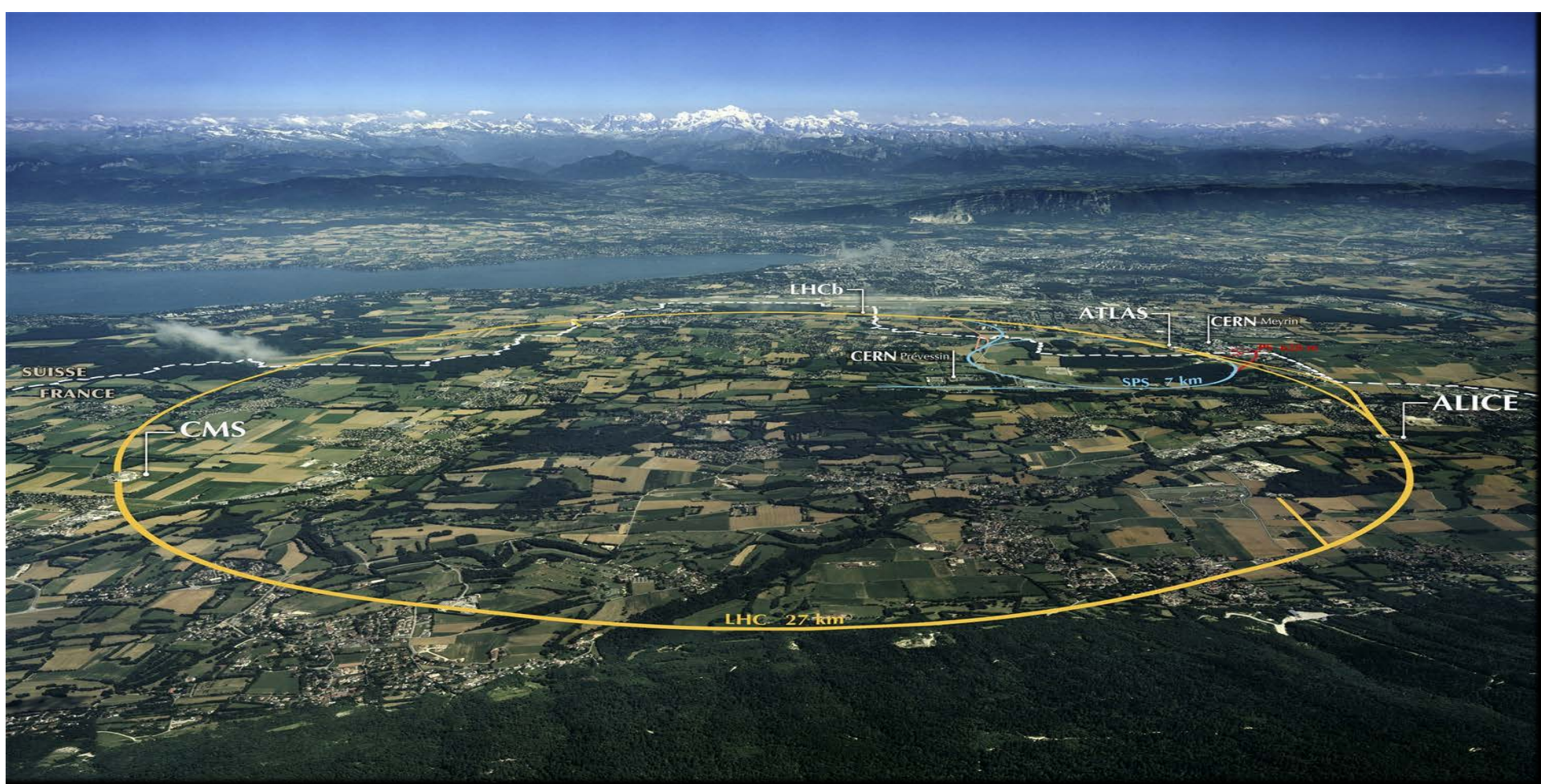




Frontiers of Science and Clouds



Tim Bell – tim.bell@cern.ch - [@noggin143](https://twitter.com/noggin143)

About Tim



CERN: founded in 1954: 12 European States

“Science for Peace”

Today: 21 Member States

~ 2300 staff

~ 1400 other paid personnel

~ 12500 scientific users

Budget (2016) ~1000 MCHF



Member States: Austria, Belgium, Bulgaria, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland and United Kingdom

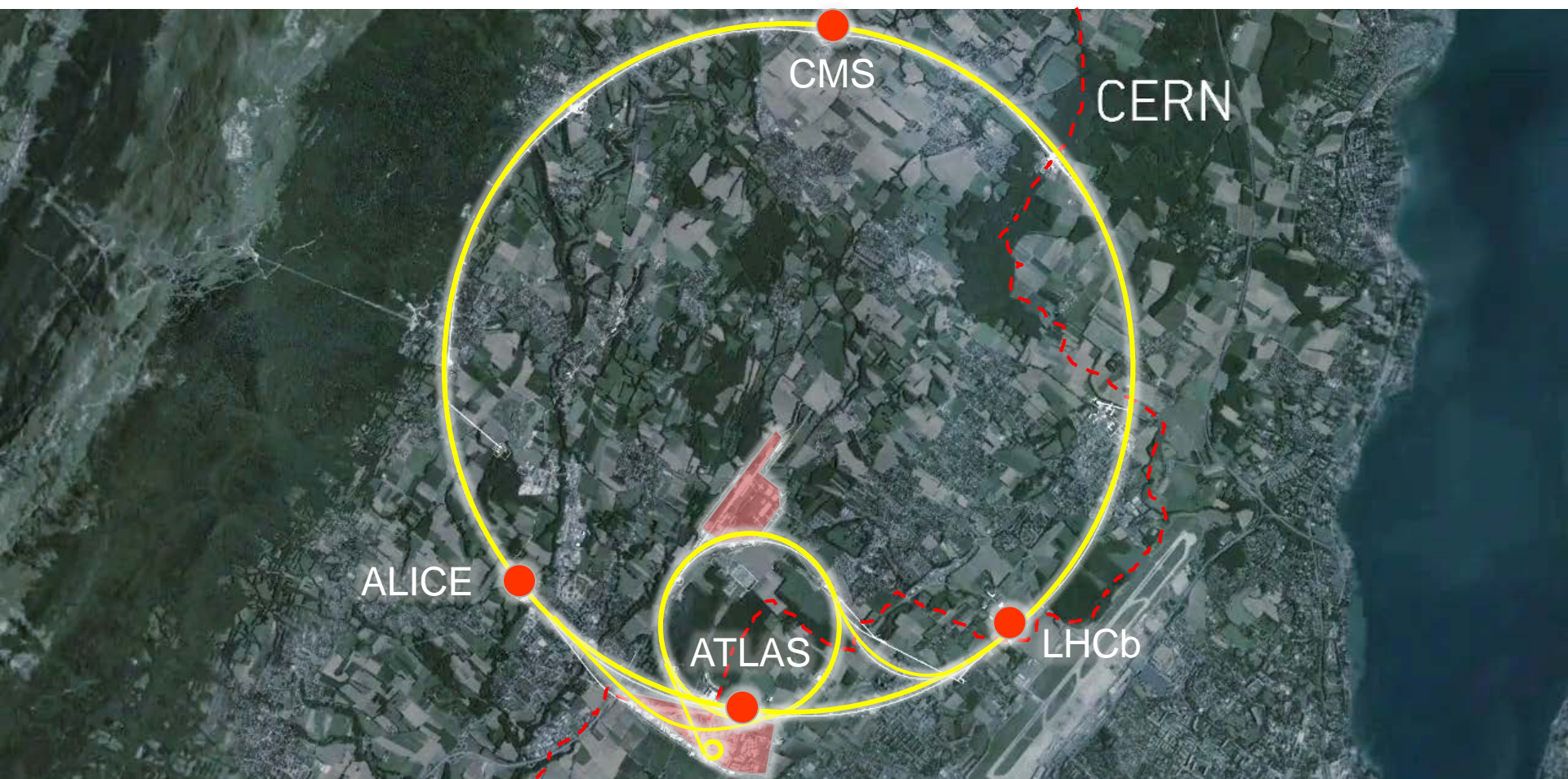
Associate Member States: Pakistan, Turkey

States in accession to Membership: Romania, Serbia

Applications for Membership or Associate Membership: Azerbaijan, Brazil, Croatia, Cyprus, India, Russia, Slovenia, Ukraine

Observers to Council: India, Japan, Russia, United States of America; European Union, CERN and UNESCO

