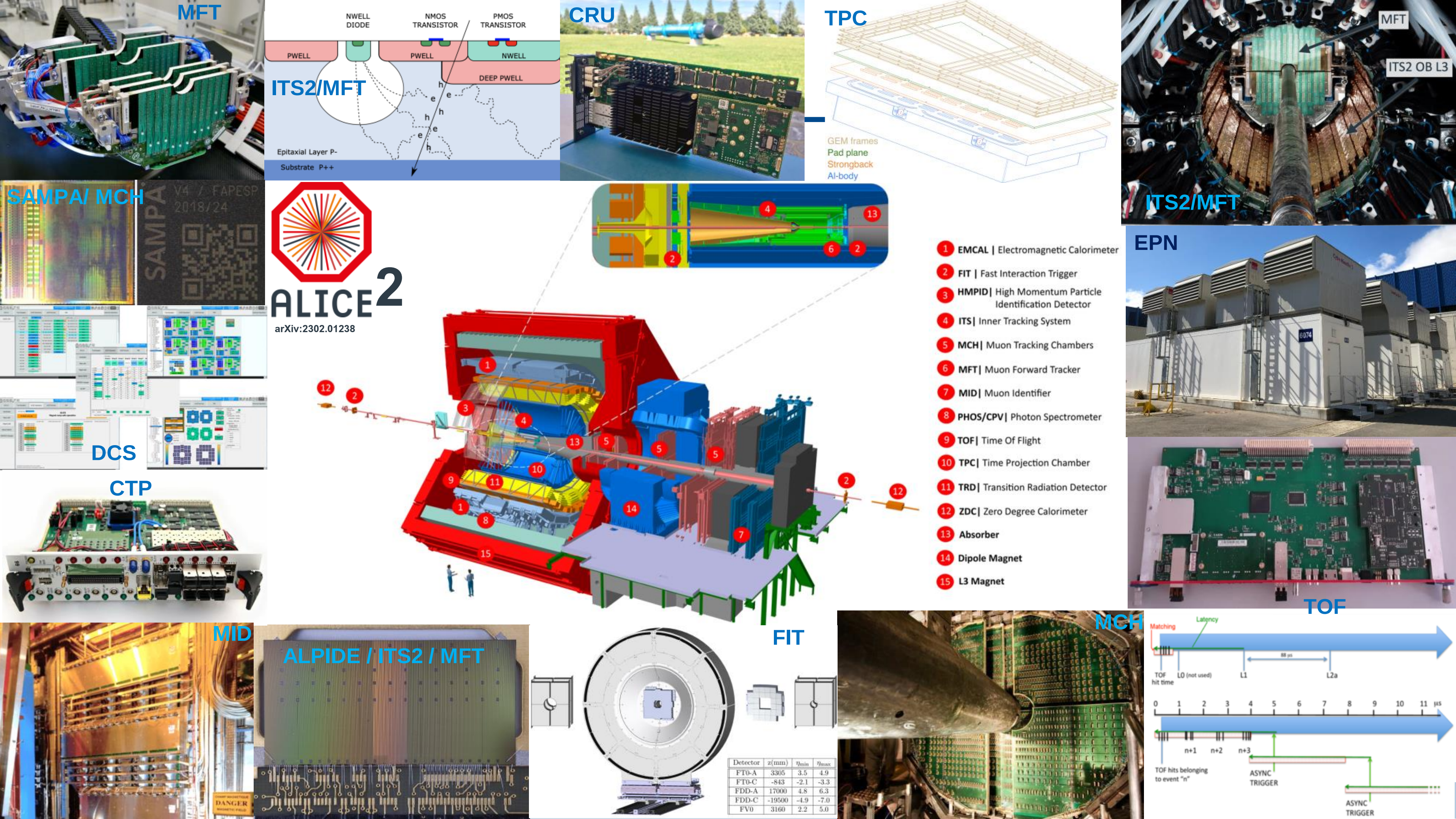
The ALICE Upgrades in LS2

Stefania Beole, Christian Lippmann, David Rohr

CERN Detector seminar

15.9.2023

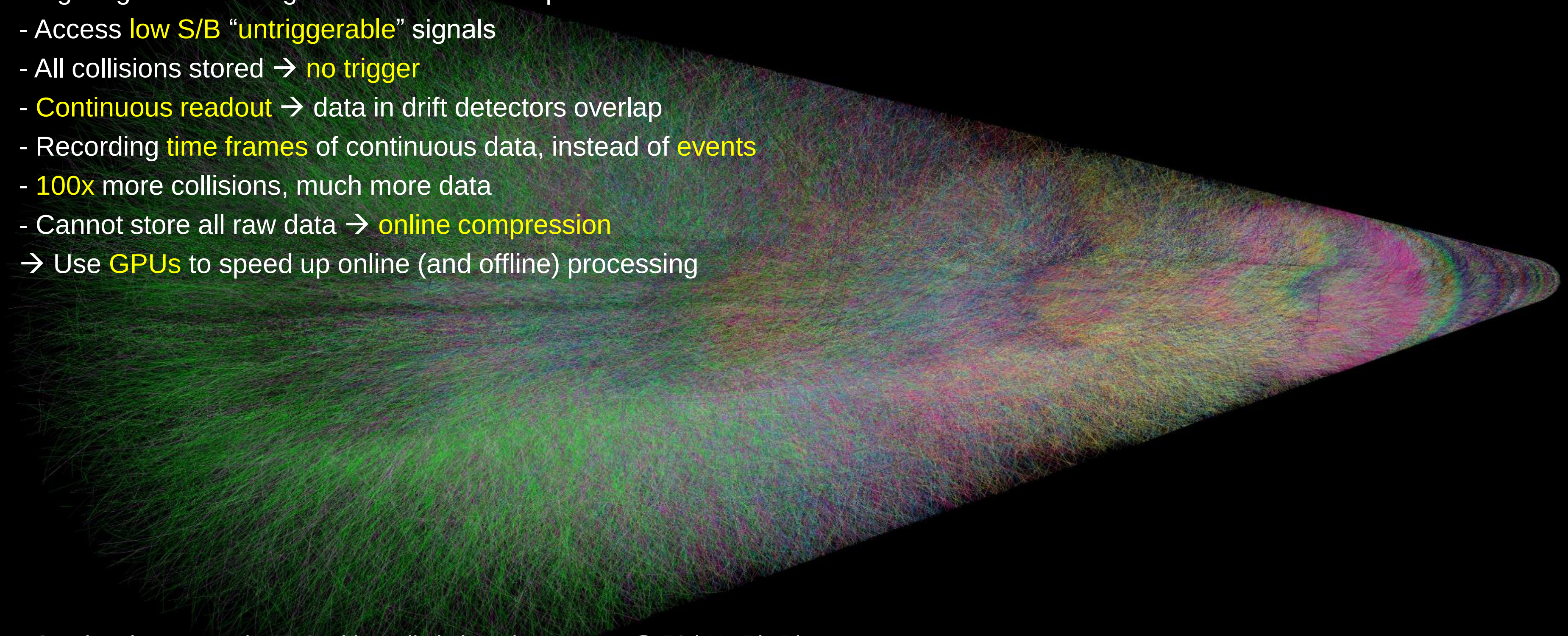

MFT
NWELL DIODE
NMOS TRANSISTOR
PMOS TRANSISTOR
PWELL
PWELL
NWELL
DEEP PWELL
ITS2/MFT
Epitaxial Layer P-
Substrate P++
CRU
TPC
GEM frames
Pad plane
Strongback
Al-body
MFT
ITS2 OB L3
ITS2/MFT
SAMPA/ MCH
V4 / FAPESP 2018/24
SAMPA
ALICE 2
arXiv:2302.01238
DCS
EPN
CTP
1 EMCAL | Electromagnetic Calorimeter
2 FIT | Fast Interaction Trigger
3 HMPID| High Momentum Particle Identification Detector
4 ITS| Inner Tracking System
5 MCH| Muon Tracking Chambers
6 MFT| Muon Forward Tracker
7 MID| Muon Identifier
8 PHOS/CPV| Photon Spectrometer
9 TOF| Time Of Flight
10 TPC| Time Projection Chamber
11 TRD| Transition Radiation Detector
12 ZDC| Zero Degree Calorimeter
13 Absorber
14 Dipole Magnet
15 L3 Magnet
TOF
MID
ALPIDE / ITS2 / MFT
FIT
Detector | z(mm) | ηmin | ηmax
FT0-A | 3305 | 3.5 | 4.9
FT0-C | -843 | -2.1 | -3.3
FDD-A | 17000 | 4.8 | 6.3
FDD-C | -19500 | -4.9 | -7.0
FV0 | 3160 | 2.2 | 5.0
MCH
Matching
Latency
TOF hit time
L0 (not used)
L1
L2a
0 1 2 3 4 5 6 7 8 9 10 11 μs
n+1
n+2
n+3
TOF hits belonging to event "n"
ASYNC TRIGGER
ASYNC TRIGGER
DANGER

ALICE in Run 3

Targeting to record large minimum bias sample.

- Access low S/B “untriggerable” signals
- All collisions stored → no trigger
- Continuous readout → data in drift detectors overlap
- Recording time frames of continuous data, instead of events
- 100x more collisions, much more data
- Cannot store all raw data → online compression

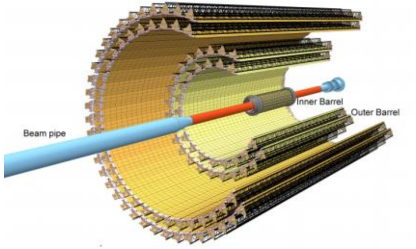
→ Use GPUs to speed up online (and offline) processing

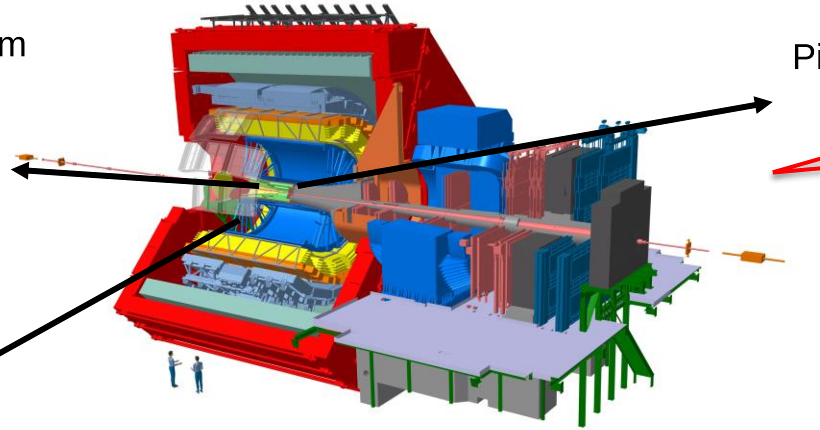


- Overlapping events in TPC with realistic bunch structure @ 50 kHz Pb-Pb.
- Timeframe of 2 ms shown (will be 10 – 20 ms in production).
- Tracks of different collisions shown in different colors.

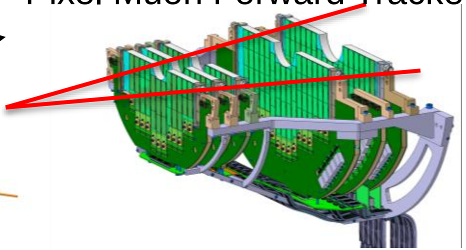
LS2 ALICE Upgrades

All-pixel Inner Tracking System

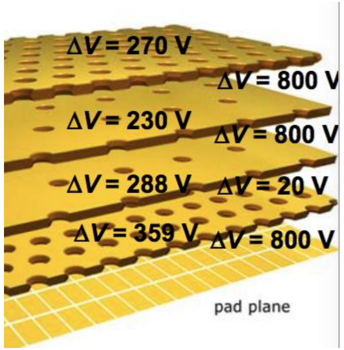




Pixel Muon Forward Tracker



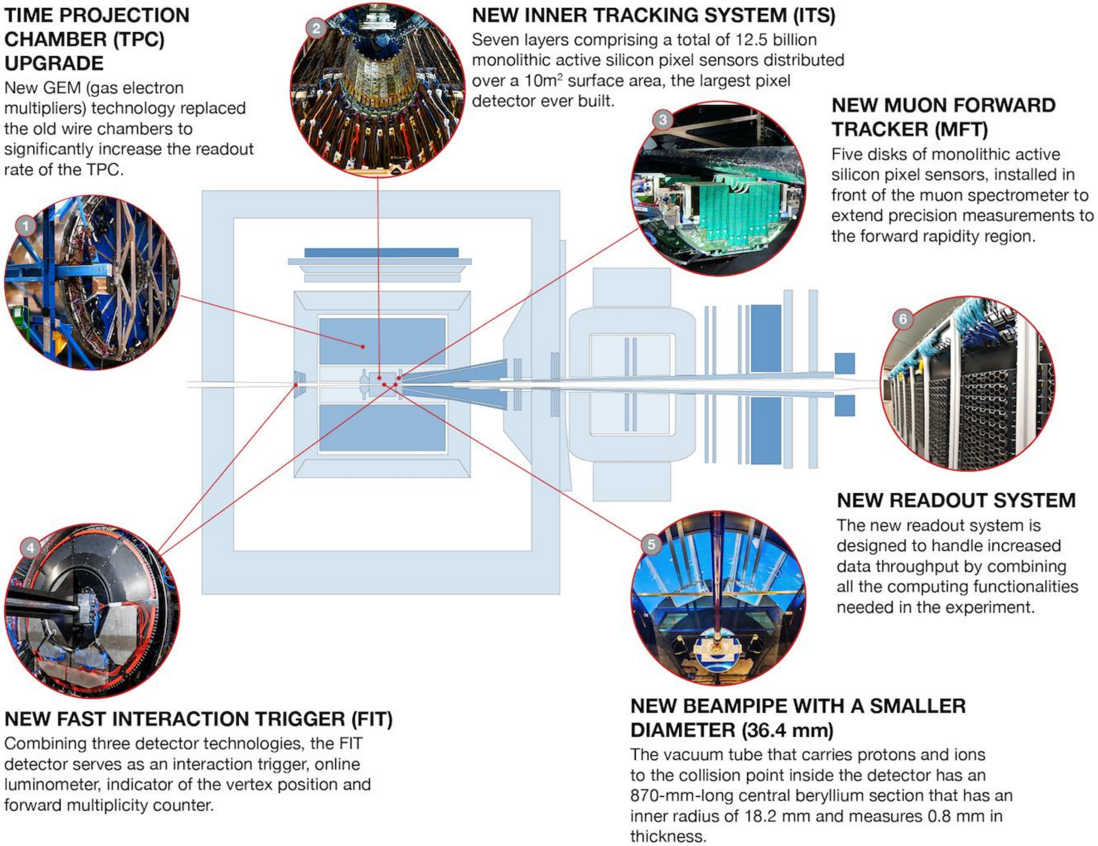
GEM-based
TPC readout



- **New detectors:**
  - Improve tracking resolution at low p_T
    → thinner, more granular
  - Enable continuous read-out
- **New online-offline computing system for synchronous and asynchronous processing**

… and much more:
- Fast Interaction Trigger
- New 50x faster readout system
- Readout upgrade of MUON, TOF, EMCAL, PHOS

