Experience with online processing of the ALICE TPC in LHC Run 3

David Rohr for the ALICE Collaboration, CERN

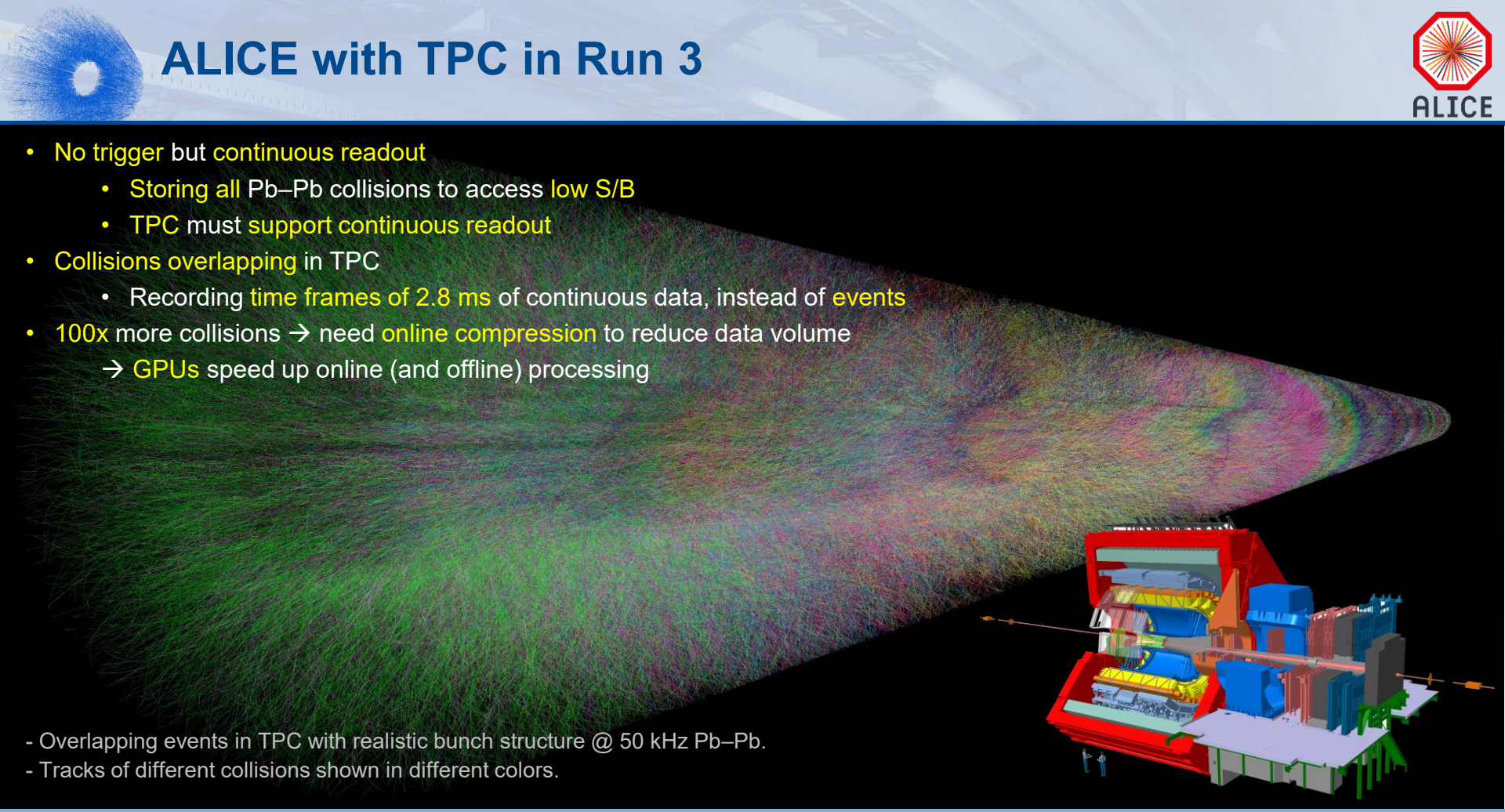
Workshop on the High Energy Circular Electron Positron Collider

26.10.2026

drohr@cern.ch

ALICE with TPC in Run 3

- No trigger but continuous readout
  - Storing all Pb–Pb collisions to access low S/B
  - TPC must support continuous readout
- Collisions overlapping in TPC
  - Recording time frames of 2.8 ms of continuous data, instead of events
- 100x more collisions → need online compression to reduce data volume
  → GPUs speed up online (and offline) processing

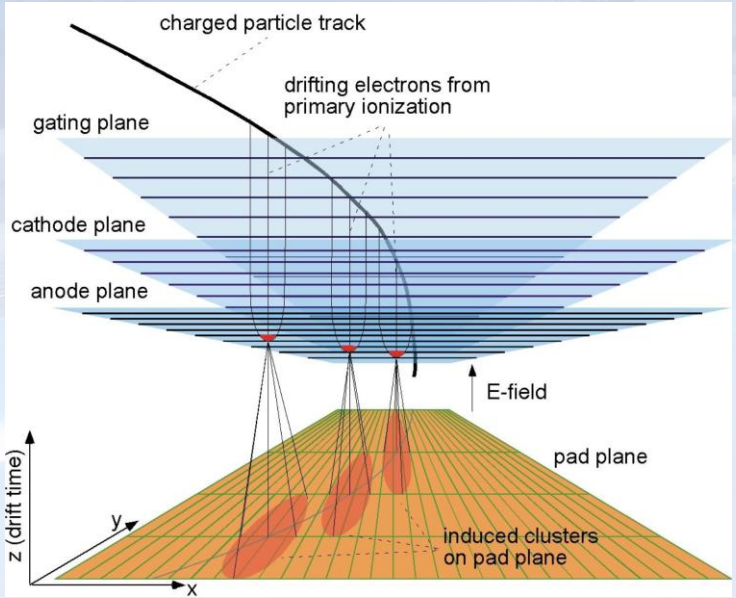


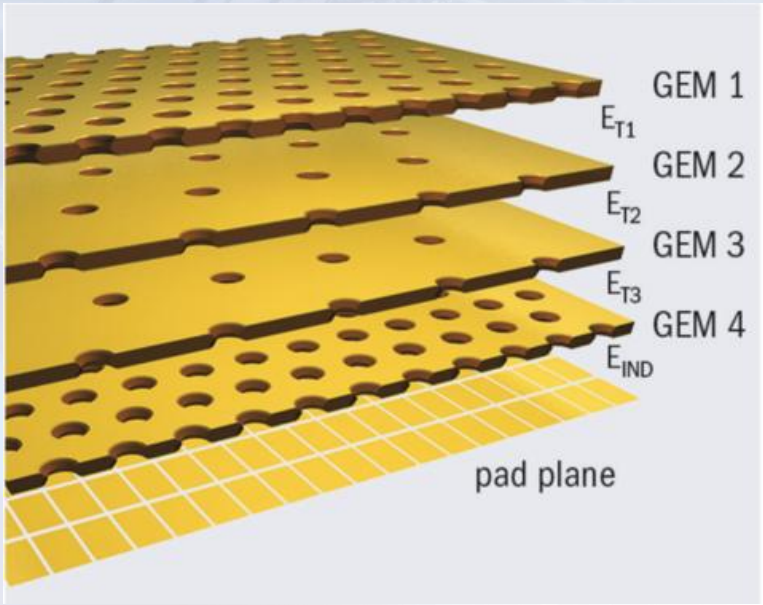
- Overlapping events in TPC with realistic bunch structure @ 50 kHz Pb–Pb.
- Tracks of different collisions shown in different colors.

ALICE TPC upgrades and implications

- **Need continuous TPC (Time Projection Chamber) readout to store full minimum bias sample.**
  - TPC of **Run 1 and 2** used **MWPC** (Multi Wire Proportional Chambers) readout and **gating grid** to **suppress ion back flow**.
  - Gating grid **limits readout to ~3 kHz**, prevents continuous readout.

  → **Replace MWPCs with GEMs** (Gas Electron Multiplier),
  Intrinsic ion back flow blocking (**99%**), no gating grid.



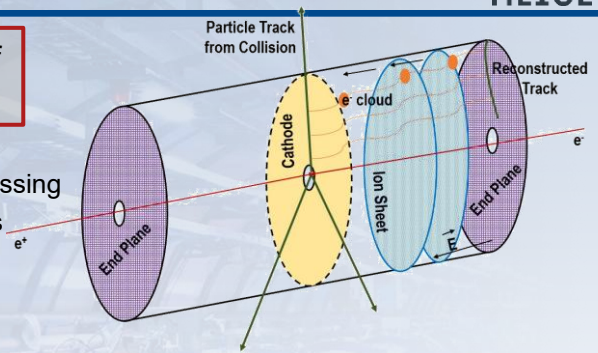


O^2 Processing steps

- **Online processing:**
  - 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

See details later!

