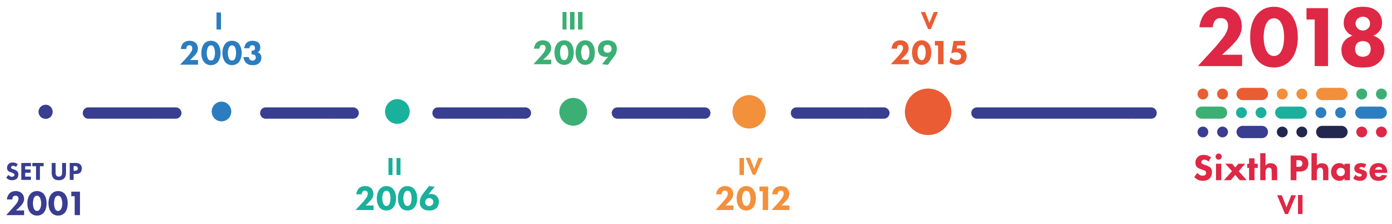
Partnering with Industry for Machine Learning at HL-LHC

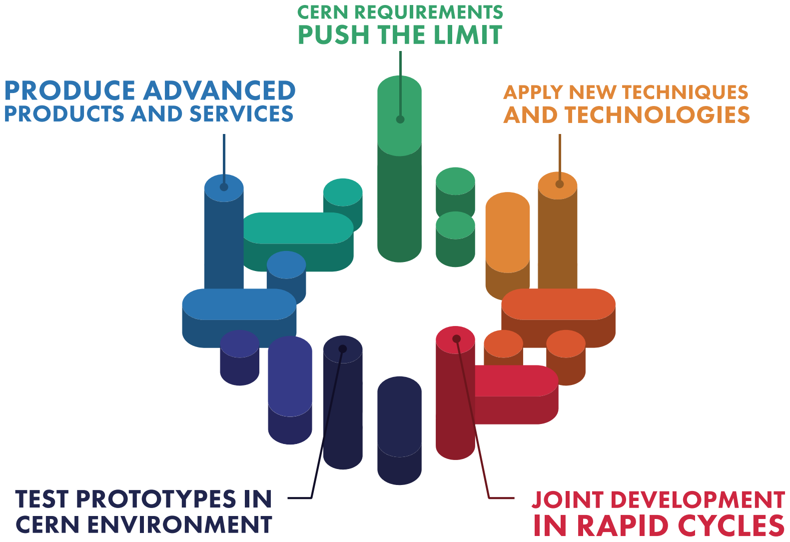
Maria Girone, CERN

CHEP18 Sofia: Track 6

CERN OPENLAB

A public-private partnership between the research community and industry, fostering innovativation



