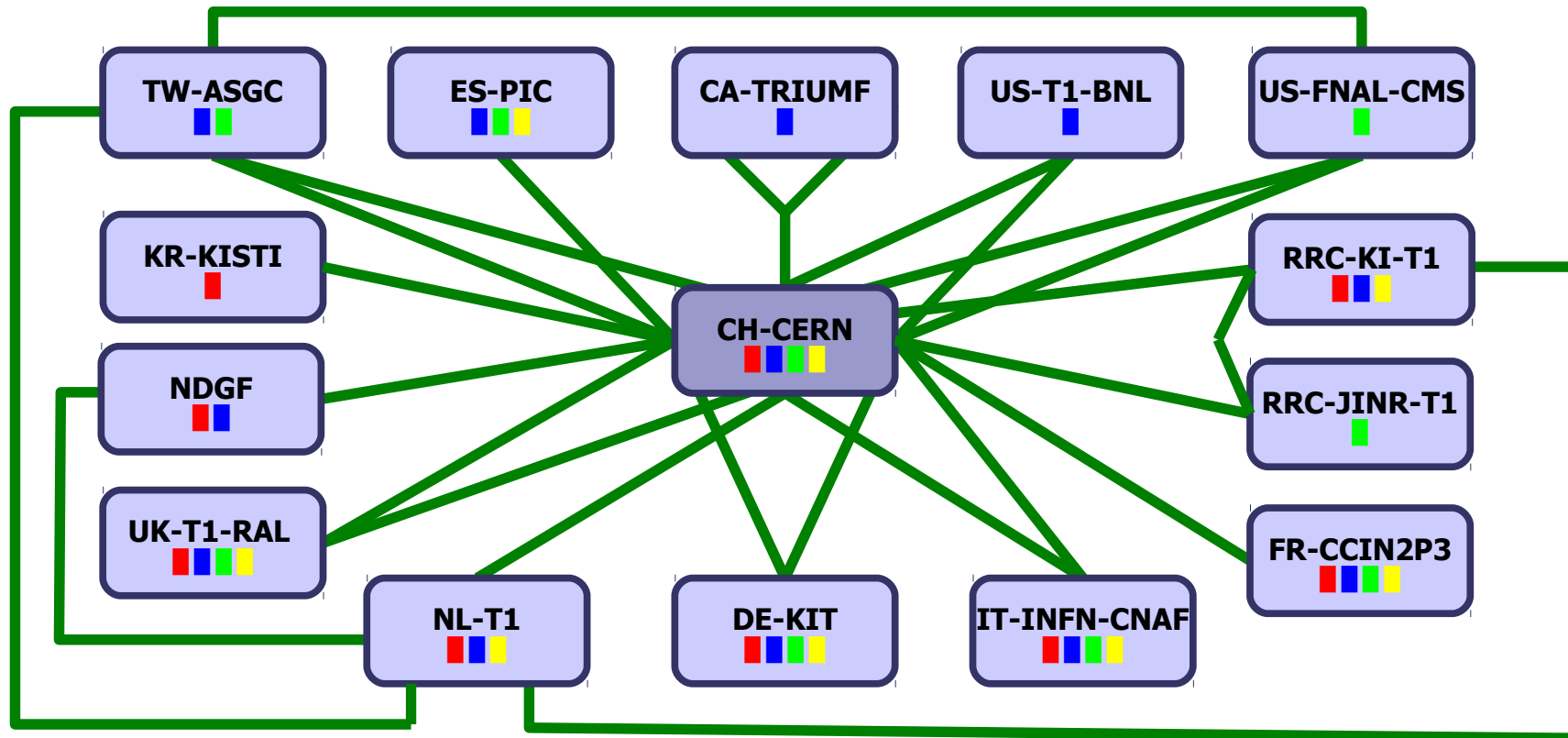


Network designed to fit the Tier0-Tier1 LHC data distribution:

