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

Performance Analysis of Congestion in Commodity Networks for Exascale Computing Platforms

Advisors: Arnaud Legrand, Augustin Degomme
Required Skills:

- C programming, UNIX, shell, ssh, MPI, git
- Good understanding of the TCP protocols and ability to recompile the linux kernel
- 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) interconnected by a custom 5-dimensional torus topology. 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 interconnected with a Gemini three-dimensional torus. More recently, the Tianhe-2 was built with 32,000 Intel Xeon (12 cores) and 48,000 Xeon Phi 31S1P interconnected through a custom (TH Express-2) fat-tree topology.

Recent evolutions amongst the world's fastest machines confirm the trend of massive level of hardware parallelism and heterogeneity. Most high end systems rely on *custom* or infiniband interconnect while lower hand systems often rely on 1G or 10G Ethernet. 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.

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 commodity low-power hardware such as the ARM and the Ethernet technology. It will not be able to reach Exascale but is a first step toward this direction.

Although there are numerous articles explaining how to tune or modify TCP to optimize end-to-end bandwidth in a HPC context, most of them completely ignore flow control and congestion management aspects. Our initial experiments with TCP/Ethernet revealed some non-trivial network congestion phenomenon in the presence of intensive HPC workload [BDG+13]. More precisely, some MPI collective operations and the NAS PB conjugate gradient can lead, even at moderate scale, to critical congestion situations that force the whole application to freeze until a TCP timeout is triggered. Such blocking situations only last for 200 milliseconds but such this is unacceptable on a platform where typical communication times are of the order of a few microseconds. Indeed, if such timeouts are harmless in a wide area context over the Internet they can occur systematically in such HPC workloads and slow down the whole execution by more than 30%!!! Beyond a simple application slowdown, it should be understood that the whole set of processors, even when forced to idleness, keeps consuming energy uselessly. Given the scale evolution of HPC platforms, such issue is likely to quickly become impossible to overlook.

There are several options to circumvent this issue:

1. Such timeouts seem to be link to the retransmission timeout (TCP RTO), which could be decreased. The value of this timeout (200ms) is unfortunately hardcoded in the kernel. Although having kernels specifically tuned for HPC workloads is perfectly legitimate, it is quite difficult to know how to set such value. Indeed, aggressively decreasing this timeout may trigger massive retransmissions that could make the network collapse even more.

2. Since most TCP variants are likely to run into the same kind of issues, another option could be to use a completely different protocol like SCTP that manages congestion in a very different way.

3. Another option could be to have user space deamons that smartly inject packets in the system to prevent such timeouts to occur.

The goal of this project is to evaluate these different options. Experiments will be conducted on the Grid5000 platform (probably on Nancy's Graphene cluster). In a first step, the student will familiarize with the experimental procedures (custom image deployment, MPI execution) and on the techniques to control their experimental environment (e.g., recompiling the kernel and preparing custom images with Kameleon).

If time allows, similar experiments will be conducted on the Mont-Blanc prototype. Such study should provide invaluable feedback to Exascale platform architects in particular with respect to the network and CPU provisioning aspect.

References

[BDG$^+$13] Paul Bedaride, Augustin Degomme, Stéphane Genaud, Arnaud Legrand, George Markomanolis, Martin Quinson, Mark Stillwell, Lee, Frédéric Suter, and Brice Videau. Toward better simulation of MPI applications on Ethernet/TCP networks. In *4th International Workshop on Performance Modeling, Benchmarking and Simulation of HPC Systems (PMBS)*, November 2013.