SAND2007-3673P

A User's Guide to Extreme Computing: Running on the Top 10% of the Top500 List

Robert A. Ballance, Sandia National Laboratories
(raballa@sandia.gov)
John T. Daly, Los Alamos National Laboratory
(jtd@lanl.gov)
LA-UR-07-2452
Mahesh Rajan, Sandia National Laboratories
(mrajan@sandia.gov)

Abstract

This tutorial originates from experiences gained from grappling with large simulations on DOE/ASC systems. This tutorial will provide foundations for users of and support personnel for capability class simulations, anticipating the issues that they are likely to encounter as they embark upon pushing new frontiers in high-end computing (HEC). It provides guidance in planning and executing simulations that use the major portion of the resources on a large system and suggests ways to maximize their chances for success while expending less effort. Tutorial attendees will learn tips, techniques, and best practices for running very large capability simulations on thousands of processors; lessons that are directly applicable to managing simulation of any size!

Overview

As systems and simulations have grown over the past 10 years, we have come to realize that there is a body of knowledge about running applications that is (i) normally learned from experience and (ii) could usefully be codified and transmitted to newer users of the systems. From our vantage point of using (Daly, Rajan) or administering (Ballance) some of the largest systems on the Top500 list, we have assembled a tutorial that will transmit some of our knowledge and experience.

One problem with this tutorial is that, at present, the working set of analysts is small. Not many people have multi-terascale systems available, and the type of work we do is specialized even for that. For example, Google has lots of CPUs, but uses them very differently. What makes the capability domain unique is that we are working with large coupled calculations, meaning that if one part of the calculation stops then everything else has to stop and start over again. Add to that the requirement to provide multiple runs, and the issues scale by another order of magnitude.

Our audience consists of programmers, analysts, computational scientists, HPC system administrators and user support staff. While many of these users are not yet running jobs that use thousands of processors, run for weeks of compute time, or generate terabytes of data, they will be doing so soon. Working efficiently and effectively on large platforms will become a critical job skill.

An important aspect of running simulations —data and metadata management—will be stressed. Often these simulations require multiple restarts before completion.Strategies for maximizing the 'application availability' and integration of huge volumes of data, in the context of still evolving robust file systems will be discussed. Another focus of this tutorial is to consider the full cycle of model generation, execution and post-processing, with new, large, models that have never been run before. Individual jobs are no longer the unit of workload on capability systems. Instead, we frequently see users running packages of jobs that involve multiple run of multiple applications, often for long periods of wall time. Managing the package becomes essential. Each of the workshop presenters has confronted, and successfully solved this problem, but in different ways. Our various points of view add to the depth of presentation.

This tutorial is intended for intermediate to advanced HPC users, who are familiar with parallel applications and use of high-end compute systems. The overall profile is 20% introductory, 50% intermediate, and 30% advanced. We expect that attendees will have experience running moderate to large-scale parallel applications on clusters or other HPC machines. At the end of the day, the attendees will leave with a greater appreciation of the problem, and numerous tips, ideas, and pointers that can be immediately applied in their current work environments.

The workshop authors work closely together, communicate frequently, and are not expecting major difficulties in merging their materials.

Outline

1) Introductions (5 min)
2) Capability class simulation as defined by ASC program managers (15 minutes)
   a) Large processor counts
   b) Long runtimes
   c) Huge datasets
   d) Multiple runs within a study: the work-package
3) Capability class systems and unique features (40 minutes)
   a) Example Systems
      ASC Red Storm
      ASC Purple
      BG/L
   b) System elements
      Lightweight-kernels
      File Systems
      Programming Models
      Usage Models
      Program development tools
      RAS
4) Steps in bringing up and testing Capability class systems (20 minutes)
   a) Network tests
   b) Hardware tests
   c) File system tests
   d) RAS Metrics

BREAK (15 minutes)

5) Environments where System Reliability and Availability Affect Performance (30 minutes)
   a) Defining application throughput: system performance from the application's perspective
   b) Monitoring and measuring: assessing the impact of system interrupts on application throughput
   c) Automated job control: keeping the application moving forward despite interrupts
   d) Shifting paradigms: ideas for further improving application throughput
6) Experience with supporting large capability simulations (15 minutes)
   a) Interaction with analysts
   b) Interaction with code developers
   c) Preparation and checklist for capability simulation
   d) Scaling Studies on Red Storm and Purple
   e) The Red Storm 7x benchmark suite; management tools and metadata for managing large studies
7) System support issues (15 minutes)
   a) Logging and metadata
   b) Getting the most from the peopleware; interaction with the sysops
   c) Getting the most from the scheduler: queuing hacks, restarts, and pushing it through
8) Successes and difficulties encountered (15 minutes)
   a) Sandia Capability class simulation movies
   b) I/O optimization
   c) Data movement
   d) Visualization of large petabyte class data and challenges

9) Wrap-up (10 minutes)

Robert A. Ballance, Ph.D.