LS2 ALICE Upgrades

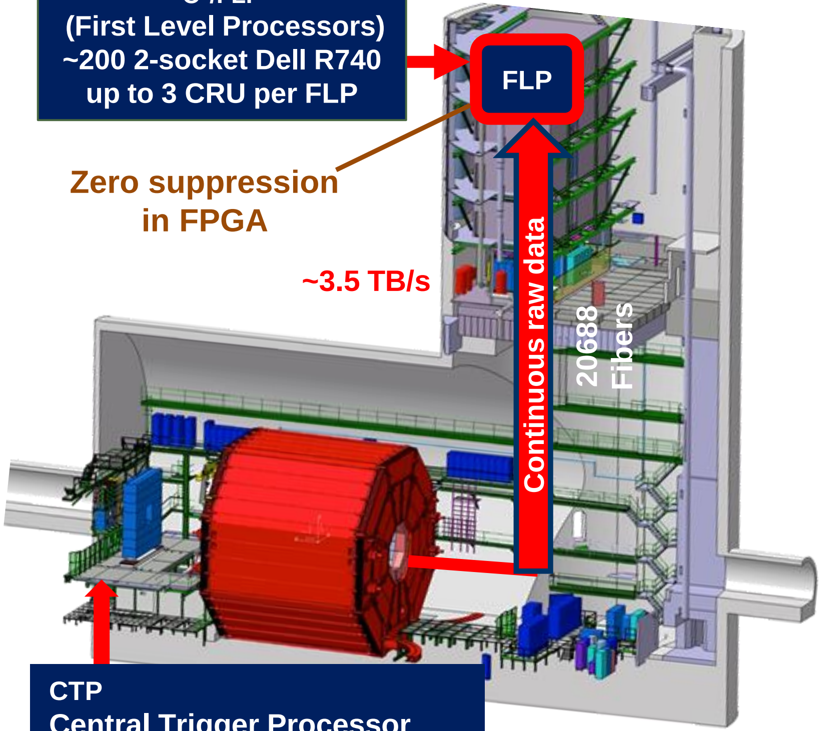


ALICE Data Flow in Run 3

O^2/FLP
(First Level Processors)
~200 2-socket Dell R740
up to 3 CRU per FLP

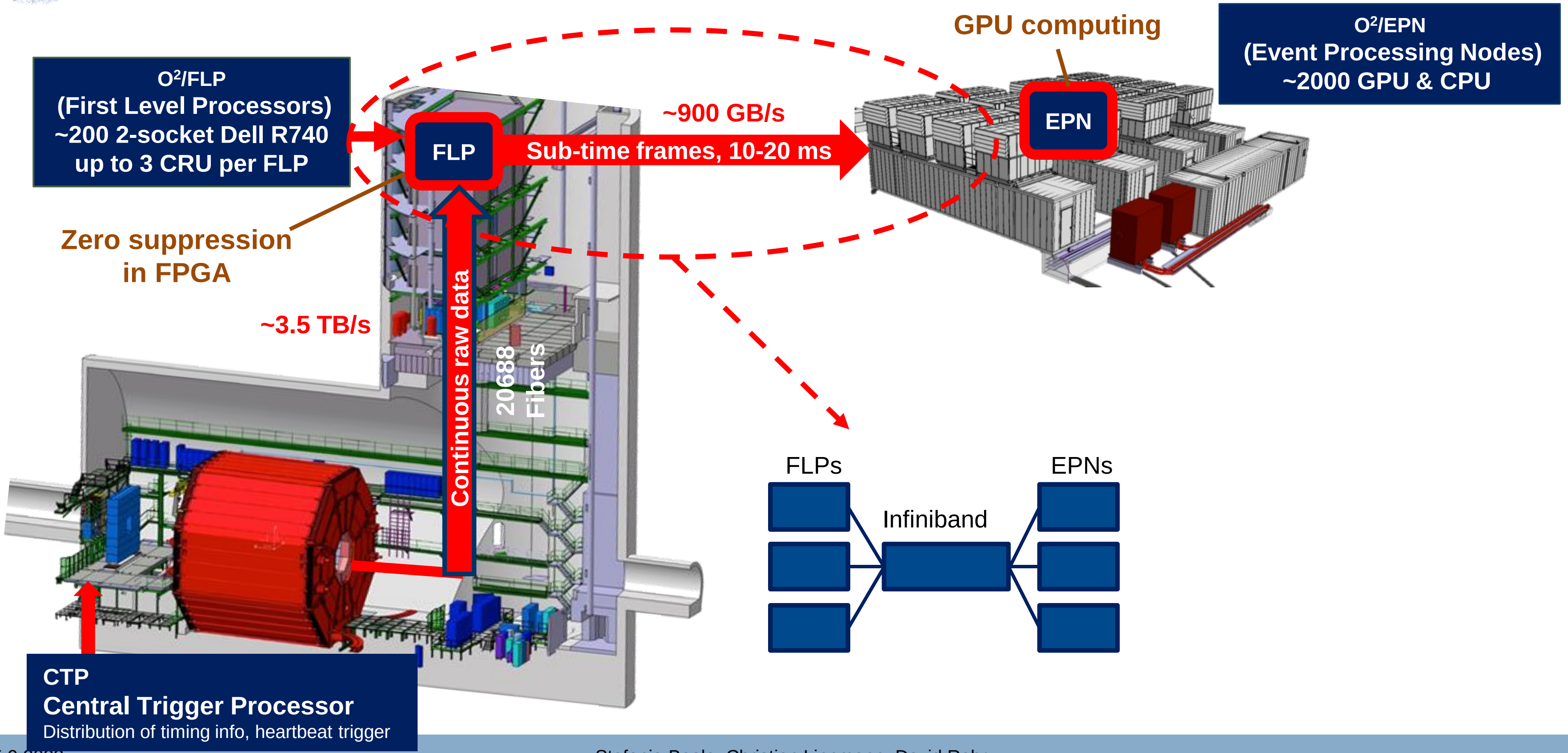
Zero suppression
in FPGA

~3.5 TB/s

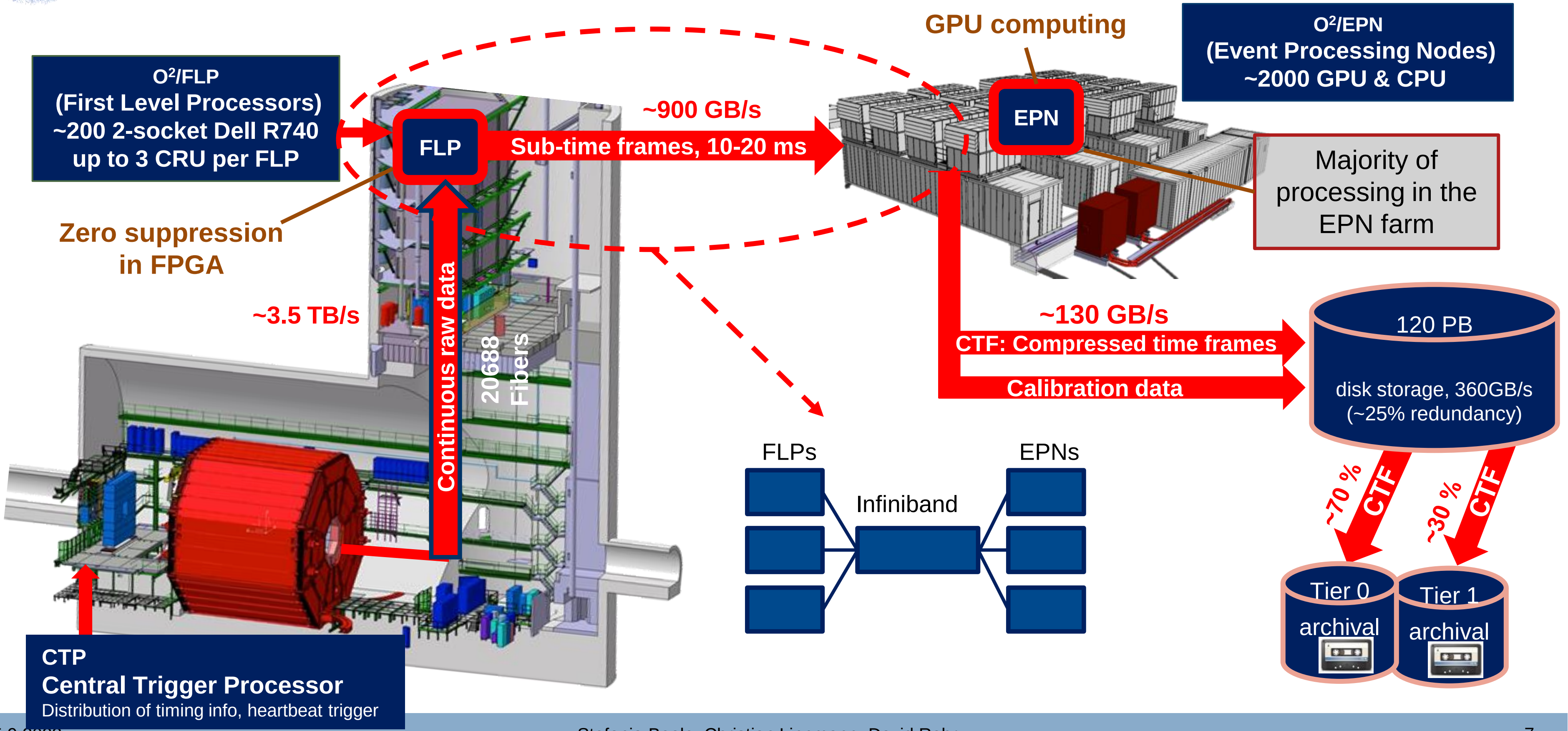


CTP
Central Trigger Processor
Distribution of timing info, heartbeat trigger

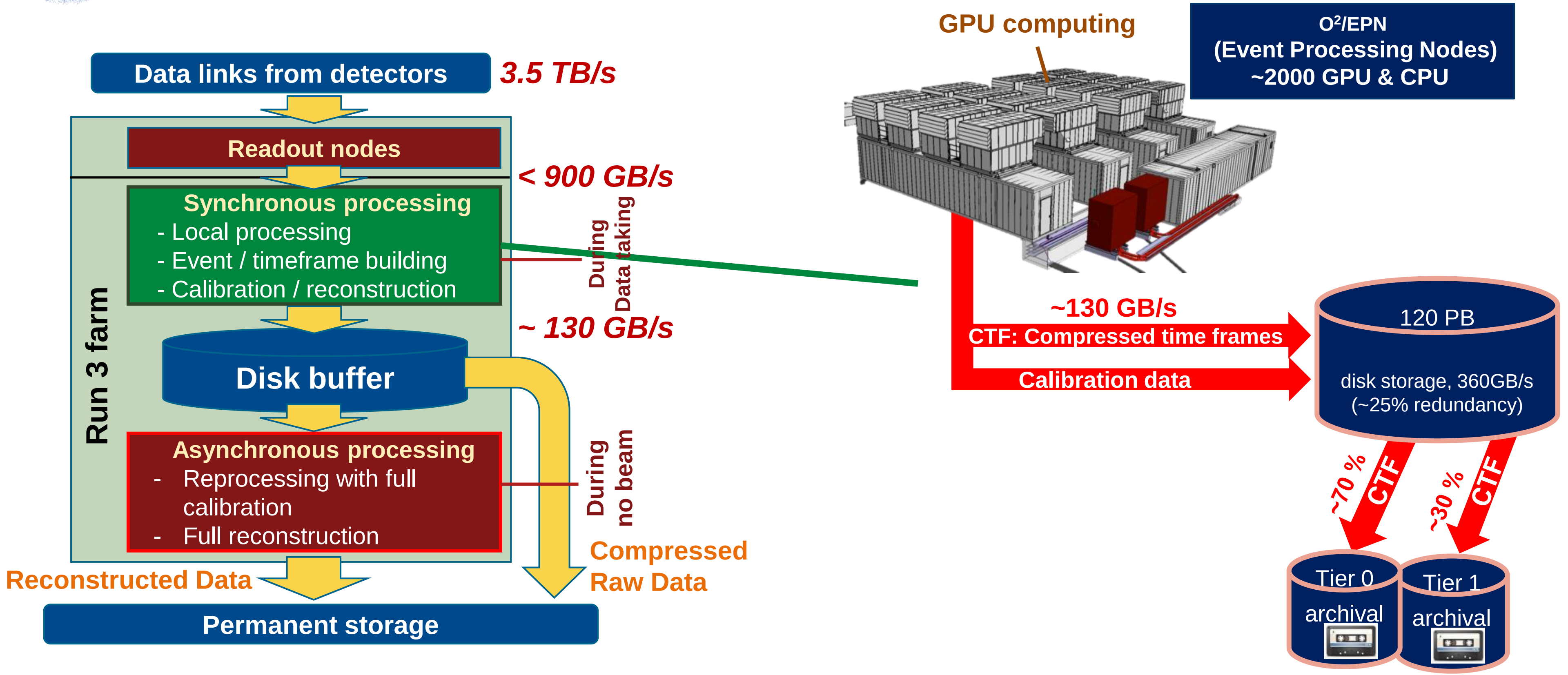
ALICE Data Flow in Run 3



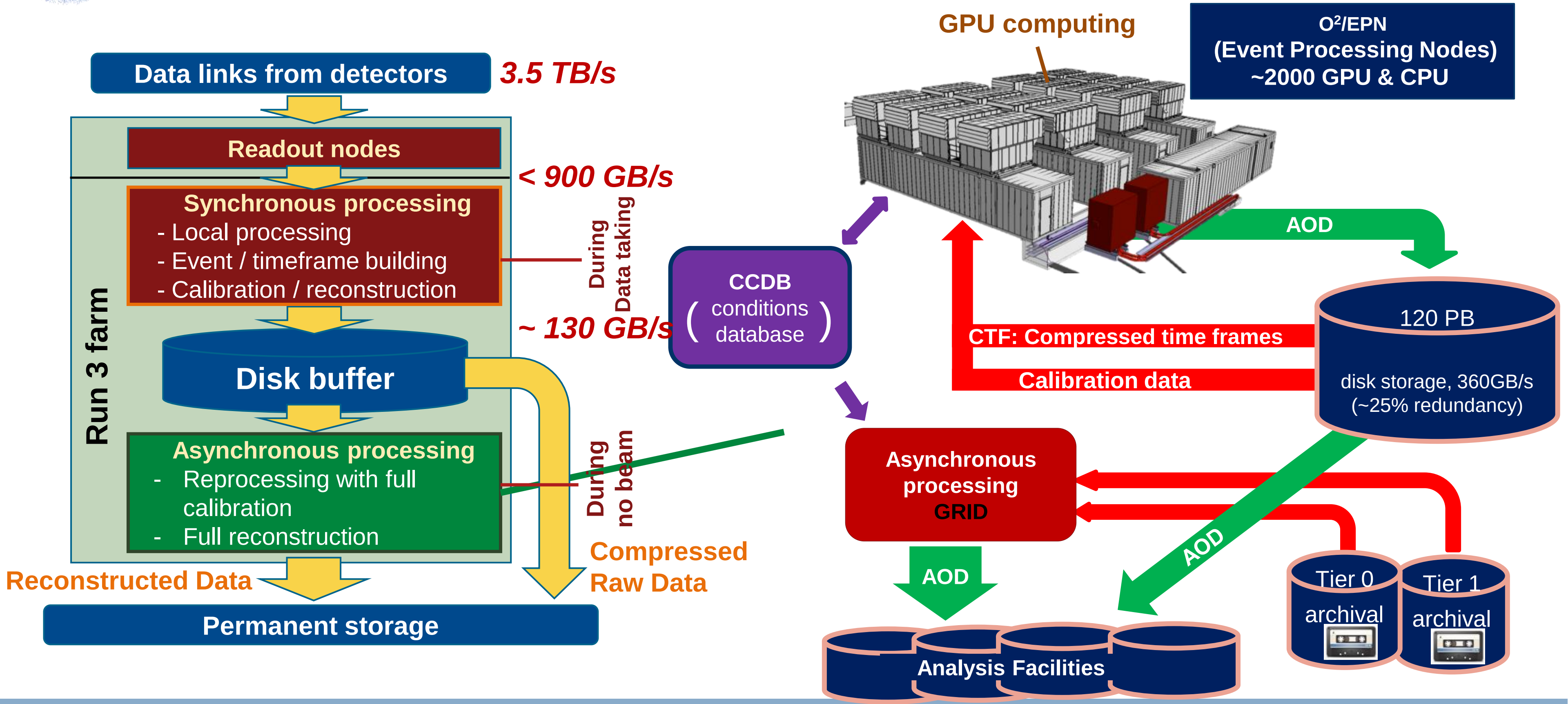
ALICE Data Flow in Run 3



Synchronous and Asynchronous Processing



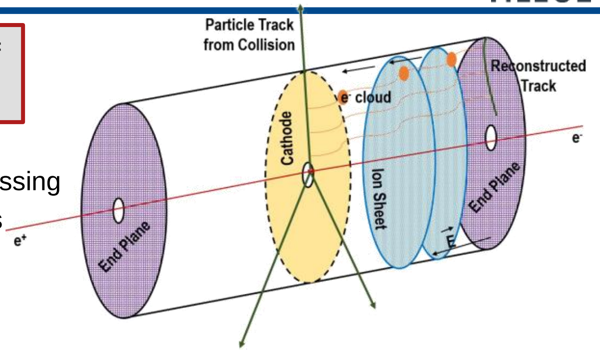
Synchronous and Asynchronous Processing



O^2 Processing steps

- **Synchronous processing (what we called online before):**
  - Extract information for **detector calibration**:
    - Previously performed in 2 offline passes over the data after the data taking
    - Run 3 avoids / reduces extra passes over the data but extracts all information in the sync. processing
    - An intermediate step between sync. and async. processing produces the final calibration objects
    - The most complicated calibration is the correction for the TPC space charge distortions

Needs tracking of 1% of tracks

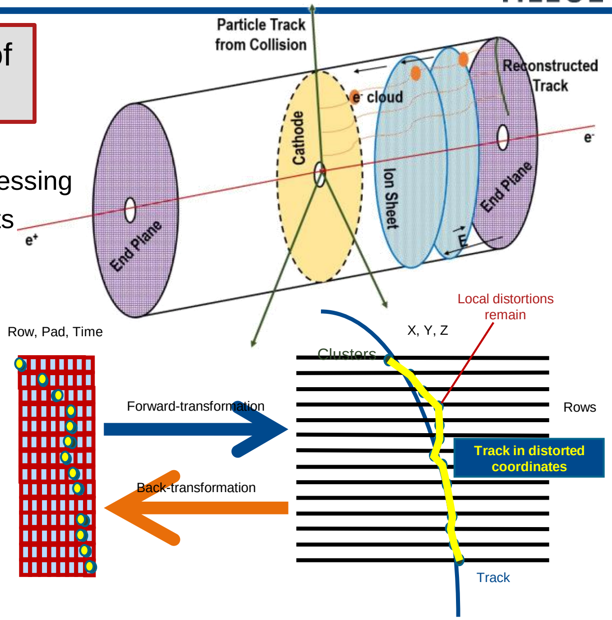


O² Processing steps

- **Synchronous processing (what we called online before):**
  - Extract information for **detector calibration**:
    - Previously performed in 2 offline passes over the data after the data taking
    - Run 3 avoids / reduces extra passes over the data but extracts all information in the sync. processing
    - An intermediate step between sync. and async. processing produces the final calibration objects
    - The most complicated calibration is the correction for the TPC space charge distortions
  - **Data compression**:
    - TPC is the largest contributor of raw data, and we employ sophisticated algorithms like storing space point coordinates as residuals to tracks to reduce the entropy and remove hits not attached to physics tracks
    - We use ANS entropy encoding for all detectors

Needs tracking of 1% of tracks

Needs 100% TPC tracking

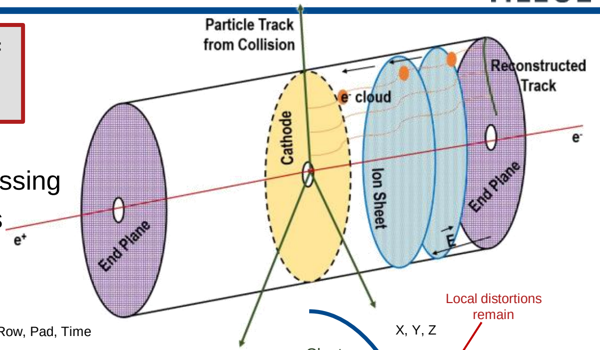


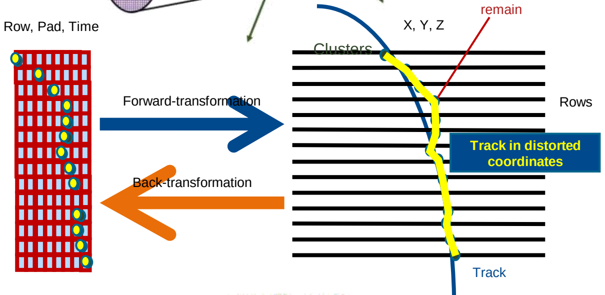
O^2 Processing steps

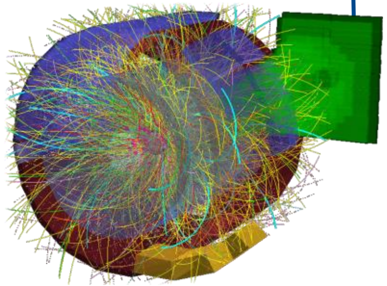
- **Synchronous processing (what we called online before):**
  - Extract information for **detector calibration**:
    - Previously performed in 2 offline passes over the data after the data taking
    - Run 3 avoids / reduces extra passes over the data but extracts all information in the sync. processing
    - An intermediate step between sync. and async. processing produces the final calibration objects
    - The most complicated calibration is the correction for the TPC space charge distortions
  - **Data compression**:
    - TPC is the largest contributor of raw data, and we employ sophisticated algorithms like storing space point coordinates as residuals to tracks to reduce the entropy and remove hits not attached to physics tracks
    - We use ANS entropy encoding for all detectors
  - **Event reconstruction** (tracking, etc.):
    - Required for calibration, compression, and online quality control
    - Need **full TPC tracking** for data compression
    - Need tracking in all detectors for ~1% of the tracks for calibration
    - → **TPC tracking dominant part**, **rest** almost negligible (**< 5%**)

Needs tracking of 1% of tracks

Needs 100% TPC tracking





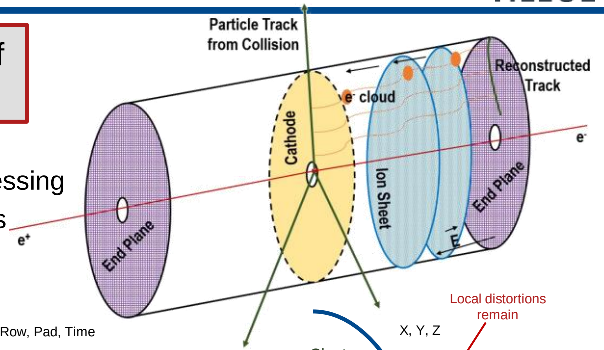


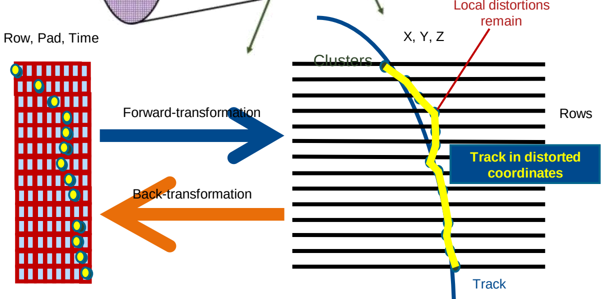
O² Processing steps

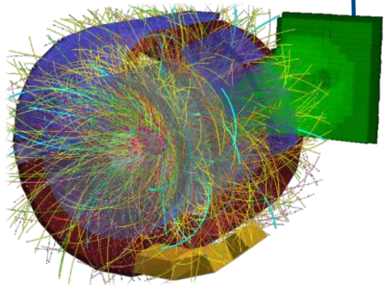
- **Synchronous processing (what we called online before):**
  - Extract information for **detector calibration**:
    - Previously performed in 2 offline passes over the data after the data taking
    - Run 3 avoids / reduces extra passes over the data but extracts all information in the sync. processing
    - An intermediate step between sync. and async. processing produces the final calibration objects
    - The most complicated calibration is the correction for the TPC space charge distortions
  - **Data compression**:
    - TPC is the largest contributor of raw data, and we employ sophisticated algorithms like storing space point coordinates as residuals to tracks to reduce the entropy and remove hits not attached to physics tracks
    - We use ANS entropy encoding for all detectors
  - **Event reconstruction** (tracking, etc.):
    - Required for calibration, compression, and online quality control
    - Need **full TPC tracking** for data compression
    - Need tracking in all detectors for ~1% of the tracks for calibration
    - → **TPC tracking dominant part**, **rest** almost negligible (**< 5%**)

Needs tracking of 1% of tracks

Needs 100% TPC tracking







- **Asynchronous processing (what we called offline before):**
  - **Full reconstruction**, **full calibration**, **all detectors**
  - TPC part faster than in synchronous processing (less hits, no clustering, no compression)
  - → **Different relative importance of GPU / CPU** algorithms compared to synchronous processing

Overview of compute time of reconstruction steps

- The table below shows the relative compute time (linux cpu time) of the processing steps running on the processor.

Synchronous processing
(50 kHz Pb-Pb, MC data)

Processing step	% of time
TPC Processing (Tracking, Clustering, Compression)	99.37 %
EMCAL Processing	0.20 %
ITS Processing (Clustering + Tracking)	0.10 %
TPC Entropy Encoder	0.10 %
ITS-TPC Matching	0.09 %
MFT Processing	0.02 %
TOF Processing	0.01 %
TOF Global Matching	0.01 %
PHOS / CPV Entropy Coder	0.01 %
ITS Entropy Coder	0.01 %
Rest	0.08 %

Only data processing steps
Quality control, calibration, event building excluded!

Asynchronous processing
(650 kHz pp, real data, calorimeters not in run)

Processing step	% of time
TPC Processing (Tracking)	61.41 %
ITS TPC Matching	6.13 %
MCH Clusterization	6.13 %
TPC Entropy Decoder	4.65 %
ITS Tracking	4.16 %
TOF Matching	4.12 %
TRD Tracking	3.95 %
MCH Tracking	2.02 %
AOD Production	0.88 %
Quality Control	4.00 %
Rest	2.32 %

Overview of compute time of reconstruction steps

- The table below shows the relative compute time (linux cpu time) of the processing steps running on the processor.

Synchronous processing
(50 kHz Pb-Pb, MC data)

Totally dominated by TPC: >99%

Processing step	% of time
TPC Processing (Tracking, Clustering, Compression)	99.37 %
EMCAL Processing	0.20 %
ITS Processing (Clustering + Tracking)	0.10 %
TPC Entropy Encoder	0.10 %
ITS-TPC Matching	0.09 %
MFT Processing	0.02 %
TOF Processing	0.01 %
TOF Global Matching	0.01 %
PHOS / CPV Entropy Coder	0.01 %
ITS Entropy Coder	0.01 %
Rest	0.08 %

Only data processing steps
Quality control, calibration, event building excluded!

Asynchronous processing
(650 kHz pp, real data, calorimeters not in run)

Processing step	% of time
TPC Processing (Tracking)	61.41 %
ITS TPC Matching	6.13 %
MCH Clusterization	6.13 %
TPC Entropy Decoder	4.65 %
ITS Tracking	4.16 %
TOF Matching	4.12 %
TRD Tracking	3.95 %
MCH Tracking	2.02 %
AOD Production	0.88 %
Quality Control	4.00 %
Rest	2.32 %

Overview of compute time of reconstruction steps

- **Synchronous processing** :
  - **99%** of compute time spent for **TPC**.
  - **EPN farm build for synchronous processing!**
- **Asynchronous reprocessing :**
  - More detectors with significant computing contribution.
  - To be kept in mind, as EPNS also run async. Reco.

Synchronous processing
(50 kHz Pb-Pb, MC data)

Processing step	% of time
TPC Processing (Tracking, Clustering, Compression)	**99.37 %**
EMCAL Processing	0.20 %
ITS Processing (Clustering + Tracking)	0.10 %
TPC Entropy Encoder	0.10 %
ITS-TPC Matching	0.09 %
MFT Processing	0.02 %
TOF Processing	0.01 %
TOF Global Matching	0.01 %
PHOS / CPV Entropy Coder	0.01 %
ITS Entropy Coder	0.01 %
Rest	0.08 %

Only data processing steps
Quality control, calibration, event building excluded!

