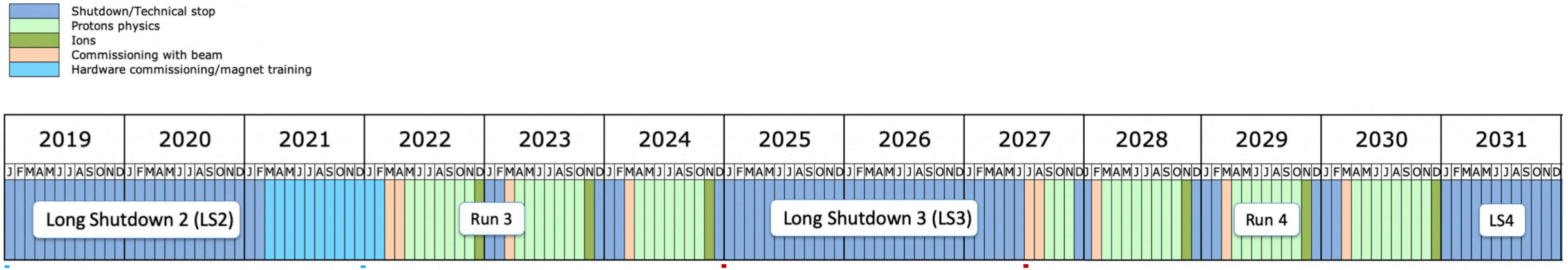
Brief introduction of ALICE O2 framework

