Networks for High Energy Physics: LHCOPN and LHCONE

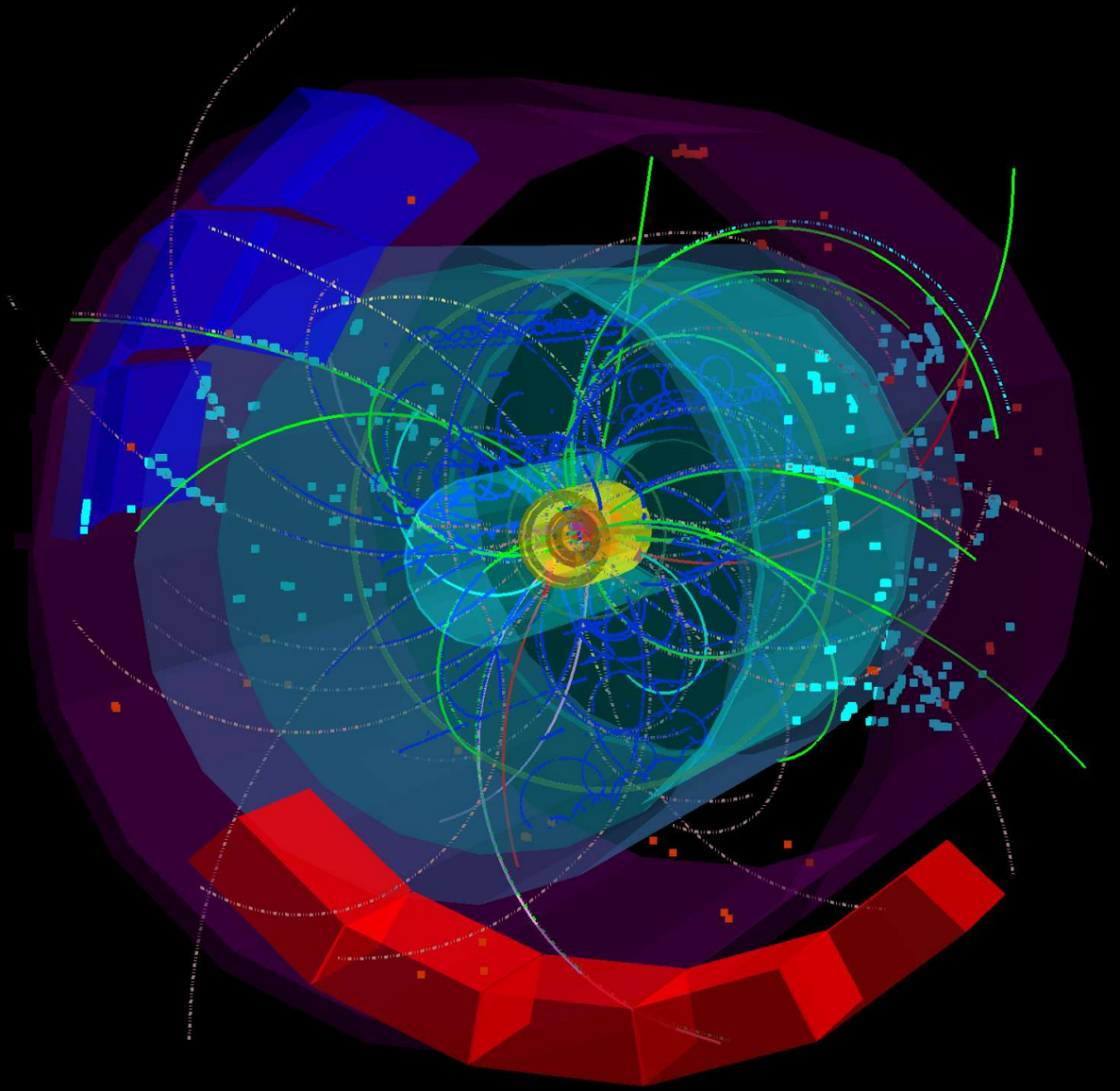
AFAD2018
Daejeon, South Korea

29th January 2018

edoardo.martelli@cern.ch

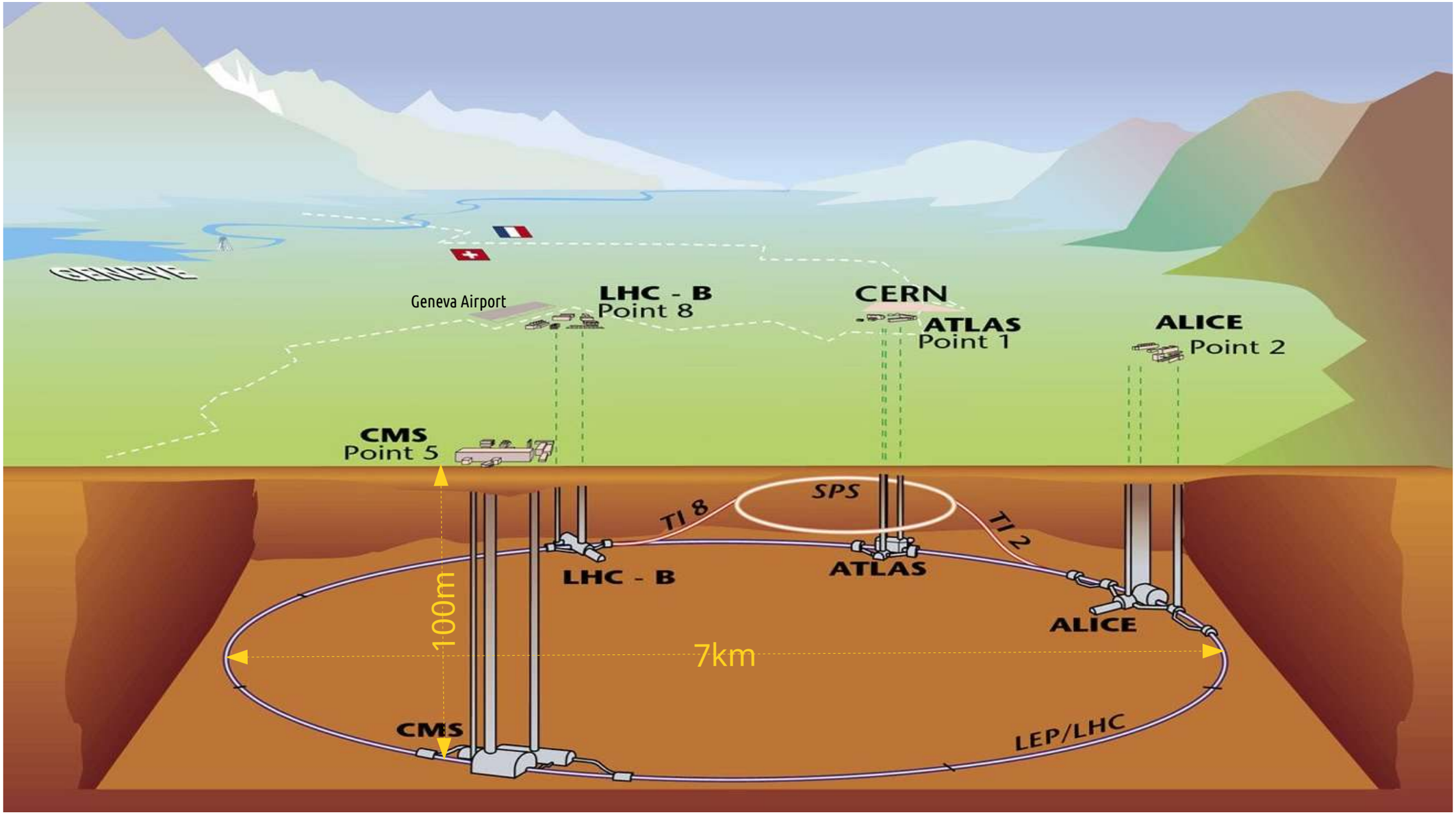
Agenda

- LHC Data Challenge
- WLCG
- LHCOPN
- LHCONE
- Future developments



LHC Data Challenge

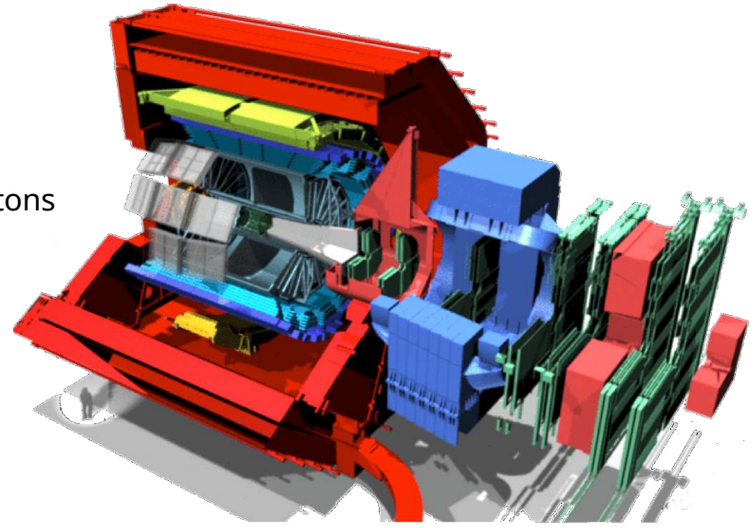
LHC facility



LHC major experiments

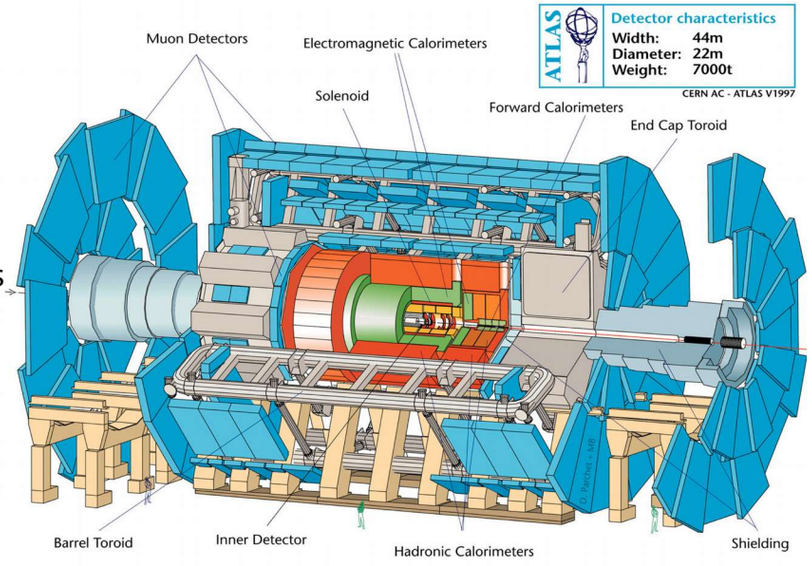
ALICE

Weight: 10,000 tons
Length: 26 m
Diameter 16 m



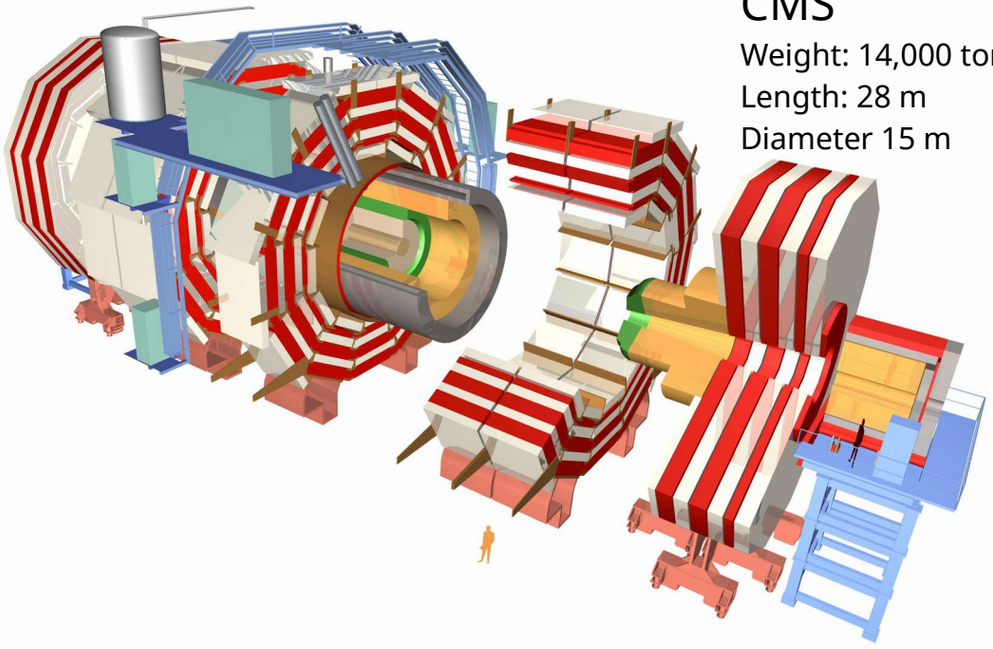
ATLAS

