Journal of Computational Physics 237 (2013) 132–150

Contents lists available at SciVerse ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Energy efficiency vs. performance of the numerical solution of PDEs: An application study on a low-power ARM-based cluster

Dominik Göddeke [a,*], Dimitri Komatitsch [b], Markus Geveler [a], Dirk Ribbrock [a], Nikola Rajovic [c,d], Nikola Puzovic [c], Alex Ramirez [c,d]

[a] Institut für Angewandte Mathematik, Fakultät für Mathematik, TU Dortmund, Germany
[b] Laboratory of Mechanics and Acoustics, CNRS/University of Aix-Marseille, France
[c] Computer Sciences Department, Barcelona Supercomputing Center, Barcelona, Spain
[d] Departament d'Arquitectura de Computadors, Universitat Politècnica de Catalunya, BarcelonaTech, Barcelona, Spain

ARTICLE INFO

Article history:
Received 28 July 2012
Received in revised form 14 October 2012
Accepted 21 November 2012
Available online 7 December 2012

Keywords:
High performance computing
Energy efficiency
Low-power processors
ARM processors
Parallel scalability
Finite elements
Multigrid
Wave propagation
Lattice-Boltzmann

ABSTRACT

Power consumption and energy efficiency are becoming critical aspects in the design and operation of large scale HPC facilities, and it is unanimously recognised that future exascale supercomputers will be strongly constrained by their power requirements. At current electricity costs, operating an HPC system over its lifetime can already be on par with the initial deployment cost. These power consumption constraints, and the benefits a more energy-efficient HPC platform may have on other societal areas, have motivated the HPC research community to investigate the use of energy-efficient technologies originally developed for the embedded and especially mobile markets. However, lower power does not always mean lower energy consumption, since execution time often also increases. In order to achieve competitive performance, applications then need to efficiently exploit a larger number of processors. In this article, we discuss how applications can efficiently exploit this new class of low-power architectures to achieve competitive performance. We evaluate if they can benefit from the increased energy efficiency that the architecture is supposed to achieve. The applications that we consider cover three different classes of numerical solution methods for partial differential equations, namely a low-order finite element multigrid solver for huge sparse linear systems of equations, a Lattice-Boltzmann code for fluid simulation, and a high-order spectral element method for acoustic or seismic wave propagation modelling. We evaluate weak and strong scalability on a cluster of 96 ARM Cortex-A9 dual-core processors and demonstrate that the ARM-based cluster can be more efficient in terms of energy to solution when executing the three applications compared to an x86-based reference machine.

1. Introduction and motivation

Energy efficiency and power consumption have become one of the most critical issues regarding the design and deployment of a high performance computing (HPC) facility, or a data centre in general. While the combined power consumption of HPC systems worldwide continues to be small in relative terms (2% of the total CO_2 emissions), in absolute terms the cost is

* Corresponding author.
E-mail addresses: dominik.goeddeke@math.tu-dortmund.de (D. Göddeke), komatitsch@lma.cnrs-mrs.fr (D. Komatitsch), markus.geveler@math.tu-dortmund.de (M. Geveler), dirk.ribbrock@math.tu-dortmund.de (D. Ribbrock), nikola.rajovic@bsc.es (N. Rajovic), nikola.puzovic@bsc.es (N. Puzovic), alex.ramirez@bsc.es (A. Ramirez).
URLs: http://www.mathematik.tu-dortmund.de/~goeddeke (D. Göddeke), http://komatitsch.free.fr/ (D. Komatitsch).

0021-9991/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jcp.2012.11.031

already extremely high, around 200–300 billion kW h [1,2]. This trend will further increase in the near future, and the cost of energy is also increasing, which will soon lead to a critical situation owing to energy efficiency limits [3].

The most immediate concern is that the energy cost of an HPC installation over its lifetime (5–7 years) is already comparable to its initial cost of acquisition and deployment, see Section 3.2. This is important on all scales, and in both academia and industry. In academia for instance, it is common – at least for medium-scale installations hosted by a single university – that (governmental or regional) funding is only provided for the initial deployment of a machine, and the university or research institution has to cover the total cost of running the machine for its lifetime. For instance it was recently announced that the state of New Mexico, USA, would discontinue operations of its 'Encanto' system, which ranked #3 in the TOP500 list when it was deployed in 2007, due to lack of maintenance funds.[1]

Since 2007, the Green500 list [4] ranks supercomputers similar to the well known TOP500 list, but based on the energy efficiency (performance per watt) of the systems. Extrapolating the energy efficiency of the #1 system in the current Green500 list, a postulated 1 exaflops system, i.e., 10^{18} operations per second, would require 500 MW of power, corresponding to more than 500 million dollars per year for electricity alone. Many studies predict a feasible power envelope of 20 MW for exascale machines, which requires a 25× efficiency improvement over the current Green500 leader and a 50× improvement over current standard x86-based systems [5,6].

