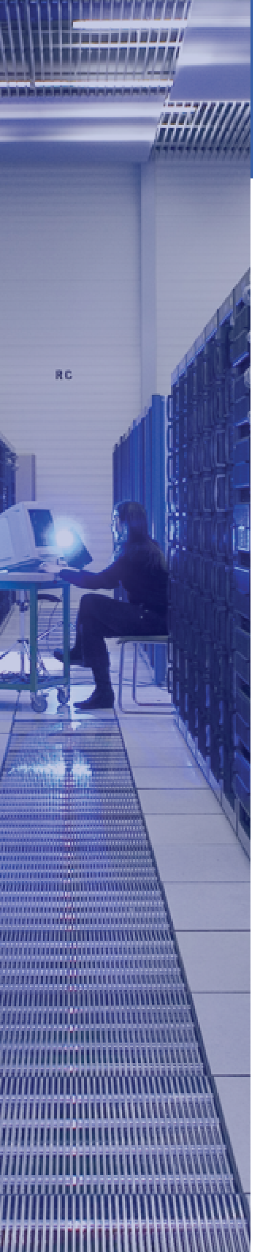
LHC Networking
and Distributed Science

David Foster

India
March 2012

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



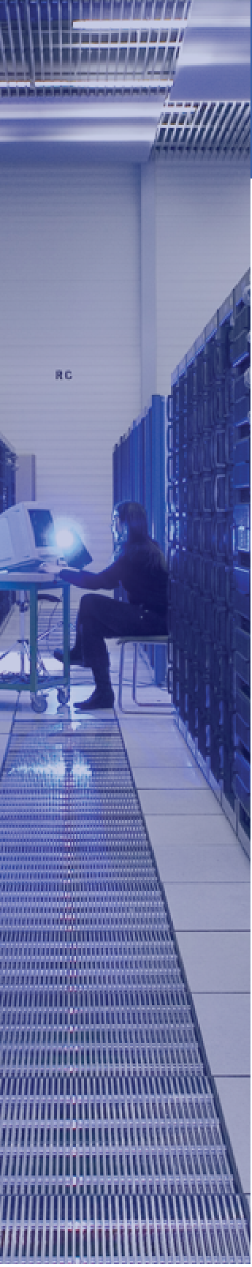
Agenda

- LCG Evolution
- LHCOPN 2004-Today
- LHCONE 2011-Today
- Distributed Science beyond the LHC

With thanks to many people for material

Bob Jones, Alberto Di Meglio, Bill Johnston, Harvey Newman, Ian Bird, Artur Barczyk and others.

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

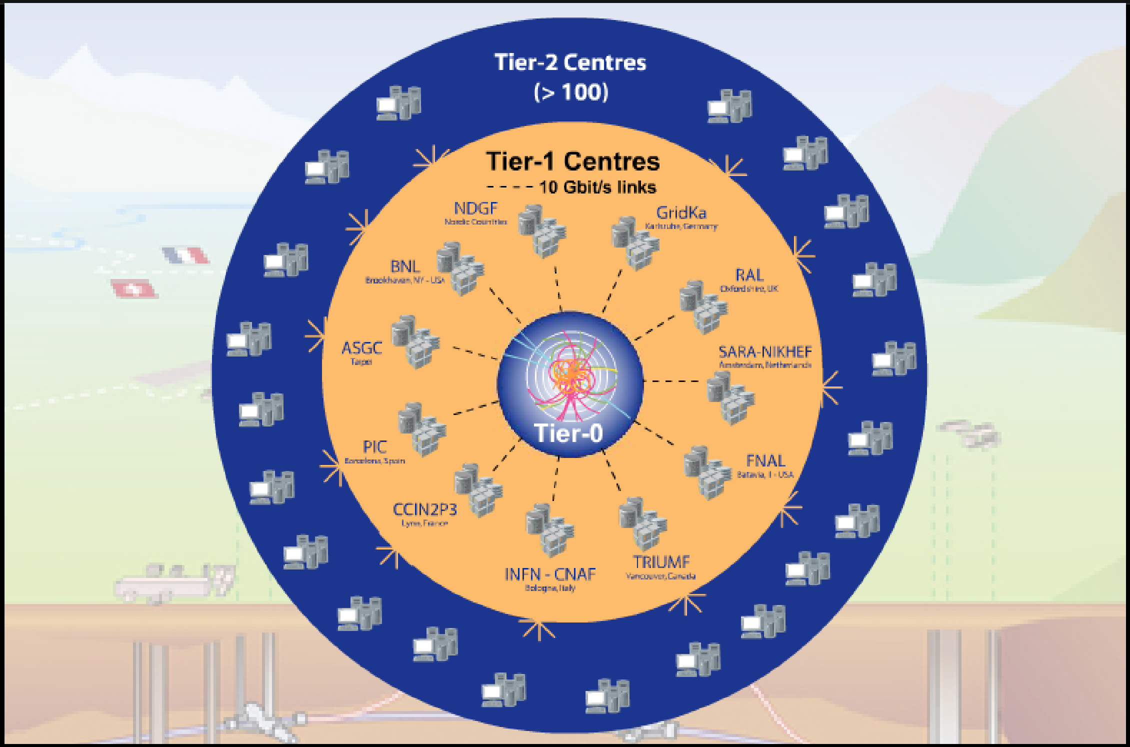


LCG Evolution

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

WLCG – what and why?

- A distributed computing infrastructure to provide the production and analysis environments for the LHC experiments
- Managed and operated by a worldwide collaboration between the experiments and the participating computer centres
- The resources are distributed – for funding and sociological reasons
- Our task was to make use of the resources available to us – no matter where they are located



Tier-0 (CERN):
- Data recording
- Initial data reconstruction
- Data distribution

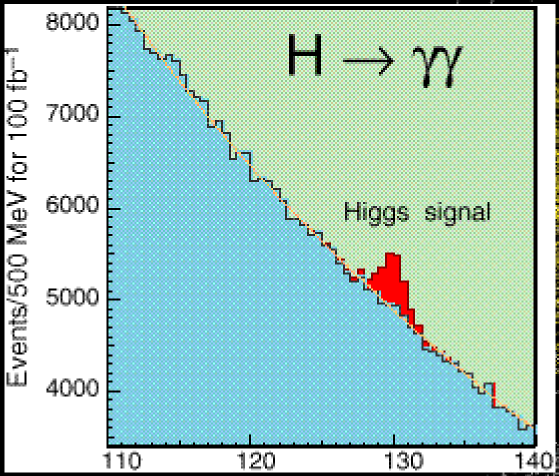
Tier-1 (11 centres):
- Permanent storage
- Re-processing
- Analysis

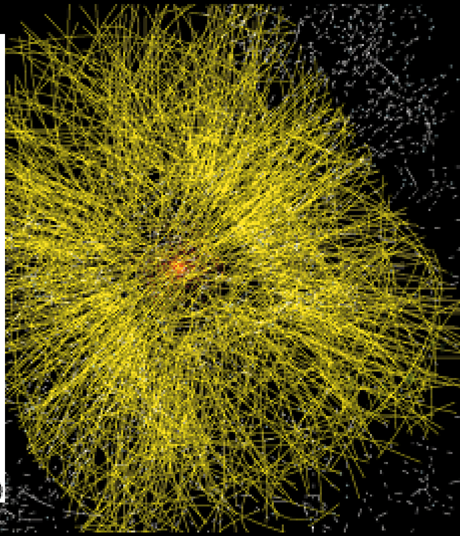
Tier-2 (~130 centres):
- Simulation
- End-user analysis

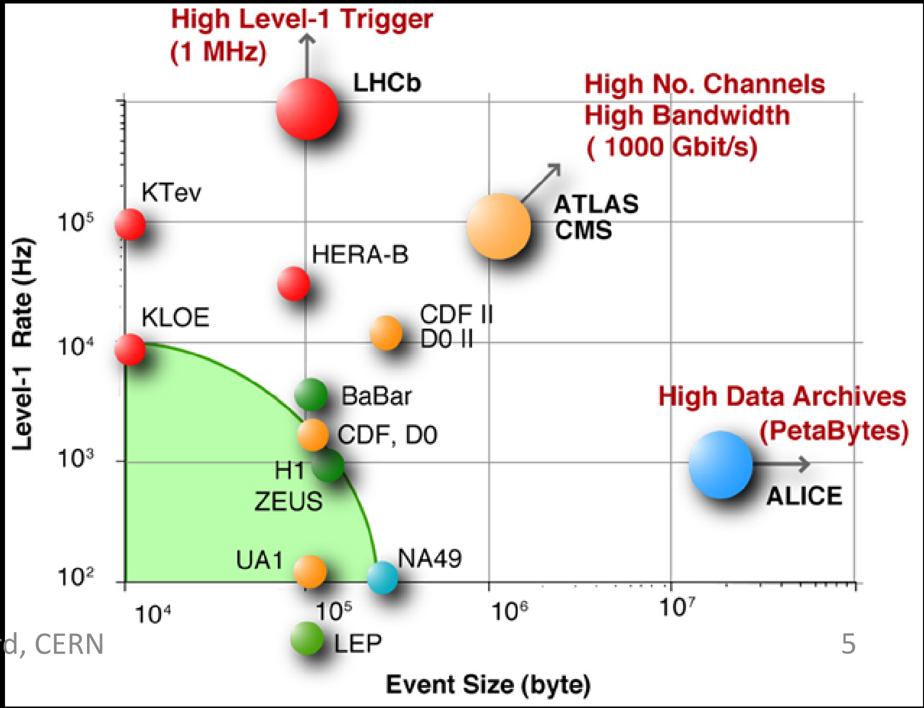
WLCG Worldwide LHC Computing Grid

The LHC Computing Challenge

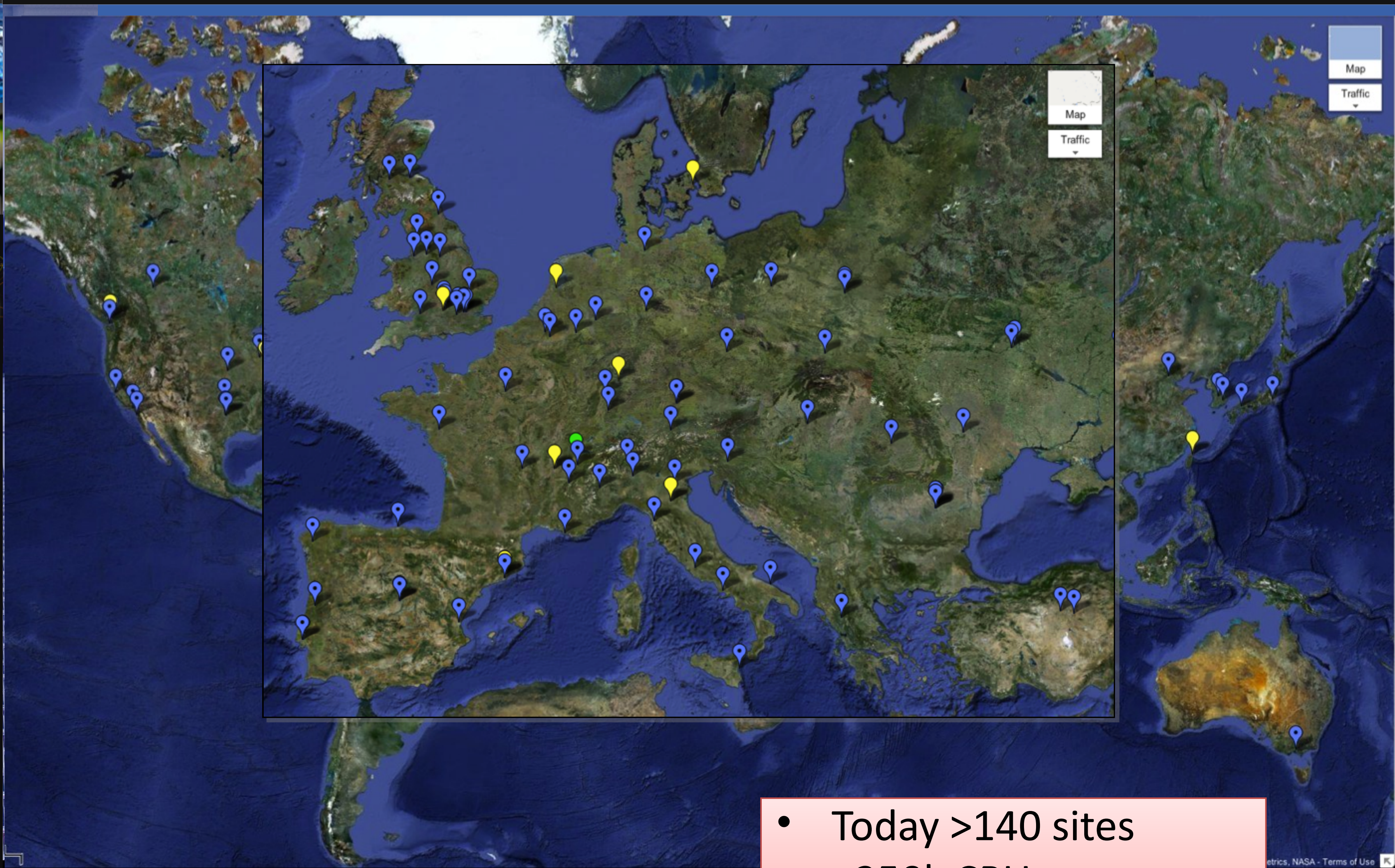
- Signal/Noise: 10^{-13} (10^{-9} offline)
- Data volume
  - High rate * large number of channels * 4 experiments
  - → 15 PetaBytes of new data each year → 23 PB in 2011
- Compute power
  - Event complexity * Nb. events * thousands users
  - → 200 k CPUs → 250 k CPU
  - → 45 PB of disk storage → 150 PB
- Worldwide analysis & funding
  - Computing funding locally in major regions & countries
  - Efficient analysis
  - → GRID technology







WLCG Grid Sites



Tier 0 · Tier 1 · Tier 2

- Today >140 sites
- >250k CPU cores
- >150 PB disk

WLCG
Worldwide LHC Computing Grid
Department
Genève 23
Switzerland
w.cern.ch/it

WLCG Collaboration Status

Tier 0; 11 Tier 1s; 68 Tier 2 federations

Today we have 50 MoU signatories, representing 35 countries:

Australia, Austria, Belgium, Brazil, Canada, China, Czech Rep, Denmark, Estonia, Finland, France, Germany, Hungary, Italy, India, Israel, Japan, Rep. Korea, Netherlands, Norway, Pakistan, Poland, Portugal, Romania, Russia, Slovenia, Spain, Sweden, Switzerland, Taipei, Turkey, UK, Ukraine, USA.

March 2012: Accepted KISTI (S.Korea) as first new associate Tier 1

CPU – around the Tiers

- The grid really works
- All sites, large and small can contribute
  - And their contributions are needed!

CPU delivered - January 2011

CERN
BNL
CNAF
KIT
NL LHC/Tier-1
RAL
FNAL
CC-IN2P3
ASGC
PIC
NDGF
TRIUMF
Tier 2

Tier 2 CPU delivered by country - January 2011

USA
France
Italy
Spain
Poland
Slovenia
China
Japan
Israel
Belgium
Hungary
Australia
Norway
Ukraine
India
Estonia
Greece
UK
Germany
Russian Federation
Canada
Switzerland
Czech Republic
Portugal
Sweden
Romania
Austria
Taipei
Republic of Korea
Turkey
Finland
Pakistan
Brazil

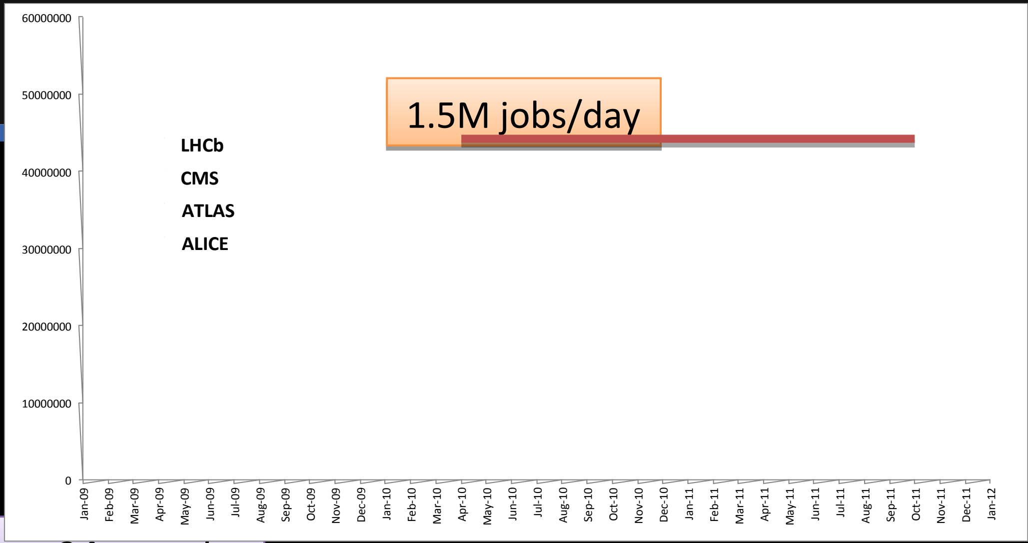
WLCG
Worldwide LHC Computing Grid

Department
Genève 23
Switzerland
w.cern.ch/it

Grid Usage

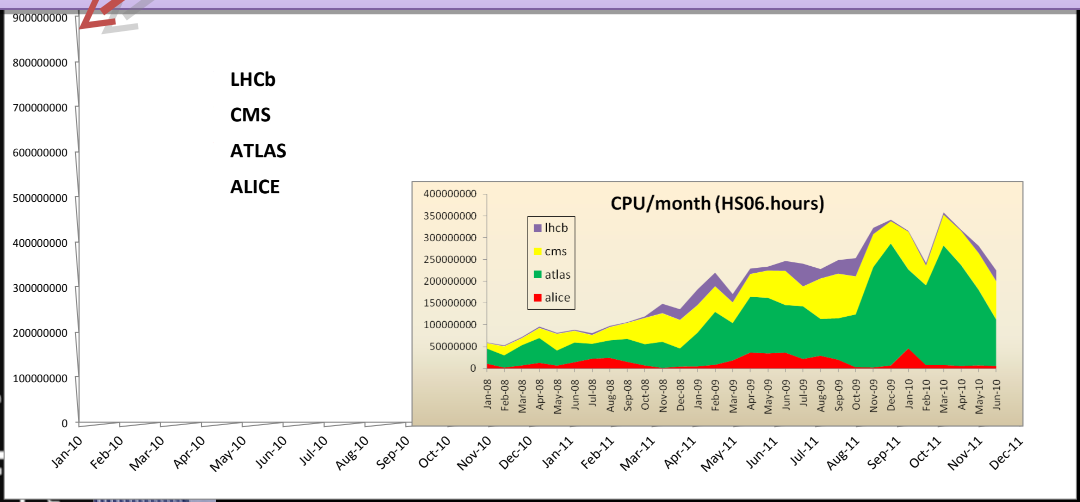
Use remains consistently high:

- >1.5 M jobs/day;
- ~150k CPU



10^9 HEPSPEC-hours/month (~150 k CPU continuous use)

CPU used at Tier 1s + Tier 2s (HS06.hrs/month) – last 24 months



At the end of 2010 we saw all Tier 1 and Tier 2 job slots being filled

CPU usage now >> double that of mid-2010 (inset shows build up over previous years)

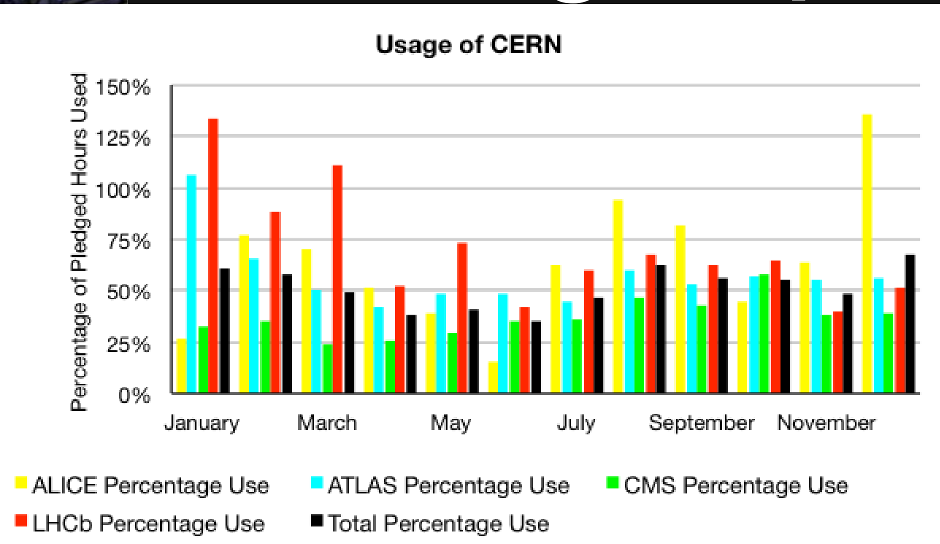
As well as LHC data, large simulation productions always ongoing

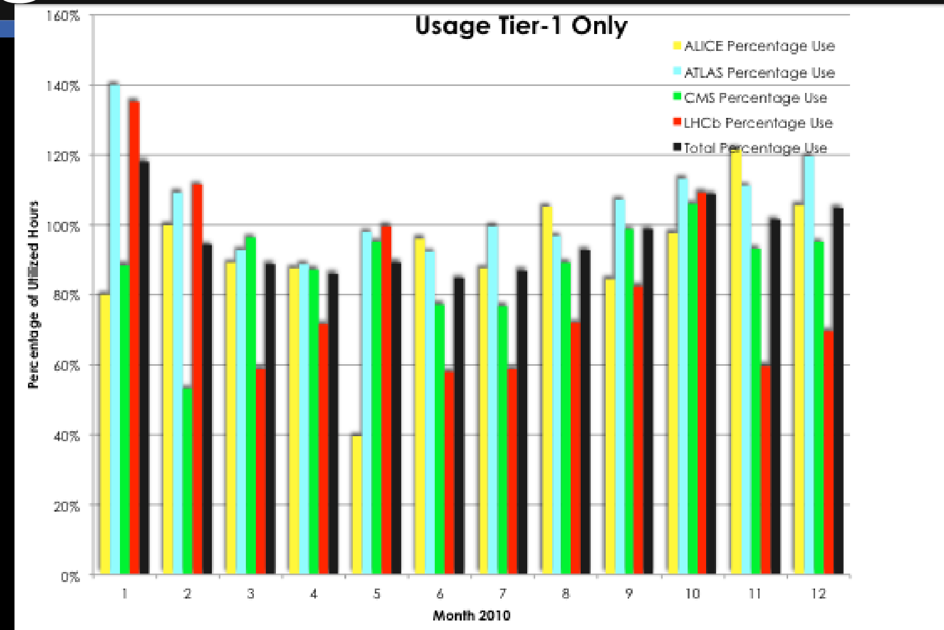
Large numbers of analysis users:

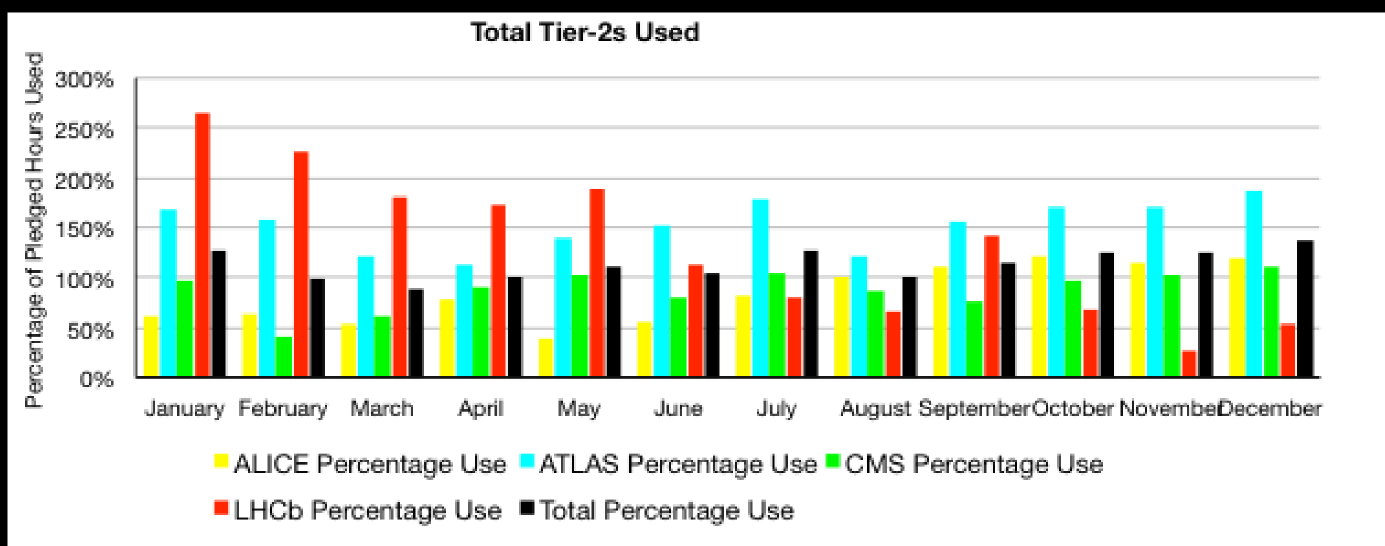
- ATLAS, CMS ~1000
- LHCb,ALICE ~250

In 2011 WLCG delivered ~ 150 CPU-millennia!

