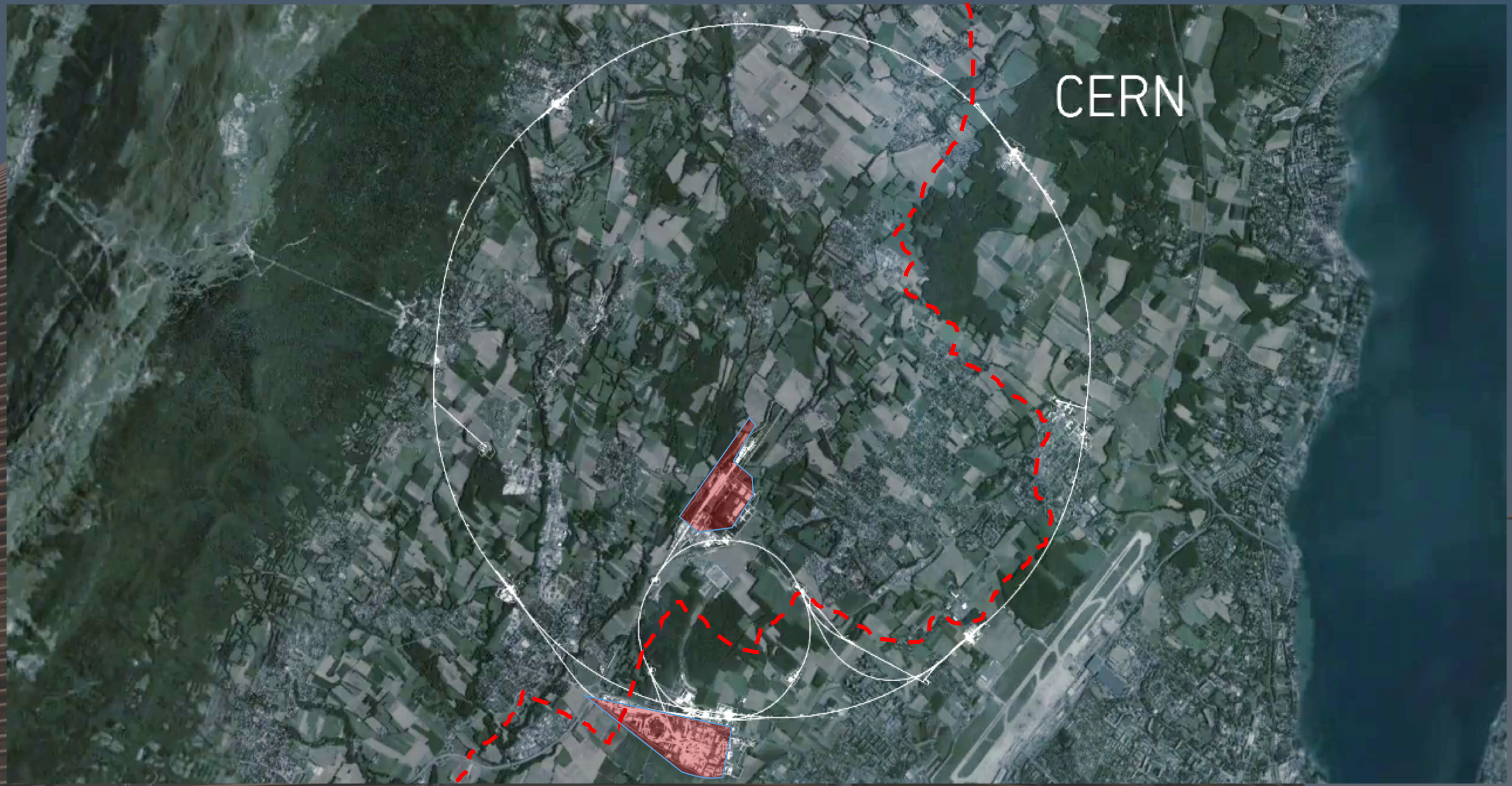
Computing Challenges at the Large Hadron Collider

Maria Girone
CERN openlab CTO

Outline

- Introduction to CERN
- Scale of the current system
- The upcoming LHC Upgrade
- New resource needs
- Explorations of HPC

Maria Girone
CERN openlab CTO



CERN is the European Laboratory for Particle Physics and straddles the Franco-Swiss border near Geneva.

Maria Girone
CERN openlab CTO

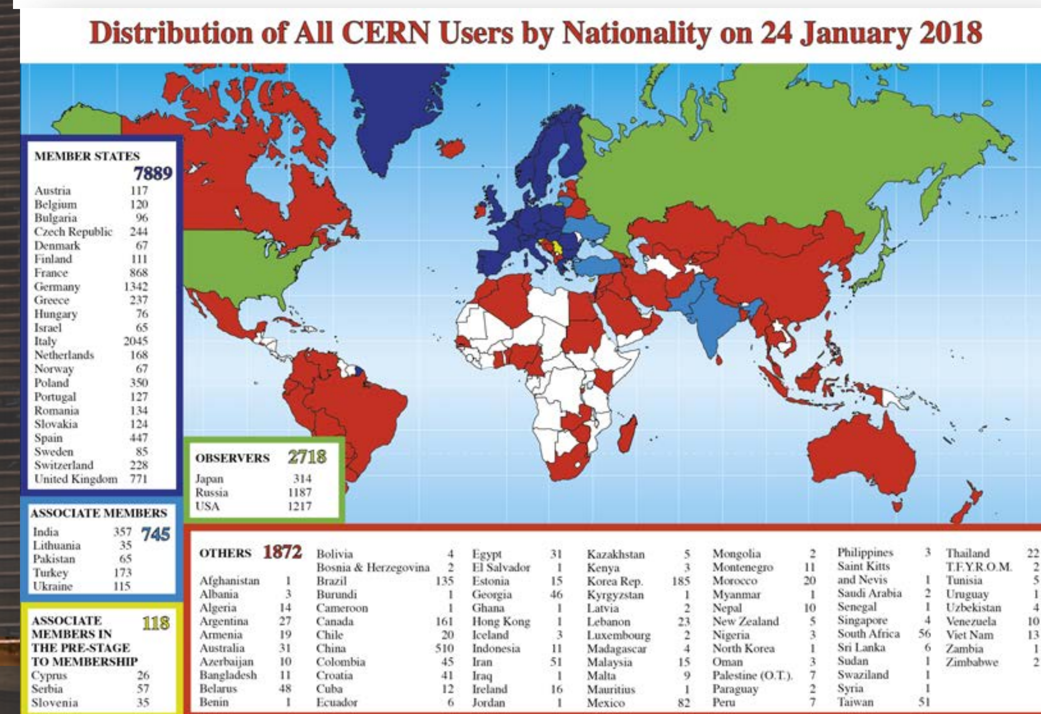
Member states

Associate member states in the pre-stage to membership

Associate member states

Observers

A world-wide endevour



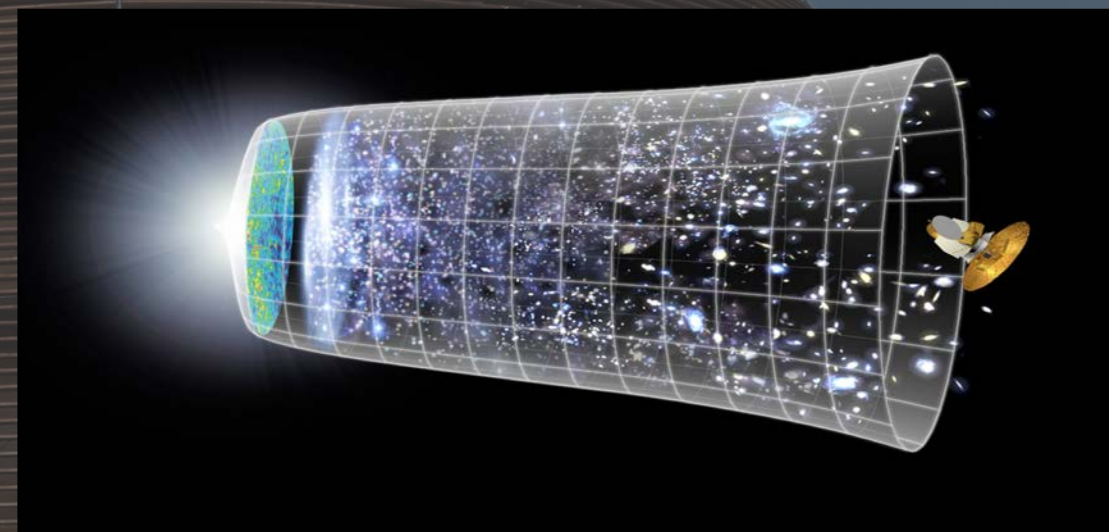
It has 22 member states and supports a global community of 15,000 researchers.

Maria Girone
CERN openlab CTO

Looking for Antimatter

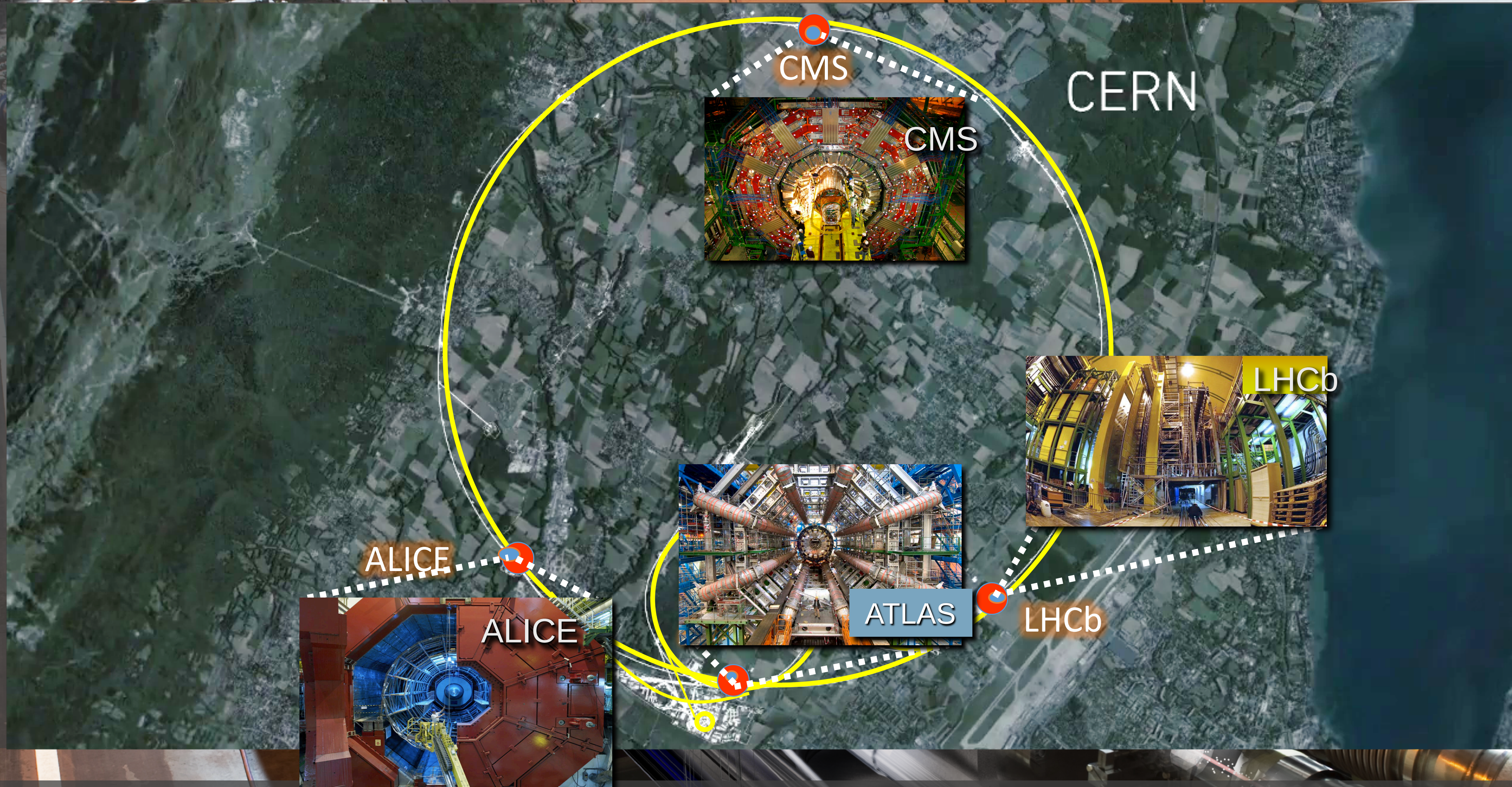
Understanding the very first moments of our Universe after the Big Bang

Understanding Dark Matter



Fundamental and cross domain research questions

Maria Girone
CERN openlab CTO



The Large Hadron Collider (LHC) is 27km circumference near Geneva and has 4 main detectors around the ring.

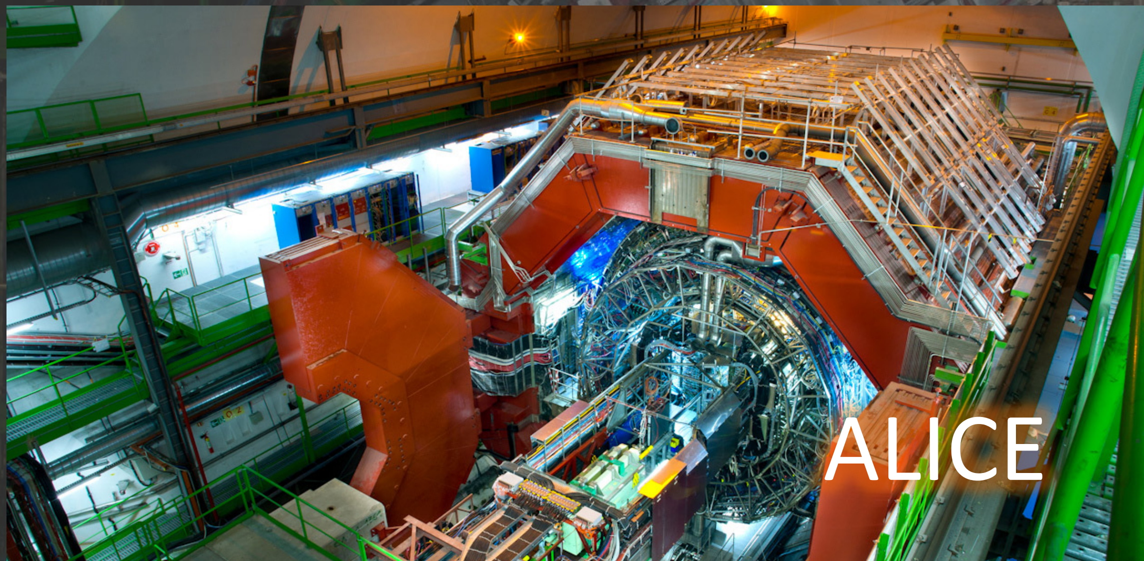
Maria Girone
CERN openlab CTO

46m long, 25m diameter
weights 7’000 tonnes
100 million electronic channels , 3 000 km of cables

22m long, 15m diameter
weights 14’000 tonnes
Most powerful superconducting solenoid ever built

Two general-purpose detectors cross-confirm discoveries, such as the Higgs boson.

