



CERN-Turkey Industry Day

Ankara

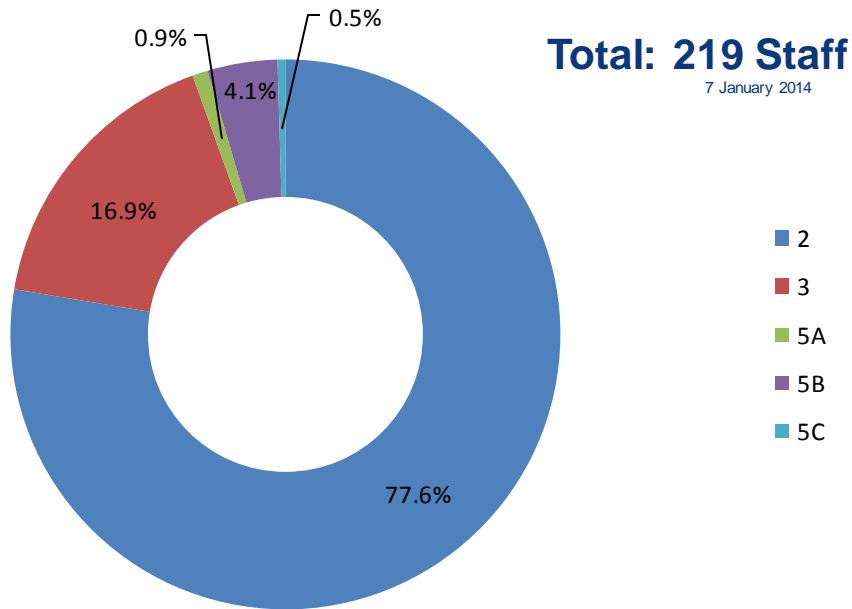
5 October 2015

CERN Information Technology Department

Overview

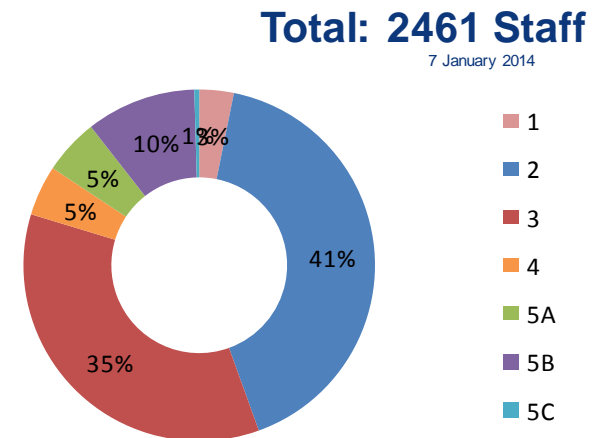


IT Staff compared to CERN



- 2
- 3
- 5A
- 5B
- 5C

1	Experimental & Theoretical Physics
2	Scientific & Engineering Work
3	Technical Work
4	Manual Work, Crafts & Trades
5A	Administrative Work
5B	Office & Administrative Work
5C	Office Work

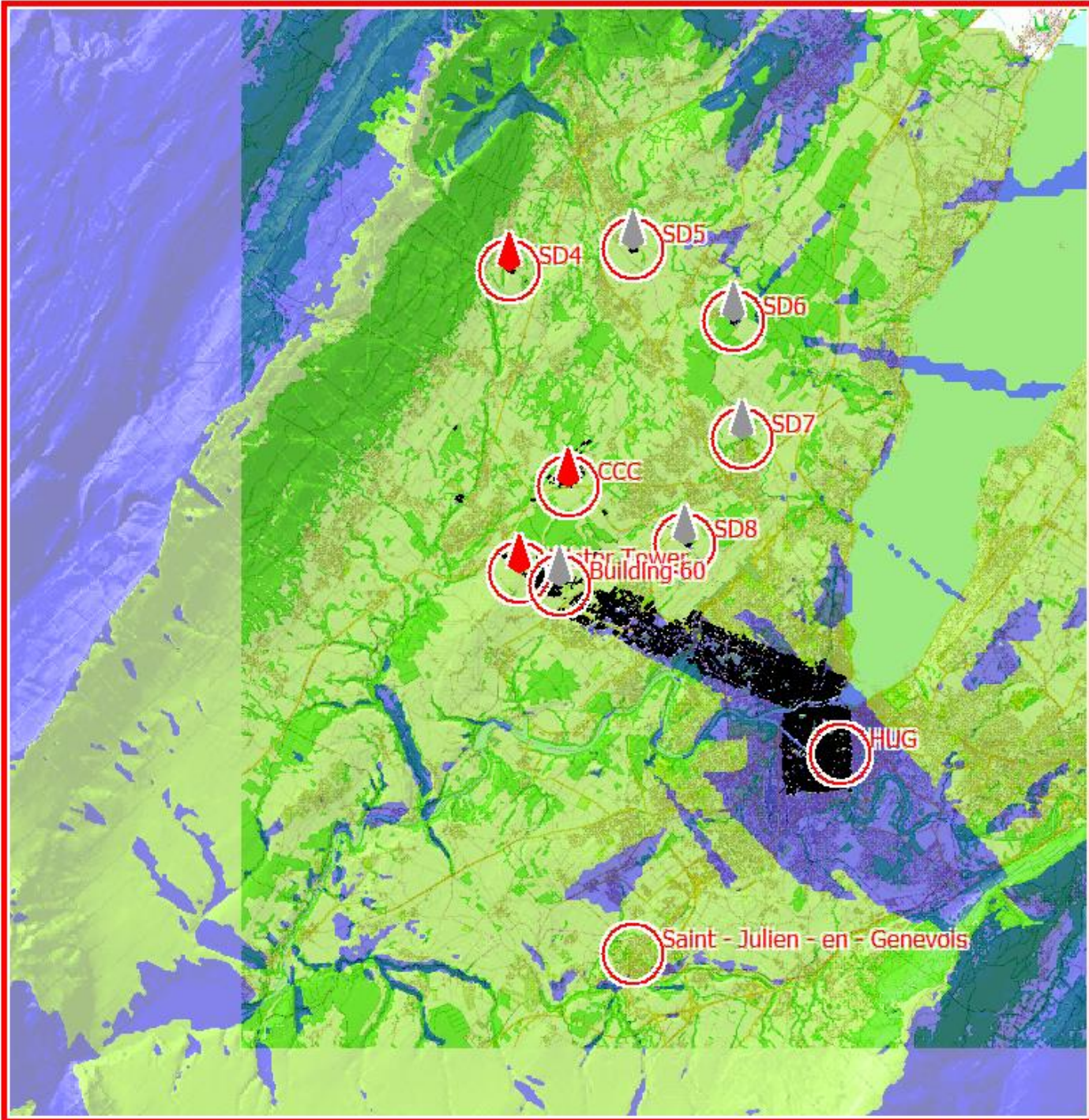


- 1
- 2
- 3
- 4
- 5A
- 5B
- 5C

IT Department Services

- **E-Mail and Distribution Lists**
 - Up to 230000 incoming messages/day, 70% detected as spam
 - 33'000 mailboxes
- **AFS (worldwide) Distributed File System**
 - 2.6 B files, 340TB, 4 B access/day, 15000 clients
- **Backup Services (files, databases)**
 - 8.5 PB, 2.5 B files
- **EOS**
 - 140 PB, 202 M files (+18%), but 120 JBOD PB deployed in preparation for LHC run 2
 - Geolocalization (Wigner / Meyrin) implemented (40% / 60 %)
- **Network, Telephony**
 - More than 400 star points and ~80 000 UTP sockets
 - ~450 wireless base stations
 - GSM (Site, Tunnels & Experimental halls) – 4300 subscriptions
 - Multi-10Gb/s External Networking Capacities
 - TETRA replaced VHF communications for Fire Brigade
 - MS Lync
- **Web Services**
 - 12000 Web sites
 - IIS, AFS hosted, Drupal, Sharepoint
- **Active Directory & SSO**
 - Central authentication service for Linux and Windows computers and (Web) applications
 - Online X509 Certificate Authority (~8'000 User certificates, 40'000 host certificates)
 - Multifactor authentication introduced
 - Federated authentication
- **Windows Services**
 - 290 TB of DFS workspaces
 - ~ 13'000 active PCs managed by CMF
 - 3600 Macs
- **CVI Virtualisation Infrastructure**
 - 3400 VM's (60% Linux, 40% Windows)
 - 355 Hypervisors
- **Cloud Infrastructure**
 - 5900 VM's, 2790 Nodes, 63500 Cores





