

CPU POWER ESTIMATION USING PMCs, AND ITS APPLICATION IN gem5

Dr Geoff Merrett

Arm Research Summit, 11 September 2017

OVERVIEW

Introduction and Background

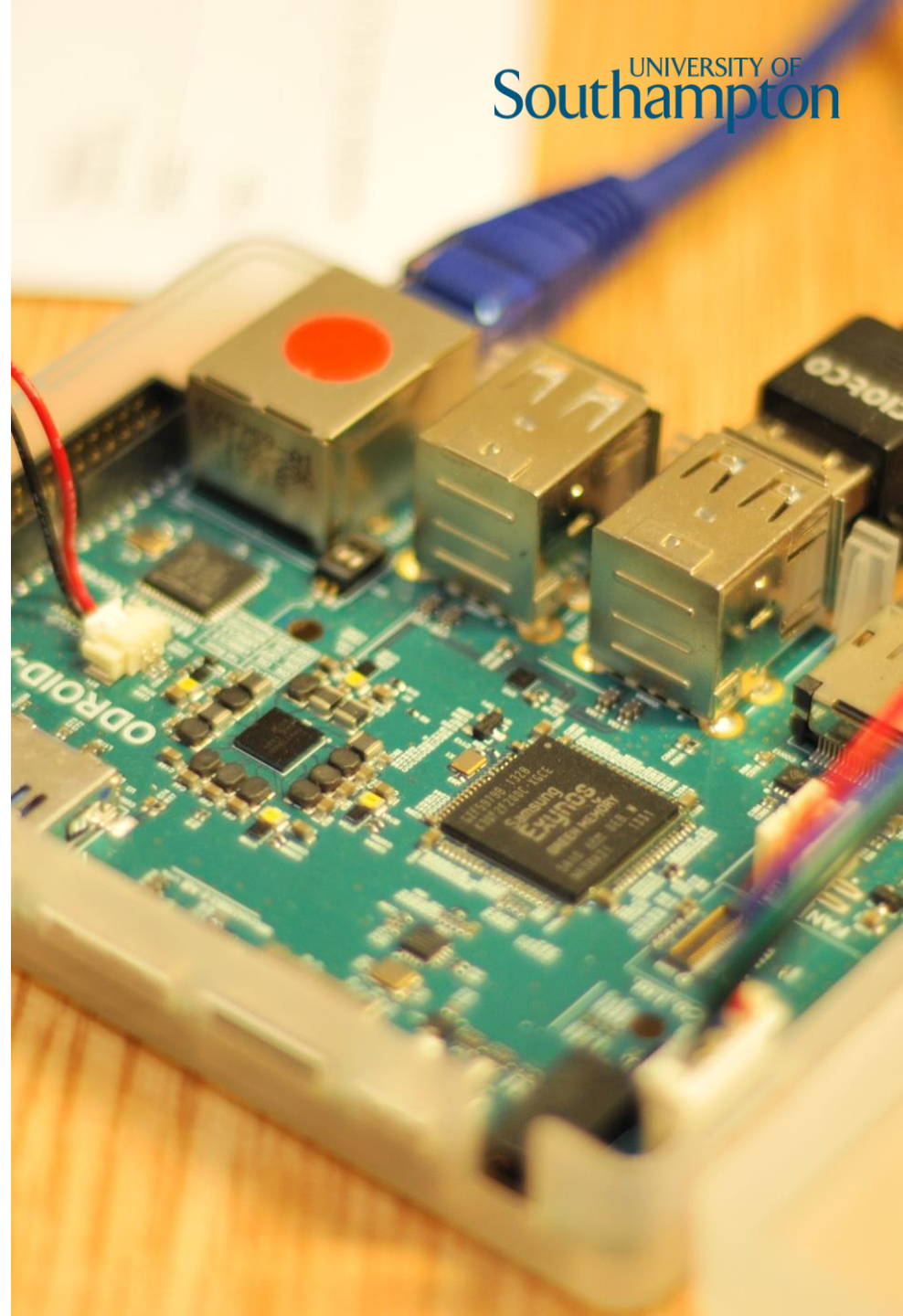
Power Estimation on Hardware

- Our Accurate and Robust Approach
- Open Source Tools

Power Estimation in gem5

- PMCs vs gem5 Statistics
- Power Estimation

Conclusions



WHY POWER ESTIMATION?

