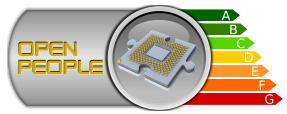
Modelling and Optimization of Power Consumption in Reconfigurable Devices

Robin BONAMY

PhD since Oct. 2009 (ANR)
CAIRN - IRISA

Advisor : Daniel CHILLET (IRISA)
Co-advisors : Sébastien BILAVARN and Olivier SENTIEYS (IRISA)

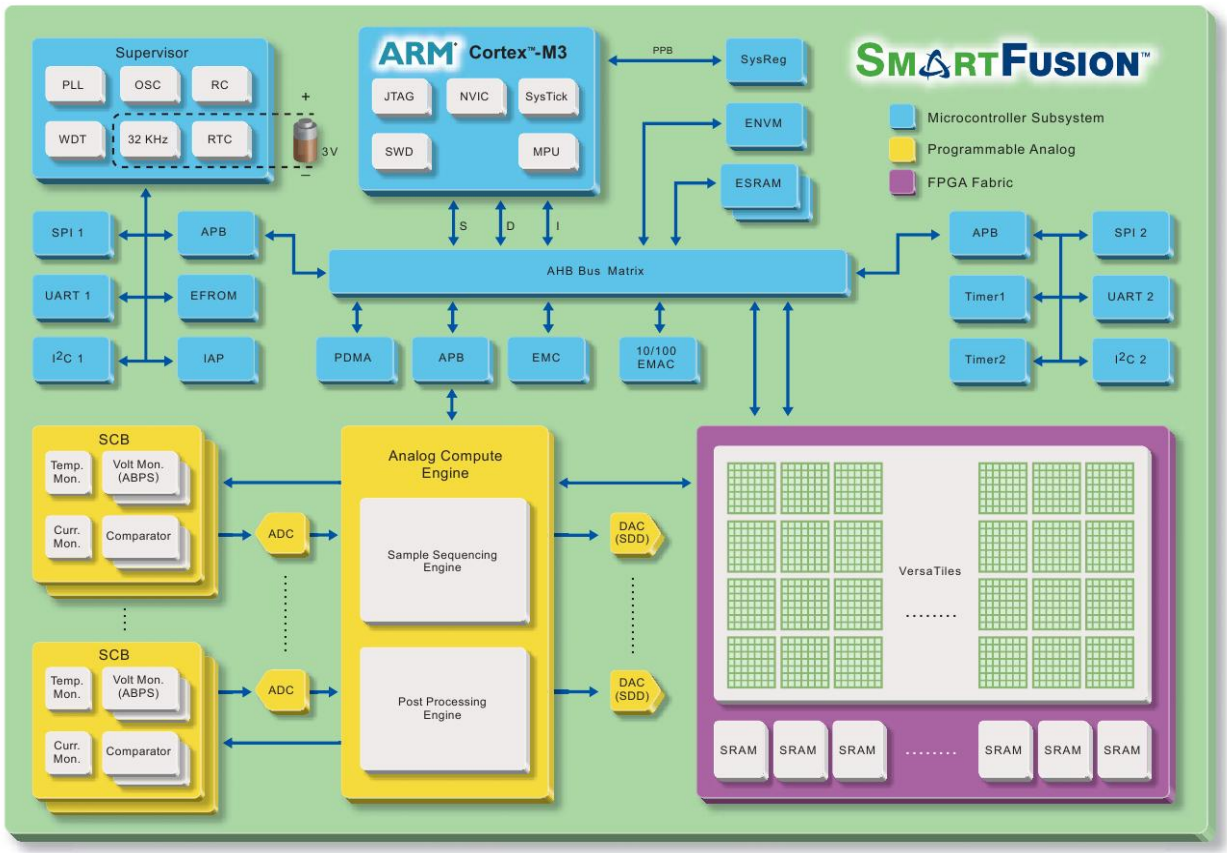
Journées Scientifiques MCSOC, June 30th 2011



System on Chip

SoC : Chip including various functions

- Processor(s)
- Configurable area(s)
- DSP(s)
- Peripheral(s)
- Memory(s)
- Analog



TI: OMAP
Xilinx: Zynq-7000
Actel: SmartFusion

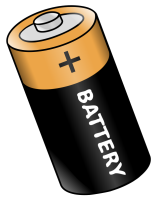
System on Chip

SoC are more and more used

- Size
- Cost

Power consumption

- Battery size/weight
- Dissipation
- Power rails design



Need to have early power estimation for heterogeneous systems

Outline

OpenPEOPLE: Who?

Consortium:

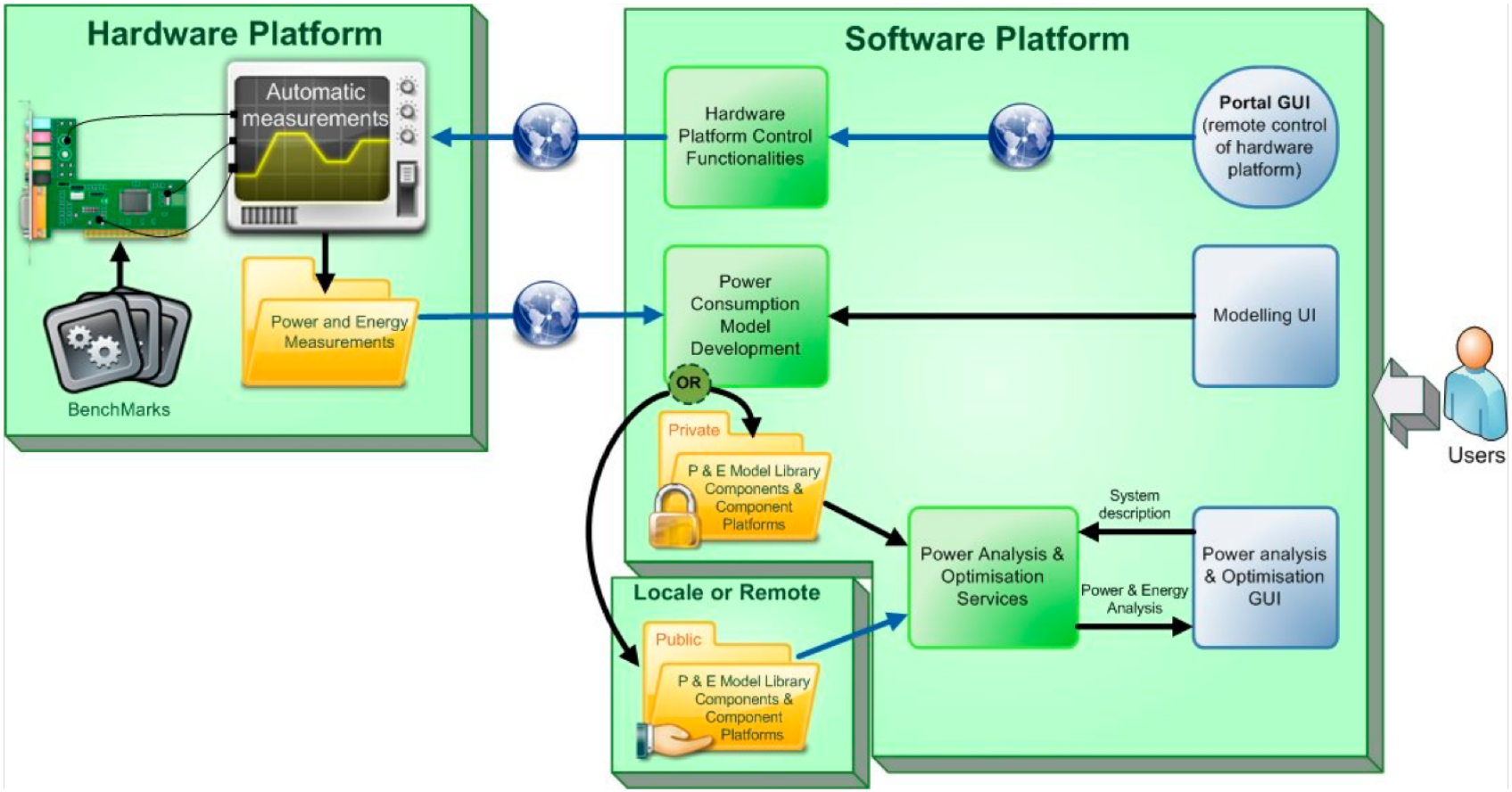
- Lab-STICC - UBS
- LORIA - INRIA Nancy
- Dart - INRIA Lille
- LEAT - UNSA
- IRISA-Cairn - UR1
- THALES Coms. - Colombes
- InPixal - Rennes

Project funded by the ANR

OpenPEOPLE: What is it?

Open-PEOPLE :

Open-Power and **E**nergy **O**ptimization **PL**atform and **E**stimator



OpenPEOPLE: What for?

