

CLAS12 Run Groups

November 14, 2018



Run Group coordination

- Run Group A/First Experiment has been the main “customer” so far...
- Multiple run groups will be active shortly: need to establish good coordination between the activities of different Run Groups for an efficient usage of available resources
 - Periodic meetings of all Run Groups to review and communicate work plans
 - Participation of Run Group representatives to CLAS12 technical WG meetings (Software, CalCom, ...)
 - Common organizational structure
- First “Run Group session” on Wednesday afternoon with presentations from all Run Groups and discussion
 - to be organized regularly in the future

Common organization structure

- **Run Group Coordinator:** representative of Run Group experiments spokesperson
- **Analysis Coordinator:**
 - point of contact with Software and CalCom Groups,
 - in charge of defining the calibration, data processing and analysis work plan in collaboration with Software and CalCom Groups
 - in charge of orchestrating the work of the calibration team and chef
- **Chef:**
 - in charge of running reconstruction and analysis trains of the Jlab batch farm under the supervision of the Software Group
- **Calibration team:**
 - Calibrators expected to be active for extended time working on multiple data sets
 - Work plan defined by Analysis Coordinator, supervision provided by CalCom group

* Not yet addressing simulations since workflow is still being developed

Computing resources

- New estimates of computing resources to process CLAS12 data based on improved reconstruction processing speed:
 - Reconstruction of 2018-2020 data expected to take approximately 14 months on our current farm allocation, not including tape access or inefficiencies
 - Factor two contingency to be “used well” (currently we have a 9 months backlog)
- Disk space currently needs to be expanded to support processing of multiple data sets:
 - Estimated data volumes from 2018-2020:
 - DSTs: 340 TB (to cache/tape)
 - Skims: 78 TB (2 versions, each <10% of DSTs)
 - Calibration and data monitoring: 150 TB assuming two data sets actively being calibrated in parallel
 - File staging: 100 TB
 - Users: 50 TB
 - Total work area needed: 330 TB with respect to 130 TB currently available
 - Request for increase presented to SciComp
- Estimates being refined in view of JLAB 12 GeV software and computing review scheduled for November 27-28 2018

Computing resources for processing on 2018/2019/2020 CLAS12 data

Batch farm usage

Assumptions

Data processing will involve primarily farm18 (87 nodes, 80 job slots/node), farm16 (46 nodes, 72 job slots/node), farm14 (83 nodes, 48 job slots/node). The total number of nodes and job slots is 216 and 14256, respectively. With a fair-share of 45%, CLAS12 can utilize about 92 nodes and 6415 job slots.

Estimate for RG-A-Spring run

The following estimate assumes the data set consists of 500k raw files (1 PB) with ~45k events each and uses one file from run 4013 as example.

- **Decoding:** the time to decode one file on a clara node is 12 min. This is likely less than when running on the farm occupying a full node. Assuming a factor two larger time, the estimate for the decoding of the full data set is 8400 core*days that would be less than two days with the fair share indicated above. This does not account for the time to retrieve files from tape, which will be largely dominant. Estimating this time for the Spring data will require more studies, since the raw data files are distributed over a large number of tapes: this is partially due to the way the files were initially stored on tape but has been magnified with the migration from LTO6 to LTO8 tapes. As an example, the 2078 files of run 4013 are currently spread over 9 tape volumes. Presently, the option that is being considered to speed-up the decoding process is to process files tape-by-tape instead of run-by-run, possibly staging the output on disk to save the hipo files from the same run on tape in sequential order.
- **Reconstruction:** according to the latest benchmark, the reconstruction rate per node, computed as the average on farm14, farm16 and farm18 nodes weighted by the corresponding node number, is of the order of 92 Hz. Assuming the fair share indicated above, this results in an overall reconstruction rate of about 9 kHz. At this rate, reconstruction of the spring data (22.5 billion events) would require about 30 days. This does not include tape access, failures etc. More precise numbers will be obtained with the planned tests on Slurm: for now, assuming a safety factor of 2, we could estimate a total of 60 days for the reconstruction of the RG-A Spring data.