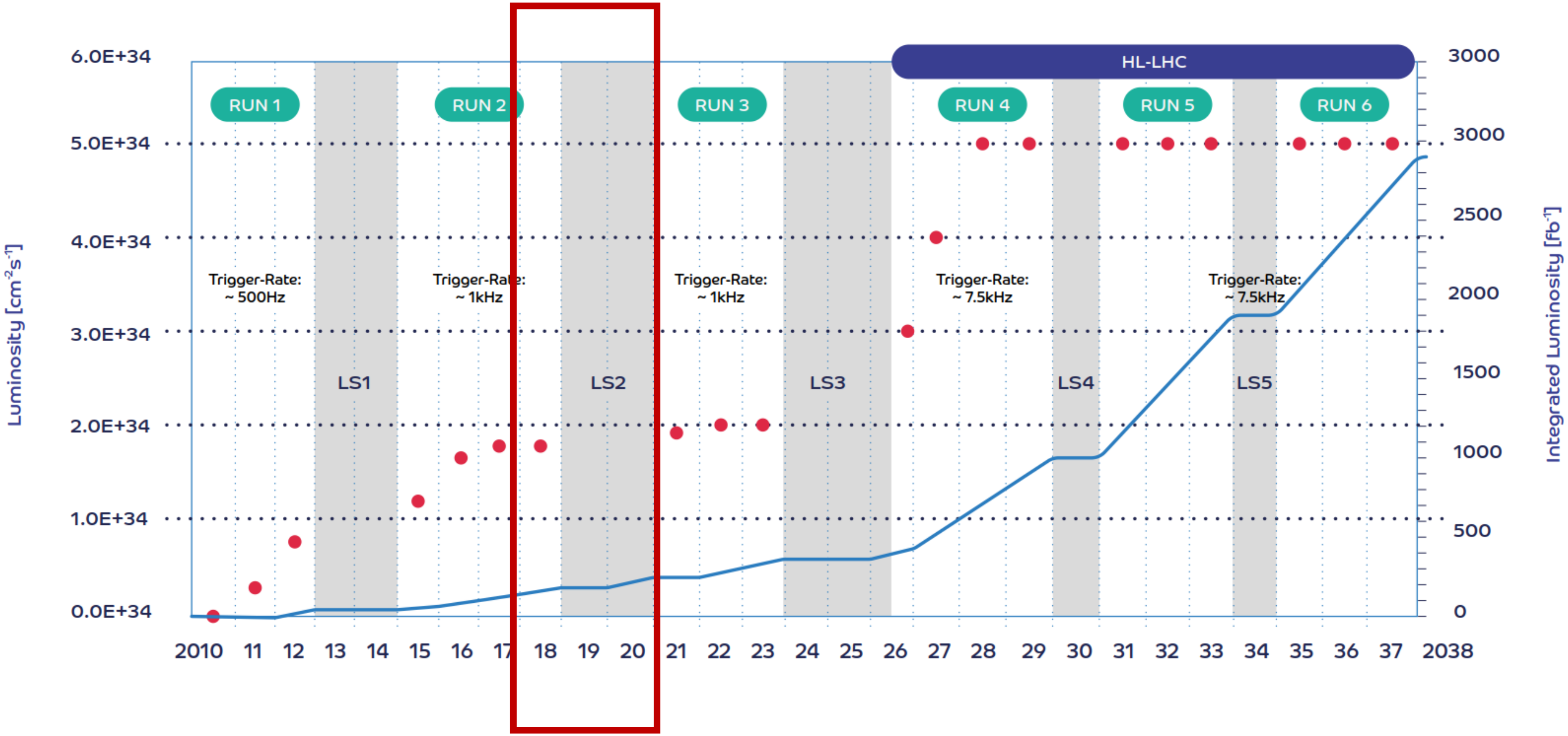


CERN openlab Collaboration

PARTNERS	CONTRIBUTORS	ASSOCIATES	RESEARCH
HUAWEI	rackspace	COMTRADE	INFN, KING'S College LONDON
intel	IBM	Yandex	Fermilab, Scimpulse foundation
ORACLE	E4 COMPUTER ENGINEERING	open systems	Newcastle University, EMBL-EBI
SIEMENS	Extreme		GSI, IVI INNOVATION VALUE INSTITUTE

Upgrades of LHC

A carefully set out programme of upgrades to increase the scientific reach.