Largest Machine on Earth



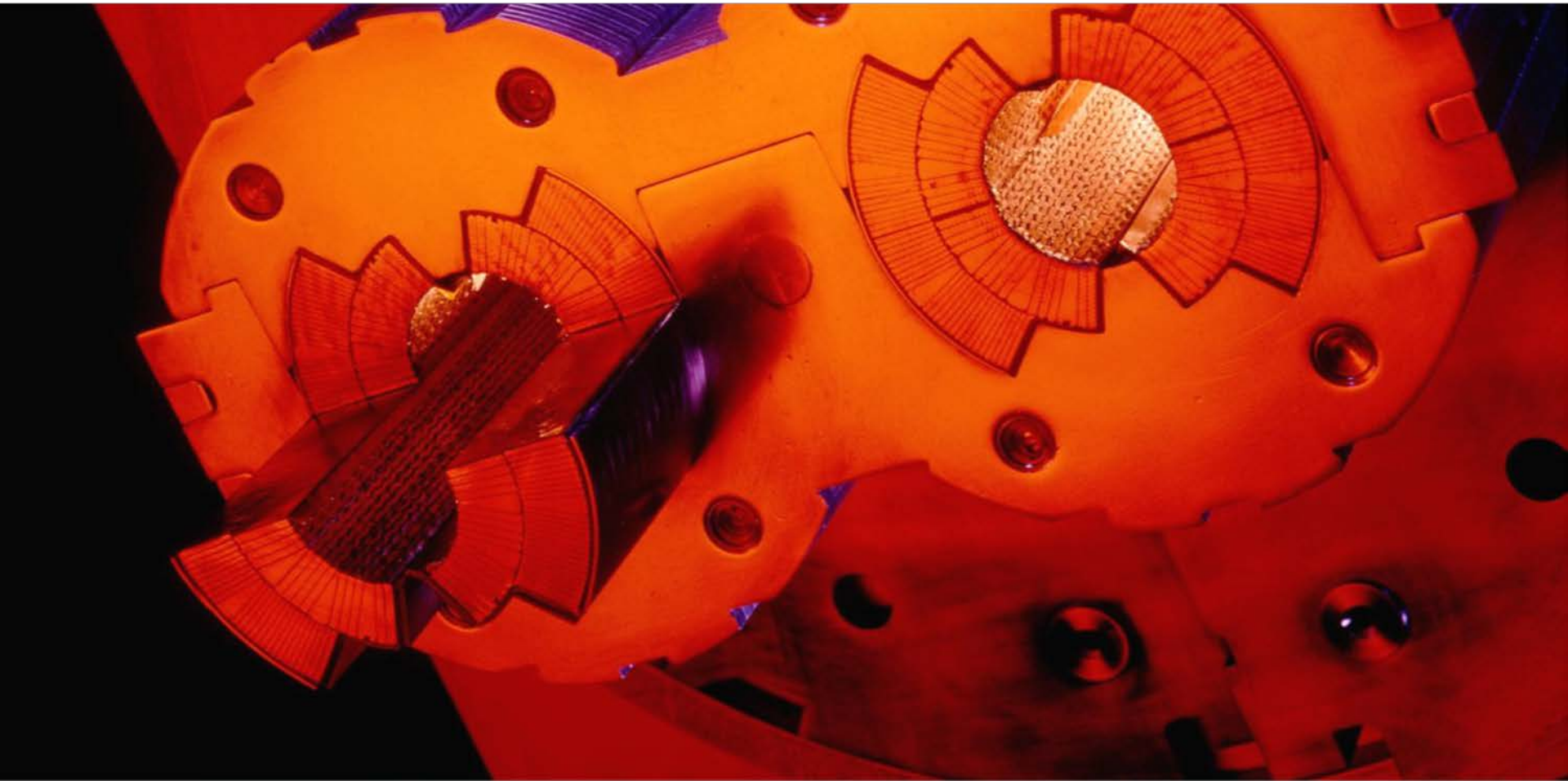
27 Km circumference, 100m underground

Most Powerful Magnets



9,600 superconducting magnets direct the beam

Highest Vacuum



Pressure is 10 times less than on the moon

Coldest Temperature



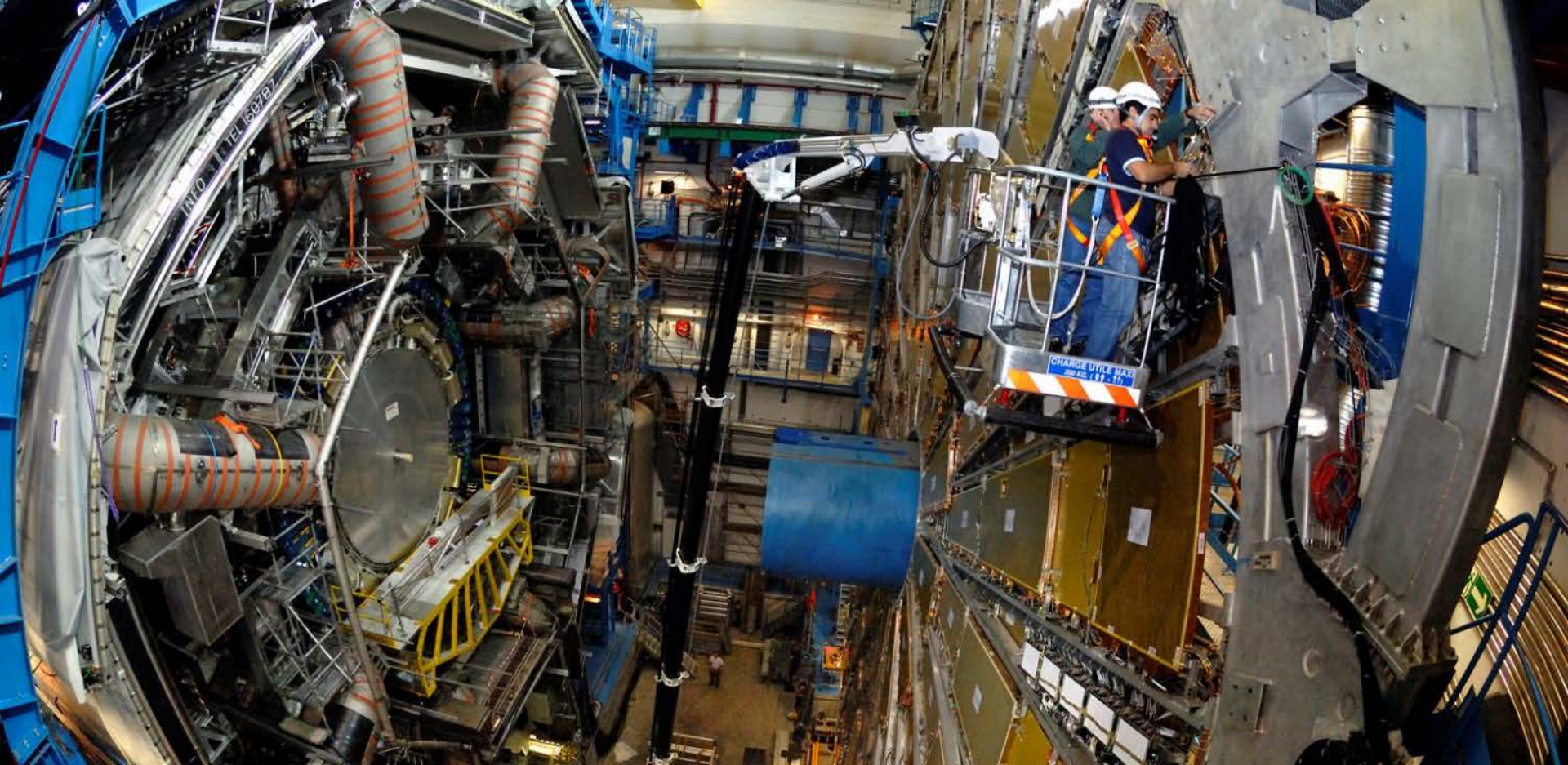
120 tonnes of liquid helium for cooling to -271°C

