Aalto University
School of Science
Degree Programme in Computer Science and Engineering

Gonçalo Marques Pestana

Energy Efficiency in High Throughput Computing

Tools, techniques and experiments

Master's Thesis
Espoo, 1 December, 2014

DRAFT! — October 27, 2014 — DRAFT!

Supervisors: Professor Jukka K. Nurminen
Advisor: Zhonghong Ou (Post-Doc.)

Aalto University
School of Science
Degree Programme in Computer Science and Engineering

ABSTRACT OF
MASTER'S THESIS

Author:	Gonçalo Marques Pestana		
Title:	Energy Efficiency in High Throughput Computing Tools, techniques and experiments		
Date:	1 December, 2014	**Pages:**	26
Major:	Data Communication Software	**Code:**	T-110
Supervisors:	Professor Jukka K. Nurminen		
Advisor:	Zhonghong Ou (Post-Doc.)		
abstract			
Keywords:	energy efficiency, scientific computing, ARM, Intel, RAPL, tools, techiques		
Language:	English		

Acknowledgements

I wish to thank all students who use LaTeX for formatting their theses, because theses formatted with LaTeX are just so nice.

Thank you, and keep up the good work!

Espoo, 1 December, 2014

Gonçalo Marques Pestana

Abbreviations and Acronyms

2k/4k/8k mode	COFDM operation modes
3GPP	3rd Generation Partnership Project
ESP	Encapsulating Security Payload; An IPsec security protocol
FLUTE	The File Delivery over Unidirectional Transport protocol
e.g.	for example (do not list here this kind of common acronymbs or abbreviations, but only those that are essential for understanding the content of your thesis.
note	Note also, that this list is not compulsory, and should be omitted if you have only few abbreviations

Contents

Abbreviations and Acronyms **4**

1 Introduction **7**

2 Background **8**
2.1 High Throughput Computing 8
2.1.1 Literature review 8
2.2 CERN and the LHC experiment 8
2.2.1 Literature review 8
2.3 Energy performance and measurement 8
2.3.1 Literature review 8
2.4 ARM architecture 9
2.4.1 Literature review 9

3 Tools and techniques for measuring energy efficiency of scientific software applications **10**

4 Experiments **11**
4.1 Experiments methodology 11
4.1.1 First Set of Experiments 11
4.1.2 Second Set of Experiments 11
4.1.3 Third Set of Experiments: RAPL in NUMA environment 11
4.2 Results 12

5 Analysis **21**
5.1 First Set of Experiments 21
5.1.1 Comparison ARM and Intel architecures 21
5.1.2 Tools and techniques 21
5.2 Second Set of Experiments 21
5.2.1 Comparison ARM and Intel architecures 22

5.2.2 Tools and techniques 22
5.3 Third Set of Experiments 22

6 Future Work **23**

7 Conclusions **24**

A First appendix **26**

Chapter 1

Introduction

Chapter 2

Background

2.1 High Throughput Computing

2.1.1 Literature review

2.2 CERN and the LHC experiment

2.2.1 Literature review

2.3 Energy performance and measurement

2.3.1 Literature review

on importance of energy consumption for engineers and scientists
The study conducted by [2], shows that engineers have been considering energy consumption as an important factor when developing software. It consists on an empirical study that aims to understand the opinions and problems of software developers about energy efficiency. The data that sustain the conclusions are mined from a well-known technical forum (*StackOverflow* [1]). Although the study is focused in an application-level energy efficiency, it shows that developers are aware of the importance of energy efficiency in computational systems. When trying to understand in depth what questions arise more frequently, it is shown that measurement techniques is amongst the most asked questions by developers. In addition, the study ascertains that the *"lack of tool support"* is an important handicap for the development of energy efficient software.

2.4 ARM architecture

2.4.1 Literature review

Chapter 3

Tools and techniques for measuring energy efficiency of scientific software applications

Chapter 4

Experiments

4.1 Experiments methodology

The experiments were performed in different sets. Whereas the first two sets of experiments aim to provide a straightforward comparison between ARM and Intel technologies, the third set of experiments aims to study the influence of a NUMA environment in hight performance computing from an energy consumption perspective. In each set, we used different techniques and tools to perform the energy measurements. The techniques and tools used to perform the measurements are described and analyzed in depth in Section 3.
The first part of this chapter outlines the scope, methodology and measurement tools used for each set. The latest part shows the results of the experiments, which are analyzed in the next chapter.
Throughout the rest of the document, the different experiments will be termed as first (FSE), second (SSE) and third set of experiments (TSE).

