GEMC

A database-driven Monte Carlo simulation program

Homepage clas12Tags

CLAS Collaboration Meeting

Jefferson Lab, Newport News, VA, USA

Nov 17-20 2026

GEMC updates and next steps

Distributed computing, and GEMC3

Jefferson Lab

U.S. Department of ENERGY

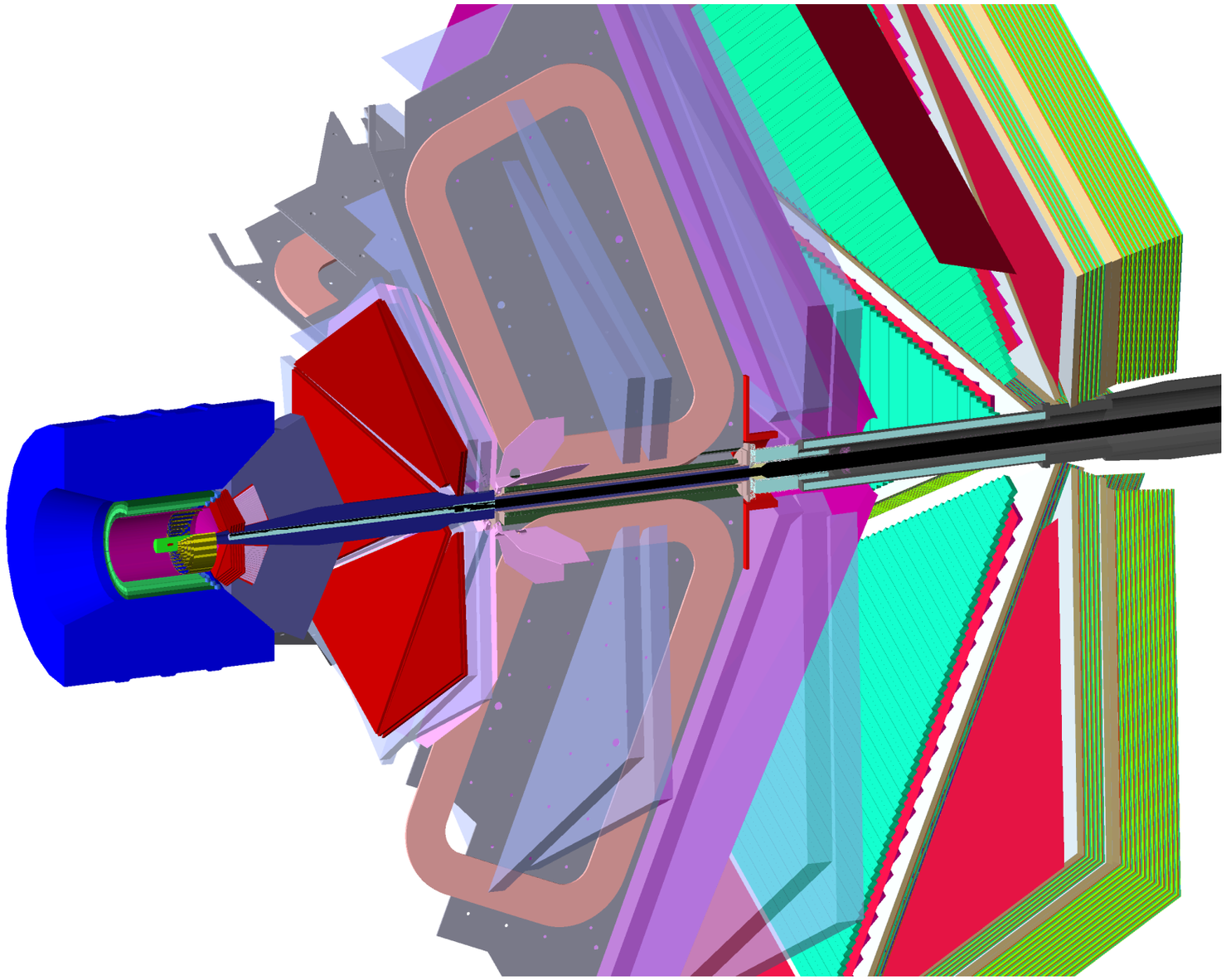
Maurizio Ungaro,
Jefferson Lab

GEMC Updates

GEMC 6.1 (August 7)

