Software Development: short and medium term plans

CLAS Collaboration Meeting

July 10, 2018

Software development priorities

During the next 6 months:

- Support to preparations for the fall-winter data taking
- Improvement/completion of reconstruction software
- Tuning/speed-up/portability of simulation software
- Development of analysis framework and tools
- Support to the First Experiment data processing (calibration, reconstruction, analysis, …)
- Documentation

Online software upgrades

Goals: reduce data volume, improve data handling efficiency and data quality monitoring

- Data rate reduction by bit-packing of fADC/MM data
- Event rate reduction by implementing full tracking trigger with geometry matching
- Output file size
- Online reconstruction farm

When: before begin of data taking

See Sergey's talk tomorrow

Fall run configuration

Goal: support studies to optimize upcoming run configuration

- Development of FMT reconstruction
- Update LTCC in reconstruction/simulations
- Give support to tracking efficiency to optimize operating luminosity

When: ongoing, to be completed by mid to end of July

See talk(s) in First Exp. Session

Reconstruction software upgrades

- Improvements to existing code:
  - Handling of magnetic field and swimming optimization
  - Tracking code speed-up (code optimization, algorithmic improvements, roads,…)
  - Code cleanup and optimization

- Completion of reconstruction chain:
  - Debugging of existing reconstruction services
  - Completion or implementation of partial or missing services

When: field-handling improvements presently ongoing, code-speed within 2-3 months, code optimization and reconstruction complete by the end of the year

Simulations: tuning, speed-up and portability

- Continue GEMC development/improvements:
  - Code speed-up
  - Magnetic field handling: rotations, non symmetric maps
  - Support for run dependent geometries
  - Tuning of detector response
  - Hipo output

- Running simulations on offsite farms:
  - Use of "containers" (docker, singularity)
  - Workflow development
  - Identification of offsite resources

When: code speed-up, field handling upgrades and deployment via "containers" ongoing, run-dependent geometries starting, support for running offsite and gemc tuning continuous

See Maurizio's talk tomorrow

Development of analysis tools

Goals: define data format for analysis and provide tools for easy and efficient data access and first analysis

- DSTs:
  - Optimization of DST content
  - Format (HIPO)
  - Event tagging mechanism
- Analysis trains:
  - Framework development
  - "Wagon" template for skimming
  - Development of workflow and usage policies

When: ongoing, first release available to support analysis of RG-A data during the summer

Detector calibration and geometry

- Calibration:
  - Continue support for calibration software
  - Ability for batch processing

- Detector geometry:
  - Complete/extend geometry services to handle run-dependent shifts and rotations from survey/alignment
  - Support development of alignment software

When: improved support for detector geometry within summer, support development of calibration or alignment software continuous, batch farm calibration processing with low priority

Offline software improvements

- Offline code:
  - Code management and build
  - Validation

- Common tools maintenance and development:
  - Upgrades to event decoder to handle bit-packed information
  - Fast-MC
  - Database tools
  - GROOT upgrades

When: upgrades to event decoder ongoing, other tasks with medium priority

Timeline

	Online:	Offline:	Simulation:	Analysis:	Calibration:
July	▪Complete implementation of bit-packing and trigger upgrades	▪Complete software developments in support of RG-A optimization studies ▪Complete magnetic field handling improvements ▪Decoding of bit-packed fADCs	▪Deployment via containers ▪Handling of magnetic fields	▪DST content	▪Improve/extend geometry service to handle run dependent rotation and shifts
August	▪Switch to large data files	▪Continue reconstruction code speed-up work (x3 speed increase)	▪Code speed-up ▪Run-dependent geometries	▪First release of analysis trains framework and template	▪Continue support to calibration and alignment software
September	▪Online farm	▪Code optimization	▪Workflow for simulations at offsite farms ▪Continue tuning of detector response ▪Hipo output	▪Continue developments of analysis trains and tools	
		▪Complete reconstruction services (to be continued…)		▪Event tagging in DSTs (to be continued…)	

Summary

- Focus of software development in upcoming months on:
  - Addressing critical areas (data rate, event reconstruction processing time, ..)
  - Consolidate offline software (complete/improve reconstruction, calibration, simulation, common tools, …)
  - Develop analysis tools
  - Documentation !!!
- Key software developers already committed at 100%
- Stronger Collaboration involvement and support needed!