4.1.1 First Set of Experiments

Done at Aalto. Explain methodology and scope.

4.1.2 Second Set of Experiments

Done at CERN. Explain methodology and scope.

4.1.3 Third Set of Experiments: RAPL in NUMA environment

Done at CERN. Explain methodology and scope.

4.2 Results

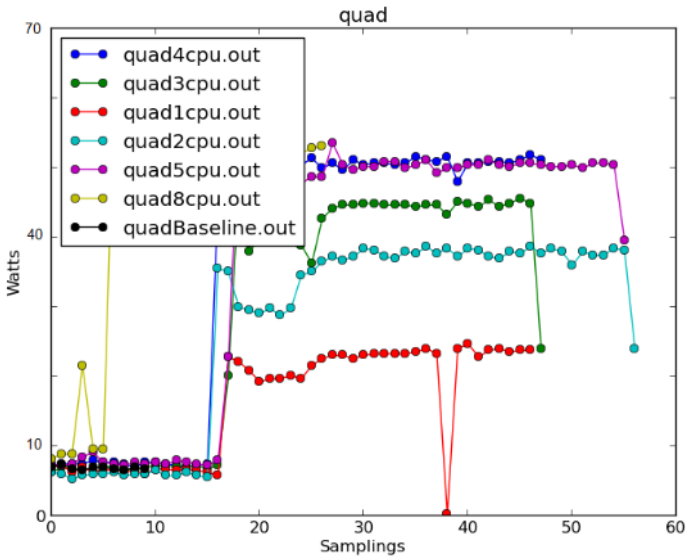

Figure 4.1: Full single threading CMS experiments on Intel Quad

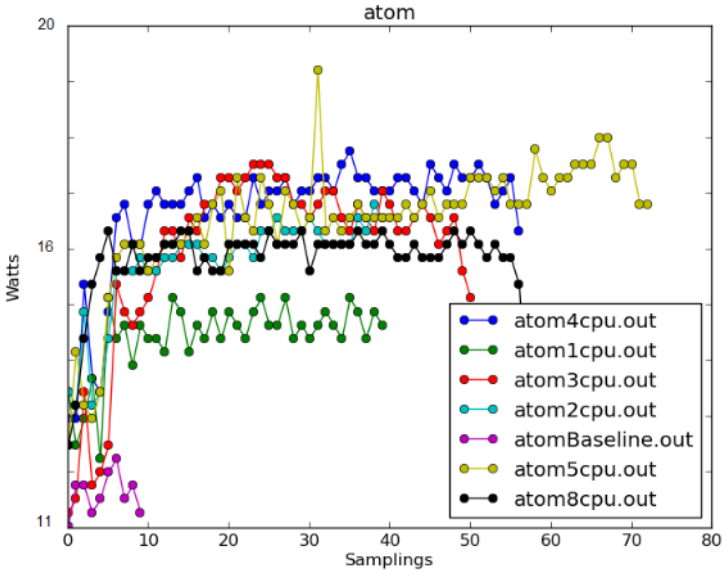

Figure 4.2: Full single threading CMS experiments on Intel Atom

