Green HPC
From Nice to Necessity

As power consumption of supercomputers grows out of control, energy efficiency will move from desirable to mandatory.

The idea of green high-performance computing (HPC) has been gaining traction over the past five years. In November 2008, the Green 500 list (www.green500.org) was introduced to “raise awareness about power consumption, promote alternative total cost of ownership performance metrics, and ensure that supercomputers only simulate climate change and not create it.”

The Green 500 lists the world’s most energy-efficient supercomputers, based on a floating point operations per second (flops) per watt metric. Although the list has helped raise awareness of energy efficiency for HPC, since its inception, more dramatic drivers for energy-efficient HPC have arisen.

Historically, HPC system power consumption has been of secondary importance, but as the HPC community looks from petascale to exascale, power will become a first-order concern.
DARPA’s exascale report forecasts a greater than 100 megawatt (MW) power budget for a 1 exaflops machine if current trends continue. This power budget is unsustainable from both an environmental and financial standpoint (a 100-MW machine would result in a $100 million power bill each year). So, while green HPC might have been merely desirable to this point, going forward it will be a requirement.

**Exascale Requires Green Computing**

The push for exascale computing will provide a driver for an unprecedented level of energy efficiency. DARPA’s Ubiquitous HPC program has as its goal a 2018 rack-level prototype that achieves 50 gflops/watt compute efficiency. By comparison, the current number one on the Green 500 list has an efficiency of only 773 mflops/watt—a difference of more than 60 times. The actual problem might be even more worrisome: eight of the top 10 machines on the Green 500 list are based on accelerator technologies, which are difficult to program and unsuitable for some workloads. In addition, the IBM PowerXCell, which powers six of the top 10, has no public roadmap going forward. For more information on accelerators, see CISE’s recent special issue.

Accelerator architectures tend to be either heavily specialized, such as the PowerXCell and graphics processing units (GPUs), or very generic, such as field programmable gate arrays (FPGAs). Both of these models present real challenges when it comes to programming.

After attending a weeklong class on Compute Unified Device Architecture (CUDA; a popular C-based language for programming graphics processors), one of my coworkers noted, “how easy it is to turn a GPU into a decelerator instead of an accelerator.” With applications commonly exceeding hundreds of thousands to millions of lines of code, programmability is a real issue. More important, however, is the inability of many applications to map to accelerator architectures. At the small scale, this isn’t a concern; a domain-specific platform is perfectly acceptable and possibly even desirable. At the high end, where machine costs can run into the hundreds of millions of dollars, suitability for a wide range of applications is a must.

**System Balance is Key**

These two issues call into question accelerator viability in the largest exascale machines. But if not accelerators, then what? This is still very much an open question and tends to dominate exascale computing discussions. However, as challenging as the node architecture will be, other architecture areas will have an important impact on the efficiency of real applications. This is where both the traditional Top 500 and the Green 500 show their weakness. Both lists are based on the Linpack benchmark, which is notorious for requiring high compute capability, moderate memory performance, and only nominal network performance. The drive to be at the top of these lists will inevitably result in system architectures that favor peak flops over system balance.

Research using Sandia National Laboratories’ Red Storm supercomputer suggests that a machine with a higher peak compute per watt might actually lower energy efficiency for real applications, particularly when the cost of higher peak compute efficiency is an unbalanced system.

The study looked at application performance on two nearly identical Red Storm configurations. The one difference between the two configurations was in the interconnect bandwidth: one configuration used full bandwidth and the other only one-quarter bandwidth. Although the quarter bandwidth configuration was 20 percent more energy efficient when looking at peak flops, it was 10 percent less energy efficient when running CTH, a shock physics code developed at Sandia. This illustrates the need for understanding application requirements and making architectural decisions based on appropriate metrics.

To help meet this need, the US Department of Energy (DOE) Office of Science has established several exascale codesign centers focused on the codesign of exascale applications and architectures. The codesign centers are part of a broader DOE-wide program; one of the program’s main goals is to enable supercomputing systems that can efficiently run DOE mission-critical science and national security applications. The results from these codesign centers could provide important information on how we can modify supercomputers to dramatically improve energy efficiency.

**In This Issue**

The push to exascale and unprecedented levels of energy efficiency will require Herculean efforts, and it’s impossible to cover all aspects of the topic in a single issue. As such, this issue will look at areas whose importance is often overlooked.

In “Money for Research, Not Energy Bills: Finding Energy and Cost Savings in High-Performance Computer Facility Designs,” Dale Sartor and Mark Wilson discuss how thoughtful facility design can greatly reduce energy demands in the machine room and beyond. Their article
describes approaches for improving datacenter efficiency and is a prime example of looking beyond the supercomputer itself to its broader context.

David W. Jensen and Arun F. Rodrigues wrote “Embedded Systems and Exascale Computing” from the perspective of both embedded and high-performance computing. Although energy efficiency is a relatively new driver for HPC, the embedded space has long sought the best balance of energy and capability. As the authors discuss, embedded computing is beginning to require compute capacities that are causing it to tackle many issues historically limited to the HPC domain, while HPC is beginning to have many of the constraints that the embedded world has dealt with for years.

In “Software and Hardware Techniques for Power-Efficient HPC Networking,” Torsten Hoefler looks at opportunities to improve the efficiency of high-performance interconnects. Power will inevitably be a limiting factor to interconnect performance at the largest scale; it’s therefore vital to make the network transport as efficient as possible. Hoefler’s article reports on a power study done on modern interconnects that points to areas of improvement for truly energy-efficient networks.

Finally, in “Advanced Architectures and Execution Models to Support Green Computing,” Richard Murphy, Thomas Sterling, and Chirag Dekate describe research that will be done as part of the DARPA Ubiquitous HPC project. Specifically, it looks at new execution models that can reduce system overheads to the lowest possible levels to increase scalability—and, in so doing, improve the energy efficiency of real applications.

Many other issues must be resolved to enable efficient exascale computing. One of the primary areas for improvement is in making memory systems more energy efficient. Providing sufficient memory performance at reasonable power will require significant changes to both the memory interface and core memory technology. Another aspect of system efficiency is resilience, which impacts how much time is spent running the application versus preparing for and recovering from faults. Methods currently used to deal with system faults are generally considered to be insufficient for exascale machines. Indeed, exascale computing will drive what have been second-order concerns into the spotlight, but none so dramatically as energy efficiency.

References


Scott Hemmert is a senior member of the technical staff at Sandia National Laboratories, where he leads the advanced supercomputer interconnect research. He’s also a member of the joint Sandia/Los Alamos National Laboratory Alliance for Computing at Extreme Scale (ACES) design team, where he serves as co-lead for interconnect architectures. His research interests include supercomputer interconnects and exascale architectures. Hemmert has a PhD in electrical engineering from Brigham Young University. Contact him at kshemme@sandia.gov.