The Low-power Architecture Approach Towards Exascale Computing

[Extended Abstract]

Nikola Rajovic, Nikola Puzovic, Lluis Vilanova, Carlos Villavieja, Alex Ramirez
Barcelona Supercomputing Center
Barcelona, Spain
first.last@bsc.es

ABSTRACT

Energy efficiency is a first-order concern when deploying any computer system. From battery-operated mobile devices, to data centers and supercomputers, energy consumption limits the performance that can be offered.

We are exploring an alternative to current supercomputers that builds on the small energy-efficient mobile processors. We present results from the prototype system based on ARM Cortex-A9 and make projections about the possibilities to increase energy efficiency.

Categories and Subject Descriptors

C.1 [**Processor Architectures**]: Parallel Architectures—*Mobile Processors*

General Terms

Design, Performance

Keywords

Exascale, Embedded, ARM

1. INTRODUCTION

Over time, supercomputers have shown a constant exponential growth in performance: according to the Top500 list of supercomputers [1], an improvement of 10x in performance is observed every 3.6 years. The Roadrunner Supercomputer achieved 1 PFLOPS (10^{15} Floating Point Operations per Second) in 2008 [3] on a power budget of 2.3 MW, and the current number one, the K Computer, achieves 10 PFLOPS at the cost of 12 MW.

Following this trend, exascale performance should be easily reached in 2018, but the power requirement will be up to 500 MW.[1] A realistic power budget for an exascale system is 20 MW [7], which requires a 50 GFLOPS/W energy efficiency.

For a long time, the only metric that was used for assessing supercomputer performance was their speed. The Top500 list ranks supercomputers based on their performance when

---

[1] For comparison, the total power of all the supercomputers in the Top 500 list today is around 353 MW [2].

running the Linpack benchmark [4]. However, performance per watt is now as important as raw computing performance: system performance is nowadays limited by power consumption and power density. The Green500 list [2] ranks supercomputers based on their power efficiency. A quick look at this list shows that the most efficient systems today only achieve 2 GFLOPS/W, and that most of the top 50 power-efficient systems are built using heterogeneous CPU + GPU platforms.

A quick estimation, based on using 16 GFLOPS processors (such as those in BlueGene/Q and the Fujitsu K), shows that a 1 EFLOPS system would require 62.5 million such processors. If we observe that only 35-50% of the 20 MW allocated to the whole computer is actually spent on the CPUs, we can also conclude that each one of those processors has a power budget of only 0.15 Watts, including caches, network-on-chip, etc.

Current high-end multicore architectures are one or two orders of magnitude away from that mark. The streaming cores used in GPU accelerators are in the required range, but lack general purpose computing capabilities. A third design alternative is to build a high performance system from energy-efficient components originally designed for mobile and embedded systems.

In this talk we evaluate the feasibility of developing a high performance compute cluster based on the current leader in mobile processors, the ARM Cortex-A9.

First, we describe the architecture of our HPC cluster, built from dual-core Nvidia Tegra2 processors and a 1 Gigabit Ethernet interconnection network. To our knowledge, this is the first large-scale HPC cluster built from ARM multicore processors.

Then, we compare the per-core performance of the Cortex-A9 with a contemporary power-optimized Intel i7[2], and evaluate the scalability and performance per watt of our ARM cluster using the Linpack benchmark.

2. POWER EFFICIENT ARCHITECTURES

As already mentioned, supercomputers already have benefit from an exponential growth in performance, at the cost of an exponential growth in power consumption. Recent years have also seen a dramatic increase in the number, performance and power consumption of servers and data-centers. This market, which includes companies such as Google, Amazon and Facebook, is also concerned with power

---

[2] Both the Nvidia Tegra2 and Intel i7 M640 were released on Q1 2010

Table 1: Power efficiency of several supercomputing systems

SuperComputer system	GFLOPS/W
Blue Gene/Q	1.9
Degima	1.34
Minotauro-Bullx B505 (Xeon+GPU)	1.23
K Computer	0.8
IBM RoadRunner	0.376
Jaguar	0.253
Red Sky - Sun Blade x6275	0.177

efficiency. Frachtenberg et al. [5] present an exhaustive description of how Facebook builds efficient servers for their data-centers, achieving a 38% reduction in power consumption due to improved cooling and power distribution.

