The CERN Network

Openlab Summer 2012
CERN, 6th August 2012

edoardo.martelli@cern.ch

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

Summary

- IT-CS
- CERN networks
- LHC Data Challenge
- WLCG
- LHCOPN and LHCONE
- Openlab
- Conclusions

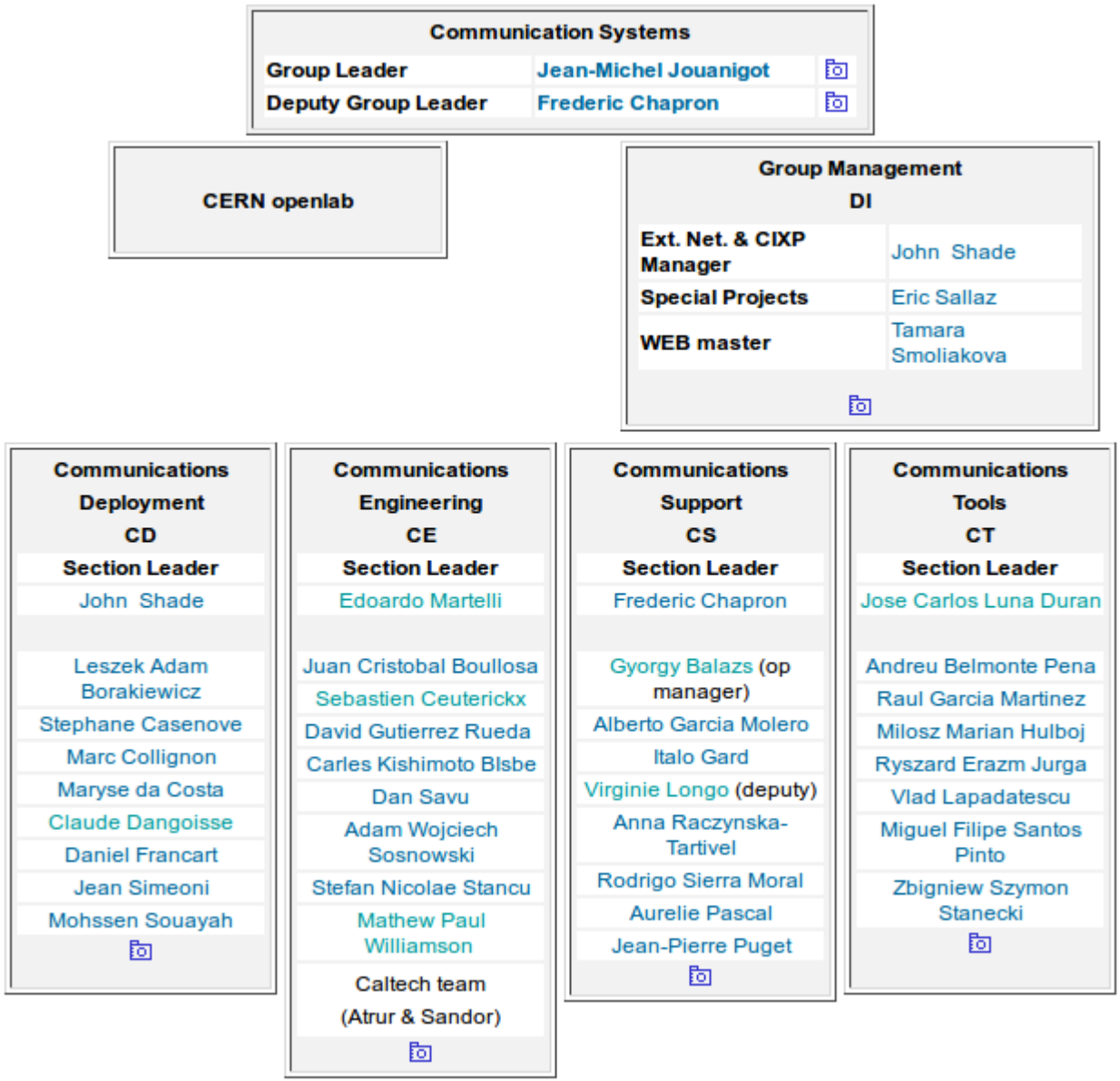
IT-CS
Communication systems

IT-CS

The IT-CS group is responsible for all communication services in use at CERN for data, voice and video

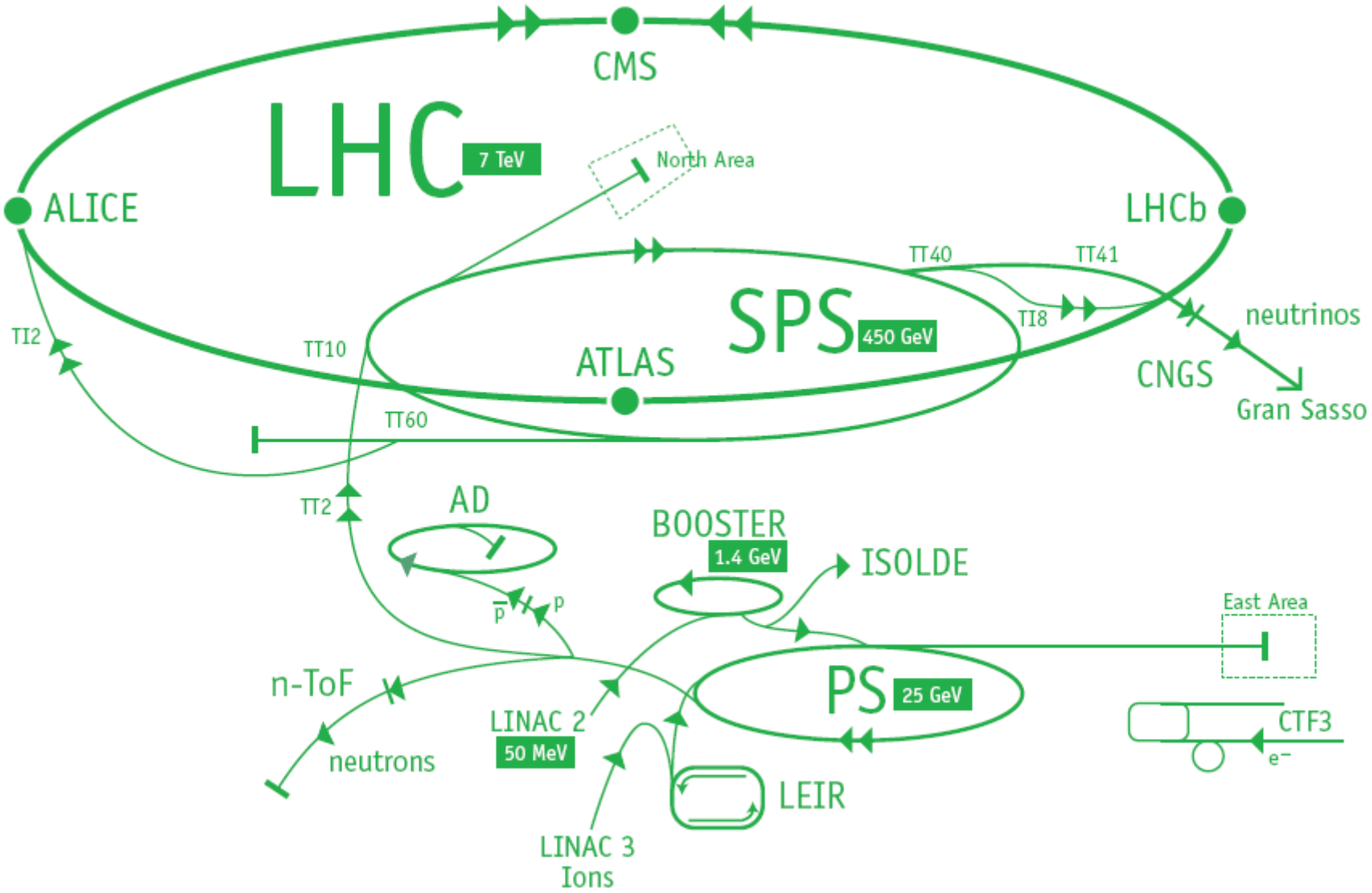
http://it-cs.web.cern.ch/it-cs/

IT-CS organization



Networks at CERN

CERN accelerator complex

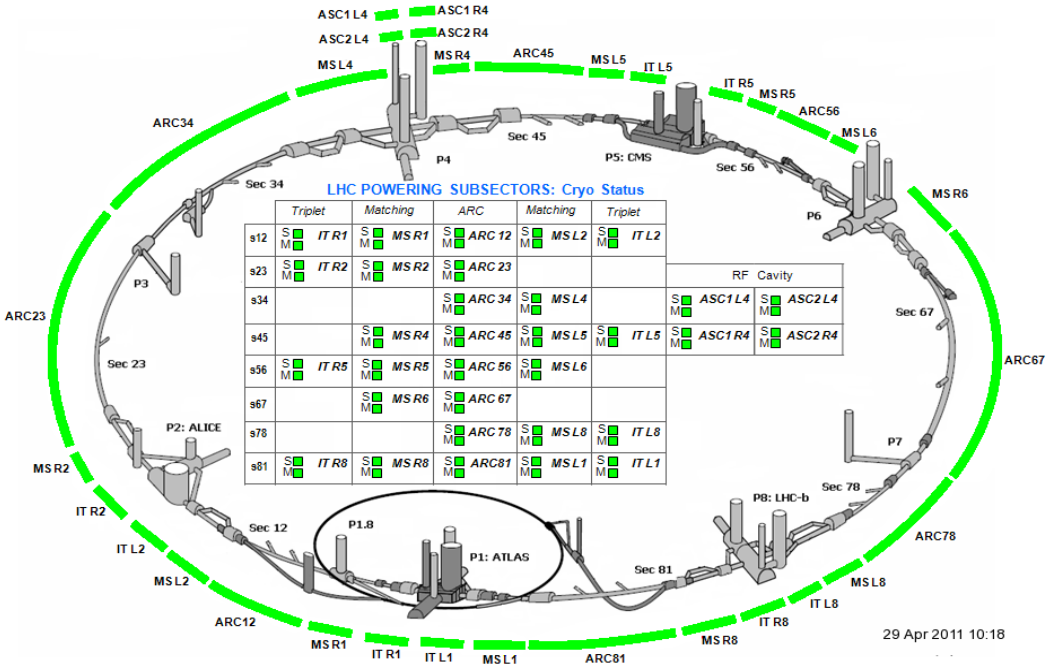


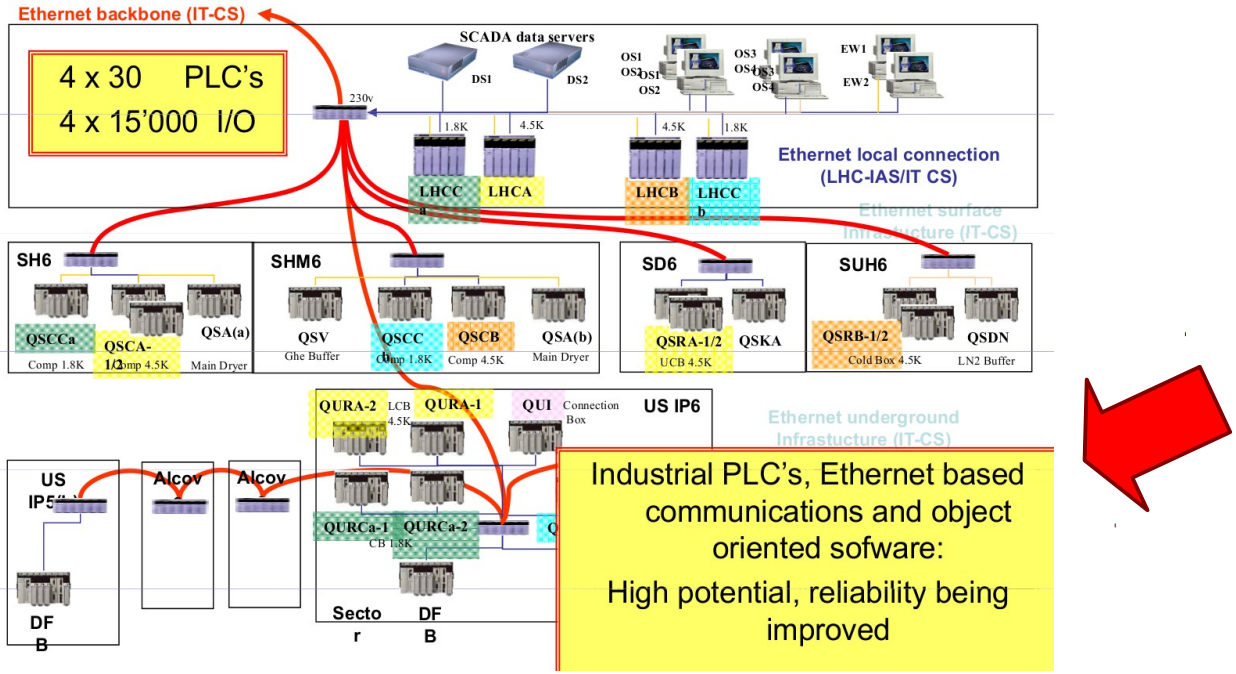
High Energy Physics over IP

Most of the CERN infrastructure is controlled and managed over a pervasive **IP network**

Cryogenics

27Km of pipes at
-271.11° C by means of
700.000 litres of Helium:
controlled over IP

