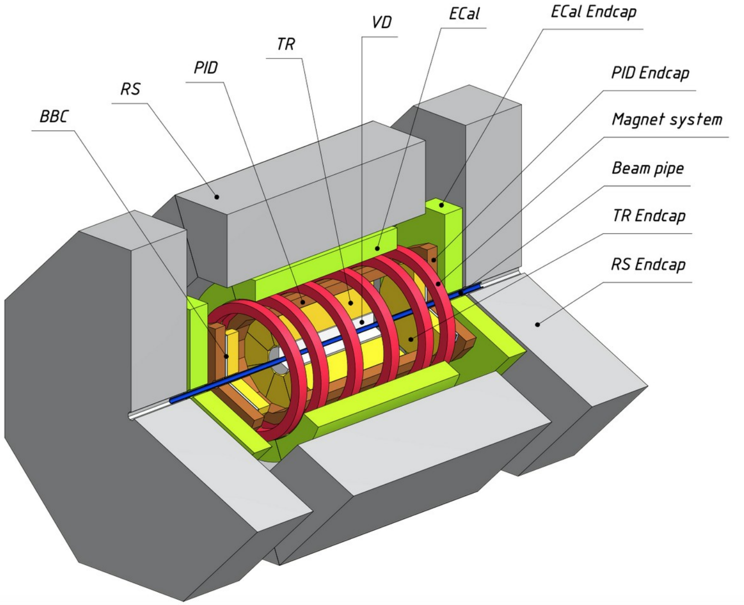
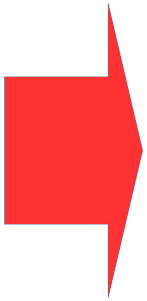
Offline Software and Computing for the SPD experiment

