MAGIC Meeting Minutes
October 1, 2014

Attendees

Jim Basney	NCSA
Bob Bohn	NIST
Bob Bonneau	DoD
Rich Carlson	DOE
Heidi Dempsey	GENI
Shantenu Jha	Rutgers Un.
Dan Katz	NSF
Jim Kirby	NRL
Ken Klingenstein	Internet2
Miron Livny	OSG
David Martin	Argonne
Grant Miller	NCO
Inder Monga	ESnet
Lavanya Ramakrishnan	LBL
Alan Sill	TTU/TACC, OGF
Brian Tierney	ESnet

Action Items

1. Grant Miller will inform MAGIC of the date/time of the MAGIC meeting at SC14.
2. MAGIC members should send Grant Miller (miller@nitrd.gov), Rich Carlson (richard.carlson@science.doe.gov) and Dan Katz (dkatz@nsf.gov) topics for discussion in MAGIC and conveners for those topics

Proceedings

This MAGIC Meeting was chaired by Rich Carlson of DOE and Dan Katz of the NSF... The primary objective of the meeting was to identify specific focus topics for MAGIC to pursue during FY15.

LSN is expected to task MAGIC for FY15 to:
Take an in-depth look at how distributed computing will evolve over the next 5 – 10 years. Some specific issues are:
1) Multi-core processors requiring codes become more parallel
2) Memory to flop ratios changing
3) I/O rates not keeping up with either parallel processing or core network speeds
4) Emerging experimental / observational science communities
5) Instruments and detectors generating more data
6) Integrating knowledge bases into Distributed Computing infrastructures
7) Visualization and analysis services for multiple and/or remote users
8) Container and virtualization technologies

Topics of interest might include:

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

- Existing/developing virtual environments: OSG, OGF, GENI, FutureGrid, Internet2 Net+ environment
- Convene the OSG, CERN, OGF… communities to discuss their different approaches and what has worked/what has not worked.
- Engage commercial providers to discuss their existing/developing capabilities that can be applied to research communities and to inform them of science environment capabilities
- Bring the NSF funded cloud environments into the MAGIC discussions to represent academic community interests
- University community researchers and providers to identify what capabilities they have and what additional capabilities they will need.
- SDN developers to identify how their developing technology might impact virtual environments and distributed resources/distributed processing
- Multi-core processing and its implications for research communities
- Memory to flops ratios are changing. How will this affect research communities?

Discussion among the MAGIC members identified that:
- One outcome of a MAGIC discussion might be to coalesce an open workshop to bring in all interested constituencies to discuss cooperation on resolving difficult issues and how to move forward.
- Human factors are important. For emerging science communities, training and workforce development to enable them to interface to collaboration resources more effectively
- NSF has funded two cloud programs recently. We need to be able to stand-up and run federated clouds across domains with heterogeneous resources. We need to address control planes across infrastructures and administrative domains. Is there a way to assess the needs and requirements across all layers and interfaces?

Virtualization and containerization

Containerization addresses many areas of virtualization. Interoperability and standards are important issues. It would be good to have commercial cloud providers talk about what they are providing and where they are going. How can the science community leverage the resources of the commercial community. Where are there gaps such that we need research programs to address the gaps? Slicing and virtualization over mobile infrastructures is increasingly important. We need a broad brush meeting to identify specific issues to enhance and enable user capabilities.

In general, how do we build an end-to-end system over the inherent variance in systems and resources? We need to build applications that are less sensitive to variance in the resources. Distributed systems increasingly are focused on getting data resources to compute resources or compute resources to data resources. How do we accomplish this in a predictable way? Reliable networking is a basic requirement to achieve reliability and lower variance.

Ian Foster added, after the meeting that he is interested in the question of how science cloud services can be encouraged and sustained: e.g., Globus, Red Cloud at Cornell, nanoHUB.

Open source software is increasingly important to building and maintaining user applications and services. Software underlies all middleware. We need to address:

- In an era of rapidly changing compute storage, and communications resources how do we provide middleware that will evolve graciously with the changing resources
- How do we run applications on resources we don't own?
- How do we develop and deploy software for distributed and remote computing
- If I deploy my application and it doesn't work, how do I debug the entire system to identify problems?

**Organize a MAGIC meeting with short presentations on virtualization technology. Identify what is needed in the future to provide workable virtualized resources for users.

Multicore processing

I/O rates are not keeping up with processing speeds. A group in San Antonio is working with OpenCompute and could provide a current snapshot of capabilities. If computing is migrating to your iPhone, how does this change the paradigm of what you can compute and the software constraints imposed? This is an issue of how you adapt to and manage the complexity of the resources available to address an application. The capabilities are constantly changing. Further, if you eliminate a bottleneck, you then proceed to the next bottleneck. There is self-configuring, self-modulating software that senses the resources available and adapts to improve performance with those resources. . How do we manage the increasing complexity? TACC is deploying XEON5 on Stampede. Dan Stanzione could provide lessons-learned.

**Software sustainability within the total framework of changing/evolving infrastructure: How do we deploy software on a dynamic and heterogeneous infrastructure and how does the software evolve as the infrastructure changes?

As the hardware changes how do middleware providers, users, and the infrastructure adapt to support applications?   How do software environments need to change? How do we move continuously from one system update to another without taking down the applications for an extended period? How do we graciously evolve our environments?

Data issues

Data is produced by instruments, data archives/resources, sensor networks, and other sources. Resources are needed to move, store, and manipulate the data. RDA could be invited to talk to MAGIC about research problems that have arisen in their engagement with the user communities, Fran Berman of the RDA could coordinate a session on data issues that we should be aware of.
- We should focus on data as it is interacting with applications and computing, not on aspects of semantics, metadata and data storage/retrieval.
- The biggest challenge is dealing with the magnitude of the data storage. What data do you keep for posterity? How many resources do you dedicate to storage? Standards is one approach.
- We should focus on policy issues and less on resources for data. What is the gap between science practices and industry? We have previously focused on storage integrated with computing resources.

AI: Grant Miller will inform MAGIC of the date/time of the MAGIC meeting at SC14.

AI: MAGIC members should send Grant Miller (miller@nitrd.gov), Rich Carlson (richard.carlson@science.doe.gov) and Dan Katz (dkatz@nsf.gov) topics for discussion in MAGIC and conveners for those topics

Upcoming Meetings:
November, SC14, New Orleans:
November: WSSSPE2 (workshop focused on software) at SC14

Next MAGIC Meetings:
November 18, 1:30-3:30, SC14
January 7, 2:00-4:00 NSF, Room II-565