



Laboratoire d'Informatique de Grenoble

UMR 5217 - CNRS, INPG, INRIA, UJF, UPMF

Téléphone : (+33) 4 76 61 20 64
Télécopie : (+33) 4 76 61 20 99
Adresse électronique : <mailto:arnaud.legrand@imag.fr>

Adresse : Inria Grenoble Rhône-Alpes
Inovallée
655 avenue de l'Europe
Montbonnot Saint Martin
38334 St Ismier Cedex
France

Arnaud Legrand, CNRS researcher
INRIA **MESCAL** project, LIG Laboratory

Proposal for a M2R internship Analysis and Modeling of Cache Performance for Exascale Computing Platforms

Advisors: Arnaud Legrand, Luka Stanisic and Brice Videau

Required Skills:

- C programming, UNIX, shell, ssh, git
- Knowledge of ruby and basics of experiment analysis with R is a plus

1 Context

There is a continued need for higher compute performance: scientific grand challenges, engineering, geophysics, bioinformatics, etc. Such studies used to be carried out on large *ad hoc* supercomputers, which, for economical reasons, were replaced by commodity clusters, i.e., sets of off-the-shelf computers interconnected by fast switches. Indeed, the technological advances driven by the home PC market have contributed to achieving high performance in commodity components. For decades, computer performance had doubled every 18 months merely by increasing the clock frequency of the processors. This trend stopped last decade for reason of electricity consumption and heat. Indeed, the computational power of a computer increases nearly sub-linearly with clock frequency while the energy consumption increases more than quadratically.

As an answer to the power and heat challenges, processor constructors have increased the amount of computing units (or cores) per processor. Modern High Performance Computing (HPC) systems comprise thousands of nodes, each of them holding several multi-core processors. For example, one of the world fastest computers, the [IBM Sequoia system](#) Laurence Livermoor National Laboratory (USA), contains 96 racks of 98,304 nodes comprising 16-core each, for a total of 1,572,864 cores. The [Cray Titan system](#) at Oak Ridge National Laboratory is made of 18,688 AMD Opteron (16-core CPUs) and 18,688 Nvidia Tesla K20X GPUs. More recently, the [Tianhe-2](#) was built with 32,000 Intel Xeon (12 cores) and 48,000 Xeon Phi 31S1P.

Recent evolutions amongst the world's fastest machines¹The top 500 ranking. The [top 500](#) confirm the trend of massive level of hardware parallelism and heterogeneity. Researchers envision systems with billions of cores (called [ExaScale](#) systems) for as early as the next decade, which will tackle through simulation major issues such as the characterization of the abrupt climate changes, understanding the interactions of dark matter and dark energy or improving the safety and economics of nuclear fission.

Despite all these efforts, energy is increasingly becoming one of the most expensive resources and the dominant cost item for running a large supercomputing facility. In fact the total energy cost of a few years of operation can almost equal the cost of the hardware infrastructure. It is unanimously recognized that Exascale systems will be strongly constrained by energy efficiency.

The analysis of the performance of HPC systems since 1993 shows exponential improvements at the rate of one order of magnitude every 3 years: One petaflops was achieved in 2008, one exaflops is expected in 2020. Based on a 20 MW power budget, this requires an efficiency of 50 GFLOPS/Watt. However, according the [Green 500](#), the current leader in energy efficiency achieves only 4.3 GFLOPS / Watt. Thus, a 12x improvement is required.

¹{

2 Environment

The members of the MESCAL team focus their research on large scale systems and parallel applications. They have a strong expertise regarding parallel applications and environment for parallel programming, performance evaluation of large scale distributed systems, middleware for clusters and grids and scheduling.

Some of MESCAL members are also involved in the [Joint Laboratory for Petascale Computing](#) between [University of Illinois at Urbana-Champaign](#), [Inria](#), [Argonne National Laboratory](#), [Illinois' Center for Extreme-Scale Computation](#), and the [National Center for Supercomputing Applications](#), and the [Barcelona Supercomputer Center](#).

Some of their members are also involved in the European [Mont-Blanc](#) (European scalable and power efficient HPC platform based on low-power embedded technology). Indeed, HPC systems developed from today's energy-efficient solutions used in embedded and mobile devices are an interesting alternative to accelerators such as GPUs or Intel Xeon Phi. As of today, the CPUs of these devices are mostly designed by ARM. However, ARM processors have not been designed for HPC, and ARM chips have never been used in HPC systems before, leading to a number of significant challenges. One envisioned possibility for designing such exascale platforms is the use of 100,000+ ARM processors connected through hierarchical Ethernet networks.

3 Goal

The Mont-Blanc project aims at building a super-computer based on the ARM technology. It will not be able to reach Exascale but is a first step toward this direction. ARM processors have almost no memory hierarchy and may thus be easier to model than more recent multicore architectures used in PetaScale architectures.

The goal of this internship is thus to experiment with multi-core ARM processors similar to the ones that may be used in future Exascale platforms and to derive from such experimentation a performance and power consumption model. Measuring the parameters of such model will require to design specific micro-benchmarks that reach the peak performance of these processors, which in turns requires exploring a wide parameter space of code compiling options, loop unrolling and vectorizing options. To ease such code generation, we will rely on the [BOAST](#) framework that allows to meta-program computing kernels and their optimizations.

Such model will then be implemented in the [SimGrid](#) toolkit to try to evaluate the performance of simple benchmarks like LINPACK on such platforms. Such study should provide invaluable feedback to Exascale platform architect in particular with respect to the network and CPU provisioning aspect.