RUN 3

ALICE & LHCb upgrades

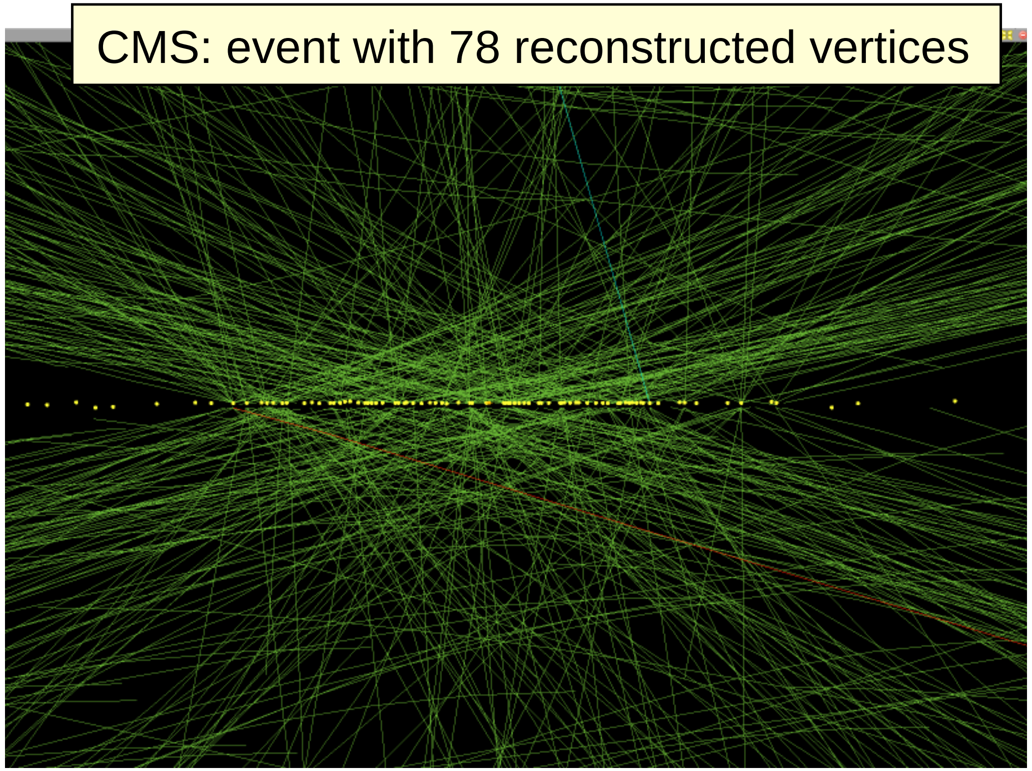
RUN 4

ATLAS & CMS upgrades

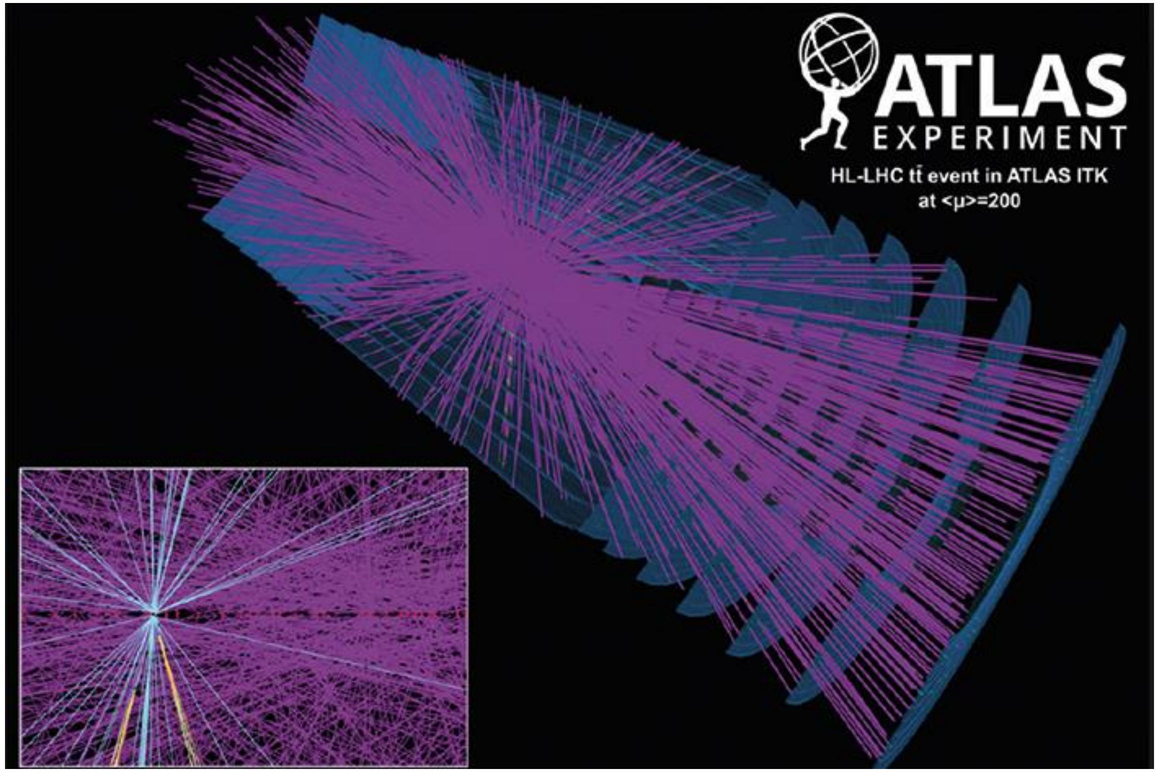
Challenges of Complexity

More collisions and more complex data at HL-LHC.

CMS: event from 2017 with 78 reconstructed vertices

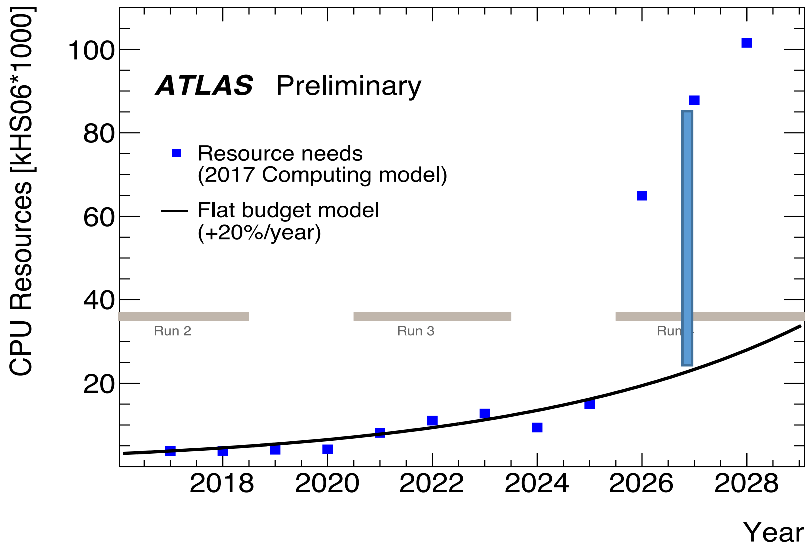


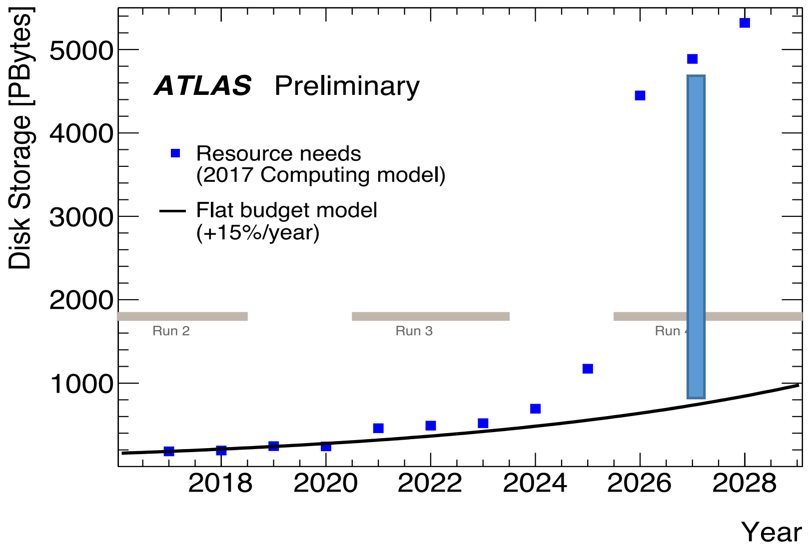
ATLAS: simulation for HL-LHC with 200 vertices



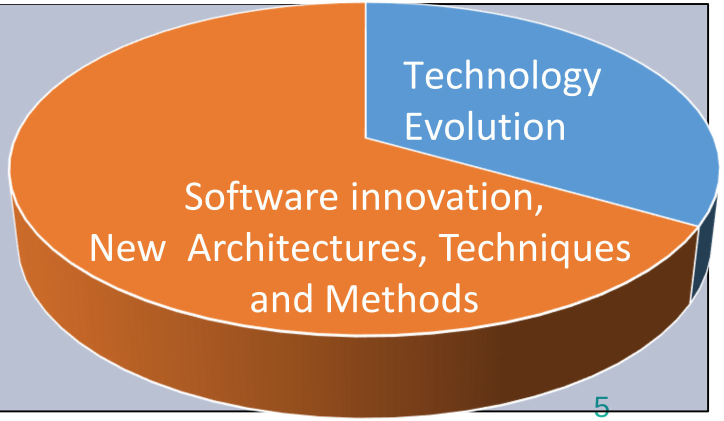
Resource Gap

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



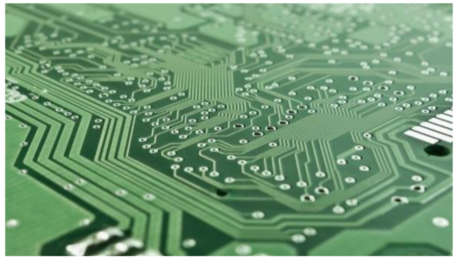


Closing the resource gap in the next decade requires close collaboration with industry.