Complete platform to

- allow rapid power and energy estimation and measurement for complex heterogeneous systems
- test the effects of different optimizations on power consumption

CAIRN: Power consumption models for hardware tasks and reconfiguration

LEAT: Scheduler, OS aspects

Outline

CMOS Power Consumption [Julien06], [Garcia99], [TI97]

- Dynamic Power
  - Voltage ^{2}
  - Frequency
  - Activity
  - Load (Nets Capacitance)

$$P_d = {V_{cc}}^2 \times F \times \alpha \times C$$

CMOS Power Consumption [Julien06], [Garcia99], [TI97]

- Dynamic Power
  - Voltage ^{2}
  - Frequency
  - Activity
  - Load (Nets Capacitance)

$$P_d = {V_{cc}}^2 \times F \times \alpha \times C$$

- Static Power (Leakage)
  - Area, Occupation Rate
  - Voltage

$$P_s = V_{cc} \times A$$

Power Modelling

- Fine Grain
  - Activity
  - Net length
  - Flip-flops
  - Operators
- Coarse Grain
  - Application parameters
  - RAM
  - Area
  - Delay

Pb: Power estimation done late

Pb: Bad power estimation accuracy

Power Modelling

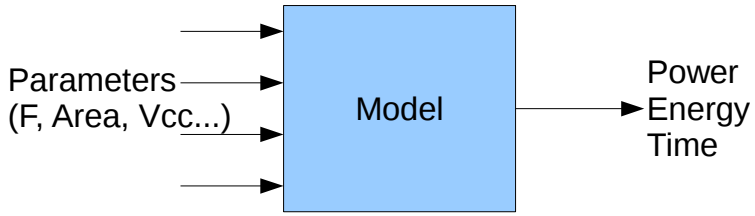
- Fine Grain
  - Activity
  - Net length
  - Flip-flops
  - Operators
- Coarse Grain
  - Application parameters
  - RAM
  - Area
  - Delay

Pb: Power estimation done late

Pb: Bad power estimation accuracy

- State Machine models [Benini00]
- Peak consumption [Gupta03]
- Glitches [Ragh96]
- ...

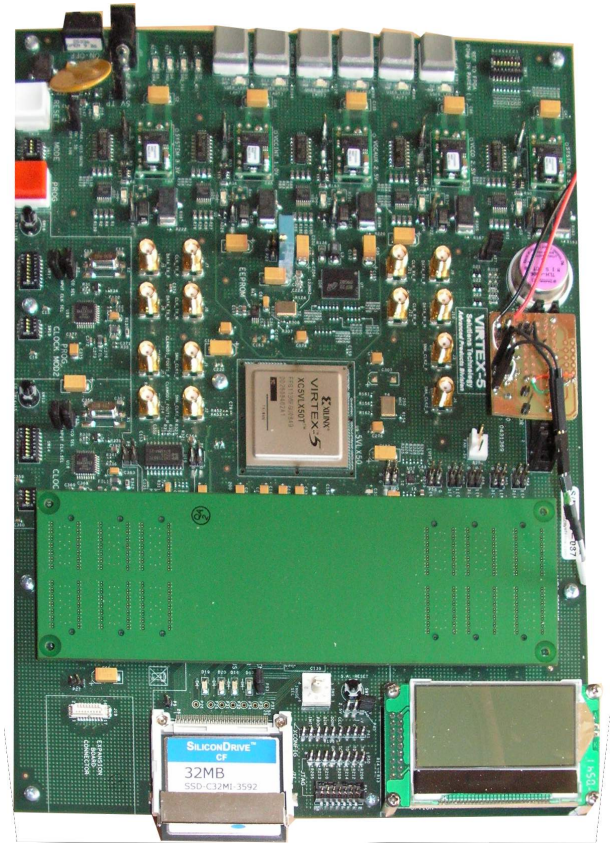
Model?



- What is a model?
  - P/E = f(parameters)
- How to build a model?
  - Measurements following parameters
  - Analysis, Statistics
  - Verification

Xilinx ML550 Board

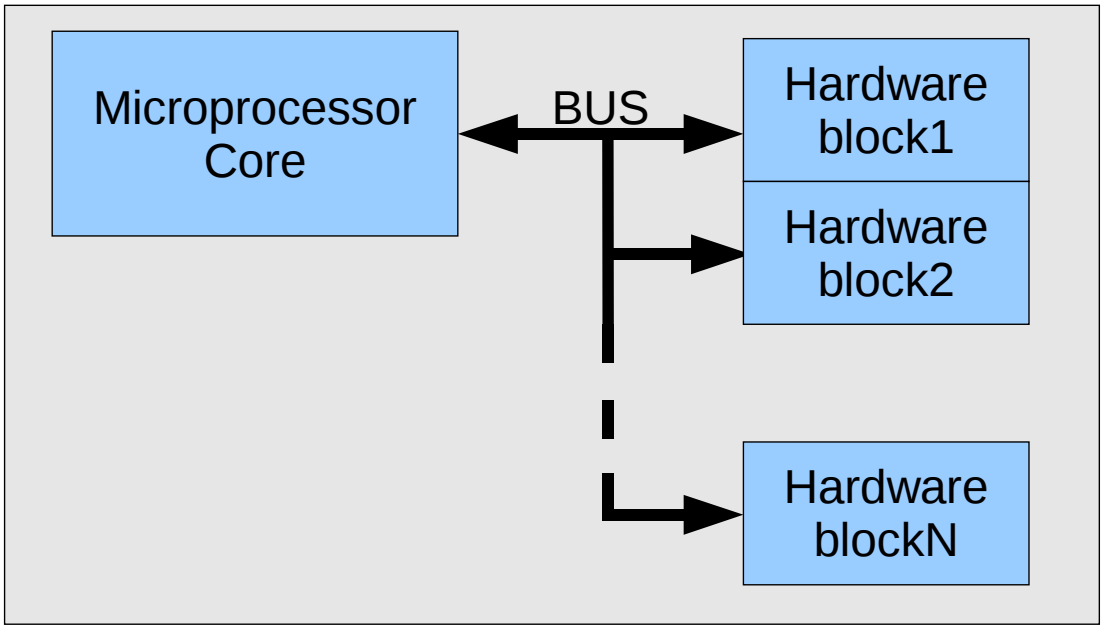
- FPGA Virtex-5 VLX50T (7200 slices)
- SystemACE CompactFlash controller
- 5 power rails (core, IOs, peripherals)
- Current sense resistors



Previous work on Actel IGLOO.
Future work planned on Virtex 6, Altera Stratix.

Outline

Hardware Blocks

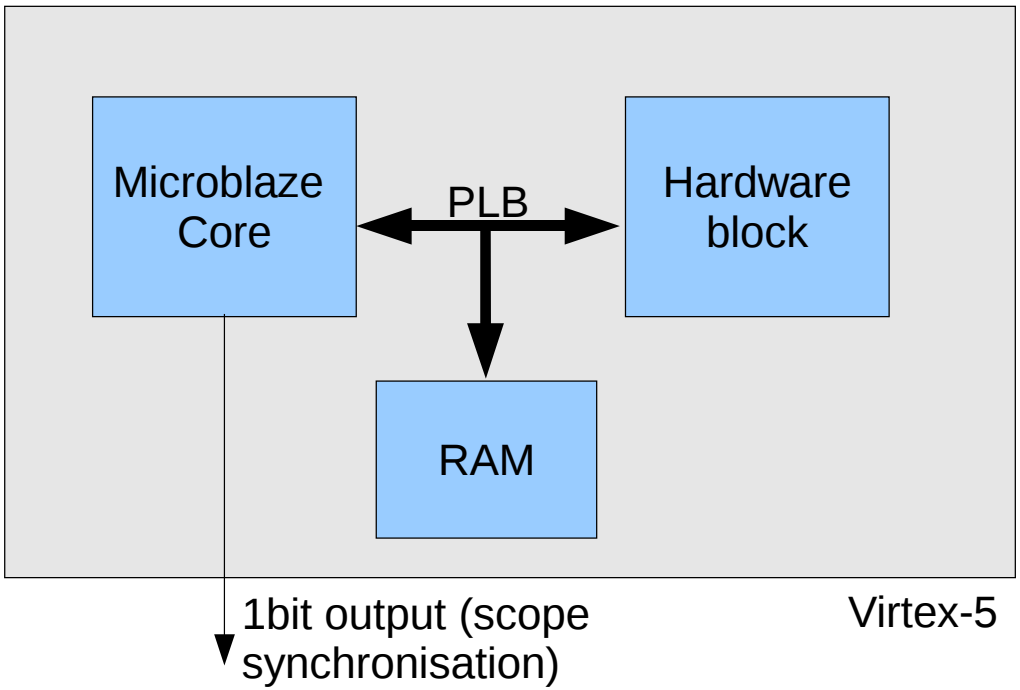


- unload processor core
- Power/Energy/Throughput efficiency [AlteraAN531]
- parallelism level