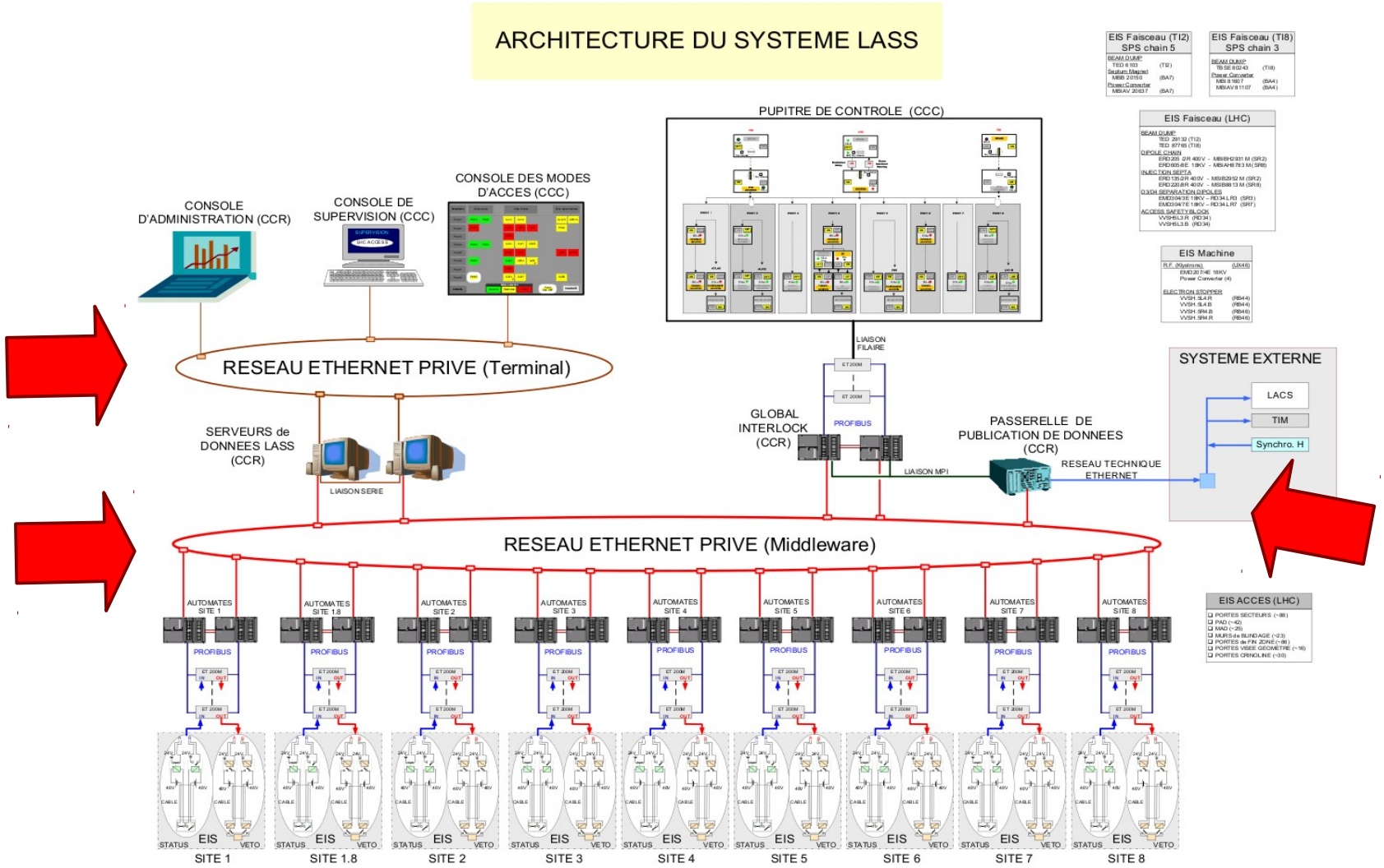




Source: http://te-dep-crg-oa.web.cern.ch/te-dep-crg-oa/te-crg-oa_fichiers/cryolhc/LHC%20Cryo_BEOP_lectures2009.pdf

Access control

Safety and Security: **made over IP**



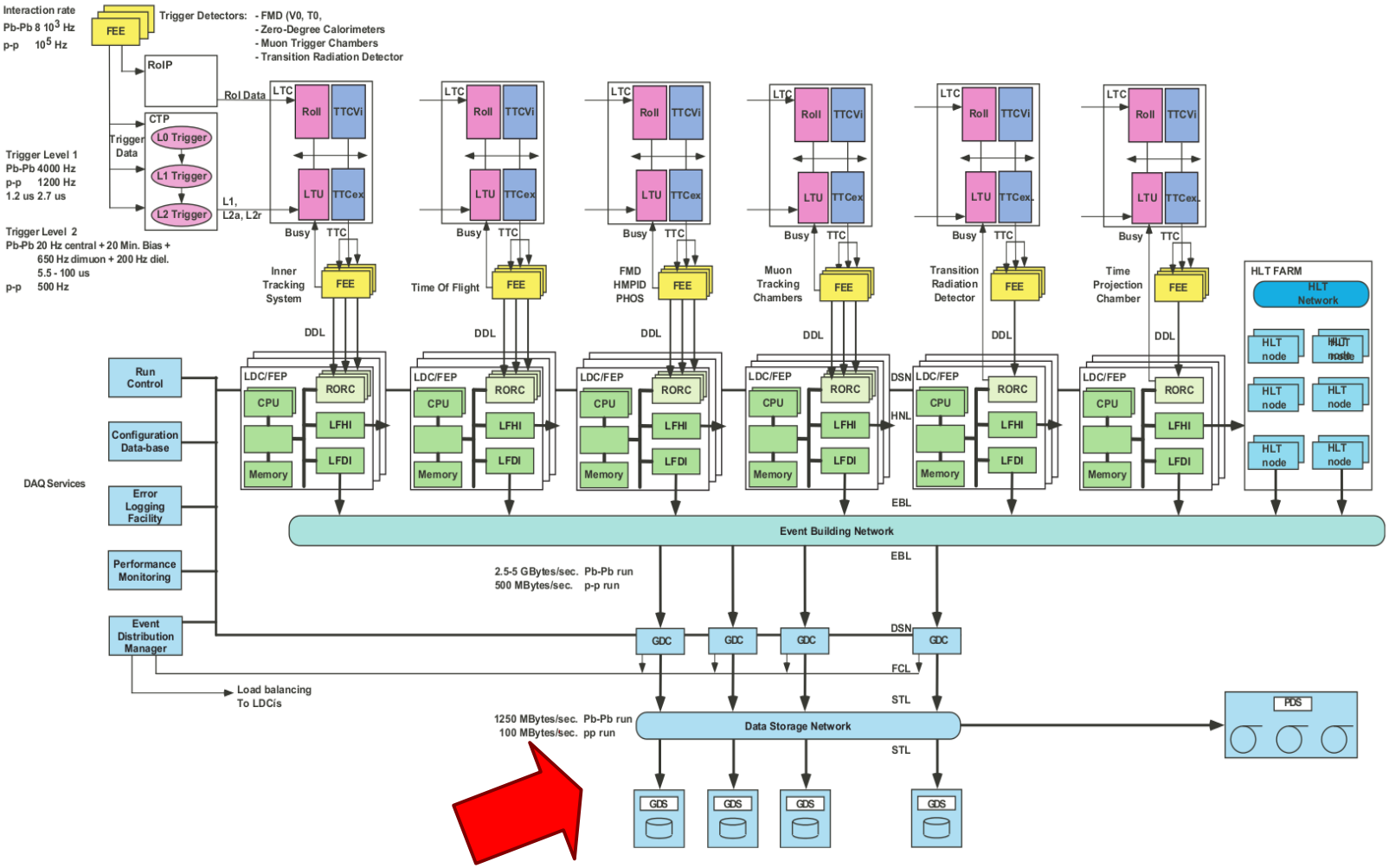
Source:https://edms.cern.ch/file/931641/1/LASS-LACS_IHM.pdf

Remote inspections

Remote inspection of dangerous areas: robots controlled and giving feedback over **WiFi and GSM IP networks**

DAQ: Data Acquisition

A constant stream of data from the four Detectors to disk storage

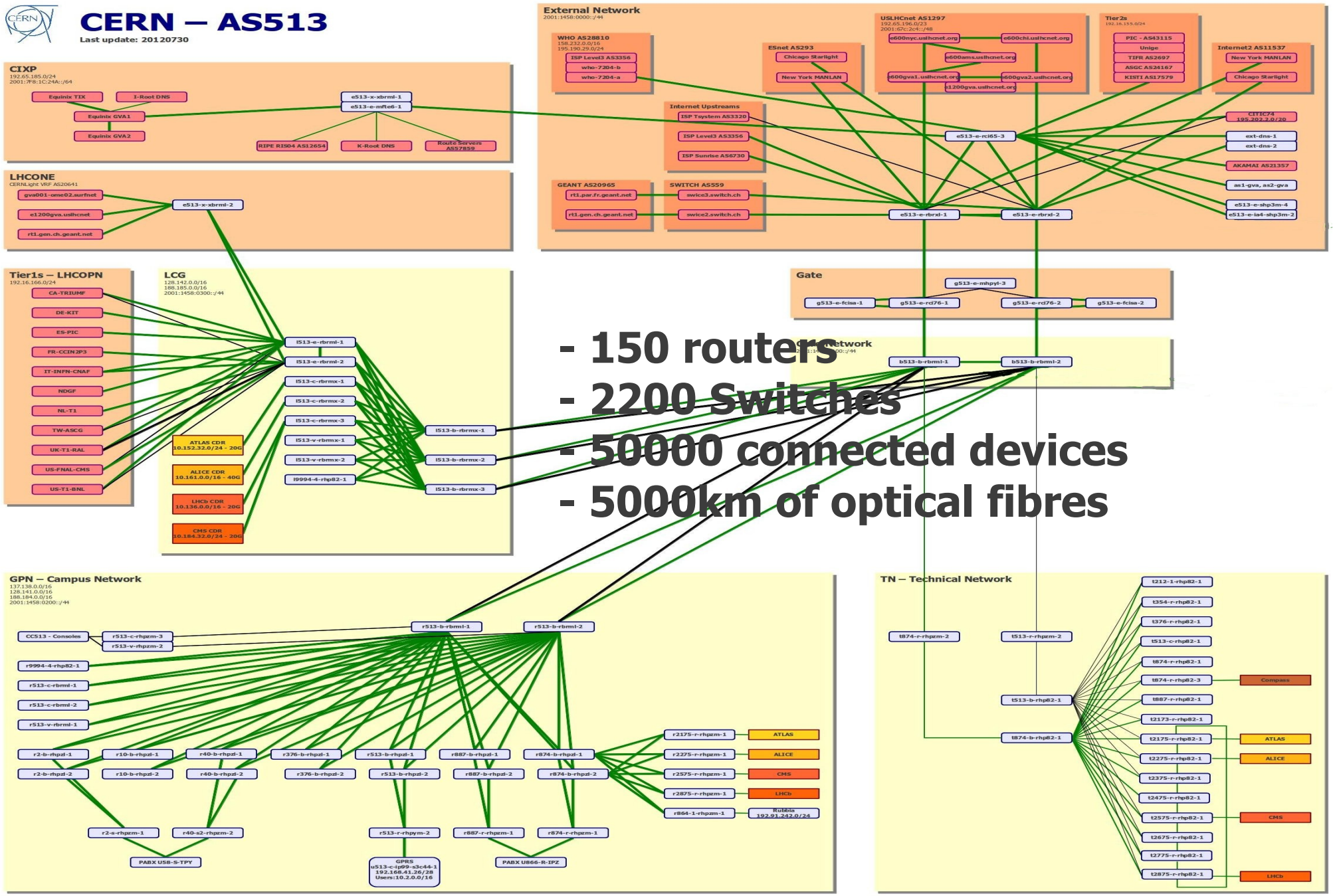


Source: http://aliceinfo.cern.ch/Public/Objects/Chapter2/DetectorComponents/daq_architecture.pdf

CCC: CERN Control Centre

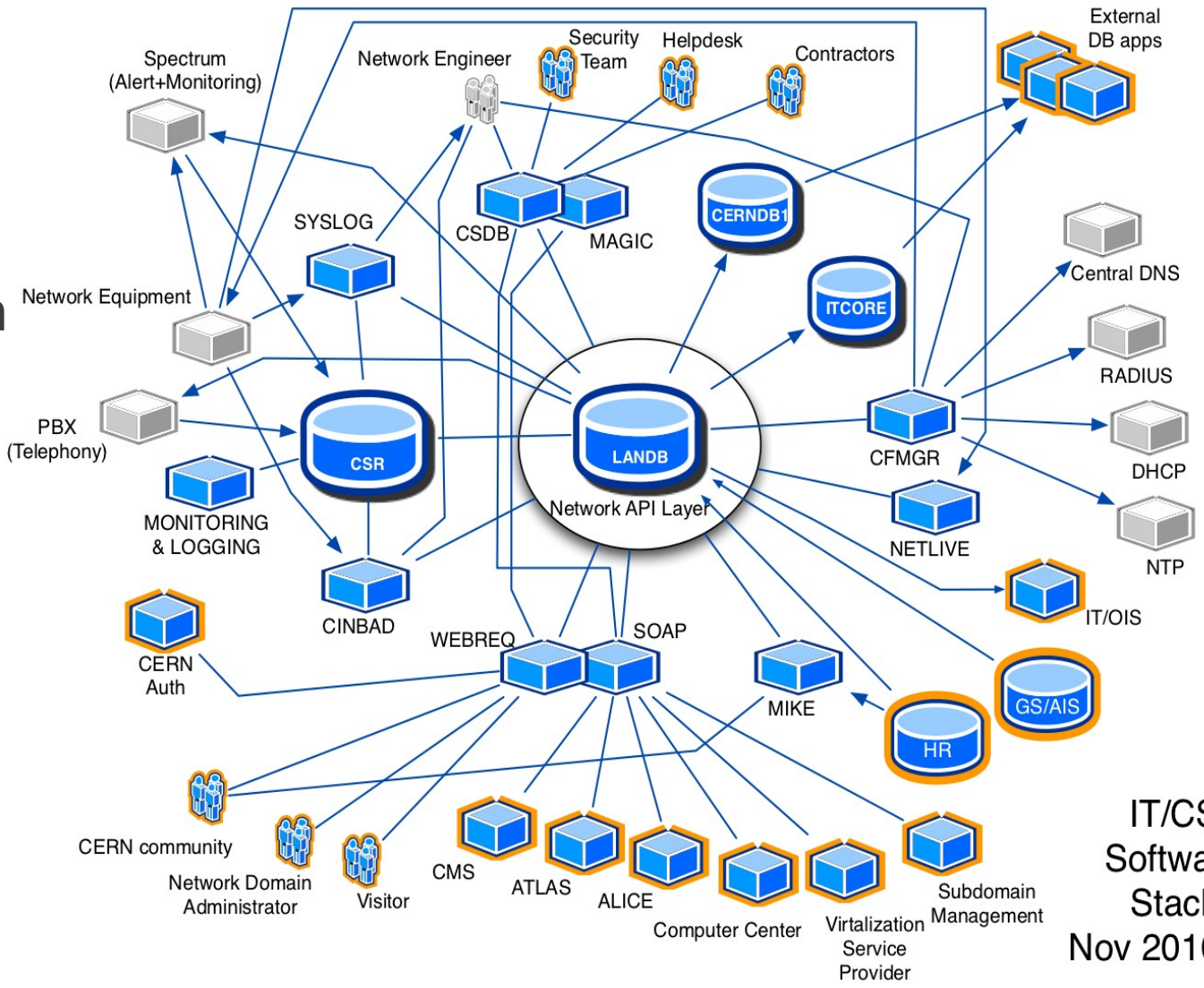
The neuralgic centre of the particle accelerator: over IP