Received Power
[dBm]

- > -95.00
- ≤ -95.00



IT Department Services

- **Database and Application Deployment Services**

- Mainly based on Oracle software
- AIS DBs and Applications, EDMS, Accelerator DBs, IT DBs, CASTOR DBs, Physics databases (Calibration, Alignment, etc...), Public J2EE Service, etc...
- 120 General Purpose Databases, 500 TB of NAS storage
- 130 Web /Application Servers with 700 virtual hosts
- 50 Terabytes of worldwide replicated Physics databases
- Database on Demand Service (Oracle, MySQL)

- **Engineering and Software Development Services**

- Mechanical and electronic CAE, field calculations, structural analysis, simulations, mathematics, etc
- 50 packages, 1000 users
- Twiki Service
- 12'000 users, 60'000 pages updated per month
- Version Control Services (CVS/SVN)
- 2'500 users, 400 projects

- **Audiovisual Service**

- support, record and archive official committees and events

- **Video Conference Service**

- video conferencing in rooms across site

- **Conference Management System (Indico)**

- Distributed and used worldwide

- **CDS-Invenio**, a Digital Library Open Source Software produced, used and maintained at CERN

- free support via mailing lists
- commercial-like support via a maintenance contract

- **Computer Security**



Higgs Update Seminar 4.7.2012

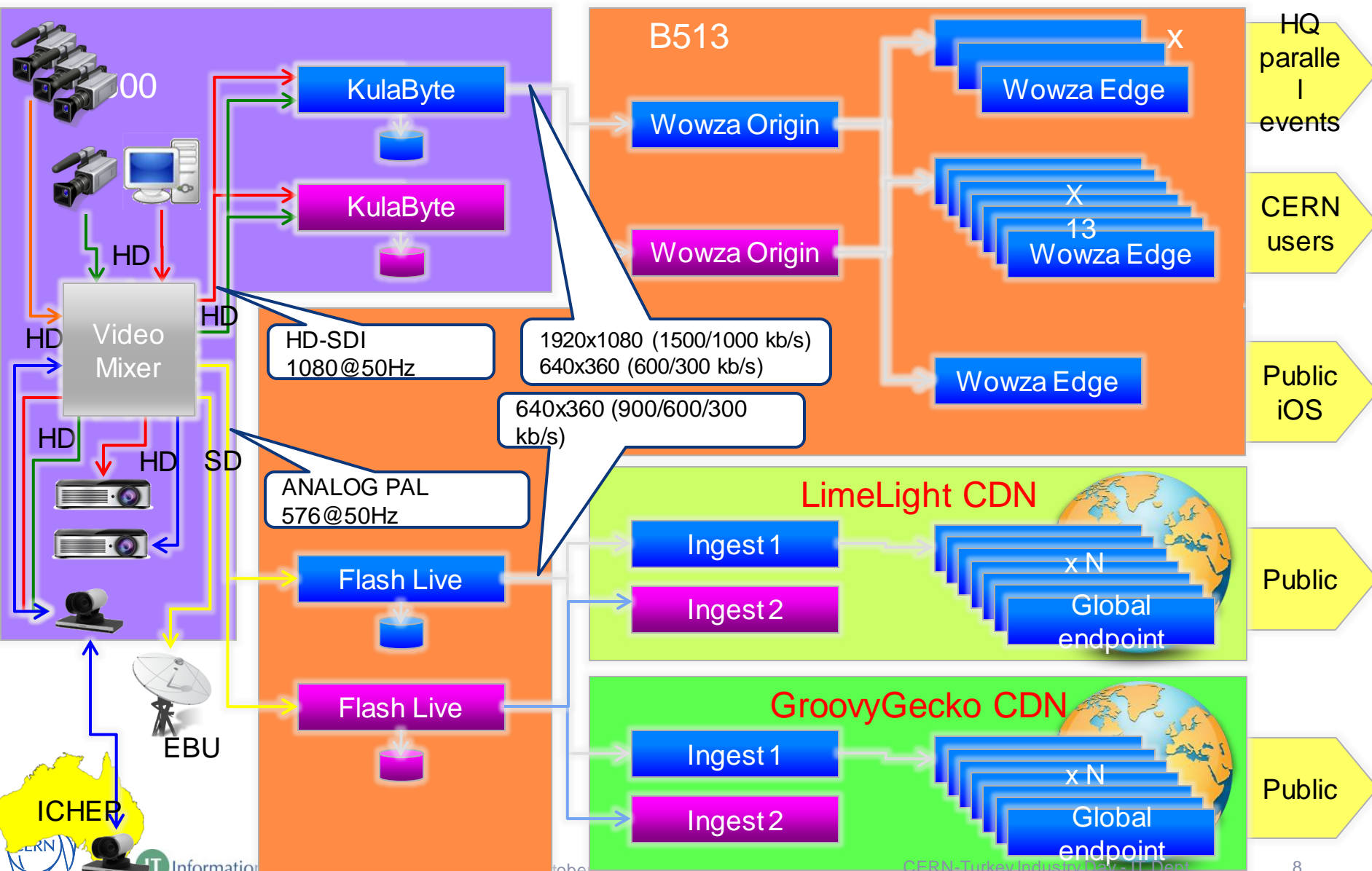
- Very short notice given
- Worldwide visibility
- A/V Team (and others) worked days & nights!
- Estimated 1 billion people reached
 - Including Antarctica ...
 - <http://avc-dashboard.web.cern.ch/node/3>
- Main auditorium upgraded with HQ just, just in time ...



Encoding

Ingestion

Streaming



Invenio

Invenio is an open source software suite enabling you to run your own technology offered by the software covers all aspects of digital classification, indexing, and curation to dissemination. Invenio compl metadata harvesting protocol (OAI-PMH) and uses MARC 21 as its u

- ~ 30 official installations
 - <http://invenio-software.org/wiki/General/Demo>
- 7 on-going support contracts:
 - Smithsonian Astrophysical Observatory
 - University Library in Bratislava
 - Centre ERL TIC

The screenshot shows the top navigation bar with links: HOME, REPOSITORY, LIBRARY SYSTEM, SERVICE, ABOUT, BLOG, NEWS. The 'ABOUT' link is highlighted. Below the navigation, there are two columns of text. The left column is titled 'ABOUT TIND.' and discusses Invenio's history since 2002 and its status as a professional spin-off. The right column is titled 'ABOUT INVENIO.' and describes its use at CERN and its open source community. At the bottom of the content area, there are two logos: 'spin-off' with a CERN logo and 'technology' with a CERN logo.

This is a smaller version of the website header, showing the 'spin-off' logo, the 'TIND' logo, and the navigation menu: HOME, REPOSITORY, LIBRARY SYSTEM, SERVICE, ABOUT, BLOG, NEWS.

This section features a blue background with the text 'MANAGE, SHOWCASE AND PRESERVE ALL DIGITAL ASSETS.' Below this, there are four icons representing different types of digital assets: a document for 'RESEARCH OUTPUT', a bar chart for 'RESEARCH DATA', a play button for 'MULTIMEDIA', and a bookshelf for 'CIRCULATION'. Each icon is accompanied by a brief description of the asset type. At the bottom of this section, it states 'CERN open source software provided as a professional cloud service.'