CODE

- Real Run Numbers: Fixed SQLITE bug that did not pick the correct fallback run
- Corrected ATOF and ALERT helium material definitions and time digitizations



GEOMETRY

- Updated and improved DDVCS geometry digitization variation for high luminosity proposal
- Added BAND to Real Run Number (RRR)
- Added RGK 2023/24 Targets to RRR

GEMC Updates

GEMC 6.2 (Nov 17)

CODE

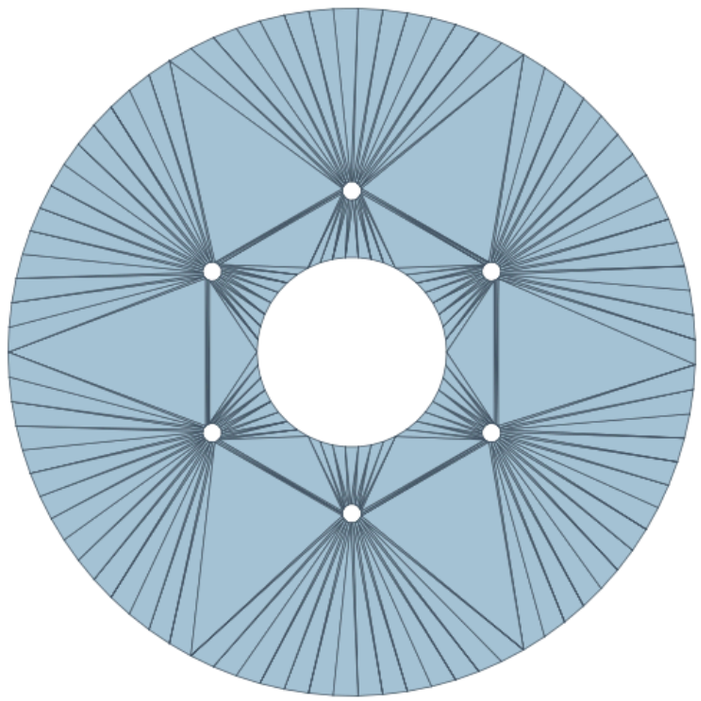
- Fixed EXEC_MACRO so command-line macros are executed in both GUI and batch modes
- Turn Off FT calorimeter and hodoscope sensitivity for FT-off configurations
- Consolidated RGD / RGE gcards for RRR
- RGE/RGD double targets accounted for in background files handling

GEOMETRY

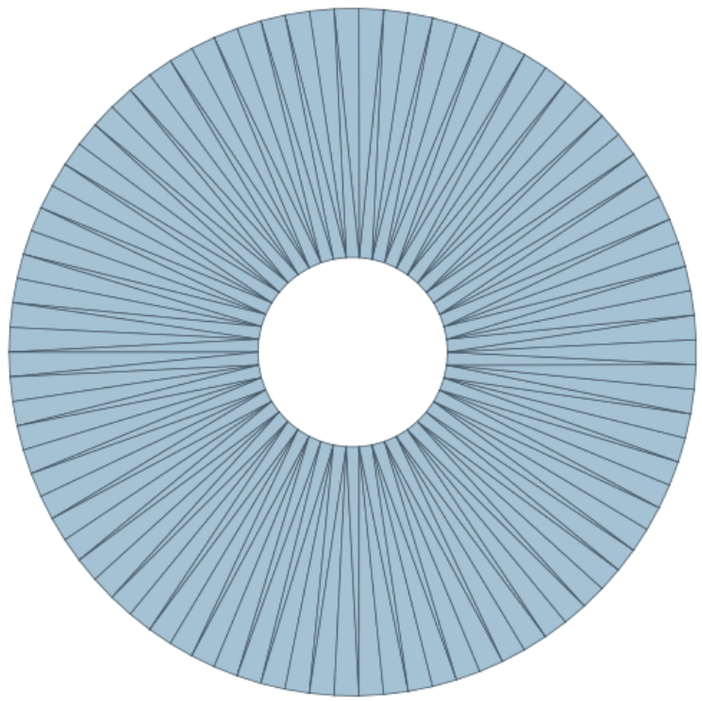
- Added the DDVCS 15deg variation
- Added the elmo7 high-luminosity beamline variation
- **Several CAD repairs and optimizations**

