CoolEmAll
A focus on Power Consumption of Applications

Leandro Fontoura Cupertino, Georges Da Costa,
Amal Sayah, **Jean-Marc Pierson**

SEPIA Team
IRIT - Toulouse Institute of Computer Science Research
UPS - University of Toulouse (Paul Sabatier)

Outline

Outline

IRIT, some numbers

- 1st French informatics lab [1]
- 250 PhD
- 250 Researchers

[1] in number of researchers and PhD

Themes and strategic axis

- Theme 1 : Information Analysis and Synthesis
- Theme 2 : Indexing and Information Search
- Theme 3 : Interaction, Autonomy, Dialogue and Cooperation
- Theme 4 : Reasoning and Decision
- Theme 5 : Modelization, Algorithms and HPC
- **Theme 6 : Architecture, Systems and Networks**
- Theme 7 : Safety of Software Development
- SA1: Computer Science for Health
- SA2: Data Mass and Calculus
- SA3: Ambient Socio-technical Systems
- SA4: Critical Embedded Systems

Paul Sabatier, Toulouse III University

- Informatics, Mathematics, Physics, Chemistry, Biology
- Pharmacy, Medicine, Dentistry

On site, around 28 000 students (about 100000 in all Toulouse's Universities)

Outline

Cool'Em All

- European Co-funded project (INFSO-ICT-288701)
- FP7 ICT Call 7 (FP7-ICT-2011-7)
- Budget: €3,614,210 (funded: €2,645,000)
- Duration: 30 months
- Start date: 1st Oct 2011

SEVENTH FRAMEWORK PROGRAMME

Cool'Em All

- European Co-funded project (INFSO-ICT-288701)
- FP7 ICT Call 7 (FP7-ICT-2011-7)
- Budget: €3,614,210 (funded: €2,645,000)
- Duration: 30 months
- Start date: 1st Oct 2011

- Consortium

Cool'Em All

Goals

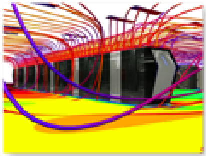
- Improve energy-efficiency of modular data centres by **optimization** of their **design** and **operation** for a wide range of **workloads**, **IT equipment** and **cooling** options

Cool'Em All

Goals

- Improve energy-efficiency of modular data centres by **optimization** of their **design** and **operation** for a wide range of **workloads**, **IT equipment** and **cooling** options

- Define open designs of computing building blocks (ComputeBox Blueprints)
- Develop an open source Simulation, Visualization and Support (SVD) toolkit
  - Inputs: Data Centre Architecture, Cooling Approaches and Energy-aware Management
  - Outputs: Efficient Airflow, Thermal Distribution and Optimal Arrangement

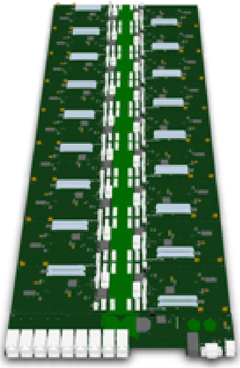


CoolEmAll Work Packages

- WP 1 Project Management
- WP 2 Simulation, Visualisation and Decision Support Toolkit
- WP 3 ComputeBox Prototype
- WP 4 Workload and Resource Management Policies
- WP 5 Energy-efficiency Metrics (leader: IRIT)
  - Metrics
  - Monitoring of applications
- WP 6 Requirements, Verification and Validation Scenarios
- WP 7 Dissemination, Exploitation and RTD Standardization

WP 3: CoolEmAll testbed

RECS: Resource Efficient Computing Systems
18 nodes on 1U. Highly configurable, can be Intel i7 or Atom, Amd Fusion, soon ARM.
3 testbeds: UPS, PSNC, HLRS



- Using Timacs API for accessing several measurements on the system (HW and SW).
- Developing new metrics to consider Heat and dynamic of the system

WP 5: Metrics, monitoring, benchmarking and Application characterization

- Derive energy-efficiency metrics for computing modules extending existing power related metrics to energy related metrics (i.e. including time) taking also into account the runtime environment of the data centre (ambient temperature, heat re-use capacities)
- Design and develop a monitoring infrastructure adapted to energy- and heat-aware scheduling
- Design a methodology for profiling applications in respect with their energy consumption
- Develop benchmarks to evaluate derived metrics