Robert A. (Bob) Ballance is a Principal Member of Technical Staff at Sandia National Laboratories, where he has been deeply involved in the delivery, acceptance testing, and production operation of Red Storm, Sandia's newest capability high performance computing (HPC) platform. As System Manager, he stands in the intersection of system design, system administration, operations, and user support.

Dr. Ballance honed his production HPC operations skills while serving as the Manager of Systems and Systems Research at the Center for High-Performance Computing at the University of New Mexico (HPC@UNM). There, he oversaw all of the high-performance computing systems associated with the Center including the 512-processor Los Lobos cluster. Los Lobos was among the first three Linux clusters to break the top 100 list of supercomputers. He also played a guiding role in the early adoption and evolution of Access Grid technology.

From 1993 to 1999 Bob was employed in the private sector, first as President of Object Science Corporation and later as an independent consultant serving both government and private sector clients. During this time, he worked as a software architect, designer, and developer for products related to scientific computing and to the Internet.

From 1988-1993, he served as an Assistant Professor of the University of New Mexico Computer Science Department. At UNM, His compiler research resulted in the definition and

application of the Program Dependence Web as a data structure for optimizing compilers and program translators.

From 1981 to 1984, while at HP Laboratories, he worked on the design and implementation of the first optimizing compiler for the C programming language developed for the HP Precision Architecture, one of the first production RISC chips. Also at HP, he wrote the first device driver for any HP flatbed scanner. From 1978 to 1980, he was a member of the technical staff at Bell Telephone.

Bob received his Ph.D. from University of California, Berkeley, in 1989. He currently serves as a member of the Linux Clusters Institute Steering Committee, chair of the XT3/XT4 Special Interest Group of the Cray User's Group, and is a member of the Poster Committee for SC07.

Recent publications

"Bringing up Red Storm: Lessons to Remember", with John P. Noe, Cray User's Group, May 2007, in preparation.

"7X Preliminary Performance Results: ASCI Red vs. Red Storm," with Karen Haskell, John P. Noe, and Joel O. Stevenson , SAND 2006-5808P, Cray Technical workshop Oct 10-12, 2006, Oxford, England.

"Red Storm: An Update on the Upgrade," with Douglas Doerfler, Suzanne M. Kelly, James L. Tomkins, and Joel Stevenson, SAND 2006-7146P, SC06 ASC Booth Presentation.

"A Preliminary Report on Red Storm RAS", with Jon Stearley, Cray User Group Meeting, 2006, Lugano, Switzerland,

"Raising Red Storm," SAND 2005-7263 P, University of New Mexico, November, 2005.

"Challenges in Production Computing at Sandia National Laboratories," with John P. Noe, SAND 2004-2149 C, The 5th LCI International Conference on Linux Clusters: The HPC Revolution 2004, Austin TX.

John T. Daly

John Daly is a technical staff member in the HPC (high performance computing) division at the Los Alamos National Laboratory, where he accumulates in excess of a half-million processor hours a day running large-scale simulations on the Red Storm, Purple, and BG/L platforms. The remainder of his time is devoted to researching and developing metrics and methodologies for predicting and optimizing supercomputer reliability and throughput performance from the perspective of the application.

Prior to that he worked as a software engineer and application analyst for Raytheon Intelligence and Information Systems, where he ported and optimized scientific applications to run on a variety of HPC platforms. He holds degrees in engineering from Caltech and Princeton University, where he studied computational fluid dynamics under Antony Jameson.

Recent publications

"The Path to Resilience: Measuring Application Throughput in the Presence of System Failure", Los Alamos Technical Report (LA-UR-07-2285).

"Accomplishments and Goals in the Age of Petascale", Los Alamos Technical Report (LA-UR-07-1011), presented at the ASC Principal Investigators Meeting, Las Vegas, NV, 2007.

"Facilitating High-Throughput ASC Calculations", to be published in NWP Highlights, 2007.