CAD repairs: improving simulation geometry

- Several CLAS12 volumes converted from the STEP engineering models to STL using FreeCAD
- Incorrect tessellation: several Geant4 warnings for overlaps and imperfections. Affects transportation.
- Defects present even when holes were closed or edges were smoothed in the original STEP files
- Use of agents (Codex, Claude) resulted in a significantly improved optimization of the tessellation: fewer triangles, faster speed, zero Geant4 warnings.
- This was took several weeks of agent guidance and geometry/simulations checks. It includes things like bolt holes fills and edges improvements and pre-processing using pymeshlab



Original — Separate parts, original axial spacing
3,772 triangles



Optimized — Separate parts, original axial spacing
1,064 triangles

CAD repairs: LTCC

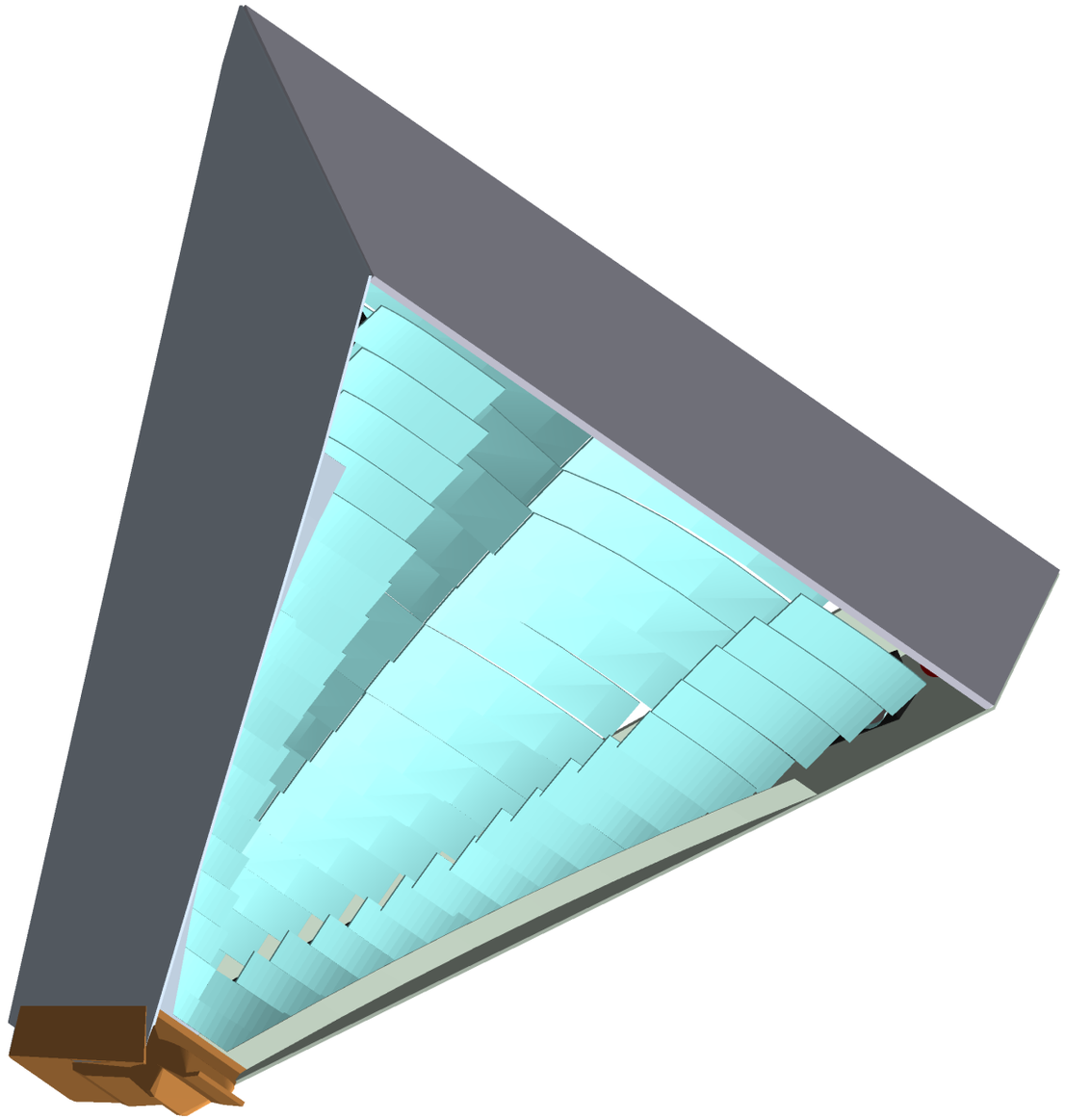
18 meshes, most could not be used because of overlaps