Outline

Introduction

Motivation

Category	Power cons. 2008 (GW)	Growth rate (p.a.)	2020 prediction (GW)
Data centers	29	12%	113
PCs	30	7.5%	71
Networking Equipment	25	12%	97
TVs	44	5%	79
Other	40	5%	72
Total	168	8.3%	443
Worldwide Electricity	2350	2.0%	2970
ICT fraction	**7.15%**		**14.57%**

Table: Worldwide ICT power consumption. [15]

Introduction

Motivation

Category	Power cons. 2008 (GW)	Growth rate (p.a.)	2020 prediction (GW)
Data centers	29	12%	113
PCs	30	7.5%	71
Networking Equipment	25	12%	97
TVs	44	5%	79
Other	40	5%	72
Total	168	8.3%	443
Worldwide Electricity	2350	2.0%	2970
ICT fraction	**7.15%**		**14.57%**

Table: Worldwide ICT power consumption. [15]

- Many data centers do not operate at full load all the time
- Power consumptions on a node are application dependent
  - Workload classes differ depending on center type
  - HPC applications, high throughput jobs, virtualization, services

Introduction

Motivation

- Important application related information are needed to make relevant decisions towards energy efficiency in large scale distributed systems.

App Monitor → App Profiler → Resource Manager → Energy Efficiency

Introduction

Motivation

- Important application related information are needed to make relevant decisions towards energy efficiency in large scale distributed systems.

App Monitor → App Profiler → Resource Manager → Energy Efficiency

- Several power models were proposed and their results depend on the hardware and benchmark used during their constructions.

Level	Type	References	Avg. error
Systemwide	Analitical (device)	[2, 3, 5, 10, 11, 12]	5%
	Gate Level Sim (global)	[4]	–
Application	Analitical (device)	[13]	0.5%
	Analitical (global)	[9, 14]	4–30%
	Statistical (global)	[8]	1.0%

Table: Power estimators for computers

Our proposal

- Requirements
  - Monitor power consumption of application
  - Compare power estimators in any environment
  - Lightweight (low overhead)
  - Modular / Easy to use / Open source

Our proposal

- Requirements
  - Monitor power consumption of application
  - Compare power estimators in any environment
  - Lightweight (low overhead)
  - Modular / Easy to use / Open source
  - Create new power models

Our proposal

- Requirements
  - Monitor power consumption of application
  - Compare power estimators in any environment
  - Lightweight (low overhead)
  - Modular / Easy to use / Open source
  - Create new power models

- Solution
  - Energy consumption library (libec)
  - Data Acquisition tool (ecdaq)
  - Data Monitoring tool (ectop)

Sensors

libec

Definition: "a device that detects or measures a physical property and records, indicates, or otherwise responds to it" (Oxford dictionary)

Sensors

libec

Definition: "a device that detects or measures a physical property and records, indicates, or otherwise responds to it" (Oxford dictionary)

- Direct measure (hardware):
  - E.g. wattmeter: measures node's electric power in watts
- Logical estimator (software):
  - Require one or more hardware sensors
  - E.g. process wattmeter: estimates the process power

Sensors

libec

Hardware sensors[2]:

- Performance Counters
- ACPI Powermeter
- Grid'5000 PDU
- Networking

Software sensors:

- CPU Usage
- Memory Usage
- Inverse CPU PE
- MinMax CPU PE

[2] Testbed: notebook, Grid5000 [6], RECS [7]

Sensors

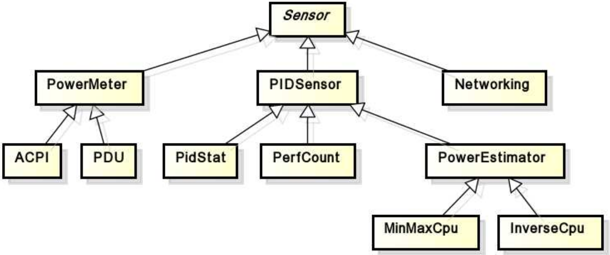
libec

Hardware sensors[2]:

- Performance Counters
- ACPI Powermeter
- Grid'5000 PDU
- Networking

Software sensors:

- CPU Usage
- Memory Usage
- Inverse CPU PE
- MinMax CPU PE



[2] Testbed: notebook, Grid5000 [6], RECS [7]

