Hessisches Kompetenzzentrum
für Hochleistungsrechnen

# *Performance Modeling at a Discount*



Project Manager
Marcus Ritter

Principal Investigator
Prof. Dr. Felix Wolf

Project Term
2019 - 2020

Clusters
Lichtenberg Cluster Darmstadt

Software
FASTEST, TensorFlow

Additional Software
Relearn

Institute
Department of Computer Science

University
Technische Universität Darmstadt

## Introduction

Identifying scalability bugs in parallel applications is a vital but also laborious and expensive task. Empirical performance models have proven to be helpful to find such limitations, though they usually require a series of small-scale experiments in order to gain valuable insights. Therefore, the experiment design determines the quality of the model as well as the overall cost of the modeling process. The current state of the art requires measurements representing all combinations of all parameter values and therefore an exponential number of samples per parameter. For specific applications, this makes it impractical to even create empirical performance models. The performance modeling at a discount project aims to create a new sparse modeling approach that allows for a more flexible design of the empirical performance experiments, while also reducing the required modeling cost.

## Methods

We use deep reinforcement learning to train an agent at the task of empirical performance experiment design to identify generally valid parameter value selection strategies for the application's configuration parameters. Furthermore, we also investigate ideal selection strategies for specific scenarios and applications types. To evaluate our new sparse modeling approach we use millions of synthetically generated performance experiments/ measurements as input to our sparse modeler to investigate its performance when using fewer measurement data and less expensive measurements.

## Results

The result of the project was a new empirical performance modeling approach, which we call sparse modeling. Together with a novel heuristic for the parameter value selection for the necessary empirical performance experiments, which we derived from the results of the deep reinforcement learning agent, we can reduce the average modeling cost by 93% while retaining 99% model accuracy.

## Discussion

The results show that we were able to reduce the modeling cost depending on the investigated application by up to a factor of 100. This is a game-changer for conducting empirical performance experiments. Since we can now measure the performance of applications more flexibly and require much less costly measurements to achieve almost the exact same results the application area of empirical performance modeling and our approach Extra-P is vastly increased.

## Outlook

In the future, we plan to further investigate if there are more possible optimizations for finding even better strategies for parameter value selection to reduce the modeling cost even further for specific applications and scenarios.

## Publications

Calotoiu, A.; Copik, M.; Hoefler, T.; Ritter, M.; Shudler, S.; Wolf, F. : "Software for Exascale Computing" - SPPEXA 2016-2019, chapter ExtraPeak: Advanced Automatic Performance Modeling for HPC Applications. Springer, pages 453–482, 2020 https://doi.org/10.1007/978-3-030-47956-5_15

Ritter, M.; Calotoiu, A.; Rinke, S.; Reimann, T.; Hoefler, T.; Wolf, F. : "Learning Cost-Effective Sampling Strategies for Empirical Performance Modeling." In Proc. of the 34th IEEE International Parallel and Distributed Processing Symposium (IPDPS), New Orleans, LA, USA, pages 884–895, IEEE, May 2020 https://doi.org/10.1109/IPDPS47924.2020.00095

## *Last Update:* 2022-03-27 15:49