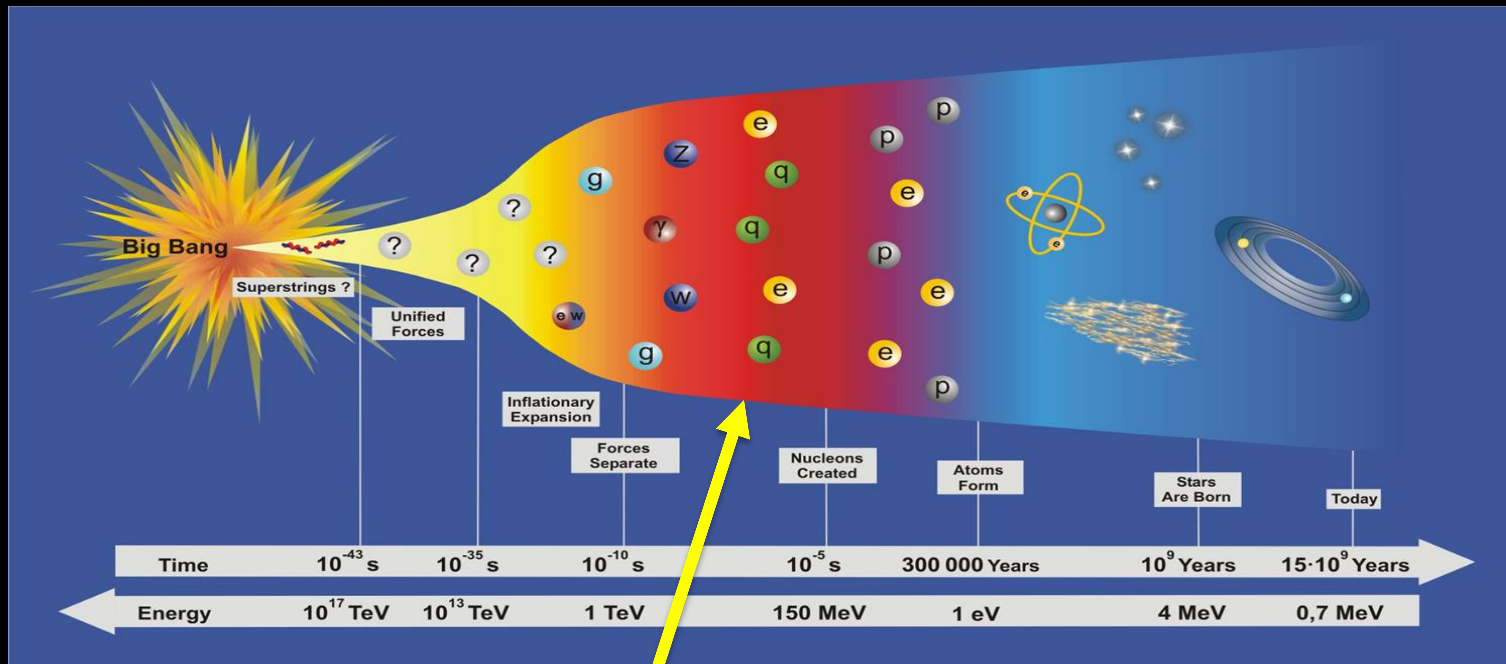
High Performance Data Processing with ALICE at the LHC

- Introduction
- ALICE Case
- ALICE Future
- Summary and Outlook

Jacek Otwinowski

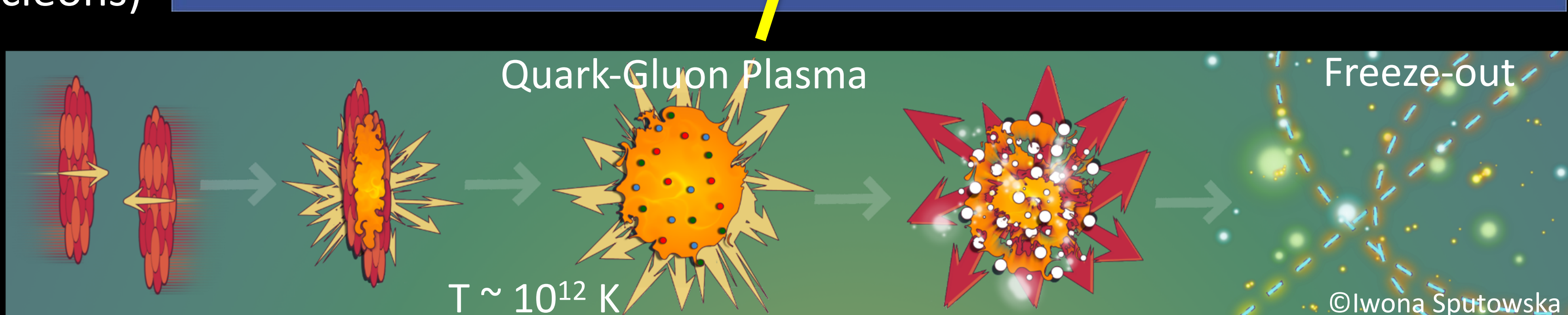
AGH, 26.11.2020

Big Bang in Laboratory

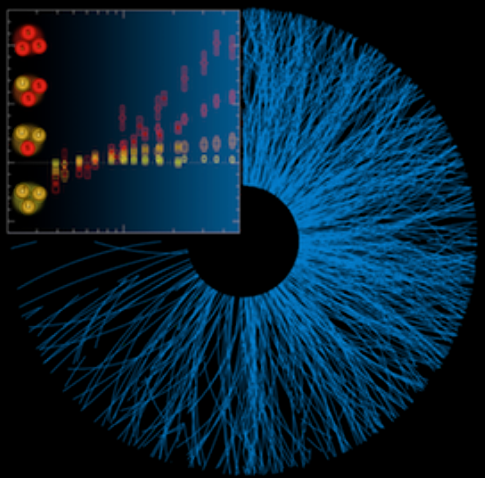


Heavy-ion collisions

$^{208}Pb + ^{208}Pb$
(208 nucleons)