Tiers usage vs pledges



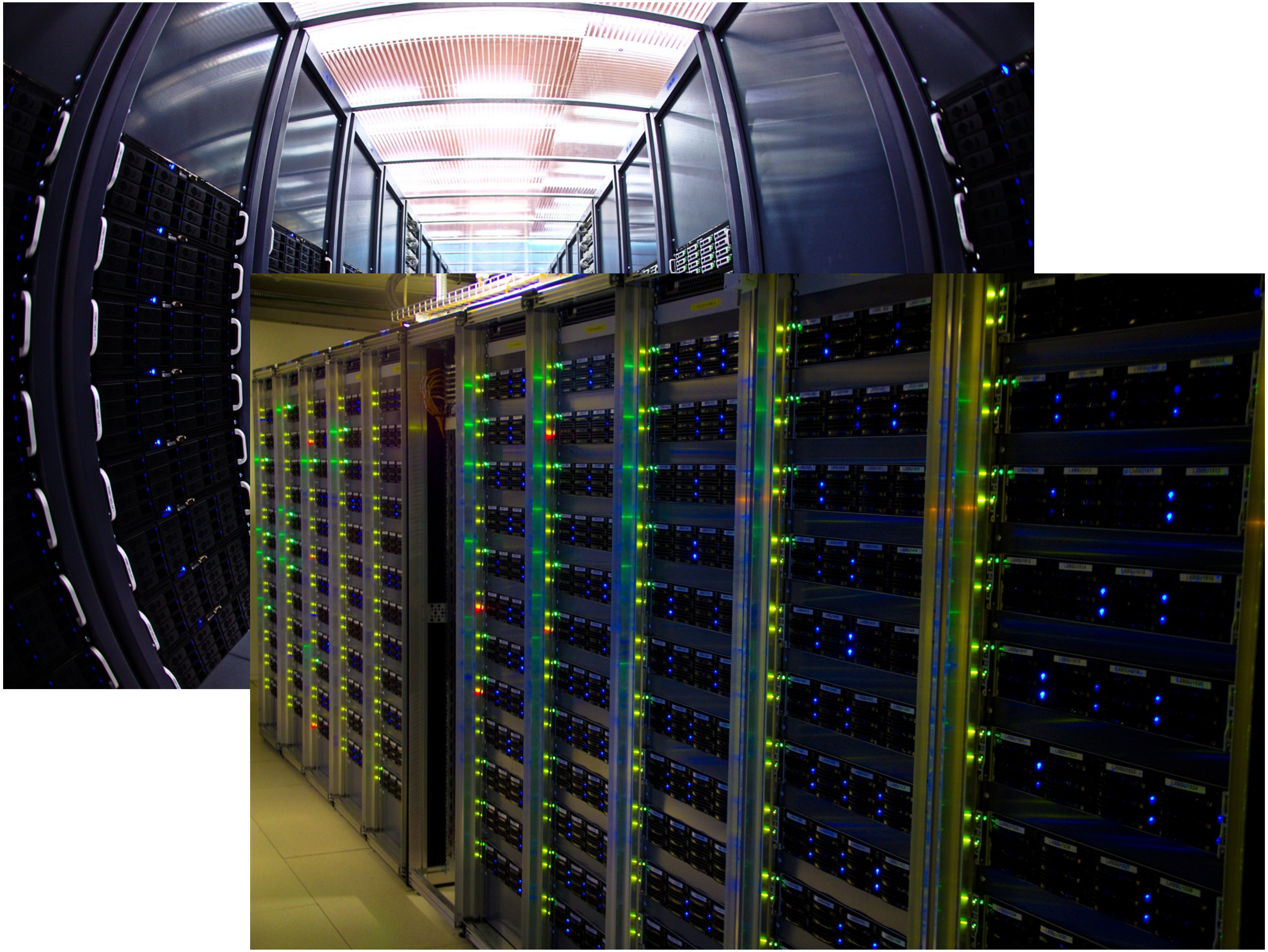




WLCG
Worldwide LHC Computing Grid

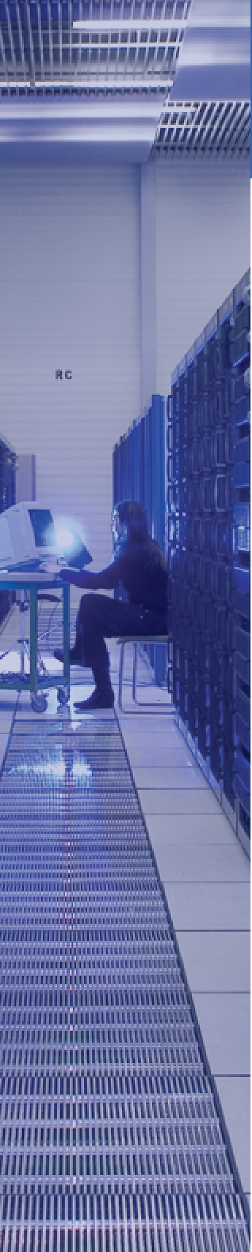
Department
Genève 23
Switzerland
w.cern.ch/it

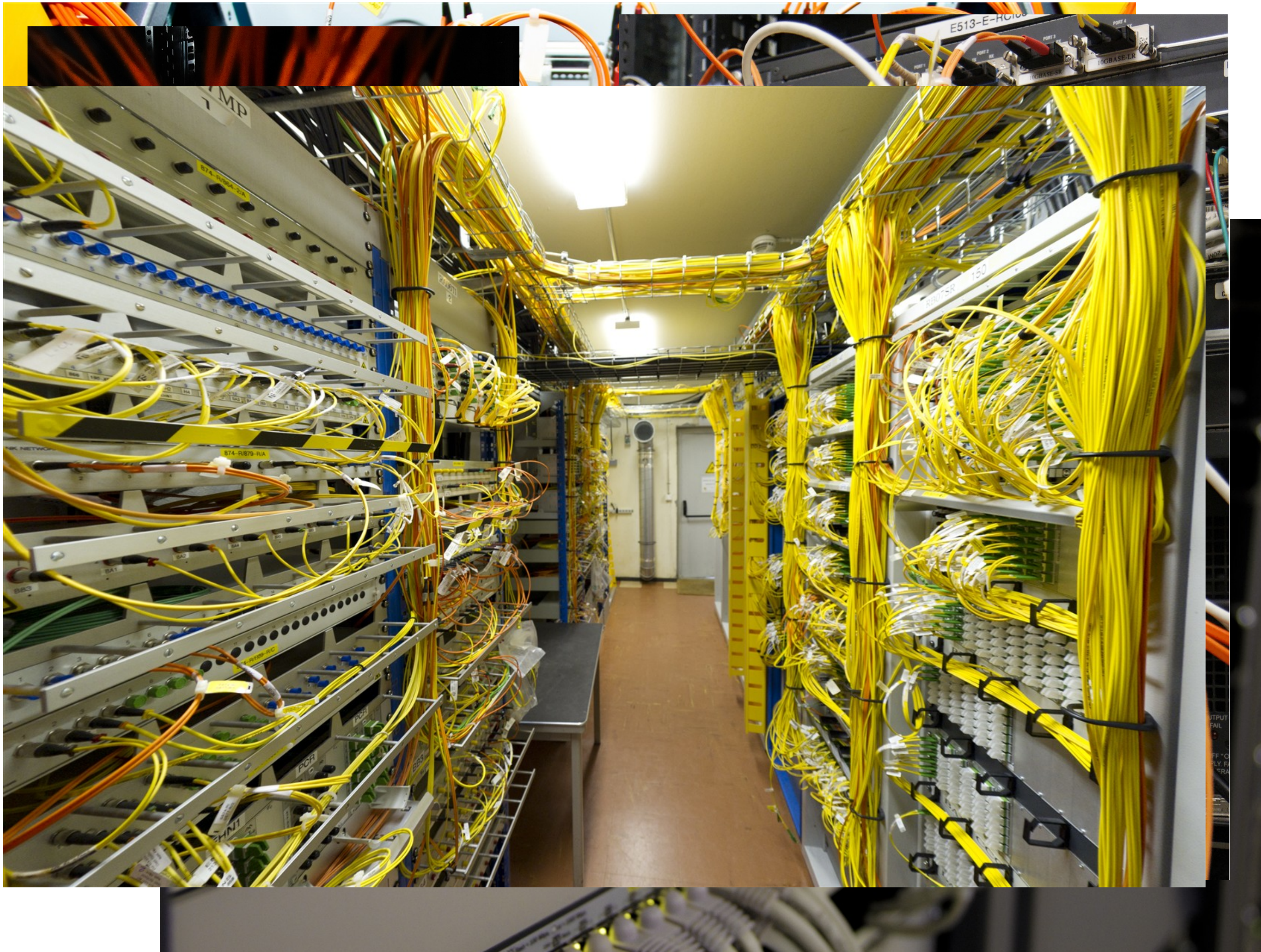
Growth of CPU and Disk capacity



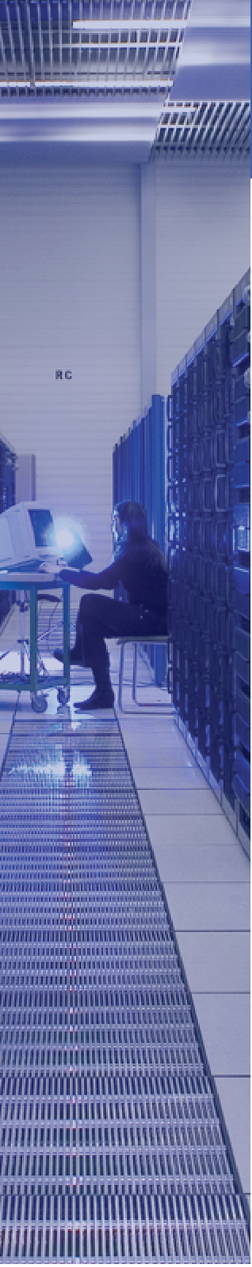
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Networks are fundamental





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

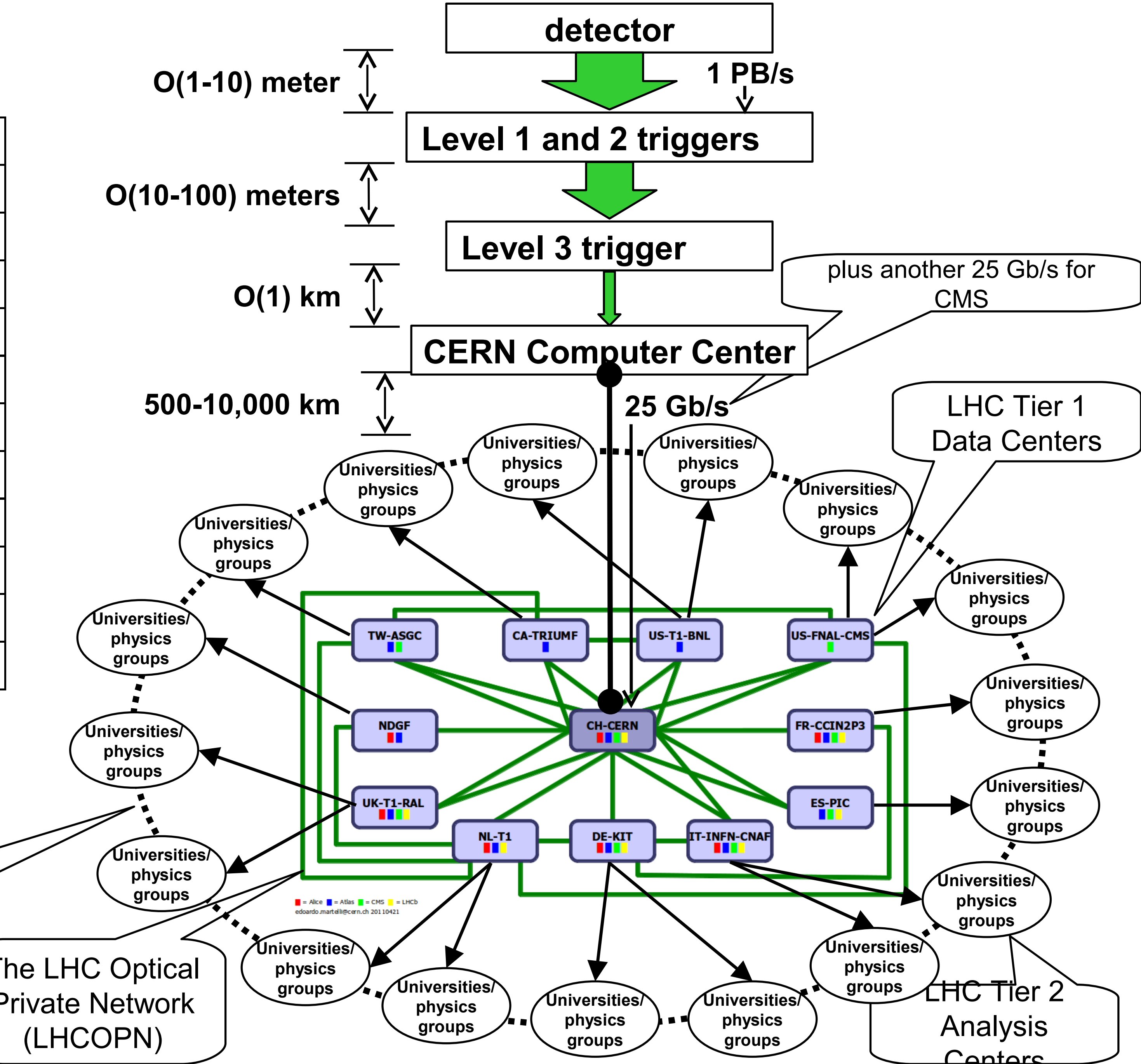


LHCOPN

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

LHC/ATLAS data flow model

CERN →T1	miles	kms
France	350	565
Italy	570	920
UK	625	1000
Netherlands	625	1000
Germany	700	1185
Spain	850	1400
Nordic	1300	2100
USA – New York	3900	6300
USA - Chicago	4400	7100
Canada – BC	5200	8400
Taiwan	6100	9850



LHCOPN Architecture 2005

CERN IT Department

- From the LHCOPN Architecture Document

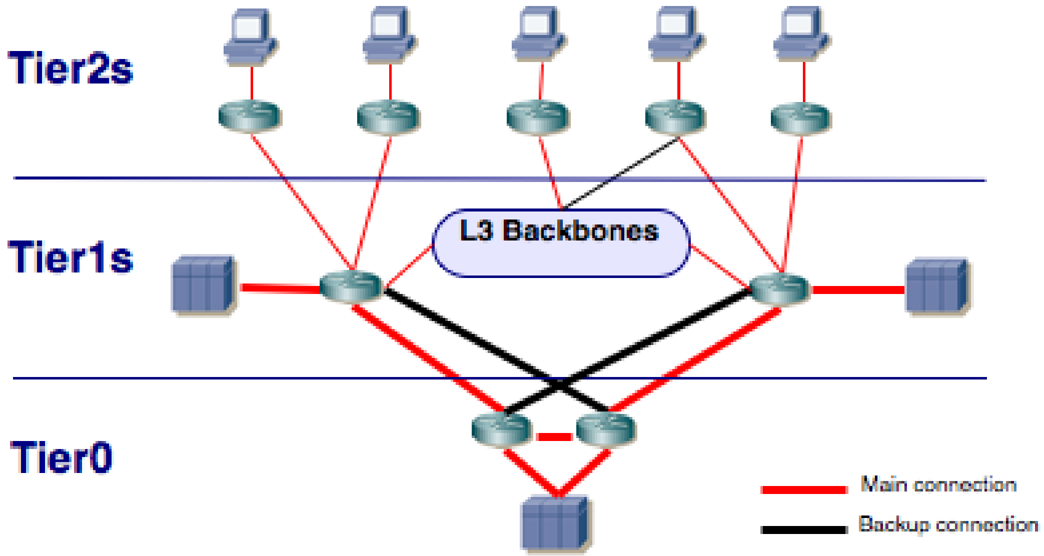
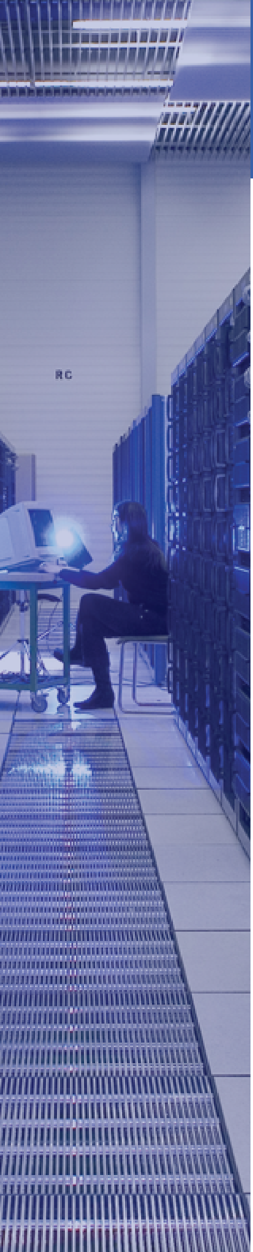


Figure 2. Network architecture

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

Some of the early documents 2004-2005

CERN IT Department



Community Network Infrastructure for LHC Data Movement
David Foster CERN IT-CS

Introduction

The document provides a summary of the requirements for the LHC data transport according to the information known at the present. It does not provide an exhaustive justification or explanation of the various options. These have been covered in many other related documents.

The Distributed Data Center

The LHC data processing infrastructure is based on a grid paradigm where resources for LHC data processing is provided by collaborating institutes world wide. These resources comprise CPU and Data storage primarily. Units of work, jobs, will be distributed to these remote locations, be executed and the data stored. This clearly raises a number of complex policy issues that will not be discussed here. However, for the purposes of this document we can assume that data movement and replication will

A COMMUNITY NETWORK FOR LHC DATA TRANSFER

A DOCUMENT SUBMITTED TO THE LCG AND TIER-1 CENTERS FOR DISCUSSION

DOCUMENT VERSION: DRAFT 1.6

DAVID FOSTER CERN
OLIVIER MARTIN CERN
PETER CLARKE UCL
HARVEY NEWMAN CALTECH

LHC high-level network architecture

produced by LHC high-level architecture group
with contribution and comments from many

Authors:
Erik-Jan Bos (SURFnet),
Edoardo Martelli (CERN), Paolo Moroni (CERN)

Editor:
David Foster (CERN)

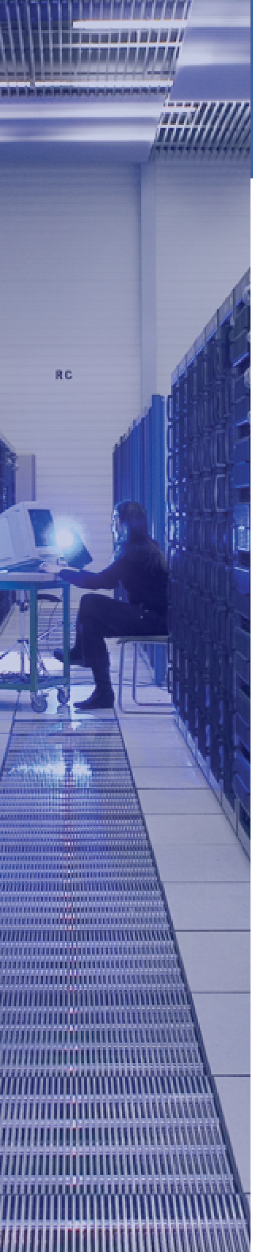
Grid Deployment Board - GDB

Summary of the first Tier 0/1 Network Meeting Held January 20/21 2005

David Foster
david.foster@cern.ch

Original Milestones 2004-2006

CERN IT Department



	Midyear	Endyear
2004	• 10Gbit "end-to-end" tests with a US Tier-1 (Fermilab?) • First version of the LHC Community Network proposal	• 10Gbit "end-to-end" test complete with European Partner • Measure performance variability and understand H/W and S/W Issues to ALL sites. • Document circuit switched options and costs, first real test if possible.
2005	• Circuit/Packet switch design completed. • LHC Community network proposal completed. • All T1 Fabric architecture documents completed. • LCG TDR completed	• Sustained throughput test achieved to some sites: 2-4 Gb/sec for 2 months. H/W and S/W problems solved.
2006	• All CERN b/w provisioned. • All T1 bandwidth in production (10Gb links) • Sustained throughput tests achieved to most sites.	• Verified performance to all sites for at least 2 months.

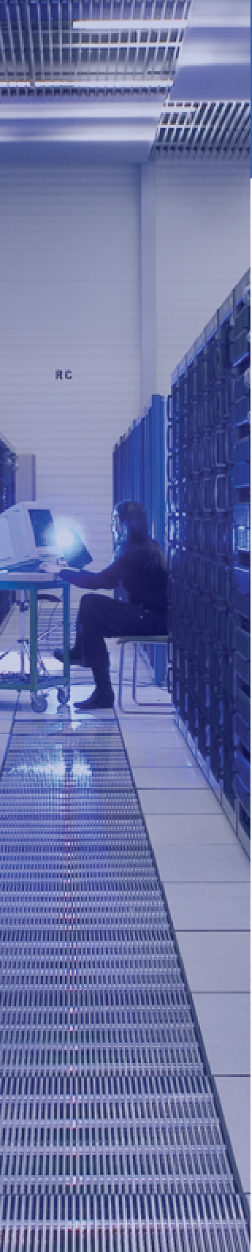
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

HEP Bandwidth Roadmap for Major Links (in Gbps): US LHCNet Example

Harvey Newman Caltech

Year	Production	Experimental	Remarks
2001	0.155	0.622-2.5	SONET/SDH
2002	0.622	2.5	SONET/SDH DWDM; GigE Integ.
2003	2.5	10-20	DWDM; 1 + 10 GigE Integration
2005-6	10-20	2-10 X 10	λ Switch; λ Provisioning
2007-8	3-4 X 10	~10 X 10; 100 Gbps	1st Gen. λ Grids
2009-10	6-8 X 10	~20 X 10 or ~2 X 100	100 Gbps λ Switching
2011-12	~20 X 10 or 2 X 100	~10 X 100	2nd Gen λ Grids Terabit Networks
2013-5	~Terabit	~MultiTbps	~Fill One Fiber

Paralleled by ESnet Roadmap for Data Intensive Sciences

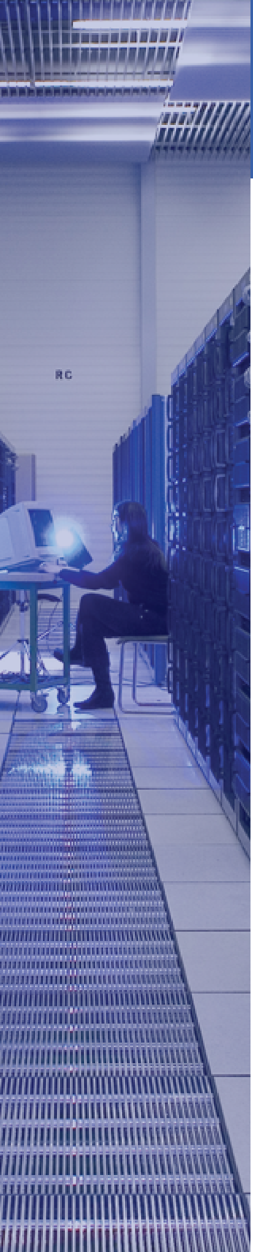


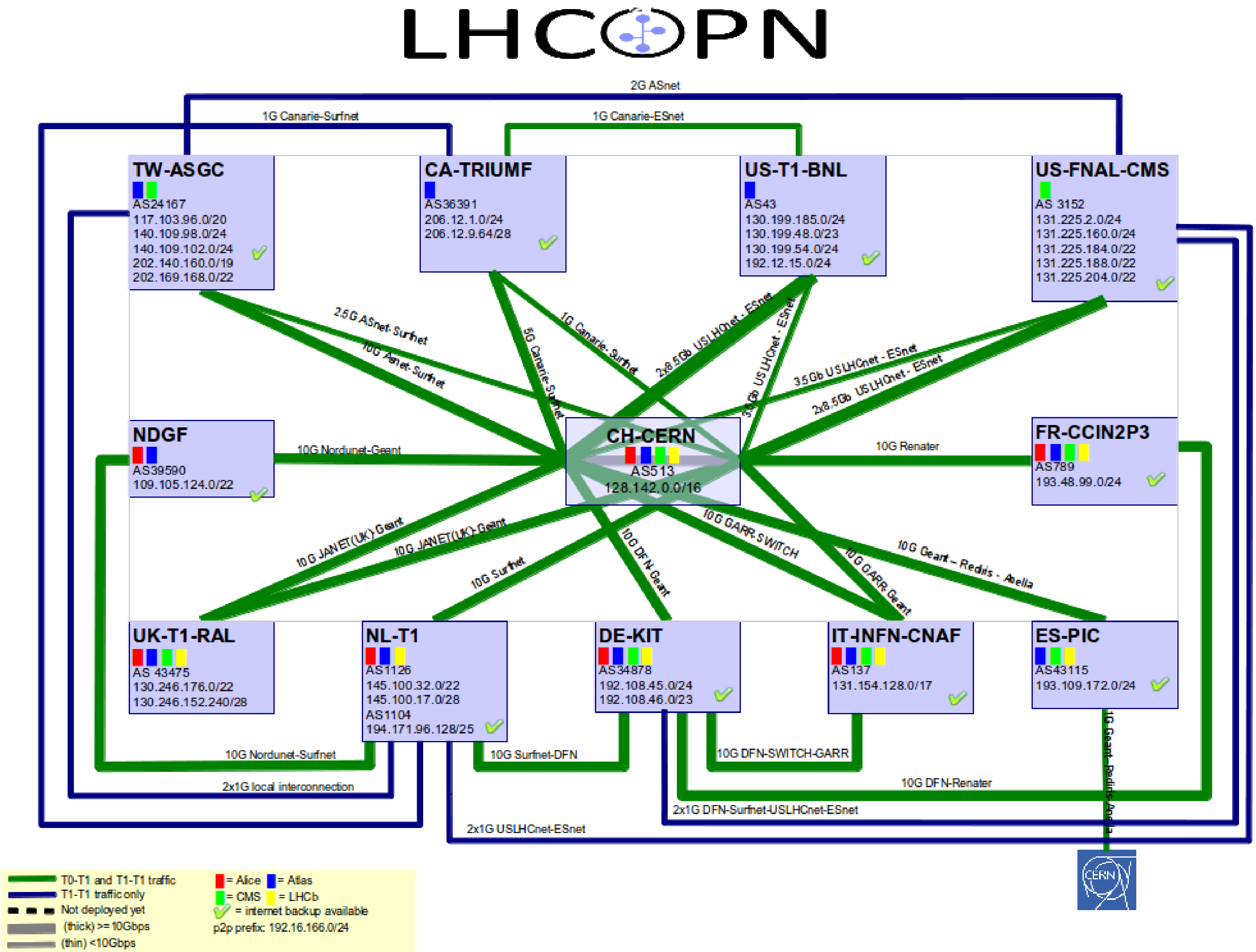
The Birth of LHCOPN

- The LHC Optical Private Network (OPN) is a dedicated network that moves the continuous stream of data from CERN to the Tier 1 national data centers.
  - Only T0→T1 traffic is allowed on the OPN.
    - Tier 2 traffic is not allowed in order to guarantee no interference with the CERN to the T1s data flow.
    - Historically to make the problem “tractable” and provide a “complete” solution.
  - The OPN has a well defined security model that was agreed upon by all of the Tier 1 centers.
    - The OPN has to come into the site without going through the site firewall due to the very high bandwidth, high volume data transfers and this necessitates a strict usage and security model for the OPN.
- The LHC networking model of the OPN has evolved considerably since it was first conceived.