Studies)
 establish institutional repositories.
 anda 2009 and Morocco 2010,
 with Kumasi University (Ghana)
 s authorizations



Indico



20-23 October 2010 Nürnberg, Germany
Europe/Berlin timezone

Please check the current list of BOFs and General Conference Information

- Overview
- Program Overview**
- Schedule
- Presentations
- Author index

Support

Platin Sponsor:



Bronze Sponsors:



Media Partner:



Collaboration Across Borders

A big part of the conference will be presentations aimed to FOSS contributors and users. Another part will be Birds of a Feather Sessions where people work on specific topics in an interactive discussion style. Hacking sessions of various special interest teams within the community make the real FOSS conference spirit and will give you very short nights.

Below you find the tracks in which the program is organized.

Packaging and openSUSE Buildservice (OBS)

The openSUSE Build Service provides software developers with a convenient and easy to use tool to build open source software for openSUSE and other Linux distributions. This track contains presentations about effective packaging technologies, collaboration in distribution maintenance and development.

Distribution Technology

Take a look into the distribution factory and discover which technologies will drive distributions tomorrow. Key topics are Quality Assurance, more Usable configuration tools and desktops.

Free Desktop

Free Desktops are the pillar of every Linux distribution. This track offers an overview over the plans of the GNOME and KDE desktops and their bleeding edge applications.

Distributions

Join this track to learn about the making of the openSUSE Distribution and welcome key people of Fedora and Debian to learn from them. New ways of distribution creation with SUSE Studio spice this track up.

Server

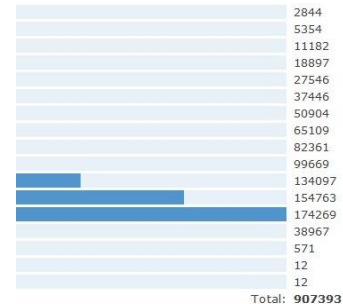
Servers are big machines in responsible positions and openSUSE is your partner for this difficult job. Here we discuss how different worlds get connected with Samba and the Invis Server as well as Management and Monitoring topics.

Community

Community is key! Successful communities have far more pillars than tools and technology: People, Communication and Spirit make a great community out of a good one.

Upstream Technology

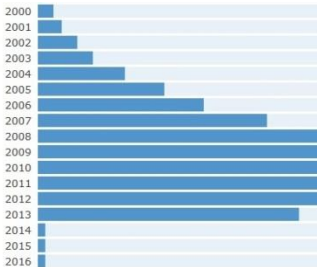
All starts upstream, a successful distribution needs to know and honor that. We are happy welcome some upstream projects sharing their ideas and plans with us.



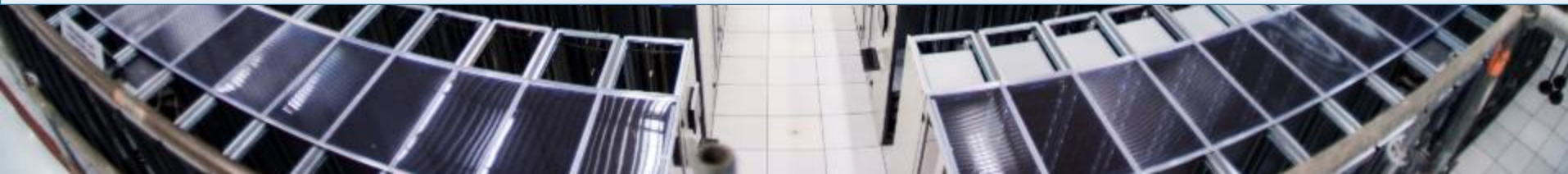
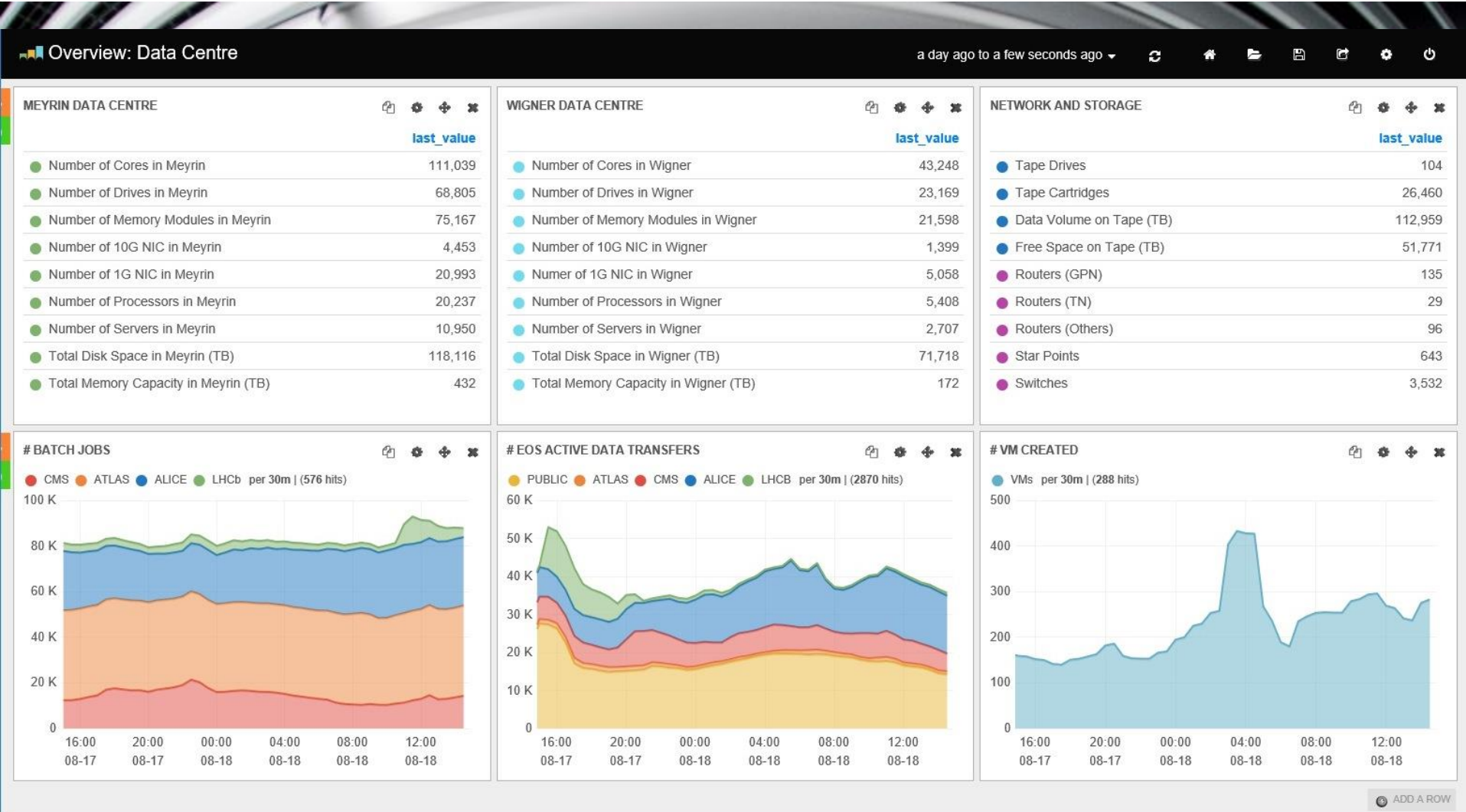
Indico
the
con
mee

