MAGIC Meeting Minutes
March 4, 2015

Attendees

Rich Carlson	DOE
Kaushik De	UTA
Ewa Deelman	USC
Shantenu Ja	Rutgers Un.
Dan Katz	NSF
Grant Miller	NCO
Lavanya Ramakrishnan	LBL
David Schissal	General Atomics
Alan Sill	TTU
Derek Simmel	PSC
Jon Weissman	Un of Minnesota

Action Items

Proceedings

This MAGIC Meeting was chaired by Rich Carlson of DOE and Dan Katz of the NSF... The meeting heard presentations on four DOE funded infrastructures to support cloud computing and big data applications.

Abstractions and Integrated Middleware for Extreme-Scale Science (AIMES): Jon Weissman

Distributed extreme-scale science applications require access to large distributed data resources including sensors, instruments and archives. They also require diverse resource platforms including leadership class computation, grids, clusters, and clouds. Applications are increasingly characterized by extreme scale and extreme heterogeneity. Federating distributed resources is essential to extreme-scale computing and that requires abstractions for applications, resources, middleware, and infrastructure. Models are also needed for applications, resources, middleware, infrastructure, and their composition.

Within this context AIMES provides:

- Abstractions => models => knowledge discovery
- How the application performs on the distributed computational infrastructure (DCI)
- How to best adapt the application to the DCI
- How the DCI can be adapted to the application
- Why the system allocated the DCI to the application.

AIMES components can be used by tools, systems, and applications in an embedded sense. AIMES enables reasoning about executing distributed workloads by using models, execution strategies and skeletons. It can help identify what variables matter most and which matter least. AIMEs is deployed in a federated environment to explore integration across layers and to collect qualitative and quantitative evidence towards models. AIMES hides application complexity while capturing essential characteristics. Skeletons are simple tools to build synthetic applications that represent real applications with similar performance. In practice

c/o National Coordination Office for Networking and Information Technology Research and Development
Suite II-405 · 4201 Wilson Boulevard · Arlington, Virginia 22230
Phone: (703) 292-4873 · Fax: (703) 292-9097 · Email: nco@nitrd.gov · Web site: www.nitrd.gov

they are accurate in predicting performance. Bundles characterize heterogeneous resource aggregates to enable automatic selection of resources for an application. Compute bundles include NSF XSEDE and FutureGrid. AIMES develops pilots to characterize workload description and execution requirements on federated resources. They define key choices for executing a given workload.

Engineering success to date has indicated that AIMES is on the right track. This year AIMES will provide multi-site OSG, expand to network and storage resources, address data-dependent workflows and implement deeper integration.

In future years AIMES will address how to compose models, resource discovery and actuation and how to architect the infrastructure for specific performance and requirements.

The complete briefing may be accessed on the MAGIC Website at: https://www.nitrd.gov/nitrdgroups/index.php?title=MAGIC_Meetings_2015 for the March 2015 meeting.

dV/dt: Accelerating the Rate of progress towards extreme scale collaborative science: Ewa Deelman

The goal of dV/dt is to make it easier for scientists to execute large-scale computational tasks that use distributed computing resources and applications. dV/dt helps estimate the application resource needs, allocate the needed resources and manage the applications and resources during a computational run. Resources of dV/dt to support distributed applications include: real-world applications, state-of-the-art computation facilities (Argonne leadership systems and OSG), campus resources at ND, UCSD, and UW, commercial cloud resources, submitting requests locally for computation globally and managing the costs of resources and manpower. Elements of dV/dt include: workload characterization, workload estimation, resource allocation, execution, monitoring resource usage, execution traces and workload archive. Monitoring includes HTC monitoring to collect information on runtime, peak disk usage, memory usage, CPU usage, etc. It provides job information from the scheduler (Cobalt). I/O information is provided using Darshan. Performance counters use AutoPerf... Resource workload archives are maintained. These archives are used to model workloads and to characterize them. Task characteristics (runtime, disk space, memory consumption) are used in algorithms to provide scheduling and resource provisioning information. Trace analysis was used for 5 workflow applications to estimate the accuracy of the workflow estimation model. Resource allocation can be provided on the basis of one task per node or multiple tasks per node. One task per node provides quick results but may underutilize resources; multiple tasks per node may take longer but can provide efficient use of resources. Flexible allocation of nodes can dynamically allocate resources based on demand to provide a compromise between run time and optimal use of resources.

dV/dt products include monitoring tools, workflow archive and methods for estimating online resource needs.