Maria Girone
CERN openlab CTO

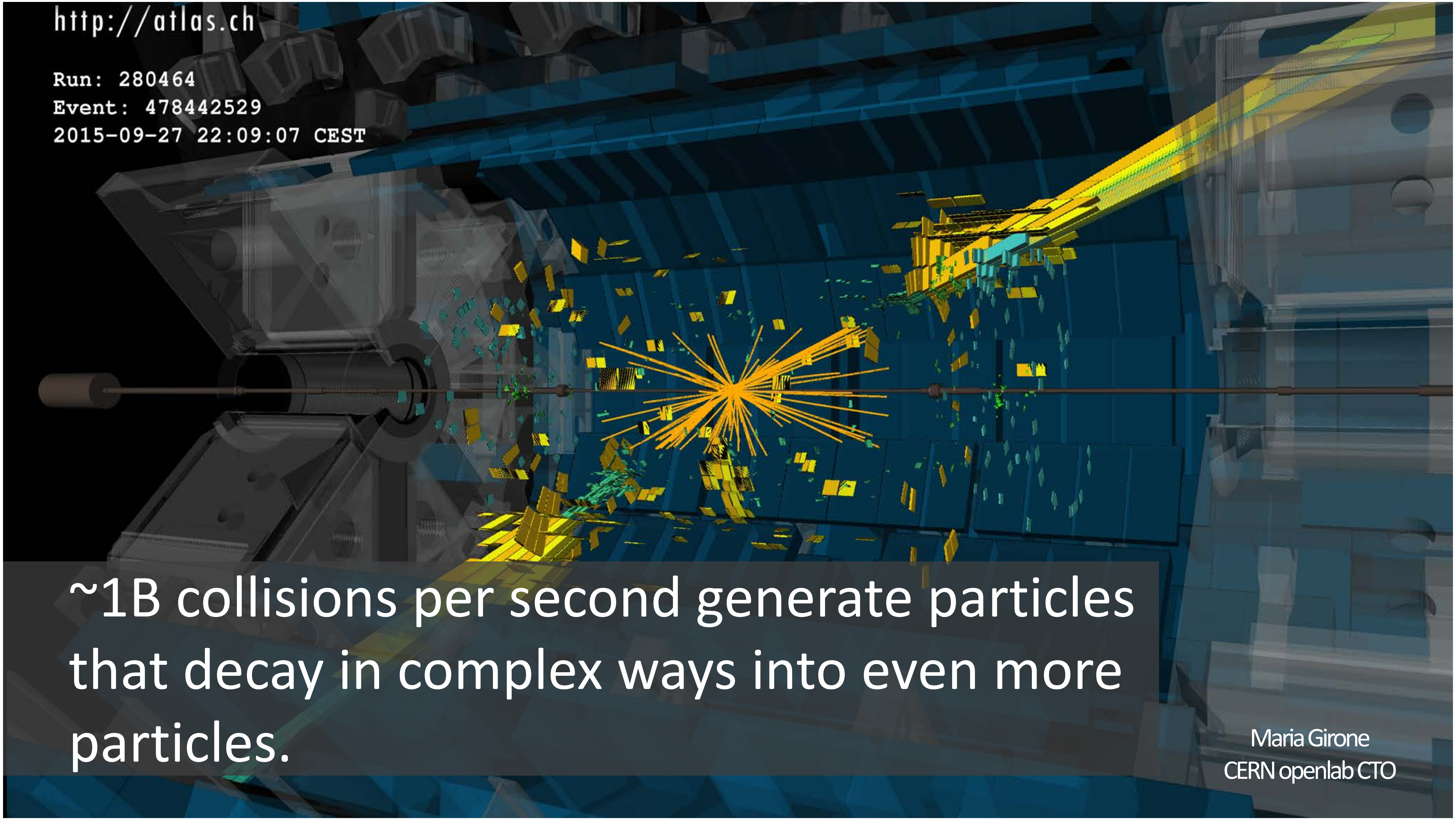


Studies the «Quark Gluon Plasma», state of matter which existed moments after the Big Bang.

Studies the behaviour difference between the b quark and the anti-b quark to explain the matter-antimatter asymmetry in the Universe.

ALICE and LHCb experiments have detectors specialised on studying specific phenomena.

Maria Girone
CERN openlab CTO



~1B collisions per second generate particles that decay in complex ways into even more particles.

Maria Girone
CERN openlab CTO

http://atlas.ch

Data generated 40 million times per second

PB/s

100,000 selections per second

TB/s

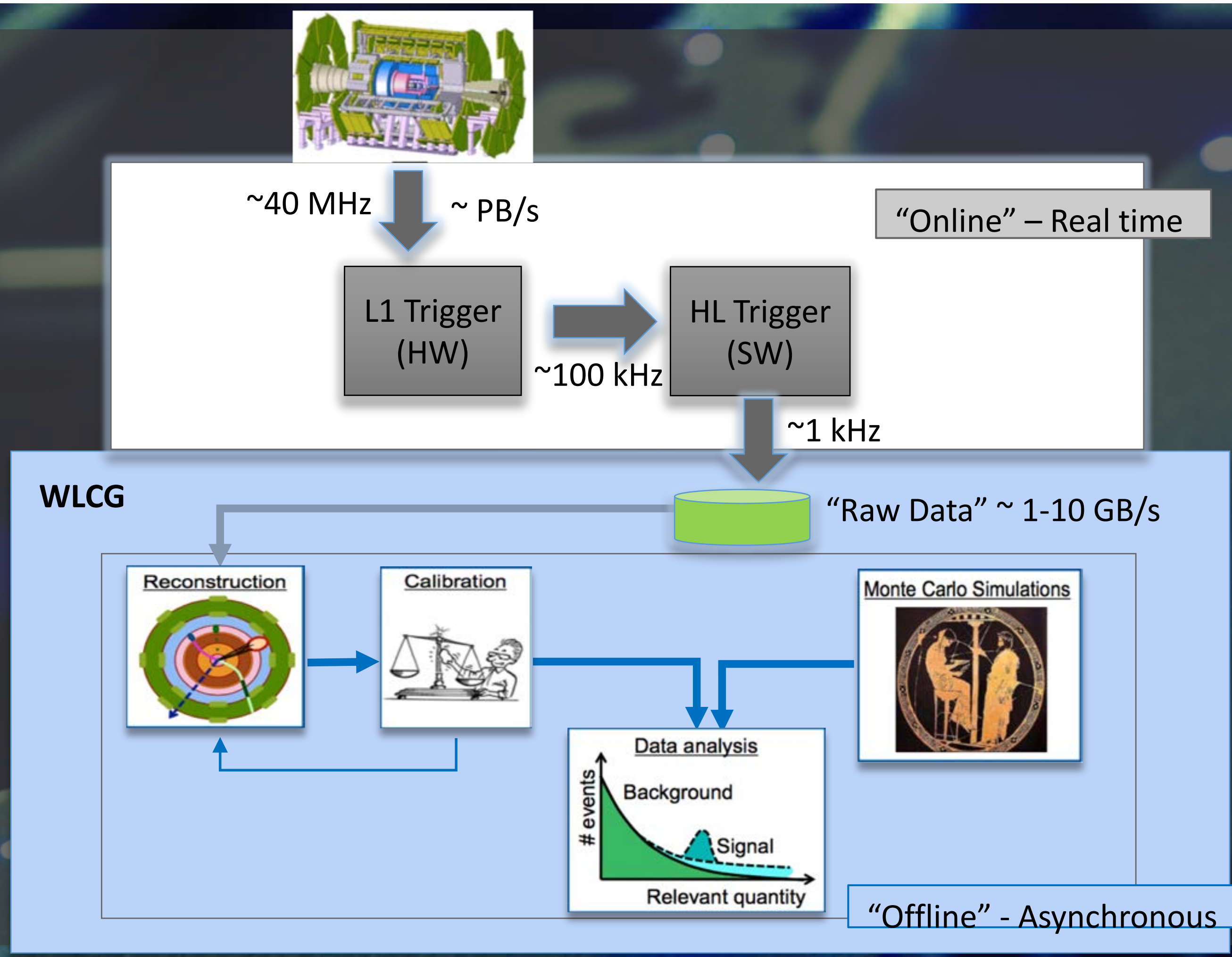
1,000 selections per second

GB/s

This can generate up to a petabyte of data per second.

Filtering the data in real time, selecting potentially interesting events (trigger).

Maria Girone
CERN openlab CTO

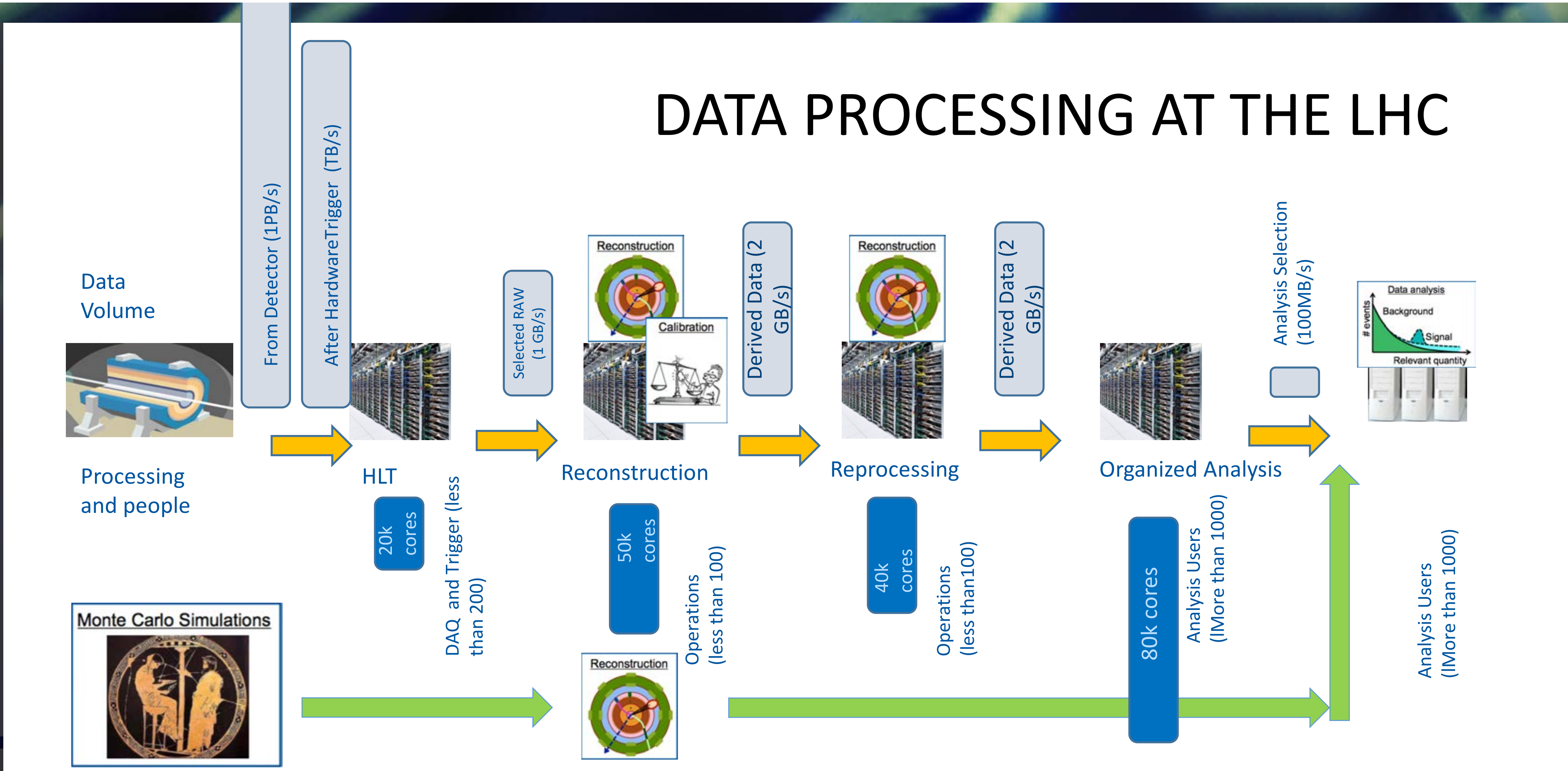


Simulation and data reconstruction are similar in their use of processing and storage resources

Physicists must sift through the 30-50 PBs produced annually by the LHC experiments.

Maria Girone
CERN openlab CTO

DATA PROCESSING AT THE LHC



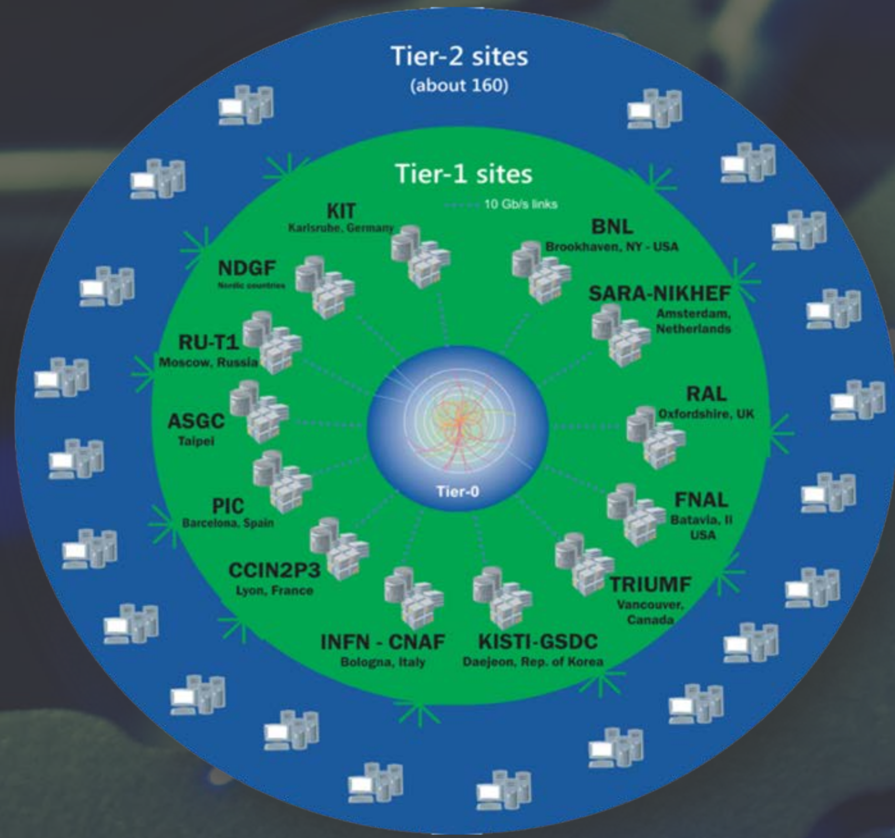
The process to transform raw data into useful physics datasets is a complicated series of steps at the LHC (Run2).

