

Present : Alexei, Danila, Jack, Kaushik, Matteo, Mikhail, Pavlo, Ruslan, Sergey, Shantenu

Discussion topics :

Decisions from TIM@UTA

- meeting's slot @3pm Eastern
- tasks assignment
- weekly progress reports
- 2018 TIMs : BNL April 25-27, Sep 17-19 @ ORNL
 - BNL TIM request extent on cost
- MAGIC meetings (The 1st Wed, monthly) - next one Feb 7th
Shantenu will send info to the list
- slides/memo (progress report and accomplishments) to BH & RC
- Lol of new project
- ATLAS SW&C week @DESY and Jun technical discussion at CERN
- meeting with Google on Mon Feb 5th will start at 15:00 Eastern, BigPanDA weekly will start at 14:30 E

Danila : weekly reports about ATLAS @Titan (part of hot topics), Kaushik : it would be a part of operations report

Danila : Mon Feb 5th. ATLAS@Titan

Pavlo : Mon Feb 5th : PanDA instance @OLCF

Sergey: Mon Feb 5th : containers

Danila : Mon Feb 12th : Harvester for ATLAS status @Titan

Danila. ATLAS production on Titan. Jan 2018 : 25.7M (ALCC : 9.9M, Backfill : 15.8M)

Jack : ATLAS plans for new ALCC proposal (deadline Feb 1st), Kaushik will follow it with US ATLAS folks

Jack : it is essential to have highlights from ATLAS to acknowledge TITAN contribution to ATLAS

Shantenu : Sarp initiated discussion of performance section of the paper. We should have a brief discussion of the paper at every meeting, however brief.

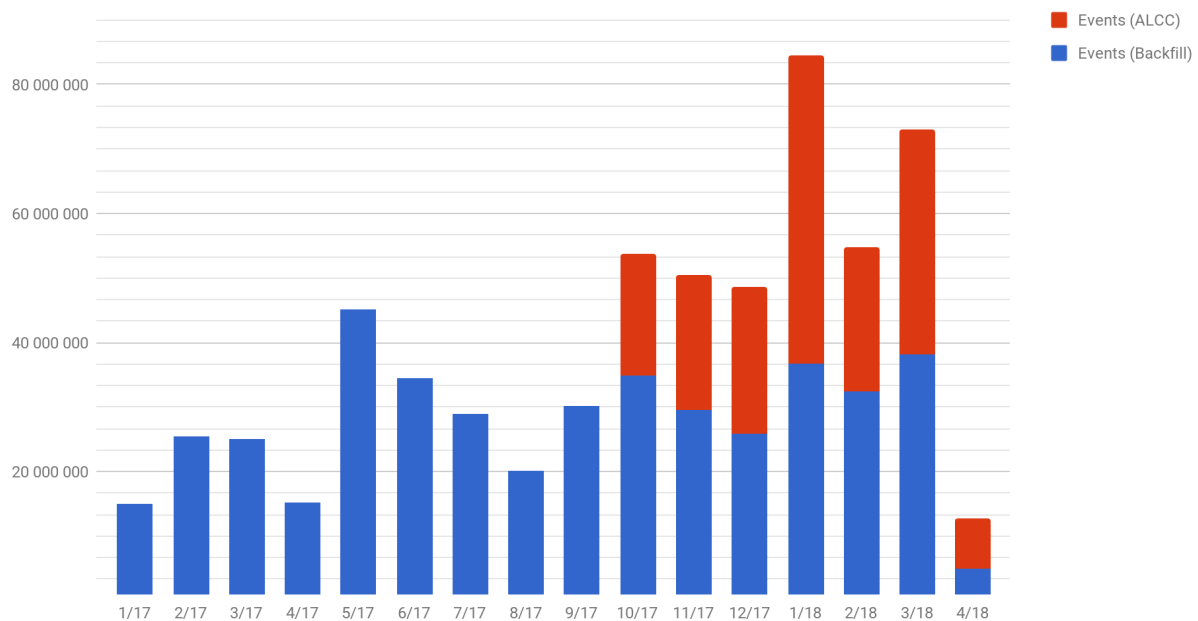
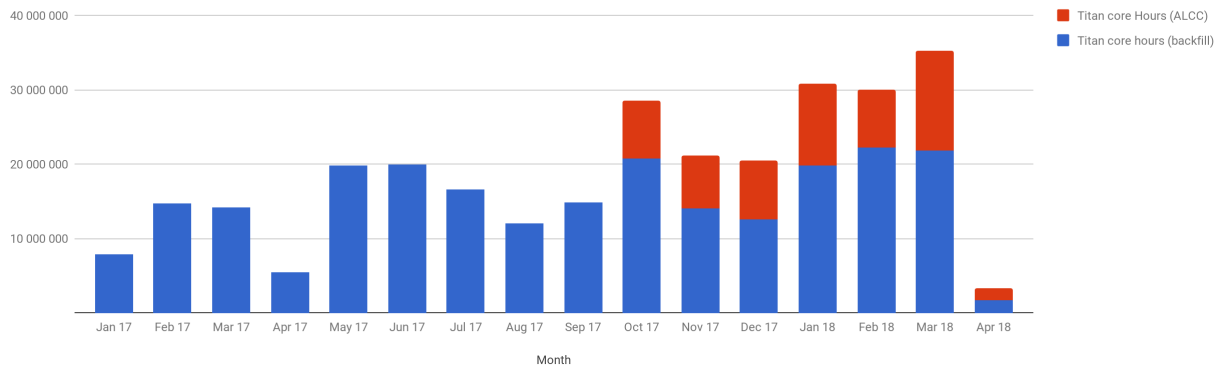
Projects, achievements and goals in 2018

