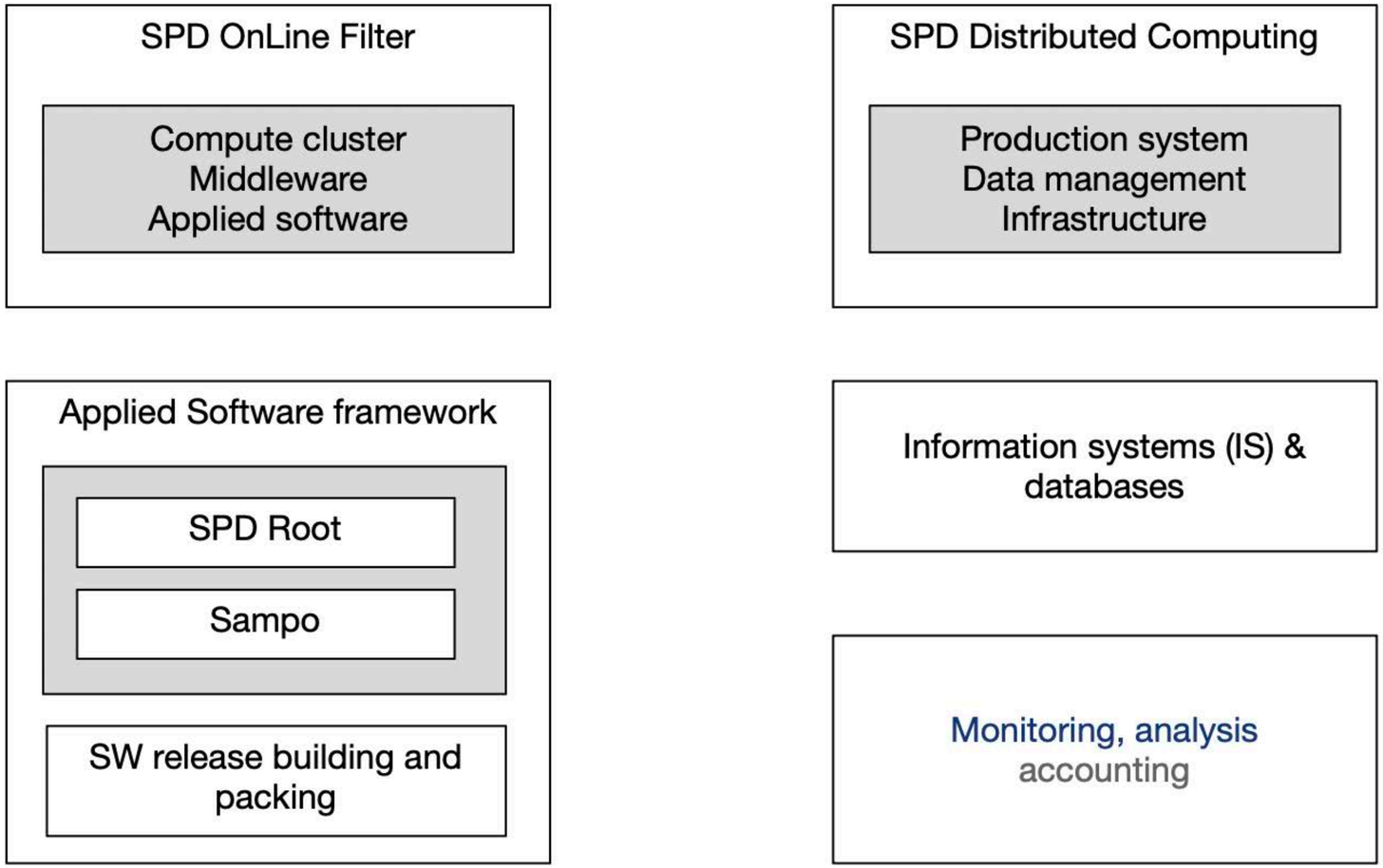
Report of the software coordinator

IX SPD collaboration meeting

Danila Oleynik, 12.05.2025

SPD Software&Computing



SPD Online filter

- Steady progress with middleware development:
  - Integration testing of subsystems working in co-operation, provides more input for tune of code and architecture of micro-services
  - “Load testing” (quite regular JINR Cloud VM’s, simplified synthetic payload)
    - 100 concurrently running pilots
    - ~2100 jobs completed in 7 min
    - Pilot overhead ~15 seconds (with stage-in, stage-out)
- In progress purchasing of hardware for prototyping of compute cluster
  - Not as planned with future NICA Computing data center, but in consideration with SPD DAQ testbed.
  - We expect to have 256 CPU Cores, 1TB RAM, 120TB HDD across four quite fat servers.
- No significant progress with applied software and algorithms
  - Collecting of requirements for applied software should be started ASAP