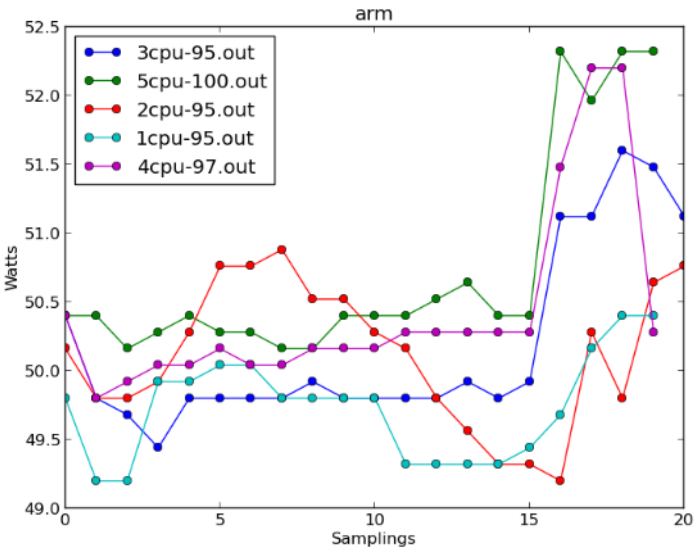

Figure 4.3: Full single threading CMS experiments on ARMv7 server

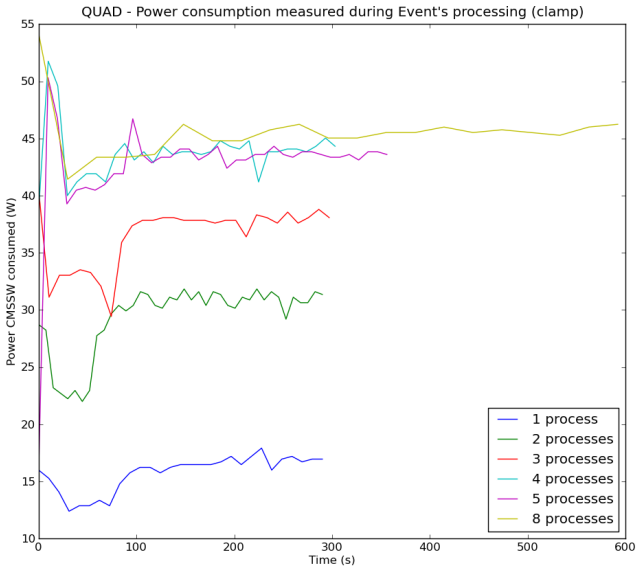

Figure 4.4: Full single threading CMS experiments on Intel Quad - event processing only

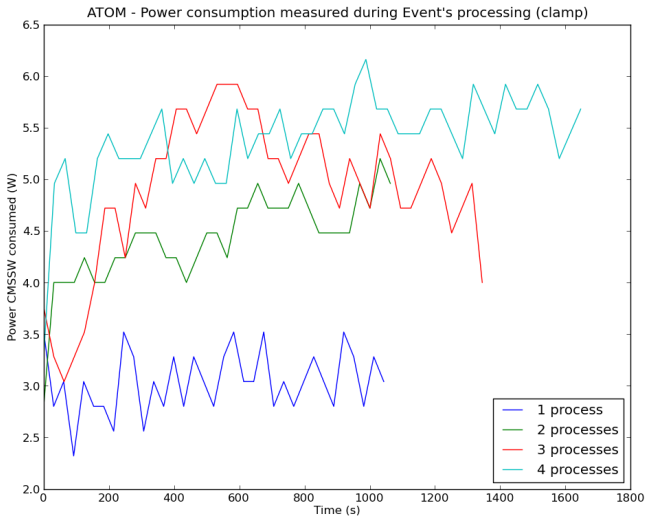

Figure 4.5: Full single threading CMS experiments on Intel Atom - event processing only

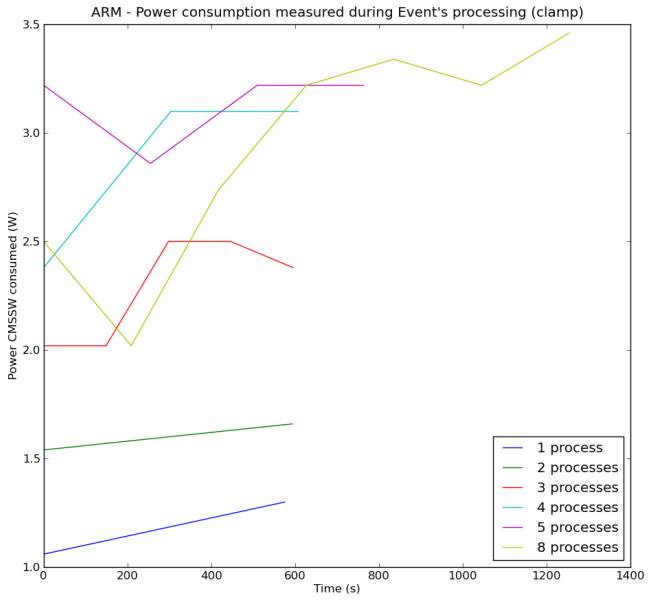

Figure 4.6: Full single threading CMS experiments on ARMv7 server - event processing only

