A Proposal for a Simulation of PCAD Network Topology Using SimGrid

Rayan Raddatz de Matos[1], **Lucas Mello Schnorr**[1]

[1] Institute of Informatics, Federal University of Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

{rayan.raddatz, schnorr}@inf.ufrgs.br

Abstract. *Simulation is a well-established methodology for evaluating High-Performance Computing (HPC) systems, ensuring reproducibility and reduced energy footprints. This work proposes a programmatic network model of the PCAD infrastructure, a heterogeneous institutional cluster, using the SimGrid S4U API. Focusing on four core partitions, we map the physical architecture into a simulated structure and calibrate the platform properties with experimentation. The main goal is to enable local researchers to replicate experiments and study scalability even with limited physical hardware.*

1. Introduction

Computer systems performance analysis can be done following three different methods, analytical modeling, simulation or measurement. The choice of which method to use when performing analysis depends on a variety of factors such as the time available, the desired accuracy and several others [Jain 1991]. While measurement tends to show the most accurate results when properly configured, it requires access to the actual system, which typically leads to higher energy consumption and limits the ability to extrapolate to new scenarios. On the other hand, analytical modeling offers great extrapolation capabilities at a considerably lower energy consumption, but often suffers from low accuracy results. Simulation acts as a middle term, maintaining good accuracy with a reduced energy footprint, while also enabling the exploration of new what-if scenarios. Furthermore, employing simulations also tackles the problem of the lack of reproducibility in High-Performance Computing [Antunes and Hill 2024], ensuring a controlled and reproducible environment for experiments.

Therefore, the simulation of High-Performance Computing (HPC) infrastructures has become a well-established methodology for evaluating system behavior. Among the available tools, SimGrid emerges as a free software capable of simulating several HPC scenarios with reproducibility and versatility, allowing researchers to simulate large platforms while maintaining accuracy [Casanova et al. 2025].

In this work, we propose a model of the network of the "Parque Computacional de Alto Desempenho" infrastructure [INF/UFRGS 2025] through the SimGrid programmatic interface Simgrid For You (S4U). We show preliminary performance results between the proposed platform created through simulation and the real cluster.

Software and Data Availability. We endeavor to make our analysis reproducible for a better science. We made available a companion material hosted in a public GitHub repository at `https://github.com/rddtz/simgrid-pcad`. Our companion contains the source code of this article and the programmatic description of the platform. We also include instructions to run the real experiment, simulations and to recreate the figures.

2. Related Work

The evaluation of High-Performance Computing (HPC) systems often requires balancing accuracy, time, and energy consumption. To achieve this balance and simultaneously address the lack of reproducibility in HPC experiments, simulation has become a well-established methodology. Frameworks like SimGrid offer the versatility needed for these tasks. Specifically, SimGrid provides the S4U API, which enables the programmatic creation of complex simulators. Recent works demonstrate the power of this API for modeling state-of-the-art environments. As an example, [Melo 2025] successfully used the S4U API to model and simulate the network topology of the Frontier supercomputer [Top500 2022]. However, while existing works often focus on modeling massive systems, there is a need to apply and validate these programmatic creation methodologies in the context of institutional HPC clusters. Moreover, by creating a validated model of the PCAD infrastructure, it enables local researchers to easily replicate experiments and create what-if scenarios, as well as study scalability even with limited physical hardware.

3. Platform Proposal and Initial Implementation

3.1. Platform Proposal

Figure 1 illustrates the PCAD network topology. The PCAD infrastructure is composed of four racks and contains a total of 47 machines separated into 19 different partitions. Furthermore, the distribution of network capacities varies across the infrastructure. For instance, while Rack 1, Rack 3, and Rack 4 utilize 1Gb switches, Rack 2 presents a mixed network capacity, containing both a 1Gb and a 10Gb switch. While the complete PCAD platform is composed of these multiple different partitions, we initially selected only a few partitions to simulate, these being the `cei`, `draco`, `tupi`, and `poti` partitions. The first three partitions have six nodes each, and the last one has only five. We chose these partitions as they are the most used ones, aiming to first create a generic approach that can later be applied to the whole cluster if necessary.

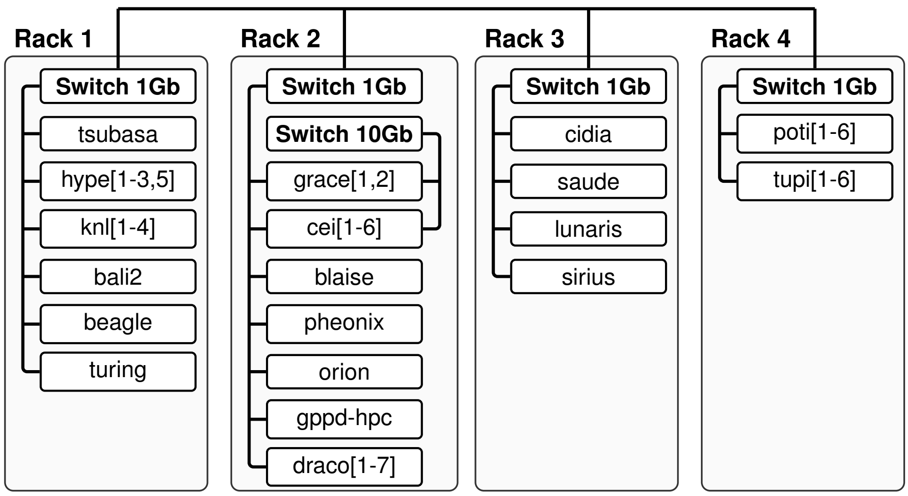

Figure 1. Network topology of PCAD.