Needs tracking of 1% of tracks

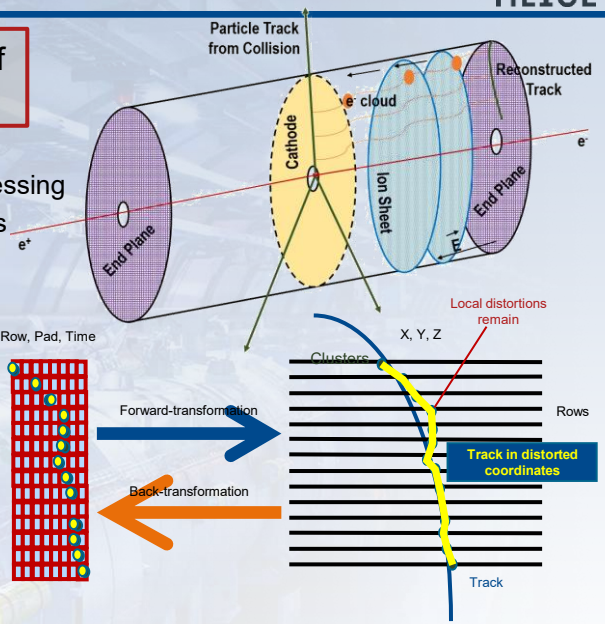


O^2 Processing steps

- **Online processing:**
  - 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

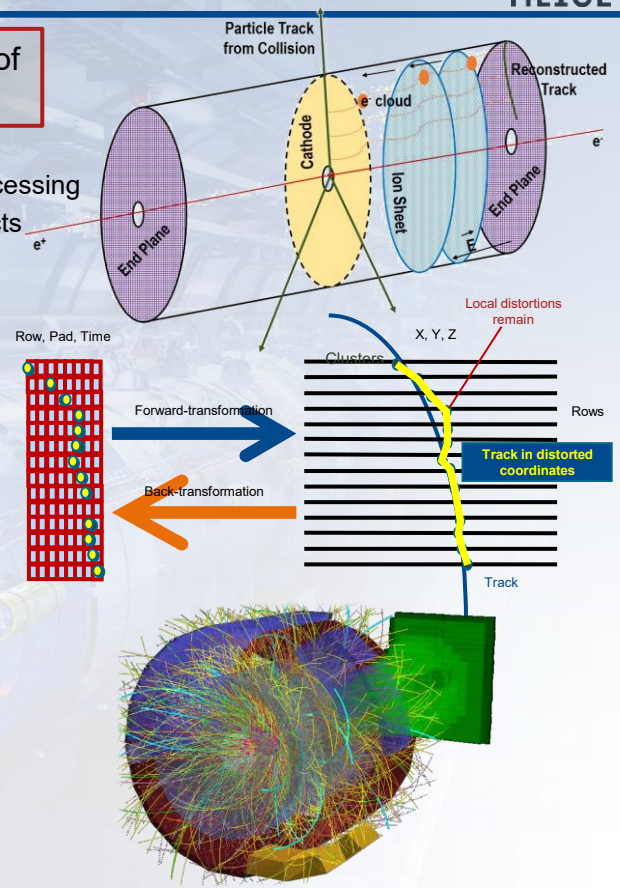


O² Processing steps

- **Online processing:**
  - 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



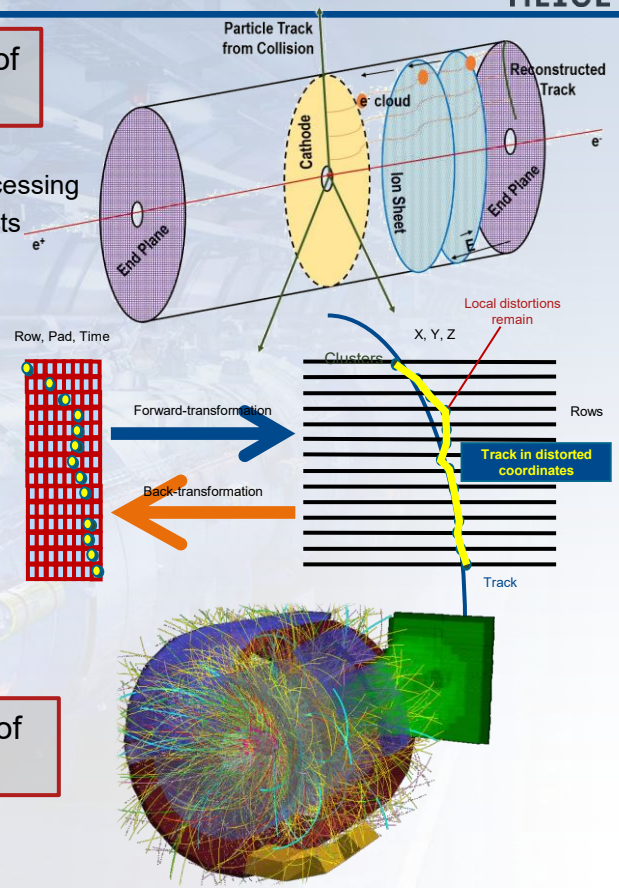
O^2 Processing steps

- **Online processing:**
  - 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%**)
- **Offline processing:**
  - **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

Needs tracking of 1% of tracks

Needs 100% TPC tracking

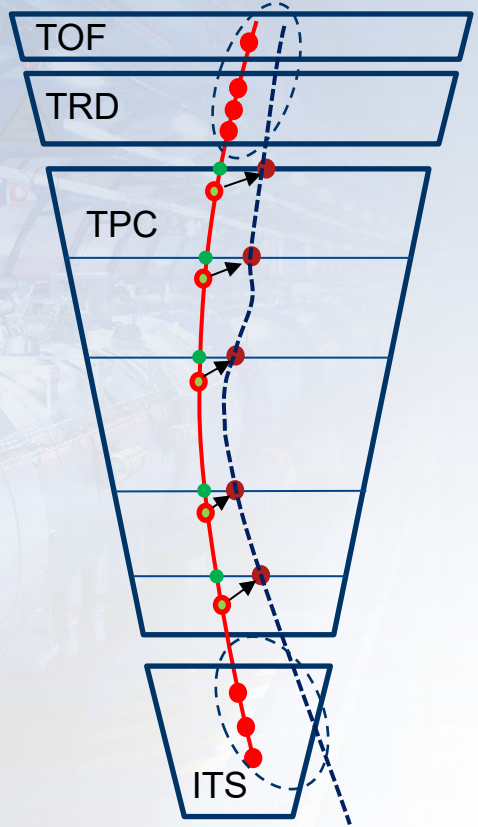
Needs 100% of everything



TPC Calibration

- **Most complicated TPC calibration is for space charge distortions (SCD).**
  - Other calibrations (like d*E*/d*x* run in parallel).
  - General idea: **Track based**:
    - TPC Tracks reconstructed with relaxed cuts, matched to inner / outer detectors.
    - Track refit with only ITS / TRD / TOF information.
    - Collect residuals of TPC hits wrt. refitted tracks.
    - TPC volume voxelized, correction per voxel calculated.
    - Corrections smoothed compensating for bad TRD chambers, holes in acceptance, …
  - This approach corrects for the average distortions
  - Yields one correction map for all corrections (including exb, etc.)
  - In principle, will need multiple maps eventually for fluctuations of the distortions
  - Using inner + outer detectors for interpolatiomn is ideal. But in principal, extrapolation from inner detector is enough.

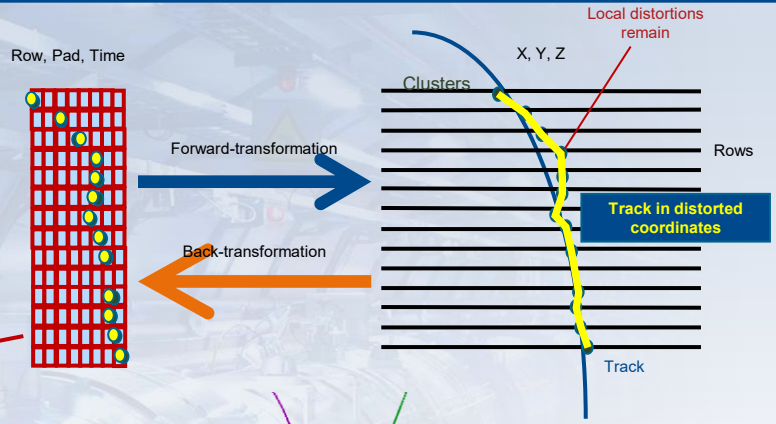
  - For details, see LHCP talk by Matthias Kleiner!

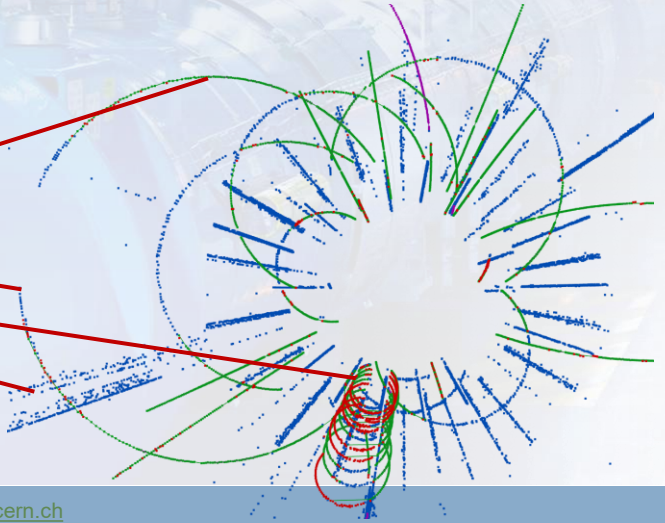


TPC Data Compression

- **TPC is the largest data-contributor**
  - Reduced from 3.4 TB/s raw data to ~170 GB/s to storage.
  - At the 170 GB/s, other detectors contribute as well.
  - → All detectors entropy-compressed, sophisticated solution for TPC.

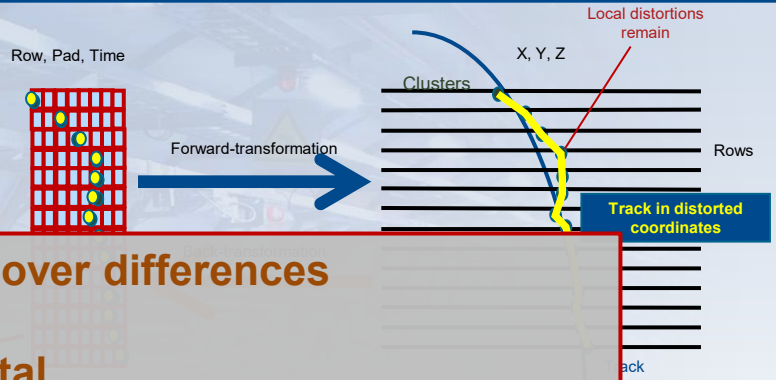
- **TPC data compression is composed of several steps:**
  - Raw data clusterized.
  - Custom data format (int / float) number of bits matching the TPC resolution.
  - Entropy reduction:
    - Clusters attached to tracks stored as track residuals.
    - Unattached clusters sorted, stored as difference wrt. previous cluster.
    - Correlated values treated together.
  - Clusters of tracks not used for physics removed:
    - Strategy A: Identification and removal of clusters:
      - Secondary leg of track with p_T < 200 MeV/c.
      - Track segment with high inclination angle.
      - Low-p_T track below 50 MeV/c.
      - Noisy pads, charge clouds from low-p_T protons.
    - Strategy B:
      - Remove everything except for identified good tracks.
  - Remaining clusters entropy-compressed with ANS encoding.





