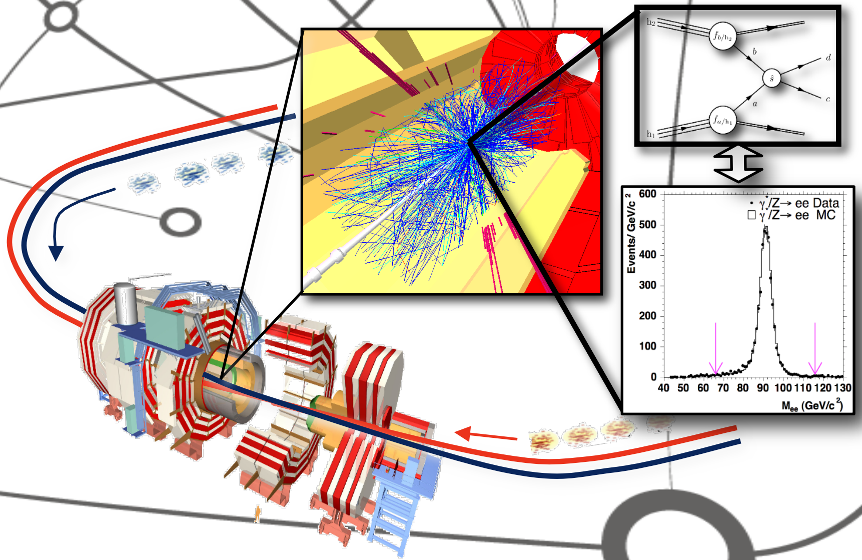
Big Data Analytics and the LHC

Maria Girone

CERN openlab CTO



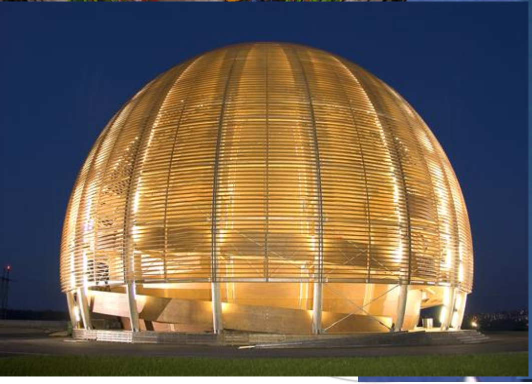
Computing Frontiers 2016, Como, May 2016

DOI: 10.5281/zenodo.45449,

European Organization for Nuclear Research, founded in 1954 by 12 European countries

"Science for Peace"

~ 2300 staff
~ 1600 other paid personnel
~ 10500 scientific users
Budget (2014) ~1000 MCHF



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

Candidate for Accession: Romania

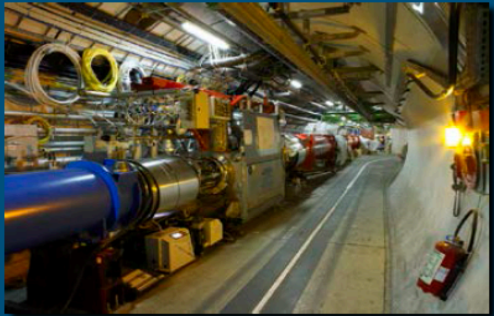
Associate Member in Pre-Stage to Membership: Serbia

Applicant States for Membership or Associate Membership:
Brazil, Croatia, Cyprus, Pakistan, Russia, Slovenia, Turkey, Ukraine

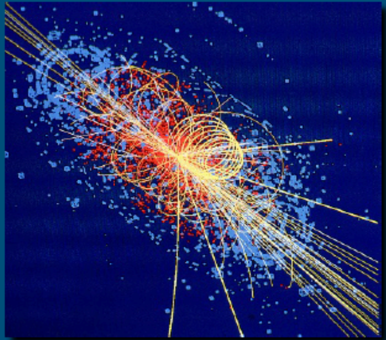
Observers to Council: India, Japan, Russia, Turkey, United States of America; European Commission and UNESCO

➔ Interfacing between fundamental science and key technological developments

➔ CERN Technologies and applications



Accelerating particle beams

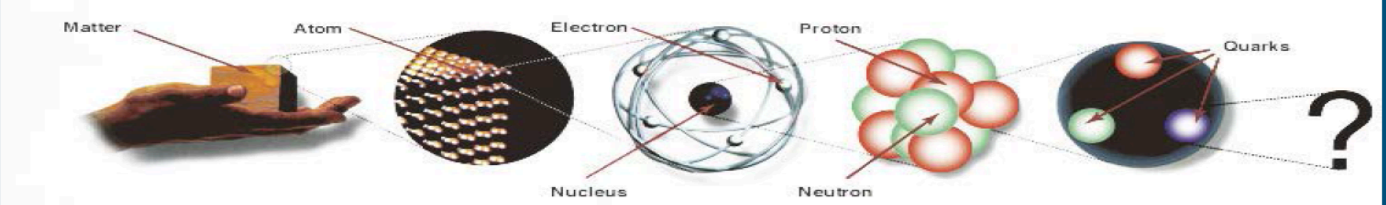


Detecting particles

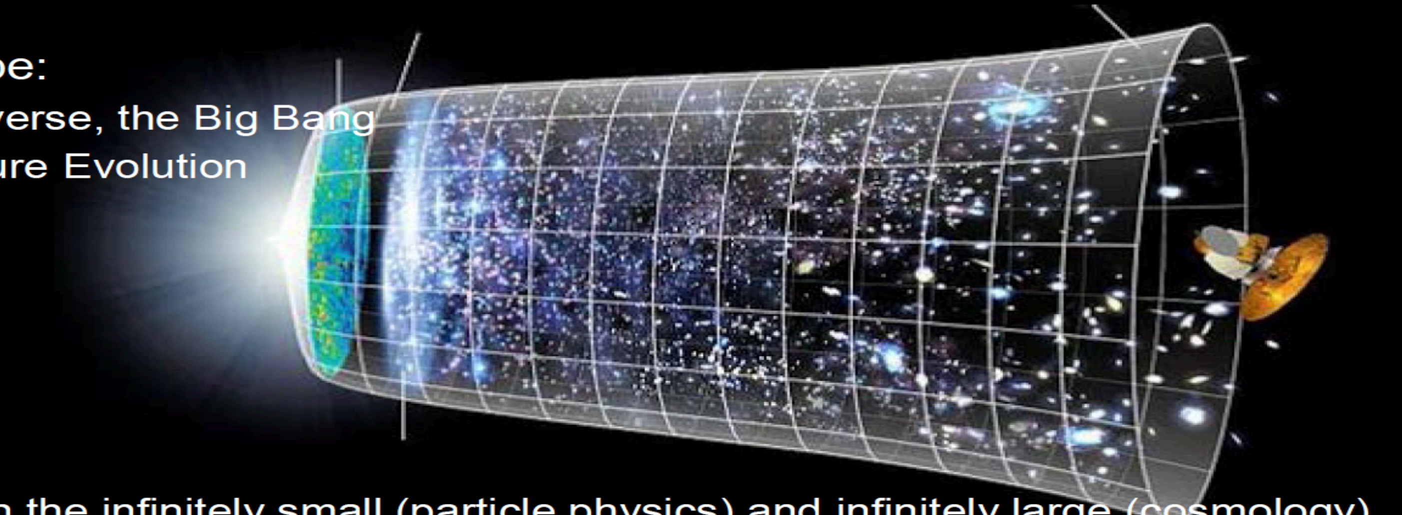
Large-scale computing (Grid)

Number 1: "Particle physics"

- Quest to understand:
  - Fundamental constituents of matter - Matter particles
  - Interactions with which particles act on each other - Interactions
  - Particles propagating the interactions - Messenger particles



- Ultimately describe:
  - Birth of the Universe, the Big Bang
  - Passed and future Evolution



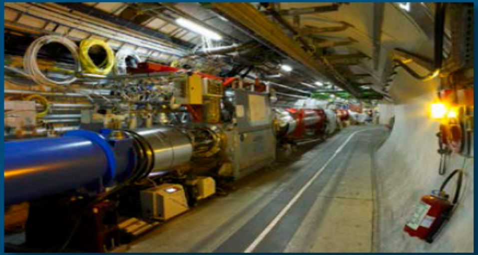
- Strong link between the infinitely small (particle physics) and infinitely large (cosmology)

"Big bang" in the laboratory

- We gain insight by *colliding particles at the highest energies* possible to measure:
  - Production rates
  - Masses & lifetimes
  - Decay rates
- From this we *derive the "spectroscopy" as well as the "dynamics" of elementary particles*.
- Progress is made by going to higher energies and brighter beams.
  - Higher energies allow higher mass particles to be produced
  - Brighter beams allow rarer phenomena to be probed
    - brighter beams means more collisions per event => more complex events

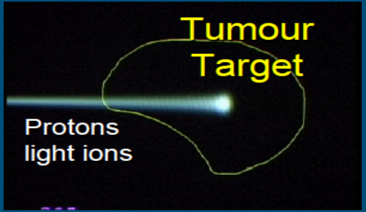
Number 2: Innovation

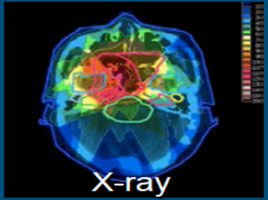
Medical applications

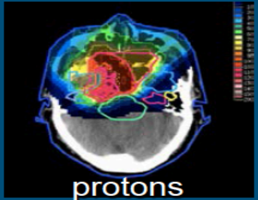


Accelerating particle beams
~30’000 accelerators worldwide
~17’000 used for medicine