30M USD/year Electricity Bill

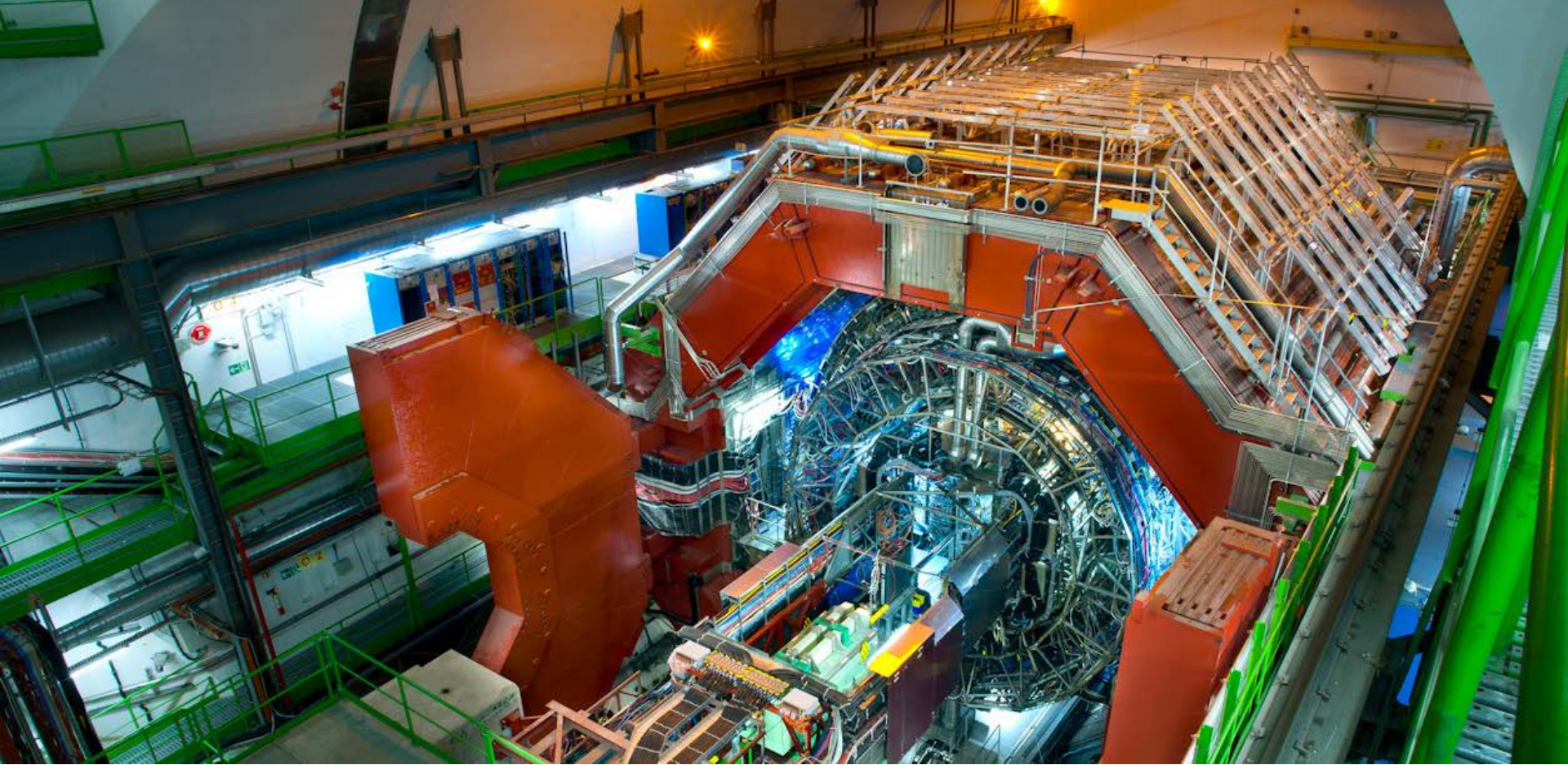


Up to 200MW at peak utilisation

ATLAS Experiment



ALICE Experiment



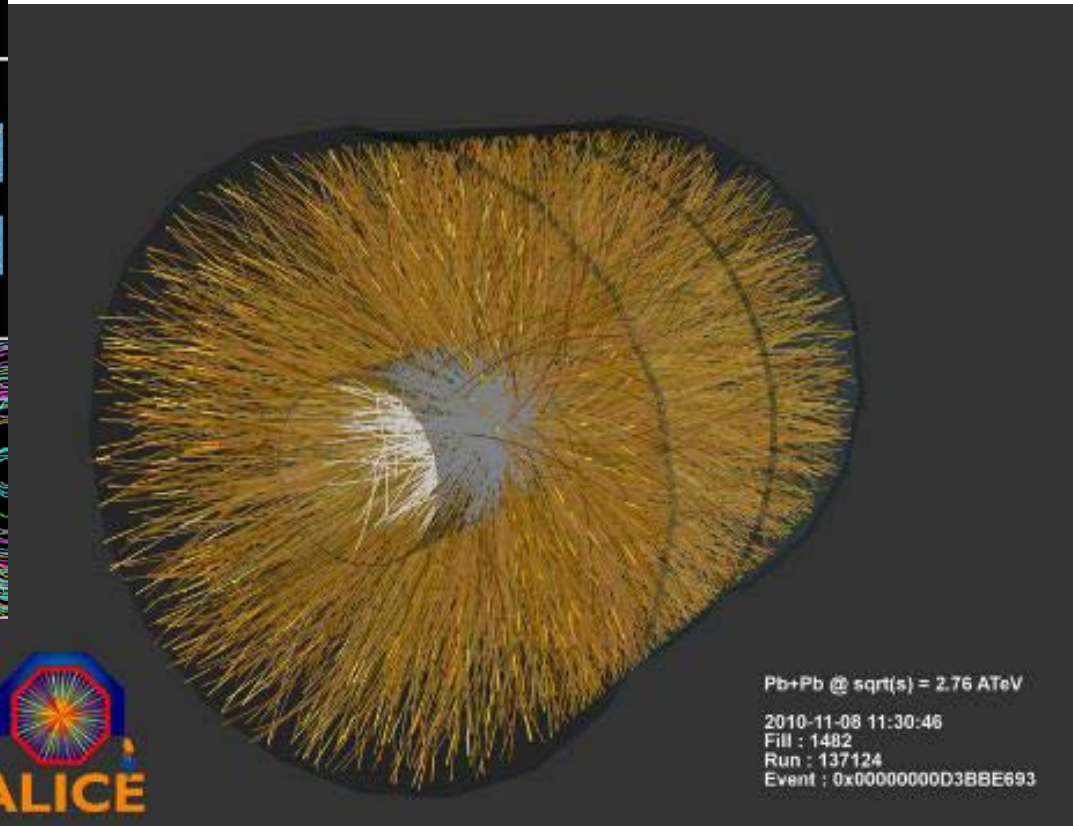
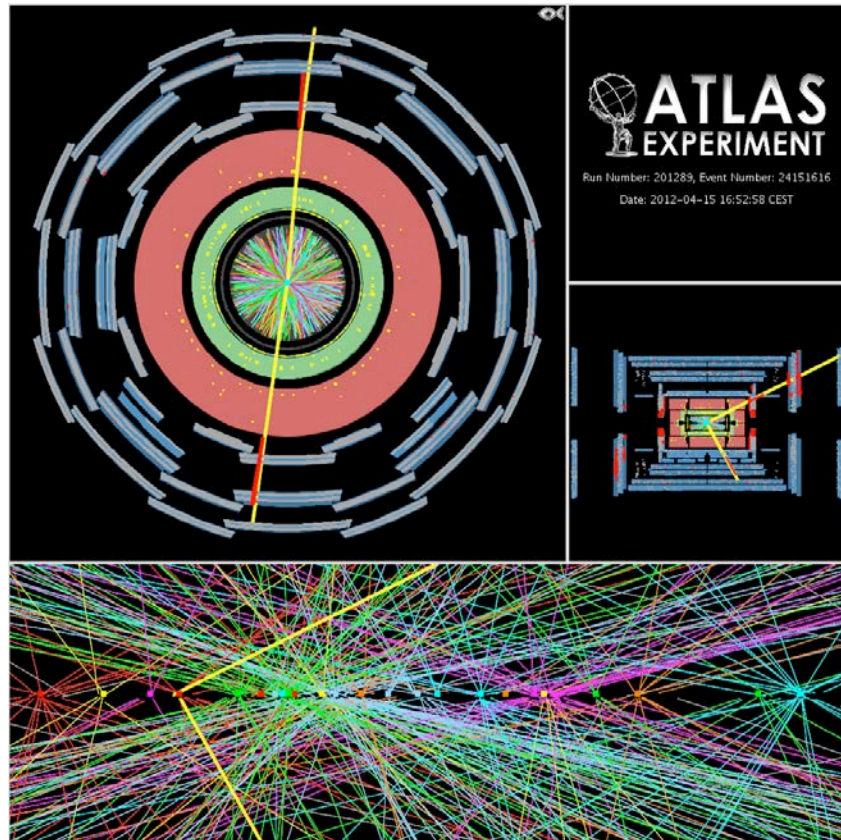
CMS Experiment



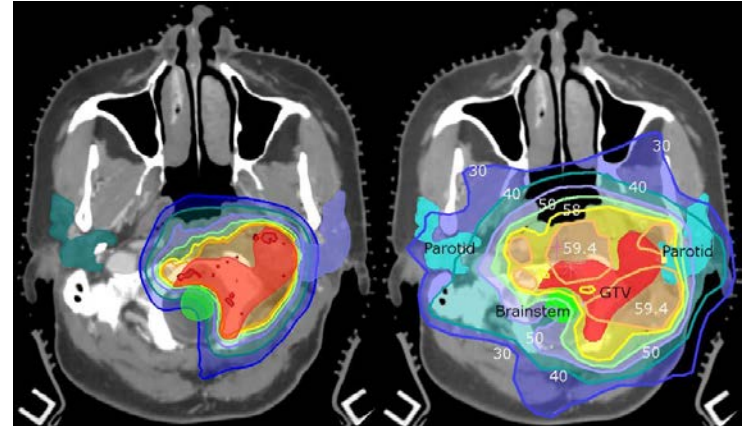
LHCb Experiment



Collisions Produce 1PB/s



Innovations

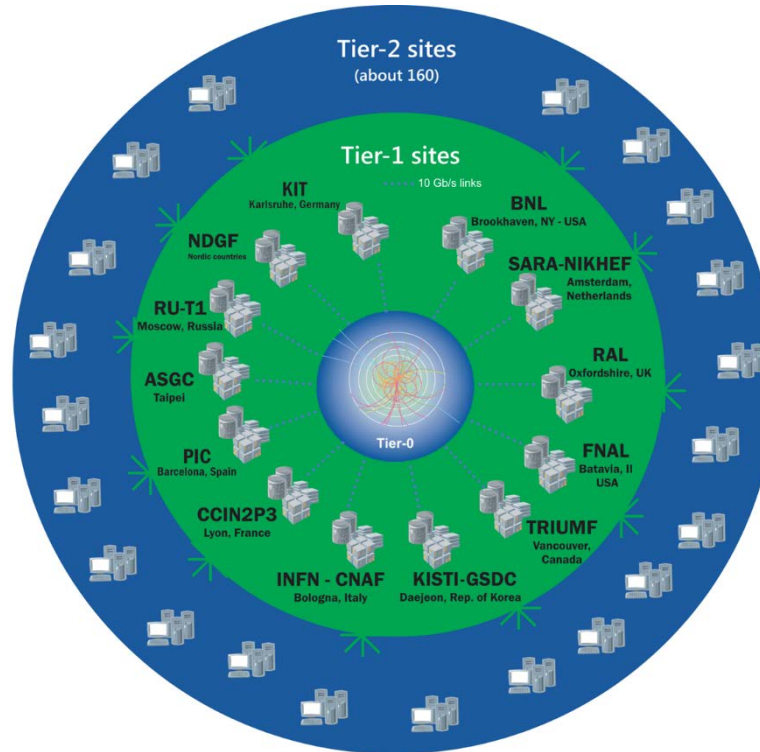


Worldwide LHC Computing Grid

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

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

TIER-2:
Simulation,
end-user analysis



nearly 170 sites,
40 countries

~350'000 cores

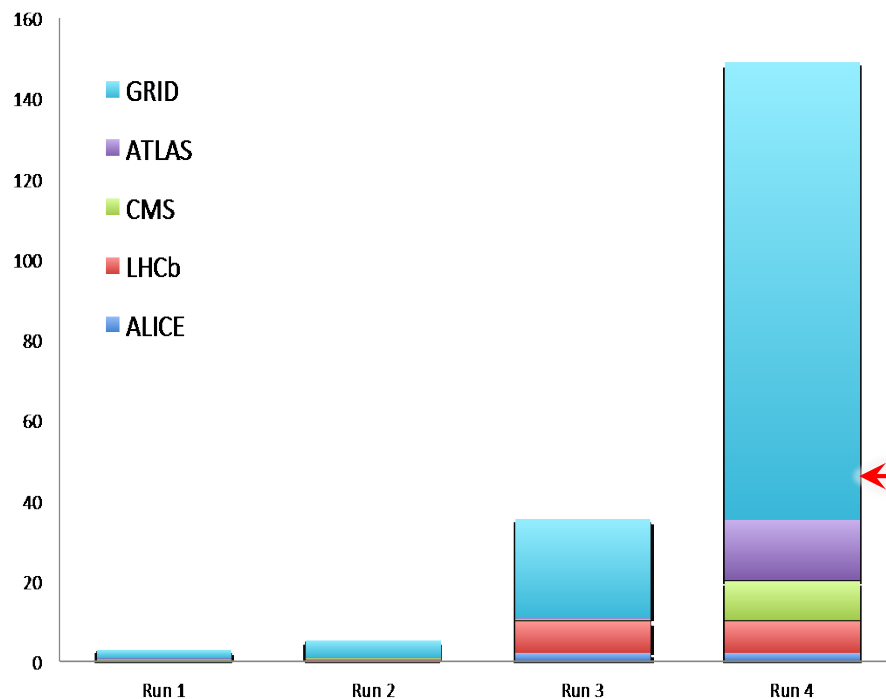
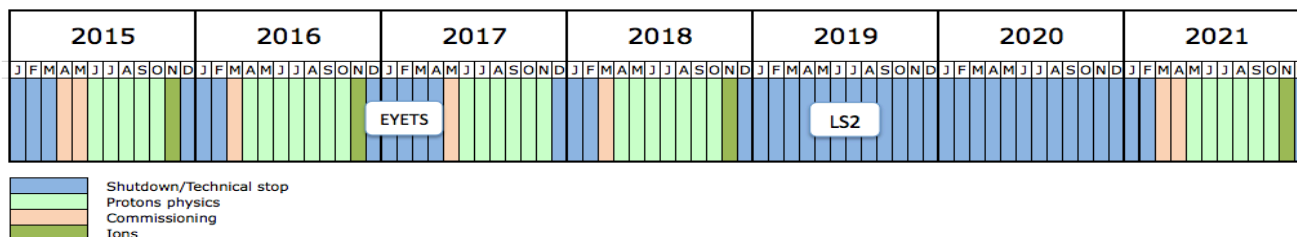
500 PB of storage

> 2 million jobs/day

10-100 Gb links

Compute Growth Outlook

The outline LHC schedule out to 2035 presented by Frederick Bordry to the SPC and FC June 2015 can be found [here](#)