Estimates for subsequent runs (RG-A-Fall, RG-K, RG-B, RG-A-2019, RG-F)

Based on the beam time allocation for 2018 and 2020, we can expect to accumulate the following statistics:

- RG-A-Fall (59 days at 13 kHz trigger rate and 50% efficiency = 33 billion events)
- RG-K (18 days at 13 kHz trigger rate and 50% efficiency = 10 billion events)

https://www.jlab.org/Hall-B/secure/claschair/software/CLAS12_computing_2018-2020.pdf

“Readiness review” for pass1

- Computing resources are becoming as “precious” as beam time...
- Run Groups should assume a single pass1

- Readiness review before pass1:
 - Assessment of the status of calibration and alignment
 - Review of data processing plans
 - Workflow
 - Bookkeeping
 - Data quality monitoring
 - Review of analysis plans:
 - Skims
 - Analysis trains
 - ...
 - Review panel?

Planning the run...

Each Run Group should define:

- Experiment configuration: detector, magnets, target, ...
- DAQ & trigger requirements
- ...
- To be coordinated with technical WGs
- More central coordination desirable

“Administrative” organization:

- COO, ESAD, RSAD, ERG
- PDL assigned by Hall B
- Run Coordinator by the Run Group

Scheduling

- Summary of 2016 discussion at <https://clasweb.jlab.org/wiki/index.php/RunGroup-17-May-18-report>
- Procedure:
 - The final scheduling is done by NPES (Nuclear Physics Experiment Scheduling committee)
 - Aspects considered:
 - Experiment readiness (ERR)
 - Scientific rating, high-impact
 - Compatibility with other Halls
 - Infrastructure/setup requirements
 - Publication records
 - ...
 - Making recommendations about scheduling is the responsibility of the Coordinating Committee. The relevant section of the charter is below:
 - IX. (M) PROPOSALS FOR RUN TIME AND SCHEDULING
 - C. The Coordinating Committee is responsible for communicating recommendations for the scheduling of CLAS12 experiments to the Jefferson Lab management. These recommendations shall be based primarily on readiness to run the experiment, the natural progression of the technical equipment, the contributions of the experimenters to the Collaboration, and any priorities attached to the experiments by the PAC.
 - https://wiki.jlab.org/clas_chair/index.php/CLAS_Collaboration_Information#CLAS_Charter_Updated_October_2018

Run Groups

Hall B

Proposal	Physics	Contact	Rating	Days	Group	Equipment	Energy	Run Group	Target
E12-06-108	Hard exclusive electro-production of π^0, η	Stoler	B	80	139	RICH (1 sector) Forward tagger	11	A F. Sabatié	liquid H ₂
E12-06-108A	Exclusive N*->KY Studies with CLAS12	Carman		(60)					
E12-06-108B	Transition Form Factor of the η' Meson with CLAS12	Kunkel		(80)					
E12-06-112	Proton's quark dynamics in SIDIS pion production	Avakian	A	60					
E12-06-112A	SIDIS Λ production in target fragmentation region	Mirazita		(60)					
E12-06-112B	Collinear nucleon structure at twist-3	Pisano		(60)					
E12-06-119(a)	Deeply Virtual Compton Scattering	Sabatie	A	80					
E12-09-003	Excitation of nucleon resonances at high Q ²	Gothe	B+	40					
E12-11-005	Hadron spectroscopy with forward tagger	Battaglieri	A-	119					
E12-11-005A	Photoproduction of the very strangest baryon	Guo		(120)					
E12-12-001	Timelike Compton Scatt. & J/ ψ production in e+e	Nadel-Turonski	A-	120					
E12-12-001A	J/ ψ Photoproduction & study of LHCb pentaquarks	Stepanyan		(120)					
E12-12-007	Exclusive ϕ meson electroproduction with CLAS12	Girod	B+	60					
E12-07-104	Neutron magnetic form factor	Gilfoyle	A-	30	90	Neutron detector RICH (1 sector) Forward tagger	11	B K. Hafidi	liquid D ₂ target
E12-09-007(a)	Study of partonic distributions in SIDIS kaon production	Hafidi	A-	30					
E12-09-008	Boer-Mulders asymmetry in K SIDIS w/ H and D targets	Contalbrigo	A-	56					
E12-09-008A	Hadron production in target fragmentation region	Mirazita		(60)					
E12-09-008B	Collinear nucleon structure at twist-3	Pisano		(60)					
E12-11-003	DVCS on neutron target	Niccolai	A	90					
E12-11-003A	In medium structure functions, SRC, and the EMC effect	Hen		(90)					
E12-003B	J/Phsi production on deuterium	Ilieva	N/A	(80)					
Beam time partial sum				765 (1555)	229				