Alexey Zhemchugov
JINR

SPD Collaboration Meeting
13 December 2021

SPD as a data source

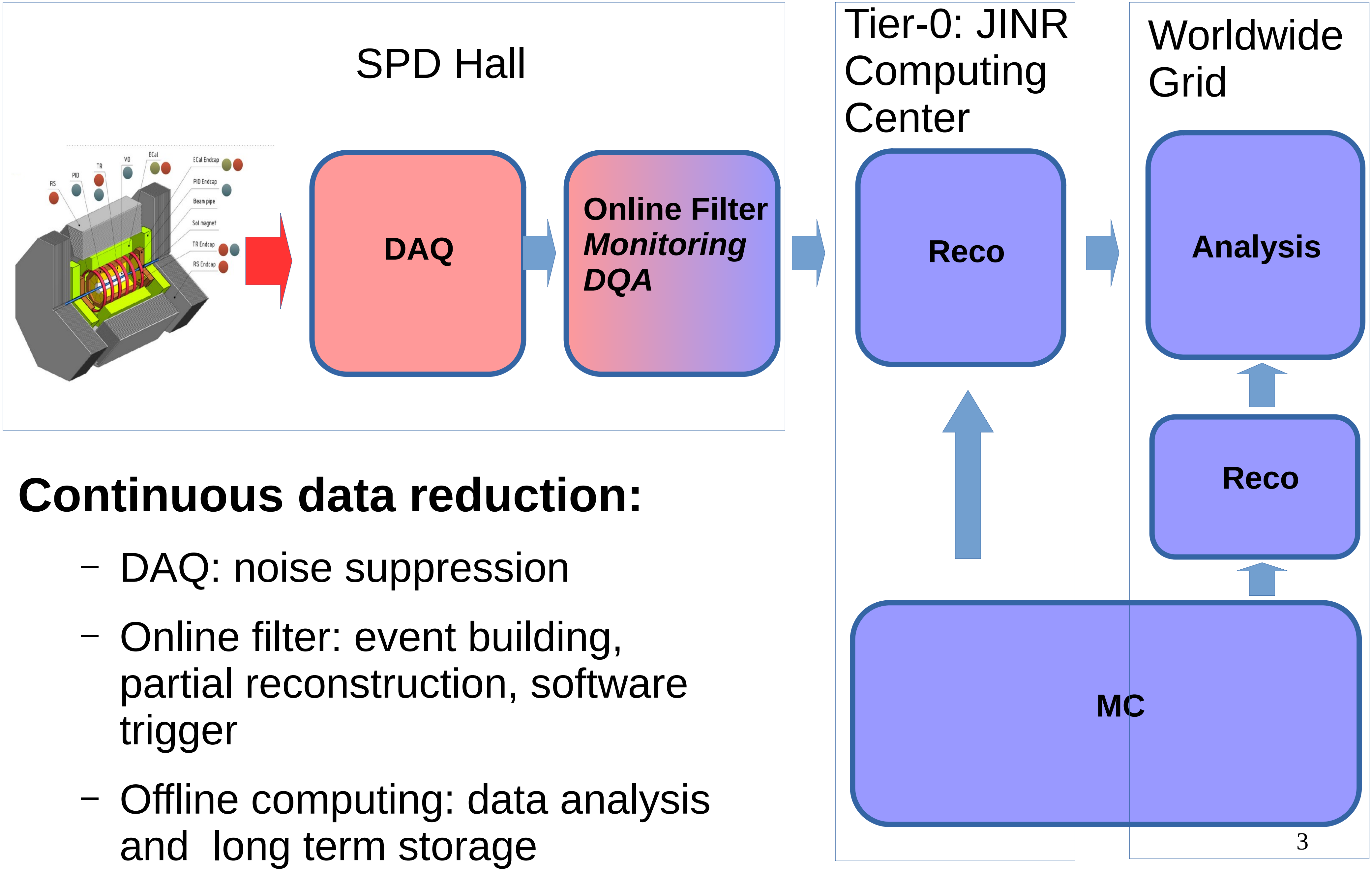




- Bunch crossing every 76 ns = crossing rate 12.5 MHz
- ~ 3 MHz event rate (at 10^{32} $cm^{-2}s^{-1}$ design luminosity) = pileups
- 20 GB/s (or 200 PB/year (raw data), $3*10^{13}$ events/year)
- Selection of physics signal requires momentum and vertex reconstruction → no simple trigger is possible

The SPD detector is a medium scale setup in size, but a large scale one in data rate!

Data workflow



SPD Hall

DAQ

Online Filter
Monitoring
DQA

Tier-0: JINR Computing Center

Reco

Worldwide Grid

Analysis

Reco

MC

Continuous data reduction:

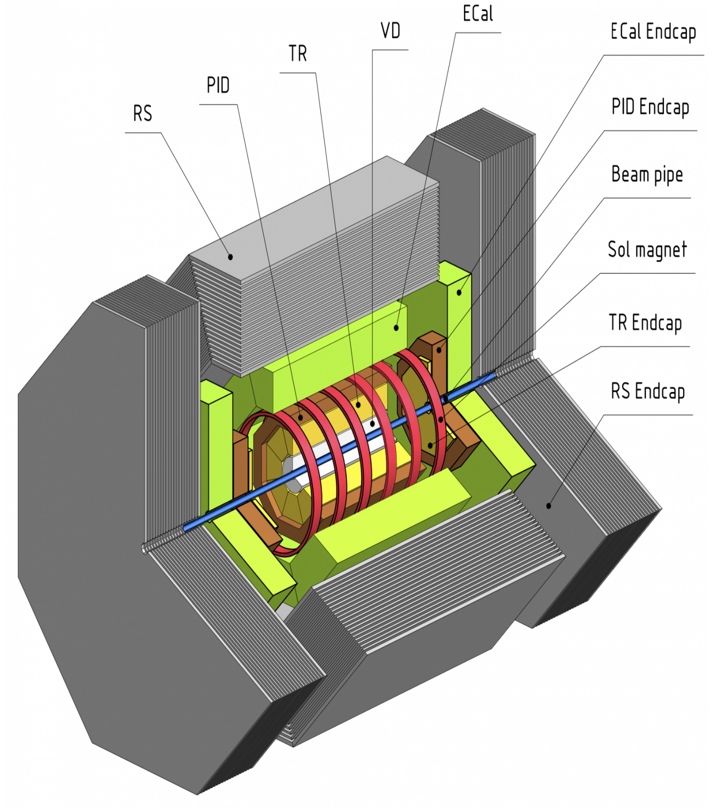
- DAQ: noise suppression
- Online filter: event building, partial reconstruction, software trigger
- Offline computing: data analysis and long term storage