Compute: Growth > x50
Moore's law only x16

← What we can afford

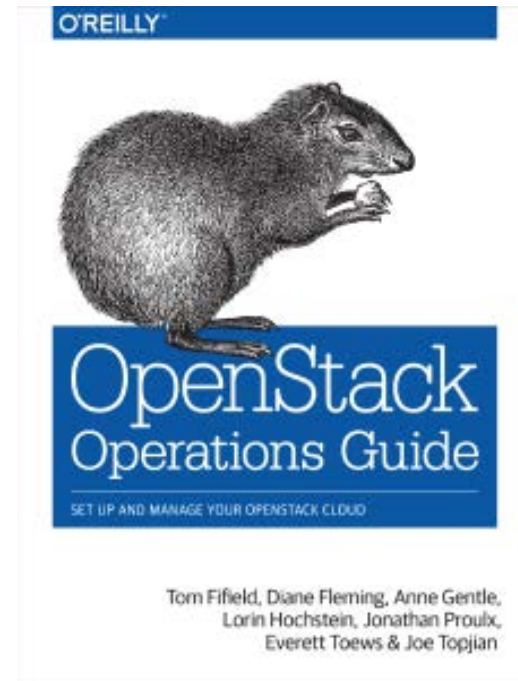
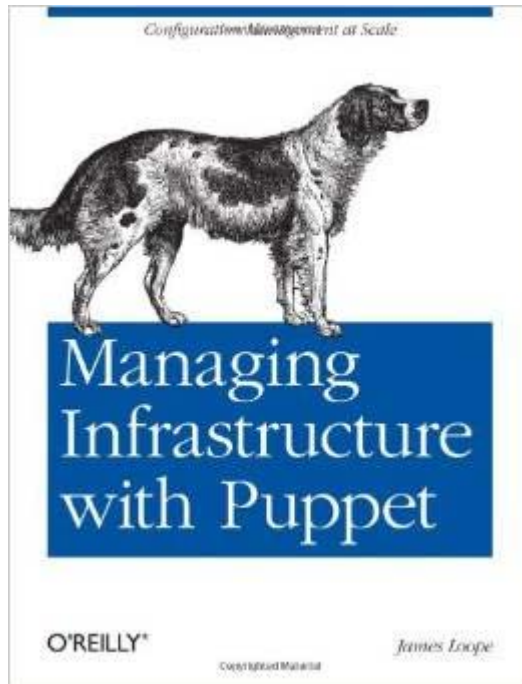
... and 400PB/year by 2023



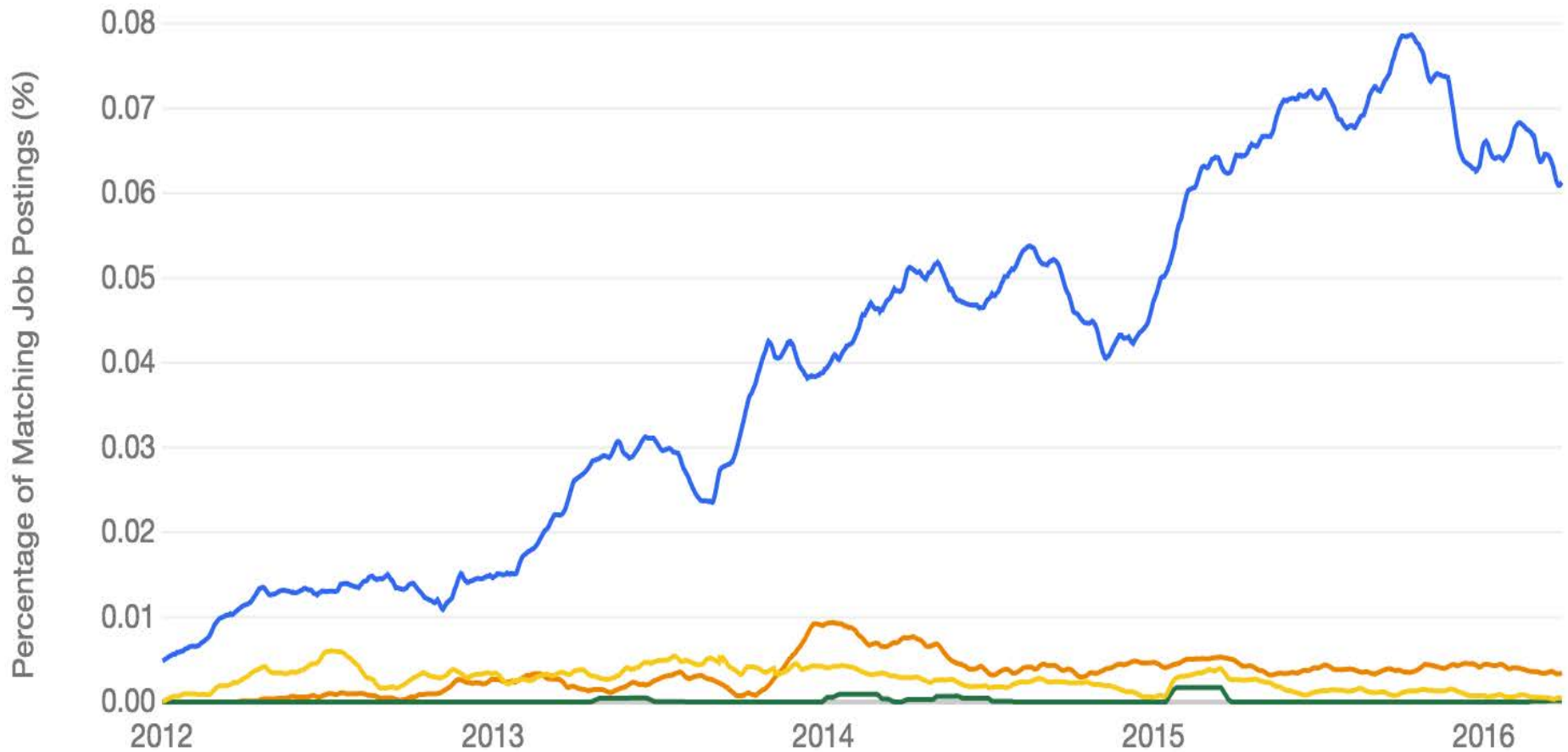
Expanded Facilities in Hungary



O'Reilly Consideration



Job Trends Consideration



From Indeed.com

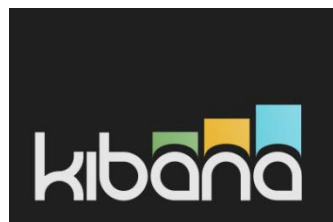
OpenStack x

CloudStack x

OpenNebula x

Eucalyptus x

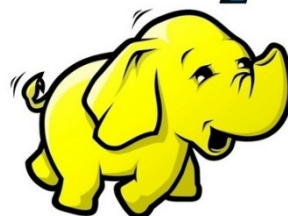
CERN Tool Chain



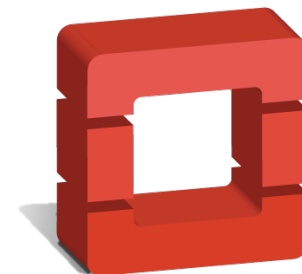
FOREMAN



hadoop



 RUNDECK

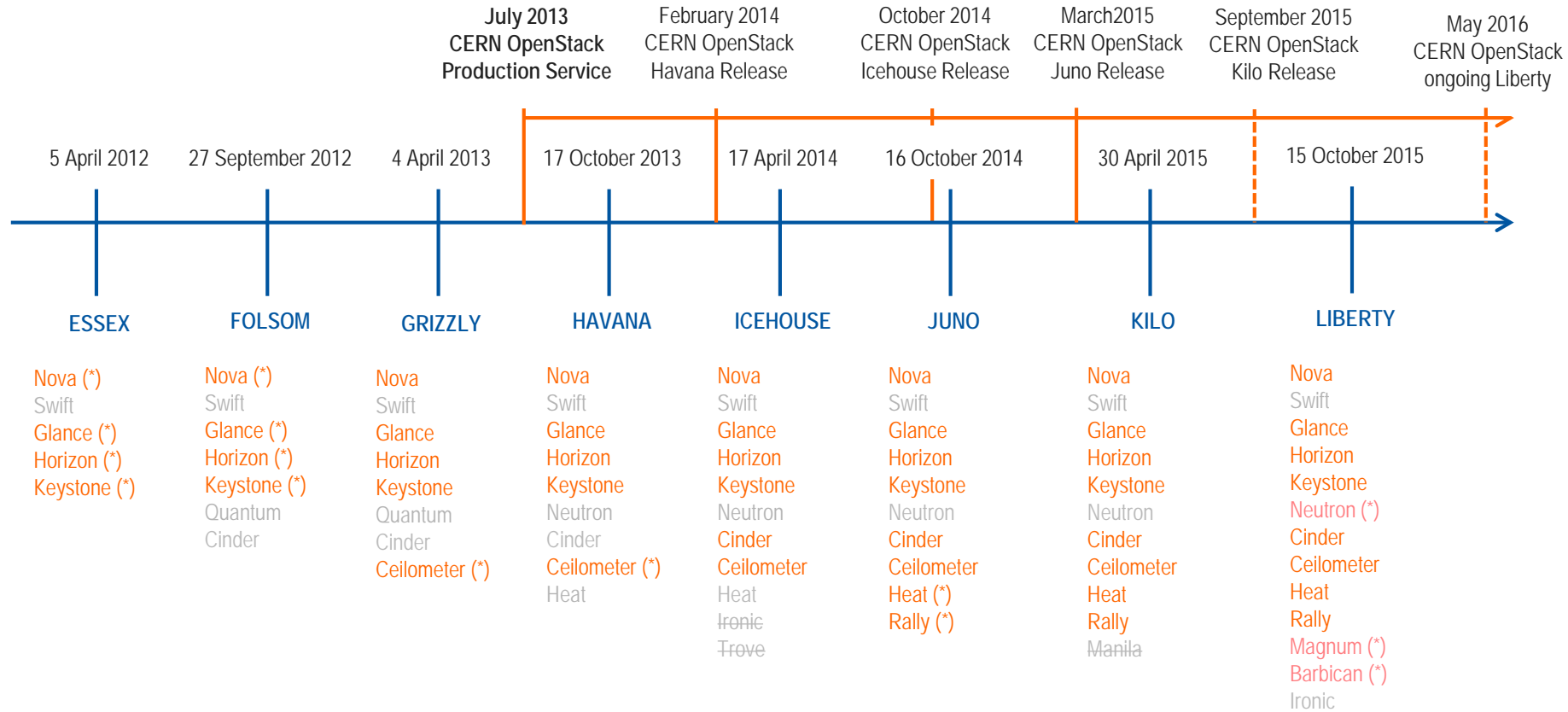


openstack™
CLOUD SOFTWARE



Jenkins

CERN OpenStack Project



(*) Pilot

Not Just The Software

Upstream OpenStack on its own does not give you a cloud service

e.g.