LHCOPN 2012

CERN IT Department

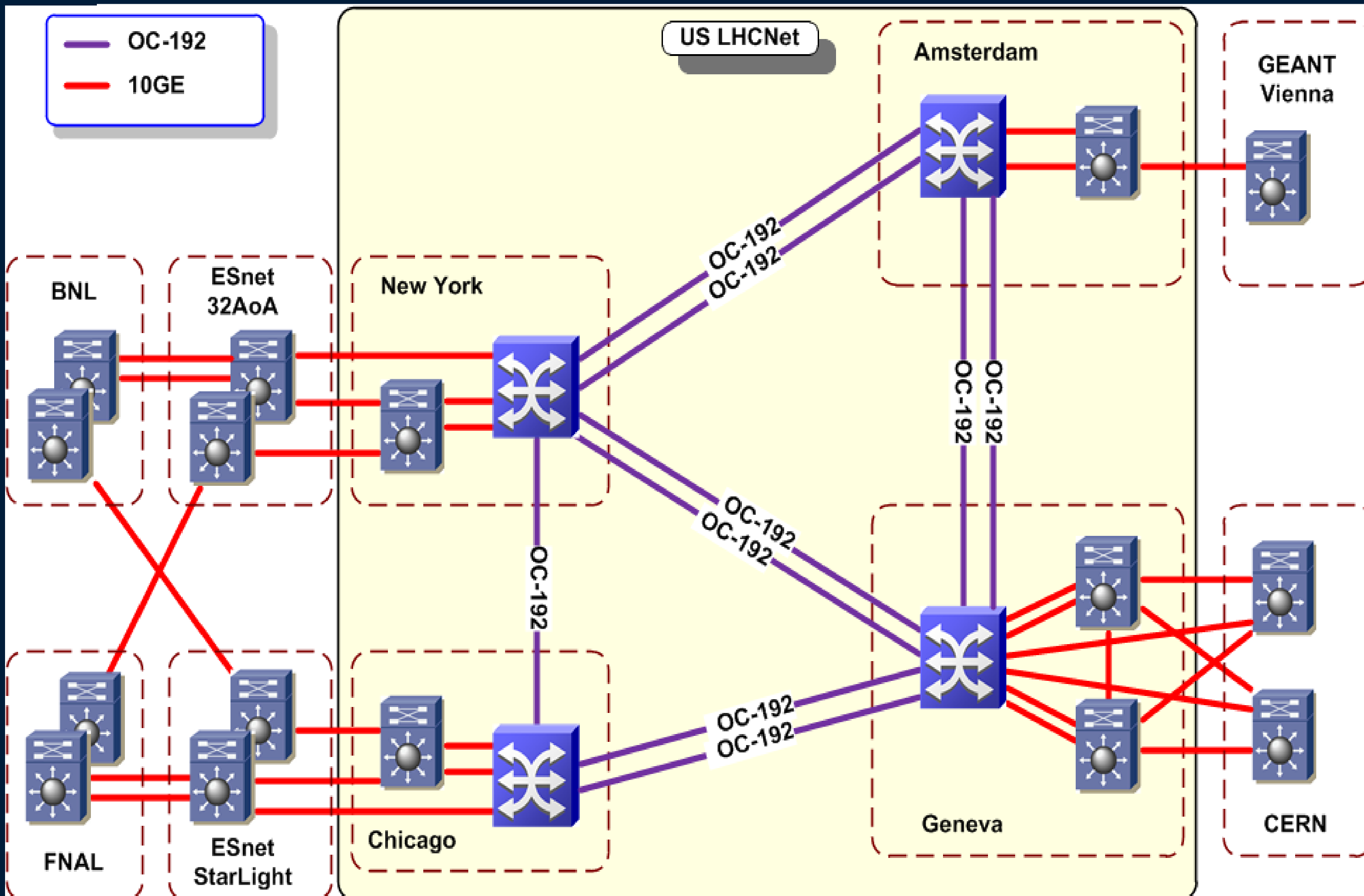




CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

US LHCNet in 2011

Non-stop Operation; Circuit-oriented Services



Core: Optical multiservice Switches [*] that provide resilience

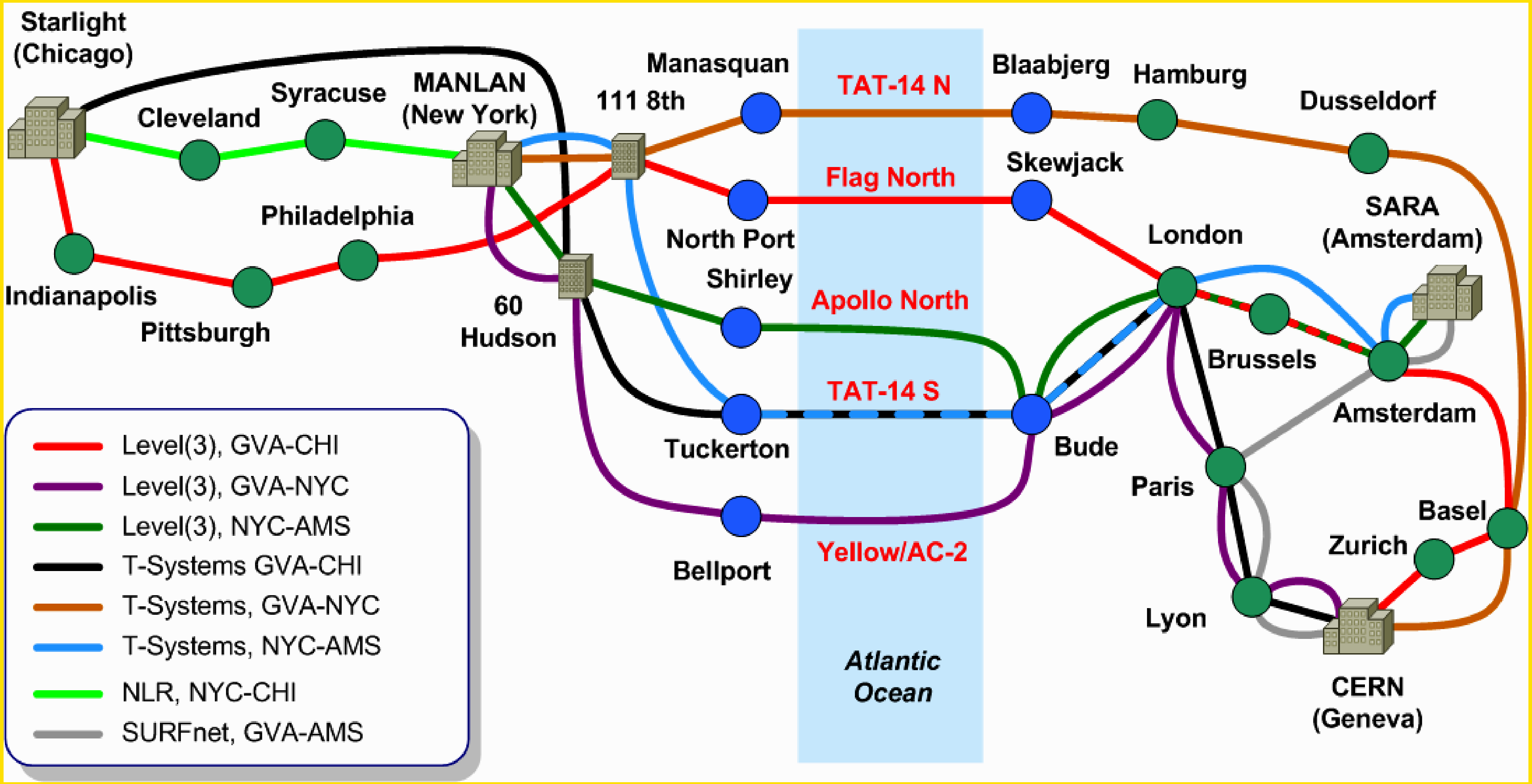
Performance enhancing Standard Extensions: VCAT, LCAS

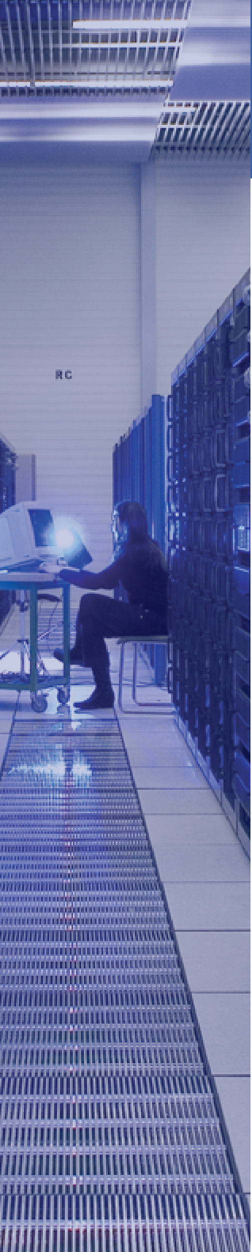
USLHCNet, ESnet, BNL & FNAL: equipment and link redundancy

[*] *Dynamic* circuit-oriented network services with BW guarantees, with robust fallback at layer 1: Hybrid optical network

Path Redundancy for Resilience

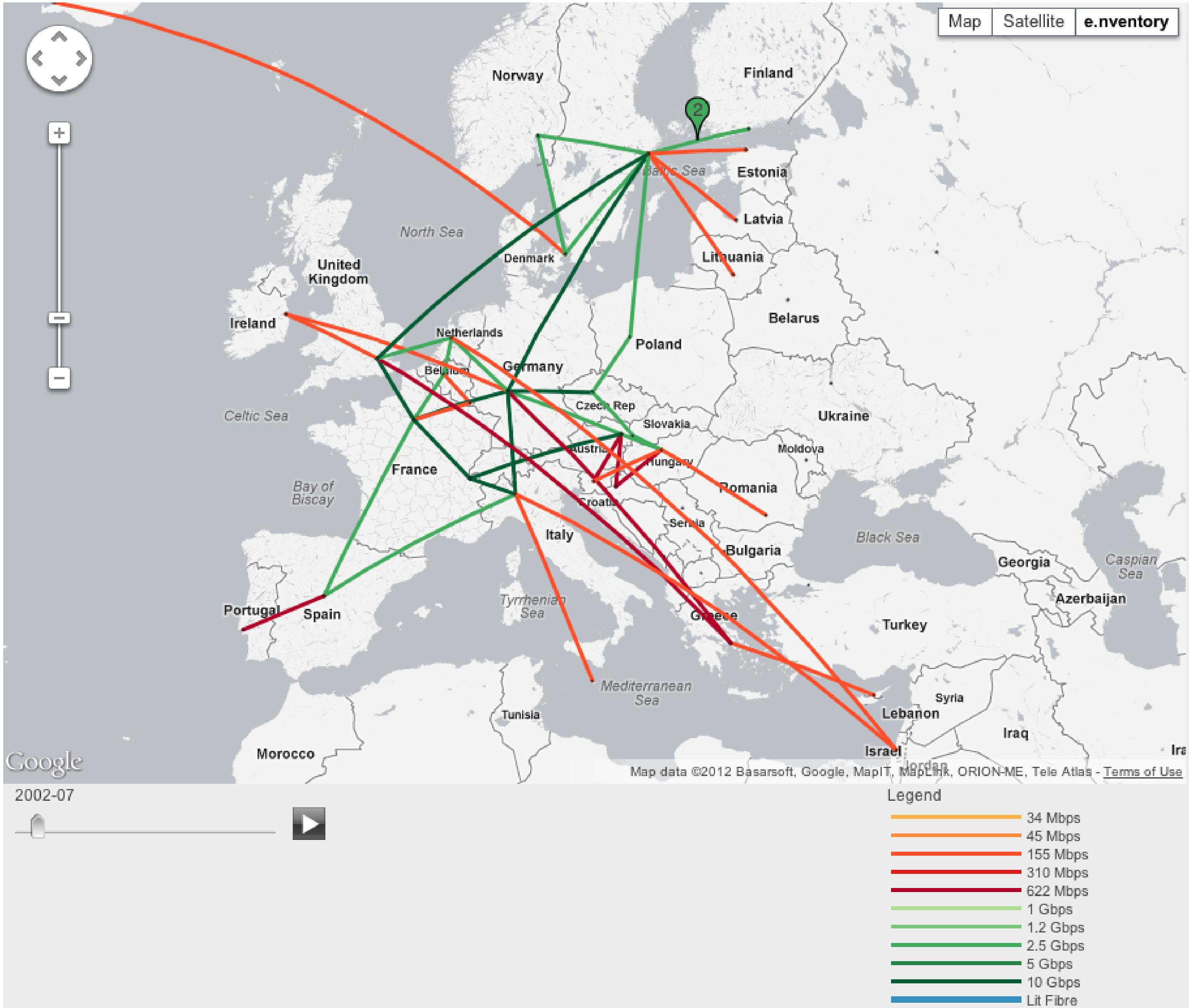
- 6 circuits spread over 5 transatlantic cable systems
- Overlaps on some segments are planned so as to avoid a double hit on the connectivity to a US Tier 1

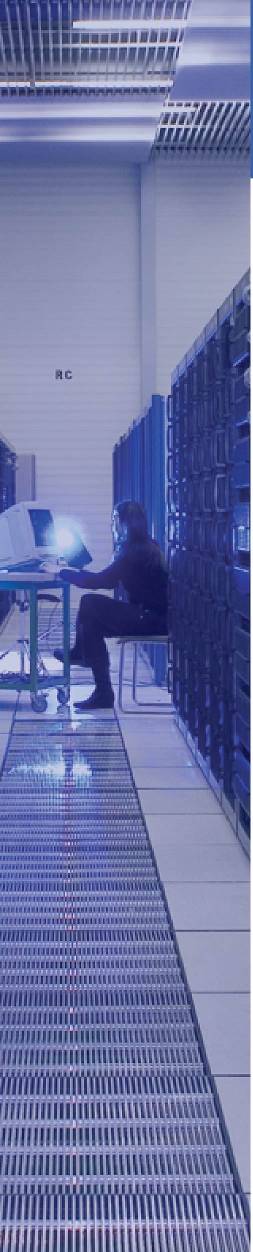




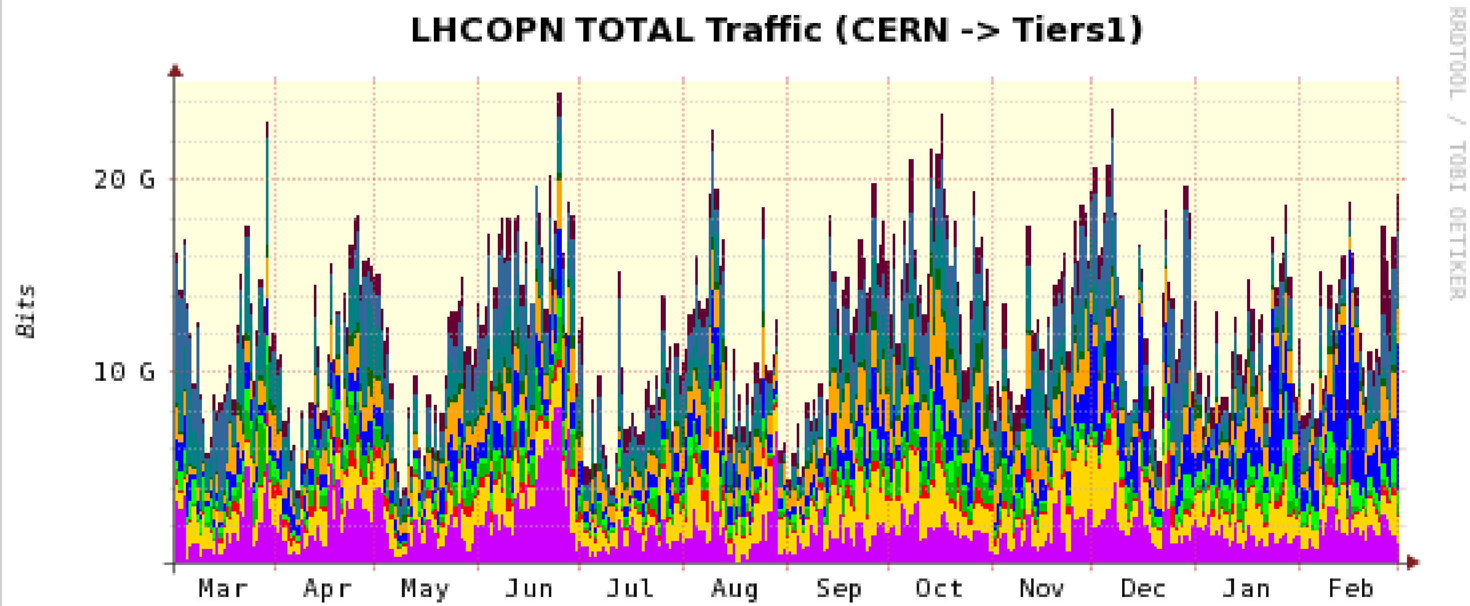
Evolution of Geant
www.enventory.eu

CERN IT Department





LHCOPN in 2011



	Avg	Max	Peak	Curr
To DE-KIT	1.93G	8.17G	9.61G	1.06G
To IN2P3	1.53G	5.63G	9.55G	1.40G
To NDGF	329.41M	1.90G	9.72G	196.71 M
To NLT1	672.78M	3.69G	9.26G	542.63 M
To ASGC	485.91M	3.20G	7.66G	614.07M
To CNAF	1.48G	8.74G	15.41G	7.65G
To RAL	1.37G	4.62G	15.22G	1.78G
To TRIUMF	359.36M	1.75G	3.97G	146.47M
To BNL	1.50G	6.64G	13.68G	850.84M
To FNAL	1.47G	8.92G	16.58G	3.00G
To ES-PIC	1.02G	4.48G	9.72G	1.85G

Total to Tiers1 Avg: 12.13G Max: 24.39G Curr: 19.10G

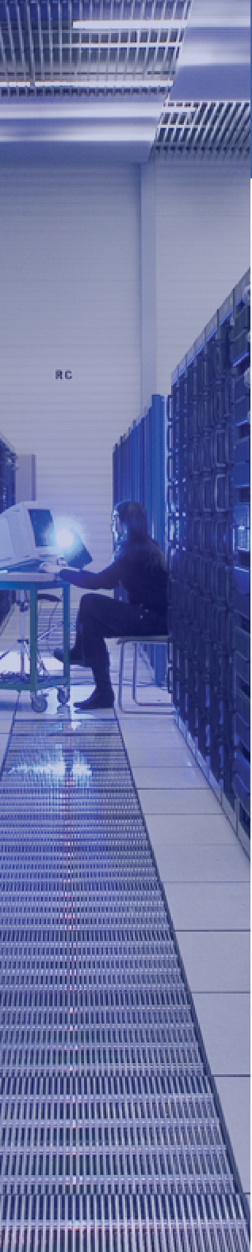
Last update: Thu Mar 01 2012 14:58:17

CERN 2011

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

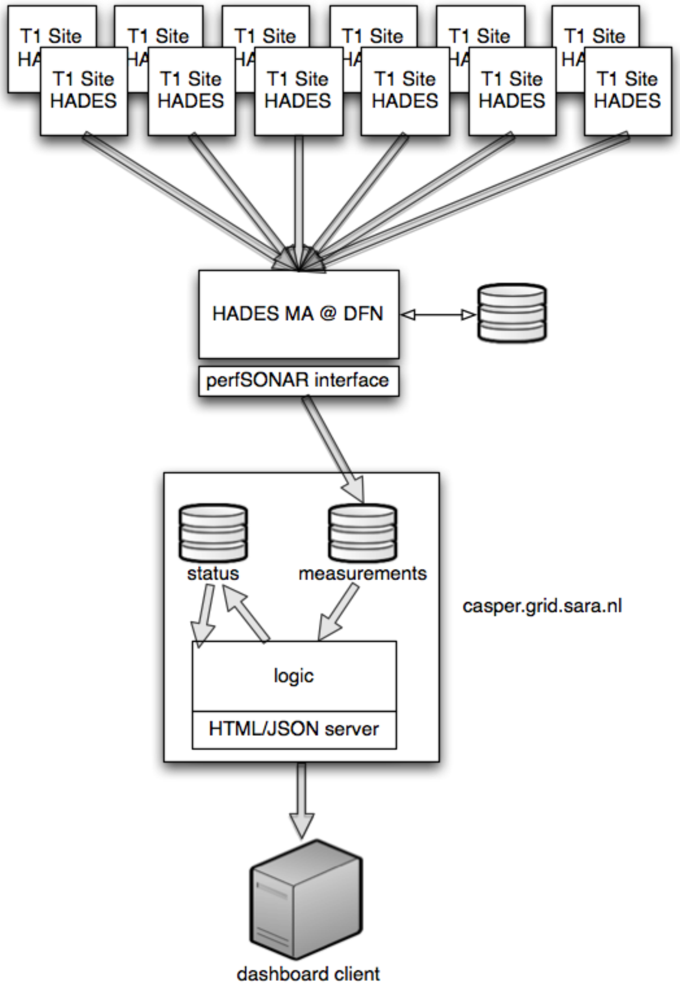
TCP is a “fragile workhorse”

- TCP is a “fragile workhorse.” It will not move very large volumes of data over international distances unless the network is error-free. (Very small packet loss rates result in large decreases in performance.)
  - Case study:
    - On a 10 Gb/s link a 0.0046% loss (1 packet in 22,000) was observed
      - In a LAN or metropolitan area network, this level of loss is barely noticeable
      - In a continental-scale network – 88 ms round trip time path (about that of across the US) – this results in an 80x throughput decrease
  - The only way to keep multi-domain, international scale networks error free is to test and monitor continuously end-to-end.
    - The perfSONAR test-monitor system is deployed extensively throughout the OPN network and at the end sites.
    - PerfSONAR is designed for federated operation
      - Each domain maintains control over what data is published
      - Published data is federated in Measurement Archives that tools can use to produce end-to-end, multi-domain views of network performance



Perfsonar Deployment LHCOPN

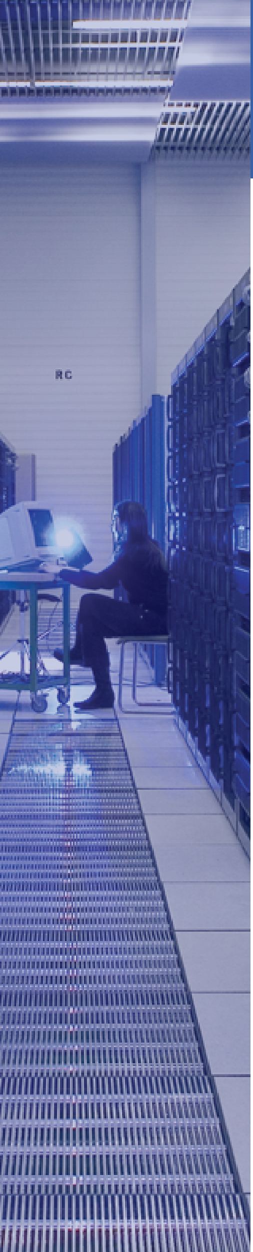
CERN IT Department

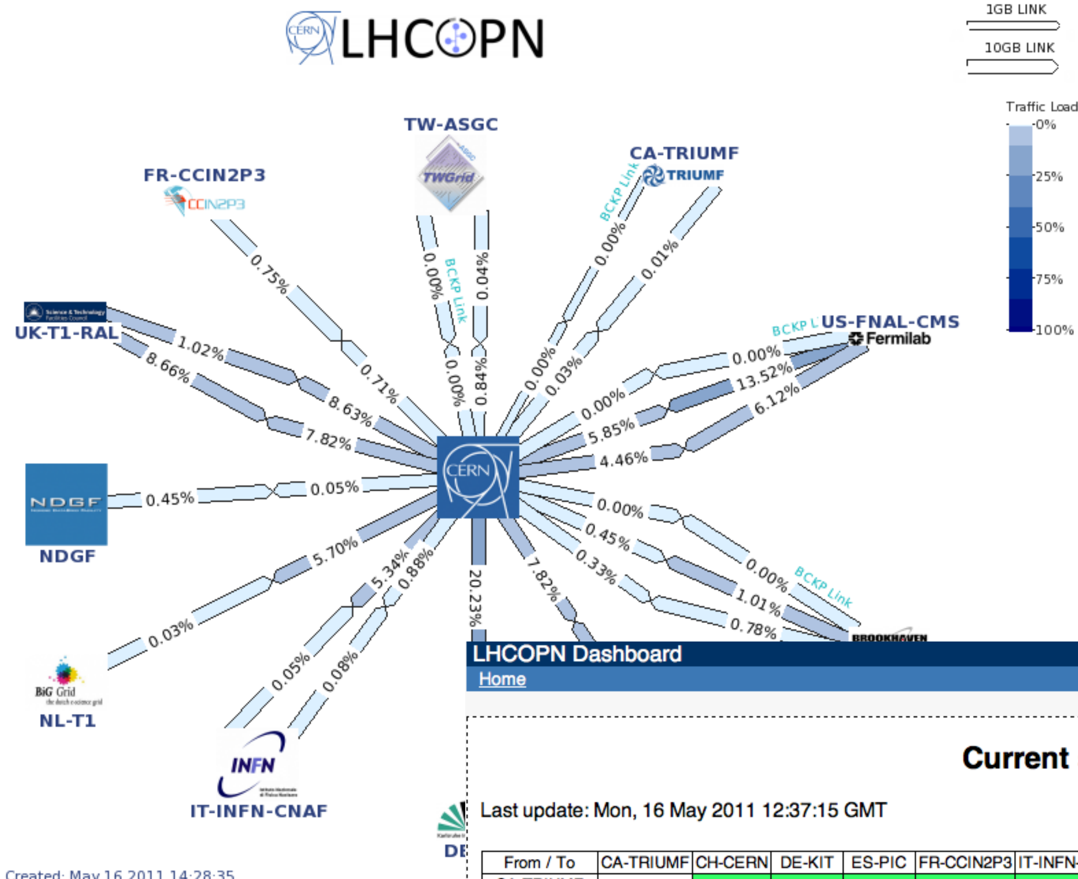


CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

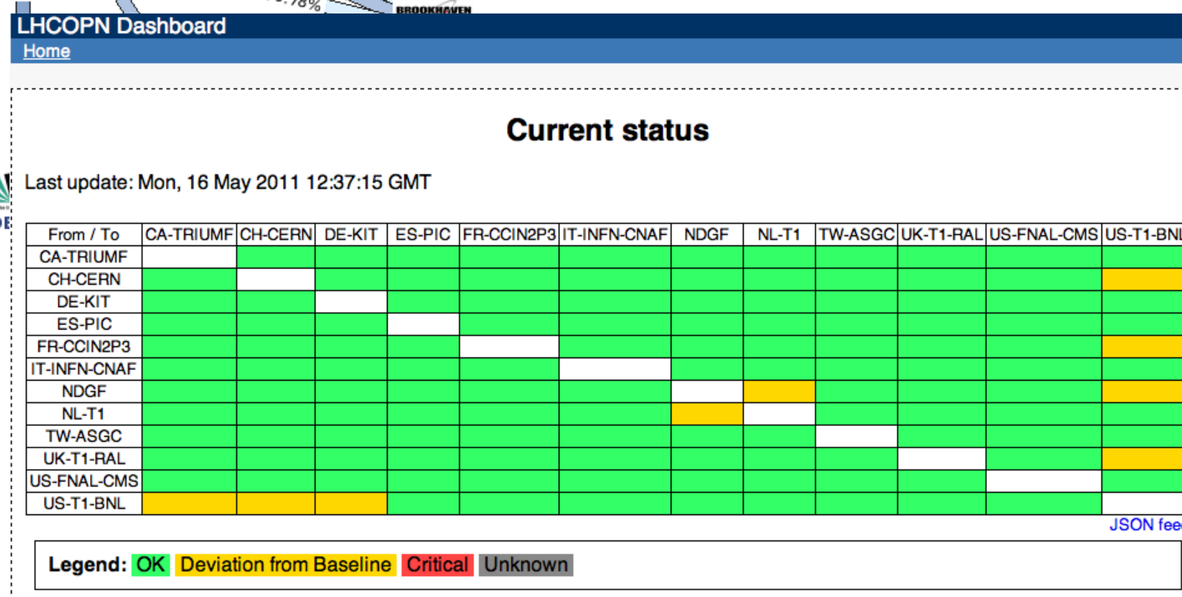
Monitoring is Important

CERN IT Department





Distributed operations is still a complex issue involving multiple providers and sites.