Biao Zhang

Central China Normal University

Nov 26, 2021

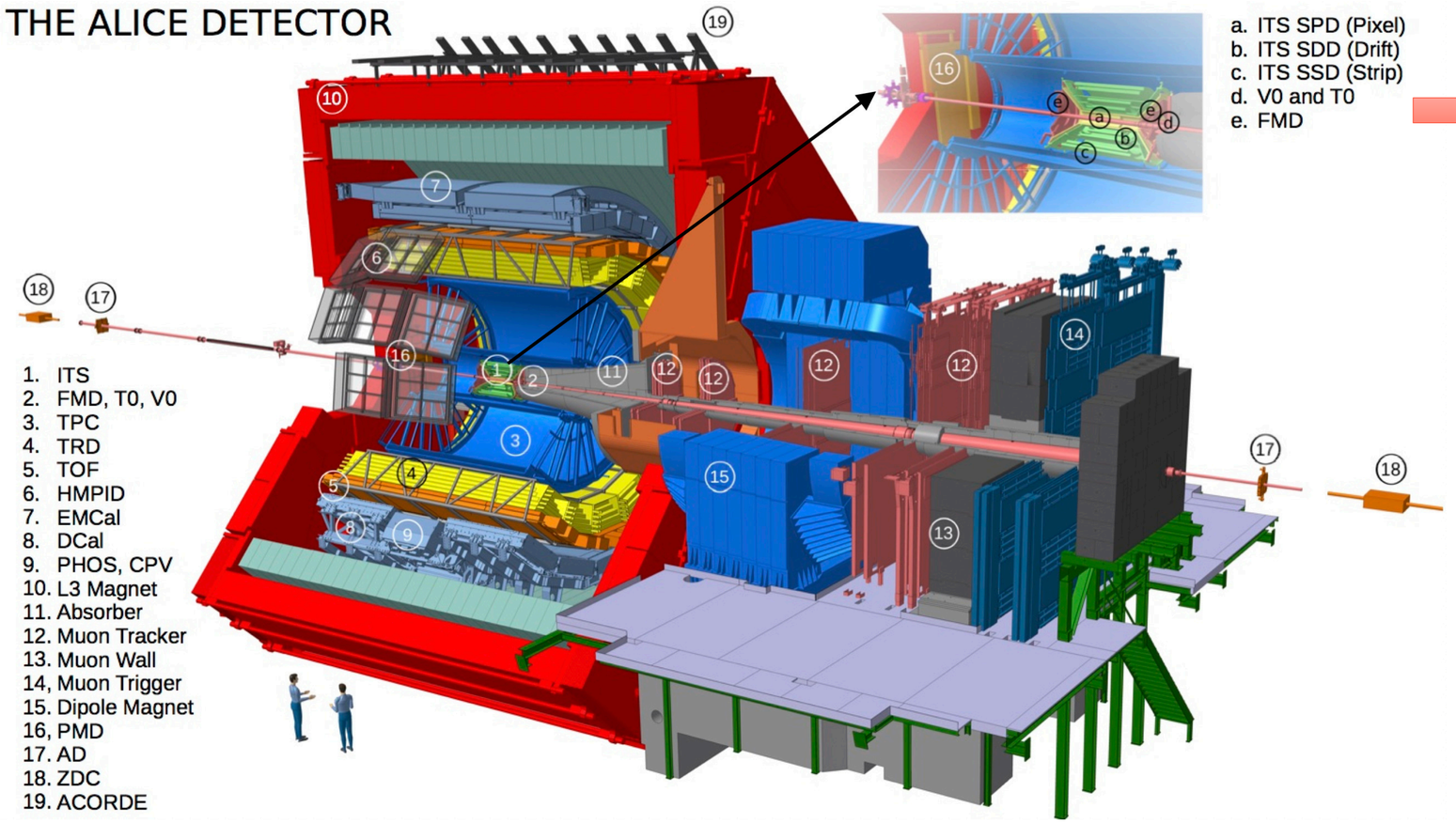
Motivation

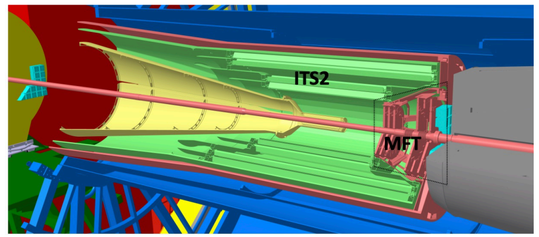


- **Physics goal in ALICE Run3**
  - Precise measurements of heavy-flavour hadrons down to very low pt
- **How to achieve?**
  - Improve the spatial resolution of detector
  - 100 times larger than the statistics collected in Run 1 and Run 2

More statistics
+ High resolution !!

ALICE upgrade in LS2

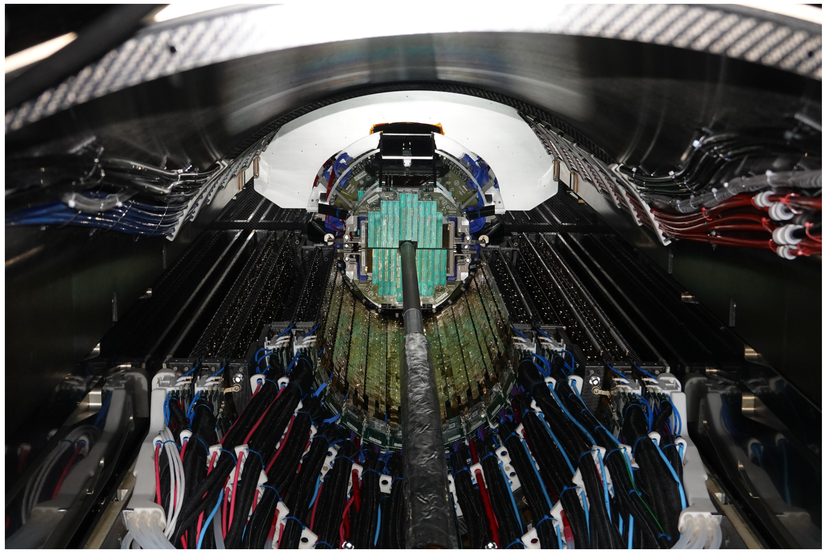




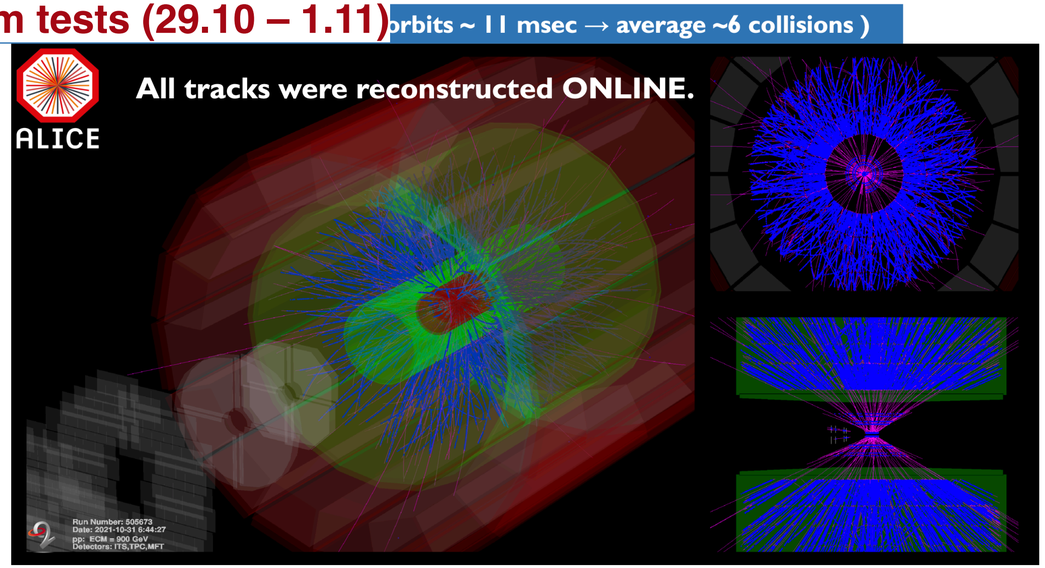
- New silicon tracker (ITS2 & MFT)
- TPC Readout planes using GEM
- New Fast Interaction Trigger (FIT)
- Upgrade readout of all other detectors
- New Online/Offline system (O2)

ALICE upgrade in LS2

Have installed successfully in May, 2021!!