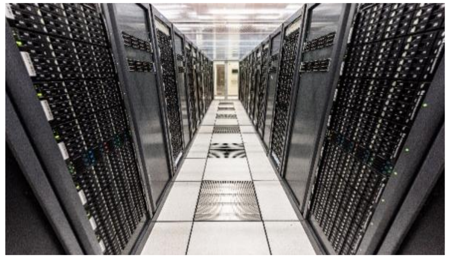


R&D Areas towards Run3 and Run4

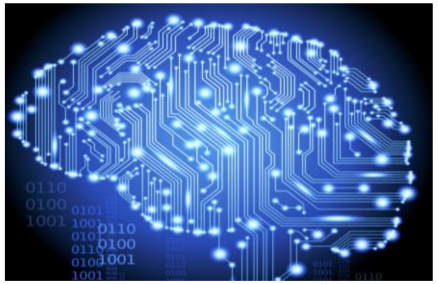
The LHC experiments have a multi-pronged approach to closing the resource gap



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



Scale out capacity with public clouds, HPC, new architectures



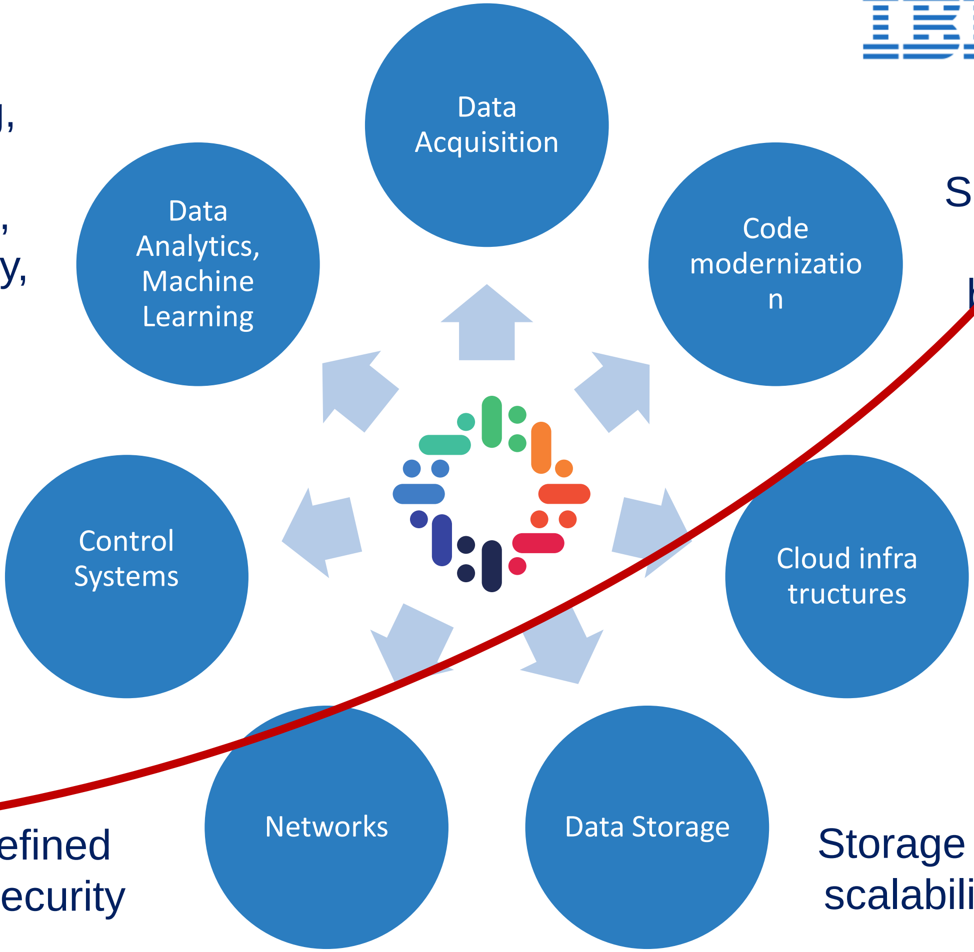
New techniques with Machine Learning, Deep Learning, Advanced Data Analytics

JOINT R&D PROJECTS

High-bandwidth fabrics, accelerated platforms

Data quality monitoring, anomaly detection, physics data reduction, benchmarking/scalability, systems biology and large-scale multi-disciplinary platforms

Simulation, HPC on the Cloud, benchmarking



Predictive/proactive maintenance and operations

Cloud federations, containers, scalability

Software Defined Networks, Security

Storage architectures, scalability, monitoring

Machine Learning in HEP

- ML is an area with heavy investment within industry
- The LHC experiments are working closely with industry via CERN openlab
  - 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**

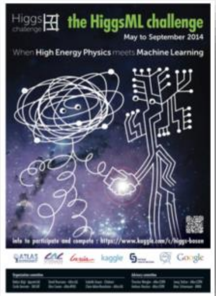
Data Reconstruction
- **Calorimeter reconstruction**
- **Boosted object jet tagging**