Applied Software frameworks

SPD Root

- Integrated with Production system
  - Bug fixes mostly to cope critical issues with reliability
- Despite clear direction for decommission, some development still appears
  - Required functionality in successor (Sampo) should be considered
- Scalability limitations work against efficiency
  - Simple example: maximum number of event per job ~4000 case usually quite short processing time, but produces a lot of files. That will affect efficiency of data management and final analysis

Applied Software frameworks

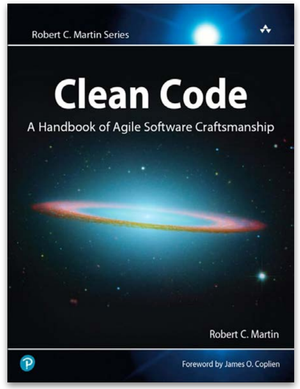
Sampo

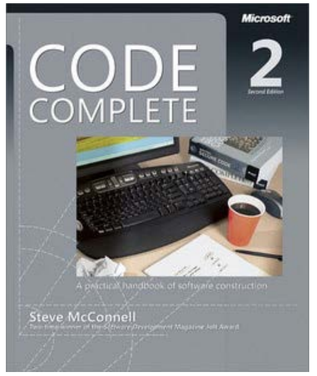
- HepMC3 as main format for produced events (tested with already integrated Pythia8 as primary vertex generator)
- Geant4 support added, including HepMC3 events as primary generator action, magnetic fields and GeoModel for detector geometry
- Gaudi Docker image on top of almalinux9 for developers (ready and tested), also Sampo Docker image for end users (raw, CI/CD configuration in process).

Applied Software

Sampo development

- We are almost ready to involve more developers for implementation of detector specific parts in Sampo
  - Sensitive detectors and hit production
  - Digitization and reconstruction
- Basic knowledges of software engineering is required for developers
  - Sampo should obtain required functionality, correctness and reliability firstly, but efficiency and maintainability also very important.



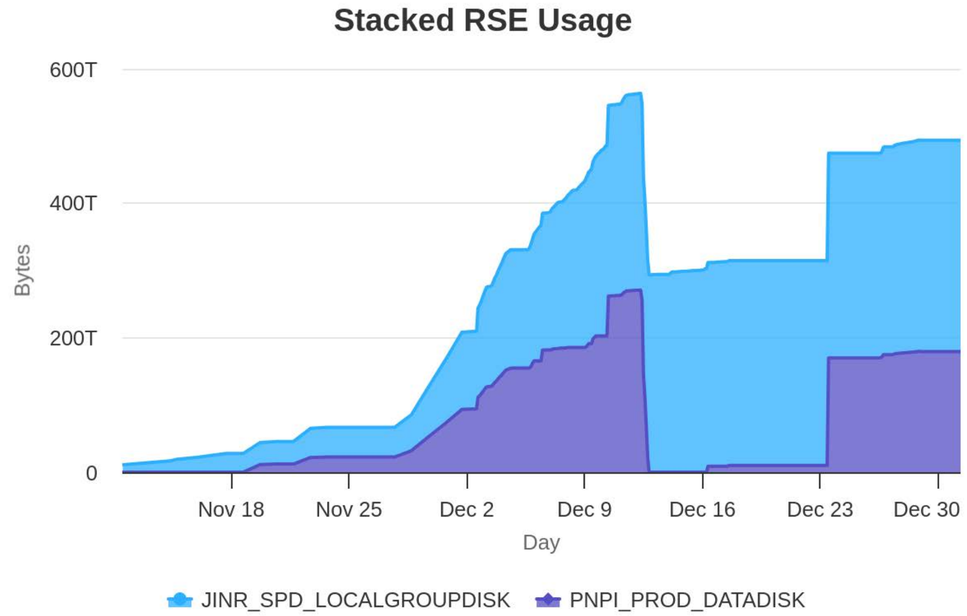


more details in Lev's talk: https://indico.jinr.ru/event/5000/contributions/30815/