F

Indico



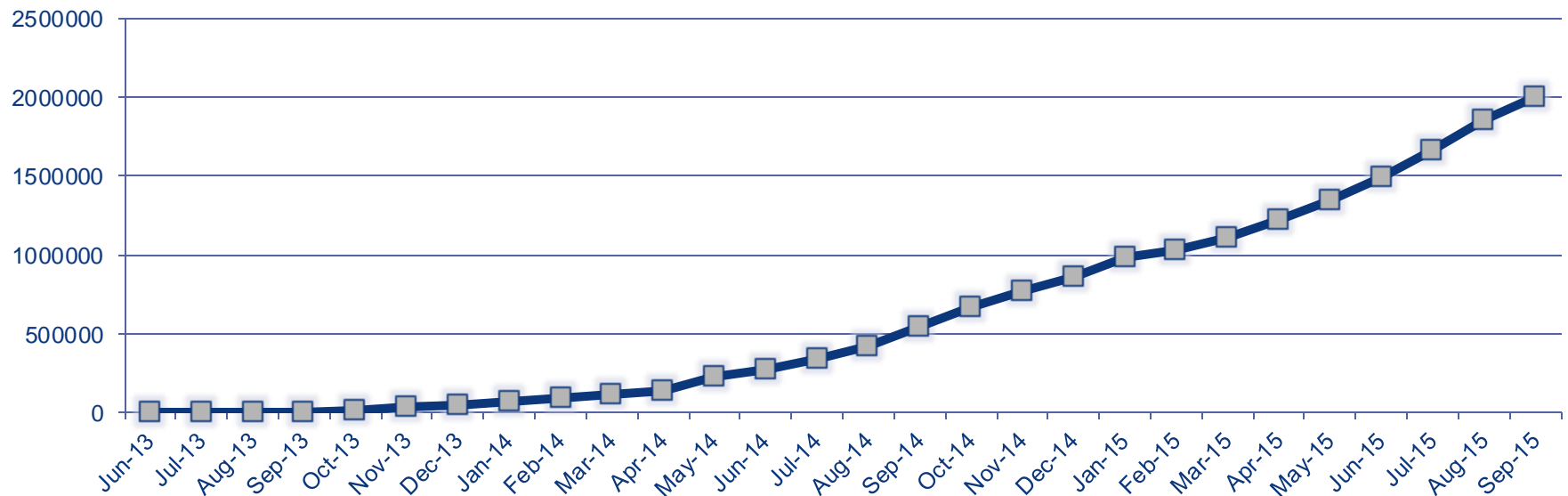
The CERN Data Centre in GVA



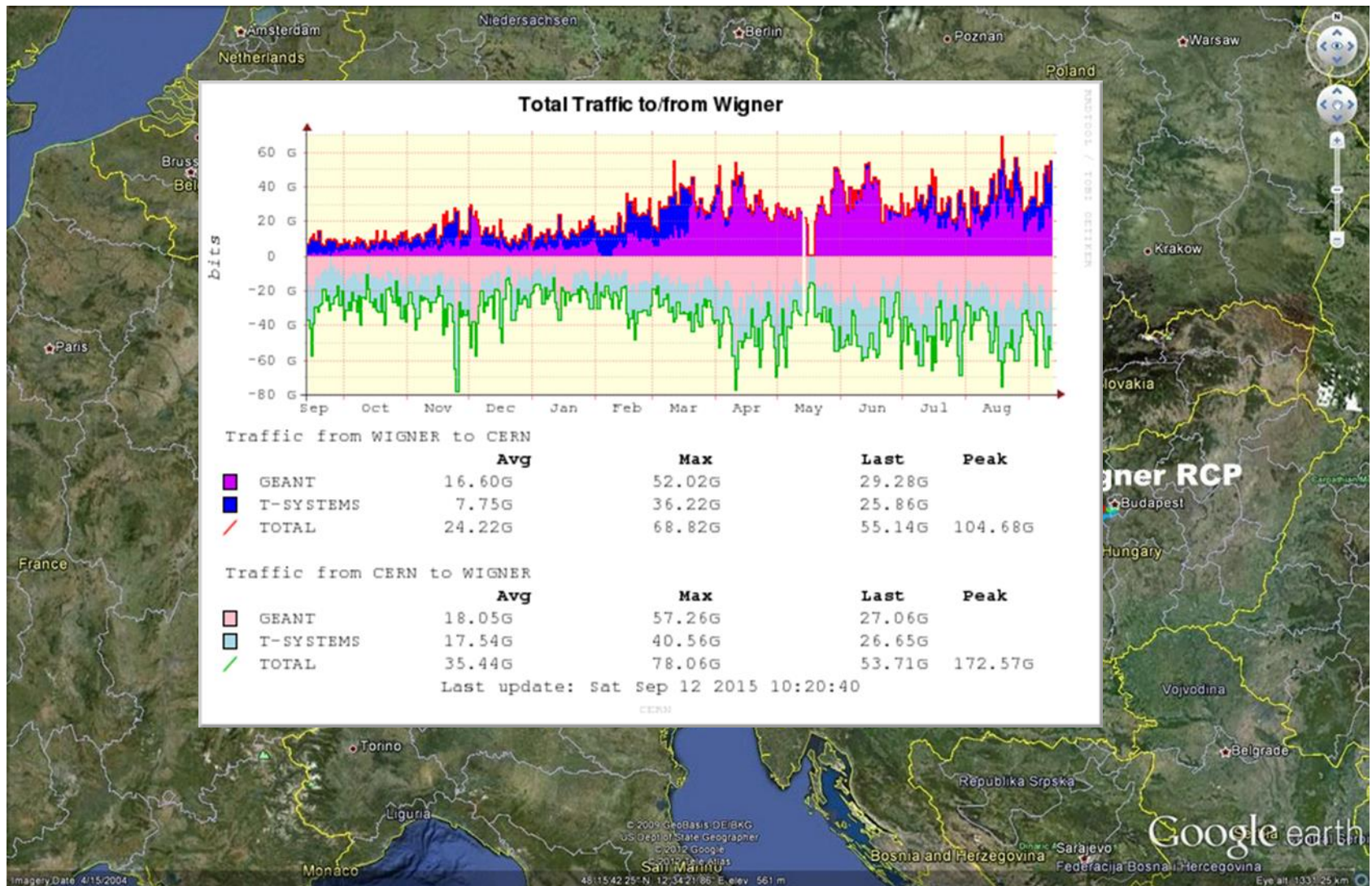
Evolution of the CERN private cloud service

- The graph below shows the evolution in Virtual Machines in the last year

Number of VMs created (cumulative)



Linking the Data Centers



Collaboration - Education

- CERN openlab
 - Intel, Huawei, Oracle, Rackspace, Siemens, Yandex

<http://cern.ch/openlab>

- CERN School of Computing

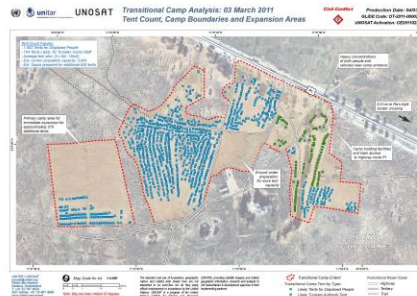
<http://cern.ch/csc>

- UNOSAT

<http://cern.ch/unosat>

- Citizen Cyber Science Collaboration

- Involving the General Public



Helping developing countries

Invenio is an open source software suite enabling you to run your own digital library or document repository on the web. The technology offered by the software covers all aspects of digital library management from document ingestion through classification, indexing, and curation to dissemination.

Invenio is the tool used for LHC publications workflow approval at CERN.