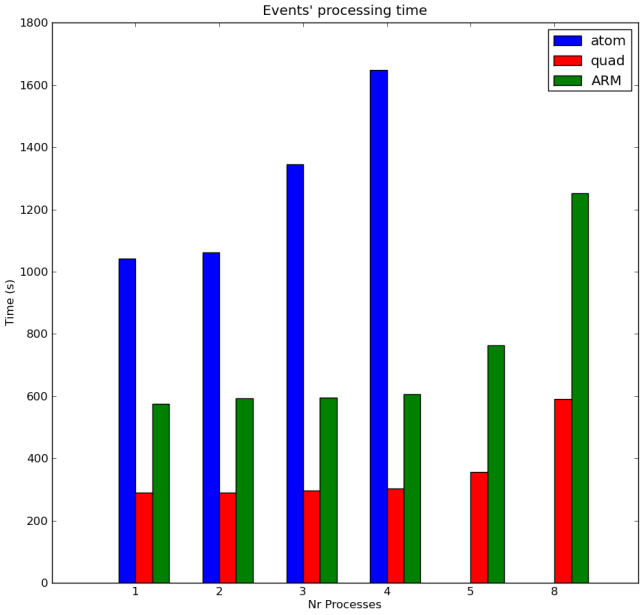

Figure 4.7: Processing time comparison

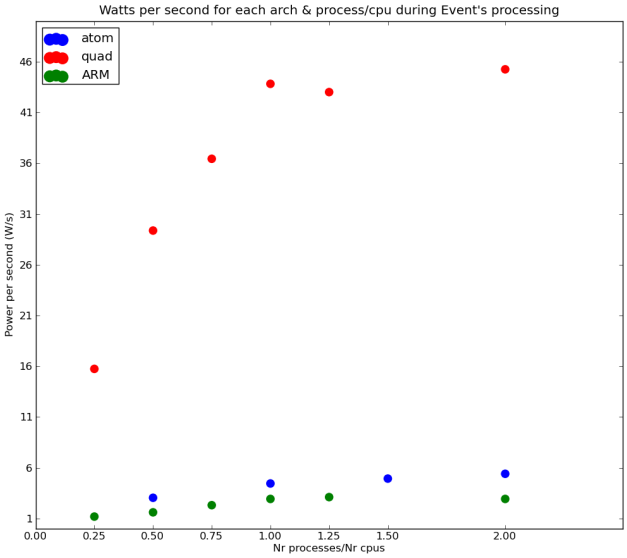

Figure 4.8: Processing stage comparison between architectures

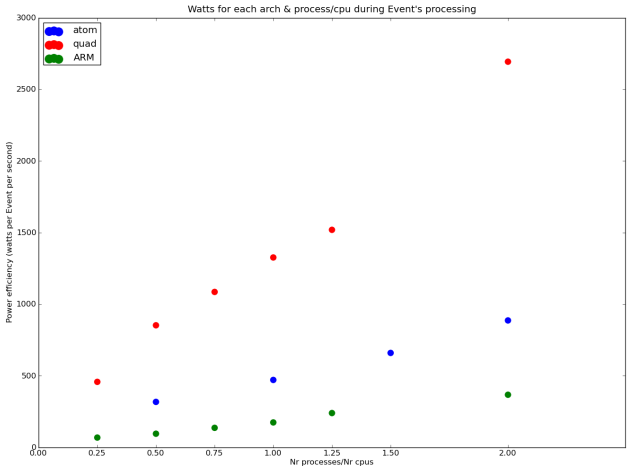

Figure 4.9: Processing stage comparison between architecturesi - 2