Maria Girone
CERN openlab CTO

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

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

Tier-2: Simulation, end-user analysis



1 PB/s of data generated by the detectors
Up to **60 PB/year** of stored data

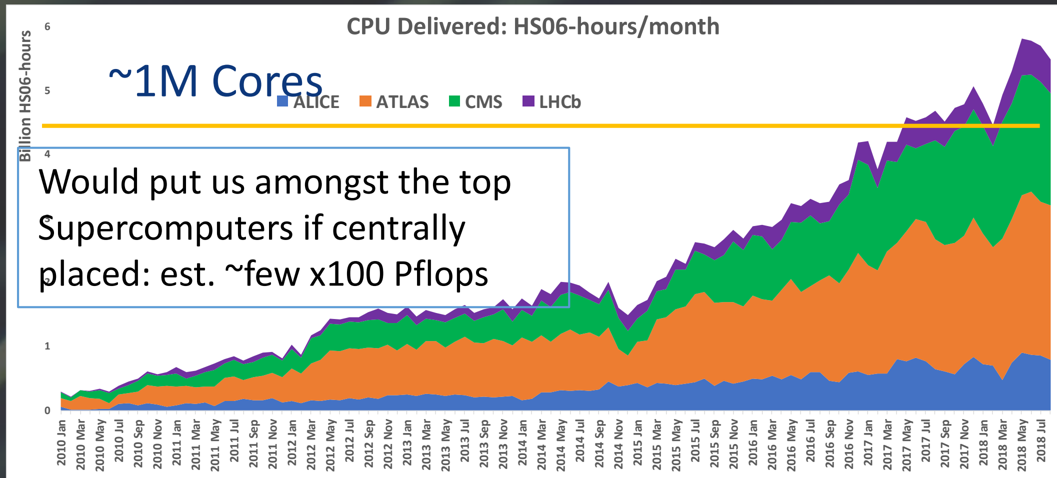
Large experiments have *managed* data sets of **>200 PB**

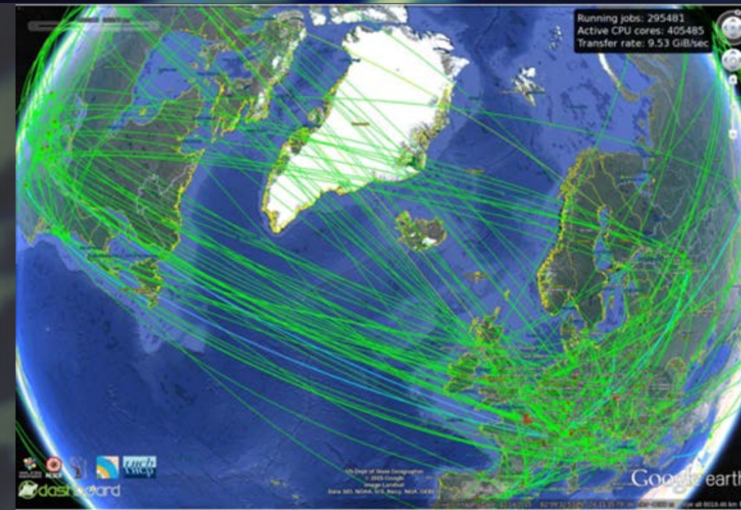
A distributed computing infrastructure of order of a **million cores** working 24/7
An average of 60M jobs/month

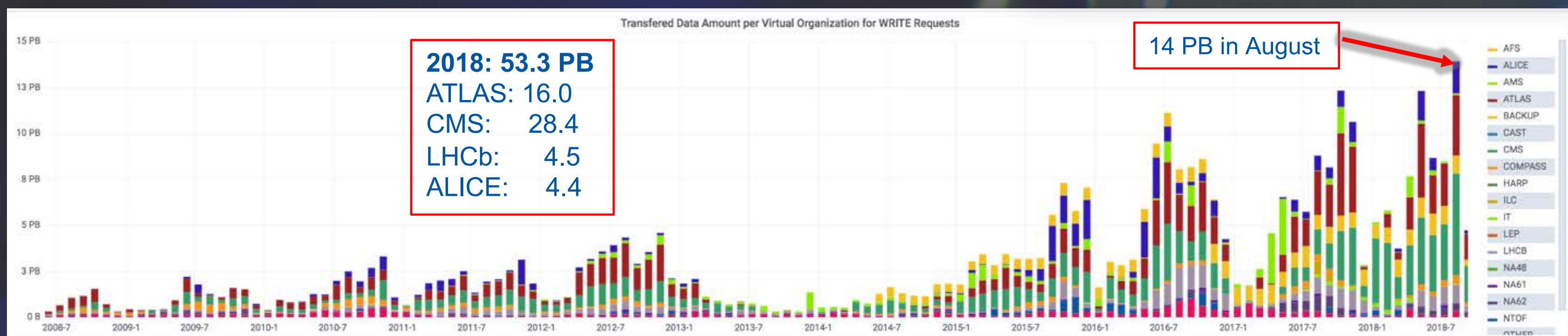
An continuous data transfer rate of 35-45 GB/s
(**3 PB/day**) across the Worldwide LHC Grid (WLCG)

Computing is handled by the Worldwide LHC Computing Grid, a hierarchical distributed computing infrastructure

Maria Girone
CERN openlab CTO



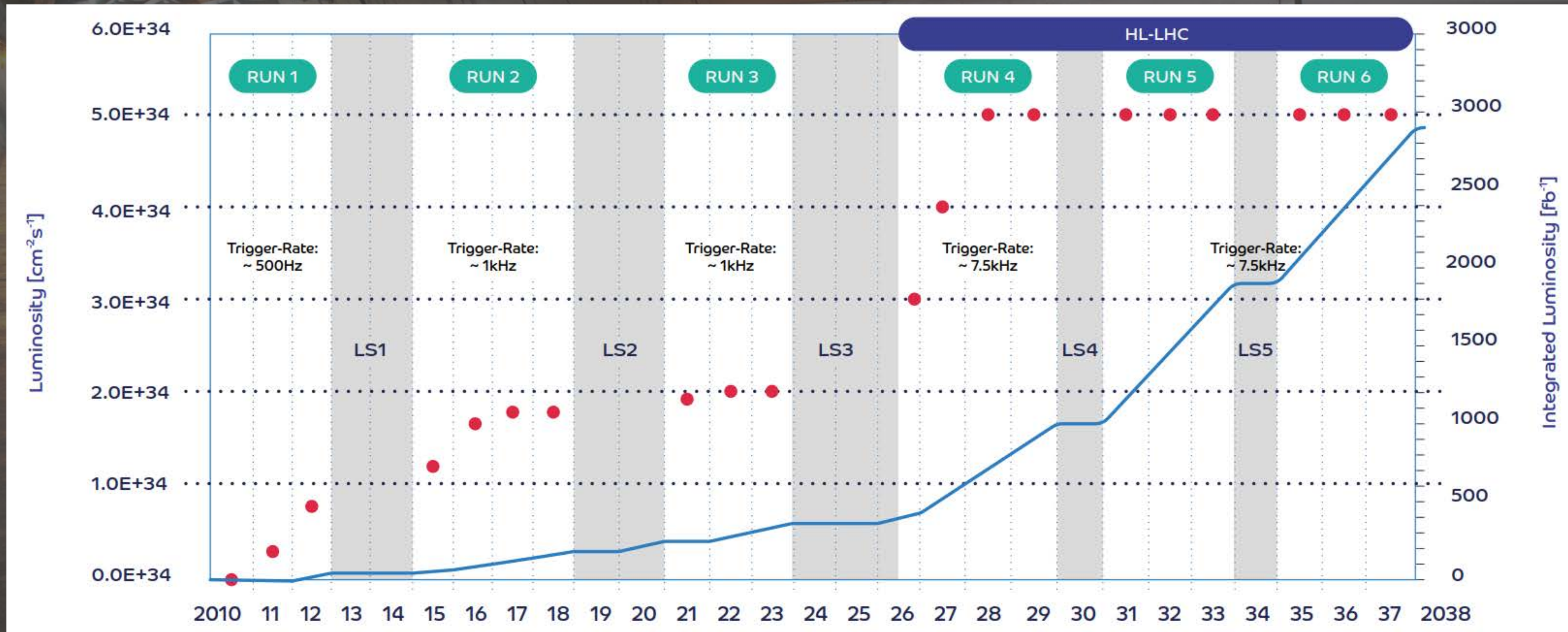




- CPU:
  - ~ 1 million cores fully occupied ("x86")
- Storage
  - ~ 1 EB (~500 PB disk, >500 PB tape)
- Network
  - 100Gb/s links between major centers and 10Gb/s between smaller sites

170 computing centres in 42 countries
Hundreds of PBs and more than 1M cores

Maria Girone
CERN openlab CTO



RUN 3

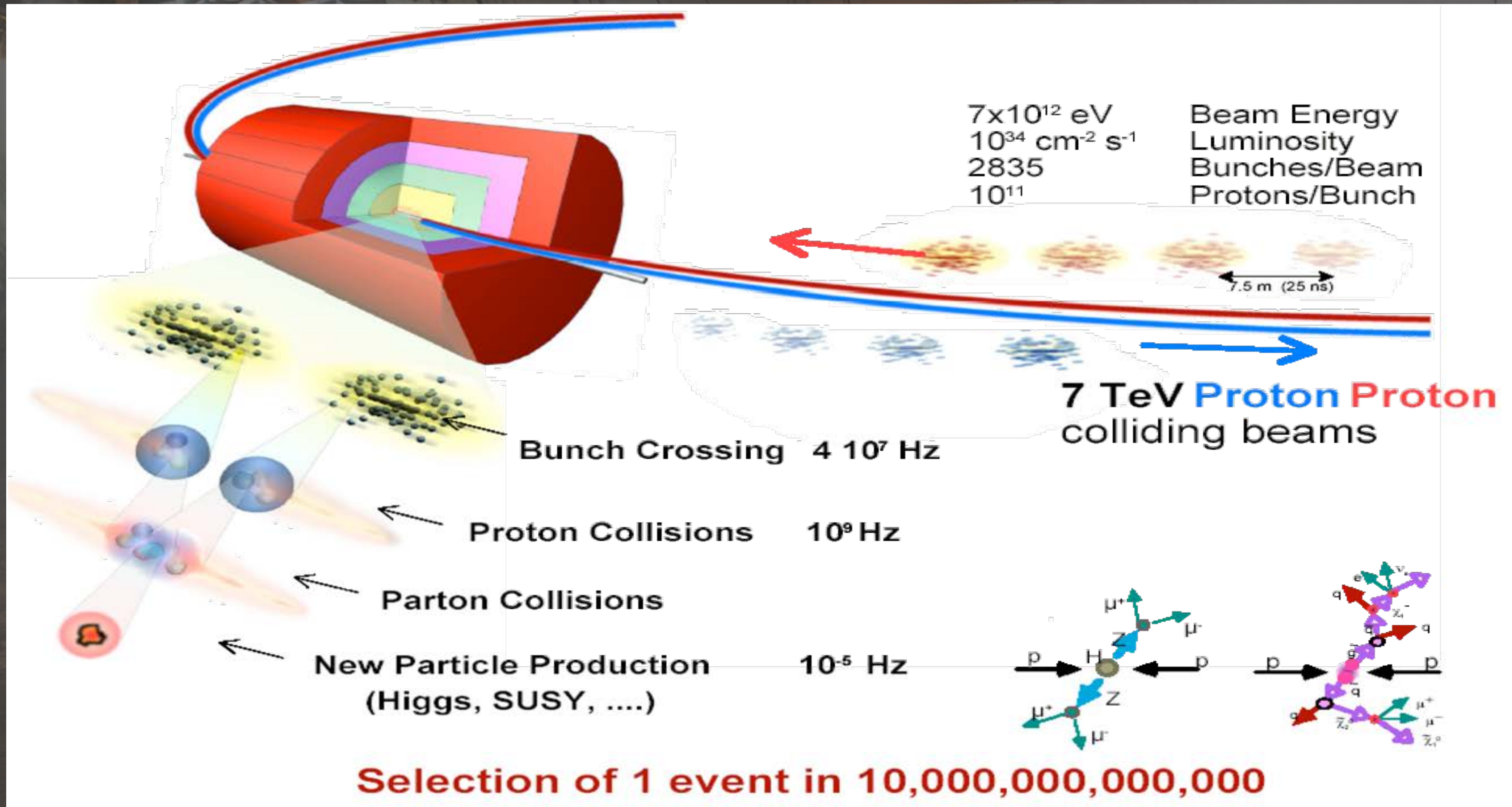
ALICE & LHCb upgrades

RUN 4

ATLAS & CMS upgrades

The LHC has been designed to follow a carefully set out programme of upgrades that will greatly increase the scientific reach.