Run-Time Management Approaches

- Make energy-savings by controlling operation.
 - DVFS (dynamic-voltage frequency scaling) and DPM
 - Task scheduling and mapping
- Make decisions based on real-time power ‘measurements’



System Research

- Design-space exploration
- Evaluating new power management strategies
- Research power-optimized software (microcode to applications)
- SOC architecture & design balancing for power and performance

POWER MODELLING APPROACHES

“Bottom–Up” Power Models

- Take a design specification (e.g. pipeline stages, ROB size etc.)
- Simulate gates and toggle rates
- Uses statistics from an architectural simulator (e.g. gem5)
- **Advantages:** flexibility to specify any design; cache size, etc.
- **Disadvantages:** large errors, slow, limited validation

“Top–Down” Power Models

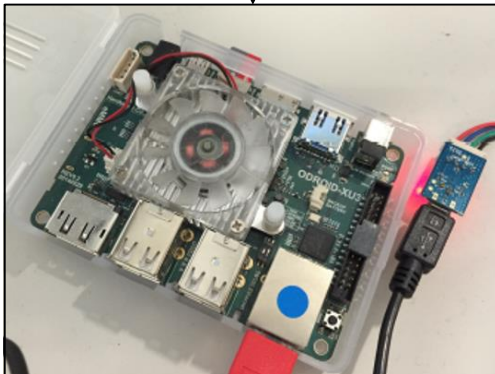
- Characterise a specific device
- Estimate relationship between measured power and stats, e.g. PMCs
- **Advantages:** accurate and lightweight
- **Disadvantages:** specific to the device they were built on

POWMON METHODOLOGY

1. Run workloads

@ different DVFS levels

39 workloads used: MiBench, LMBench, Roy Longbottom, ParMiBench and ALPBench



ODROID-XU3

Exynos-5422
4x Cortex-A7
4x Cortex-A5

5. Validate

- K-fold cross validation
- $R^2 : > 0.99$
- Error: 2.8 – 3.7%



4. Build Model

- OLS multiple regression
- Considers collinearity and heteroscedasticity
- “sensible” equation

2. Record

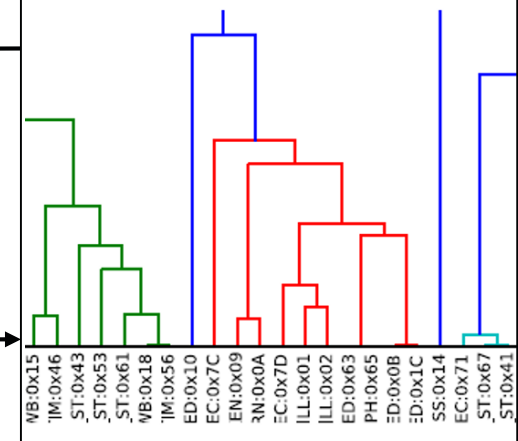
- PMCs
- Power, Voltage, Temperature, etc.

6. Uses

- OS Run-time management
- Reference for research
- gem5 add-on

3. Choose PMCs

Hierarchical cluster analysis,
Correlation matrix analysis,
Exhaustive search, etc.



THE POWMON APPROACH

A power model's stability is more important than its average error

Unstable model

- Appears accurate, but performs poorly with diverse workloads

Stable model

- Remains accurate across a diverse range of workloads and scenarios
- Requires careful choice of inputs (PMCs) & observations (workloads)

Eg: choose 3 sensors and appropriate training data to estimate colour:

Input Colour Channel A:



Training Dataset A:



Unstable

Input Colour Channel B:



Training Dataset B:



Stable

PERFORMANCE MONITORING COUNTERS

CPU Registers that count architectural and microarchitectural events

- E.g. L2 cache miss, TLB access, integer instruction, etc.

Positives

- Available on several platforms (e.g. ARM, Intel, AMD); low overhead
- Many different available events (>70)...