LHCOPN Dashboard

Home

Current status

Last update: Mon, 16 May 2011 12:37:15 GMT

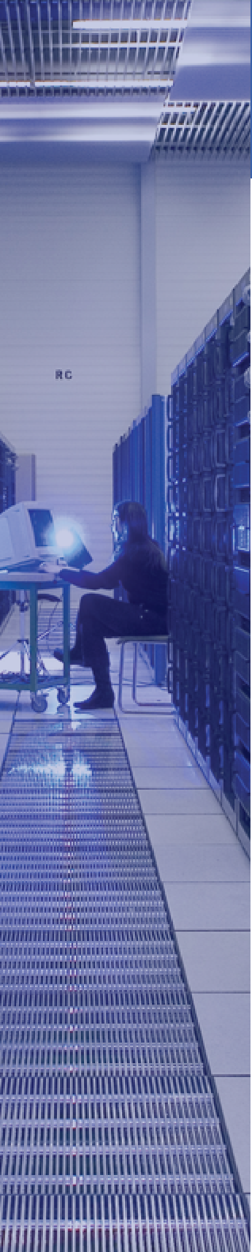
From / To	CA-TRIUMF	CH-CERN	DE-KIT	ES-PIC	FR-CCIN2P3	IT-INFN-CNAF	NDGF	NL-T1	TW-ASGC	UK-T1-RAL	US-FNAL-CMS	US-T1-BNL
CA-TRIUMF												
CH-CERN												
DE-KIT												
ES-PIC												
FR-CCIN2P3												
IT-INFN-CNAF												
NDGF												
NL-T1												
TW-ASGC												
UK-T1-RAL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

JSON feed

Legend: OK Deviation from Baseline Critical Unknown

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

LHCOPN Backup Link Test Status



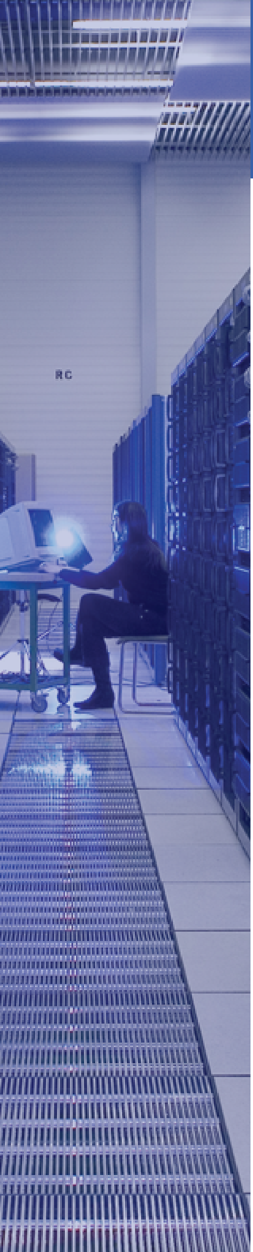
Regular backup tests:

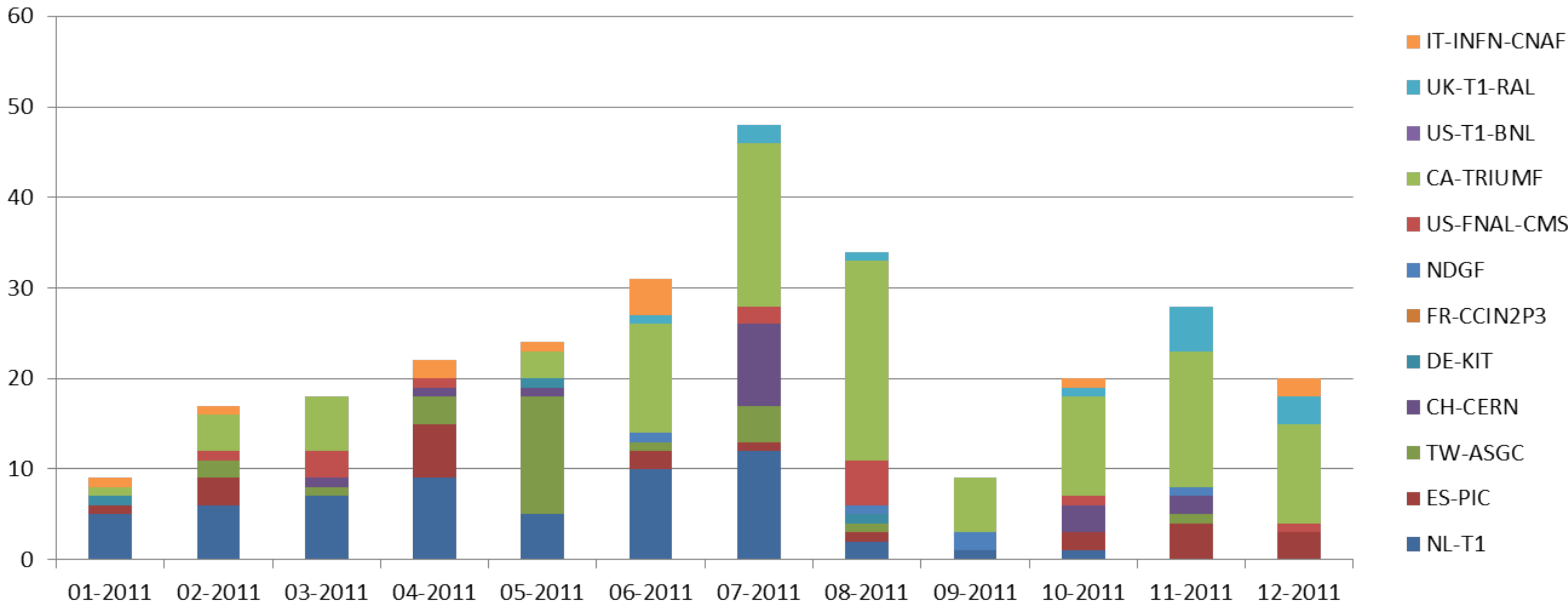
Site	Date of last backup test	Have we a report since 1 year?
CA-TRIUMF	2008-06-03	KO
CH-CERN	2011-10-15	OK
DE-KIT	2011-08-16	OK (partly)
ES-PIC	2011-05-05	OK (partly)
FR-CCIN2P3	2010-03-08	KO
IT-INFN-CNAF	2008-04-09	KO
NDGF	2008-04-09	KO
NL-T1	2009-02-10	KO
TW-ASGC	2010-12-28	KO
UK-T1-RAL	2010-08-24	KO
US-FNAL-CMS	2008-04-24	KO
US-T1-BNL	2008-06-03	KO

More systematic tests planned

Ticket distribution over Tier-1

CERN IT Department





- 2011 total-# of tickets: 318
- Approx. average of 26,5 per month
- Last quarter : total TTS : 68 → approx 22,5 per month

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

LHCONE

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

Change of the data model…

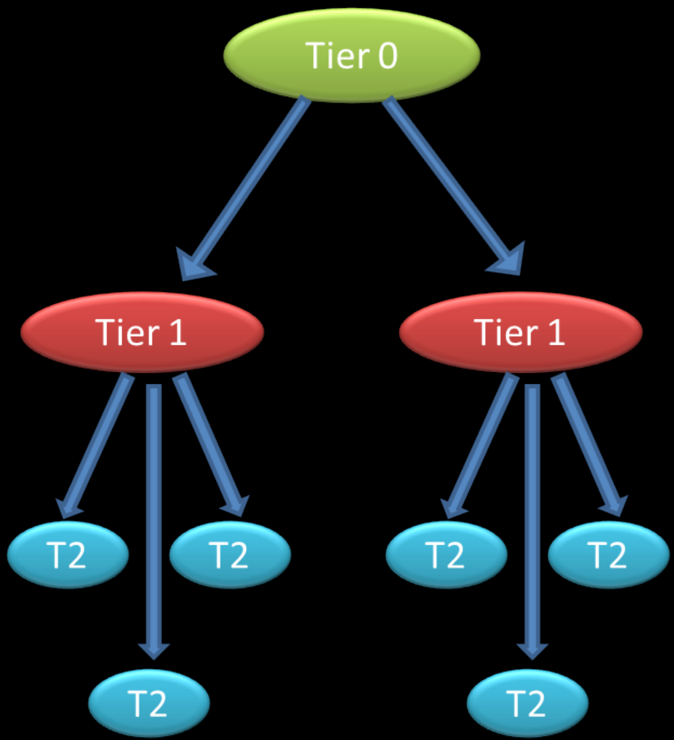
- Data placement will now be based on
  - Organised placement as before where the need is clear
  - Dynamic placement when jobs are sent to a site
  - Data popularity – popular data is replicated – unused data is removed
- Analysis disk becomes a more dynamic cache
- Also start to use remote (WAN) I/O:
  - Fetch a file missing from a dataset
  - Read a file remotely over the network
  - Often means *less* network traffic

WLCG Worldwide LHC Computing Grid

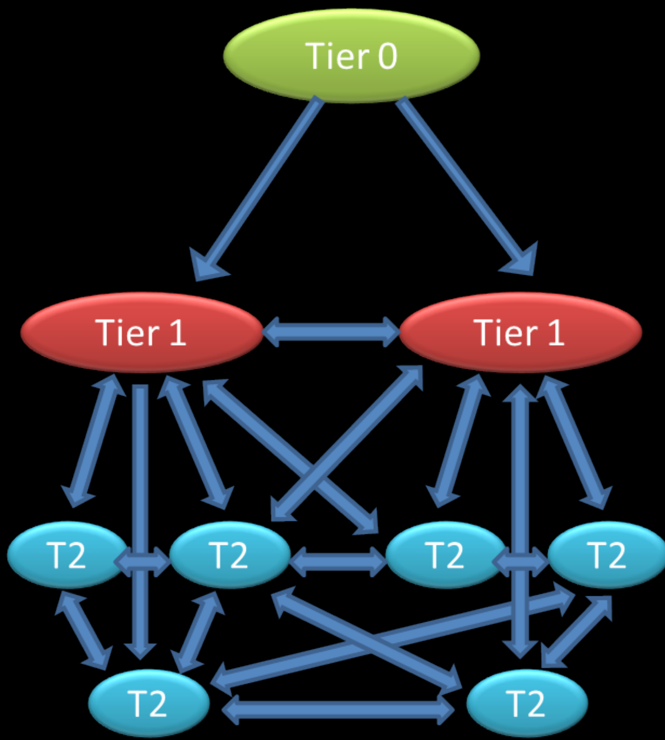
LCG

Computing model evolution

Evolution of computing models

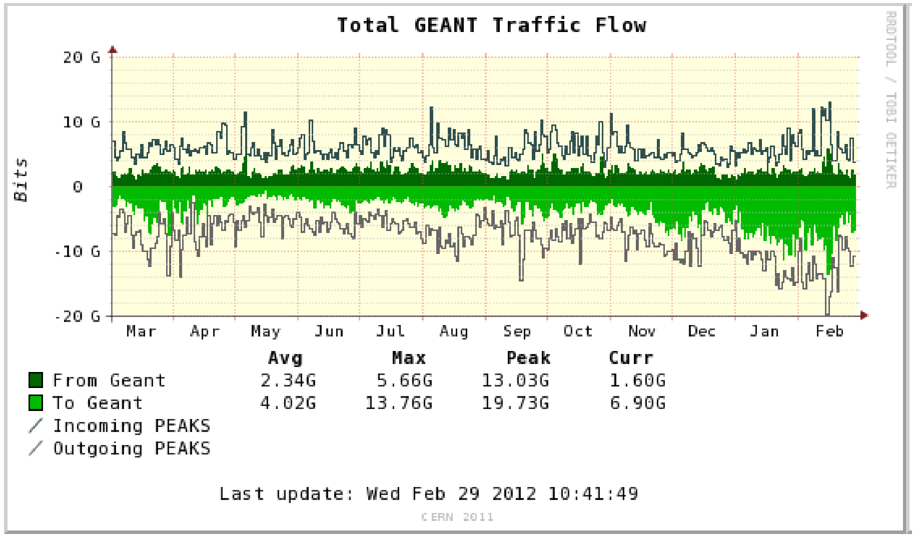


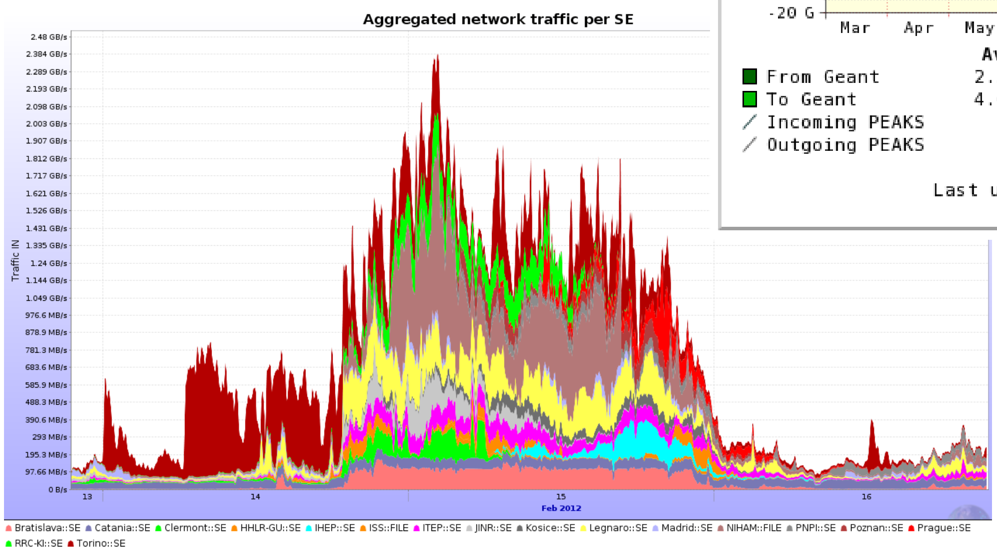
Hierarchy



Mesh

ALICE Traffic – Special Case





The cause of IP saturation turned out to have been the tail of the ALICE Pb data production with unusual circumstances: the number of concurrent jobs was a factor 2 higher than usual, each dealing with half the normal amount of data, which caused the outgoing traffic to be higher by the same factor and the run to finish twice as fast.

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

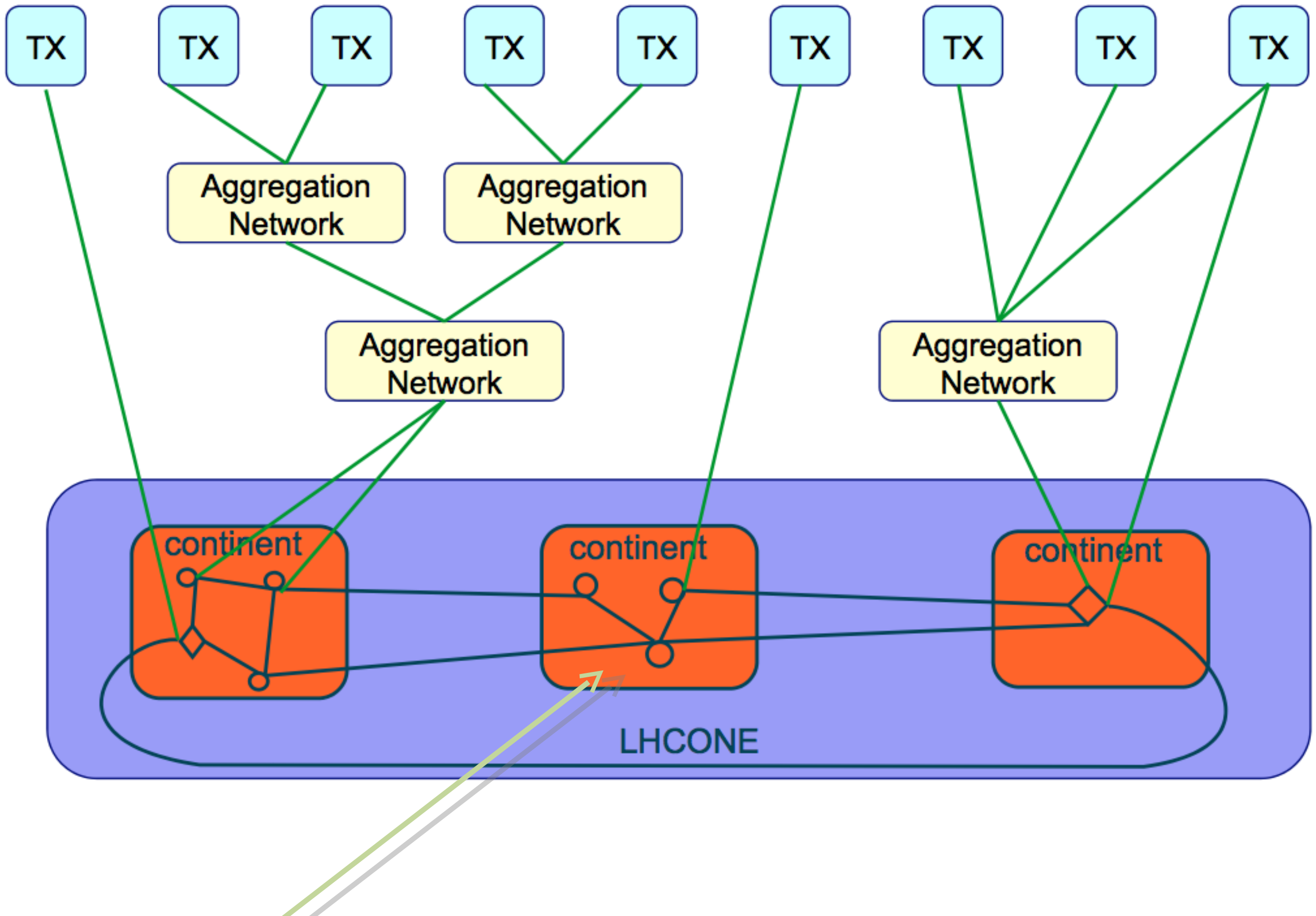
The Birth of LHCONE

CERN IT Department

- Initially the Tier 2 centers were expected to get data only from their associated national Tier 1 center and networks were planned with this in mind
  - The initial projection of 1 Gb/s T1→T2 is now closer to 10 Gb/s for the large T2s.
  - This hierarchical data flow model broke down fairly quickly when real data analysis started.
    - At that time the T2s started going to any T1 – and even other T2s – that had data that was useful to them.
  - This resulted in the T2s using the general R&E network infrastructure even more extensively and in a more "chaotic" manner.
    - It was not engineered to support frequent, very large data transfers, especially across the Atlantic.
  - To address this issue, the LHCONE – largely an overlay on the existing national networks (a science "VPN") – was designed and built so that the LHC traffic could be isolated, managed, and provided for without interfering with other traffic.
    - http://lhcone.net