Production System testing

November - December 2024

- JINR & PNPI resources used
  - x*10^5 performed jobs n*10^5 produced files
- >500M events produced
- Data distribution on 2 sites (automatically)
- 0,5 PB of data managed by Rucio



☑ Review of naming convention

☑ “Productions registry” required

☑ Operational monitoring required

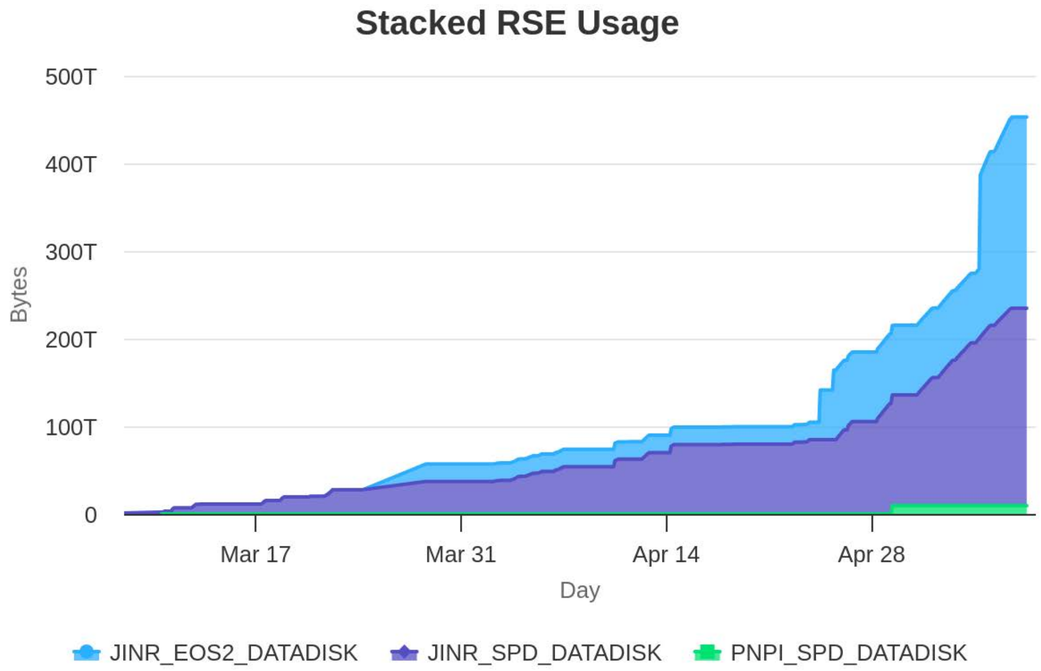
☑ Brokerage tune required

Production system in operation

“Production” - an amount of MC data produced with fixed set of parameters and measured in number of events.

- More or less routine starting from March 2025
  - > 400M events produced (> 800M events processed due to 2-step processing chain)
  - > 200TB of data produced
- Output datasets have size up to 20TB
  - too big to be processed/analysed with desktop workstation
- Please participate in round table on Wednesday to discuss possible solutions for analysis facilities

Production name/ID	Status	Description					Software type/version	Short description (for datasets naming)	Number of events
		Stage	Collision type	Geometry type	Energy	Polarization			
PROD2025-001	Done	S1	pp	Micromegas, TS, ECal, RS, BBC, ZDC (sketch)	10 Gev	UU	spdroot417-dev	minbias-P8-spdroot417-dev test	5 000 000
PROD2025-002	Done	S1	pp	Micromegas, TS, ECal, RS, BBC, ZDC (sketch)	10 GeV	UU	spdroot417-dev	minbias-P8-spdroot417-dev	20 000 000
PROD2025-003	Done	S1	pp	Micromegas, TS, ECal, RS, BBC, ZDC (sketch)	10 GeV	UU	spdroot417-dev	minbias-P8-spdroot417-dev	20 000 000
PROD2025-004	Done	S1	pp	Micromegas, TS, ECal, RS, BBC, ZDC (sketch)	10 GeV	UU	spdroot417-dev	minbias-P8-spdroot417-dev	40 000 000
PROD2025-005	Done	S1	pp	Micromegas, TS, ECal, RS, BBC, ZDC (sketch)	5 GeV	UU	spdroot-dev-4.1.7.1	minbias-FTF-spdroot4171-dev	5 000 000

