Energy Models for DVFS Processors

Thomas Rauber[1] Gudula Rünger[2] Michael Schwind[2] Haibin Xu[2]
Simon Melzner[1]

1) Universität Bayreuth 2) TU Chemnitz

9th Scheduling for Large Scale Systems Workshop
July 1-4, 2014
Lyon, France

Outline

Introduction and Motivation

- **Energy consumption** is an important concern in today's consideration of **parallel programs** especially for **HPC**.
- Several different **energy acquisition methods** based on hardware, software and simulation approaches have been proposed in a large variety of different setups.
- Current commodity processors provide the **dynamic voltage-frequency scaling (DVFS) technique**.
  processors can dynamically **adjust voltage and frequencies** of cores to reduce power consumption
- **Reducing the frequency** leads to a **smaller power consumption**. However, **longer computation times** result due to the reduced frequency.
- It would be valuable to be able to **choose a suitable frequency before running** a larger HPC program.

Introduction and Contribution

- We investigate **two energy measurement techniques** for DVFS processors
  hardware-based measurement with **power-meters** and **RAPL sensors** accessing MSR hardware counters.
- As application programs, we have chosen the **SPEC CPU2006**, the **PARSEC benchmarks** and the **SPLASH benchmarks**, which represent a **broad range of sequential and multithreaded application codes**.
- We also compare **three different energy models** for DVFS concerning their ability to **capture** the energy consumption of the benchmarks.
  physical energy models and a new heuristic model
- An **experimental investigation** is provided comparing the **energy prediction capabilities** of the energy models.

Outline

DVFS processors

- Modern microprocessors such as the Intel Core i7 processors incorporate a **sophisticated power management technology** performance states (**P-states**), throttle states (**T-states**), idle states (**C-states**) and sleep states (**S-states**)
- **P-states** are predefined sets of frequency and voltage combinations at which an active core can operate.
- A **C-state** is an idle state in which parts of the processor are **powered down** to save energy.

Outline

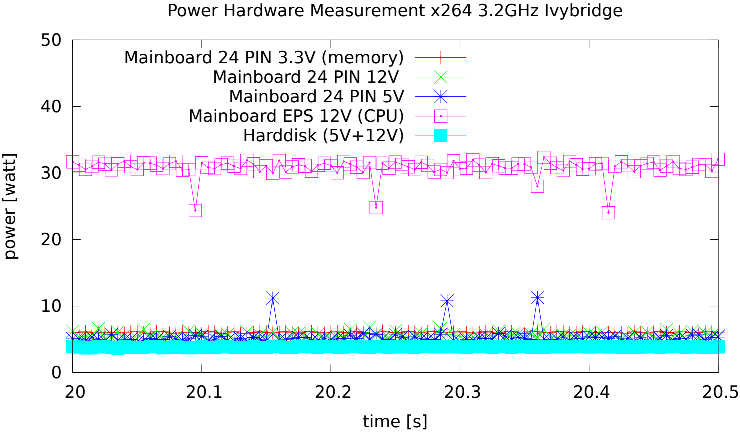
Power measurement with power-meters (NI9205 device)

The NI9205 enables a fine-grain power measurement of different components of a computer system.

Detailed power measurement using the NI9205 device

- Power acquisition and profiling with **LabView**
- **Challenge**: relate the power data measured to the application program whose energy consumption is to be determined;
- **User-configured modules** operating in a client-server fashion had to be written
- Detailed measurement for **different pins** supplying **different components** of the computer system.

Example: PARSEC benchmark x264 on Core i7 Ivy Bridge



time interval between 20 and 20.5 sec

Platforms for Experimental Evaluation

	Core i7-2600	**Xeon E3-1225V2**	**Core i7 4770**
architecture	Sandy Bridge	Ivybridge	Haswell
min. frequency	1.6 GHz	1.6 GHz	0.8 GHz
max. frequency	3.2 GHz	3.2 GHz	3.4 GHz
TDP	95 W	77 W	84 W
step size freq.	100 MHz	100 MHz	200 MHz
physical cores	4	4	4
hyberthreading	yes	no	yes
virtual cores	8	4	8
L1 data cache	32 KByte	32 KByte	32 KByte
L2 cache	256 KByte	256 KByte	256 KByte
L3 shared cache	8 MByte	8 Mbytes	8 MByte
RAM size	8 GByte	8 GByte	8 GByte