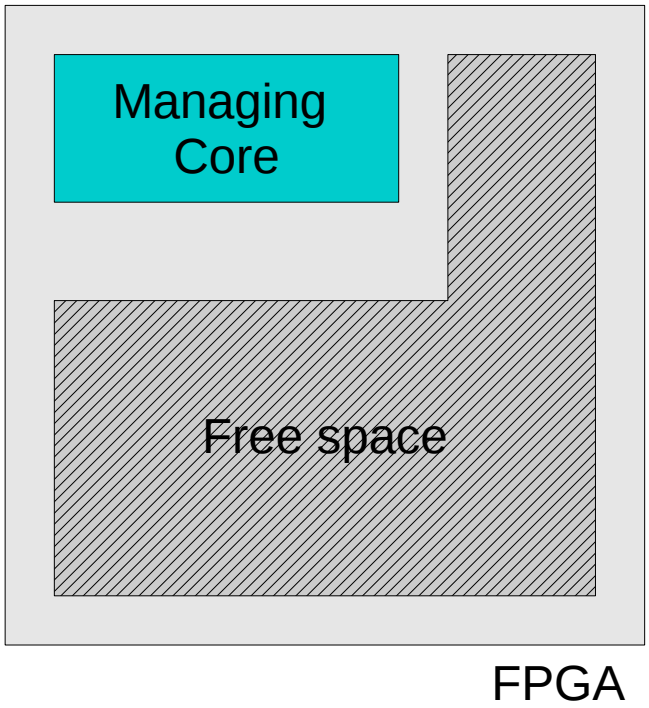
C to VHDL

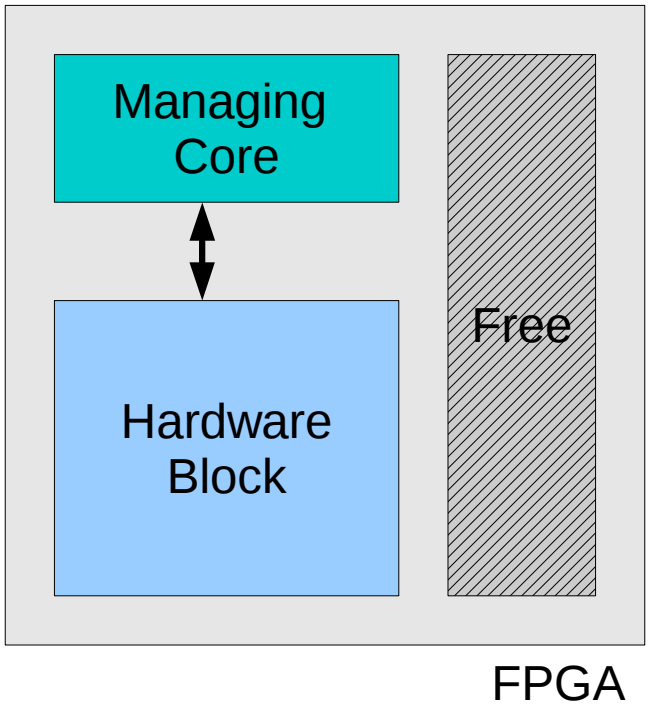
Effect of parallelism level on Power consumption

- Generation of hardware blocks
- High level synthesis (C to VHDL)
- Loop unrolling
- PLB block, Microblaze at 100MHz



Measurement Protocol





- Idle power consumption
- Active power consumption
- Execution Time

Matrix Multiplication C code

```
for (i = 0; i < M; i++)
    for (j = 0; j < P; j++)
        _C[i][j] = 0;

for (i = 0; i < M; i++)                          //(loop 1)
    for (k = 0; k < N; k++)                      //(loop 2)
        for (j = 0; j < P; j++)                  //(loop 3)
            _C[i][j] += _A[i][k]*_B[k][j];
```

Loop Unrolling Index (LUI) notation :
LUI_1, LUI_2, LUI_3

Energy vs. Execution Time

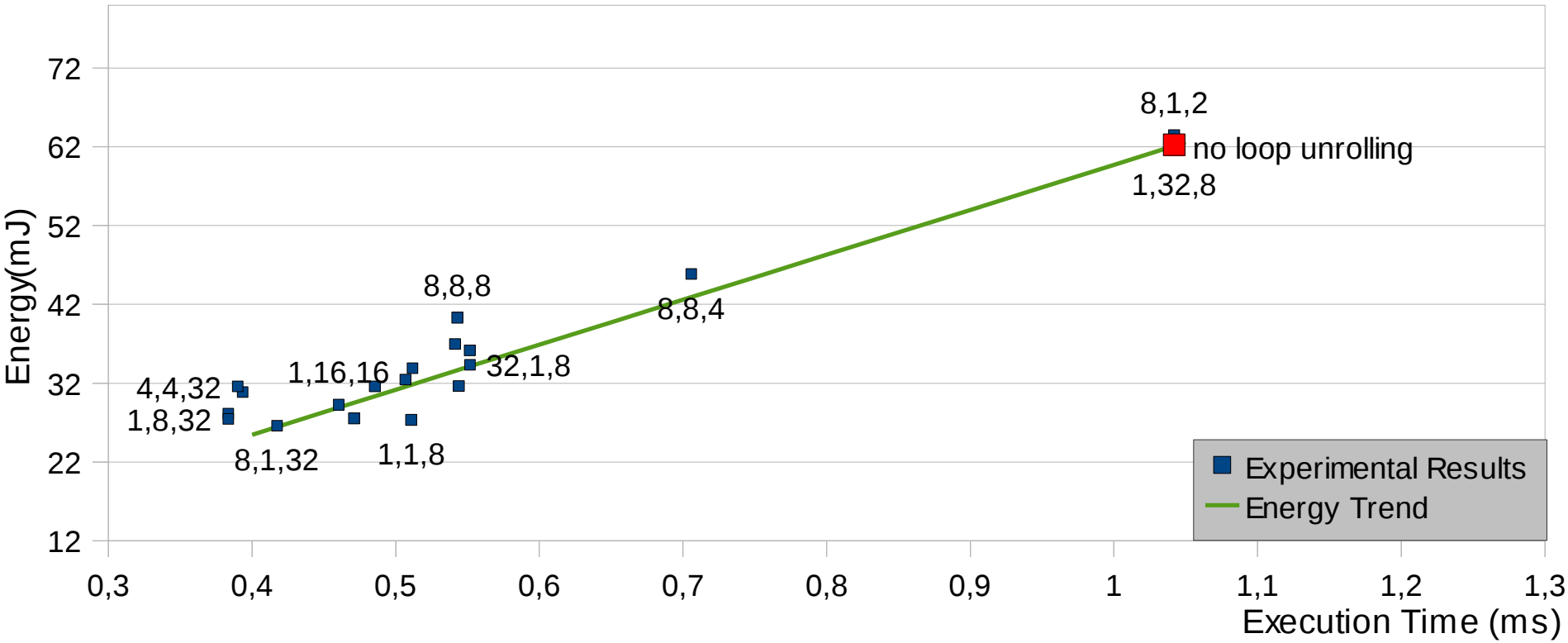


Figure 1: Energy of the matrix multiplication versus execution time.$(LUI1, LUI2, LUI3)$. Trend $E = A + B \times t$.

Two Other Algorithms

Table 1: Comparison of execution time and energy between three different implementations of each algorithm.

	Matrix mult.		Full Search		Deblock. filter	
	Time	Energy	Time	Energy	Time	Energy
Soft (ms, mJ)	10.75	244.79	0.4786	18.2	0.5742	26.09
HardS (ms, mJ)	1.04	61.99	0.0369	2.01	0.0529	3.68
HardP (ms, mJ)	0.38	27.48	0.0246	1.05	0.0417	2.74
Soft/HardP **Ratio**	**28.29**	**8.91**	**19.37**	**17.33**	**13.77**	**9.52**
HardS/HardP **Ratio**	**2.74**	**2.26**	**1.50**	**1.91**	**1.26**	**1.34**

Soft represents an execution on the microblaze core,
HardS represents a sequential implementation of the hardware task,
HardP represents the best paralleled solution in terms of time.

Energy is not a constant

→ When space is available, exploit parallelism of each task

Energy is not a constant

→ When space is available, exploit parallelism of each task

↗ parallelism level
↗ area
↗ bitstream
→ Impact of reconfiguration ?