After defining the platform scope, the SimGrid for You (S4U) C++ API with SimGrid v4.1 was used to programmatically implement the network model. The physical architecture was mapped into a hierarchical simulated structure using S4U's Star Zones. Each rack was therefore defined as a distinct network zone and within these zones individual host nodes are instantiated. The modeling explicitly defines hardware characteristics for each node, including physical core counts and computational speed measured in Gflops. Network capabilities are established by attaching dedicated links to each host, specifying both bandwidth and latency.

3.2. Platform Calibration Methodology

After having a programmatic model of the network topology, the next stop is to calibrate it. The calibration consists in the execution of several experiments in the target system to gather information about its capabilities. The essential information for each host is the host computing power and connection bandwidth and latency. The information about the computing power can be achieved by various ways, we chose to use the High-Performance Linpack (HPL) [Dongarra et al. 2003]. For each partition, the HPL was executed 10 times across two different machines, and the mean result in Gflops was taken as the speed of a node in that partition. For the network properties, we used simple network tools such as `iperf3` for bandwidth and `ping` for latency. Official SimGrid material was also used to calibrate the network and MPI factors for each partition [Cornebize 2021].

4. Preliminary Results

We used the LULESH proxy app [Karlin et al. 2013], a common benchmark for MPI communication and network performance, to evaluate an initial implementation. Figure 2 shows the mean time for execution of different experiment sizes (represented as a bigger circle) as a function of the number of nodes. Small circles represent our observations. The time is given by the application as output at the end of the computation. The black bars centered around the mean represent the standard error with a CI of 99%. The figure shows the results faceted by partition, with the colors identifying where the execution takes place, in the real cluster (`real`) or in the SimGrid (`sim`). The simulation was executed in a node of the `cei` partition.

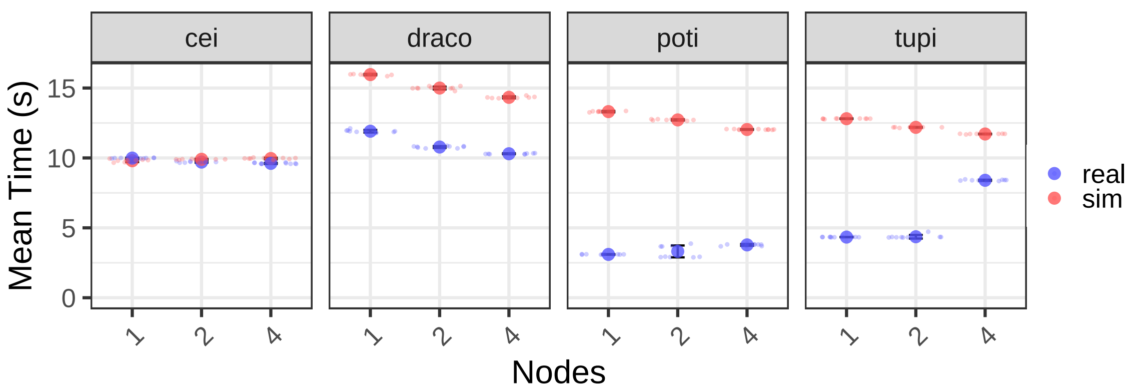

Figure 2. Mean execution time of the LULESH benchmark for different partitions.

As observed in the results, there is currently a noticeable gap between the simulated execution times and the real cluster performance across the partitions. This initial discrepancy is an expected behavior during the early stages of modeling a heterogeneous environment. While the initial calibration successfully captured host computing power and link limits for some partitions, the complex behavior of the network requires further refinement. Consequently, the next necessary step is to perform a deeper, more granular calibration, particularly focusing on adjusting the network and MPI factors to accurately simulate communication overhead and contention.

5. Conclusion

This work proposes a programmatic methodology to model the heterogeneous PCAD infrastructure using the SimGrid S4U API. We established a hierarchical baseline and an initial calibration workflow. Preliminary results show that while static hardware capabilities are represented, the network behavior still diverges from the physical cluster. Future work focuses on an iterative fine-grained calibration of MPI and network parameters.

Acknowledgements

We would like to thank the PCAD at INF/UFRGS for making infrastructure and hardware used in the experiments available. We also acknowledge the Brazilian National Council for Scientific Technological Development (CNPq) for their financial support through the PIBIC-UFRGS scholarship.

References

Antunes, B. and Hill, D. R. (2024). Reproducibility, replicability and repeatability: A survey of reproducible research with a focus on high performance computing. *Computer Science Review*, 53:100655.

Casanova, H., Giersch, A., Legrand, A., Quinson, M., and Suter, F. (2025). Lowering entry barriers to developing custom simulators of distributed applications and platforms with SimGrid. *Parallel Computing*, 123:103–125.

Cornebize, T. (2021). *High Performance Computing : towards better Performance Predictions and Experiments*. Theses, Université Grenoble Alpes [2020-....].

Dongarra, J. J., Luszczek, P., and Petitet, A. (2003). The LINPACK benchmark: past, present and future. *Concurrency and Computation: Practice and Experience*, 15(9):803–820.

INF/UFRGS (2025). Laboratório de processamento paralelo e distribuído (lppd). `https://pcad.inf.ufrgs.br/`. Accessed: 2026-02-16.

Jain, R. (1991). *The Art of Computer Systems Performance Analysis: Techniques For Experimental Design, Measurement, Simulation, and Modeling, NY: Wiley*.

Karlin, I., Keasler, J., and Neely, J. R. (2013). Lulesh 2.0 updates and changes. Technical report, Lawrence Livermore National Laboratory (LLNL), Livermore, CA (United States).

Melo, A. A. F. (2025). Um simulador da rede do supercomputador frontier utilizando simgrid. Bachelor's thesis.

Top500 (2022). Frontier. `https://www.top500.org/system/180047/`. Accessed: 2026-02-22.