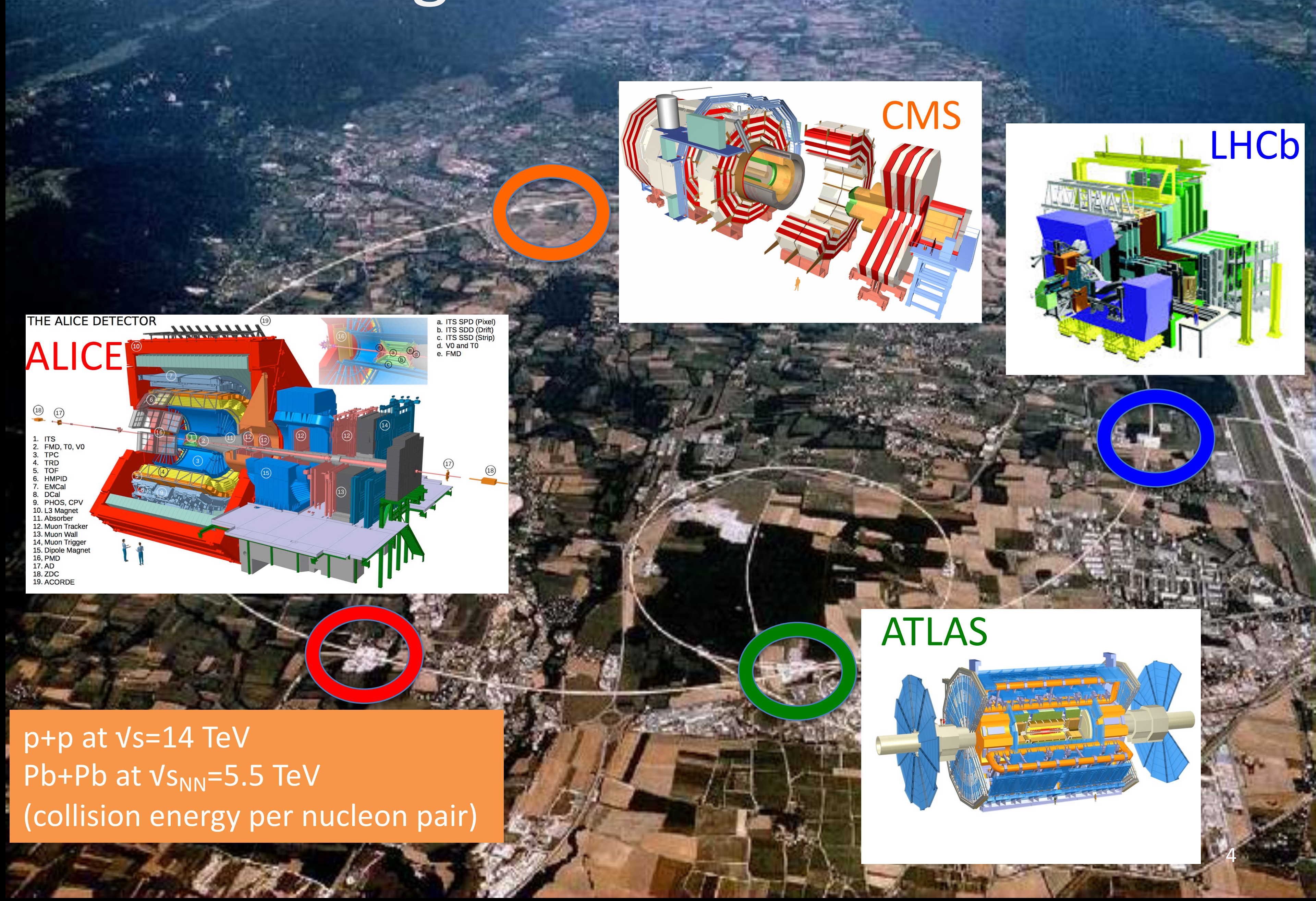


Time



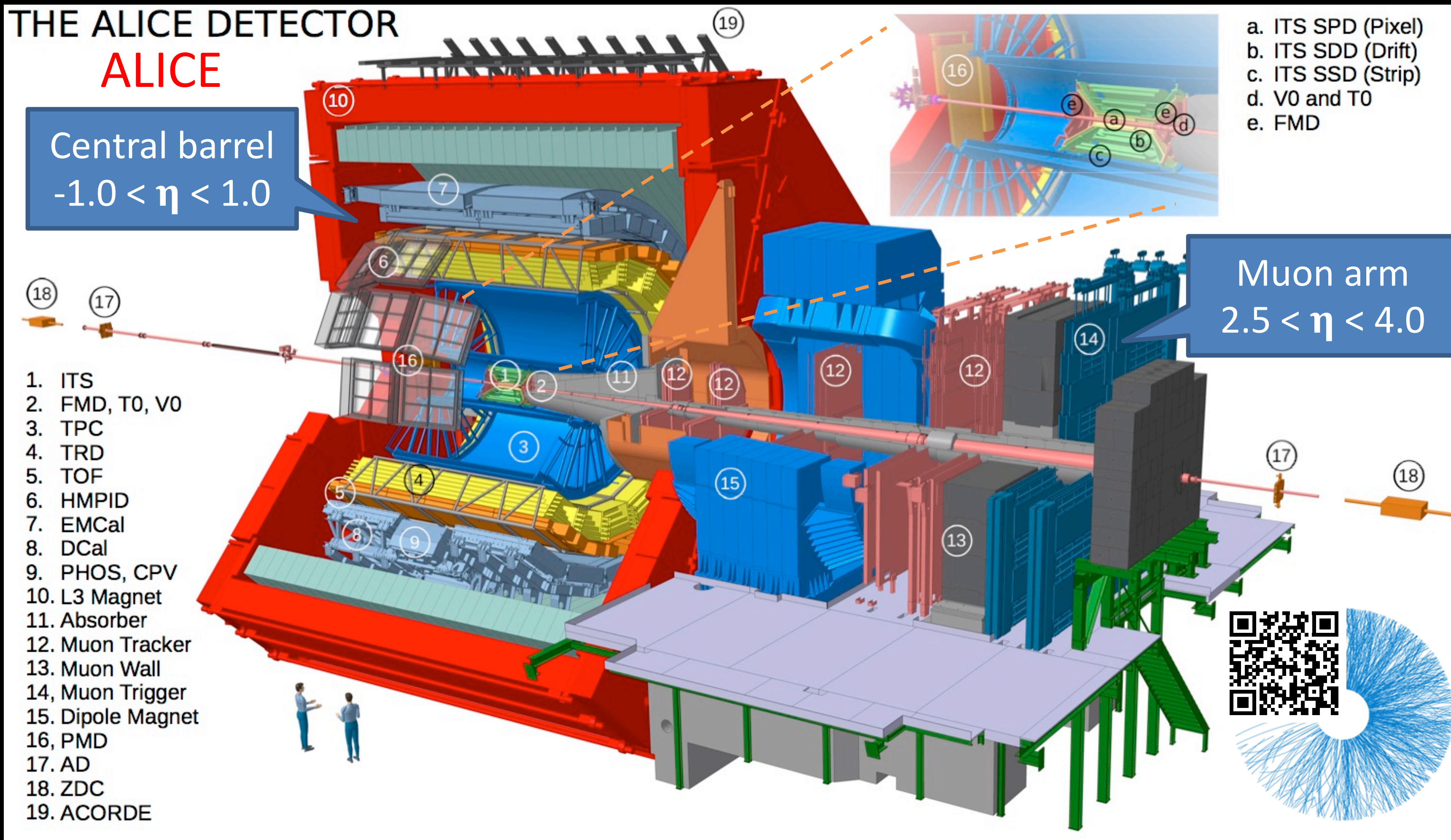
ALICE CASE

Large Hadron Collider



p+p at √s=14 TeV
Pb+Pb at $\sqrt{s_{NN}}$=5.5 TeV
(collision energy per nucleon pair)

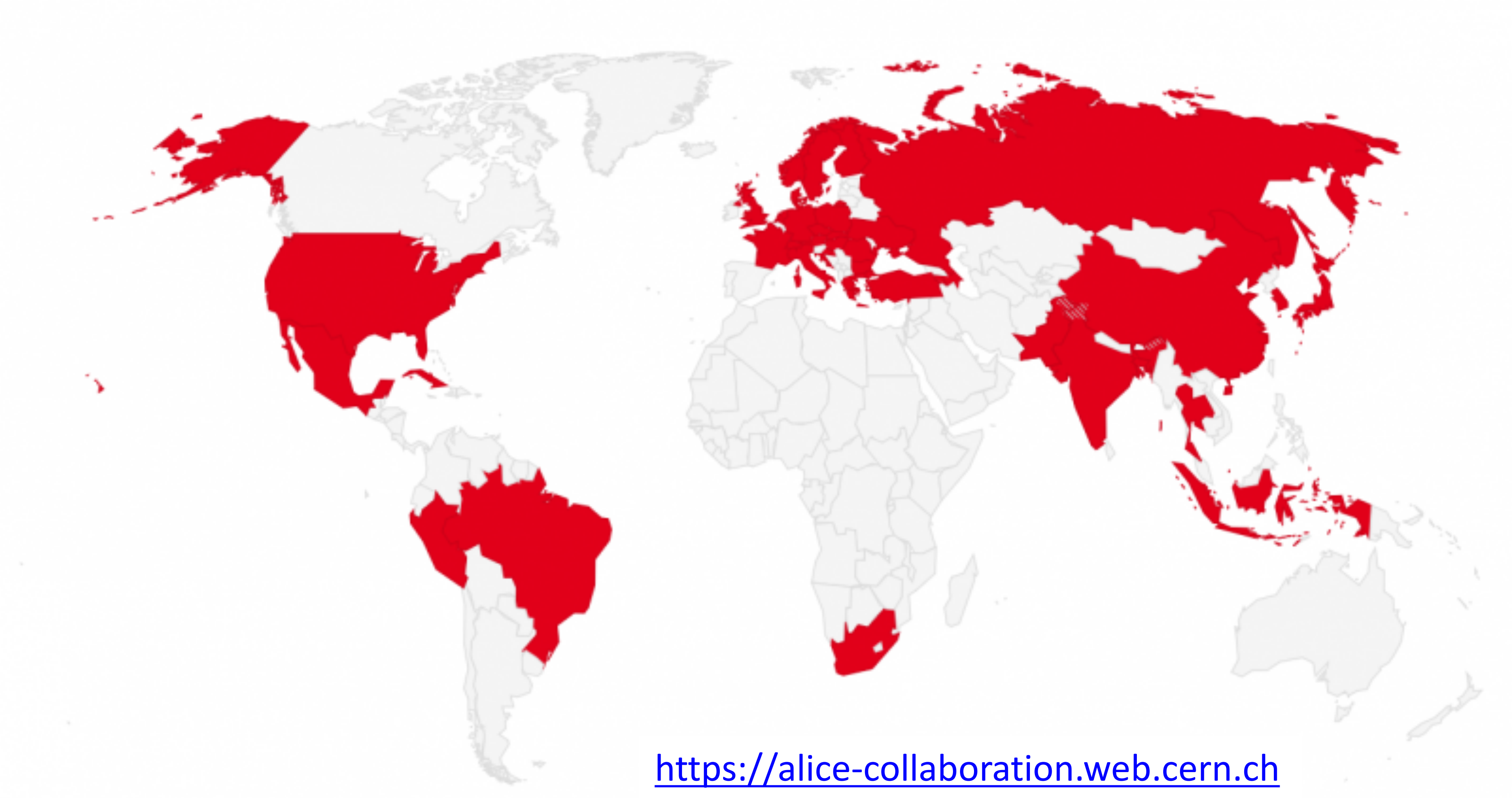
A Large Ion Collider Experiment



Physics observables, simulations, calibration and reconstruction, data quality control

ALICE Collaboration

39 countries, 175 institutes, 1927 members



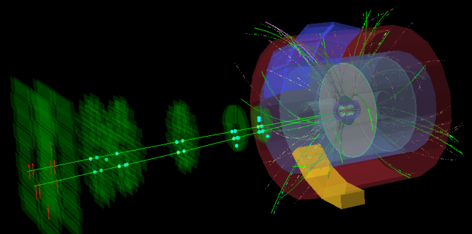
https://alice-collaboration.web.cern.ch

4 institutes from Poland: IFJ PAN, AGH, PW, NCBJ (~50 members, including students)

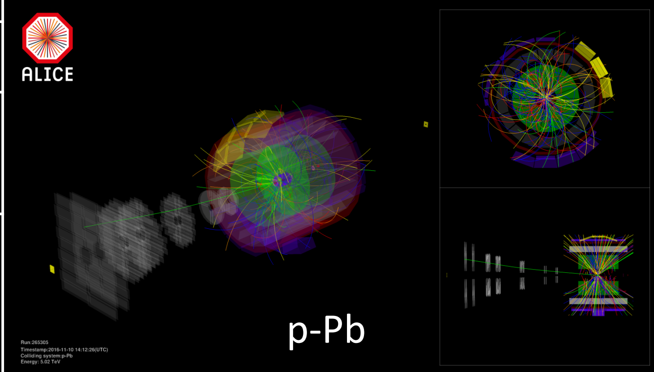
ALICE at work since 2009

LHC Run-1 (2009-2014) and Run-2 (2015-2018)

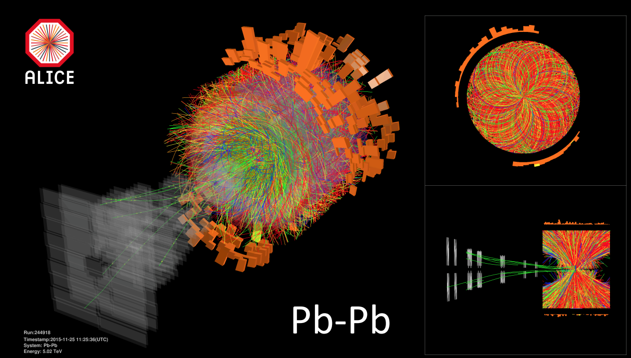
System	Year	$\sqrt{s_{NN}}$ (TeV)	L_{int}
Pb-Pb	2010-2011 2015 2018	2.76 5.02 5.02	~75 μb^{-1} ~250 μb^{-1} ~0.55 nb^{-1}
Xe-Xe	2017	5.44	~0.3 μb^{-1}
p-Pb	2013 2016	5.02 5.02, 8.16	~15 nb^{-1} ~3 nb^{-1}, ~25 nb^{-1}
pp	2009-2013 2015-2018	0.9, 2.76, 7, 8 5.02, 13	~200 μb^{-1}, ~100 μb^{-1}, ~1.5 pb^{-1}, ~2.5 pb^{-1} ~1.3 pb^{-1}, ~36 pb^{-1}



pp



p-Pb

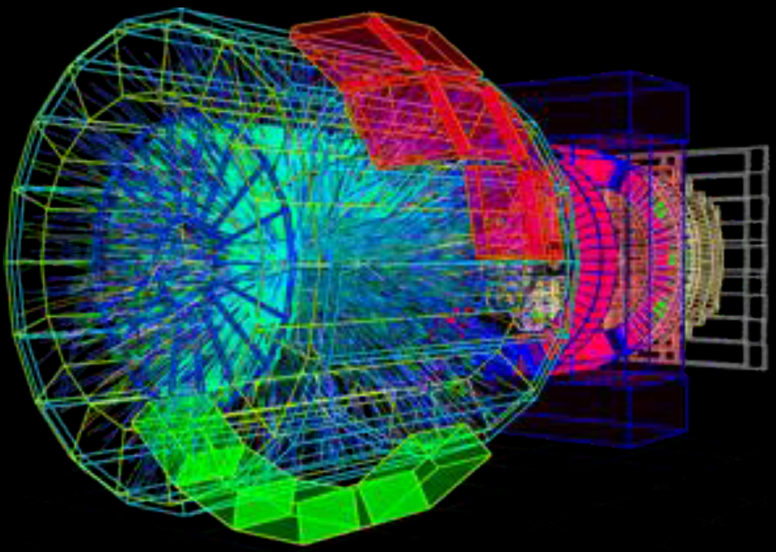


Pb-Pb

- Energy and system dependence studies of particle production are possible
- Large statistics of pp, p-Pb and Pb-Pb collisions at the same $\sqrt{s_{NN}}$

→ precise comparison studies
~ 300 publications till now

Data Processing in ALICE (past)

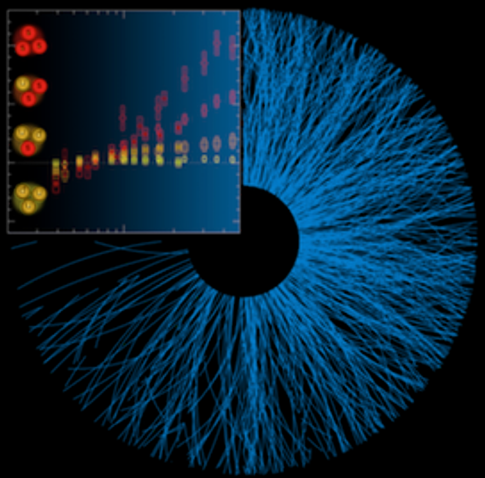


~12 GB/s

DAQ and HLT (High Level Trigger)
- ~1000 CPUs and FPGAs
- Data acquisition, online reconstruction and compression