Negatives

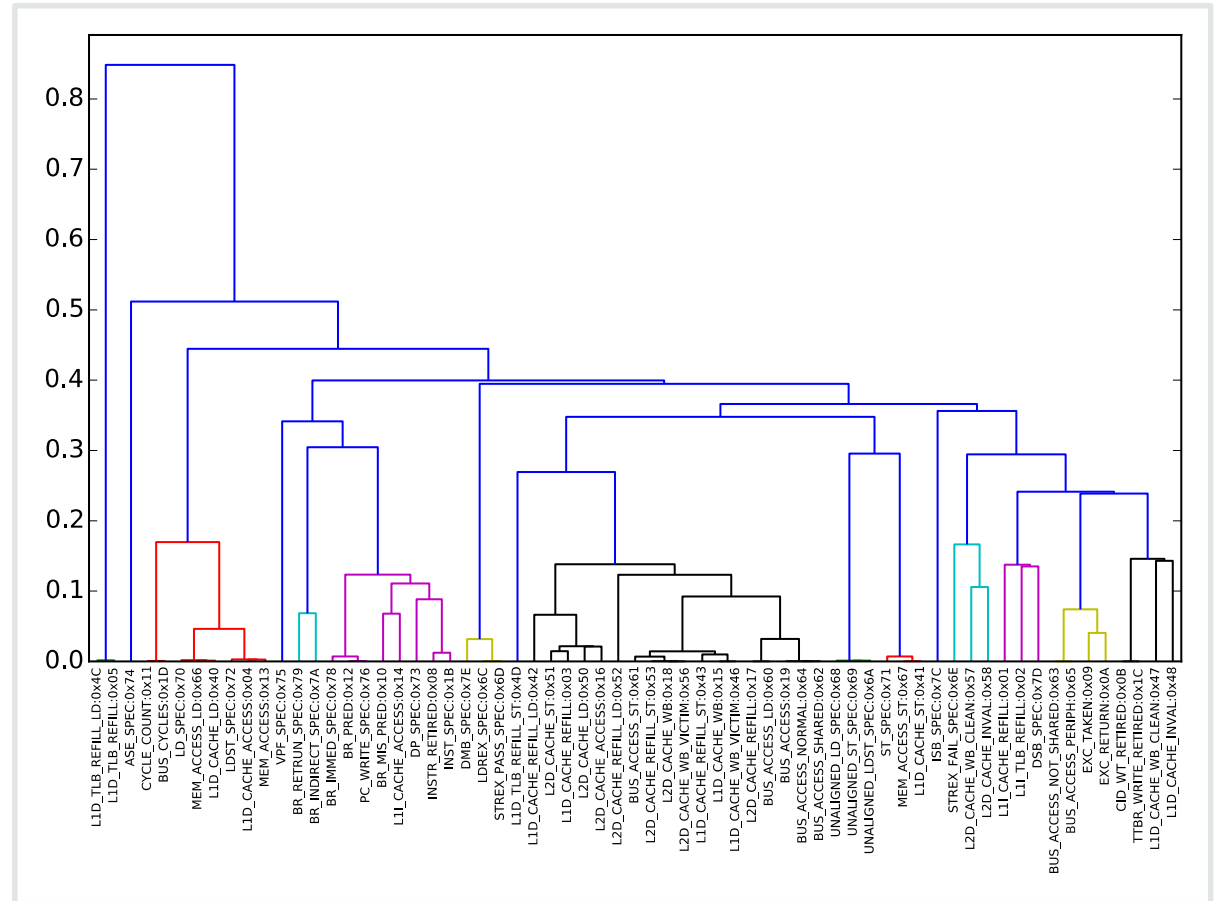
- ...but a small number (e.g. 4-6) can be monitored simultaneously

PMCs are often selected using intuition – e.g. try to split PMCs into different sub-architectural units. However can be problematic as:

- They may not gather enough information
- Different PMCs are correlated (can make a model unstable)