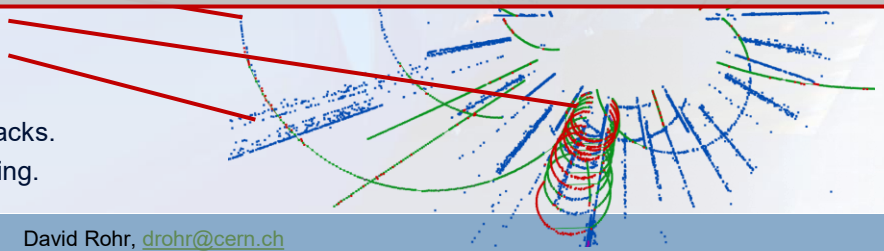
TPC Data Compression

- **TPC is the largest data-contributor**
  - Reduced from 3.4 TB/s raw data to ~70 GB/s to storage.
  - At the 70 GB/s, other detectors contribute as well.
  - → All detectors entropy-compressed, sophisticated solution for TPC.

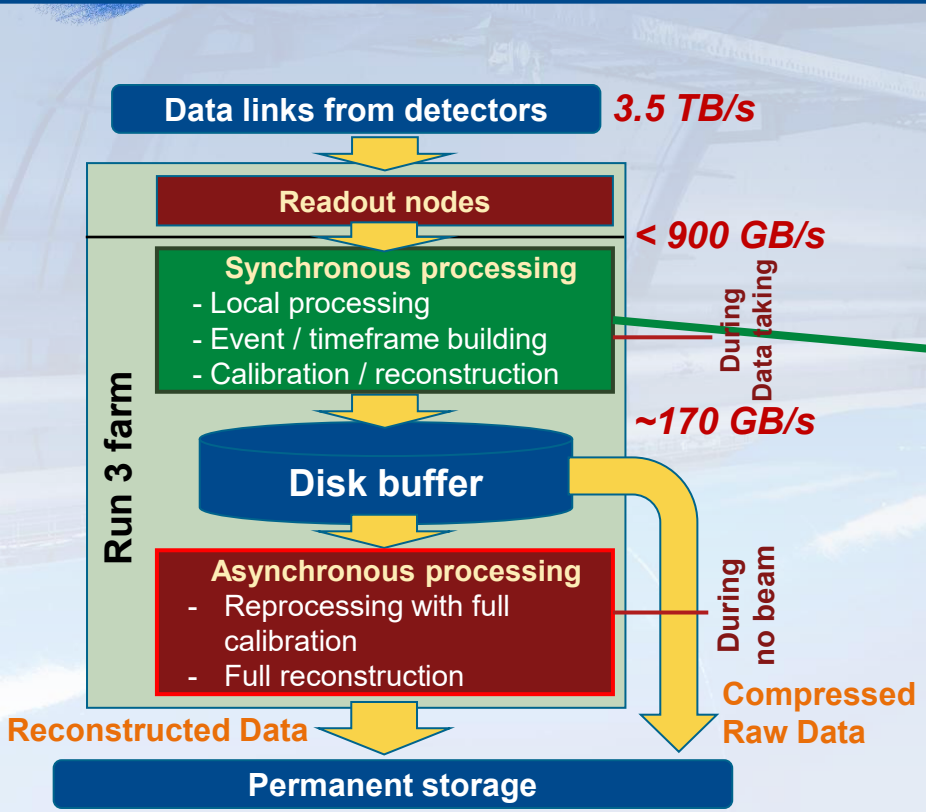


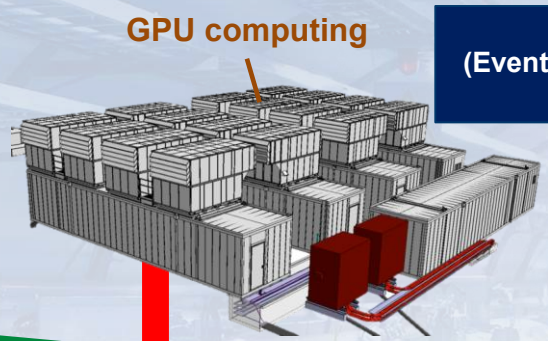
- **TPC data c**
  - Raw data c
  - Custom dat
  - Entropy red
    - Cluste
    - Unatta
    - Correla
  - Clusters of
    - Strateg
      - S
      - T
      - Low-p_T track below 50 MeV/c.
      - Noisy pads, charge clouds from low-p_T protons.
    - Strategy B:
      - Remove everything except for identified good tracks.
  - Remaining clusters entropy-compressed with ANS encoding.

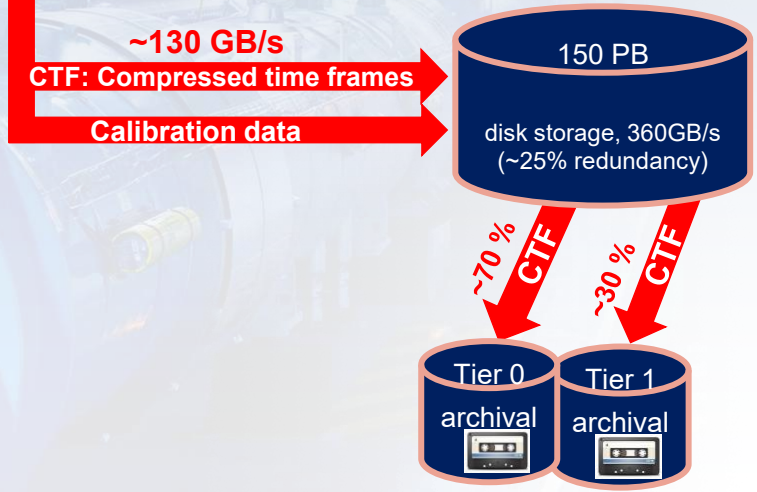
- **Advantage of Track Model Compression over differences compression is higher for sparse data**
  - **At 50 kHz Pb-Pb, it is only ~5% in total.**
- **Strategy A rejection in ALICE is around 15%, Strategy B could reach up to 35%, but has not yet been used.**
- **Advantage of AND entropy compression of Huffman is ~5%**
  - **Using modern tailored compression, going beyond pure entropy compression, can yield up to 15% improvement (WIP).**



