Chapter 8

On the Path to Exascale

Ken Alvin
Sandia National Laboratories, USA

Brian Barrett
Sandia National Laboratories, USA

Ron Brightwell
Sandia National Laboratories, USA

Sudip Dosanjh
Sandia National Laboratories, USA

Al Geist
Oak Ridge National Laboratory, USA

Scott Hemmert
Sandia National Laboratories, USA

Michael Heroux
Sandia National Laboratories, USA

Doug Kothe
Oak Ridge National Laboratory, USA

Richard Murphy
Sandia National Laboratories, USA

Jeff Nichols
Oak Ridge National Laboratory, USA

Ron Oldfield
Sandia National Laboratories, USA

Arun Rodrigues
Sandia National Laboratories, USA

Jeffrey S. Vetter
Oak Ridge National Laboratory, USA

ABSTRACT

There is considerable interest in achieving a 1000 fold increase in supercomputing power in the next decade, but the challenges are formidable. In this chapter, the authors discuss some of the driving science and security applications that require Exascale computing (a million, trillion operations per second). Key architectural challenges include power, memory, interconnection networks and resilience. The chapter summarizes ongoing research aimed at overcoming these hurdles. Topics of interest are architecture aware and scalable algorithms, system simulation, 3D integration, new approaches to system-directed resilience and new benchmarks. Although significant progress is being made, a broader international program is needed.

DOI: 10.4018/978-1-4666-0906-8.ch008

INTRODUCTION

In 1997 Intel's ASCI Red broke the Teraflops barrier, achieving over 1 trillion floating point operations per second (Heermann, 1998). Last year both IBM's Roadrunner and Cray's Jaguar system surpassed 1 PetaFLOPS or 1,000 TeraFLOPS (Feldman, 2008). The next factor of 1000 improvement in supercomputing performance (Exascale) will be even more challenging. The primary driver for the architectural change underway is that clock speeds are increasingly constrained by power and cooling limits. All of the major chip manufactures are moving to multi-core architectures, which results in the addition of hierarchical parallelism to supercomputers. Oak Ridge National Laboratory's Cray Jaguar system has 18,688 nodes, each comprised of two quad core AMD Opterons. Roadrunner is a one of a kind supercomputer composed of 6480 dual core AMD Opterons, each connected to two IBM PowerXCell 8i processors, which are similar to the processor used in Sony's Playstation 3 (Kahle, 2005). Both of these systems demonstrate that the transition to multi-core is already a key design challenge for supercomputer architectures.

The memory wall is defined as the mismatch between CPU and memory performance (latency, chip I/O capabilities, etc.), and will continue to plague processor design. Today's applications are primarily limited by data movement, represented by the statement "FLOPS are free" (Shiva, 2005). Power budgets continue to increase, which will result in power requirements in excess of 100 MW if existing Petascale design methodologies are used for an Exascale system. It is unlikely that such a system would be built unless power demands can be decreased significantly. Additionally, resilience will limit the availability of such systems. The current method for recovering from faults is check-pointing – at various times during a calculation a restart file is written to disk. Given weak scaling, I/O bandwidth requirements scale with the memory size and inversely with the mean time between interrupts (which decreases with increasing parts counts). Since Exascale will require millions of nodes, the mean time between interrupts will decrease and check-pointing will become an impractical mechanism for fault recovery.

These challenges require new approaches to applications, algorithms, system software, and computer architecture, which has been noted by the numerous workshops and reports devoted to the problem (Kogge et al., 2008; Simon et al., 2007). The remainder of the chapter discusses science and security applications that require Exascale computing, architectural challenges, and ongoing work aimed at overcoming some of these obstacles.

APPLICATIONS

Scientific computing is essential to the advancement of numerous fields of study. Never before have we been able to accurately anticipate, analyze, and plan for complex events that have not occurred—from the operation of a reactor running at 100 million degrees to future changes in climate. Combined with the more traditional approaches of theory and experiment, scientific computing provides a profound tool for insight as we look at complex systems containing billions of components.

SCIENCE

Key areas of scientific research, including materials science, Earth science, energy assurance, fundamental science, biology and medicine, engineering design, and security can benefit from continued growth in high performance computing (to Exascale and beyond). Table 1 summarizes scientific opportunities that can be enabled by Exascale computing, key application areas, and the goals and associated benefits.

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/path-exascale/64445

Related Content

High Performance Computing Design by Code Migration for Distributed Desktop Computing Grids
Makoto Yoshidaand Kazumine Kojima (2011). *International Journal of Grid and High Performance Computing (pp. 53-70).*
www.irma-international.org/article/high-performance-computing-design-code/60255

Discovering Knowledge in Data Using Formal Concept Analysis
Simon Andrewsand Constantinos Orphanides (2013). *International Journal of Distributed Systems and Technologies (pp. 31-50).*
www.irma-international.org/article/discovering-knowledge-data-using-formal/78152

Testing-Effort Dependent Software Reliability Model for Distributed Systems
Omar Shatnawi (2013). *International Journal of Distributed Systems and Technologies (pp. 1-14).*
www.irma-international.org/article/testing-effort-dependent-software-reliability/78150

Improving Energy-Efficiency of Computational Grids via Scheduling
Ziliang Zong, Xiaojun Ruan, Adam Manzanares, Kiranmai Bellamand Xiao Qin (2010). *Handbook of Research on P2P and Grid Systems for Service-Oriented Computing: Models, Methodologies and Applications (pp. 519-542).*
www.irma-international.org/chapter/improving-energy-efficiency-computational-grids/40816

Improved Compact Routing Schemes for Random Interconnects
Chi-Hieu Nguyen, Chung T. Kieuand Khanh-Van Nguyen (2020). *International Journal of Distributed Systems and Technologies (pp. 89-109).*
www.irma-international.org/article/improved-compact-routing-schemes-for-random-interconnects/256208