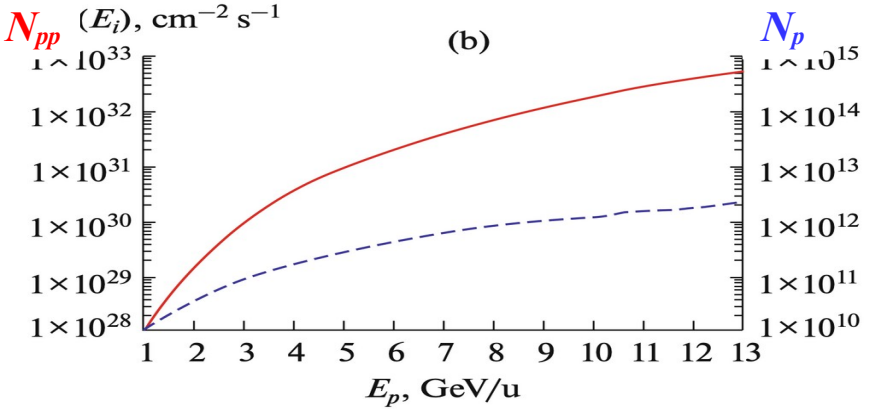
SPD
DAQ & Computing & Software

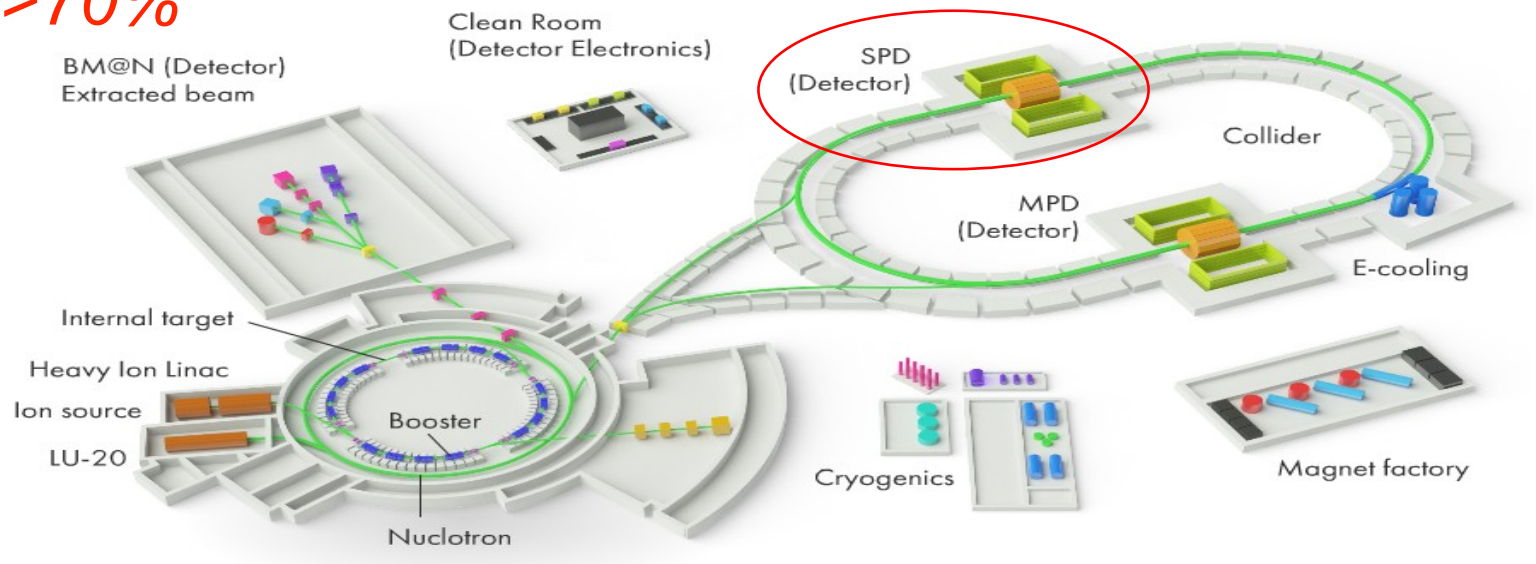
Alexey Zhemchugov
SPD Software Coordinator
zhemchugov@jinr.ru

SPD at NICA

p ↑ p ↑ : $\sqrt{s} \leq 27$ GeV
d ↑d ↑ : $\sqrt{s} \leq 13.5$ GeV
d ↑p ↑ : $\sqrt{s} \leq 19$ GeV