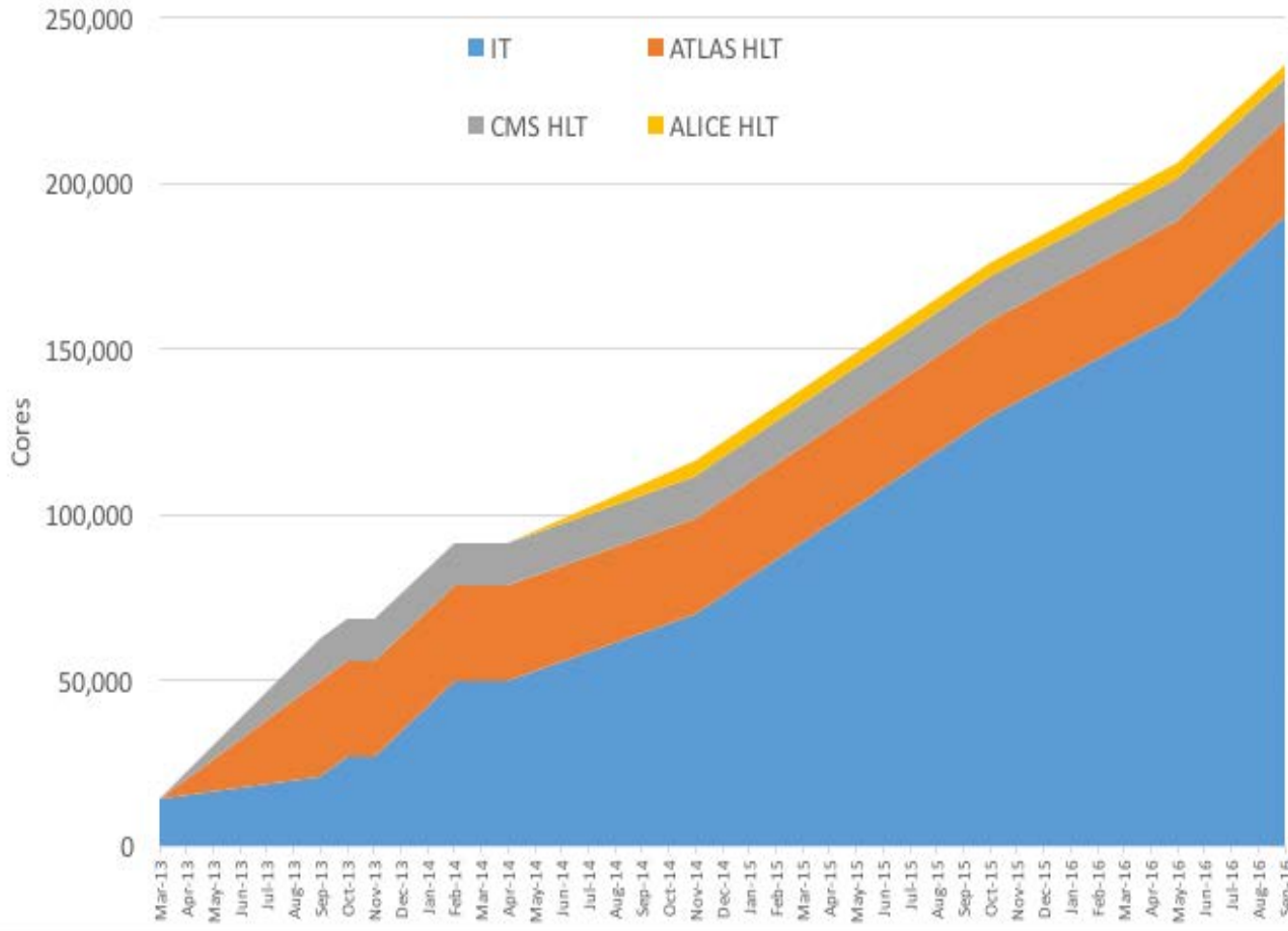
- 200 people arrive and leave CERN / month
- User skill levels vary widely
- Application range from server consolidation, dev/test to production compute



Subbu Allamaraju @ eBay

OpenStack@CERN Status

Total Cores in OpenStack Clouds at CERN



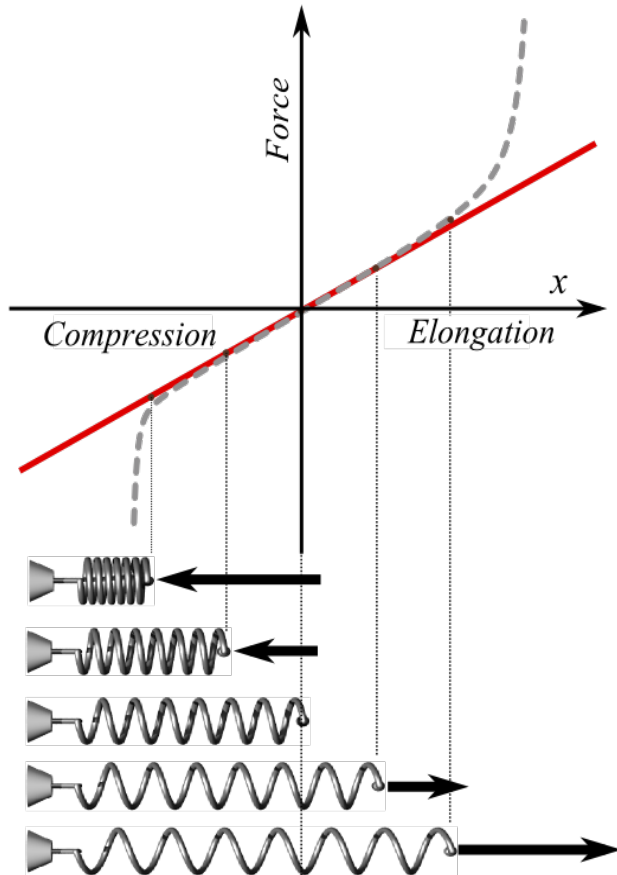
In production:

- 4 clouds
- >200K cores
- >8,000 hypervisors

~60,000 additional cores being installed in Q2 2016

90% of CERN's compute resources are now delivered on top of OpenStack

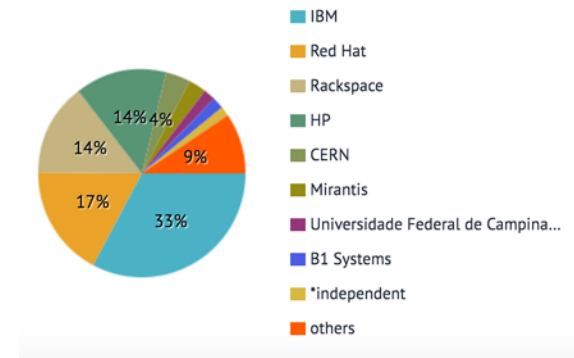
Cultural Change Impact



- Move to Agile technologies brings great benefits
 - Reduced deployment time
 - Continuous integration validation
 - Flexibility
- Don't forget Hooke's Law (adapted)
 - Under load, an organisation can extend proportional to external force
 - Too much load leads to permanent deformation
- Ensure the tail is moving fast as well as the head
 - Application support
 - Cultural challenges
 - Process change

Communities Encourage Change

- Open source collaboration sets model for in-house teams
- External recognition by the community is highly rewarding for contributors
- Reviews and being reviewed is a constant learning experience
- Operator sharing is much more than just OpenStack



Keystone 2013



Paris 2014

Deployment Models



- Skills
- Culture
- Need for Speed
- Risk Appetite

Keeping Up with Releases



Users and Technology



When it's not on the menu?



