ULPHPC -- Ultra Low Power High Performance Computing towards Exascale

Satoshi Matsuoka, Takayuki Aoki
Tokyo Institute of Technology, Tokyo, Japan
matsu@is.titech.ac.jp, taoki@gsic.titech.ac.jp

Reiji Suda
The University of Tokyo, Tokyo, Japan
reiji@is.s.u-tokyo.ac.jp

Hiroki Honda
Unveristy of Electro-Communications, Tokyo, Japan
honda@is.uec.ac.jp

Michihiro Koibuchi
Natioal Institute of Informatics, Tokyo, Japan
koibuchi@nii.ac.jp

***Abstract*—ULPHPC is a JST-CREST sponsored project where we seek to achieve 1000-fold increase in power efficiency of high performance computing over a 10-year span, which is 10 times beyond what Moore's law provides. We achieve this by employing new devices such as SSDs and architectural components such as the GPU that are much more power efficient than the past aggressively into supercomputers; develop system software and programming substrates, and algorithms to effectively utilize such new devices; and properly model their power-performance characteristics; and finally employing advanced auto-tuning techniques to optimize for best power and energy consumption metrics. Some of the results of our research are already incorporated into TSUBAME2.0, a supercomputer that was developed and deployed at the Tokyo Institute of Technology in Nov. 2010, which was awarded the "The Greenest Production Supercomputer in the World" award by the Green500 committee.**

Supercomputers of the past were "performance at all cost" including power consumption. This is no longer true, i.e. supercomputers require even higher power-performance efficiencies than normal computers. In fact, power is considered to be the biggest impediment for supercomputers as we move forward for more performance to satisfy the requirements of scientific advances being made in simulations.

Our new project "JST-CREST ULP-HPC" aims to resolve this problem by establishing 1000-fold increase in power efficiency of high performance computing over a 10-year span, which is 10 times beyond what Moore's law provides. We achieve this in the following way.

1. We aggressively employ new devices such as non-volatile memory for memory/storage and silicon photonics for networks, and new architectural components such as (General-Purpose) GPUs and SSDs; they are are much more power efficient than what have been employed in the past in traditional supercomputers.
2. Such components often cannot be simply a "drop-in" for existing components; rather they require new system software and programming substrates, and algorithms to effectively utilize such new devices. We have made significant strides with multitudes of results in such regards, such as extensive sets of middleware to employ GPUs at large scale such as attaining high reliability through checkpointing (which is difficult with GPUs) and power-aware scheduling in heterogeneous CPU-GPU environment; programming of GPUs using traditional languages such as OpenMP rather than specialized languages such as CUDA; and power-scheduling of networks to maintain performance while achieving low power. Also basic libraries such as FFT and large-scale BLAS as well as in real applications we have achieved the fastest speed in the world in many instances while being extremely power efficient.
3. Properly model the power-performance characteristics of such software and application substrates with appropriate power-performance models, and develop a framework to combine their metrics for full-machine optimization. Also, develop systems to properly monitor and control the power and other important metrics.
4. And finally employing advanced auto-tuning techniques to optimize for best power and energy consumption metrics. We have developed both theoretical underpinnings of auto-tuning as well as have applied auto-tuning to libraries and applications in a most novel fashion, allowing us to achieve x10 improvement over Moore's law.

Some of the results of our research are already incorporated into TSUBAME2.0, a supercomputer that was developed and deployed at the Tokyo Institute of Technology in Nov. 2010, which become not only the 4th fastest supercomputer in the world on the Top500 list, but also awarded the "The Greenest Production Supercomputer in the World" award by the Green500 committee. It was consecutively re-awarded with the same prize on June, 2011. The achieved metrics of 958 Megaflops/W is 3 times more power efficient than an ordinary laptop. Moreover, large-scale applications that fully utilized TSUBAME2.0 became the two of the five finalists for the ACM Gordon Bell Prize to be decided in Nov., 2011, each of which achieving petascale performance at unprecedented metrics at 1Teraflop/W including cooling.

These power-performance metrics are approximately 30-fold improvements over Tsubame1.0, the previous generation supercomputer we had designed and deployed early 2006 and serving as the base metric. Collectively they demonstrate that we are well on track for 1000-fold improvement as we strive for exascale performance in 2018-2020 timeframe.