Hadron Therapy

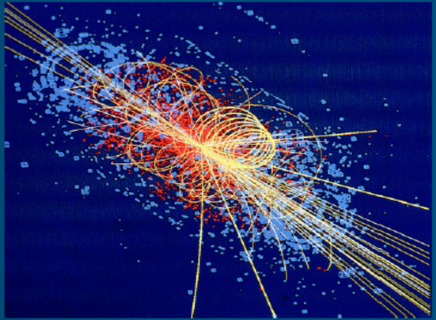






Leadership in Ion Beam Therapy no[w] in Europe and Japan

>100’000 patients treated worldwide (45 facilities)
>50’000 patients treated in Europe (14 facilities)

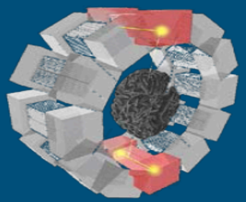


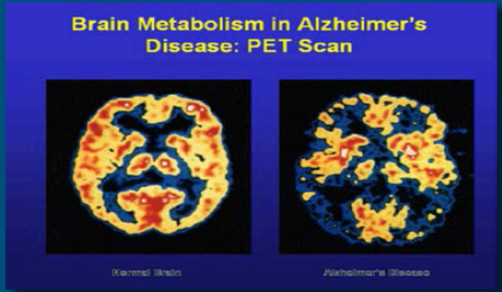
Detecting particles

Imaging

Clinical trial in Portugal, France and Italy for new breast imaging system (ClearPEM)

PET Scanner

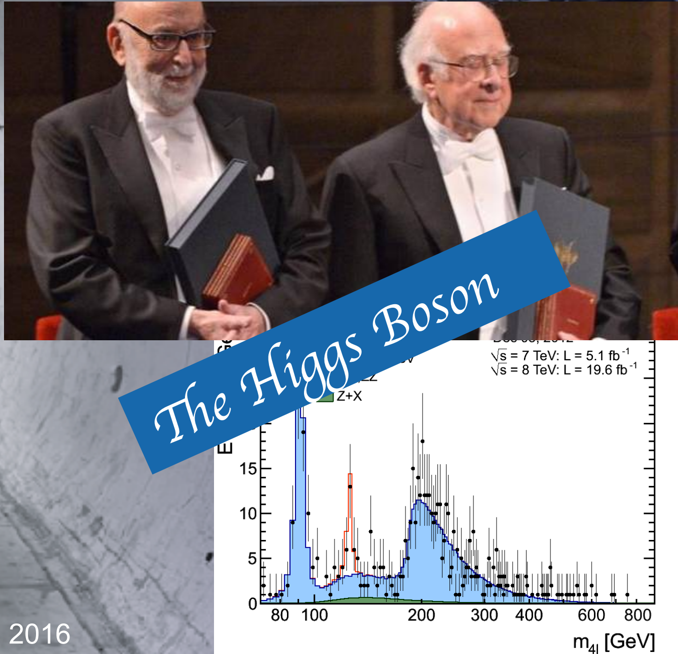




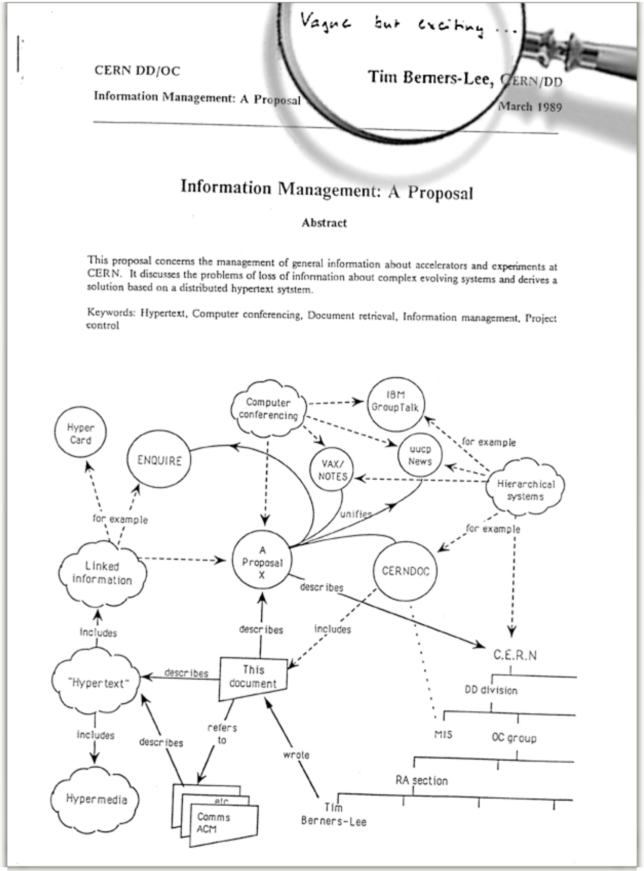
CERN: A UNIQUE ENVIRONMENT TO PUSH TECHNOLOGIES TO THEIR LIMITS

In its 60 year life CERN has made some of the important discoveries in particle physics

- Observation of the W and Z Bosons
- The number of neutrino families
- The Higgs Boson Discovery



CERN – Where the Web was Born



Vague but exciting ...