Nova Cells

Top level cell

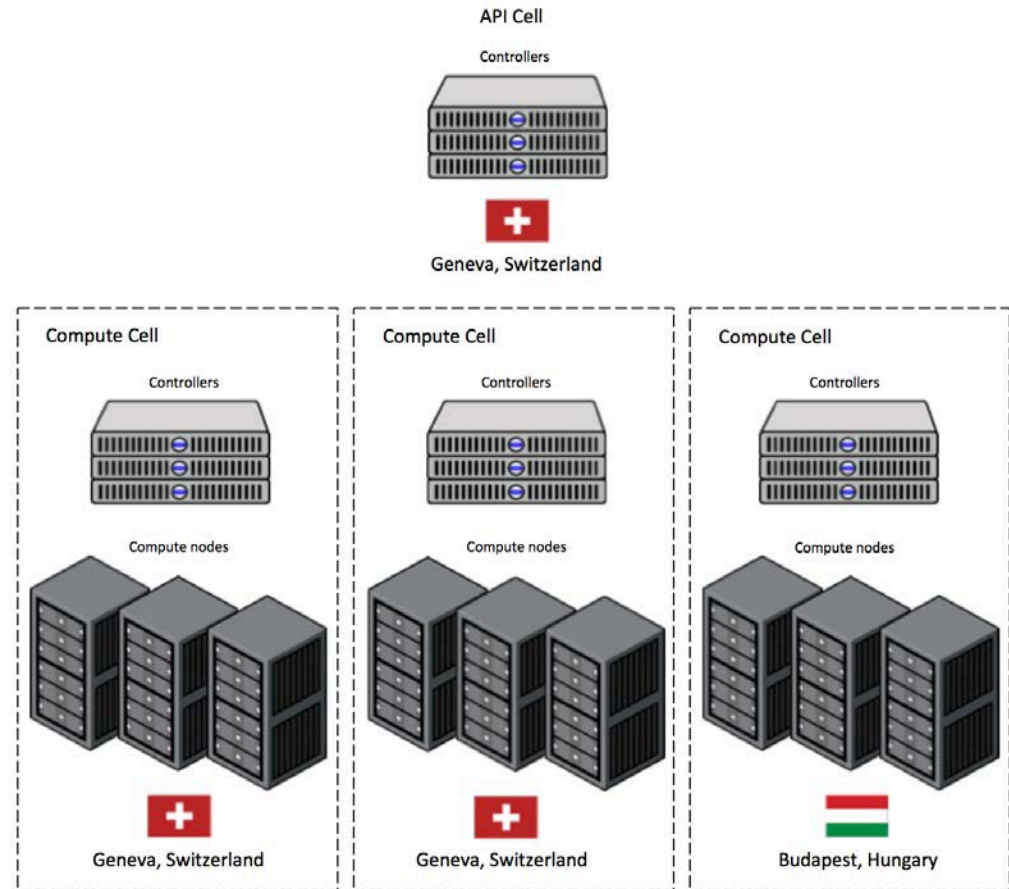
- Runs API service
- Top cell scheduler

Child cells run

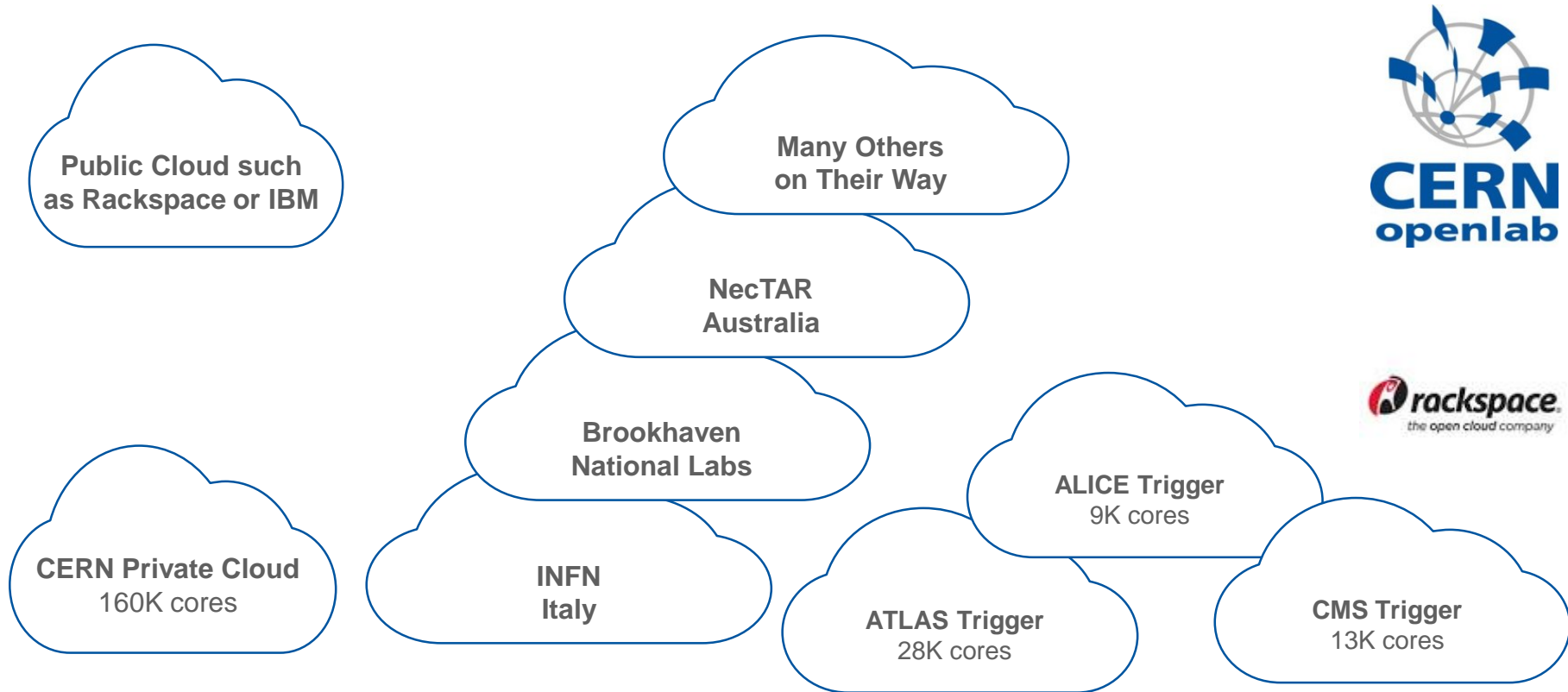
- Compute nodes
- Nova network
- Scheduler
- Conductor

Version 2 coming

- Default for all



Onwards Federated Clouds



Available in standard OpenStack since Kilo

Containers on Clouds



For the user

- Interactive
- Dynamic
- Choice

For IT

- **Timely!**
- Secure
- Managed
- Integrated

Higgs decay to two photons

The Standard Model predicted the decay of the [Higgs bosons](#) into photons. The process is depicted by the diagrams below:

(a) (b) (c)

At the [Large Hadron Collider](#), this process has been measured. This figure shows how a Higgs boson decay looks in the CMS detector:

This ROOTbook illustrates a simplified fitting procedure aiming to identify the peak due to the Higgs boson decay over the exponentially falling background.

Importing input data into a ROOT file

First of all we import the input data, here simplistically stored into a text file, into a [ROOT file](#).

```
In [1]: TTree tree("higgsTree", "The tree cont");
auto next = tree.Read11c("Hgg.txt", ">");
if (next != 0) {
```

CERN now runs Magnum on the production cloud



Outlook for next 12 months

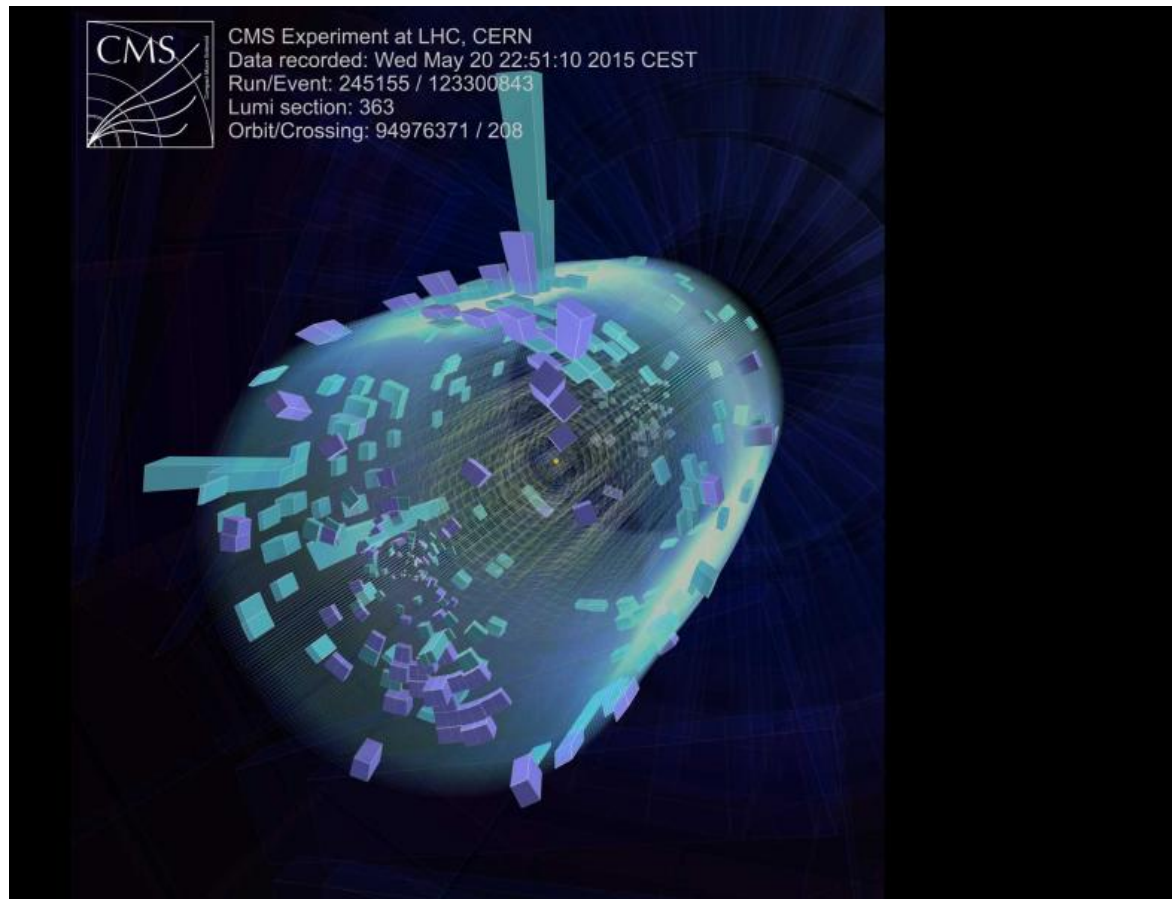
- Scale out
 - Another ~100K cores end 2016
 - Further significant retirements and migrations
- Consolidate
 - EC2 API project
 - Nova network to Neutron
- Enhance
 - Kubernetes, Mesos, Swarm containers for all
 - Bare metal (Ironic) and Workflow (Mistral) to pilot
 - Accounting / Fleet Management
- Investigate further capacity options
 - Public clouds

Summary

- OpenStack at CERN has been in production for 3 years
- Major cultural and technology changes have been successfully addressed
- Contributing back upstream has led to sustainable tools and effective technology transfer

This transformation would not have been possible without community contribution

Thanks to all of you for contributing!



Technical details at <http://openstack-in-production.blogspot.fr>

Thanks to the CERN IT team and
Francois Briard from CERN
communications

Some history of scale...

Date	Collaboration sizes	Data volume, archive technology
Late 1950's	2-3	Kilobits, notebooks
1960's	10-15	kB, punchcards
1970's	~35	MB, tape
1980's	~100	GB, tape, disk
1990's	~750	TB, tape, disk
2010's	~3000	PB, tape, disk

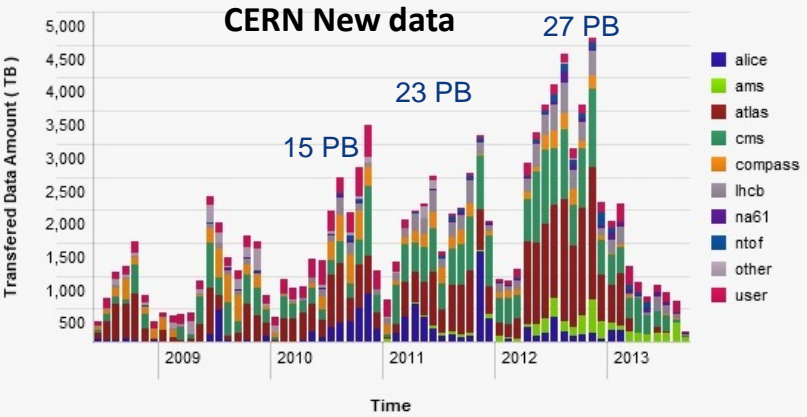
For comparison:

1990's: Total LEP data set
~few TB

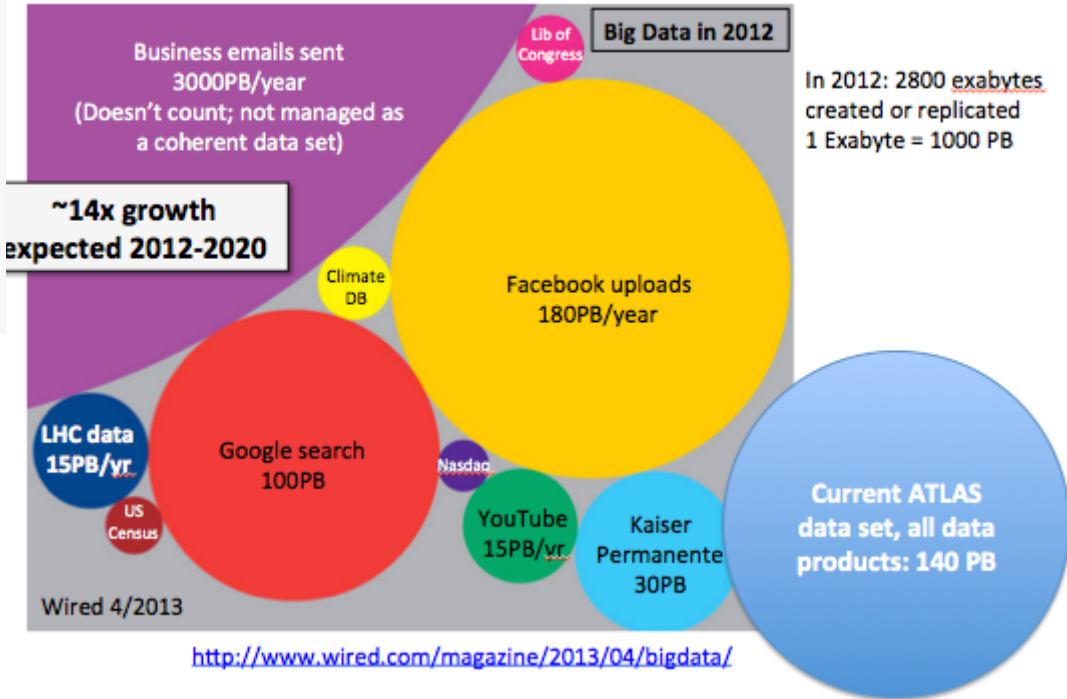
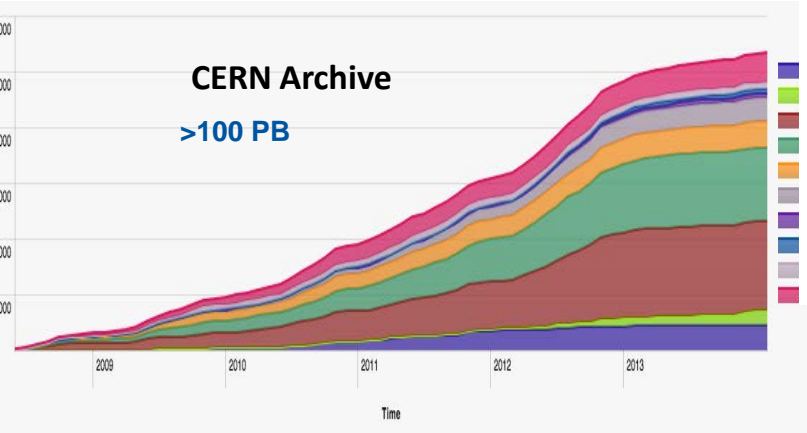
Would fit on 1 tape today