~3 GB/s

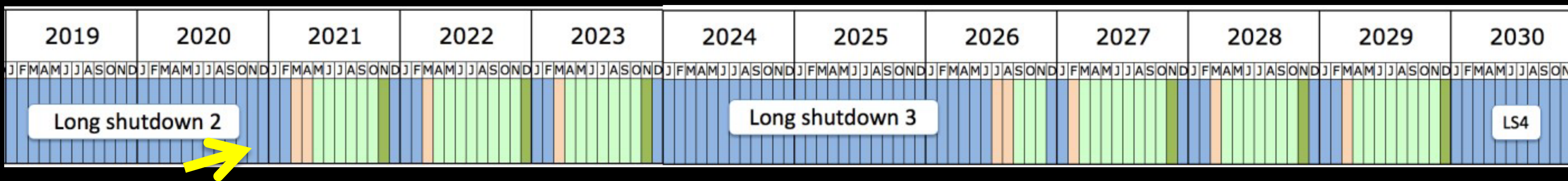
ALICE grid (AliEn)
- ~ 50 PB disk storage
- ~ 60000 CPUs
- Offline data calibration, reconstruction and analysis
- Monte Carlo simulations



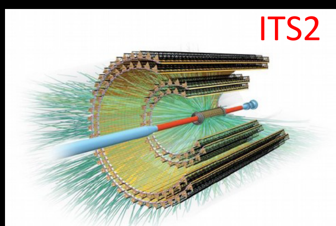
ALICE FUTURE

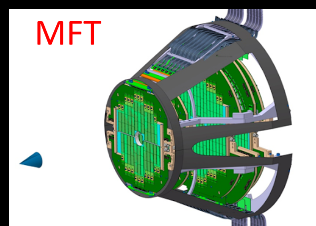
ALICE Future

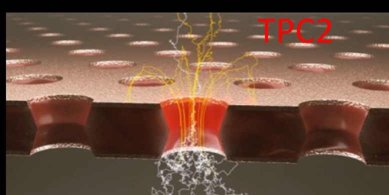
Run-3 COVID-19 → ~ a year of delay Run-4

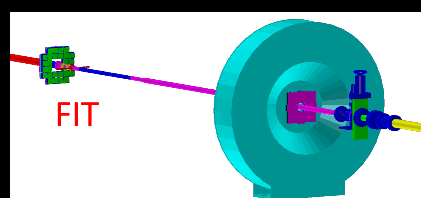


LS2: upgrades

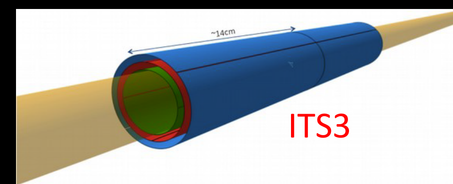


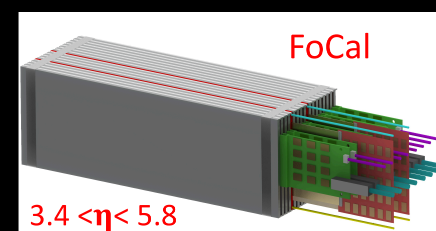






LS3: upgrades

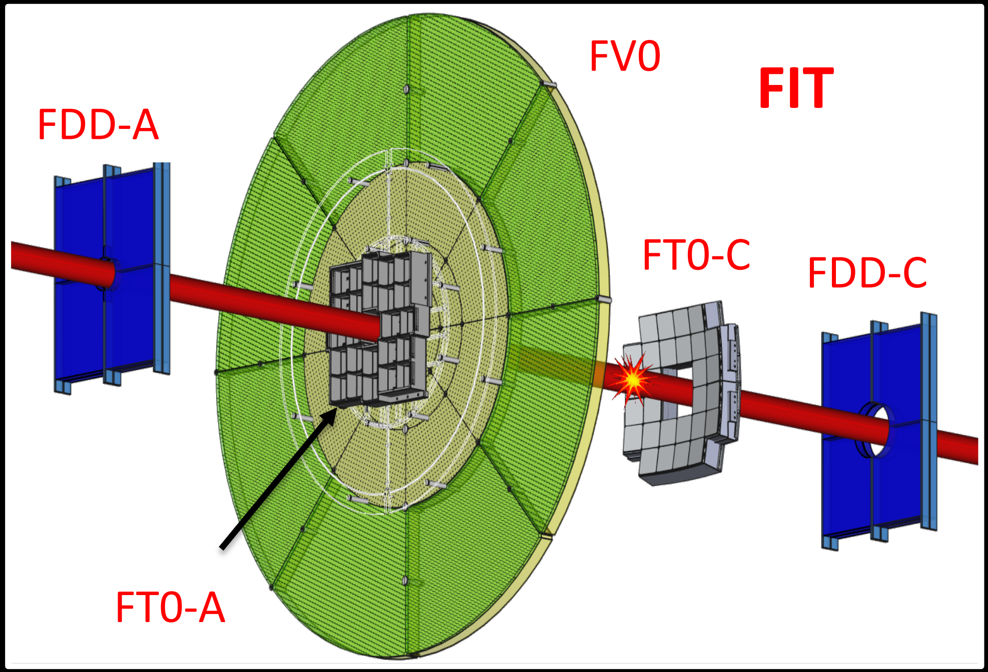


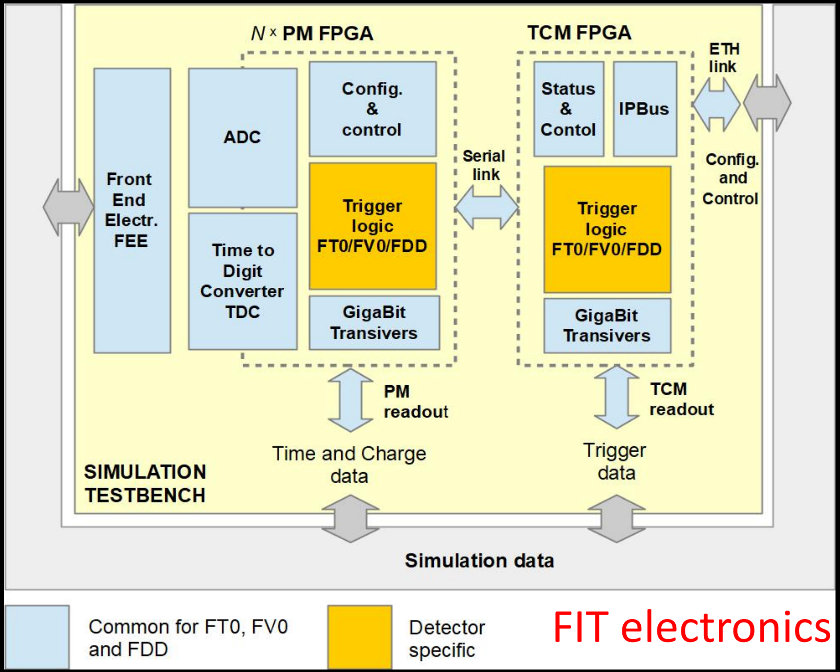


- Continuous data taking
  - Detector upgrade
  - Online-offline computing system upgrade
  - Readout electronics and trigger upgrade

Polish groups currently involved in the FIT, TPC2, FoCal, computing system and readout electronics developments

Fast Interaction Trigger (FIT)



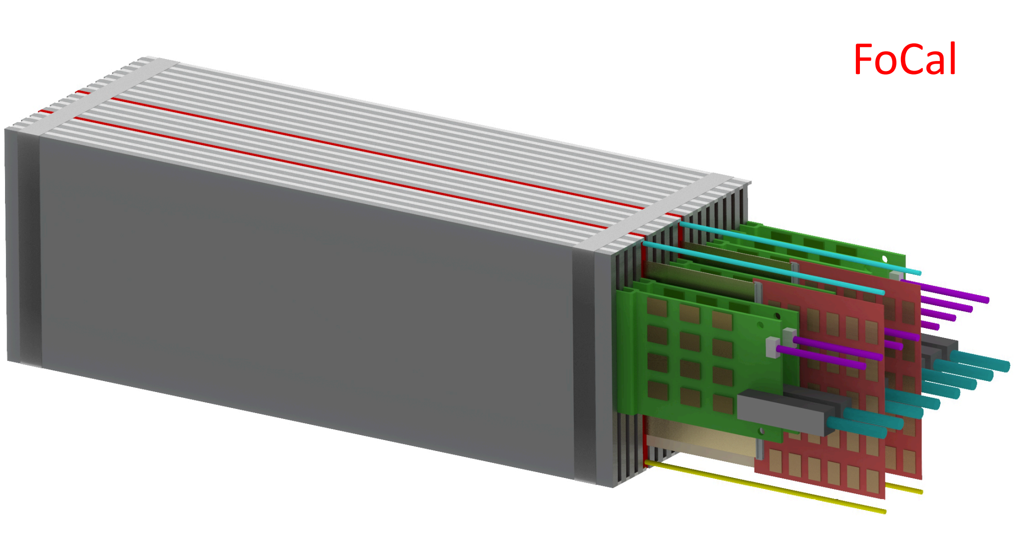


- Interaction time and vertex
- Particle multiplicity and reaction plane
- Particle production in ultra peripheral collisions and diffractive processes
- LHC luminosity and background monitoring
- Fast trigger (<425 ns) based on FPGA processing

IFJ PAN and AGH involved in the FIT detector and trigger developments!

Forward Calorimeter (FoCal)

To be ready for measurements in 2026-2030



- Forward electromagnetic and hadronic calorimeter
- High energetic particle jets measurements

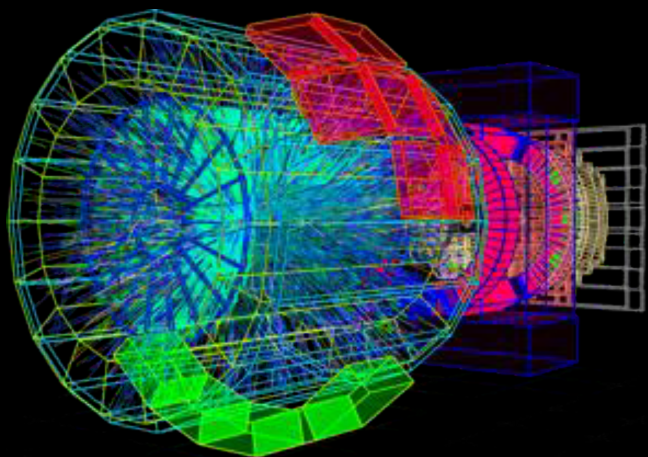
- A big opportunity to participate in all stages of detector developments
- Electronics R&D is still ongoing
- Fast trigger and data compression based on FPGA processing required
- Detector software to be developed
- First prototype tests with beam in 2021

IFJ PAN and AGH has recently joined the project!