"Failure Analysis From the Application's Point of View", Los Alamos Technical Report (LA-UR-06-8089).

"Methodology and Metrics for Quantifying Application Throughput", to be published in Proceedings of the Nuclear Explosives Code Developers Conference, 2006.

"A Higher Order Estimate of the Optimum Checkpoint Interval for Restart Dumps", Future Generation Computer Systems 22, 2006, pp. 300-312.

"CIS External Review: A Customer Perspective on Red Storm", Los Alamos Technical Report (LA-UR-06-3802).

"Evaluating the Performance of a Checkpointing Application Given the Number and Types of Interrupts," Proceedings of the Workshop on High Performance Computing Reliability Issues, HPCA, San Francisco, USA, February, 2005.

"A Strategy for Running Large Scale Applications Based on a Model that Optimizes the Checkpoint Interval for Restart Dumps," Proceedings of the First International Workshop on Software Engineering for High Performance Computing System Applications, ICSE, Edinburgh, Scotland, May, 2004, pp. 70-74.

Mahesh Rajan, Ph.D.

Dr. Mahesh Rajan is serving Sandia National Laboratories as Principal Member of the Technical Staff, providing user support and consultation on High Performance Computer (HPC) applications on Red Storm and other high-end computing clusters. During his five year service at Sandia, he has defined and measured parallel application performance of several key applications in support of Sandia's platform strategy in national level reviews - (1) the eminent JASONS group, (2) the National Research Council - and (3) the national High End Computing Revitalization Task Force review. He also developed a multi-physics linking software for coupling fire simulation radiation potentials with weapon-in-a-fire thermal uncertainty quantification analysis. He has an active research program in performance analysis and modeling, funded through Sandia-CSRF.

Prior to joining Sandia, Mahesh was Senior Technical Consultant at Caltech/JPL for Hewlett Packard Co. (HP) serving as the on-site staff for application porting, development and tuning for the largest HP parallel computer installed in the world at that time. He advised researchers and scientists on issues relating to successful parallel application development using message passing and multi-threaded programming models. He also coordinated product performance issues with HP sales and technical teams. He conducted seminars and workshops on the HP Exemplar/Superdome architecture and on issues impacting effective use of the system.

His other industrial experience includes serving as the Segment Leader, Scientific & Technical Systems for International Business Machines for the western region for two years, and, at Intel Supercomputer Systems Division as an Field Applications Engineer, for three years. He was a member of the Intel proposal response team for Sandia National Laboratories TeraFlops

initiative providing system architecture solutions. At Intel he was also a member of the field technical team at Caltech and SDSC providing technical assistance in state-of-the-art HPC application development, and supported JPL in developing high performance I/O interfaces using Parallel File Systems. Prior to Intel Mahesh worked for HPC startup companies Maspar (3 years) and Supercomputing Solutions (4 years), developing parallel applications, and math libraries.

Before joining the industry, Mahesh was a tenured faculty at Arizona State University for six years in the department of Mechanical and Aerospace engineering. He advised five masters and two PhD students and had NSF, NASA funded research programs in engineering mechanics. Mahesh obtained his Ph.D. in Engineering Science and Mechanics from Virginia Tech with Prof. John L. Junkins (presently with Texas A&M) in 1981.

Recent Publications

"Experiences with the use of CrayPat in Performance Analysis", Cray Users Group, May 2007, in Preparation, Seattle, WA.

"Performance Analysis in Support of Capability Computing," First Annual Cray Workshop-North America, Feb 26-28, 2007, Nashville, TN

"Supercomputer and Cluster Performance Modeling and Analysis", Co-author, Sandia Report, SAND2007-0601, Feb. 2007.

"Performance Analysis, Modeling and Enhancement of Sandia's Integrated TIGER Series(ITS) Coupled Electron/Photon Monte Carlo Transport Code, ", with Franke, B., Benner, R., Kensek, R., and Laub, T., Proceedings of the LACSI Symposium, Oct. 11-13, 2005, Santa Fe, NM

"What Jumbo Mode did for Alegra, Salinas, Fuego and Others", with R.E. Benner", ASC PI Meeting, Feb. 21-24, 2005, San Antonio, TX

"Investigation of Scaling Performance and Mesh Convergence with Sandia's ASC SIERRA/Fuego Code for Fire Model Predictions of Heat Flux," With Amalia Black and Stefan Domino, Cray User Group Meeting, 2006, Lugano, Switzerland.