How to Create a new sensor

libec[3]

demos/ex1/MySensor.h

```
#ifndef MYSENSOR_H__
#define MYSENSOR_H__

#include <libec/sensor/SensorPid.h>

class MySensor : public cea::PIDSensor
{
public:
  MySensor();

  cea::sensor_t getValue(pid_t pid);

  void update(pid_t pid);
};

#endif
```

demos/ex1/MySensor.cpp

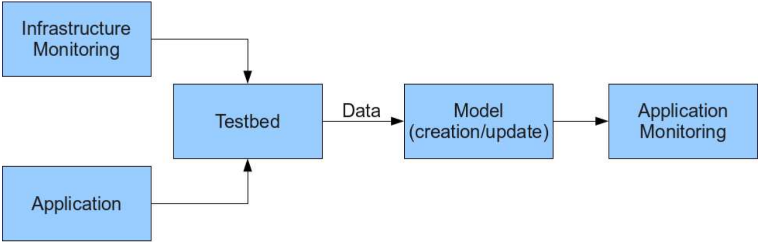
```
#include "MySensor.h"
#include <libec/tools/Tools.h>

MySensor::MySensor() {
  _name = "MySensor";
  _alias = "MS";
  _type = cea::U64;
  _isActive = true;
}

cea::sensor_t MySensor::getValue(pid_t pid) {
  update(pid);
  return _cValue;
}

void MySensor::update(pid_t pid) {
  _cValue.U64 = pid * cea::Tools::rnd(1, 10);
}
```

[3] Documentation available [html]

ECDAQ

Data Acquisition Tool

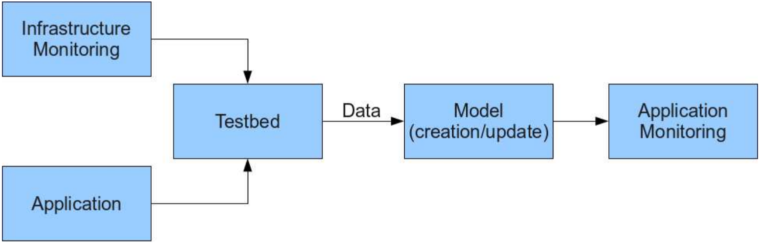
- Environment



ECDAQ

Data Acquisition Tool

- Environment



- Input: command line benchmark
- Output: gnuplot compatible file

- Demos: standard/customized

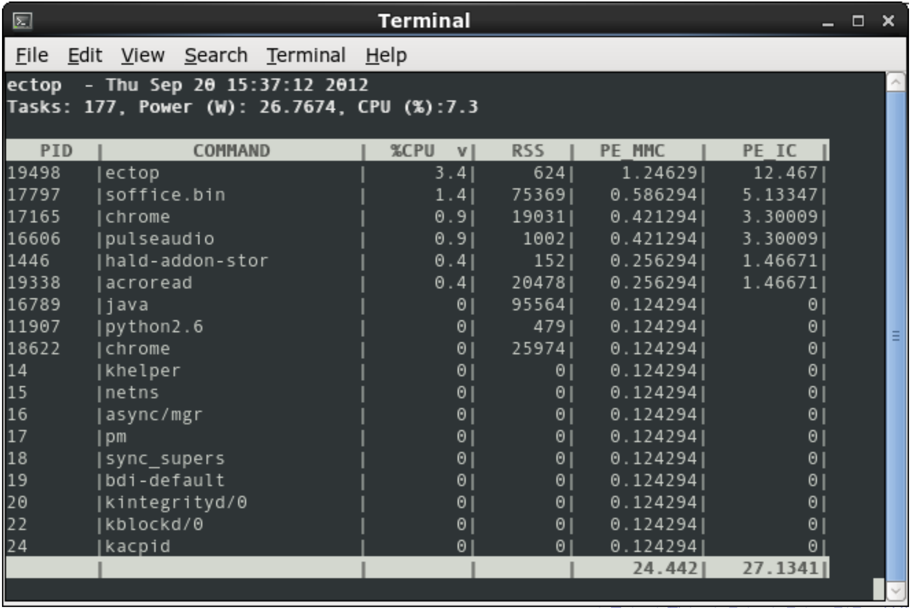
ECTOP

Data Monitoring Tool

- Functionality
  - Sort data by columns (ascending/descending)
  - Show accumulated values (sum bar)
  - Pan view
- Lightweight
  - Memory: 3Kb
  - CPU[4]:
    - 0.3% (unsorted)
    - 0.7% (sorted/sum bar)
    - 2.0% (sorted/sum bar/html)
- Documentation available [html]
- Testbed: notebook, Grid5000 [6], RECS [7]