- 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

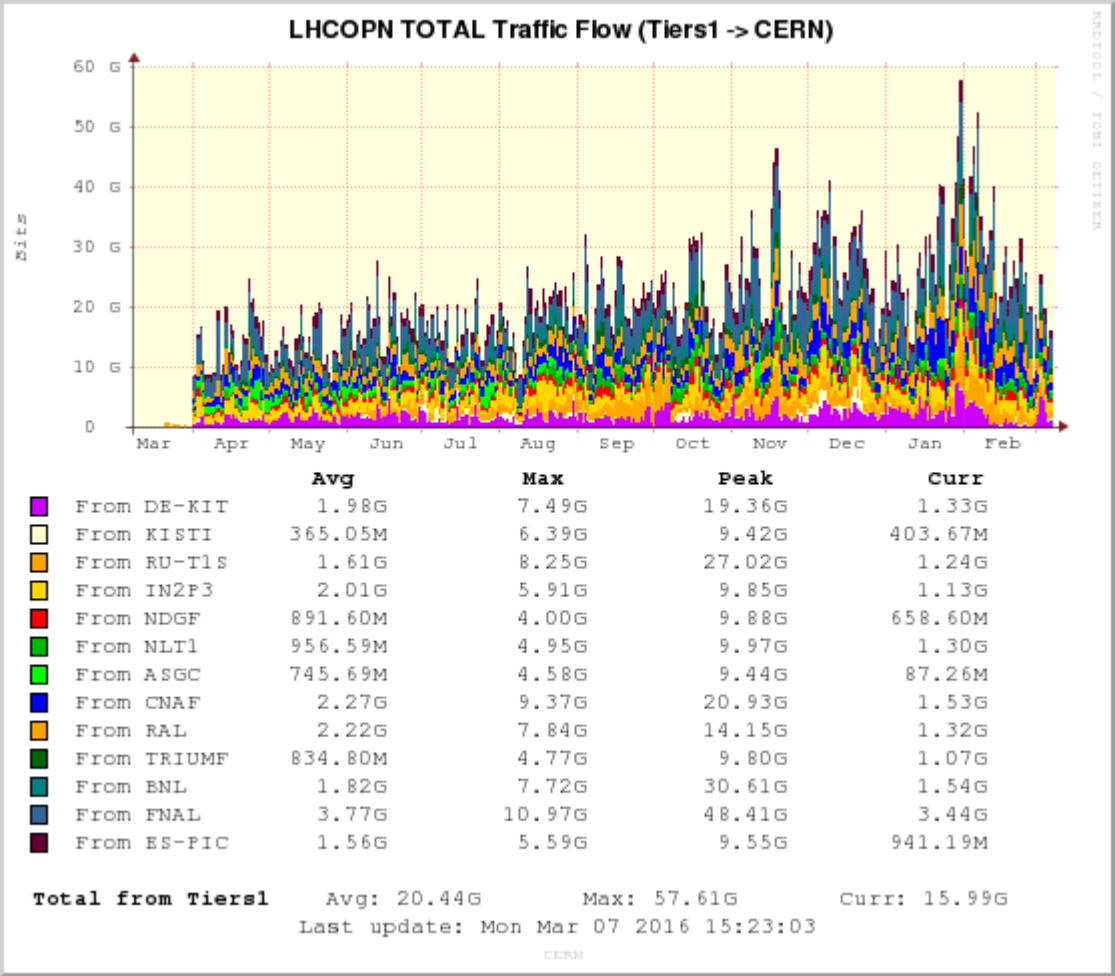
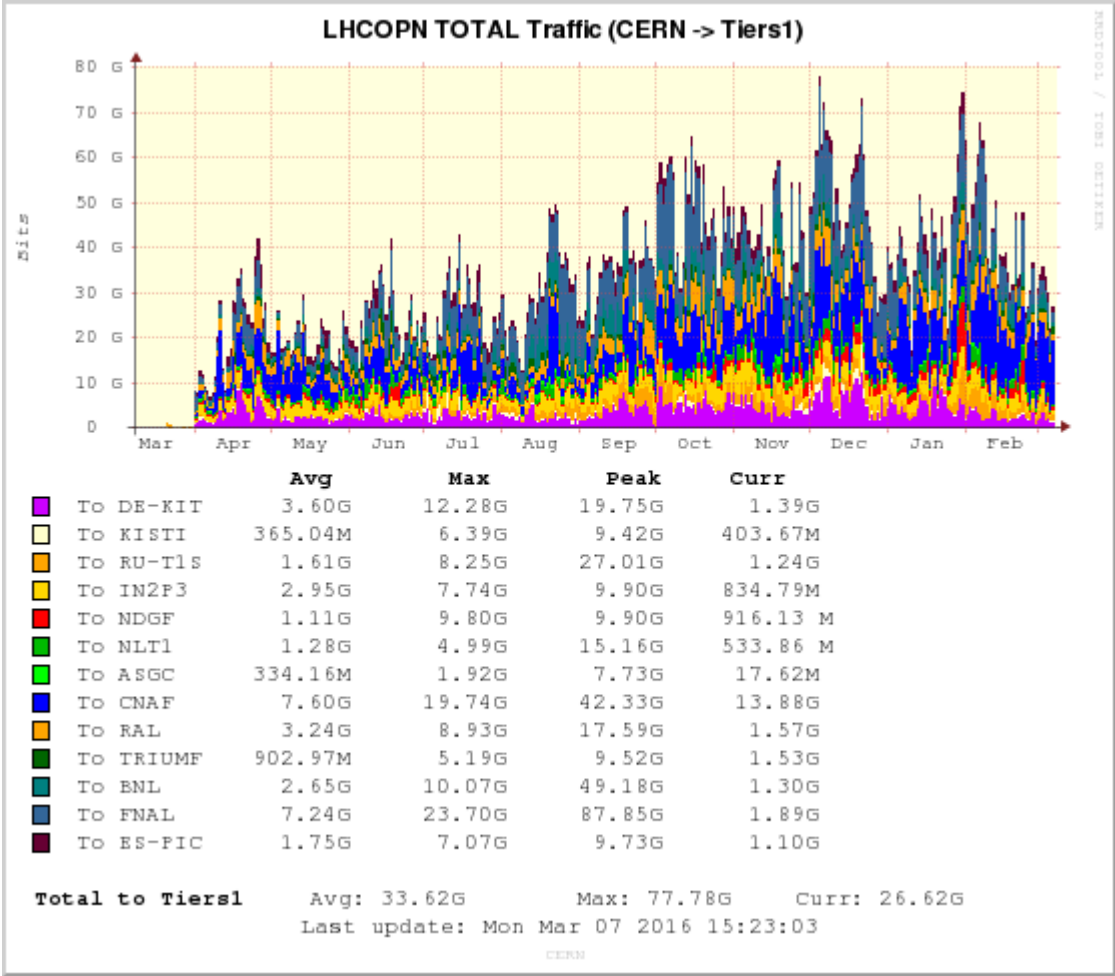


■ = Alice ■ = Atlas ■ = CMS ■ = LHCb
eduardo.martelli@cern.ch 20150821

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



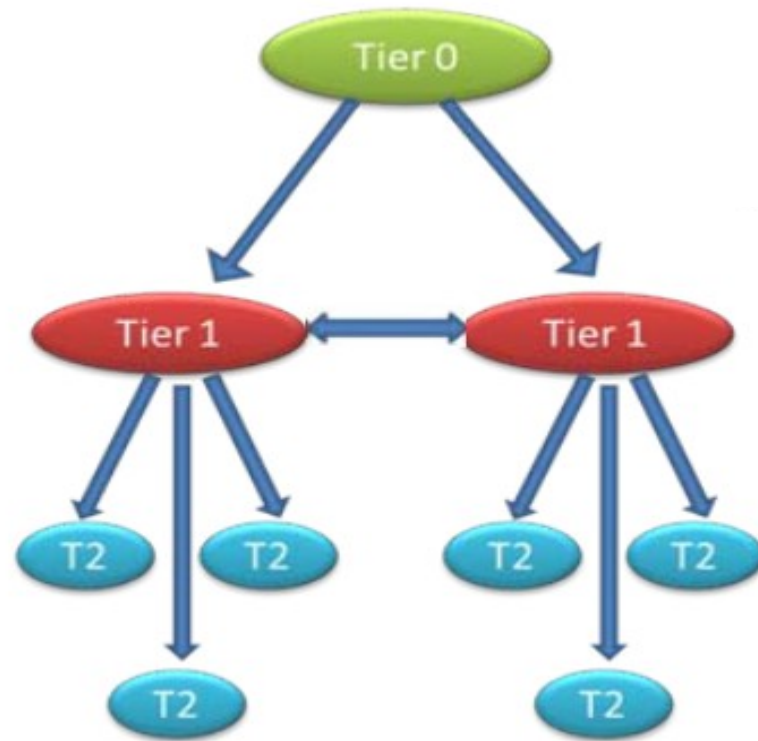
Source: <https://netstat.cern.ch/monitoring/network-statistics/ext/?q=LHCOPN&p=LHCOPN&mn=00-Total-Traffic&t=all>