LHCONE Architecture 2011

CERN IT Department

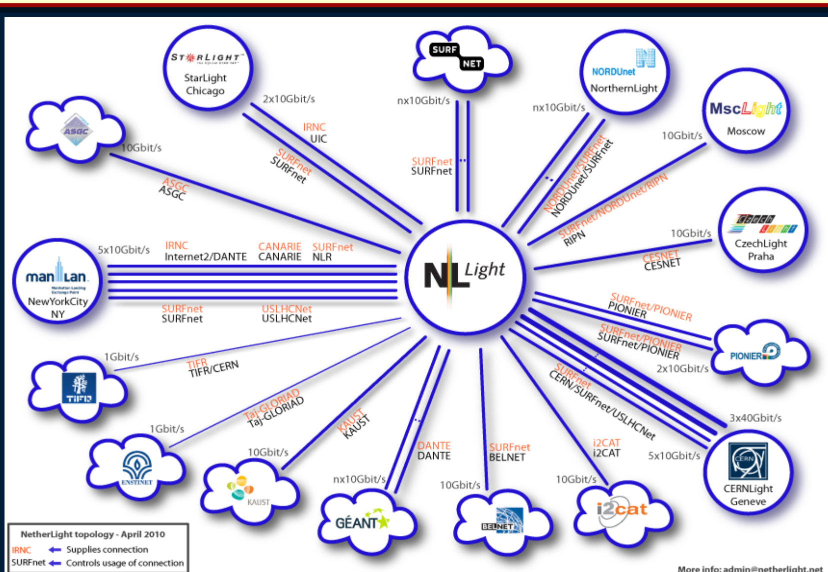


Open Exchanges

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

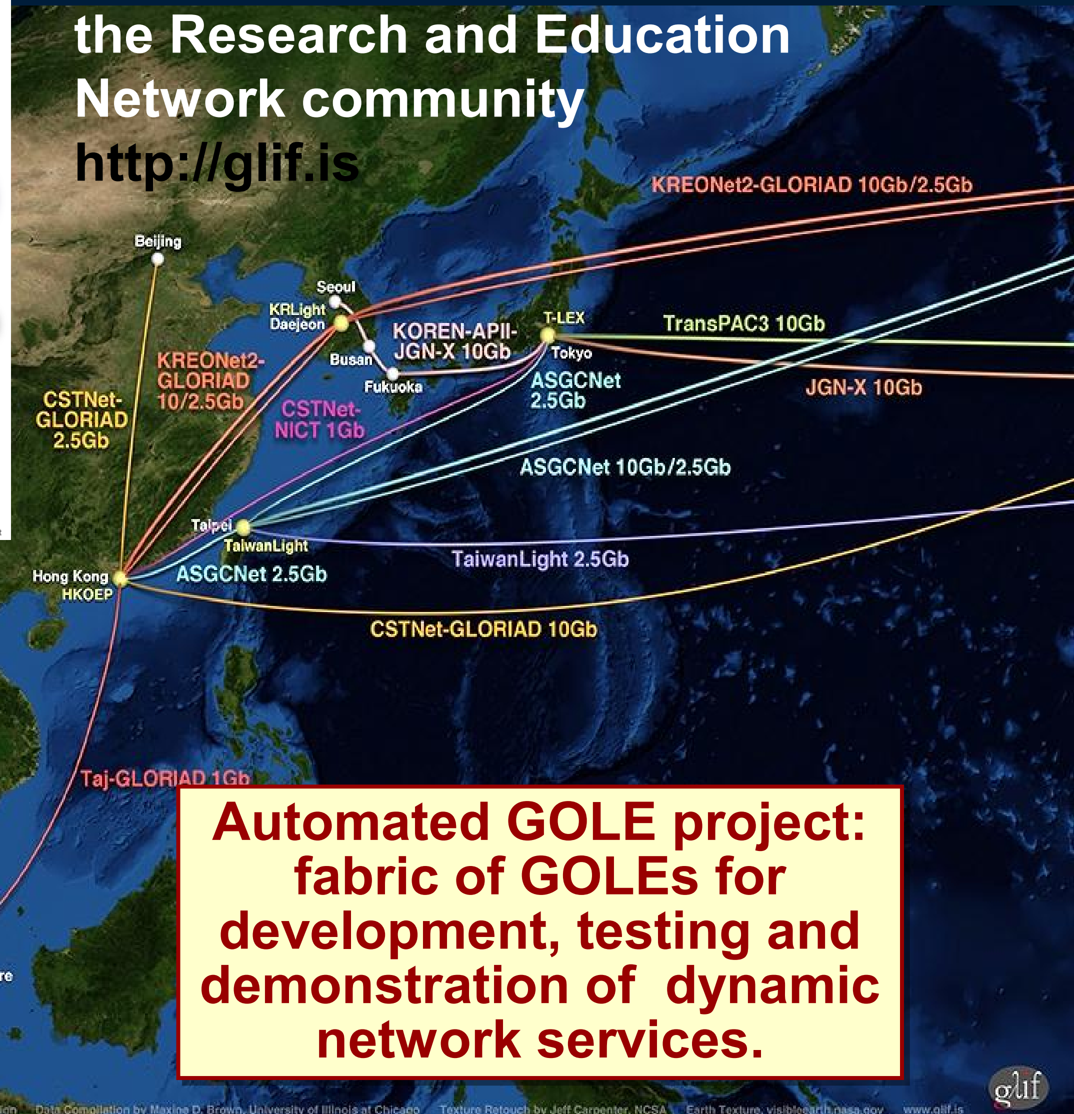
GLIF Open Lightpath Exchanges

GOLE Example: Netherlight in Amsterdam



Exchange Points operated by the Research and Education Network community

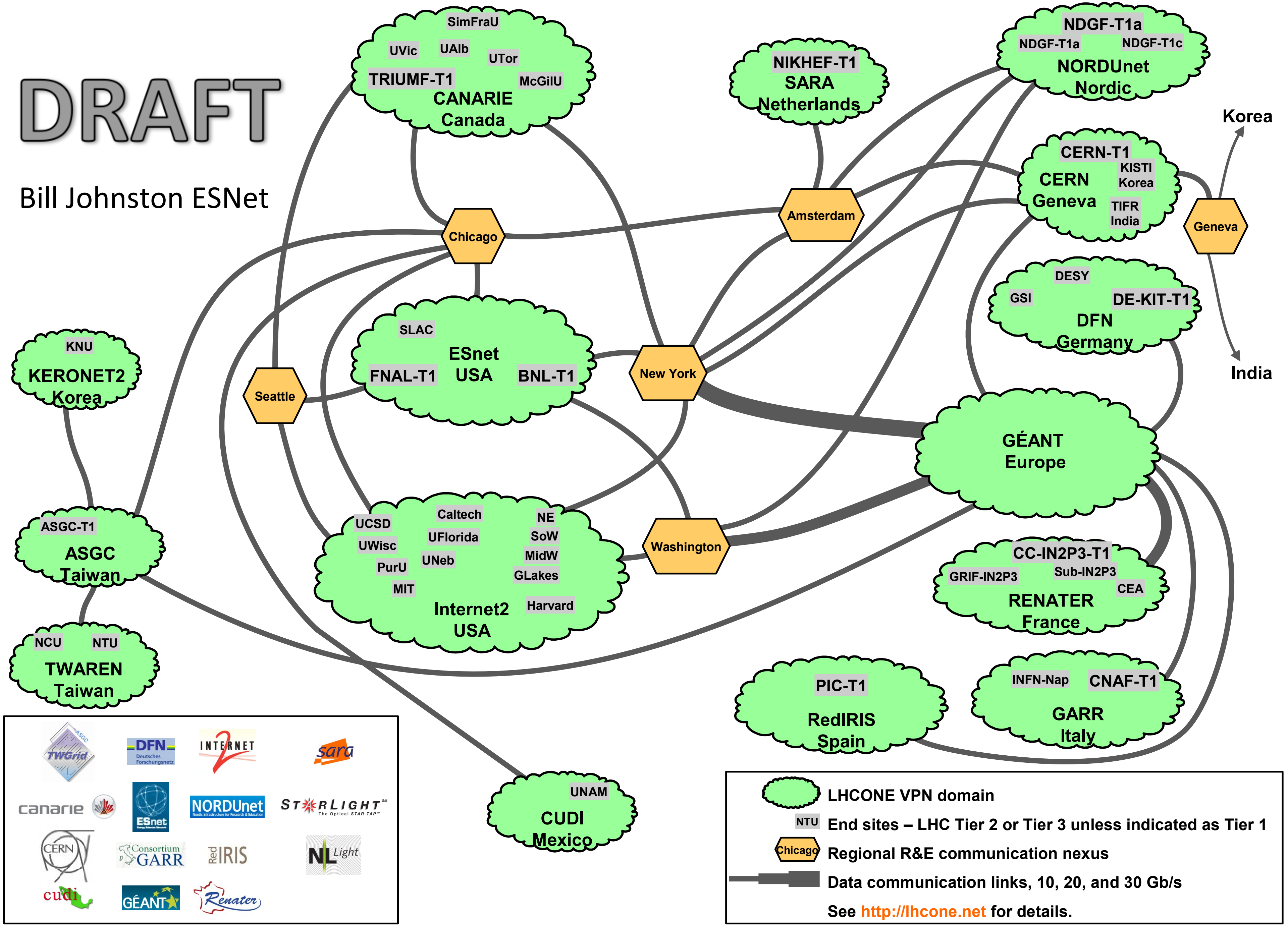
http://glif.is

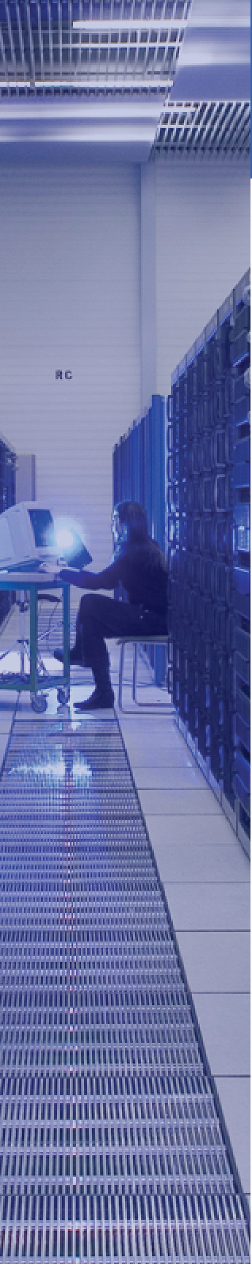


Automated GOLE project: fabric of GOLEs for development, testing and demonstration of dynamic network services.

http://glif.is

LHCONE: A global infrastructure for the LHC Tier1 data center – Tier 2 analysis center connectivity





Connecting TIFR

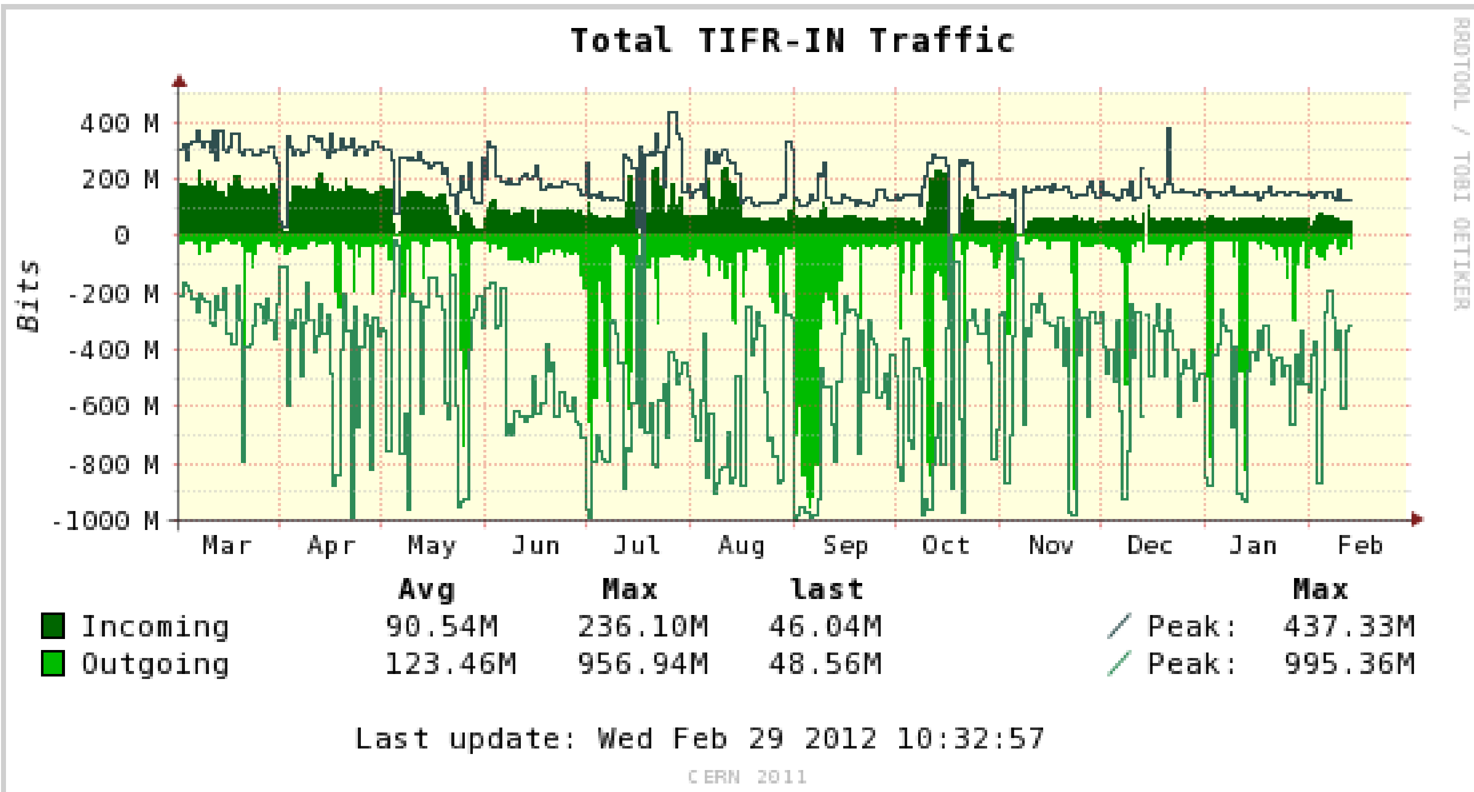
TIFR Link History

- Contract X106/IT signed 26 July 2007
- 2007-2008: 300M for 18hrs 1G for 6hrs
- 2008-2009: 400M for 18hrs 1G for 6hrs
- 2009-2010: 600M for 18hrs 1G for 6hrs
- 2010-2011: 1G for 24hrs
- 2011-2012: 1G for 18Hrs, 1.5G for 6hrs
  - 1.5G between 10 am and 4 pm Indian Standard Time.
  - 10G interfaces installed to provide for future upgrades.

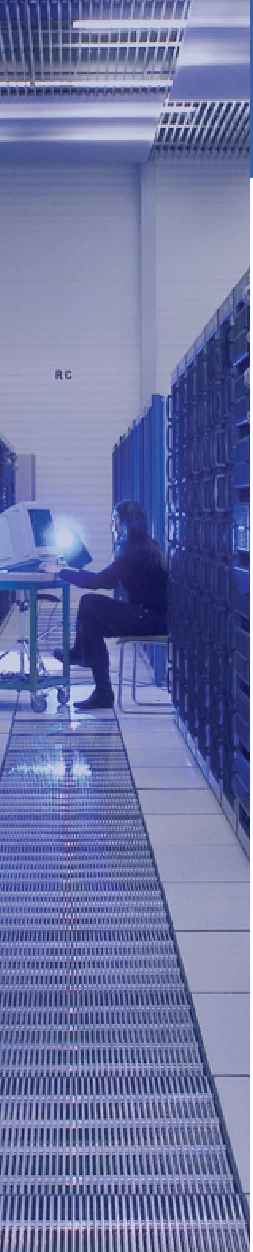
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

TIFR – CERN Traffic 2011

CERN IT Department

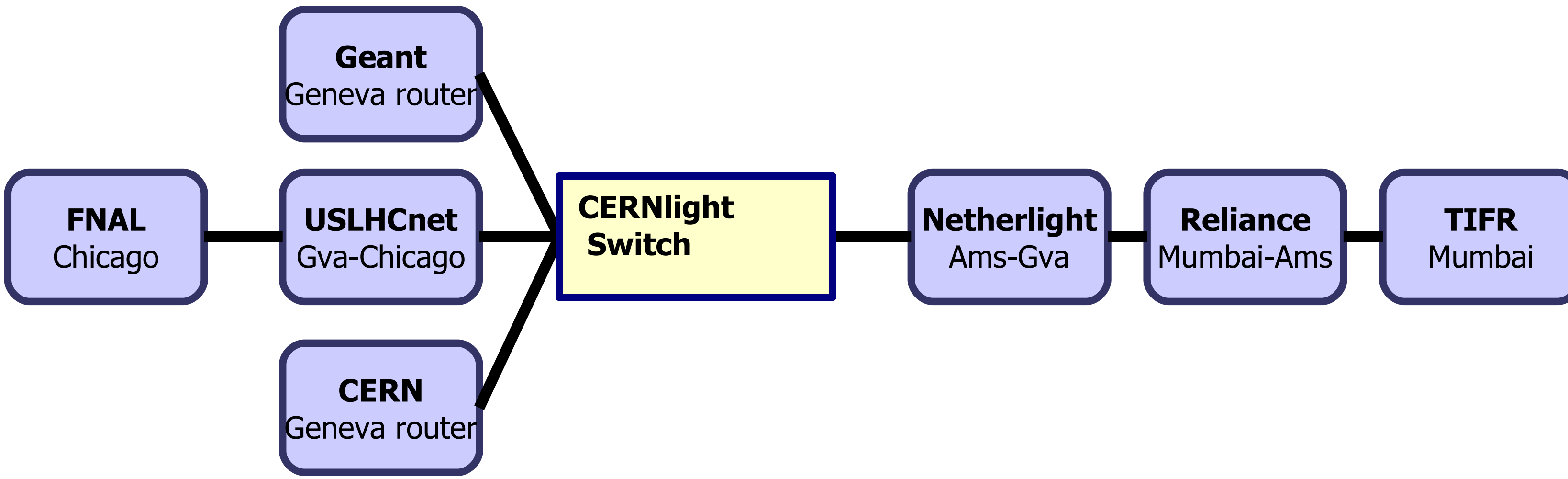


CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it



TIFR Connection

CERN IT Department



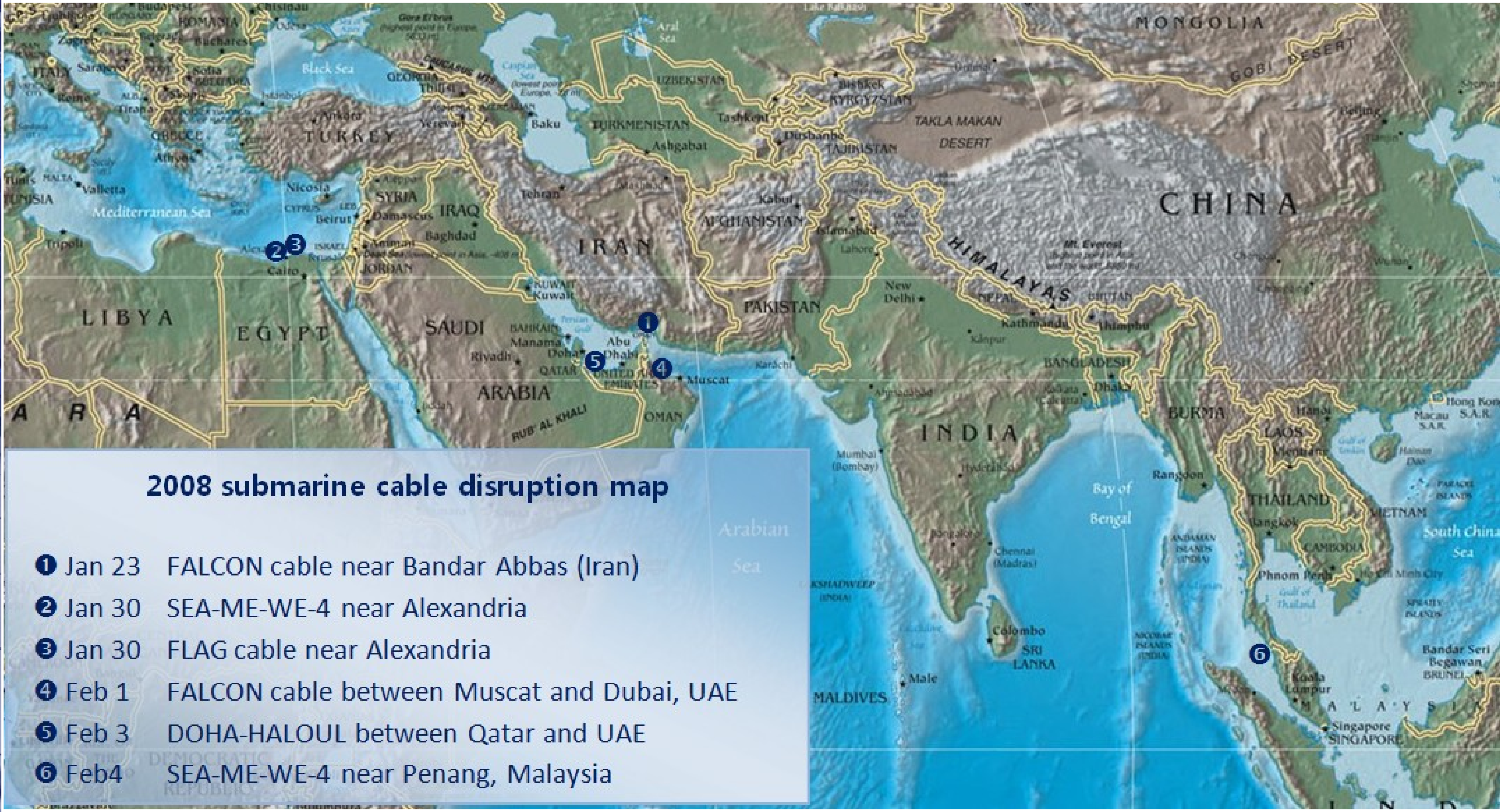
Vlan's between TIFR and FNAL, TIFR and CERN, TIFR and GEANT

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

Submarine Cable Cuts

CERN IT Department

The value of having a protected circuit

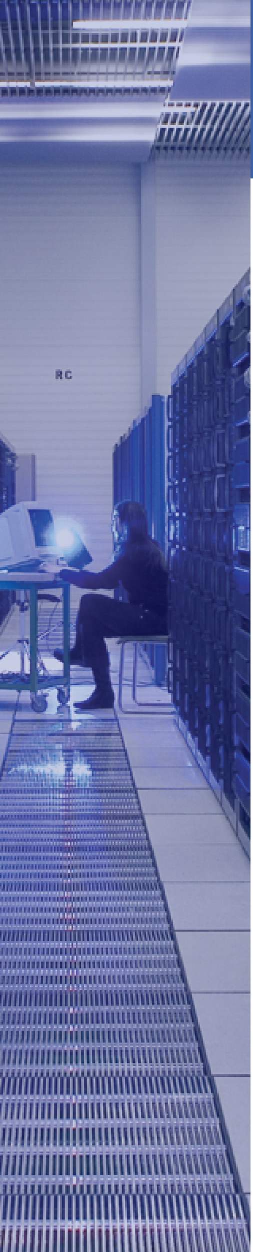


CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Another Challenge

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

New CERN Computer Center



Networking Requirements

CERN IT Department

- Should be an extension of the CERN center
  - Service will be increasingly virtualised.
  - Operations will be remote
  - Essentially a “private cloud”
- 100G circuits by 2 diverse paths
- Could explore wide-area Tb networking in the coming (few) years.
  - Would provide “seamless” integration.
  - 10x100G initially.

HEP Bandwidth Roadmap for Major Links (in Gbps): US LHCNet Example

CERN IT Department

Harvey Newman Caltech

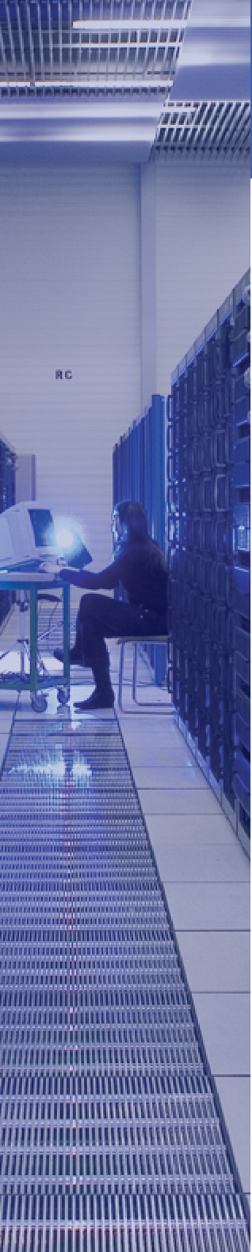
Year	Production	Experimental	Remarks
2001	0.155	0.622-2.5	SONET/SDH
2002	0.622	2.5	SONET/SDH DWDM; GigE Integ.
2003	2.5	10-20	DWDM; 1 + 10 GigE Integration
2005-6	10-20	2-10 X 10	λ Switch; λ Provisioning
2007-8	3-4 X 10	~10 X 10; 100 Gbps	1st Gen. λ Grids
2009-10	6-8 X 10	~20 X 10 or ~2 X 100	100 Gbps λ Switching
2011-12	~20 X 10 or 2 X 100	~10 X 100	2nd Gen λ Grids Terabit Networks
2013-5	~Terabit	~MultiTbps	~Fill One Fiber

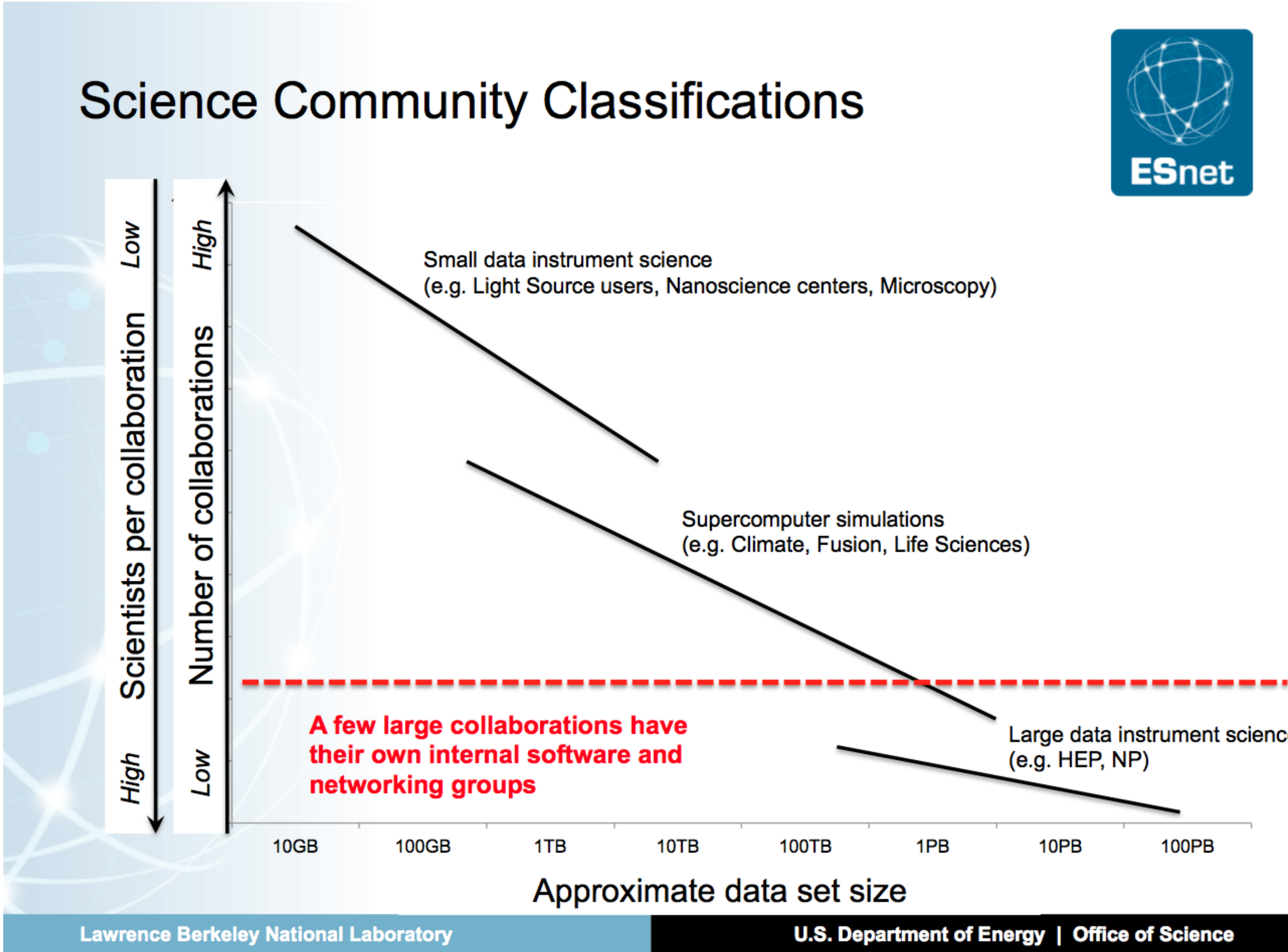
Paralleled by ESnet Roadmap for Data Intensive Sciences