CERN DD/OC

Information Management: A Proposal

Tim Berners-Lee, CERN/DD

March 1989

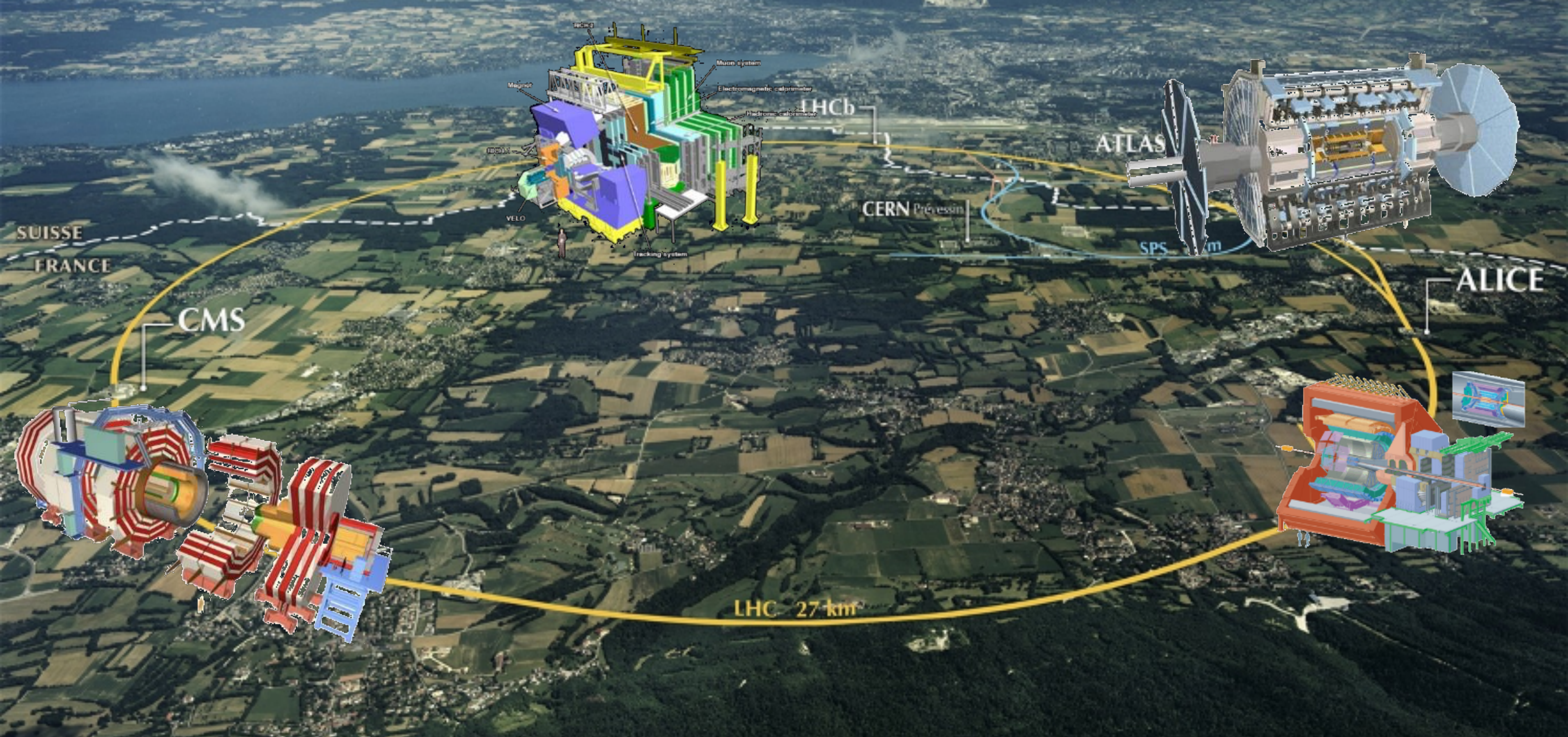
Information Management: A Proposal

Abstract

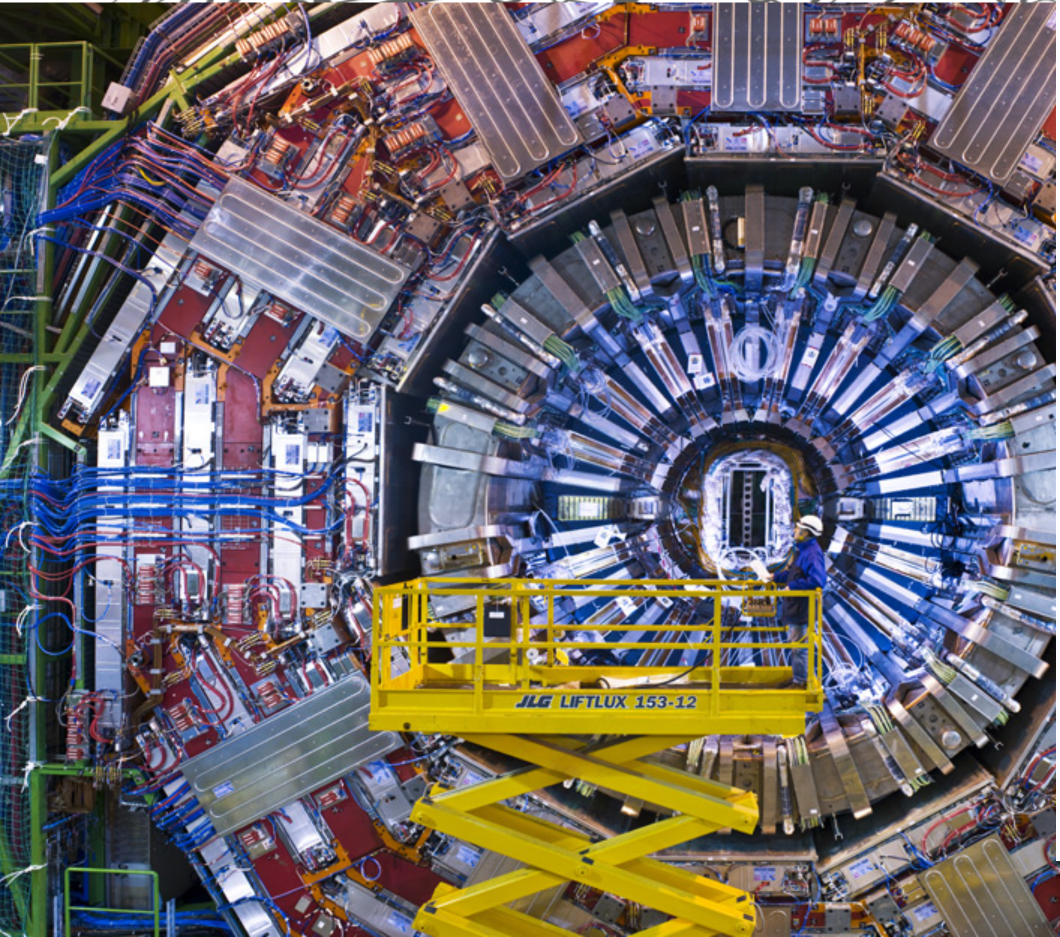
This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext sytstem.

Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control

The Large Hadron Collider (LHC)



The CMS Experiment



- **80 Million electronic channels**
  x 4 bytes
  x 40MHz

  ~ **10 Petabytes/sec** of information
  x 1/1000 zero-suppression
  x 1/100,000 online event filtering

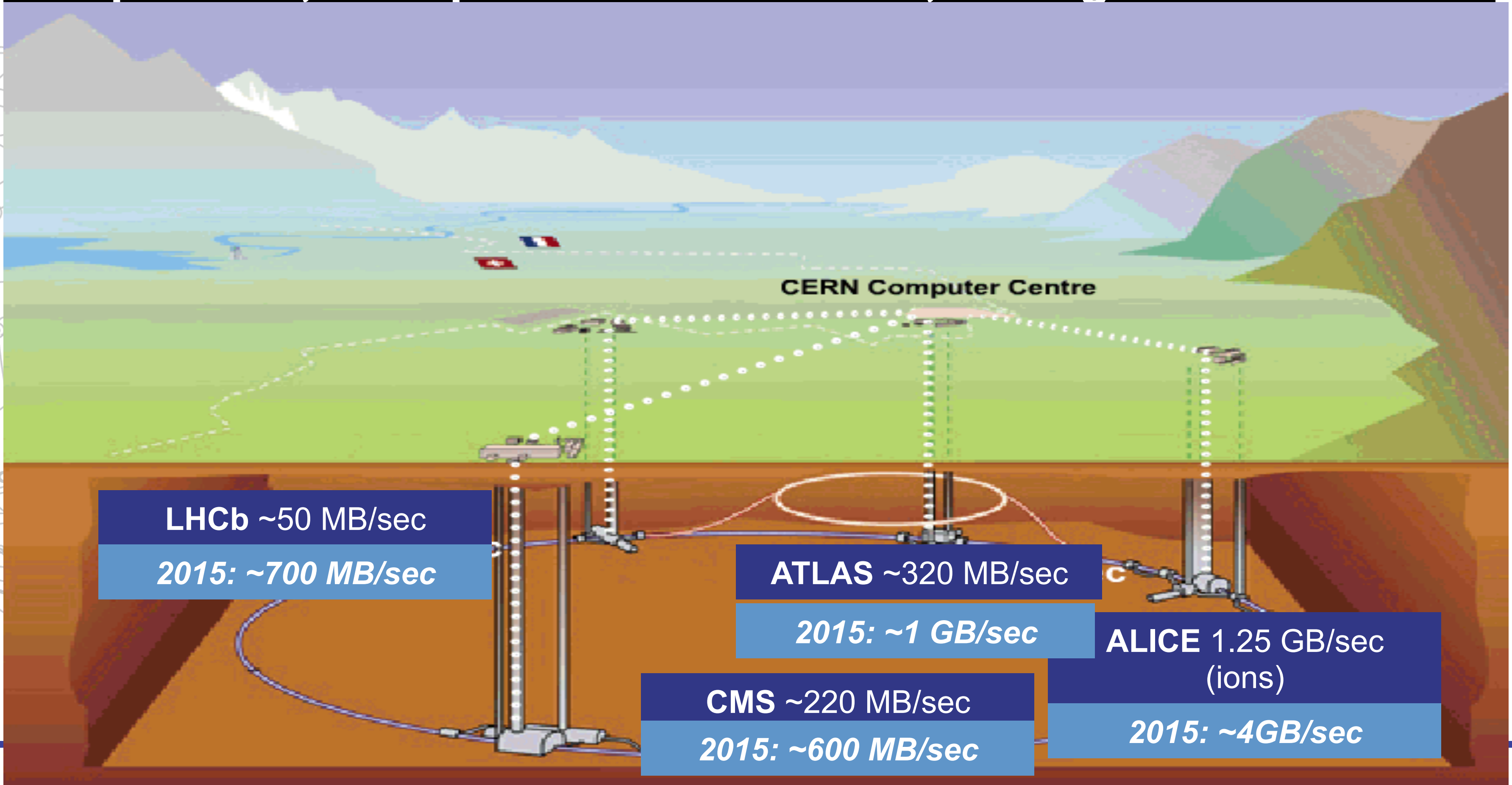
  ~ 1000 Megabytes/sec raw data to tape
  50 Petabytes of data per year
  written to tape, incl. simulations.

- 2000 Scientists (1200 Ph.D. in physics)
  - ~ 180 Institutions
  - ~ 40 countries

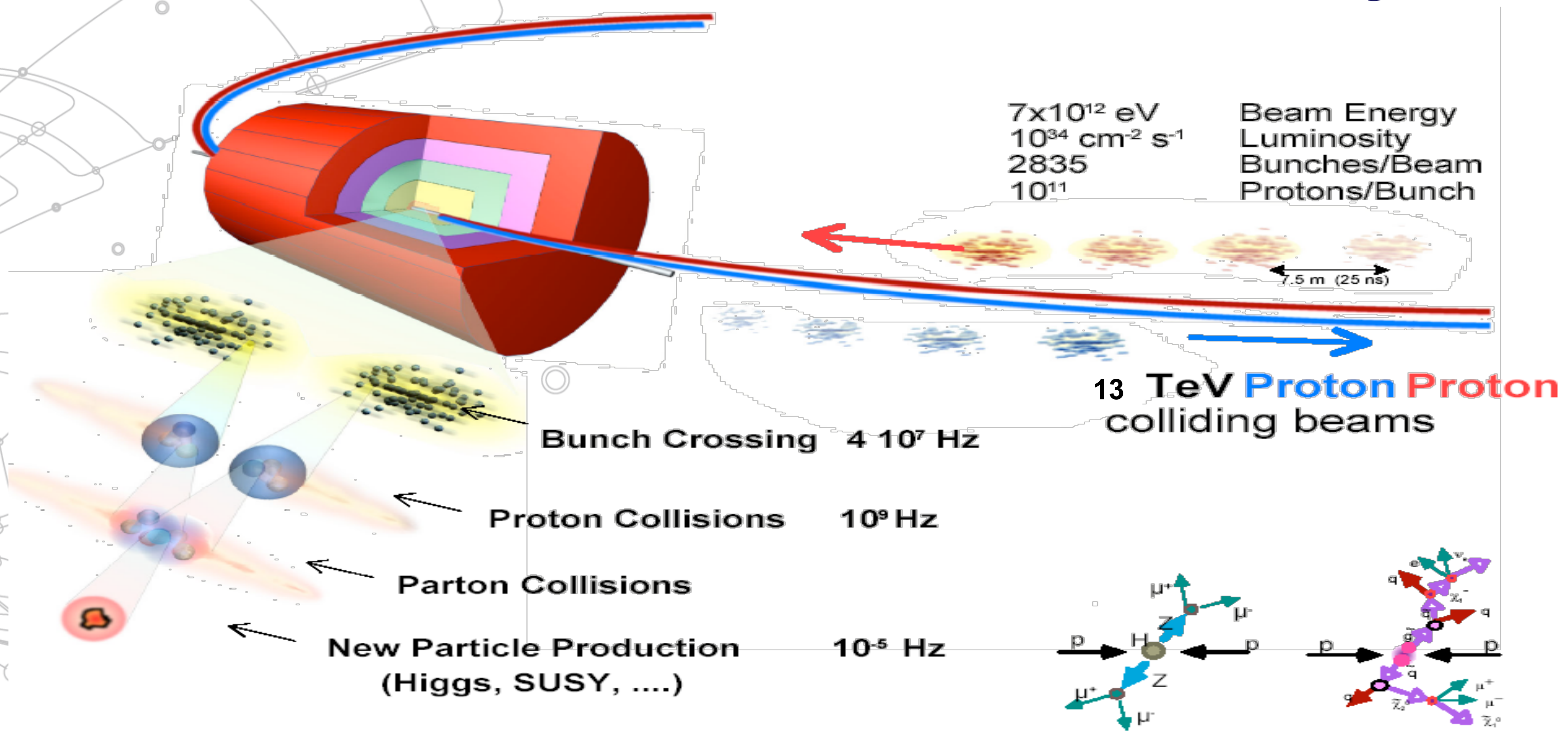
- 12,500 tons, 21m long, 16m diameter

CERN openlab

CERN Computer Centre (Tier-0): Acquisition, First pass reconstruction, Storage & Distribution