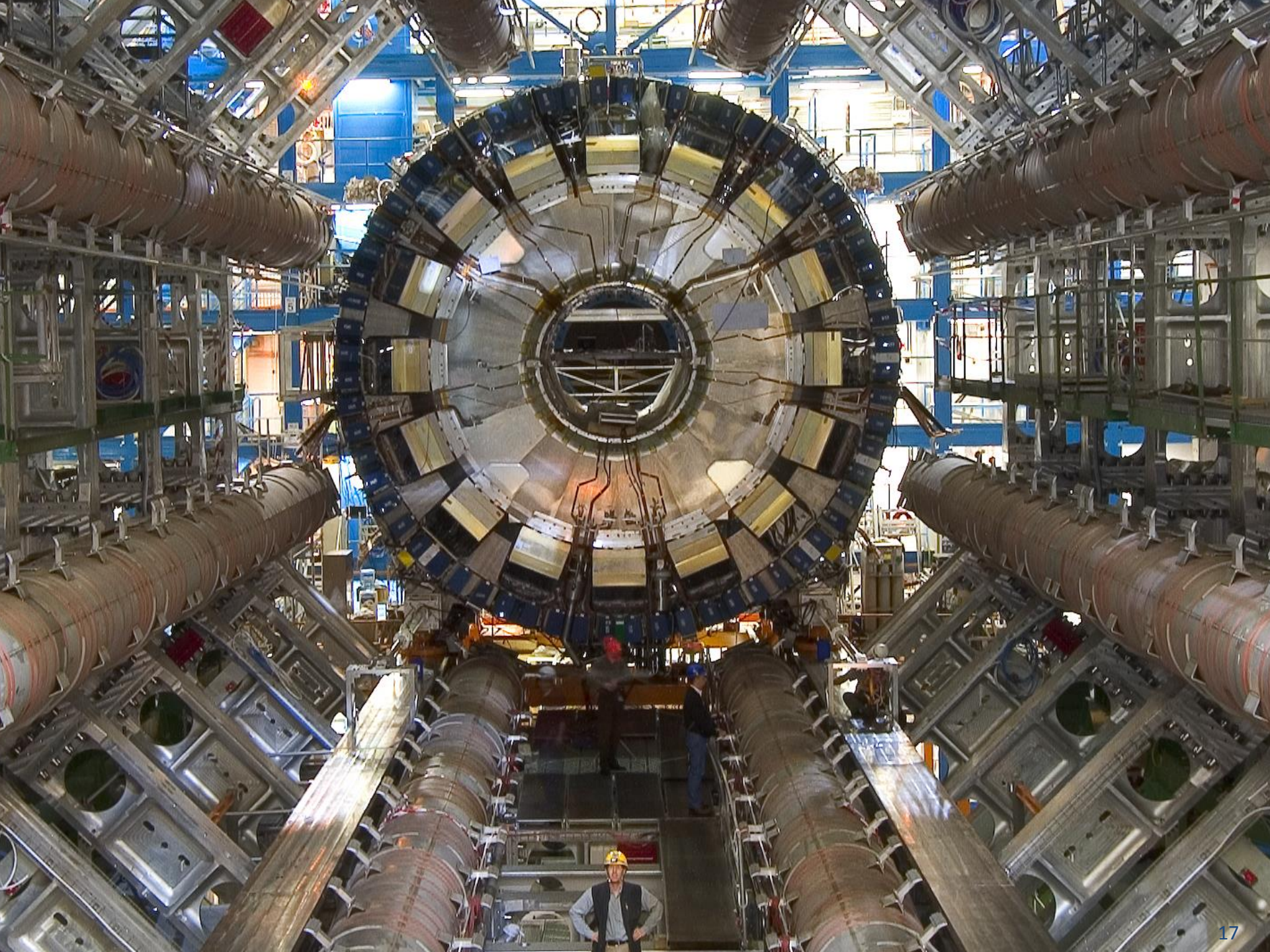
UNESCO & CERN have sponsored digital library schools in Africa (Rwanda 2009, Morocco 2010, Senegal 2011)

CERN has donated computers to Africa (Morocco, Ghana, Egypt, Senegal) as well as Philippines & Pakistan to help capacity building.



