Solving Scalability Issues in SimGrid Network Simulation

Master Internship

1. Context and Contact

Simulation is a classical approach for scientifically evaluating the quality of algorithms and heuristics. It offers complete control over experimental conditions, allowing reproducible experiments by which one can compare competing designs and solutions. The widespread use of simulation in several scientific communities has led to accepted tools and methodologies, for example in the areas of microprocessor design and network protocol design.

The emergence of methodological tools for simulation is unfortunately still to be seen in the communities of high-performance computing (HPC), parallel computing (encompassing Grid and Cloud Computing), and of distributed computing (encompassing Volunteer Computing and Peer-to-Peer research). Since no standard tool exists, most people build their own "ad-hoc" solutions. This approach was certainly reasonable in the past, e.g., for simple and deterministic platforms (such as supercomputers) or when timing realism was not mandatory (as in most works in distributed computing). However, this approach becomes hardly justifiable for today's systems. Indeed, these systems exhibit very complex behaviors because, for example, of their heterogeneity (both quantitative and qualitative), and of their scale. Likewise, as the amount of data exchanged in distributed algorithms increases, it becomes mandatory to obtain simulation timings that match real-world timings accurately.

SimGrid [CLQ08] is recognized in the HPC community as one of the most prominent simulation environments as shown by its large community of users and the number of publications that use it.

The member of the SimGrid core team are Henri Casanova (University of Hawai'i, USA), Arnaud Legrand (INRIA-MESCAL, Grenoble), Martin Quinson (INRIA-Algorille, Nancy), and Frédéric Suter (CC IN2P3, Lyon). The internship will take place in the Algorille INRIA project team under the direction of Martin Quinson and Arnaud Legrand. It will last from 4 to 6 months.

2. Description

The core component of a simulator is the underlying model used to represent the real world. Over the years, several classes of distributed computation platform models have been proposed. **Microscopic models** are used very commonly to represent the network. In these models, the movements of each and every packets sent over the network are computed and simulated [ns, Ril03]. The accuracy of such models is rarely disputed, but they suffer of poor simulation performance and thus lead to very large simulation times [FC07]. Several **macroscopic models** have been proposed [BG87, MR99]. These models provide a good compromise between accuracy and simulation speed [VL09]. At the other end of the spectrum, **simplistic models** are often used in distributed and P2P projects. For example, PlanetSim [GPM+05] uses a model by which any communication takes the same amount of time. Pastry [RD01] uses a model by which the nodes are placed on a plane and the communication time between two nodes is proportional to the Euclidean distance between the involved nodes. While these models allow for impressive scalability, it is clear that they are not realistic.

The SimGrid project aims at providing a simulation framework that provides several models, allowing the users to run simulations at various scales, using several models with different levels of accuracy.

SimGrid was initially developed by people from the parallel algorithms and high performance computing community. Thus, the current models in SimGrid enable to account for complex topology and network sharing strategies. To this end, a complex routing table is stored and thus it is currently impossible to simulate a platform with more than a few thousands of nodes, regardless of the simulated application. The purpose of this internship is twofold:

- First, we need to get rid of the comprehensive routing table. De Munck *et al.* have proposed [MVB09] to recompute dynamically this routing using shortest-path algorithms. However, this approach suffers both from performance and scalability issues. Recently, a new platform description formalism to handle very large platforms has been proposed and implemented within SimGrid [FQS08]. This formalism enables to hierarchically describe the platform and the routing through rules. It is possible to take advantage of such a description to build a small memory footprint description of the platform that would replace the current comprehensive routing table.

This improvement will enable to keep using the same realistic flow-based models as earlier but with platforms that are many orders of magnitude larger than what is currently possible.

- Second, such "precise" flow-based models make sense for grid platform but it is likely that they do not make sense anymore at a very large scale (e.g., for a Volunteer Computing platform). It is thus essential to design more simple models more in the spirit of those that can be found in simulators that have been developed by researchers from the distributed algorithms community.

References

[BG87] Dimitri P. Bertsekas and Robert G. Gallager. *Data Networks*. Prenctice Hall, 1987.

[CLQ08] Henri Casanova, Arnaud Legrand, and Martin Quinson. SimGrid: a Generic Framework for Large-Scale Distributed Experiments. In *10th IEEE International Conference on Computer Modeling and Simulation*. IEEE Computer Society Press, March 2008.

[FC07] Kayo Fujiwara and Henri Casanova. Speed and accuracy of network simulation in the SimGrid framework. In *Proceedings of the First International Workshop on Network Simulation Tools (NSTools)*, 2007.

[FQS08] Marc-Eduard Frincu, Martin Quinson, and Frédéric Suter. Handling Very Large Platforms with the New SimGrid Platform Description Formalism. Technical Report RT-0348, INRIA, 2008.

[GPM+05] Pedro García, Carles Pairot, Rubén Mondéjar, Jordi Pujol, Helio Tejedor, and Robert Rallo. Planetsim: A new overlay network simulation framework. In *Software Engineering and Middleware, SEM 2004*, volume 3437 of *LNCS*, pages 123–137, Linz, Austria, March 2005.

[MR99] Laurent Massoulié and James Roberts. Bandwidth Sharing: Objectives and Algorithms. In *INFOCOM*, volume 3, pages 1395–1403, 1999.

[MVB09] S. De Munck, K. Vanmechelen, and J. Broeckhove. Improving the scalability of SimGrid using dynamic routing. In *ICCS*, 2009.

[ns] The network simulator - ns-2. http://www.isi.edu/nsnam/ns.

[RD01] Antony Rowstron and Peter Druschel. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In *IFIP/ACM International Conference on Distributed Systems Platforms (Middleware)*, pages 329–350, November 2001.

[Ril03] George F. Riley. The Georgia Tech Network Simulator. In *ACM SIGCOMM workshop on Models, Methods and Tools for Reproducible Network Research*, pages 5–12, 2003.

[VL09] Pedro Velho and Arnaud Legrand. Accuracy study and improvement of network simulation in the SimGrid framework. In *SIMUTools'09, 2nd International Conference on Simulation Tools and Techniques*, 2009.