LHCOPN evolution

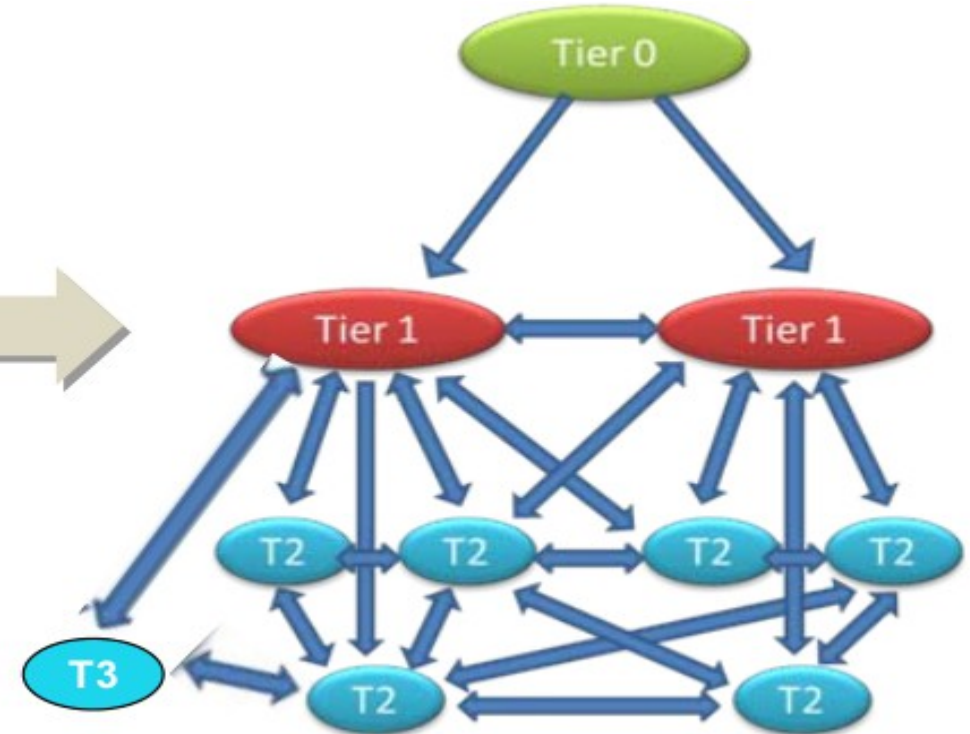
- The LHCOPN is kept as the main network to exchange data among Tier0 and Tier1s
- Tier1-Tier1 traffic and backup functionalities progressively moving to LHCONE (the Tier1/2/3 network)
- Links being upgraded to multiple 10Gbps and 100G

LHCONE - Tier1/2/3 network

Computing model evolution



Original MONARCH model



Model evolution

LHCONE 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

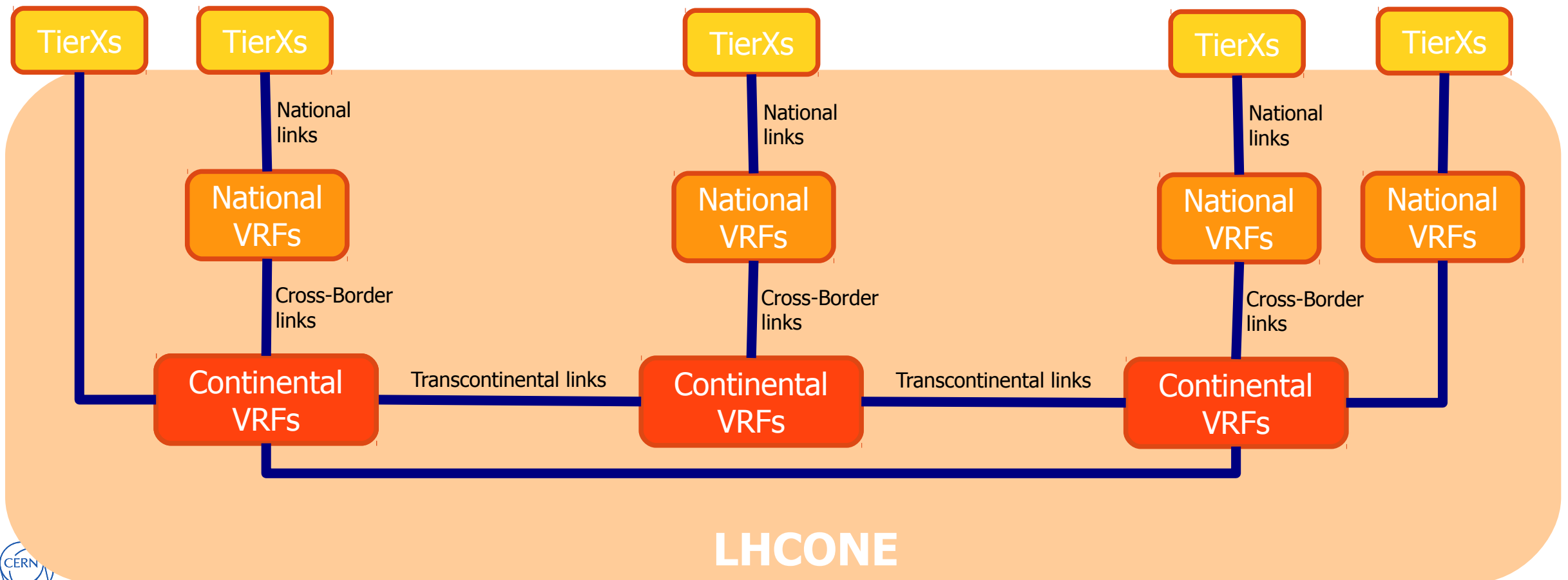
Closed network, only trusted traffic that can bypass perimeter firewalls