An analysis of current HPC systems shows that 40–60% of the energy consumption can be attributed to the compute nodes (processors and memories), 10% to the interconnect and storage systems, and the remainder (up to 50%) is consumed by the infrastructure itself, including lighting, power supply, and most of all, cooling [7,8]. Consequently, the largest returns can be expected from improving the energy efficiency of the compute nodes, since any improvement there will also translate into reduced cooling requirements.

It is also possible to improve the energy efficiency of a computer system by offloading parts of the application to a more specialised hardware accelerator that achieves higher performance at a lower energy cost for a particular type of computation. Current representatives of such heterogeneous compute systems couple conventional x86-based CPUs with GPU accelerators, or the recently introduced Xeon Phi. However, such improvements come at the cost of major changes in the application codes, which must be partitioned to offload computation to the accelerator and partly rewritten in the accelerator specific programming model. While these heterogeneous architectures offer significant improvements in energy efficiency, the investment required to rewrite production codes that have been used for years or decades is often difficult to handle, which increases the pressure to find more efficient general purpose solutions.

There are several ongoing research proposals and industrial initiatives that suggest the use of devices originally designed for the embedded and mobile markets in a high performance computing environment. These devices have been designed from scratch to operate in energy and temperature constrained environments such as cell phones, which operate on batteries and are stored in our pockets. Recent developments in mobile processors have included the addition of double-precision floating point units, which is a necessary condition for them to become accepted in HPC. Their higher energy efficiency as well as their large market volume, which drives prices down, make these devices a promising candidate in terms of performance per watt and performance per dollar. Nowadays, the majority of such mobile devices are built on the intellectual property of ARM Holdings plc. For instance, ARM's 2011 annual report to investors [9] quotes constant market shares of 95% for handheld devices (smartphones and tablet computers) since 2008.

For these devices and systems built from them, the software environment is favourable to porting production codes: The GNU/Linux tool chain has been supported on ARM processors for a long time, together with all major programming and scripting languages such as Fortran, C/C++ or Python, as well as parallel programming models such as MPI and OpenMP. The effort of getting an application up and running on an ARM-based system is thus identical to the effort on any classical architecture, i.e., usually small. The required investment in manual tuning for specific (micro-) architectural features of these processors is also comparable between classical and ARM designs, as their architecture is not as radically different as for instance that of GPUs. In this article, we discuss a range of tuning strategies that are particularly important.

Low power also often means low performance: Mobile devices offer 10 to 100 times lower peak performance than their high-end counterparts, but they do so at 100 to 1000 times lower power. As an example, the chips that we use in this work deliver a theoretical peak performance of 2 GFLOP/s at roughly 0.5 W, while our reference x86 designs require a TDP (thermal design power) of 95 W for a theoretical peak performance of 42.6 GFLOP/s. Thus, owing to the longer execution time, lower power may not translate into lower energy. To achieve the degree of performance required by actual applications, many low-power cluster nodes have to be used in one system. Very good weak and strong scalability of the (numerical) methodology *and* the implementation are required simultaneously. Only then is it possible to distribute large 'real-world' problems over more nodes with less memory each and to compensate for lower single-node performance by using more nodes.

For most non-trivial PDE (partial differential equations) applications this is highly challenging to achieve simultaneously, and the behaviour is strongly application dependent. In this article, we employ a prototype ARM-based cluster that, given its size, allows us to quantitatively analyse the power-performance trade-off (measured in time vs. energy to solution) for the first time at a scale of interest for future HPC systems and for production codes solving real-world problems. We perform this

[1] HPC Wire, July 12, 2012, 'New Mexico to Pull Plug on Encanto, Former Top 5 Supercomputer', http://www.hpcwire.com/hpcwire/2012-07-12/new_mexico_to_pull_plug_on_encanto_former_top_5_supercomputer.html.