PANDA QoS Architecture

Quality-of-Service Levels

Based on the processing stages and produced data objects the following Quality of Services (QoS) objects are defined:

1. **RAW**: De-contextualized ("unpacked") data extracted from the data stream as delivered by the PANDA data acquisition after the online event selection. Typical objects are digis (digital single-channel information), typical containers are event candidates. In addition this data also contains the information which trigger line has selected the data.
2. **FULL_OLINE**: For a fraction of the filtered events the full online reconstructed data set is written out to monitor and optimize the online event filter. The data types match to ESD plus AOD and is transient.
3. **SIM**: Simulated data with full MC information. PANDA intends to provide simulation statistics equivalent to twice the triggered experiment data statistics. Data objects are MC tracks and MC points, containers are events.
4. **RAW_SIM**: Data looking like coming from the experiment but based on simulated events. Typical objects are digis with additional information from which MC data (SIM) they are coming from.
5. **ESD**: Event Summary Data contains the detailed output of the detector reconstruction and will be produced from raw(_sim) data. It will contain sufficient information to allow particle identification, track re-fitting, energy calibration etc. thus allowing for the rapid tuning of reconstruction algorithms and calibrations. In addition it contains a link to the RAW data it was produced from. This data is transient.
6. **AOD**: Analysis Object Data serving as input for high-level user analysis. AOD are derived from ESD by combining the different detector results for a single particle and skimming. Different AOD types may be defined serving different physics analysis objectives. Typical objects are charged or neutral particle candidates, typical containers are reconstructed events.
7. **PAR_RUN**: Parameter data required for the production of the ESDs which have been set for a run or measured during a run. Typical parameter sets comprise the experiment configuration (geometry, settings) defined at the start-up of the experimental run, the running conditions monitored and recorded during the experiment operation. This data does not change anymore after the run is finished and is as important as the RAW data. A reconstruction of the RAW data without the parameters is not possible with the required precision. Parameters are typically managed through appropriate data bases.
8. **PAR_CAL**: Calibration parameters obtained through an analysis of RAW data needed for the production of ESDs. Parameters are typically managed through appropriate data bases.
9. **PHY**: Physics-level results, usually in a binned and inclusive format. These are derived from AOD and constitute the experiment results to be made public.

Data Operations

The anticipated workflow can be outlined by the following steps:

1. Experiment data taking: RAW data are delivered to on-site storage media (RAW_HOT). PANDA aims at taking data for 100 effective days per year. The estimated raw data volume is 1 PB/a.
2. PAR_RUN is stored with each run in addition to RAW.
3. RAW and PAR_RUN are copied to a long-term archive (RAW_COLD). Two copies of RAW have to be archived in two distinct data centers, one being FAIR/GSI. The archival process can already be started during experiment operation.
4. Offline-level calibration parameters PAR_CAL are derived by analysis of a subset of RAW. This happens at the host lab, where RAW_HOT is available.
5. ESD are created from RAW and PAR. ESD data is transient.
6. AOD are created from ESD.
7. AOD are transferred to the participating data centers, serving a regional PANDA user community. The type of AOD may depend on the principal analysis aims of the respective community. A full copy of all AOD data is kept at the host laboratory.
8. AOD are subjected to high-level physics analysis which lead to PHY data for publication.
9. It is planned to regenerate PAR_CAL and reprocess the ESD and AOD data from RAW data in 4 consecutive years with improved detector calibrations and reconstruction algorithms.
10. SIM data are generated in correspondence with the experiment settings, conditions and statistics. RAW_SIM, ESD and AOD production from SIM data is performed on-the-fly. In addition to the AOD data the MC truth of simulated tracks and the generated DIGI data are stored. Simulation can be performed on any of the participating data centers.
11. AOD_SIM produced from SIM are subjected to physics analysis, complementing the analysis of the corresponding experiment data.

Quality-of-Service Policy

The following service levels can be identified on the base of the workflow outlined above. They do not include PAR and PHY data because of their negligible size compared to experiment and simulated data.

Name	Usage	Volume	Characteristics	Example media
RAW_COLD	Long-term storage prime experiment data.	1 PB/a	High reliability and long-term stability	Tape
RAW_HOT	Availability for calibration and production of ESD for 4 years after data taking	4 PB	Low latency	Disc
ESD	Transient data for AOD production	1 PB	Low latency, transient	Disc
AOD	Availability for user-	23 PB	Low latency,	Disc

	level physics analysis up to 10 years after data taking		distributed access	
SIM	Availability for MC-level analysis up to 4 years after production	11 PB	Low latency	Disc
AOD_SIM	Availability for user-level physics analysis up to 10 years after data taking	12 PB	Low latency, distributed access	Disc
PAR_RAW	Needed for RAW and ESD processing		Low latency, Database	Disc
PAR_COLD	Long-term storage prime experiment data.		High reliability and long-term stability, Database	Tape