U, L, T
|P|>70%





SPD Collaboration

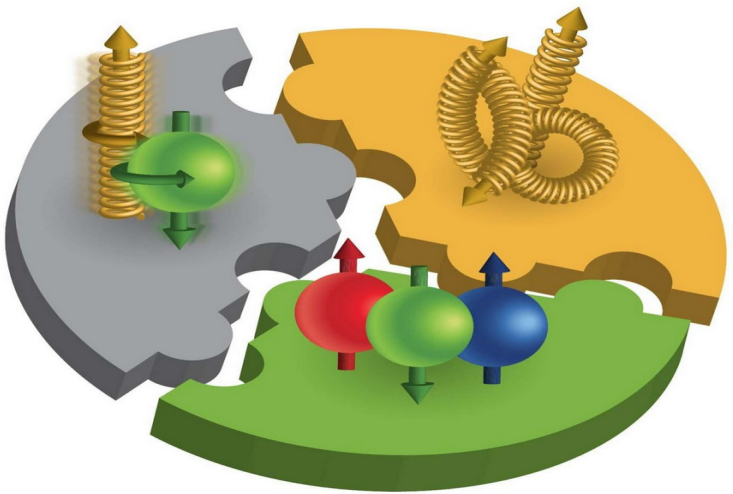
36 institutes from 15 countries, ~400 members

Alexey Zhemchugov on behalf of SPD Collaboration, 14.06.2024

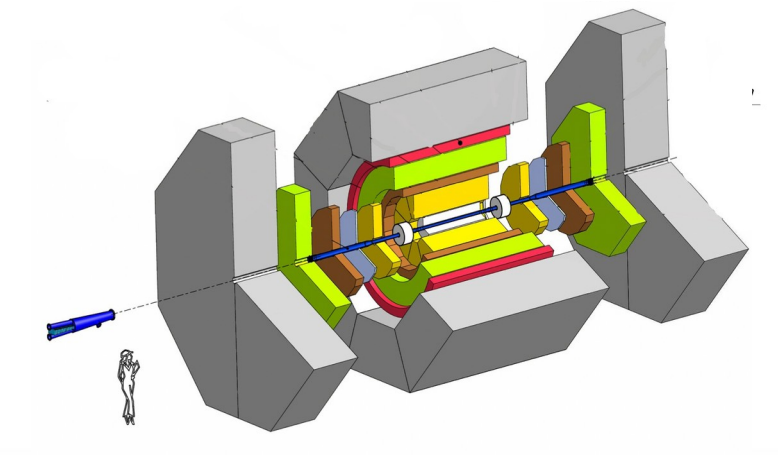
Physics program

- SPD - a universal facility for comprehensive study of gluon content in proton and deuteron at large x
  - Prompt photons
  - Charmonia
  - Open charm
- Other spin-related phenomena
- Other physics

More details:
Prog.Part.Nucl.Phys. 119 (2021) 103858
arXiv:2011.15005



SPD as a data source

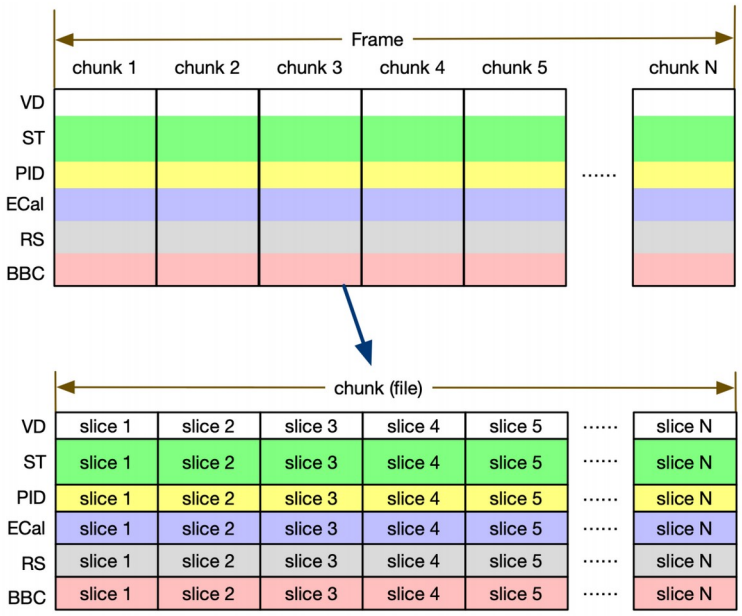


- Bunch crossing every 76.3 ns = crossing rate 13 MHz
- ~ 3 MHz event rate (at 10^{32} $cm^{-2}s^{-1}$ design luminosity)
- 20 GB/s (or 200 PB/year (raw data), $3*10^{13}$ events/year)
- Selection of physics signal requires momentum and vertex reconstruction → no simple trigger is possible