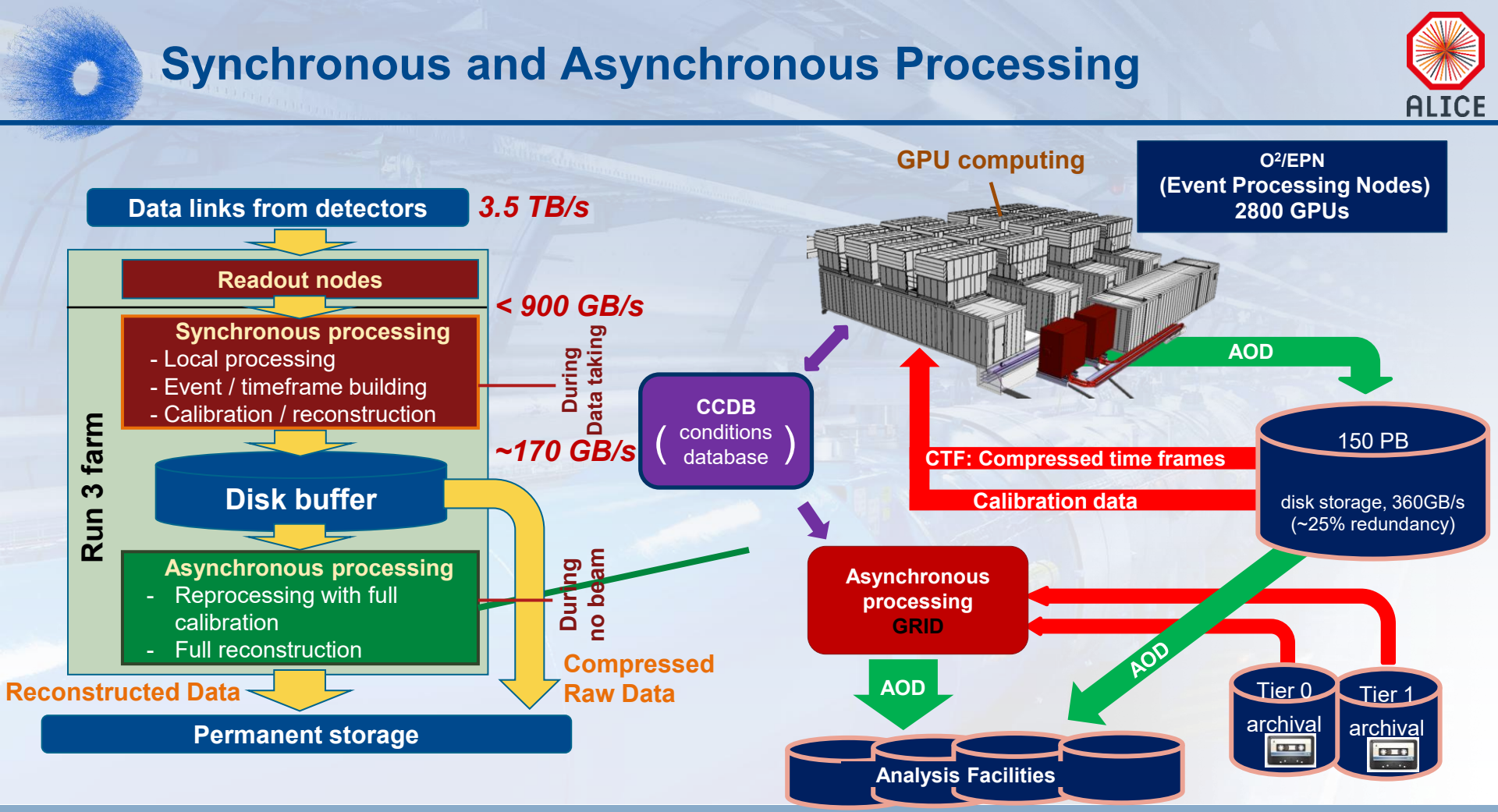
Synchronous and Asynchronous Processing



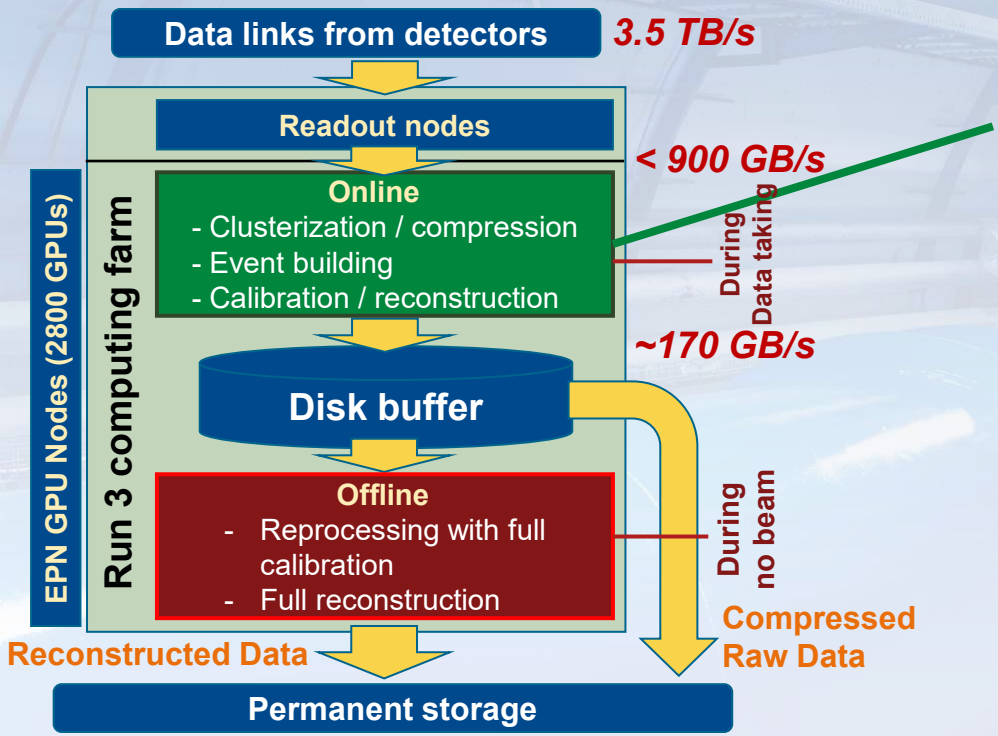




Synchronous and Asynchronous Processing



ALICE Run 3 Online and Offline Processing Scheme



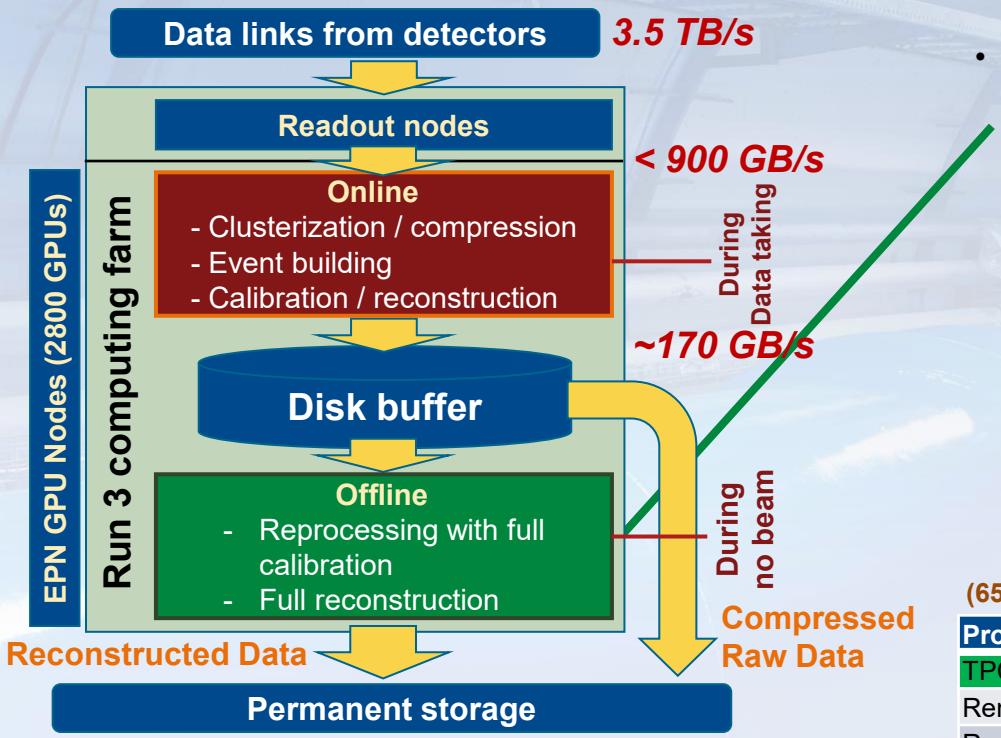
- **Online processing totally dominated by the TPC**
  - Full TPC processing runs on GPU
  - 1 AMD MI50 GPU replaces ~80 CPU cores (AMD Epyc 7452)
  - Data transfer, postprocessing, etc. fully pipelined
    → 100% GPU utilization
  - Full GPU speedup available in online processing
  - In operation since beginning of Run 3

Online reconstruction (50 kHz Pb–Pb, MC data, no QA / calib)

	% of time
TPC Processing (Tracking, Clustering, Compression)	99.37 %
Rest	0.63 %

Relative CPU time (linux cputime) with full processing on CPU

ALICE Run 3 Online and Offline Processing Scheme



- **Offline processing much more heterogeneous**
  - **TPC** significant (**50%** - **60%**), but not fully dominant
- **Should use GPUs also when no beam in LHC for offline**

Online reconstruction (50 kHz Pb–Pb, MC data, no QA / calib)

	% of time
TPC Processing (Tracking, Clustering, Compression)	99.37 %
Rest	0.63 %

Relative CPU time (linux cputime) with full processing on CPU

Offline processing (650 kHz pp, 2022, no Calorimeters)

Processing step	% of time
TPC Processing (Tracking)	61.41 %
Remaining Barrel Tracking	18.36 %
Rest	20.23 %

Offline processing (47 kHz Pb–Pb, 2024)

Processing step	% of time
TPC Processing (Tracking)	52.39 %
Remaining Barrel Tracking	31.50 %
Rest	16.11 %

ALICE GPU Offline Processing on EPNs since 2023

- **90% of EPN compute power in GPUs.** (depends on workload, 90% measured in TPC online processing)
- **EPN limited to ~10% of performance with CPU-only code, up to ~10x speedup possible with GPU.**
- **Since 2023, TPC offline reconstruction on GPU:**
  - **Offloads 50% - 60%**, still **CPU-bound**, high **GPU idle time** (**> 66%**).

  → **60%** of **compute load removed from CPU**, **40% remain**, possible **speedup 2.5x**.

Online reconstruction (50 kHz Pb–Pb, MC data, no QA / calib)

	% of time
TPC Processing (Tracking, Clustering, Compression)	99.37 %
Rest	0.63 %

Relative CPU time (linux cputime) with full processing on CPU

Offline processing (650 kHz pp, 2022, no Calorimeters)

Processing step	% of time
TPC Processing (Tracking)	61.41 %
Remaining Barrel Tracking	18.36 %
Rest	20.23 %

Offline processing (47 kHz Pb–Pb, 2024)

Processing step	% of time
TPC Processing (Tracking)	52.39 %
Remaining Barrel Tracking	31.50 %
Rest	16.11 %

ALICE GPU Offline Processing on EPNs since 2023

- **90% of EPN compute power in GPUs.** (depends on workload, 90% measured in TPC online processing)
- **EPN limited to ~10% of performance with CPU-only code, up to ~10x speedup possible with GPU.**
- **Since 2023, TPC offline reconstruction on GPU:**
  - **Offloads 50% - 60%**, still **CPU-bound**, high **GPU idle time** (**> 66%**).
  - → **60%** of **compute load removed from CPU**, **40% remain**, possible **speedup 2.5x**.

- **Can split the EPN server in different ways**
  - Allows for different number of parallel offline jobs
  - Tuned for the most efficient configurations (see later!) (separately for CPU and GPU)

Configuration (2022 pp, 650 kHz)	Time per TF (11ms TF, 1 instance)	Time per TF (11ms TF, 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**

Achieved exactly expected speedup
Factor 2.51, matches precisely offloading 60%

Online reconstruction (50 kHz Pb–Pb, MC data, no QA / calib)

	% of time
TPC Processing (Tracking, Clustering, Compression)	99.37 %
Rest	0.63 %

Relative CPU time (linux cputime) with full processing on CPU

Offline processing (650 kHz pp, 2022, no Calorimeters)

Processing step	% of time
TPC Processing (Tracking)	61.41 %
Remaining Barrel Tracking	18.36 %
Rest	20.23 %

Offline processing (47 kHz Pb–Pb, 2024)

Processing step	% of time
TPC Processing (Tracking)	52.39 %
Remaining Barrel Tracking	31.50 %
Rest	16.11 %

ALICE GPU Offline Processing on EPNs since 2023

- **90% of EPN compute power in GPUs.** (depends on workload, 90% measured in TPC online processing)
- **EPN limited to ~10% of performance with CPU-only code, up to ~10x speedup possible with GPU.**
- **Since 2023, TPC offline reconstruction on GPU:**
  - **Offloads 50% - 60%**, still **CPU-bound**, high **GPU idle time** (**> 66%**).
  - → **60%** of compute load removed from CPU, 40% remain, possible speedup 2.5x
- **Can split th**
  - Allows fo
  - Tuned fo
    (separate

- **Important difference to online**:
  - **Throughput** is the important metric
  - Pure GPU "**speedup**" **irrelevant** without 100% GPU utilization
  - If **GPU** has **idle** time → **CPU-bound**
  - Only way to speed up: **Offload more tasks to GPU**
- **Online**: **CPU must be fast enough to fully load the GPU**
  - **Incoming time frame rate defined from outside**
- **Offline**: **Aiming for constant 100% CPU utilization**
  - **Process time frames as fast as necessary to reach 100% load**
  - **But keep memory constraints in mind!**

Configuration (2022 pp, 650 kHz)
CPU 8 core
CPU 16 core
1 GPU + 16 CPU co
1 NUMA domain (4 GPUs + 64 cores)

