Simple DSL for Power-Performance Modeling with Segmented Linear Models

Yuan He
Shenyang University of Technology
isaacyhe@acm.org

Yasutaka Wada
Meisei University
yasutaka.wada@meisei-u.ac.jp

Guanqin Pan
Shenyang University of Technology
panguanqin@gmail.com

Masaaki Kondo
The University of Tokyo / RIKEN
kondo@hal.ipc.i.u-tokyo.ac.jp

ABSTRACT

It is required to utilize the given power budget effectively to make future HPC systems more sustainable and to keep them scale in performance. For an energy-efficient HPC system, its users are required to optimize the applications considering both their performance and power consumption when launched. We have been developing the PomPP tools and libraries to provide an easy-to-use software framework to obtain higher application performance under a given power budget. In this poster, we extend the DSL used in the PomPP tools to better represent the power-performance relationship of the target system. In this extension, we employ a multi-segment linear model to describe the complicated non-linear relationship between CPU performance and its power consumption.

KEYWORDS

HPC, Performance, Power, Optimization, DSL

ACM Reference Format:
Yuan He, Yasutaka Wada, Guanqin Pan, and Masaaki Kondo. 2019. Simple DSL for Power-Performance Modeling with Segmented Linear Models. In *Proceedings of The 48th International Conference on Parallel Processing (ICPP 2019).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Power consumption is one of the main issues to realize future HPC systems with Exa-scale performance and even further. A hardware over-provisioned system can be one of the solutions to make full use of the given power budget [1, 3]. However, optimizing user applications considering both their performance and power consumption at the same time is not an easy task. To address this issue, we have been developing the PomPP tools, which is a versatile power management framework targeting at power constrained large scale HPC systems. This framework provides an extensible hardware/software interface to existing/future power knobs and tools like RAPL [2], and a simple domain-specific language (DSL) as a front-end to many utilities and tools in it [4].

We have succeeded to realize effective power-performance optimization with relatively simple models [4], and still we need to improve this framework to make it possible to accept more complicated and realistic models for more precise power-performance optimization. In this poster, we extend the DSL, a front-end for the PomPP framework, to accept a power-performance model represented as a segmented linear model.

ICPP 2019, August 2019, Kyoto, Japan
2019. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

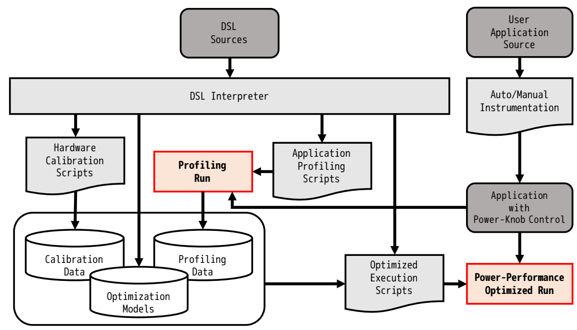

Figure 1: Overview of the Performance and Power Optimization Framework (Fig.2 in [4])

2 POMPP TOOLS: A POWER MANAGEMENT FRAMEWORK

To apply power-performance optimization and power management to an HPC application, the PomPP framework has been developed to support the following functionalities [1, 3, 4]:

- Specifying configurations of the target machine;
- Calibrating power consumption of the hardware;
- Measuring and controlling application power consumption;
- Analyzing and instrumenting source code;
- Optimizing applications under specified power budgets and/or power-performance models.

Figure 1 shows the overall structure of the PomPP framework. This framework requires two sets of input, the DSL source and the user application source code. Based on them, it provides power-performance optimizations for the application. In our previous work, we have evaluated the PomPP tools with simple case studies [4], and we extend the DSL to make it possible to represent more complicated power-performance models in this poster.

3 DSL EXTENSIONS FOR ADDITIONAL POWER-PERFORMANCE MODELS

Our DSL extension aims to support more complicated models in addition to the simple linear model in our previous work [4]. Such complicated models include segmented linear models and look-up tables. In Listing 1, a segmented linear model with 9 segments is specified from 10 power-performance profiles for BT in NPB. These