Weight: 7,000 tons
Length: 44 m
Diameter 22 m



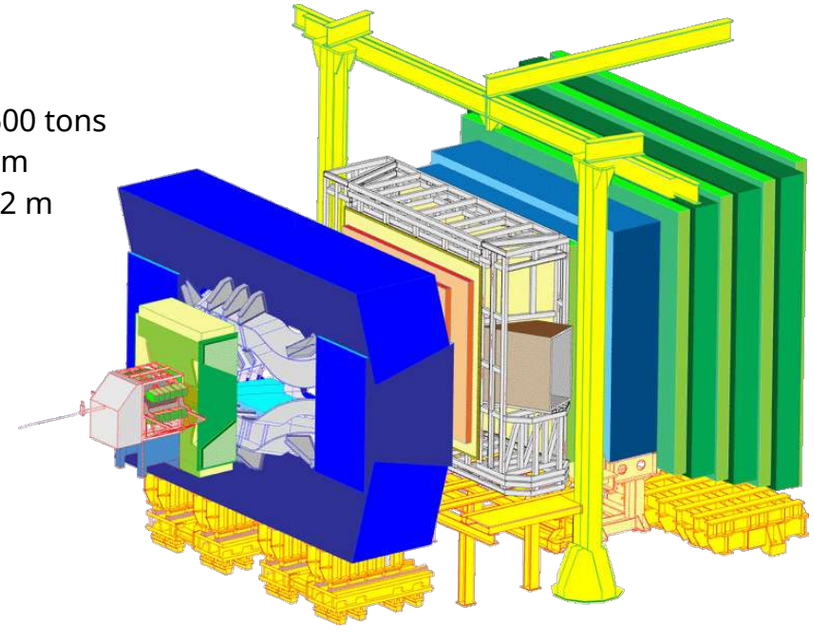
CMS

Weight: 14,000 tons
Length: 28 m
Diameter 15 m



LHCb

Weight: 5,600 tons
Length: 21 m
Diameter 12 m

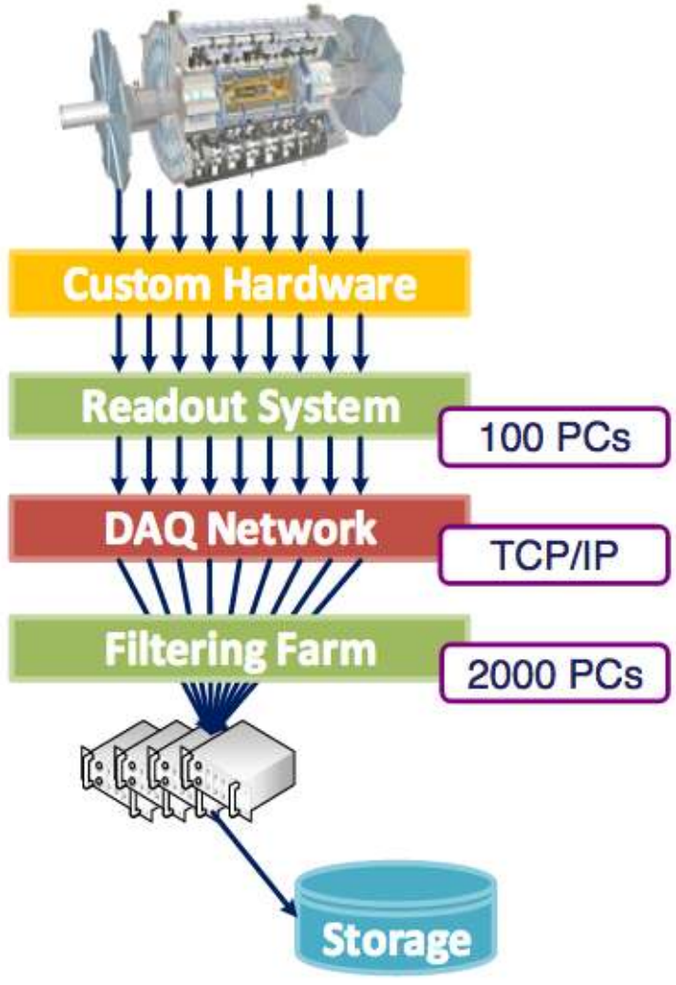


Experiments' Data Acquisition

40Mhz

100Khz

1Khz



Data Challenge

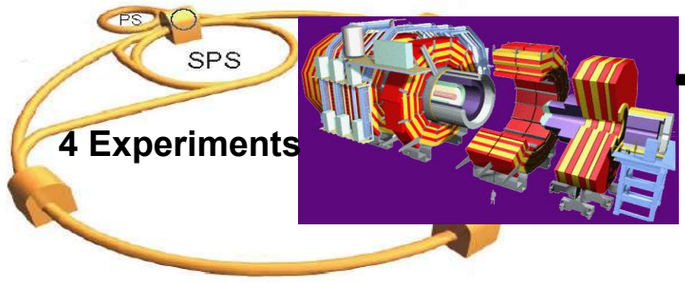
40 million collisions per second

After filtering, 1000 collisions of interest per second

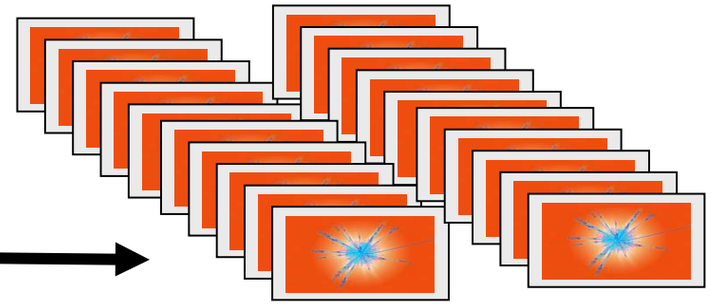
>10^{10} collisions recorded each year

> 40 Petabytes/year of data!!