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

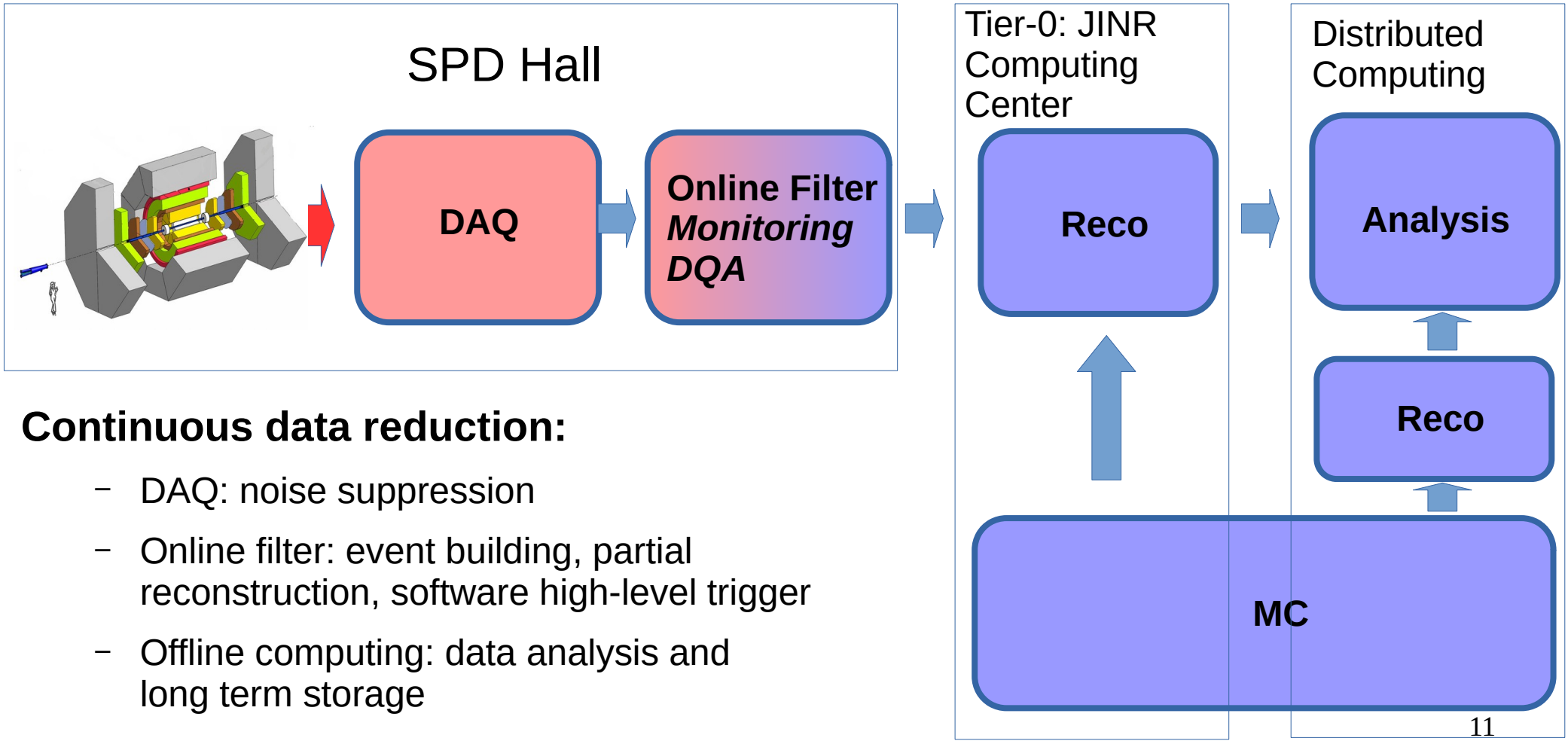
Free running DAQ

No trigger = No classical events anymore



- Primary data unit: **time slice** (1 us — 8.3 ms)
- Time slices combined in **time frames** (up to 549 s, 16 GB max, < 160 MB to fullfil 20 GB/s limit)
- Intermediate units — **time chunks** of 0.1-0.2 s (2-4 GB or ~10^5-10^6 events) are being discussed now
- Every time slices will contain signals from a few to many collisions (events)
- Event building have to unscramble events from a series of time slices.

