Next Generation Workload Management and Analysis System For Big Data: Big PanDA

P. Svirin, S. Wilkinson, J. Kincl, R. Mashinistov, A.Merzky, D. Oleynik, S. Oral, S. Panitkin, M. Turilli,
K. De, S. Jha, A. Klimentov, J. Wells, T. Wenaus

Project Vision

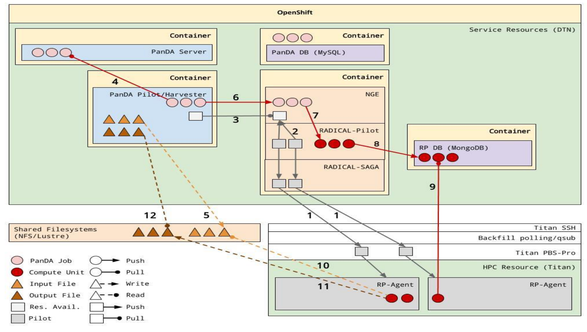
- Translate the R&D artifacts and accomplishments from the BigPanDA and AIMES projects into LCF operational advances and enhancement.
- Utilize compute cycles that would be otherwise unusable and the ability to increase the overall utilization on Titan.
- Extend and generalize the ability of using HPC for HTC for ATLAS and other communities.

Accomplishment, Impact and Future

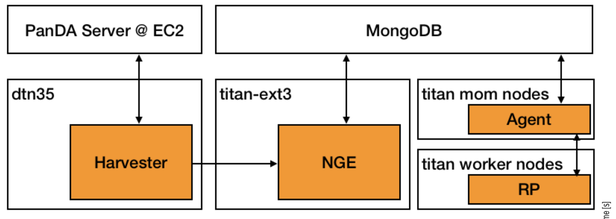
- OLCF resources are utilized using the latest PanDA software versions.
- Optimized execution for payloads on Titan using Next-Generation Executor (NGE). New payload types to be tested with Harvester/NGE integration.
- Continue to develop current software to support heterogeneous resources and diverse workflows for multiple experiments and projects beyond ATLAS.

Adaptive Execution

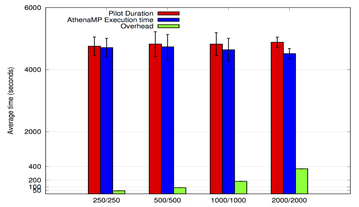
- Next-Generation executor(NGE) conforms to model of Pilot Abstraction: https://arxiv.org/abs/1508.04180
- Flexible Execution: No assumption about number of events per simulation. Better utilization via multiple execution generations
- Investigate the advantages of scale, flexible execution and interoperability outside production constraints on HPC resources.
- Enables PanDA to be independent from: Workload heterogeneity:
  - No need of tailoring MPI scripts to the specifics of each type of workload.
  - Resource heterogeneity: No need of implementing tailored support for each HPC machine flavour
- Modeling probability of allocation utilization (ALCC) as a function of execution strategy.
- Supporting backfill/regular queueing on Titan; concurrent execution of workloads on the same pilot; distribution of jobs on Titan and Summit supercomputers.
- Future plans:
  - Extending profiling and analytics capabilities of NGE stack to Harvester stack executions.
  - Port design features to Yoda.



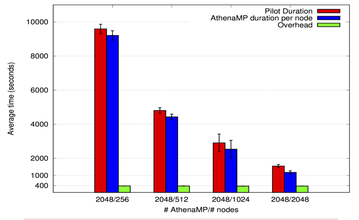
PanDA-NGE Integration: Architecture



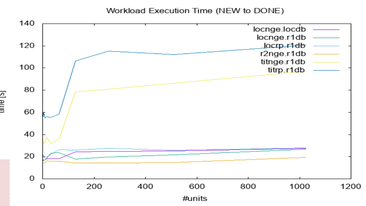
PanDA-NGE Integration: deployment on Titan



Weak scaling behaviour using NGE with AthenaMP workload



Strong scaling behaviour for NGE with AthenaMP workload using 100 events per AthenaMP task.