Pilot beam tests (29.10 – 1.11)

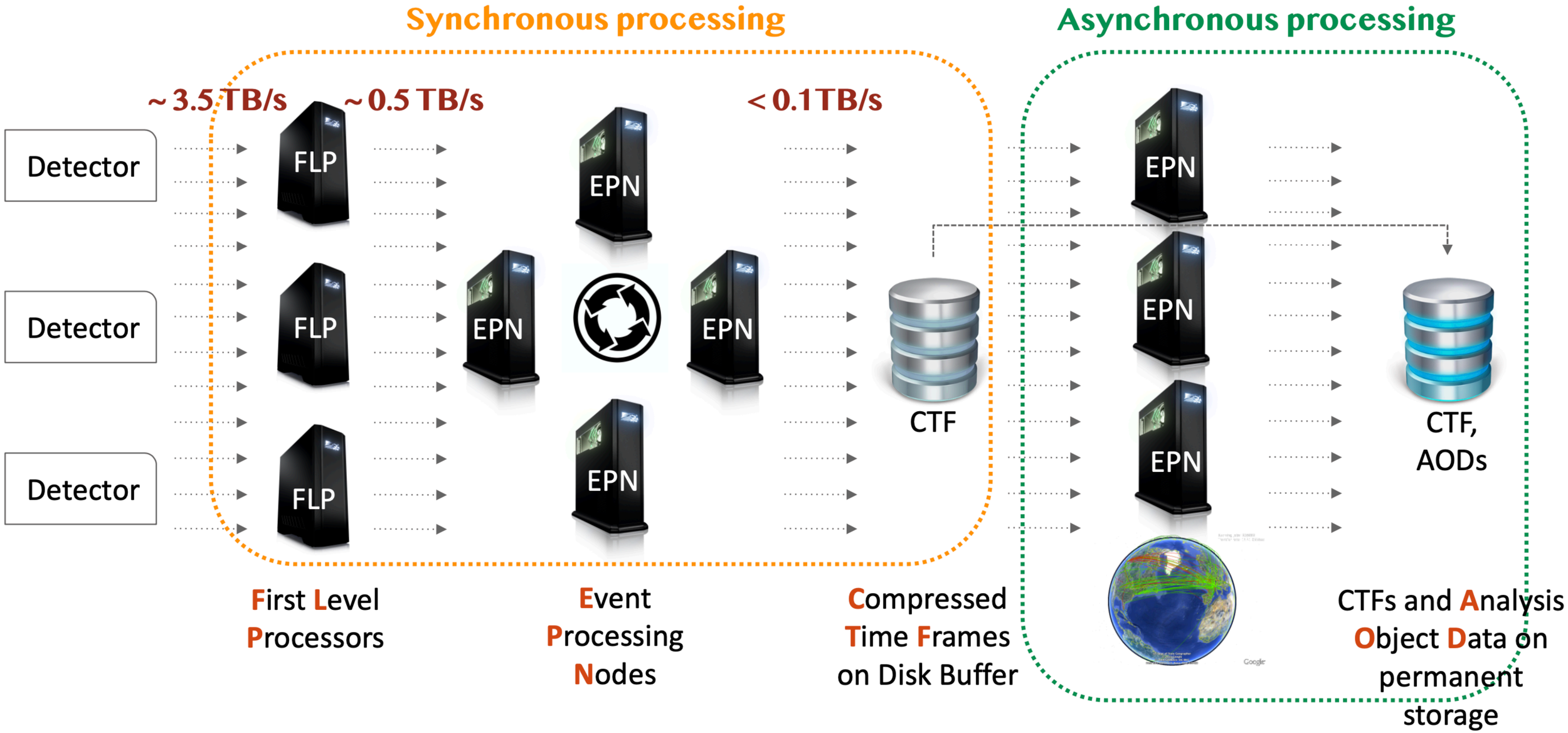


- **Strategy in ALICE Run3**

  ✓ Continuous readout and online data reconstruction

  ➡ **New Online/Offline system (O2)**

Data processing for run 3 and 4



Analysis framework for run 3 and 4

Challenge:

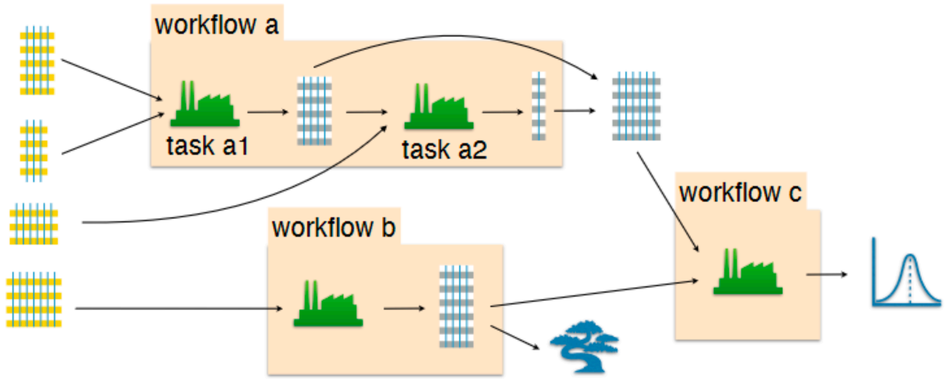
process 100 times more data with just 4 times the resources (wrt Run 2)

- **How can we do?**

✓ Modern data structures and modern C++

✓Parallelized execution (process a full time frame, not an event)

✓ Combine (highly optimizable) declarative and (flexible) imperative concepts



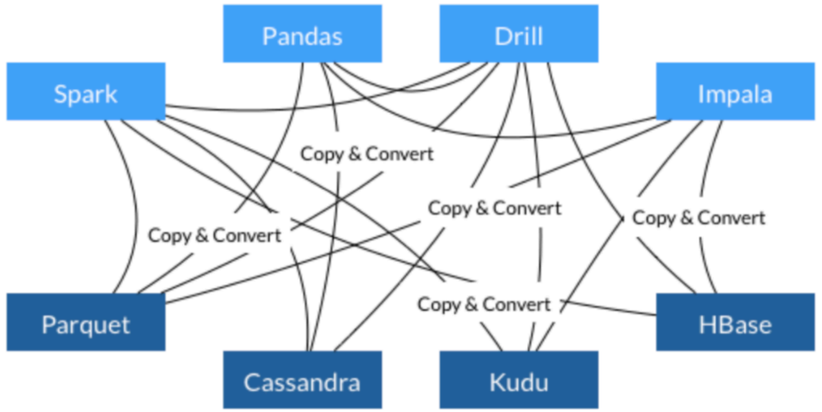
➡ O2 Analysis Framework built on O2 Data Processing Layer (DPL)

- Analysis split into blocks
- Each block consumes trees/tables, produces histograms or trees/tables

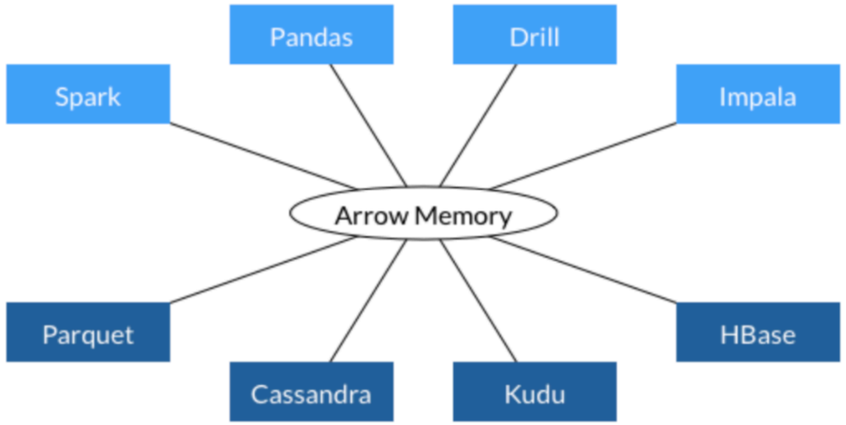
Apache arrow

Analysis tool

Data base



apache arrow



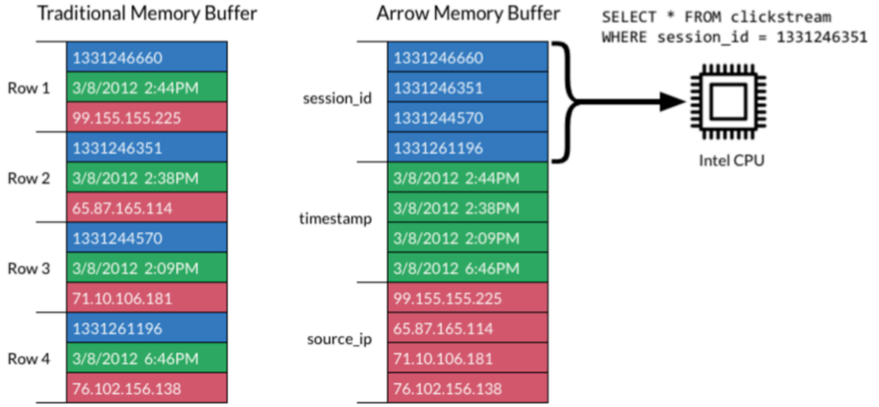
70% ~80% CPU cost for the copy and convert!