Outline

Dynamic Partial Reconfiguration

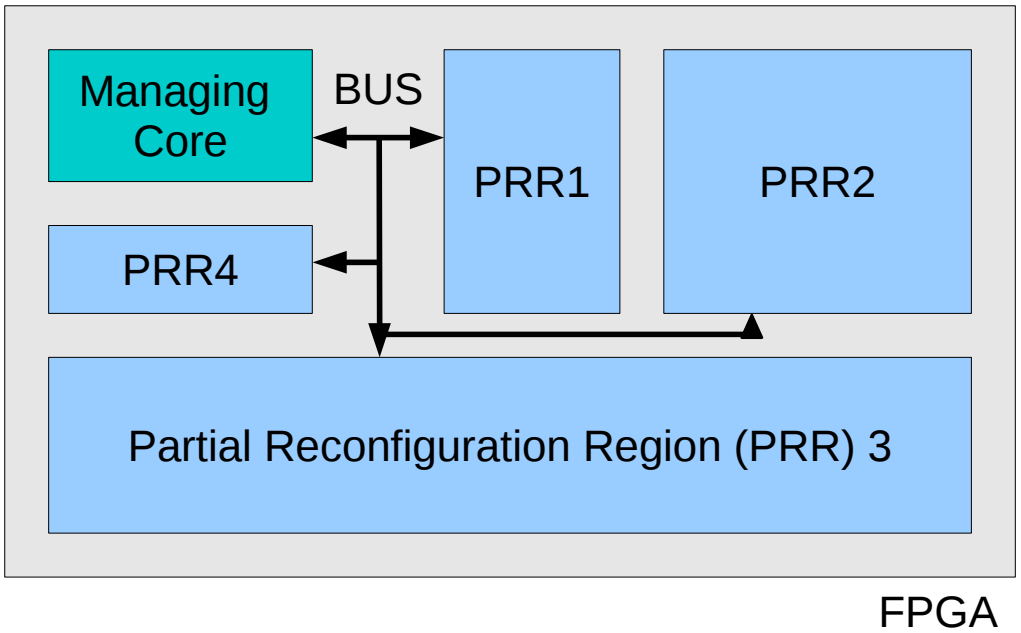
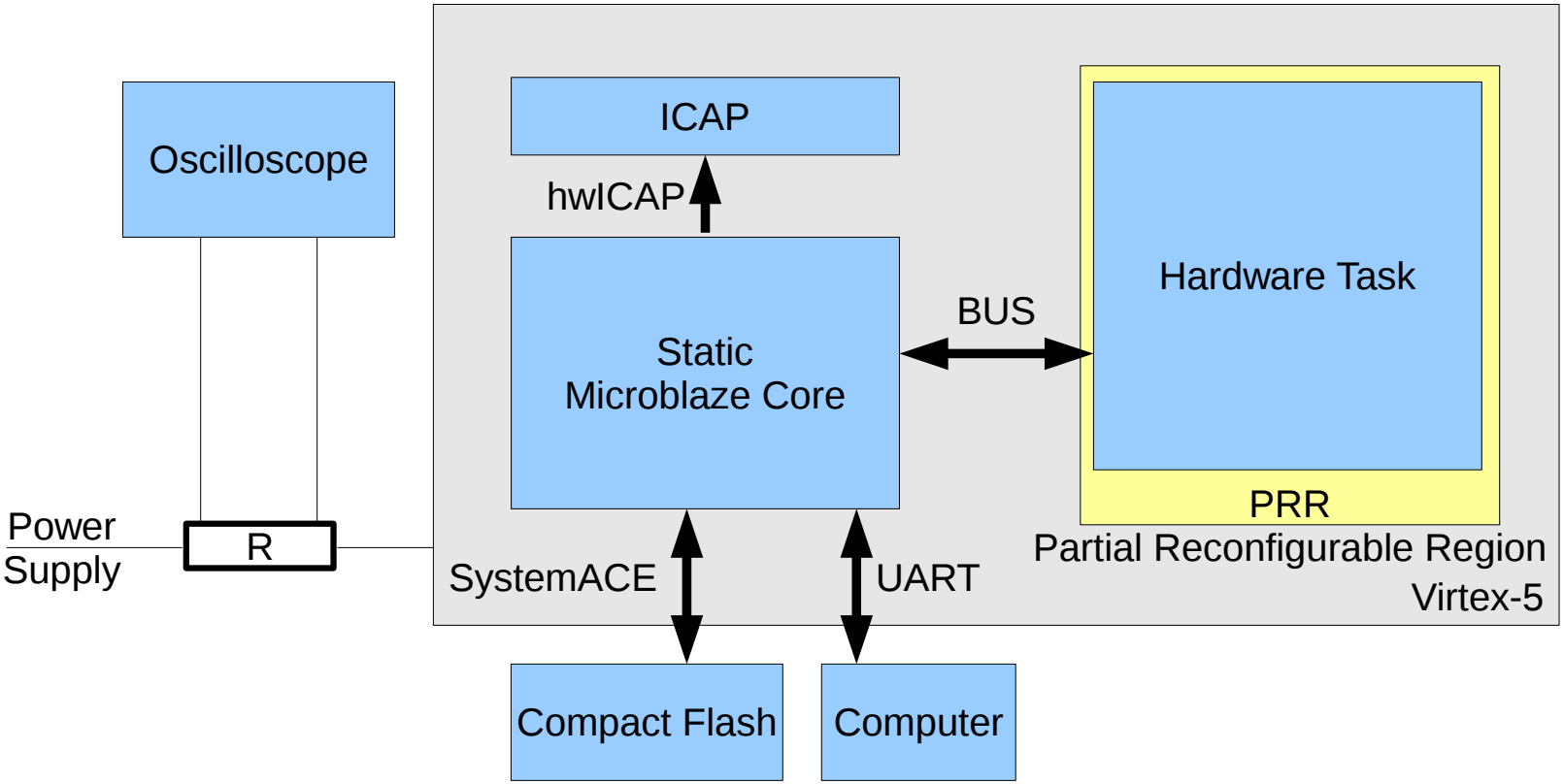


Figure : Partial Reconfiguration segmentation example

Partial reconfiguration (PR) is studied for space and energy saving. [Savary07], [Becker03]
No detailed model for PR power consumption

Dynamic Partial Reconfiguration

A Microblaze soft. core and a hardware task.