Aluminum Plates and Wall merged in single volumes

All nose parts merged

WC simplified

Quantity	Original	Full port
STL source meshes	18	7
Frame triangles	112,640	1,660
Nose triangles	12,012	1,986
Winston-cone triangles	9,966	4,500
All triangles	134,618	8,146
STL bytes	6,732,412	407,888
Frame/nose placements	90	24
TEXT geometry rows	2,438	2,377
Component rows	668	662



CAD repairs: LTCC

Initialization (seconds)

Measurement	Original	Full Port
CPU Time	1.87	0.68
CAD Only CPU Time	1.3	0.098

EM, Optical, Hadronic Transport (seconds, 1,000 events)

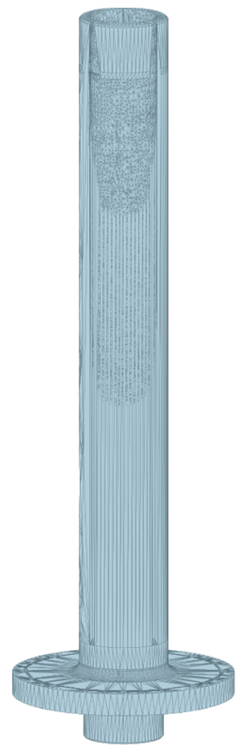
Workload	Original	Full Port	Gain
Electrons + Photons	26.1	45.1	0.6
Electrons	4.7	3.7	1.3
Protons	1.0	1.5	0.7
Pions	1.1	1.4	0.7
Neutrons	0.6	0.9	0.7

The increased CPU time is consistent with more complete transport in the optimized geometry

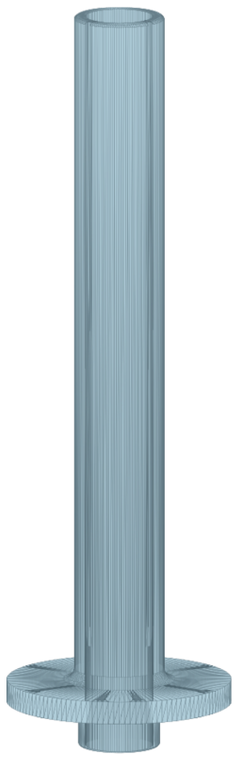
Lower CPU time in the defective original model reflects incorrect or incomplete transport

CAD repairs: FT

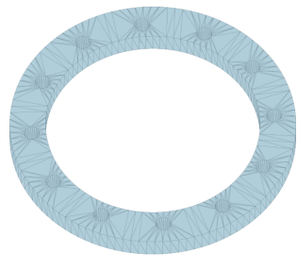
FT inner shield: original
44,122 triangles



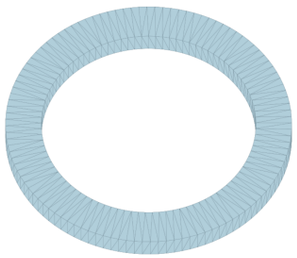
FT inner shield: optimized
13,500 triangles



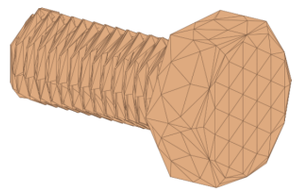
FT outer flange: original
2,132 triangles



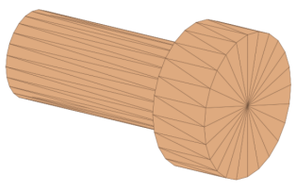
FT outer flange: optimized
836 triangles



Bolt 4: original
2,224 triangles



Bolt 4: smooth
192 triangles



CAD repairs: Beamline

- FT geometry cleanup: renamed/optimized outer flange, inner shield, bolts, mounts, adjusters, all synced CAD → runtime → run-dependent DB
- Moller cone and downstream beam tube: hole repairs, socket relief, mesh reduction, joint clearances across all run placements
- All 17 beamline configurations verified; tracking benchmarked before/after each optimization
- CAD optimization note extended all week with mesh-quality, tracking benchmarks, and reproducible evidence

CAD repairs: HTCC, DC

- Fixed and updated production cuts across 50 GEMC dev cards (all 46 volumes resolve)
- Torus model improvement meeting with colleagues
- CAD study archives organized and hash-verified

CAD repairs: magnets and beamline

- Big CAD mesh-sharing pass on magnets: vacuum jackets, DC supports/membranes, hub shields, torus coils — large facet reductions
- Fixed and updated production cuts across 50 GEMC dev cards (all 46 volumes resolve)
- CAD study archives organized and hash-verified

Using Meson

Easier, Modern, Flexible installation. Long term sustainability.