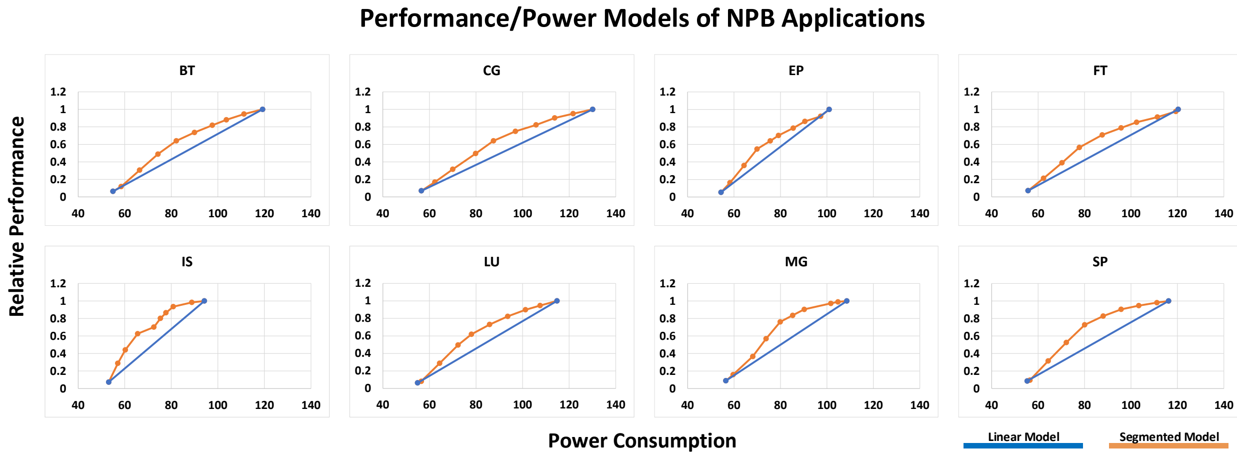

Figure 2: Linear and Segmented Power-Performance Models for NPB Applications

Listing 1: DSL Code Snippets for Model Specification

```
CREATE MODEL BT_SEG
SET BT_SEG(MODEL_TYPE) LINEAR
SET BT_SEG(Y_TYPE) Power_Cap
SET BT_SEG(X_TYPE) Relative_Performance
SET BT_SEG(X_MAX) 1
SET BT_SEG(X_MIN) 0.06356
SET BT_SEG(SET_OF_TUPLES) {(1.0,119.207)
    ,(0.94611,111.213),(0.88059,103.646)
    ,(0.81796,97.587),(0.73608,89.959)
    ,(0.64037,82.169),(0.48857,74.264)
    ,(0.30555,66.382),(0.11831,58.472)
    ,(0.06356,54.948)}
```

10 profiles can be found in BT of Figure 2 and this model can be used to guide power-performance optimization for BT.

4 EVALUATIONS WITH A CASE STUDY ON THE SEGMENTED MODEL

We provide a case study to demonstrate the usage of our DSL extension on the segmented model. This case study is programmed with the DSL and interpreted on a test platform with two Intel Xeon E5-2640 v2 processors and 32GB of main memory. The OS and Compiler we use are CentOS 7.6 and GCC 8.2.1, respectively. The applications we chose are from NPB 3.4 and they are compiled with OpenMPI 1.10.7. We employ RAPL interface as the power knob, and consider only CPU power to be controlled through RAPL under the assumption that DRAM power consumption has strong correlations with CPU performance and power.

To understand performance and power characteristics of the applications and to construct both the linear and the segmented models, profiling under different power caps is necessary and the results are shown in Figure 2. These profiling processes are also specified with our DSL. Each linear model requires two pairs of performance/power values while each 9-segment linear model requires 10 pairs of them.

In this case study, a linear model and a segmented model of the applications are compared for two user desired performance expectations, 50% and 80%. Using the models shown in Figure 2, performance demands, such as 50% and 80% in this case study, can be set from the users when they submit their jobs through our DSL. Figure 3 presents errors for such performance demands under different models, we can clearly see that errors with segmented models are much smaller (and are smaller in all cases) when compared to the simple linear model.

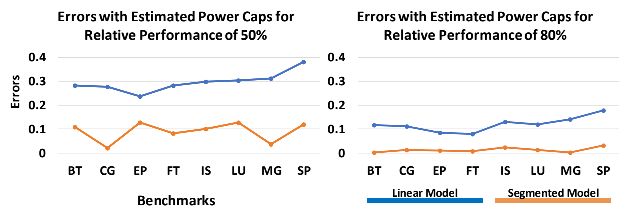

Figure 3: Errors with Estimated Power Caps under Performance Demands

5 CONCLUSIONS

This poster extends the PomPP tools, which provide a simple software framework to optimize an HPC application for its performance and power. Our simple DSL, the front-end of the framework, is extended to accept segmented linear models and evaluated to show its effectiveness.

ACKNOWLEDGMENTS

This work was supported in part by JSPS KAKENHI under Grant Number 17K12665 and by the Natural Science Foundation of Liaoning Province under Grant Number 20180550194.

REFERENCES

[1] Yuichi Inadomi et al. 2015. Analyzing and Mitigating the Impact of Manufacturing Variability in Power-constrained Supercomputing. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC15)*. 78:1–78:12.
[2] Intel Corporation. 2016. Intel® 64 and IA-32 Architectures Software Developer's Manual.
[3] PomPP Library and Tools. [n.d.]. https://github.com/pompp/pompp_tools.
[4] Yasutaka Wada et al. 2018. A Power Management Framework with Simple DSL for Automatic Power-Performance Optimization on Power-Constrained HPC Systems. In *Supercomputing Frontiers*, Rio Yokota and Weigang Wu (Eds.). Springer International Publishing, Cham, 199–218.