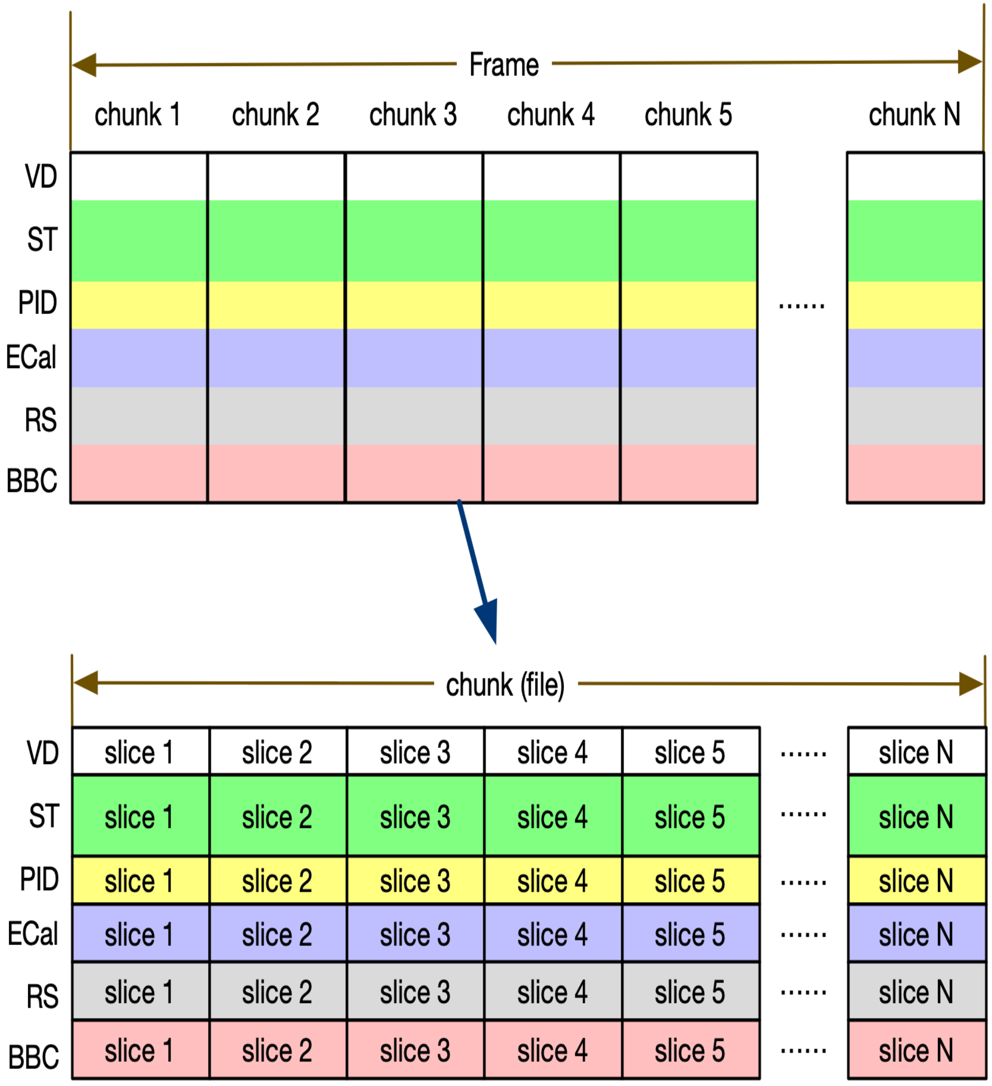
Work packages

- Online data filter
- Offline software
- Offline computing

Input data structure

No trigger = No classical events anymore





Frame

	chunk 1	chunk 2	chunk 3	chunk 4	chunk 5		chunk N
VD							
ST							
PID						......	
ECal							
RS							
BBC							

chunk (file)