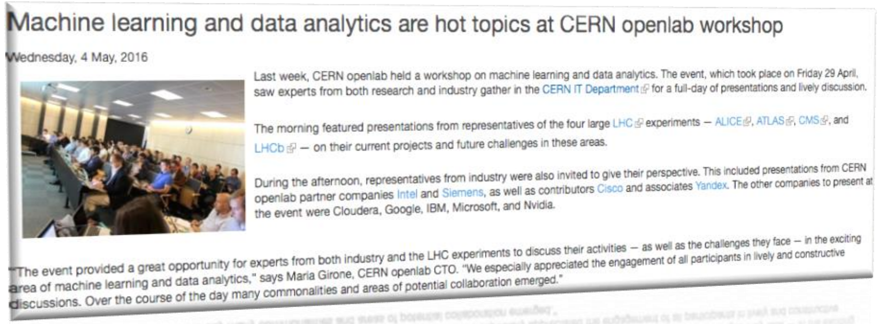
Data Processing
- **Computing resource optimization**
- **Predicting data popularity**
- **Intelligent networking**

Data Simulation
- **Adversarial networks**
- **Fast simulation**

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

SIEMENS intel E4 COMPUTER ENGINEERING IBM Yandex




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

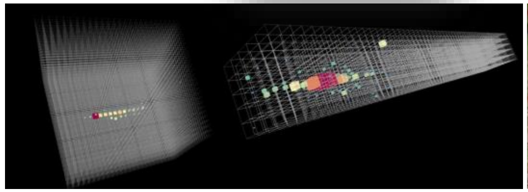
Wednesday, 4 May, 2016

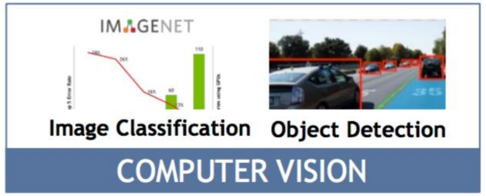
Last week, CERN openlab held a workshop on machine learning and data analytics. The event, which took place on Friday 29 April, saw experts from both research and industry gather in the CERN IT Department for a full-day of presentations and lively discussion.

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.

"The event provided a great opportunity for experts from both industry and the LHC experiments to discuss their activities — as well as the challenges they face — in the exciting area of machine learning and data analytics," says Maria Girone, CERN openlab CTO. "We especially appreciated the engagement of all participants in lively and constructive discussions. Over the course of the day many commonalities and areas of potential collaboration emerged."





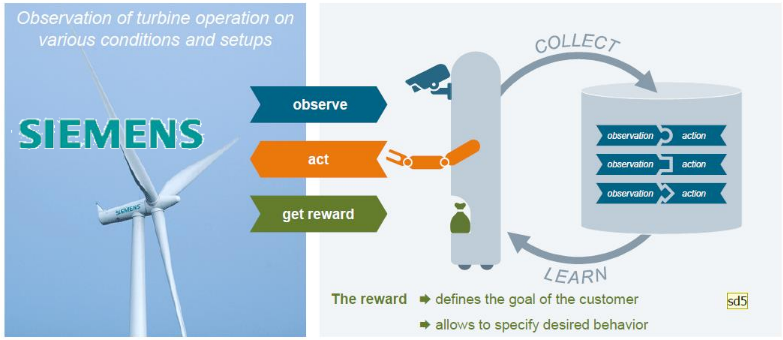
Monitoring, Automation, Anomaly Detection

SIEMENS intel Yandex

- 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)

Reinforcement learning



SIEMENS

A multitude of Industrial Control Systems

Cooling & Ventilation   VACUUM

Cryogenics   GAS

Electric Grid   LHC Circuit, QPS, WIC, PIC, …

Data Quality Monitoring

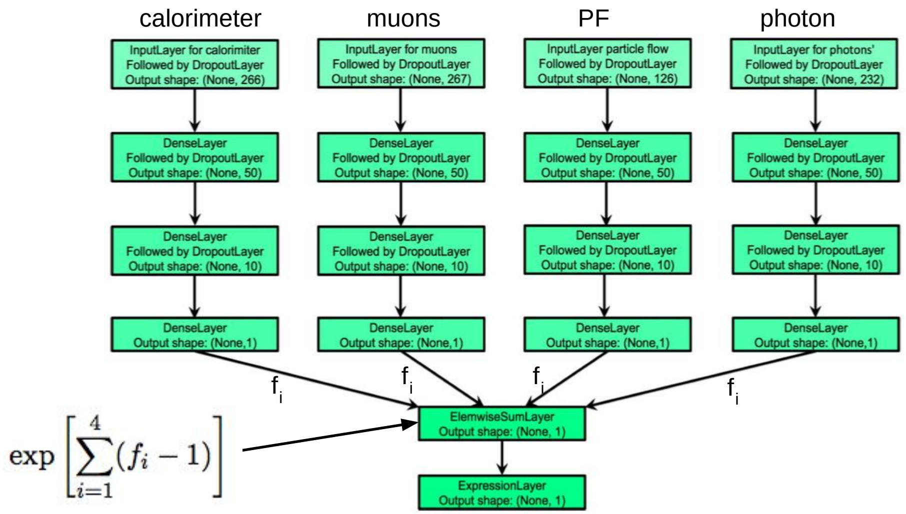
Monitoring the data continuously is effort intensive and critical

- data is monitored by shift teams looking for anomalies in distributions

CMS uses supervised learning on reconstructed data with multi-head NN to predict a probability of anomaly in separated channels