[4] Data for a single core from top application for a Intel(R) Core(TM)2 Duo CPU T8300 @ 2.40GHz

ECTOP - Command line tool

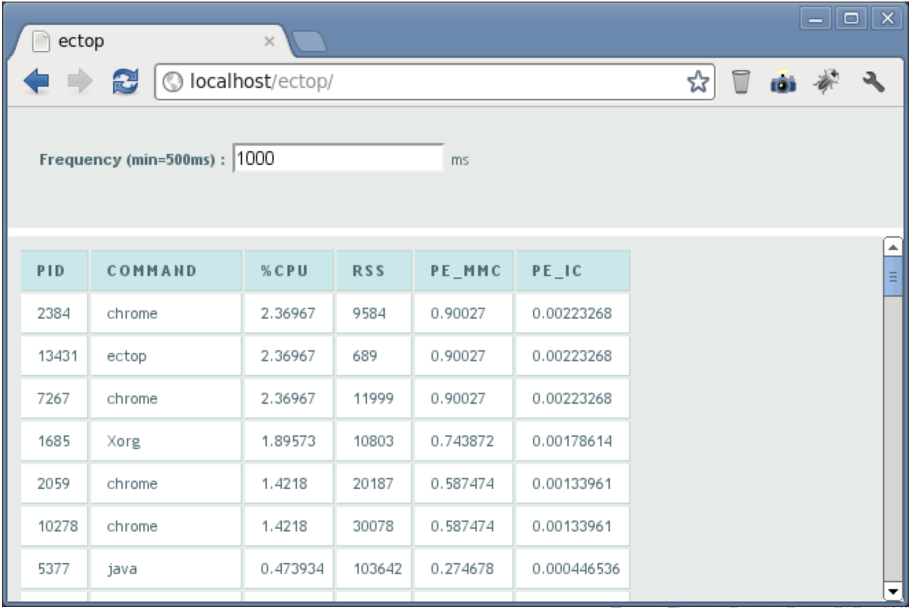

Terminal

File Edit View Search Terminal Help

ectop - Thu Sep 20 15:37:12 2012
Tasks: 177, Power (W): 26.7674, CPU (%):7.3

PID	COMMAND	%CPU v	RSS	PE_MMC	PE_IC
19498	ectop	3.4	624	1.24629	12.467
17797	soffice.bin	1.4	75369	0.586294	5.13347
17165	chrome	0.9	19031	0.421294	3.30009
16606	pulseaudio	0.9	1002	0.421294	3.30009
1446	hald-addon-stor	0.4	152	0.256294	1.46671
19338	acroread	0.4	20478	0.256294	1.46671
16789	java	0	95564	0.124294	0
11907	python2.6	0	479	0.124294	0
18622	chrome	0	25974	0.124294	0
14	khelper	0	0	0.124294	0
15	netns	0	0	0.124294	0
16	async/mgr	0	0	0.124294	0
17	pm	0	0	0.124294	0
18	sync_supers	0	0	0.124294	0
19	bdi-default	0	0	0.124294	0
20	kintegrityd/0	0	0	0.124294	0
22	kblockd/0	0	0	0.124294	0
24	kacpid	0	0	0.124294	0
				24.442	27.1341

ECTOP - HTML version



PID	COMMAND	%CPU	RSS	PE_MMC	PE_IC
2384	chrome	2.36967	9584	0.90027	0.00223268
13431	ectop	2.36967	689	0.90027	0.00223268
7267	chrome	2.36967	11999	0.90027	0.00223268
1685	Xorg	1.89573	10803	0.743872	0.00178614
2059	chrome	1.4218	20187	0.587474	0.00133961
10278	chrome	1.4218	30078	0.587474	0.00133961
5377	java	0.473934	103642	0.274678	0.000446536

Outline

PhD thesis of Leandro Fontoura Cupertino

Global model

- Entire system power usage model per application
- Embraces all components (CPU, memory, disk, bus)
- Can be extended to new architectures

PhD thesis of Leandro Fontoura Cupertino

Global model

- Entire system power usage model per application
- Embraces all components (CPU, memory, disk, bus)
- Can be extended to new architectures