VD	slice 1	slice 2	slice 3	slice 4	slice 5	......	slice N
ST	slice 1	slice 2	slice 3	slice 4	slice 5	......	slice N
PID	slice 1	slice 2	slice 3	slice 4	slice 5	......	slice N
ECal	slice 1	slice 2	slice 3	slice 4	slice 5	......	slice N
RS	slice 1	slice 2	slice 3	slice 4	slice 5	......	slice N
BBC	slice 1	slice 2	slice 3	slice 4	slice 5	......	slice N

Online Data Filter ingredients

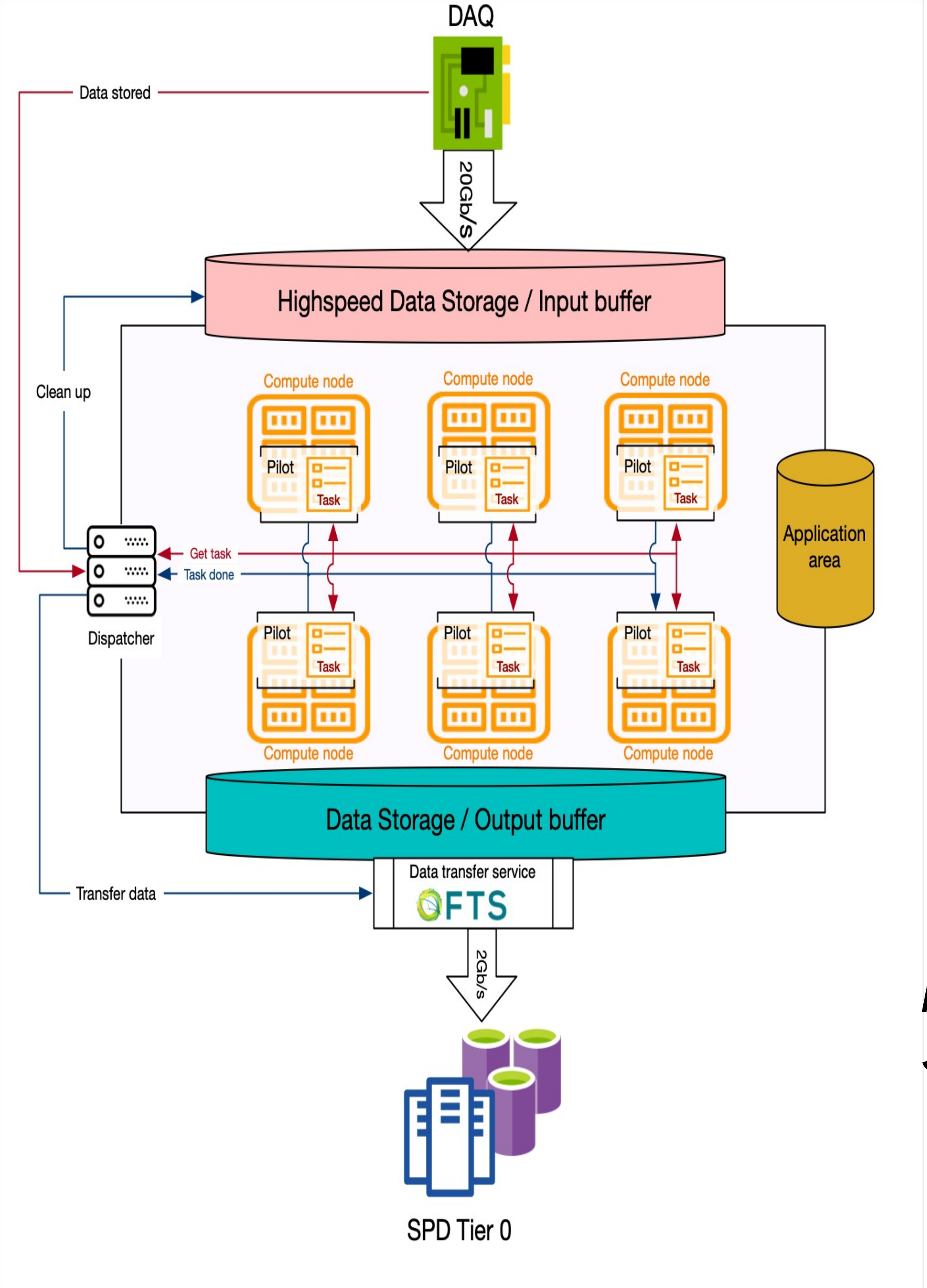
High-performance heterogeneous computing cluster