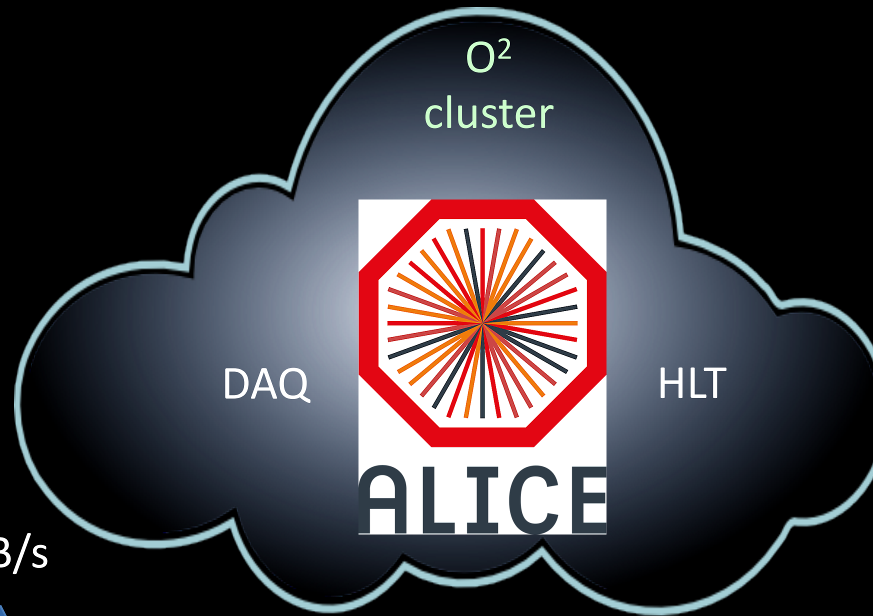
https://cds.cern.ch/record/2696471

ALICE O^2 Online-Offline Computing

- Continuous data readout
- ~100x more data in Run 3-4 than in Run 1-2



~3 TB/s



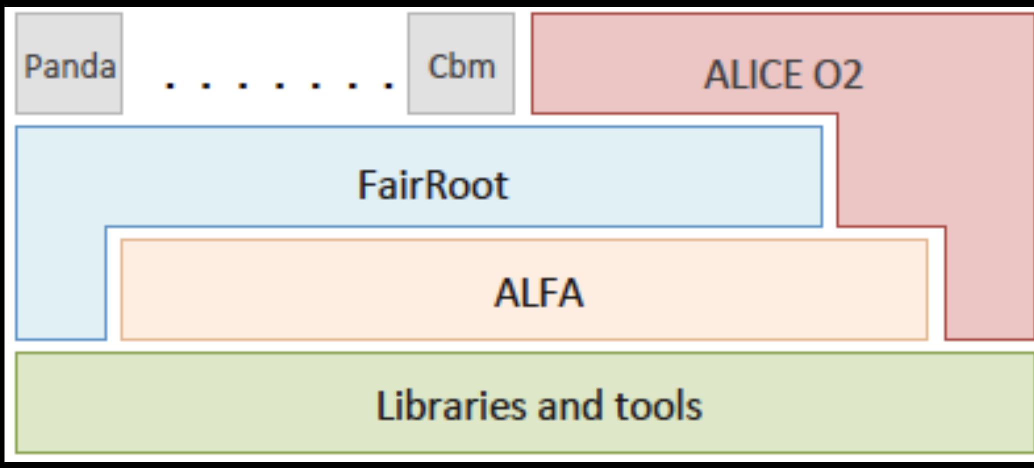
~100 GB/s

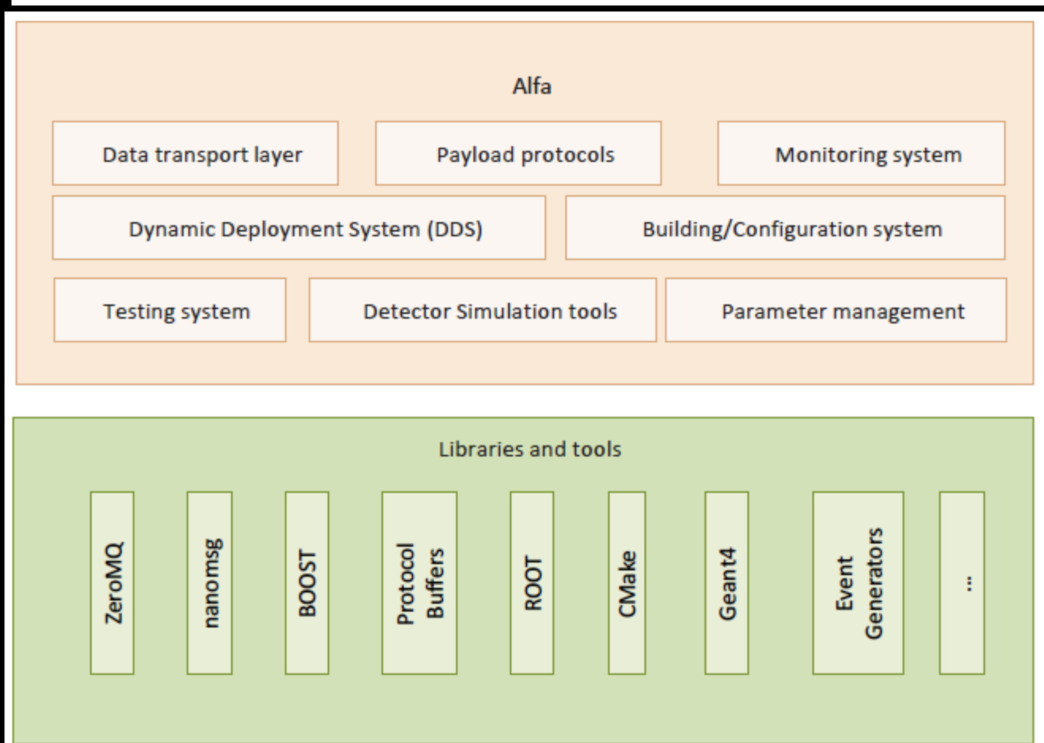
ALICE grid (AliEn 2.)

- ~ 0.5 Exabyte of data
- Offline data calibration, reconstruction and analysis
- Monte Carlo simulations

- New software framework O^2 and data model for parallel data processing
- Heterogeneous system FPGAs, GPUs, CPUs
- Optimized I/O

ALICE O^2 Software Ecosystem



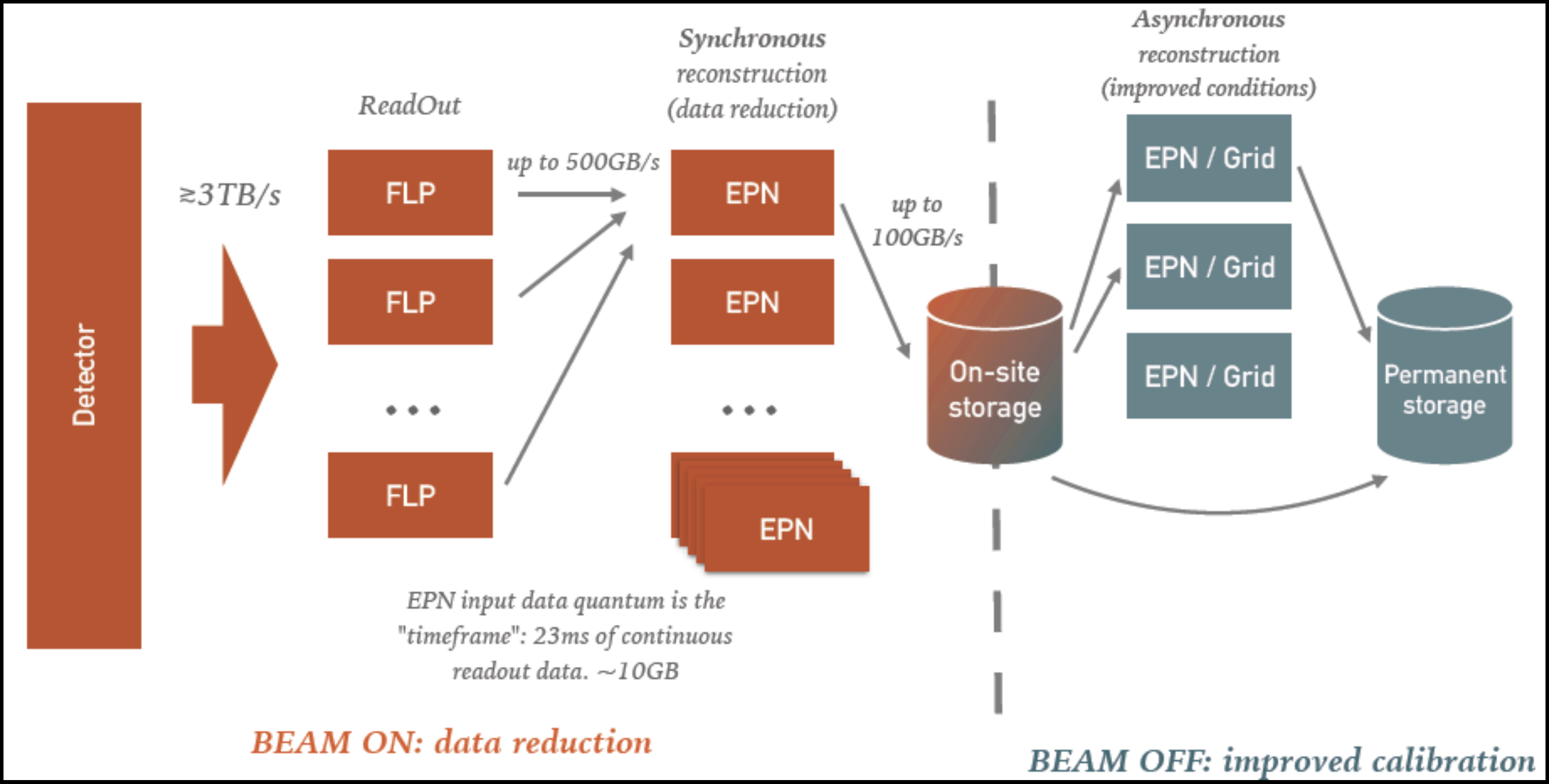


ALFA (ALICE-FAIR) concurrency framework for efficient parallel data processing on heterogeneous systems

- Data transport layer based on ZeroMQ/nanomsg
- Several data serialization/deserialization standards
  - BOOST serialization
  - ROOT streamers
  - …

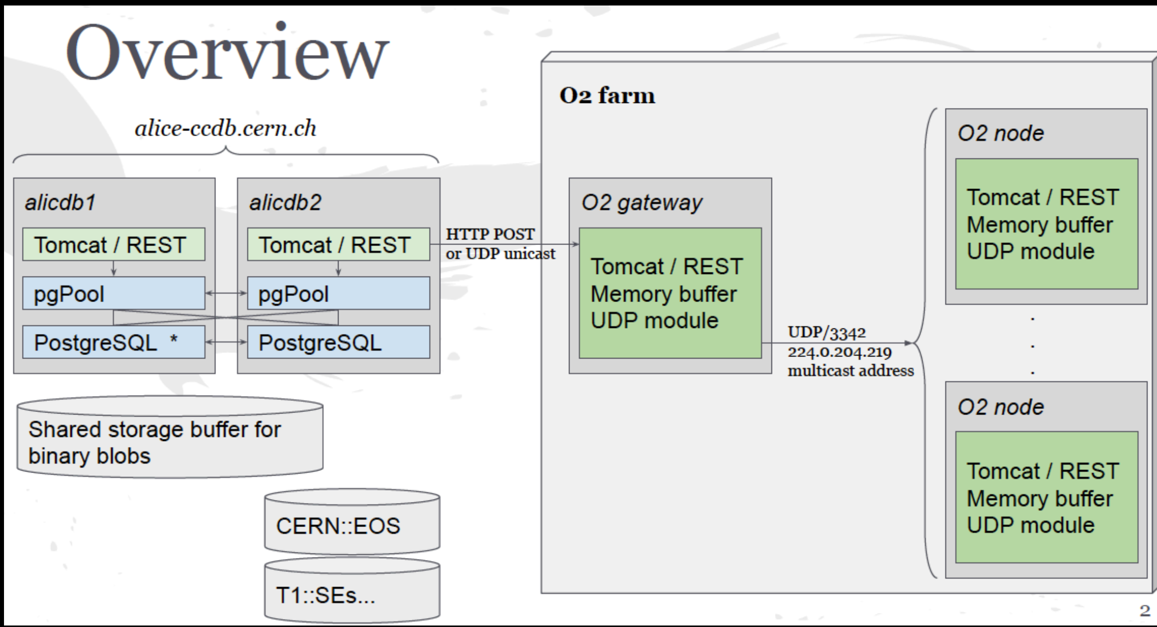
https://cds.cern.ch/record/2011297

ALICE Data Processing



- Raw data in time frames
- O^2 cluster (FLPs ~150 nodes, EPNs ~ 250 nodes)
- ~1000 modern GPUs on EPN nodes will be used for online tracking