Maria Girone
CERN openlab CTO



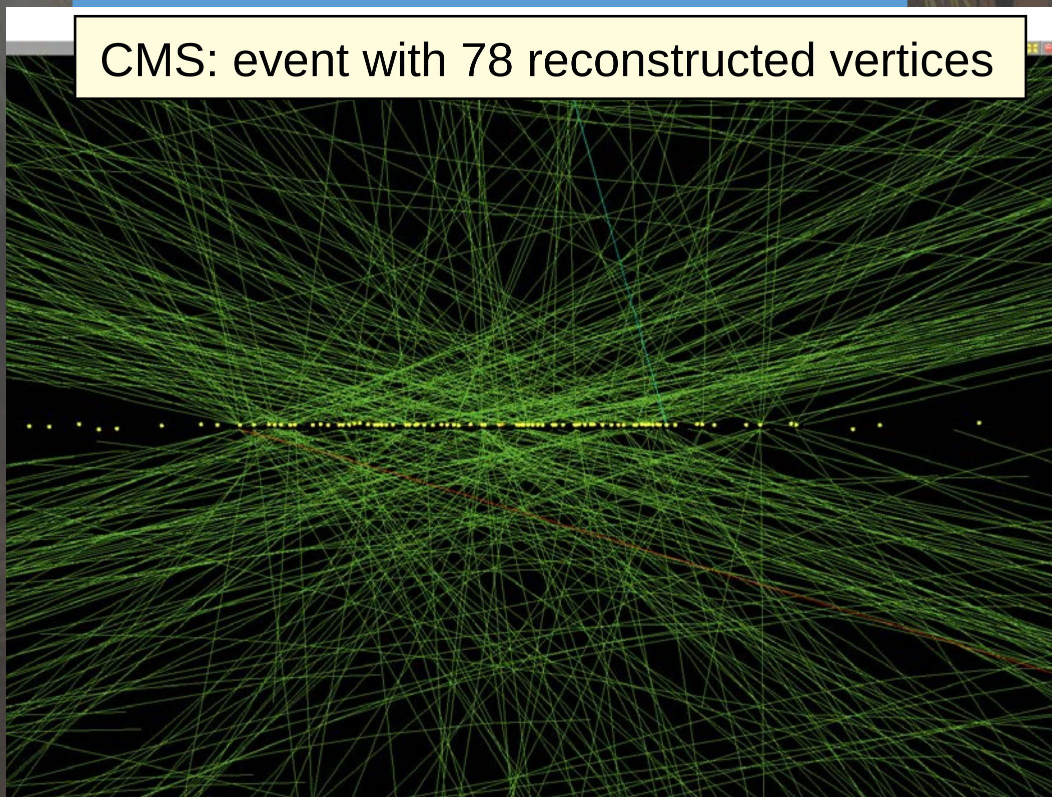
Rate of new physics is 1 event in 10^{12}

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

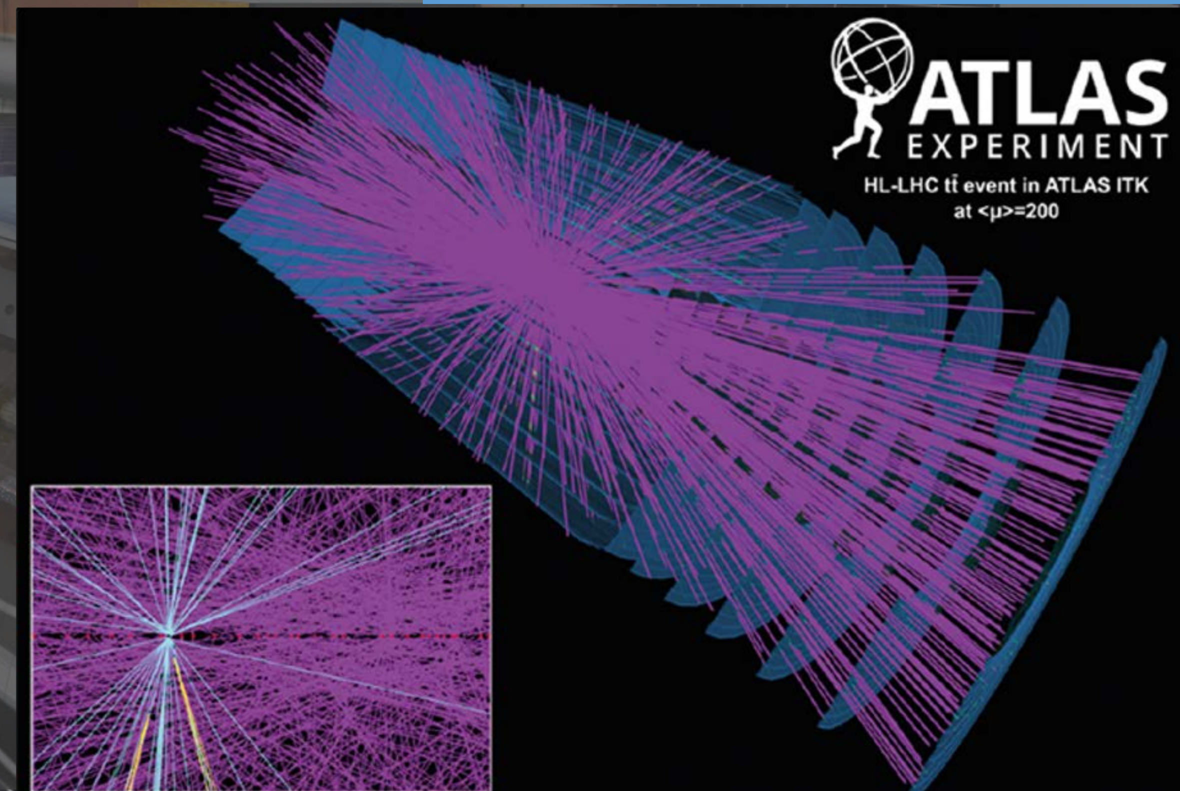
More collisions help physicists to observe rare processes and study with greater precision.

Maria Girone
CERN openlab CTO

CMS: event from 2017 with 78 reconstructed vertices

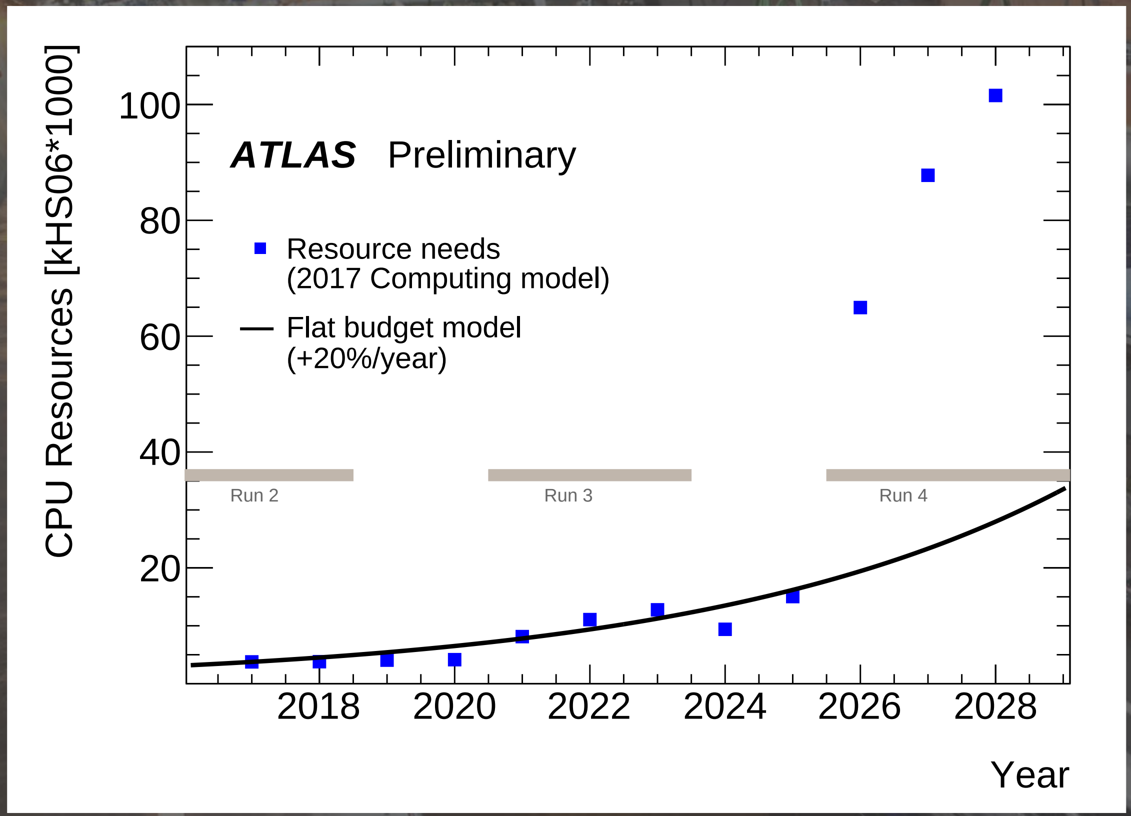


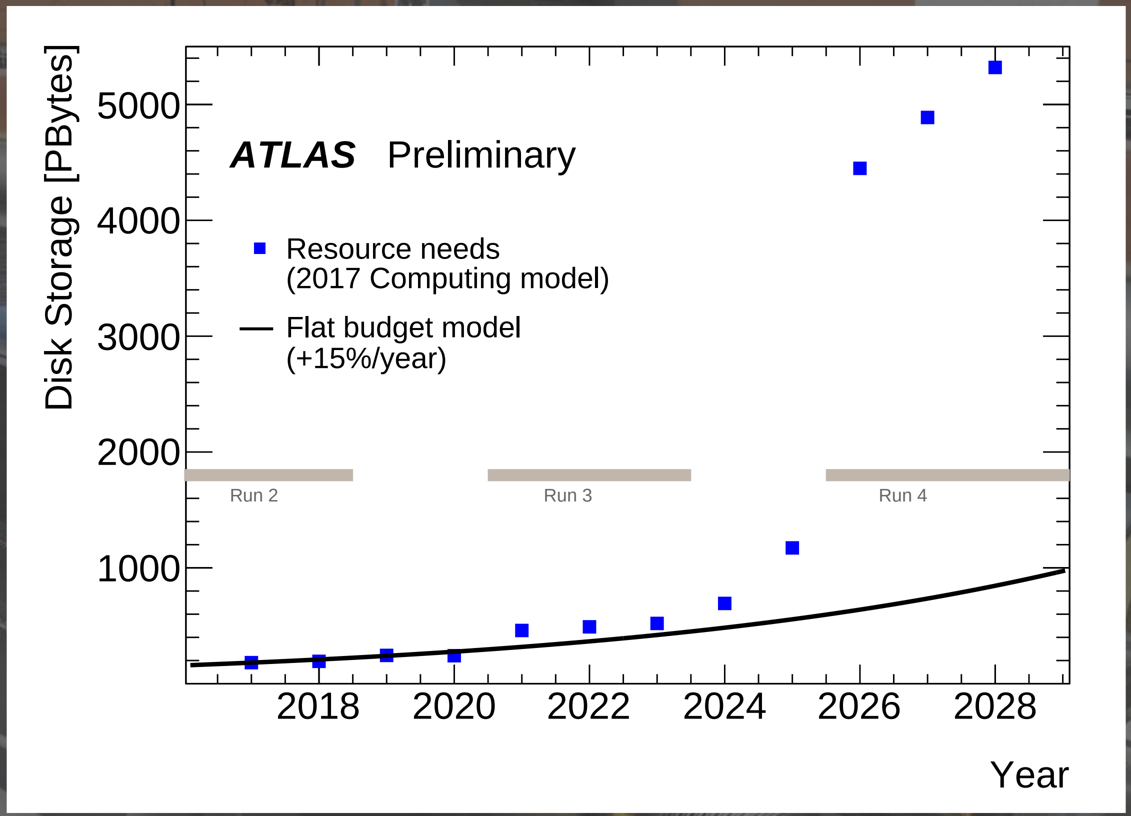
ATLAS: simulation for HL-LHC with 200 vertices



The HL-LHC will come online around 2026.
More collisions and more complex data.