- Partial reconstruction
  - Fast tracking and vertex reconstruction
  - Fast ECAL clustering
  - Fast RS clustering
- Event unscrambling
- Software trigger
  - several data streams

Machine learning is a key technology

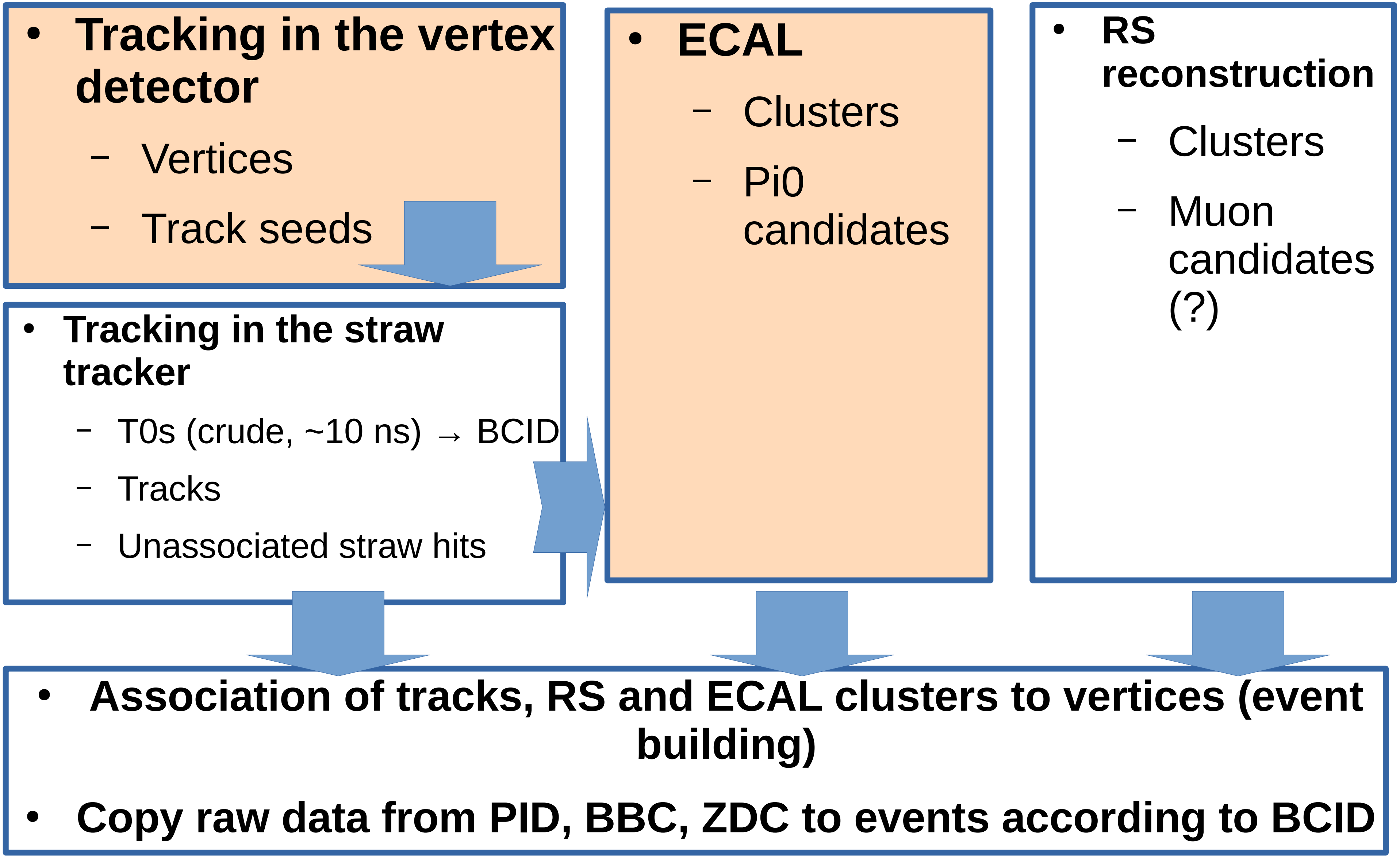
- Monitoring and Data quality assessment
- Local polarimetry
- *More details in my talk in the afternoon*