ALICE Data Base - CCDB

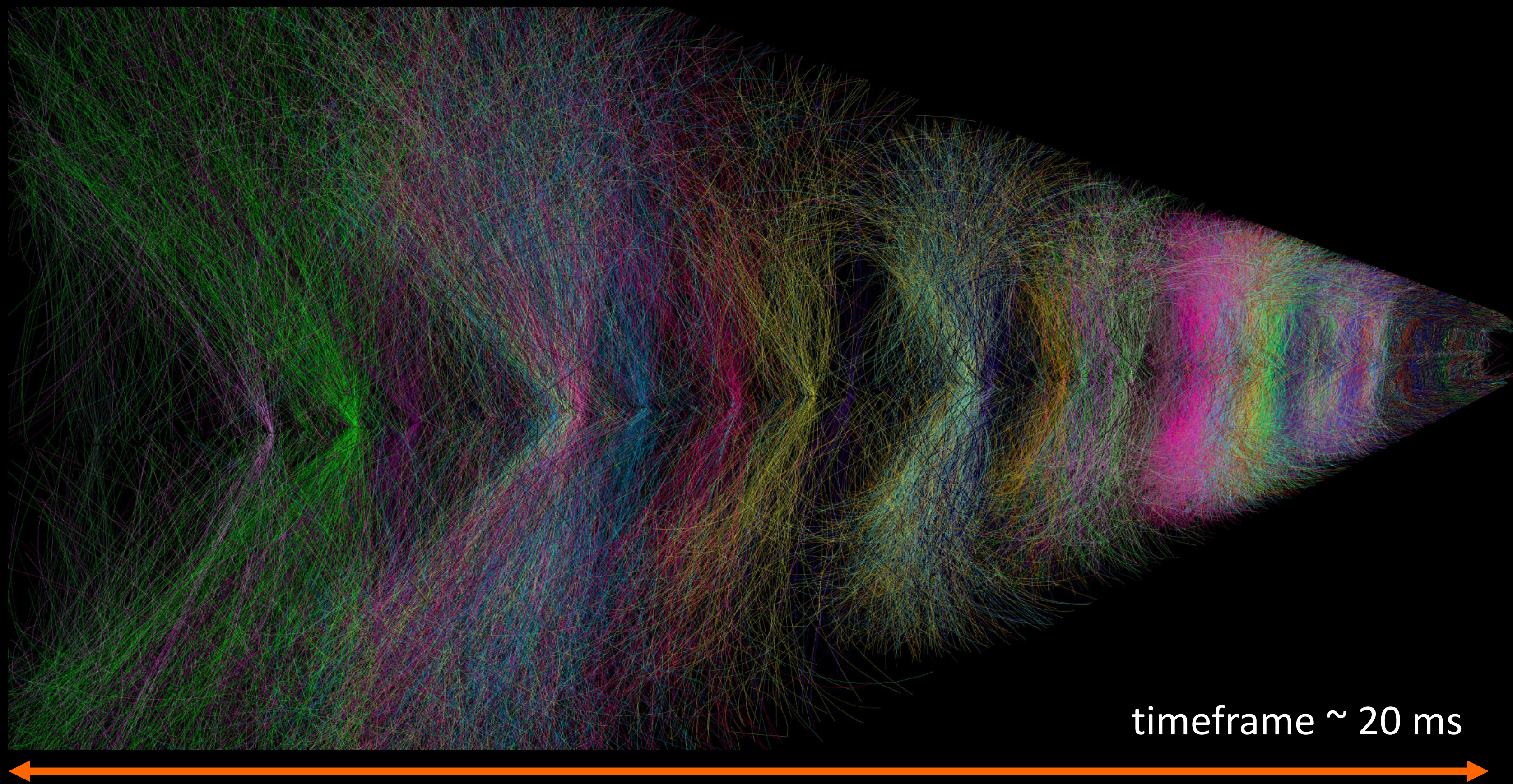


- General purpose data base for calibration and quality control objects
- Combination of PostgreSQL (for metadata) and CERN EOS distributed storage (for actual data)
- HTTP REST API hides the actual backend from the clients

AGH involved in the CCDB developments using test cluster at AGH Cyfronet

ALICE Tracking in Run-3 and Run-4

Continues data readout



- Several collision events in one timeframe
- Tracking of 20 ms timeframe needs ~20 s on modern GPU

ALICE Tracking Algorithms

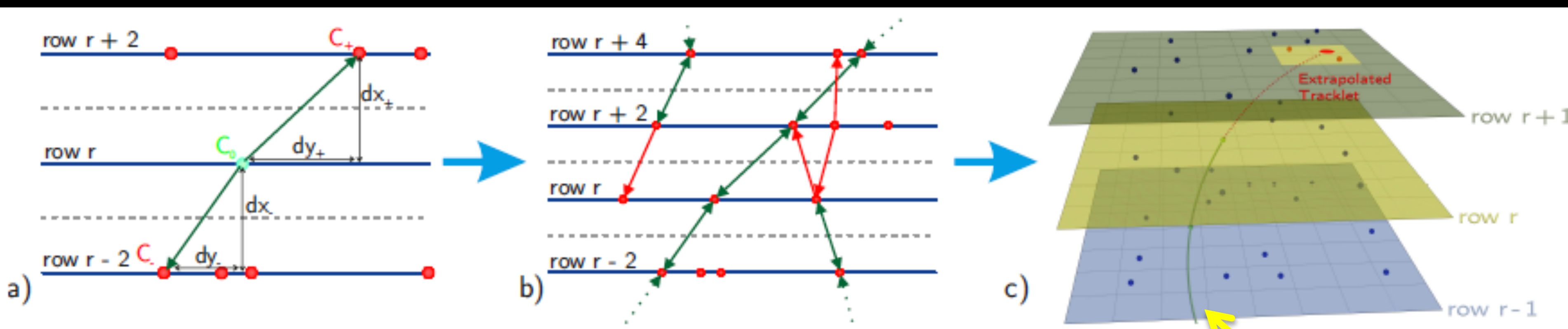
Cellular Automaton

Kalman filter

track forming

track concatenating

track following



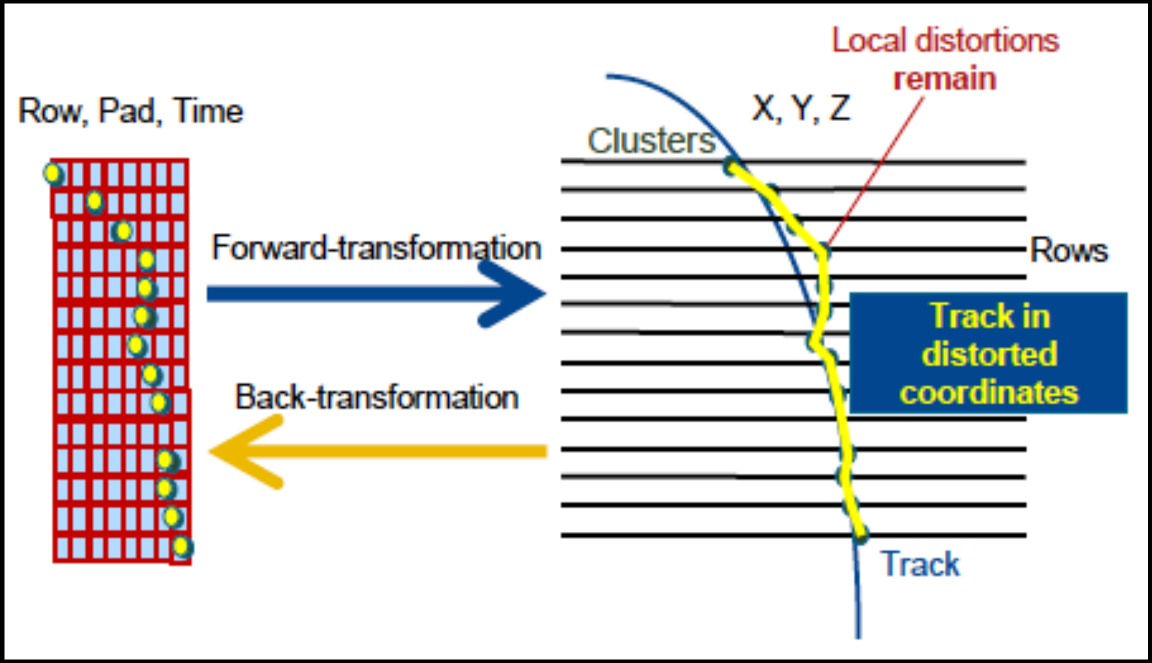
Helix (track model)

- Cellular Automaton for finding track candidates (track forming and concatenating)
- Kalman filter for track fitting and extrapolation (track following)

arXiv:1709.00618

ALICE Data Compression

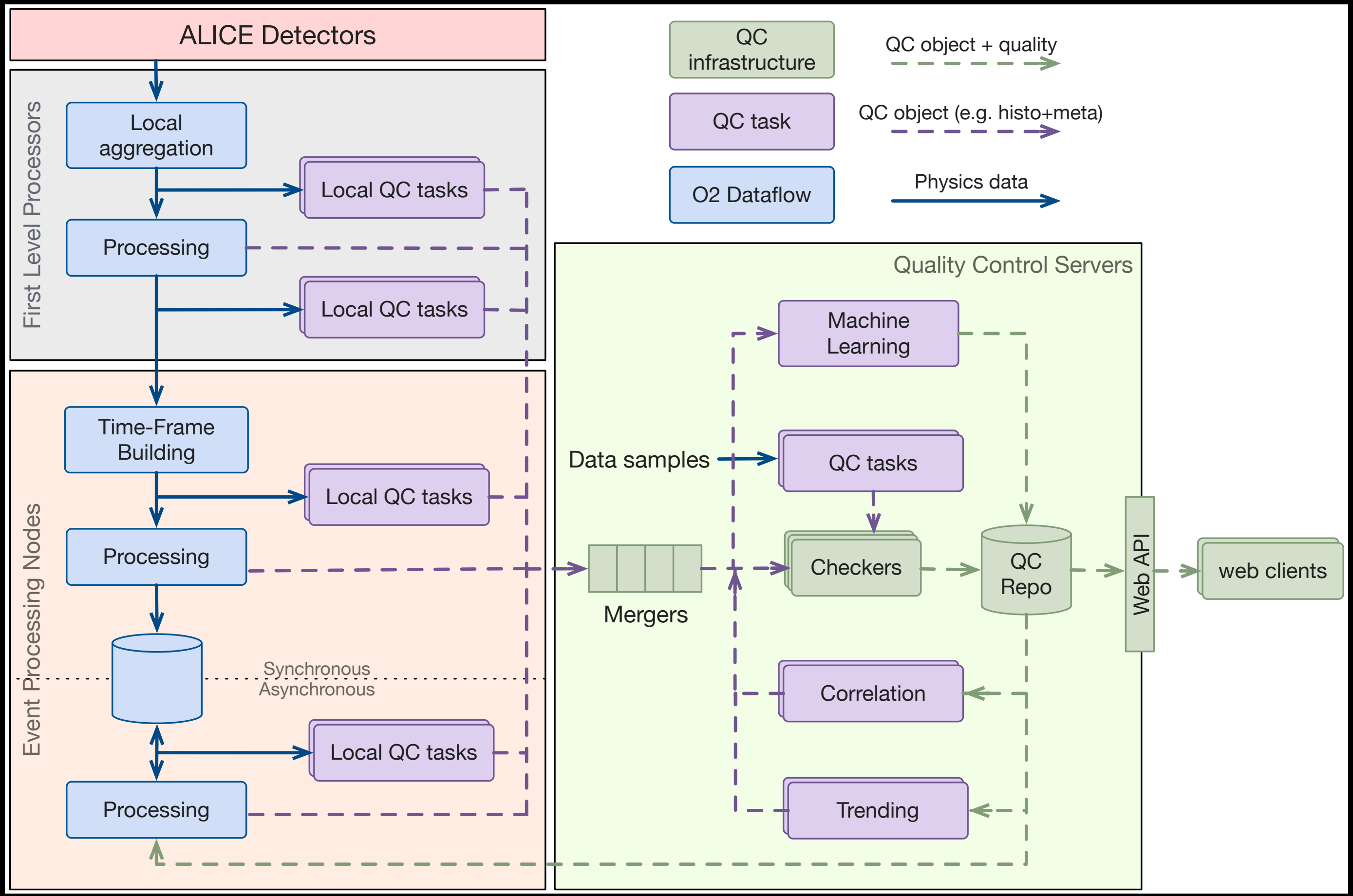
Required online data compression: ~3 TB/s → ~100 GB/s