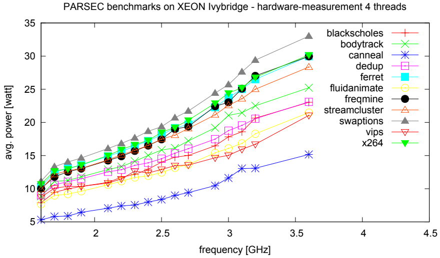
Outline

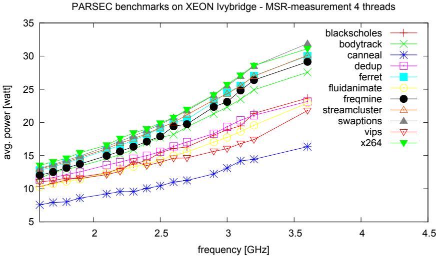
Power measurement with RAPL sensors

- Runtime and energy measurements for different **Intel Core i7 processors (Sandy Bridge, Ivy Bridge, Haswell)**.
  access to **Model Specific Registers (MSRs)** via **rdmsr** and **wrmsr** instructions
- The **RAPL (Running Average Power Limit)** interface provides mechanisms to control power consumption;
- The MSRs provide information about the energy status of the **PP0 and PP1 power planes** via specific registers.
- **likwid-powermeter** from the **likwid tool-set** (Version 3.0) to access the MSRs.
- The **cpufreq_set** tool has been used to set the core frequencies.

Outline

Comparison of the measurement techniques





only the +12VDC EPS connector power is shown left

observation: the two alternative measurement techniques coincide qualitatively and quantitatively for a **wide range of frequencies**

small difference as the 24 PIN 5V connector **also supplies the CPU** (and other mainboard devices)

in the following: measurement with RAPL

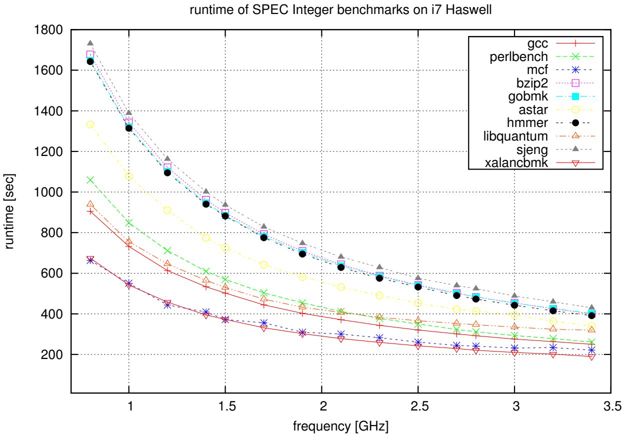
Outline

Outline

SPEC CPU2006 benchmarks

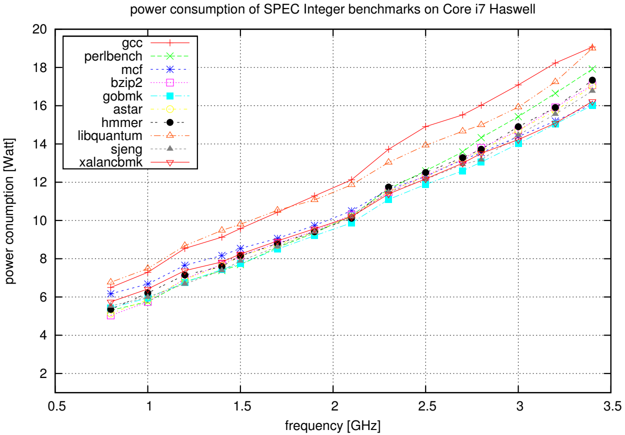
integer and floating-point benchmarks from different application areas

runtimes on Core i7 Haswell for integer benchmarks using different frequencies:



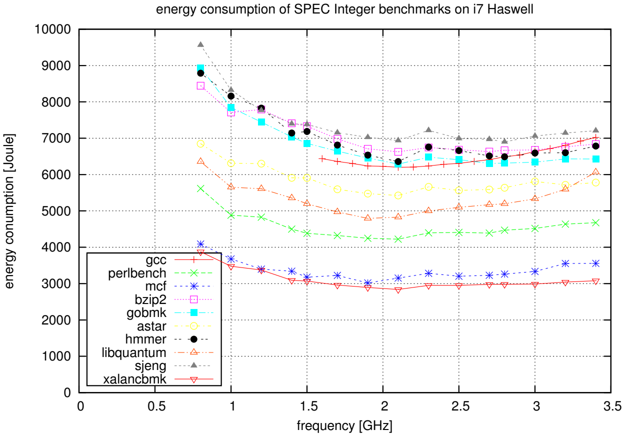
more-than-linear increase of the execution time for smaller frequencies

SPEC CPU2006 integer benchmarks: power consumption on Haswell



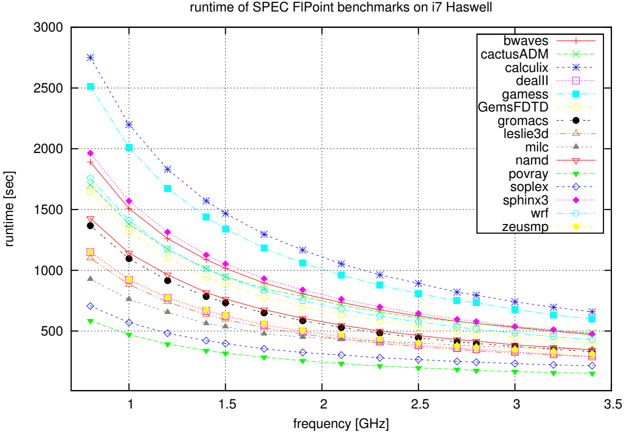
different applications lead to different power consumption

SPEC CPU2006 integer benchmarks: energy consumption on Haswell



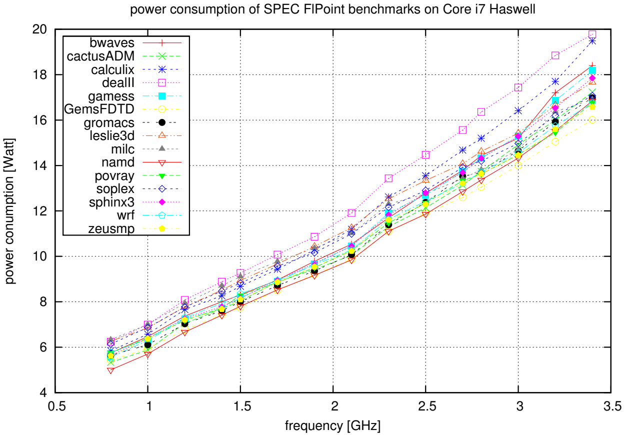
no large variation of the energy consumption with the frequency

SPEC CPU2006 floating point benchmarks: runtime on Haswell



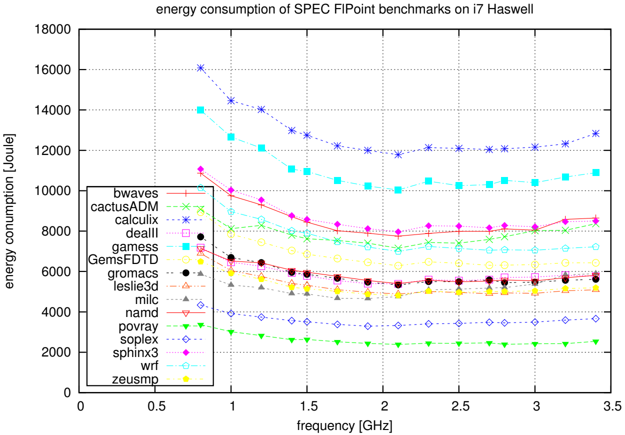
more-than-linear increase of the execution time for smaller frequencies

SPEC CPU2006 floating point benchmarks: power consumption on Haswell



different applications lead to different power consumption
slightly larger power consumption as for the integer benchmarks