Online filter operation



https://git.jinr.ru/SPD_online_filter

Fast reconstruction workflow



- **Tracking in the vertex detector**
  - Vertices
  - Track seeds

- **Tracking in the straw tracker**
  - T0s (crude, ~10 ns) → BCID
  - Tracks
  - Unassociated straw hits

- **ECAL**
  - Clusters
  - Pi0 candidates

- **RS reconstruction**
  - Clusters
  - Muon candidates (?)

- **Association of tracks, RS and ECAL clusters to vertices (event building)**
- **Copy raw data from PID, BBC, ZDC to events according to BCID**

SpdRoot: SPD Offline Software

- Core Framework (FairRoot)
- Detector Description
- Event Generators
- Simulation
- Reconstruction

Git repository: http://git.jinr.ru/nica/spdroot

Documentation Wiki:
https://git.jinr.ru/nica/spdroot/-/wikis/home

More details in the talk of Artur Tkachenko in the afternoon

SpdRoot releases in 2021

v 4.1.0 [30.03.2021]

- Compatible with FairRoot 18.2.0 and 18.4.2
- New event structure (particle-track-vertex)
- Physics examples
- Cleanup of obsolete code, geometry options etc.
- ECAL reconstruction

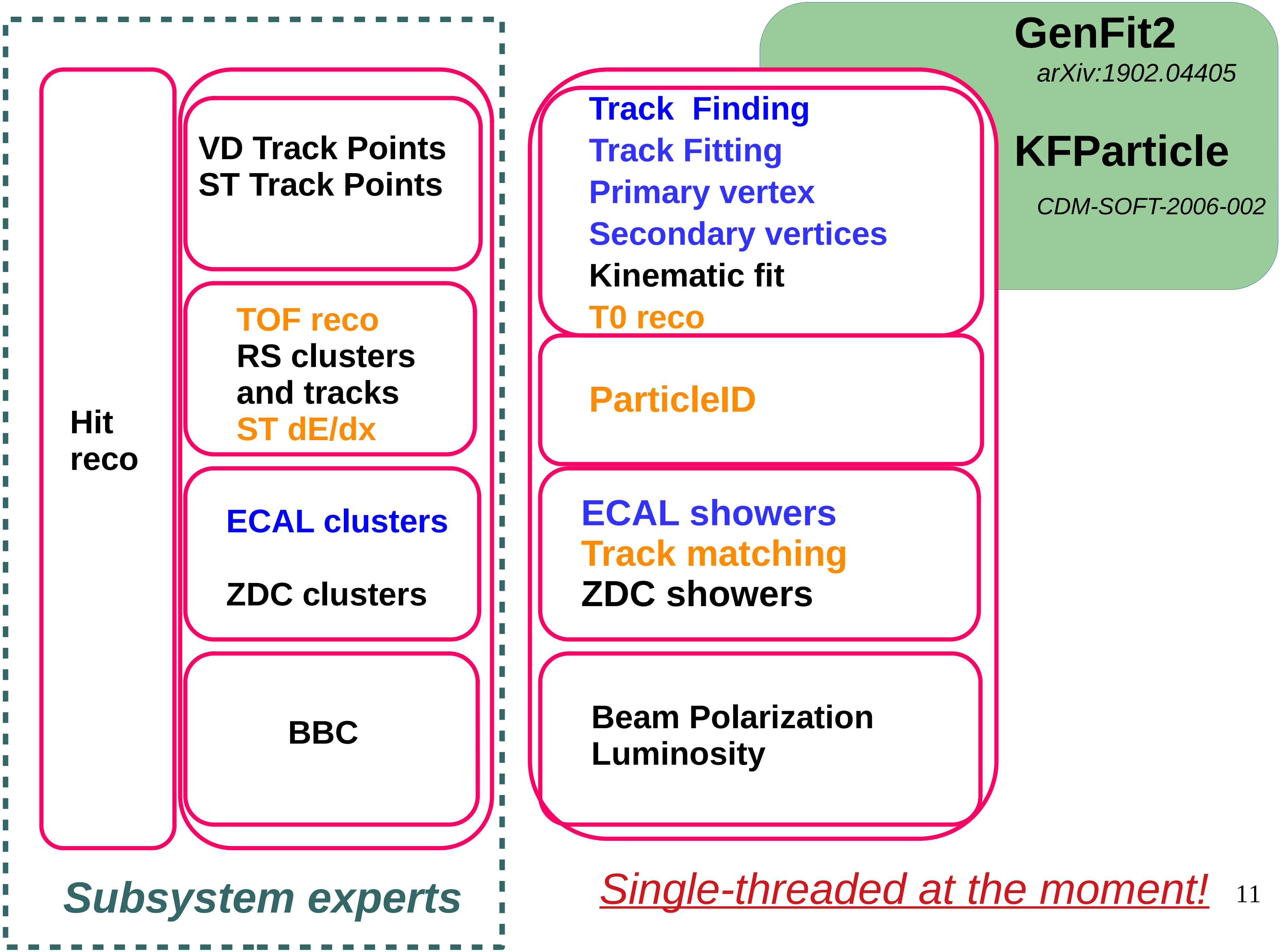
v 4.1.1 [07.09.2021]