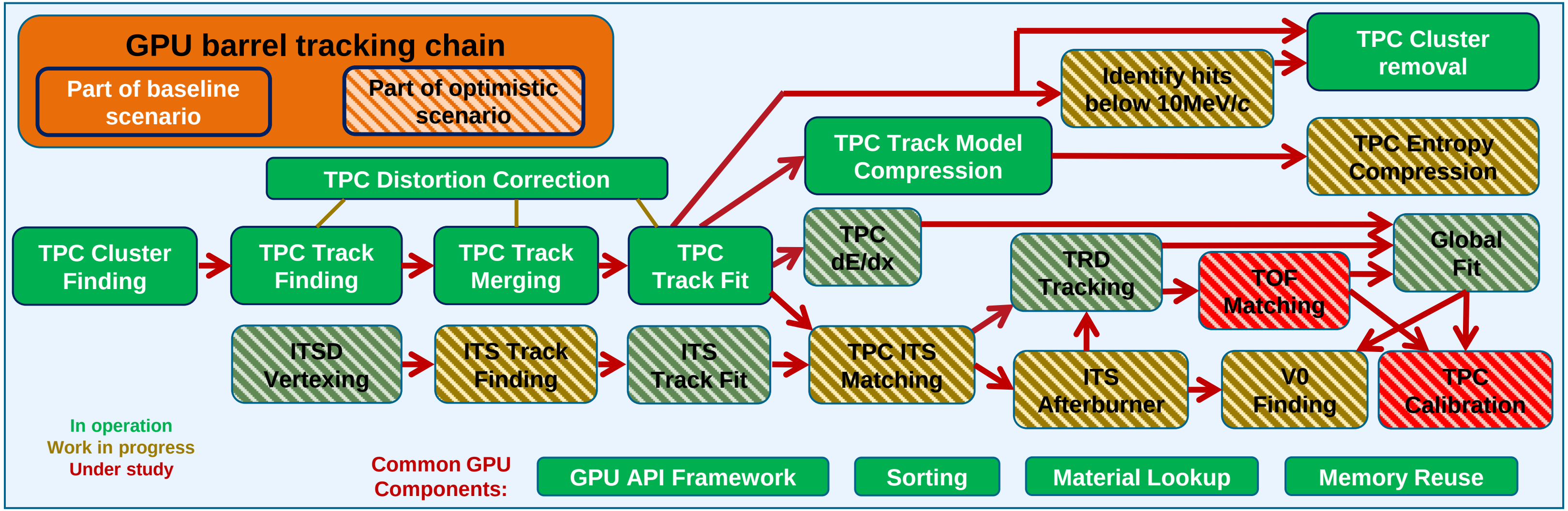
- **GPUs** well suited for **TPC** reco (from Run 1 and 2 experience).
- **GPUs** provide the **required compute power**.
  - Time frame concepts yields large enough GPU data chunks.
- Following up **2 scenarios** for EPN GPU processing:

Baseline solution (available today):
- Mandatory for synchronous processing TPC sync. reco on GPU

Optimistic solution (under development):
- Achieve best GPU usage in async phase
- Run most of tracking + X on GPU

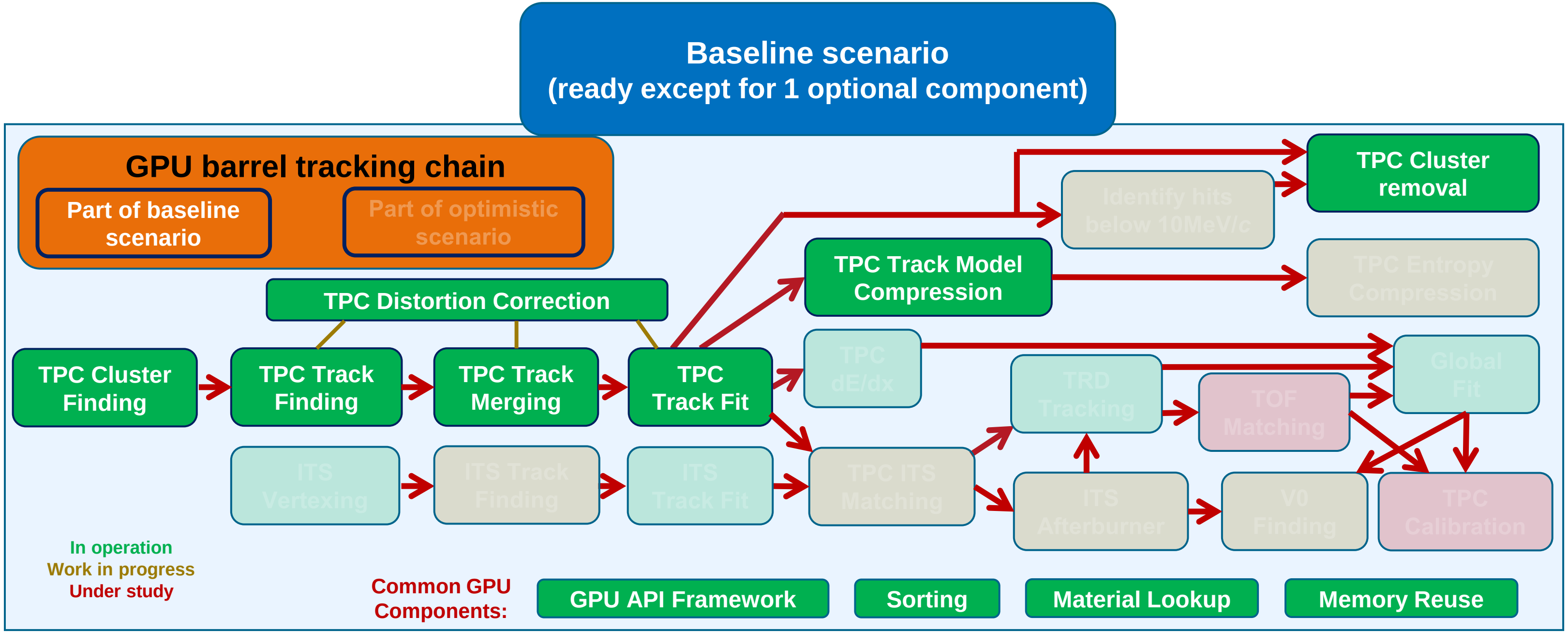
Central barrel global tracking chain

- **Central barrel tracking chosen as best candidate for optimistic scenario for asynchronous reco:**
  - Mandatory **baseline scenario** includes everything that must run on the GPU during synchronous reconstruction.
  - **Optimistic scenario** includes everything related to the barrel tracking.



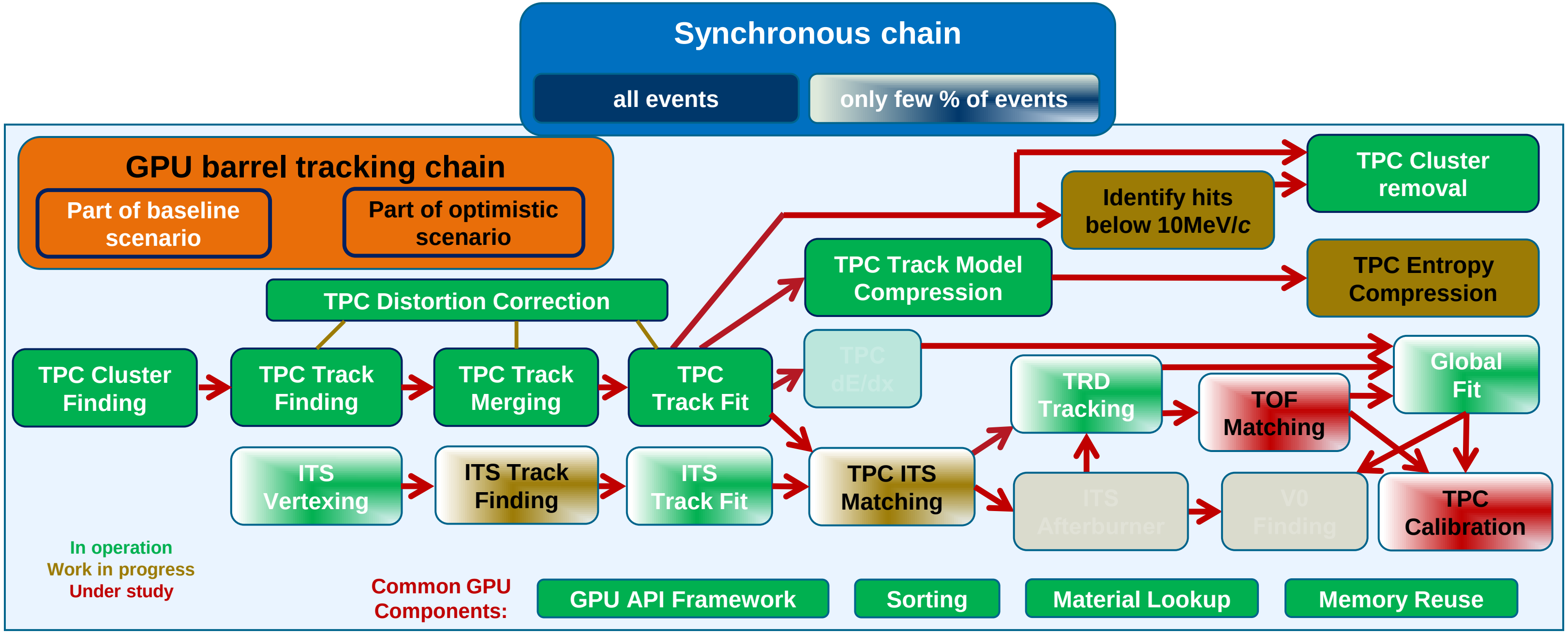
Central barrel global tracking chain

- **Baseline scenario fully implemented.**
  - Not mandatory to speed up the synchronous GPU code further.



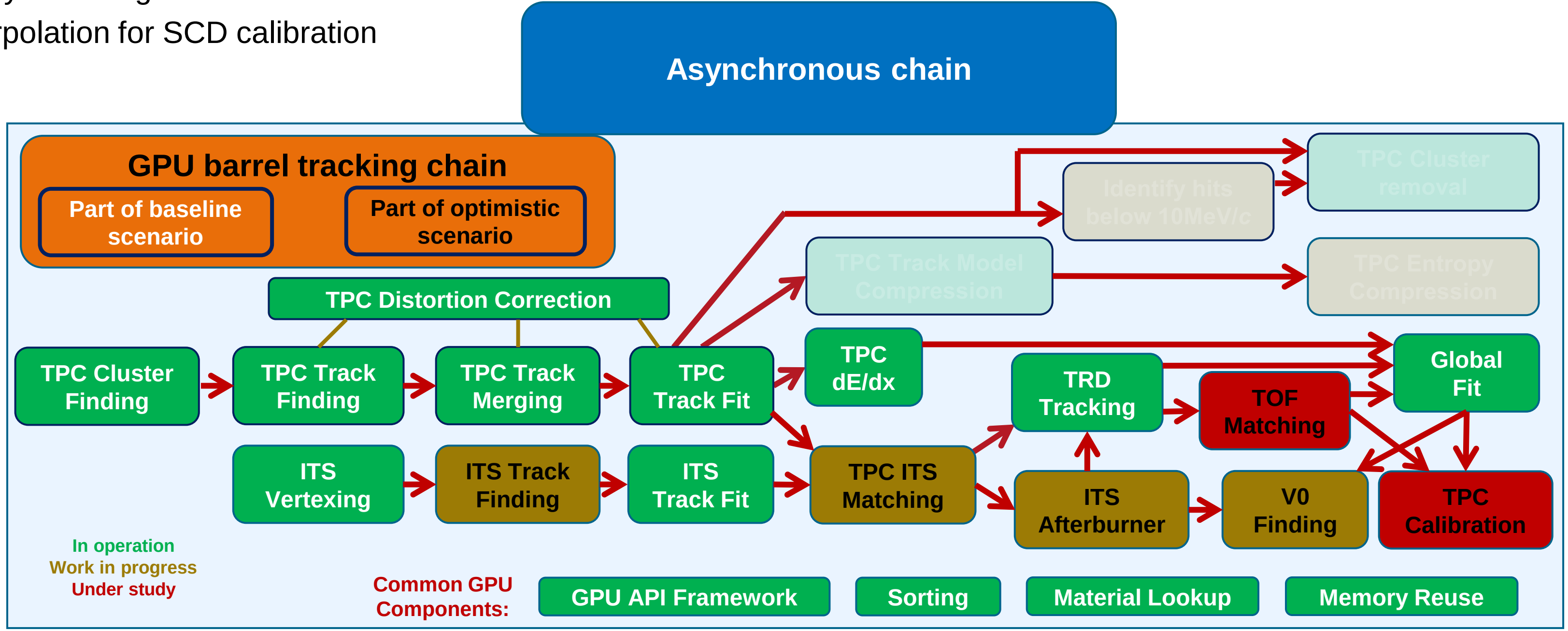
Central barrel global tracking chain

- **TPC synchronous processing almost fully on the GPU.**
  - 2 optional parts still being investigated for sync. reco on GPU: TPC entropy encoding / Looper identification < 10 MeV.



Central barrel global tracking chain

- **Several steps missing in asynchronous reconstruction:**
  - Matching to ITS
  - Matching to TOF
  - Secondary vertexing
  - TPC interpolation for SCD calibration



Overview of compute time of reconstruction steps

- **The table below shows the relative compute time (linux cpu time) of the processing steps running on the processor.**
  - **Synchronous reconstruction** fully dominated by the **TPC** (**99%**), no reason to offload anything else to the GPU.
  - In async reco, currently the **61.4% TPC** are on the **GPU**, with the **full optimistic scenario** (full barrel tracking) it will be **79.77**%.
    - **Offloading 60%** of the workload to GPU, should yield a **speed-up** of **2.5x** (since async reco is CPU-bound).

Synchronous processing
(50 kHz Pb-Pb, MC data, processing only)

Processing step	% of time
TPC Processing (Tracking, Clustering, Compression)	99.37 %
EMCAL Processing	0.20 %
ITS Processing (Clustering + Tracking)	0.10 %
TPC Entropy Encoder	0.10 %
ITS-TPC Matching	0.09 %
MFT Processing	0.02 %
TOF Processing	0.01 %
TOF Global Matching	0.01 %
PHOS / CPV Entropy Coder	0.01 %
ITS Entropy Coder	0.01 %
Rest	0.08 %

Asynchronous processing
(650 kHz pp, real data, calorimeters not in run)

Processing step	% of time
TPC Processing (Tracking)	61.41 %
ITS TPC Matching	6.13 %
MCH Clusterization	6.13 %
TPC Entropy Decoder	4.65 %
ITS Tracking	4.16 %
TOF Matching	4.12 %
TRD Tracking	3.95 %
MCH Tracking	2.02 %
AOD Production	0.88 %
Quality Control	4.00 %
Rest	2.32 %

Running on GPU in baseline scenario **Running on GPU in optimistic scenario**

GPU Processing Design principles

1. **GPU code should be modular, such that individual parts can run independently.**
   - Multiple consecutive components on the GPU should operate with as little host interaction as possible.
2. **GPU code should be generic C++ and not depend on one particular vendor or API. (O^2 supports CUDA, HIP, OpenCL)**
   - No usage of special features that are not portable.
3. **GPU usage should be optional and transparent: running O^2 should not require any vendor libraries installed.**
   - All GPU code is contained in plugins, with a common interface.
   - Even multiple plugins (GPU backends) can run on the same node.

For details on GPU implementation see CERN compute accelerator talk: https://indico.cern.ch/event/1264298/

4. **Minimize time spent for memory management.**
   - We allocate one large memory segment, and then distribute memory chunks internally.
5. **Processing on GPU and data transfer should overlap, such that the GPU does not idle while waiting for data.**
   - This is implemented via a pipelined processing within time frames, and we also overlap consecutive time frames.
6. **Data chunks processed by the GPU must be large enough to exploit the full parallelism.**
   - Fulfilled by design with TFs containing > 100 collisions.
7. **GPU and CPU output should be as close as possible.**
   - But small differences due to concurrency or non-associative floating point arithmetic cannot be avoided.

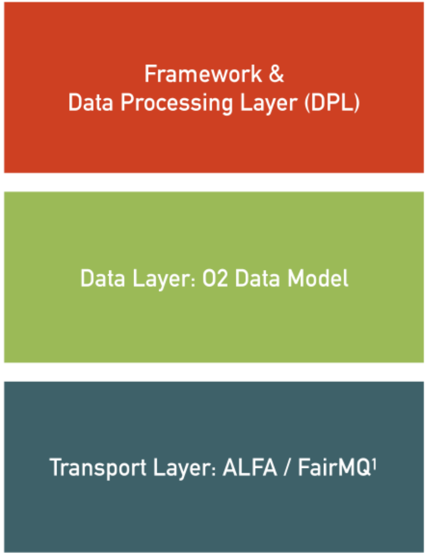
Usage of EPNs for sync/async reconstruction

- **For asynchronous reconstruction, EPN nodes are used as GRID nodes.**
  - **Identical workflow** as on other **GRID** sites, only different configuration using GPU, more memory, more CPU cores.
  - EPN farm split in **2 scheduling pools**: synchronous and asynchronous.
    - Unused nodes in the synchronous pool are moved to the asynchronous pool.
    - As needed for data-taking, nodes are moved to the synchronous pool with lead time to let the current jobs finished.
      - If needed immediately, GRID jobs are killed and nodes moved immediately.

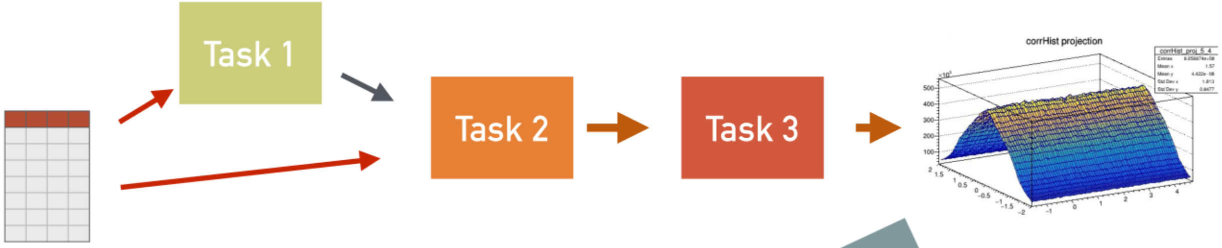
Usage of EPNs for sync/async reconstruction

- **For asynchronous reconstruction, EPN nodes are used as GRID nodes.**
  - **Identical workflow** as on other **GRID** sites, only different configuration using GPU, more memory, more CPU cores.
  - EPN farm split in **2 scheduling pools**: synchronous and asynchronous.
    - Unused nodes in the synchronous pool are moved to the asynchronous pool.
    - As needed for data-taking, nodes are moved to the synchronous pool with lead time to let the current jobs finished.
      - If needed immediately, GRID jobs are killed and nodes moved immediately.

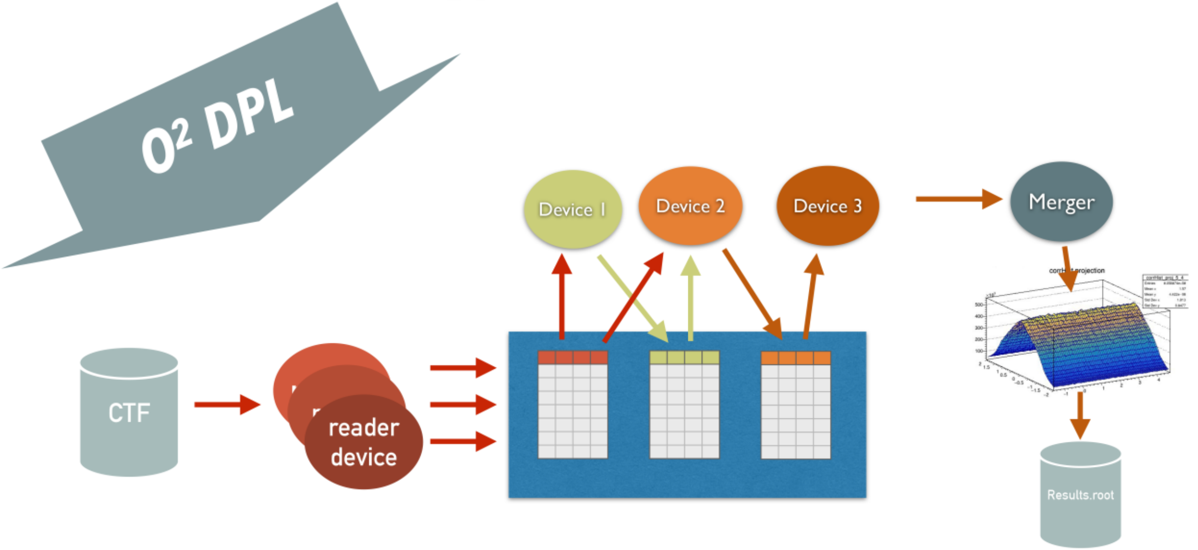
- **A new common software framework was developed for O^2 processing.**
  - Same software for synchronous and asynchronous reconstruction.
  - Same framework runs on the online computing farm, in the GRID, and on the laptop.
  - Layered approach developed jointly with GSI.
  - **Reconstruction steps** in the **processing graph** are independent operating system processes called **devices**.



O^2 Data Processing Layer - Declarative approach

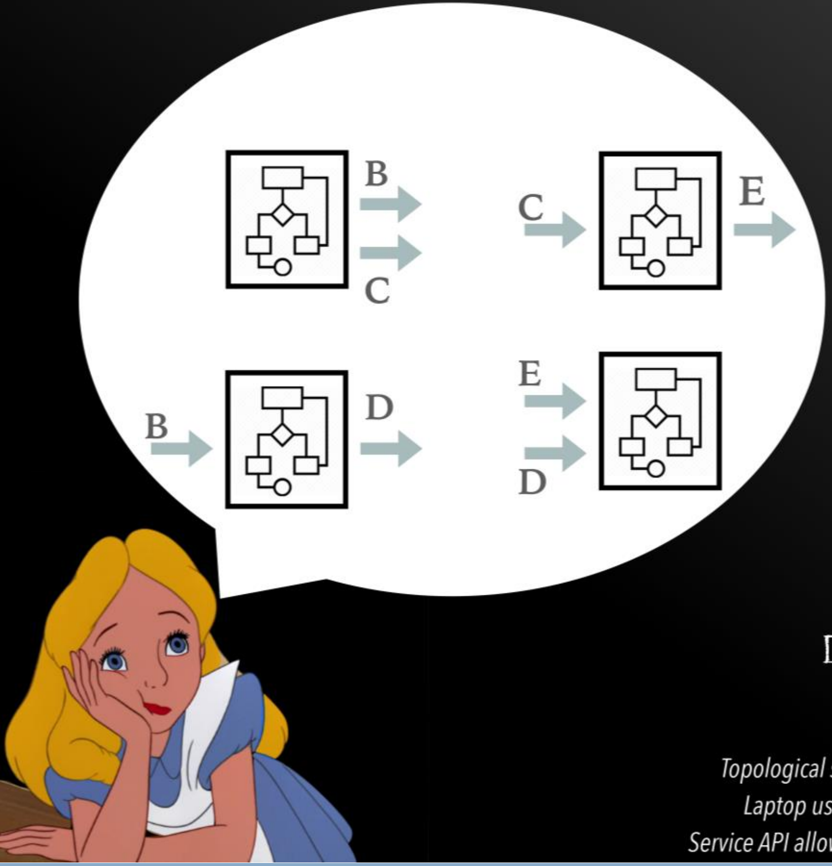


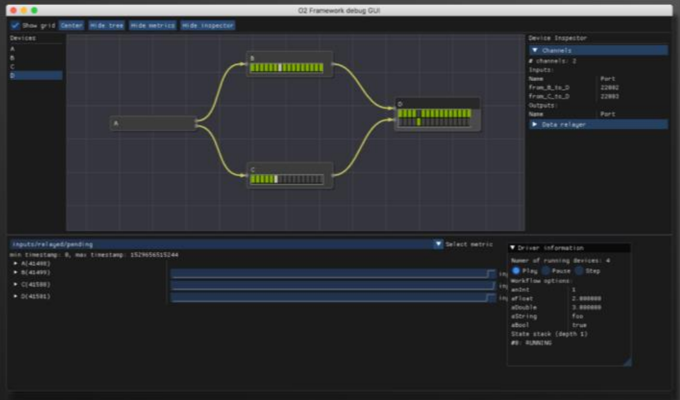
User provides a description in terms of tasks and physics quantities.