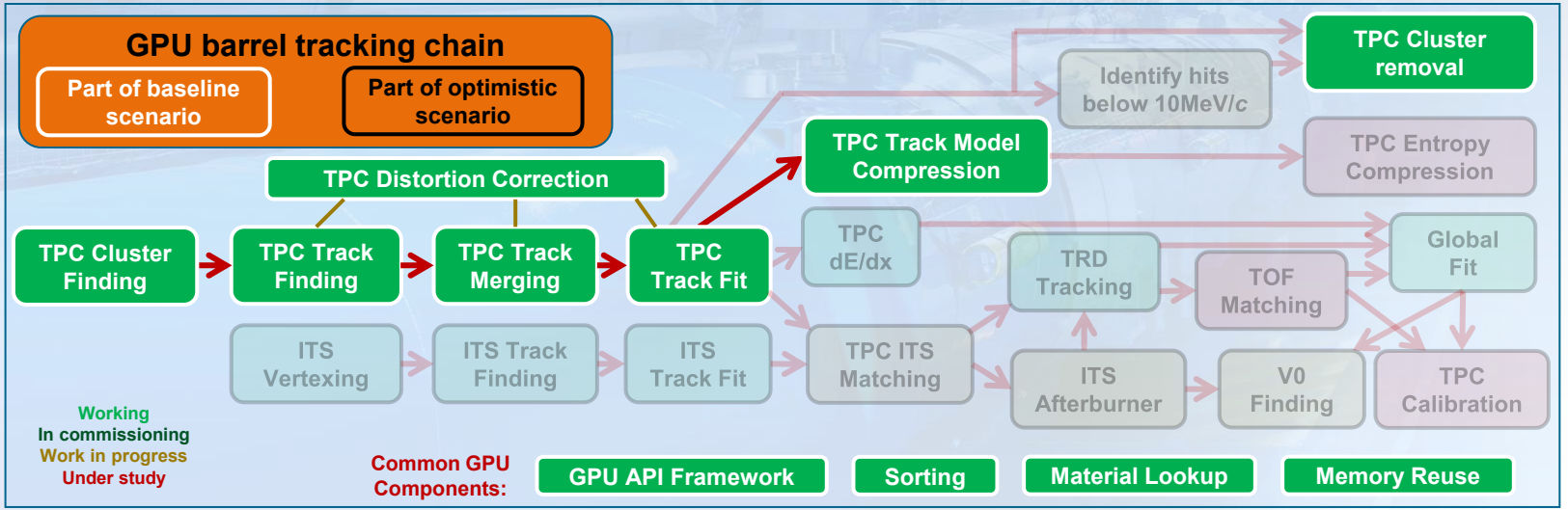
Achieved exactly expected speedup
Factor 2.51, matches precisely offloading 60%

Processing step	% of time
Remaining Barrel Tracking	18.36 %
Rest	20.23 %

Processing step	% of time
	52.39 %
Remaining Barrel Tracking	31.50 %
Rest	16.11 %

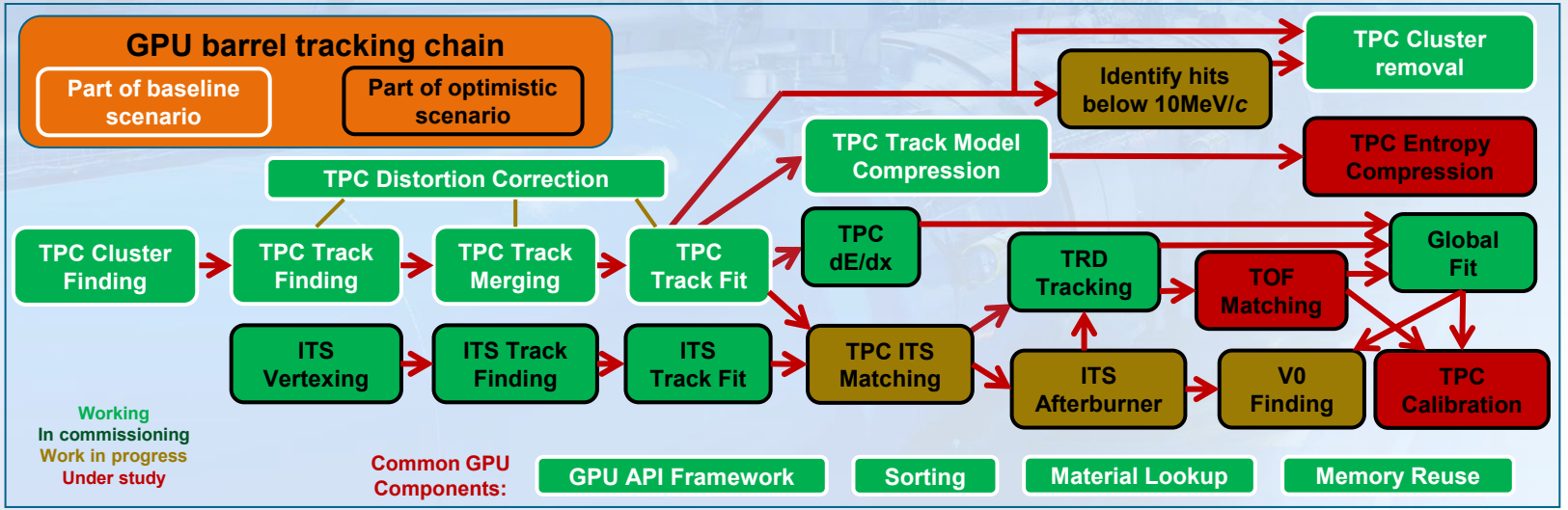
Improving offline GPU usage – Optimistic Scenario

- **Work on moving the full barrel tracking to GPU is ongoing:**
  - **Baseline scenario** with only **TPC tracking** in operation since **2023**: **2x - 2.5x** throughout increase on EPNs (white labels).



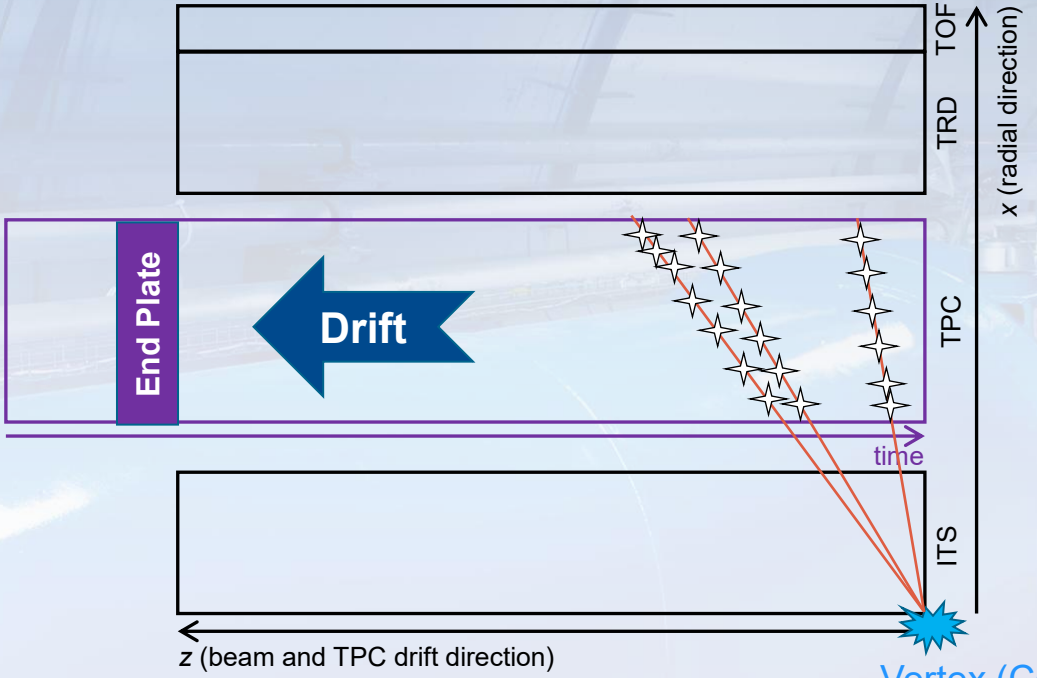
Improving offline GPU usage – Optimistic Scenario

- **Work on moving the full barrel tracking to GPU is ongoing:**
  - **Baseline scenario** with only **TPC tracking** in operation since **2023**: **2x - 2.5x** throughout increase on EPNs (white labels).
  - **TPC track model decoding** moved to **GPU** in **2024** for **~3%** speedup (LHCP 2024 talk).
  - **ITS tracking** on **GPU** commissioned end of **2025** for **26%** speedup (CHEP 2026 talk).
  - **Next important milestone**: **TPC ITS matching** as central task, will allow full GPU processing without data transfer.

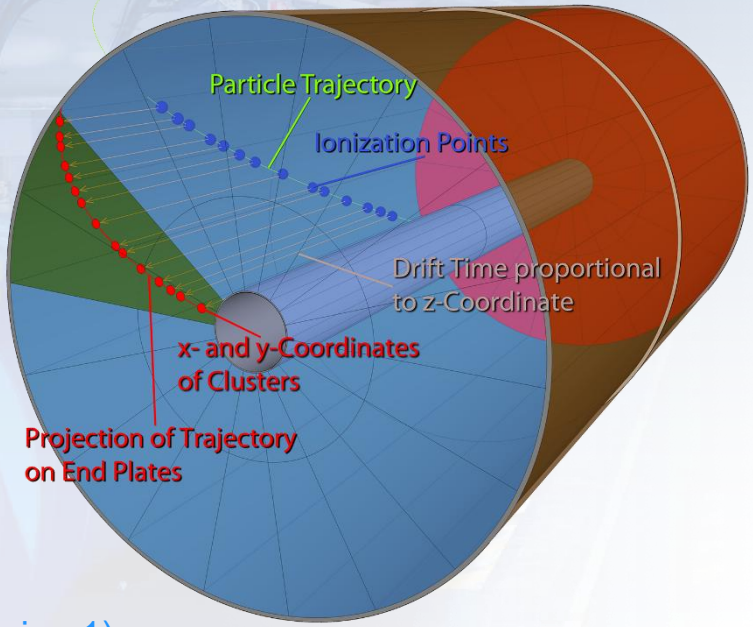


The tracking challenge

- **Tracking continuous data…**
  - The TPC sees **multiple overlapped collisions** (shifted in time).
  - Other detectors know the (rough) time of the collision.