Data's journey

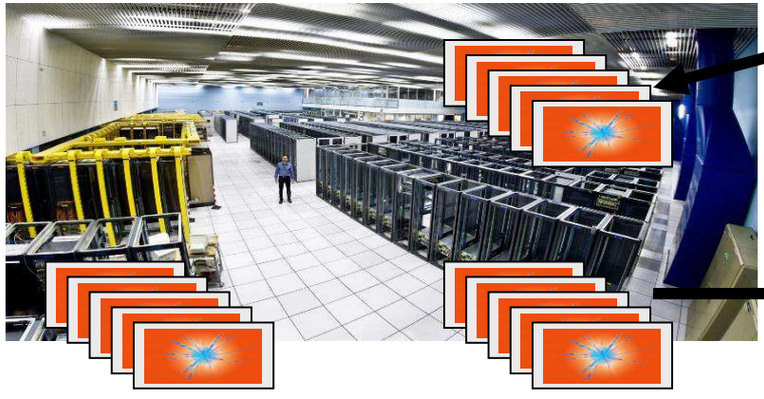


24 Pbit/s

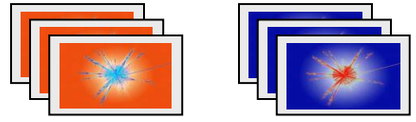


Filter and first selection

200 Gbit/s
to the CERN computer center



Create sub-samples



10 Tbit/s
Distributed + local

World-Wide Analysis

Physics
Explanation of nature

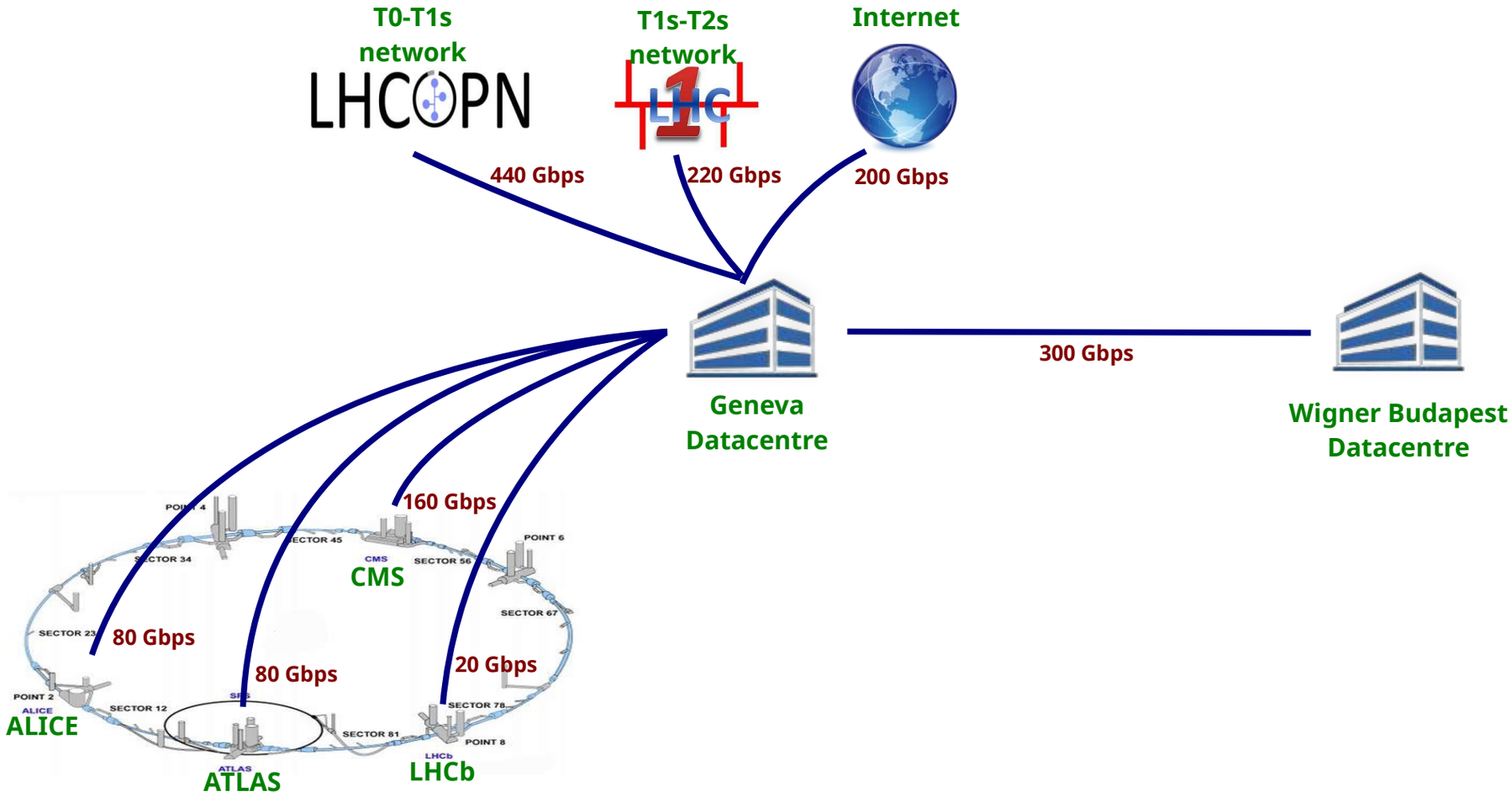
200 Gbit/s

40 Gbit/s

Store on disk and tape

Export copies

Connectivity for Experiments' data



WLCG

Worldwide LHC Computing Grid

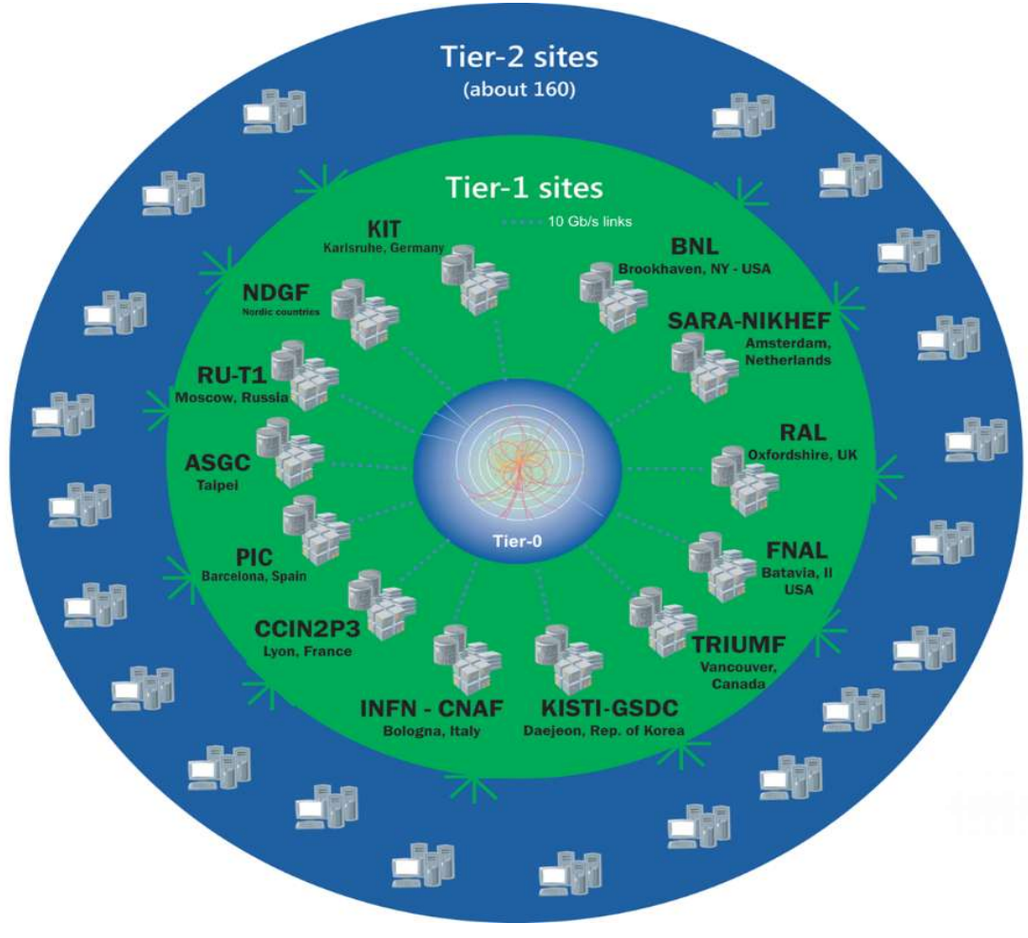
WLCG

The **Worldwide LHC Computing Grid**:
distributed computing infrastructure for LHC experiments

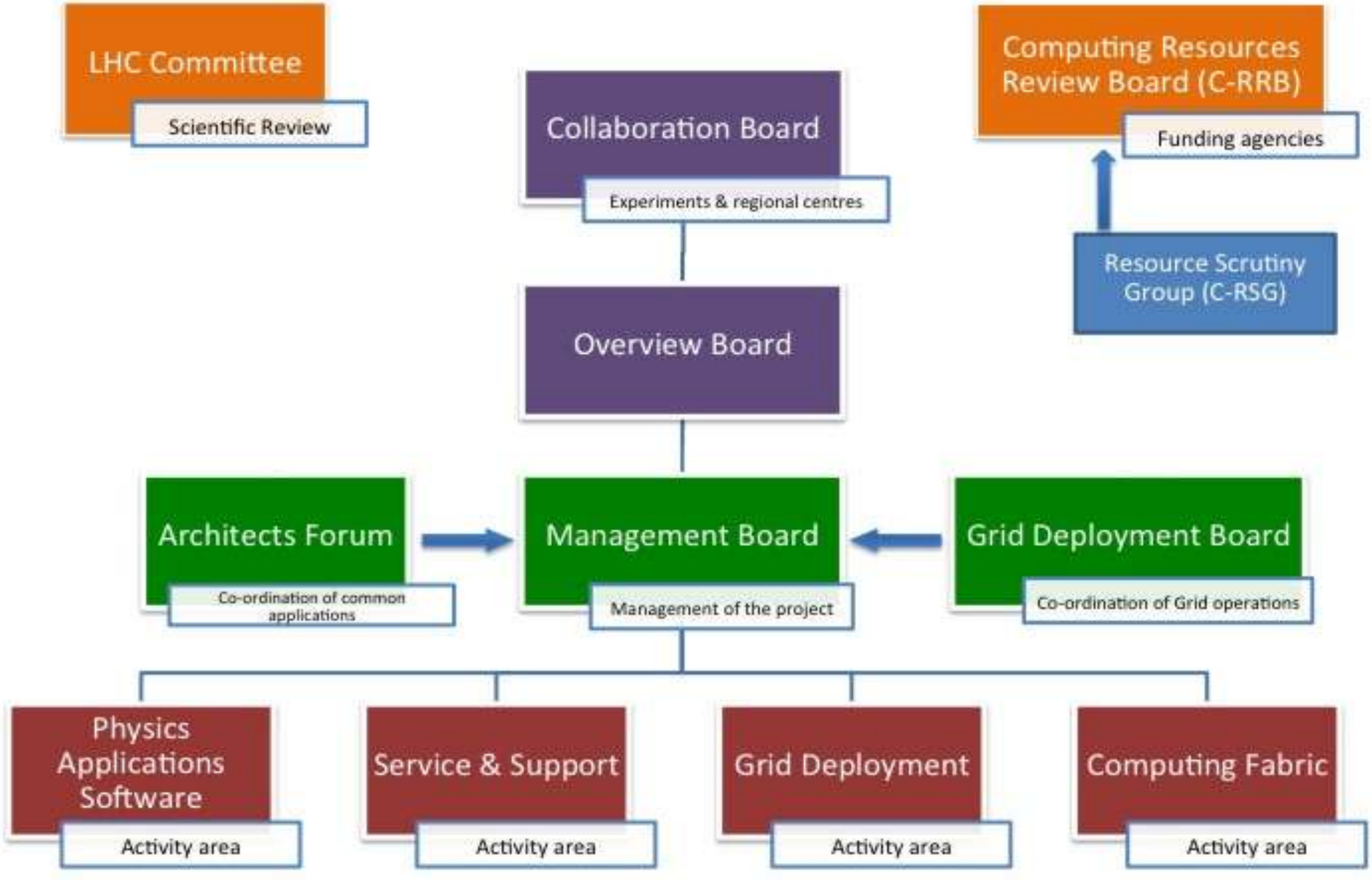
Tier-0 (CERN):
data recording, reconstruction and distribution