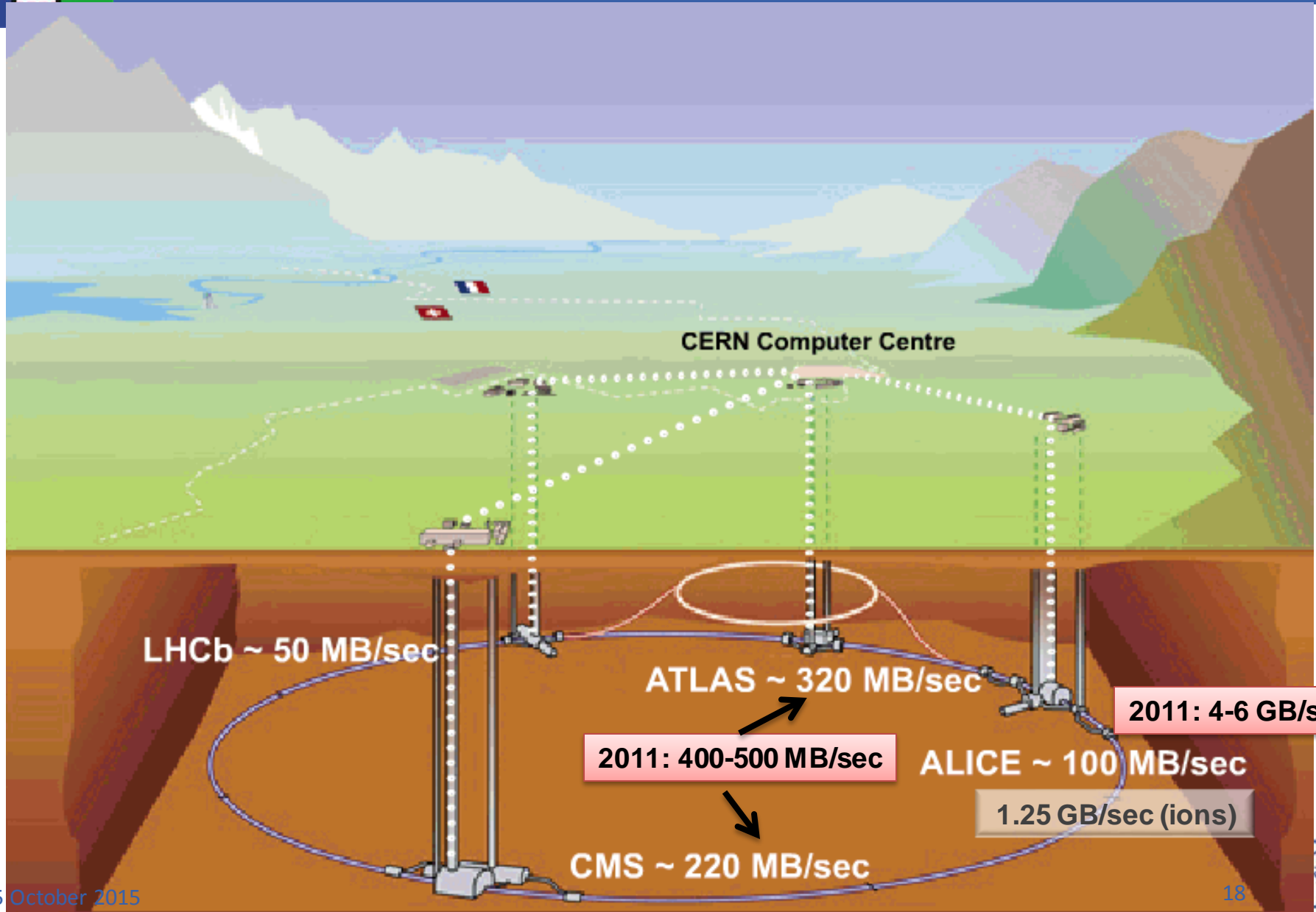
The Worldwide LHC Computing Grid





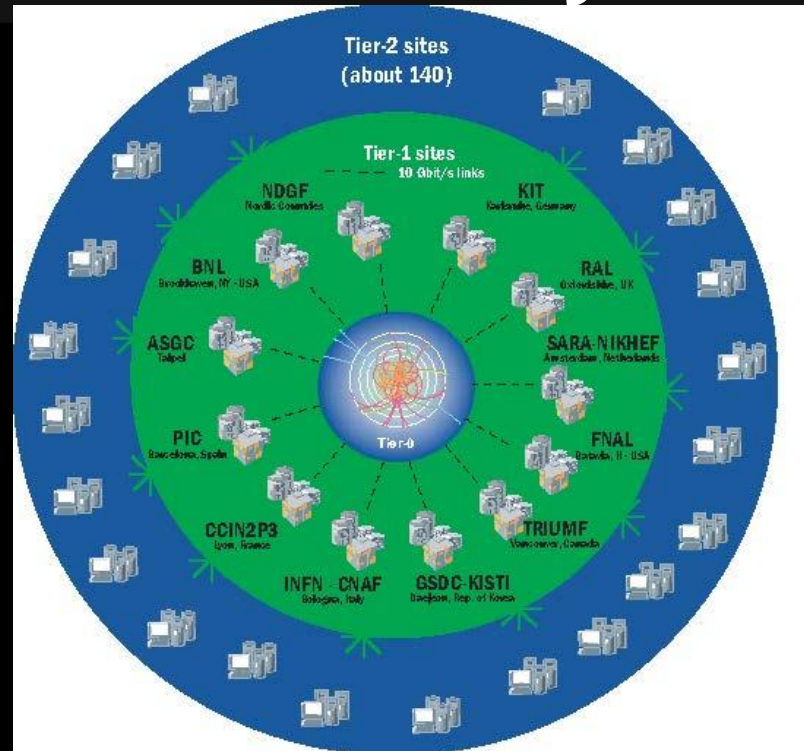


Tier 0 at CERN: Acquisition, First pass reconstruction, Storage & Distribution



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):

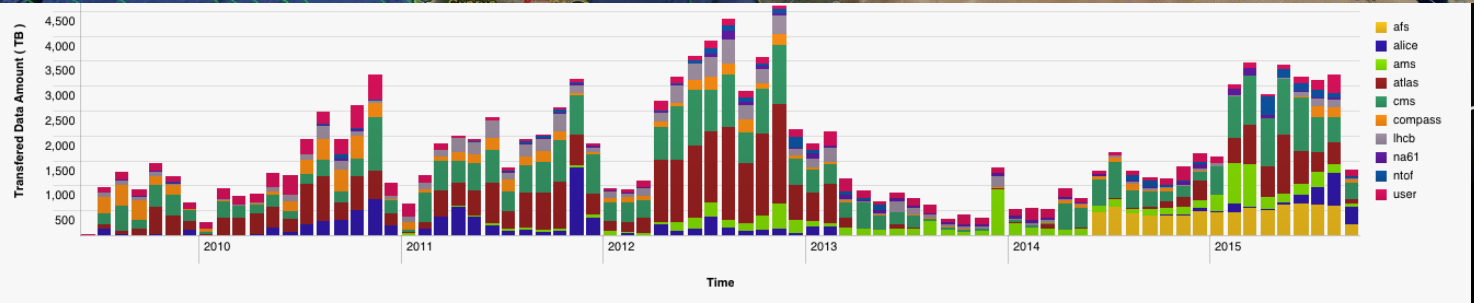
- Data recording
- Initial data reconstruction
- Data distribution

Tier-1 (12 centres + Russia):

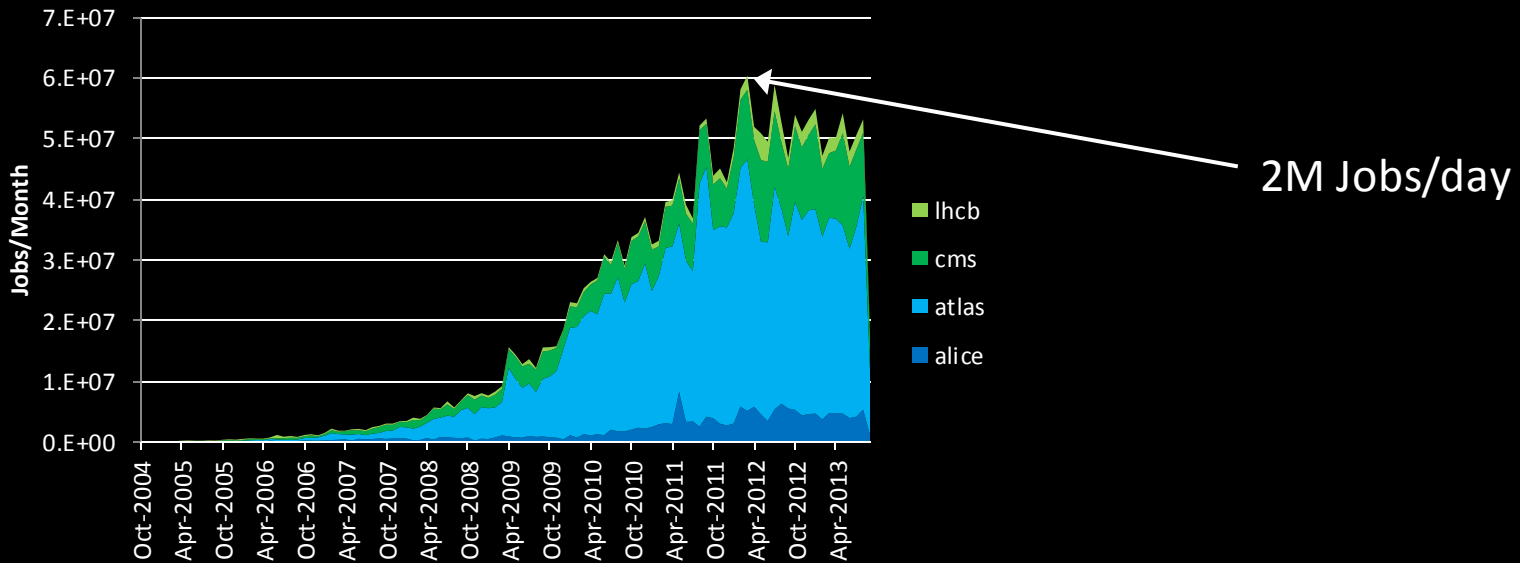
- Permanent storage
- Re-processing
- Analysis

Tier-2 (~140 centres):

- Simulation
- End-user analysis
- ~ 160 sites, 35 countries
- 300000 cores
- 200 PB of storage
- 2 Million jobs/day
- 10 Gbps links

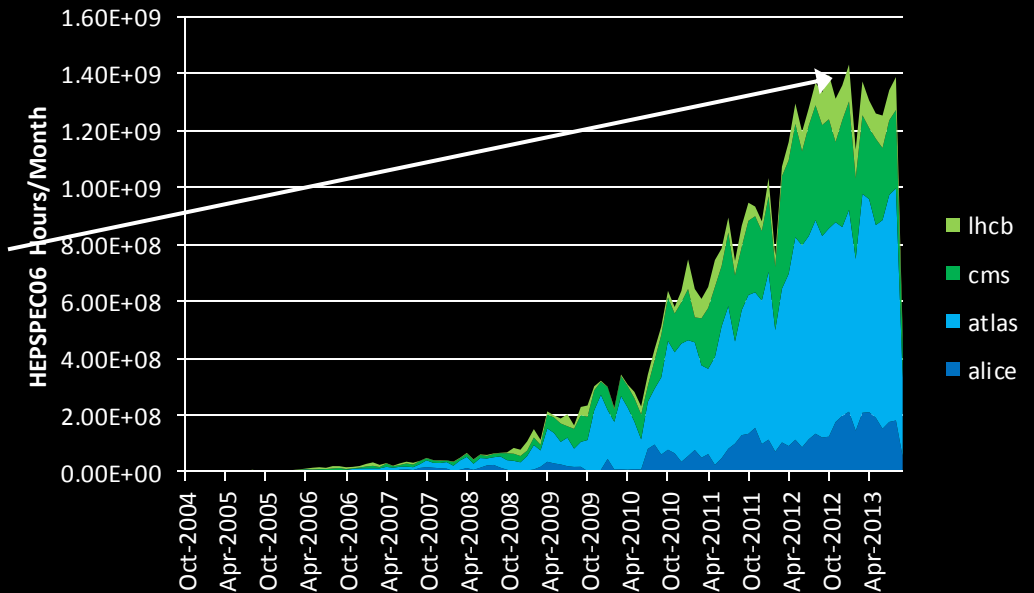


Processing on the Grid



1.4 10^9 HEPSPROC06/Month
(210 K CPU continuous use)

