

A user perspective on energy profiling tools in HPC environments

Fahimeh Alizadeh Moghaddam^{1,2}, Thomas Geenen³, Patricia Lago², Paola Grosso¹

¹ System and Network Engineering group, University of Amsterdam

² Software and Services group, VU University Amsterdam

³ SURFsara, Amsterdam, The Netherlands

Introduction

Different energy profiling tools allow to manage the power consumption of applications running in HPC environments. The choice for a tool depends on the final purpose of users for the information retrieved.

We categorize two types of users:

- **The generic user:** interested in summary results
- **The software developer:** interested in behavioral-related energy information.

Research Question

- Which tool is more suitable for the different types of users?
- What are the consequences in terms of accuracy and overhead of this choice?

Infrastructure Setup

- We use the Cartesius system at SURFsara.
- We run HPC Challenge benchmark as our experimental application.
- SLURM and Score-P are two available tools in the infrastructure.
- The PAPI/RAPL software power model is supported by both tools.

Conclusion

Both tools provide required information to both types of users but they differ in granularity and accuracy:

The generic user:

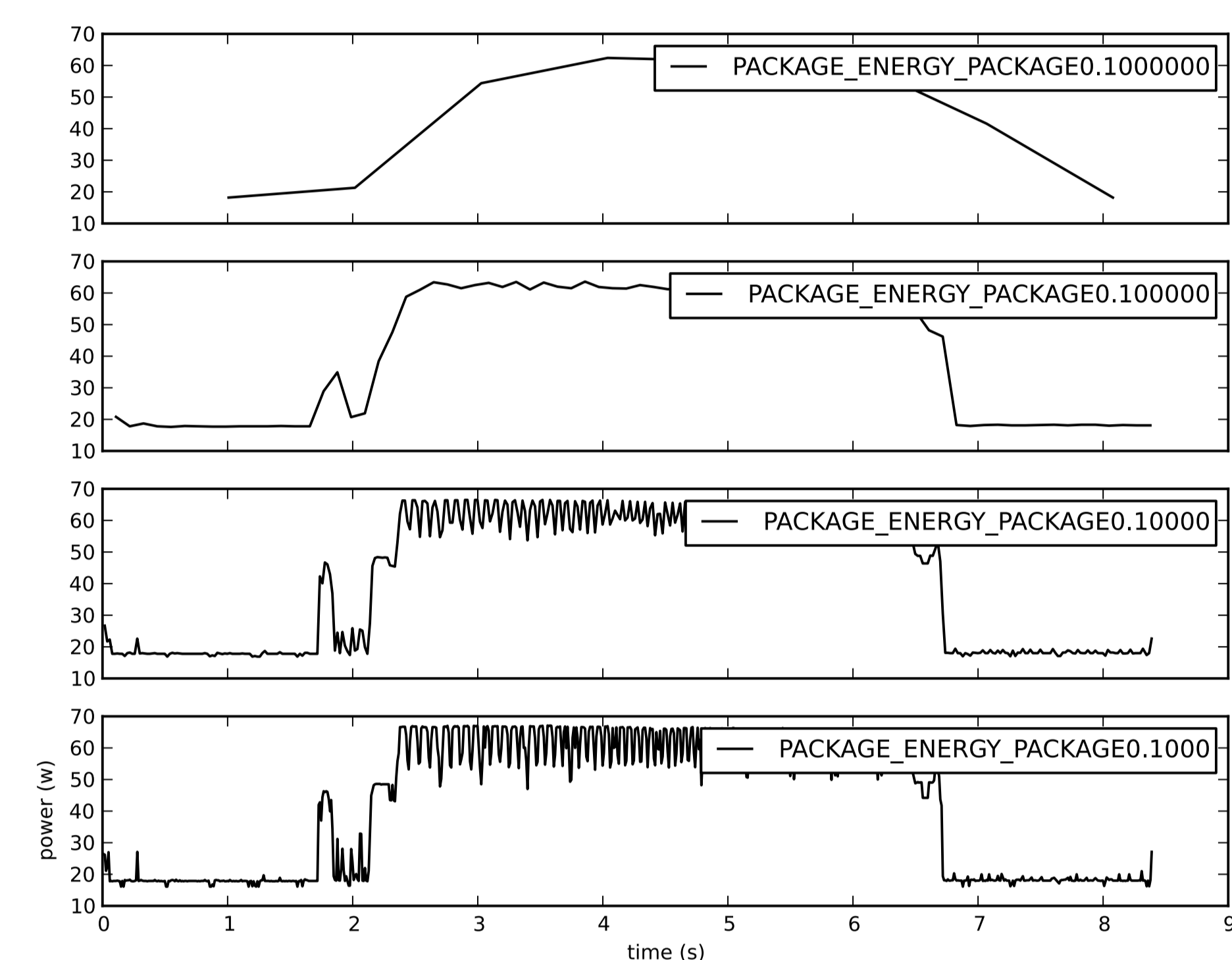
- + SLURM provides accurate summary info
- + fine-grained summary data from Score-P
- very coarse-grained data from SLURM
- inaccurate reported data by Score-P

The software developer:

- + time series of power consumption using SLURM
- + total power consumption of function calls using Score-P
- no time series provided by the visualization tool of Score-P (CUBE)