CERN data network



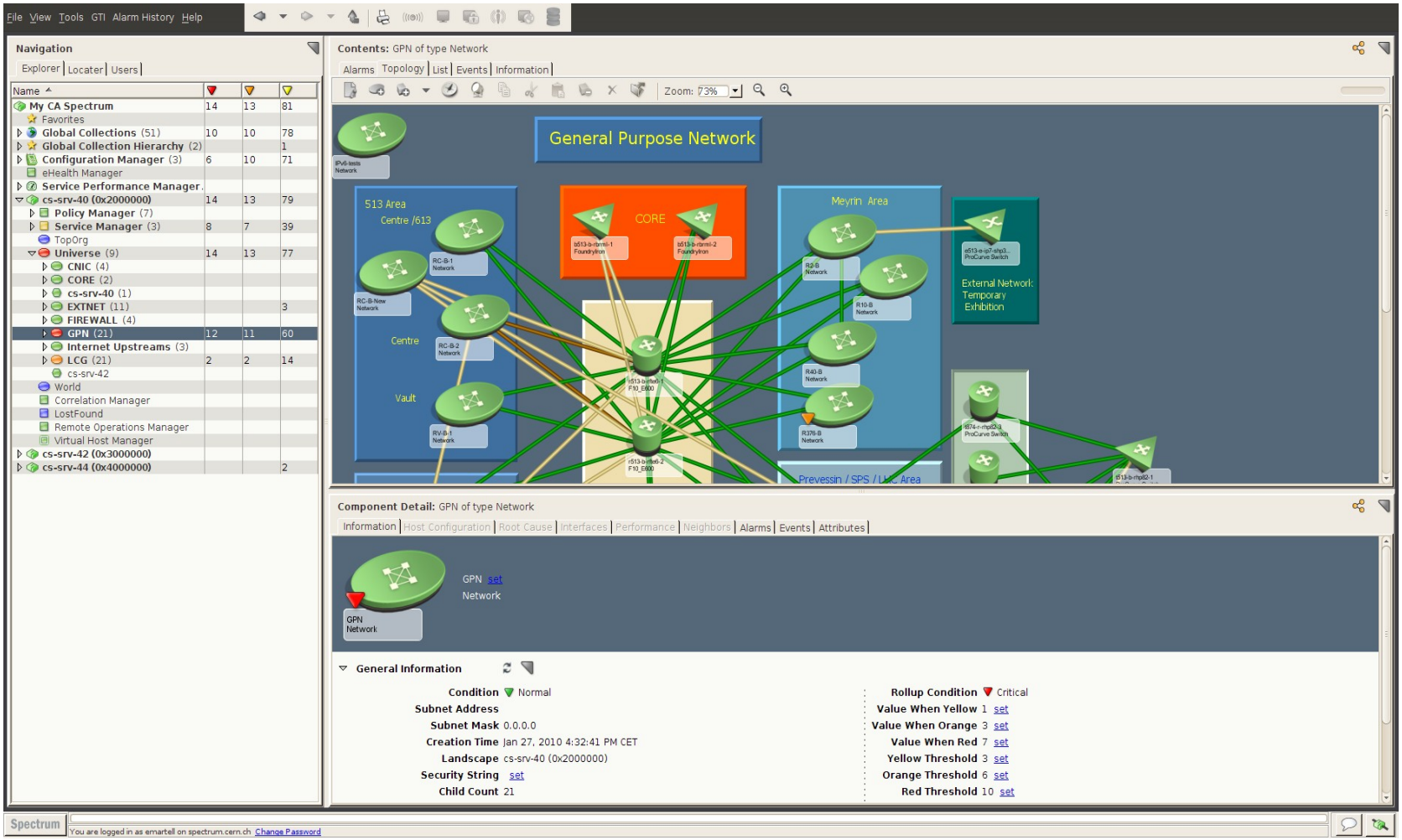
Network Provisioning and Management System

- 250 Database tables
- 100,000 Registered devices
- 50,000 hits/day on web user interface
- 1,000,000 lines of codes
- 11 years of development



Monitoring and Operations

The whole network is monitored and operated by the CERN NOC (Network Operation Centre)



IPv6

IPv6 dual stack network deployment on going: ready in 2013

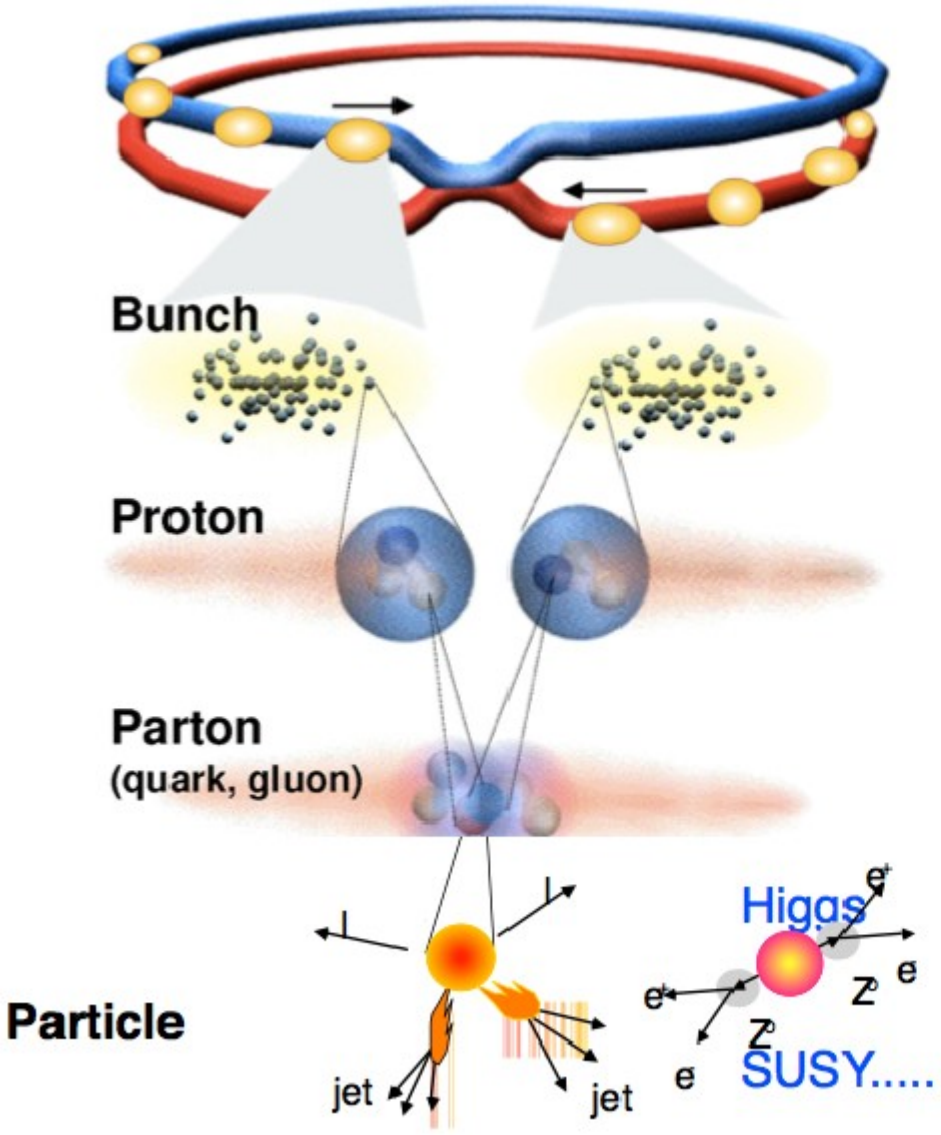
Already available: dual-stack testbed

More information:
http://cern.ch/ipv6

almost 

LHC Data Challenge

Collisions in the LHC

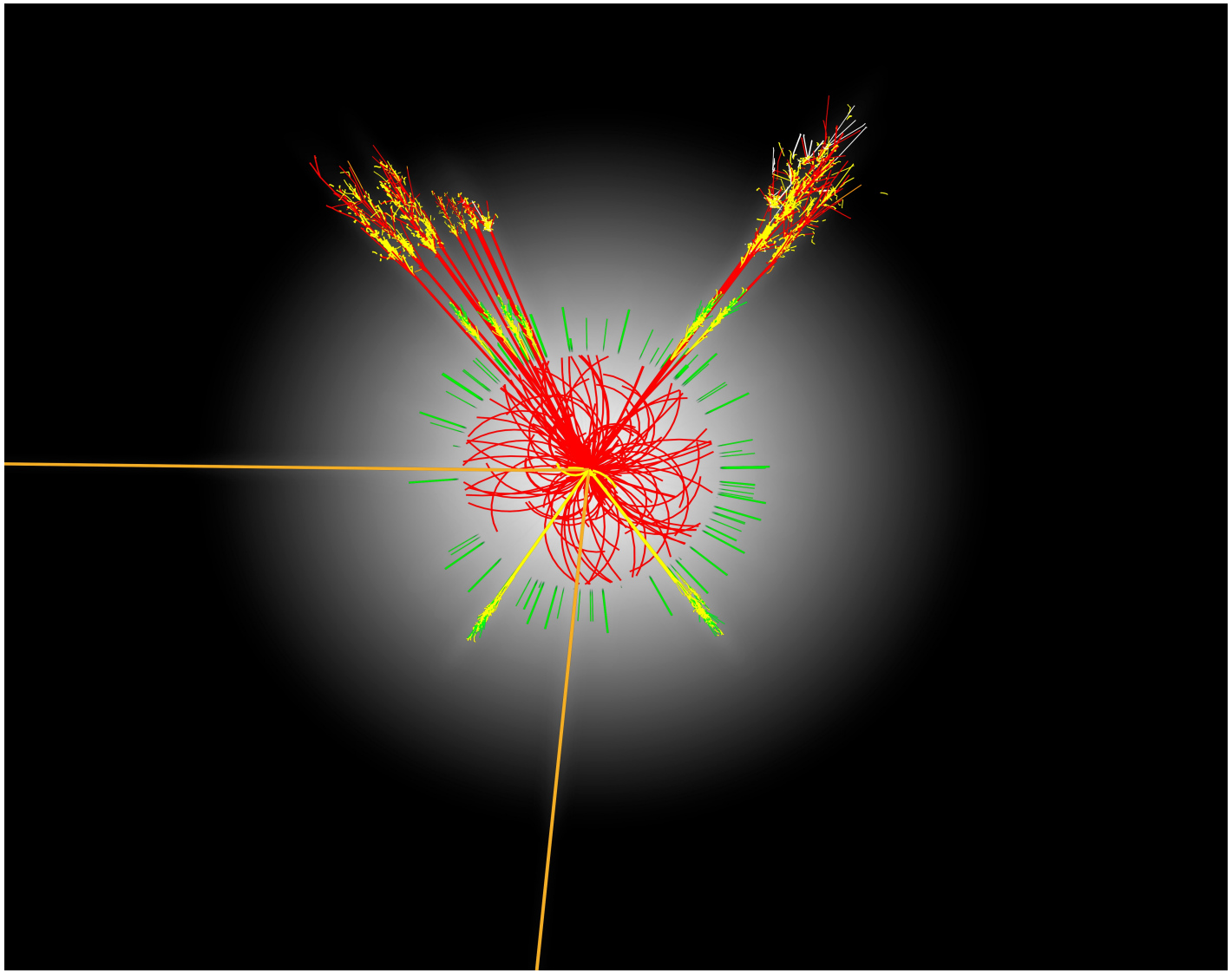


Proton - Proton	2808 bunch/beam
Protons/bunch	10^{11}
Beam energy	7 TeV (7×10^{12} eV)
Luminosity	$10^{34} cm^{-2} s^{-1}$
Crossing rate	40 MHz
Collision rate ≈	10^{7}-10^{9}

New physics rate ≈ .00001 Hz

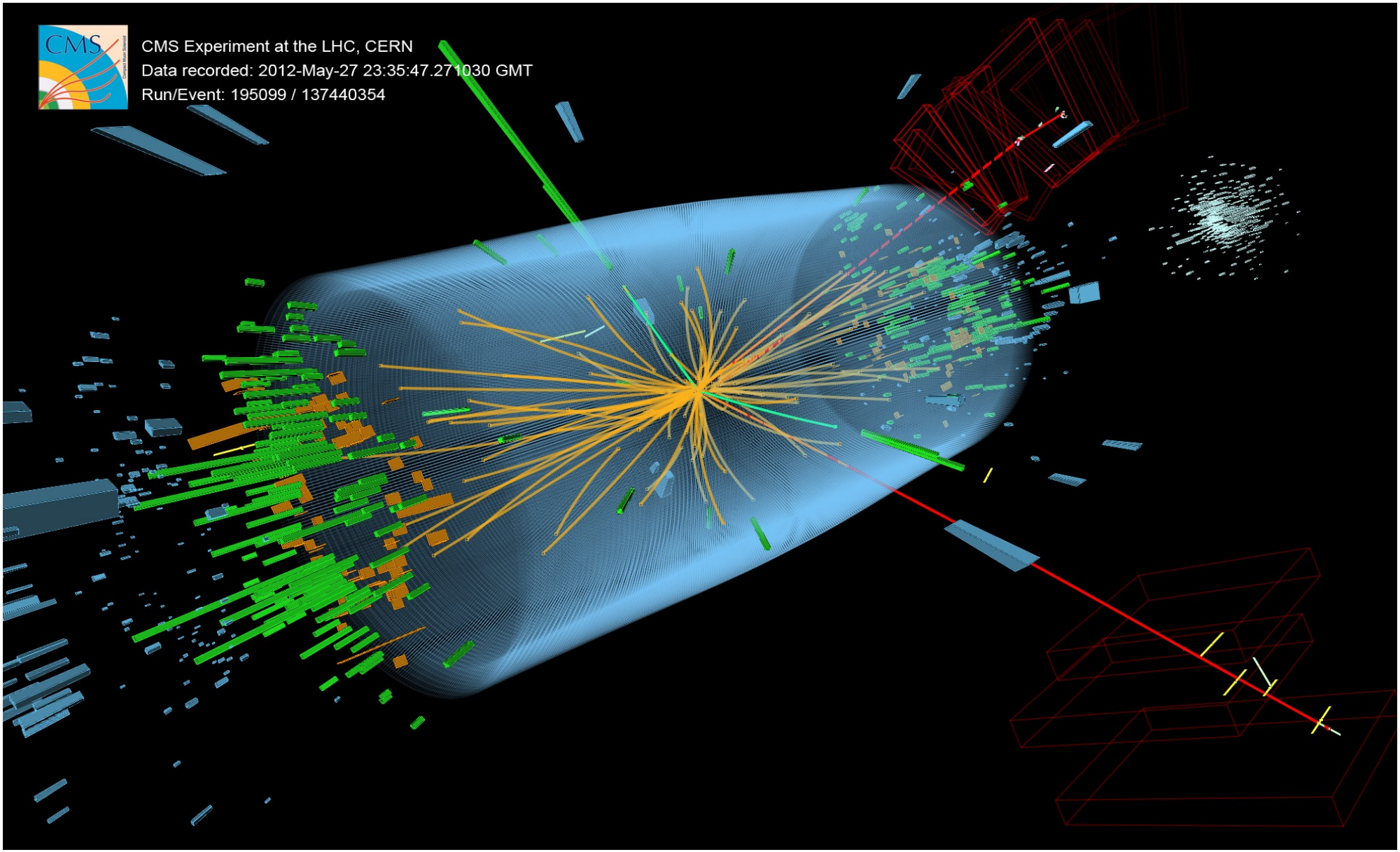
Event selection:
1 in 10,000,000,000,000

Comparing theory...