Tier-1:
permanent storage, re-processing, analysis

Tier-2:
Simulation, end-user analysis

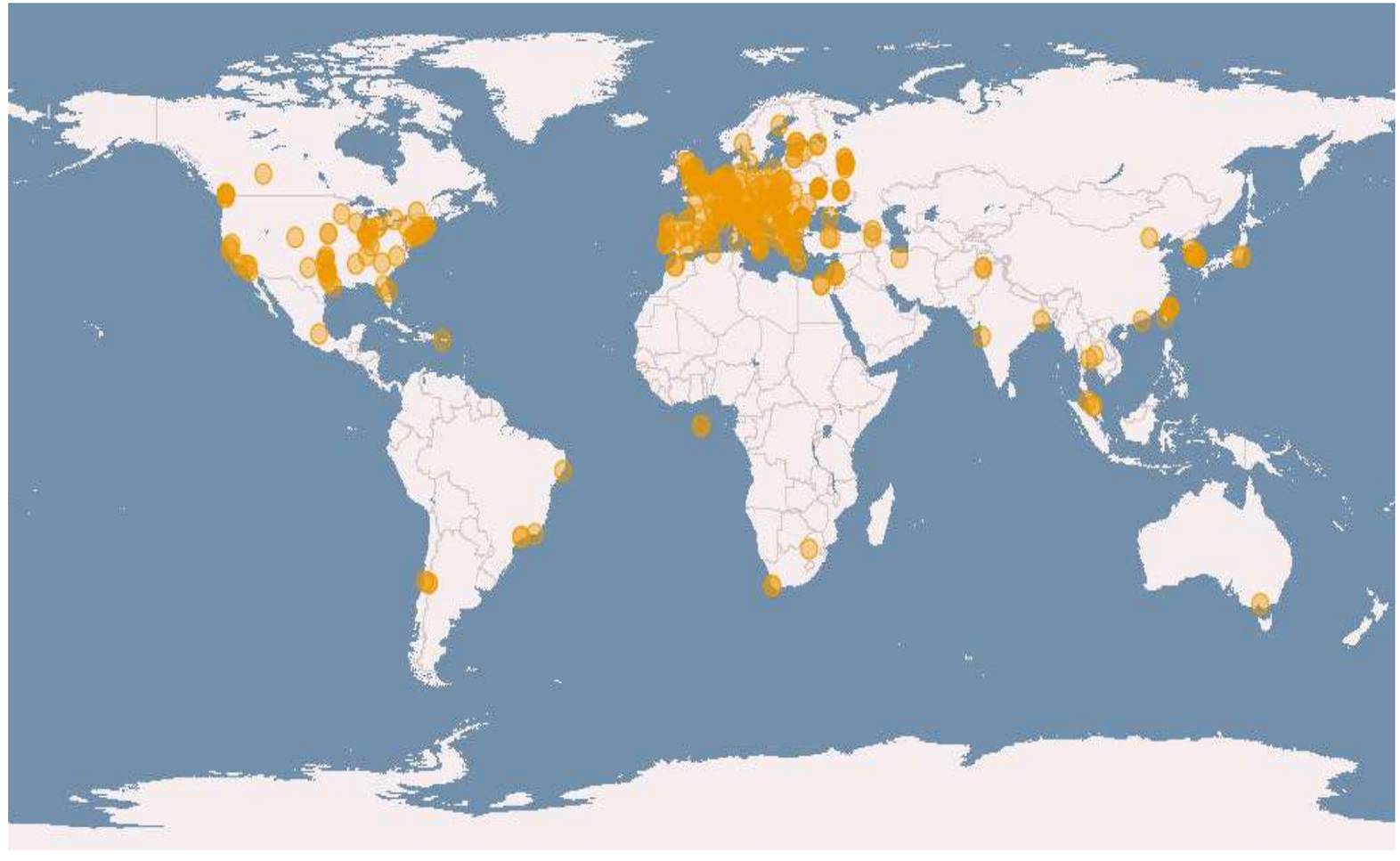


WLCG organization

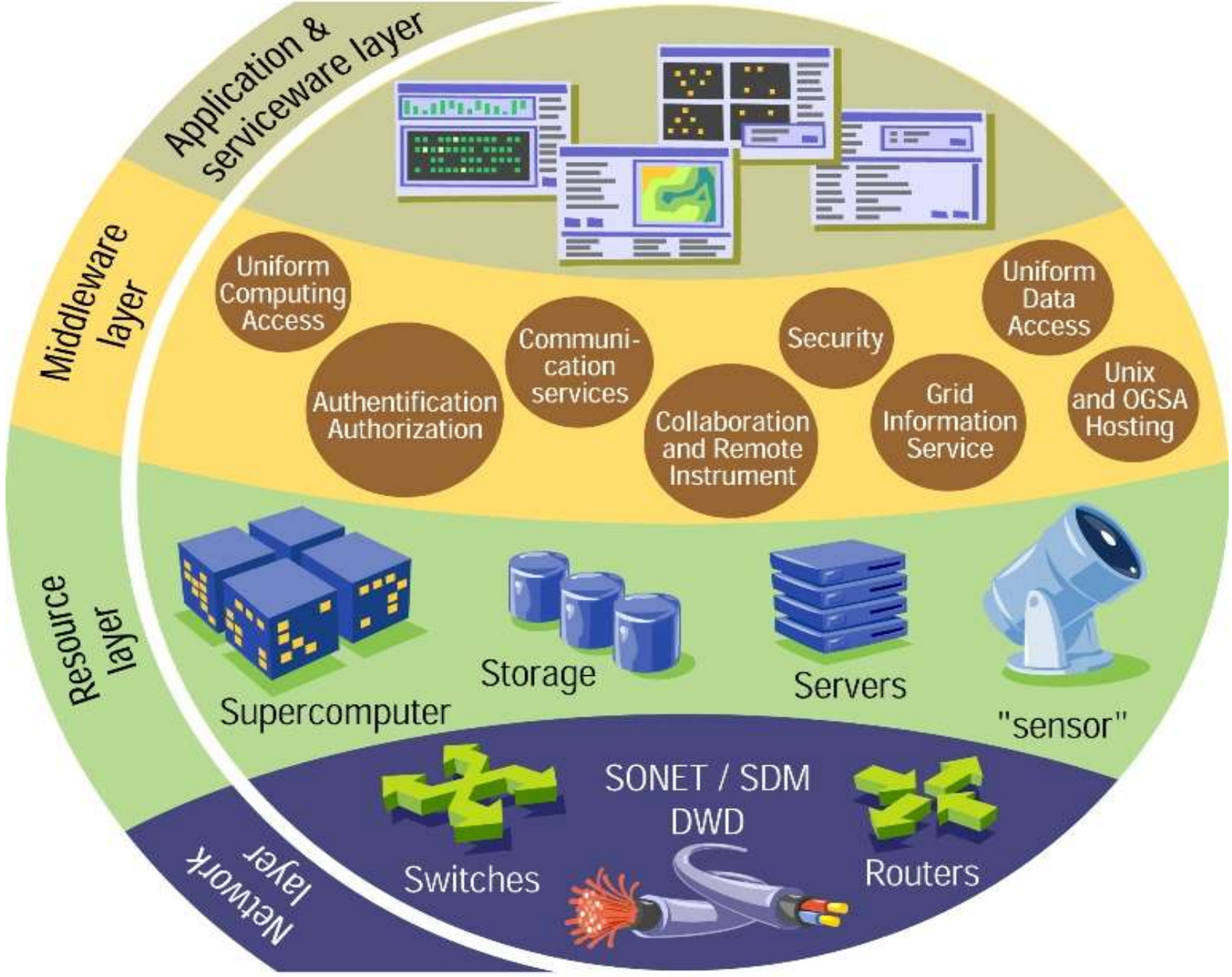


WLCG resources

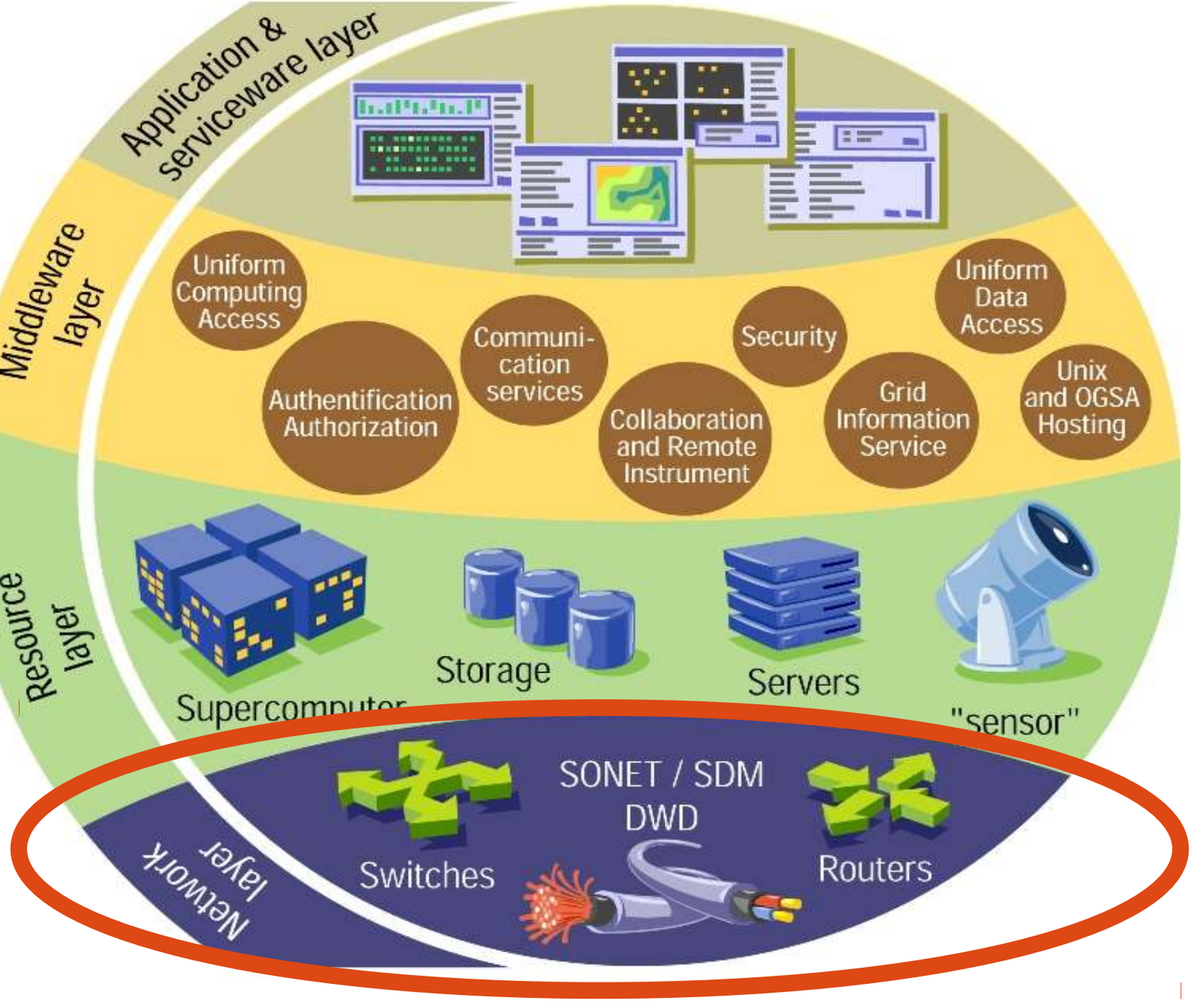
- 1 Tier0 (CERN)
- 13 Tier1s
- ~170 Tier2s
- >300 Tier3s worldwide
- ~350,000 cores
- ~500PB of disk space
- 2 millions jobs per day