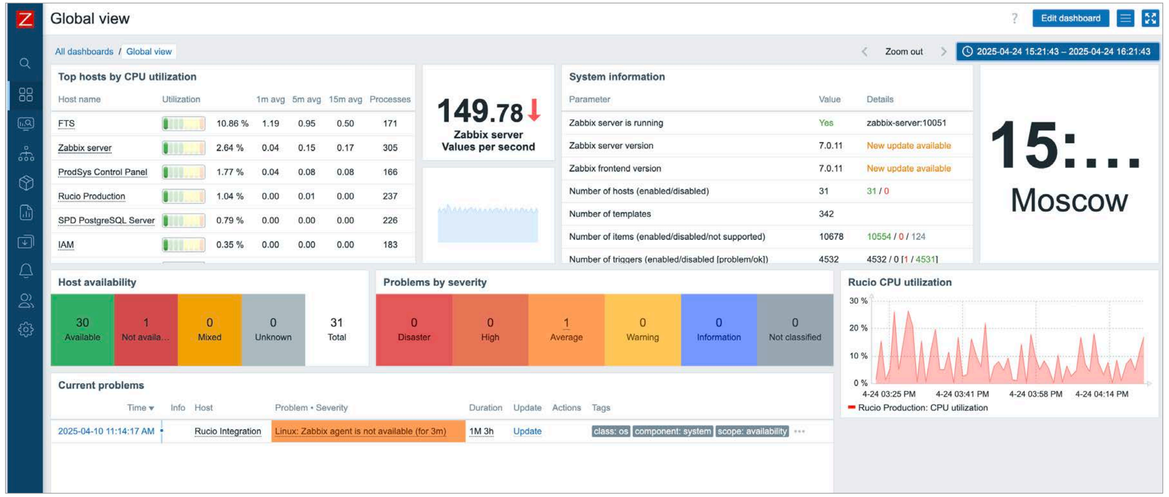


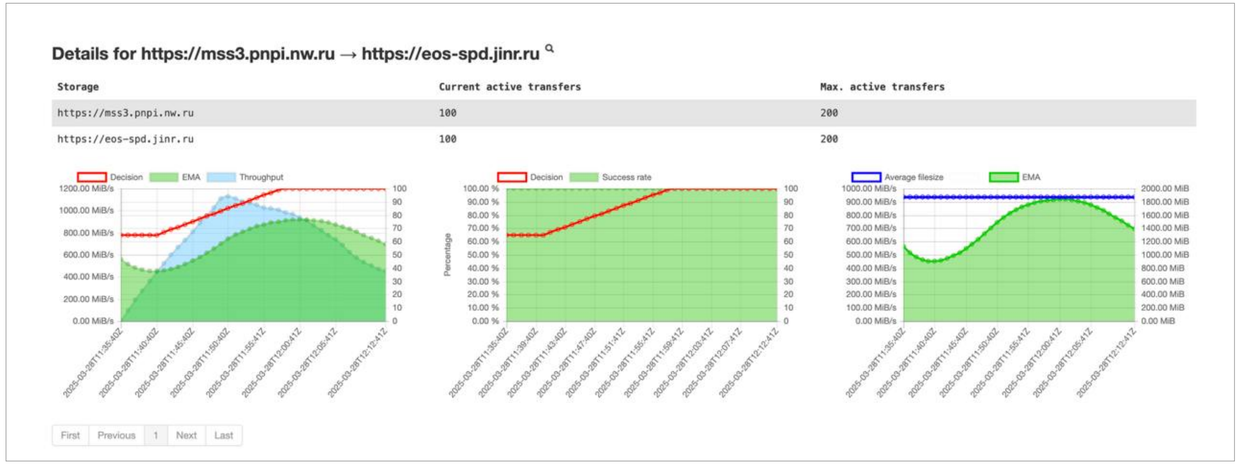
SPD EOS in JINR

- There were some technical and operational issues with central EOS instance eos.jinr.ru at the end of 2024
  - After discussion with the MLIT Director it was decided to deploy a **dedicated storage for SPD at JINR MLIT**
  - Equipment was secured around mid-February 2025:
    - 18 servers
      - 3 head (metadata) nodes
      - 15 disk nodes (24 drives of 20TB each)
    - 7.2 PB raw capacity (**5.3 PB of usable space**)
- Disk volume at JINR should be enough for next 3-4 years
  - Longterm data storage policy should be agreed

Monitoring and analysis

- Definitely required for operation and proper maintains
- For the moment, it’s some uncoordinated solutions
- Requirements should be collected and analysed
- Technology stack should be considered





Software & Computing session