Next steps with dV/dt include:

- Enhancing monitoring and modeling
- Using resource predictions for provisioning and scheduling
- Productizing tools
-

The complete briefing may be accessed on the MAGIC Website at:

https://www.nitrd.gov/nitrdgroups/index.php?title=MAGIC_Meetings_2015 for the March 2015 meeting.

Big PanDA: Kaushik De

Big Panda is a workload management system developed in support of the ATLAS experiment of the LHC. It processes hundreds of petabytes per year for thousands of users worldwide. It provides the structure for a hierarchy of computing centers working together. The main challenge is to provide efficient automated performance while making resource =s available to all users. It makes hundreds of distributed sites appear as local to users and provides a central queue for users similar to a batch system. It also:

- Reduces site related errors
- Hides middleware
- Hides infrastructure variations
- Production and Analysis users see the same PanDA system

PanDA provides seamless integration of distributed resources, access by all users to the same worldwide resources, uniform fair-share policies, and automation and error-handling. PanDA is deployed on the WLCG infrastructure and provides a standards-based implementation using REST, Oracle or MySQL, CondorG, Python packages, and command-line and GUI/Web interfaces. See; https://twiki.cern.ch/twiki/bin/view/PanDA/PanDA
the next generation of Big PanDA will provide:

- Generalization of PanDA for HEP and other data-intensive sciences
- Factoring the core
- Extending the scope
- Leveraging intelligent networks
- Improving usability and monitoring

BigPanDA is configured to run ATLAS on Titan at Oak Ridge. It improves CPU performance and provides the needed storage facilities for LHC computation. It enables finer granularity going from files to events. It is being expanded for use by HEP and other sciences. It utilizes facilities and services worldwide including Europe and Asian resources.

For the complete briefing please see:
https://www.nitrd.gov/nitrdgroups/index.php?title=MAGIC_Meetings_2015 for the March 2015 meeting.

Metadata Provenance and Ontology (MPO) Project: David Schissel

Magnetic fusion research has a long history of collaboration on data management for experiments and electronic logbook which has 500 thousand entries. The MPO project has the objective to preserve the meaning of data by documenting the data, their process, and their provenance. It identifies where a particular piece of data came from and where the data was used. It documents how particular published data were developed and it enables the capability to revisit the data and recalculate it to correct for errors. The MPO supports experimental and computational data. Users can record as much or as little data s they want. It functions in a heterogeneous environment and interoperates with existing user workflow tools. It is designed to work as autonomously as possible. The MPO system has servers for the database, the API, the Web and an event server. Directed Acyclic Graphs (DAGs) are used to define and show the processing chain. Parent/child relationships are stored. Each object has a name and description and is tagged with a time and user's name via workflow connections. The data model stores structured data, activities that create, moves, or transmutes data, connections between inputs, actions and results, comments, and collections (defined by the user). The data store is permanent but MPO doesn't dictate the specific implementation of the permanent store. Each object has a globally unique numerical identifier and is provided with a pointer using a Uniform Resource Indicator (URI). Searching is enhanced by defining a controlled vocabulary; an MPO ontology. MPO is a We service, Postgress SQL is used for the current implementation. A Twitter "Bootstrap" creates a standardized Web front end. Authentication uses x.509 certificates.

The basic components re all built. It is in the process of Beta testing. They are doing outreach to the potential user communities. This is the final year of development and they are working to expand the system's depth and expand the tools into other sciences. The complete briefing may be found at:
https://www.nitrd.gov/nitrdgroups/index.php?title=MAGIC_Meetings_2015 for the March 2015 meeting.

Upcoming Meetings:
March 23-26: GEC22, US Ignite, OGF43, GLIF, Washington DC
May 21-22: OGF44, Lisbon, Portugal
June 16, 2015: Portland, Oregon Workshop: Science of Cyberinfrastructure: Research, Experience, Applications and Models (SCREAM) in conjunction with HPDC'15.
https://sites.google.com/site/scream15workshop/

Next MAGIC Meetings:
April 1, 2015, NSF
May 6, 2015, NSF