Apache Arrow:

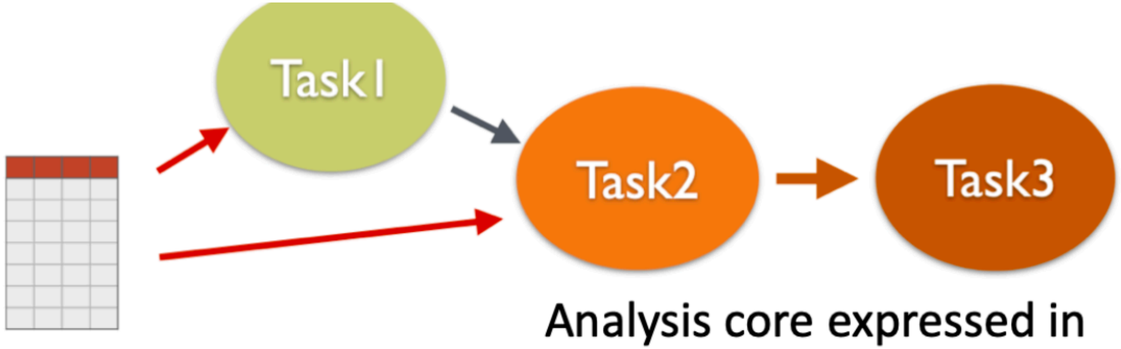
- Cross-language development platform for in-memory analytics
- shared in-memory columnar data and zero-copy

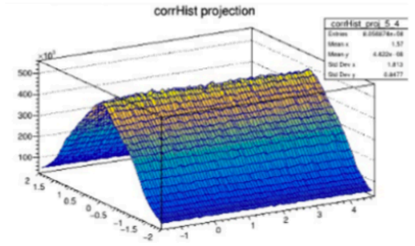
☑ Enable big data systems to process and move data fast. Exploited by ALICE experiment to analyse the unprecedented amount of data for run3 in O2

	session_id	timestamp	source_ip
Row 1	1331246660	3/8/2012 2:44PM	99.155.155.225
Row 2	1331246351	3/8/2012 2:38PM	65.87.165.114
Row 3	1331244570	3/8/2012 2:09PM	71.10.106.181
Row 4	1331261196	3/8/2012 6:46PM	76.102.156.138



Analysis chain





ROOT serialized output

Data model for analysis based on **flat tables** arranged in a relational-database-like manner:

- minimise I/O cost
- improve vectorisation / parallelism

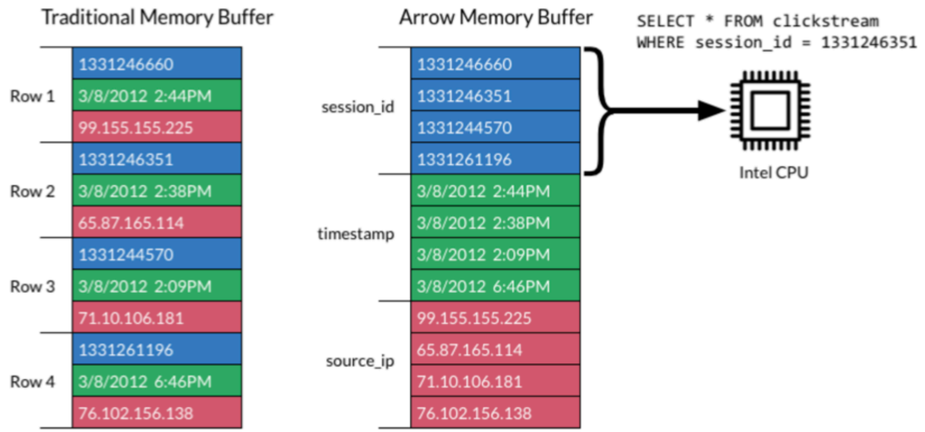
Apache Arrow hidden behind a classic C++ API

