CLAS12 "Ready for Science" Review
Hall B Internal Review
Dates: September 25-26, 2017, CC Rm L102/F326-327
Final Report, October 2, 2017.

Review committee:
Keith Griffioen, William and Mary
Kyungseon Joo, Univ. of Connecticut
David Lawrence, Hall D
Bryan Hess, Scientific Computing
Elton Smith, Hall D (co-chair)
Stepan Stepanyan, Hall B (co-chair)
Beni Zihlmann, Hall D

We would like to commend the "CLAS12 First Experiment" effort for preparing and planning for the efficient collection, calibration and analysis of the first experiments with 12 GeV beams in Hall B. The committee thanks all presenters for working together to provide all the information required to address the charge and for clear and careful presentations. We would also like to thank everyone for open and frank assessment of issues that came up during the review. The agenda and list of talks presented at the meeting are given in the Appendix.

We respond to each question in the charge point-by-point and make a list suggestions or recommendations for improving or optimizing the process toward publication of first results.

Specific Charge Items

1. *Is the presented **commissioning plan** for CLAS12 comprehensive and developed in sufficient detail to ensure that upon completion the CLAS12 system will be ready for production data taking? Is **the timeline reasonable** and optimized, both in terms of duration of the study and the order of activities?*

   Yes, a comprehensive plan has been presented with sufficient detail. However, considering possible inefficiencies in CEBAF running, especially at 5-pass with 750 MHz RF separator, this program must be prioritized. Activities for preparation of a successful running of RG-A must have the highest priority (e.g. commissioning of detectors/triggers with RG-A run conditions). We recommend taking sufficient data in RG-A conditions (magnet fields, luminosity) in December to allow for detailed studies of test reactions during the break. Other commissioning items can be done during later stages of the engineering run, if time allows (e.g. multiple combinations of magnet field settings).

   1. Various commissioning and calibration steps must be supported by simulations. Clear goals and deliverables must be defined for each step.
   2. Break down commissioning plan into must/should/like categories to ease prioritization.

2. *Have the necessary **production triggers** been developed that are needed for the physics run, and are plans in place to test their efficiency?*

   The data acquisition electronics and software is on course to be fully functional and ready by December 1, 2017. Some electronics is still being installed, but there are no identified problems. The system should allow for 10k events per second with minimal DAQ failures.

   A long list of triggers was presented. Commissioning of only two trigger types were outlined with no details. One of the triggers in the validation list was absent from the trigger bit table (PCAL_ECAL). There are many FT triggers discussed, but it was unclear in which order FT triggers would be validated.

1. Areas of concern: There is insufficient time allocated for trigger commissioning. There should be a clear plan as to what trigger configurations are necessary, what pre-scale factors are required to accommodate all experiments, and how to compromise if this cannot be achieved. Simulations of the various trigger rates will allow the collaboration to prioritize.
2. A GUI to modify the trigger sounds ill-advised. Although it makes sense for shift workers to adjust the pre-scale factors on the triggers, designing the trigger configuration should be left to an expert.
3. There needs to be a second person during the experiment who is qualified to carry the DAQ pager.
4. Until the trigger configuration and data-rates are understood and stable, compromises may be required that globally optimize the statistics for multiple experiments.

3. *Are the presented **monitoring and software tools** adequate for the efficient commissioning of all CLAS12 systems?*

Offline Software Tools

The offline software package coatjava is well developed and appears mature enough to handle processing of the first physics data. Integration into the existing batch farm system has been tested and looks well suited for processing the data from the RG-A run.

Online/offline Monitoring

1. A responsible person should be identified to lead the online monitoring effort.
2. Numerous expert monitoring tools exist for different subsystems. The mon12 tool and CED appear to be the only existing tools that will be utilized by shift takers to monitor data quality. It is recommended that a system be implemented that makes it easy to compare a screen from mon12 side by side with a reference plot. This should be integrated into mon12 if possible, but can be as simple as a button that opens a web browser to the appropriate URL.
3. Monitoring plots should be published periodically to the web. This could be as often as once per run, but should be at least once per day. Having plots available on the web will allow offsite collaborators to help with monitoring duties. There was some discussion at the review

about inserting monitoring plots into the e-log and this would be one possible solution.

