MAGIC Meeting Minutes
June 3, 2015

Attendees

Richard Carlson	Richard.carlson@science.doe.gov
Shantenu Jha	Rutgers Un.
Dan Katz	NSF
Kirsten Kleese-van Dam	PNNL
David Martin	Northwestern Un.
Grant Miller	NCO
John Towns	NCSA

Action Items

Proceedings

The meeting was held June 3, 2015 at Stafford II-515 in Arlington, VA, and chaired by Rich Carlson, DOE and Dan Katz, NSF..

The May MAGIC minutes were approved. Rich Carlson asked that Grant Miller post them to the MAGIC website.

XSEDE: John Towns

Towns, John, *Overview of XSEDE Systems Engineering,* June 2015

XSEDE followed onto the TeraGrid Phase III: eXtreme Digital Resources for Science and Engineering. It enables digitally-enabled researchers and engineers to participate in multidisciplinary collaborations with access to computing resources and data. XSEDE has been funded for over $120 million over a 5 year period . It has no funding for major hardware. About 250 individuals are funded over 20 partner institutions. N IH, DOD, DOE, NASA, and other agencies.

XSEDE is an ecosystem of advanced digital services to accelerate scientific discovery: advanced computing, high-end visualization, data analysis and other resources and services. It is interoperable with other infrastructures. XSEDE provides a virtual organization with a dynamic distributed infrastructure. The networking foundation for XSEDE is the Internet2 AL2S service. It has over 7500 open accounts, over 5500 active accounts with gateways and over 2600 active users.

XSEDE resources include:
- Compute engines
- Data: simulation, input, output, instrument data,…
- Instruments: telescopes, beam lines, sensor nets, microscopes,…
- Infrastructure: local networks, wide-area networks,…

Services include: collaboration such as wikis, forums,…; data including data transport, data management; access (authentication, authorization, accounting); user support and other services such as education and outreach.

XSEDE resources are configured to satisfy user needs including the collection of resources to enable the applications including changes over time. XSEDE system functions include:
- Identity management

FOR OFFICIAL GOVERNMENT USE ONLY
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

- Interactive login
- Accessing remote files
- Submitting and managing computations
- Transferring datasets
- Discovering and providing resource information

User requirements are used to develop use cases that include:
- Data analysis, management and visualization
- Campus bridging
- Computing
- Connecting instrumentation
- Federation and interoperability
- Science gateways.

XSEDE is highly documented to enable users in developing and identifying XSEDE use cases.

Identity management has been reconfigured to Globus Authorization that links identities from many sources.

XSEDE components are configured into systems to address user requirements-
- Campus Bridging and Service Provider coordination
- End-to-end engineering process
- XSEDE Resource Allocation Service (XRAS) is used to select elements of the engineering process
- Engineering services and software are used to manage the use cases.

Ongoing challenges include:
- Managing engineering documents
- Stakeholder visibility into the engineering process
- Pace of getting components through the engineering process
- Timely stakeholder participation
- Resources to address big data, clouds, advanced SaaS and other emerging needs.

XSEDE will be continually upgraded to provide sustainability and improved user services

For the complete briefing please see the MAGIC Website, June 2015.

PNNL : Distributed Workflows, Current and Future: Kirsten Kleese-van Dam

PNNL supports the DOE BER Atmospheric Radiation Measurement (ARM) program. The ARM has 360 instruments at multiple remote field sites. Monthly, these instruments generate 18 TB of data and 1 million files. The core focus is to intake the raw data, provide quality assurance, produce data products across a number of sites and to provide the data to users.. Much of this process is provided through cloud services. Sites are supported by wireless networks, phone networks and all other access modes. It is challenging to connect to the remote sites at adequate data rates. Modeling simulations on the data produce large amounts of data that have to be managed.

The EERE Atmosphere to Electrons project has a goal of making wind farms more efficient. Data collection and modeling address the operational environment, collecting data, quality assurance on the data, and production of data products. Remote sites are addressed in a cloud environment.

PNNL operates projects for DOE EE on the electrical infrastructure through the Pacific NorthWest demonstration project. 60,000 metered customers provide data for collection, quality assurance ,and analysis.

PNNL supports the DOE HEP office for the Belle II Experiment at the KEK in Tsukuba, Japan. It provides 32 TB/day of ESnet data. The data is used to detect off-normal functioning in near-real-time.

Other PNNL distributed projects include:
- VELO
- ASCEM: a collaborative computing environment
- ACME: Accelerated Climate Model for Energy that accelerates modeling of climate events for faster prediction
- Premier Network: Electron microscopy informatics, education, and research network among about 30 partners. Cooperating together provides new capabilities. A portal enables uploading data, sharing analytic tools and sharing visualization products.

Moving forward, PNNL is addressing:
- Enabling more of the emerging collaborative science environments
- On-site data management and publication services
- Integration of resources into virtual laboratory infrastructures
- SubTER: monitoring subterranean sites for structures, earthquakes, and energy storage capabilities
- Low dose STEM: modification of scan speed and magnification to optimize observations

New tools and infrastructure are enabling support for smaller and more numerous research collaborations.

Meetings:
August 13-14: BER virtual laboratory meeting, Washington DC: How to design a virtual laboratory
September 28-29, WSSPE meeting, Boulder, CO: Software sustainability
October 15-16: Software management workshop, Washington DC
Current: INCITE is soliciting requests for time on their machines

<u>Next MAGIC Meeting</u>
July 1, 2015, NSF, Room TBD.