Maria Girone
CERN openlab CTO

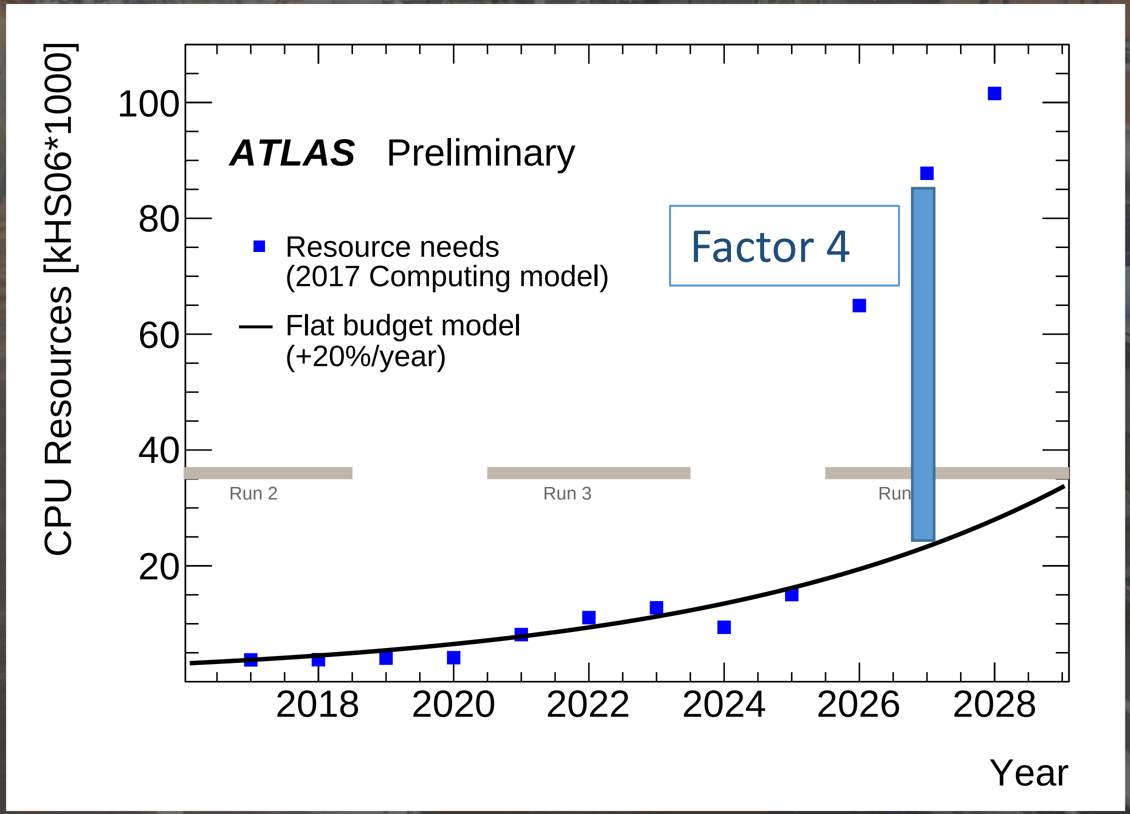


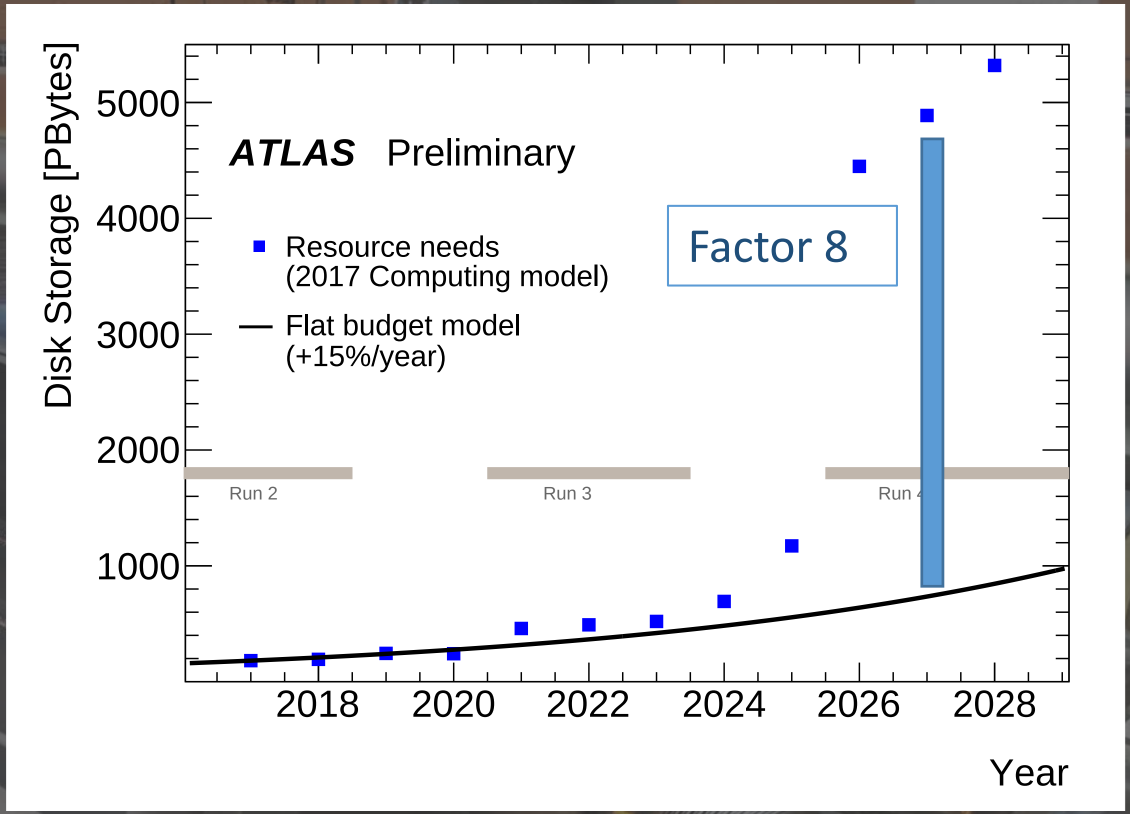


Using current techniques, required computing capacity increases 50-100 times

Data storage needs are expected to be in the order of Exabytes by this time.

Maria Girone
CERN openlab CTO





We have to overcome a factor of 4 to 8 even after technology improvements

- We need innovation and new techniques

Maria Girone
CERN openlab CTO

Three main areas of research and development.

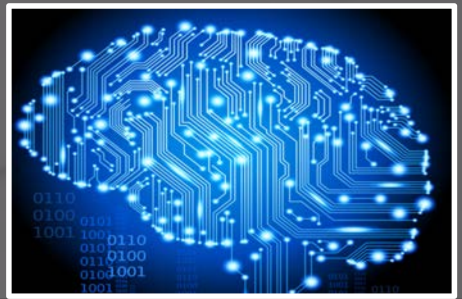
High Luminosity LHC

COMPUTING CHALLENGES

Scale out capacity with public clouds, HPC, new architectures

Increase **data centre performance** with hardware accelerators (FPGAs, GPUs, ..) optimized software

New techniques with Machine Learning, Deep Learning, Advanced Data Analytics



To close the resource gap we are improving our resource utilization, working on new techniques and finding new resources (Clouds, HPC, new architectures)

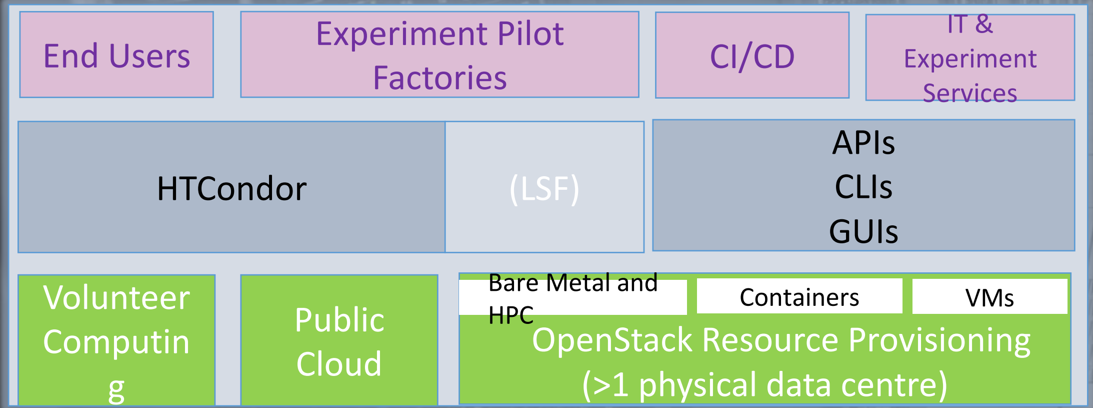
Maria Girone
CERN openlab CTO

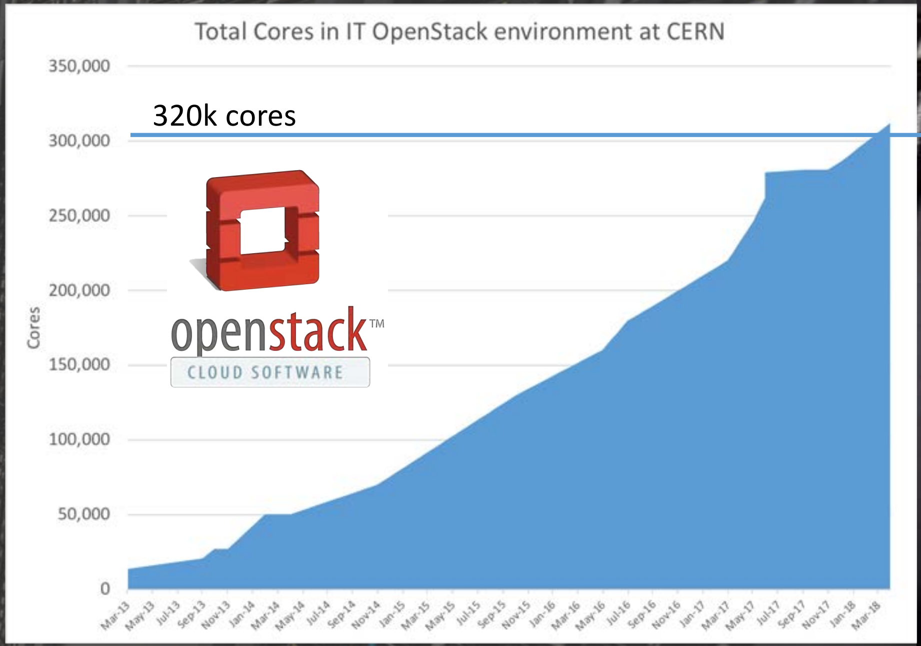
CERN is one of the early adopters and largest contributors to OpenStack

- 90% of the resources are provided through a private cloud
- Allows for flexible and dynamic deployment

Moving to containers for even more flexibility

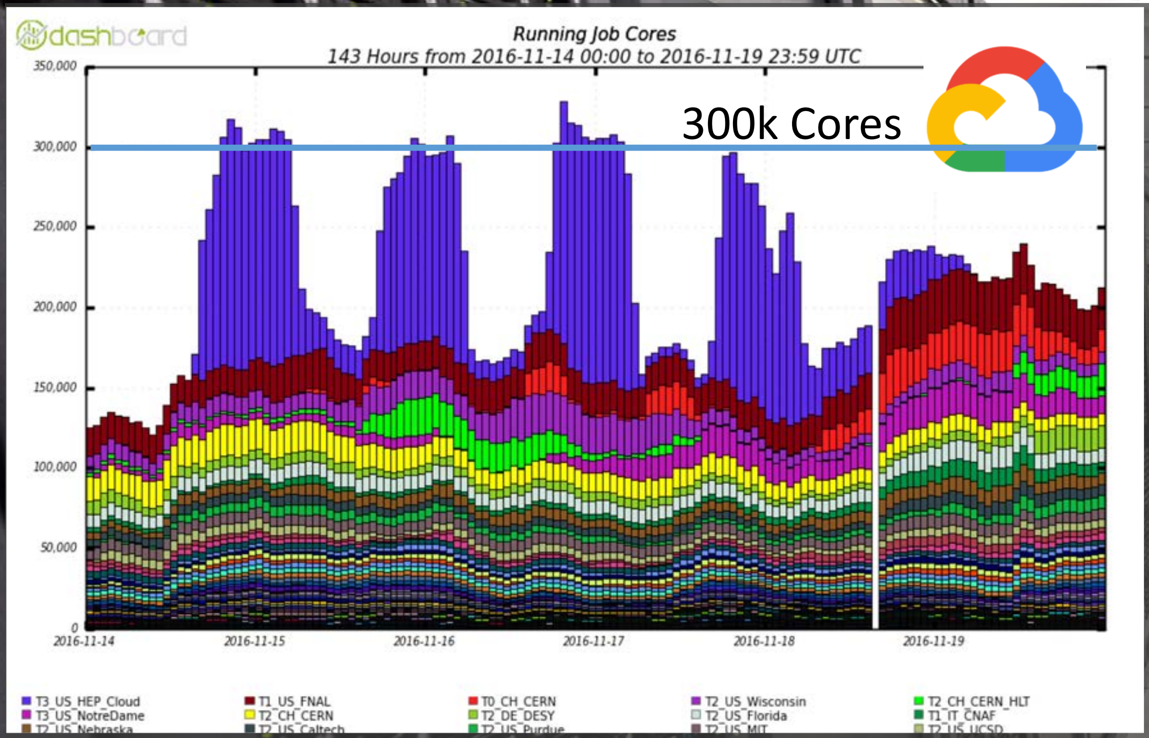
- Current investigations within CERN openlab



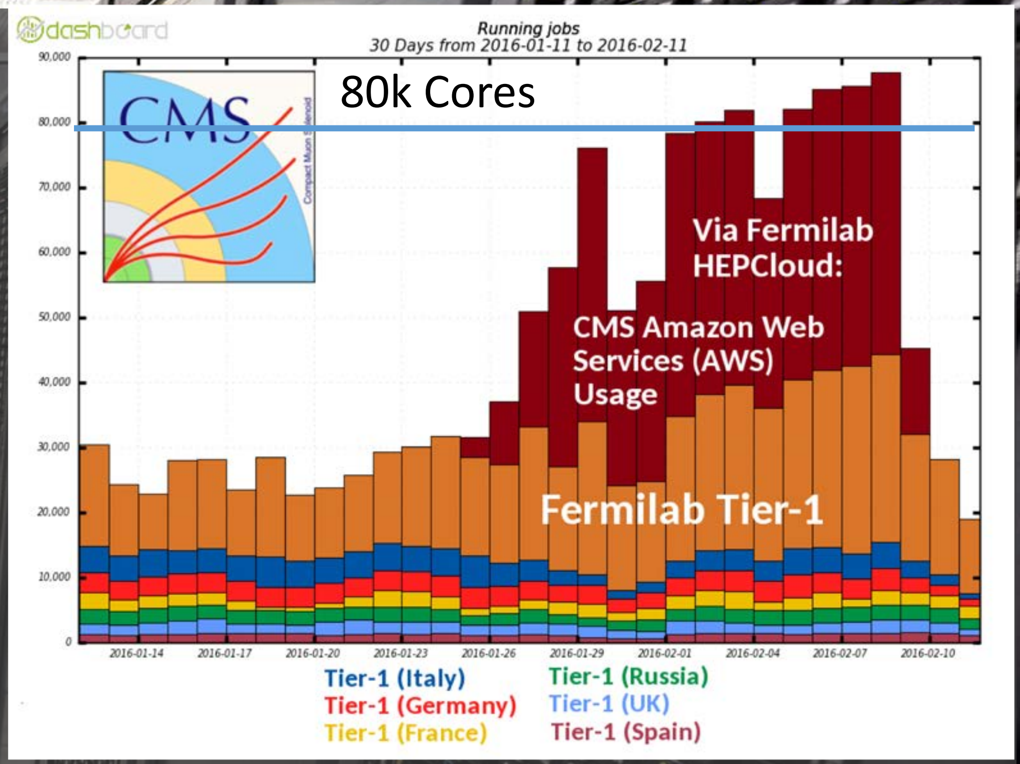


Layered, virtualized services provide flexibility and efficiency.

Maria Girone
CERN openlab CTO



Fermilab



Experiments have demonstrated that it is possible to elastically and dynamically expand production resources to commercial clouds.

Large-scale tests with commercial clouds.

There is a big push in the community to utilize HPC

- There is a substantial public investment
- Optimized for tightly coupled calculations
- Each resource is huge but independent; authorization, access, and interfaces are all specific to the site
- HPC centers are often early adopters of powerful new architectures

HPC Centers are huge computational resources.