The building blocks of the Grid



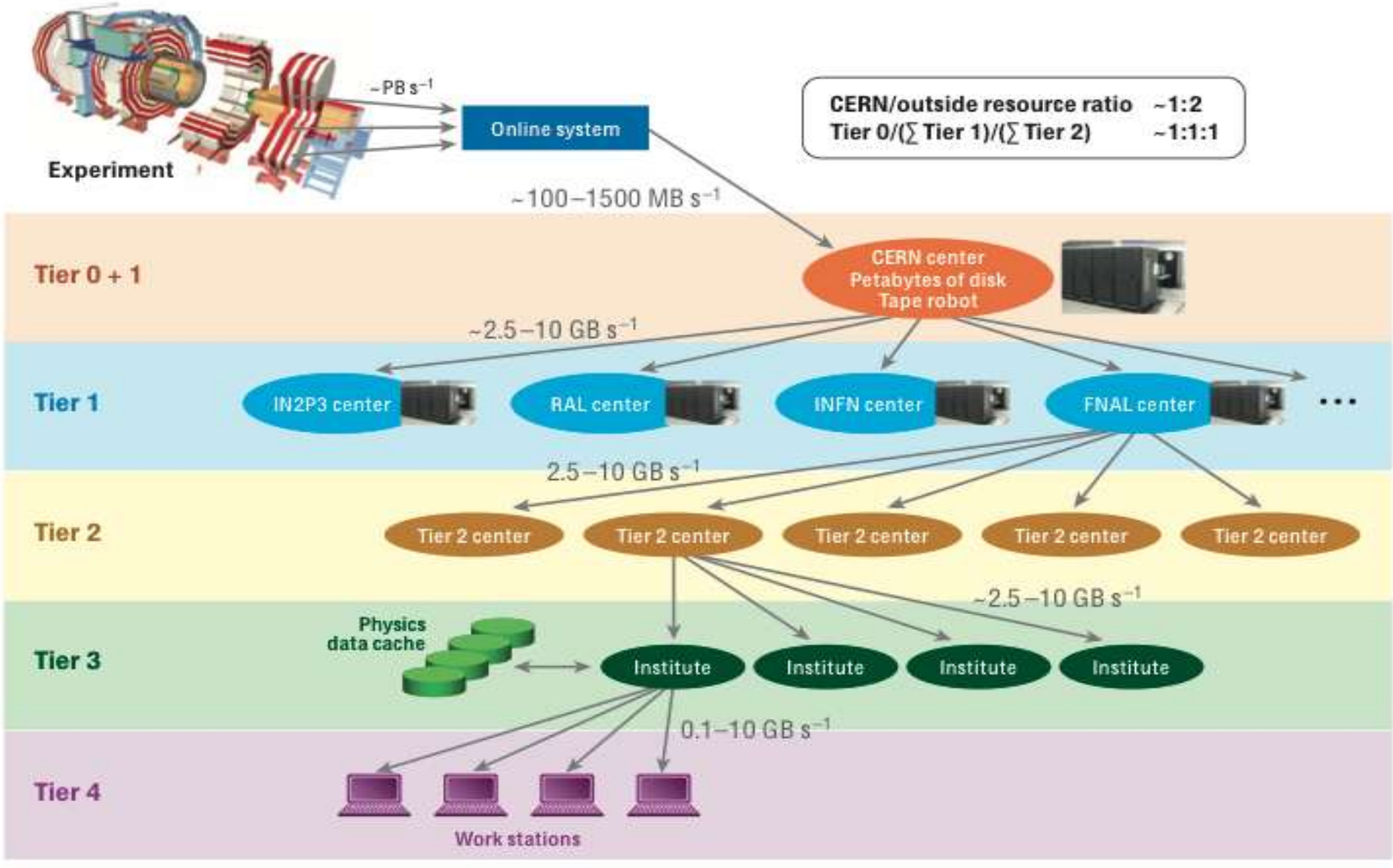
The building blocks of the Grid



LHCOPN

T0-T1s network

Original Computing Model



14 Tier1s all around the world

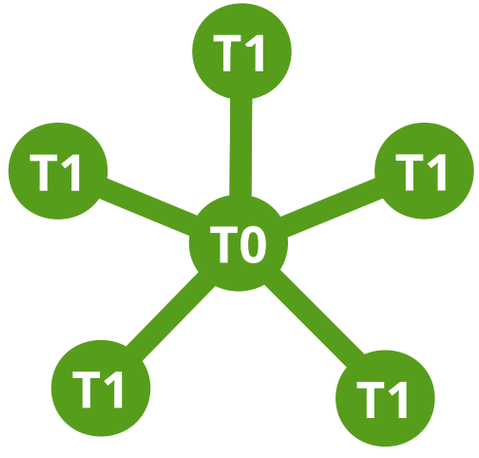
LHCOPN

Private network connecting Tier0 and Tier1s

- Dedicated to LHC data transfers and analysis

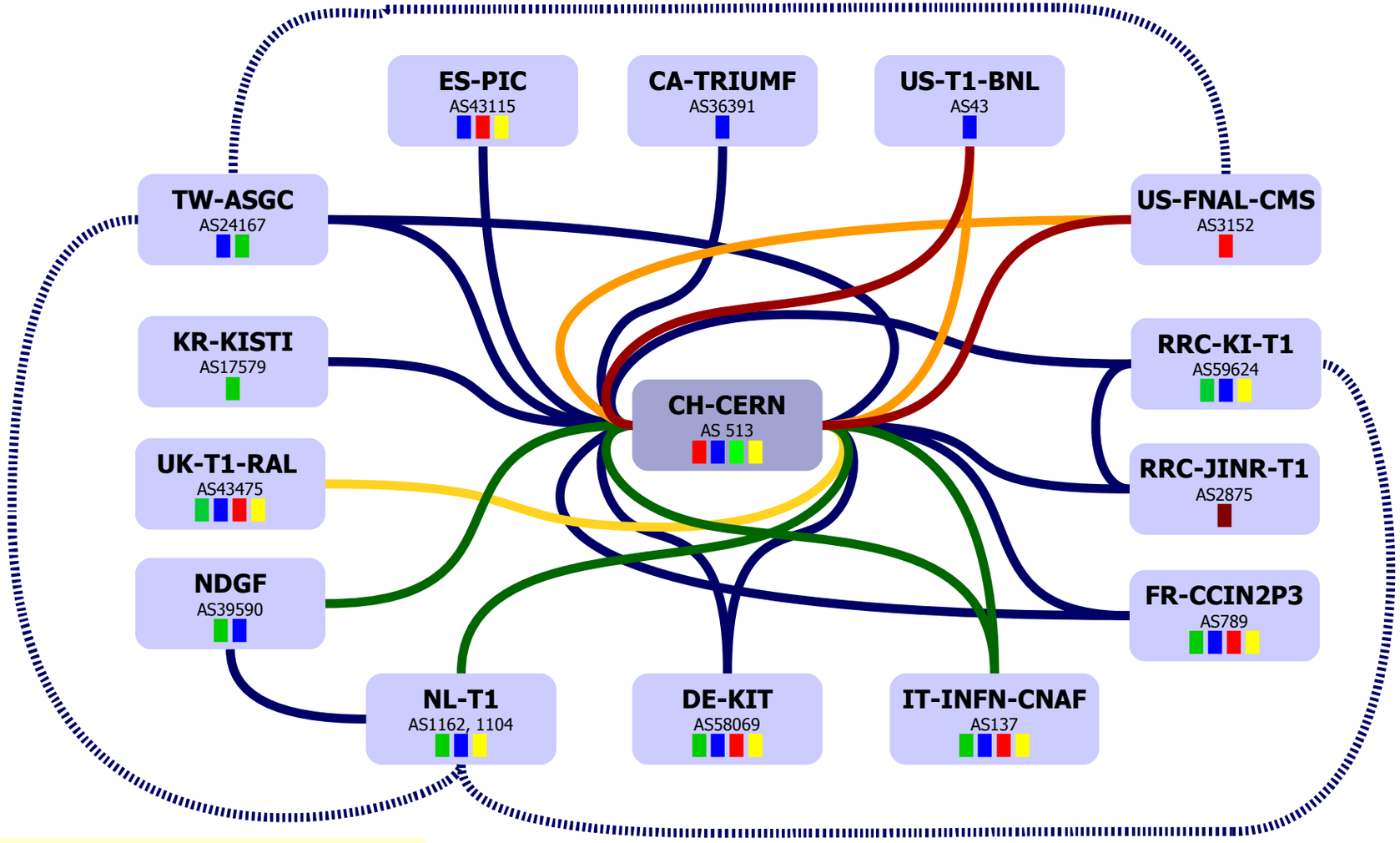
- Secured: only declared IP prefixes can exchange traffic

- Advanced routing: communities for traffic engineering, load balancing.



LHCOPN

LHCOPN



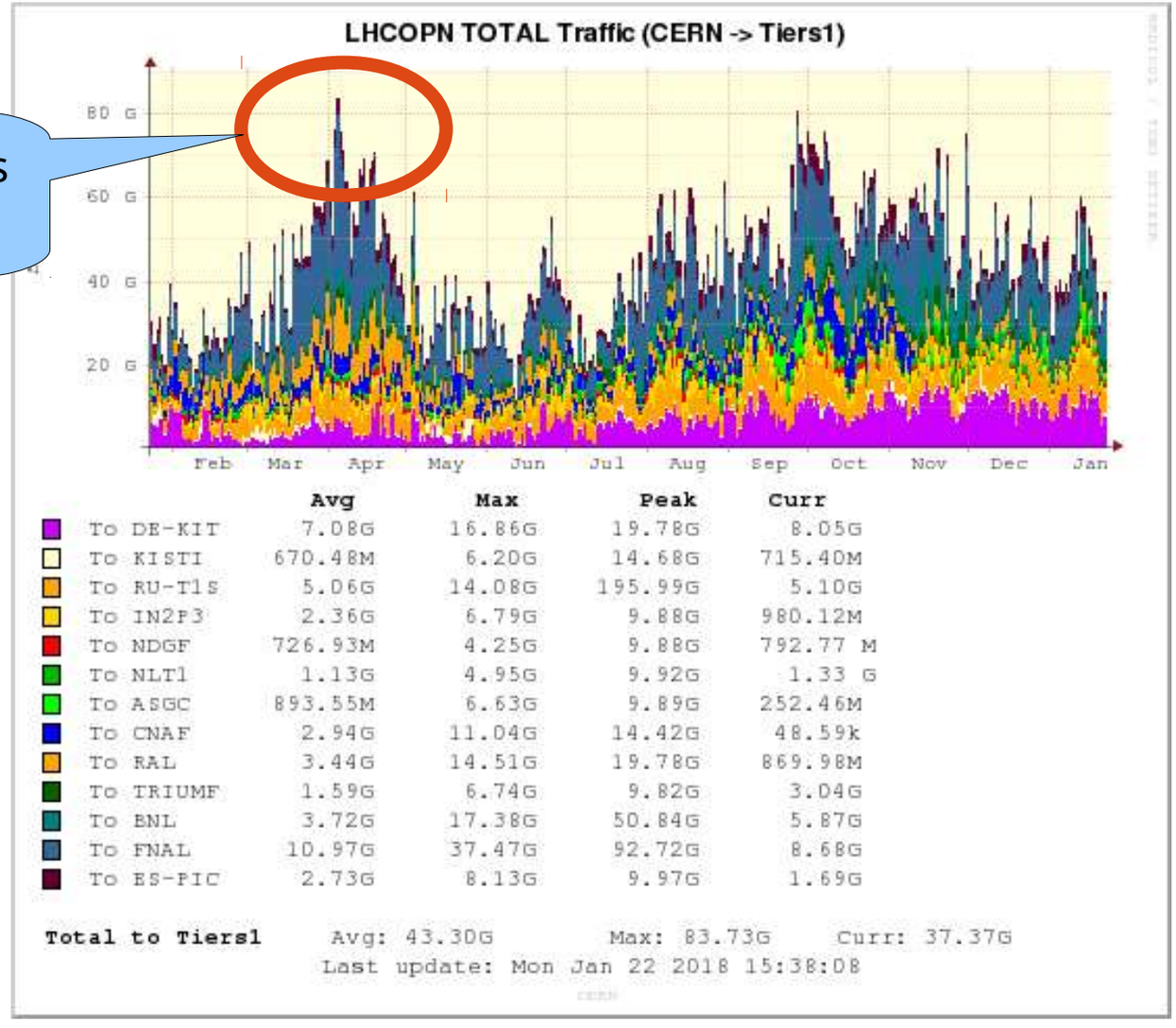
https://twiki.cern.ch/twiki/bin/view/LHCOPN/OverallNetworkMaps

Numbers

- 14 Tier1s + 1 Tier0
- 12 countries in 3 continents
- Dual stack IPv4-IPv6
- 440 Gbps to the Tier0
- Moved ~160 PB in the last year

LHCOPN Traffic (last 12 months)

Peaking at 80Gbps
Average 43Gbps



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

Latest developments

Traffic volume constantly growing
- already increased of 200% since the beginning of Run2

Tier1s are considering to upgrade links to 100Gbs
- already done for US sites
- waiting deployment of new network router at CERN for affordable 100G interfaces

Almost fully dual-stack IPv4-IPv6
- 13 Tier1s and the Tier0 peering over IPv6
- dual-stack perfSONAR installed in all of them

LHCOPN

Plans

Deploy more 100G links
- now more cost effective than multiple 10G

Complete IPv6 deployment:
- connect all Tier1s
- use IPv6 for production data transfers

LHCOPN

LHCOPN meetings

First meeting on January 2005 at SARA, Amsterdam

- Tier0/1 network meeting chaired by David Foster and Kors Bos

Set design principle for connectivity of WLCG Tier0 and Tier1s

- first Tier1s connected with 10G links already in 2005

Community regularly meeting every 6 months since

Forum to discuss operations and progresses

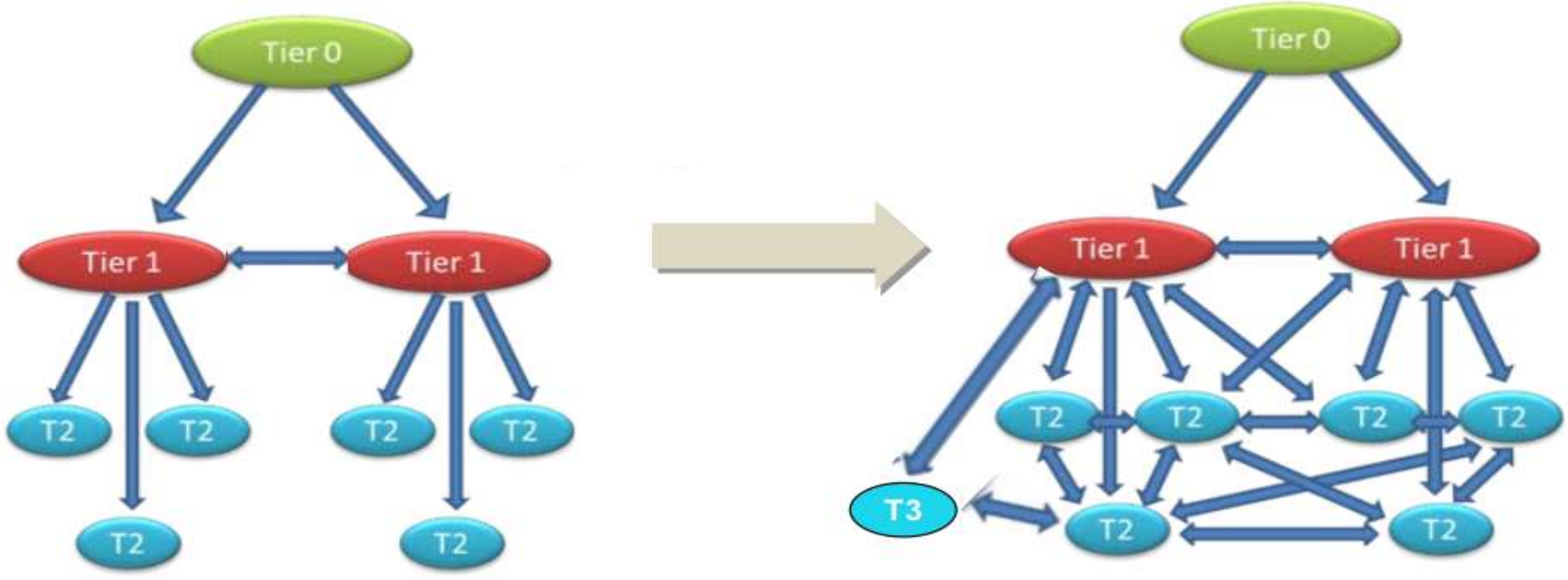
- for WLCG site operators, LHC Experiments, R&E network operators

LHCOPN

LHCONE

T1s-T2s network

Computing Model evolution for Run2



Original MONARCH model

Model evolution

LHCONE

- Network serving HEP sites according to their needs and allowing them to grow
- Sharing the cost of expensive resources for common benefit
- Traffic separation: no clash with other data transfers, resource allocated for and funded by the HEP community
- Trusted peers: common security policies