Trends and Lessons

Collaborations

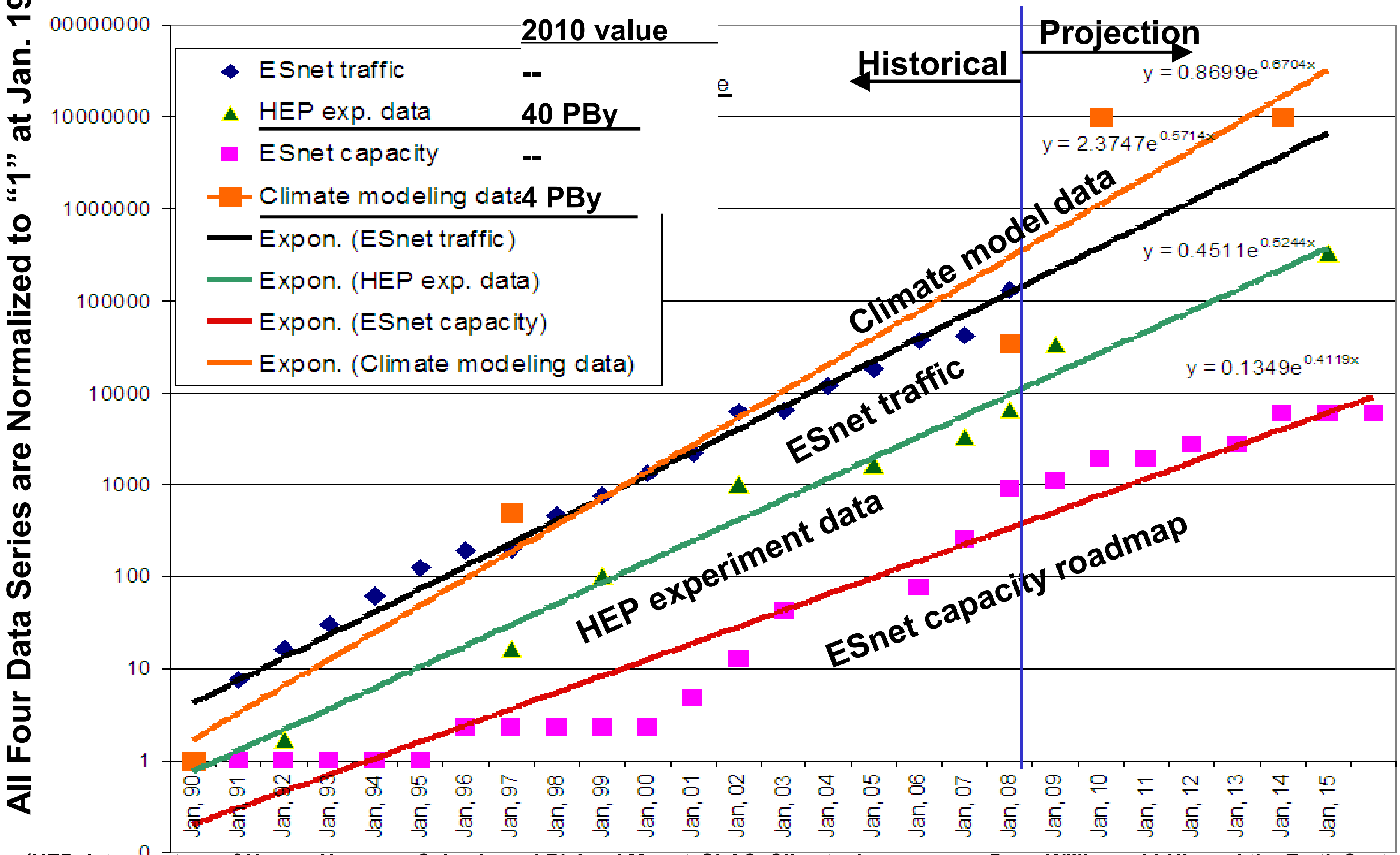
CERN IT Department





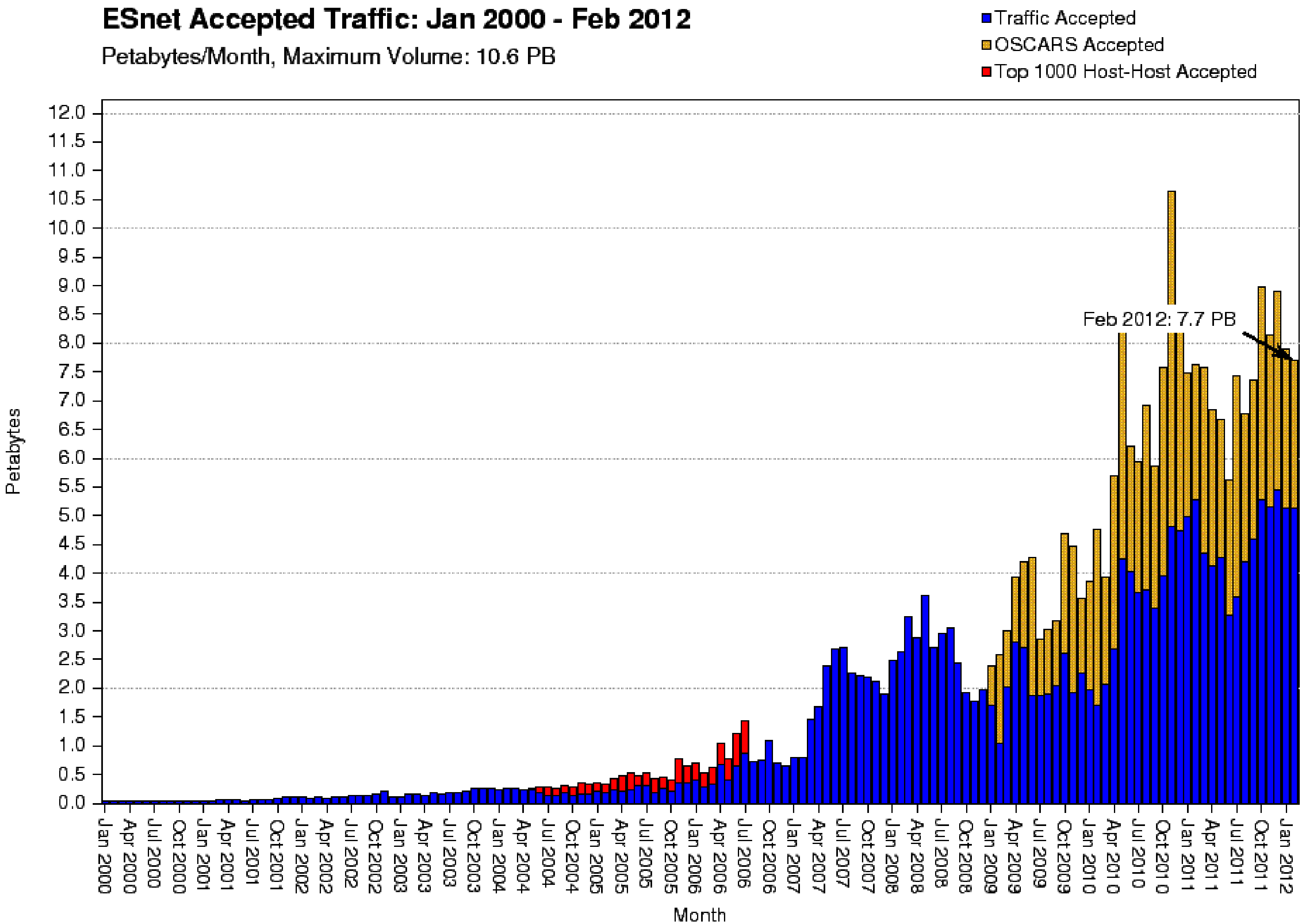
Network Traffic, Science Data, and Network Capacity

Ignore the units of the quantities being graphed they are normalized to 1 in 1990, just look at the long-term trends: ***All of the "ground truth" measures are growing significantly faster than ESnet projected capacity based on stated requirements***

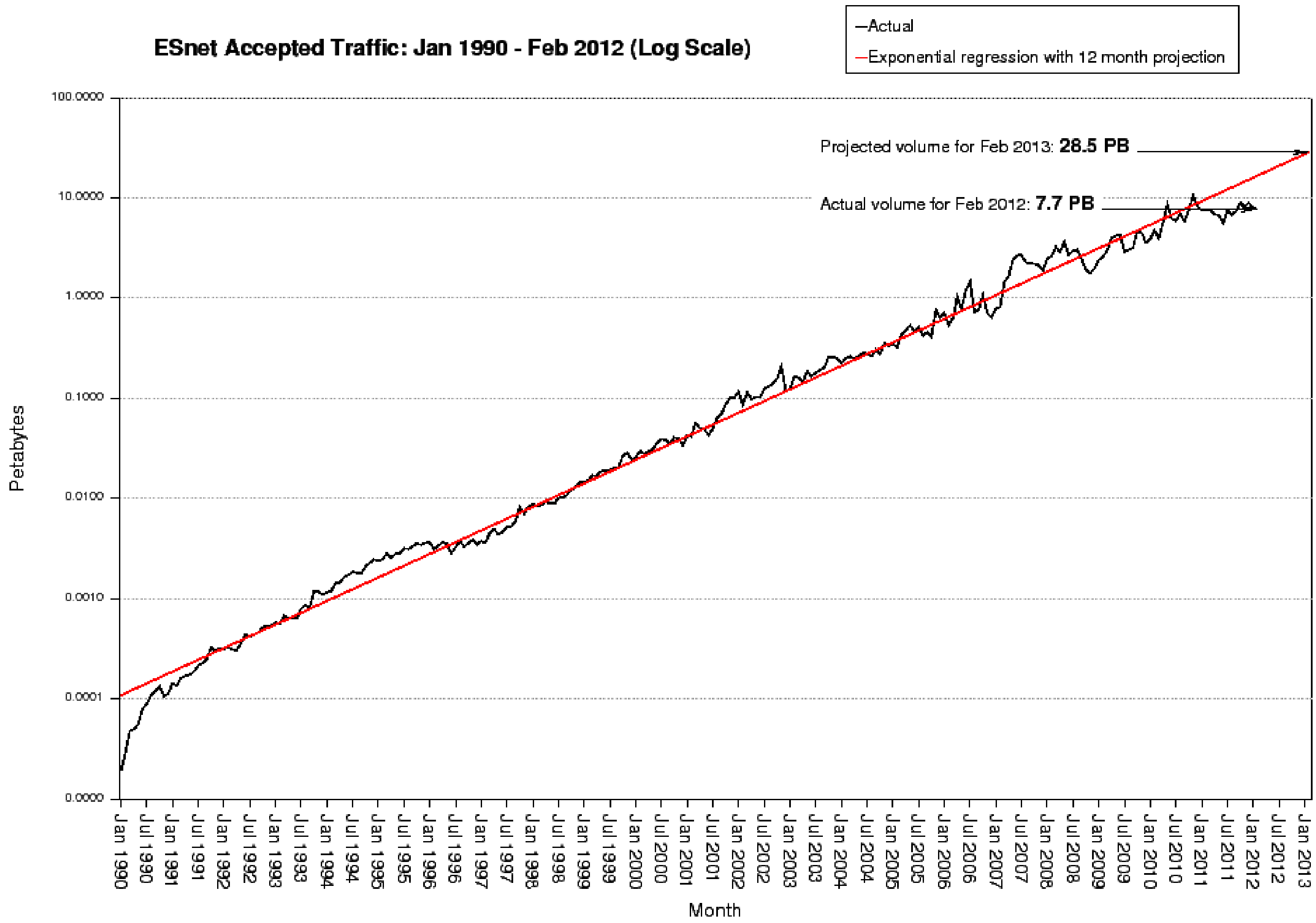


(HEP data courtesy of Harvey Newman, Caltech, and Richard Mount, SLAC. Climate data courtesy Dean Williams, LLNL, and the Earth Systems Grid Development Team.)

ESNet Traffic



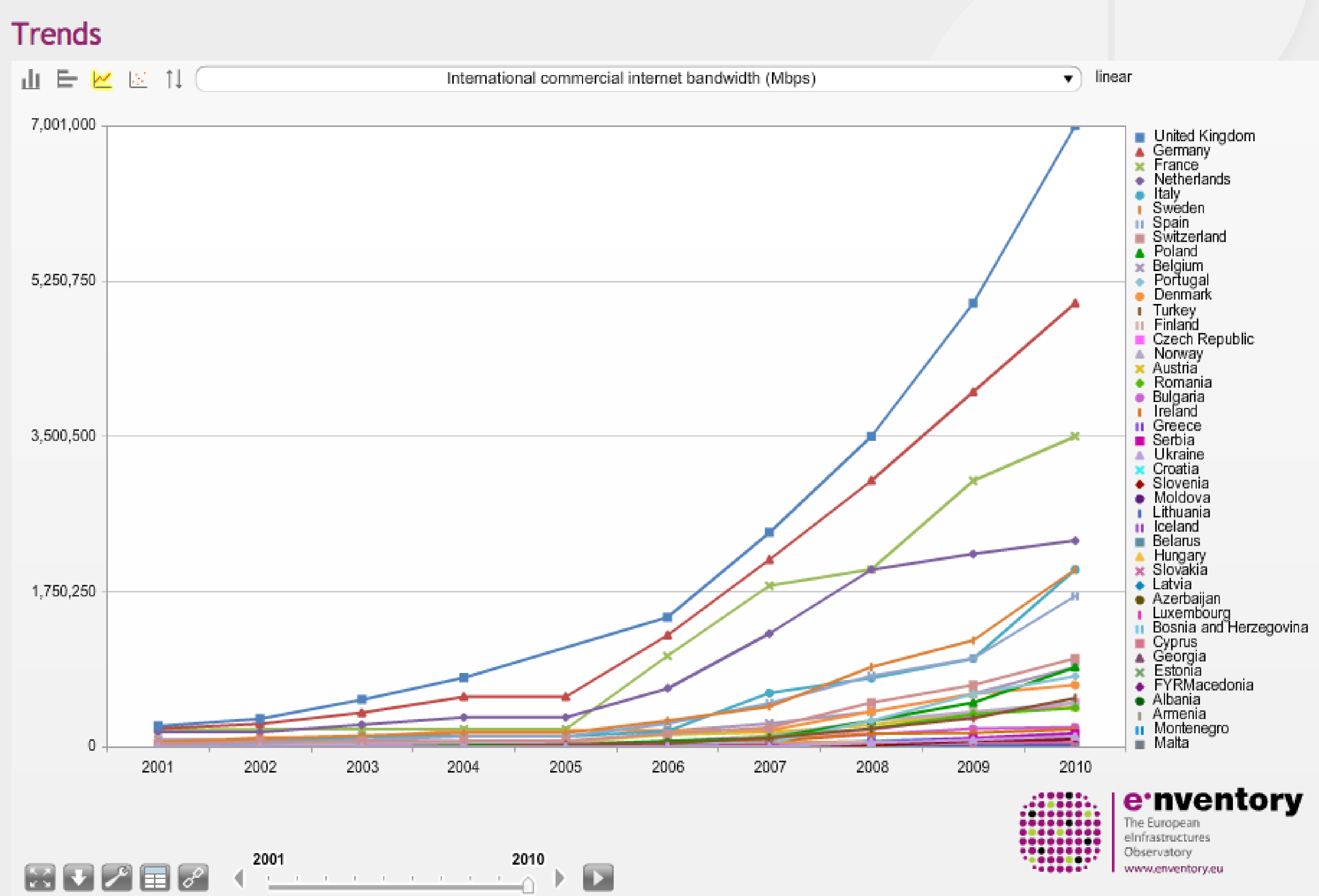
ESNet Traffic Evolution



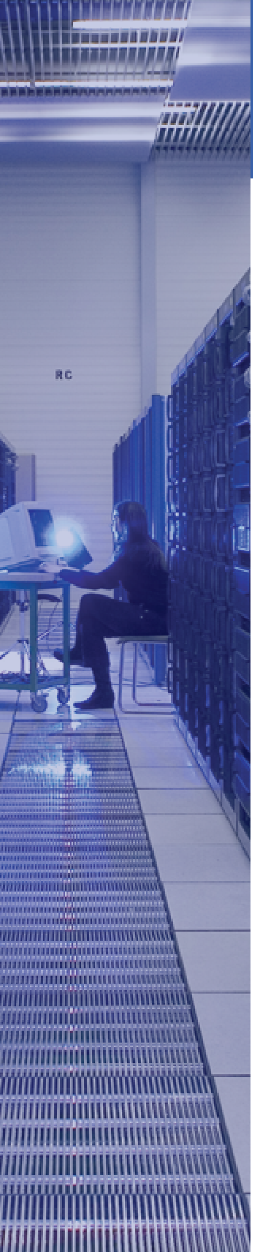
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Commercial Internet Bandwidth

CERN IT Department

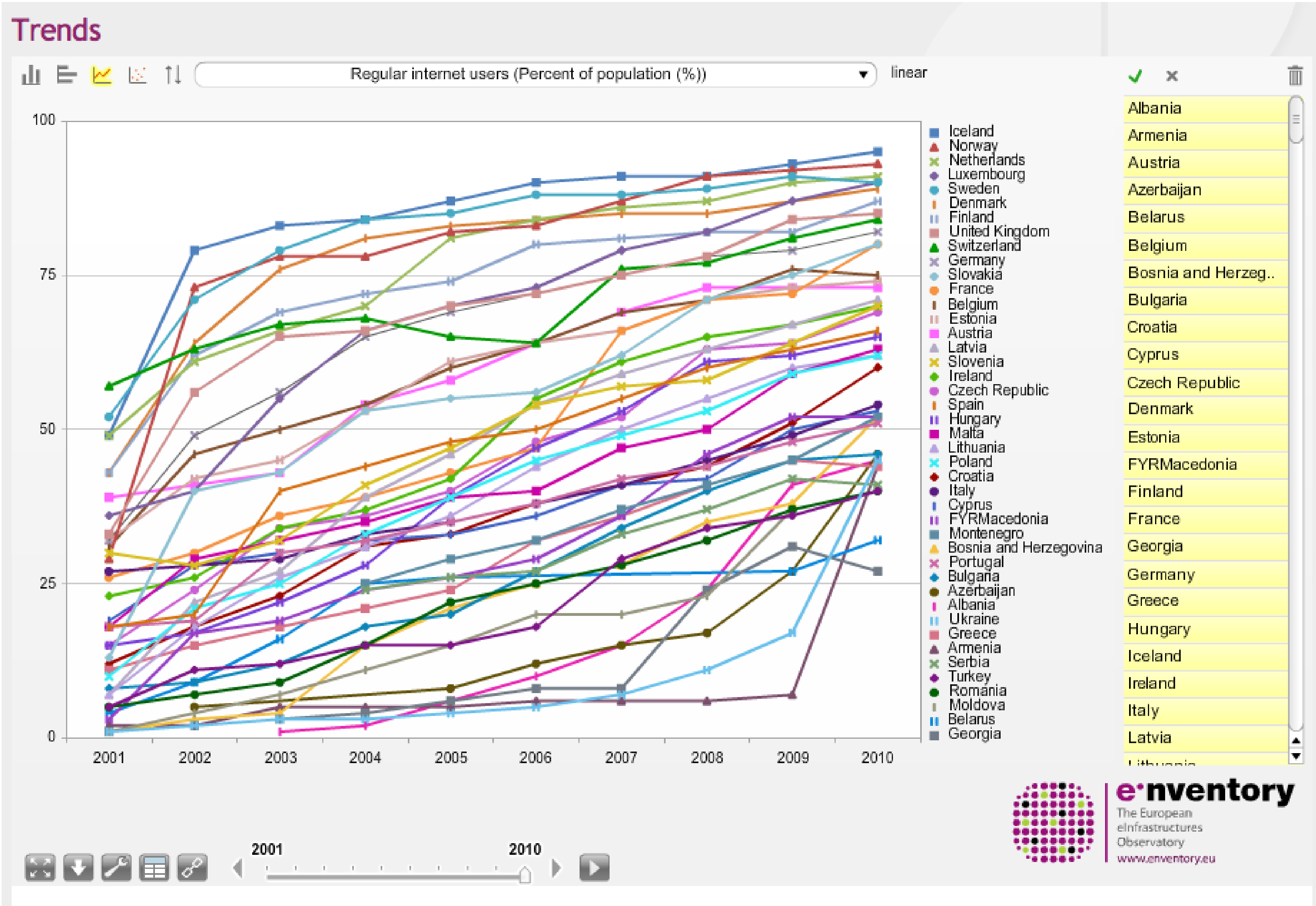


CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it



People and Science

CERN IT Department



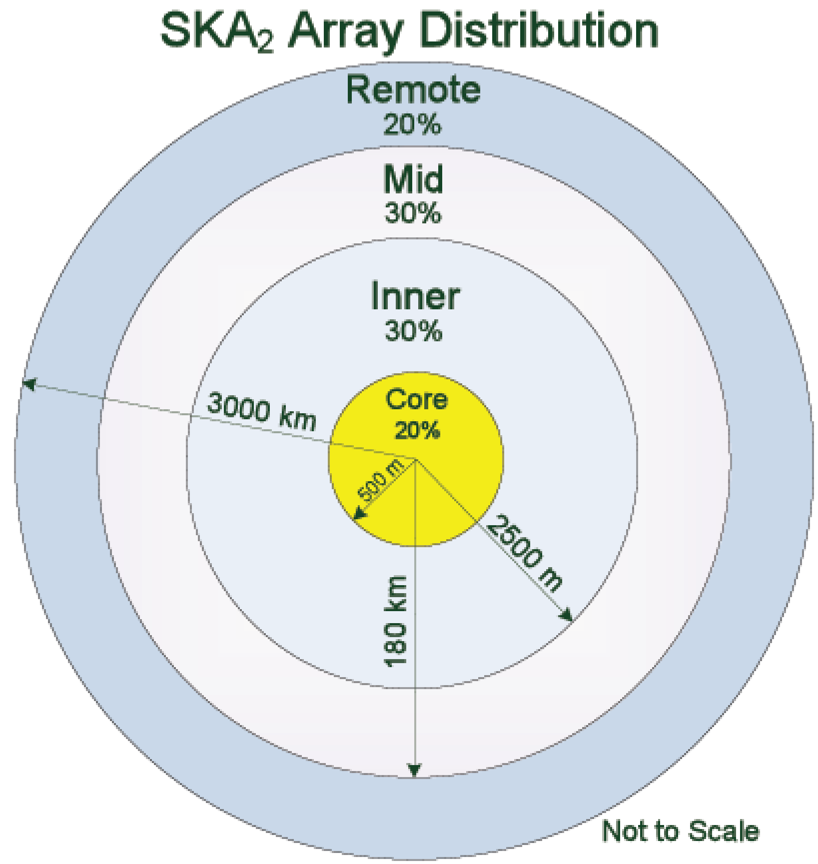
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Distribution of SKA collecting area

Diagram showing the generic distribution of SKA collecting area in the core, inner, mid and remote zones for the dish array. [1]

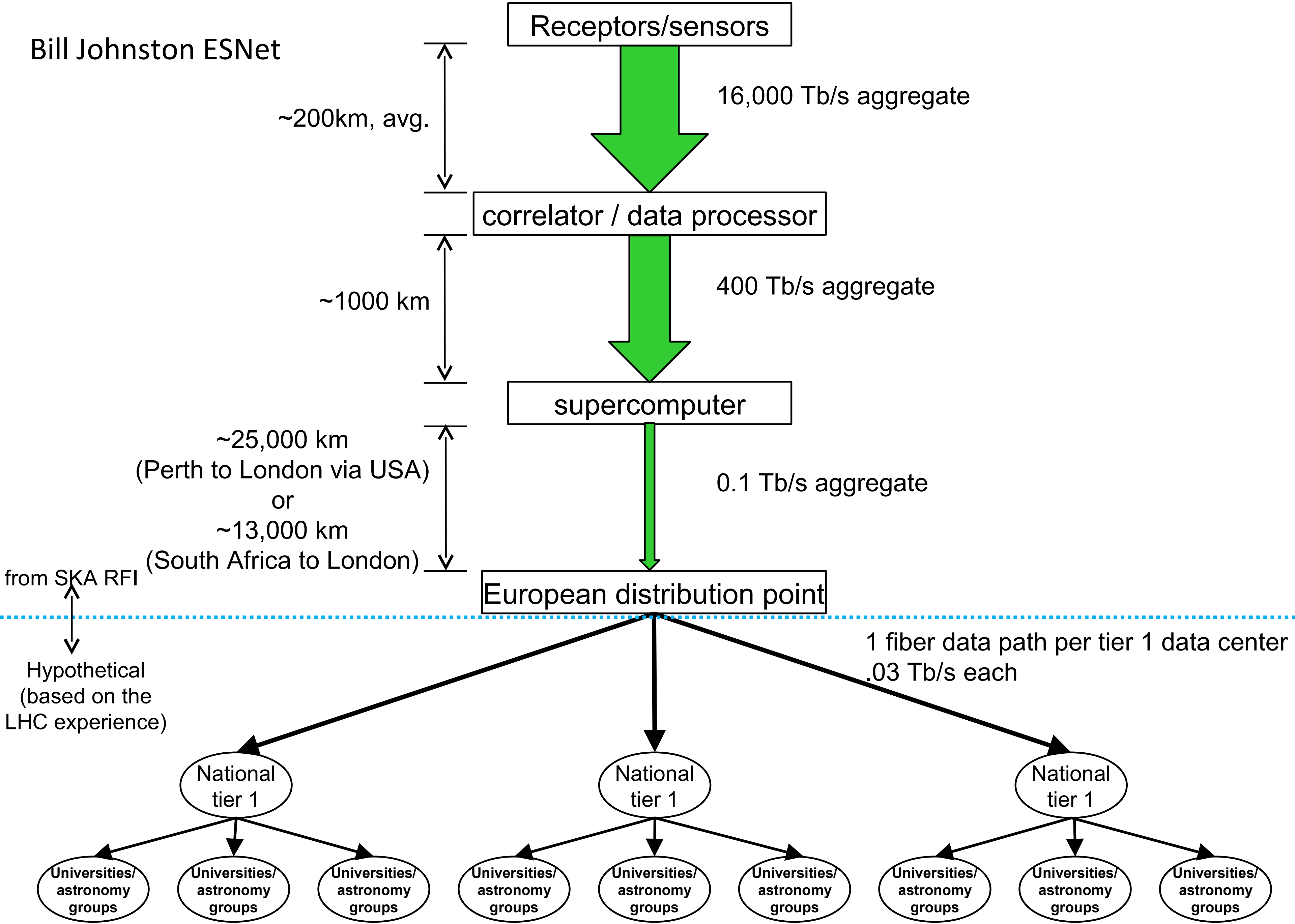
- 700 antennae in a 1km diameter **core** area,
- 1050 antennae outside the core in a 5km diameter **inner** area,
- 1050 antennae outside the inner area in a 360km diameter **mid** area, and
- 700 antennae outside the mid area in a **remote** area that extends out as far as 3000km

The core + inner + mid areas are collectively referred to as the **central** area

