Emulation of Arbitrary Applications with SimGrid

Advisor: Martin Quinson (IRISA, team Myriads). Martin.Quinson@ens-rennes.fr
Duration: 4 to 6 months. Follow-up possible in PhD thesis.

Executive summary: The goal of this project is to design an evaluation environment for distributed applications and infrastructures where the instances of the real application are executed in a virtual environment simulated by the SimGrid simulator. Typical target applications include the Storm Event Processing Infrastructure and the Ceph Distributed Storage.

Context

Distributed systems such as grids, peer-to-peer systems, cloud computing infrastructures or desktop computing environments, benefit of an ever increasing popularity nowadays. Distributed applications (such as decentralized data sharing solutions, games, scientific applications, high-traffic web applications) are executed routinely on these systems.

By nature, the resulting environments and applications are extremely complex and dynamic because they aggregate thousands of elements that are heterogeneous and shared among several users. This make these systems very challenging to study, test, and evaluate. Purely theoretical studies rely on assumptions that are at best simplistic and often unrealistic. Most of the studies are thus done through experiments, often on dedicated facilities such as Grid'5000.

But the recent evolution of the target systems in size, dynamicity and complexity makes it difficult to even test the infrastructures in a reliable and reproducible manner. An appealing alternative is to rely on simulation to evaluate the algorithms but this approach is not suited to real applications.

SimGrid (developed by the tutor in nation-wide collaboration) is a toolkit providing core functionality for the simulation of distributed applications in heterogeneous distributed environments. The specific goal of the project is to facilitate research in the area of distributed and parallel application scheduling on distributed computing platforms ranging from simple network of workstations to Computational Grids. It is possible to study MPI applications directly on SimGrid as the standard is reimplemented on top of the simulator. For other interfaces, users have to extract the logic of their applications and rewrite them using the specific interfaces of SimGrid.

The Simterpose project, which is the core of this proposal, tries to intercept the actions of real applications to mediate them through the simulator. This would allow to run unmodified arbitrary applications on top of SimGrid. For that, Simterpose intercepts all system calls and mediate the communications according to the results computed by the simulator while the computations are only benchmarked so that their timings can be injected into the simulator.

Description

Several proof of concepts were developed by previous interns, but many problems remain to be solved, both on the theoretical and practical sides. The proposed thesis is expected to lead to several contributions that fall into two main domains.

First, the applicant is expected to contribute a solid framework, properly interfaced in the
Linux operating system, to mediate any actions of arbitrary distributed applications. The ex-
isting proof of concepts (based on ptrace) will certainly guide the work, but other novel ap-
proaches (such as uprobe, seccomp/BPF or others) should be evaluated after an adequate bibliographical work on virtualization techniques. The designed infrastructure should then be thoughtfully implemented to enable a practical evaluation of the possible approaches.

Then, the SimGrid models will have to be reviewed and improved to improve the simulation accuracy. In particular, the existing network models are already accurate enough to run MPI applications on high performance LAN, but they still need to be carefully assessed on the targeted WAN infrastructures.

This work is motivated by the need to assess modern distributed infrastructures, such as the Ceph distributed storage solution, the Storm event processing system, the Samba networked disk server, or others. As such, it should be assessed on these use cases that are representative of the typical workload experienced in the IT industry. Other use cases from the HPC community in which SimGrid is already used will also be considered.

Goals of the Internship

- Conduct a throughout bibliographical study of the field. The goal is to gain a better understanding of the system parts that Simterpose should emulate (CPU, communication, DNS, threads, etc), and the possible approaches for each aspect.

- Implement the selected approaches. This development is necessary to the practical evaluation of the contribution. The goal is to only develop a proof of concept, that will be technically reinforced by an engineer afterward.

- Evaluate the feasibility of the contribution. An evaluation framework should be designed, with the selection beforehand of the applications and workload. The goal is to study which approaches are possible to lightly virtualize distributed applications. The emulation performance (that is, the overhead of the virtualization) should also be assessed.

- Evaluate the correctness of our models in this context. The simulation results should be compared to measurements on real platforms such as Grid’5000, and (if time permits) the SimGrid models should be modified if they prove unadapted to this new context.

Skills required

In addition to the skills that can reasonably be expected from Master-level students, the applicant should have a very strong knowledge of system programming in C, and of Linux and other modern Unix-based Operating Systems.

Links