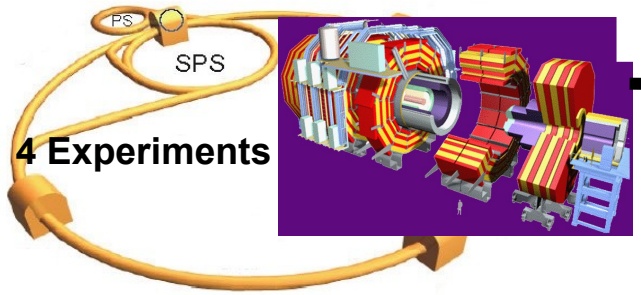
Simulated production of a Higgs event in ATLAS

.. to real events

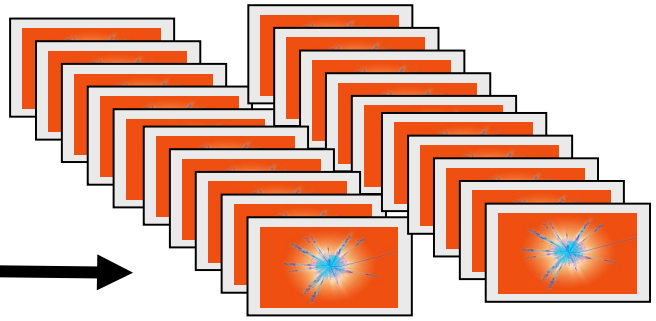


Higgs event in CMS

Data flow

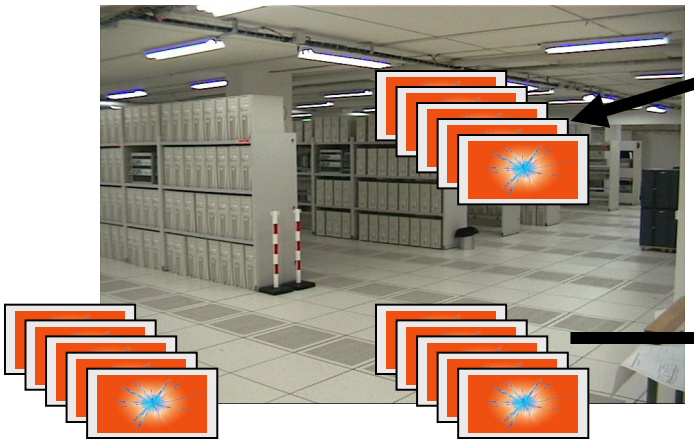


3 PBytes/s

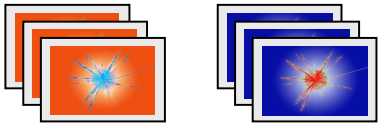


Filter and first selection

2 GBytes/s
to the CERN computer center



Create sub-samples



1 TByte/s ?
Distributed + local

World-Wide Analysis

Physics
Explanation of nature

10 GBytes/s

4 GBytes/s

Store on disk and tape

Export copies

$$\sigma_{f\bar{f}} \approx \sigma^0_{f\bar{f}} \times \frac{s\Gamma_Z^2}{(s-m_Z^2)^2 + s^2\Gamma_Z^2/m_Z^2} \quad \text{with}$$

$$\sigma^0_{f\bar{f}} = \frac{12\pi}{m_Z^2}\frac{\Gamma_{ee}\Gamma_{ff}}{\Gamma_Z^2} \quad \text{and} \quad \Gamma_{ff} = \frac{G_F m_Z^3}{6\pi\sqrt{2}} \times (v_f^2 + a_f^2)$$

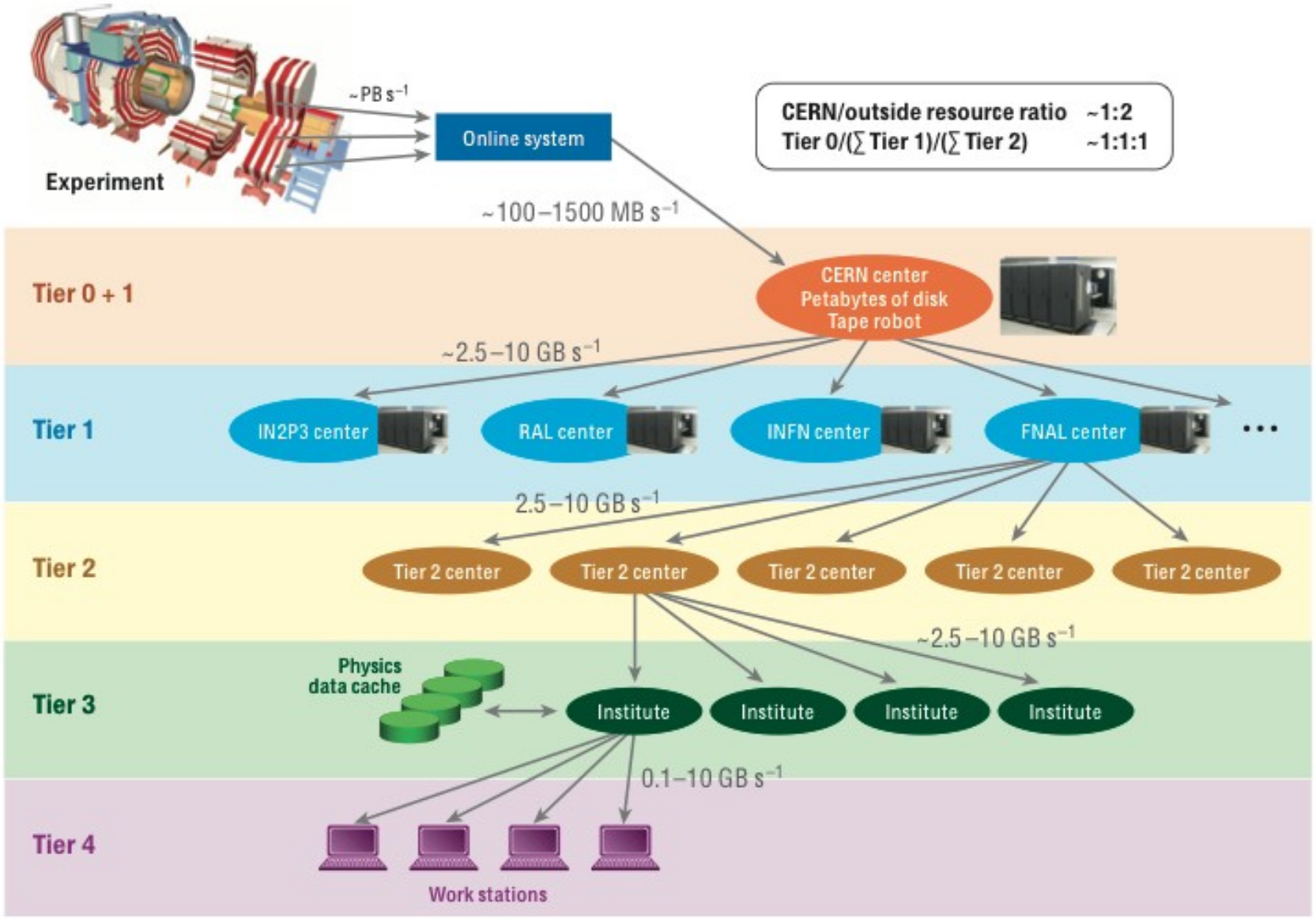
Data Challenge

- 40 million collisions per second

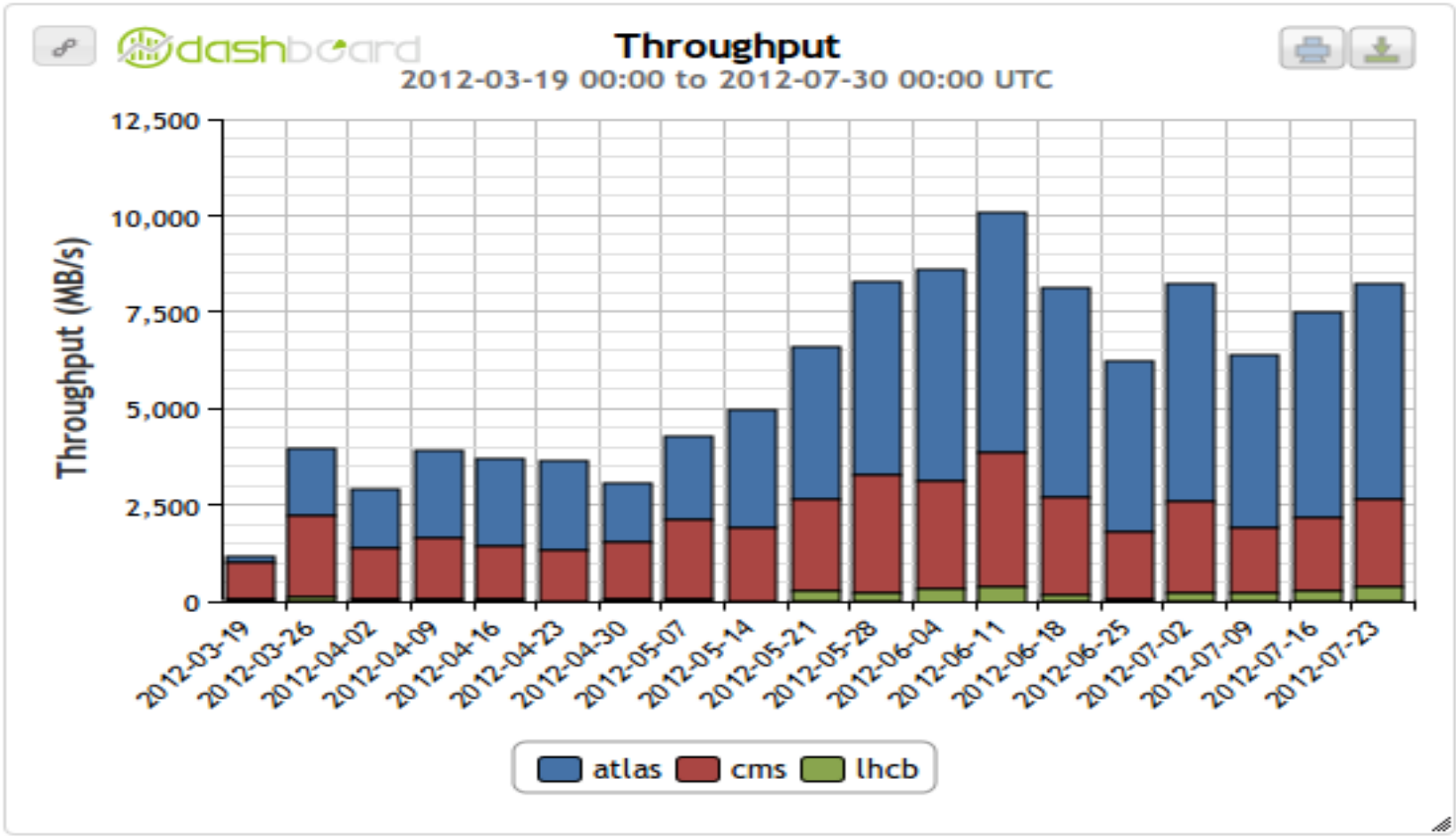
- After filtering, 100 collisions of interest per second