NGE performance : null workload execution

Operations

Achievements:

- A VO-independent PanDA Server has been set up in Amazon EC2. It is used for non-ATLAS projects that utilise heterogeneous Grid and HPC resources.
- In March 2017 a new PanDA server instance has been set up at ORNL to serve various experiments. This installation is using a container cluster management and orchestration system, Red Hat OpenShift Origin.

nEDM Precision measurements of the properties of the neutron present an opportunity to search for violations of fundamental symmetries and to make critical tests of the validity of the Standard Model of electroweak interactions. The goal of the nEDM experiment at the Fundamental Neutron Physics Beamline at the Spallation Neutron Source (ORNL) is to further improve the precision of this measurement by a factor of 100. nEDM experiment requires detailed simulation of the detector.

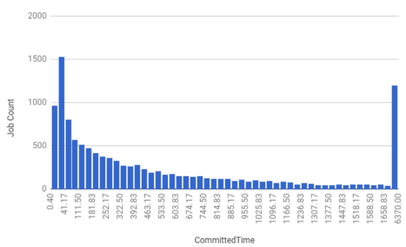
Achievements:

- Detailed nEDM detector simulations were executed on Titan via PanDA Server at OLCF
- Currently nEDM prepares for a future computational campaign

IceCube IceCube is the world's largest neutrino detector, encompassing a cubic kilometer of ice at the South Pole. It searches for neutrinos from the most violent astrophysical sources: exploding stars, gamma-ray bursts, and cataclysmic phenomena involving black holes and neutron stars. The IceCube telescope is a powerful tool to search for dark matter and could reveal the physical processes associated with the enigmatic origin of the highest energy particles in nature. IceCube studies the neutrinos at energies that far exceed those produced by accelerator beams. In order to understand experimental details of neutrino-event observations, large-scale detector simulations are needed.

Achievements:

- IceCube payloads were tested in Singularity containers on Titan using job shaping.
- Input and output data exchange via GridFTP tested for IceCube jobs on Titan.



Walltime distribution for IceCube payloads on Titan

PMI Current development efforts of the project are focused on advancing protein simulations where chemistry is linked with large-scale protein conformational dynamics and free energy computations. The application involves, but not limited to, kinases, ATP hydrolases, and DNA repair enzymes. These biomolecules present a unique set of challenges that present experimental techniques have difficulty addressing, while theory can provide detailed understanding at the atomic level.

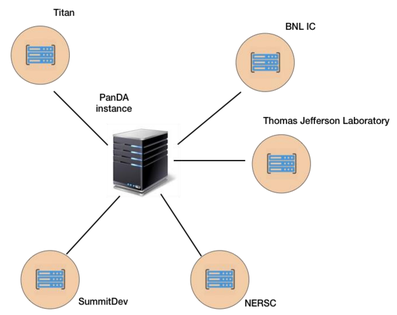
Achievements:

- CHARMM payload (hybrid MPI/OpenMP/GPU) example built and executed on Titan
- Depending on the type of projects, payloads can expand beyond 500 nodes on Titan; currently, it uses 60-124 nodes for each project
- Payloads have been also tested with Harvester/NGE integration

USQCD Lattice QCD (LQCD) is a well-established non-perturbative approach to solving the quantum chromodynamics theory of quarks and gluons. Current LQCD payloads can be characterized as massively parallel, occupying thousands of nodes on leadership-class supercomputers. It is understood that future LQCD calculations will require exascale computing capacities and new a workload management system in order to manage throughput efficiently.

Achievements:

- Large LQCD payloads have been successfully tested with PanDA on Titan together with automated data transfer.
- New kinds of payloads will be available for Summit in terms of Early Science Program.



LQCD Future computing infrastructure

- Data to be stored at OLCF and will be transferred to and from execution sites via Globus Online. Automated data transfers from OLCF have been tested with Globus Online modules for Harvester

LSST A goal of LSST (Large Synoptic Survey Telescope) project is to conduct a 10-year survey of the sky that is expected to deliver 200 petabytes of data after it begins full science operations in 2022. The project will address some of the most pressing questions about the structure and evolution of the universe and the objects in it. It will require a large amount of simulations, which model the atmosphere, optics and camera to understand the collected data.

Achievements:

- Phosim simulations were run on Titan using PanDA Server at OLCF
- Phosim long-running jobs required exploration of checkpointing capabilities on Titan
- Grid environment with 36 endpoints for LSST was configured and tested with PanDA Server in Amazon Cloud