- Lossless compression
  - Asymmetric Numeral Systems (J. Duda arXiv:1311.2540)
  - Storing only residuals to the clusters but not all cluster coordinates
- Non-lossless compression
  - Cluster finding
  - Removal of clusters of low momentum tracks

arXiv:1709.00618

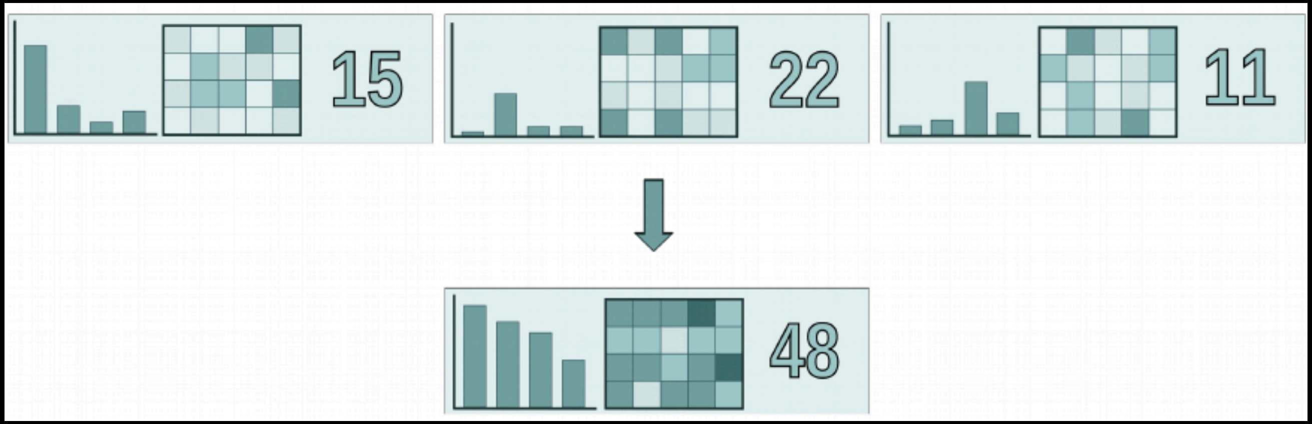
ALICE O^2 Data Quality Control

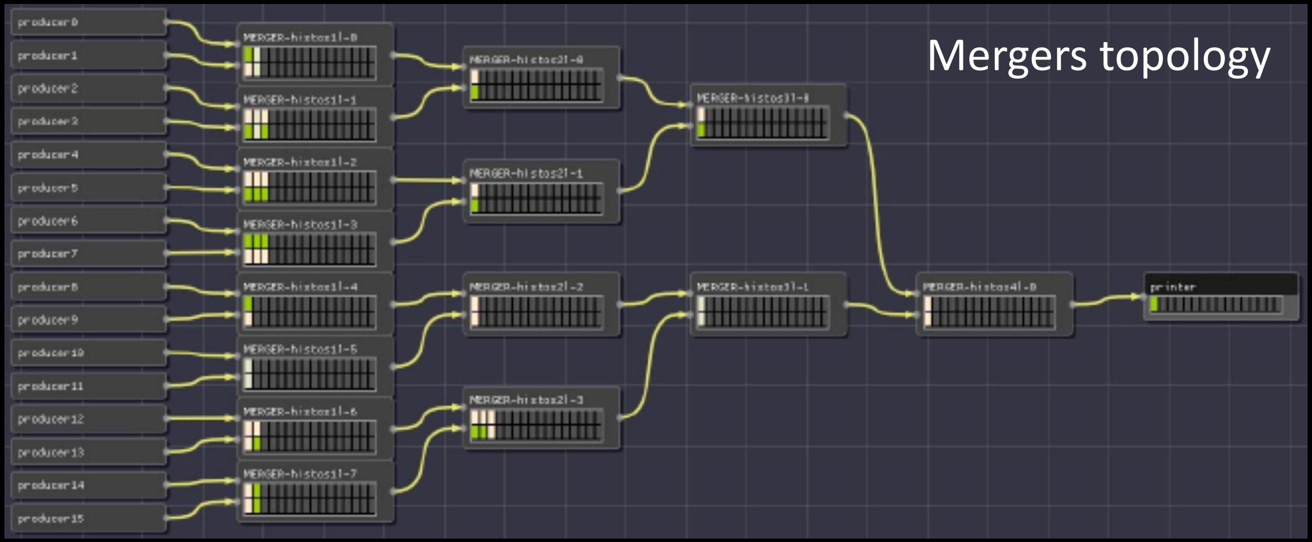


- ~150000 parallel processes
- ~400 QC objects updated per second

Quality Control system developed by IFJ PAN and AGH!

Data Mergers





- Fast data merging in ALICE O^2
- Merged data are checked for anomalies

Data Quality Control and Machine Learning

Classification of anomalies (needed: labeled dataset)

Regression of one value which may indicate anomalies (needed: dataset with known values)

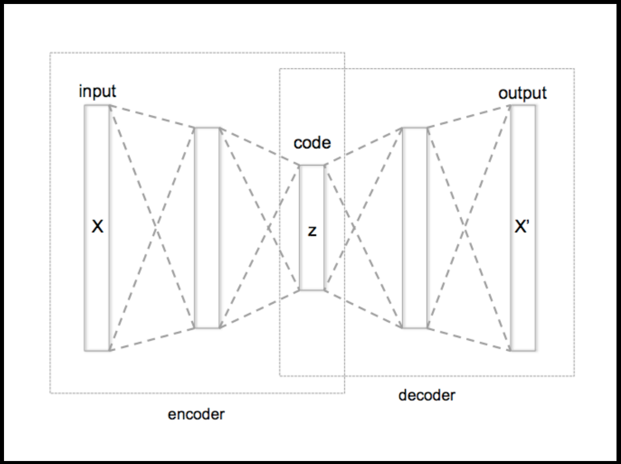
Clustering of unknown data and searching for outliers (needed: noisy data)

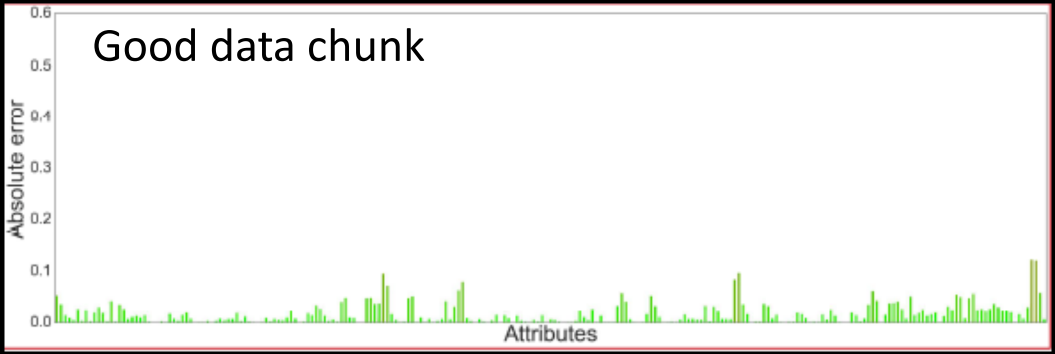
Dimensionality reduction for sparse data representation and searching for outliers (needed: high dimensional data)

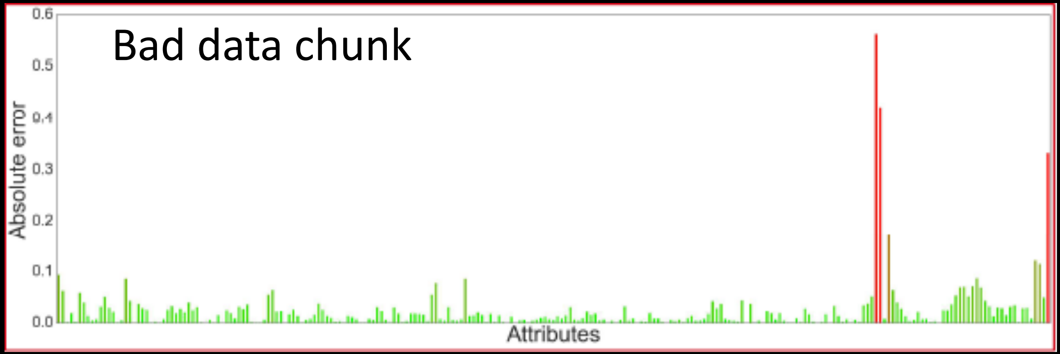
Unsupervised Learning with Autoencoders (example)

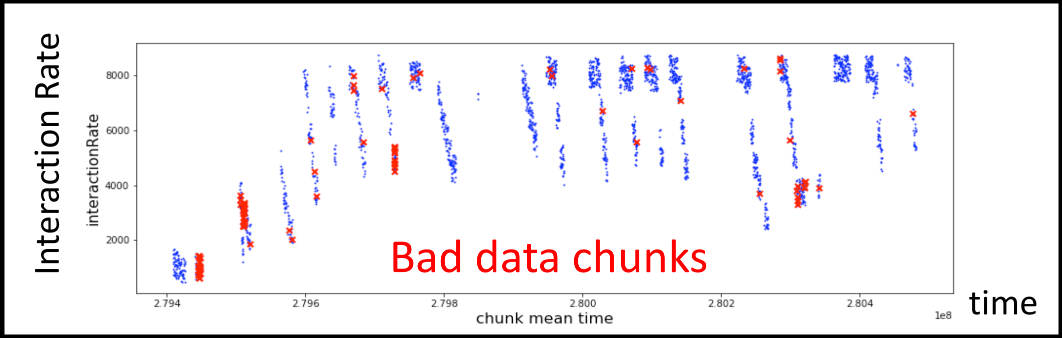
- Thousands of data chunks
- Hundreds of input parameters