Machine Learning Algorithms

- Pros
  - Architecture independent (no hardware knowledge is needed)
  - Can be updated according to the use of the data center
  - Approximates any unknown function

PhD thesis of Leandro Fontoura Cupertino

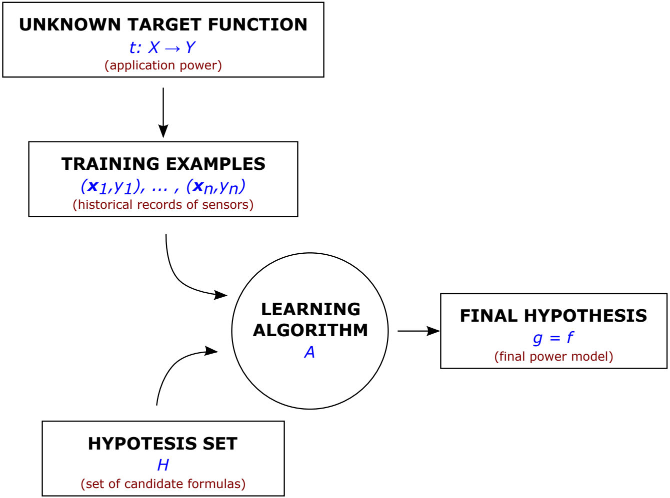
Global model

- Entire system power usage model per application
- Embraces all components (CPU, memory, disk, bus)
- Can be extended to new architectures

Machine Learning Algorithms

- Pros
  - Architecture independent (no hardware knowledge is needed)
  - Can be updated according to the use of the data center
  - Approximates any unknown function
- Cons
  - Time to learn may be high (from few seconds to several days)
  - Cannot generalize for data which was not presented to achieve the approximator

Machine Learning



Machine Learning

Hipotesis set

- Genetic Programing
- Artificial Neural Network
- Radial Basis Function

Machine Learning

Hipotesis set

- Genetic Programing
- Artificial Neural Network
- Radial Basis Function

Issues

- Need data to learn from
- Application's power cannot be measured

Machine Learning

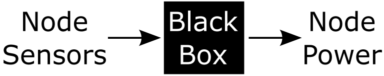
Hipotesis set

- Genetic Programing
- Artificial Neural Network
- Radial Basis Function

Issues

- Need data to learn from
- Application's power cannot be measured

How we deal with lack of information?



Machine Learning

Hipotesis set

- Genetic Programing
- Artificial Neural Network
- Radial Basis Function

Issues

- Need data to learn from
- Application's power cannot be measured

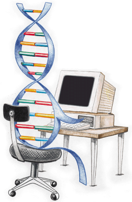
How we deal with lack of information?

Linear Genetic Programming (LGP)

- Evolutionary algorithm (population based)
- Representation: sequence of instructions from imperative programming language or machine code
- Advantage over graph representation
  - Imperative programming: flexibility (interpreted)
  - Machine code: speed (directly executed on the CPU) *
- Each program is a linear sequence of instructions
- Number of instructions: fixed or variable

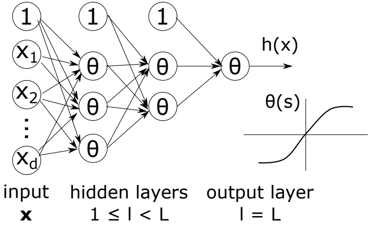
Exemple of a LGP representation

Instruction 1	Instruction 2	...	Instruction N



Artificial Neural Networks

- Network layout [1]



$$x_j^{(l)} = \theta\left(\sum_{i=0}^{d^{l-1}} w_{ij}^{(l)} x_i^{(l-1)}\right) \quad (1)$$

where $\theta(s) = \tanh(s)$

- Backpropagation *

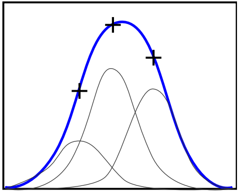
$$\Delta w_{ij}^{(l)} = -\eta x_i^{(l-1)} \delta_j^{(l)} \quad (2)$$

$$\delta_i^{(l-1)} = \left(1 - \left(x_i^{(l-1)}\right)^2\right) \sum_{j=1}^{d^{(l)}} w_{ij}^{(l)} \delta_j^{(l)} \quad (3)$$

Radial Basis Function

- Definition [1]

$$h(\mathbf{x}) = \sum_{k=1}^{K} w_k \exp\left(-\gamma ||\mathbf{x} - \mu_k||^2\right) \qquad (4)$$



Choose μ_k's: Lloyd's algorithm
Choose w_k's: Pseudo-inverse

- RBF and Regularization

$$\sum_{n=1}^{N} (h(x_n) - y_n)^2 + \lambda \sum_{k=0}^{\infty} a_k \int_{-\infty}^{\infty} \left(\frac{d^k h}{dx^k}\right) dx \qquad (5)$$

"smoothest interpolation"

Questions?