- Results are combined to establish the quality of the data

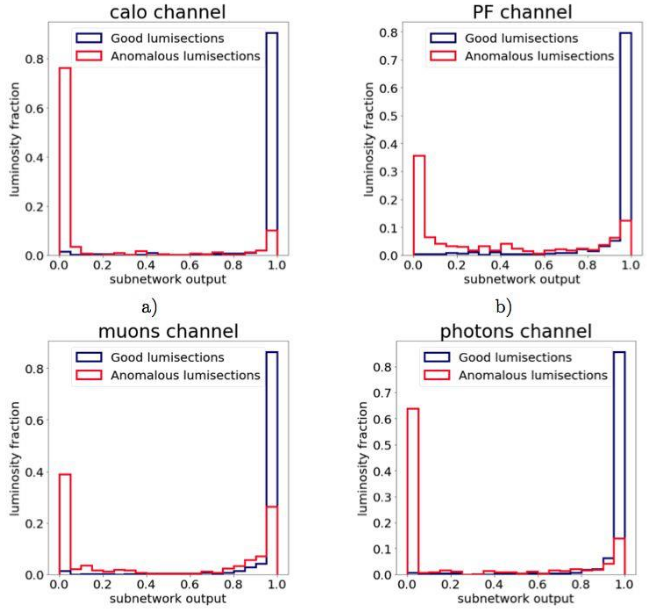
Yandex



CMS is also working with IBM on automated online monitoring at sub-detector data and metadata level

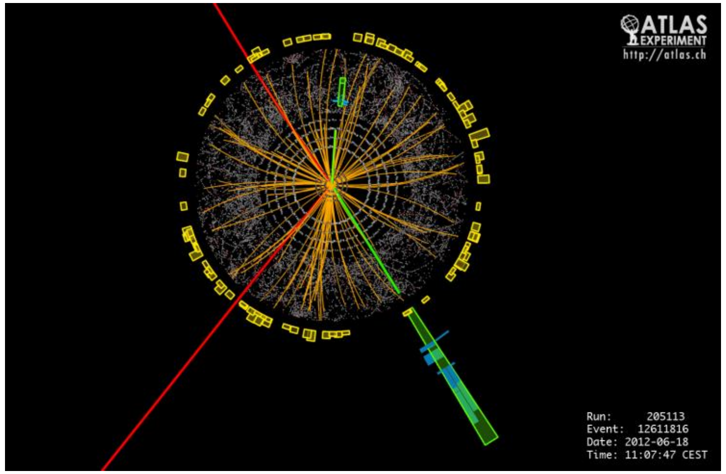
- Predict anomalies in Ecal and Hcal using deep learning recursive NN
- Goal is to integrate detector control systems for a comprehensive monitoring overview

IBM



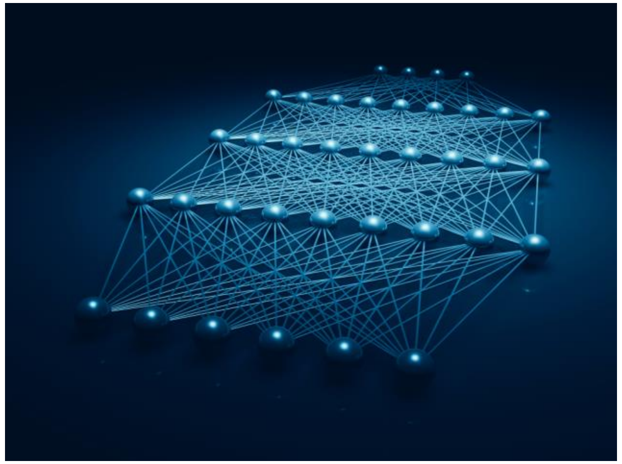
V. Azzolini: Improving the use of data quality metadata via a partnership of technologies and resources between the CMS experiment at CERN and industry

Event Reconstruction



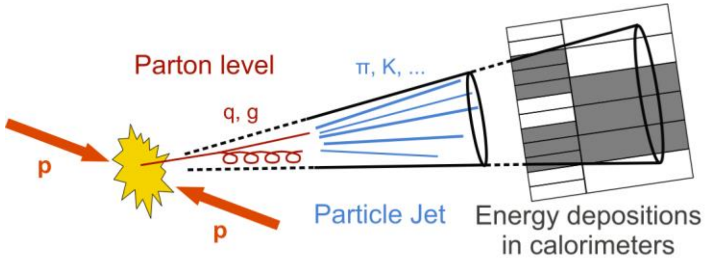
With current software and computer an event like HL-LHC takes 10s of seconds

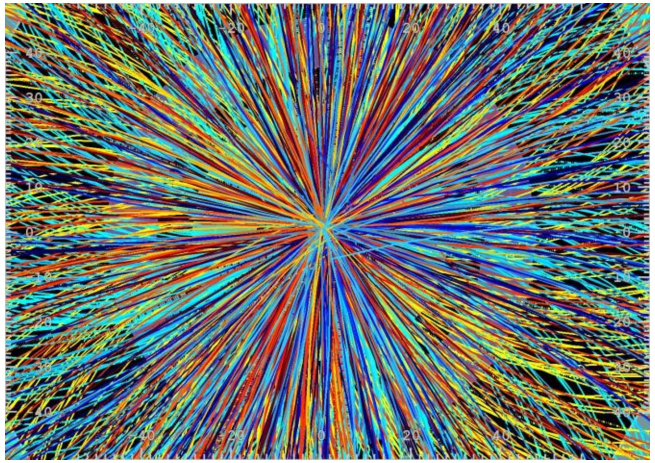
- Investigating CNN



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

- Recognize physics objects from learned patterns
  - Most ML recognition techniques are designed for regular coordinate systems