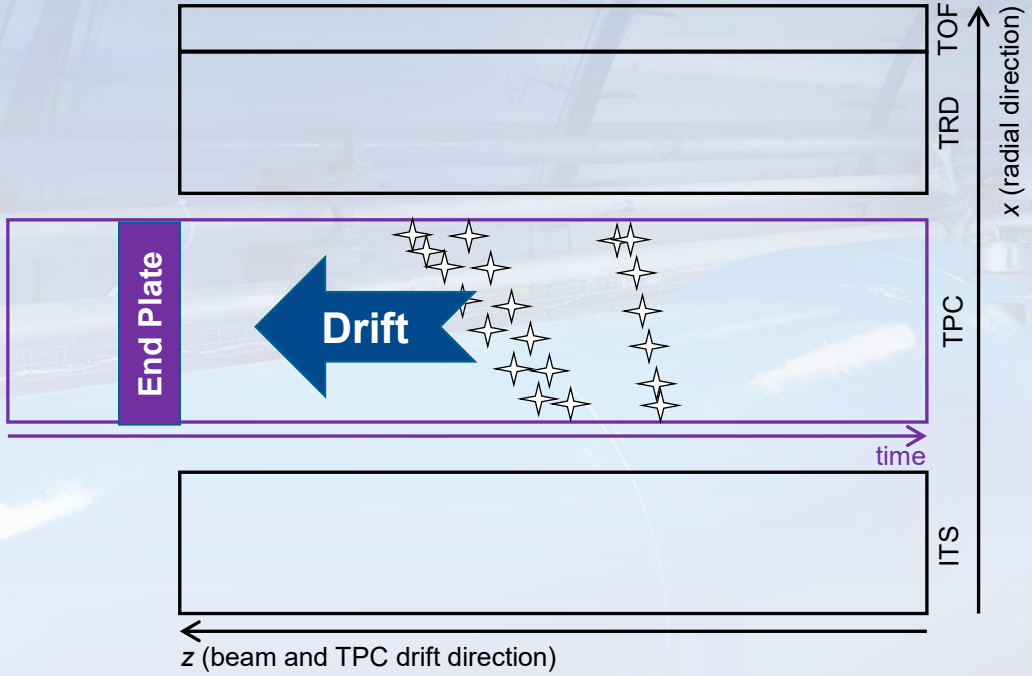


- **Problem: TPC clusters have no defined *z*-position but only a time. They can be shifted in *z* arbitrarily.**

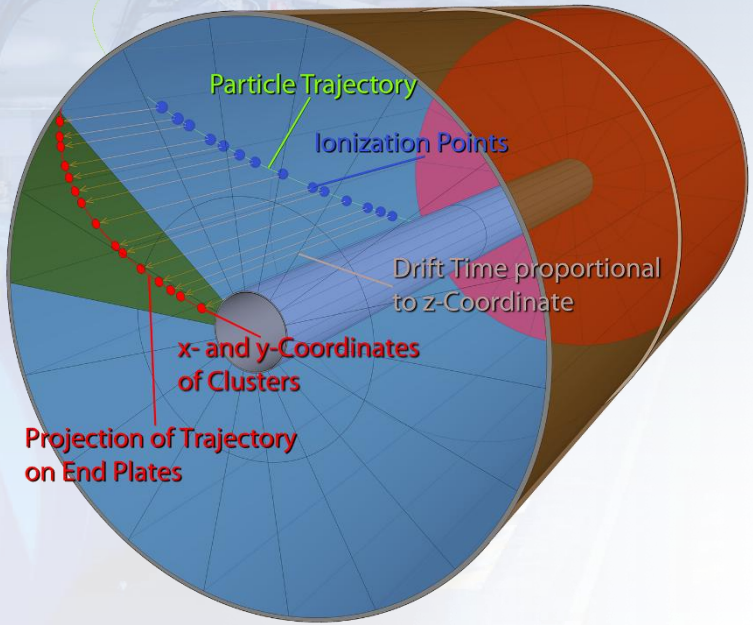


The tracking challenge

- **Tracking continuous data…**
  - The TPC sees **multiple overlapped collisions** (shifted in time).
  - Other detectors know the (rough) time of the collision.

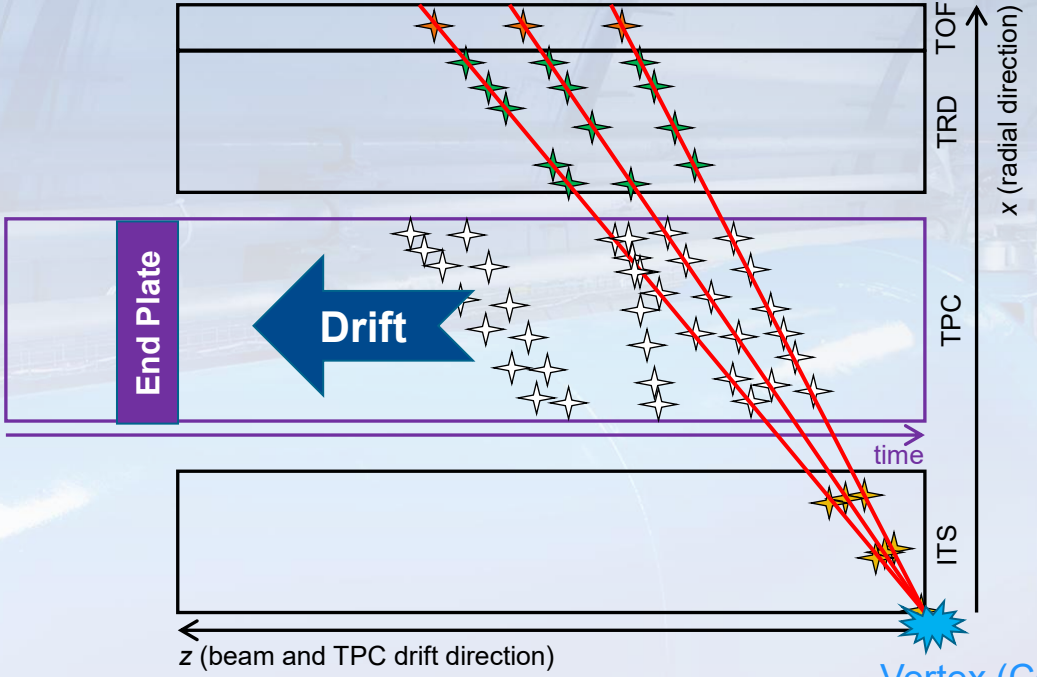


- **Problem: TPC clusters have no defined *z*-position but only a time. They can be shifted in *z* arbitrarily.**

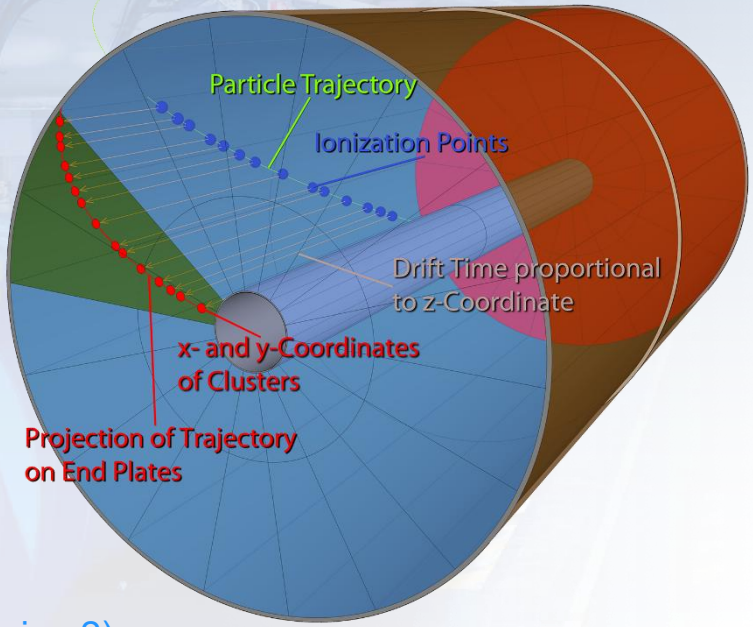


The tracking challenge

- **Tracking continuous data…**
  - The TPC sees **multiple overlapped collisions** (shifted in time).
  - Other detectors know the (rough) time of the collision.



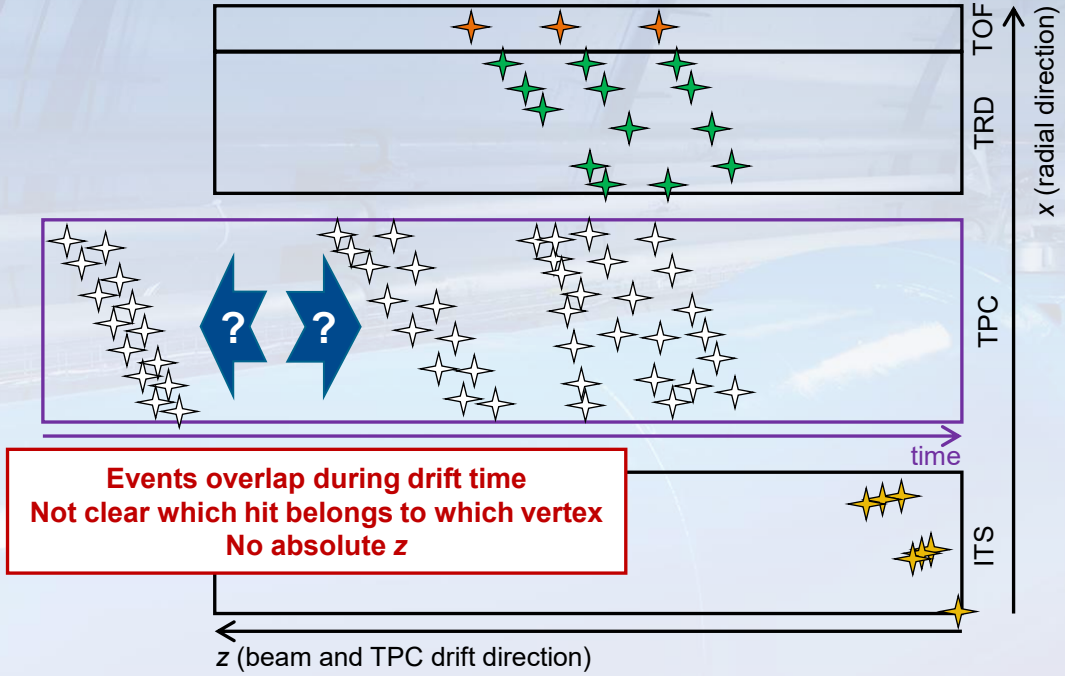
- **Problem: TPC clusters have no defined *z*-position but only a time. They can be shifted in *z* arbitrarily.**