Three Services

- L3VPN
- perfSONAR monitoring
- R&D: circuits on demand, DTNs

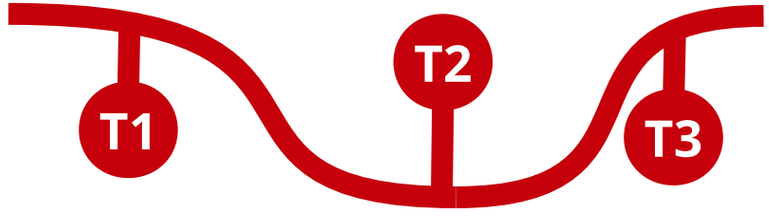
LHCONE L3VPN service

Layer3 (routed) Virtual Private Network

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

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

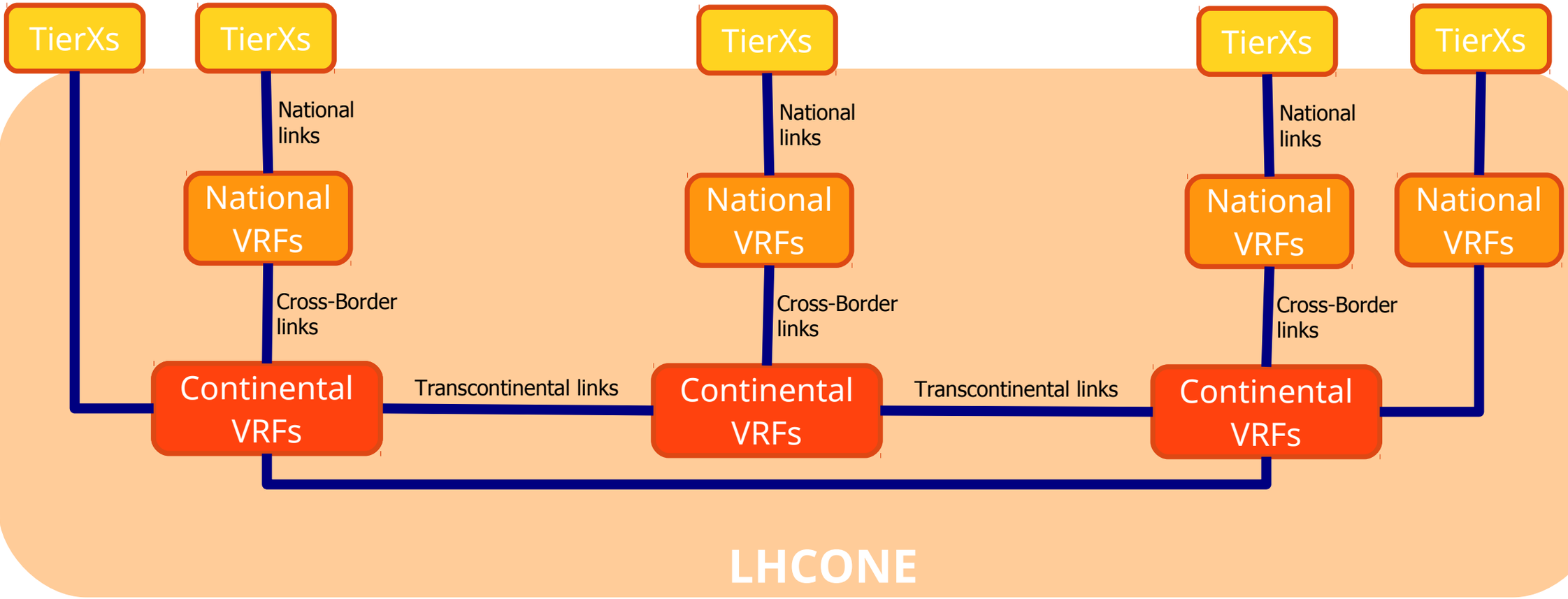
Trusted traffic that can bypass slow perimeter firewalls



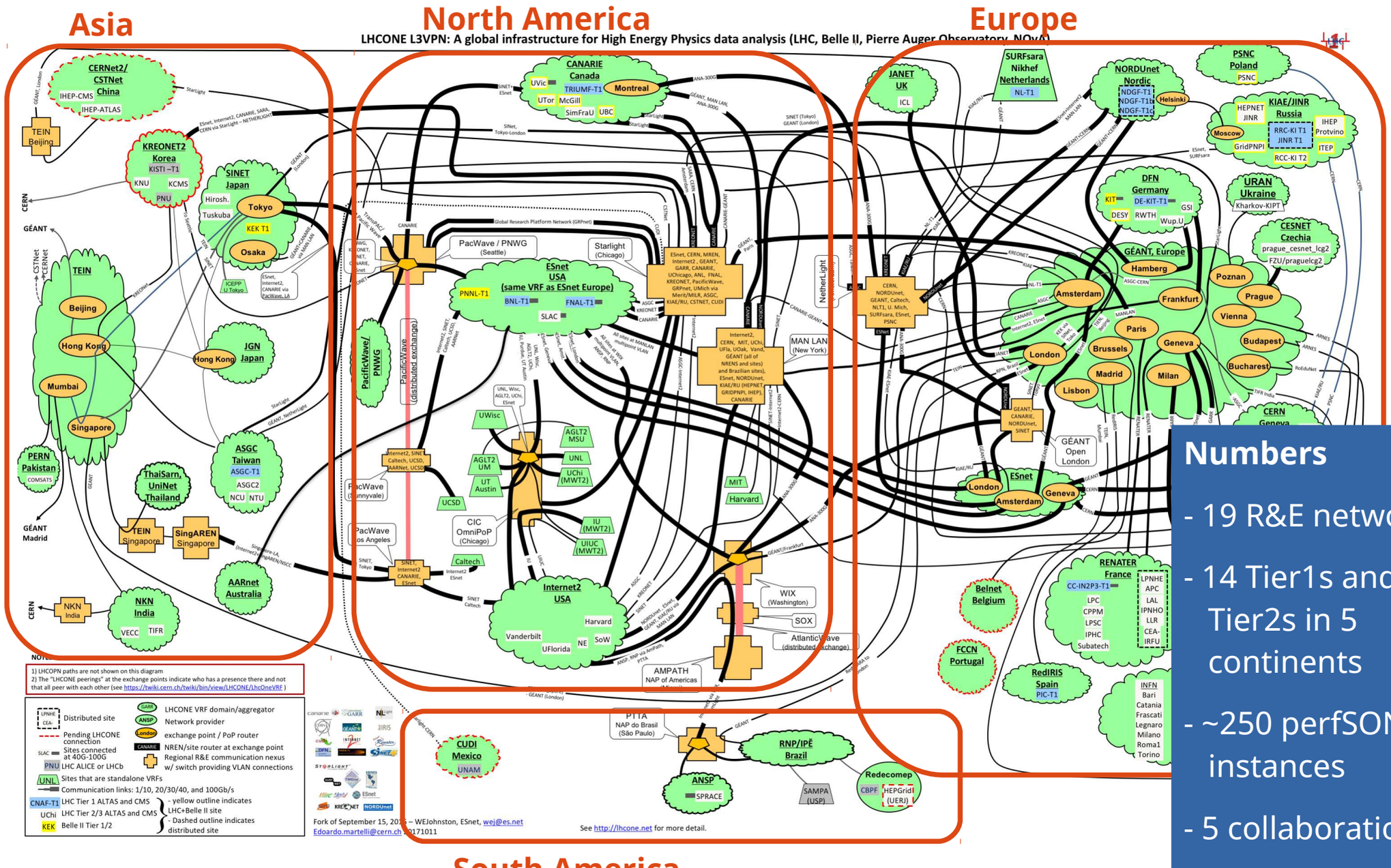
Architecture

Network dedicated to Tier1/2/3s

Distributed L3VPN, dedicated to HEP data



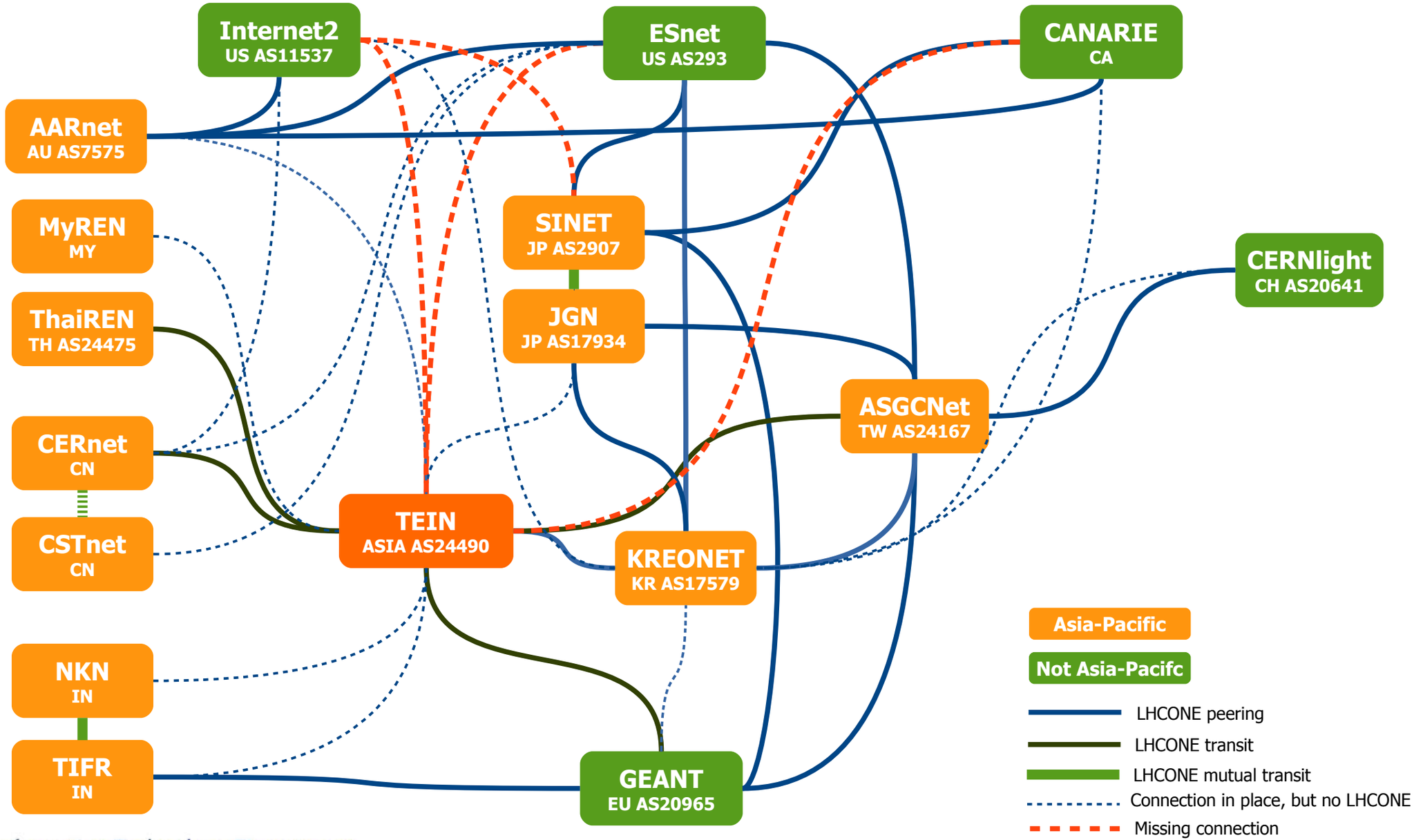
L3VPN Current topology



Numbers

- 19 R&E networks
- 14 Tier1s and ~66 Tier2s in 5 continents
- ~250 perfSONAR instances
- 5 collaborations

Asia-Pacific VRFs after the ATCF3



Open to other HEP collaborations

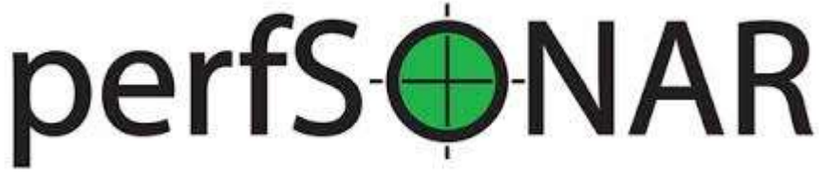
Other High Energy Physics projects can use LHCONE