- With GEMC 6.2 released, GEMC now uses the **`meson build system`**.

- Added build options to reuse installed CCDB, HIPO, clas12-cmag libraries and external magfield location
- gemc/meson does not install the subproject libraries anymore, unless requested (avoids conflicts)
- Added workflows that build and publish multi-architecture container images and binary tarballs, then test the tarballs in minimal supported environments
- Added automated deployment, binary-tarball validation, and rolling reference-HIPO generation.

Done by Continuous Integration

N. Events	N tracks	Total Time	Time / event
200	01	15.1917	.0759
200	02	13.6454	.0682
200	03	34.6691	.1733
200	05	88.6077	.4430
200	10	155.615	.7780
200	15	265.629	1.3281
200	20	356.905	1.7845
200	clasdis_acc	127.355	.6367
200	clasdis_all	144.584	.7229
200	clasdis_all_no_int	98.2024	.4910
200	clasdis_all_savemothers	97.0608	.4853

Where does simulation spend its time?

- New Callgrinder GitHub Action published: automated profiling for GEMC2, GEMC3, Geant4
- Reports separate detector digitization and other major processing costs.

rga

Program totals: **178,022.3 Mcycles** (CEst), 144,227.1 Minstr (Ir). Percentages describe estimated CPU cycles, not elapsed time.

Both tables use the same columns. Categories group the configured entry routines; the top table lists individual routines. **% of run** includes callees and overlaps, so it must not be added. **Self %** counts only direct work; overlapping category patterns can repeat that work. Names in parentheses identify the nearest project callers, walking upstream past runtime functions such as malloc, strtod, and C++ stream helpers. Up to three distinct function names are shown; +N more indicates additional names, available in JSON with full signatures and hop distances. Project ownership is inferred from runtime names or configured with project_callers. **Calls** sums recorded incoming calls, including recursion; **Direct callers** lists each caller with its recorded calls. — means no match or unavailable data.

#	Category	Routine / entry symbol(s)	Inclusive (Mcycles)	Self (Mcycles)	% of run	Self %	Calls	Direct callers (calls)
1	Track swimming (field propagation)	`G4PropagatorInField::ComputeStep` (`G4Transportation::AlongStepGetPhysicalInteractionLength`)	23,350.5	1,236.9	13.12%	0.69%	1,519,162	`G4Transportation::AlongStepGetPhysicalInteractionLength(G4Track const&, double, double, d…` (1,519,162)
2	Field evaluation: GField_Clas12BinFactory	`GField_Clas12BinFactory::GetFieldValue` (`G4DormandPrince745::Stepper` , `G4EquationOfMotion::RightHandSide` , `G4IntegrationDriver<G4MagIntegratorStepper>::AccurateAdvance` ; +2 more)	18,362.3	1,292.2	10.31%	0.73%	25,115,210	`non-virtual thunk to GField_Clas12BinFactory::GetFieldValue(double const*, double*) const` (12,537,100) `G4DormandPrince745::Stepper(double const*, double const*, double, double*, double*)` (10,758,822) `G4EquationOfMotion::RightHandSide(double const*, double*) const` (1,551,713) `G4IntegrationDriver<G4MagIntegratorStepper>::AccurateAdvance(G4FieldTrack&, double, doubl…` (226,565) `GMagneto::magnetic_field_magnitude_tesla(std::shared_ptr<GField> const&, CLHEP::Hep3Vecto…` (41,010)
3	Hit collection (sensitive detector)	`GSensitiveDetector::ProcessHits` (`G4SteppingManager::Stepping` , `GSensitiveDetector::ProcessHits`)	14,185.7	482.9	7.97%	0.27%	1,608,482	`G4SteppingManager::Stepping()` (804,241) `non-virtual thunk to GSensitiveDetector::ProcessHits(G4Step*, G4TouchableHistory*)` (804,241)
4	Digitization: DC_digitization	`DC_digitization::digitizeHit` (`GEventAction::EndOfEventAction`)	5,897.6	23.2	3.31%	0.01%	4,584	`GEventAction::EndOfEventAction(G4Event const*)` (4,584)