Close to full capacity



CERN openlab in a nutshell

- A science – industry partnership to drive R&D and innovation with over a decade of success
- Evaluate state-of-the-art technologies in a challenging environment and improve them
- Test in a research environment today what will be used in many business sectors tomorrow
- Train next generation of engineers/employees
- Disseminate results and outreach to new audiences

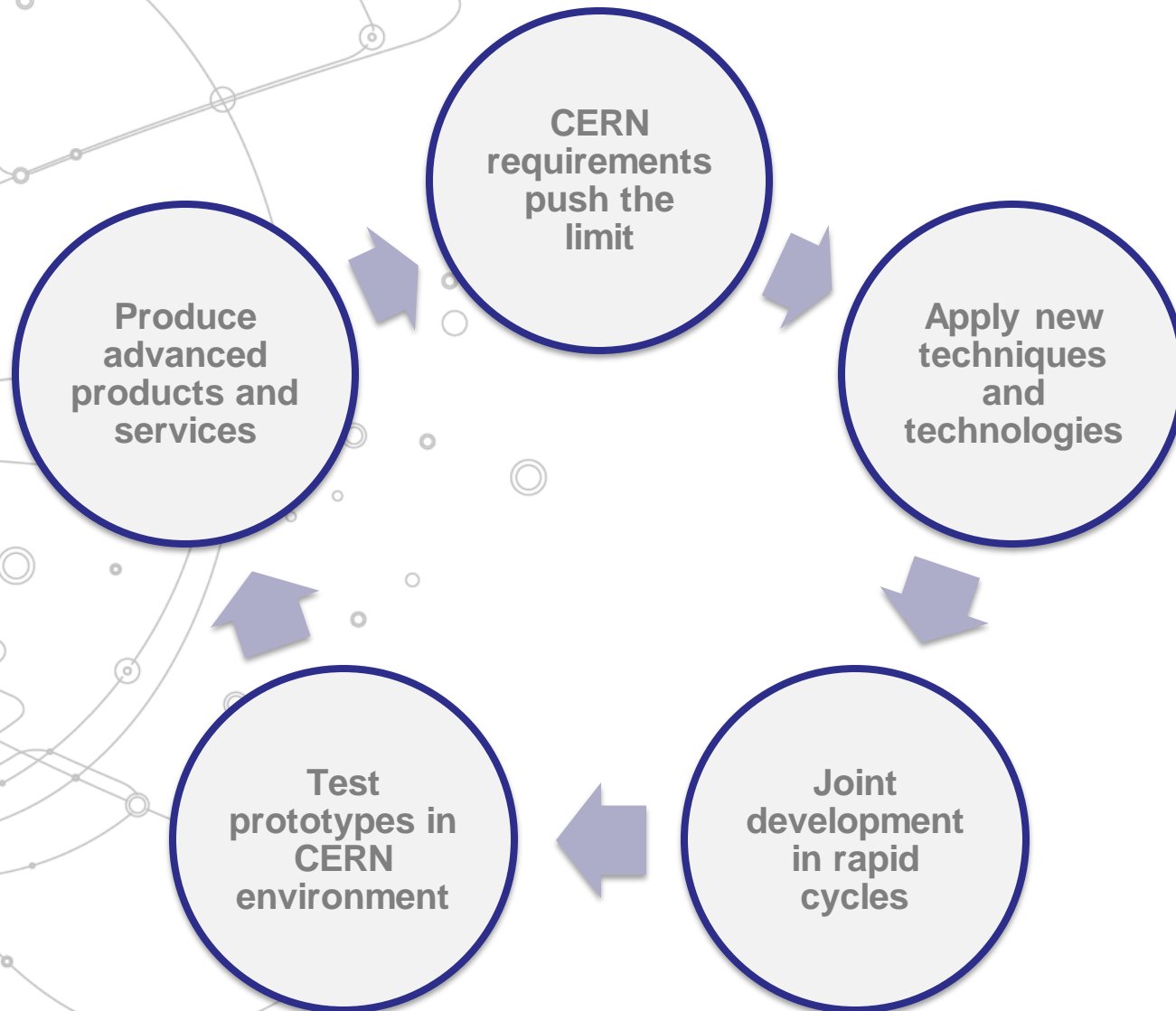


ORACLE

SIEMENS



Virtuous Cycle



A public-private partnership between the research community and industry

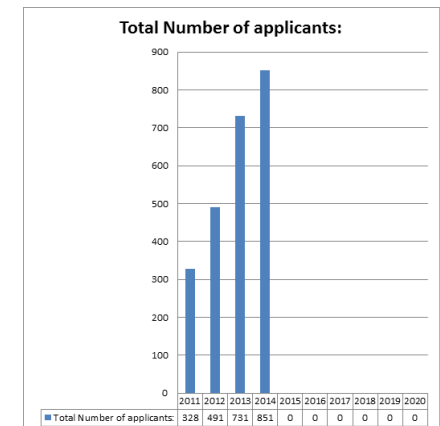
› Summer student program 2013

- 720+ applicants
- 22 selected candidates
- 13 lectures (including new lectures from external labs)
- A new lightning talks session
- 22 technical reports



› Summer student program 2014

- 850+ applicants
- 23 selected candidates
- Lectures and visits program being designed, will include contributions from external labs and companies



Challenges Ahead

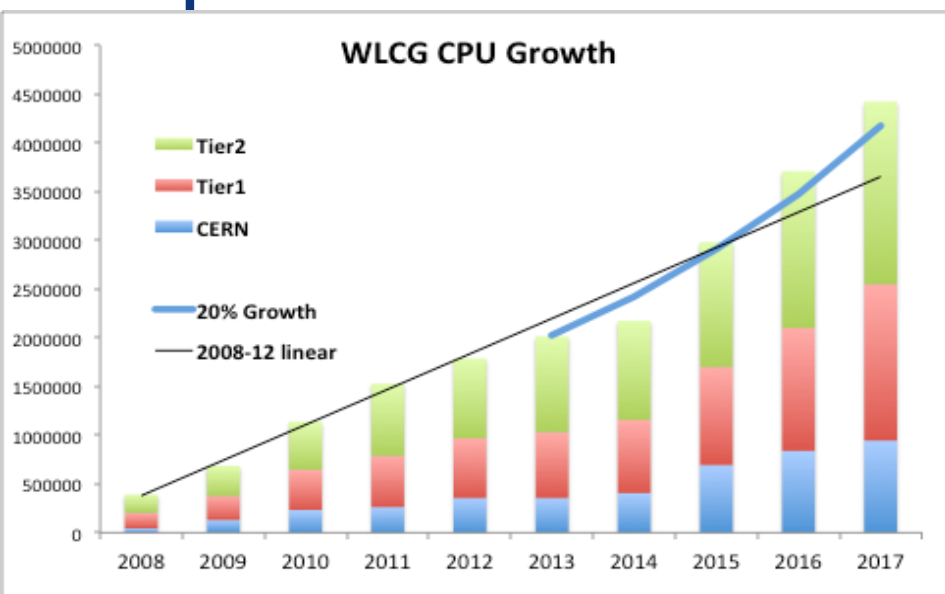
Evolution of requirements

Higher trigger (data) rates driven by physics needs

Based on understanding of likely LHC parameters;