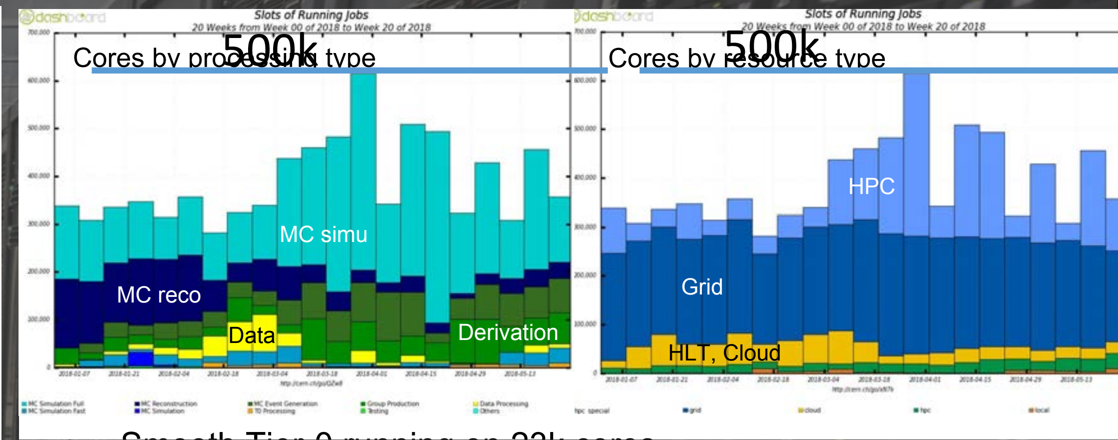
Maria Girone
CERN openlab CTO

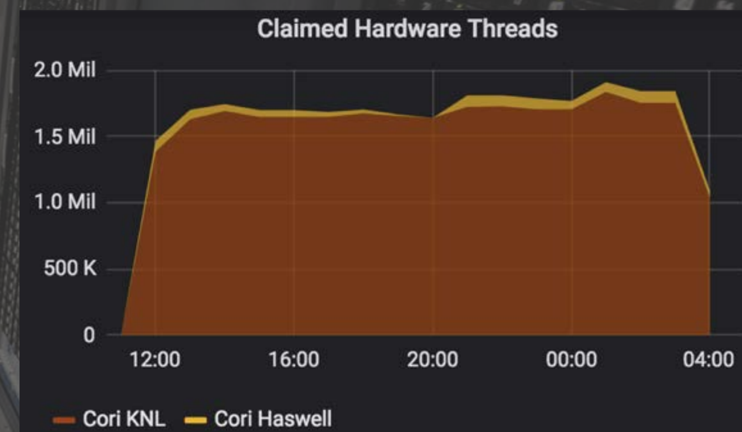
ATLAS reached more than 200k traditional x86 HPC cores for simulation workflows

The core count reaches to similar numbers to our global infrastructure

- HPC cores are not always equivalent to our commodity clusters
- Our resource efficiency on HPC is currently low

All experiments are exploring the use of heterogeneous HPC architectures



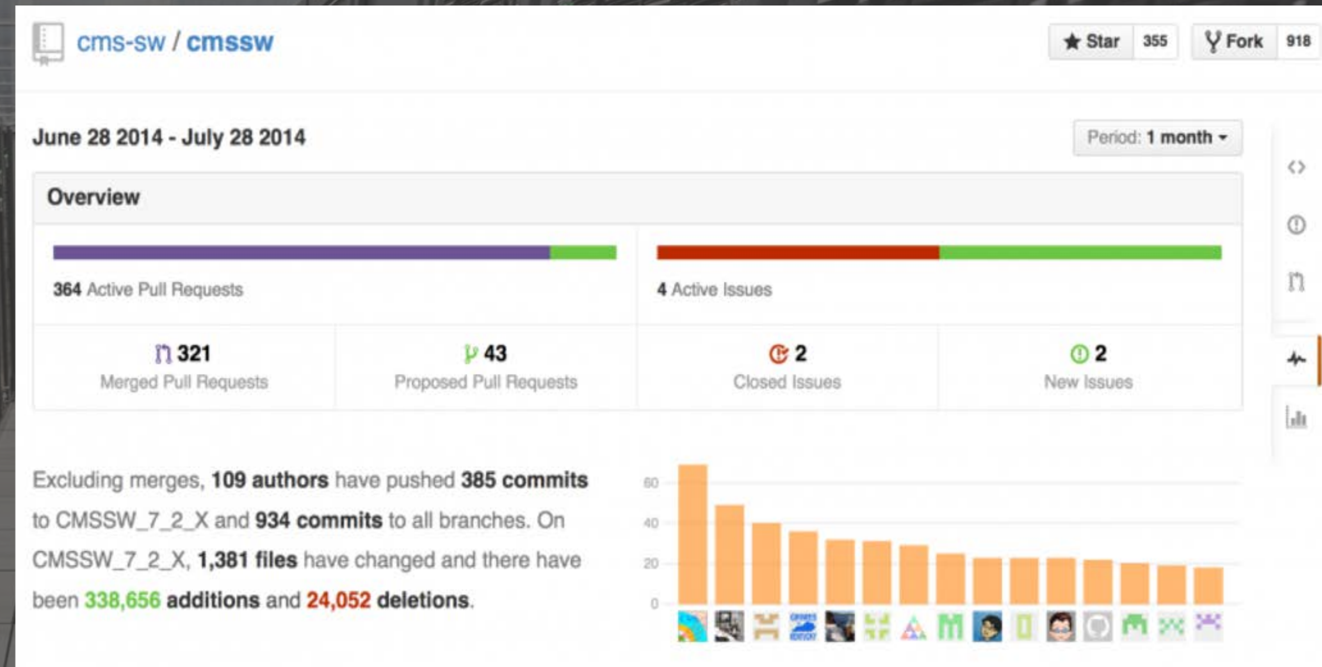


Demonstrations with large-scale, dedicated HPC resources, too, including EU HPC initiatives.

Maria Girone
CERN openlab CTO

The number of active programmers in High Energy Physics collaborations is a strength in general but a weakness for developing optimal code for HPC

- The software framework for each of the experiments is several million lines of C++ and Python code and have contributions from nearly 1000 people

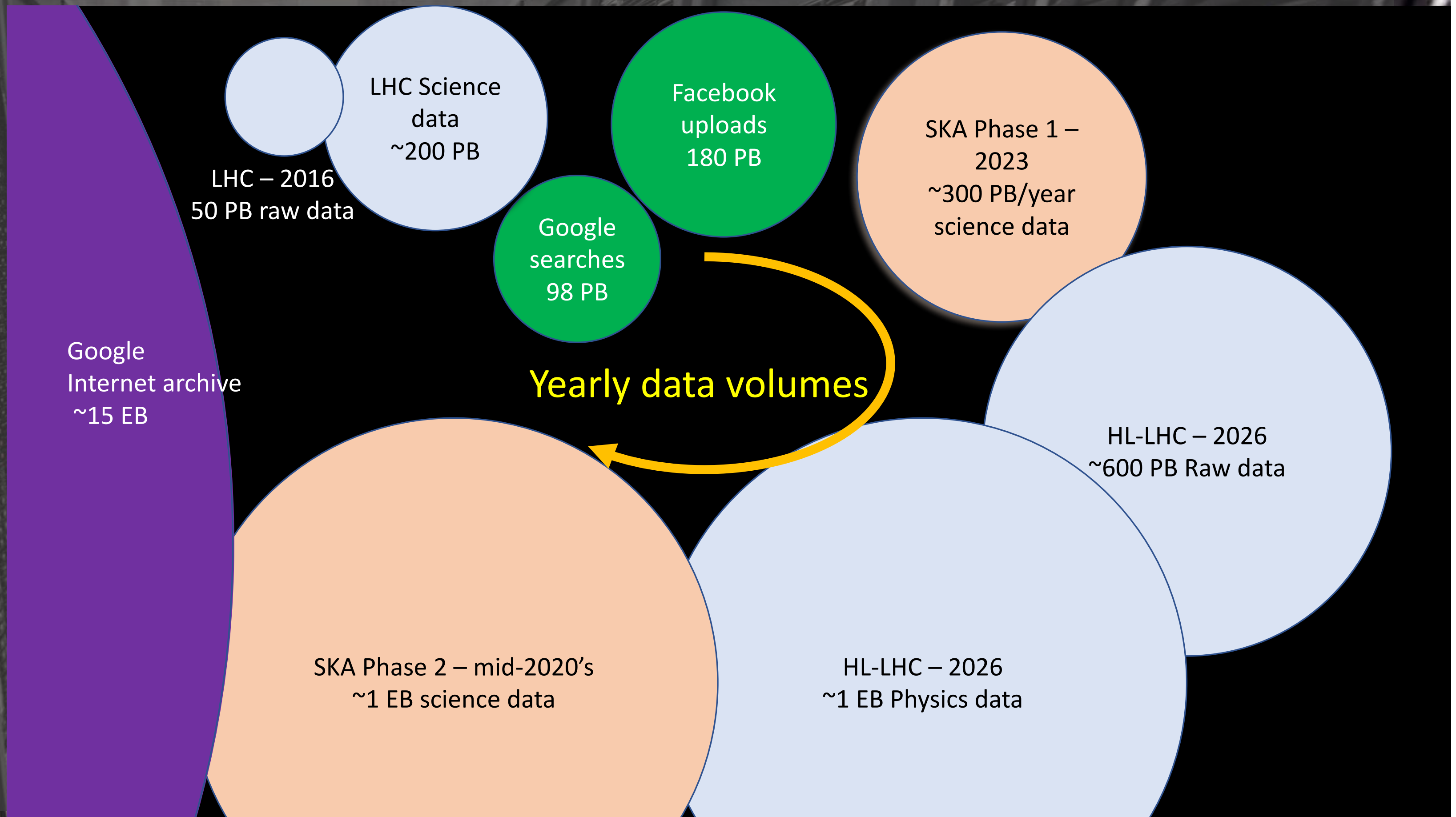


Applications are long-lived, large, and involve contributions from many people

Maria Girone
CERN openlab CTO

The other problem we have is delivering and retrieving data to HPC resources

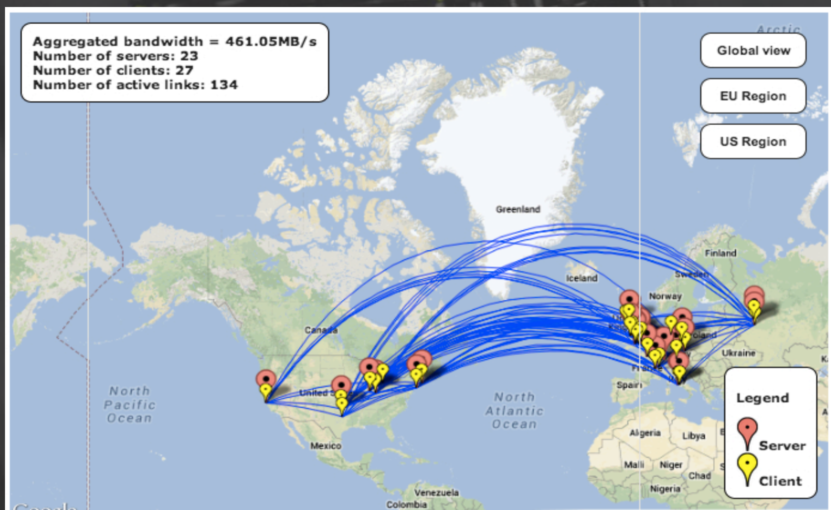
- The datasets are huge and grow over decades
- Data is frequently reprocessed as new hypotheses are investigated
- Need better mechanisms to import and export data

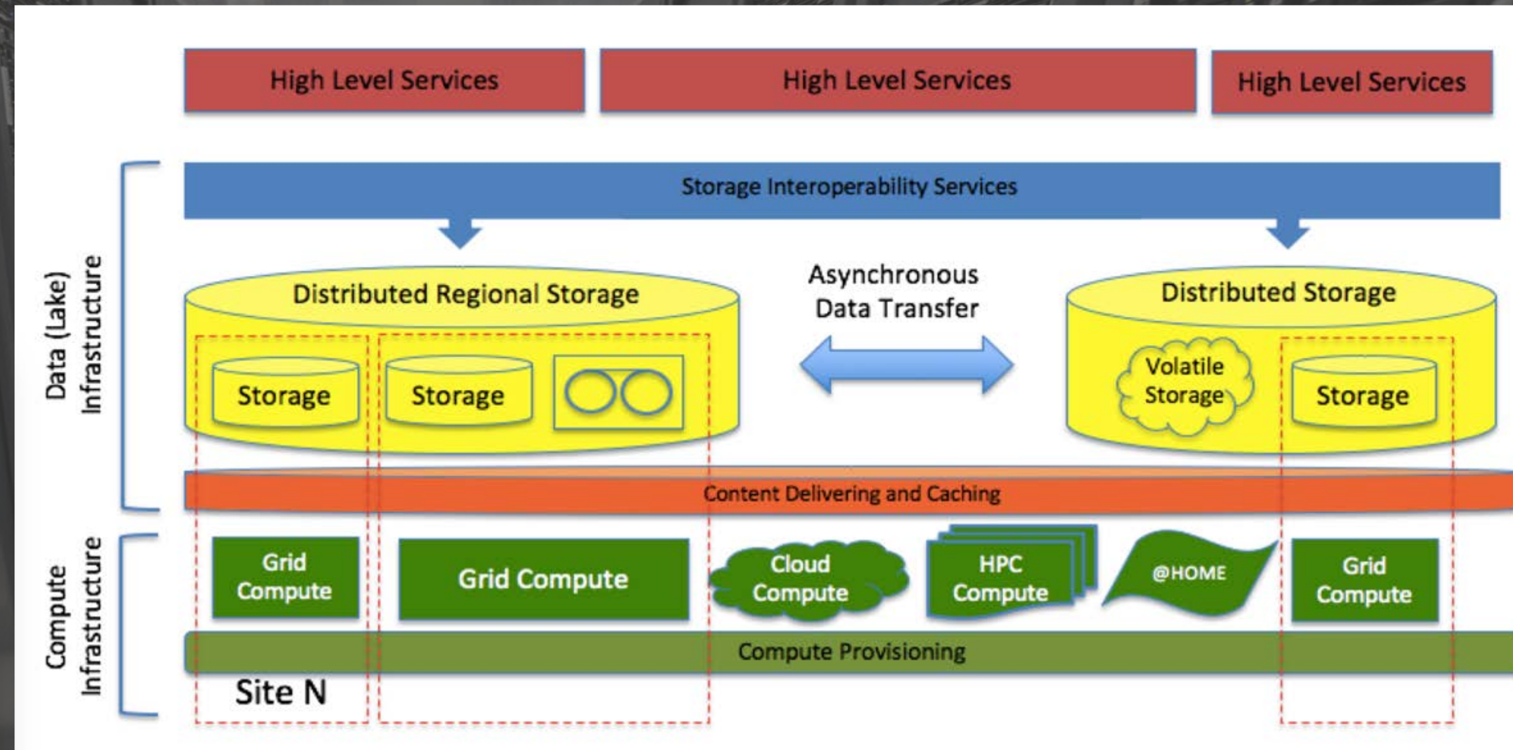


Data volumes are huge and datasets are active. HPC centers are often protected resources.

Maria Girone
CERN openlab CTO

Making hundreds of petabytes of data accessible globally to scientists is one the biggest challenges of WLCG





Data Organization, Management and Access in WLCG. Wide area streaming and caching at the exascale.