LHCONE perfSONAR monitoring service

- LHCONE Network monitoring infrastructure based on perfSONAR monitoring suite

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

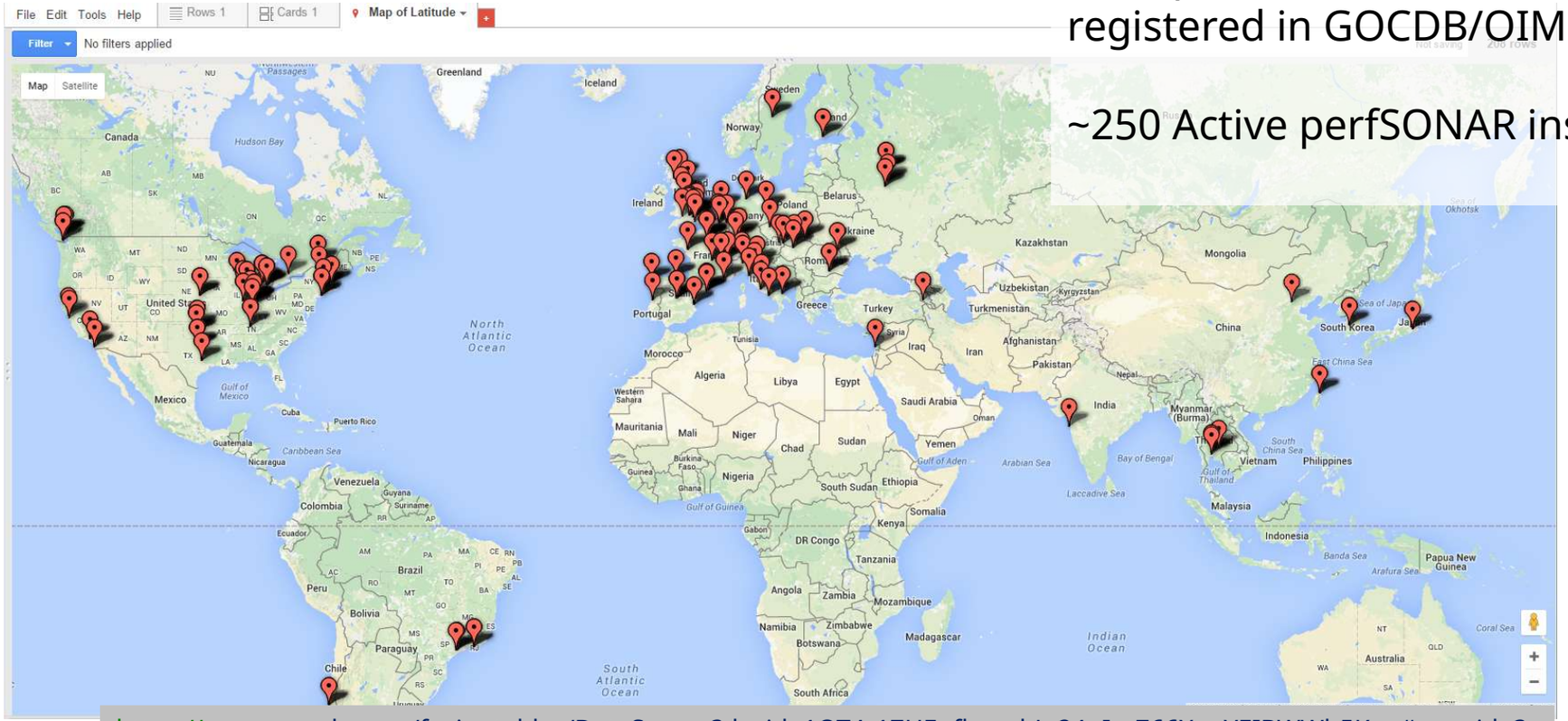
- Accessible to any TierX to check network healthiness and debug issues



LHCONE perfSONAR: status

~280 perfSONAR instances registered in GOCDB/OIM

~250 Active perfSONAR instances



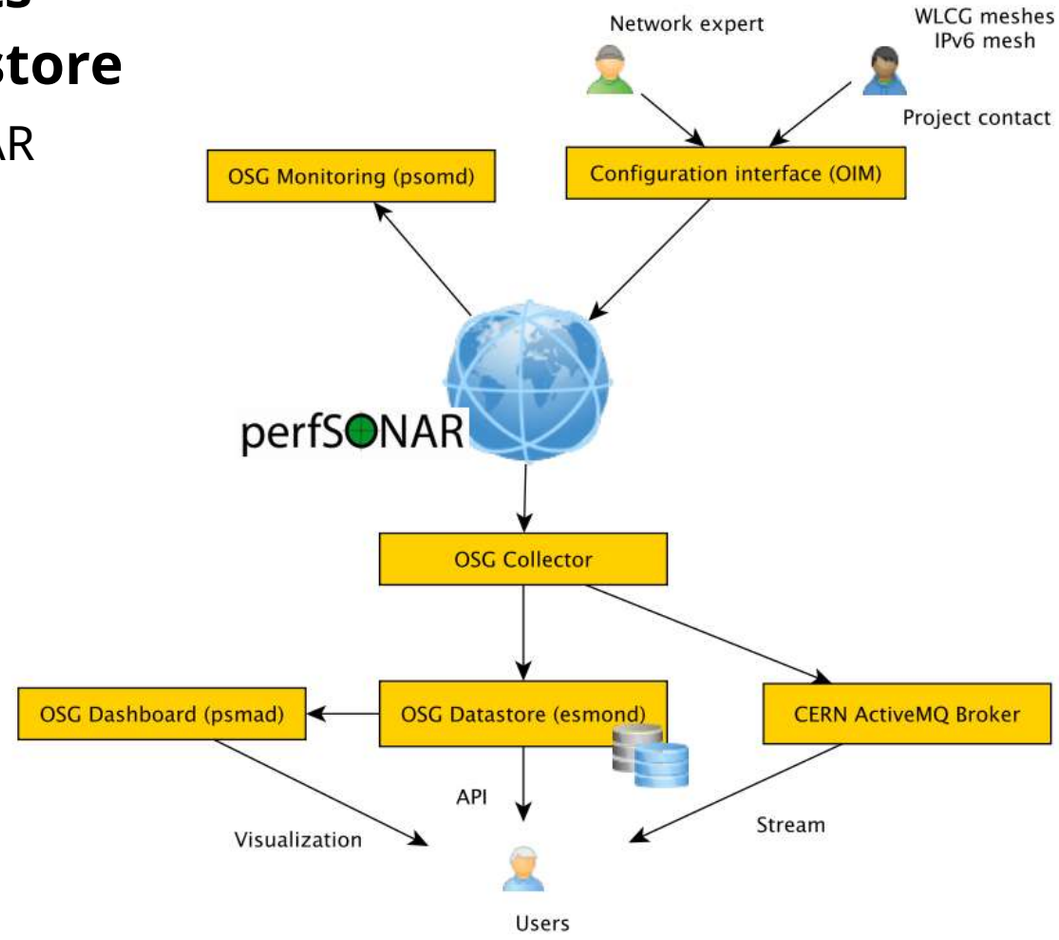
https://www.google.com/fusiontables/DataSource?docid=1QT4r17HEufkvnqhJu24nIptZ66XauYEIBWWh5Kpa#map:id=3

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

Slide credit: Shawn McKee, University of Michigan

perfSONAR: gathering and storing metrics

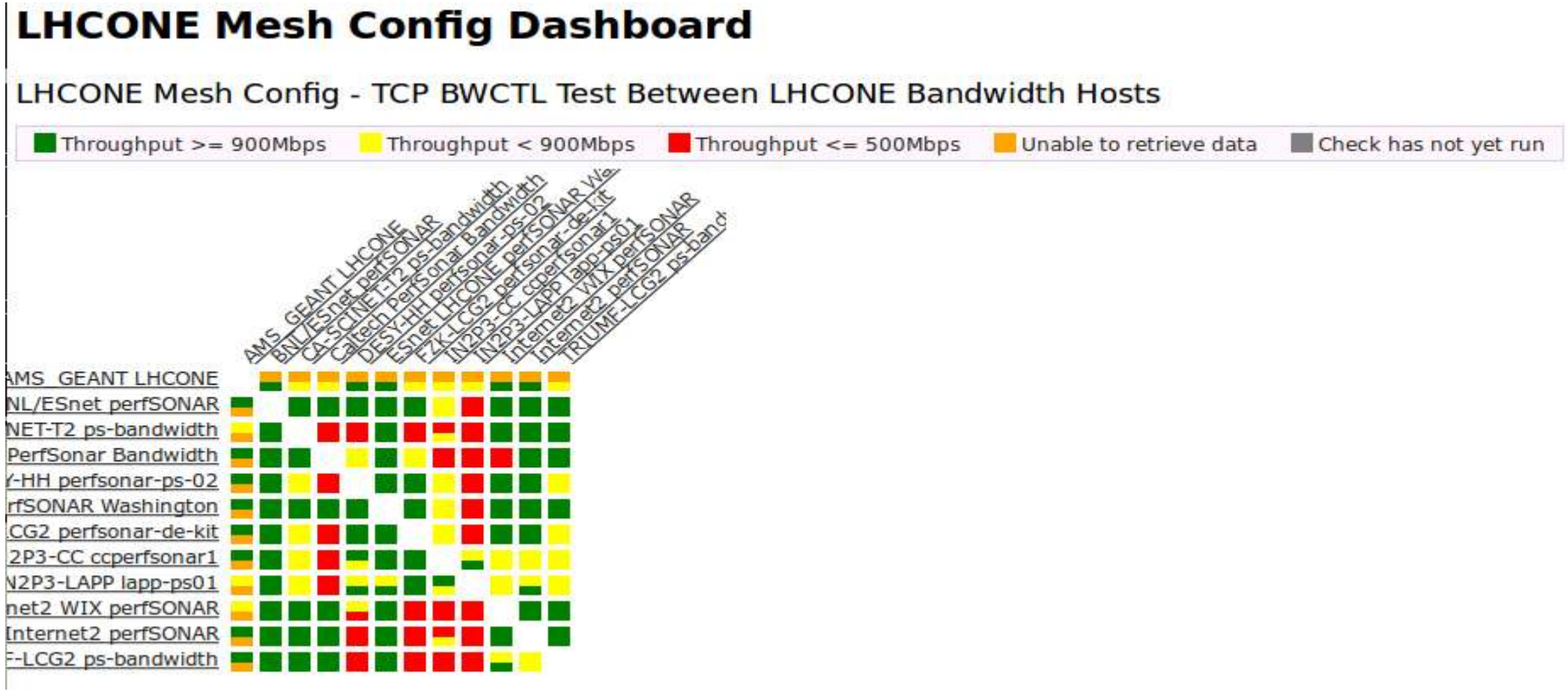
- **OSG is providing network metric data for its members and WLCG via the Network Datastore**
  - The data is gathered from all WLCG/OSG perfSONAR instances
  - Stored indefinitely on OSG hardware
  - Data available via Esmond API
  - In production since September 14th 2015
- **The primary use-cases**
  - 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 MadDash

- LHCONE Network monitoring dashboard
- Probe installed at the TierXs



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

perfSONAR update

Completed MCA (Mesh Configuration Admin) tool

WLCG is working on ETF (Experiment Test Framework) to monitor perfSONAR services

ATLAS is working on getting network metrics into an analytics platform:

- packet loss from perfSONAR
- network stats and flows from ESnet routers
- stats from CERN routers will be added

Looking for more router sources

LHCONE meetings

First meeting on June 2010 at CERN

- workshop on Transatlantic Networking for LHC Experiments

Design of Tier1/2/3 network for a changing computing model

- including advanced services and end-to-end monitoring
- implementation of services started already in 2011, ready for LHC Run2

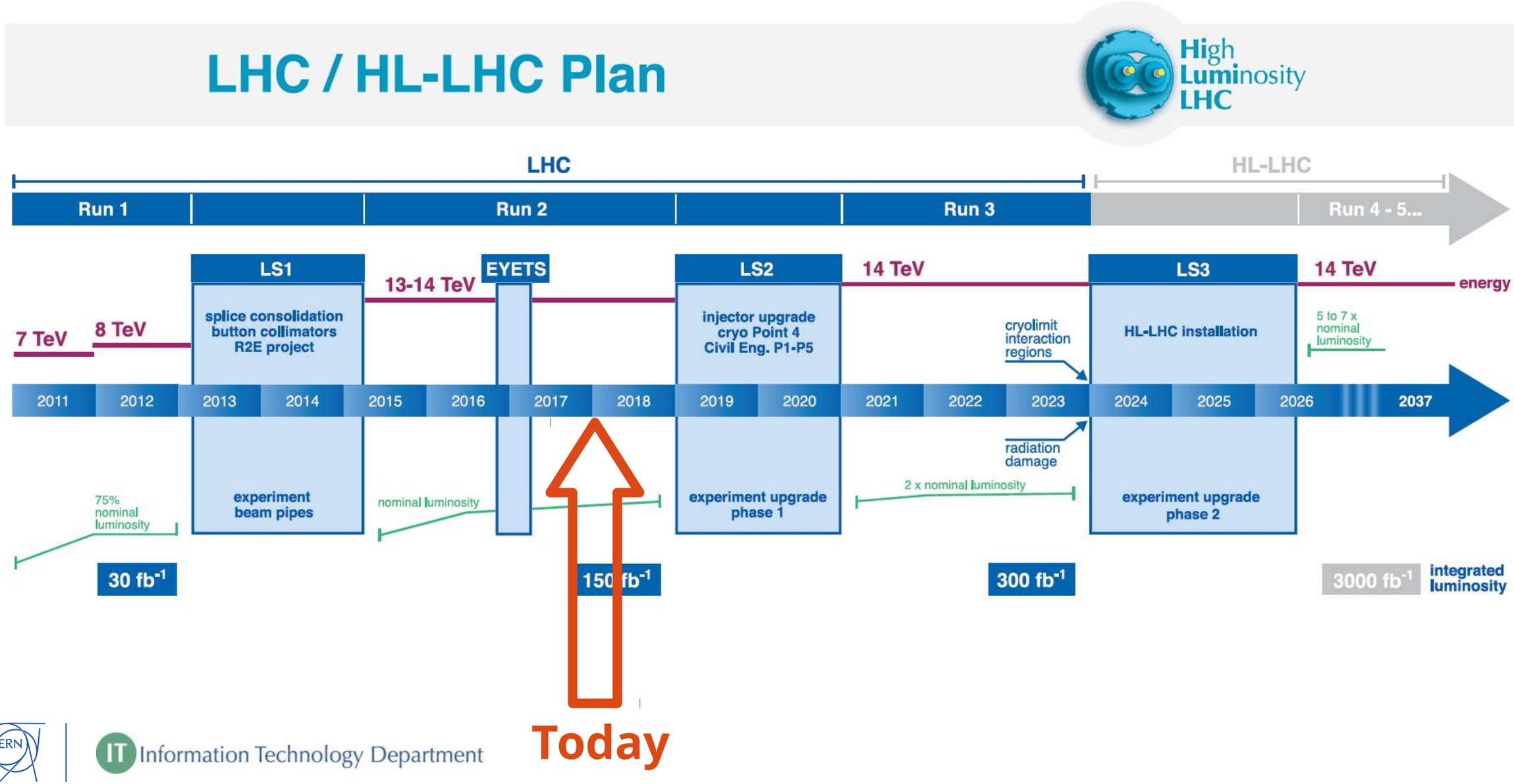
Join LHCONE + LHCOPN meetings every 6 months