GEMC v3

GEMC v3 validation

Simulation on Open Science Grid: updates

- Migration from `/cvmfs/oasis.opensciencegrid.org/jlab/` to `jlab.opensciencegrid.org` (initially problematic)

- Added portal LUND requirement: path must start with `/volatile/clas12/` - (had many /work or /home etc submissions)

- ***MCGEN*** updated to 3.22:
  - **EAVM_RGD** v0.4: Vector Meson electroproduction on RGD targets
  - **stringspinner** 2.0.0: spin effects addition to PYTHIA 8

- HTCondor transfer issue - now prevented by checking hipo file integrity with `hipo-utils -test`
- Prevention pelican bug (half transferred files)

- Portal improvements:
  - Improved priority algorithms, accounts for both running and idle jobs
  - Estimates for job completions
  - Estimated number of events / day

Simulation on Open Science Grid: updates

user	queued	on OSG	n. jobs (est. on queue)	done	running	idle	estimated time remaining **
carman	0	2	20,000 (0)	19,891	61	0	left: 109 — 0.0 days
storyf	16	4	40,000 (160,000)	23,394	2,279	14,296	left: 176,606 — 6.9 days
mkerr	0	3	30,000 (0)	18,038	1,992	9,935	left: 11,962 — 1.6 days
ishige	7	4	40,000 (70,000)	17,745	2,800	19,446	left: 92,255 — 5.7 days
asliacar	0	1	10,000 (0)	9,999	0	0	left: 1 — 0.0 days
naya	0	3	30,000 (0)	25,429	566	3,977	left: 4,571 — 0.6 days
ahmede	0	1	100 (0)	83	2	0	left: 17 — 0.0 days
thayward	0	2	20,000 (0)	18,685	629	674	left: 1,315 — 0.3 days
lanza	0	2	4,000 (0)	2,232	709	1,058	left: 1,768 — 0.3 days
aosmond	0	4	40,000 (0)	18,209	2,187	19,549	left: 21,791 — 2.1 days
totals	23	26	234,100 (230,000)	153,705	11,225	68,935	Jobs left: 310,395 — 6.9 days

Color coded: **red** users will take a long time

Completion time includes wait times based on priority

** User estimates combine processing time **tproc** and queue time **tqueue**.
tproc uses the completion rate and running jobs per submission.
tqueue uses queue position and an estimated submissions rate. A 'to osg' estimate shows the first queued submission's arrival only.
Total time left is at least the longest user completion estimate and includes the last queued submission's arrival and processing. Queue waiting and processing can overlap.
Note: Total time left includes 10,000 jobs for each queued submission.

Average concurrent jobs per OSG submission	432 jobs
Completion Rate (latest 24-hour interval)	66,350.0 jobs / day
Estimated queue to OSG rate (10,000 jobs / submission)	6.6 submissions / day

Completions / day from hourly report from HTCondor

Events in the last 10 days assumes 5,000 events / job

Completed events (millions) per 24 hours — last 10 days — assuming 5,000 events / job
Dates mark interval ends at 18:46:20 UTC.

Nov 7	Nov 8	Nov 9	Nov 10	Nov 11	Nov 12	Nov 13	Nov 14	Nov 15	Nov 16
129.5	107.1	147.0	150.5	124.2	115.7	113.2	127.8	173.0	331.8

Adding NERSC to our Simulation Resources Pool

Using Perlmutter

- ✓ User Account Granted
- ✓ NERSC AY2026 allocation requested: 250 hours of a 256 cores node until January 2027

  for exploring feasibility of using NERSC for CLAS12 simulations
- ✓ Project **`m5464`** granted. 20 TB disk space.
- ✓ CVMFS JLab access point requested and granted, effective Nov 11
- ✓ Pelican transfers to/from JLab successful: token created on the fly inside the image
- ✓ Using Our Docker image, stored on Permutter.
- ✓ Complete simulation + reconstruction workflow using either LUND / MCGen generators
- ✓ Starting trial on a full node:
  - ✓ **taskfarmer (NERSC in house scheduler)**
  - ✓ **mcdj**

If we can run at NERSC, we will ask for resources for 2027.

Plan: jobs will be allocated from the portal to NERSC. Mechanism transparent to users.

Summary, Outlook