OSG/NP

NP and prospects for OSG funding

OSG face to face council meeting
2017/10/03-04

Community past activities & outcomes

NP/Theory

- No recent running campaign but harvesting the science enabled by the OSG - latest in the 10 M hours
    - Two key publications in the community: [A data-driven analysis for the temperature dependence of the heavy quark diffusion coefficient in relativistic heavy-ion collisions](#) and [Investigating the collision energy dependence of ηη/s in RHIC beam energy scan us statistics](#)
    - 5 conference proceedings ([1594942](#), [1594915](#), [1591731](#), [1591644](#), [1591631](#))

    - As customary from NP, acknowledgement mentions the OSG explicitly.

- Unique results, High visibility for the OSG - results would not have been possible without it (and other Theorists noticed)

GlueX

- OSG GlueX experience presented [at JLab in July](#) (Computing Round Table - strategy) as well as an [OSG overview](#) (Frank)
- Usage 50 Mcore hours / year for simulations fits well within the OSG
- First usage in 2012 (glideinWMS), regain of interest in 2016
  - Visit @ a NP Computing Workshop n March, 2016
  - Discussed a GWMS submit host for JLab users - OSG setup/tested, locally supported
- GlueX now investigating the use of Singularity containers - usage possibly growing (more users) - good news and initial pays off
- Lab IT management supports JLab on OSG - some user support issues
  - Started help pages, quick start user guide and HOWTOs
- The "reboot" of GlueX on OSG was a good experience and OSG back on the table as a core strategic resource and approach
  - Resource providers: UConn, NWU, FIU, FSU, CMU, IU, MIT?

HPC (ongoing @ Cori)

Containerized environment - small footprint (< 1.4 GB / thread). Event parallelization done from a steering script - real data production (one thread starts a local DB). Dedicated space / data buffer. 25 M hours allocation.

OSG gatekeeper - end to end fully autonomous component production workflow (yes, uses Globus). File starts from HPSS and ends to HPSS (BNL). Steady ingest of jobs.

Recent push: 800 Haswell nodes for 48 hours → 800*32*48 = 1.23 M CPU hours
- 25,600 cores is ~ x2 the # of CPUs available at the RACF @ BNL for data production
- Not only a huge boost but production size larger than at Tier0
- Slot utilization efficiency 97.6% (lower limit).

Plans include test/use of CVMFS/OASIS, IO performance (Xrootd?StashCahe?) & revisit transfers
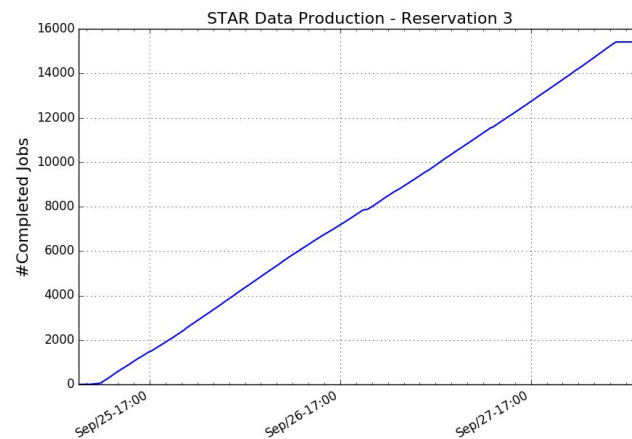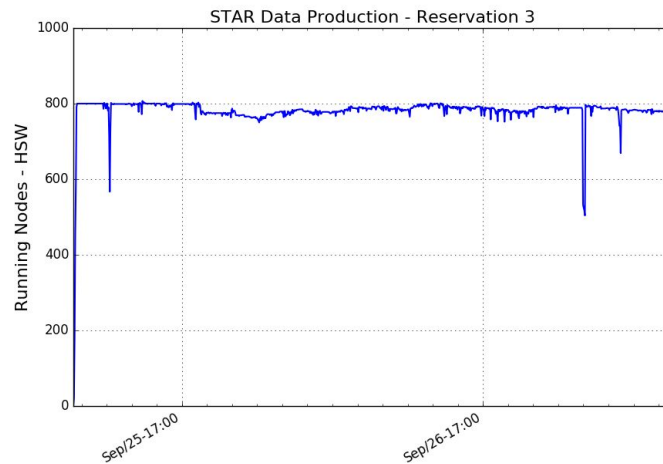
Other

- PHENIX / sPHENIX still has as plan to evaluate further use of OSG services but attention on sPHENIX at the moment
- With success story, NP/Theory likely to use OSG again (looking at HPC and parallel processing)
- HPC - Coming interest from GlueX
  - Pilot project to do production and possibly simulations on Cori, as a proof of principle for GlueX within the next 6 months
  - Opened communication within NP community

- Attempting to gather the NP communities & trying to self-organize / communicate and exchange lesson learned and best practices
  - Effort consolidation & message consistency to the agency

Proposal ...

Proposal core bullets (reminder)

1. **Foster use of OSG APIs in support of science carried at ASCR/HPC facilities**
   - Demonstrate High efficiency workflows on HPC, liaise with NERSC scientists and promote common interface, support opportunistic simulations for R&D experiments and NP/Th, export lesson learn

2. **Contribute to OSG technology & software that empower the provisioning and access of HPC resources**
   - Container, OSG related SL/RH, data transfers, etc …

3. **Support of cost effective High Performance data access solution on HPC**
   - Cost effective high performance IO layer, caching, software distribution, …

4. **Data processing intelligence – global workflow optimization**
   - Dynamic with HEP/ASCR, planner/resoner, convergence of all models from resource ownership and mix of paid/allocated/opportunistic resources

Action items were (Feb 2017)

- Agency requested to hear from BNL - heard in a few occasion + cohesive support statement from LBNL management
- Asked for a PoW with "options" - discussed a few scenari including strict minimal option (still equates to $ amount > than NP current investment in OSG)
- Agency requested investigating synergy with SCiDAC (beyond the HEP clear line of possible collaboration) - more explicitly, can buller #4 funding be "helped"

Meetings and discussions

- Plenty of EMail exchanges …
- Face to face (informal)   - June 2017 (27th?)   - reviewed the status and proposed work
- Phone (semi-formal)   - September 19th   - reviewed option and how to reach budget
- Face to face   - September 27th   - reiterated TODO bullet points

Current state

- We are continuing to demonstrate the usability of HPC for experimental real data reconstruction
  - Progress on HPC is very good (beyond all expectations) - efficiencies remain in excess of 97% 97% Eff = Slot occupancy efficiency x RunTime efficiency
  - NP/NERSC/ASCR recommend we submit further allocation requests
- STAR/ALICE considering moving $ from other program (not usable for PDSF) to complement the OSG/NP proposal (also clear HPC is a NERSC priority)
  - Community buy-in still seek: GlueX running on Cori, sPHENIX may have an interest, …
- Proposal needs some appeal to ASCR (still) to complement funding
  - NP Cori success seen as a cornerstone for further ASCR appeal - rebranding highly suggested
  - PM wants to push success story and appeal to ASCR

Timeline ?? Unclear but should be January 2018 (TBC)