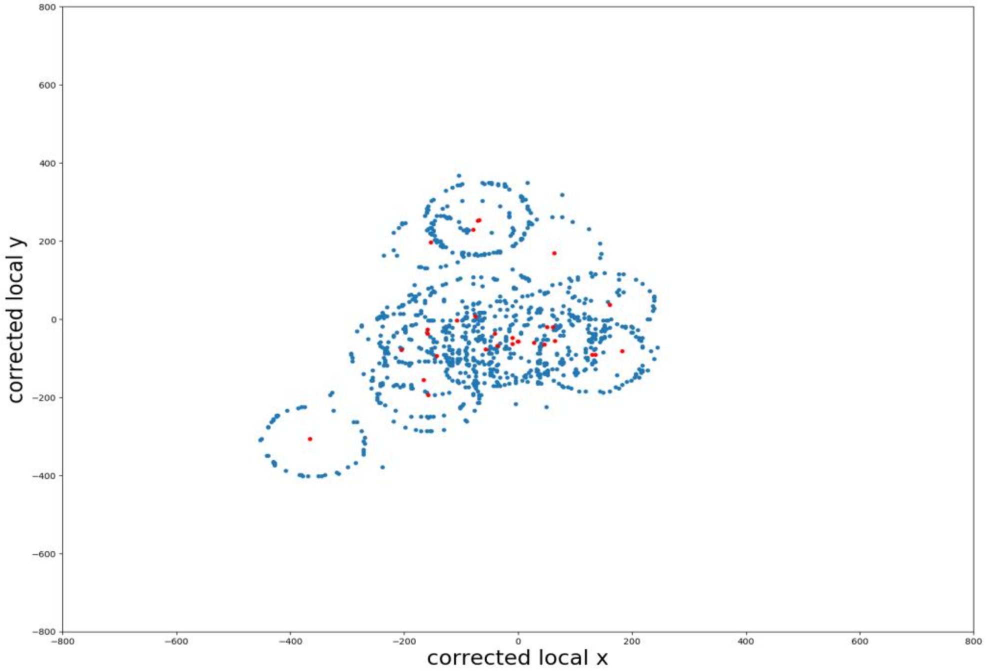
Investigate train-on-demand services to be executed on HPC

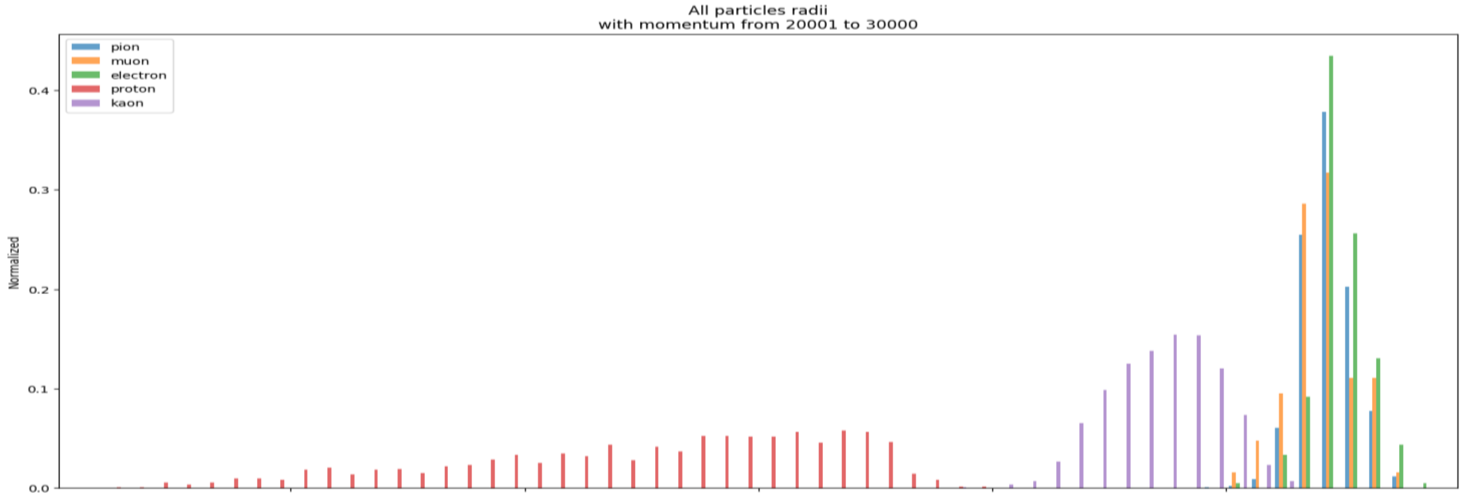
F. Pantaleo: Distributed training of deep NN models
M. Kiehn: TrackML: the Kaggle HEP tracking challenge

Object Identification

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
- Two ongoing projects in openlab with E4 COMPUTER ENGINEERING and IBM





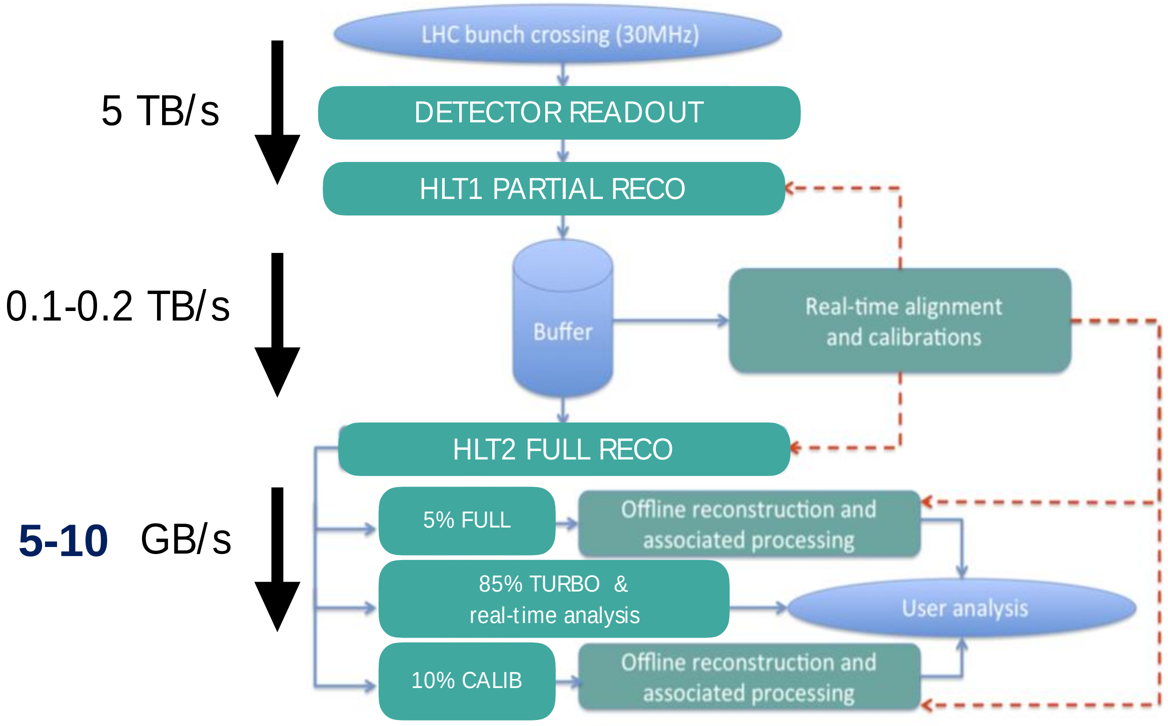
Software-based Filtering and real-time Reconstruction