Run Groups

Hall B

E12-06-109	Longitudinal Spin Structure of the Nucleon	Kuhn	A	80	185	Polarized target RICH (1 sector) Forward tagger	11	C	NH ₃ ND ₃ S. Kuhn
E12-06-109A	DVCS on the neutron with polarized deuterium target	Niccolai		(60)					
E12-06-119(b)	DVCS on longitudinally polarized proton target	Sabatie	A	120					
E12-07-107	Spin-Orbit Correl. with Longitudinally polarized target	Avakian	A-	103					
E12-09-007(b)	Study of partonic distributions using SIDIS K production	Hafidi	A-	80					
E12-09-009	Spin-Orbit correlations in K production w/ pol. targets	Avakian	B+	103					
E12-06-106	Color transparency in exclusive vector meson production	Hafidi	B+	60	60		11	D	
E12-06-117	Quark propagation and hadron formation	Brooks	A-	60	60		11	E	Nuclear
E12-06-113	Free Neutron structure at large x	Bueltman	A	42	42	Radial TPC	11	F	Gas D ₂
E12-14-001	EMC effect in spin structure functions	Brooks	B+	55	55	Pol. LiH target	11	G	LiH
TOTAL CLAS12 run time (approved experiments)				1466 (2118)	631				

Proposal	Physics	Contact	Rating	Days	Group	Equipment	Energy	Group	Target
C12-11-111	SIDIS on transverse polarized target	Contalbrigo	A	110	110	Transverse target	11	H	HD
C12-12-009	Transversity w/ di-hadron on transverse target	Avakian	A	110					
C12-12-010	DVCS with transverse polarized target in CLAS12	Elouadrhiri	A	110					
All CLAS12 transverse target proposals				330	110				
E12-11-006	Heavy Photon Search at Jefferson Lab (HPS)	Jaros	A	180	180	Setup in alcove	2.2, 6.6	I	Nuclear
E12-11-106	High Precision Measurement of the Proton Charge Radius	Gasparian	A	15	15	Primex	1.1, 2.2	J	H2 gas
Beam time request from CLAS12 C1 experiments + non-CLAS12 experiments				525	305				
Beam time from approved CLAS12 experiments (from previous table)				1466 (2118)	631				
Beam time for Hall B experiments table 1 + table 2 (incl. 110 days of C1 approved exp.)				1991 (2643)	936				

Run Groups

Hall B

Proposal	Physics	Contact	Rating	Days	Group	Equipment	Energy	Group	Target
E12-16-010	A search for Hybrid Baryons in Hall B with CLAS12	D'Angelo	A-	100	100	Forward Tagger	6.6, 8.8	K Confinement & Strong QCD	IH2
E12-16-010A	Nucleon Resonances in exc. KY electroproduction	Carman	A-	(100)					
E12-16-010B	DVCS with CLAS12 at 6.6 and 8.8 GeV	Elouadrhiri	A-	(100)					
Total Beam time of Run Group K				100 (300)	100				
E12-17-012	Partonic Structure of Light Nuclei	Meziani	A-	(35)	55	ALERT detector	2.2, 11	L	D ₂ , ⁴ He
E12-17-012A	Tagged EMC measurements on Light Nuclei	Dupre	A-	(45)					
E12-17-012B	Spectator-Tagged DVCS on Light Nuclei	Armstrong	A-	(45)					
E12-17-012C	Other Physics Opportunities w/ ALERT	Hafidi	A-	55					
Total Beam time Run Group L				55 (180)	55				
E12-17-006		Hen		45	45	Special target		M	Nuclear
E12-17-006A		Hen							
Total beam time of Run Group M				45	45				
Beam time of approved & C1 approved CLAS12 experiments from table 1 + table 2				1991 (2763)	936				
Beam time for Hall B experiments table 1 + table 2 + table 3				2191 (3288)	1136				

Proposal Count	Experiment Days	Run Groups	RG days
43	3288	13	1136