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
CERN is the European Laboratory for Particle Physics and straddles the Franco-Swiss border near Geneva.
It has 22 member states and supports a global community of 15,000 researchers.
Fundamental and cross domain research questions

Looking for Antimatter

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

Understanding Dark Matter

Maria Girone
CERN openlab CTO
The Large Hadron Collider (LHC) is 27km circumference near Geneva and has 4 main detectors around the ring.
Two general-purpose detectors cross-confirm discoveries, such as the Higgs boson.

ATLAS
- 46m long, 25m diameter
- Weights 7,000 tonnes
- 100 million electronic channels, 3,000 km of cables

CMS
- 22m long, 15m diameter
- Weights 14,000 tonnes
- Most powerful superconducting solenoid ever built

Maria Girone
CERN openlab CTO
ALICE and LHCb experiments have detectors specialised on studying specific phenomena.

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.
~1B collisions per second generate particles that decay in complex ways into even more particles.
Data generated 40 million times per second

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
Physicists must sift through the 30-50 PBs produced annually by the LHC experiments.
The process to transform raw data into useful physics datasets is a complicated series of steps at the LHC (Run2).
Computing is handled by the Worldwide LHC Computing Grid, a hierarchical distributed computing infrastructure.

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

- **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

Would put us amongst the top Supercomputers if centrally placed: est. ~few x100 Pflops
The LHC has been designed to follow a carefully set out programme of upgrades that will greatly increase the scientific reach.
More collisions help physicists to observe rare processes and study with greater precision.

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

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

Maria Girone
CERN openlab CTO
The HL-LHC will come online around 2026. More collisions and more complex data.
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.
We have to overcome a factor of 4 to 8 even after technology improvements

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

**Three main areas of research and development.**

- **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

**COMPUTING CHALLENGES**

Maria Girone
CERN openlab CTO
Layered, virtualized services provide flexibility and efficiency.

**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

---

**María Girone**
CERN openlab CTO
Experiments have demonstrated that it is possible to elastically and dynamically expand production resources to commercial clouds.

Large-scale tests with commercial clouds.
HPC Centers are huge computational resources.

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
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.
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.
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.
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
Alternative architectures like those of HPC centres are improving fast.

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
Taking advantage of GPU hardware.

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
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.
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

Using HPC Resources for Machine Learning Training as a Service.
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
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)
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.
Simulation is one of the most resource-intensive computing applications.

- 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
• 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

Monitoring, Automation and Anomaly Detection.
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
The ALICE and LHCb experiments will increase their data acceptance rates for Run 3.

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

Maria Girone
CERN openlab CTO
Quantum Computing is also on the horizon.

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