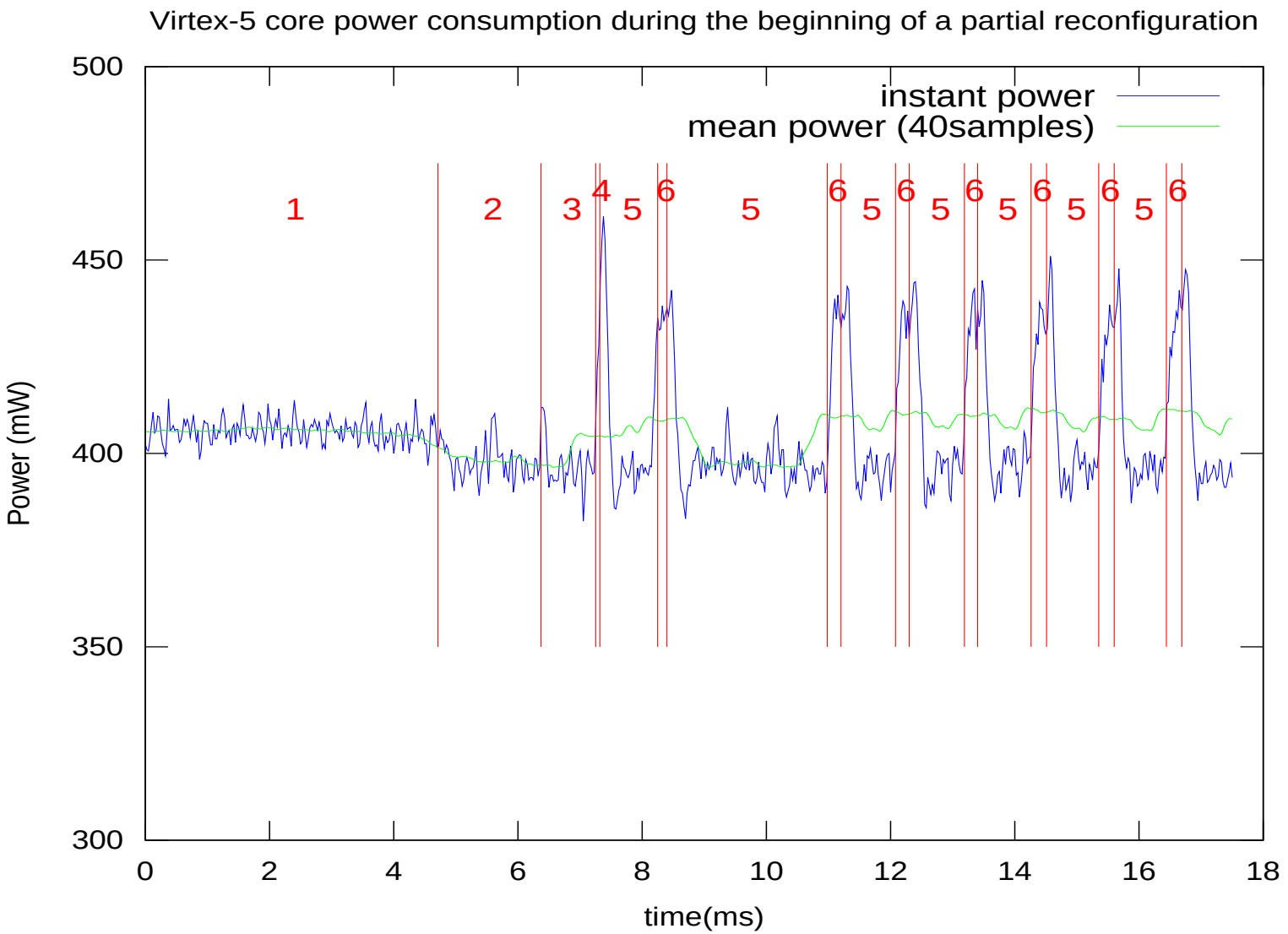


PR Steps

(1) reconfiguration **order** arrives
(2) **open** bitstream's file
(3) **read** bitstream's header
(4) **check** header's validity
(5) **read** of a bitstream sector
(6) **write** data to ICAP (Internal Configuration Access Port)
(7) repeat (5) and (6) until the end of the bitstream.

Power Consumption at the Beginning of PR

(2) **open**
(3) **read**
(4) **check**
(5) **read**
(6) **write**



Power Consumption during PR

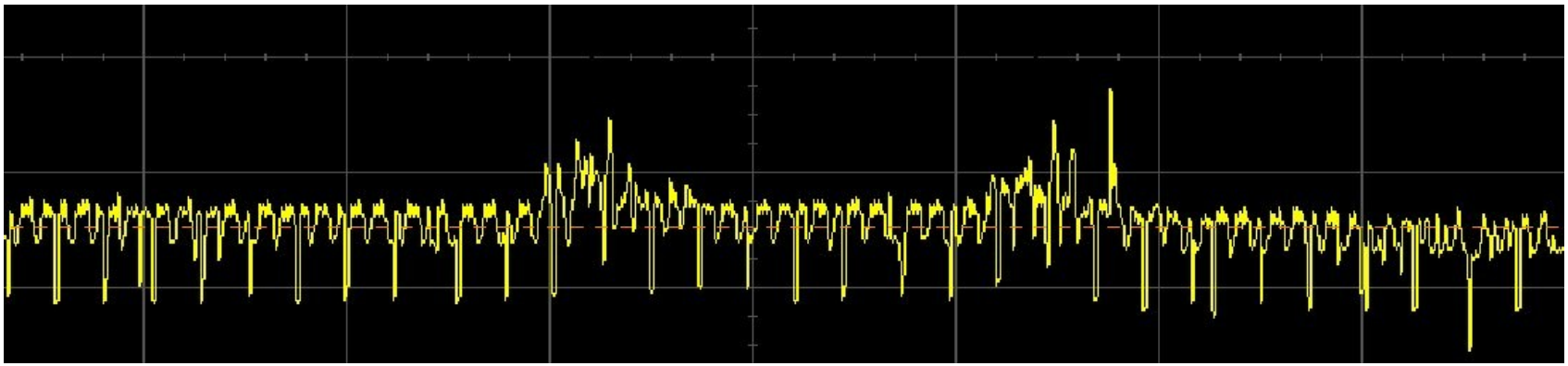


Figure : Partial Reconfiguration of PRR from a task to another one bottom: Core power consumption during PR

Power Consumption during PR

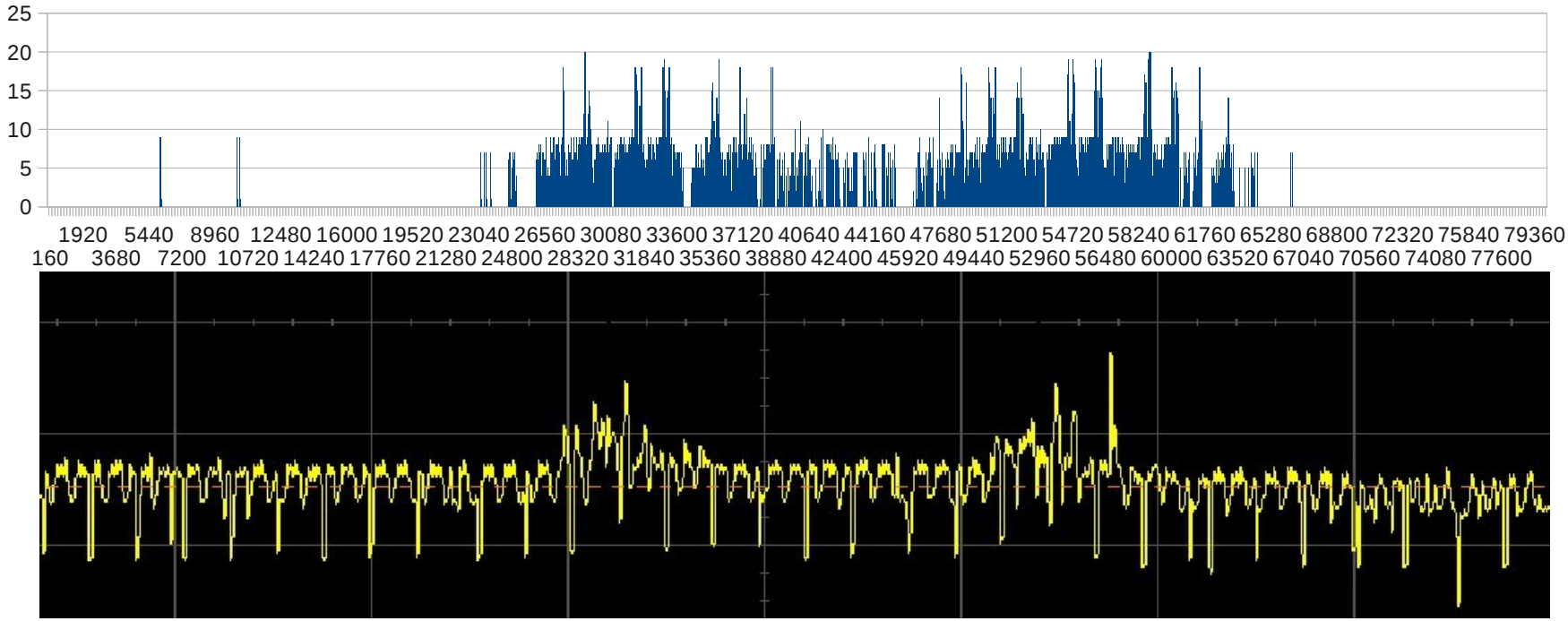


Figure : Partial Reconfiguration of PRR from a task to another one
top: Hamming distance between bitstreams of Task1 and Task2
Core power consumption during PR

Power Model from Parameters

Parameters that affect energy consumption

- Memory access
- Activity of the managing core (Read and write)
- PRR Area/Bitstream size
- Difference with the previous configuration

Parameters that don't affect energy consumption

- Shape of the Partial Reconfiguration Region
- Bitstream composition

Parameters to study

- Other: memory, management core, ICAP IP, device

Trivial model (Virtex5 VLX50T):