Analysis core expressed in the form of a **task**

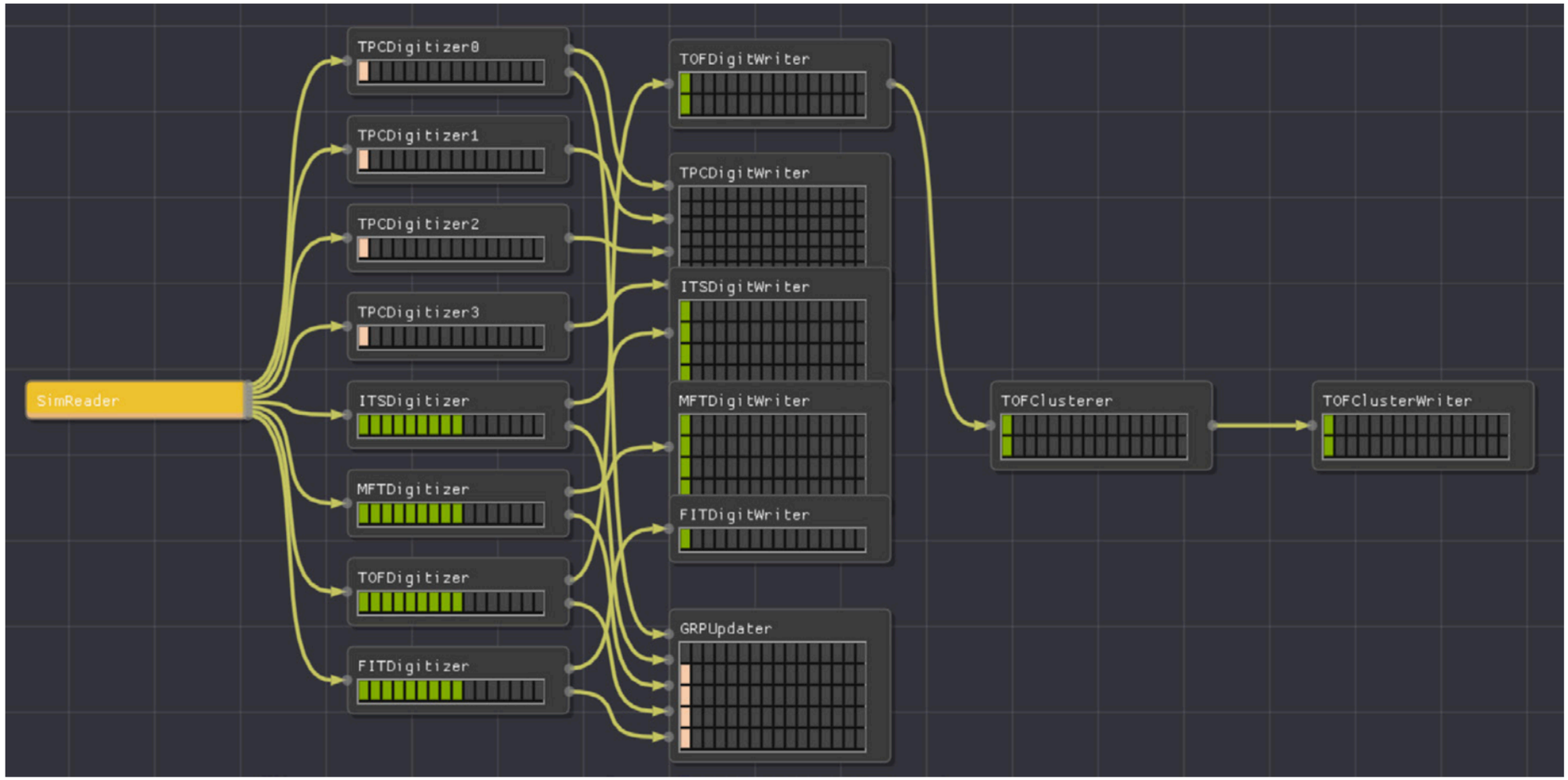
- legacy from Run 1 + 2
- filters and selections
- merging, concatenation of tables

	session_id	timestamp	source_ip
Row 1	1331246660	3/8/2012 2:44PM	99.155.155.225
Row 2	1331246351	3/8/2012 2:38PM	65.87.165.114
Row 3	1331244570	3/8/2012 2:09PM	71.10.106.181
Row 4	1331261196	3/8/2012 6:46PM	76.102.156.138

Apache Arrow !!

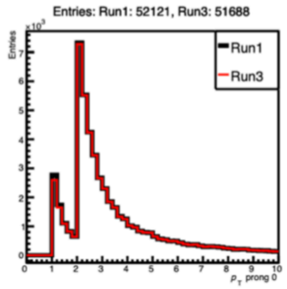


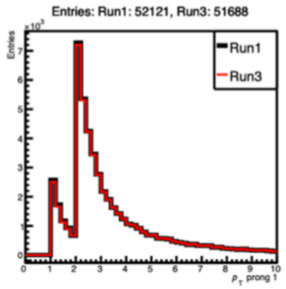
DEBUG GUI

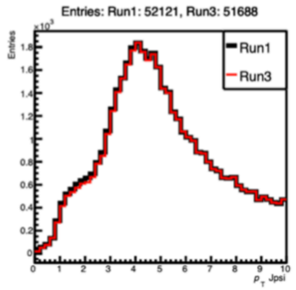


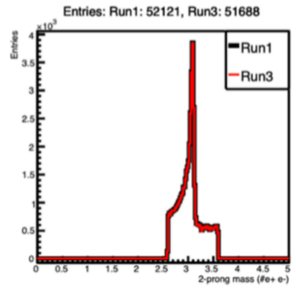
- Digitization workflow on-going as shown in the topology viewer of debug GUI