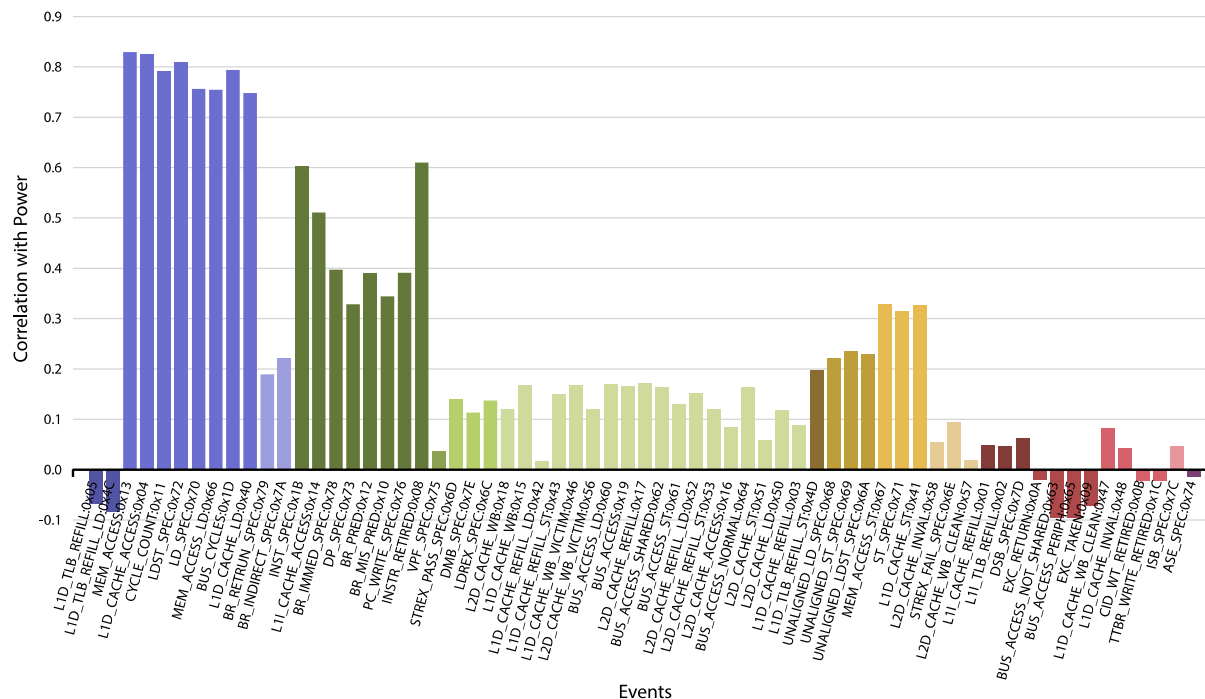
HIERARCHICAL CLUSTER ANALYSIS

- HCA groups similar events together
- Output is a dendrogram
- This allows PMCs to be grouped into clusters



HIERARCHICAL CLUSTER ANALYSIS

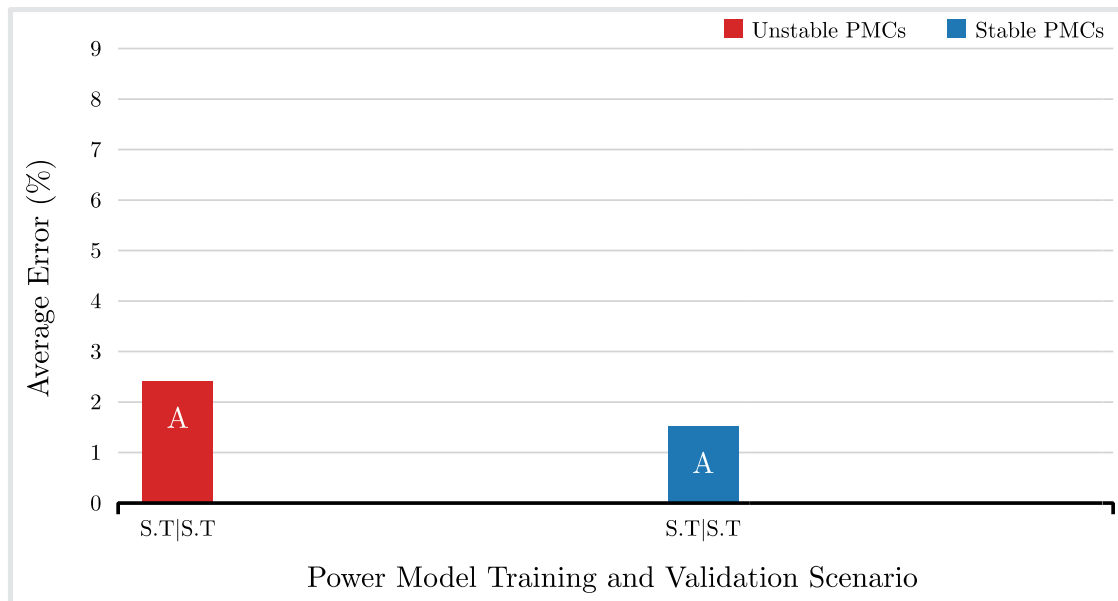
- We combine clusters with correlation of each event with CPU power