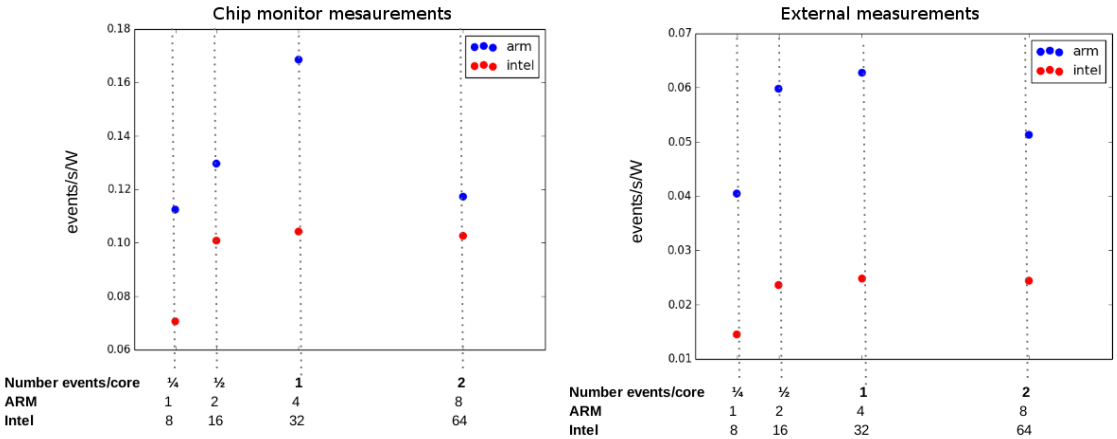

Figure 4.10: Multithreaded ParFullCMS comparison Intel Xeon vs ODROID ARMv7

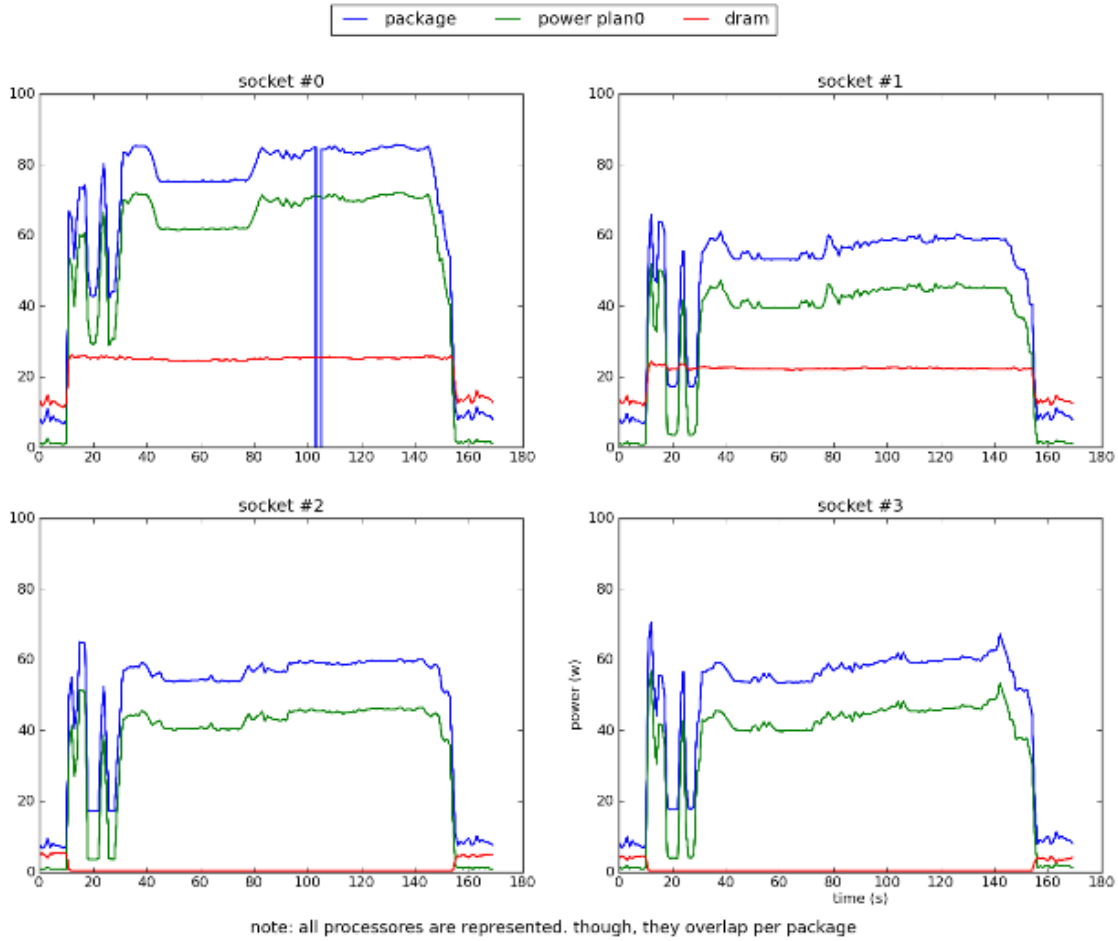

Figure 4.11: RAPL measurements of NUMA nodes - 16 processes with no explicit binding