4. Some specific physics channels were identified for purposes of online monitoring. It is recommended that you follow this up by implementing a system to produce a standard set of plots either directly from ET or at least daily by the offline shift crew. These should include fits where appropriate with values printed on the plots that can be compared to reference plots.
5. Clear instructions need to be given to the shift crew regarding their monitoring responsibilities.
6. The alarm system needs to be integrated to watch all quantities of concern for the experiment, including high level quantities.

DAQ

1. The DAQ/Trigger shift worker documentation is not quite complete. While adequately describing how to start a run and recover from a crash, it does not describe how to cleanly stop a run and start a new one. It also does not describe where the data is written to in the counting house or where it is written to in the tape library. (This would be useful for shift takers to check that the data is being saved properly.) The section on the Trigger GUI also seems incomplete.
2. The collaboration plans to take both mode 1 (full sample) and mode 7 (integrated) data using the flash ADCs during the engineering run. It was not clear that there is consensus within the collaboration on whether to take mode 1 or 7 during production running. It is recommended that the collaboration adopt a position that standard production running be done using mode 7. Emphasis should then be placed on verifying the data quality of mode 7 as soon as possible during the engineering run. This will ensure that data rates (disk write speeds and network bandwidths) do not limit choices on luminosity or trigger parameters.
3. Taking data in mode 1 and 7 simultaneously is very useful to directly compare firmware output with emulation. Firmware development takes time so you should consider starting the process now so that it will be ready for future runs after the Spring 2018 run.

4. *Are the **online and offline analysis shift staffing** plans during the commissioning period appropriate and adequate?*

An organizational scheme has been presented, which outlines the management structure of the online and offline shift organization. What is missing is a detailed task distribution in both lines. Comments for the Online have been given above. For the Offline we have the following recommendations:

1. A dedicated person needs to be assigned who will make sure a collection of the monitoring plots is available. This person does not have to make the plots but make sure they get generated/defined by the sub-detector experts. It was not mentioned but it is very useful to have a dedicated run meeting every day to discuss the run-schedule of the day *and* discuss the status of the detector and running of the past 24 hours based on the online monitoring and offline data analysis results.
2. Designate a data quality task manager who oversees the quality of the data taken and reports its status at the run meeting.
3. There are many hardware parameters for the modern electronics and the constants need to be controlled. It is recommended that a scheme be implemented that allows for verification of the correctness of these parameters.

5. *Are the **available resources** (e.g. computing manpower) sufficient to enable the implementation of the commissioning results into the production data analysis on a reasonably short time scale (weeks)?*

1. Yes. The plans described for near-term offline processing is compatible with the existing batch farm model and allocation. The disk space requirements can be met with the existing allocation given that a cleanup of existing disk space is performed. Staff have been identified for this work and a software framework is in place.
2. There are risks that could lengthen the time required for data processing: Because the uncertain data rate, and possible large simulation samples, CPU requirements could exceed the current farm capacity in future run periods. In that case, sampling and work prioritization would be required until additional compute capacity, potentially off-site, is identified.
3. It is recommended to plan now for compute and storage requirements

beyond the engineering run. This will ensure that there is sufficient lead time for procurement and integration of new storage and compute capacity. Lead time will be needed to integrate cloud- or grid-based jobs into the offline computing framework.

6. *Is the **documentation of all systems** (detector hardware, online/offline software, operating procedures, etc.) sufficiently detailed and complete to provide the required support for the shift taker and experts?*

   The documentation for the commissioning and first production experiments is extensive and includes all the required documents. The documentation is also well organized and easily accessible via web interfaces, but is so extensive it is hard to find the few essentials. We have the following recommendation:

   1. Develop 1-page summaries of the common systems that shift workers can use for guidance and place them in a central location.

7. *Is the scope of **simulation studies** that have been performed or are planned before the run period adequate to understand the expected baseline performance of the CLAS12 system?*

   Overall simulation studies are well on track for commissioning and RG-A run data taking. We have the following recommendations:

   1. Background merge mechanism needs to be implemented so that it can be compared with data and extensively tested.
   2. Perform realistic simulations for detector systems and compared with data such as trigger simulations, normalized raw TOF rates vs. lab angle etc.
   3. Verify the GEMC geometry and match to original drawings.
   4. Prepare high level representative simulation plots to be compared with data such as missing mass resolutions, pi0 mass resolutions, tracking efficiencies as function of lab angle and momentum, neutron detection

efficiencies as function of momentum, etc.