O^2 Data Processing Layer (DPL) translates the implicit workflow(s) defined by physicists to an actual FairMQ topology of devices, injecting readers and merger devices, completing the topology and taking care of parallelism & rate limiting.

O² Data Processing Layer - Generating workflows





Data Processing Layer

Topology is defined implicitly.
Topological sort ensures a viable dataflow is constructed (no cycles!).
Laptop users gets immediate feedback through the debug GUI.
Service API allows integration with non data flow components (e.g. Control)

EPN Server Architecture

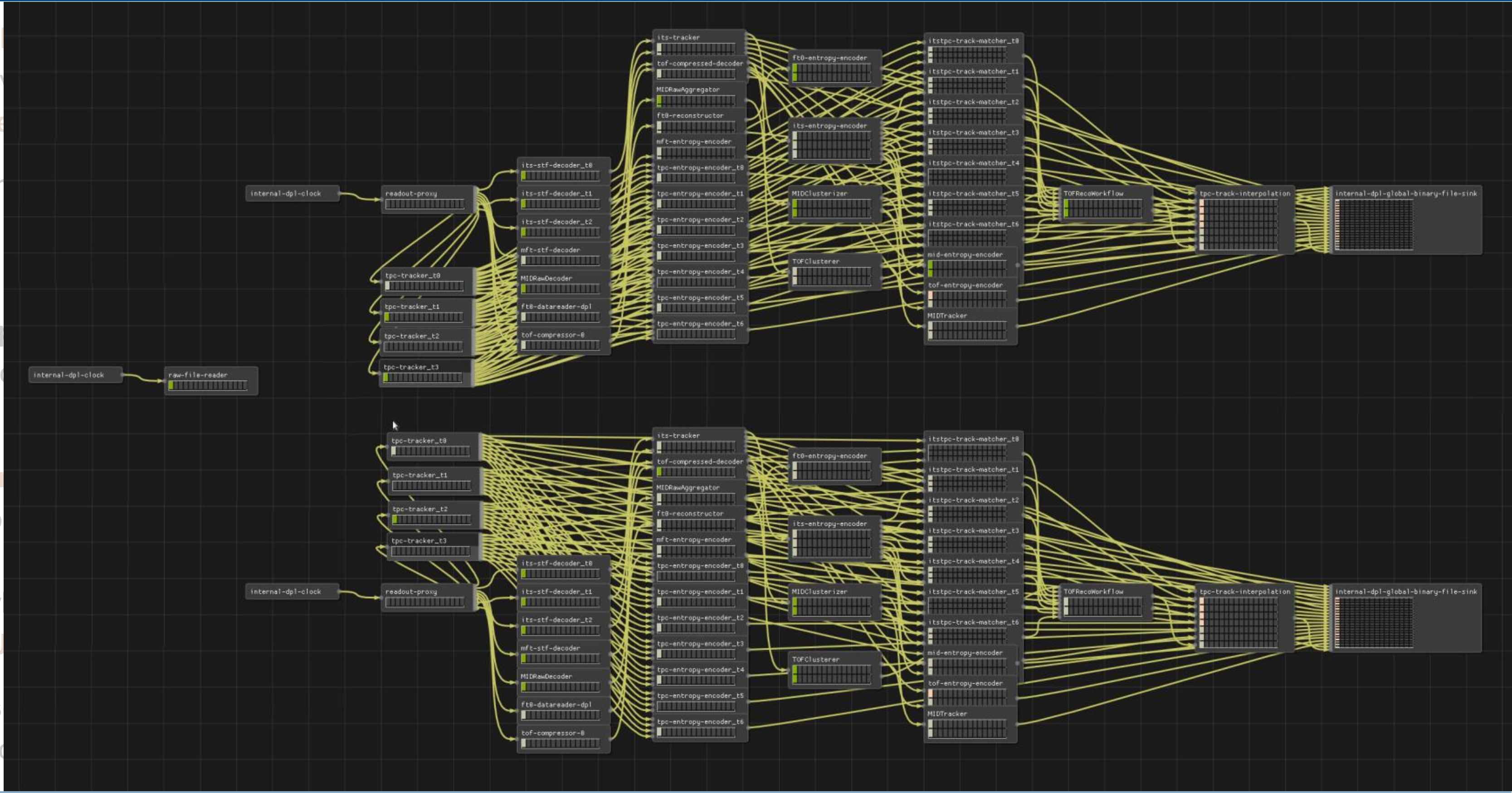
- **Multiple GPUs in a server minimize the cost.**
  - Less servers, less network.
  - **Synergies** of using the **same CPU components** for multiple GPUs, same for memory.
- **Splitting the node into 2 NUMA domains minimizes inter-socket communication**
  → **2 virtual EPNs**.
  - Still only **1 HCA** for the input → writing to shared memory segment in **interleaved memory**.
- **GPUs are processing individual time frames → no inter-GPU communication.**
  - Host processes can drive 1 GPU each, or run CPU only tasks.
- **GPUs can be shared between algorithms.**
  - With **memory reuse** if within the same process.
  - With separate memory in case of multiple processes (Not done at the moment).

EPN Server Architecture

- **Multiple GPUs in a server minimize the cost.**
  - Less servers, less network.
  - **Synergies** of using the **same CPU components** for multiple GPUs, same for memory.
- **Splitting the node into 2 NUMA domains minimizes inter-socket communication**
  → **2 virtual EPNs**.
  - Still only **1 HCA** for the input → writing to shared memory segment in **interleaved memory**.
- **GPUs are processing individual time frames → no inter-GPU communication.**
  - Host processes can drive 1 GPU each, or run CPU only tasks.
- **GPUs can be shared between algorithms.**
  - With **memory reuse** if within the same process.
  - With separate memory in case of multiple processes (Not done at the moment).
- **Benchmarked with MC data: For 100% utilization of 8 GPUs** (AMD MI50)**, we need:**
  - ~**50 CPU cores**, ~**400 GB** of memory, **30 GB/s** network input speed, GPU PCIe negligible.
- **Selected server:**
  - Supermicro AS-4124GS-TNR, **8 * MI50** GPU, **2 * 32 core** AMD Rome 7452 CPU (2.35 GHz), **512 GB RAM** (16 * 32GB)
  - Infiniband HDR / HDR100 network.

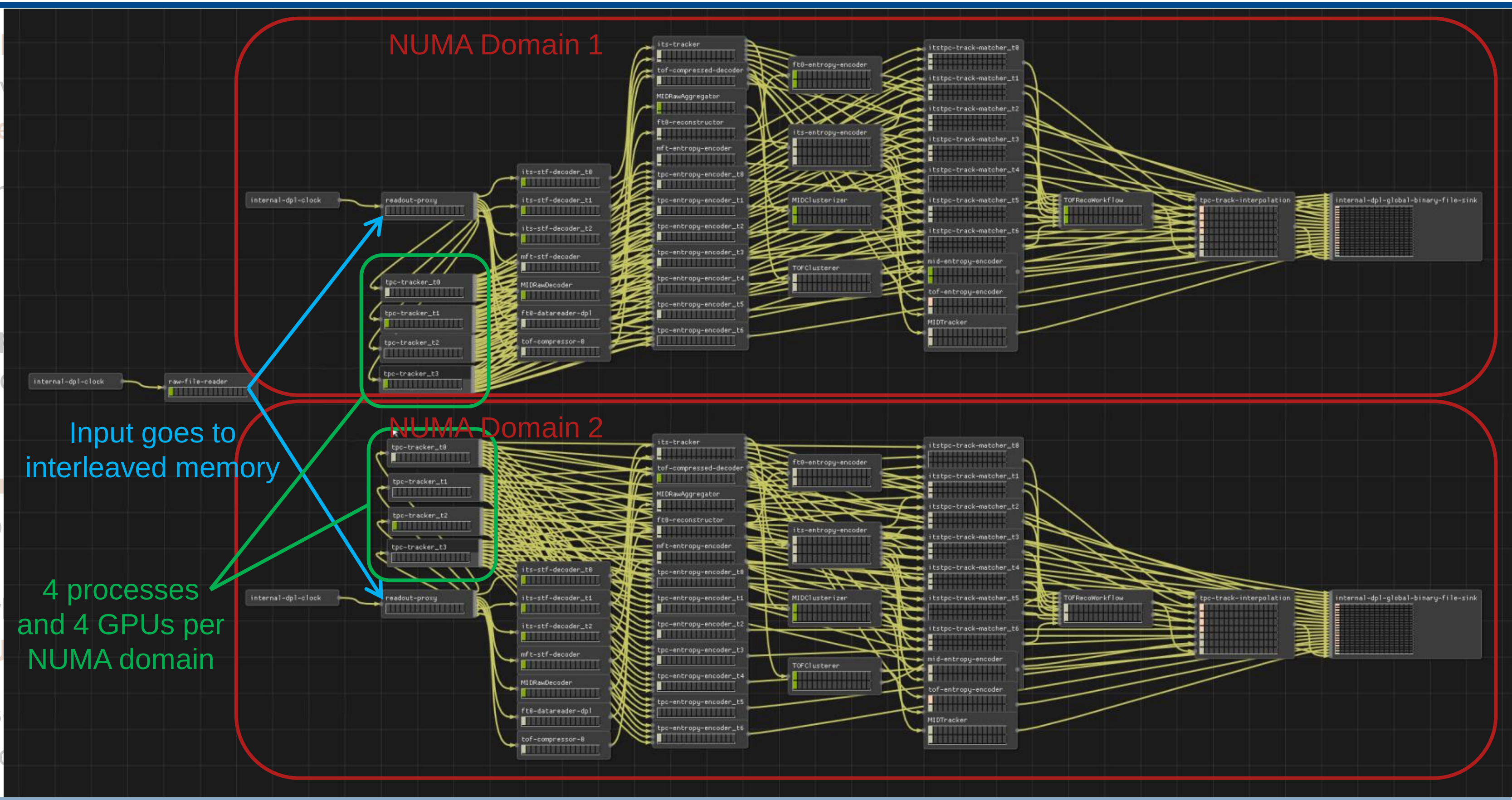
Synchronous processing DPL workflow

- Multiple GI
  - Less ser
  - Synergie
- Splitting th
  → 2 virtual
  - Still only
- GPUs are
  - Host proc
- GPUs can
  - With me
  - With sep
- Benchmar
  - ~50 CPU
- Selected s
  - Supermic



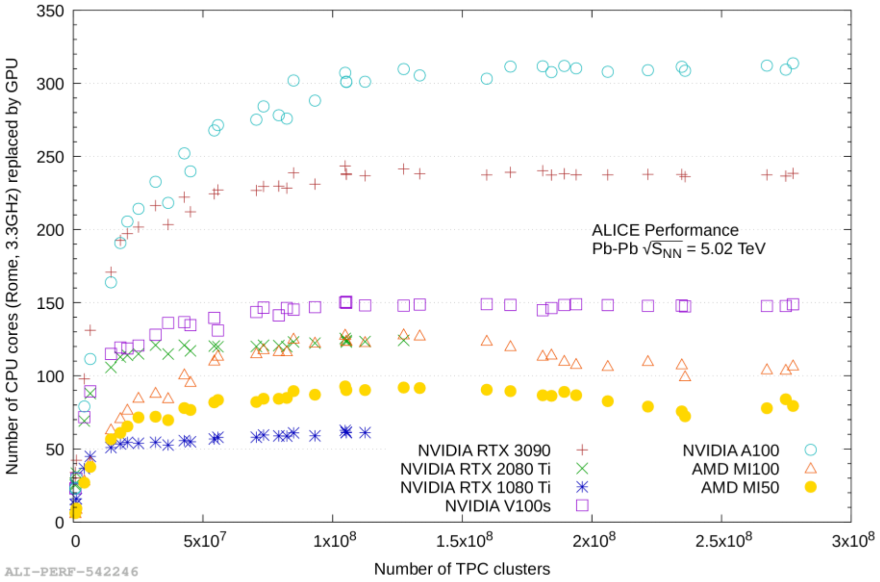
Synchronous processing DPL workflow

- Multiple GI
  - Less ser
  - Synergie
- Splitting th
  → 2 virtual
  - Still only
- GPUs are
  - Host proc
- GPUs can
  - With me
  - With sep
- Benchmar
  - ~50 CPU
- Selected s
  - Supermi



Performance of EPN servers

- **Performance of EPN servers in synchronous reconstruction mostly evaluated processing 50 kHz Pb-Pb MC data.**
  - **GPU-bound**, CPU resource usage is ~**44 cores** (out of 64 cores available).
  - Evaluated several GPU models, current farm has 280 EPN servers with MI50 GPUs and 70 newer servers with MI100 GPUs.
  - **1 MI50 GPU** replaces ~**80 CPU cores** in **synchronous reconstruction** and ~**55 cores** in **asynchronous reconstruction**. (measured against the AMD Rome 7452 cores in the EPN server).
  - Current EPN farm has the capacity to process the estimated 50 kHz Pb-Pb rate with a **30% margin**.
  - **Without GPUs**, would need **>2000 64-core servers**.



Performance of EPN servers

- **Performance of EPN servers in synchronous reconstruction mostly evaluated processing 50 kHz Pb-Pb MC data.**
  - **GPU-bound**, CPU resource usage is ~**44 cores** (out of 64 cores available).
  - Evaluated several GPU models, current farm has 280 EPN servers with MI50 GPUs and 70 newer servers with MI100 GPUs.
  - **1 MI50 GPU** replaces ~**80 CPU cores** in **synchronous reconstruction** and ~**55 cores** in **asynchronous reconstruction**. (measured against the AMD Rome 7452 cores in the EPN server).
  - Current EPN farm has the capacity to process the estimated 50 kHz Pb-Pb rate with a **30% margin**.
  - **Without GPUs**, would need **>2000 64-core servers**.
- **Asynchronous Performance benchmarks cover multiple cases (In all cases server fully loaded with identical jobs):**
  - EPN split into 16 * **8 cores**, or into 8 * **16 cores**, ignoring the GPU : to compare CPUs and GPUs.
  - EPN split into 8 or 2 identical fractions: **1 NUMA** domain (4 GPUs) or **1 GPU**.

Configuration (2022 pp, 650 kHz)	Time per TF (11ms, 1 instance)	Time per TF (11ms, full server)
CPU 8 core	76.91s	4.81s
CPU 16 core	34.18s	**4.27s**
1 GPU + 16 CPU cores	14.60s	1.83s
1 NUMA domain (4 GPUs + 64 cores)	3.5s	**1.70s**

Factor 2.51

Matches expected factor 2.5

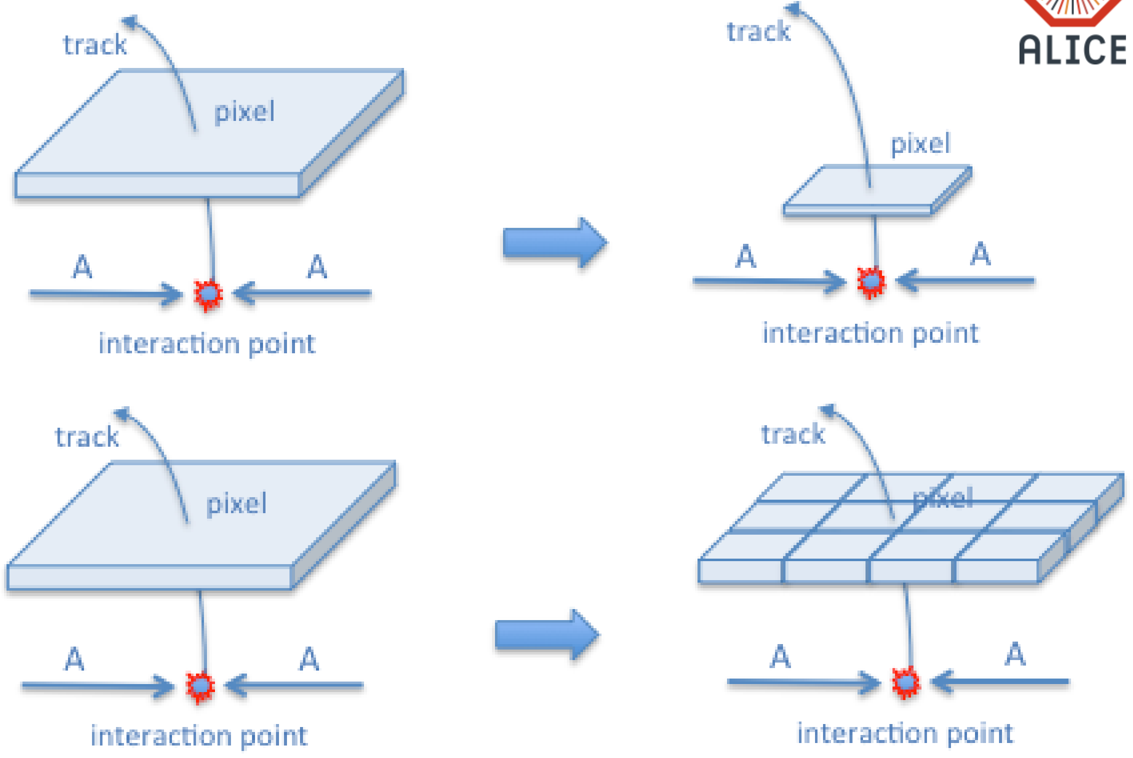
Configuration used for production

ITS2

ITS2 design objectives

Improve **impact parameter resolution** by a factor ~3(5) in rφ(z) at p_T = 500 MeV/c

- get closer to IP: 39 mm → 23 mm (innermost layer)
- reduce material budget: ~ 1.14% X_0 : → 0.36% X_0 per layer (for the inner layers)
- reduce pixel size: 50 µm ✕ 425 µm → O(30 ✕ 30 µm²)

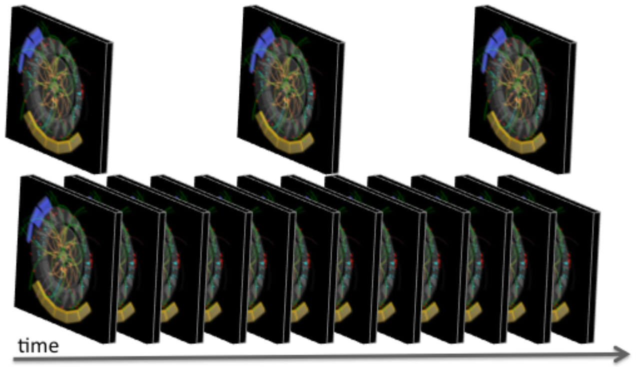


Improve tracking efficiency and p_T resolution at low p_T

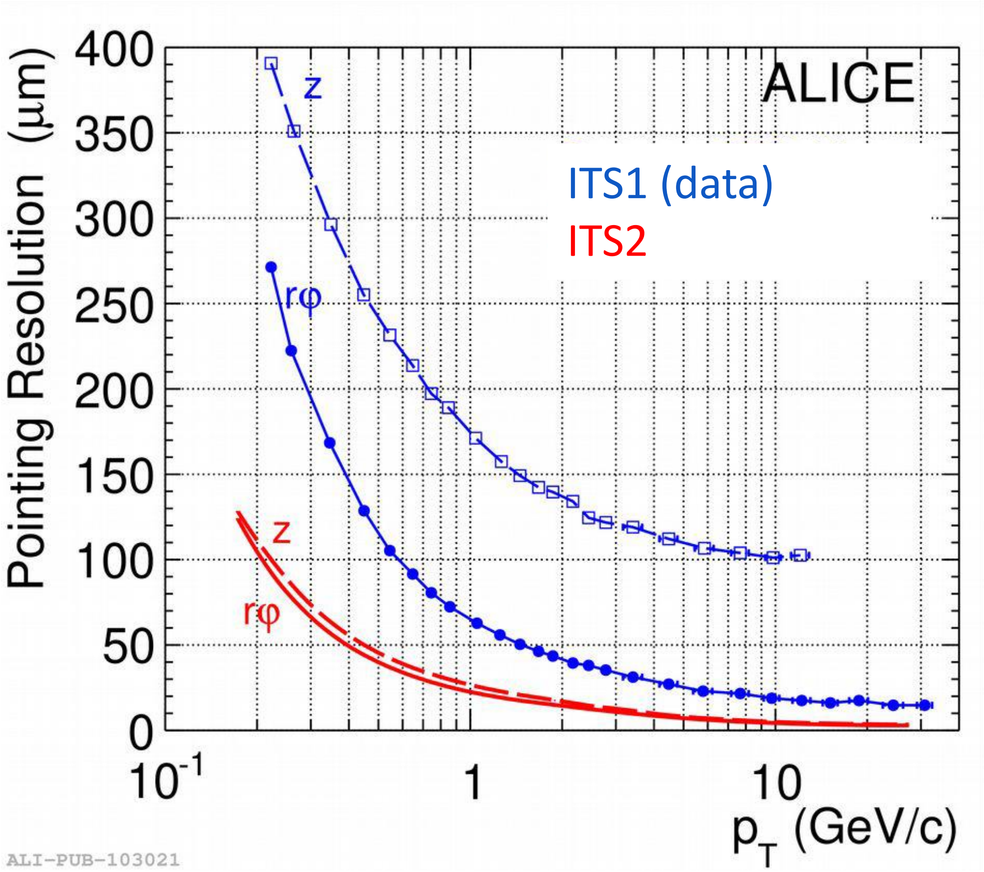
- Increase granularity: 6 layers ➔ 7 pixel layers