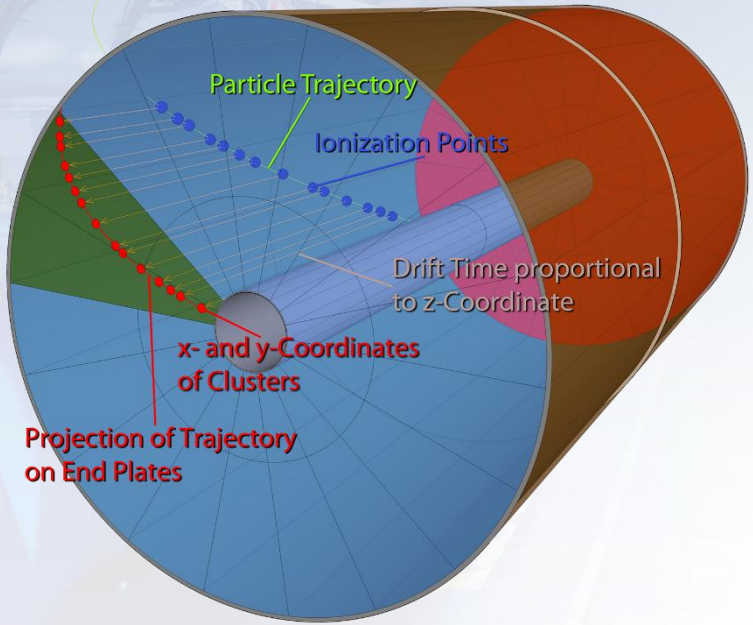
Vertex (Collision 2)

The tracking challenge

- **There are 2 main challenges caused by continuous readout / space charge distortions**
  - **Cluster overlap**?
  - How to **assign the z-coordinate** and **apply cluster corrections**.

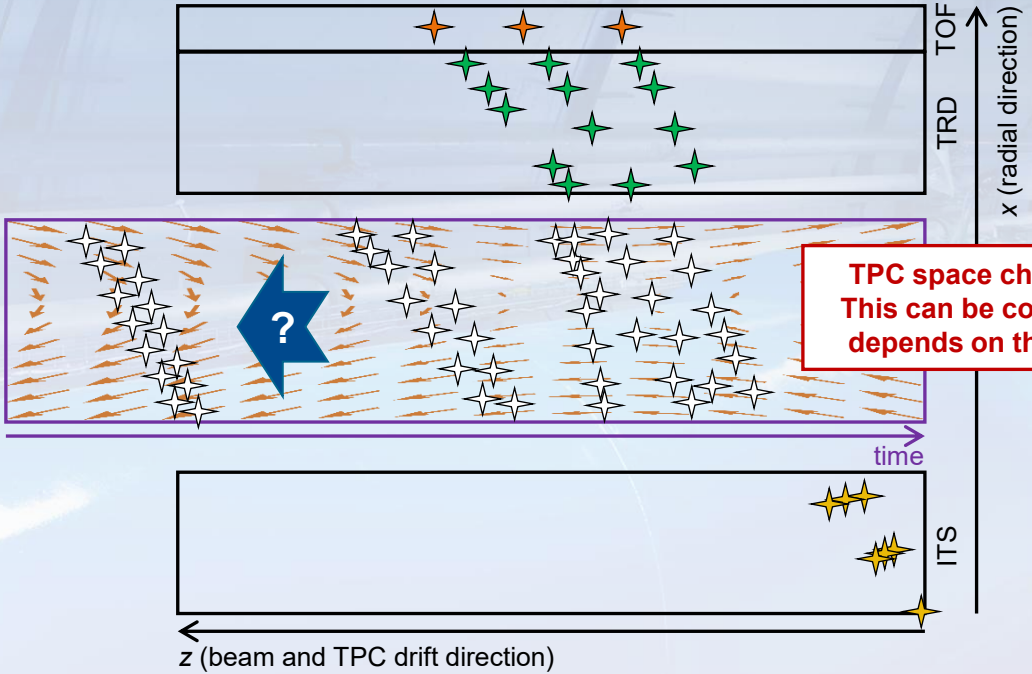


Events overlap during drift time
Not clear which hit belongs to which vertex
No absolute *z*



The tracking challenge

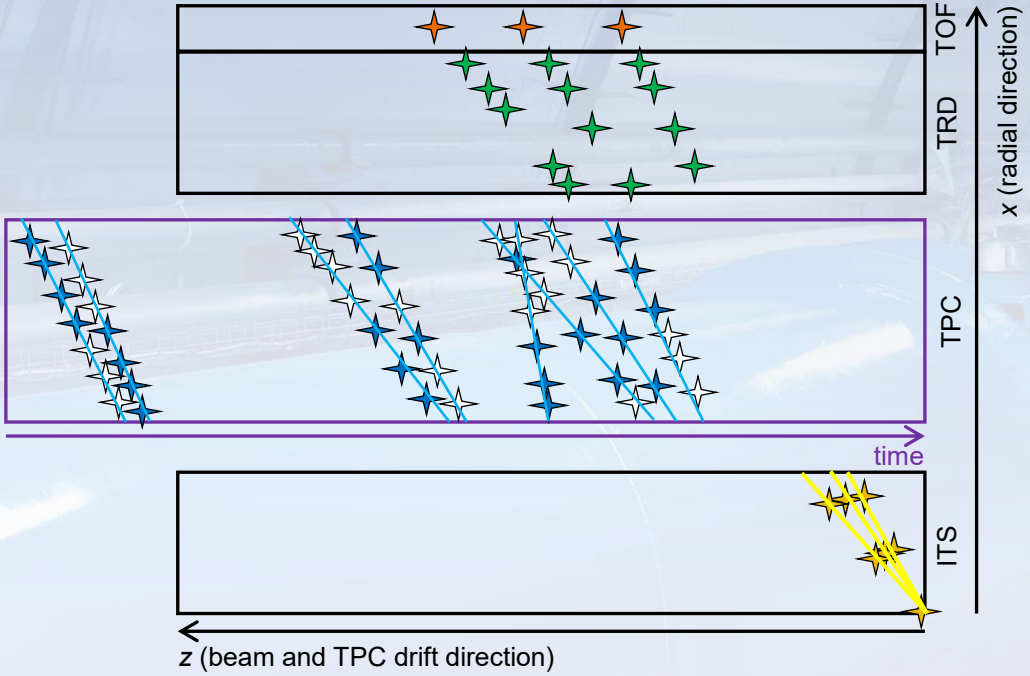
- **There are 2 main challenges caused by continuous readout / space charge distortions**
  - **Cluster overlap**?
  - How to **assign the z-coordinate** and **apply cluster corrections**.



TPC space charge distorts the clusters. This can be corrected, but the correction depends on the *z* position of the cluster

The tracking challenge – standalone detector tracking

- **There are 2 main challenges caused by continuous readout / space charge distortions**
  - **Cluster overlap**?
  - How to **assign the z-coordinate** and **apply cluster corrections**.



- Standalone ITS tracking.
- Standalone TPC tracking, scaling *t* linearly to an arbitrary *z*.

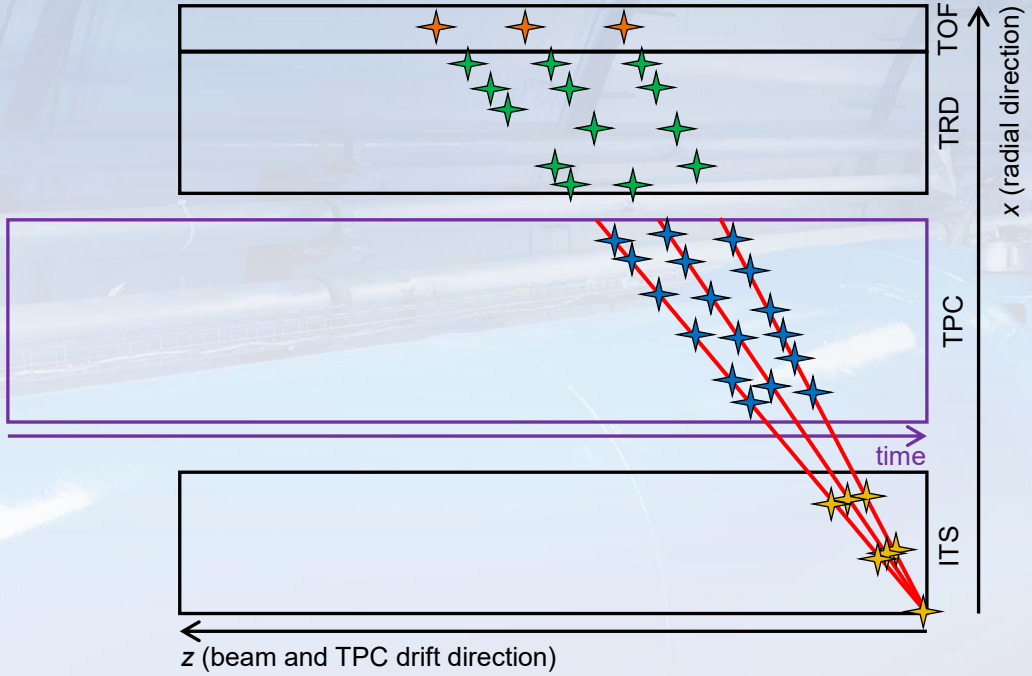
Precise tracking needs *z* for:

- Cluster error parameterization
- Inhomogeneous B-field
- Distortion correction

Effects smooth →
irrelevant for initial trackletting

The tracking challenge – track merging

- **There are 2 main challenges caused by continuous readout / space charge distortions**
  - **Cluster overlap**?
  - How to **assign the z-coordinate** and **apply cluster corrections**.