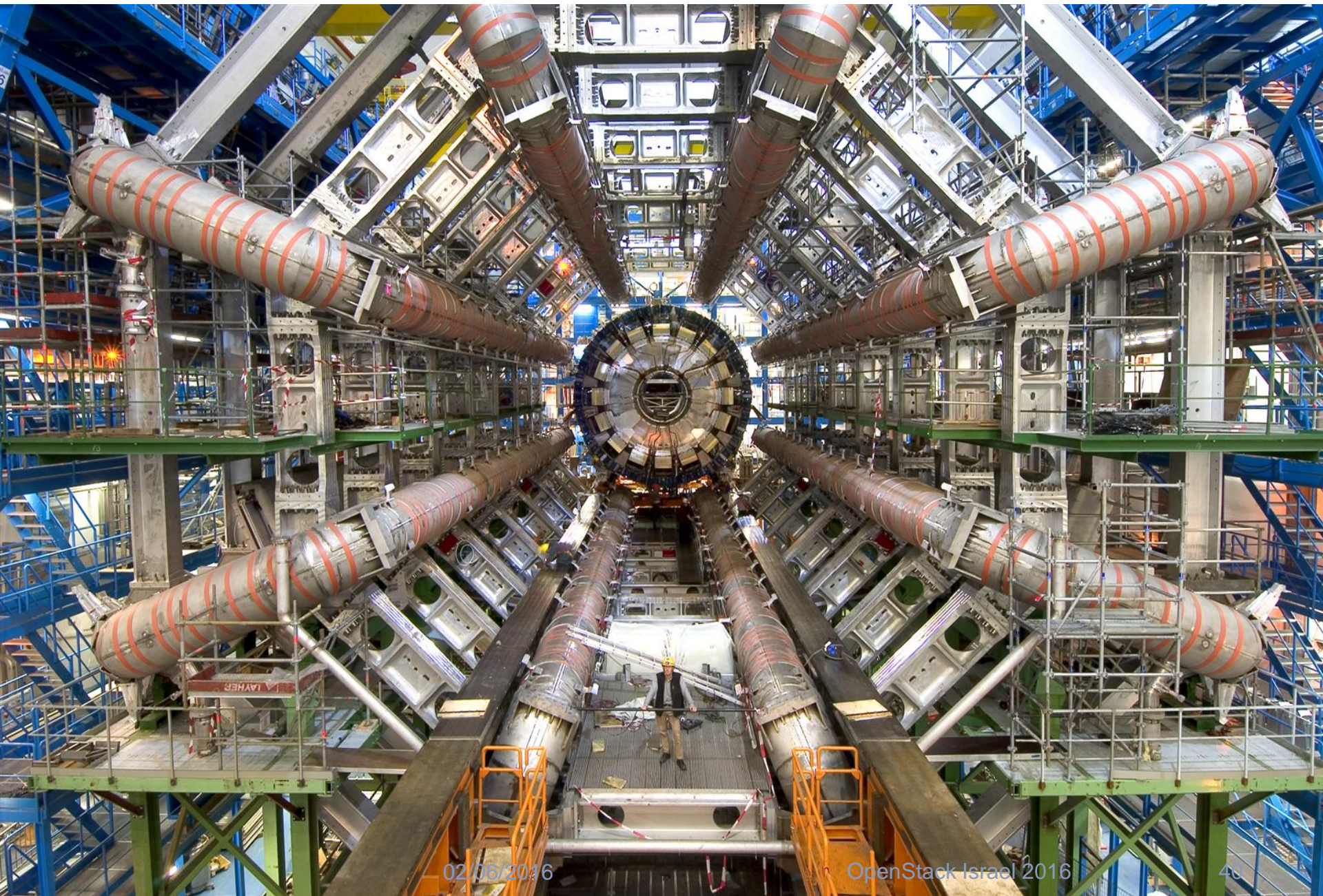
Today: 1 year of LHC data
~27 PB

CERN New data



CERN Archive

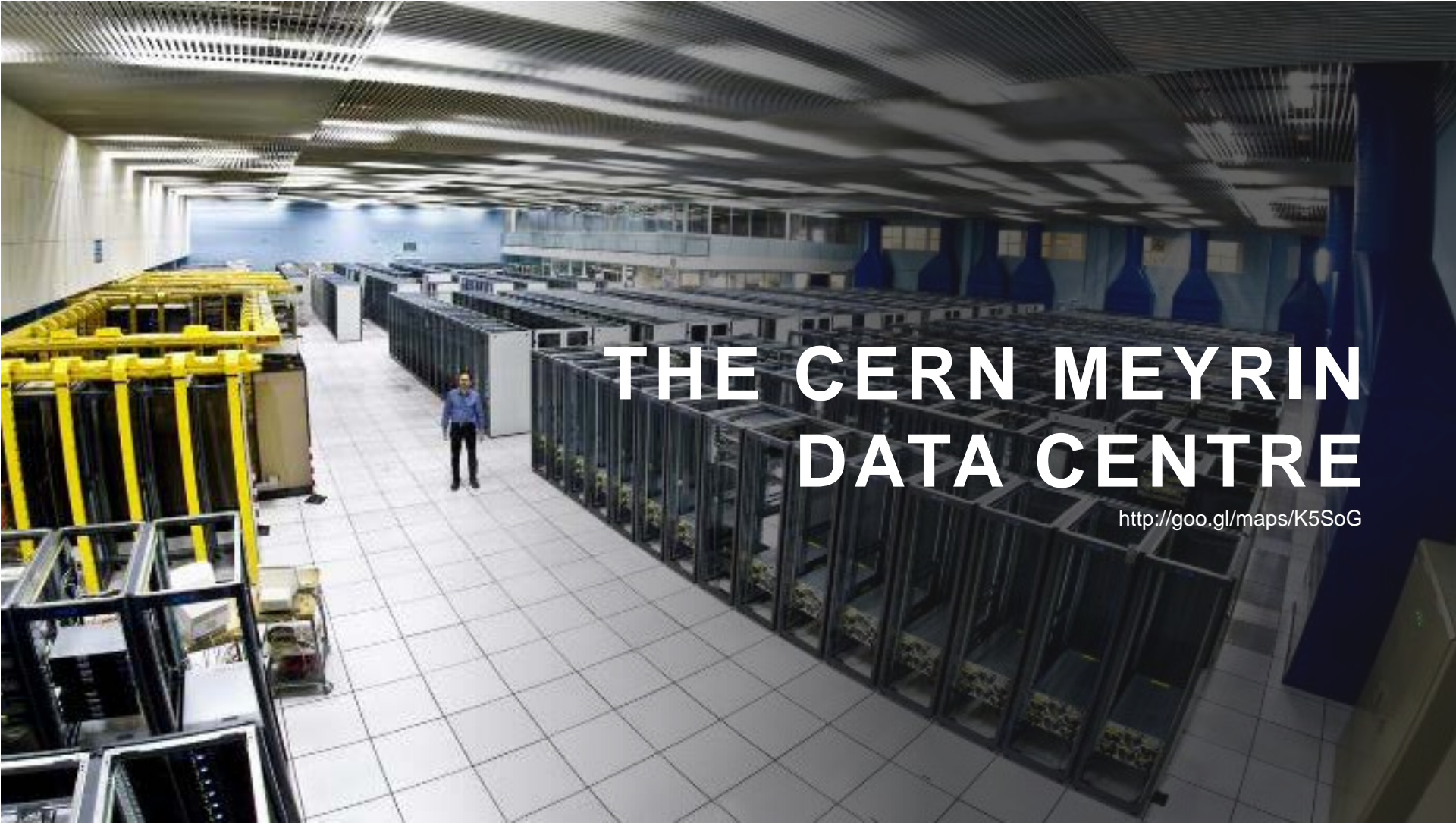




02/06/2016

OpenStack Israel 2016

46



THE CERN MEYRIN DATA CENTRE

<http://goo.gl/maps/K5SoG>

Public Procurement Cycle

Step	Time (Days)	Elapsed (Days)
User expresses requirement		0
Market Survey prepared	15	15
Market Survey for possible vendors	30	45
Specifications prepared	15	60
Vendor responses	30	90
Test systems evaluated	30	120
Offers adjudicated	10	130
Finance committee	30	160
Hardware delivered	90	250
Burn in and acceptance	30 days typical with 380 worst case	280
Total		280+ Days

Good News, Bad News

- Additional data centre in Budapest now online
- Increasing use of facilities as data rates increase

But...

- Staff numbers are fixed, no more people
- Materials budget decreasing, no more money
- Legacy tools are high maintenance and brittle
- User expectations are for fast self-service

Innovation Dilemma

- How can we avoid the sustainability trap ?
 - Define requirements
 - No solution available that meets those requirements
 - Develop our own new solution
 - Accumulate technical debt
- How can we learn from others and share ?
 - Find compatible open source communities
 - Contribute back where there is missing functionality
 - Stay mainstream

Are CERN computing needs really special ?