- Aim:** Choose PMCs with a high correlation with power, avoiding ones from the same cluster

STABLE vs UNSTABLE MODELS

1) Training and validating the model with a 'typical' set of workloads



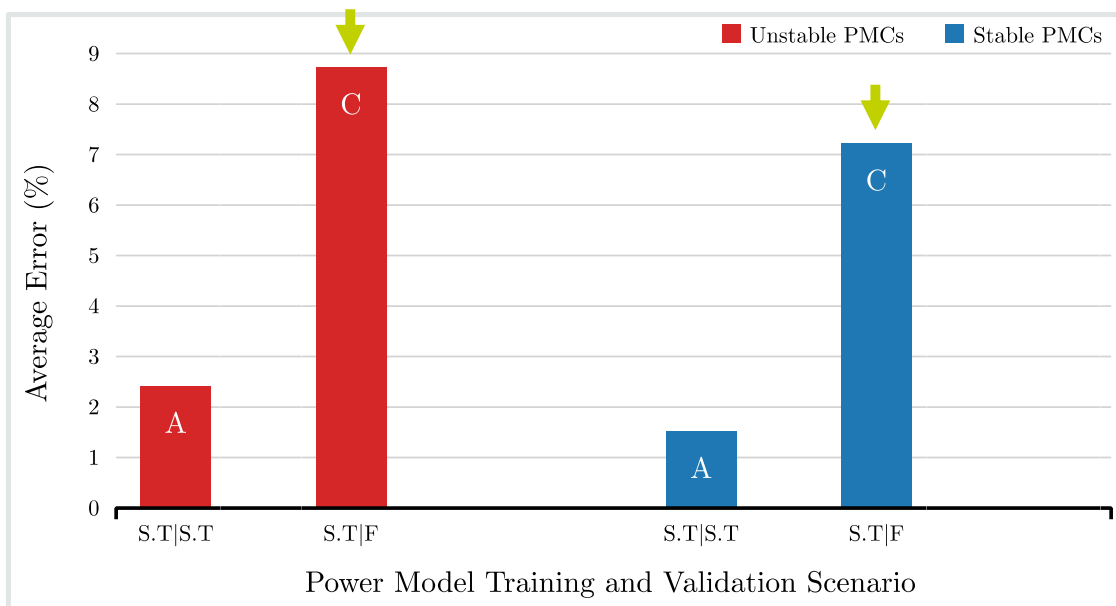
Training: Small set of 20 typical workloads (S.T), e.g. MiBench

Testing: Small set of 20 typical workloads (S.T), e.g. MiBench

Both unstable and stable model seem good (<2.5%)