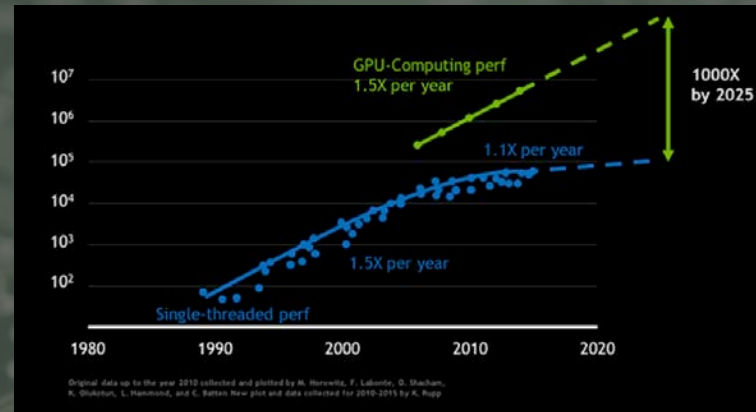
Maria Girone
CERN openlab CTO

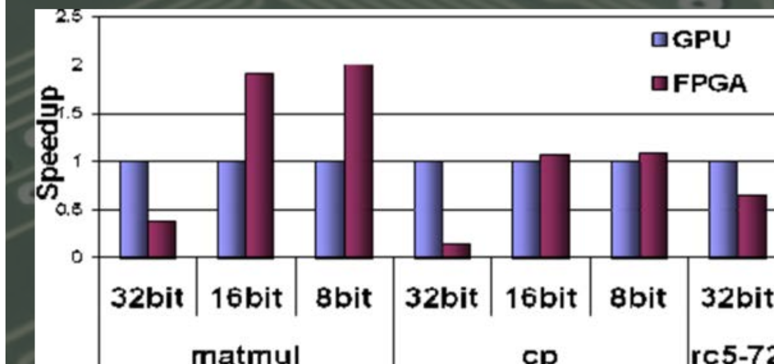
GPUs are currently experiencing a ~18 month doubling cycle

- As compared to ~36 months for CPUs
- If this trend continues the separation becomes very significant on the time scale of HL-LHC

We are actively working on benchmarking our applications on GPUs

- Currently GPUs are also more expensive than CPUs





Another promising technology for reconstruction and data transformations is FPGAs

- Very low power
- Improving development environment for facilitating adoption
- A number of R&D activities in HEP in high data rate applications

Alternative architectures like those of HPC centres are improving fast.

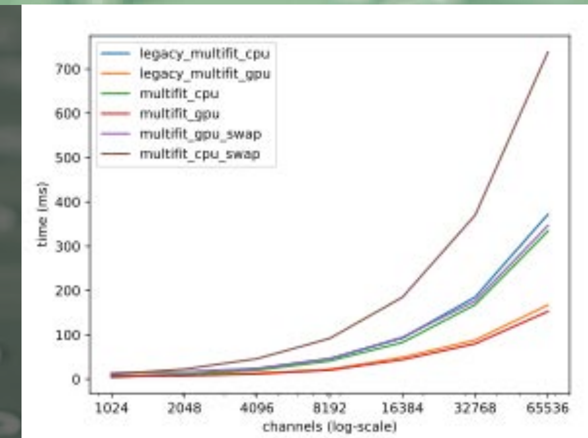
Maria Girone
CERN openlab CTO

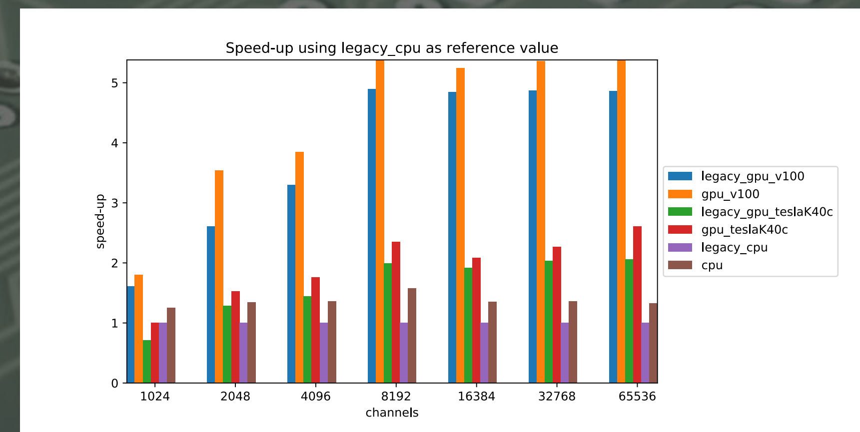
DEEP-EST: Dynamical Exascale Entry Platform - Extreme Scale Technologies

Work within the DEEP-EST project demonstrated the usefulness and advantages of heterogeneous architectures (GPUs)

- The idea is to offload some of the reconstruction workload from CPU to GPU
- The algorithm performs a regression for every calorimeter channel
  - the optimized algorithm is 30% faster on CPU
  - achieved a factor up to 5 parallelizing on GPUs
  - Need to benchmark further

Results achieved comparing a 3.5GHz Haswell CPU and a V100 GPU





Taking advantage of GPU hardware.

Maria Girone
CERN openlab CTO

Industry has invested heavily in optimizing machine learning libraries on GPUs (training and inference) and FPGAs (high speed inference)

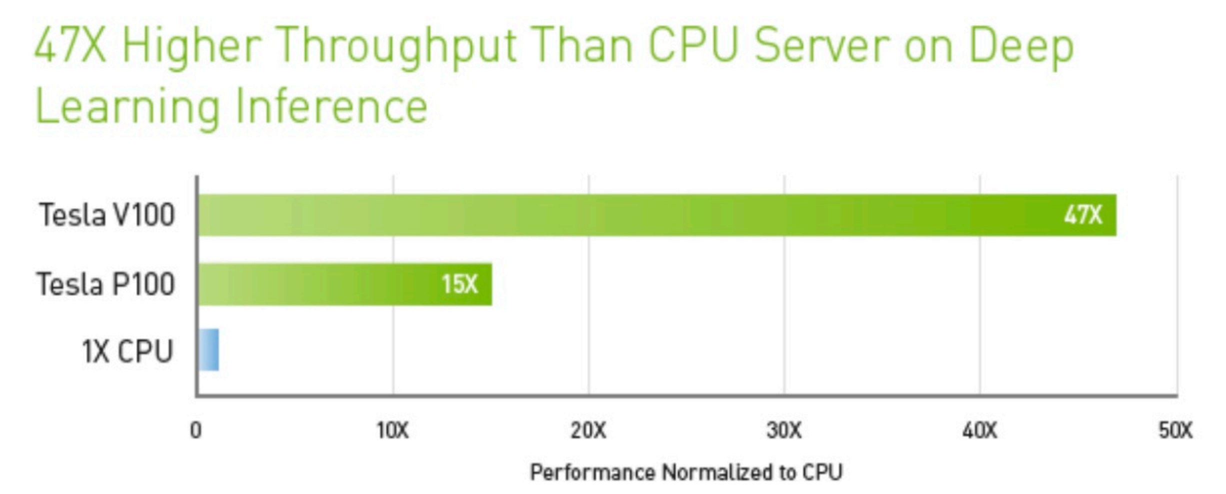
- The large user base ensures support on all recent hardware

A way to capitalize on the capabilities of HPC is to rethink our applications as a Machine Learning problem

- Rely on optimized code on accelerated hardware for training

There are ongoing R&D efforts in HEP to capitalize on this investment by exploiting machine learning techniques

- Filtering
- Object Identification and Event Classification
- Reconstruction
- Simulation



Maximizing the effectiveness of GPUs in ML.

Maria Girone
CERN openlab CTO

5 of the top 10 machines on the "Top 500" list get significant fractions of the computing power from GPUs

- The most powerful machine "Summit" has most of the capability coming from GPUs and applications are required to use the GPUs for an allocation

There is pressure in many countries to make effective use of national HPC resources in data intensive sciences

- A potential efficient area would be Machine Learning Training as a Service relying on the GPU resources located at HPC facilities

TOP 500

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , **IBM** DOE/SC/Oak Ridge National Laboratory United States	2,282,544	122,300.0	187,659.3	8,806
2	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , **NRCPC** National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
3	**Sierra** - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , **IBM** DOE/NNSA/LLNL United States	1,572,480	71,610.0	119,193.6	
4	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 , **NUDT** National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
5	**AI Bridging Cloud Infrastructure (ABCI)** - PRIMERGY CX2550 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , **Fujitsu** National Institute of [illegible] Industrial Science and Technology (AIST) Japan	391,680	19,880.0	32,576.6	1,649
6	**[illegible]** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , **Cray Inc.** Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
7	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , **Cray Inc.** DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
8	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , **IBM** DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890
9	**Trinity** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , **Cray Inc.** DOE/NNSA/LANL/SNL United States	979,968	14,137.3	43,902.6	3,844
10	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , **Cray Inc.** DOE/SC/LBNL/NERSC United States	622,336	14,014.7	27,880.7	3,939

Using HPC Resources for Machine Learning Training as a Service.

Maria Girone
CERN openlab CTO

- The LHC experiments are working closely with industry via CERN openlab on machine learning
  - Focus on adoption of accelerators (GPUs, FPGAs)
  - Engineering resources dedicated to support the application porting and increase knowhow on deep learning techniques

Data acquisition
- Real time event categorization
- Data monitoring & certification
- Fast inference for trigger systems

Data Reconstruction
- Calorimeter reconstruction
- Boosted object jet tagging

Data Processing
- Computing resource optimization
- Predicting data popularity
- Intelligent networking

Data Simulation
- Adversarial networks
- Fast simulation

Data Analysis
- Knowledge base
- Data reduction
- Searches for new physics

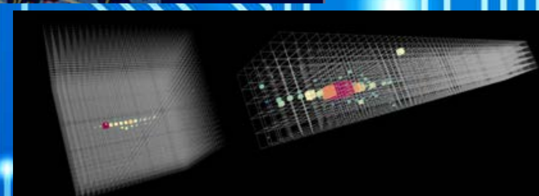

Machine learning and data analytics are hot topics at CERN openlab workshop

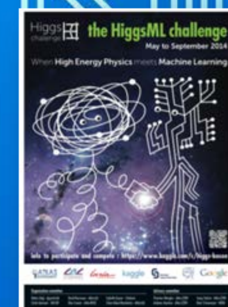
Wednesday, 4 May, 2016

Last week, CERN openlab held a workshop on machine learning and data analytics. The event, which took place on Friday 29 April, saw experts from both research and industry gather in the CERN IT Department for a full-day of presentations and lively discussion.

The morning featured presentations from representatives of the four large LHC experiments — ALICE, ATLAS, CMS, and LHCb — on their current projects and future challenges in these areas.

During the afternoon, representatives from industry were also invited to give their perspective. This included presentations from CERN openlab partner companies Intel and Siemens, as well as contributors Cisco and associates Yandex. The other companies to present at the event were Cloudera, Google, IBM, Microsoft, and Nvidia.





Exploration areas.

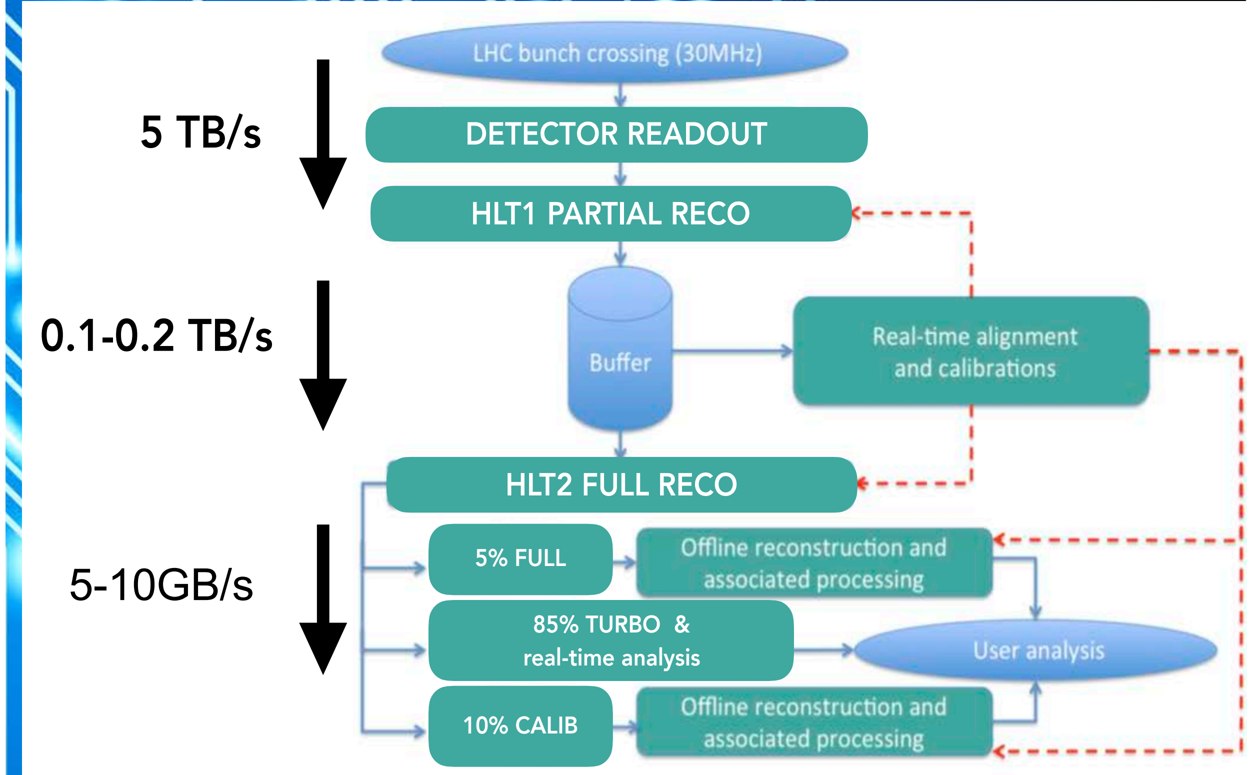
Maria Girone
CERN openlab CTO

The LHC experiments will increase their data acceptance rates for the upgrades

- Investigating FPGAs and GPUs to allow selection using Machine Learning in data intensive environments
    - Complex decisions very close to real time (micro-seconds)

Accelerated hardware can reduce the reconstruction time by orders of magnitude

- Integrate GPUs in the HLT farm to give high-quality reconstruction in 200 msec latency (as opposed to tens of sec)



Investigating Accelerated Architectures for for filtering and reconstruction.

