Report of the software coordinator

Alexey Zhemchugov
JINR

SPD Collaboration Meeting
20 May 2024

SPD Software & Computing Project

Online filter

Offline computing system

Offline software

- SpdRoot
- Gaudi-based framework

Databases

SPD Online Filter

- **Step-by-step development of the middleware:**
  - State models of job and file processing are developed
  - Significant progress in the workload management system realization, both server part and the pilot application *(More details in a talk by Nikita Greben on Thursday)*
  - SOF MW integration infrastructure smoothly migrated to Almalinux9 - diminishes issues with python3.
  - Dedicated DBMS (PostgreSQL 16) deployed for all services
- **Not many updates about the payload**
  - The framework development is now stalled due to the lack of manpower
  - Using common (Gaudi based) framework both for offline and for online data processing is being studied
  - No progress in time slice reconstruction from Samara
  - Efficient C++ implementation of TrackNET inference is available

A DAQ-SOF testbed

SpdRoot

Current release 4.1.6 (21 October 2023)

No updates since last Collaboration Meeting

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

Containers: easy start with SpdRoot

- **Docker**

  docker pull jemtchou/spdroot:4.1.6

  docker run -it jemtchou/spdroot:4.1.6

- **Singularity**

  singularity run -H /my/workdir spdroot-4.1.6.sif "run_spdroot_sim.C(10,\"run_555.root\",\"param_555.root\",77777777)"

  spdroot-4.1.6.sif is available at */cvmfs/spd.jinr.ru/images*

This is the recommended way of using SpdRoot!

No not try to install from source unless you really need it!

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 is an alternative solution.

- There is a good progress in making the new framework *(more details on Thursday)*
- Moved from CentOS7 to AlmaLinux9. External packages and Gaudi were rebuilt.
- The building infrastructure is ready
- Pythia generator, GeoModelSvc and basic EDM are implemented. Geant4Svc and MagFieldSvc are at work
- The plan is to have the main simulation chain ready by the end of summer

Offline computing system (1 of 2)

- **Virtual organization management service**
  - ✔ Transition from VOMS to Indigo IAM was performed
  - ✔ All user records were transferred from VOMS to IAM
  - ✔ IAM service provides JWT support, now users can work without using X.509 certificates
  - ✔ But, if needed, IAM provides proxy generation mechanism as well
  - ✔ SSO authentication mechanism configuration is now ongoing. As soon as done, the service will be fully integrated with JINR computing security
  - ✔ We plan to integrate IAM with JINR HR database to enable “auto registration” for members of the SPD collaboration
- **Data transfer service**
  - ✔ A FTS service was deployed in December at JINR
  - ✔ Several connectivity and performance tests between EOS instances at JINR and PNPI done during winter

Offline computing system (2 of 2)

- **Data management service**
  - ✔ Storage elements of JINR and PNPI were defined in Rucio
  - ✔ SSO, IAM integration: users can log in using any preferable authentication method
  - ✔ Definition of rules, deletion and replication were performed
  - ✔ Rucio clients were deployed on CVMFS, now available to all users; users can use either voms-proxy-init or JWT to authenticate
  - ✔ Web UI is also available, integrated with JINR SSO and SPD IAM service
  - ✔ Naming convention for production datasets to be stored in Rucio was defined
- **Workflow management**
  - ✔ Successfully generated several samples of full chain MC productions. Result is stored on EOS, metadata and data discovery available through Rucio
  - ✔ Generation of larger samples in order to get more statistics is now ongoing

Databases (or Information Systems)

A number of databases are being developed to effectively handle experimental data and detector information

- **Hardware Database** – catalog of hardware components: Prototype: data storage, back-end API, Web front
- **Mapping Database** – connection scheme for DAQ:

  Prototype: data storage, back-end API
- **Physics Metadata:**

  Prototype with tables for runs, datasets , software and MC tags: data schemes, back-end, deployment and testing
- **Event Index** – catalog of events from detector and MC

  Prototype: data storage, API back-end, web front, data loader
- **Personal and publication databases:** looking forward for a common solution for NICA

Software & Computing workplan (1 of 2)

	2024	2025	2026	2027	2028
Core services support & DevOps					
Core services support & DevOps					
Offline computing					
Offline middleware development					
Offline midleware support					
Data challenges and testing					
Tier-0 deployment					
Tier-1 deployment					
Offline software					
Core framework					
Simulation					
Reconstruction					
Integration with the offline computing and deployment					
Data challenges and testing					
SpdRoot support					

Software & Computing workplan (2 of 2)

	2024	2025	2026	2027	2028
Online filter					
Online Filter Middleware					
Online Middleware Support					
Online Filter Application					
Hardware purchase and installation					
Software testing and deployment					
AI/ML					
Neural network R&D					
Integration with Online Filter Application					
Integration with Offline Computing					
Databases					
Experiment DBs					
EventIndex					
CRIC					

(Minimum) manpower request

- Core services + DevOps — 5 FTE
  - core services (git, vm, containers…), devops, online filter hardware
- Online Filter — 6 FTE
  - *middleware (4), application (2)*
- Offline Software — 10 FTE
  - *core fw (1), DD+MagF (1), simu (1), reco (6), tests (1)*
- Offline Computing — 12 FTE (+ 1-2 FTE/Tier-2)
  - Panda (1), Rucio (1), Workflow (1), Monitoring (1), Ops (2), Tier-0 (4), Tier-1 (2)
- AI/ML — 2 FTE
  - Tracking (1), FARICH (1)
- Databases — 5 FTE
  - Exp Dbs (2), EventIndex (2), CRIC (1)

TOTAL — **40 FTE (+ 1-2 FTE/Tier-2)**

~14 FTE available now