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:

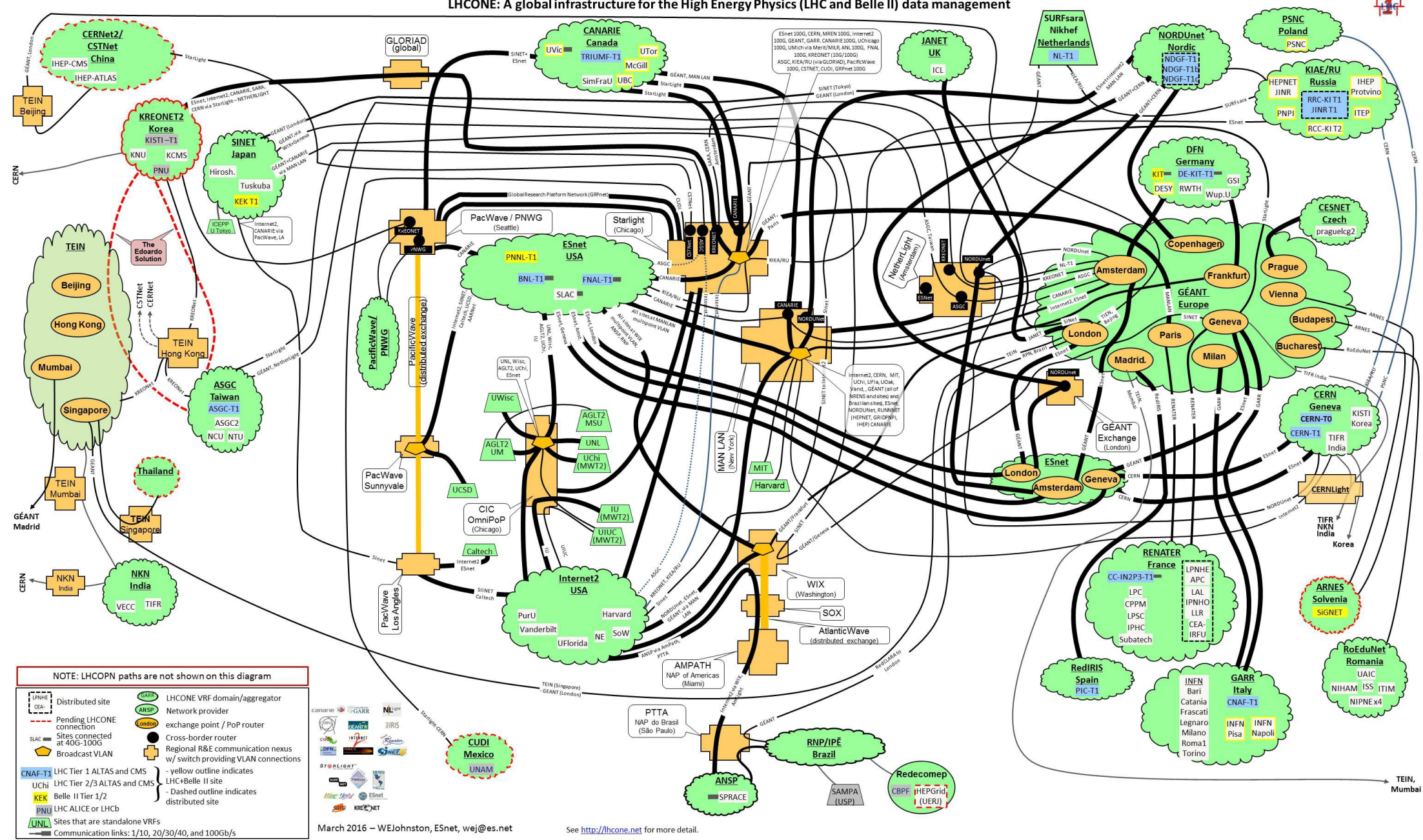
- 10 Tier1s
- 45 Tier2s

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

LHCONE: A global infrastructure for the High Energy Physics (LHC and Belle II) data management