As already said, power envelope of future exascale system is 20 MW [7], which leads to a required power efficiency of 50 GFLOPS/W. The current number one supercomputer, the K Computer, achieves only 0.83 GFLOPS/W [2], so a substantial increase in efficiency is required. According to [2], among the most power efficient supercomputers (Table 1) are (a) those based on processors designed with supercomputing in mind, or (b) those based on general purpose CPUs with accelerators (GPU or Cell). An example of the first type is Blue Gene/Q (Power A2, 2.1 GFLOPS/W). Examples of the second type are the Degima Cluster (Intel i5 and ATI Radeon GPU, 1.4 GFLOPS/W) and QPACE SFB TR Cluster - uses PowerXCell 8i and achieves 0.77 GFLOPS/W.

3. PROTOTYPE

The prototype that we are building consists of 256 nodes divided into 32 containers. Each node has one NVIDIA Tegra2 SoC at 1 GHz (dual-core ARM Cortex-A9) and 1 GB of DDR2-667 RAM memory. Nodes are connected using 1 GbE Ethernet in a Tree-alike topology.

3.1 Initial Results

First we evaluated the single-core performance of ARM Cortex-A9 with Dhrystone [9], STREAM [8] and SPEC CPU 2006 [6] benchmark suites. The experiments are executed both on the ARM platform and also on a power-optimized Intel Core i7 laptop.

In terms of performance, on all the benchmarks, the Intel Core i7 outperforms Tegra 2 - on Dhrystone by a factor of nine (factor of five in STREAM). However, due to lower power consumption, Tegra uses less energy to execute the benchmarks due to lower power consumption. The ARM CPU that is used runs at 1 GHz, while Core i7 runs at 2.83 GHz (2.8x bigger). When frequency is factored out from the performance (assuming that, at 1 GHz frequency Core i7 would be 2.8 times slower), the difference in performance becomes smaller: Core i7 would be up to 3.2 times faster. Given the obvious difference in design complexity, the difference in performance is not so big.

Furthermore, we have run the High Performance Linpack benchmark, achieving power efficiency of 200 MFLOPS/W on a single node.

4. CONCLUSIONS

To the best of our knowledge, we are the first to deploy and evaluate a cluster for High Performance Comput-

ing built from commodity energy-efficient embedded components, such as ARM multicore processors. Unlike heterogeneous systems based on GPU accelerators, which require code restructuring and special-purpose programming models like CUDA or OpenCL, our system can be programmed using well known MPI + SMP programming models.

Our results show that the ARM Cortex-A9 in the Nvidia Tegra2 SoC is up to 9 times slower than a mobile Intel i7 processor, but still achieves a competitive energy efficiency. Our results also show that HPC applications can scale up to a high number of processors to compensate for the slower processors with higher parallelism.

Given that Tegra2 is among the first ARM multicore products to implement double-precision floating point, we consider it very encouraging that such an early platform, built from off-the-shelf components achieves competitive energy efficiency results compared to contemporary multicore systems in the Green500 list.

A properly designed system, not a developer board built towards software development for mobile platforms, would save energy on unnecessary peripherals like HDMI and USB ports, and integrate a more efficient memory and network interface, making it a serius contender.

If we account for an upcoming quad-core ARM Cortex-A15 SoC, it could achieve an energy efficiency similar to current GPU systems (1.8 GFLOPS/W). This would provide the same energy efficiency, but using a homogeneous multicore architecture, which appears as a very competitive solution for the next generation of high performance computing systems.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] TOP500: TOP 500 Supercomputer Sites. http://www.top500.org.

[2] The Green 500 list: Environmentally Responsible Supercomputing. http://www.green500.org/, 2011.

[3] K. J. Barker and et al. Entering the petaflop era: the architecture and performance of roadrunner. In *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*, SC '08, pages 1:1–1:11, Piscataway, NJ, USA, 2008. IEEE Press.

[4] J. Dongarra, P. Luszczek, and A. Petitet. The LINPACK Benchmark: past, present and future. *Concurrency and Computation: Practice and Experience*, 15(9):803–820, 2003.

[5] E. Frachtenberg, A. Heydari, H. Li, A. Michael, J. Na, A. Nisbet, and P. Sarti. High-Efficiency Server Design. In *Proceedings of the 2011 ACM/IEEE conference on Supercomputing*. IEEE Computer Society, 2011.

[6] J. Henning. SPEC CPU2006 benchmark descriptions. *ACM SIGARCH Computer Architecture News*, 34(4):1–17, 2006.

[7] K. Bergman et al. Exascale Computing Study: Technology Challenges in Achieving Exascale Systems. In *DARPA Technical Report*, 2008.

[8] J. McCalpin. A survey of memory bandwidth and machine balance in current high performance computers. *IEEE TCCA Newsletter*, pages 19–25, 1995.

[9] R. Weicker. Dhrystone: a synthetic systems programming benchmark. *Communications of the ACM*, 27(10):1013–1030, 1984.