- added ECAL and RS MC-truth classes and algorithms
- TOF added

v 4.1.2 [08.12.2021]

- The new particle decayer compatible with Pythia8 generator: SpdPythia8Decayer.
- BBC, ZDC and AEG added
- Detailed description for TOF (barrel and endcaps) modules
- First version of track finding

Reconstruction



Hit reco

VD Track Points
ST Track Points

TOF reco
RS clusters
and tracks
ST dE/dx

ECAL clusters

ZDC clusters

BBC

Subsystem experts

Track Finding
Track Fitting
Primary vertex
Secondary vertices
Kinematic fit
T0 reco

ParticleID

ECAL showers
Track matching
ZDC showers

Beam Polarization
Luminosity

GenFit2

arXiv:1902.04405

KFParticle

CDM-SOFT-2006-002

Single-threaded at the moment!

Missing pieces

- Simulation of the raw data structure to train neural networks
- Detailed detector response simulation
- Reconstruction of RS clusters and muon identification
- Improvement of the detector description
- Calibration, Alignement, Conditions
- *Multithreaded reconstruction*
- *Overall computing performance should be improved in 5-10 times*

An alternative Gaudi-based framework

SpdRoot remains the main tool for the physics studies in the coming years. However, it requires significant developments before it can be used for the real data processing after SPD starts operation.

A new Gaudi-based framework could be an alternative.

- **What we gain:**
  - Robust multithreaded framework with a good user support
  - It is likely we can largely keep and transfer the EDM and key algorithms from the SpdRoot to the new framework
- **What we lose:**
  - Simplicity and smaller size of the SpdRoot
  - Existing expertise in the FairRoot internals

It is a kind of strategic decision. We need more external information and own studies before we make or do not make it.

Offline computing system

All basic components are already available from LHC experiments:

- Workload management: PANDA
- Data management: RUCIO and FTS
- Software distribution: CVMFS + docker containers
- Adaptation to the SPD computing model is ongoing. A prototype exists
- JINR, PNPI, SpbSU and Samara University are willing to participate. Other computing centers are very welcome.
- *More details in the talk of Artem Petrosyan in the afternoon*

Software & Computing R&D in scope of the TDR preparation

- ML-based event reconstruction and an Online Filter prototype
- HDF5 as a data format
- Multithreading and alternative architectures
- FairRoot vs Gaudi
- Conditions DB, Calib&Align
- Computing system prototype and a mock-up test

Summary

- Computing and software of the SPD experiment is making a good progress.

- The online data filter is designed.

- SpdRoot functionality grows steadily

- A prototype of the offline computing system exists already

- Still, we need to do a lot for the TDR preparation:

  - Working prototypes of the fast reconstruction chain based on machine learning

  - Decision on the offline software framework

  - Reilable estimate of the computing performance