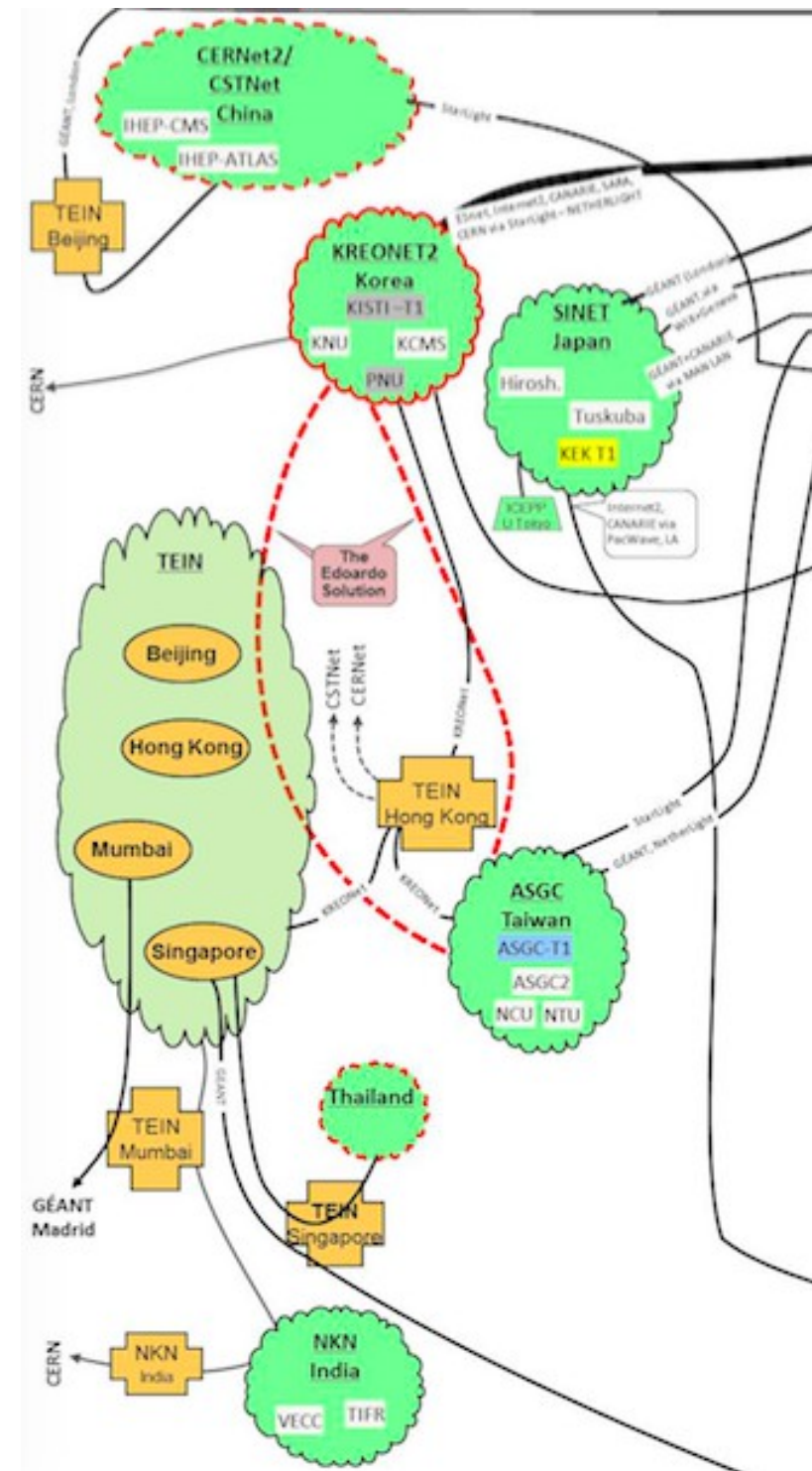
credits: Bill Johnston, ESnet

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

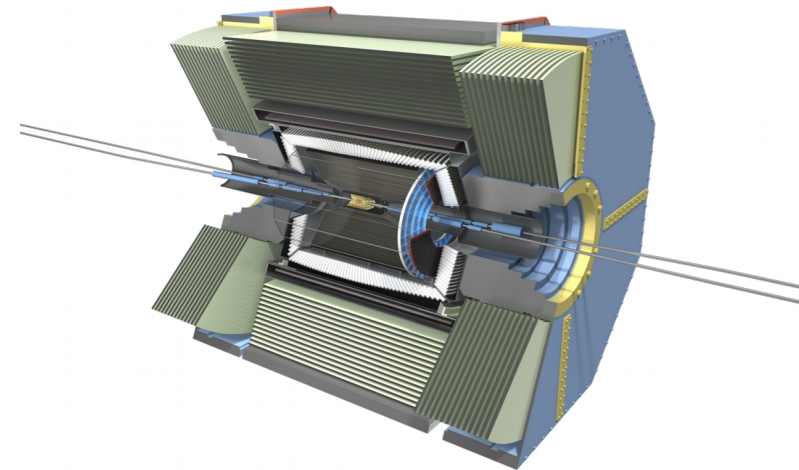
SINET (JP), CERnet/CSTnet (CN) and others are planning their own VRFs



Open to other HEP collaborations

The L3VPN is now used also by:

- **Belle II experiment**

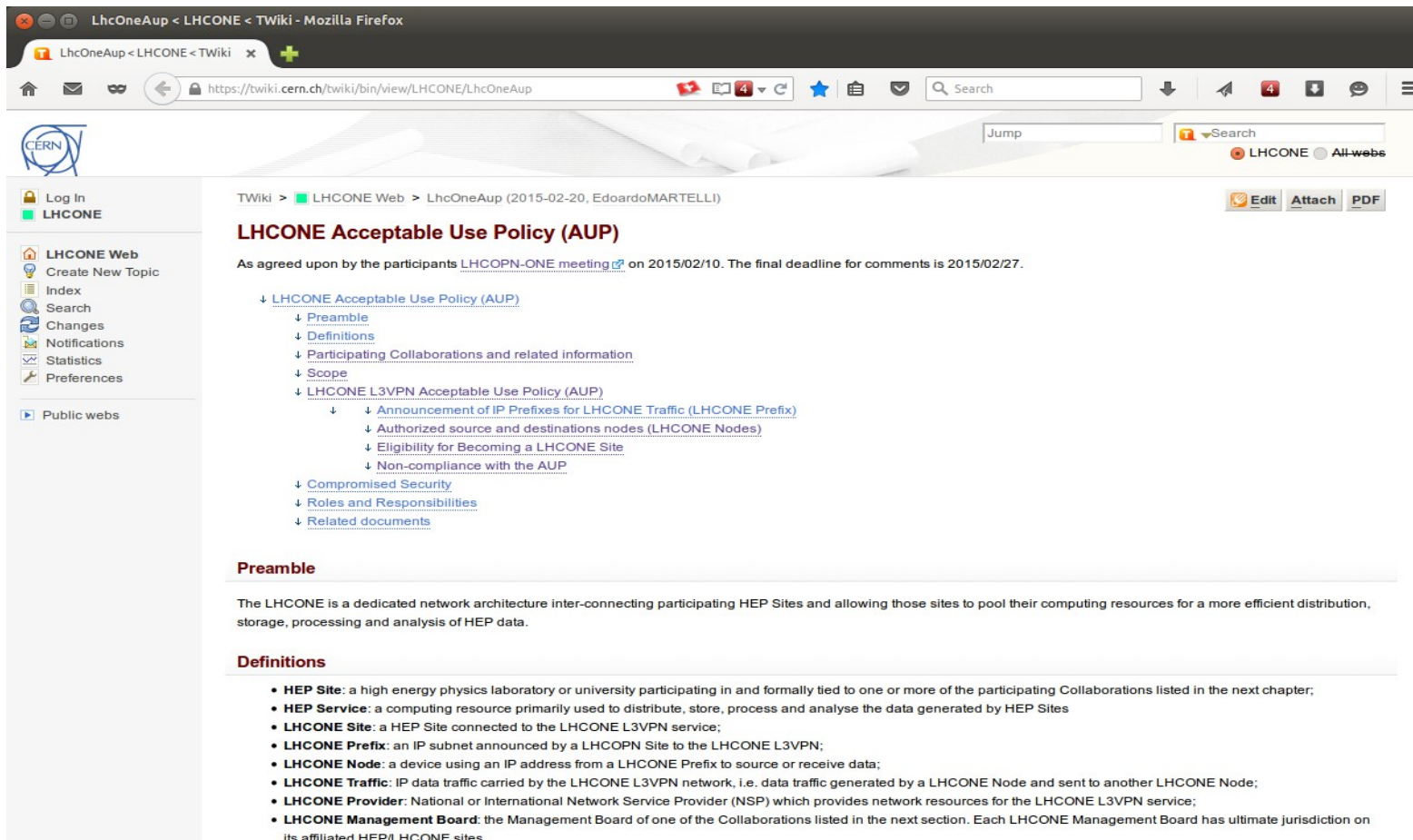


- **Pierre Auger Observatory**



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>)