- Autoencoder with 5 fully connected layers

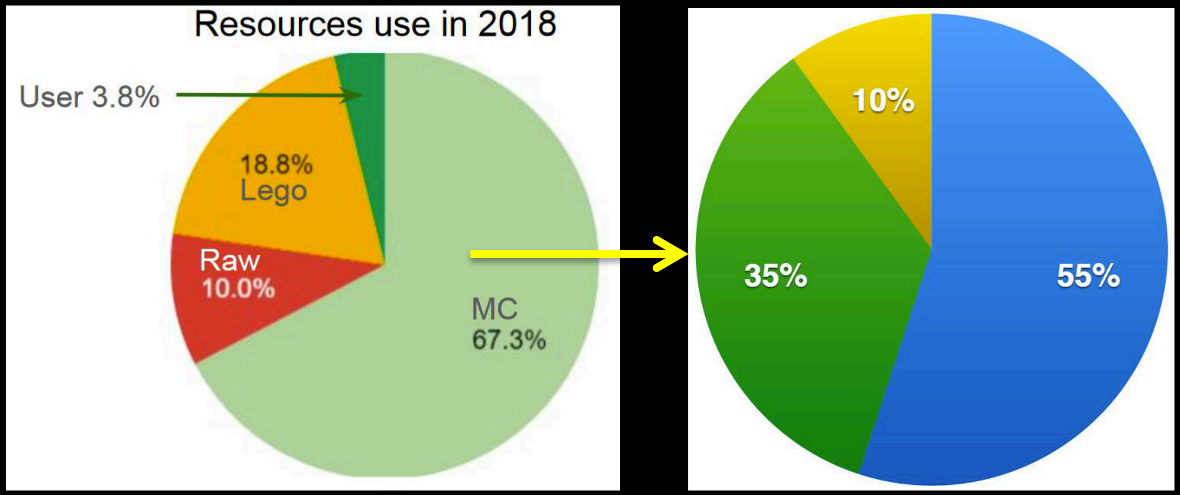








Monte-Carlo Simulations in ALICE



- particle propagation through detector
- digitization (detector response)
- rest (rec., QA/QC,…)

- More than 2/3 resources spend on MC simulations (Geant3/Geant4) in Run-2
- Expected 100 times more data in Run-3 and Run-4

  → Cannot be covered with the current simulation software

- Possible directions: fast detector simulations, embedding, optimizing current software…

Summary

- Plenty of interesting physics to be measured in next ~ 10 years
- Excellent environment for scientific and other developments
- Challenging tasks to be solved in collaboration between physicists and engineers
- Custom hardware and software developments
- CERN career programmes and common grant applications
- Strong overlap with industry

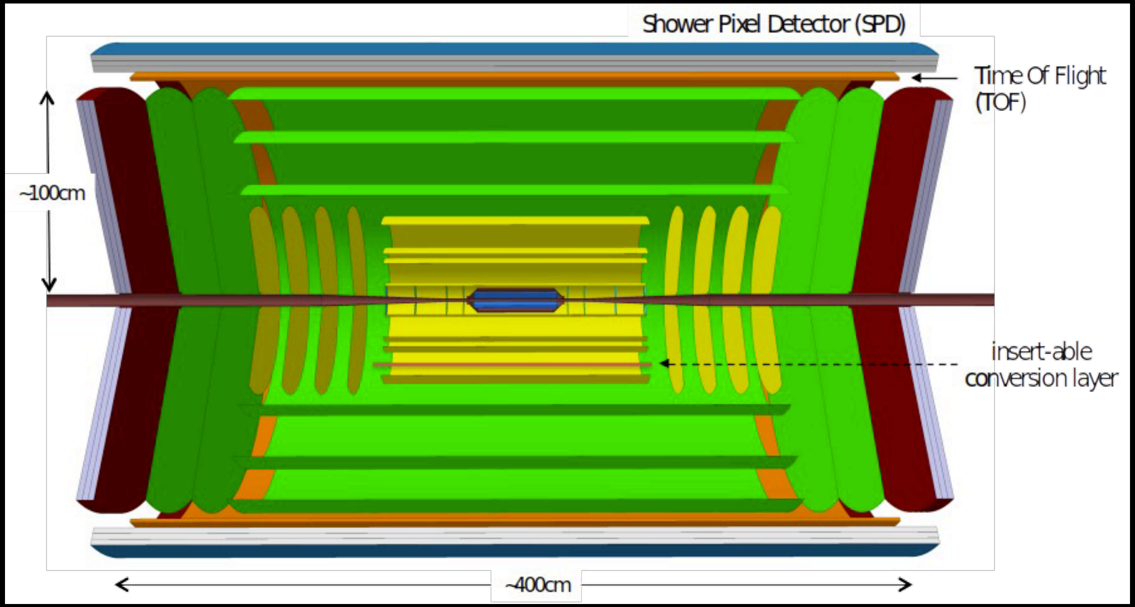
You are welcome to join and participate in the developments!

Contact:
Jacek Otwinowski (IFJ PAN): Jacek.Otwinowski@ifj.edu.pl
Jacek Kitowski (AGH): kito@agh.edu.pl

Outlook

Possibility to extend heavy ion measurements at the LHC beyond 2030

- Future Heavy Ion Detector [arXiv:1902.01211]



Physics:

- Heavy flavor and quarkonia
- Low mass dielectrons
- Soft photons and hadrons
- BSM

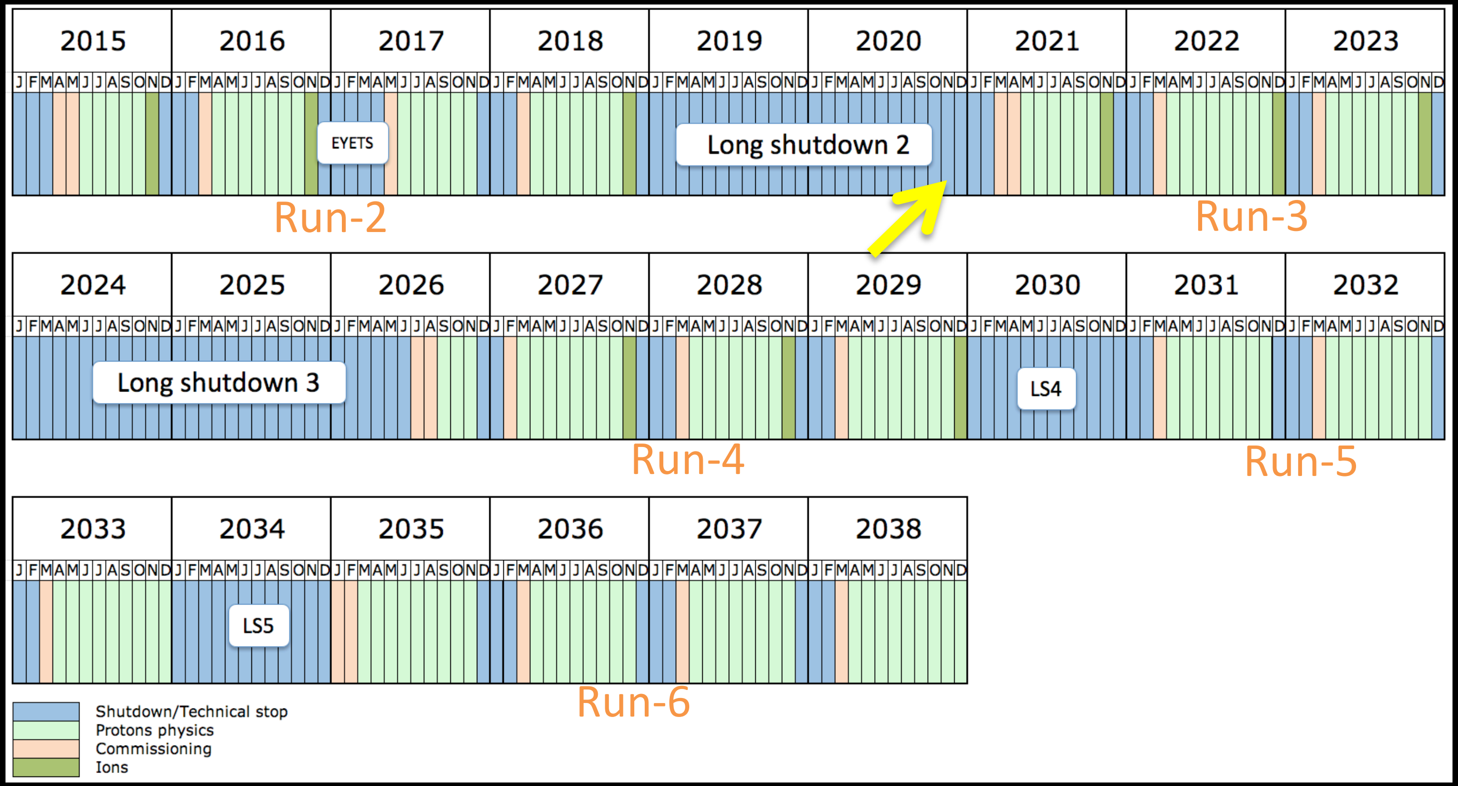
Design guidelines:

- All silicon detector
- High-rate capability: ~10^{34}/cm^2/s (~50x Run3-4)
- Vertex spatial resolution: ~1 μm
- Tracking over wide kinematic range
  - 30 MeV/c < p_T < 10 GeV/c
  - $|\eta| < 4.0$

You are welcome to join and participate in the developments!

Backup

LHC Running Program

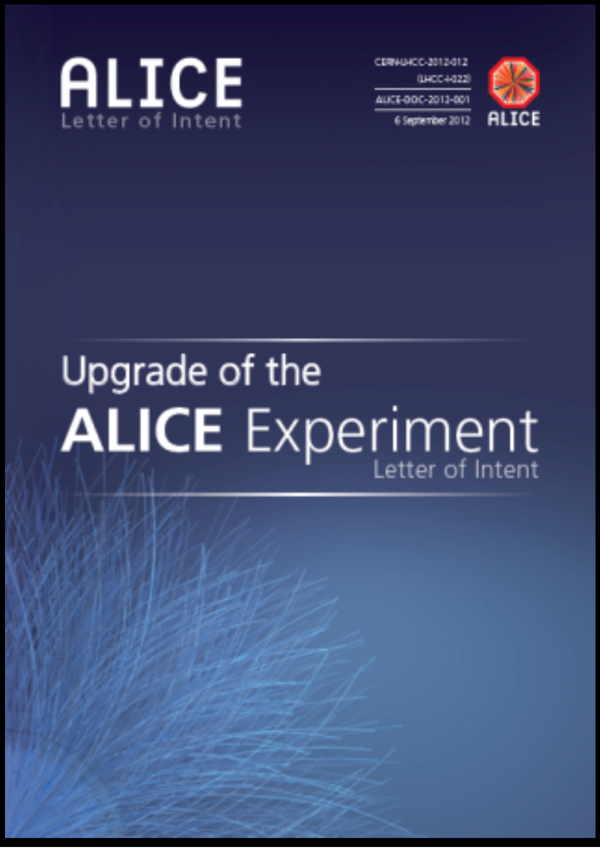


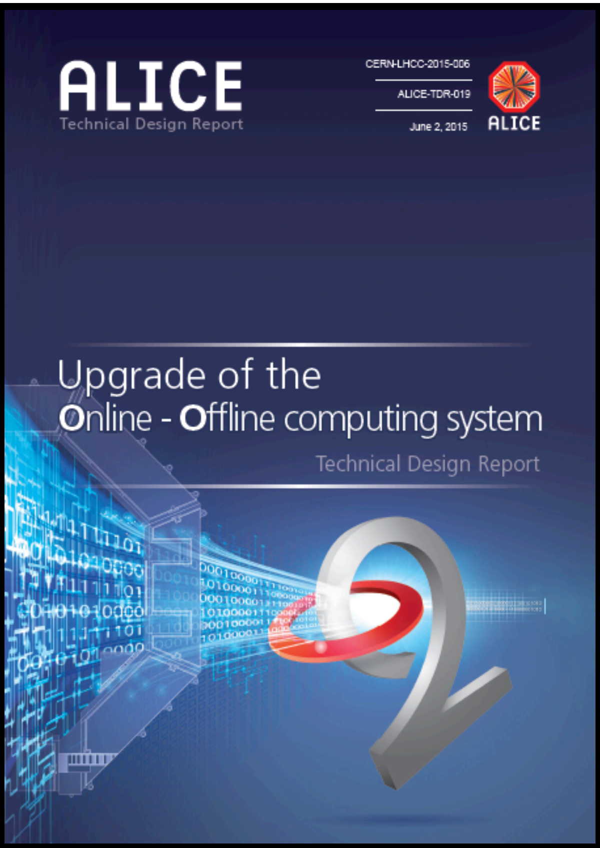
ALICE works on preparations for Run-3 and Run-4