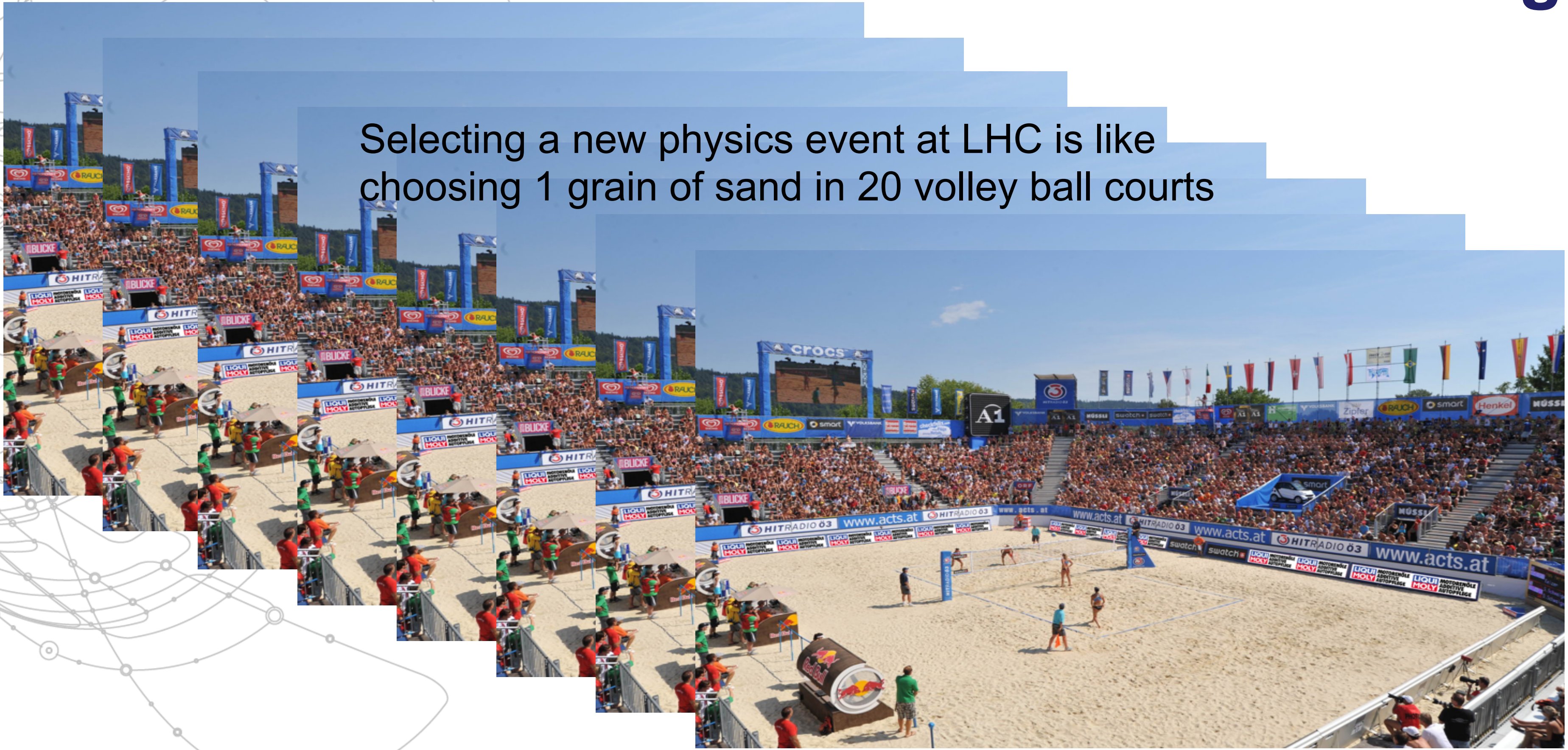
A Needle in a Haystack



Selection of 1 event in 10,000,000,000,000

Data Mining

Selecting a new physics event at LHC is like choosing 1 grain of sand in 20 volley ball courts



LHC Computing Scale

A sample equivalent to the accumulated data/simulation of the 10 year LEP program is produced 5 times a day

1 PB/s of data generated by the detectors

Up to **100 PB/year** of stored data

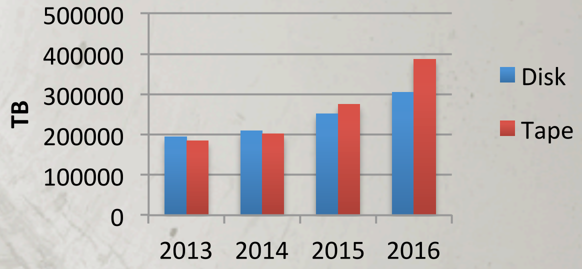
Would put us amongst the top Supercomputers if centrally placed

A distributed computing infrastructure of half a million cores working 24/7

An average of 40M jobs/month

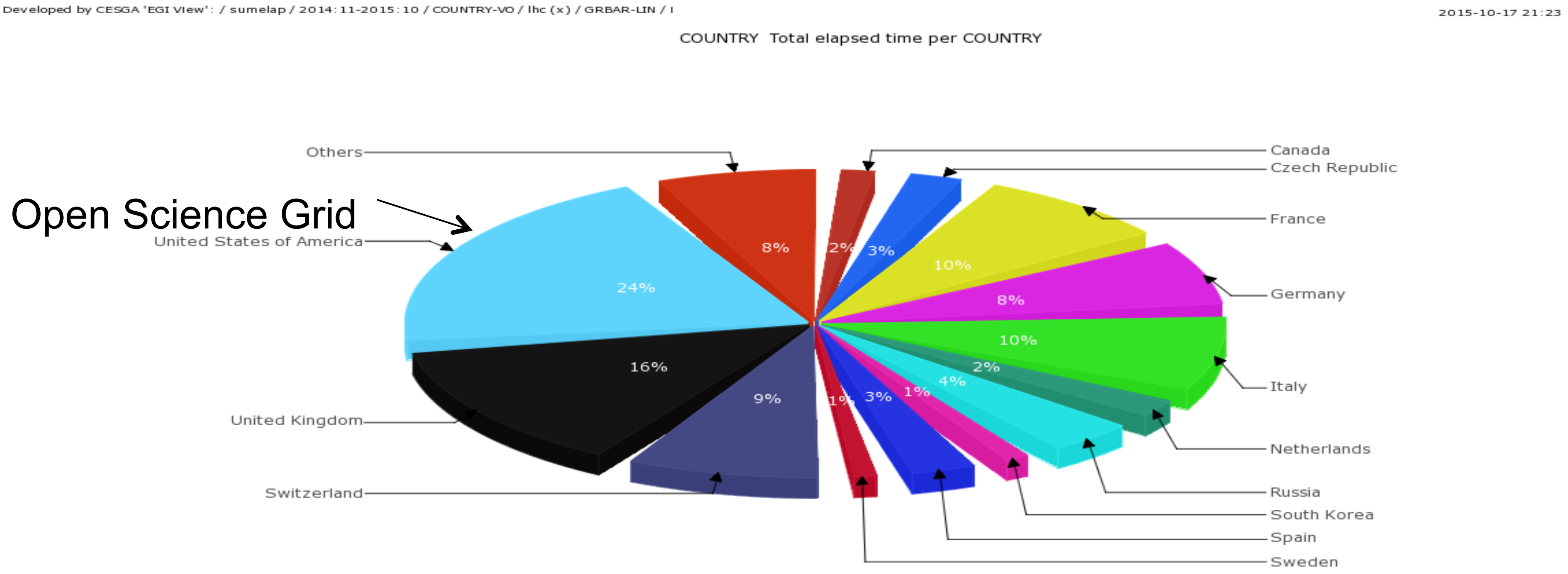
More than 100PB moved and accessed by 10k people

An continuous data transfer rate of 6 GB/s (**600TB/day**) across the Worldwide LHC Grid (WLCG)