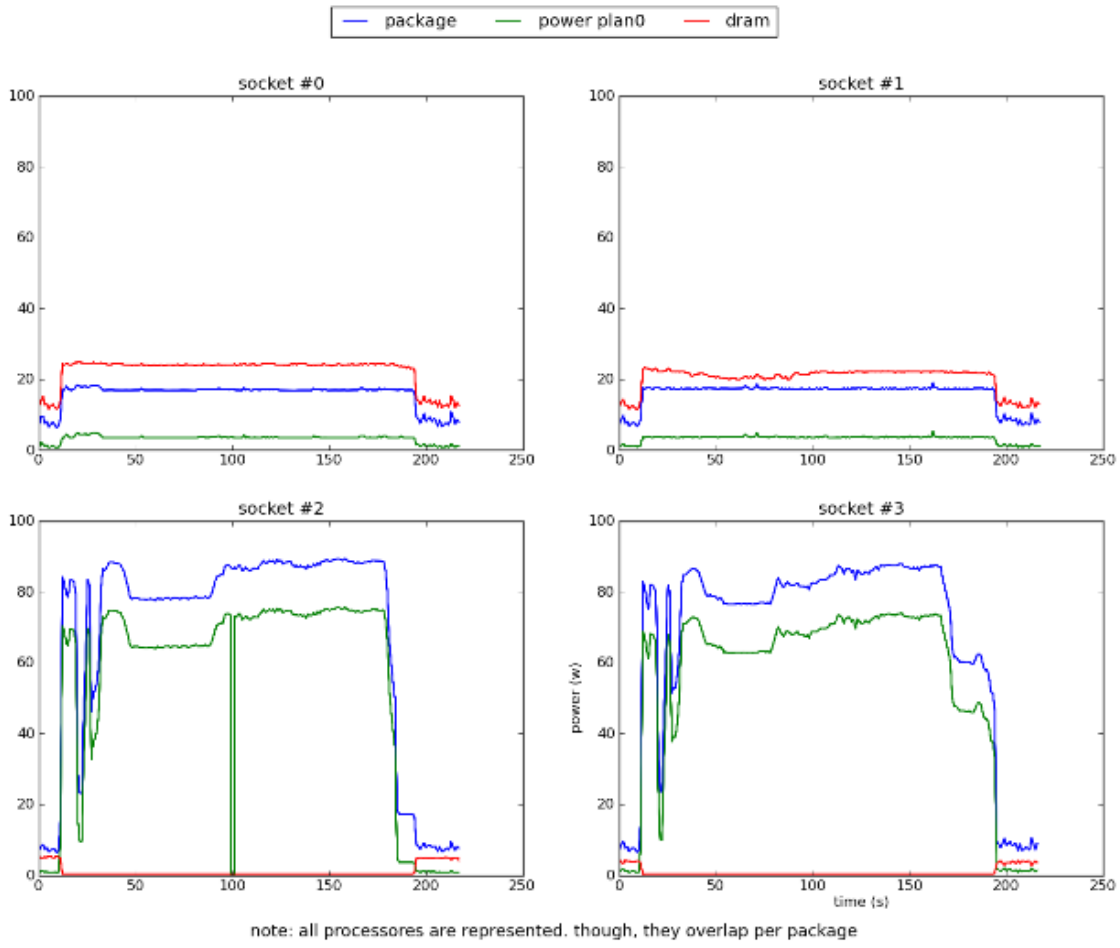

Figure 4.12: RAPL measurements of NUMA nodes - 16 processes. Explicit binding on node #2 and node #3 binding