STABLE vs UNSTABLE MODELS

2) Validating the same model with a 'full' set of workloads



Training: Small set of 20 typical workloads (S.T), e.g. MiBench

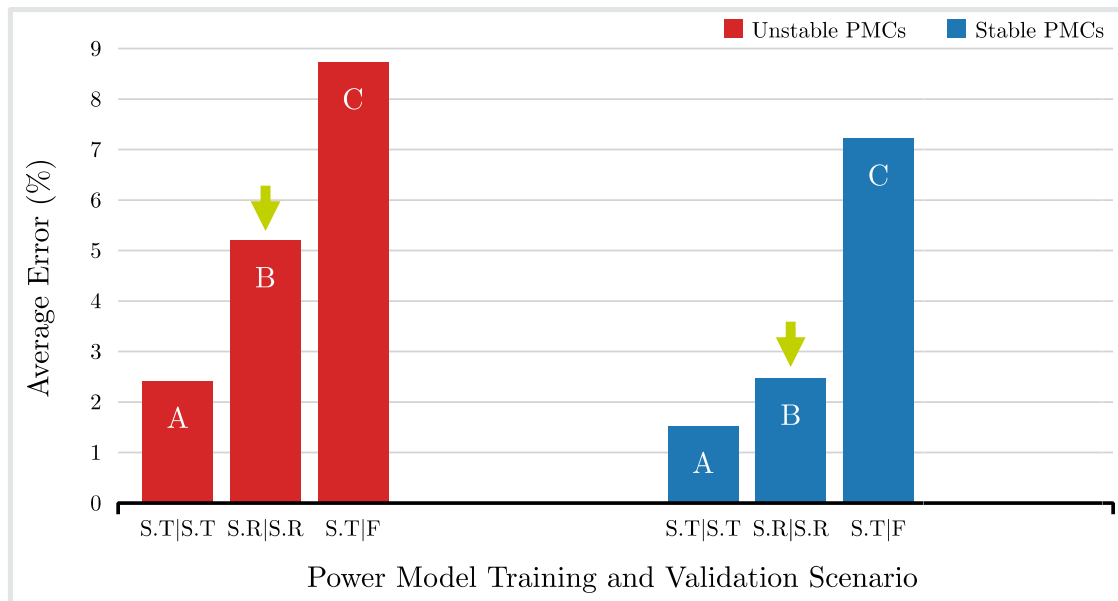
Testing: Full set of 60 diverse workloads (F)

Both models perform poorly, errors > 7%; not enough information from training workloads.

Testing with a small set of workloads results an optimistic reported error

STABLE vs UNSTABLE MODELS

3) Training and validating the model with a 'random' set of workloads



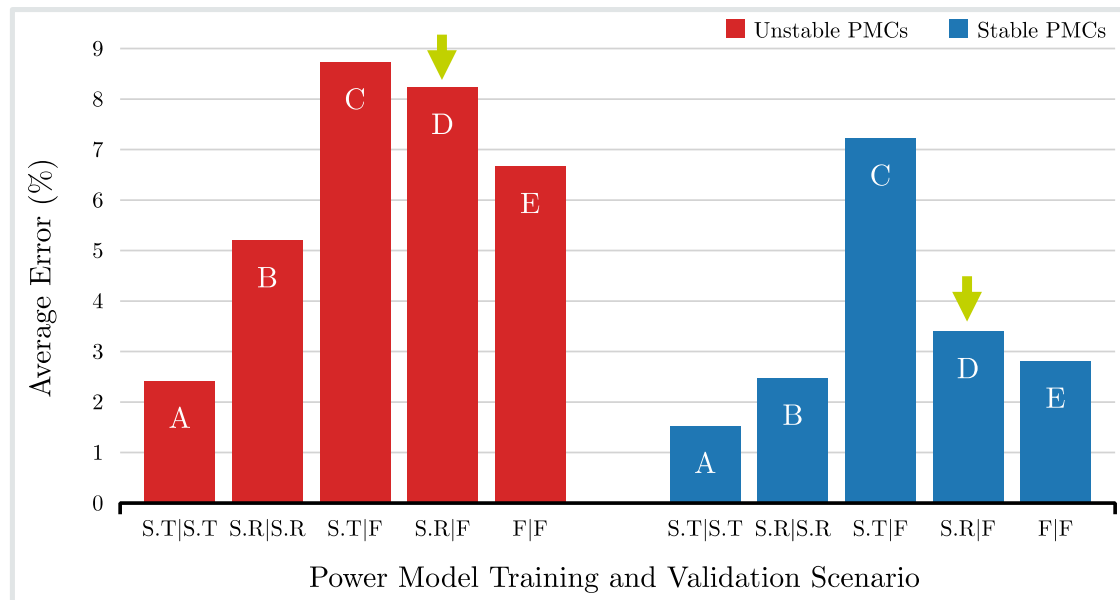
Training: Small set of 20 random workloads (S.R)