ATLAS

achievements:

- ATLAS Data processing in production mode at OLCF
- 80M Titan core hours ALCC award
- By using backfill resources more than 198M Titan core hours were consumed starting from January 2017
- 36M Titan core hours were consumed from ALCC as of today
- About 500M simulated events was delivered for ATLAS
- Important studies for efficient execution of high-throughput payloads on HPC systems was required during increasing of loading. This studies are used during integration of other LCFs to ATLAS production chain and induced development of next generation of PanDA software

Titan core hours



- ATLAS software was successfully ported to Titan and compiled locally from scratch. This builds a foundation for porting ATLAS software to Power9 architecture of Summit supercomputer

Goals:

- Increase of deliverables for ATLAS with ongoing reducing of operations effort
- Migrate to next generation of PanDA software (Harvester, Pilot 2) to increase automation level and get rid of shortcomings of the current architecture
- Port ATLAS software to Summit supercomputer at OLCF

Containers on Titan

Achievements:

- Developed recipes for installation of ATLAS software releases in Singularity containers
- Created customized Singularity containers to run ATLAS detector simulations on Titan
- Studied performance and I/O properties of ATLAS simulations running in Singularity containers on Titan. Concluded that use of containers shows both performance and I/O improvements compared to running ATLAS simulations with software installed on Titan's shared file system.

Goals:

- Perform scalability tests of ATLAS simulations in Singularity containers on Titan (in progress)
- Test use of containers with current PanDA job submission system on Titan
- Integrate ATLAS containers with Harvester edge service on Titan
- Test usage of containerized workloads with NGE

NGE

- Code at: <https://github.com/ATLAS-Titan/PanDA-NGE>

achievements:

- Execution up to 65K 32-core, 15-minutes long tasks on up to 131K cores on Titan.
- CPU and GPU executions on Titan.
- Submission of pilots to Titan using backfill capabilities.
- Optimization of scheduling algorithms of tasks on pilots on Titan.
- Prototype of integration with PanDA Pilot on Titan. Demo at: <https://drive.google.com/file/d/1dnxxSz26tG2qe14md6zEXUH2KKpihz1b/view?usp=sharing>

goals:

- Integration with Harvester on Titan
- Concurrent execution of tasks on CPU and GPU
- Deployment of NGE via a container
- Execution of (non-production) ATLAS and Molecular Dynamics workloads/workflows
- Support Summit-dev
- Characterization of multiple execution configurations (pilot size/duration; concurrent backfill/queues; concurrent CPU/GPU; compute- or data-heavy workflows; etc.)

LSST DESC:

achievements:

- A dedicated queue configured for LSST DESC on PanDA Server deployed on Amazon EC2 cloud
- PanDA edge services were installed on Titan and NERSC for running LSST DESC simulations
- Demonstrated LSST job submission via PanDA to Titan and Cori supercomputers

- Several Grid sites from Open Science Grid (USA) and GridPP (UK) were integrated in the setup creating a heterogeneous Grid-HPC infrastructure with PanDA as job meta-scheduler
- PanDA Pilot Factory configured to work with LSST DESC Grid sites
- Implemented automated data transfers between Grid sites

goals:

- extend infrastructure with new sites
- test user analysis for LSST/DESC on Grid and Titan

