Cost Optimization for Simulations in the Cloud

Oleksandr Shcherbakov[1], Kai Polsterer[1], and Volodymyr Svjatnyj[2]

[1]Heidelberg Institute for Theoretical Studies
[2]Donetsk National Technical University

The growing popularity of cloud service providers is driven by their affordability and flexibility. For scientific purposes the possibility of obtaining unused resources with a substantial discount (spot instances), for computational intensive tasks that do not require real-time simulation, seems to be attractive. We are developing an approach for cost optimization and a framework supporting all stages of running simulations in a cloud in the most efficient way.

1 Introduction to the DPSE framework

Distributed Parallel Simulation Environment (DPSE) [1, 2, 3] is a system organization that specifies how high performance computing resources and system / modeling software can be brought together. It focuses on a user-friendly way to develop and implement computational models that operate in distributed parallel environments.

2 Cost optimization

Assume, we want to launch a simulation on the cloud cluster and minimizing its computation costs. The simulation has been previously benchmarked with test data on different types and numbers of instances. This benchmark serves as a basis for deciding which resources are more economical/best performing for the simulation at hand. For simplification, we take into account only two types of Amazon Web Services (AWS) EC2 spot-instances: *c4.large* and *c4.8xlarge*. *Frankfurt* is chosen as an AWS region. The execution time of this simple benchmark model is shown in table 1. Table 2 shows a price comparison for one simulation running on the *onDemand*- and on *spot*-instances. Full utilization is assumed. Average spot prices were computed by the spot-market analysis tool of DPSE. Values for 6 *c4.large* instances are removed as they are redundant and will wrongly increase a calculation of the resulting ratio.

When comparing the price per run on on-demand instances with the price per run on spot-instances, the latter turn out to be 6.5 times cheaper. Running a benchmark on different types of instances an additional cost improvement of factor 1.4 can be achieved when choosing the best performing one (see tables 1, 2). In combination this yields a cost improvement of factor 9.1.

Number of instances	Instance type	Duration (min)	Simulation runs per hour
1	c4.large	14	4.3
2	c4.large	8	7.5
3	c4.large	6	10
4	c4.large	5	12
6	c4.large	5	12
1	c4.8xlarge	1	60

Table 1: Simulation benchmarking for different types of instances

Number of instances	Instance type	*OnDemand* price/run *(USD)*	*Spot* price/run *(USD)*
1	c4.large	0.0312	0.0048
2	c4.large	0.0357	0.0055
3	c4.large	0.0402	0.0062
4	c4.large	0.0447	0.0068
1	c4.8xlarge	0.0356	0.0066

Table 2: Prices per run for the benchmarked model

3 Conclusions

A prototype of DPSE providing support for all phases of model and simulation development of dynamic network objects and other models was developed. Running DPSE entirely on cloud resources (as opposed to using in-house resources) provided us with further experiments on whether using cloud resources for different classes of simulations can deliver an economical and performance benefit.

An example of saving up to 910% of simulation costs when running a certain simulation in a cloud, shows the potential of our research. This was achieved by choosing an optimal type of instance for the presented simulation model and by automatically monitoring the prices on the spot-market. This benefit will differ for other types of models.

Calculations were made without taking into account the prices for EBS storage, EBS I/Os and data transfer for downloading the results of the simulations. This will be done in future research.

References

[1] L.P. Feldmann, V.A. Svjatnyj, M. Resch, and M. Zeitz. “Forschungsgebiet: Parallele Simulationstechnik.” In ASIM 2014 22. Symposium Simulationstechnik (2014): 3-7.

[2] V. Svjatnyj, V. Kushnarenko, O. Shcherbakov, and M. Resch. “Dekomposition der verteilten parallelen Simulationsumgebung.” Scientific papers of Donetsk National Technical University. Series “Problems of modeling and design automation” (2012): 227-234.

[3] O. Shcherbakov and V. Svjatnyj. “Decomposition into subsystems and organization of work of distributed parallel simulation environment on the web.” Informatics and Computer Technologies (2010): 192-194