SPEC CPU2006 floating point benchmarks: energy consumption on Haswell



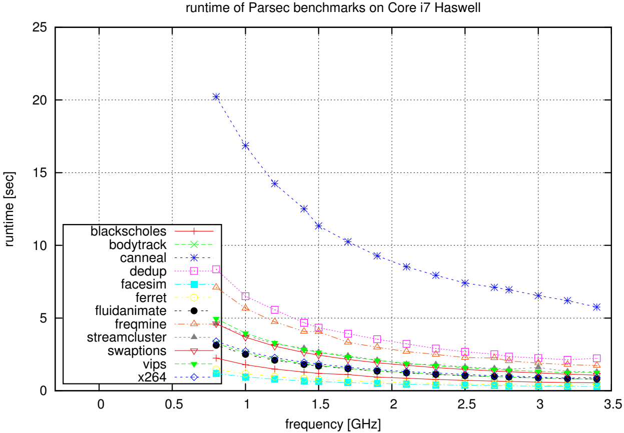
no large variation of the energy consumption with the frequency

Outline

PARSEC benchmarks – runtime Haswell

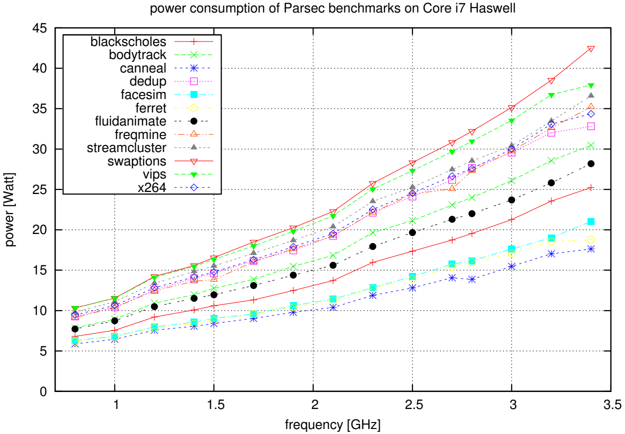
12 programs from different application areas

different parallel models for shared address spaces are used



Execution time increases **more than linearly for smaller frequencies** (below about 1.7 GHz).

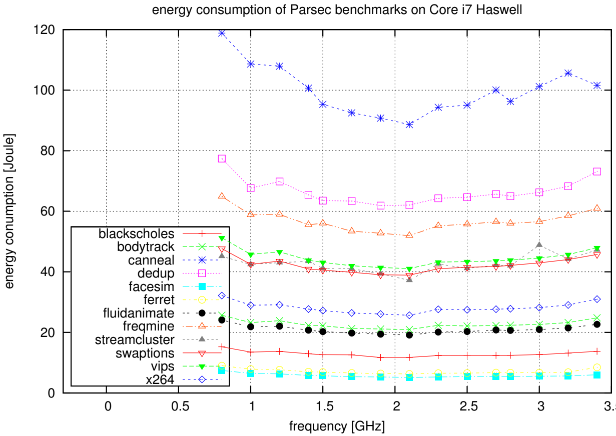
PARSEC – power consumption with varying frequency



large variation of the power consumption for different benchmarks

Benchmarks with a **sequential workload** typically lead to smaller power values

PARSEC benchmarks – energy consumption Haswell



smallest energy consumption between 2 GHz and 2.5 GHz

Outline

Energy models with frequency scaling

- Energy models usually take the **dynamic power consumption** and the **static power consumption** into consideration.
- The **dynamic power consumption** is related to the supply voltage and the **switching activity** during the computing activity of the processor.
- The **static power consumption** is intended to capture the **leakage power consumption** as well as the power consumption of **peripheral devices**.
- The **total power consumption** of the CPU is obtained as the sum of these two components.
- For **DVFS processors**, the power consumption depends on the **operational frequency** f.

Outline

Physical energy models

- The **energy consumption** of an application program can be described as $E = \int_{t=t_0}^{t_{max}} P(t) \cdot dt$.
- The **dynamic power consumption** is often approximated by
  $P_{dyn} = \alpha \cdot C_L \cdot V^2 \cdot f$
  α: switching probability; C_L: load capacitance; V: supply voltage;
  f: operational frequency.
- Modeling of the **static power consumption due to leakage power**:
  $P_{static} = V \cdot N \cdot k_{\text{design}} \cdot I_{\text{leak}}$
  N: number of transistors; k_{design}: design-dependent parameter;
  I_{leak}: technology-dependent parameter.
- The frequency scaling can be expressed by a **dimensionless scaling factor** $s \geq 1$, which describes $\tilde{f} < f_{max}$ as $\tilde{f} = f_{max}/s$.
- The frequency f depends linearly on the supply voltage V: $V = \beta \cdot f$.
- Thus, the dependence of the **dynamic power** on f is approximated by $P_{dyn} = \gamma \cdot f^3$ with $\gamma = \alpha \cdot C_L \cdot \beta^2$.