The screenshot shows a web browser window displaying the LHCONE Acceptable Use Policy (AUP) page. The browser's address bar shows the URL <https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneAup>. The page header includes the CERN logo and navigation links. The main content area is titled "LHCONE Acceptable Use Policy (AUP)" and includes a preamble, definitions, and a list of participating collaborations. The page is structured with a left sidebar for navigation and a main content area for the policy text.

Log In
LHCONE

LHCONE Web
Create New Topic
Index
Search
Changes
Notifications
Statistics
Preferences

Public webs

Twiki > LHCONE Web > LhcOneAup (2015-02-20, EdoardoMARTELLI)

LHCONE Acceptable Use Policy (AUP)

As agreed upon by the participants [LHCOPN-ONE meeting](#) on 2015/02/10. The final deadline for comments is 2015/02/27.

- ↓ [LHCONE Acceptable Use Policy \(AUP\)](#)
 - ↓ [Preamble](#)
 - ↓ [Definitions](#)
 - ↓ [Participating Collaborations and related information](#)
 - ↓ [Scope](#)
 - ↓ [LHCONE L3VPN Acceptable Use Policy \(AUP\)](#)
 - ↓ [Announcement of IP Prefixes for LHCONE Traffic \(LHCONE Prefix\)](#)
 - ↓ [Authorized source and destinations nodes \(LHCONE Nodes\)](#)
 - ↓ [Eligibility for Becoming a LHCONE Site](#)
 - ↓ [Non-compliance with the AUP](#)
 - ↓ [Compromised Security](#)
 - ↓ [Roles and Responsibilities](#)
 - ↓ [Related documents](#)

Preamble

The LHCONE is a dedicated network architecture inter-connecting participating HEP Sites and allowing those sites to pool their computing resources for a more efficient distribution, storage, processing and analysis of HEP data.

Definitions

- **HEP Site:** a high energy physics laboratory or university participating in and formally tied to one or more of the participating Collaborations listed in the next chapter;
- **HEP Service:** a computing resource primarily used to distribute, store, process and analyse the data generated by HEP Sites
- **LHCONE Site:** a HEP Site connected to the LHCONE L3VPN service;
- **LHCONE Prefix:** an IP subnet announced by a LHCOPN Site to the LHCONE L3VPN;
- **LHCONE Node:** a device using an IP address from a LHCONE Prefix to source or receive data;
- **LHCONE Traffic:** IP data traffic carried by the LHCONE L3VPN network, i.e. data traffic generated by a LHCONE Node and sent to another LHCONE Node;
- **LHCONE Provider:** National or International Network Service Provider (NSP) which provides network resources for the LHCONE L3VPN service;
- **LHCONE Management Board:** the Management Board of one of the Collaborations listed in the next section. Each LHCONE Management Board has ultimate jurisdiction on its affiliated HEP/LHCONE sites.

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:

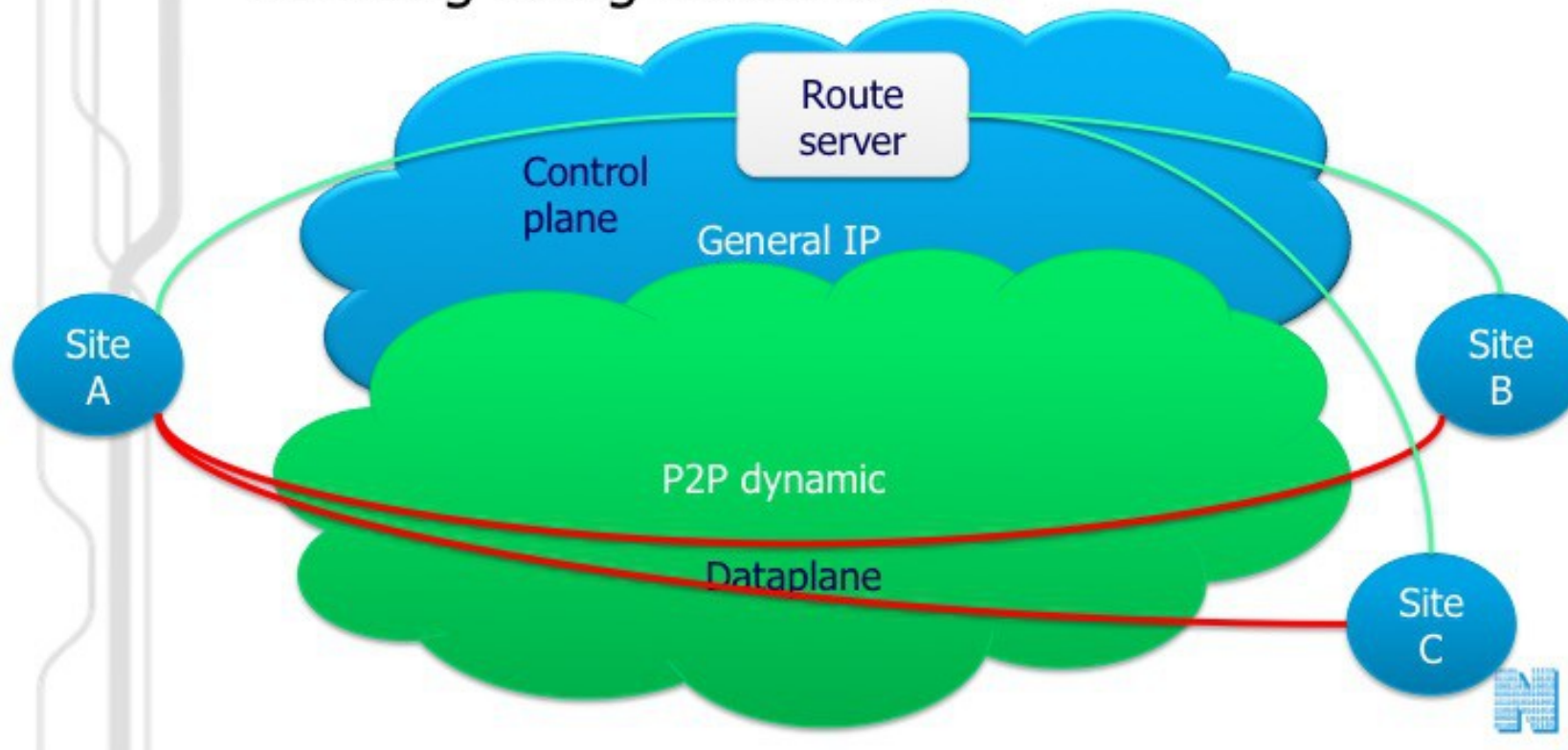
- multi-domain provisioning system
- intra-TierX connectivity
- TierX-TierY routing
- integration with WLCG software
- bandwidth allocation and protection