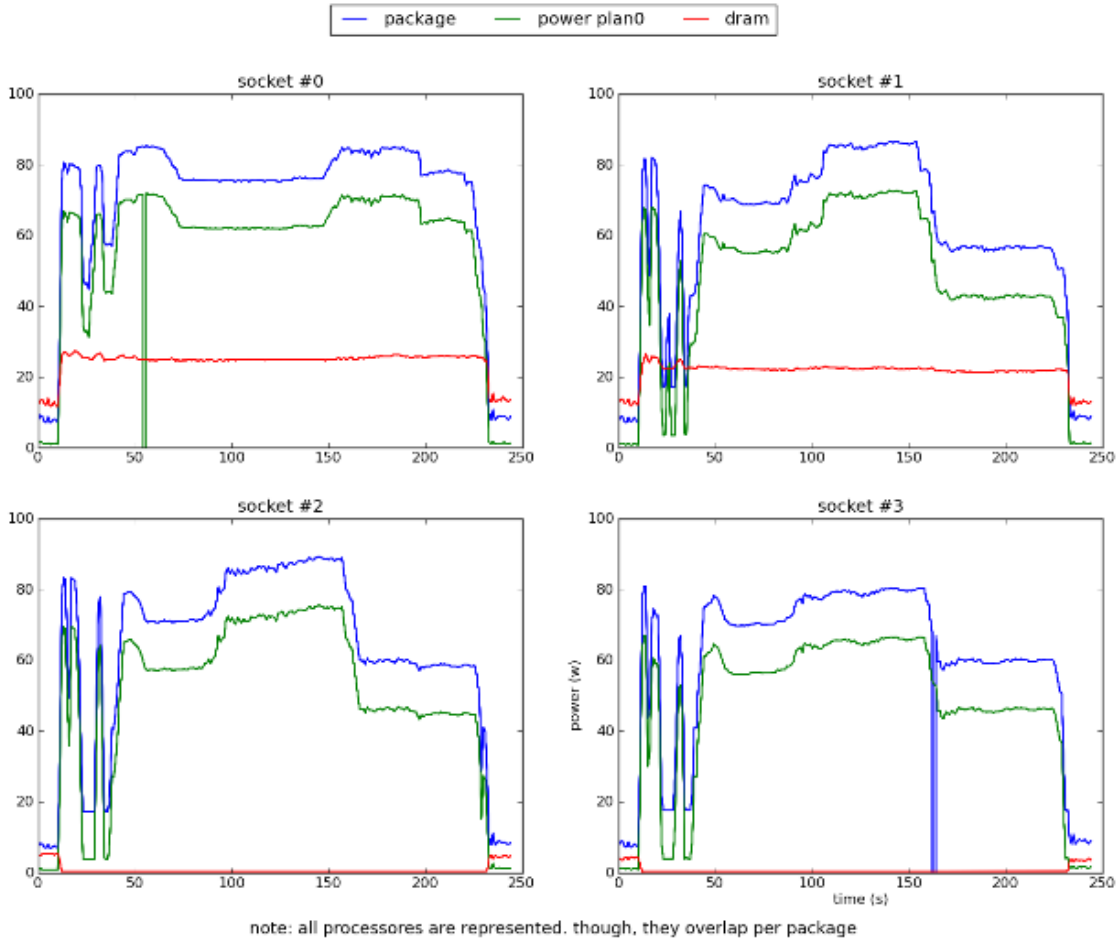

Figure 4.13: RAPL measurements of NUMA nodes - 32 processes with no explicit binding