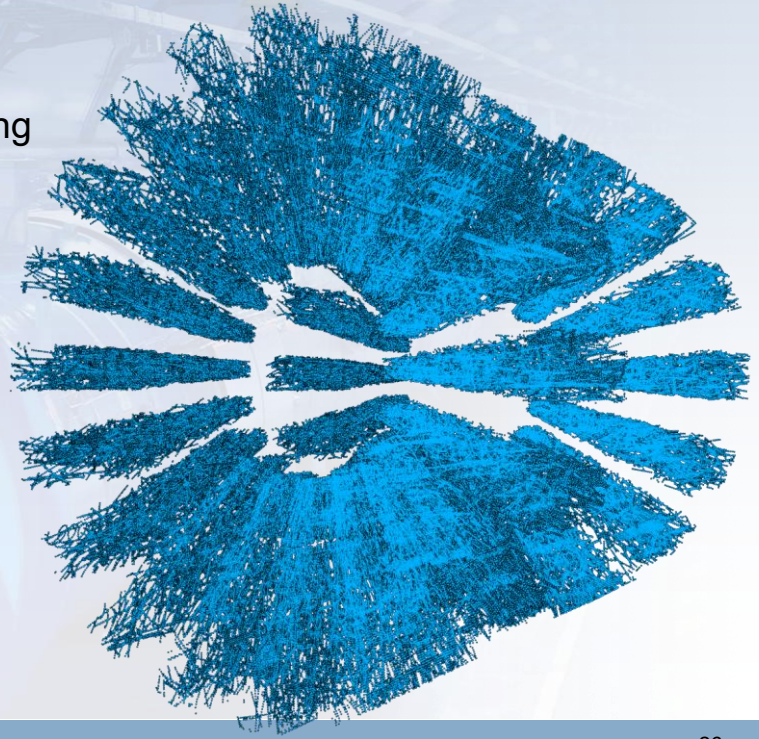
See CTD 2018 talk for more details

- Standalone ITS tracking.
- Standalone TPC tracking, scaling *t* linearly to an arbitrary *z*.
- Extrapolate to $x = 0$, define $z = 0$ as if the track was primary.
- Track following to find missing clusters. For cluster error parameterization, distortions, and B-field, shift the track such that $z = 0$ at $x = 0$.
- Refine $z = 0$ estimate, refit track with best precision
- For the tracks in one ITS readout frame, select all TPC tracks with a compatible time (from $z = 0$ estimate).
- Match TPC track to ITS track, fixing z-position and time of the TPC track.
- Refit ITS + TPC track outwards.

Tracking Algorithm

- **ALICE TPC Tracking algorithm:**
  1. Tracking inside the sectors
  2. Merging of tracks inside and between sectors
  3. Track refit in 3 iterations

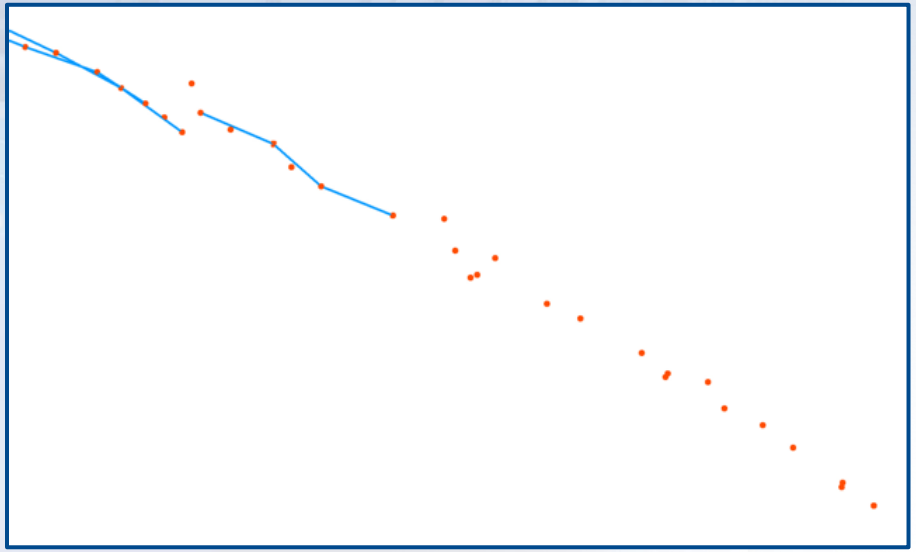
- Note: ALICE Run 3 TPC tracking derived from Run 2 HLT tracking
  - Performance more important than maximum efficiency.



Tracking Algorithm – Sector Tracking

- **ALICE TPC Tracking algorithm:**
  1. **Tracking inside the sectors**
  2. Merging of tracks inside and between sectors
  3. Track refit in 3 iterations

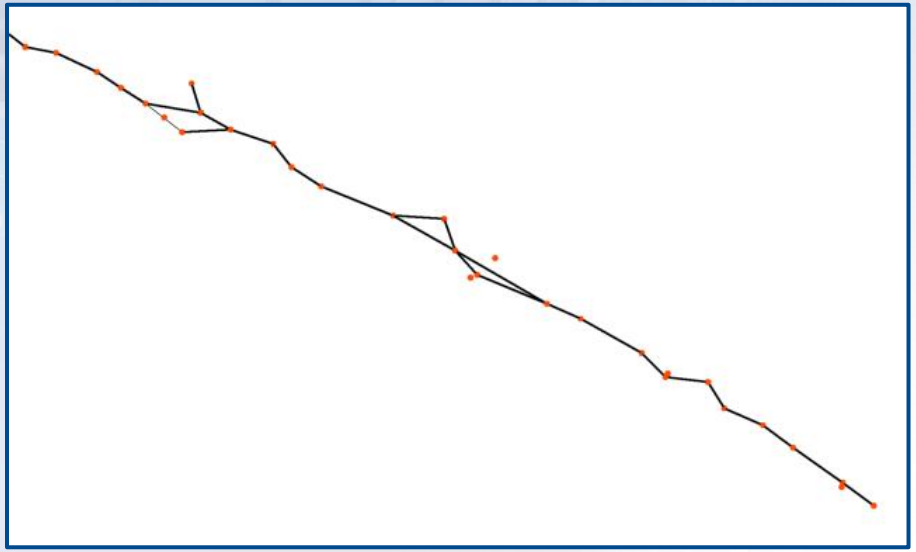
- **Sector Tracking**:
  1. **Combinatorial seeding** (cellular automaton)
     - **Multiple seed** per **track** possible.



Tracking Algorithm – Sector Tracking

- **ALICE TPC Tracking algorithm:**
  1. **Tracking inside the sectors**
  2. Merging of tracks inside and between sectors
  3. Track refit in 3 iterations

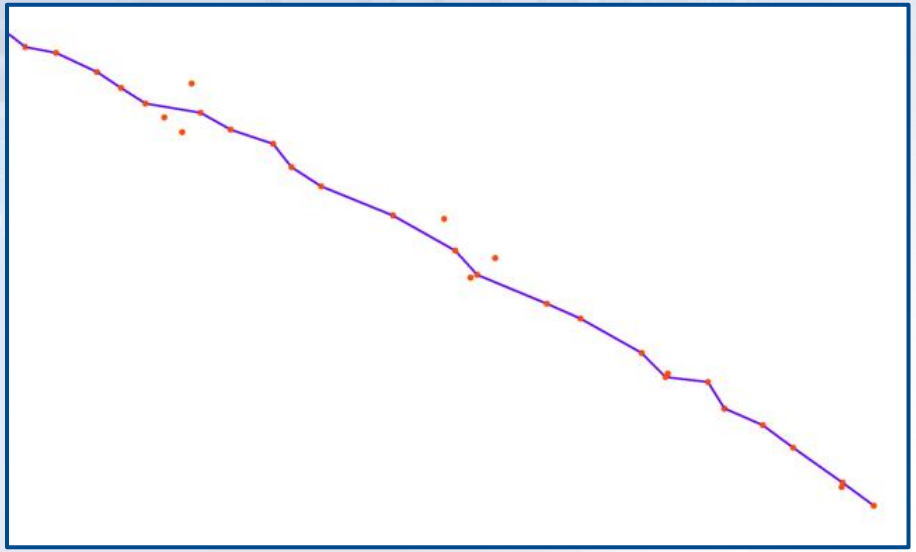
- **Sector Tracking**:
  1. **Combinatorial seeding** (cellular automaton)
     - **Multiple seed** per **track** possible
  2. **Track following** (Kalman filter)
     - **Track** can be found **multiple times**



Tracking Algorithm – Sector Tracking

- **ALICE TPC Tracking algorithm:**
  1. **Tracking inside the sectors**
  2. Merging of tracks inside and between sectors
  3. Track refit in 3 iterations

- **Sector Tracking**:
  1. **Combinatorial seeding** (cellular automaton)
     - **Multiple seed** per **track** possible
  2. **Track following** (Kalman filter)
     - **Track** can be found **multiple times**
  3. **Ambiguity solving**
     - Find **best instance** of the track
     - Assign **all clusters** to **this instance**
       - Some **shared clusters** allowed for crossing tracks.
     - **Remove clones**
       - Due to many seeds, all these steps must be fast!
       - Cannot run complex / itereative steps.



Tracking Algorithm – Sector Tracking

- **ALICE TPC Tracking algorithm:**
  1. Tracking inside the sectors
  2. **Merging of tracks inside and between sectors**
  3. Track refit in 3 iterations

- **Track Merging**:
  - Tracks **transported** to **reference positions** inside or between sectors
  - **Merged** by **track parameters**

Tracking Algorithm – Track Fit

- **ALICE TPC Tracking algorithm:**
  1. Tracking inside the sectors
  2. Merging of tracks inside and between sectors
  3. **Track refit in 3 iterations**

- **Track Fit**:
  - Goes in **3 iterations inward** (**1**) – **outward** (**2**) – **inward** (**3**)
  - 3 Iterations yield the best estimate of the track parameters at the beginning of the 3rd fit, needed for linearization.
  - No additional cluster assignment, only rejection.
    - Rejection based on interpolated cluster positions between 1st and 2nd fit.
    - Clusters themselves are only rejected in the 3rd fit.

Conclusions

- ALICE GEM TPC can operate in continuous readout in Pb-Pb up to 5Khz
- Despite GEM ion back flow suppression, space charge distortions of several cm.
- TPC Distortions are the most complicated calibration.
  - Correction map created residuals of TPC hits to tracks extrapolated from inner / outer detectors.

- TPC detector with the dominant data volume, online data compression mandatory.
  - Using among others track model compression, rejection of clusters not used in physics, AND encoding.

- Tracking complicated in absence of absolute z-position.
  - In particular, SCD corrections can only be applied once a z-estimation is available.
- Tracking algorithm based on Cellular Automaton and Kalman filter.

- TPC processing by far the most time-consuming task online
  - Full TPC online processing running on GPUs.

- TPC offline processing can run on GPUs when the online computing farm is used for offline.
  - TPC represents ~60 of offline reconstruction workload, must run other detectors on GPU for full GPU utilization.
  - Currently 2x – 2.5x throughput increase with GPUs, aiming for full barrel tracking on GPU with 5x speedup.