INL/EXT-10-19942

Process for Selecting NEAMS Applications for Access to Idaho National Laboratory High Performance Computing Resources

Michael Pernice

September 2010

The INL is a U.S. Department of Energy National Laboratory
operated by Battelle Energy Alliance

INL/EXT-10-19942

Process for Selecting NEAMS Applications for Access to Idaho National Laboratory High Performance Computing Resources

Michael Pernice

September 2010

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Process for selecting NEAMS applications for access to Idaho National Laboratory high performance computing resources

Michael Pernice*

September 16, 2010

1 Introduction

INL has agreed to provide participants in the Nuclear Energy Advanced Modeling and Simulation (NEAMS) program with access to its high performance computing (HPC) resources under sponsorship of the Enabling Computational Technologies (ECT) program element. This report documents the process used to select applications and the software stack in place at INL.

2 Resources

At INL, the Center for Advanced Modeling and Simulation (CAMS) manages INL's modeling and simulation cyberinfrastructure in partnership with the High Performance Computing group in the Information Management directorate. The HPC center features Icestorm, an SGI Altix/ICE system with 2,048 processors, dual InfiniBand® interconnects, 4 Tbytes of distributed memory, and a peak performance rate of nearly 22 Tflops. One of the InfiniBand® interconnects is dedicated to interprocessor communication needed by parallel jobs; the other provides global access to 70 TBytes of distributed storage. Icestorm is complemented by two smaller systems. Helios is a heterogeneous cluster with 840 processors, a Gigabit Ethernet interconnect, 1.7 Tbytes of distributed memory, and a peak performance rate of 8.9 Tflops. Eos is a fat-node cluster with 88 processors, 640 Gbytes of memory, and a peak performance rate of 921 Gflops.

3 Application selection process

The application selection process began with contacting members of the Integrated Performance Safety Code (IPSC) teams to understand their needs and

*Center for Advanced Modeling and Simulation, Idaho National Laboratory, Idaho Falls, ID 83415 email: `michael.pernice@inl.gov`.

to determine their interest. This began in December 2009 via electronic mail. Project leaders for the Reactors, Fuels, and Waste Forms IPSCs[1] were contacted. Randy Summers , project lead for the Waste Forms IPSC team, indicated that the software development effort was just getting started and that available computing resources at Sandia National Laboratories were adequate. Both the Reactors and Fuels IPSC teams expressed interest in obtaining access to additional HPC cycles.

These initial contacts were followed up by visits to obtain more detail on the computational requirements and to provide descriptions of the available resources. In February 2010 I attended a modeling and simulation planning meeting that preceded the Advanced Fuel Cycle Workshop on Small Scale Experiments in Albuquerque, NM. Several participants at that meeting expressed interest in using INL HPC resources.

- David Andersson (LANL) is studying Xe redistribution in UO_{2+x} using density functional theory, molecular, and continuum scale calculations. The long term objective is to improve understanding of fission gas redistribution as a function of operating conditions and initial microstructure.

- Jeff Rest (ANL) is developing a new model based on rate theory to simulate gas-bubble nucleation in nuclear fuels.

- Rad Radhakrishnan (ORNL) is using a Monte Carlo/Potts model to perform mesoscale simulation of pore/bubble evolution in nuclear fuel.

- Chris Stanek (LANL) is performing atomistic simulations to improve understanding of fission gas release in oxide fuels. Their simulation capability can now be used to consider grain boundary structure.

Also in February 2010, I traveled to ORNL and ANL along with other members of the ECT team to meet with members of the Fuels and Reactors IPSC teams. At ORNL, Kevin Clarno gave an overview of `AMP`. INL already provides HPC access to INL members of the Fuels IPSC team, and we have been looking into ways to support the distributed software development efforts of this IPSC. Before this can occur, export control issues need to be resolved. At ANL, we met with Tim Tautges, Paul Fischer, and Micheal Smith. Fischer indicated that there were some flow studies that would be suitable for using INL resources.

Given these responses, our expectation was that needs would exceed the available resources, so we decided to request short proposals to assist in the downselection process. In early March 2010, I solicited short HPC resource requests from Andersson, Fischer, Radhakrishnan, Rest, Smith, and Stanek. By mid-March I had received responses from Fischer and Stanek. Fischer seeks to analyse thermal striping in a T-junction, while Stanek will perform atomistic simulations of fission gas behavior and thermal conductivity in oxide fuels.

[1] The Separations and Safeguards IPSC team was not contacted at that time; subsequent follow-up at the NEAMS Spring 2010 Principal Investigators meeting identified additional potential candidates.

Smith indicated he had access to adequate resources at ANL and was reluctant to spread his efforts too thin across different compouting facilities. Follow-up attempts to obtain short proposals from the other researchers via phone calls were unsuccessful.

As Fischer and Stanek had each requested slightly more than 10% of INL's current HPC capacity and the deadline for selecting applications was near, we initially decided to support these two applications. By the end of March, application forms for accounts and RSA tokens were sent to Fischer and Stanek. Completed forms were not returned.

Subsequently, members of the ECT and Safe Separations program elements met at the spring NEAMS PI meeting at ANL. The Safe Separations team provided an informative overview of their efforts, which include lumped model descriptions of a processing plant, unit components of continuum models such as microflows in contactors, and lower length scale simulations that span quantum to molecular scales.