$E \simeq 39\mu J$ *per* kB with CF, microblaze @ 100MHz, Vcore=1V

Power Consumption during PR

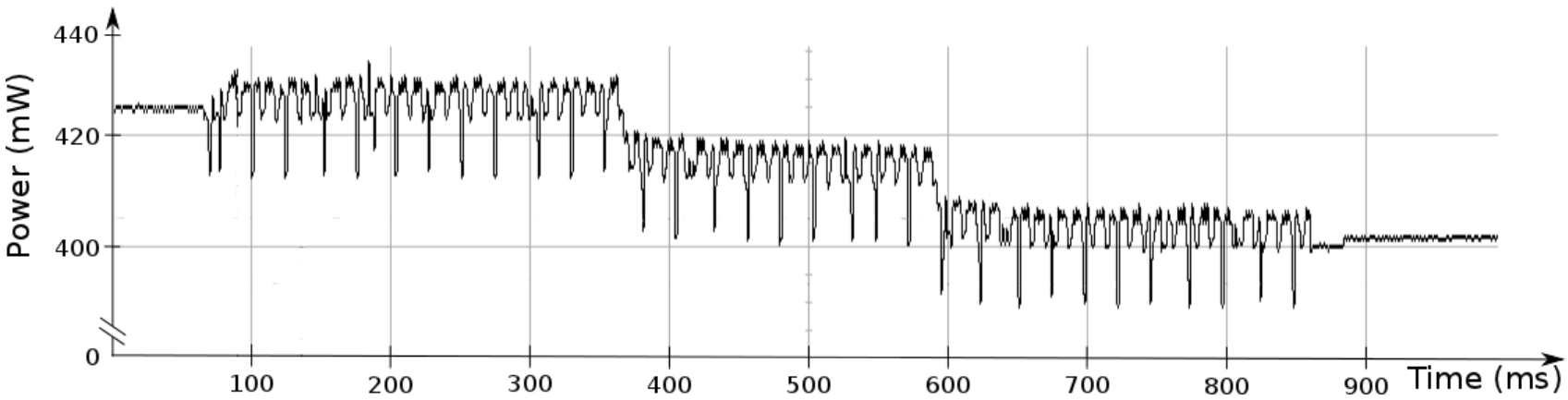
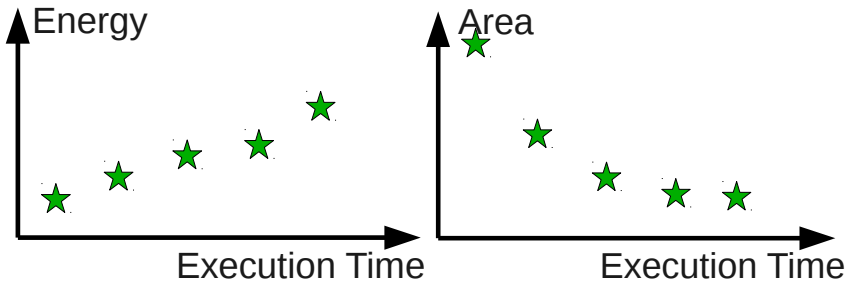


Figure : Partial Reconfiguration of PRR from a matrix multiplication hardware task to a low power task