Data flow to permanent storage: 4-6 GB/sec

CERN Computer Centre

LHCb ~ 200-400 MB/sec

ATLAS ~ 1-2 GB/sec

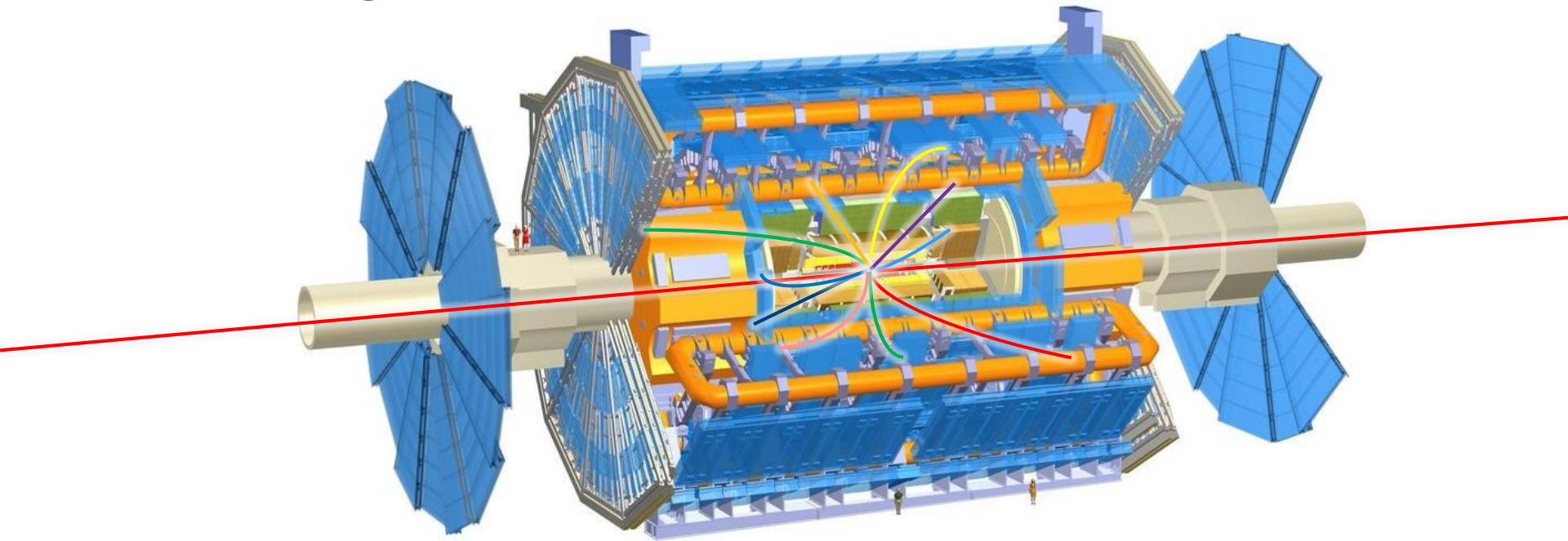
ALICE ~ 1.25 GB/sec

CMS ~ 1-2 GB/sec

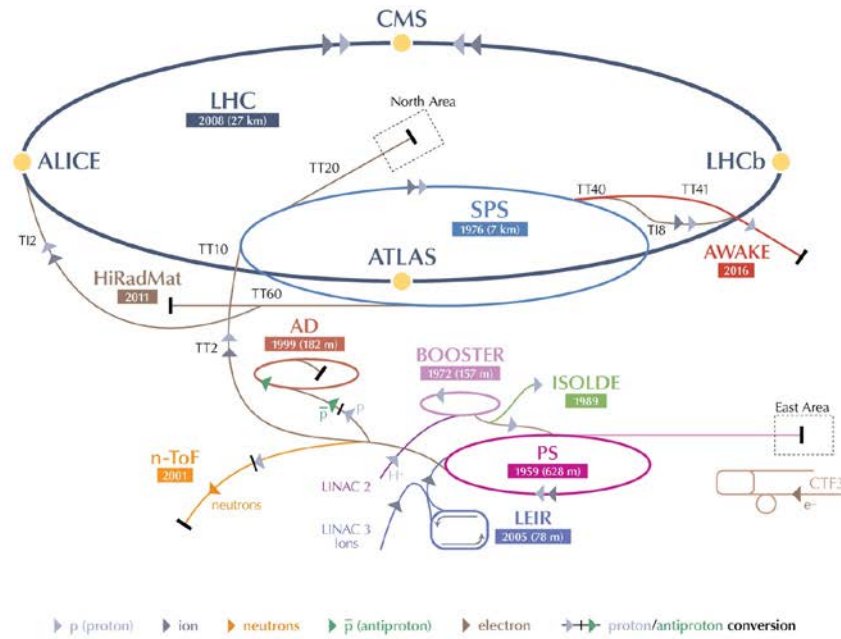
OpenStack Israel 2016



The largest detectors



CERN's Accelerator Complex



LHC Large Hadron Collider SPS Super Proton Synchrotron PS Proton Synchrotron

AD Antiproton Decelerator CTF3 Clic Test Facility AWAKE Advanced WAKEfield Experiment ISOLDE Isotope Separator OnLine DEvice

LEIR Low Energy Ion Ring LINAC LINEar ACcelerator n-ToF Neutrons Time Of Flight HiRadMat High-Radiation to Materials

© CERN 2013



Strategic Plan

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



To support the computing capacity needs for the ATLAS experiment



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



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



To improve the speed and quality of research for finding surrogate biomarkers based on brain images

Additional Users:



Suppliers

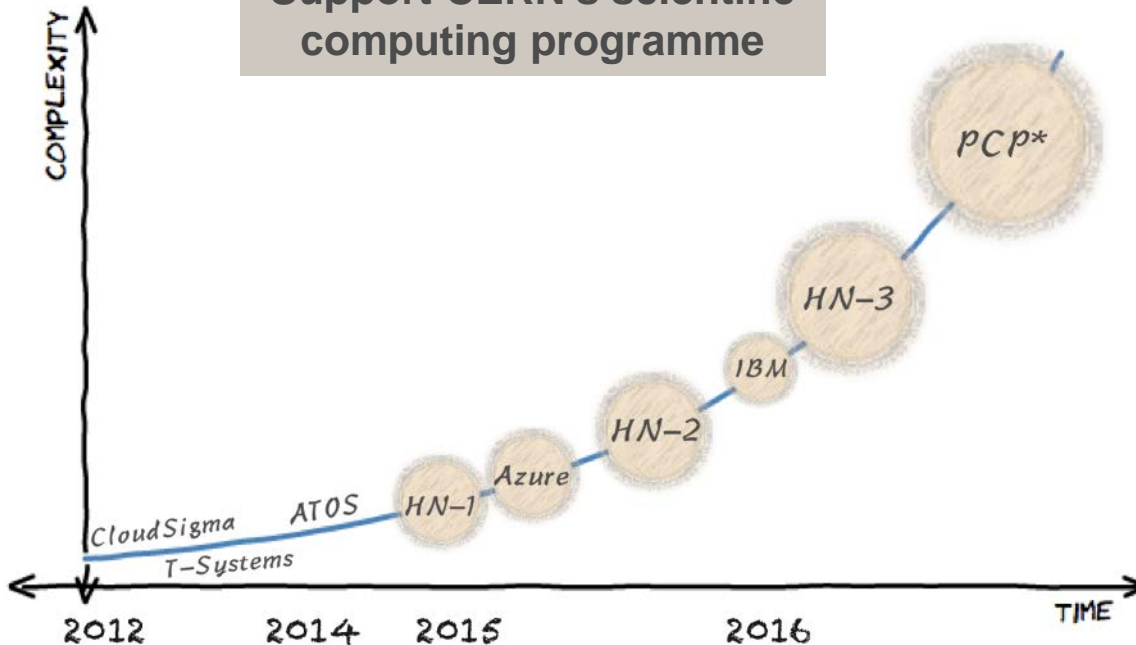


Adopters



Investigations

Support CERN's scientific computing programme



HN - Helix Nebula

- Partnership between research organization and European commercial cloud providers

* EC co-funded joint Pre-Commercial Procurement (PCP) project: <https://indico.cern.ch/event/319753>

** Other work has been conducted outside CERN, such as the Amazon Pilot project at BNL for ATLAS

Tier 0

MEYRIN DATA CENTRE

	last_value
● Number of Cores in Meyrin	151,107
● Number of Drives in Meyrin	83,702
● Number of 10G NIC in Meyrin	9,305
● Number of 1G NIC in Meyrin	23,641
● Number of Processors in Meyrin	25,207
● Number of Servers in Meyrin	13,373
● Total Disk Space in Meyrin (TB)	175,893
● Total Memory Capacity in Meyrin (TB)	613

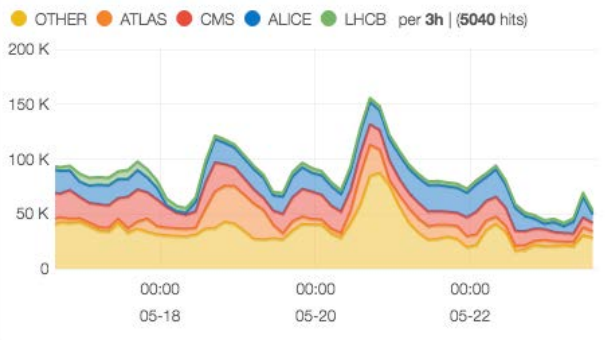
WIGNER DATA CENTRE

	last_value
● Number of Cores in Wigner	43,328
● Number of Drives in Wigner	23,180
● Number of 10G NIC in Wigner	1,399
● Number of 1G NIC in Wigner	5,067
● Number of Processors in Wigner	5,418
● Number of Servers in Wigner	2,712
● Total Disk Space in Wigner (TB)	71,738
● Total Memory Capacity in Wigner (TB)	172

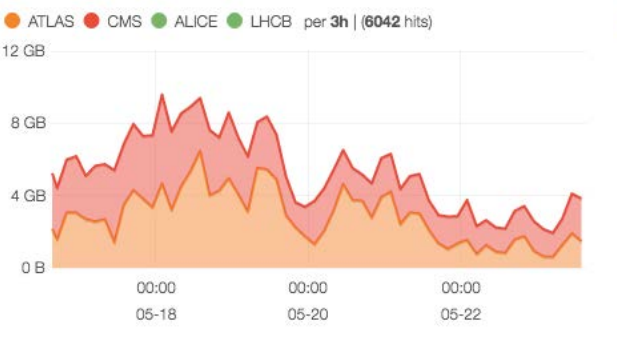
NETWORK AND STORAGE

	last_value
● Tape Drives	104
● Tape Cartridges	20,517
● Data Volume on Tape (TB)	140,606
● Free Space on Tape (TB)	44,024
● Routers (GPN)	140
● Routers (TN)	30
● Routers (Others)	107
● Switches	3,708

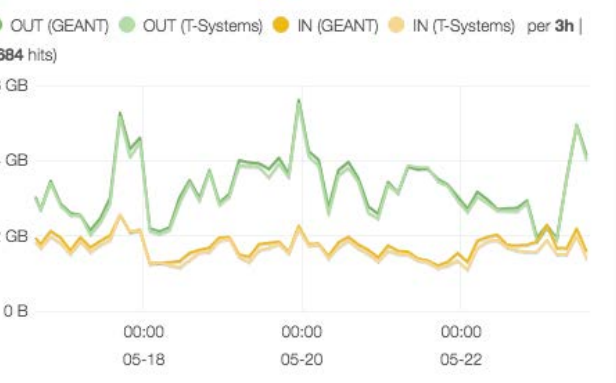
BATCH JOBS (#)



FILE TRANSFER THROUGHPUT (GB/S)



WIGNER NETWORK LINKS (GBIT/S)



Run 2 has only just started

- Hint of an excess with diphoton mass of 750 GeV
 - Seen by ATLAS and CMS – coincidence or a new signal?