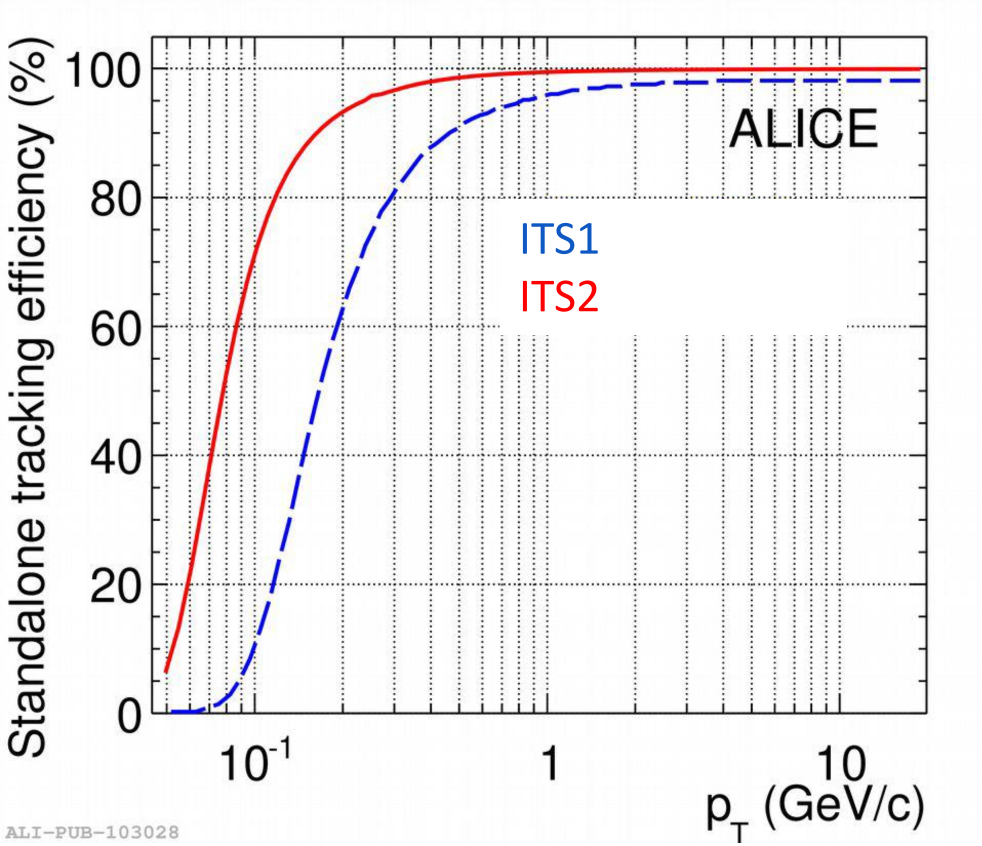
Fast readout

- Readout of Pb-Pb at up to 100 kHz (previously 1 kHz) and 400 kHz for pp



Detector performance in Run 3 — simulations





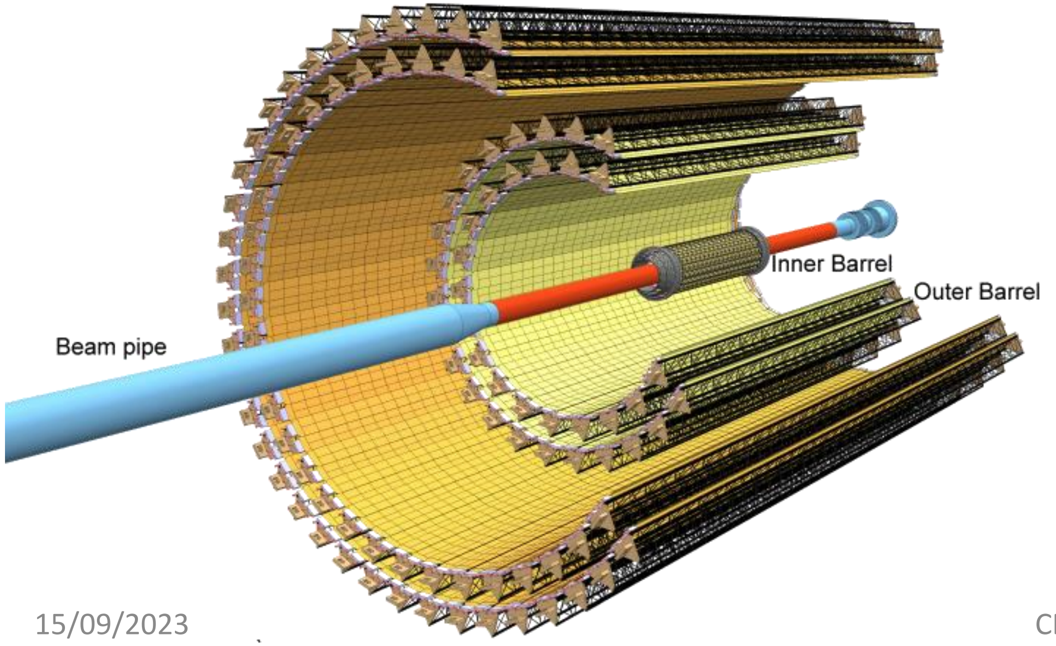
- Improved tracking efficiency (95% instead of 60% at 200 MeV/*c*)
- Pointing resolution 3x better in transverse plane (6x along beam axis) at 200 MeV/*c*

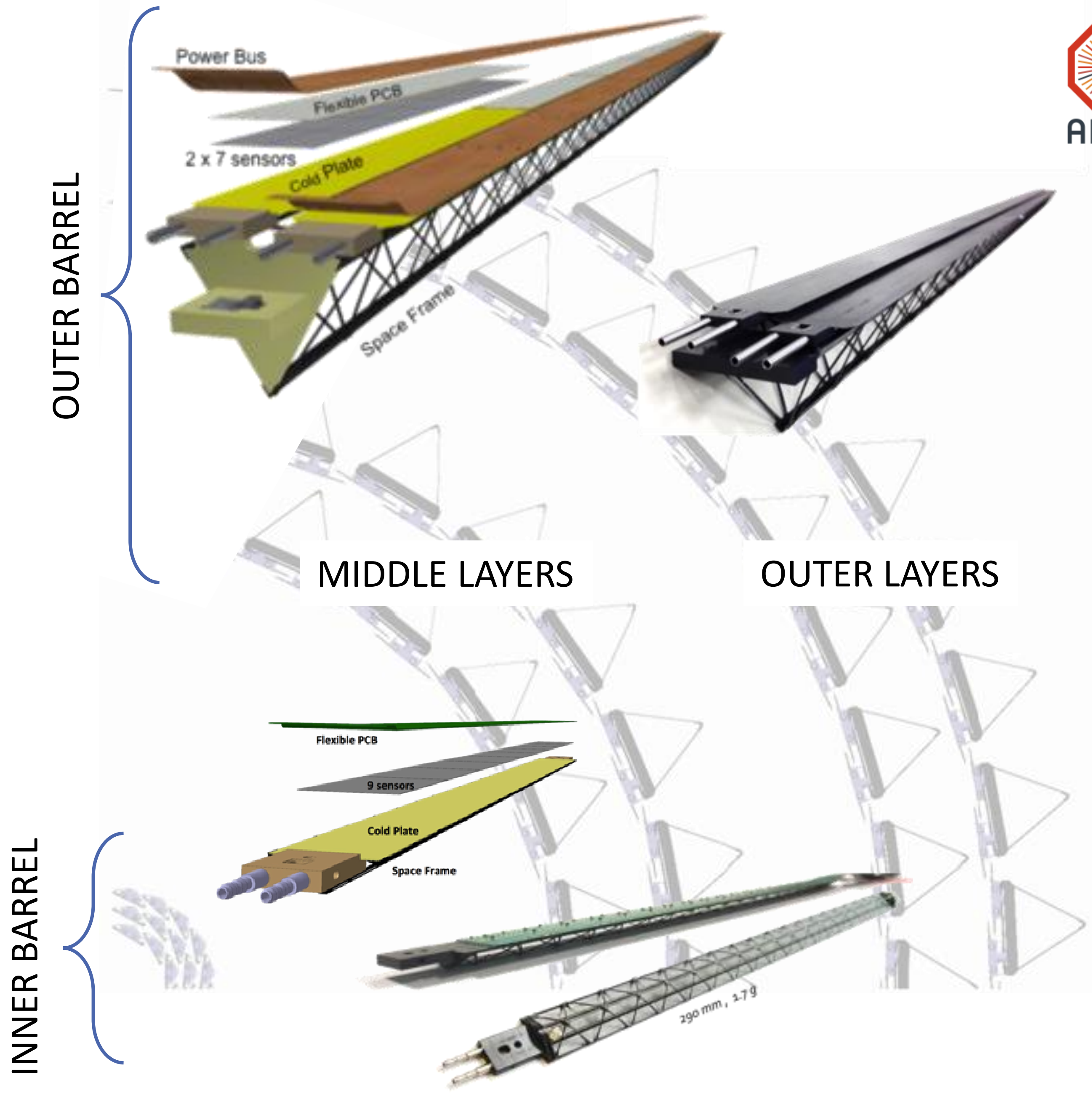
ITS2 layout

- 7 layers (inner/middle/outer): 3/2/2 from R = 23 mm to R = 400 mm
- 192 staves (IL/ML/OL): 48/54/90
- Ultra-lightweight support structure and cooling

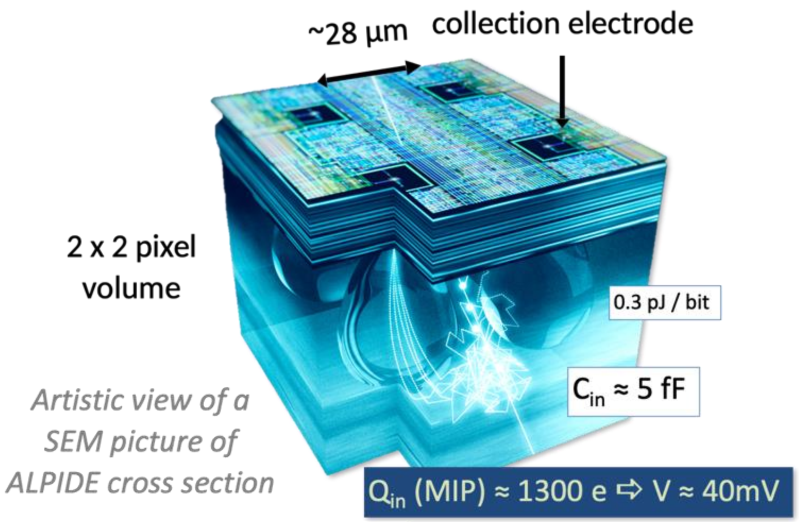
10 m^2 active silicon area, 12.5×10^9 pixels

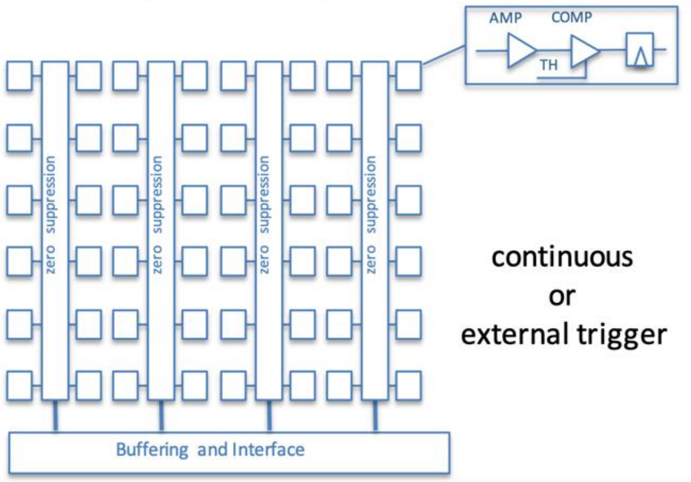
more details on construction, installation and commissioning in Felix Reidt’s seminar https://indico.cern.ch/event/1210704/





The ALPIDE chip: A Monolithic Active Pixel Sensor (MAPS) developed within the ITS2 project





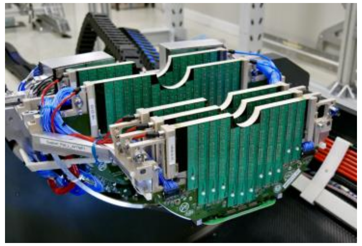
Technology

- TowerJazz 180 nm CMOS Imaging Process
- High-resistivity (> 1kΩ cm) p-type epitaxial layer (25 μm) on p-type substrate
- Small n-well diode (2 μm diameter), ~100 times smaller than pixel (~30 μm) ➔ low capacitance (~fF)
- Reverse bias voltage ($-6\ V < V_{BB} < 0\ V$) to substrate to increase depletion zone around NWELL collection diode
- Deep PWELL shields NWELL of PMOS transistors

Key features

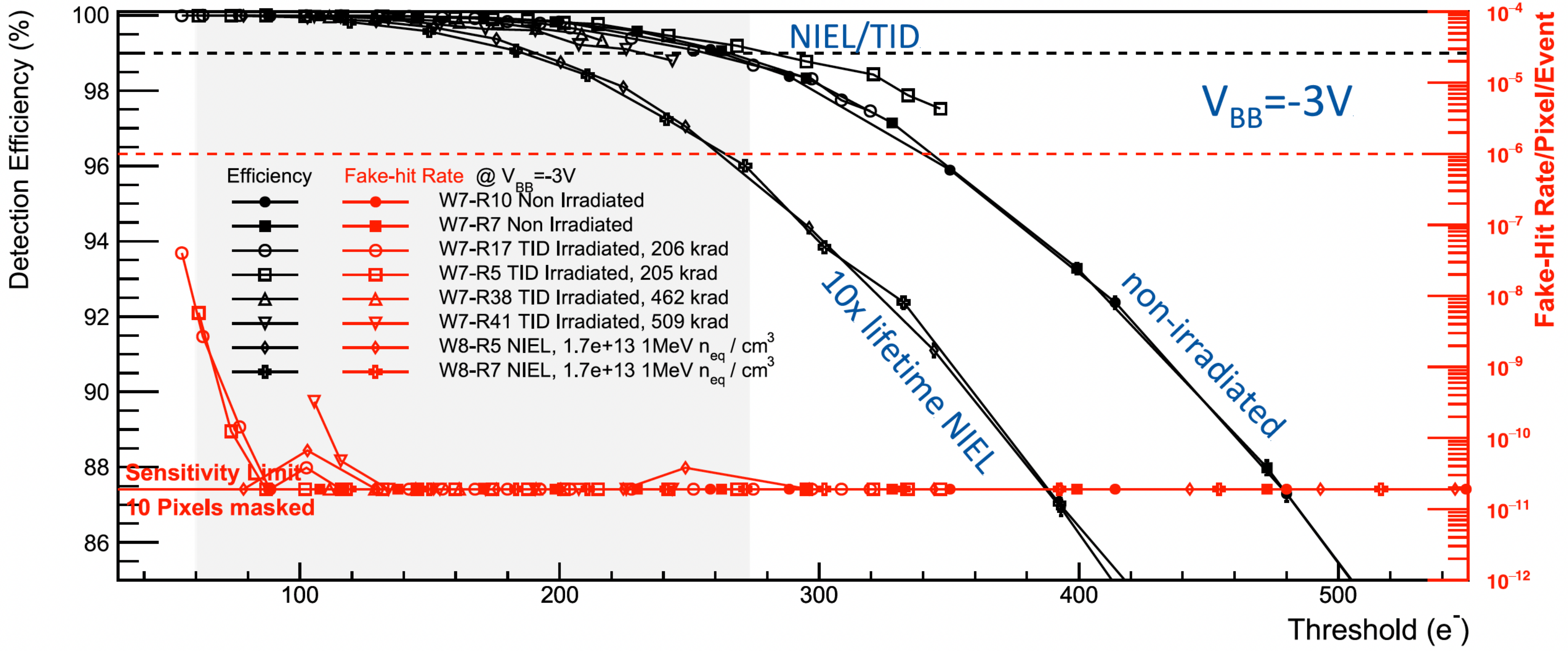
- In-pixel amplification and shaping, discrimination and Multiple-Event Buffers (MEB)
- In-matrix data sparsification
- On-chip high-speed link (1.2 Gbps)
- Low total power consumption < 47 mW/cm^2

ALPIDE chip used for many other applications: ALICE MFT, sPHENIX MVTX, etc...



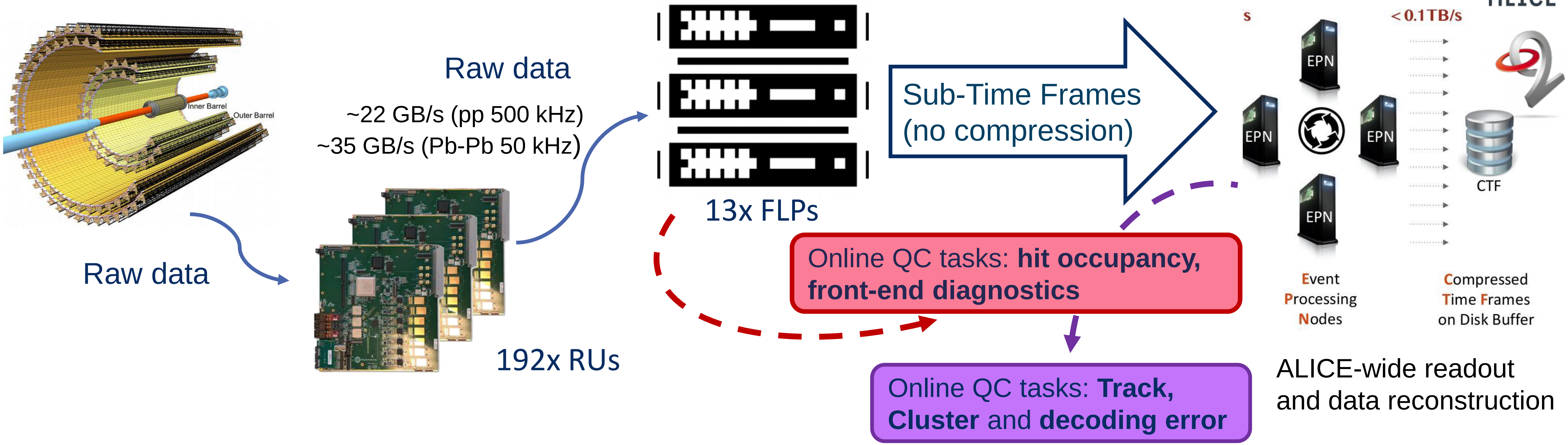
MFT

ALPIDE performance — detection efficiency and fake-hit rate



Availability and excellent support from test-beam facilities all over the world have been key for the development of this chip: BTF Frascati, CERN, DESY, LBNL, UC Louvain la Neuve, Pohang (Korea), Rez (Czech Republic), SLRI (Thailand)

Data readout architecture and quality control (QC)



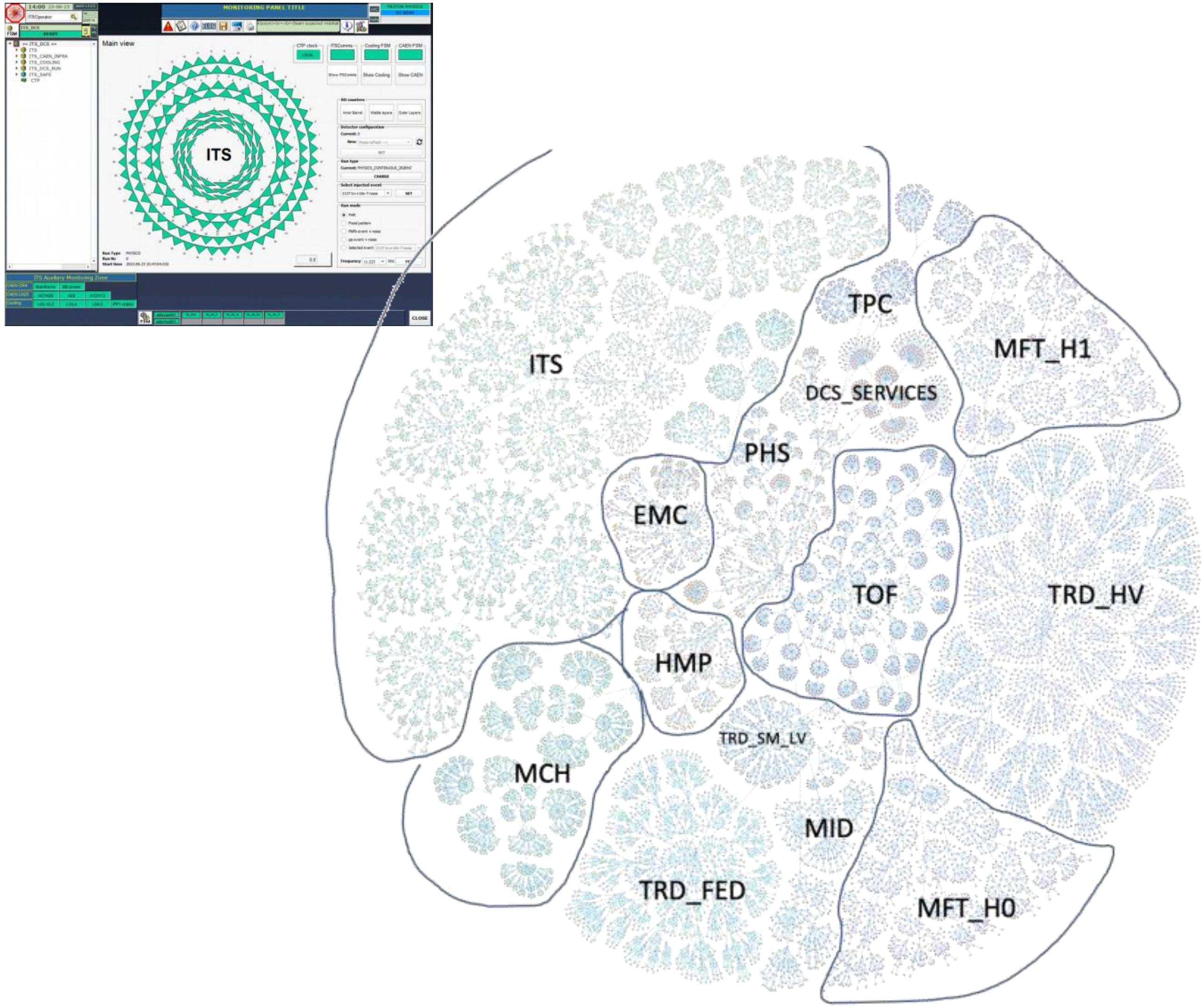
- 13 ITS First Level Processors (FLPs)
  - Online quality control tasks: hit occupancy and front-end electronics diagnostics.
- 350 Event Processing Nodes (total EPN from ALICE farm)
  - Online quality control tasks: reconstructed ITS2 tracks, clusters and decoding errors.
- Synchronous reconstruction, calibration and data compression (→ GPUs)
- Asynchronous stage: reconstruction with final calibration → final Analysis Object Data (AOD)