Two members of the Safe Separations team expressed interest in utilizing INL's systems, have completed account applications and received their RSA tokens, and are just getting started using the INL systems. Valmor de Almeida (ORNL) is using molecular dynamics simulations to study the interfacial transport of water into tri-butyl-phosphate, the most prevalent nuclear reprocessing solvent extractant. We plan to install additional software to support de Almeida's efforts, including `git` for distributed software version control and `judy`, a general purpose dynamic array management package. Marianne Francois (LANL) is investigating droplet-droplet interactions in preparation for large-scale three-dimensional simulation of multiple droplet dynamics. We have already installed a copy of `hypre`, a package of high-performance preconditioners, to support Francois' efforts. Other required software includes `netcdf`, `python`, and `tecplot`, which are already available on INL's HPC systems. In addition, a `/projects/SafeSep` directory has been created to facilitate installation of third-party libraries that are specific to this IPSC.

4 INL Software Stack

4.1 Background

INL acquired Icestorm in the fall of 2007. This description specifically applies to the software stack currently operating on that system. Icestorm will continue to be maintained as a computing resource for the Lab through mid-year 2013. INL is also in the process of acquiring a larger cluster which is expected to be delivered by late fall of 2010. Future plans include implementation of a similar software stack with some minor variations. Those variations will depend on the hardware architecture and native software stack supplied by the system provider.

There are two support teams that manage the operations of the INL software stack; they are the operations and the software consulting groups. Although their daily responsibilities can overlap, the operations group is primarily responsible for the software stack on individual nodes, monitoring performance, and maintaining many of the libraries. The software consulting group maintains the various MPI libraries and compilers, debugs and optimizes users codes, and coordinates job scheduling.

4.2 Stack Definition

The Icestorm compute cluster's software stack is currently based on SuSE Enterprise Linux (SLES) 10.3 teamed with SGIs Tempo 1.10 software stack. INL follows the release cycle of SLES in conjunction with SGIs releases of their Tempo versions. The software stack is customized by INL to allow future installation of a Panasas parallel storage client. All cluster compute nodes are diskless which allows for software stack unity between nodes.

INL uses Environment Modules to provide dynamic modification of a users environment. This allows the user to easily change their environment between different applications and compilers. Additionally, they facilitate the users ability to move seamlessly between versions of the same software, where newer libraries or compilers can be required for one application and older versions for another.

Fundamental components of the INL software stack on Icestorm are:

- A complete SuSE Enterprise Linux (SLES) distribution.
- The Open Fabrics Enterprise Distribution (OFED) InfiniBand® software stack including MVAPICH and OpenMPI libraries.
- PBS Pro resource manager and batch job submission.
- Fully integrated Panasas parallel file system clients.
- Cluster administration tools to facilitate installation, configuration (including BIOS setup/upgrade), and remote lights-out management.
- An extensible cluster monitoring solution with support for both in-band and out-of-band (e.g., Intelligent Platform Management Interface (IPMI)) methods.
- GNU C, C++, and Fortran90 compilers integrated with MVAPICH and OpenMPI.
- Intel and PGI compilers integrated with MVAPICH and OpenMPI.
- Dynamic environment modification via Modules.
- Totalview Parallel Debugger
- NIS Authentication

4.3 INL Monitoring and Configuration Management

The following open source software packages are utilized as key components of the monitoring portion of the software stack. All run in conjunction with locally developed Perl scripts.

- Nagios is used to monitor headnodes and management systems, their respective services, and provide notification to administrators in the event of potential system problems. It is currently monitoring services such as SSH, system logs, filesystems, and system load. Since these systems are critical for the proper operation of HPC clusters, monitoring their performance is imperative for the overall health of the environment.

- Ganglia is the primary mechanism for monitoring the compute nodes. It monitors various metrics of the compute nodes similar to Nagios, such as load, memory usage, and network traffic. These statistics are prepared and displayed for the cluster as a whole as well as for individual nodes.

- MRTG is a network traffic monitoring system that has been enhanced with in-house Perl and Python scripts to monitor HPC cluster resources. Our implementation, called HPC System Loading Charts, displays CPUs in use (the actual load on the system) compared to CPUs requested (queuing system requests). In addition, it displays job information such as the id, name, user, status, wall-time (time the job runs). The graphical display is very useful in ascertaining how efficiently the clusters are being utilized.

4.4 INL Evolution

INL relies on the vendor's software stack of a given large-scale compute cluster. The HPC support team expects the vendor to provide a customized operating system/software stack that is tailored and optimized for their hardware solution. This allows the system to be operational within a short period of time from delivery and installation. The goal is to keep the software stack simple and flexible while maintaining a secure environment that can be easily monitored. Although INL is a long way from exascale computing, it is participating in a DOE laboratory-wide effort to adopt monitoring techniques and technologies that can eventually scale to peta- and exascale computing.

In an effort to increase security and integrate with the overall infrastructure of the Laboratory, HPC intends to develop a new authentication strategy. This strategy will provide users with a better sense of security and continuity with the overall infrastructure at the Lab.

New monitoring solutions will continue to be investigated in order to improve the visibility of events in a cluster and in the physical data center. These new technologies will provide the capabilities of indexing and searching logs along with historical usage and environment metrics.

Acknowledgement

The description of the INL software stack was provided by INL's High Performance Computing Group. Special thanks are extended to J. Staffon for assembling this information.