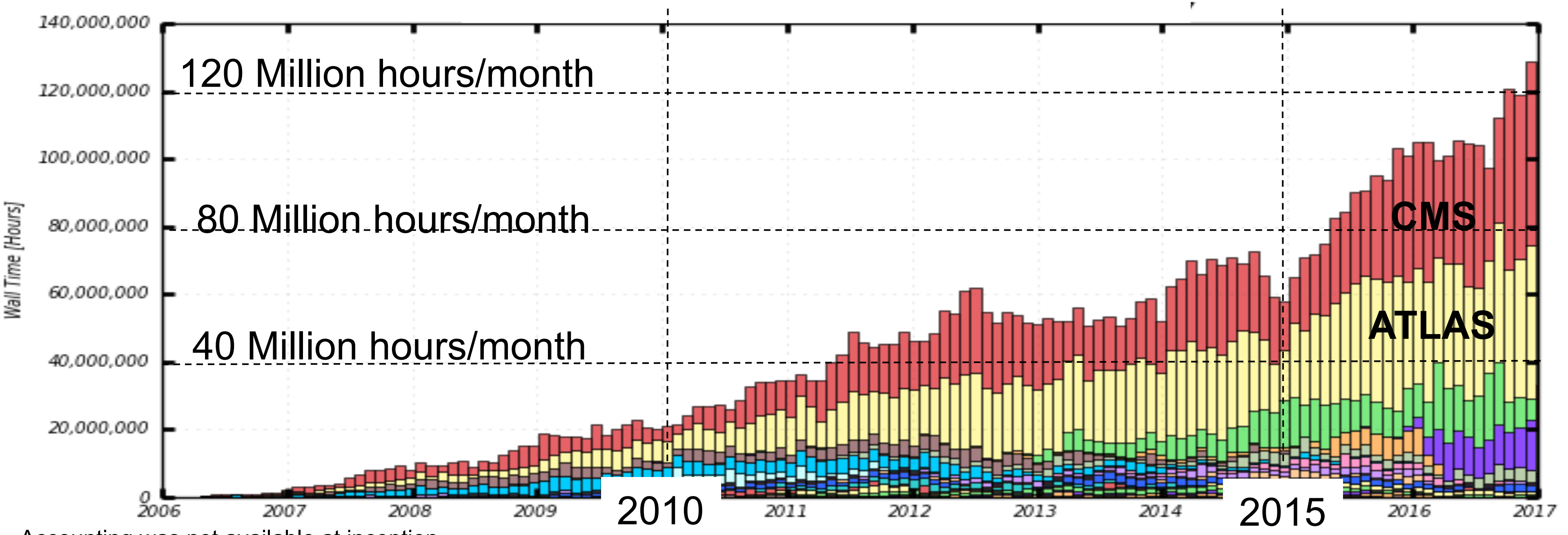
LHC Disk and Tape Storage

Global Distribution



Largest national contribution is only 24% of total resources.

OSG since Inception



Accounting was not available at inception

The Large Hadron Collider Experiments ATLAS & CMS operate on a 180,000 core distributed infrastructure in the US alone

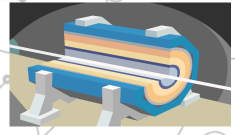
Big Data Analytics

Data Analytics at the LHC

Data Analytics (DA) is the science of examining raw data with the purpose of drawing conclusions about that information

- This is a complicated series of steps at the LHC (Run2)

Data Volume



Processing and people

From Detector (1PB/s)

After HardwareTrigger (TB/s)

HLT

20k cores

DAQ and Trigger (less than 200)

Selected RAW (1GB/s)

Reconstruction

40k cores

Operations (less than 100)

Derived Data (2GB/s)

Reprocessing

30k cores

Operations (less than100)

Derived Data (2GB/s)

Organized Analysis

60k cores

Analysis Users (IMore than 1000)

Analysis Selection (100MB/s)

Final Selection

Analysis Users (IMore than 1000)

Data to Manage

- Datasets => distributed globally
- Calibration Releases => distributed globally
- Software Releases => distributed globally

- A typical physicist doing data analysis
  - uses custom software & configs on top of a standardized software release
  - re-applies some high level calibrations
  - does so uniformly across all datasets used in the analysis.

Software & Calibrations

- Both are distributed via systems that use Squid caches.
- Calibrations:
  - Frontier system is backended by an Oracle DB
- Software:
  - CVMFS is backended by a filesystem
- Data distribution achieved via globally distributed caching infrastructure.

Calibration Data

For details see e.g.:

"Operational Experience with the Frontier System in CMS"

Journal of Physics: Conference Series **396** (2012) 052014

doi:10.1088/1742-6596/396/5/052014

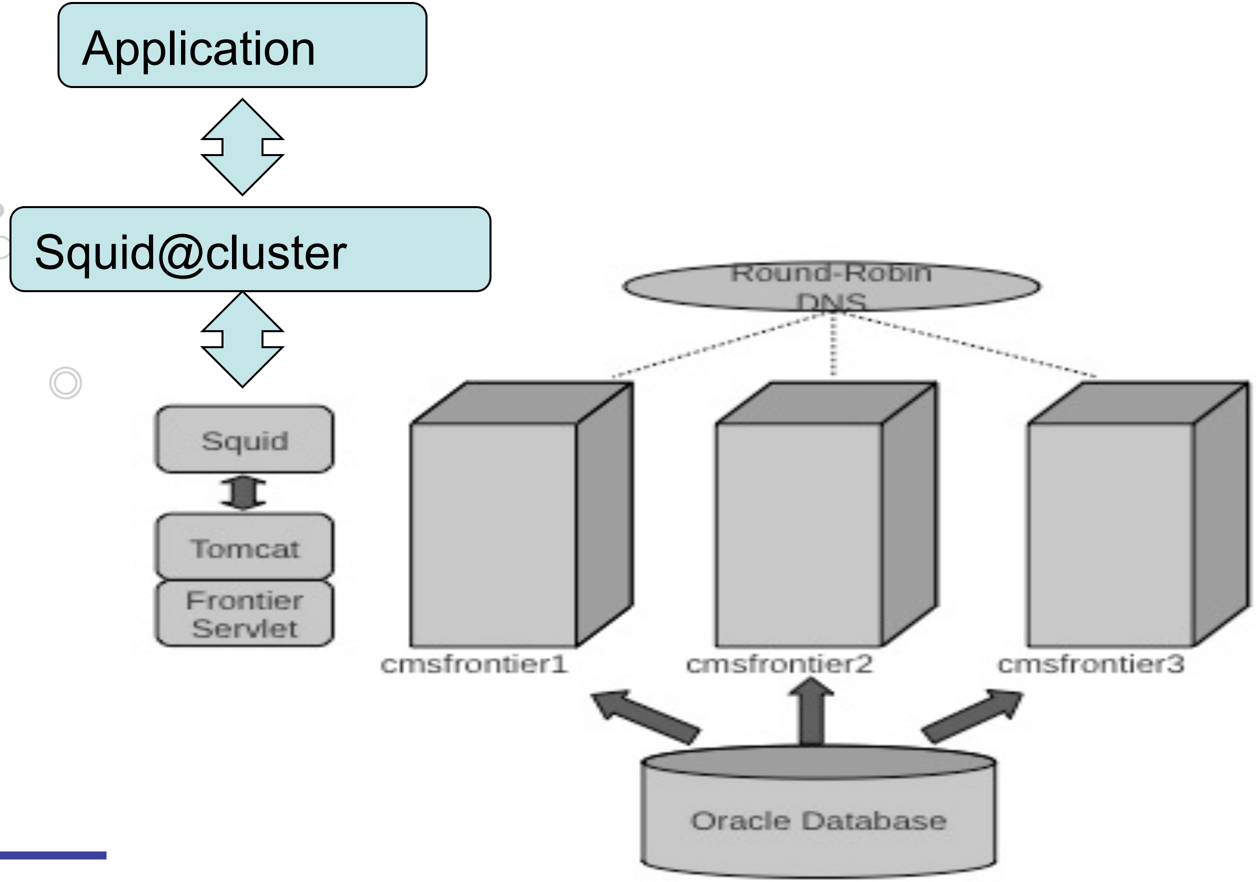
CERNopenlab

Frontier System Architecture

Infrastructure at
~ 100 clusters globally

~ 1 Billion requests
per day in aggregate
~ 40TB of data per day

Infrastructure @ CERN
~ 5 Million requests
per day serving
~ 40GB of data



Software Distribution

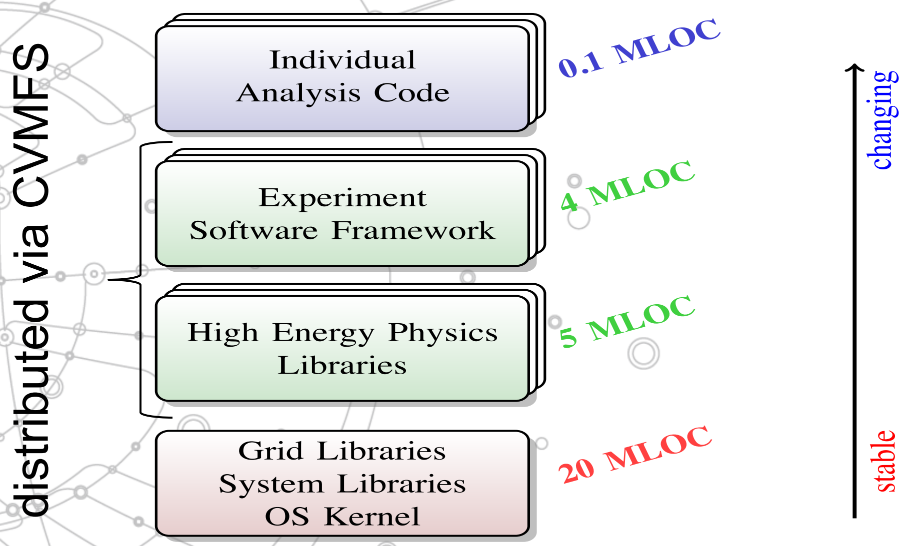
For details see e.g.:

"The need for a versioned Analysis Software Environment"

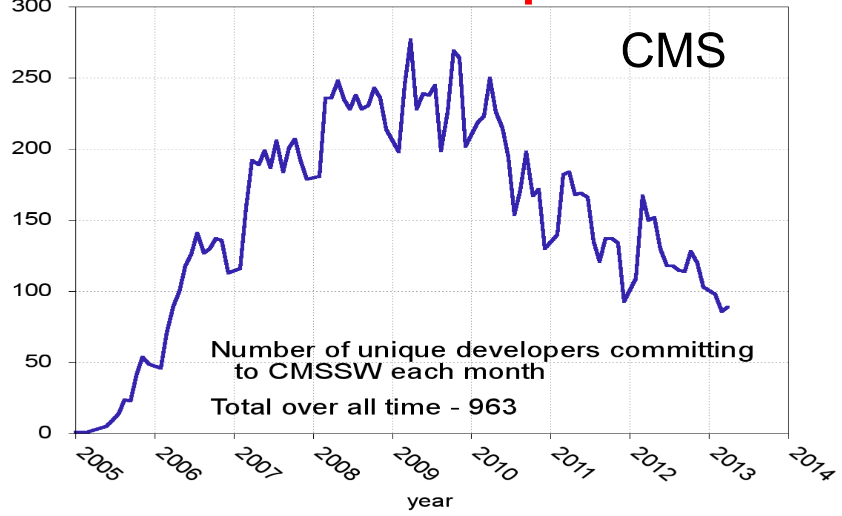
arXiv:1407.3063 [cs.SE]

and references therein.