- 10^{10} collisions recorded each year
  = 15 Petabytes/year of data

Computing model



Last months data transfers



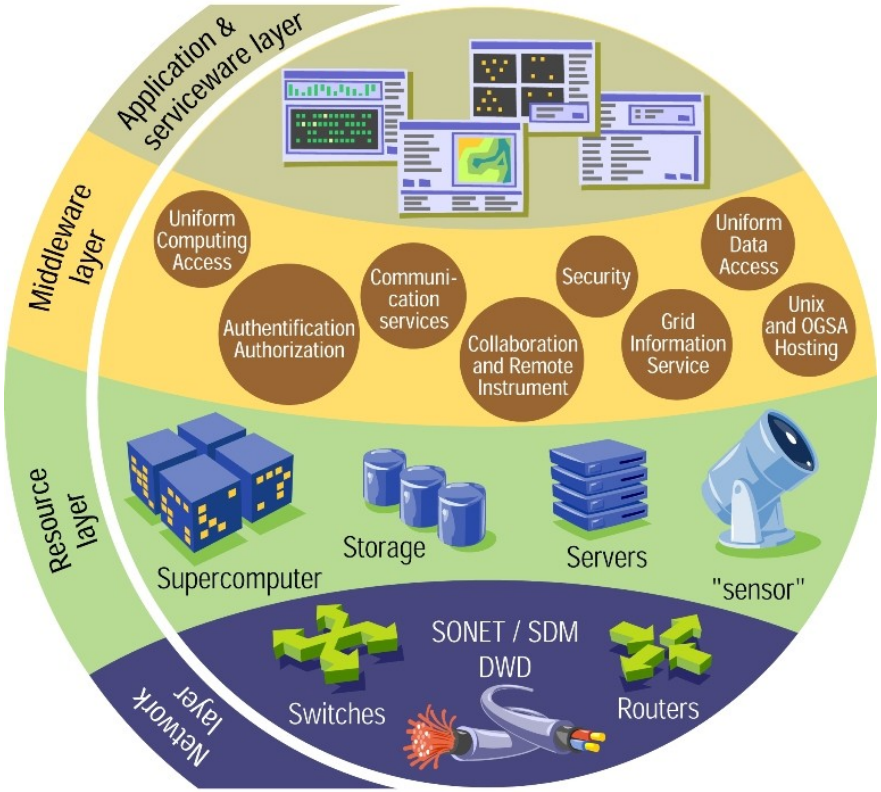
WLCG

Worldwide LHC Computing Grid

WLCG
Worldwide LHC Computing Grid

Distributed Computing Infrastructure for LHC experiments

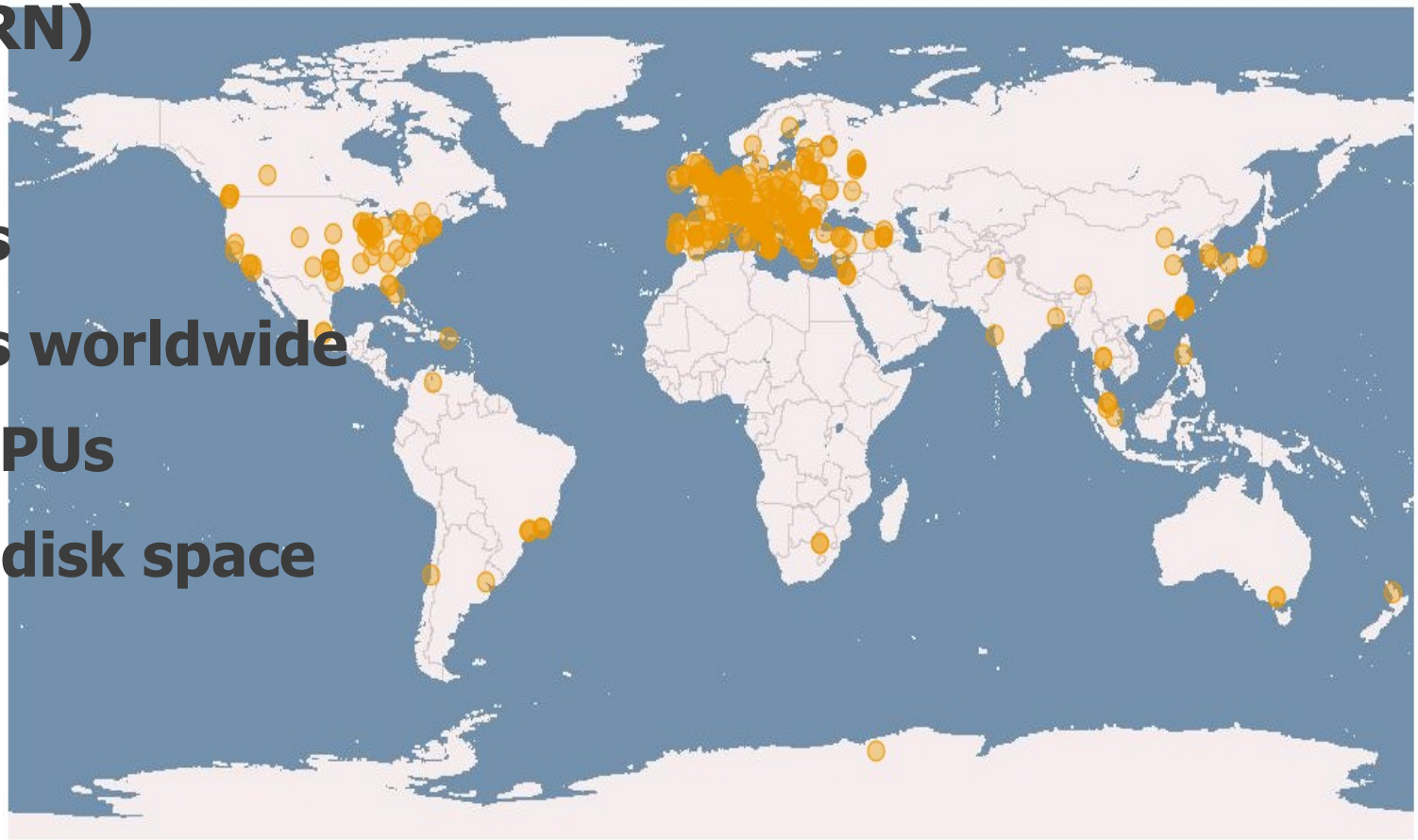
Collaborative effort of the HEP community



WLCG resources

WLCG sites:

- 1 Tier0 (CERN)
- 11 Tier1s
- ~140 Tier2s
- >300 Tier3s worldwide
- ~250,000 CPUs
- ~ 150PB of disk space



CERN Tier0 resources

Servers	11000
Processors	15000
Cores	64000
HEPspec06	480000

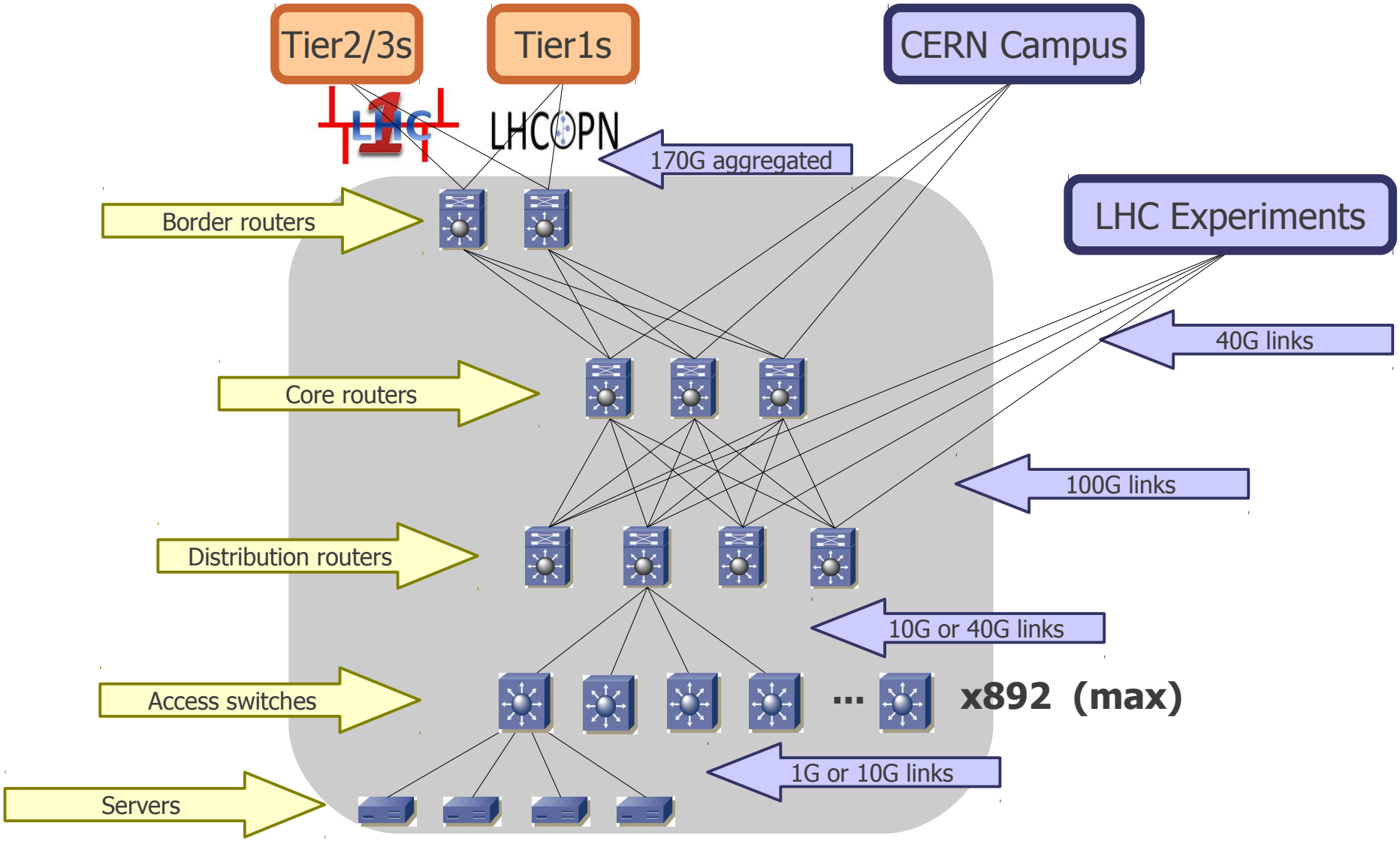
Disks	64000
Raw Disk Capacity (TB)	63000
Memory Modules	56000
RAID controllers	3750

Tape drives	160
Tape cartridges	45000
Tape slots	56000
Tape capacity(TB)	34000

High Speed routers	23
Ethernet switches	500
10Gbps ports	3000
100Gbps ports	48

March 2012

CERN Tier0 LCG network



Trends

Virtualization mobility (Software Defined Networks)

Commodity Servers with 10G NICs

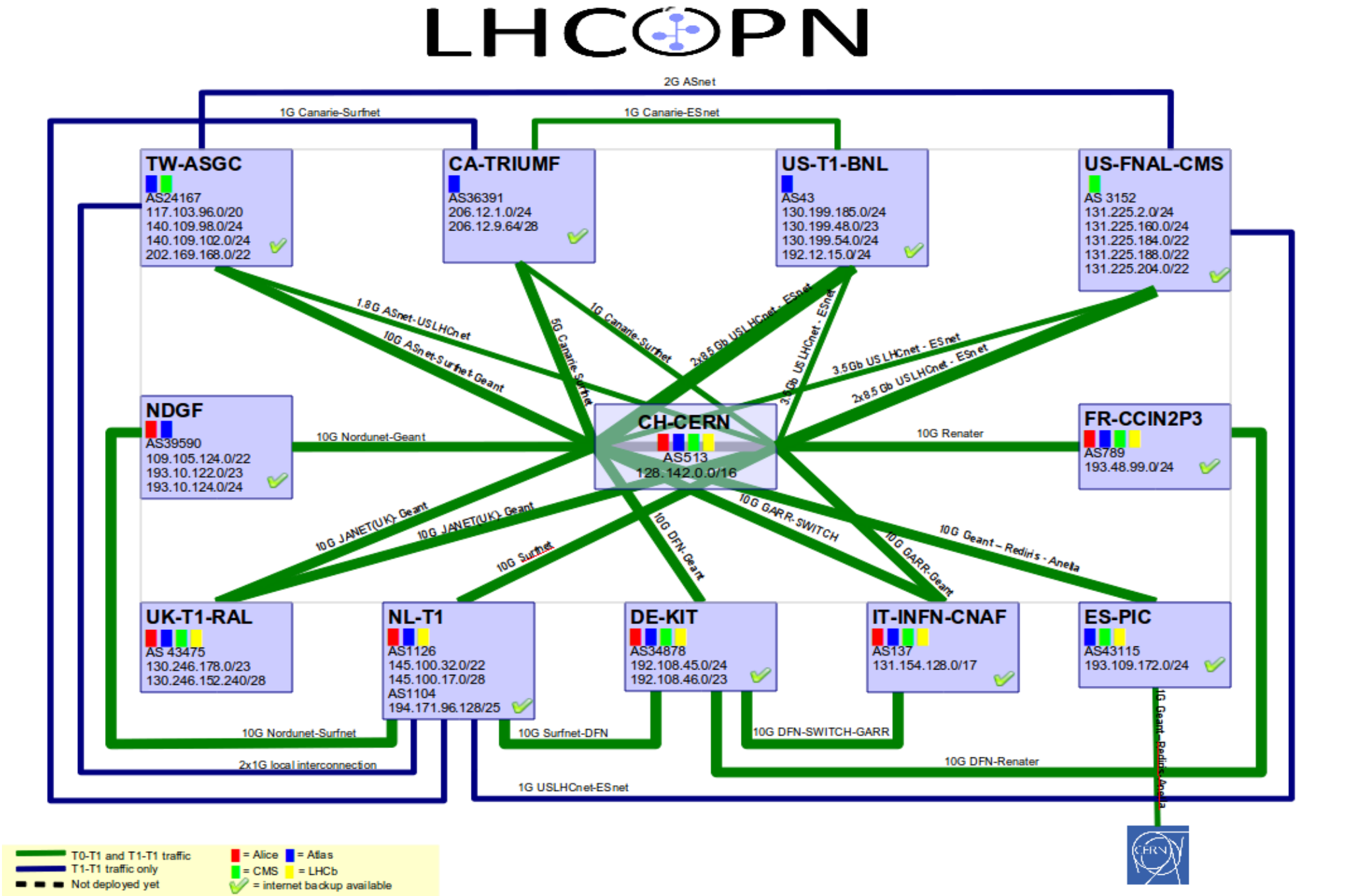
High-end Servers with 40G NICs

40G and 100G interfaces on switches and routers

LHCOPN

LHC Optical Private Network

Tier0-Tier1s network



A collaborative effort

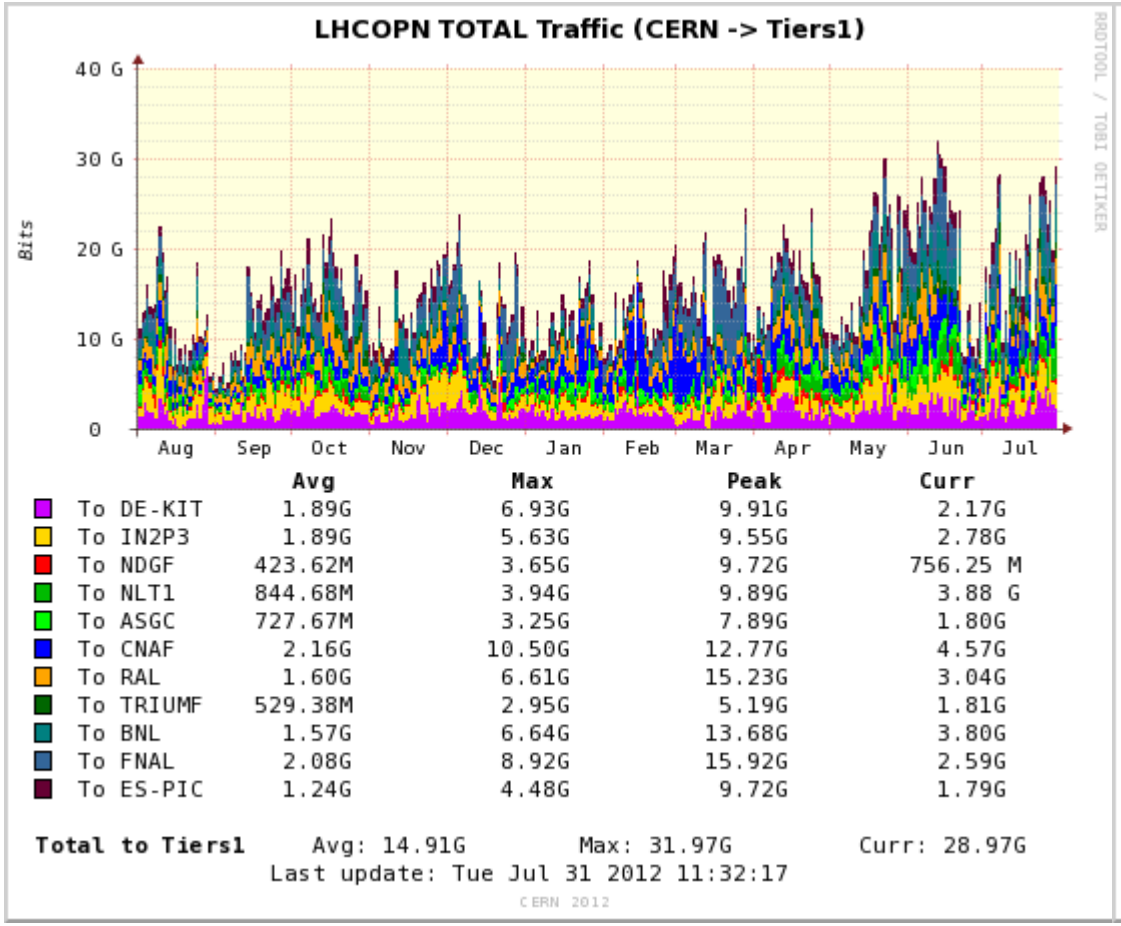
Designed, built and operated by the Tier0-Tier1s community