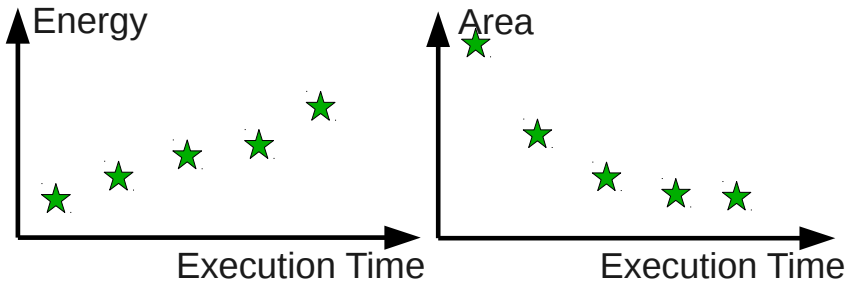
Future Work

Tasks



Future Work

Tasks



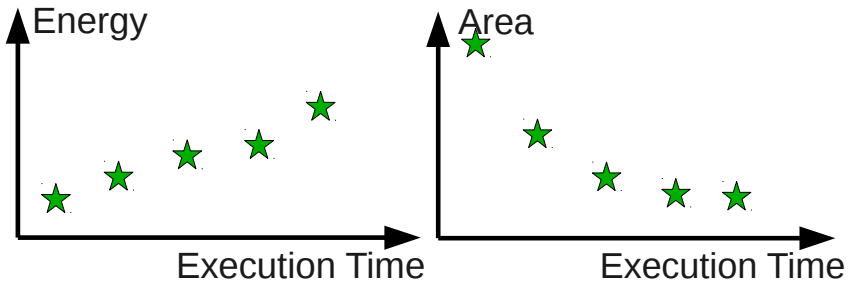
+

Partial Reconfiguration

$$Power/Energy = f(bitstream, area, F, ... V, memory, device...)$$

Future Work

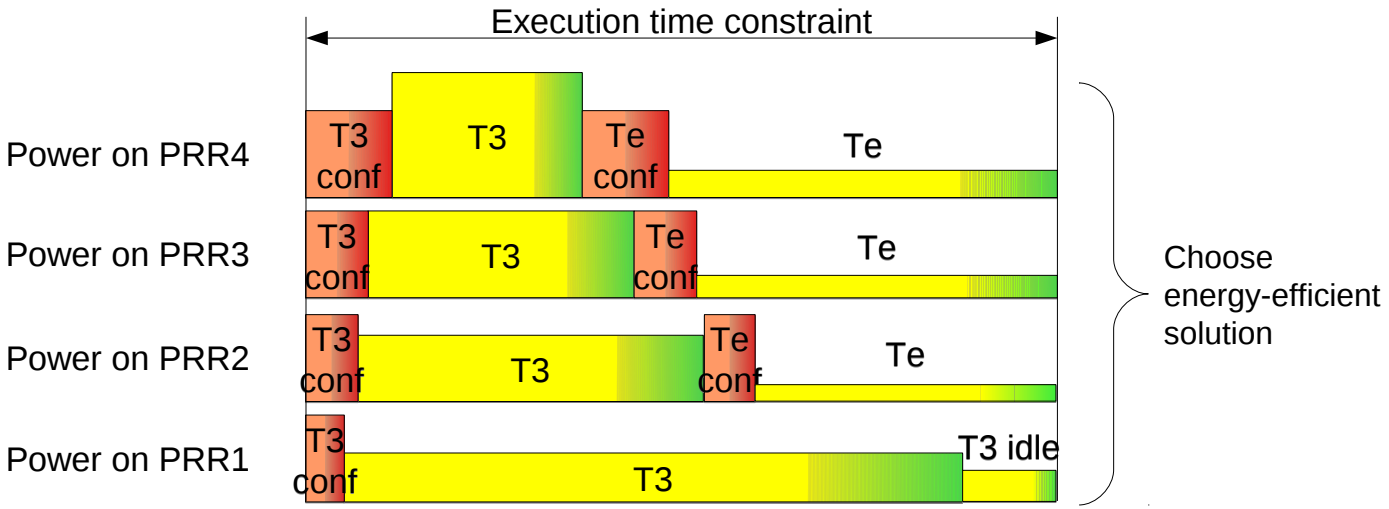
Tasks



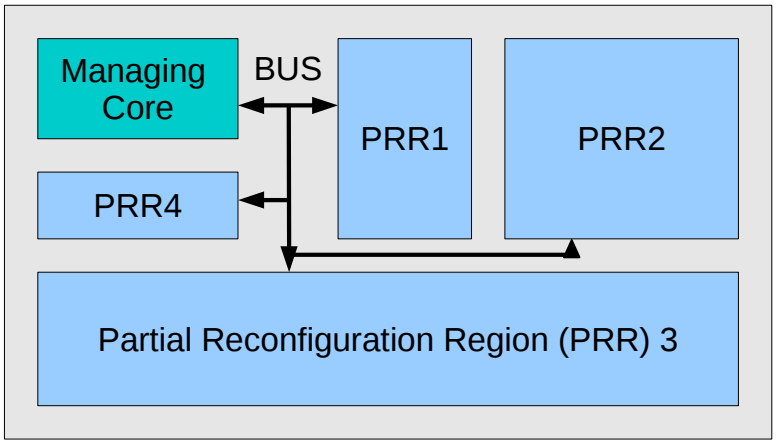
+

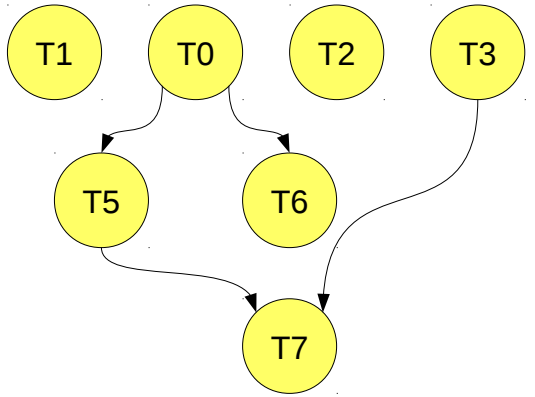
Partial Reconfiguration

$$Power/Energy = f(bitstream, area, F, ... V, memory, device...)$$



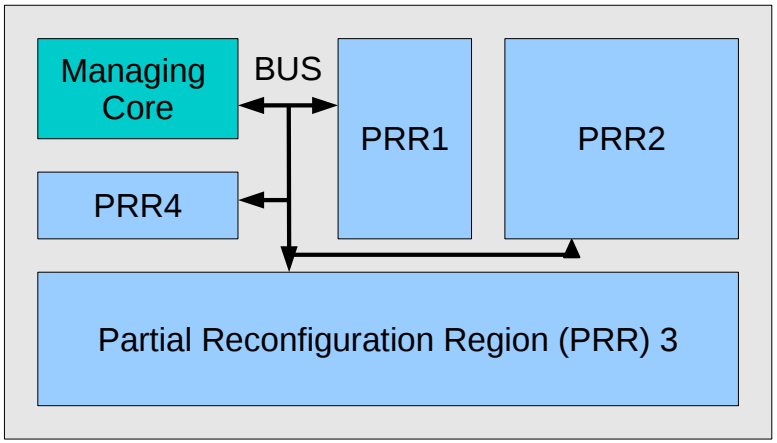
Future Work

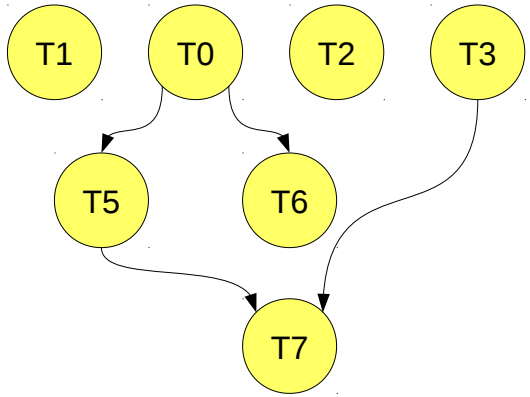




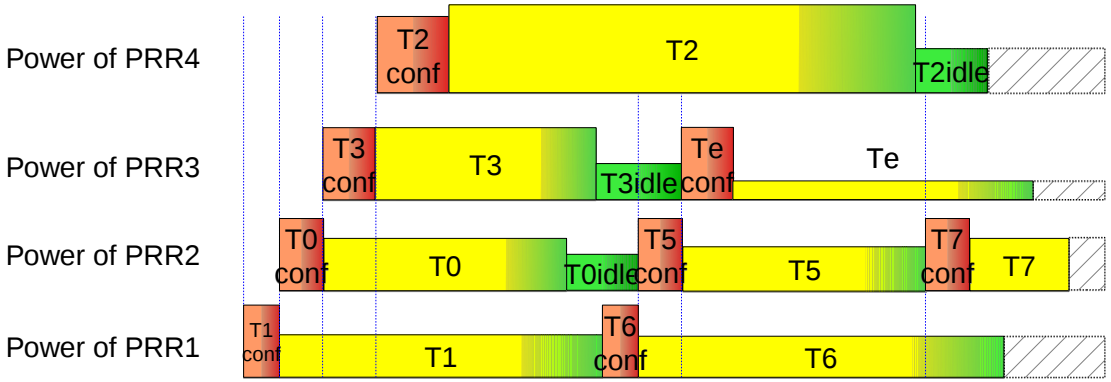
Application flow graph example.

Future Work





Application flow graph example.



Bibliography


Julien N., Caractérisation et modélisation de la consommation sur FPGA, ECOFAC 2006


Garcia A. and al., Power modelling in Field Programmable Gate Arrays (FPGA), In International Workshop on Field Programmable Logic and Applications 1999


Texas Instruments, CMOS Power Consumption and Cpd Calculation, 1997


Altera Corporation, AN 531: Reducing Power with Hardware Accelerators, 2008


Savary Y., Thèse : Etude du Potentiel des Architectures Reconfigurables pour Maîtriser la Consommation dans les Applications Embarquées, 2007


Becker J. and al., Power estimation and power measurement of Xilinx Virtex FPGAs: trade-offs and limitations, in Integrated Circuits and Systems Design, 2003


Benini L. and al., Regression-based RTL power models for controllers. In Proceedings of the 10th Great Lakes symposium on VLSI, 2000


Gupta S. and al., Energy and peak-current per-cycle estimation at RTL. Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, 2003


Raghunathan A. and al., Register-transfer level estimation techniques for switching activity and power consumption. In Proceedings of the 1996