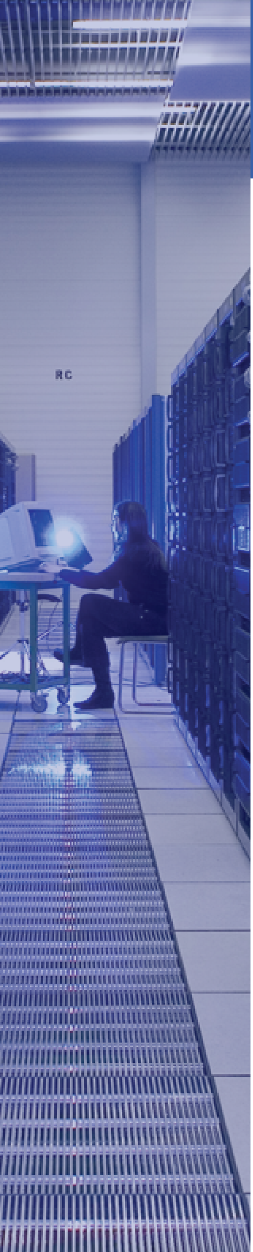


SKA data flow model

Bill Johnston ESNet

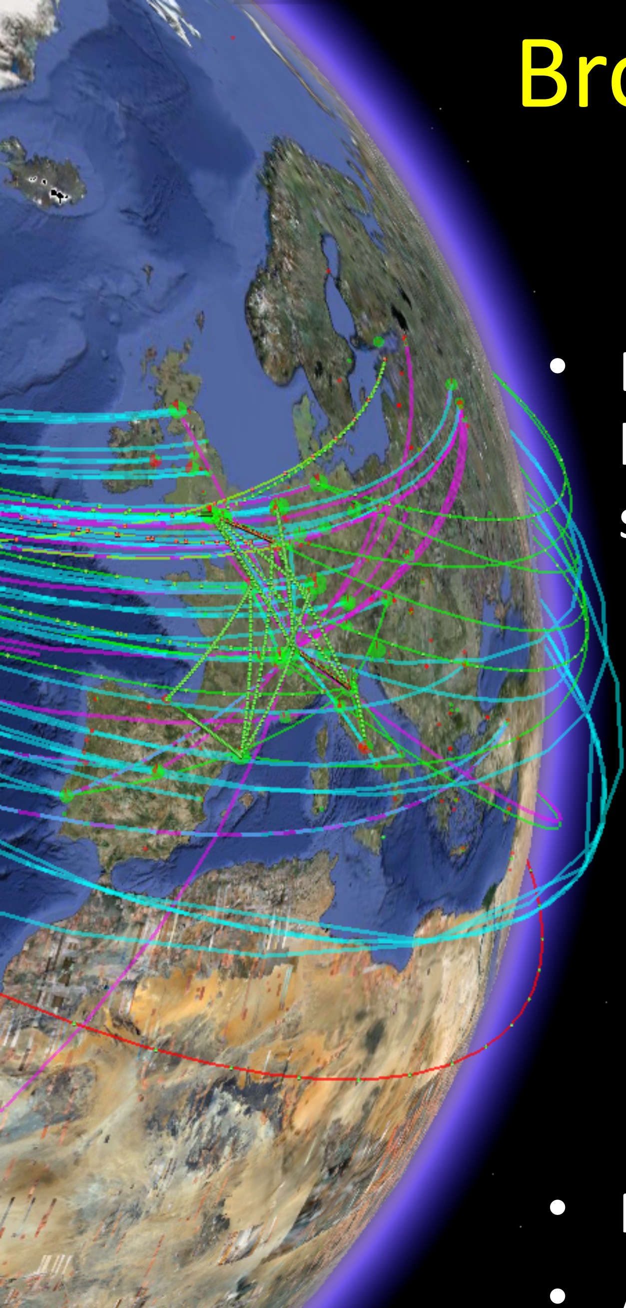


General Network Technology and Capabilities Requirements for Distribute Science



- ***Bandwidth***
  - Adequate network capacity to ensure timely movement of data produced by the facilities
- ***Reliability***
  - High reliability is required for large instruments which now depend on the network to accomplish their science
- ***Connectivity***
  - Geographic reach sufficient to connect users and collaborators and analysis systems to facilities, instruments and supercomputers.
- ***Services***
  - Guaranteed bandwidth, traffic isolation, end-to-end monitoring
  - Network ***service delivery architecture***
    - SOA / Grid / “Systems of Systems”

Broader Impact of the LHC Computing Grid



- Experience with the LHC Computing Grid has been leveraged to benefit the wider scientific community
  - Europe:
    - Enabling Grids for E-sciencE (EGEE) 2004-2010
    - European Grid Infrastructure (EGI) 2010--
  - USA:
    - Open Science Grid (OSG) 2006-2012 (+ extension?)

- Many scientific applications ➔
- Architectures for future Big Science applications

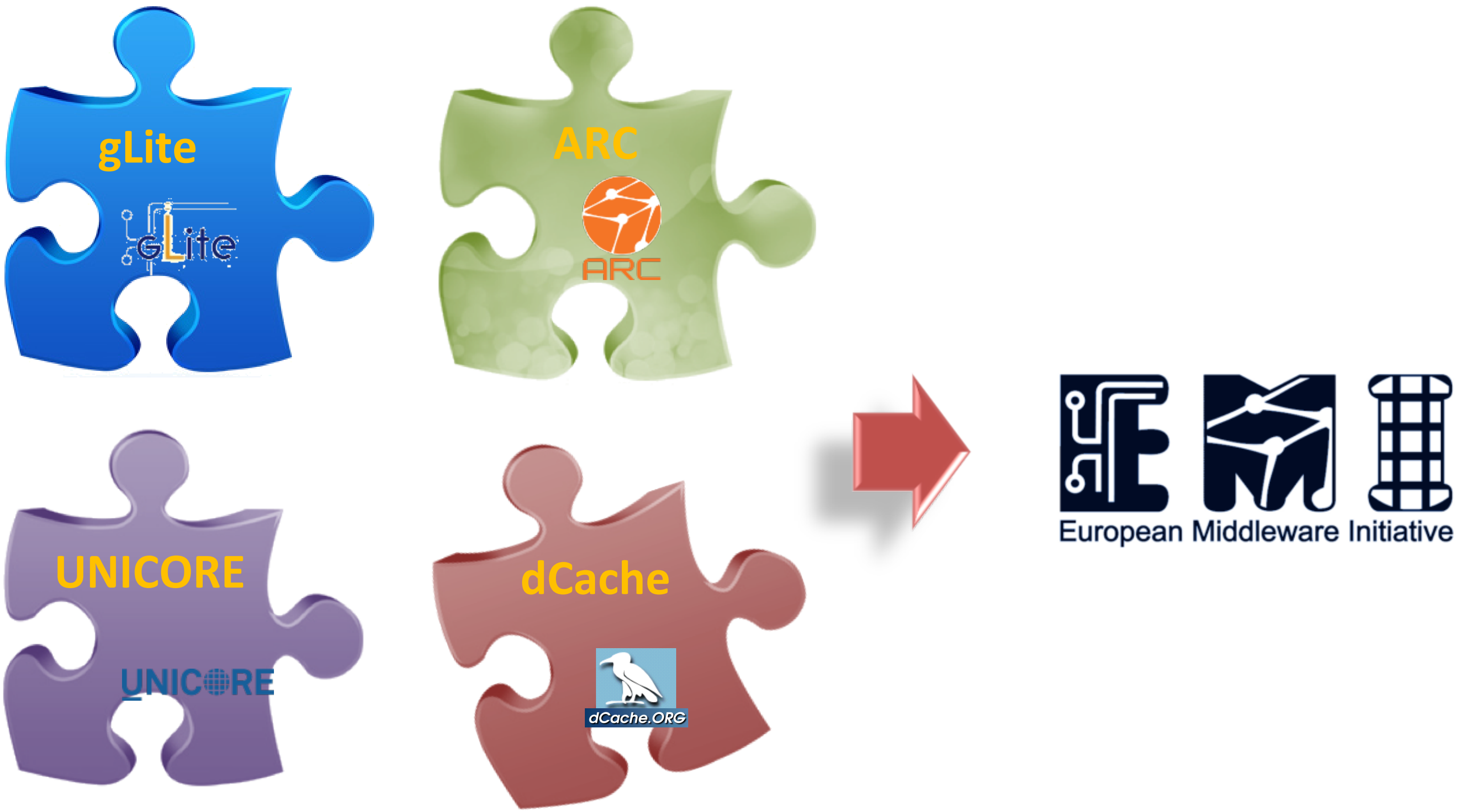
Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
Material Sciences
…

WLCG Worldwide LHC Computing Grid

Two New Initiatives

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

EMI



EMI INFSO-RI-261611

ScienceSoft

A new initiative

•Promoted by EMI in collaboration with EGI, StratusLab, iMarine, OpenAIRE and a number of other projects and SMEs

With the goal of

•Exploring the feasibility and advantages of creating an open source community for software specific to scientific communities

•Collecting community requirements, propose realistic solutions

•Making the activities of producing and using open source software for science more transparent and collaborative across communities and projects

EMI INFSO-RI-261611

Top 8 reported problems

- Lack of continuity in support, development, coordination of software
- Non-optimal communication between users and developers
- Lack of consistent real usage information
- Limited access to other users experience
- Limited or complex ways of finding what exists already
- No way of influencing the production of software
- Lack of visibility of the software activities
- No way of assessing the user “market”

EMI INFSO-RI-261611

Benefits

EUROPEAN MIDDLEWARE INITIATIVE

- More information about software and its usage
  - Categorization, usage and technical metrics
  - Assessment of costs, resource optimization
  - Supporting evidence for funding requests
  - Software licenses adoption and compliance, compatibility checks
  - More visibility for developers, more information and transparency for users
  - Peer-reviewed information

Example of Services

- Marketplace for products, services, people
  - Match demand and offer, commercial support
- Links to technical services:
  - Support, testing, deployment
  - Provided by users to users or by third-parties (including commercial companies)
- Platform integration support (third-party)
  - Definition and sharing of community-specific profiles or software stacks
  - Deployment using cloud or grid technologies

EMI INFSO-RI-261611

Examples of Services

EUROPEAN MIDDLEWARE INITIATIVE

- Support for creation of customized community and group portals
  - By technical interests, scientific domain, etc.
  - Coordination, collaboration and discussion tools
- Support for organization of technical events

EMI INFSO-RI-261611

ScienceSoft

EUROPEAN MIDDLEWARE INITIATIVE

Requirements/Gaps

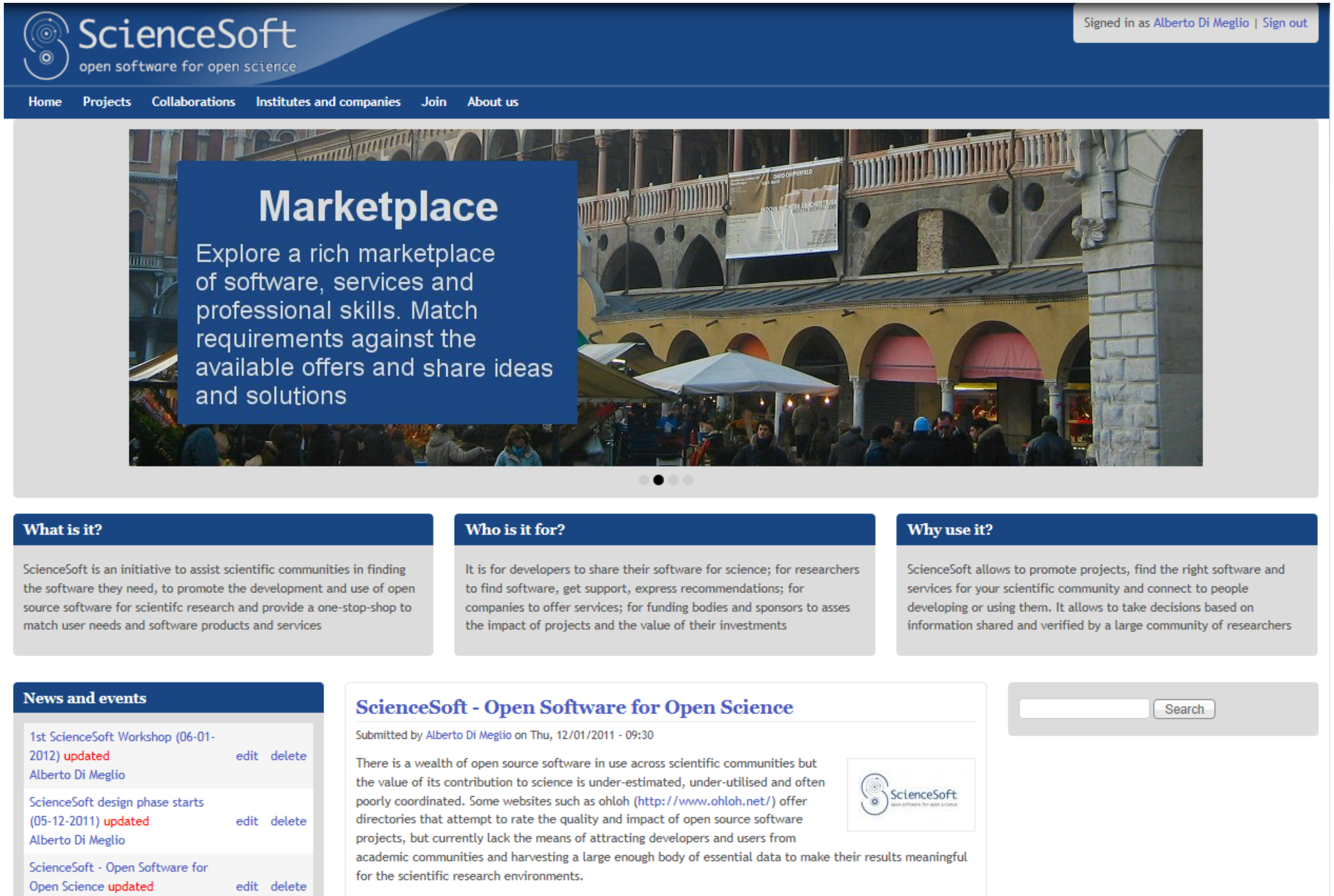
- Lack of continuity in support, development, coordination of software
- Non-optimal communication between users and developers
- Lack of consistent real usage information
- Limited access to other users' experience
- Limited or complex ways of finding what exists already
- Limited possibilities of influencing the production of software
- Lack of visibility and recognition of development activities
- No way of assessing the user "market" and potential revenues

Possible solutions

- Software and services catalogues
- Generation of usage statistics
- Honour system (Peer-reviews)
- Citation system to allow software to be referenced in papers
- Marketplace for products, services, and people to match user needs and software products and skills
- Platform integration support based on the catalogues information
- Support for creation of ad-hoc communities and groups
- Coordination, collaboration and discussion tools
- Support for organization of technical events

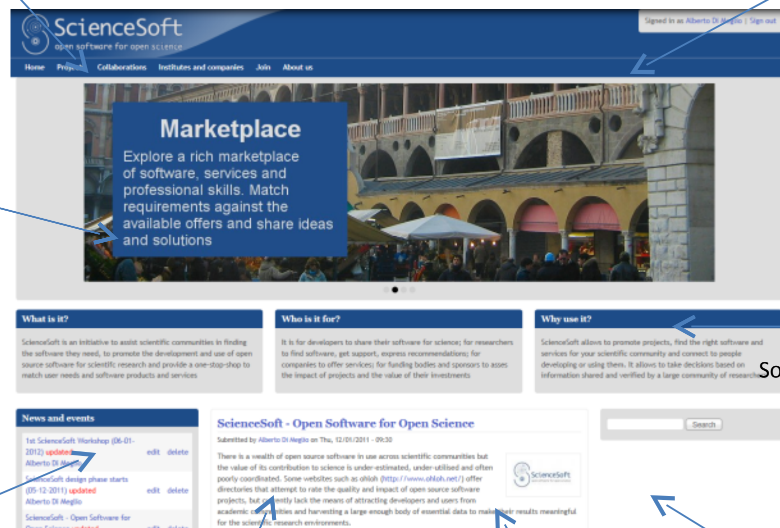
EMI INFSO-RI-261611

ScienceSoft



http://sciencesoft.org

Software inventory and metrics



Community tools and services (source code repos, image repos, testbeds, etc.)

Web and collaboration tools

Source code repositories

Software management procedures

Event management

Deployment and engineering services

The brands and projects show here have not officially endorsed in any way ScienceSoft and are used as examples of possible functionality providers

Current status

4 phases

– Concept, Design, Implementation, Operations

Concept phase:

– January 2012 to June 2012
– Discuss, share ideas, pros and cons, decide whether there is something worth pursuing or not

Design should not start later than June

EMI INFSO-RI-261611

Join

EUROPEAN MIDDLEWARE INITIATIVE

ScienceSoft

open software for open science

http://sciencesoft.org/join

EMI INFSO-RI-261611

Grid vs Cloud a Personal View

- Its like the difference between “Pay as you View” TV and a “Film Club”
  - Resource allocation and ownership
  - Economic model
  - Collaboration model
  - Operational model
  - Service model
- But they are complementary, you may use both!

Objectives of the initiative

- Set up a cloud computing infrastructure for European Research Area
- Identify and adopt policies for trust, security and privacy on a European-level
- Create a light-weight governance structure involving all stakeholders
- Define a short and medium term funding scheme

Supply-side companies: Atos Origin, BT Services, Cap Gemini, CloudSigma, Interoute, Logica, Orange, SAP, Terradue, The ServerLabs, T-Systems, SixSq, Terradue, Thales, Telefonica, EGI.eu, OpenNebula, etc.

Timeline

Set-up (2011)

Select flagships use cases,
identify service providers,
define governance model

Pilot phase (2012-2014)

Deploy flagships,
Analysis of functionality,
performance & financial model

Full-scale cloud service market (2014 …)

More applications,
More services,
More users,
More service providers

Pilot Phase

- Through the pilot phase we expect to explore/push a series of perceived barriers to Cloud adoption:
  - **Security**: Unknown or low compliance and security standards
  - **Reliability**: Availability of service for business critical tasks
  - **Data privacy**: Moving sensitive data to the Cloud
  - **Scalability/Elasticity**: Will the Cloud scale-up to our needs
  - **Network performance**: Data transfer bottleneck; QoS
  - **Integration**: Hybrid systems with in-house/legacy systems
  - **Vendor lock-in**: Dependency on vendors once data & applications have been transferred to the Cloud
  - **Legal concerns**: Such as who has legal liability
  - **Transparency**: Clarity of conditions, terms and pricing

Flagship use cases
Participating Suppliers

CloudSigma

Atos

interoute
from the ground to the cloud

logica
be brilliant together

terradue 2.0

sixsq.

T-Systems

the SERVER LABS
the IT architects

Flagship use cases

	ATLAS H.E.P. Cloud Use (CERN)	Genomic Assembly in the Cloud (EMBL)	SuperSites Exploitation Platform (ESA/CNES/DLR)
Scientific goal/society impact/photogenic	●	●	●
Scale of resources used	●	●	
Federation/Aggregation of datasets		●	●
Long-term archiving of data			●
On-demand processing	●	●	●
Impact on community & benefits	●	●	●
Potential increase of users	●	●	●
Interoperability	●	●	●
Data security	●	●	●
Maturity	●	●	●
Access to license-controlled sw			●

ATLAS Cloud Computing R&D

EFFICIENCY, ELASTICITY

- ATLAS Cloud Computing R&D is a young initiative
  - Active participation, almost 10 persons working part time on various topics
  - Goal: How we can integrate cloud resources with our current grid resources?

- **Data processing and workload management (Today's topic)**
  - PanDA queues in the cloud
    - Centrally managed, non-trivial deployment but scalable
    - Benefits ATLAS & sites, transparent to users
  - Tier3 analysis clusters: instant cloud sites
    - Institute managed, low/medium complexity
  - Personal analysis queue: one click, run my jobs
    - User managed, low complexity (almost transparent)
- **Data storage**
  - Short term data caching to accelerate above data processing use cases
    - Transient data
  - Long term data storage in the cloud
    - Integrate with DDM

Helix Nebula EC project proposal

Coordination action submitted to INFRA-2012-3.3 in November 2011

- Negotiations currently under way: if successful, the project will start 1st June 2012

no.	Organisation name	Short name	Country
1 (coord)	European Organization for Nuclear Research	CERN	CH
2	STICHTING EUROPEAN GRID INITIATIVE	EGI.eu	NE
3	European Molecular Biology Laboratory	EMBL	DE
4	ATOS ORIGIN NEDERLAND	Atos	NE
5	T-Systems International GMBH	T-Systems	DE
6	CLOUDSIGMA AG	CloudSigma	CH
7	SAP AG	SAP	DE
8	Logica Deutschland GmbH & Co KG	Logica	DE
9	CONSIGLIO NAZIONALE DELLE RICERCHE	CNR	IT
10	Cloud Security Alliance EMEA	CSA	UK

A European cloud computing partnership: big science teams up with big business

HELIX NEBULA THESCIENCECLOUD

Strategic Plan

- Establish multi-tenant, multi-provider cloud infrastructure
- Identify and adopt policies for trust, security and privacy
- Create governance structure
- Define funding schemes

CERN

To support the computing capacity needs for the ATLAS experiment

EMBL

Setting up a new service to simplify analysis of large genomes, for a deeper insight into evolution and biodiversity

esa

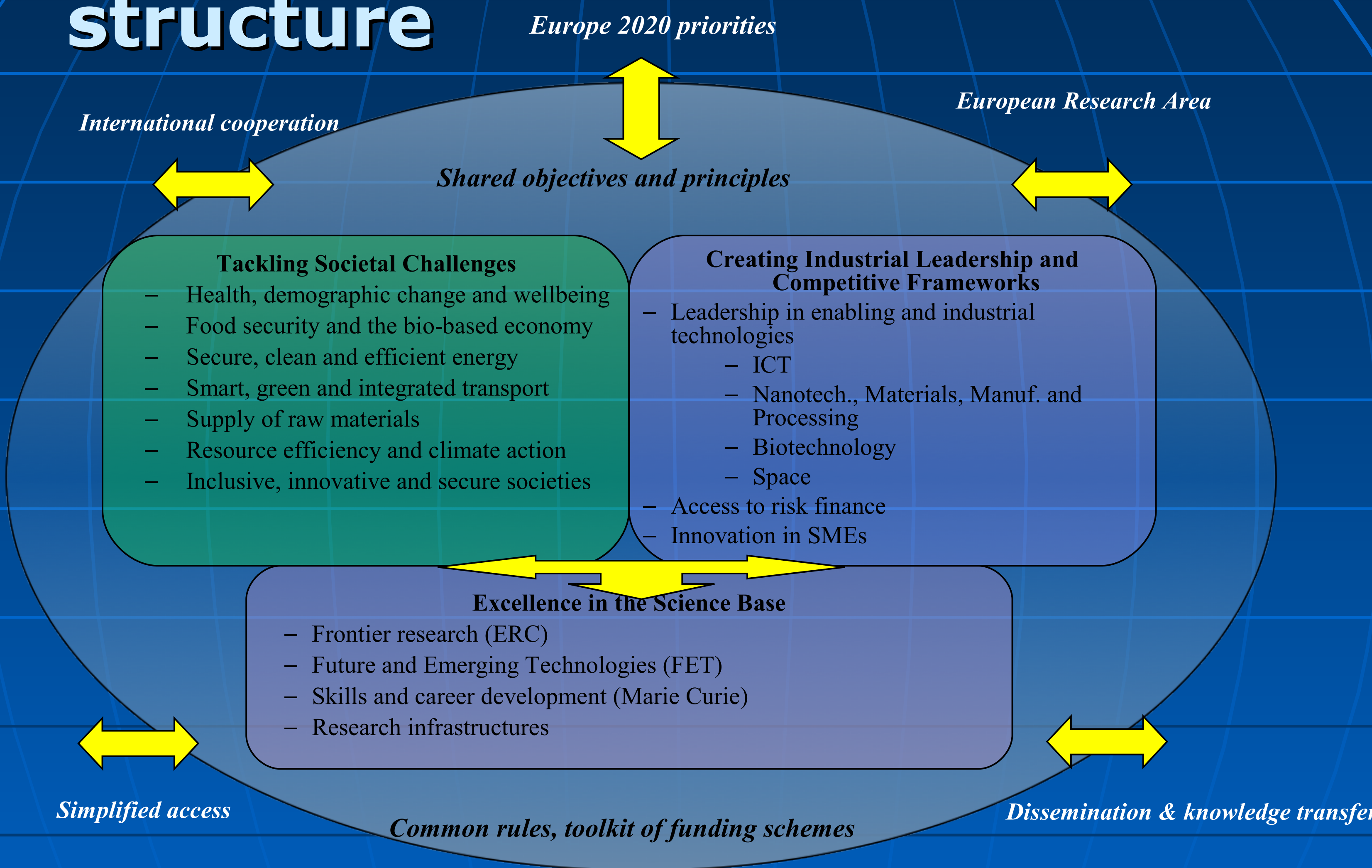
To create an Earth Observation platform, focusing on earthquake and volcano research

Atos · Capgemini (Consulting.Technology.Outsourcing) · CloudSigma · egi · interoute (from the ground to the cloud) · logica (be brilliant together) · OpenNebula.org (The Open Source Toolkit for Cloud Computing) · orange Business Services · SAP · the SERVER LABS (the IT architects) · sixsq · Telefónica · terradue 2.0 · THALES · T-Systems

Email:contact@helix-nebula.eu Twitter: HelixNebulaSC Facebook: HelixNebula.TheScienceCloud

Finally

Horizon 2020 – Objectives and structure



ICT Services of public interest
(from Guidelines, COM(2011) 657/3)

Trans-European high-speed backbone connections for public administrations

"A public trans-European backbone service infrastructure will provide very high speed and connectivity between public institutions of the EU in areas such as **public administration, culture, education and health**."