Lattice QCD (LQCD):

achievements:

- A system for LQCD jobs submission was created at Oak Ridge Leadership Facility (OLCF) using PanDA Server installed as a container in OLCF OpenShift
- A micro workflow description feature was implemented which allows to describe sequences of jobs using only PanDA tools
- Successfully tested realistic LQCD payloads

goals:

- Extend infrastructure with new Grid and HPC sites
- Develop tools for automatic data transfer between sites

nEDM:

achievements:

- a system for job submission via PanDA was set up on a front node of Titan supercomputer; it includes a dedicated PanDA queue and an edge service
- nEDM detector simulations were submitted to Titan via PanDA

goals:

- prepare and participate in future large-scale detector simulations

Computational Biology:

achievements:

- infrastructure built with PanDA instance deployed at Oak Ridge Leadership Facility (OLCF) : a PanDA queue + dedicated edge services
- GBOOST payload (GPU) example built and tested on Titan
- Client and Pilots packages provided to run it under BIF102 project
- Set of realistic tests were conducted by Biologists using provided tools (client and pilot factory)

goals:

- prepare for future data challenges

Molecular Dynamics:

achievements:

- infrastructure built with PanDA instance deployed at Oak Ridge Leadership Facility (OLCF) : a PanDA queue + dedicated edge services
- CHARMM payload (hybrid MPI/OpenMP/GPU) example built and tested on Titan

goals:

- prepare for future data challenges

IceCube:

achievements:

- infrastructure built with PanDA instance deployed at Oak Ridge Leadership Facility (OLCF) : a PanDA queue + dedicated edge services
- Benchmark application (in Singularity container) example built and tested on Titan
- Same example with data stage-in/-out to the GridFTP remote storage was tested

goals:

- Implementation of the workflow to be as close to production as possible

BlueBrainProject:

achievements:

- Pilot collaborative project PanDA@BBP
- 1st phase of 'proof-of-concept' is successfully finished
- BigPanDA panda portal is installed at BBP. Instance includes: PanDA server, set of pilots and launcher scripts, web interface.
- Diversity of resources are supported and tested via PanDA: Titan@OLCF, IBM BBP BlueGene/Q in Lugano, Intel x86-NVIDIA GPU based BBP clusters located in Geneva and Lugano, Amazon on-demand cluster.

goals:

- Pre-production is under investigation

Modeling utilization of allocation time:

achievements:

- Developed a simplified mathematical model which describes the dependence of utilization on time for allocated resources
- Developed simulator based on queueing theory, which simulates the load on the resource and traces of the execution of the jobs for a given workload (code at: <https://github.com/ATLAS-Titan/allocation-modeling>)
- Tested the developed mathematical model on the synthetic and real (partial log data from Titan) data with simulator

goals:

- Reveal the optimal set of essential job parameters which maximize the probability of complete utilization of allocated resources (i.e., optimal execution strategy)
- Estimate utilization based on provided job parameters for a given time period and workload

Mar 02, 2018 discussion

Present: Kaushik, Alexey, Alexei, Jack, Pavlo, Ruslan, Sergey, Shantenu,, Danila, Sean, Mikhail

1. ATLAS accomplishments and impact on LHC scientific program
 - a. Backfill and allocations - Danila
 - i. ALCC is a part of LHC computing planning
 - b. address summit-dev needs (ATLAS SW releases) - Sergey
2. Beyond HEP
 - a. OLCF PanDA instance - Ruslan, Jack
 - b. Diverse collection of new science use cases - Pavlo, Sergey, Jack
3. New development
 - a. NGE - Shantenu, Matteo
 - b. Containers - Sergey
 - c. Modelling - Shantenu, Mikhail, Alexey
4. Looking in the future and BigPanDA+++ (ECP, HL-LHC,...) - KD, JW, SJ, AK, All
 - a. 2019 INCITE call will emphasize data-intensive and/or learning applications.
 - b. Need to ask ECP how project will be embracing WLMS.
 - c. Current hot topic is ECP - Facilities integration plan.
 - d. Rich placed emphasis on Experimental-Observations Data (EOD). Rich is also working on Science Cloud.
 - e. Space to watch: SciDAC Partnerships Budget. Could there be a SciDAC program for EOD science?
 - f. May be we can bridge WF and DM ?

How to organize :

1. Quadchart - TBD
2. 2 pages memo (also will be used for press-releases)
 - a. Drafts by Mon Mar 12th

Papers to be published

1. "OLCF for ATLAS"
2. Beyond HEP