Software @ LHC

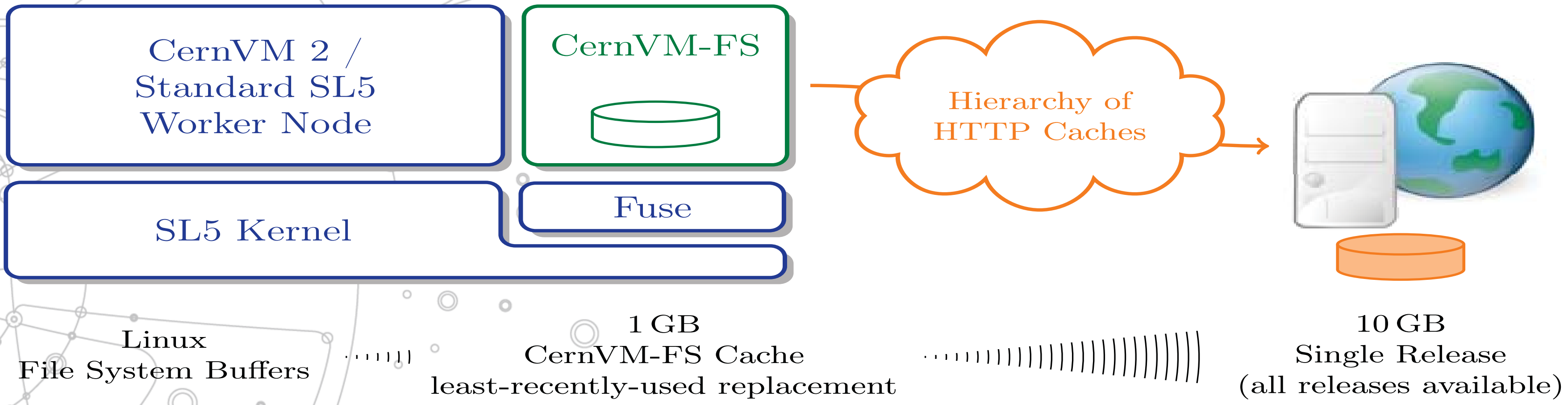


Software is a team sport



~1/4 of the CMS collaboration, scientists and engineers, contributed to the common source code of ~3.6M C++ SLOC.

CVMFS Architecture



Release Statistics for 2013:

	Releases	**Shared libraries and plug-ins per release**
ATLAS	19	3900
CMS	40	2200
ALICE	49	210

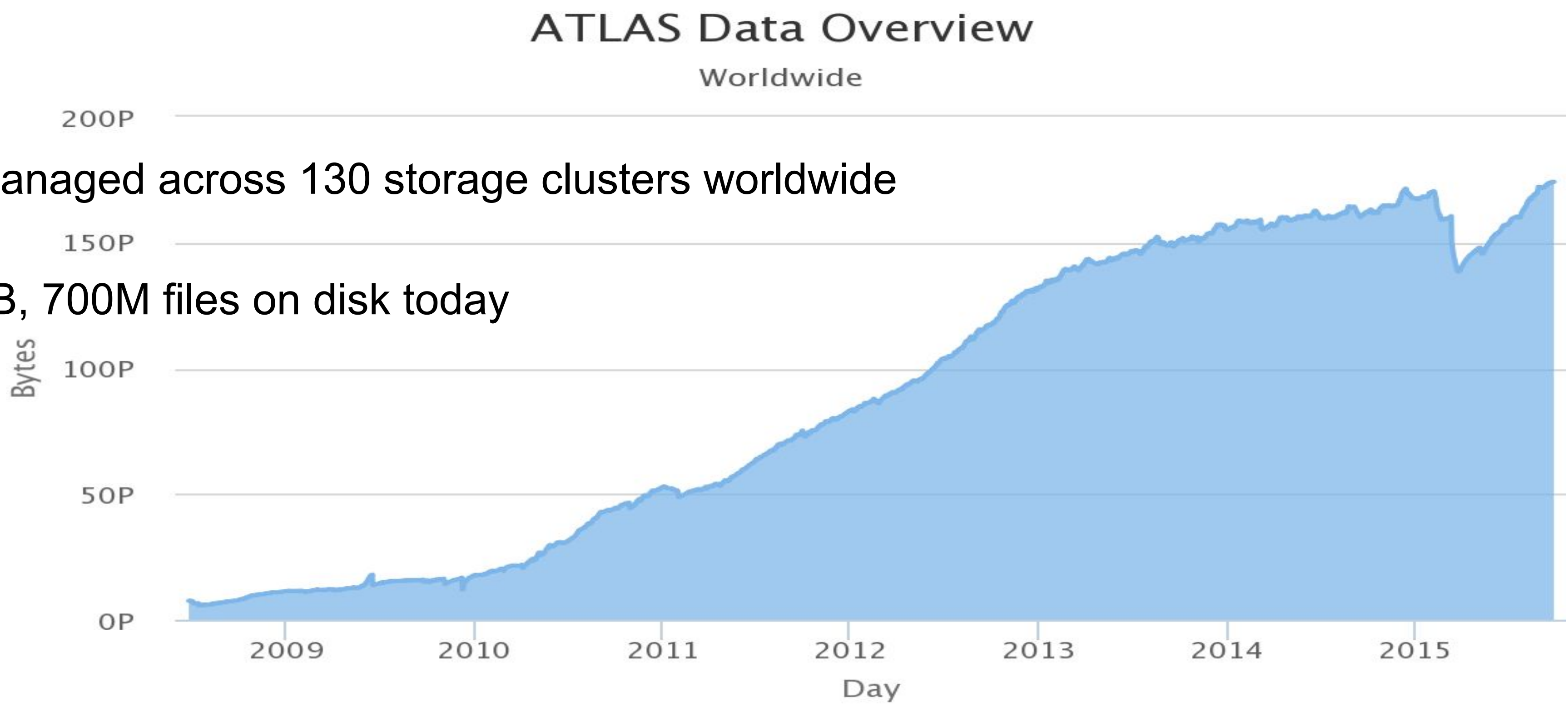
Dataset Distribution

~ 50PB per year per experiment

Dataset Distribution Strategies

- Managed Pre-staging of datasets to clusters
  - managed based on human intelligence
  - managed based on “data popularity”
- Data Transfer integrated with processing workflow
  - determine popularity dynamically based on pending workloads in WMS.
- Remote file open & reads via data federation
- Dynamic caching just like for calibrations & software.

ATLAS Data Total on Disk



Data managed across 130 storage clusters worldwide

170 PB, 700M files on disk today

Any Data, Any Time, Anywhere: Global Data Access for Science

http://arxiv.org/abs/1508.01443

… making the case for WAN reads.

CERN openlab

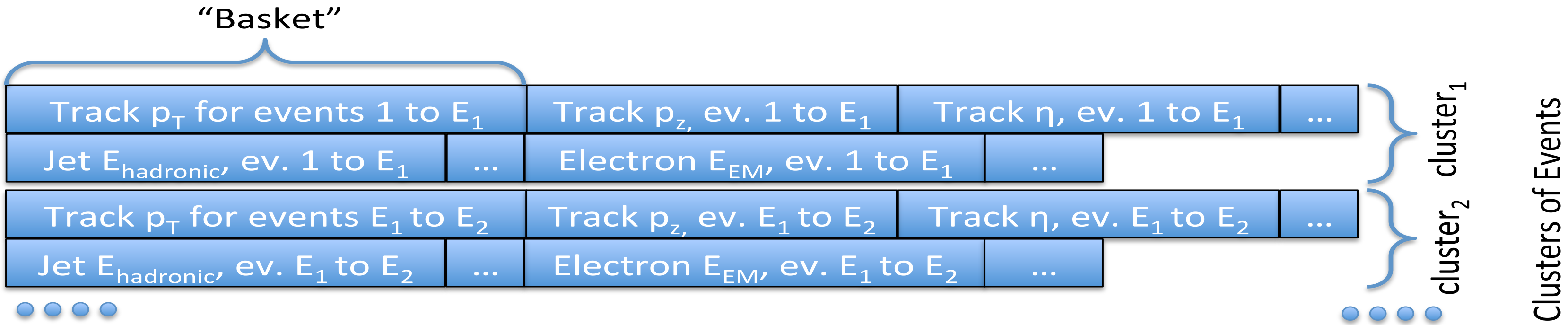
Optimize Data Structure for Partial Reads

An Event has many collections of objects with many components.