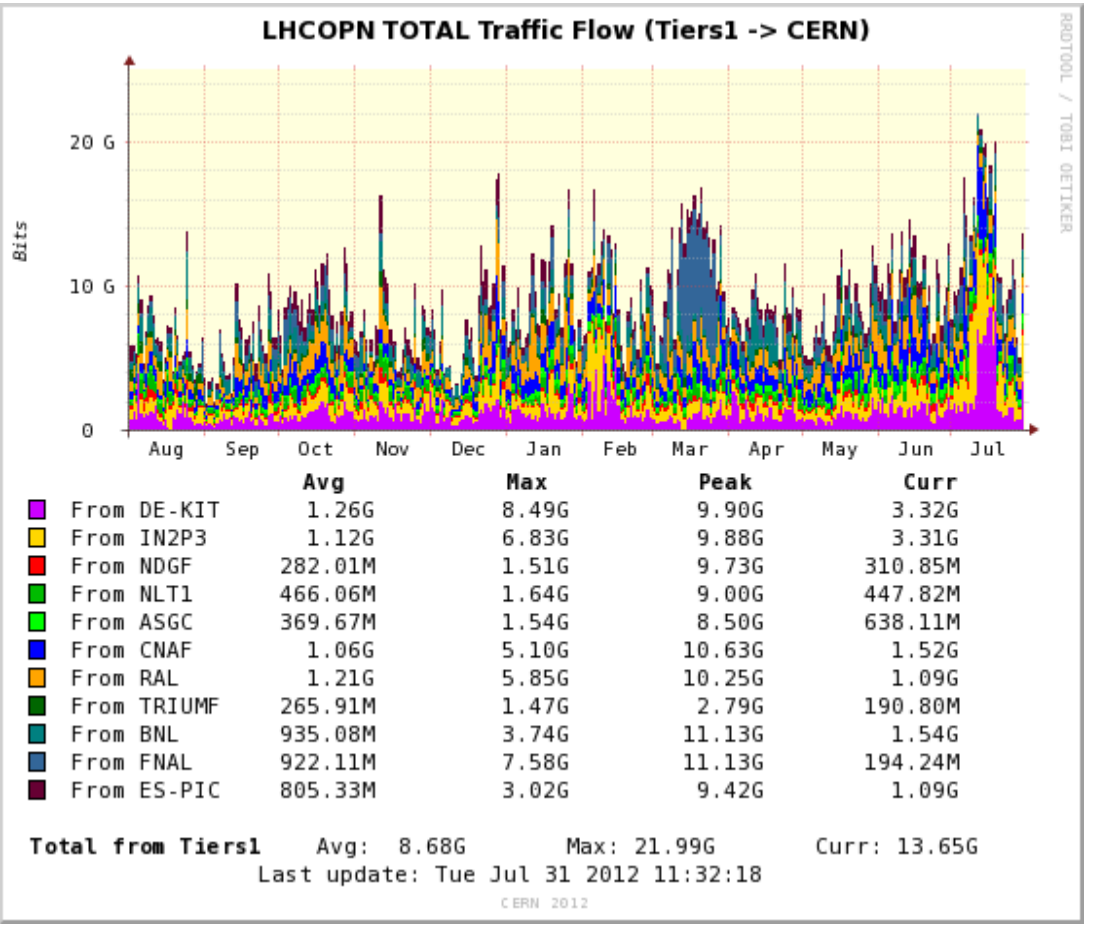
Links provided by the Research and Education network providers: Geant, USLHCnet, Esnet, Canarie, ASnet, Nordunet, Surfnet, GARR, Renater, JANET.UK, Rediris, DFN, SWITCH

Technology

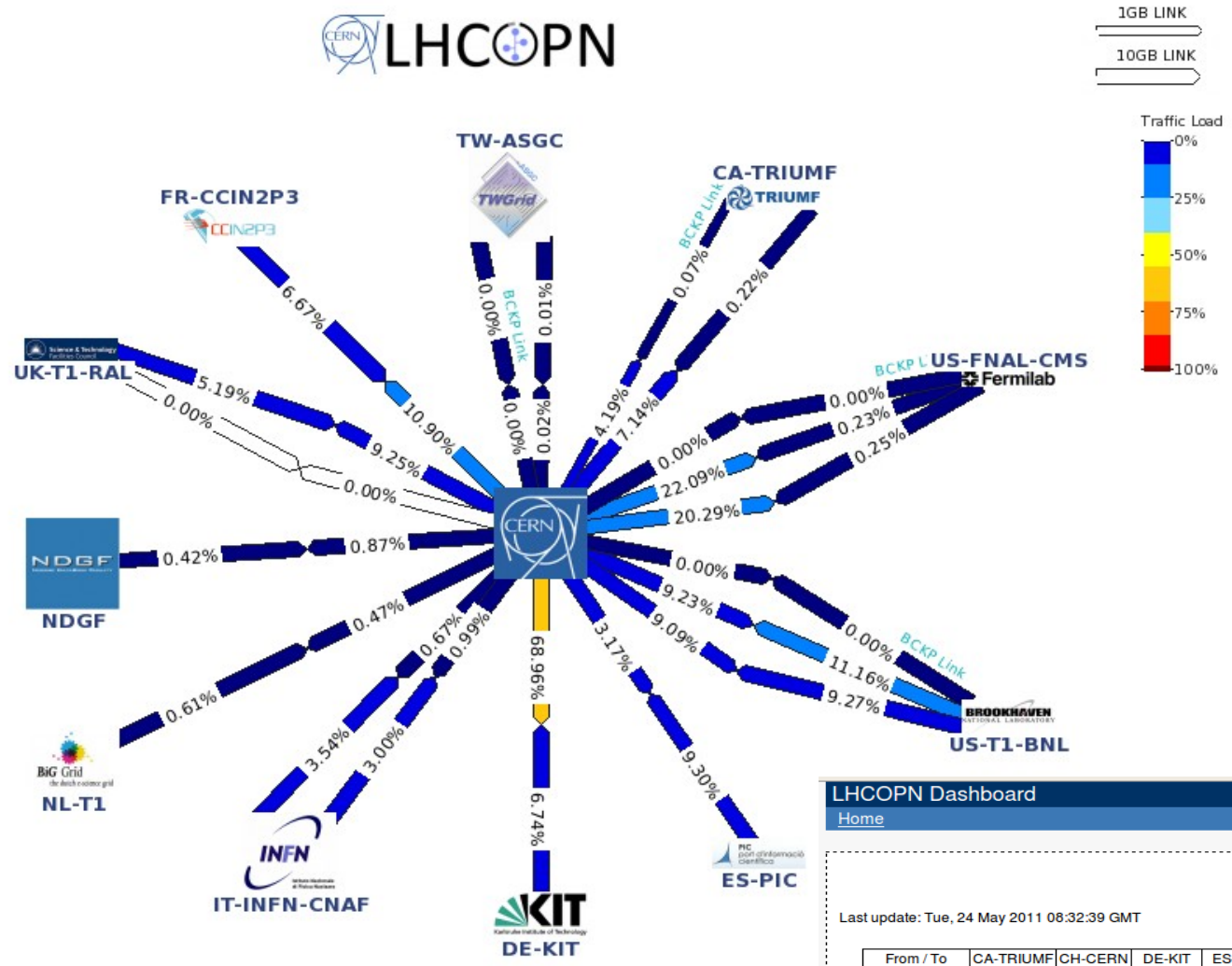
- Single and bundled long distance 10G ethernet links
- Multiple redundant paths. Star+PartialMesh topology
- BGP routing: communities for traffic engineering, load balancing.
- Security: only declared IP prefixes can exchange traffic.

Traffic to the Tier1s





Monitoring



LHCOPN Dashboard

Home

Current status

Last update: Tue, 24 May 2011 08:32:39 GMT

From / To	CA-TRIUMF	CH-CERN	DE-KIT	ES-PIC	FR-CCIN2P3	IT-INFN-CNAF	NDGF	NL-T1	TW-ASGC	UK-T1-RAL	US-FNAL-CMS	US-T1-BNL
CA-TRIUMF												
CH-CERN												
DE-KIT												
ES-PIC												
FR-CCIN2P3												
IT-INFN-CNAF												
NDGF												
NL-T1												
TW-ASGC												
UK-T1-RAL												
US-FNAL-CMS												
US-T1-BNL												

JSON feed

Legend: OK | Deviation from Baseline | Critical | Unknown

LHCONE

LHC Open Network Environment

Driving the change

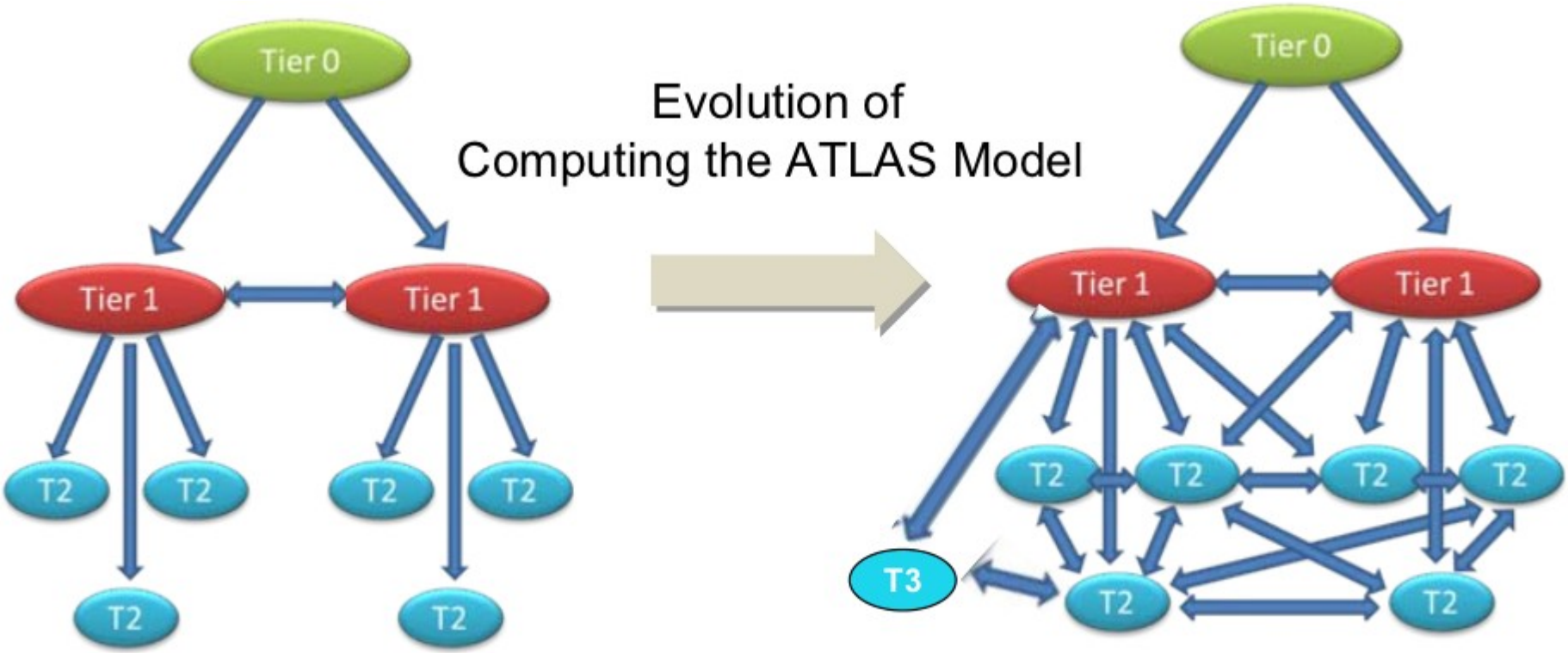
"The Network infrastructure is the most reliable service we have"

"Network Bandwidth (rather than disk) will need to scale more with users and data volume"

"Data placement will be driven by demand for analysis and not pre-placement"

Ian Bird, WLCG project leader

Change of computing model (ATLAS)



New computing model

- Better and more dynamic use of storage
- Reduce the load on the Tier1s for data serving
- Increase the speed to populate analysis facilities

Needs for a faster, predictable, pervasive network connecting Tier1s and Tier2s

Requirements from the Experiments

- Connecting any pair of sites, regardless of the continent they reside
- Bandwidth ranging from 1Gbps (Minimal), 5Gbps (Nominal), 10G and above (Leadership)
- Scalability: sites are expected to grow
- Flexibility: sites may join and leave at any time
- Predictable cost: well defined cost, and not too high

Needs for a better network

- more bandwidth by federating (existing) resources
- sharing cost of expensive resources
- accessible to any TierX site