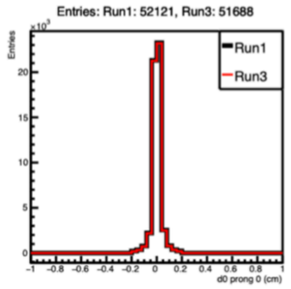
Run1/2 vs Run3

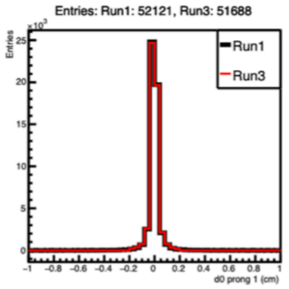


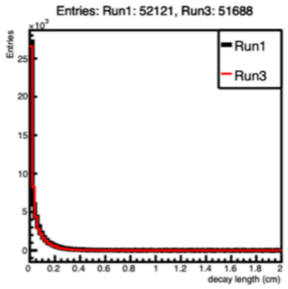


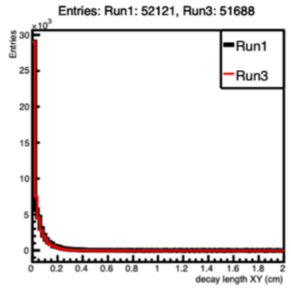






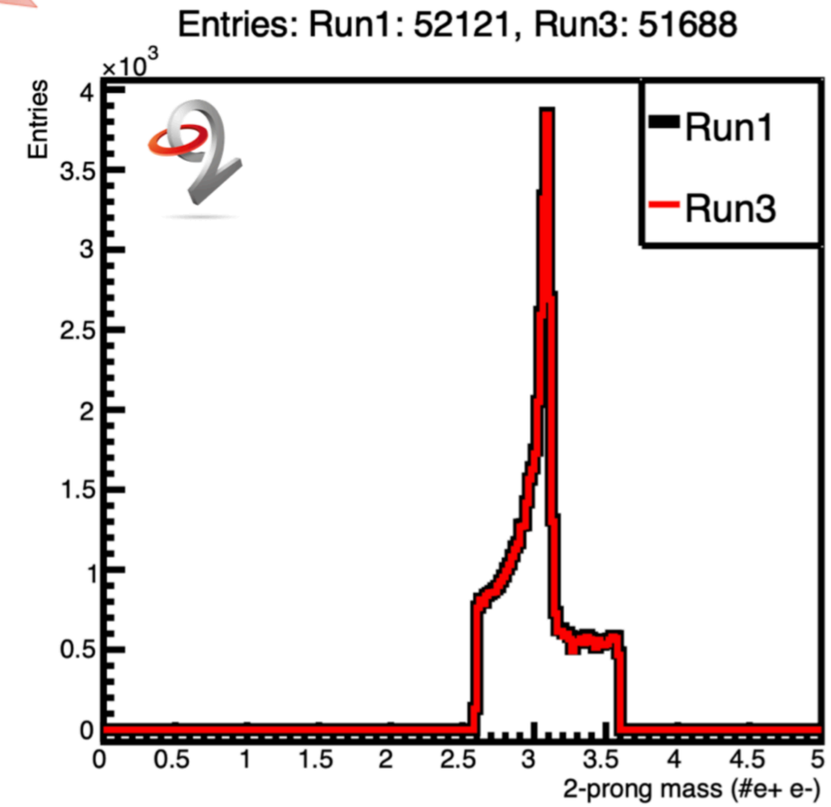




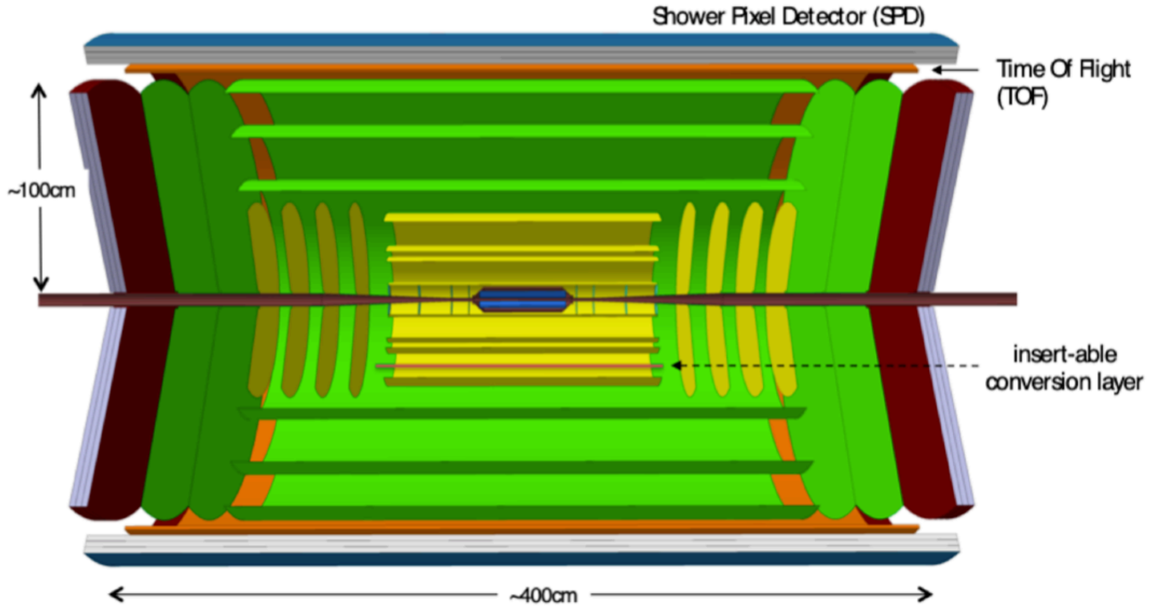


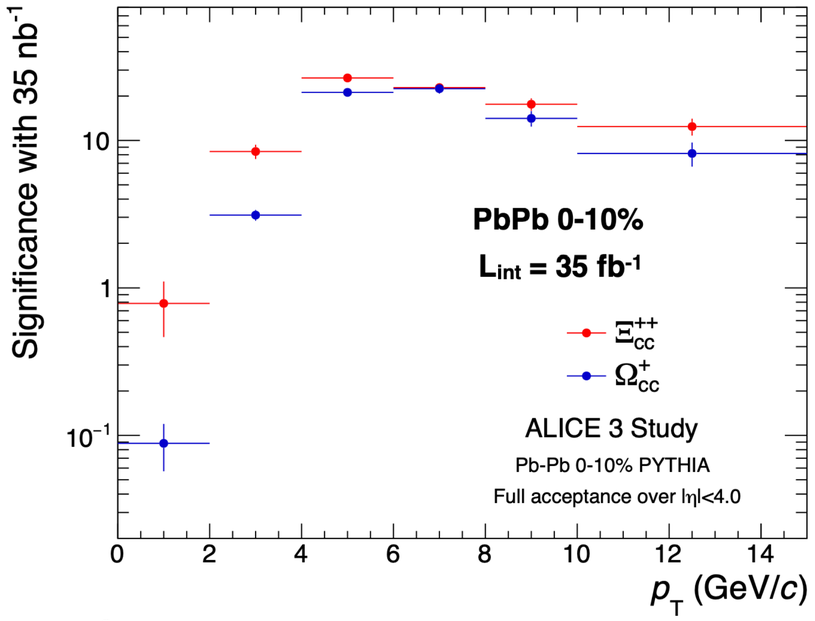
Run2 MC DATA: LHC19c6 (285013 285014 285015 285064 285106)

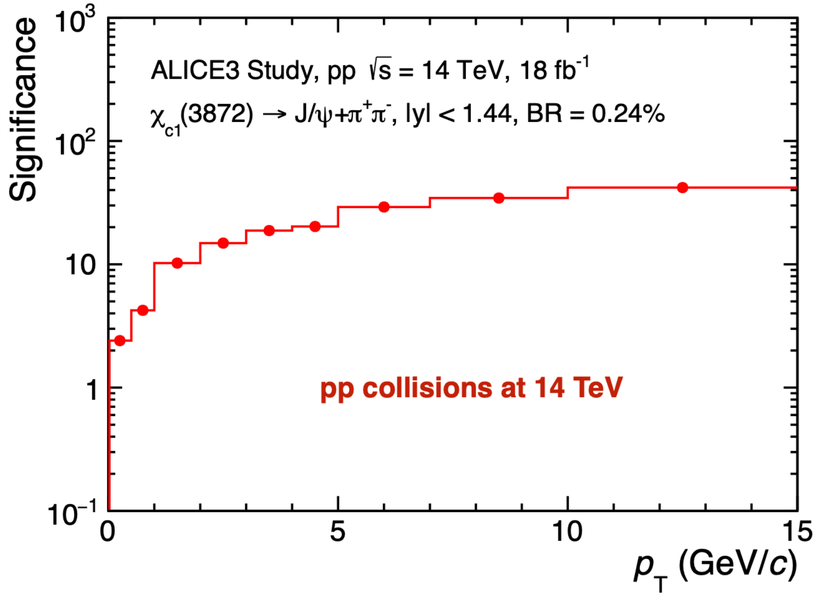
- **Good agreement with AliPhysics and O2**
- **Nice shape of Inv. Mass**



ALICE3 Run5 (based on O2)







➡ Large mount of simulation data be produced for Multiple charm hadron(Ω_{cc}^{+}, Ξ_{cc}^{++}), exotic state (X(3872)) study for run5 in ALICE3

Summary

- ALICE upgrade completely in the July of 2021, Pilot beam test have started in the Oct, 2021
- Analysis framework for run 3 and 4 (O2) have be developed, showed consistent results with Run1/2 framework
- The O2 system has been exploited for the ALICE3 performance study