Modeling the static power consumption

- For **earlier processors**, the static power consumption was considered to be **neglectable**.
- For **recent processors**, the static power consumption may be **too large** to be ignored.
- **Model 1**: static power depends **linearly on the frequency**: $P_{static} = \delta \cdot f$ with $\delta = N \cdot k_{\text{design}} \cdot I_{\text{leak}} \cdot \beta$.
- **Model 2**: static power is **constant**, independently of f.
- **Reducing the operational frequency** of a processor by a scaling factor of s, $s \geq 1$, **increases the execution time** of a program by the same factor.

Outline

Heuristic energy model

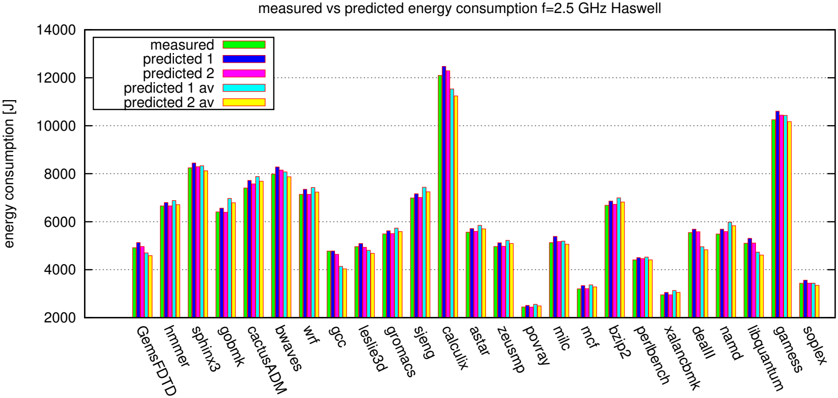
- A (new) heuristic model considers the entire power consumption and uses **least squares methods** to derive a formula describing the power consumption in **closed form**.
- **Observation** from the exeriments: there is an **almost linear dependence** of the power on the frequency f: $P_{heu}(f) = a + bf^{1+\epsilon}$
- The parameter a can be interpreted as the **static part** of the power consumption that **does not change** with the frequency.
- The parameter b captures the **dynamic part** of the power consumption that **increases with the operational frequency** of the CPU.
- For the parameter ϵ, several **fixed values** have been tested and the computation of a and b is done by the least squares method.
- **Different benchmarks** may have **different values** for these parameters a and b due to their specific computational and memory access behavior.

Outline

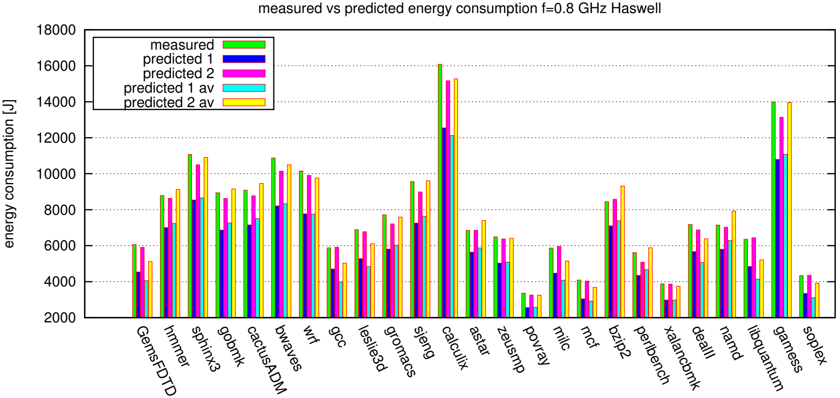
Validating the energy models

- Comparison of the **measured energy values** with energy values **predicted by the models** for different frequencies.
- For the **analytical model**, the parameters γ and δ have been determined by curve fitting using the least squares method.
- For different benchmarks, the resulting values for the parameters γ and δ are **quite similar** for most of the benchmarks on the same architecture (the difference is typically below 10 %).
- Thus, in principle, the **average of the parameters** for the different benchmarks could be used and would lead to a similar correspondence between measured and predicted values.
- For the different architectures, different values for the parameters γ and δ result.
- For the **heuristic model**, $\epsilon = 0.2$ has been used.