Track collection: { Track_1, Track_2, … Track_{Nt} } a few 100s per event

Jet collection: { Jet_1, Jet_2, … Jet_{Nj} } a few 10s per event

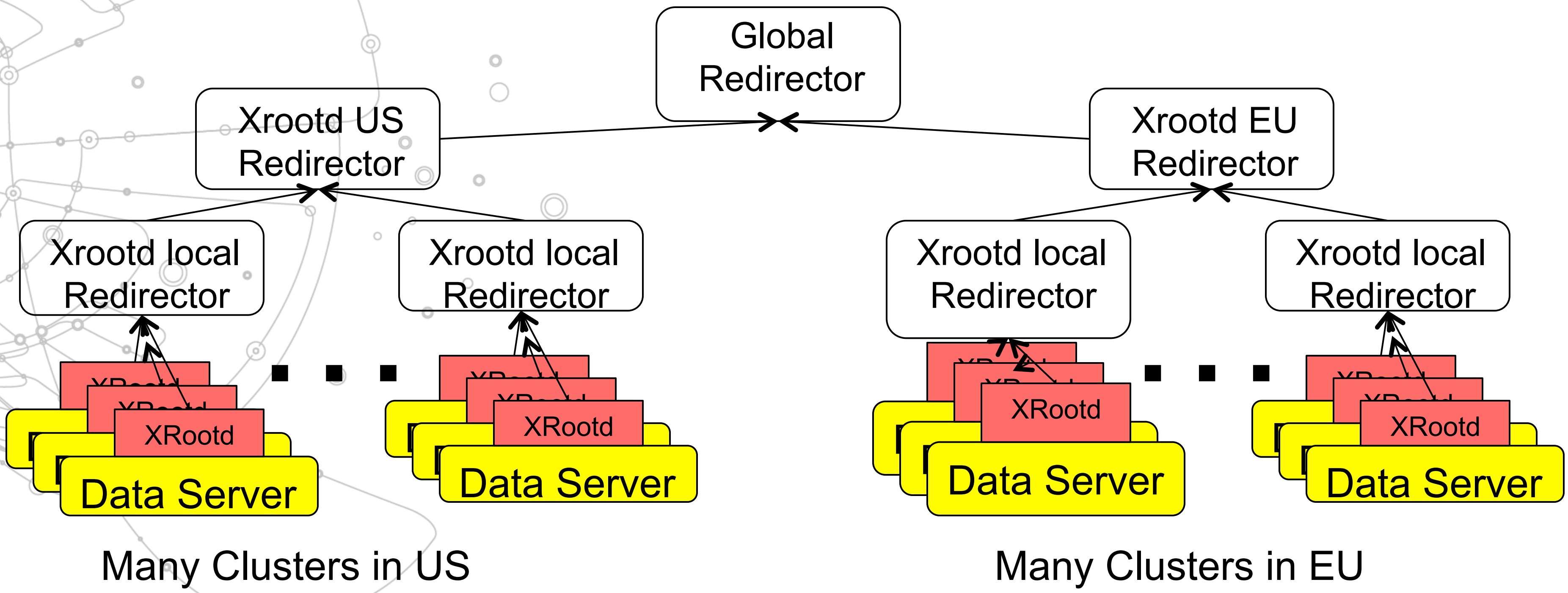
Electron collection: { Electron_1, … } a couple, if at all



- Each “basket” compressed separately => optimized to efficiently read:
  - partial event, e.g., only Tracks;
  - partial object, e.g., only Electron E_{EM} to decide if the event is interesting at all
- For CMS, cluster size on disk is 5 – 20 MB or 10 – 40 events
- Total file size from 100 MB to 10 GB

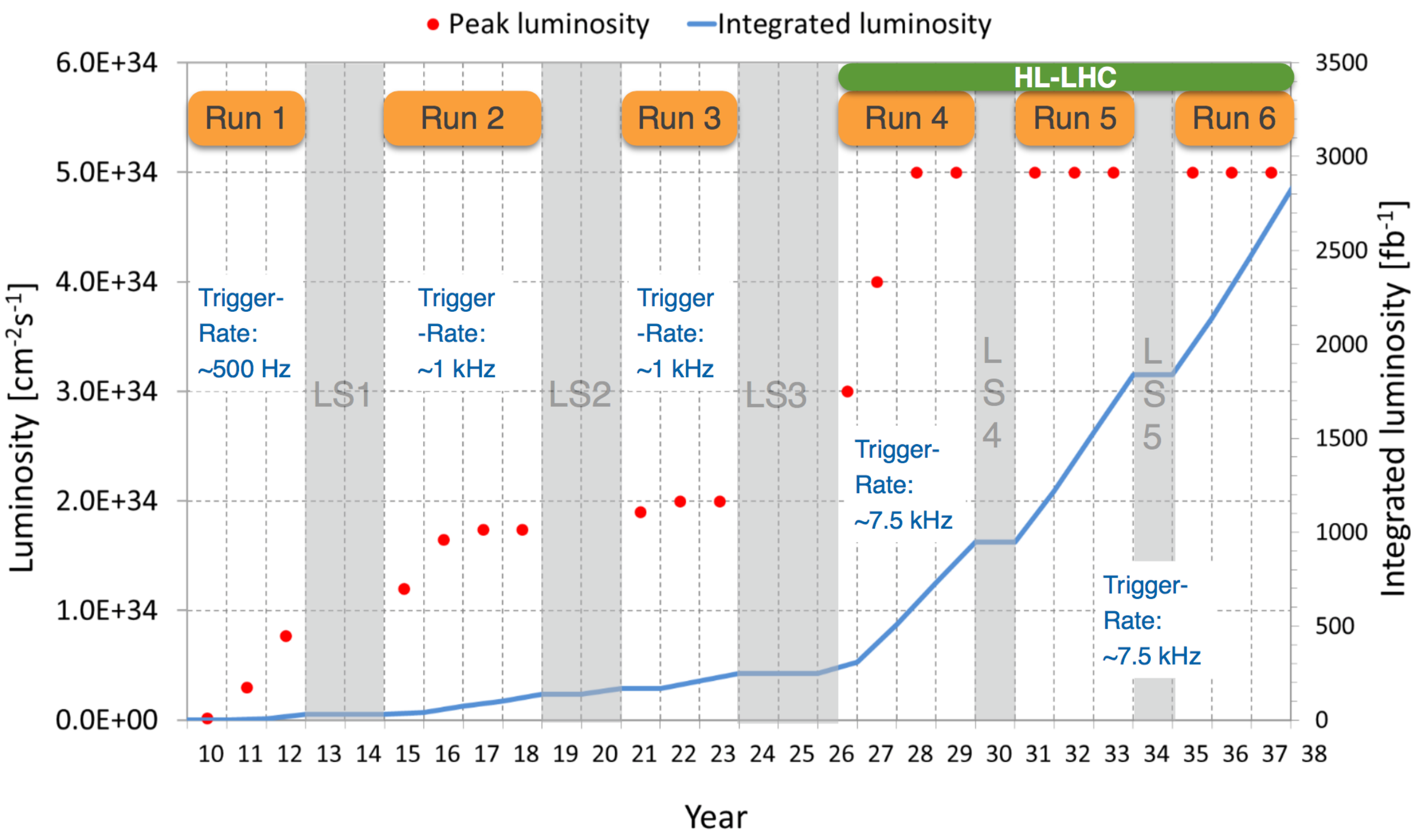
Data Federation for WAN Reads

Applications connect to local/regional redirector.
Redirect upwards only if file does not exist in tree below.
Minimizing WAN read access latency this way.



HEADING TOWARDS THE FUTURE

LHC Schedule

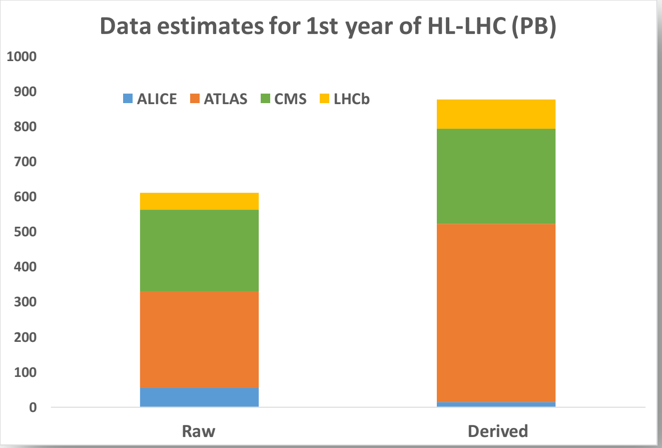


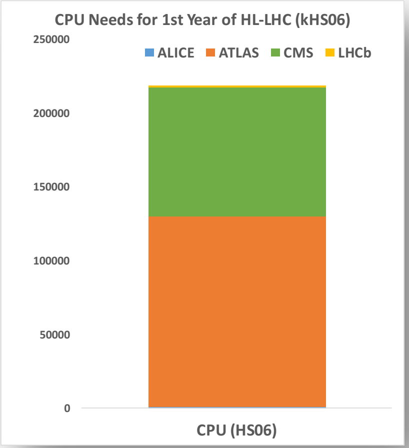
LHC Run3 and Run4 Scale and Challenges

2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 …… 2030?

First run | LS1 | Second run | LS2 | Third run | LS3 | HL-LHC | FCC?

