MAGIC Meeting Minutes
April 7, 2017

Attendees

Brian Bockelman	OSG
Dan Gunter	LBL
Shantenu Jha	Rutgers U.
Padma Krishnaswami	FCC
Miron Livny	U. Wisconsin
Paul Love	NCO
David Martin	Northwestern U.
Bogdan Mihaila	NSF/Physics
Grant Miller	NCO
Matteo Turilli	Rutgers U.

Action Items

Proceedings
This MAGIC meeting was coordinated by Grant Miller of the NCO. Shantenu Jha and Matteo Turilli provided briefings on Converging HTC-HPC: Two Case Studies that addressed Big Panda services.

Converging HTC-HPC: Two Case Studies: Shantenu Jha and Matteo Turilli
High Throughput computing comprises multiple tasks. The workload is for a single application, a single workflow or multiple workflows as part of a computational campaign. It is generally characterized by the number of tasks executed. It might also focus on minimizing the time to execute the task and/or to maximize the number of tasks executed. High throughput computing depends on:

- Number of tasks
- Number of resources
- Size of each task
- Duration of each task
- Number of jobs

High throughput computing requires abstraction and models at multiple levels:

- Application framework
- Workload management
- Task run-time
- Resource access layer

A Pilot-Abstraction model provides a scheduling overlay, decouples workload from resource management, enables fine-grained “slicing and dicing” of resources, provides tighter temporal control and builds a higher-level framework without explicit resource management. An AIMES Execution model provides dynamic integration of workload and resource information. Execution strategy is a temporally ordered set of decisions that need to be made to execute a

given workload. Implementing AIMES model and pilots supports uniform execution models over XSEDE and OSG. XSEDE provides resources on 10 supercomputers, 10 storage facilities, shared allocations process and user support, and is tailored toward grand challenge applications. OSG distributed cyberinfrastructure supports 100 sites, 40,000 cores and 1 PB of storage. It uses single core machines loosely coupled tailored toward high throughput.

Titan provides next-generation workflow management for the High Energy Physics (HEP) community. In 10 years the data from the LHC luminosity increase will increase by a factor of 10. There will be more complex events and more computational capacity. Production and Distributed Analysis (PanDA) will provide subsystems for the HEP workflows and workloads and high throughput computing on Grid infrastructures. It is a task-based and job-based execution mode to support single researchers, research groups, and ATLAS production workflows. ATLAS detector simulation is integrated with Titan which contributes a large fraction of the computing for simulation.
BigPanDA utilizes backfill slots significantly to improve the overall utilization on Titan. PanDA deployment increases the efficiency of computational resources, and avoids workload and resource heterogeneity (simulations are run on one machine instead of many machines). The PanDA Next Generation Executor (NGE) exposes capabilities and availability acquired from the PanDA pilots.

Big PanDA is also use for Biomolecular simulations supporting the NCI-DOE collaboration on large-scale computational cancer science. Big PanDA addresses the gap between weak scaling for larger biological systems and strong scaling for long time scaling. PanDA provides virtually linear scaling for the biomolecular simulations.

For the complete briefing please see the MAGIC Website at:
https://www.nitrd.gov/nitrdgroups/index.php?title=Middleware_And_Grid_Interagency_Coordination_(MAGIC)#title under April 2017 meetings.

NCO Report
The NCO is organizing a new group on the National Broadband Research Agenda to address research needs and research coordination on Broadband wireless networking supporting access and services to communities and hard to reach constituencies (rural areas, Indian reservations, prison populations,…). Their first meeting is April 17 at the NSF. . If you would like to participate, please contact Grant Miller at: miller@nitrd.gov.

Meetings
FuturePlatform workshop. The objective of the workshop is to prepare a roadmap charting out research challenges that need to be overcome to build a future platform for science.
The workshop will be held on April 4–5, 2017 at the Hyatt Regency Crystal City

Next MAGIC Meeting
May 3, 12:00-2:00 Eastern, NSF, Room TBD

Please note that the regular time for the MAGIC meetings has permanently been changed to 12;00-2:00 Eastern.