Detector Control System (DCS) - a quick view

- User Interface developed in WinCC – detector logic implemented in a Finite State Machine
- Detector operation, monitoring and archiving of detector data
- Deal with ~110000 data points (ITS only)

→ typical monitoring frequency of 1 Hz

- Built as a hierarchical system (partitioned with system of locks) → ITS occupies a big slice of the ALICE hierarchy
- An independent safety system (ITS2S) interlocks power channels based on stave temperatures and cooling status



ITS2 calibration (1)

Main ITS calibration features:

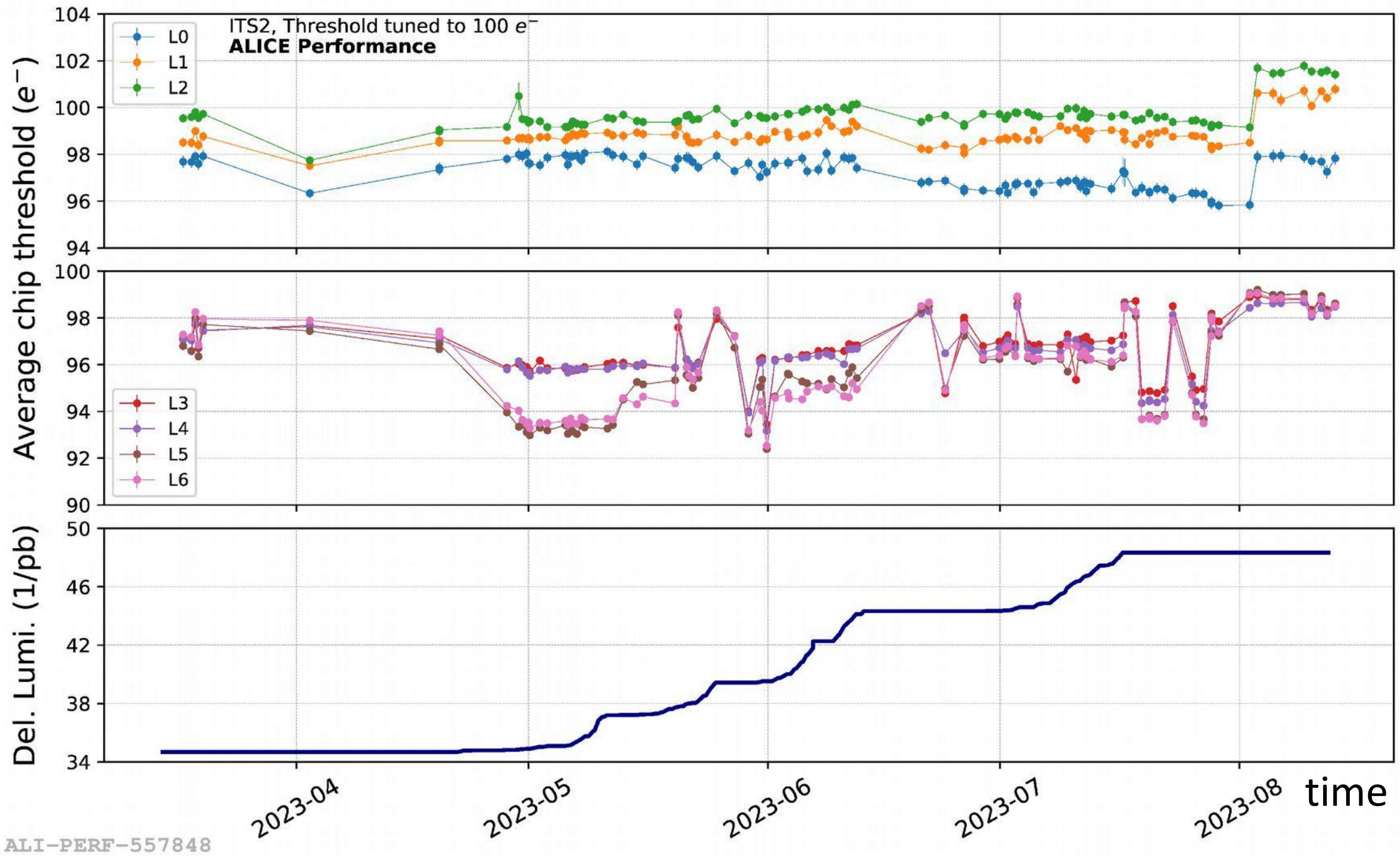
- **Masking of noisy pixels**
- **Tuning of in-pixel discriminating thresholds**
- Optimise power supply voltage
- Measure on-chip temperatures

Threshold calibration of 24120 chips is challenging:

- Online on 40 EPNs in parallel
- ~1% pixels pulsed: ~252Ghits
- Threshold target: $100e^-$ (chip-to-chip RMS < 6 e^- ; on-chip RMS ~20 e^-)

Non-working pixels: ~0.2 %

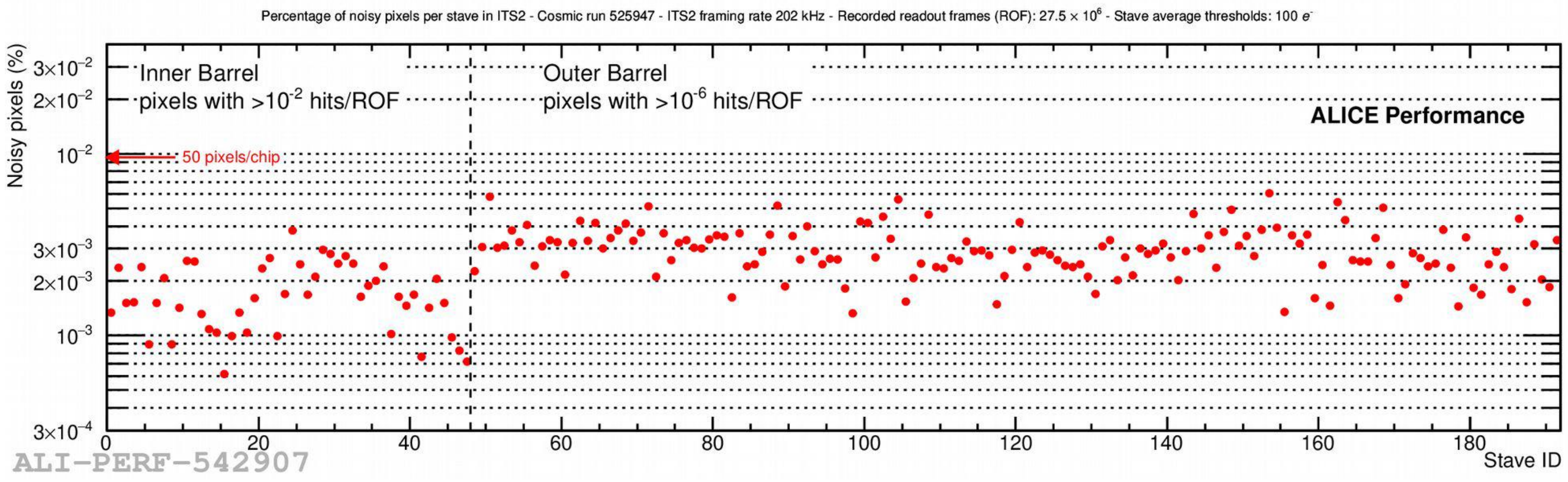
THRESHOLDS



Thresholds stable during 2023 without retuning

Fluctuations of 2-3 e^- due to optimizations of the voltage to chips

ITS2 calibration (2)



Noisy pixel definition:

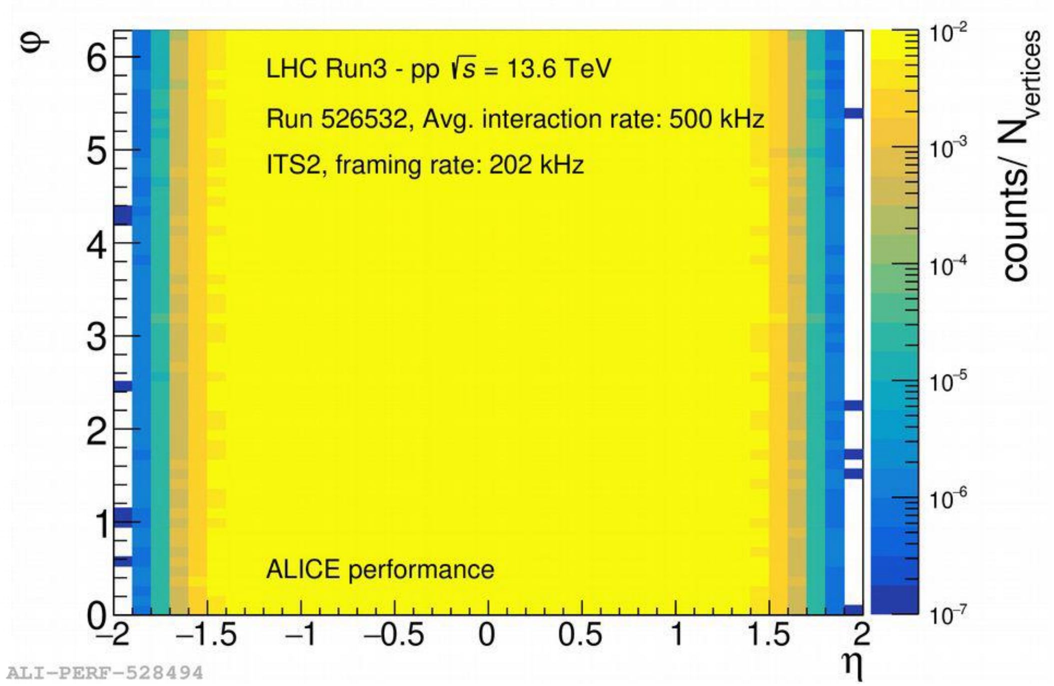
- IB: occupancy > 10^{-2} hits/event
- OB: occupancy > 10^{-6} hits/event

Percentage of noisy pixels masked per stave is extremely small: ~0.02-0.03 ‰

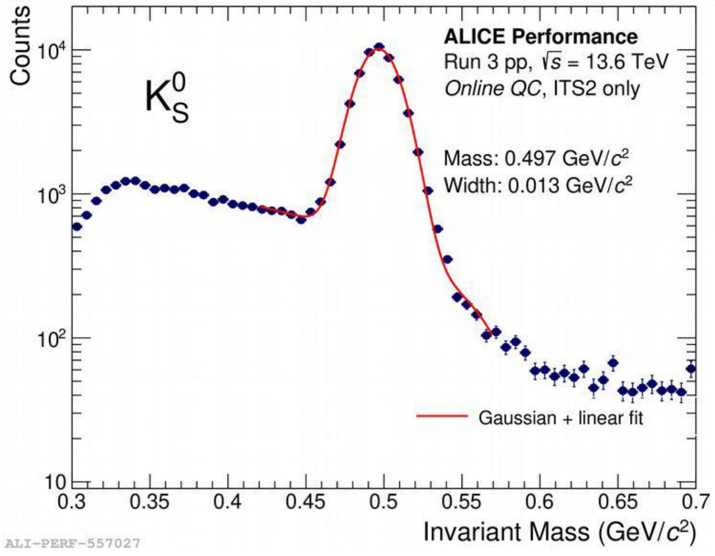
Fake-hit rate trend during cosmic runs (tuned thresholds + noise masks): stable and < 10^{-6} hits/event/pixel (design requirement) by masking only ~0.03‰ of the pixels (15 pixels)

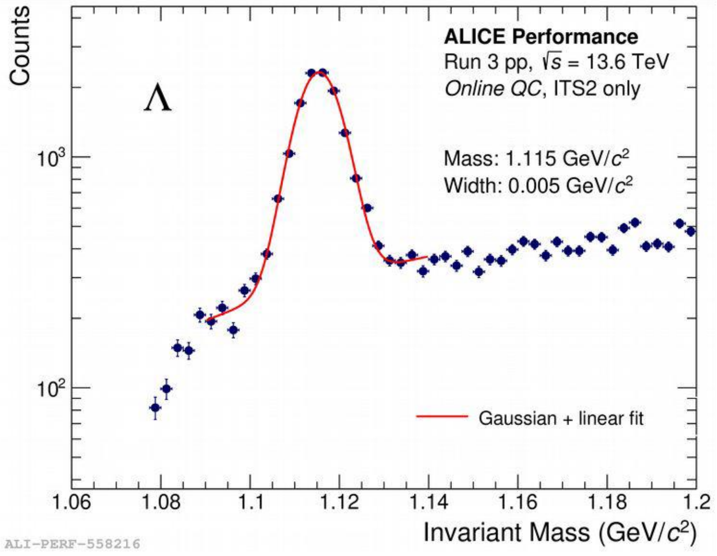
Preliminary performance

- ITS tracking: excellent performance with current detector alignment
  - Cellular automaton algorithm
  - Online tracking for quick QA of the data
  - Angular distribution of tracks of good quality → good detector acceptance



- Online physics performance from QC through Λ and $K0_S$ invariant mass peaks



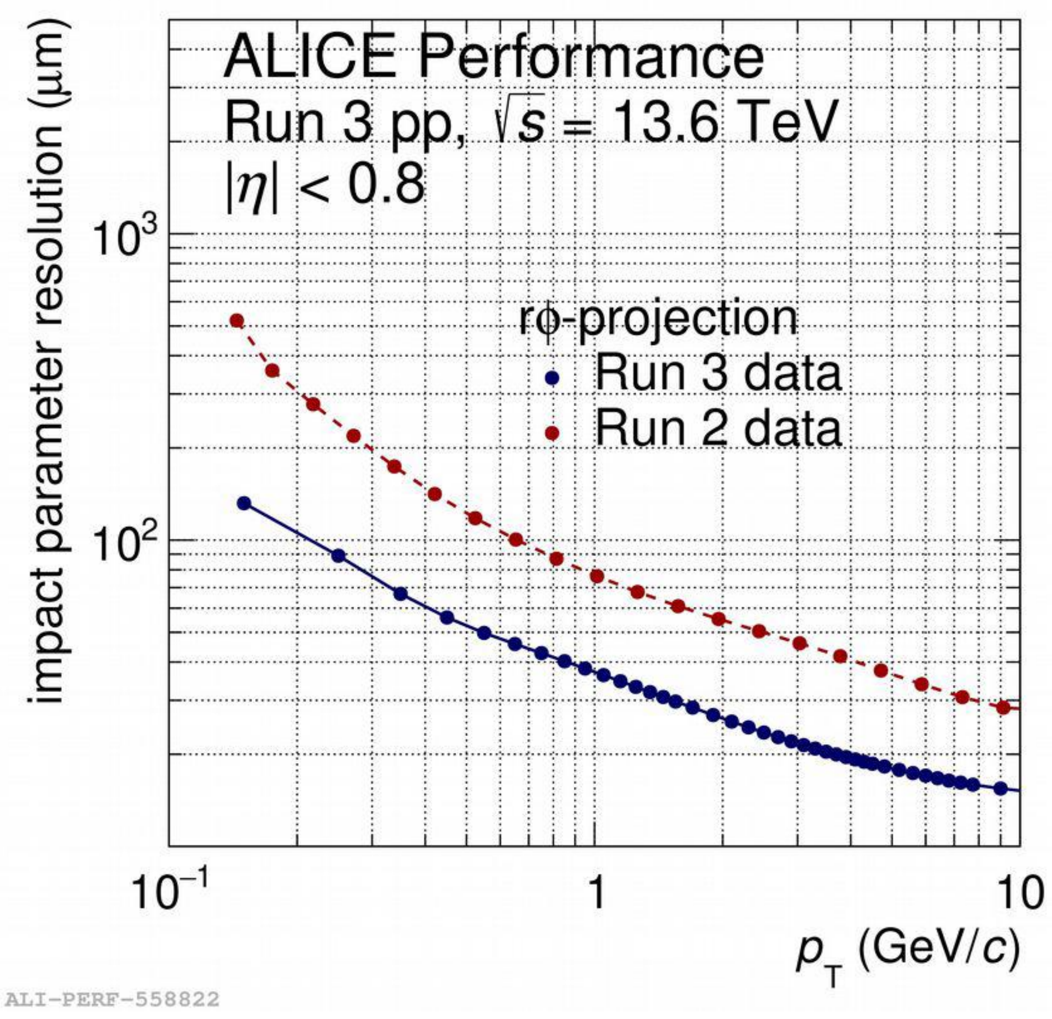


from online QC

Performance: impact parameter

Impact parameter resolution measured with Run3 pp data → excellent performance

- About 2.5x improvement at p_T = 500 MeV/c
- Detector alignment, space charge corrections and calibrations still continuously improving



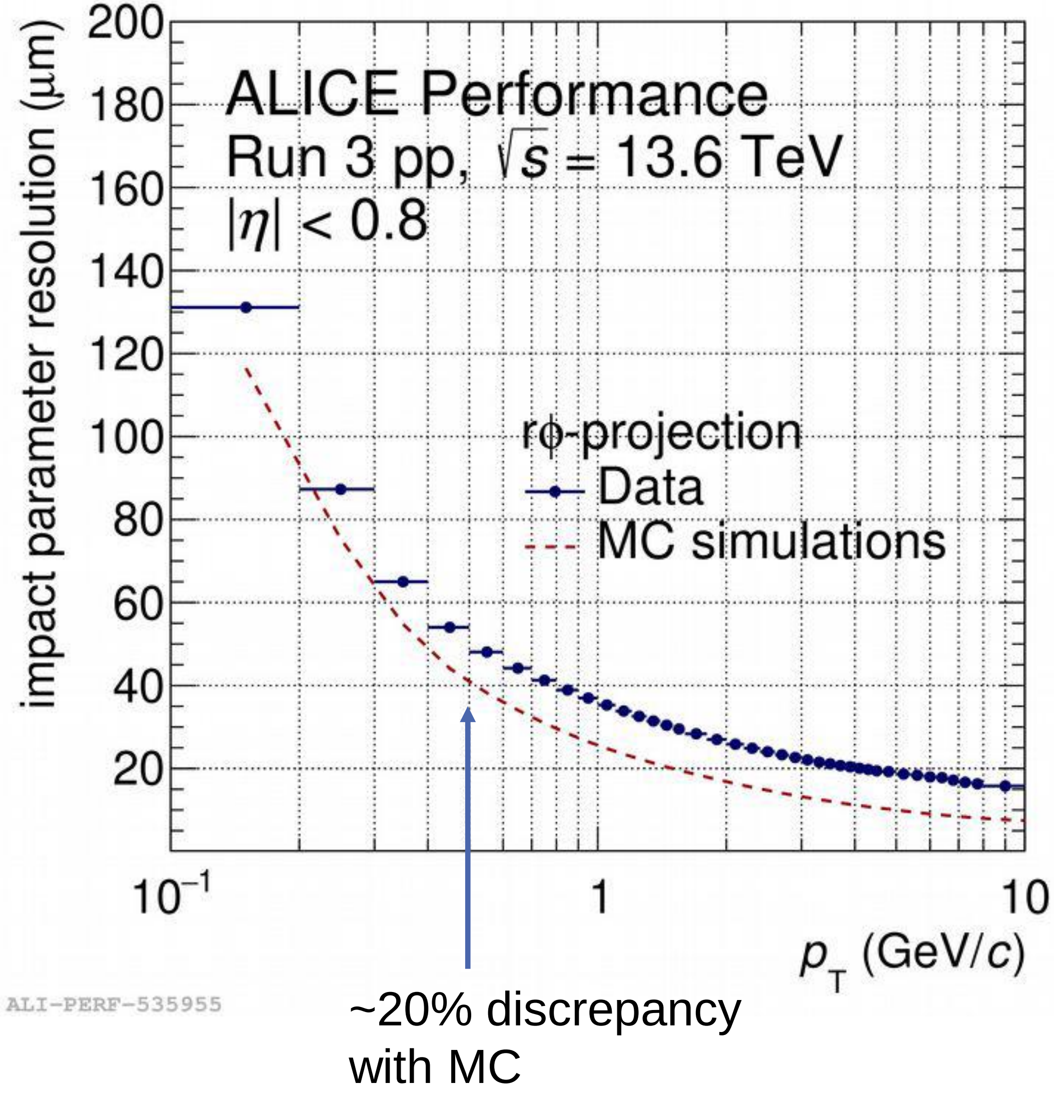
Global tracks with at least 1 hit in Inner Barrel (Run 3) or in the two innermost pixel layers (Run 2)