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

NORDUnet
Nordic Gateway for Research & Education

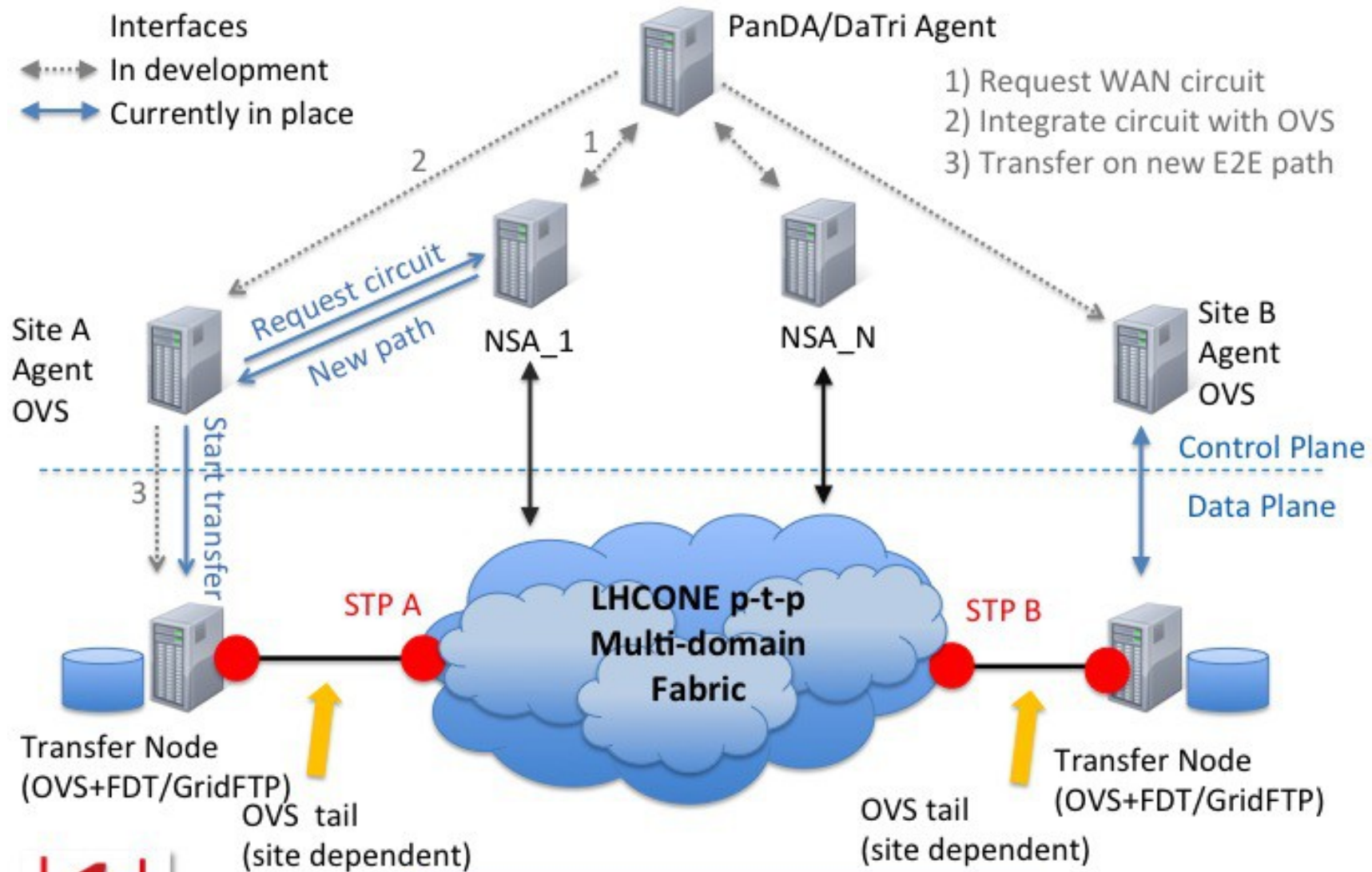
Route server

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



Solution 2 – based on SDN Openflow

Diagram of Possible Future SDN Dev-Ops Testbed



March 14, 2016

Original Slide from Ramiro/Azher, Caltech

7

LHCONE perfSONAR service

perfSONAR

- software framework for active and passive network probing
- developed by Internet2, ESnet, GEANT and others

perfSONAR

<http://www.perfsonar.net/>

The screenshot shows the perfSONAR web interface for a node named 'pS-Performance Node For CERN In Geneva, Meyrin, CH'. The browser address bar shows the URL <https://perfsonar-bw.cern.ch/toolkit/>. The interface is divided into a left sidebar and a main content area.