5. Implement mechanism to incorporate run-by-run calibrations into MC so that the simulations can match data in real time.
6. Consider using a post-processor for run-dependent conditions instead of including them directly into the simulation.
7. Actively seek offsite computing power for simulations.

8. *Are there **studies or tests missing** that should be specifically included in the plan to ensure the readiness for production data taking and processing?*

    1. Consider measurements that might be needed to determine the accuracy of beam polarization measurements.
    2. Study the effect of high luminosity and 250 MHz beam structure on accidentals.

APPENDIX

- Monday, 25 September 2017

  - 08:15 - 08:30 Executive Session: Charge to the Committee *15'*
    Speaker: Dr. Volker Burkert (Jefferson Lab)
  - 08:30 - 10:40 CLAS12 Ready for Science: Project Status: Online

  - 08:45 **CLAS12 Ready for Science – Project Overview** *20'*
    Speaker: Latifa Elouadrhiri (Jefferson Lab)
  - 09:05 **Collaboration organization** *15'*
    Speaker: Jerry Gilfoyle (University of Richmond)
  - 09:20 **Commissioning Plan** *25'*
    Speaker: Dr. Daniel Carman (JLab)
  - 09:45 **DAQ, Trigger and Online Computing Requierements** *30'*
    Speaker: Sergey Boyarinov (jlab)
  - 10:15 **Online Monitoring** *25'*
    Speaker: Dr. Cole Smith (JLAB)
  - 10:40 - 11:00 Break
  - 11:00 - 12:30 CLAS12 Ready for Science: Offline Simulations and Event Reconstruction

  - 11:00 **Geant 4 Simulations** *30'*
    Speaker: Dr. Maurizio Ungaro (Jefferson Lab)
  - 11:30 **Offline Reconstruction Software - CLARA and Offline Computing Requirements** *30'*
    Speaker: Veronique Ziegler (Jefferson Lab)
  - 12:00 **Particle Identification** *30'*
    Speaker: Nathan Baltzell (Jefferson Lab)
  - 12:30 - 13:30 Lunch
  - 13:30 - 15:15 CLAS12 Ready for Science: Calibrations and Detector Performance

  - 13:30 **Data Base and Computing Resources Management** *20'*
    Speaker: Dr. Harut Avagyan (Jefferson Lab)
  - 13:50 **Detector Calibration** *40'*
    Speaker: Raffaella De Vita (INFN - Genova)
  - 14:30 **Scope of CLAS12 Simulations Status and Plans** *15'*
    Speaker: Latifa Elouadrhiri (Jefferson Lab)
  - 14:45 **Performance of CLAS12 Detector System** *30'*

Speaker: Francois-Xavier Girod (JLab)

- 15:15 - 15:30 Break
- 15:30 - 17:50 Executive Session and Home Work Assignment

Tuesday, 26 September 2017

- 08:25 - 10:45 CLAS12 Ready for Science: Physics Analysis

  - 08:25 **MesonX** *25'*
    Speaker: Dr. Marco Battaglieri (INFN-GE)
  - 08:50 **Hadron Structure** *25'*
    Speaker: Ralf Gothe (University of South Carolina)
  - 09:15 **Deep Processes** *25'*
    Speaker: Dr. Franck Sabatié (CEA-Saclay)
  - 09:40 **Semi-Inclusive Physics** *25'*
    Speaker: Dr. Harut Avagyan (Jefferson Lab)
  - 10:05 **J/psi production near threshold region** *25'*
    Speaker: Mr. Joseph Newton (Old Dominion University)
  - 10:30 **Summary, Documentation & Timeline** *15'*
    Speaker: Latifa Elouadrhiri (Jefferson Lab)
- 10:45 - 11:00 Beak
- 11:00 - 14:55 Executive Session/Lunch
- 15:00 - 15:30 Closeou