=

LHC Open Network Environment

LHCONE concepts

- Serves any LHC sites according to their needs and allowing them to grow

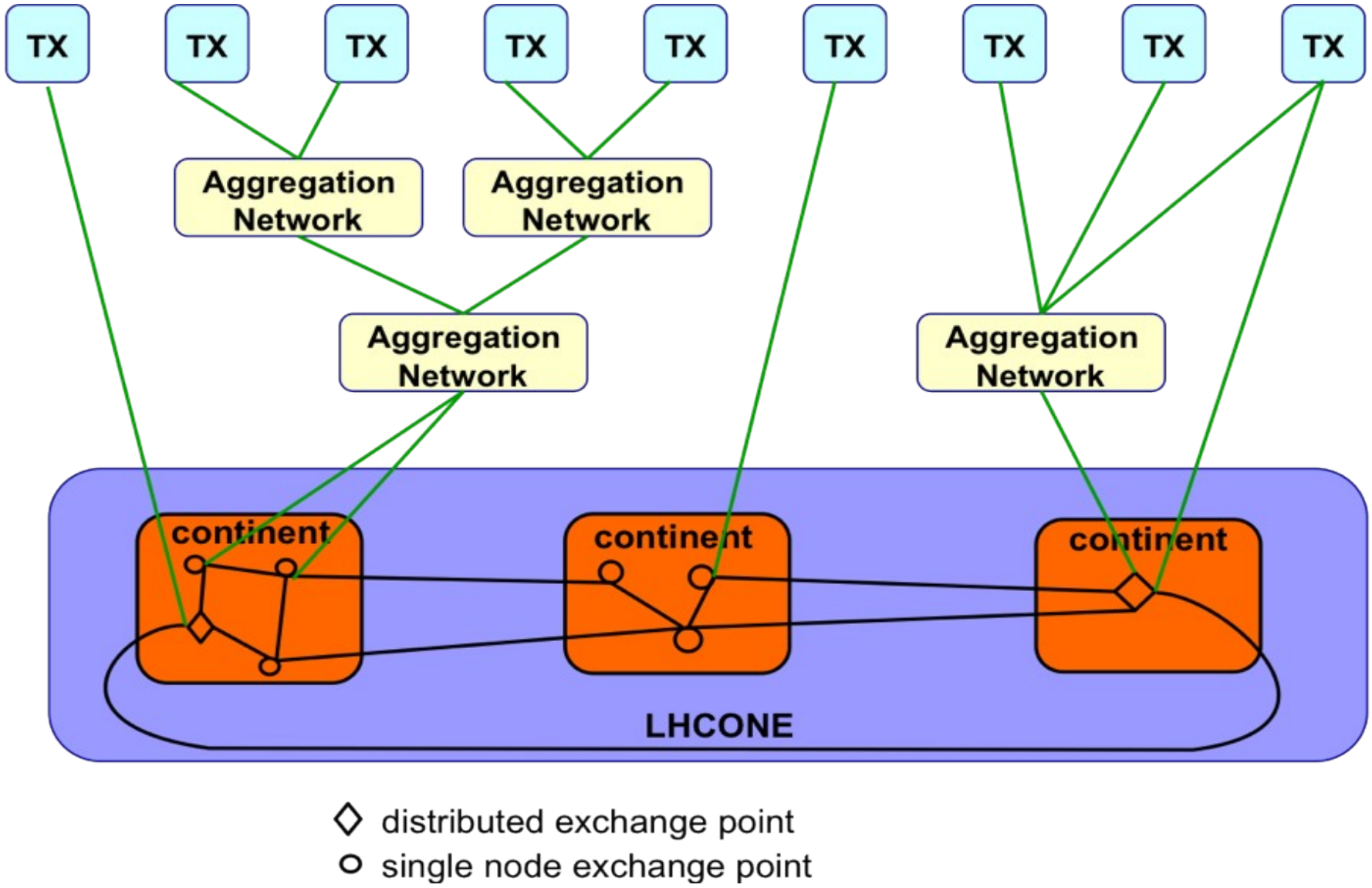
- A collaborative effort among Research & Education Network Providers

- Based on Open Exchange Points: easy to join, neutral

- Multiple services: one cannot fit all

- Traffic separation: no clash with other data transfer, resource allocated for and funded by HEP community

LHCONE architecture

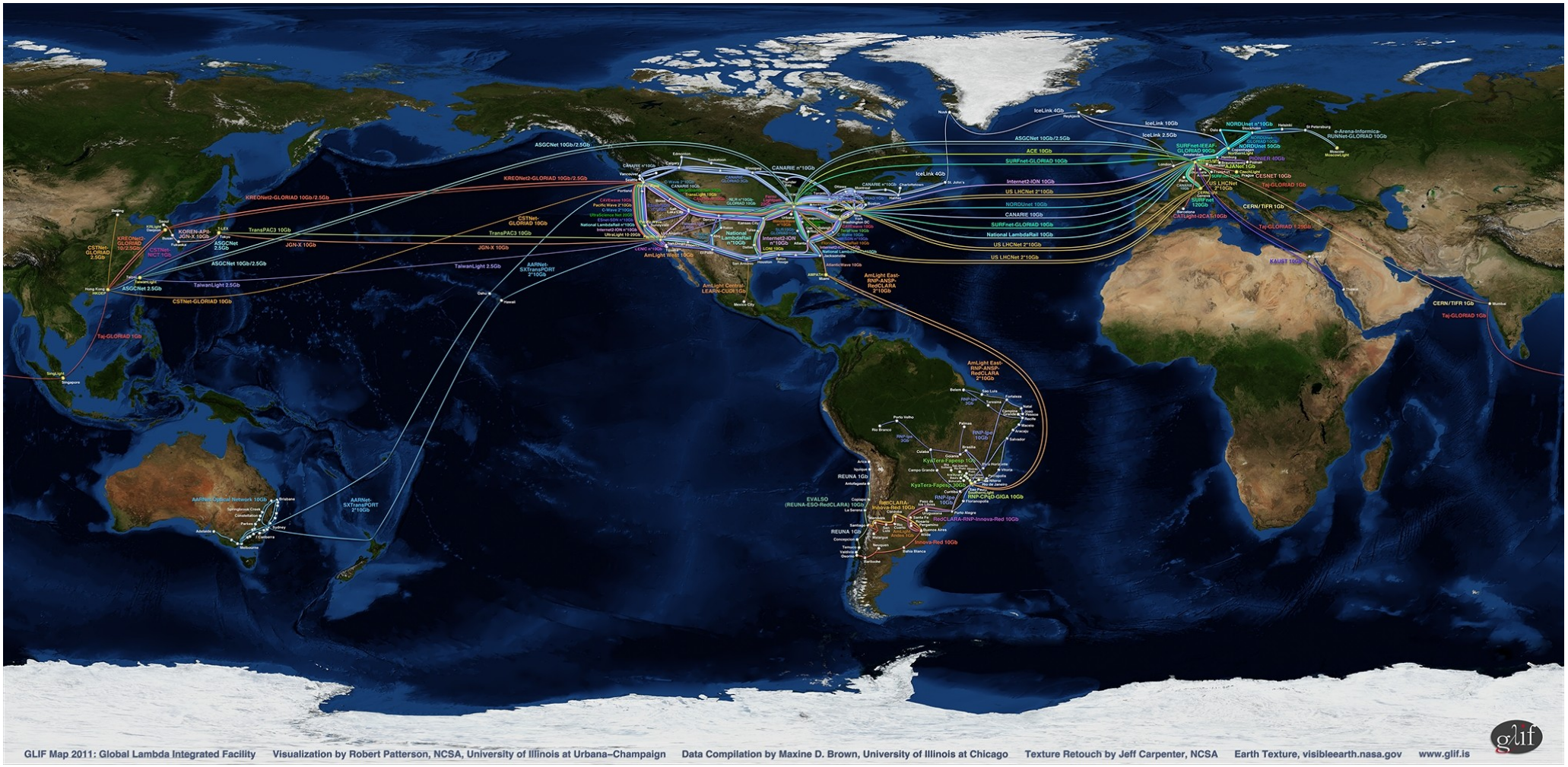


LHCONE building blocks

- **Single node Exchange Points**
- **Continental/regional Distributed Exchange Points**
- **Interconnect circuits between Exchange Points**

These exchange points and the links in between collectively provide LHCONE services and operate under a common LHCONE policy

The underlying infrastructure



LHCONE services

- Layer3 VPN
- Point-to-Point links
- Monitoring

Openlab and IT-CS

Openlab project:
CINBAD

CINBAD

CERN **I**nvestigation of **N**etwork **B**ehaviour and **A**nomaly **D**etection

Project Goals:

Understand the behaviour of large computer networks (10'000+ nodes) in High Performance Computing or large Campus installations to be able to:

- detect traffic anomalies in the system
- perform trend analysis
- automatically take counter measures
- provide post-mortem analysis facilities

Resources:

- In collaboration with HP Networking
- Two Engineers in IT-CS

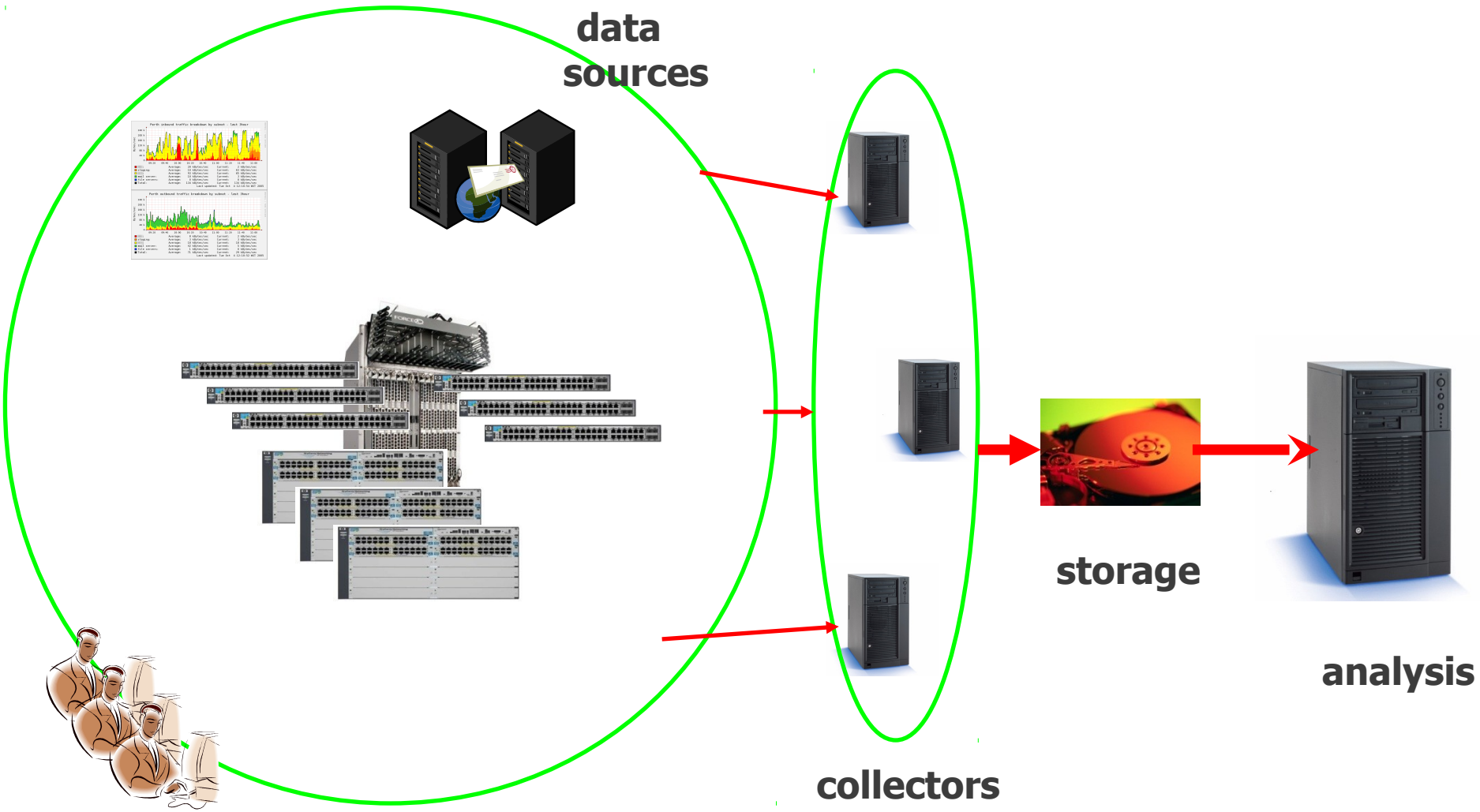
Results

Project completed in 2010

For CERN:
Designed and deployed a complete framework (hardware and software) to detect anomalies in the Campus Network (GPN)

For HP:
Intellectual properties of new technologies used in commercial products

CINBAD Architecture



Openlab project:
WIND

WIND

Wireless **I**nfrastructure **N**etwork **D**eployment

Project Goals

- Analyze the problems of large scale wireless deployments and understand the constraint
- Simulate behaviour of WLAN
- Develop new optimisation algorithms

Resources:

- In collaboration with HP Networking
- Two Engineers in IT-CS
- Started in 2010

Needs

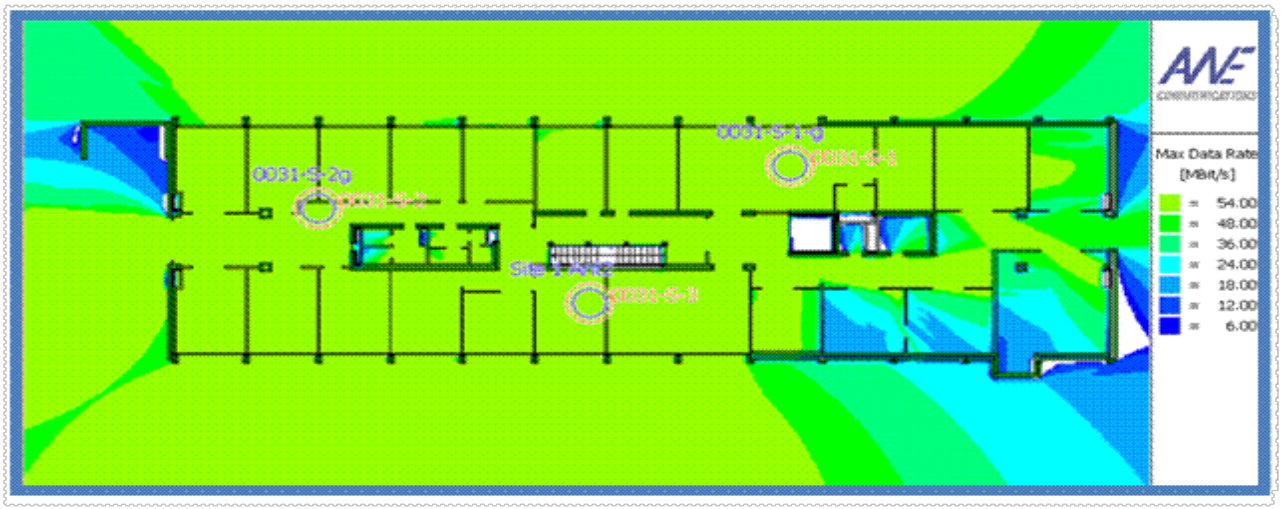
Wireless LAN (WLAN) deployments are problematic:

- Radio propagation is very difficult to predict
- Interference is an ever present danger
- WLANs are difficult to properly deploy
- Monitoring was not an issue when the first standards were developed
- When administrators are struggling just to operate the WLAN, performance optimisation is often forgotten

Example: Radio interferences

Max data rate in 0031-S: The APs work on the same channel

Max data rate in 0031-S: The APs work on 3 independent channels



Expected results

Extend monitoring and analysis tools

Act on the network
- smart load balancing
- isolating misbehaving clients
- intelligent minimum data rates

More accurate troubleshooting

Streamline WLAN design

Openlab project:
ViSION

ViSION

Project Goals:

- Develop a SDN traffic orchestrator using OpenFlow

Resources:

- In collaboration with HP Networking
- Two Engineers in IT-CS
- Started in 2012

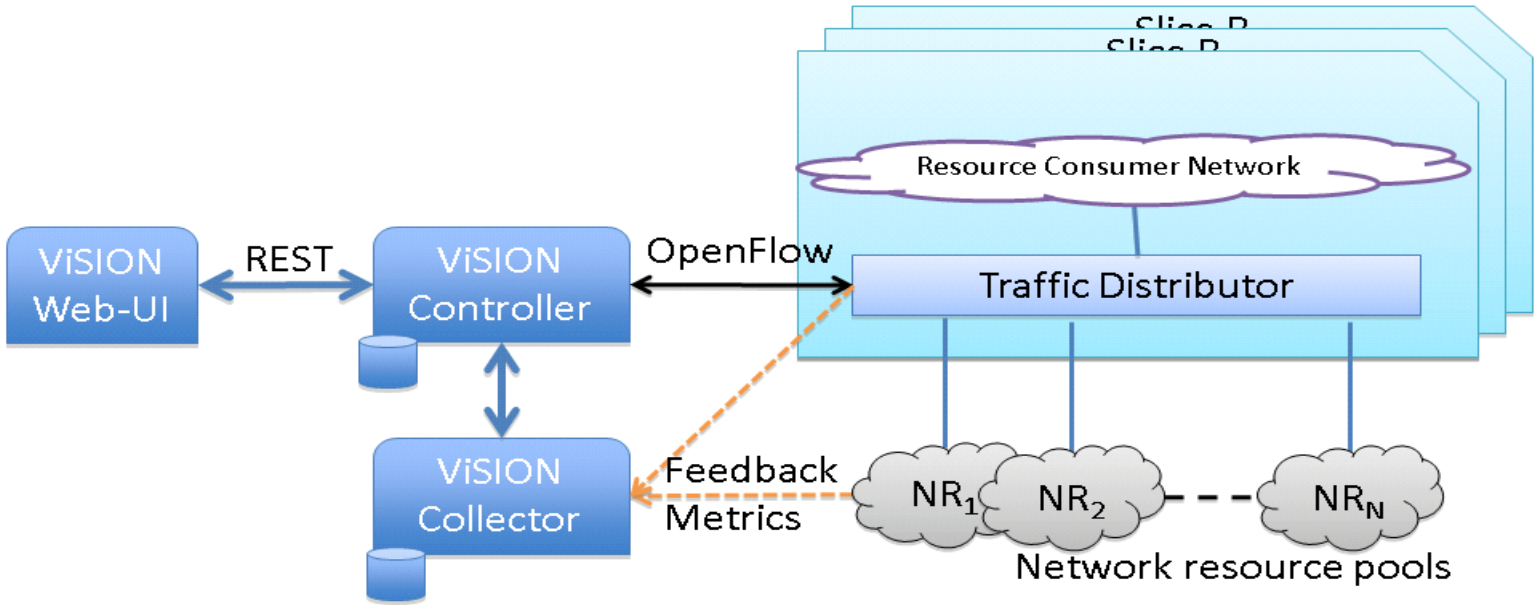
Goals

SDN traffic orchestrator using OpenFlow:

- distribute traffic over a set of network resources
- perform classification (different types of applications and resources)
- perform load sharing (similar resources).

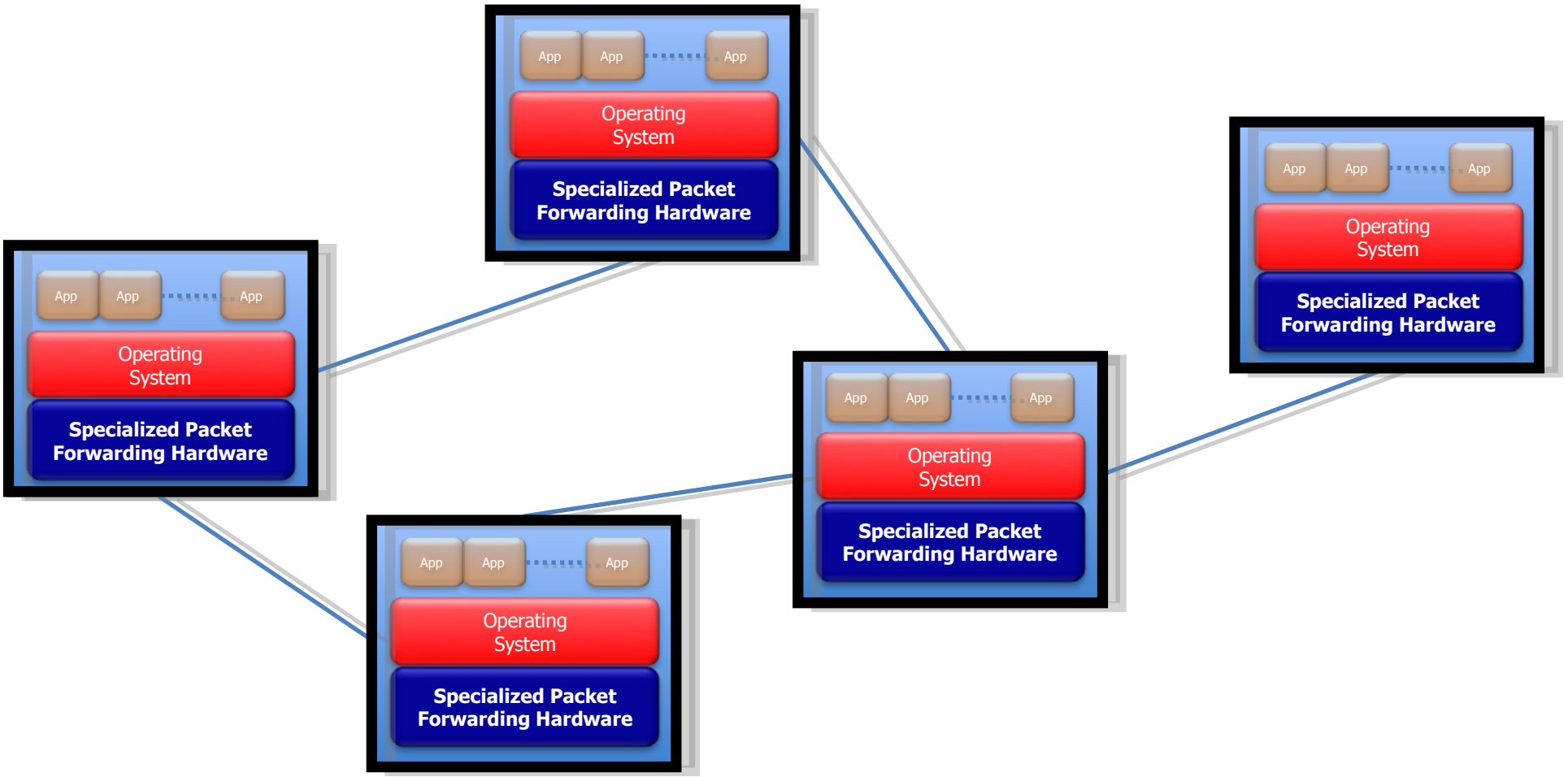
Benefits:

- improved scalability and control than traditional networking technologies

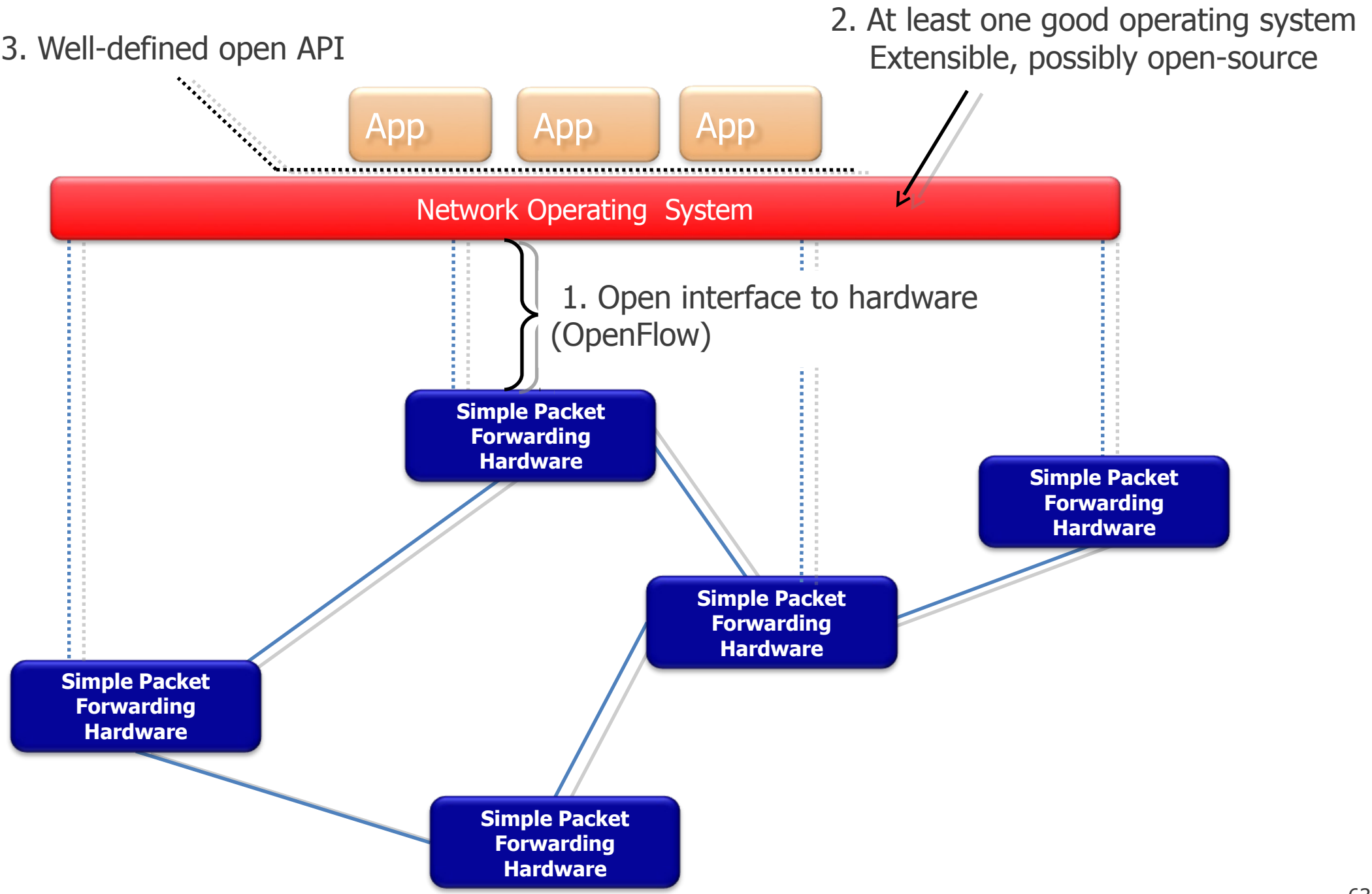


From traditional networks...

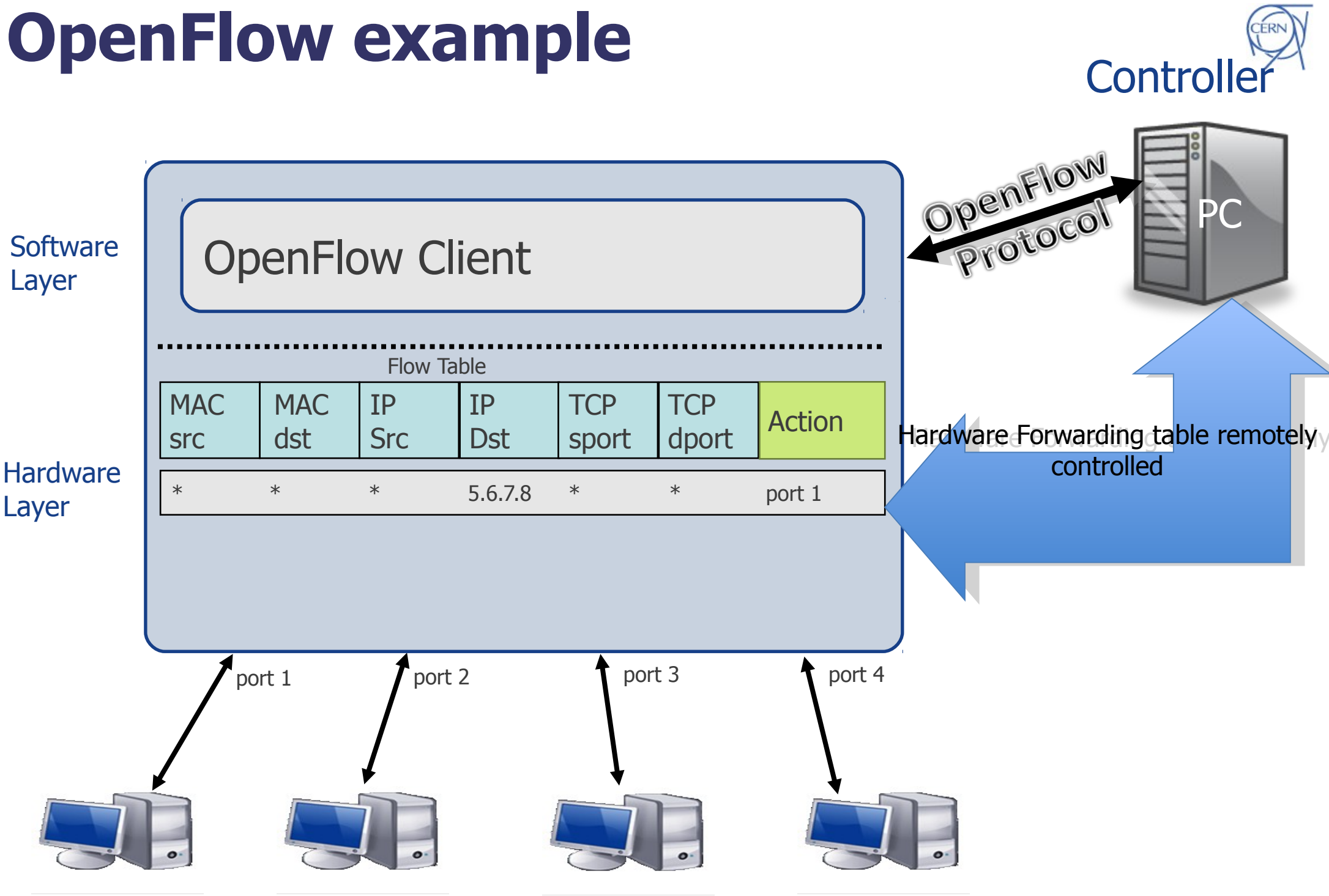
Closed boxes, fully distributed protocols



.. to Software Defined Networks (SDN)



OpenFlow example



Conclusions

Conclusions

- The Data Network is an essential component of the LHC instrument
- The Data Network is a key part of the LHC data processing and will become even more important
- More and more security and design challenges to come

Credits

Artur Barczyk (LHCONE)

Dan Savu (VISION)

Milosz Hulboj (WIND and CINBAD)

Ryszrard Jurga (CINBAD)

Sebastien Ceuterickx (WIND)

Stefan Stancu (VISION)

Vlad Lapadatescu (WIND)

What's next

SWAN: Space Wide Area Network :-)