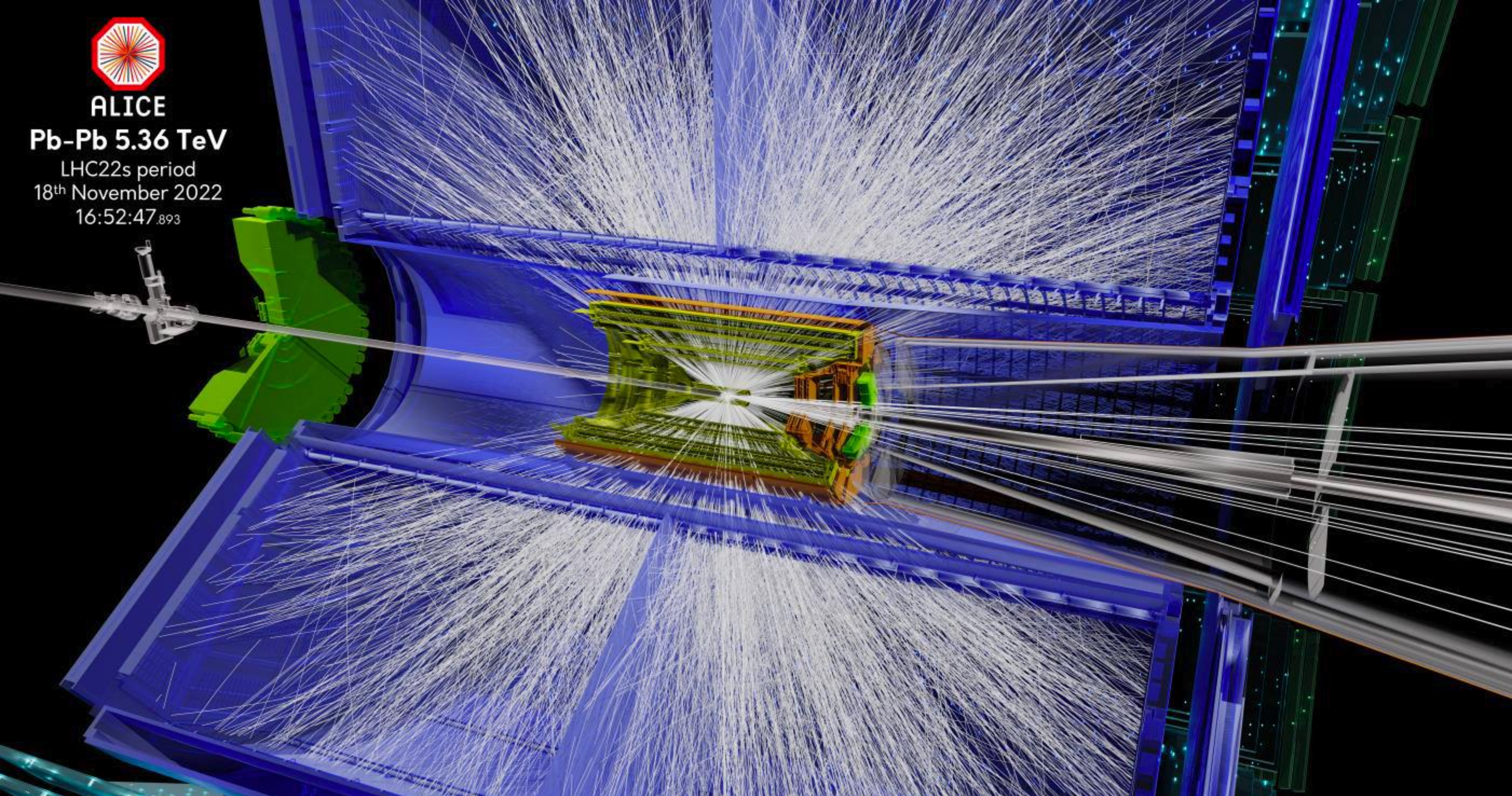
Performance: impact parameter

Impact parameter resolution measured with Run3 pp data → excellent performance

- About 2.5x improvement at p_T = 500 MeV/c
- Detector alignment, space charge corrections and calibrations still continuously improving

remaining ~20% discrepancy with MC




ALICE
Pb-Pb 5.36 TeV
LHC22s period
18th November 2022
16:52:47.893

TPC

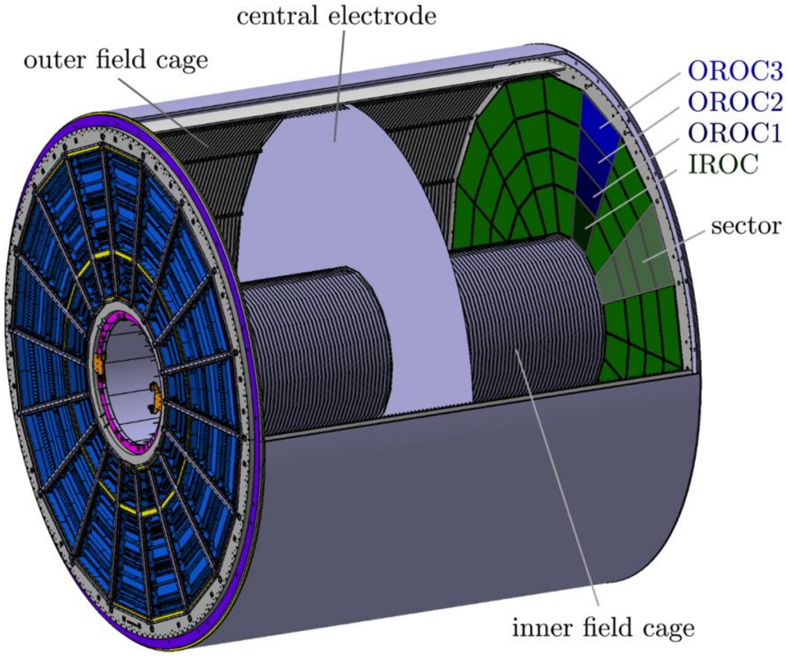
The ALICE TPC

- A large tracking and PID device in the central barrel of ALICE
  - Cylindrical drift volume, 5 m long, 5 m diameter
  - Two sides, split by central drift electrode
  - 18 azimuthal sectors with readout chambers per side
  - ~100 us electron drift time for max. drift distance

- The past: MWPC readout until 2018
  - < 2 kHz event readout rate with Pb–Pb collisions

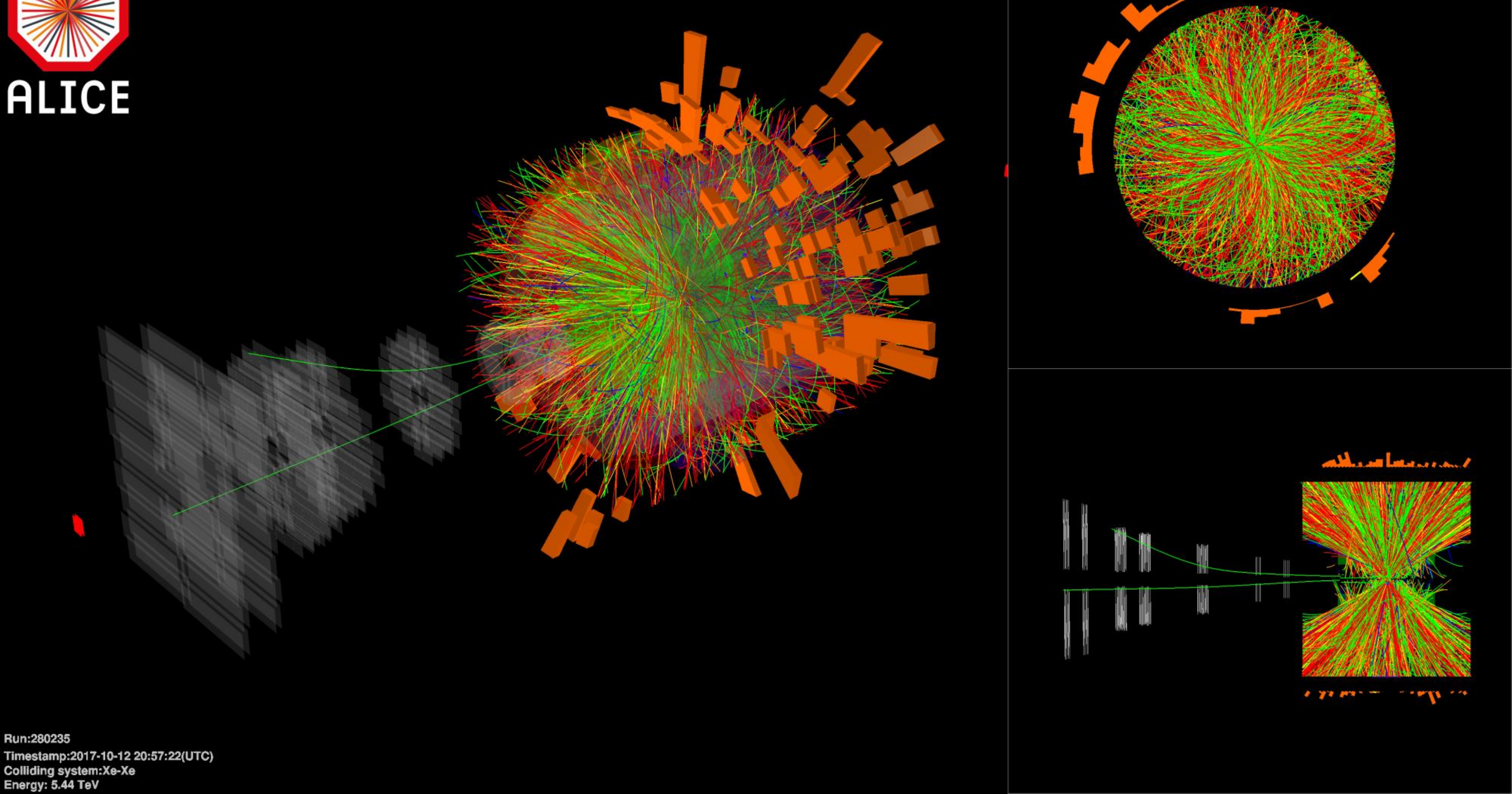
- The present: Continuous readout
  - 50 kHz collision rate with the requirement to read out ALL min. bias Pb—Pb collisions
  - No dead time allowed, no triggering, no gating

➔ **need to minimise ion backflow**

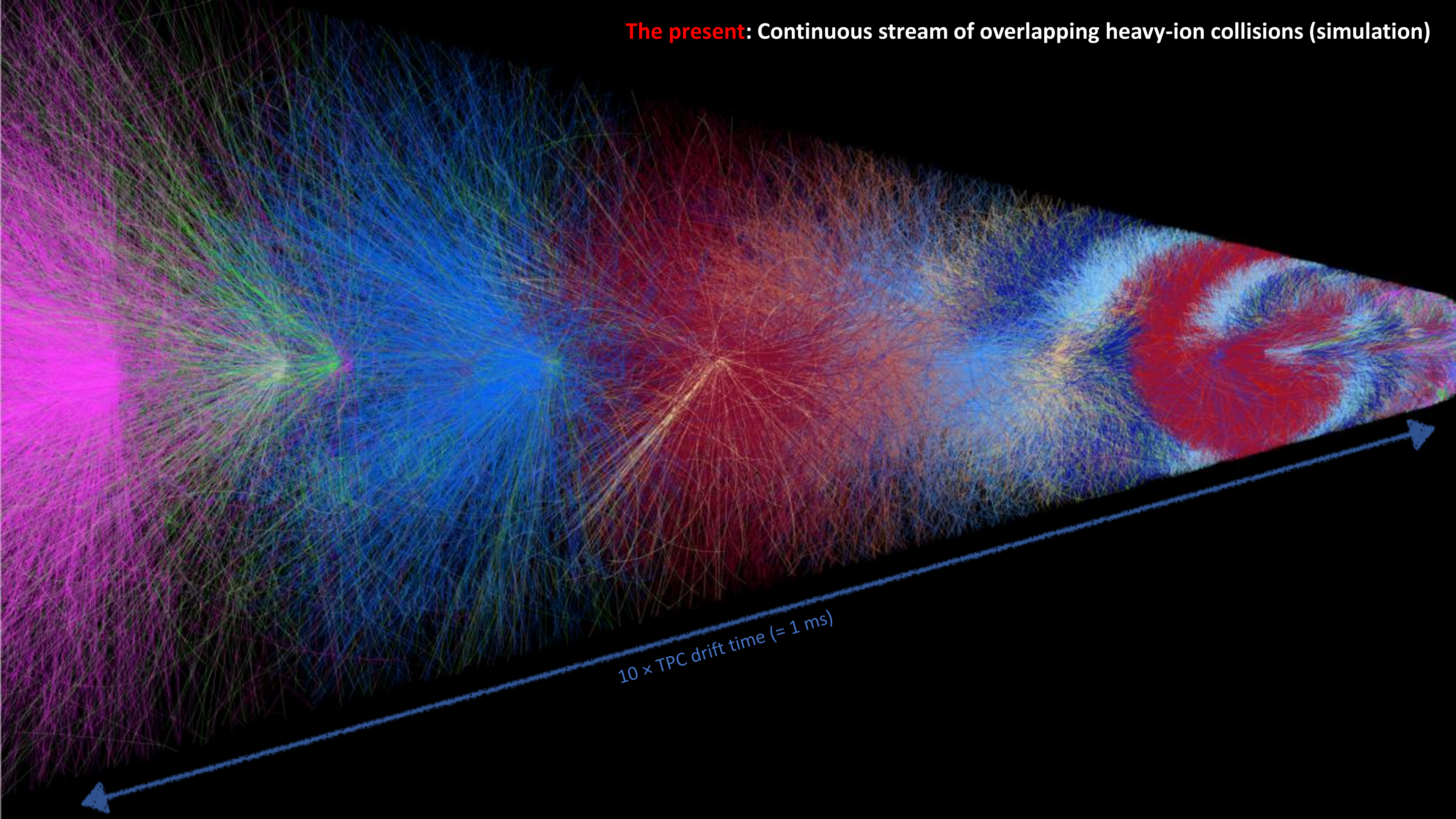


[ALICE TPC Collaboration – JINST 16 – 2021]

The past: Example heavy-ion collision. One triggered event

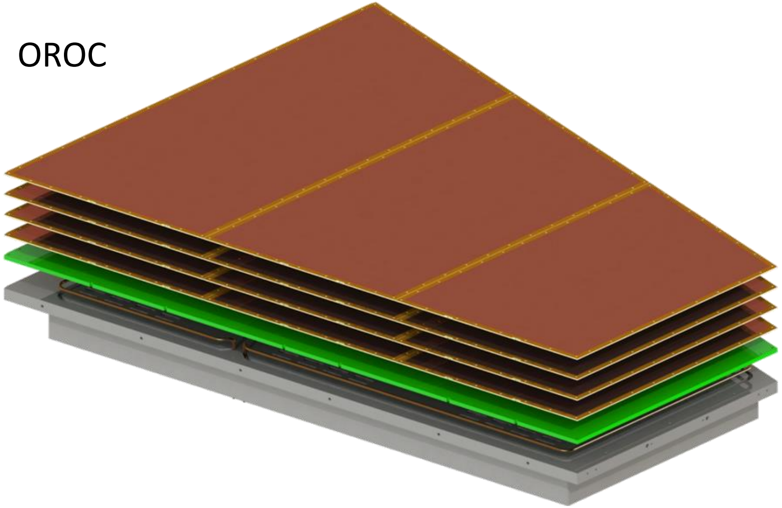


The present: Continuous stream of overlapping heavy-ion collisions (simulation)

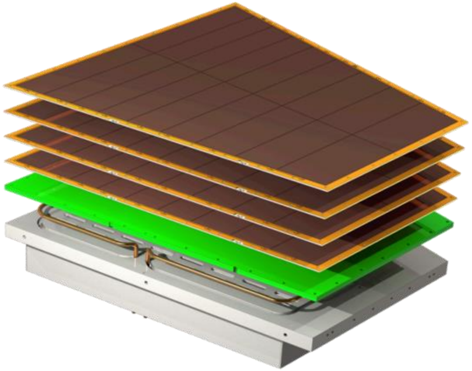


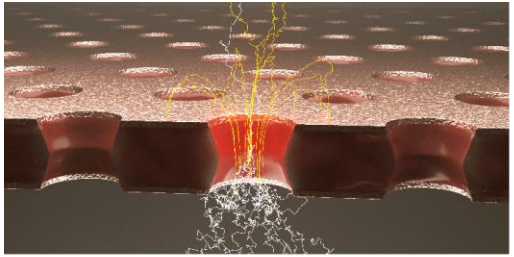
Readout chambers

OROC



IROC

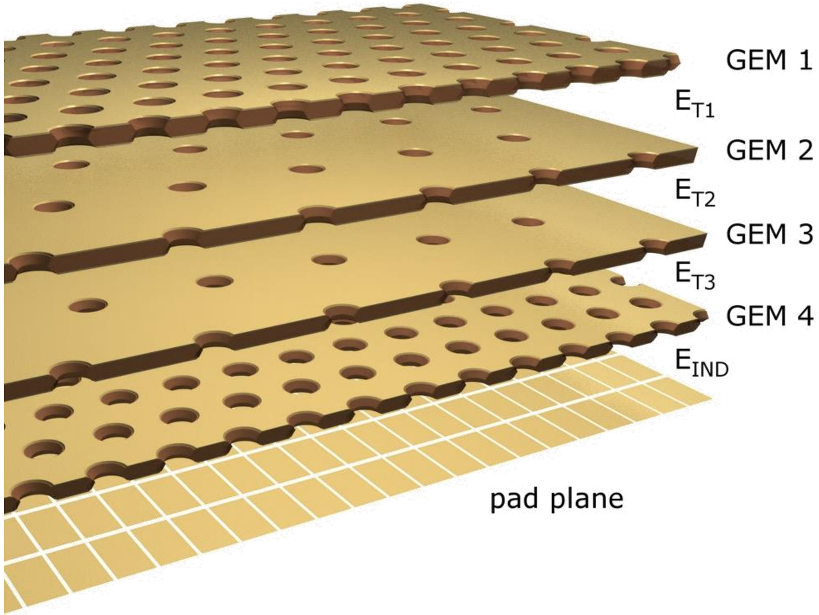




Simulated avalanche in a GEM hole

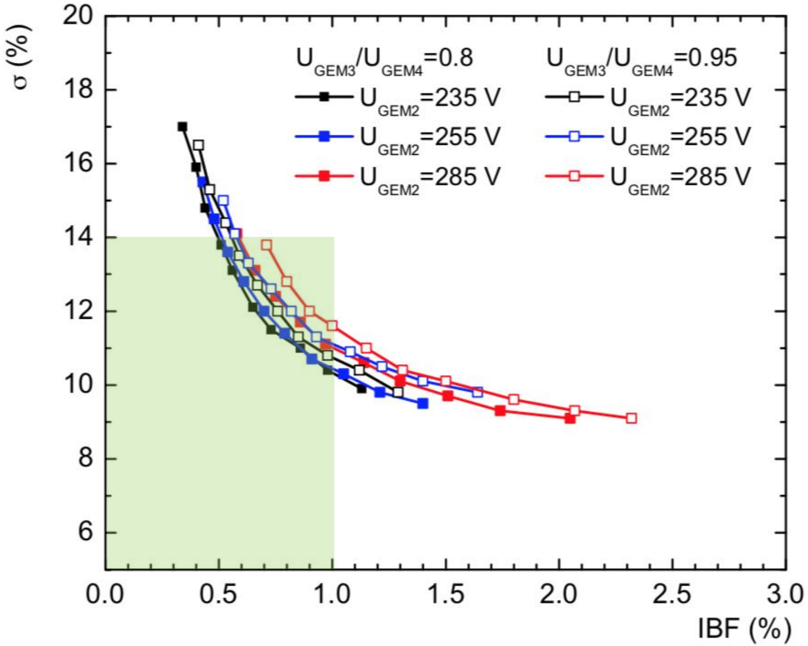
- **GEM = Gas Electron Multiplier**
- **Stacks of 4 GEM foils**
- **3 stacks for the large Outer ReadOut Chambers (OROC)**
- **1 stack for the smaller Inner ReadOut Chambers (IROC)**

4-GEM stacks



Schematic view of pad plane and 4-GEM stack
GEMs 1 and 4: Standard large-area single-mask GEM foils
GEMs 2 and 3: Large-pitch GEM foils

Highly optimized HV settings

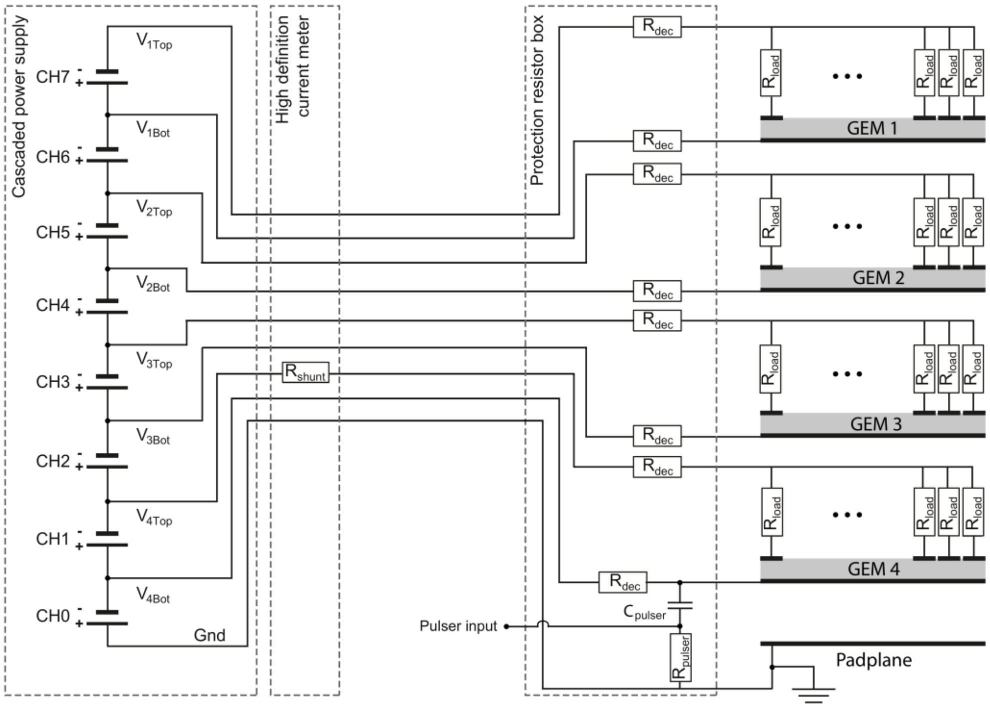


Performance with optimised HV configuration
IBF = Ion BackFlow
σ = energy resolution for ^{55}Fe

2019: GEM stacks mirrored on aluminised central drift electrode before installation of last IROC

HV system

CAEN A1515

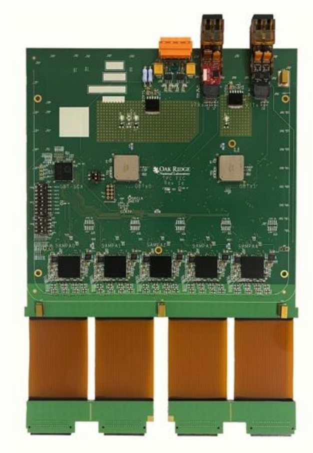


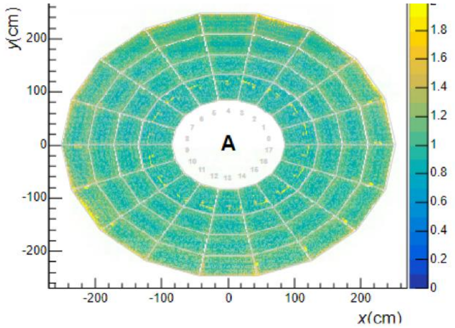
- Cascaded power supply units
- Designed for the operation of quadruple-GEM systems