More info at: http://www.irit.fr/~Jean-Marc.Pierson or pierson @ irit.fr
and http://www.coolemall.eu

References I

[1] Abu-Mostafa, Y. S.
Learning from Data (http://work.caltech.edu/).
2012.

[2] Allalouf, M., Arbitman, Y., Factor, M., Kat, R., Meth, K., and Naor, D.
Storage modeling for power estimation.
In *Proceedings of SYSTOR 2009: The Israeli Experimental Systems Conference* (2009), ACM, p. 3.

[3] Basmadjian, R., and de Meer, H.
Evaluating and modeling power consumption of multi-core processors.
In *Future Energy Systems: Where Energy, Computing and Communication Meet (e-Energy), 2012 Third International Conference on* (2012), IEEE, pp. 1–10.

[4] Beck, A., Mattos, J., Wagner, F., and Carro, L.
Caco-ps: a general purpose cycle-accurate configurable power simulator.
In *Integrated Circuits and Systems Design, 2003. SBCCI 2003. Proceedings. 16th Symposium on* (2003), IEEE, pp. 349–354.

[5] Brooks, D., Tiwari, V., and Martonosi, M.
Wattch: a framework for architectural-level power analysis and optimizations.
In *ACM SIGARCH Computer Architecture News* (2000), vol. 28, ACM, pp. 83–94.

[6] Cappello, F., Caron, E., Dayde, M., Desprez, F., Jégou, Y., Primet, P., Jeannot, E., Lanteri, S., Leduc, J., Melab, N., et al.
Grid'5000: a large scale and highly reconfigurable grid experimental testbed.
In *Grid Computing, 2005. The 6th IEEE/ACM International Workshop on* (2005), IEEE, pp. 8–pp.

[7] christmann informationstechnik + medien GmbH & Co. KG.
RECS — Server Duo 510.

[8] Da Costa, G., and Hlavacs, H.
Methodology of Measurement for Energy Consumption of Applications.
In *Grid Computing (GRID), 2010 11th IEEE/ACM International Conference on* (Oct. 2010), pp. 290–297.

References II

[9] Do, T., Rawshdeh, S., and Shi, W.
ptop: A process-level power profiling tool.
In *Proceedings of the 2nd Workshop on Power Aware Computing and Systems (HotPower09)* (2009).

[10] Joseph, R., and Martonosi, M.
Run-time power estimation in high performance microprocessors.
In *Proceedings of the 2001 international symposium on Low power electronics and design* (2001), ACM, pp. 135–140.

[11] Kadayif, I., Chinoda, T., Kandemir, M., Vijaykirsnan, N., Irwin, M., and Sivasubramaniam, A.
vEC: virtual energy counters.
In *Proceedings of the 2001 ACM SIGPLAN-SIGSOFT workshop on Program analysis for software tools and engineering* (2001), ACM, pp. 28–31.

[12] Ma, X., Dong, M., Zhong, L., and Deng, Z.
Statistical power consumption analysis and modeling for gpu-based computing.
In *Proceeding of ACM SOSP Workshop on Power Aware Computing and Systems (HotPower)* (2009).

[13] Noureddine, A., Bourdon, A., Rouvoy, R., and Seinturier, L.
A preliminary study of the impact of software engineering on GreenIT.
In *Green and Sustainable Software (GREENS), 2012 First International Workshop on* (June 2012), pp. 21–27.

[14] Rivoire, S., Ranganathan, P., and Kozyrakis, C.
A comparison of high-level full-system power models.
In *Proceedings of the 2008 conference on Power aware computing and systems* (2008), USENIX Association, pp. 3–3.

[15] Vereecken, W., Van Heddeghem, W., Colle, D., Pickavet, M., and Demeester, P.
Overall ict footprint and green communication technologies.
In *Communications, Control and Signal Processing (ISCCSP), 2010 4th International Symposium on* (march 2010), pp. 1 –6.