Foreseen technology evolution (CPU, disk, tape)

Experiments work hard to fit within constant budget scenario



Estimated evolution of requirements 2015-2017
(NB. Does not reflect outcome of current RSG scrutiny)

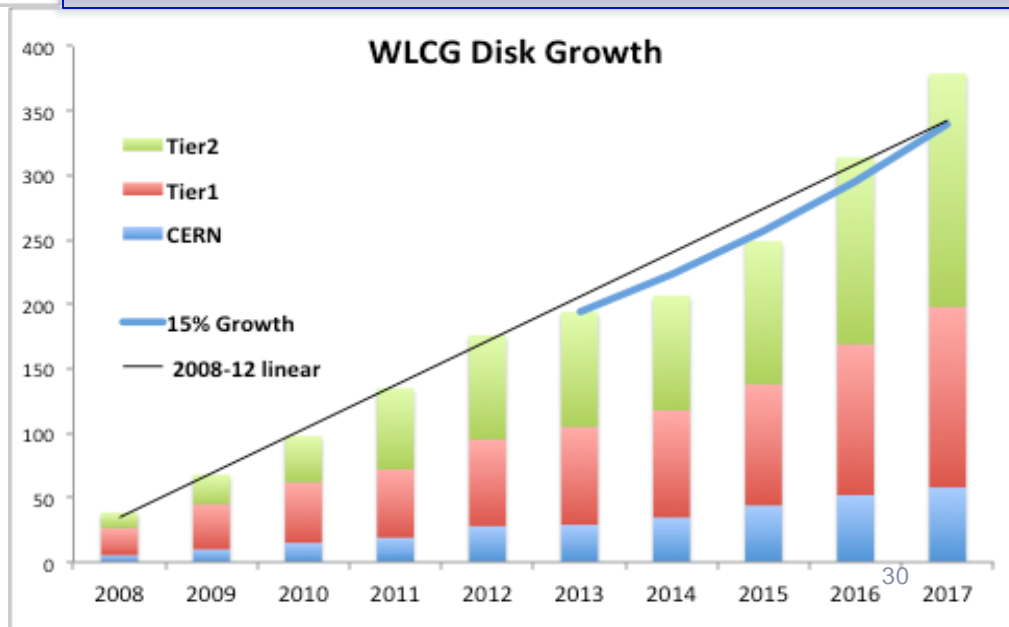
2008-2013: Actual deployed capacity

Line: extrapolation of 2008-2012 actual resources

Curves: expected potential growth of technology with a constant budget (see next)

CPU: 20% yearly growth

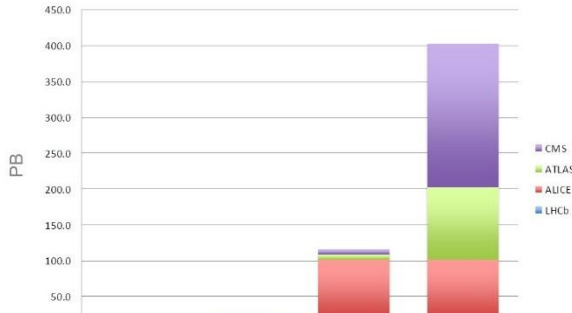
Disk: 15% yearly growth



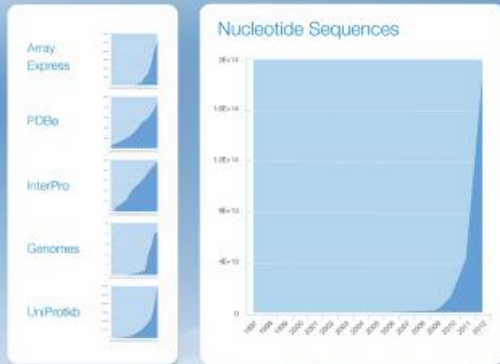
Anticipated Data Needs



Data: Outlook for HL-LHC



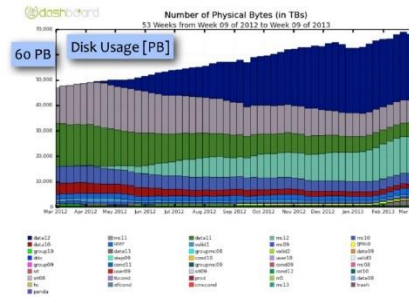
Bigger and bigger data



EMBL-EBI

Data Management: Energy Frontier

Atlas Data on Disk, across 11 Tier-I Centers



Current LHC data sizes

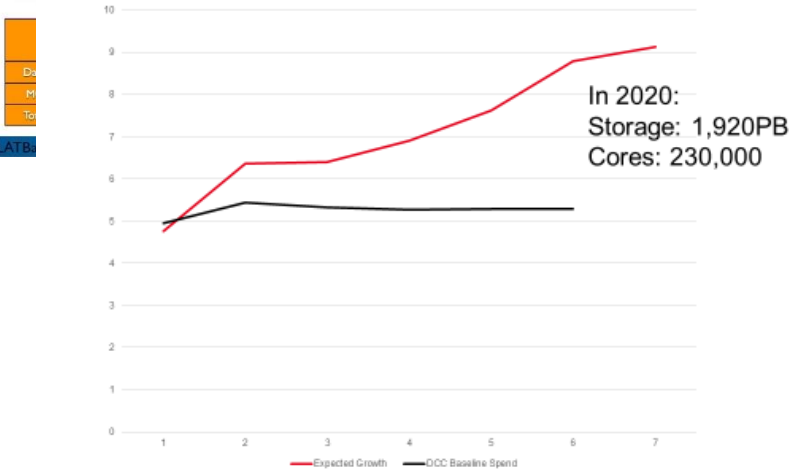
- ★ examples
 - ◆ Atlas: 70 PB on disk, world-wide
 - ◆ CMS 18.2 PB on tape, at Fermilab

Future increases

- ★ an estimate for 2021
 - ◆ ~130 PBytes detector data
 - ◆ ~350 PBytes simulated data
 - ◆ ~270 PBytes US "data library"

CI

Projected Hardware Requirements and Cost (€M)



EMBL-EBI



Further Reading...

Future IT Challenges in Scientific Research

Compute Management & Provisioning

- Data Acquisition
- Computing Platforms
- Data Storage Architectures
- Compute Management and Provisioning
- Networks and Connectivity
- Data Analytics



<http://zenodo.org/record/8765>

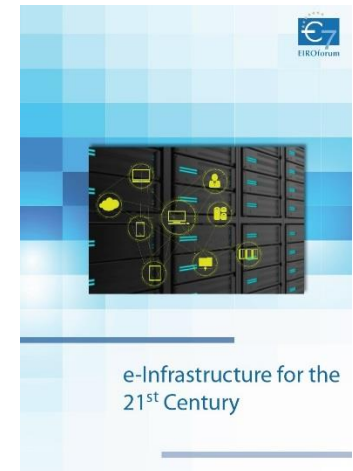
Update of the Computing Models of the WLCG and the LHC Experiments

<http://cds.cern.ch/record/1695401>



E-Infrastructure for the 21st century

<http://zenodo.org/record/7592>



Summary

- Opportunities to work with the IT department
 - Supplying computing equipment and/or capacity
 - Proving services based on IT technologies
 - R&D with the CERN openlab
 - Hosting events such as the CSC
 - ...

Thank you!





www.cern.ch