LHCOPN and LHCONE Milestones

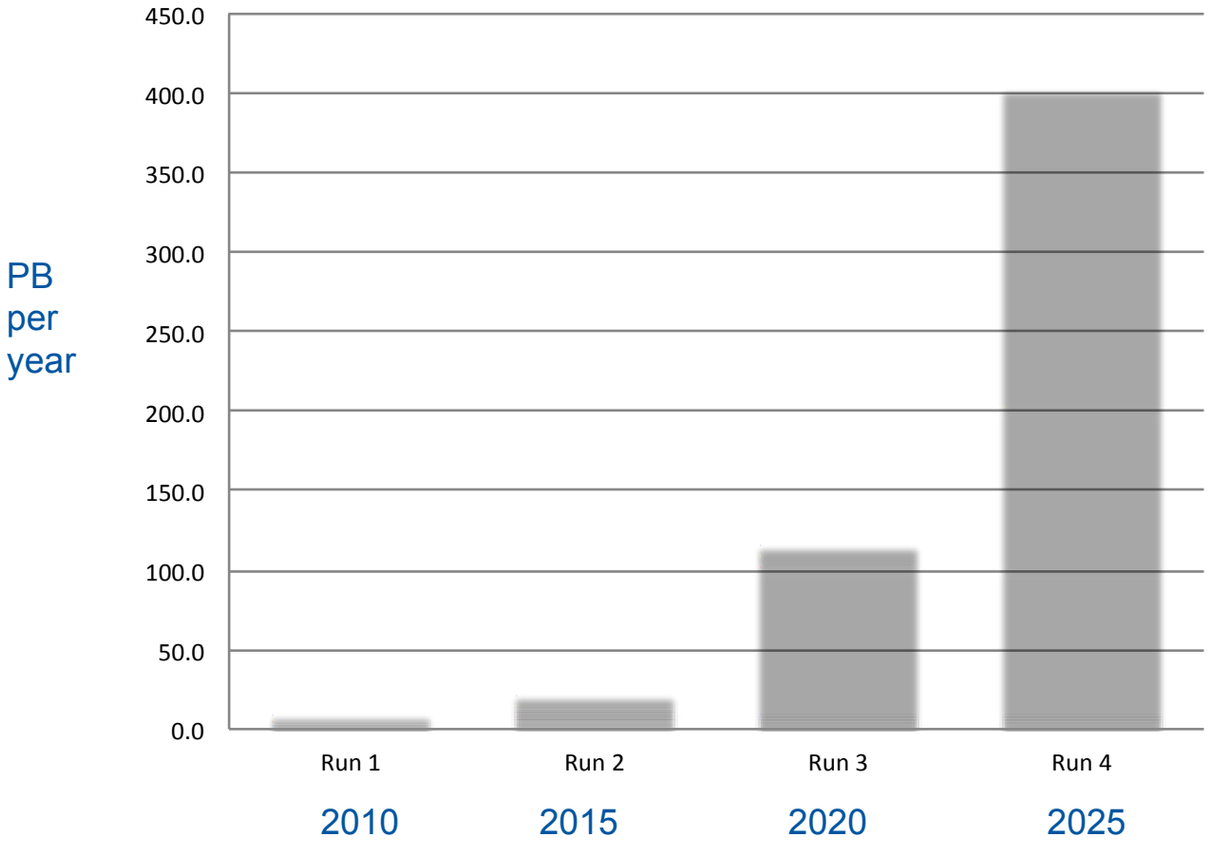
- 2005, January: First LHCOPN meeting at SARA, Amsterdam
- **2007, July: all Tier1s connected to the Tier0 in LHCOPN**
- 2008, September: LHC first beam
- 2010, June: first LHCONE meeting at CERN
- 2010, October: all LHCOPN links backed up
- 2011, June: First prototype of LHCONE implemented
- **2012, May: LHCONE L3VPN implemented in Europe and North America**
- **2013, January: LHCONE reaches ASIA**
- 2014, July: BelleII joins LHCONE
- 2015, February: LHCONE AUP agreed
- **2015, June: LHCONE reaches South America**
- 2016: Pierre Auger, NOVA, XENON joins LHCONE
- **2017, August: LHCONE reaches Australia**

Future developments

LHC Schedule



Networks have to follow LHC data growth



CMS
ATLAS
ALICE
LHCb

LHC data expected to record 400PB/year in Run4

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

Networks must grow accordingly

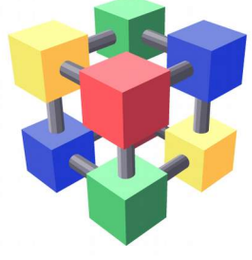
LHC experiments' requirements for Run3

pre-GDB meeting on networks held in January 2017. Gathered LHC experiments, WLCG sites and REN operators

Collected requirements for Run3:
- increase network capacity of 5-10 times
- improve monitoring and share statistics of the networks

Acknowledged the usefulness of gathering Experiments, Sites and Network operators all together. Should be repeated every 1-2 years

https://indico.cern.ch/event/571501/



WLCG
Worldwide LHC Computing Grid

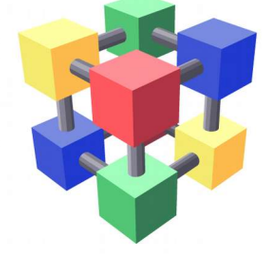
Community White Paper

The HEP community has produced a Community White Paper (CWP) which describes the community strategy and a roadmap for software and computing R&D in HEP for the 2020s.

The CWP identifies and prioritizes the software research and development investments required:

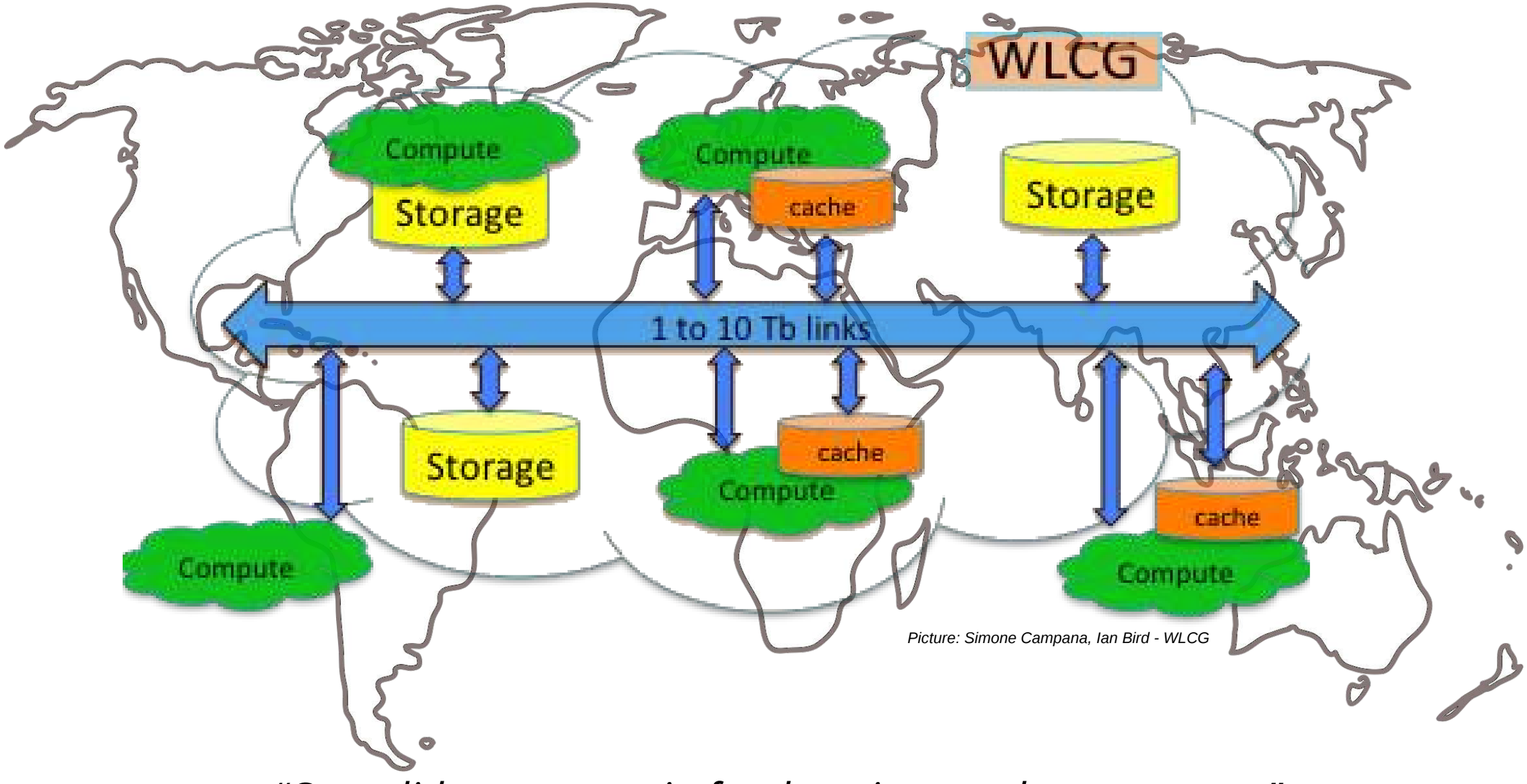
- to achieve improvements in software efficiency, scalability and performance and to make use of the advances in CPU, storage and network technologies
- to enable new approaches to computing and software that could radically extend the physics reach of the detectors
- to ensure the long term sustainability of the software through the lifetime of the HL-LHC

http://hepsoftwarefoundation.org/activities/cwp.html



WLCG
Worldwide LHC Computing Grid

Possible change of computing model



Picture: Simone Campana, Ian Bird - WLCG

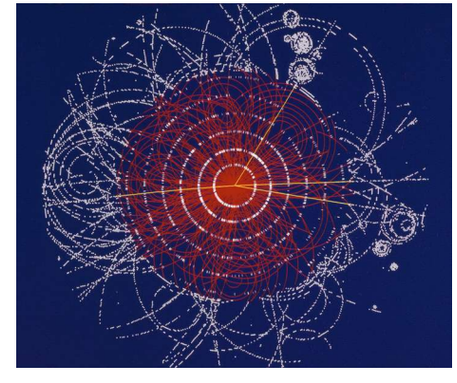
"Consolidate storage in few locations to decrease costs"

Use of Commercial Cloud Services

Evaluation of use of Commercial Cloud Services for HEP computing have been on-going

Research and Education Networks have designed and deployed 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



Conclusion

Summary

LHCOPN:

- primary purpose: T0 to T1s data distribution
- started 100G adoption

LHCONE:

- T1s-T2s data transfers
- expanding in capacity and extension
- more connections in Asia-Pacific

Future developments:

- planning for data deluge in Run4
- possible change of computing model

More information

Next LHCOPN/ONE meeting:
Date: 6-7 March 2018
Location: RAL - Abingdon, UK
https://indico.cern.ch/event/681168/

Previous LHCOPN/ONE meetings:
KEK, Tsukuba, october 2017: https://indico.cern.ch/event/646629/
BNL New York, March 2017: http://indico.cern.ch/event/581520

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