Maria Girone
CERN openlab CTO

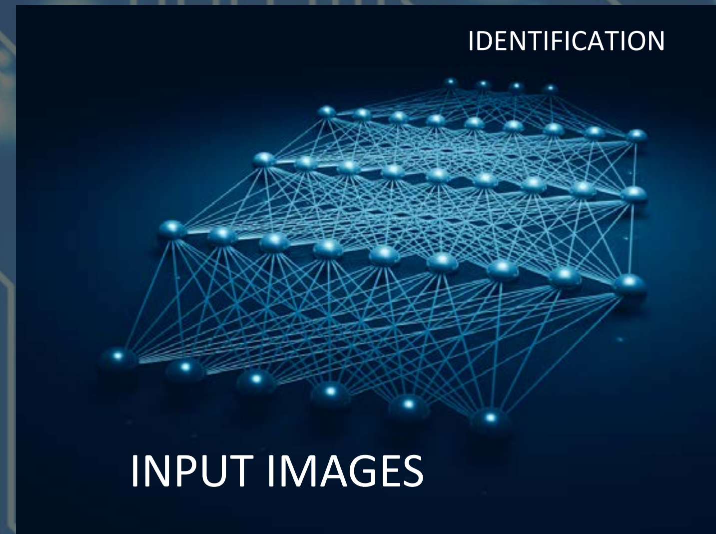
With current software and computers an event like HL-LHC takes 10s of seconds

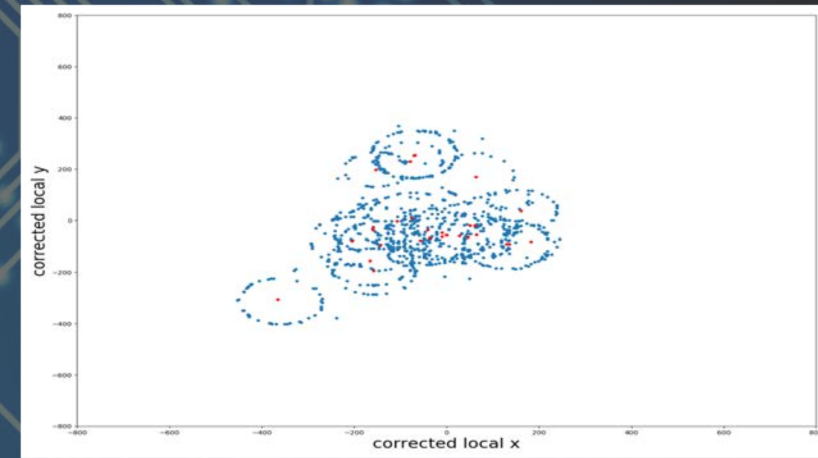
Examine the detector hit information and use 3D image recognition techniques to identify objects

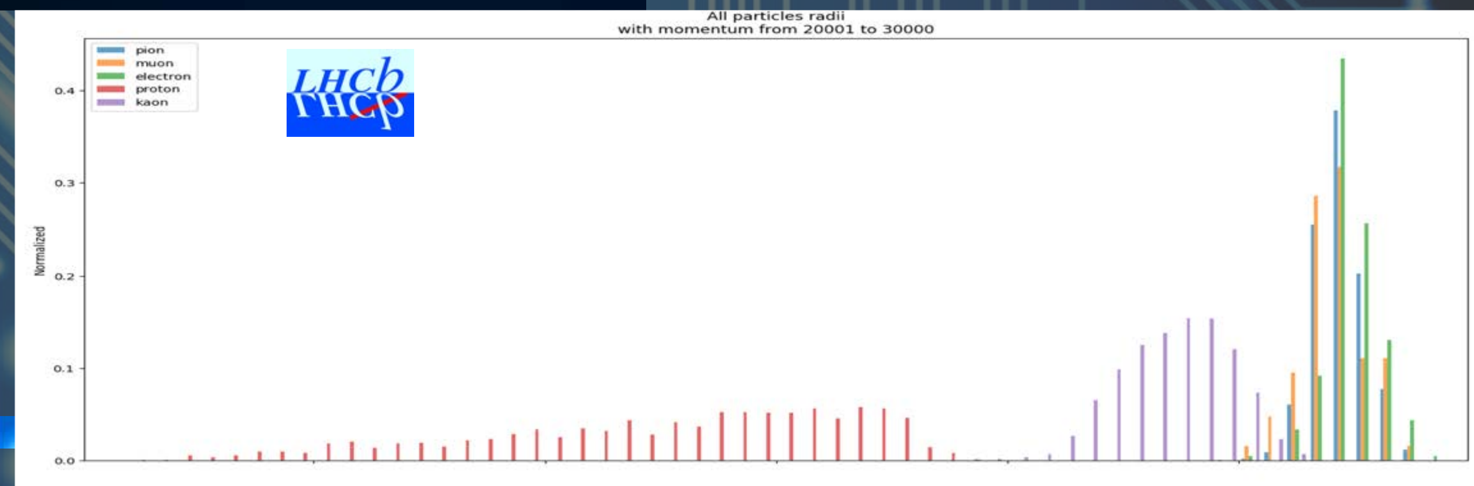
- Recognize physics objects from learned patterns

LHCb is exploring particle identification in the RICH detector

- Convolutional neural networks to classify particles based on the radius
- Comparing several modern frameworks: Keras, TensorFlow, and Caffe



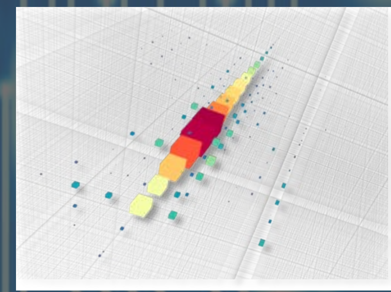


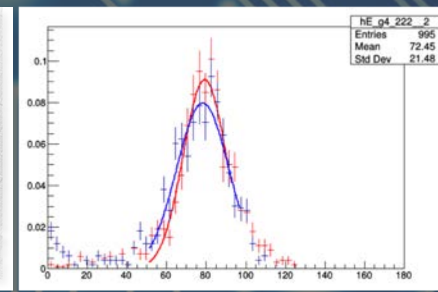


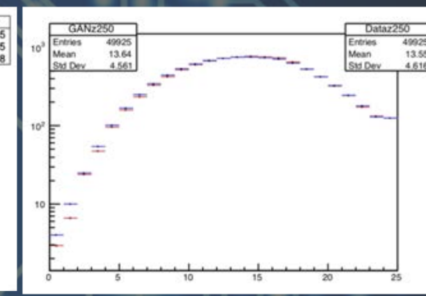
Exploring image recognition for reconstruction.

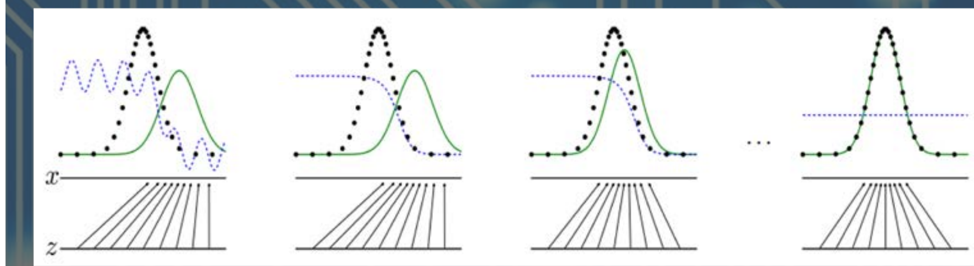
Maria Girone
CERN openlab CTO

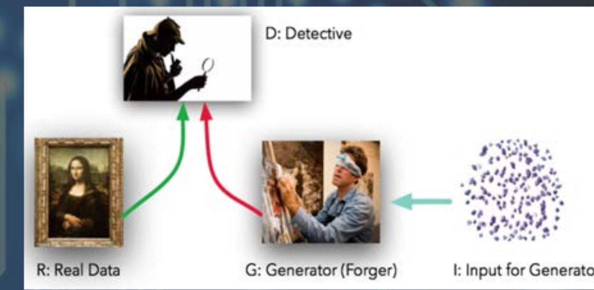
- Two main R&D areas
  - Adapting the existing code to new computing architectures
  - Replacing complex algorithms with deep-learning approaches (FAST SIMULATION)











Looking at adversarial networks to improve speed without giving up accuracy of simulated events

- One network attempts to simulate events that match a data distribution
- While a second network tries to distinguish data and simulation

Simulation is one of the most resource-intensive computing applications.

Maria Girone
CERN openlab CTO

- Network security and fraud detection
- Industrial monitoring and predictive failures
  - Looking at optimizing performance of complex systems
    - Minimize costs and improve resource utilization

- LHC magnets, industrial controls, …
- Detector Health
  - Complex system monitoring to minimize downtime and reduce operations costs
- Resource Utilization (scheduling, data placement, I/O optimization)
- Automated online monitoring at sub-detector data and metadata level with anomaly detection

Reinforcement learning

Monitoring, Automation and Anomaly Detection.

Maria Girone
CERN openlab CTO

CERN has been pushing the boundaries of knowledge and technology for more than 60 years.

The next phase of the programme will include unprecedented computing challenges.

Looking to engage HPC centres and alternative architectures to provide software and computing for the next generation of the LHC programme.

BACKUP SLIDES

- LHCb and ALICE will move offline processing closer to the online data collection chain
  - Performing processing and data analysis in near real-time
  - Solutions under investigation

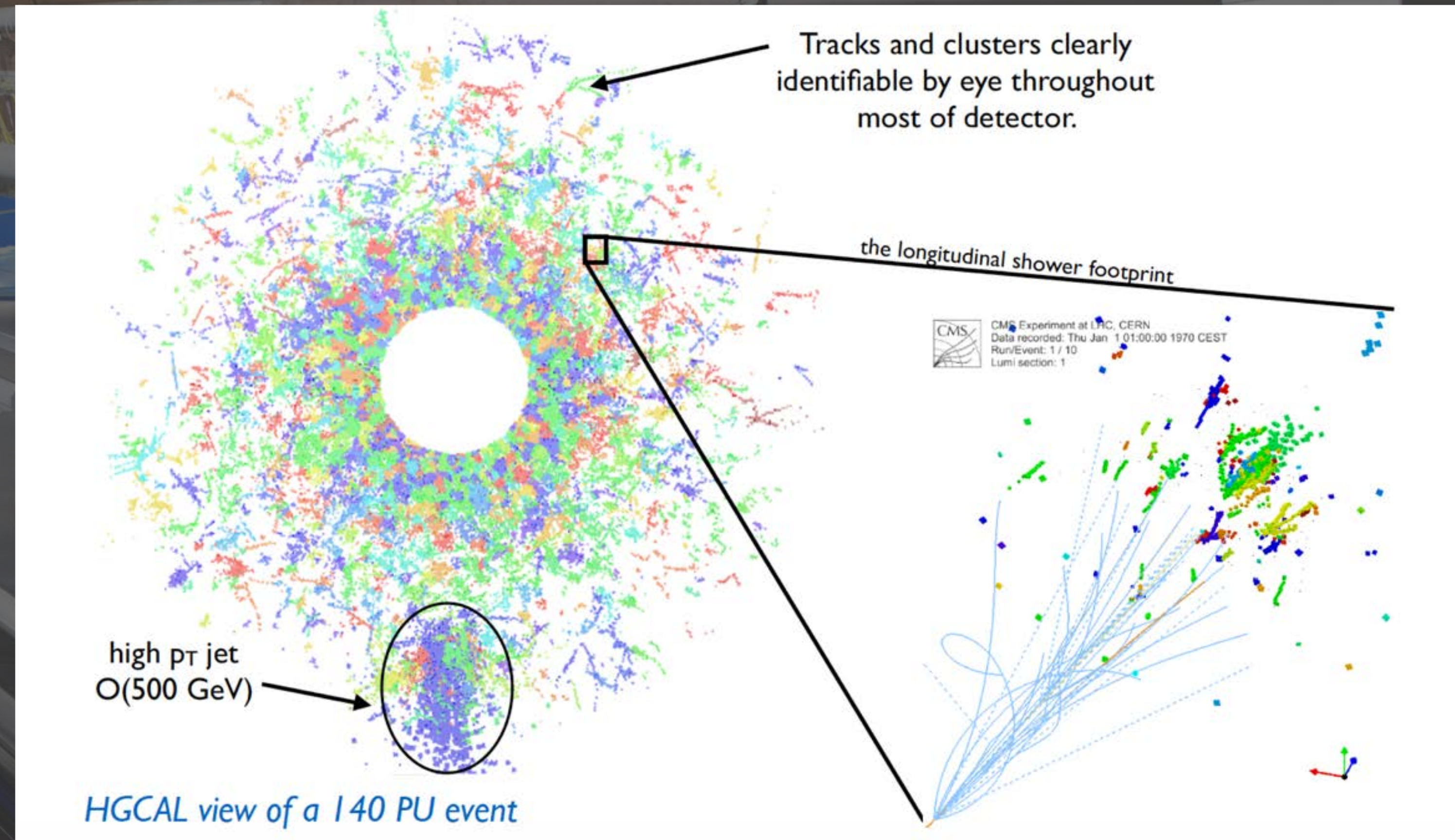
- New HLT farms for Run3
  - Flexible and efficient system with ambitious PUE ratio

Courtesy of Automation Data Center Facilities

The ALICE and LHCb experiments will increase their data acceptance rates for Run 3.

Maria Girone
CERN openlab CTO

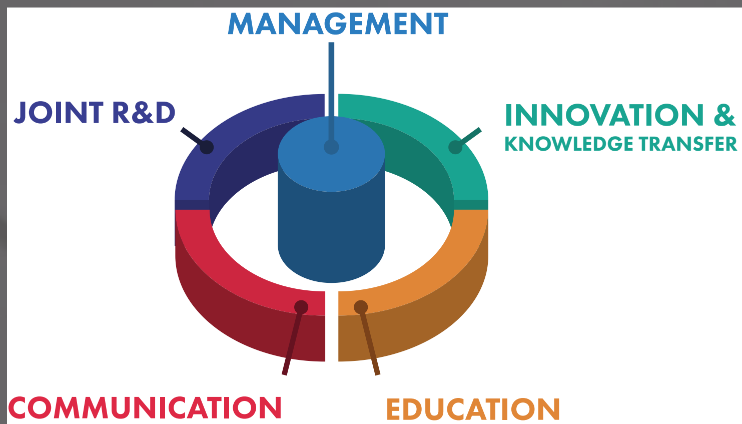
- By Run4, the detectors will become more granular and more radiation hard.
- Reconstructing more particles with more granular detectors will be computationally more expensive.



The ATLAS and CMS experiments will be significantly upgraded for the HL-LHC.

Maria Girone
CERN openlab CTO

CERN openlab



PARTNERS	CONTRIBUTORS	ASSOCIATES	RESEARCH
HUAWEI	rackspace	COMTRADE	INFN, King's College London
intel	IBM	Yandex	Fermilab, Scimpulse
ORACLE	E4 Computer Engineering	open systems	Newcastle University, EMBL-EBI
SIEMENS	Extreme		GSI, IVI Innovation Value Institute

CERN openlab is a unique science-industry partnership, fostering research and innovation.

Maria Girone
CERN openlab CTO

CERN openlab is engaging in QC with industry

- Can substantially speed-up training of deep learning and combinatorial searches
- Well suited for fitting, minimization, optimization
- Can directly describe basic interactions as well as lattice QCD calculations

Quantum Computing is also on the horizon.