

10th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID'2023)



Contribution ID: 302

Type: **not specified**

Simulation Framework for Research and Education in Distributed Computing

Friday, 7 July 2023 09:30 (30 minutes)

Modern systems and applications are increasingly distributed due to the growing performance, scalability and availability requirements. Distributed computing allows to flexibly aggregate the resources of individual machines into scalable computing infrastructures with required characteristics. However, distributed systems are hard to build, test and operate because of their asynchronous and nondeterministic nature, absence of a global clock, partial failures and large scale. Therefore there is ongoing work both in academia and industry to advance the methods and technologies for solving the related problems. There is also a growing need for education of qualified specialists in this field.

Due to the large scale of considered systems, it is generally not feasible or time-consuming and expensive to conduct experiments and evaluate the proposed solutions in a real system. Also, due to the client behavior, dynamicity and non-determinism of production environments, the experimental conditions are hard to control and the results are not reproducible, which is unsuitable for comparison of several solutions. Building a copy of a real system or even a new system solely for the purpose of research is also economically infeasible. The similar observations can be made for education in distributed computing. While it is possible to build a small real lab environment for students, such environment requires a significant effort to operate and cannot expose the students to all problems that occur in modern large-scale systems.

Replacing a real system with simulation allows to resolve these issues. Simulation significantly reduces the cost and time needed to run experiments, while requiring much less resources. For researchers, it enables studying of alternative system configurations and application scenarios, provides a full control over environment and ensures reproducibility. Simulation can also be used in education to provide students a virtual environment for practical assignments that simulates common problems, such as node crashes and network failures, and allows to deterministically execute and check student solutions.

In this report, a general-purpose software framework for simulation of distributed systems, called DSLab, is presented. The main advantages of DSLab in comparison to other similar projects are versatility and extensibility, convenient and flexible programming model, high performance and ability to simulate large-scale systems. DSLab is organized as a set of loosely coupled software modules, which allows users to flexibly assemble solutions for specific purposes. Current modules include a generic discrete-event simulation engine, models of basic system resources (compute, storage and network), reusable modeling primitives, message passing simulator and a set of domain-specific simulators for different research areas such as task scheduling and cloud resource management. The functionality of these modules, their evaluation and the use in research and educational projects are discussed. DSLab is available as an open source project on GitHub: <https://github.com/osukhoroslov/dslab>.

Summary

Primary author: SUKHOROSLOV, Oleg (IITP RAS)

Presenter: SUKHOROSLOV, Oleg (IITP RAS)

Session Classification: Plenary

Track Classification: Plenary