Data workflow



Continuous data reduction:

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

Online Data Filter

High-performance heterogeneous computing cluster

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

Machine learning is a key technology

Control of systematics?

- Monitoring and Data quality assessment
- Local polarimetry

Reconstruction workflow

- **Tracking in the vertex detector (at the second stage)**
  - Vertices
  - Track seeds

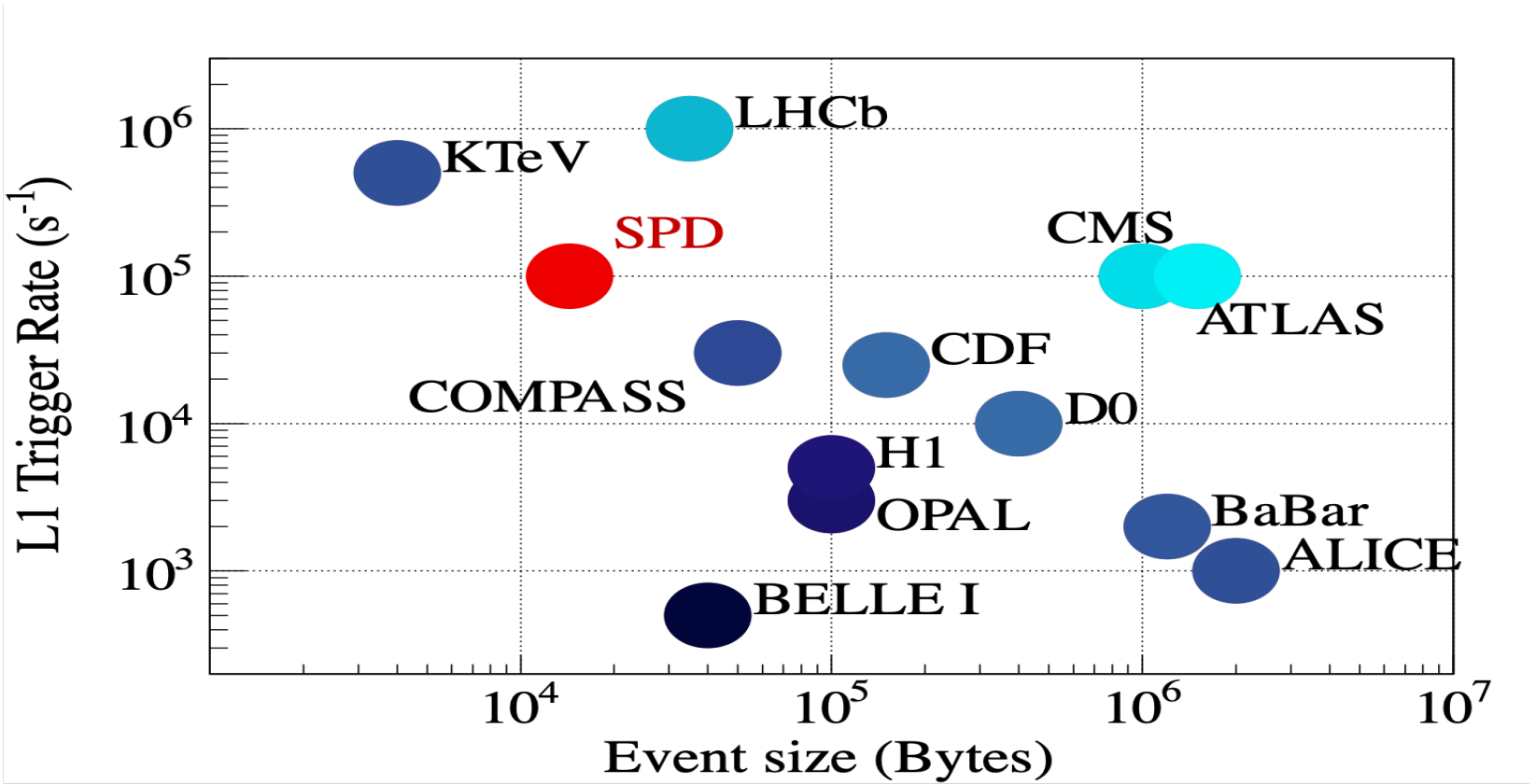
- **Tracking in the straw tracker (+ MCT at the first stage)**
  - T0s (crude, ~10 ns) → bunch crossing time
  - Tracks
  - Unassociated straw hits