Power and Parallelization

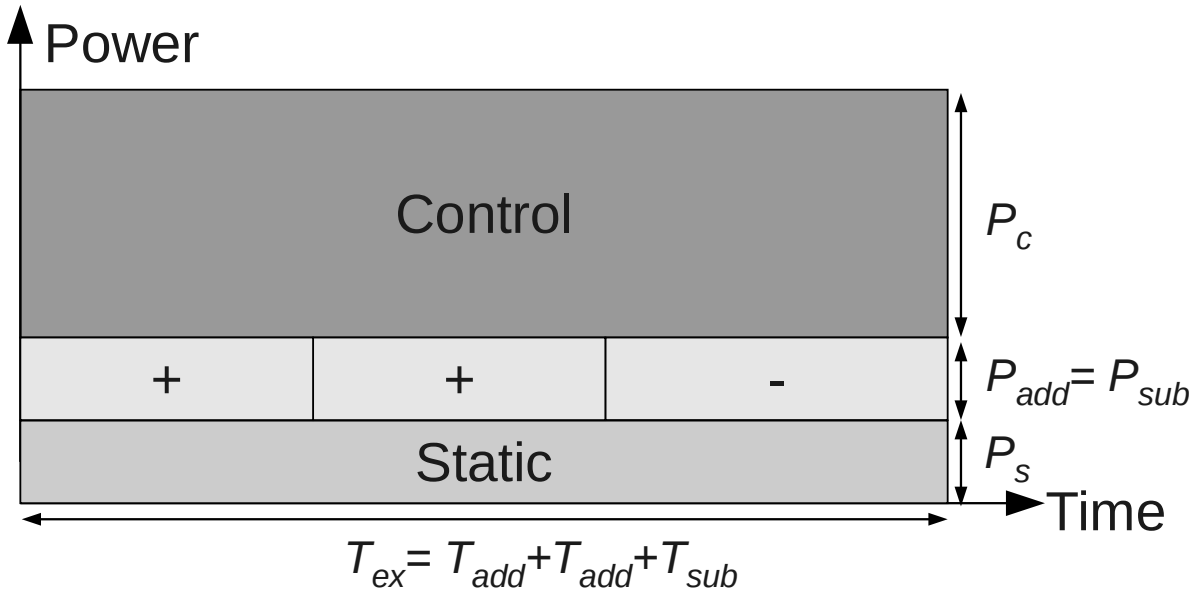


Figure 2: Representation of power consumption for a hardware task

Power and Parallelization

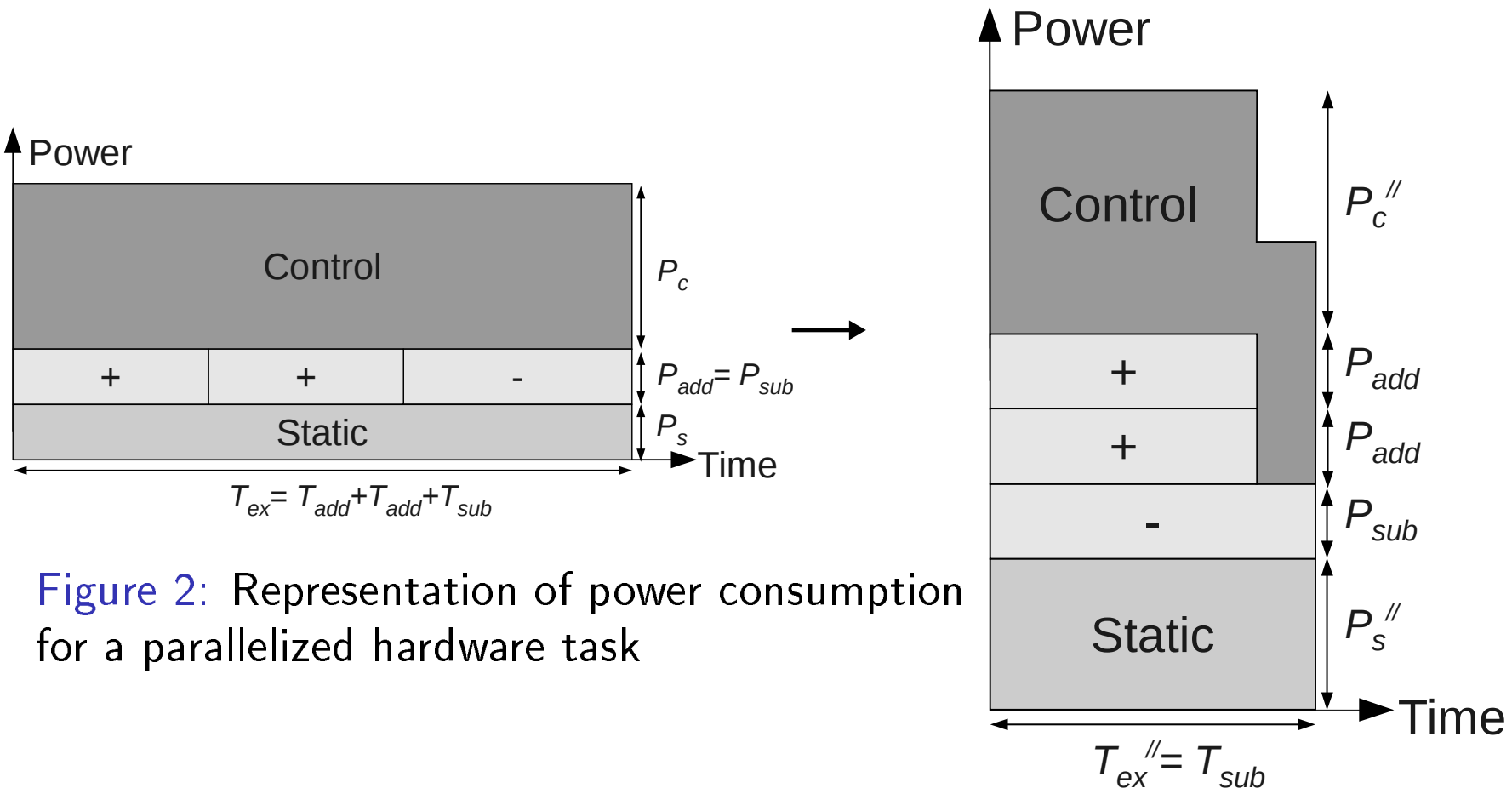


Figure 2: Representation of power consumption for a parallelized hardware task

Execution time versus Loop Unrolling Index

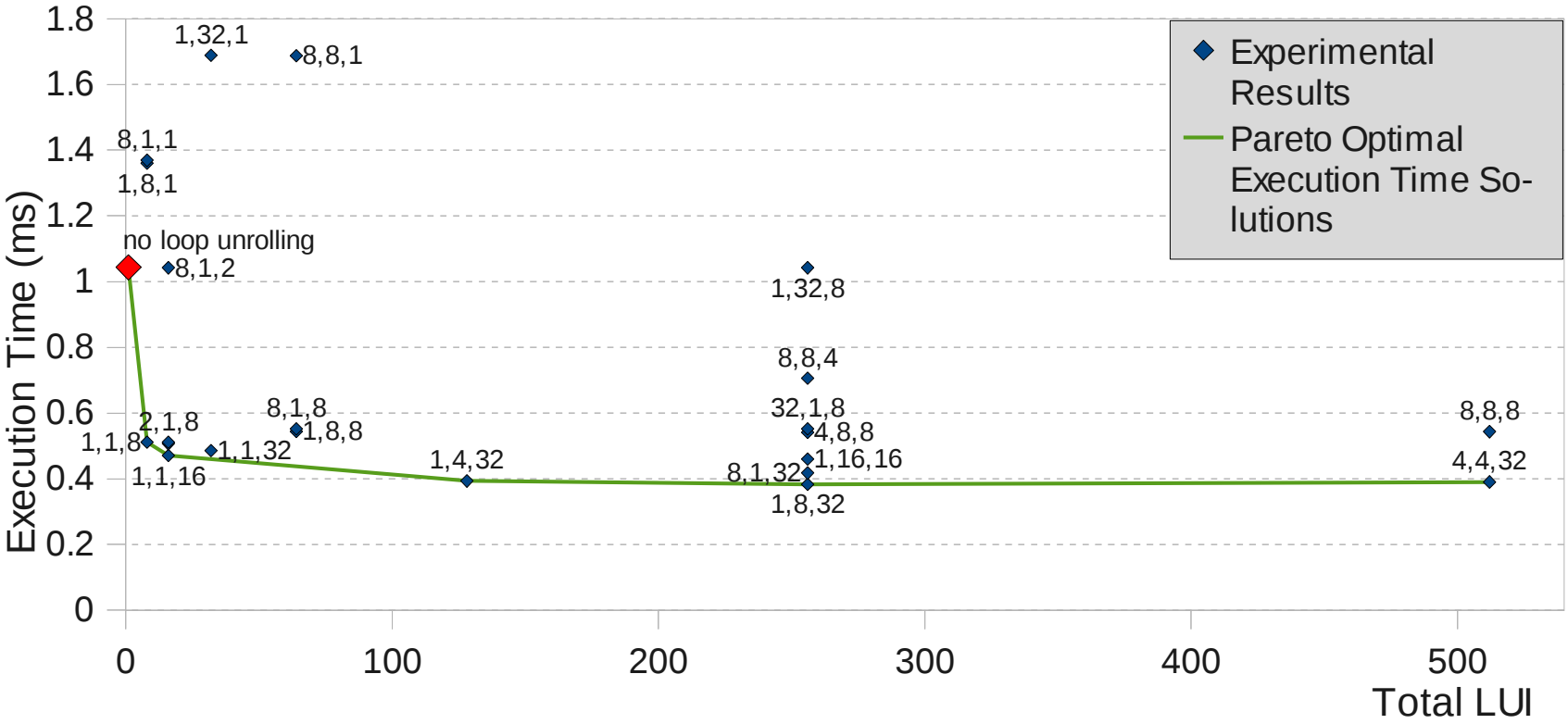


Figure 3: Matrix multiplication execution time versus total loop unrolling index.($LUI1, LUI2, LUI3$)

Energy consumption versus Loop Unrolling Index

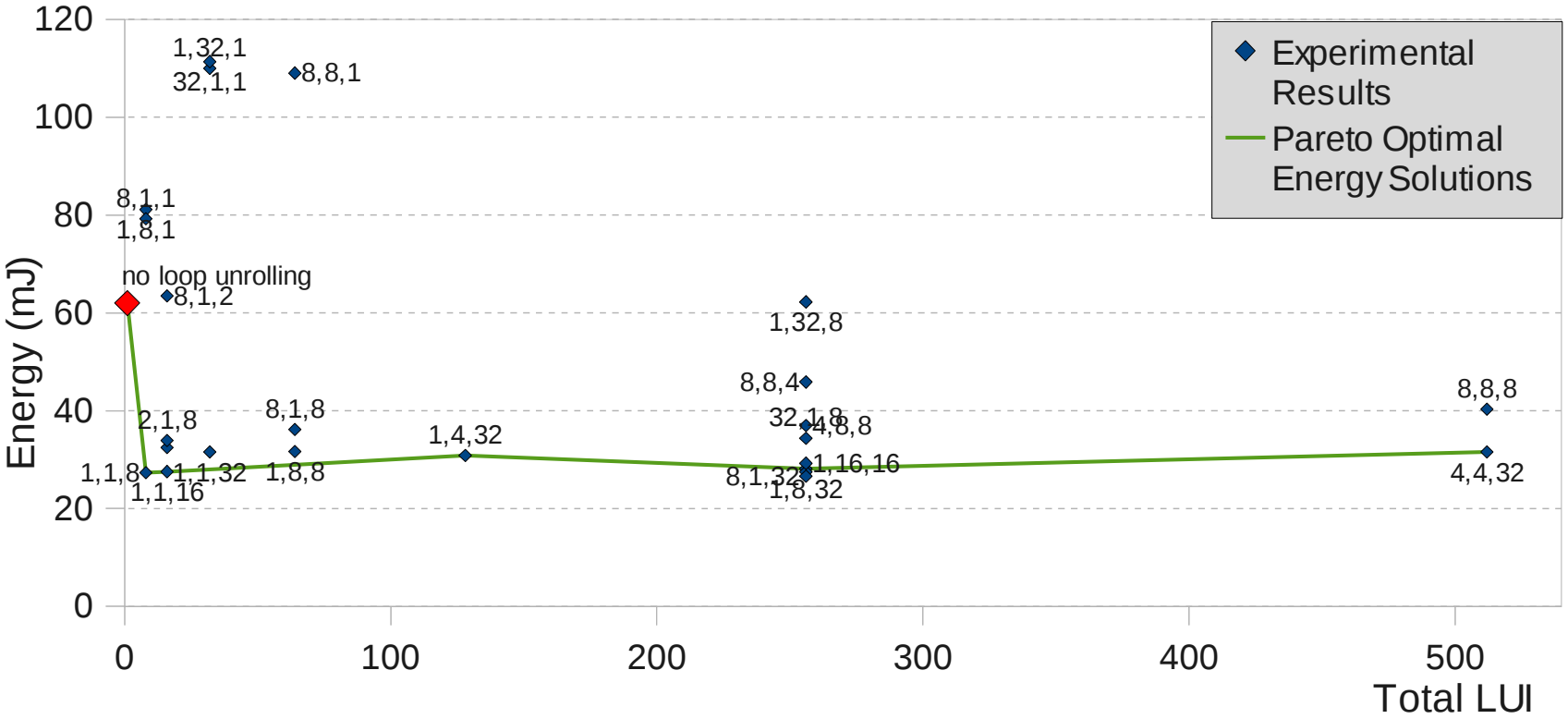


Figure 4: Power consumption measurements : matrix multiplication energy consumption versus total loop unrolling index.($LUI1, LUI2, LUI3$)