MAGIC Meeting Minutes
July 3, 2013

Attendees

Peter Beckman	ANL
Rich Carlson	DOE/SC
Rudi Eigenman	NSF
Dan Gunter	LBL
Dan Katz	NSF
Kate Keahy	ANL
David Martin	Northwestern U.
Grant Miller	NCO
Von Welch	Indiana U.

Action Items

Proceedings

This MAGIC Meeting was chaired by Rich Carlson of DOE/SC.

Pete Beckman of Argonne National Laboratory provided a presentation and discussion of Argo, an exascale operating system and runtime research project.

Argo

The DOE Operating System/Runtime Technical Council held 7 meetings in 2012 to discuss the needs and user requirements for exascale OS/R capabilities. See: https://collab.mcs.anl.gov/display/exaosr.
The Argo team includes researchers from a wide range of facilities including ANL, BU, LLNL, PNNL, UC, UIUC, UO, and UTK.

Data from Peter Kogge demonstrates that, for OS/R we have hit a power ceiling and a clock ceiling and that sockets and cores are growing in supercomputers.

A new whole-system view is needed to accommodate:

- Dynamic user environments with complex workflows, coupled applications, multi-physics, customized software stacks, Users want more functionality, dynamism, and flexibility.
- Power is a first-class resource.

Other areas needing attention include memory technology, fault tolerance and resilience, and embedded NICs. Key innovation areas include:

- Node O/S
- Lightweight runtime for concurrency
- Event, control, and performance backplanes
- Global optimization

Key Argo abstractions are tree-based hierarchy and recursive decomposition. At each level in the hierarchy there are changes in granularity of control,

communication frequency, goals, and data resolution. Embedded feedback and response mechanisms provide a self-aware, goal-based active runtime system. A meta-handle for enclaves provides for managing parallelism, task-management, and other tasks. The hierarchical, coordinated, global system can set and manage power budgets, respond to faults, support enclave components, and manage intranode parallelism

Resource management design principles provide for hierarchical resource management and the resource managers are stackable, integrated, and customizable and adaptable. Sharing is avoided when possible.

Key research areas include:

- Threads/tasks: managing exploding parallelism: the programmer cannot hand-pick granularity/resource mapping
- PLASMA: parallel linear algebra s/w for multicore architectures: Provides high utilization for each core, scaling to a large number of cores, and shared or distributed memory.
- Charm++: the computation is decomposed into natural objects of the application which are assigned to processors by Charm++
- Argo parallelism: move away from SPMD block synchronous, link lightweight task runtime into the OS, explore memory placement, support data dependency driven computation, and explore pluggable schedulers
- New memory will become available: Spin-Torque Transfer RAM, phase change RAM, resistive RAM.
- A significant portion of memory will be non-volatile: reduces power, helps resilience, decreases cost
- Power/energy trace tools: command line tool, sampling power consumption at specified intervals, summarizing total power consumption.

Summary
Future systems will provide
- Node OS
- Lightweight runtime for concurrency
- Event control and performance backplanes
- Global optimization

For the full briefing, please see the MAGIC website meeting for July 3 at:
http://www.nitrd.gov/nitrdgroups/index.php?title=Middleware_And_Grid_Interagency_Coordination_(MAGIC)#title

Upcoming Meetings
Week of July 30-Aug 2, OGF Workshop, Miami, Florida
Week of July 30-Aug 2 Federated Cloud Workshop, Germany
November 11-15 InCommon Identity Week, Silicon Valley

OSG and XSEDE are offering a summer school to provide understanding of the principles, concepts and applications. A link to this meeting is provided on the XSEDE Web page.

Next MAGIC Meetings

- August 7, 2:00-4:00, NSF, Room II-415
- September 4, 2:00-4:00, NSF, Room II-415