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

- LHCb is investigating FPGAs and GPUs to allow reconstruction of 5 TB/s of events in real time. Deep learning techniques under investigation

At HL-LHC higher data rates will require more selective triggering and faster reconstruction

- CMS is porting heavy "offline" tasks to real-time processing for HL-LHC
  - Integrate GPUs in the HLT farm to give high-quality reconstruction in 200 msec latency (as opposed to tens of sec)



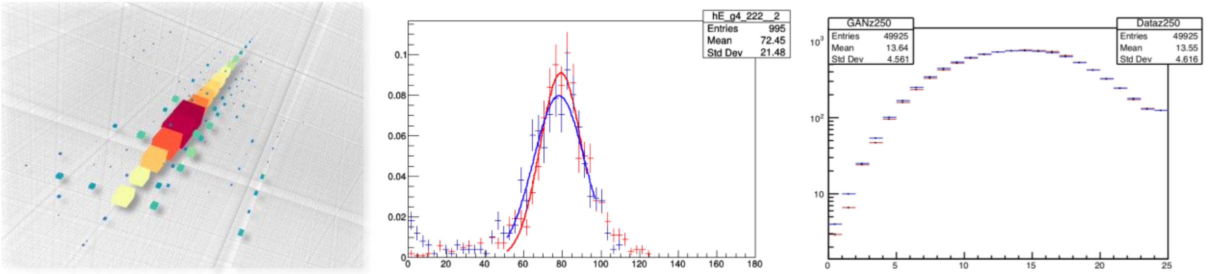
F.. Pantaleo Patatrack

Event Simulation

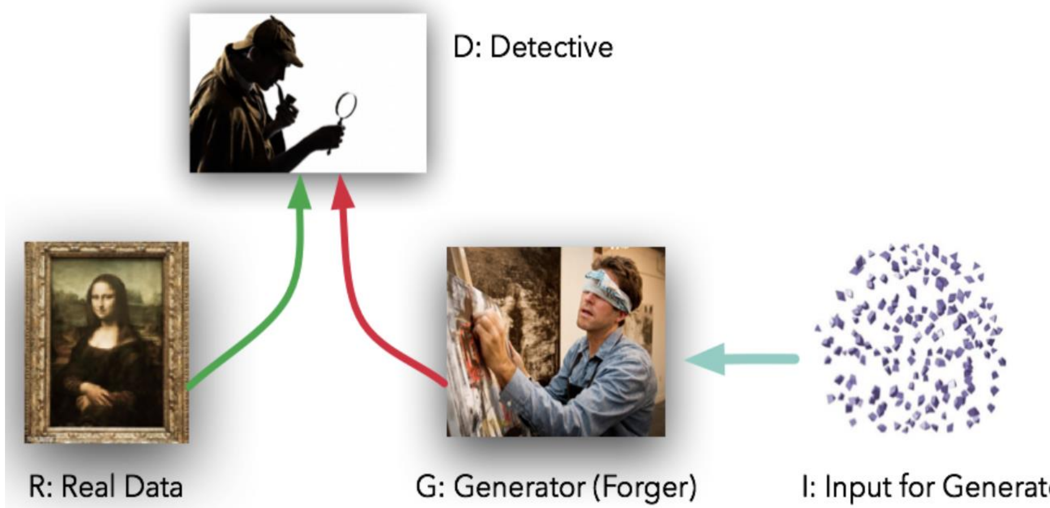
- Simulation is one of the most resource-intensive computing applications.
- Main R&D areas:
  - Adapting the existing code to new computing architectures
  - Replacing complex algorithms with deep-learning approaches (FAST SIMULATION)

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

- One network attempts to simulate events that match a data distribution (Generator G)
- While a second network tries to distinguish data and simulation (Discriminator D)



S. Vallecorsa: A machine learning tool for fast simulation



Outlook

- Machine Learning has heavy investments by industry and rapid development cycles

- CERN openlab has active projects with industry in key areas such as data acquisition, processing and analysis
  - Help closing the resource gap in the LHC Run3 and Run4
  - Machine learning is one of the primary focus activities of CERN openlab phase VI

- We have made good progress towards adopting ML for automating data quality monitoring and making faster simulation (without giving up on accuracy)
  - Object identification and full event reconstruction are active areas of investigation

- ML is a fast moving field and we are hoping for more breakthroughs