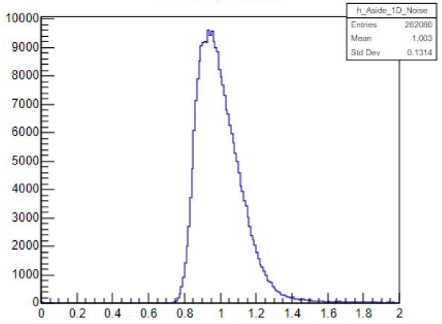
TPC readout electronics

- **SAMPA ASIC**
  - 130 nm TSMC CMOS
  - 32 channels with preamplifier, shaper, 10 bit ADC
  - Continuous (or triggered) readout

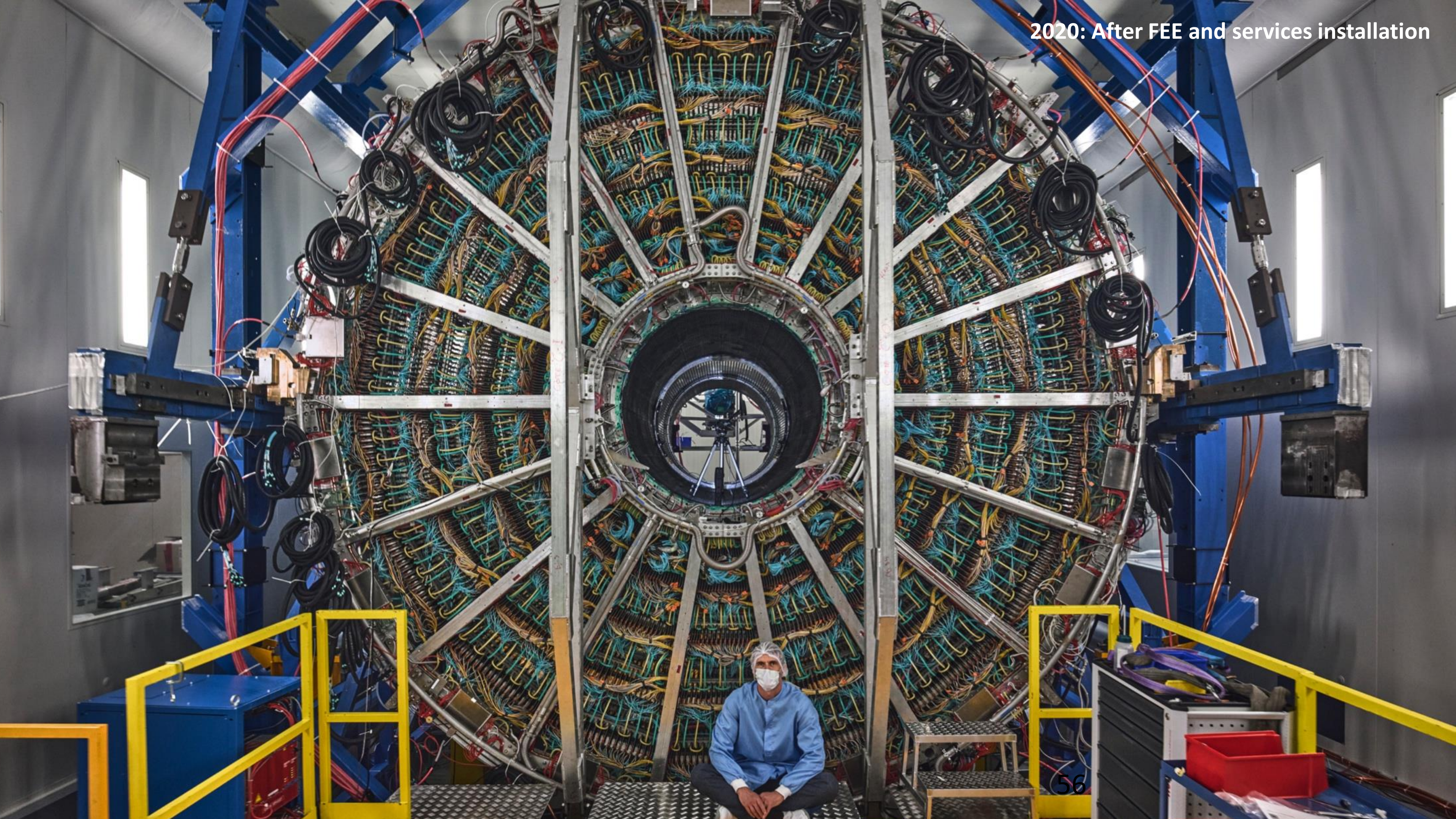
- **Front-End Cards (FECs)**
  - 5 SAMPA chips per FEC
  - Continuous sampling at 5 MHz
  - All ADC values read out
  - Readout link: CERN GBT / Versatile link system
  - 3276 FECs in total, 3.3 TB/s total data rate



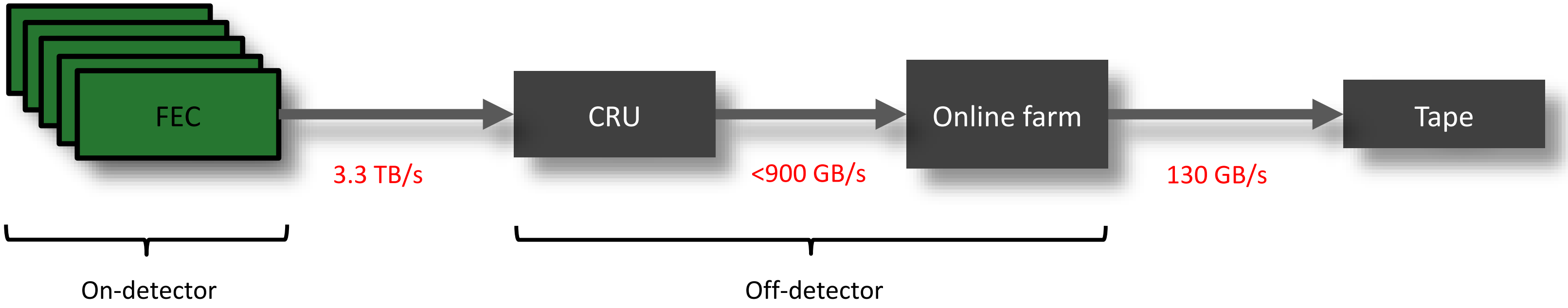




Excellent mean noise: 670 e^- @18 pF


2020: After FEE and services installation

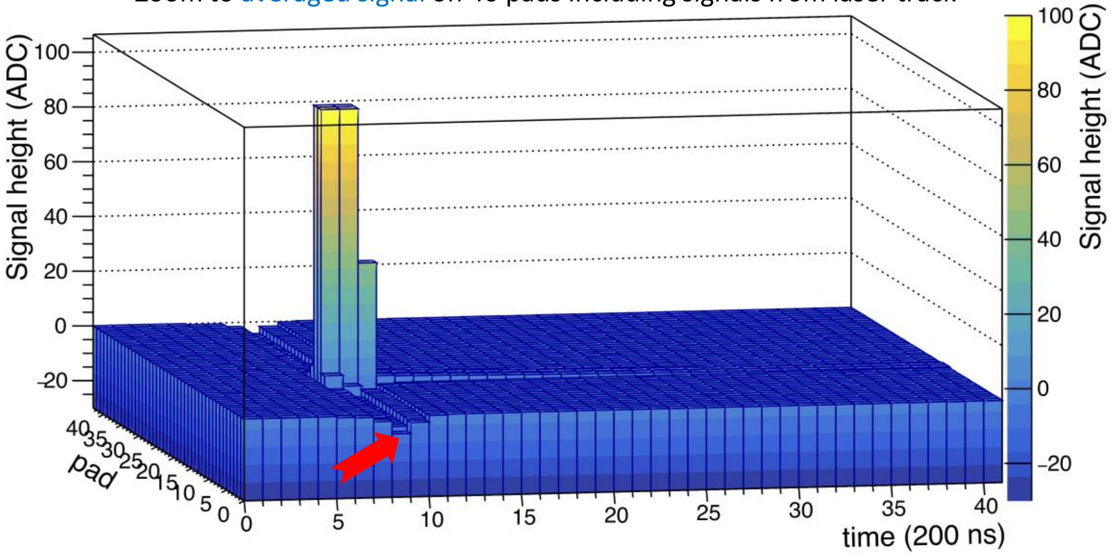
Data processing chain

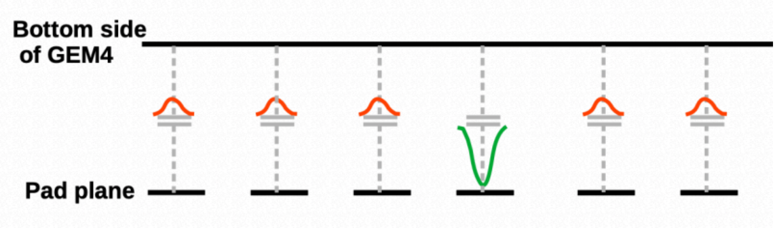


- **Continuous readout**
- **CRU (FPGA-based readout cards) installed in the FLPs**
  - **Receive the data through 6552 optical links**
  - **Data processing (see next 3 slides) and reduction (zero suppression)**
- **Further data reduction in online farm**

Common mode (CM) effect

Zoom to averaged signal on 40 pads including signals from laser track

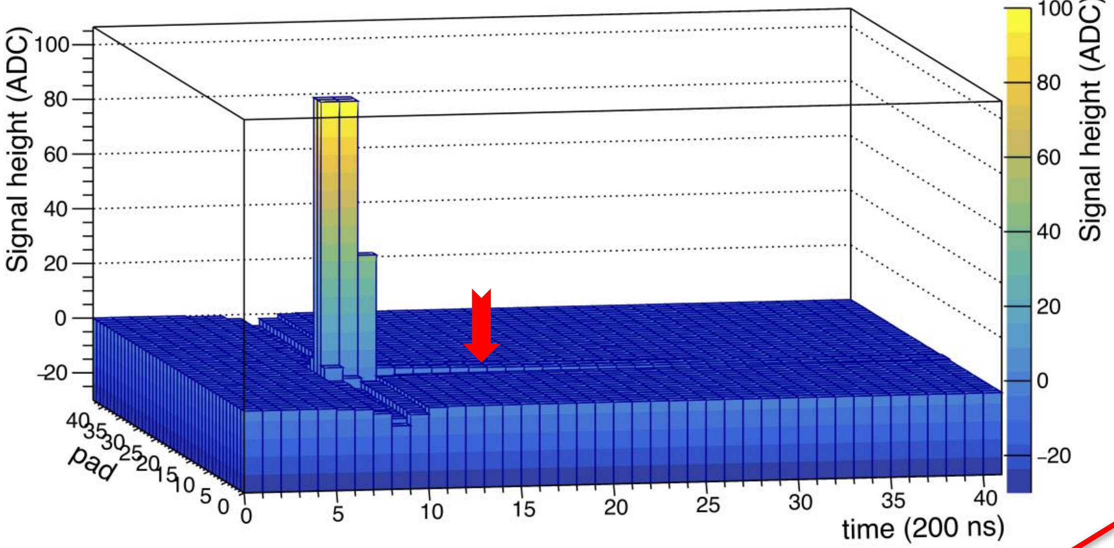


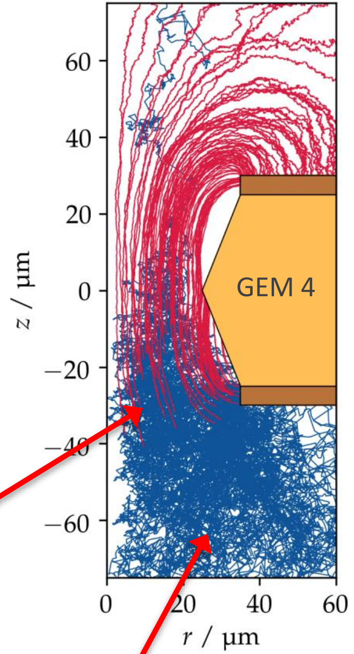


- Capacitors in HV distribution often used to reduce CM effect
- But such capacitors would lead to potential problems with discharges
- At high occupancy the CM signals from many tracks will superimpose and lead to a baseline shift

Ion tail!

Zoom to averaged signal on 40 pads including signals from laser track





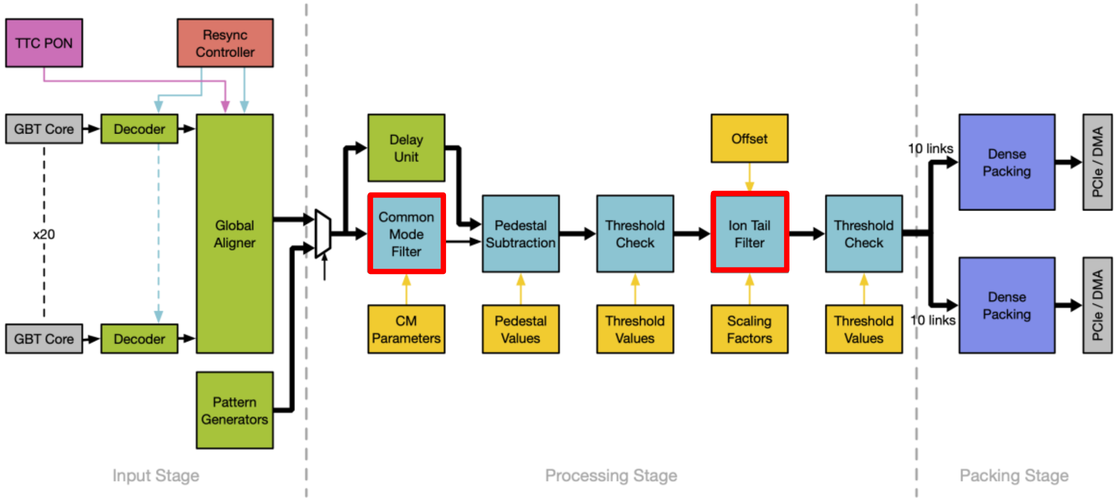
- An ion tail is visible. Two contributions:
  - Exponential contribution from ions created just below GEM 4
  - Linear component **due to marginal amplification in the induction gap (specific to our HV settings)**

Online data processing in FPGAs

- The (negative) baseline shift due to the dommon mode effect is measured and removed

- The ion tail is removed as well

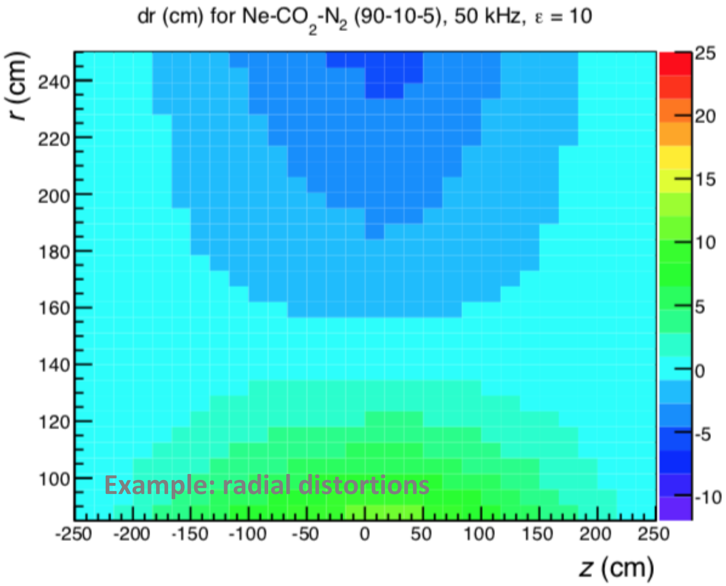
Building blocks of the data processing implemented in firmware in the CRU FPGAs



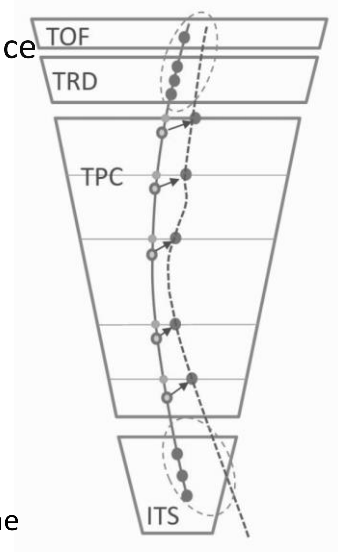
- Additional tasks: data alignment, pedestal subtraction, zero suppression, efficient packing of the data

Distortions

- With remaining ion back flow still considerable space charge distortions up to few cm



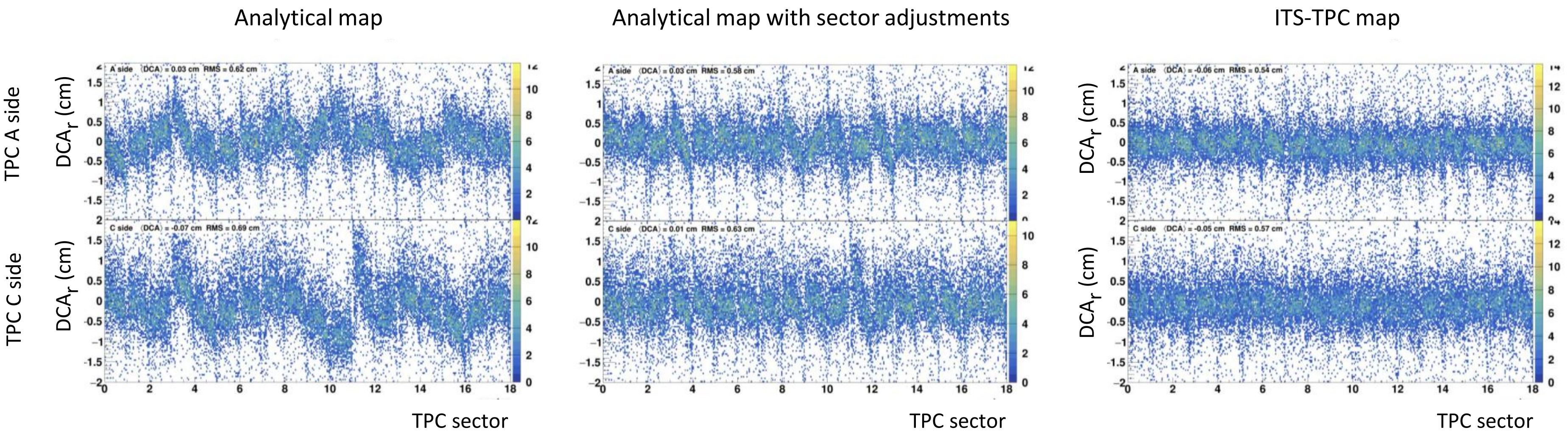
- Correction using track interpolation (experience from Runs 1 and 2)
- Calculate average distortion map which is slowly changing with collision rate



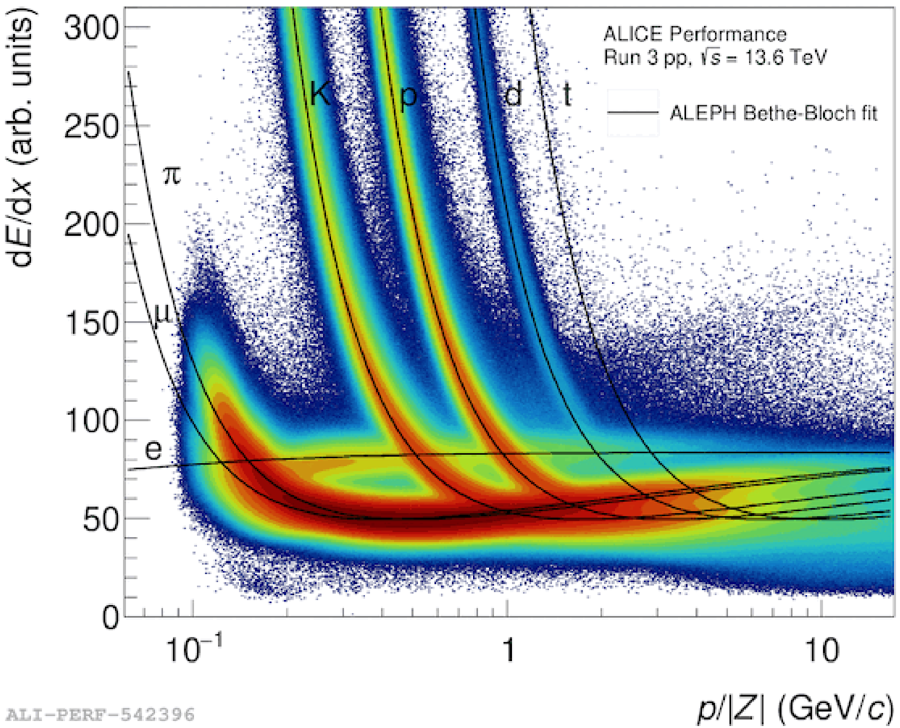
- In addition, fluctuations around the average distortions are important to reach intrinsic TPC resolution
- Fluctuations can be extracted by
  - integrating the ADC values over the ion drift time (Integrated Digital Currents) or by
  - measuring the analog currents at the GEM 4 top electrodes of all GEM stacks

Work in progress: Dist. calibrations

- Calibration of the space-charge distortions started using an analytical map
- With the envisaged interpolation method (here only ITS) the performance improves
- DCA_r = Distance of Closest Approach to beam line for reconstructed tracks
- Also time variations important



TPC performance



- TPC GEM upgrade successfully concluded. System behaving according to expectations
- The performance is according to expectations
- Remaining: Calibration of space-charge distortions **(electrostatic deflection of drifting electrons by ions present in the drift volume)**

Summary

- The construction, commissioning, installation, and start of operation of the upgraded ALICE detector and data processing chain were successfully achieved
- **ALICE employs GPUs heavily to speed up online and offline processing**
  - **99% of synchronous reconstruction on GPUs**
  - **60% of asynchronous processing on GPUs (2.5x speedup), will be ~80% in future (optimistic scenario)**
- ITS2 shows an excellent performance in RUN 3
  - **Based on ALPIDE pixel chip**
  - **Closer to beam pipe, reduced material budget, reduced pixel size**
- **The upgraded TPC is ready for Pb-Pb collisions in October!**
  - **Continuous readout**
  - **Ion backflow suppression built into GEM chambers**
  - **Extensive data processing and reduction in FPGAs**
  - **Distortions due to remaining space charge under control**
- We're looking forward to the starting of the Pb-Pb collisions

back up

Impact parameter