- **ECAL reconsruction**
  - Clusters
  - π^0 candidates

- **RS reconstruction**
  - Clusters
  - Muon candidates

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

After the online filter



Expected data volumes

- **Preparation for the experiment.**
  - Monte Carlo simulation from 2024 to 2028 will provide 2 PB per year.
  - Total per stage: 10 PB.
- **Stage I: running at low luminosity of the NICA collider.**
  - Monte Carlo simulation and real data taking from 2028 to 2030 will provide 4 PB per year. Reprocessing: 2 PB per year.
  - Total per stage: 18 PB.
- **Upgrade of the setup for operation at high luminosity.**
  - Monte Carlo simulation from 2031 to 2032 will provide 2 PB per year. Reprocessing: 2 PB per year.
  - Total per stage: 8 PB.
- **Stage II: running at maximum design luminosity of the NICA collider.**
  - Monte Carlo simulation and real data taking from 2033 to 2036 will provide 20 PB per year. Reprocessing: 10 PB per year.
  - Total per stage: 120 PB.

Total for all stages: 156 PB.

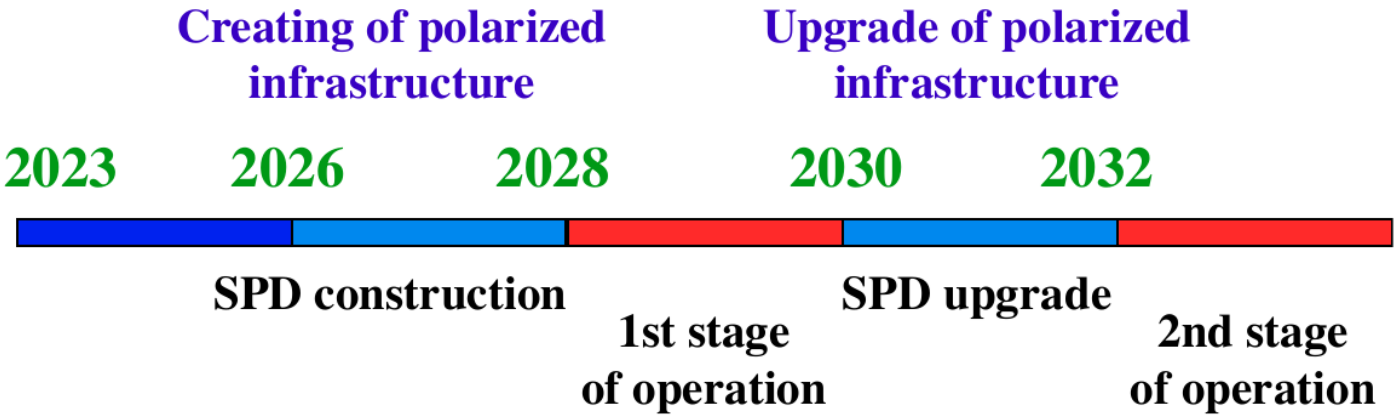
Databases

- Several databases are needed:
  - Data taking conditions and calibration data
  - Physics metadata (including MC input configurations)
  - EventIndex: catalog of physics events, both collected from the detector and simulated
  - Hardware database and mapping
  - Monitoring and logging
  - Collaboration management data.
- Designed as a complex information system that includes data collection and transfer tools, APIs for access from the production and analysis software, client software, supervisors, and monitoring.
- A PostgreSQL RDBMS is considered as a database platform

Offline Software

- A Gaudi-based software framework is being developed:
  - Geometry description: GeoModel
  - Generators: Pythia8, FTF, UrQMD + capability to add more generators
  - Simulation: Geant4
  - Reconstruction: ACTS or GenFit for tracking, Kfparticle for vertex reconstruction, own algorithms for other subsystems
- Current simulation and performance studies are done by another framework SpdRoot, based on FairRoot software

Timeline


Creating of polarized infrastructure
Upgrade of polarized infrastructure
2023
2026
2028
2030
2032
SPD construction
1st stage of operation
SPD upgrade
2nd stage of operation