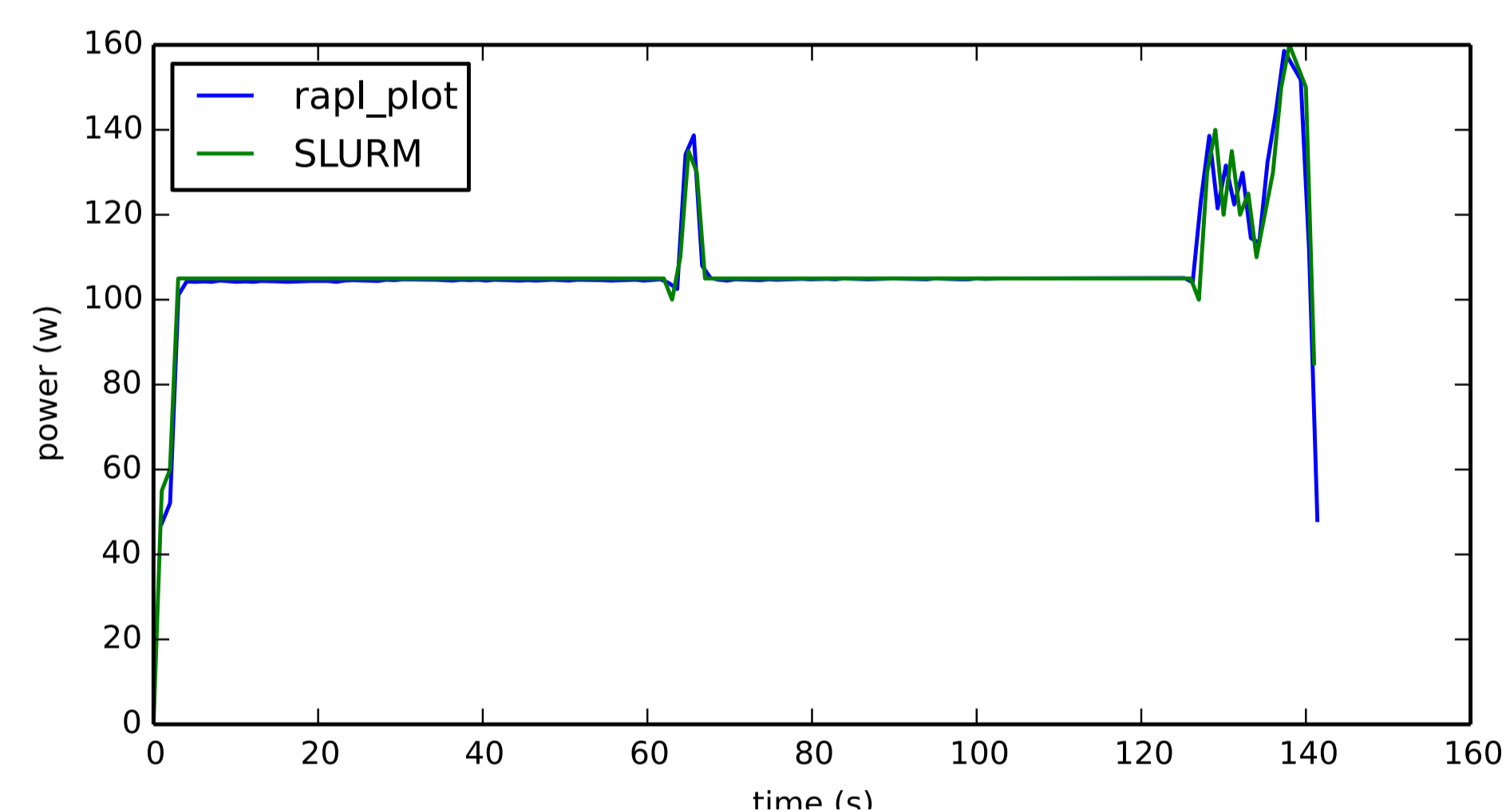
Results

Experiment 1: Collecting power measurements directly by the PAPI library using the rapl_plot application.



- 4 sample rates are considered, from every 1 second (top plot) to every 1000 microseconds (bottom plot)
- As the sample rate decreases, more details are missing from the plots.

Experiment 2: Collecting energy data through SLURM.



- The measurements from Rapl_plot and SLURM are almost identical.
- SLURM shows small delays to the sudden changes in power consumption.

Experiment 3: Collecting energy data through Score-P.

#	Test		Power Consumption Overhead
1	HPL		6%
2	DGEMM	Single	22.6%
		Star	47.3%
3	FFT	MPI	13.2%
		Star	28.9%
4	LatencyBandwidth		34.9%
5	PTRANS		29.5%
6	RandomAccess	MPI	1.7%
		Star	50.9%
7	RandomAccess_LCG	MPI	1.9%
		Star	46.1%
8	STREAM	Single	-23.3%
		Star	49.2%

- Score-P introduces varying amount of overhead for different application runs.
- MPI code path performs with the least measurement overhead compared to Single and Star variants.

Fahimeh Alizadeh Moghaddam is a joint Ph.D student at the University of Amsterdam and the VU University Amsterdam.

Her research interest is to discover the models to support energy efficient operations of distributed software across programmable networks.

Her supervisors are dr. P.Grosso (UvA) and prof. P. Lago (VUA).