" **Core service** platform… In particular it will provide connectivity for other trans-European services inter alia those mentioned in this Annex. This infrastructure will be fully **integrated in the Internet as a key capacity for trans-European public service** and will support the adoption of emerging standards (IPv6)…"

"The integration of the core platform into the European public services will be facilitated by the deployment of **generic services**: authorisation, authentication, inter-domain security and bandwidth on demand, federation of services, mobility management, quality control and performance control, integration of national infrastructures. Interoperable **'cloud computing'** …"

See: http://ec.europa.eu/budget/reform/commission-proposals-for-the-multiannual-financial-framework-2014-2020/index_en.htm

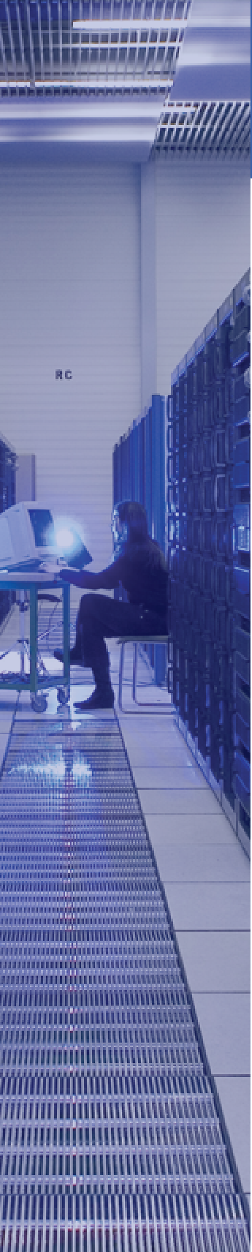
Conclusion

- We see a vibrant range of activities:
  - Science initiatives across all disciplines.
  - Software, Infrastructure and Services developing rapidly
  - Grid an established success and Cloud computing for science emerging.
- Excellent R&E networking is:
  - A major driver of economic growth.
  - **<u>Fundamental</u>** to participating in global scientific activities in the 21st century!

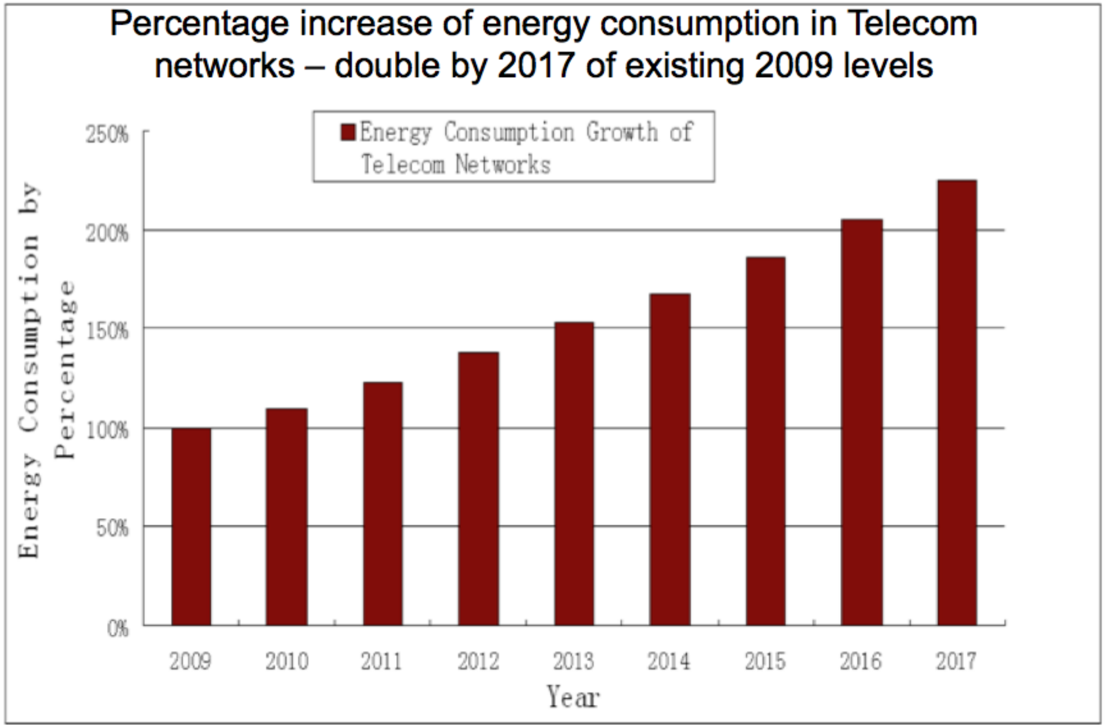
Extras

CERN IT Department

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



Energy consumption is a growing concern



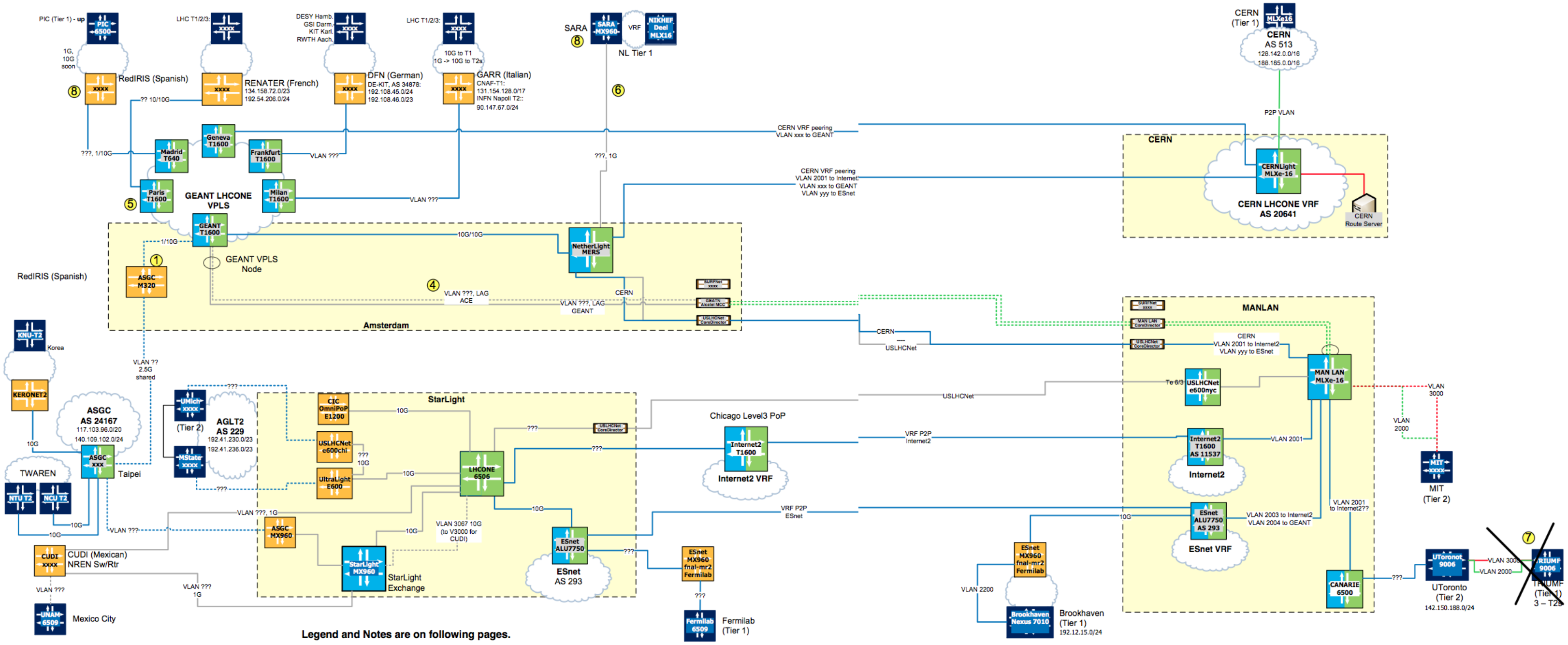
Source: Yi Zhang; Chowdhury, P.; Tornatore, M.; Mukherjee, B.; , "Energy Efficiency in Telecom Optical Networks," Communications Surveys & Tutorials, IEEE , vol.12, no.4, pp.441-458, Fourth Quarter 2010

Community Measurement, Architecture, Planning are all needed.
Nice talk here from Inder Monga (ESNet):
http://www.glif.is/meetings/2011/rap/monga-greenactivities.pdf

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

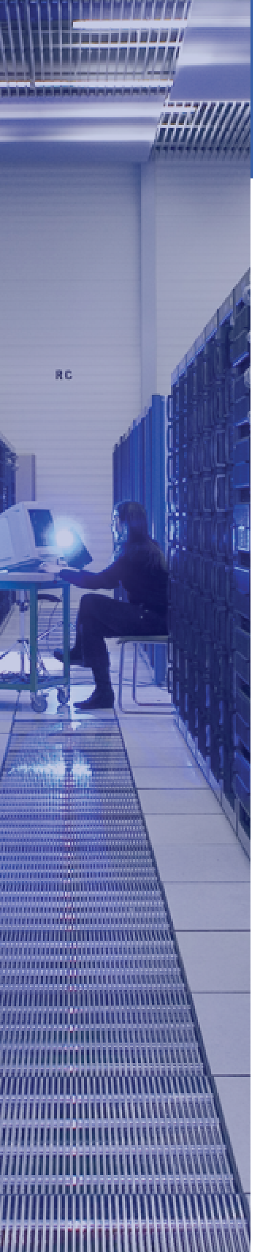
LHCONE VRF Layout

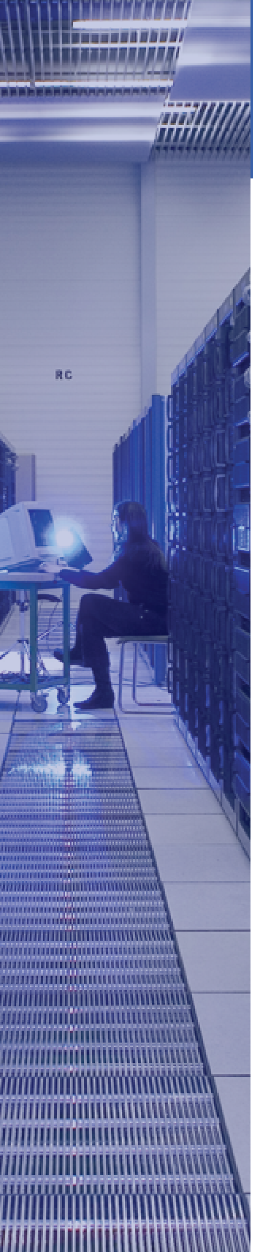
CERN IT Department



From Bill Johnston ESNet

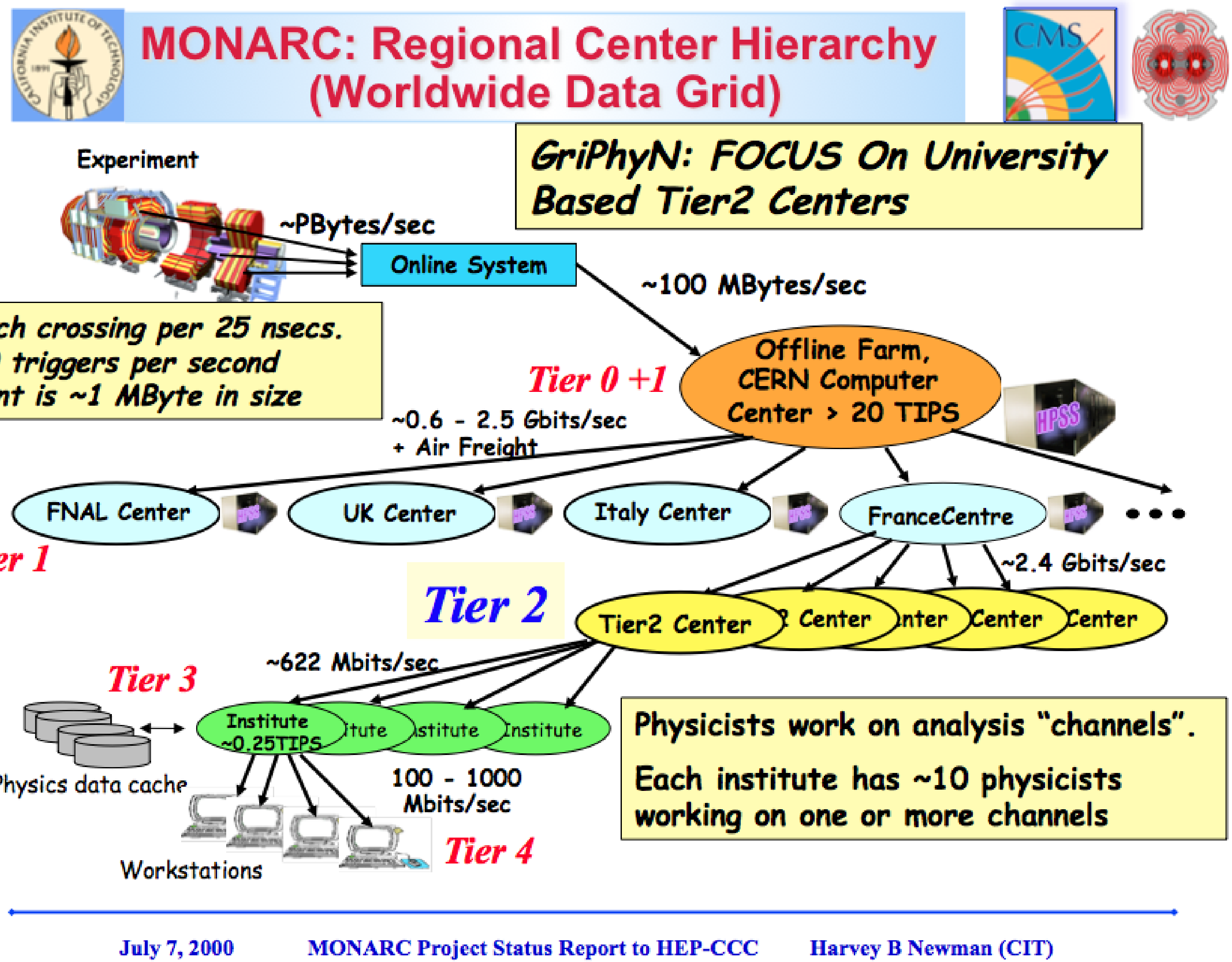
DRAFT





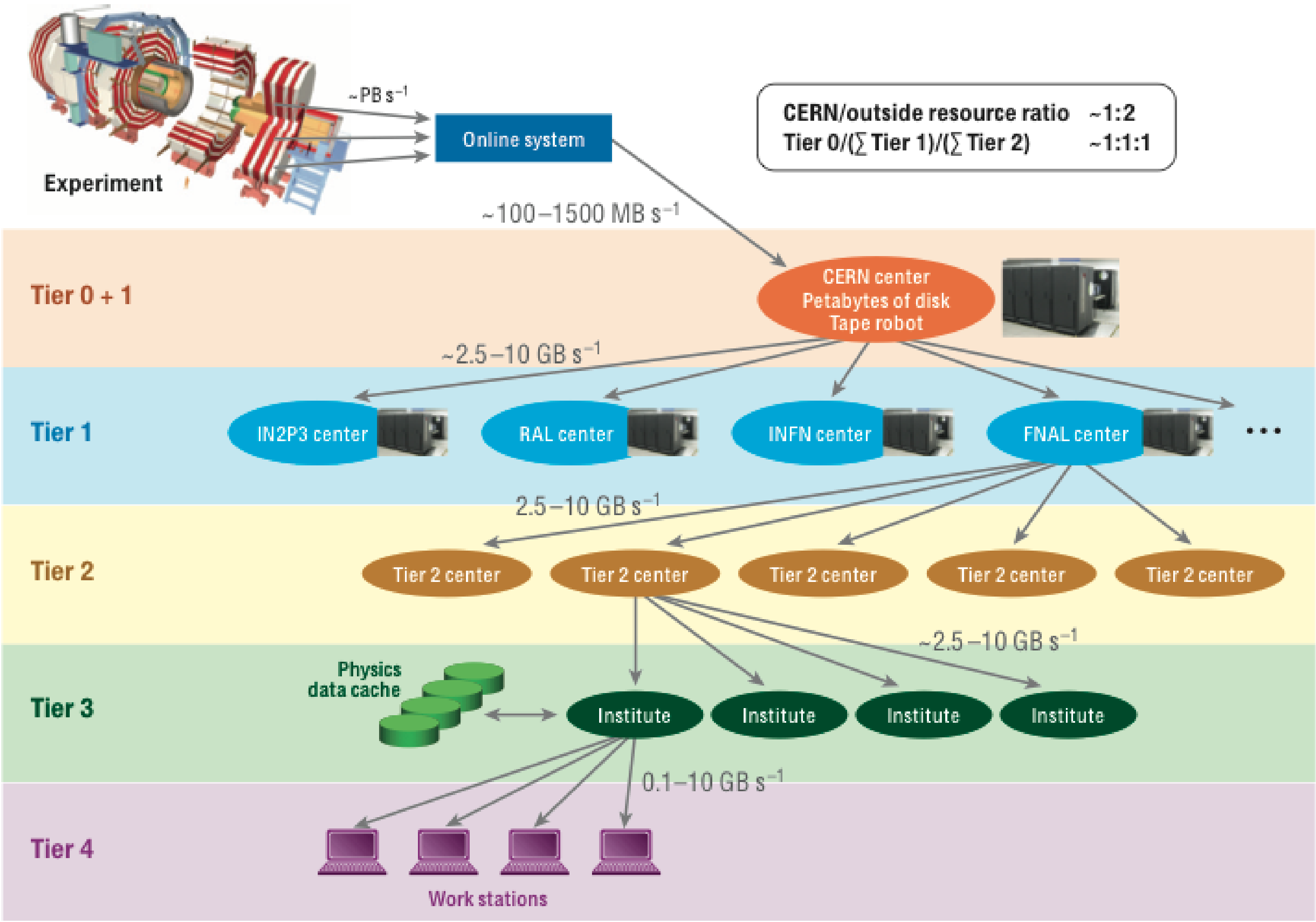
MONARC

CERN IT Department



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

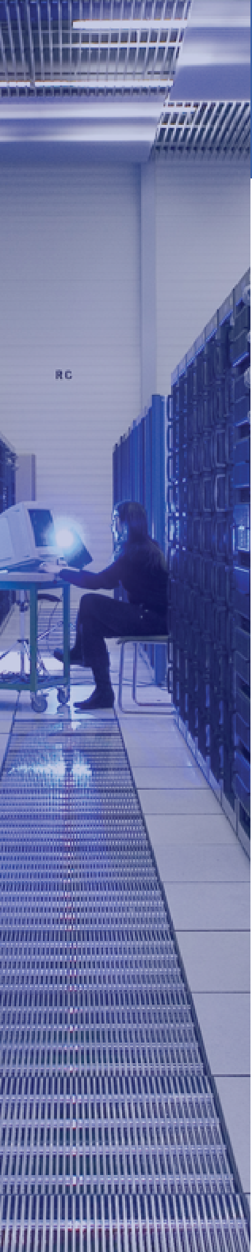
“Classical” Computing Model

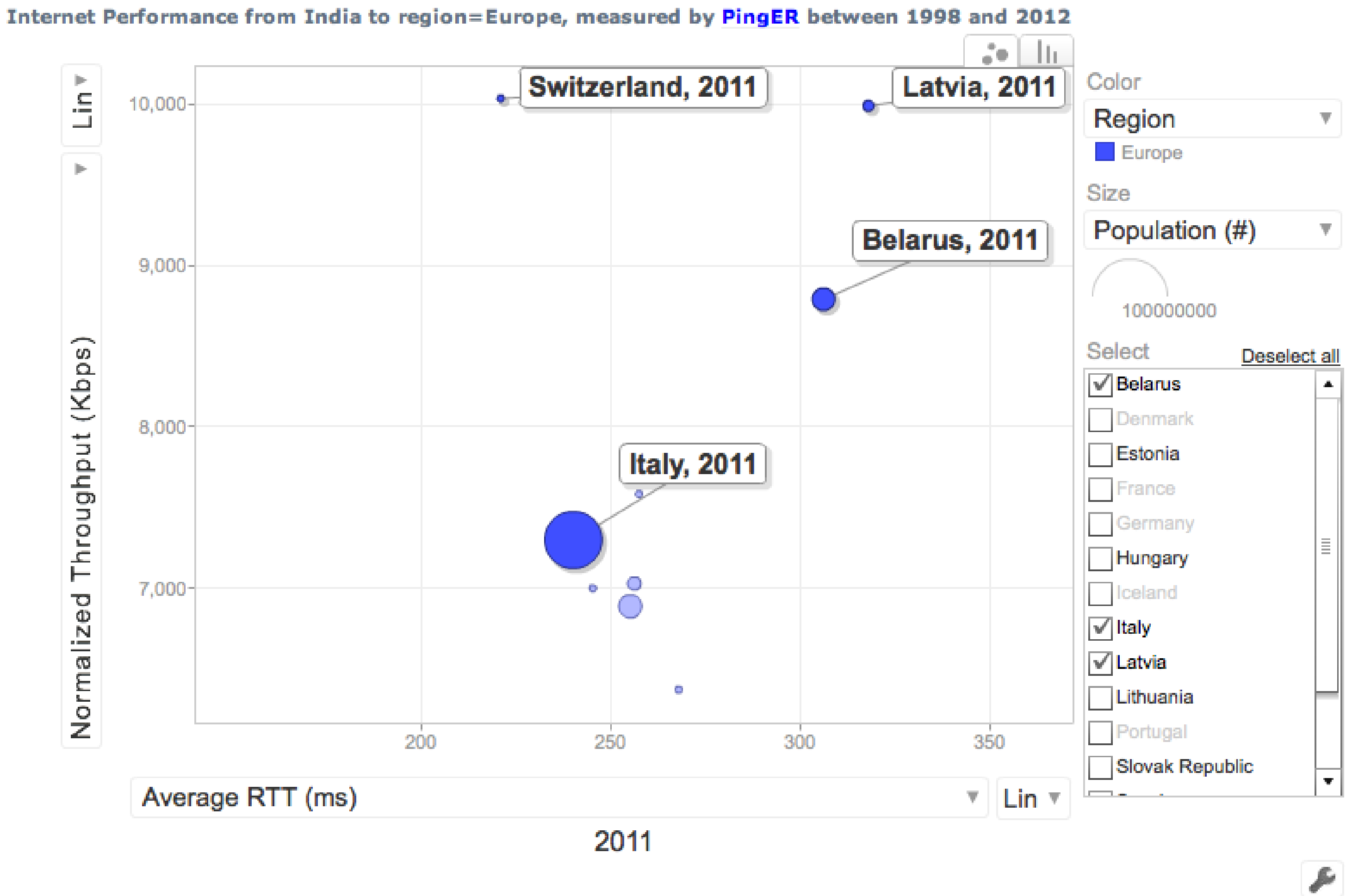


CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

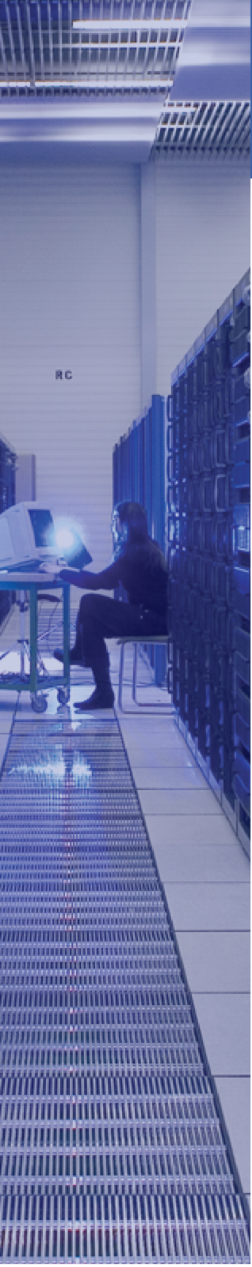
Pinger Data India Europe

CERN IT Department





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



PingER Throughput

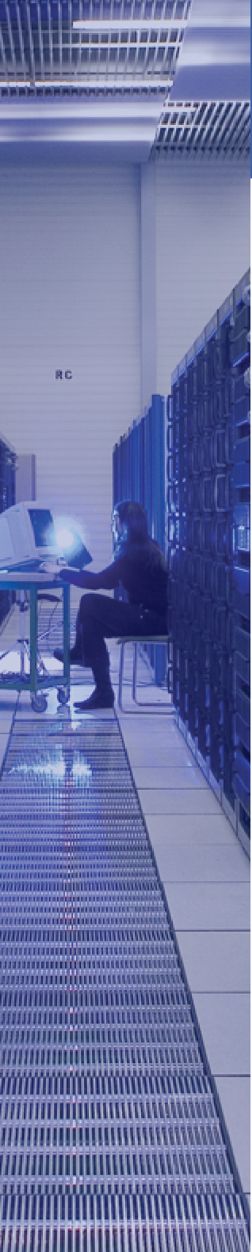
CERN IT Department

Summary for Sites in IN.CDACMUMBAI.N1 seen from CH.CERN.N20

Tick	min	25th%	avg	median	75th%	90th%	95th%	max	iqr	std dev	# pairs
Mar2012	1277.429	.	1277.429	1277.429	.	.	.	1277.429	.	.	1
Feb2012	945.003	.	945.003	945.003	.	.	.	945.003	.	.	1
Jan2012	948.680	.	948.680	948.680	.	.	.	948.680	.	.	1
Dec2011	842.238	.	842.238	842.238	.	.	.	842.238	.	.	1
Nov2011	707.585	.	707.585	707.585	.	.	.	707.585	.	.	1
Oct2011	653.122	.	653.122	653.122	.	.	.	653.122	.	.	1
Sep2011	587.445	.	587.445	587.445	.	.	.	587.445	.	.	1
Aug2011	620.559	.	620.559	620.559	.	.	.	620.559	.	.	1
Jul2011	556.555	.	556.555	556.555	.	.	.	556.555	.	.	1
Jun2011	622.824	.	622.824	622.824	.	.	.	622.824	.	.	1
May2011	689.270	.	689.270	689.270	.	.	.	689.270	.	.	1
Apr2011	597.738	.	597.738	597.738	.	.	.	597.738	.	.	1

Median value of medians is 671.196

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



The science data distribution

- Guaranteed bandwidth virtual circuit services have become increasingly important.
- The environment of large-scale science is inherently multi-domain
- The unique service characteristics that have evolved are
  - Guaranteed, reservable bandwidth with resiliency
  - Requested and managed in a Web Services framework
  - Traffic isolation and non standard transports
  - Traffic engineering (for network operations)
  - Secure connections
  - Flexible service semantics
  - Rich service semantics – e.g. to reliability through redundancy.