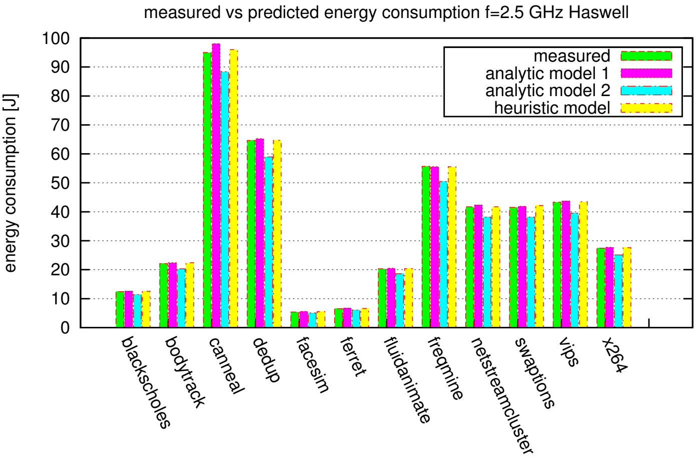
SPEC: Comparison for Haswell $f = 2.5$ GHz



SPEC: Comparison for Haswell $f = 0.8$ GHz



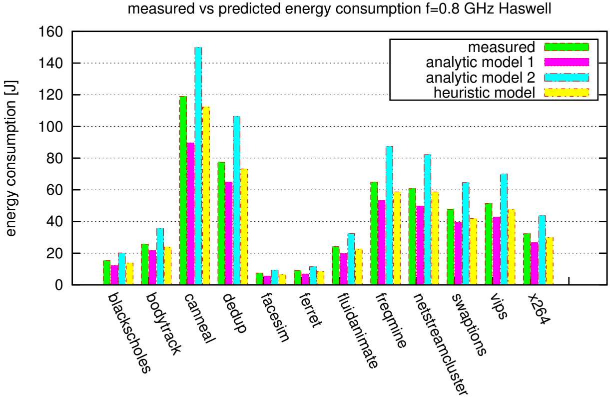
PARSEC: Comparison for Haswell $f = 2.5$ GHz



Model 1: parameter γ (dynamic part) lies between 12 and 31 for different benchmarks;

parameter δ (static part) lies between 7 and 13.5;

PARSEC: Comparison for Haswell $f = 0.8$ GHz



best predictions by the heuristic model

Observations

- For most situations, **both the analytical and the heuristic energy models** are **well suited** to describe the energy consumption of most benchmark programs.
  The **deviations** usually lie **below 10%**.
- The two **analytical models** both provide reasonable predictions with **slight advantages for Model 1**.
- Using the analytical models, **larger deviations** between the measured and predicted values can be observed for **smaller frequencies** on the Haswell architecture.
  The **heuristic model** leads to better predictions in this situation.
- Only for **smaller frequencies**, there are some **deviations** between the models. In this context, the **heuristic model** provides better predictions.
- **Summary:** the energy models are able to capture the energy consumption with **reasonable accuracy** for most situations.

Outline

Conclusions

- **Frequency scaling** provides the possibility to choose an energy and runtime efficient state for processing an application program.
- We have studied various hardware, software and simulation approaches.
- Both **measurement methods** considered (power-meters, hardware counters) provide **qualitatively and quantitatively corresponding data**.
- **Large variation** of power consumption for the different benchmarks; **speedup** plays an important role
  variations are **smaller** for **sequential workloads**
- **Energy models** are suitable for an **energy performance prediction**.

References

- Rauber, Rünger, Schwind: **Energy Measurement and Prediction for Multi-Threaded Programs** In: Proc. of the 22nd High Performance Computing Symposium 2014 (HPC 2014), Tampa, USA, 2014
- Rauber, Rünger, Schwind, Xu, Melzner: **Energy Measurement, Modeling, and Prediction for Processors with Frequency Scaling**, to appear: The Journal of Supercomputing, Springer, 2014.
- Rauber, Rünger: **Modeling and Analyzing the Energy Consumption of Fork-Join-based Task Parallel Programs**, to appear: Concurrency and Computation: Practice and Experience, Wiley, 2014