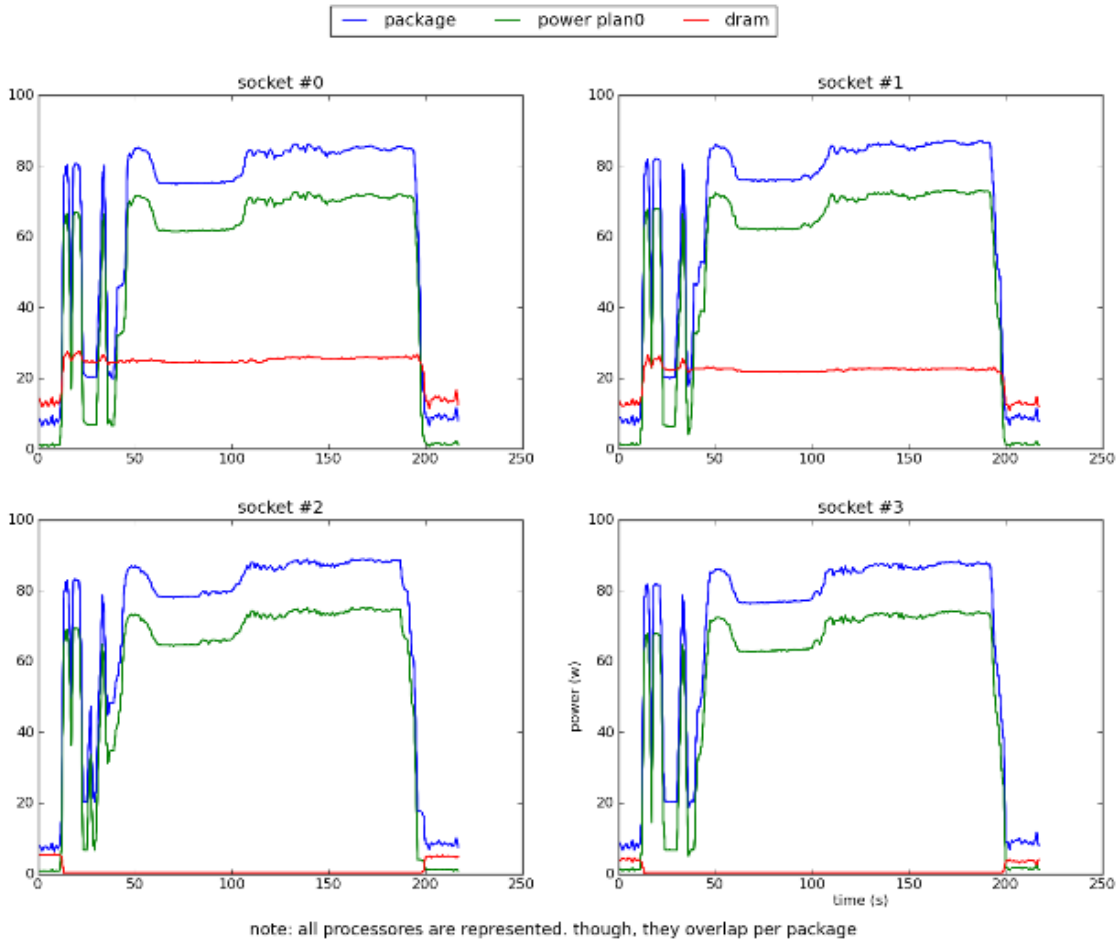

Figure 4.14: RAPL measurements of NUMA nodes - 32 processes. Processes distributed evenly explicitly - 8 processes per node.

Chapter 5

Analysis

The scope of the analysis presented in this section is twofold: to compare the energy efficiency of the studied platforms and analyze the tools and techniques used on the different experiment sets.
Whereas the first two sections analyze the energy efficiency of the platforms studied and the particularities of the tools and techniques used, the last section covers the results and issues which arose when using RAPL to measure the energy consumption of a NUMA environment.

5.1 First Set of Experiments

ARM server vs ATOM and QUAD, using clap and software-based experiments

5.1.1 Comparison ARM and Intel architecures

5.1.2 Tools and techniques

5.2 Second Set of Experiments

ARM board and Intel Xeon, using on chip and external measurements

5.2.1 Comparison ARM and Intel architecures

5.2.2 Tools and techniques

5.3 Third Set of Experiments

Intel Xeon, using RAPL to measure energy consumed by the different nodes, with different types of binding

Chapter 6

Future Work

Chapter 7

Conclusions

Bibliography

[1] Stack overflow. `http://stackoverflow.com/`. Accessed: 2014-10-27.

[2] Pinto, G., Castor, F., and Liu, Y. D. Mining questions about software energy consumption. In *Proceedings of the 11th Working Conference on Mining Software Repositories* (New York, NY, USA, 2014), MSR 2014, ACM, pp. 22–31.

Appendix A

First appendix

This is the first appendix. You could put some test images or verbose data in an appendix, if there is too much data to fit in the actual text nicely.