- Raw data volume for LHC increases exponentially and with it processing and analysis load
- Technology at ~20%/year will bring x6-10 in 10-11 years
- Estimates of resource needs at HL-LHC x10 above what is realistic to expect from technology with reasonably **constant cost**





Data:
- Raw 2016: 50 PB → 2027: 600 PB
- Derived (1 copy): 2016: 80 PB → 2027: 900 PB

CPU:
- x60 from 2016

Technology revolutions are needed to close the resource gap

Courtesy of I. Bird

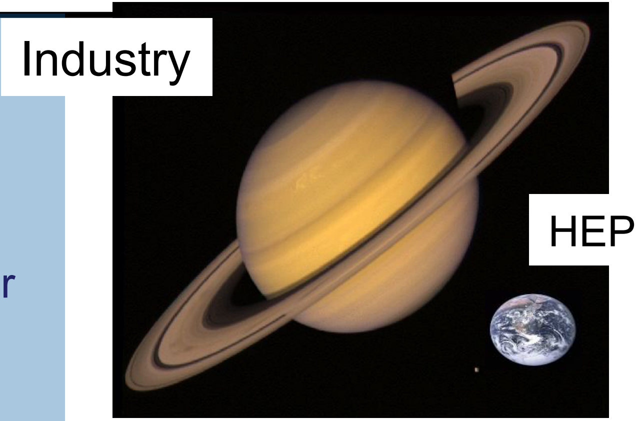
Relative Size of Things

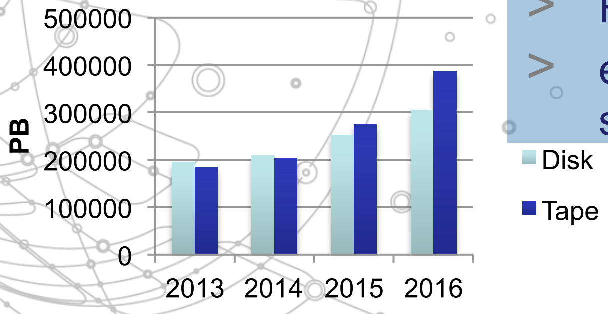
Processing

- Amazon has more than 40 million processor cores in EC2

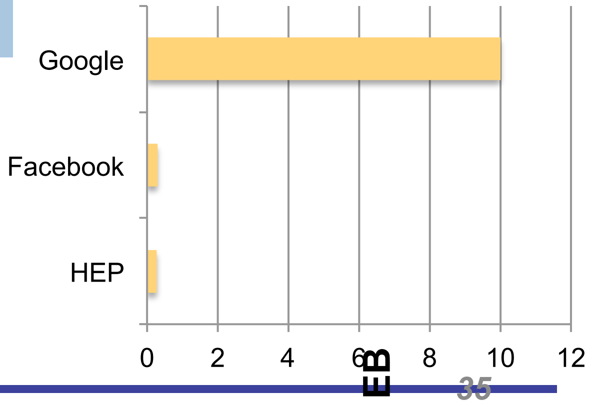
Storage

- Amazon has $2x10^{12}$ unique user objects and supports 2M queries per second
- Google has 10-15 exabytes under management
- Facebook 300PB
- eBay collected and accessed the same amount of data as LHC Run1



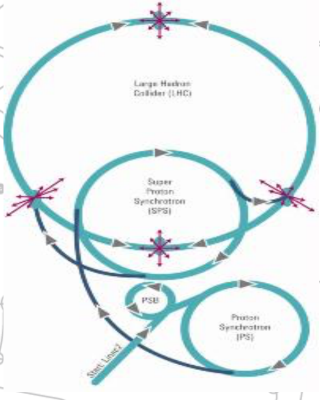


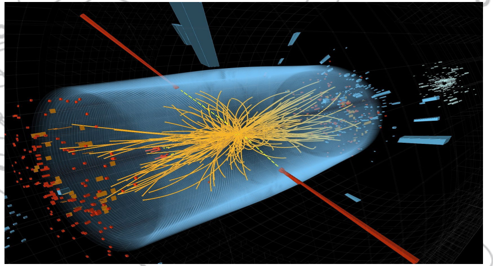
LHC Disk and Tape Storage



Our data and processing problems are ~1% the size of the largest industry problems

Data Analytics to the Rescue ?

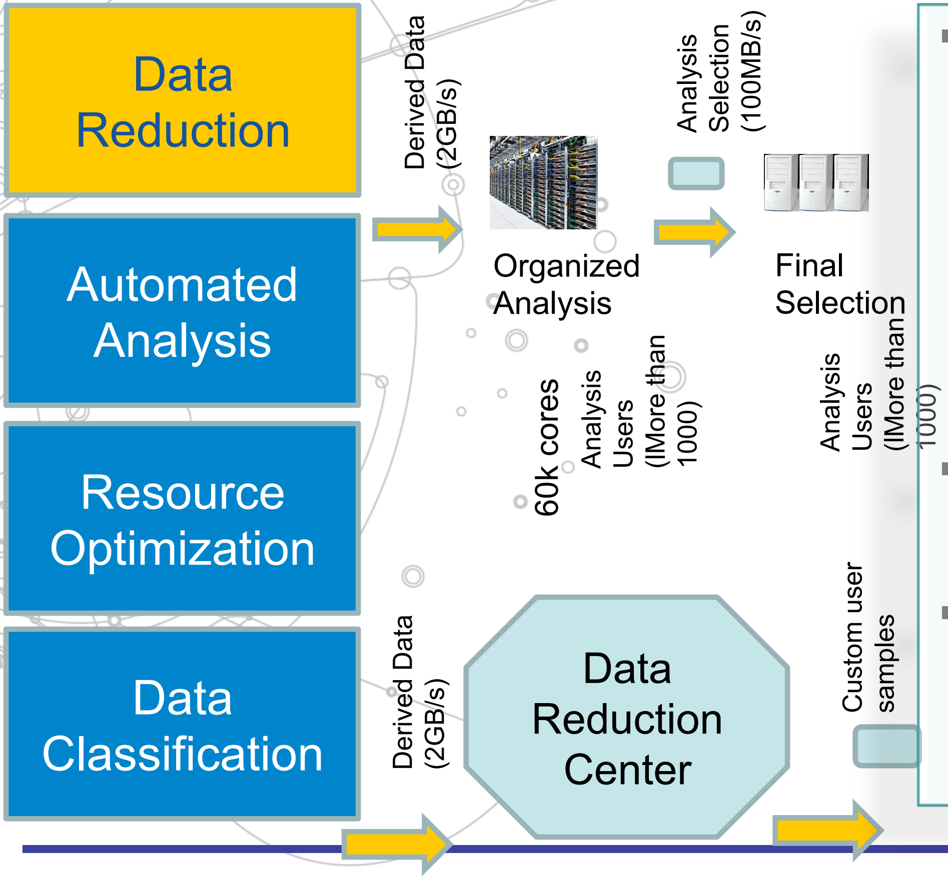




How to make more effective use of the data collected is critical to maximise scientific discovery and close the resource gap

- There are currently ongoing projects in
  - Accelerator system controls
  - Data handling and quality optimizations
  - Data reduction
  - Optimized formats
  - Investigations for machine learning for analysis and event categorization

Data Analytics



- The increase in data volume expected by the experiments in Run3 and Run4 changes the operations mode and opens up possibilities for analysis
  - In ALICE and LHCb events will leave the detector essentially reconstructed with final calibrations
    - Analysis can start immediately (maybe even online)
  - ATLAS and CMS will both have much higher triggers and a desire to streamline analysis
- Interest to look at industry tools for improved data analysis
  - SPARK, Hadoop, etc.
- Community building around analysis challenges, e.g.
  - https://www.kaggle.com/c/flavours-of-physics

Data Analytics

Data Reduction

Automated Analysis

Resource Optmisation

Data Classification

Reconstruction

20k cores

Operations (less than 50)

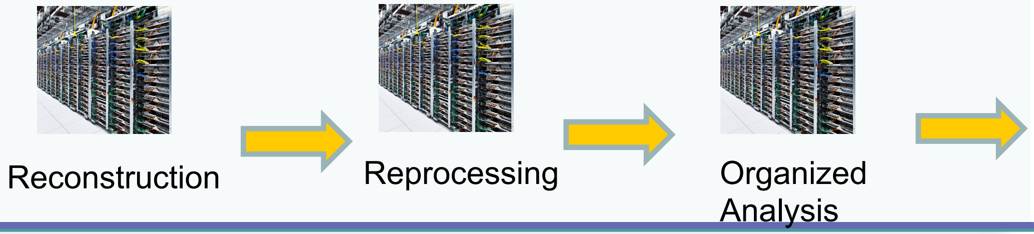
- The Data Quality Monitoring (DQM) is a key to delivering high-quality data for physics. It is used both in the online and offline environments
  - Currently involves scrutinizing of a large number of histograms by detector experts comparing them with a reference
  - Aim at applying recent progress in Machine Learning techniques to the automation of the DQM scrutiny
- The LHC is the largest piece of scientific apparatus ever built
  - There is a tremendous amount of real time monitoring information to assess health and diagnose faults
  - The volume and diversity of information makes this an interesting application of big data analytics

Data Analytics

- Data Reduction
- Automated Analysis
- Resource Optimization
- Data Classification

- Use machine learning techniques to predict how data should be placed and processing resources scheduled in order to achieve a dramatic reduction in latency for delivering data samples to analysts

- Design a system capable of using information about resource usage (disk access, CPU efficiency, job success rates, data transfer performances, and more) to make more automated decisions about resource allocation



Data Analytics

Data Reduction

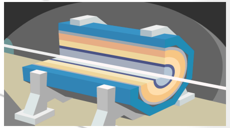
Automated Analysis

Resource Optimization

Data Classification

Data Volume

All Data (TB/s)



- Investigate the possibility of performing real-time event classification in the high-level trigger systems of LHC experiments
  - Extract information from events that would otherwise be rejected
- Uncategorized events might potentially be the most interesting, revealing the presence of new phenomena in the LHC data
  - Event classification would allow both a more efficient trigger design and an extension of the physics program, beyond the boundaries of the traditional trigger strategies

Engagement with Industry

- Industry has focused on improving data analytics techniques to solve large scale industrial problems and analyzing social media data
  - Has provided a common ground for developments and activities with industry
    - Leaders in the area of big data, machine learning and data analytics
    - High Energy Physics has long been a leader in data analysis and data handling, but industry has closed the gap in this area

Summary and Outlook

- The LHC today operates a globally distributed 500,000 X86 core infrastructure ingesting more than 100PBytes per year.
- The LHC is planning to dramatically increase the volume and complexity of data collected by Run3 and Run4
  - This results in an unprecedented computing challenge in the field of High Energy Physics
- Meeting this challenge within a realistic budget requires rethinking how we work
  - Need to shrink computing needs by roughly x10 over naïve extrapolations of what we do today.
  - Turning to industry and other sciences for improvements in data analytics
    - Data reduction and automated analysis through machine learning techniques need to be investigated