Left Sidebar:

- perfSONAR Services:** Local Services, Global Services
- View Test Results:** Throughput / Latency Graphs, Traceroute Graphs
- External Tools:** Reverse Ping, Reverse Traceroute, Reverse Tracetest, Traceroute Visualization
- Toolkit Administration:** Configure Tests, Administrative Information, Enabled Services, NTP, perfSONAR Logs, BWCTL Log Analysis, OWAMP Log Analysis, NDT Log Analysis
- Performance Toolkit:** Configuration Manual, Frequently Asked Questions, About, Credits

Main Content Area:

pS-Performance Node For CERN In Geneva, Meyrin, CH

Host Information

Organization Name	CERN
City, State, Country	Geneva, Meyrin, CH
Postal Code	1217
Latitude, Longitude	46.232498, 6.04593
Administrator Name	Stefan Stancu
Administrator Email	stefan.stancu@cern.ch

Communities This Host Participates In

HEPIX IPv6 testbed LHCOPN WLCG

Host Status

Primary Address	perfsonar-bw.cern.ch
MTU	9000
NTP Status	Synced
Memory	8GB
Globally registered	No

Services Offered

Bandwidth Test Controller (BWCTL)	Running
• tcp://perfsonar-bw.cern.ch:4823 • Testing Ports: 6001-6200(peer), 5001-5300(iperf), 5301-5600(nuttcp), 5601-5900(owamp), 5001-5900(test)	
Regular Testing	Running
One-Way Ping Service (OWAMP)	Disabled
• tcp://perfsonar-bw.cern.ch:861 • Testing Ports: 8760-9960(test)	
Network Diagnostic Tester (NDT)	Disabled
• http://perfsonar-bw.cern.ch:7123/ • tcp://perfsonar-bw.cern.ch:3001	
esmond Measurement Archive	Running
• http://perfsonar-bw.cern.ch/esmond/perfsonar/archive/	
Network Path and Application Diagnosis (NPAD)	Disabled
• http://perfsonar-bw.cern.ch:8000/ • tcp://perfsonar-bw.cern.ch:8001	

LHCONE perfSONAR service

LHCONE Network monitoring infrastructure

Probes installed at:

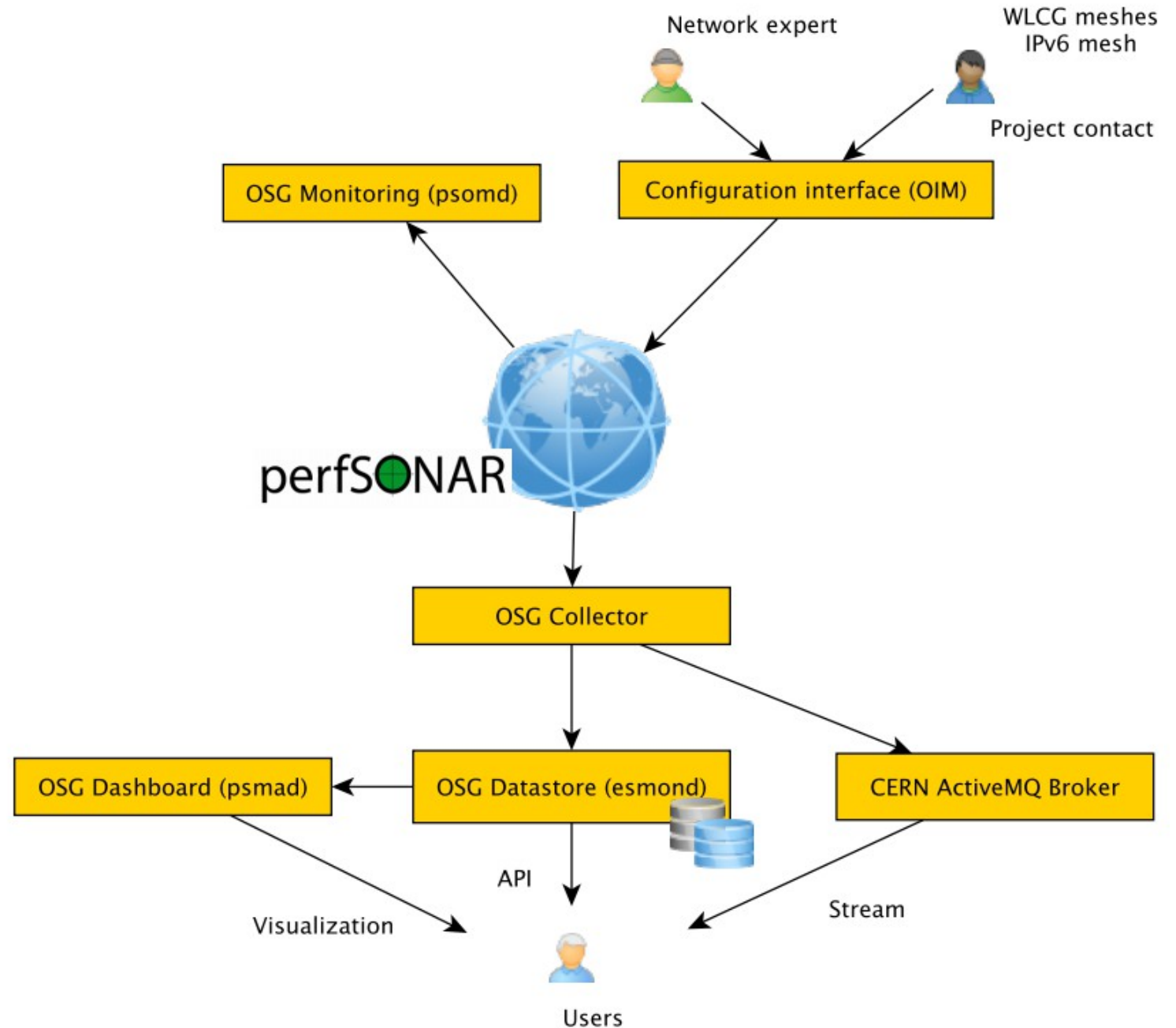
- VRFs interconnecting points
- Sites

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

LHCONe Mesh Config - TCP BWCTL Test Between L

■ Throughput ≥ 900 Mbps
 ■ Throughput < 900 Mbps
 ■ Throughput ≤ 500 M



LHCONe Mesh Config - OWAMP Test Between LHCO

■ Loss rate is ≤ 0
■ Loss rate is ≥ 0
■ Loss rate is ≥ 0.01
■ Unable



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?