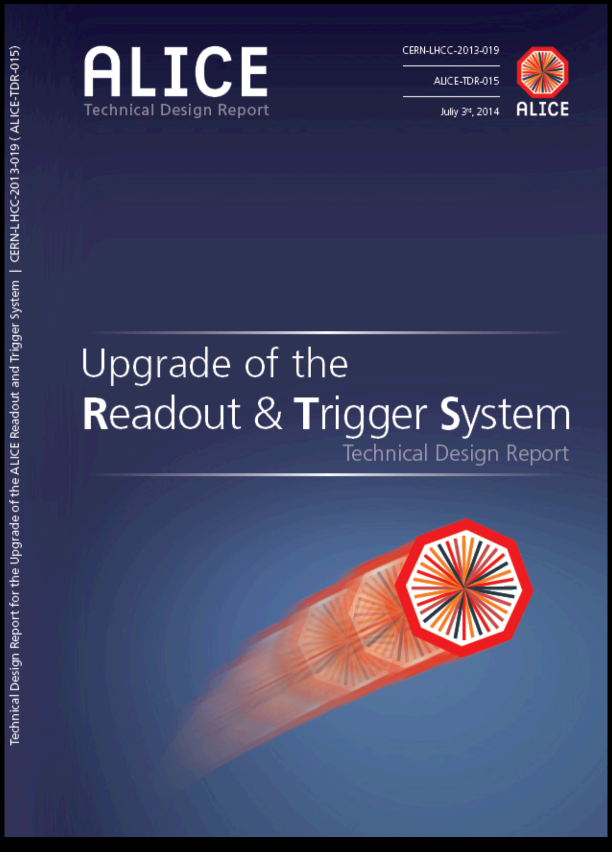
- Detector upgrade
- Online-offline computing system upgrade
- Readout electronics and trigger upgrade

ALICE Upgrade Documents

https://cds.cern.ch/record/2011297



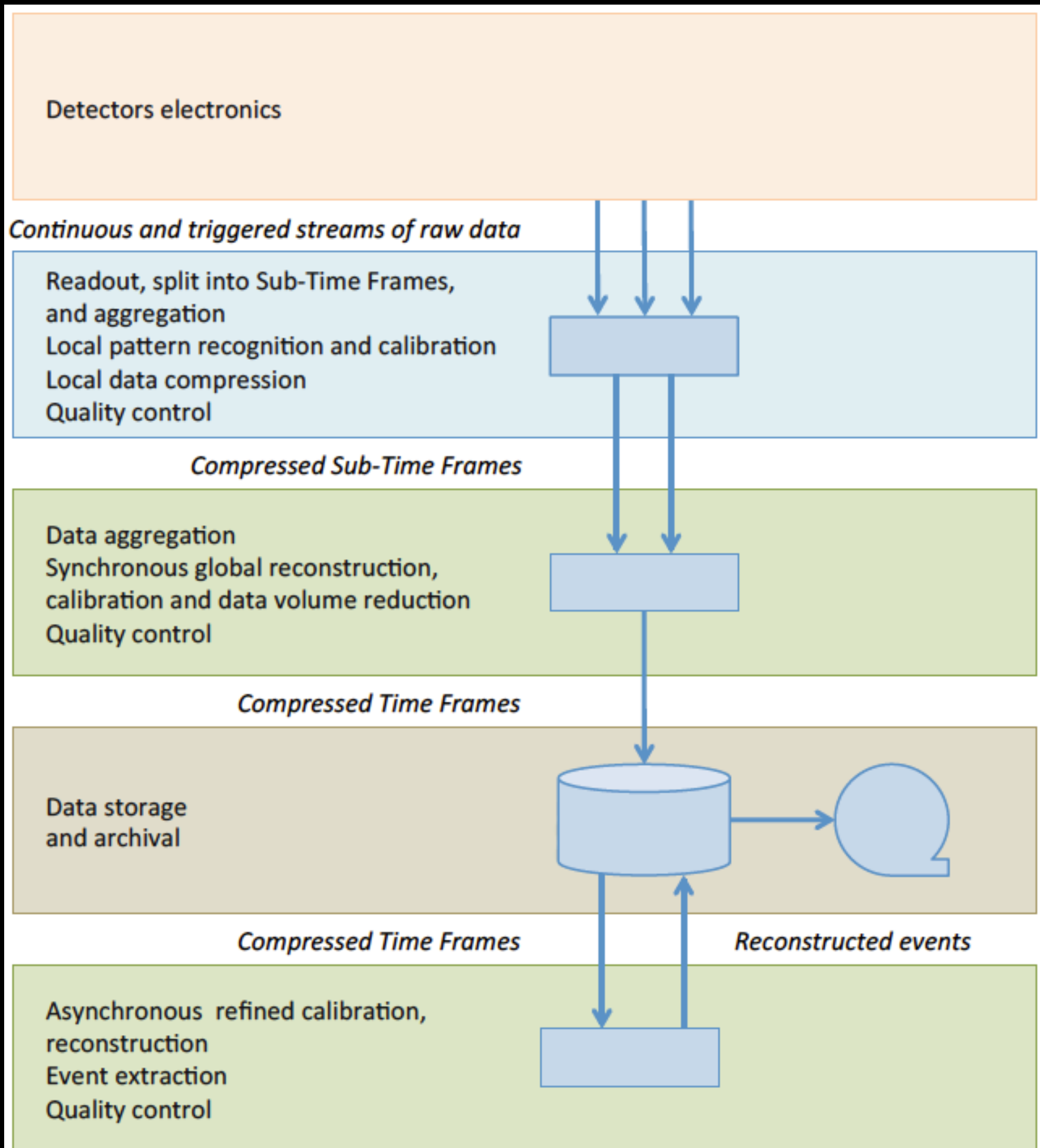




http://cds.cern.ch/record/1475243

http://cds.cern.ch/record/1603472

ALICE O^2 Functional Flow



https://cds.cern.ch/record/2011297

- Raw data in time frames
- O^2 cluster (FLPs, EPNs)

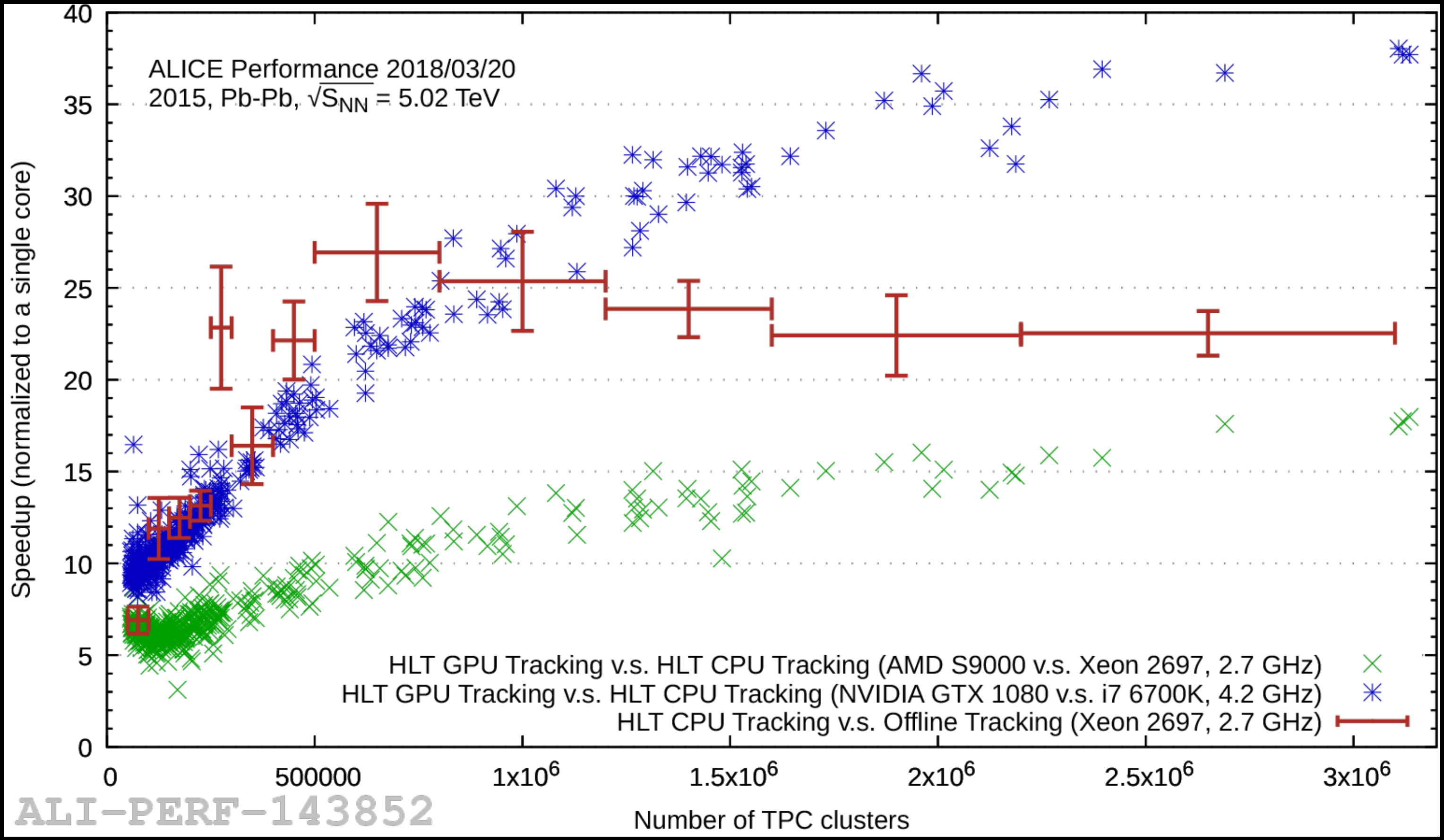
First Level Processors (FLPs)
O(150)

Event Processing Nodes (EPNs)
O(250)

Data movers (O^2 cluster -> grid)
ALICE grid storage servers
Data Base servers

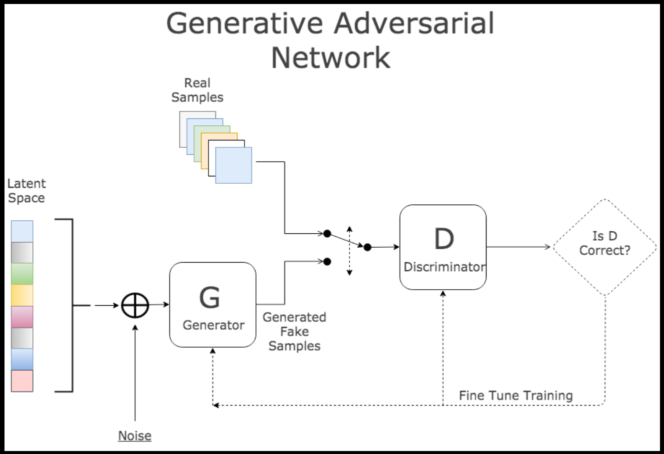
ALICE grid nodes and EPNs

ALICE Tracking Performance on GPUs



- Modern GPU replaces 40 CPU cores (4.2 GHz)
- 20 ms timeframe tracking needs ~20 s on GPU

Cluster Simulations with Generative Adversarial Networks (GANs)



- Training on original reconstructed data
- Conditional Deep Convolutional GAN

- Simulation speed-up ~25 (CPU), ~250 (GPU)
- But we are not there yet…