Testing: Small set of 20 random workloads (S.R)

Stable model copes better with workload diversity

STABLE vs UNSTABLE MODELS

4) Validating the same model with a 'full' set of workloads



Training: Small set of 20 random workloads (S.R)

Testing: Full set of 60 diverse workloads (F)

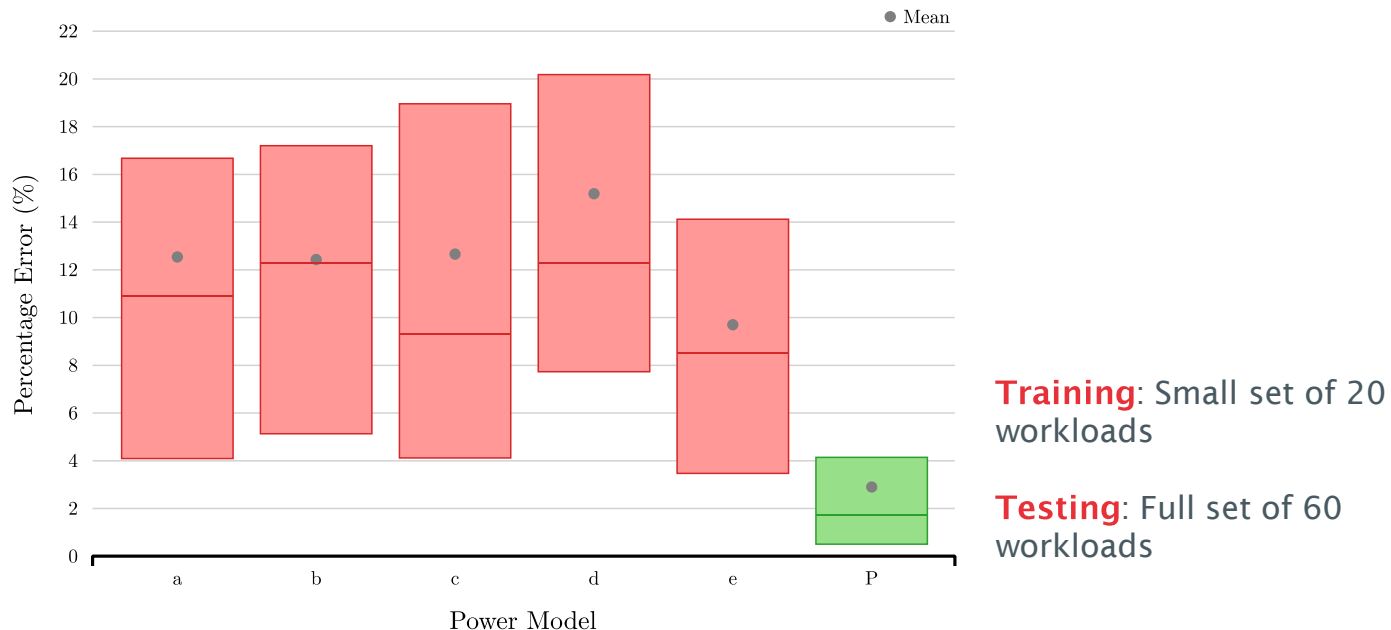
Accuracy of stable model close to full training set (E); unstable model poor

Diverse (random) training allows a stable model to gain prediction power

Stable models perform well even with limited training workloads

STABLE vs UNSTABLE MODELS

Our stable approach achieves a low average error and narrow error distribution compared to existing techniques. Models trained with 20 workloads, validated with 60.



[a] M. Pricopi, T. S. Muthukaruppan, V. Venkataramani, T. Mitra, and S. Vishin, "Power-performance modeling on asymmetric multi-cores," CASES '13.

[b] M. Walker et al., "Run-time power estimation for mobile and embedded asymmetric multi-core cpus," HIPEAC Workshop Energy Efficiency with Hetero. Comp. 2015

[c] S. K. Rethinagiri et al., "System-level power estimation tool for embedded processor based platforms," RAPIDO '14. New York, 2014.

[d], [e] R. Rodrigues et al, "A study on the use of performance counters to estimate power in microprocessors," IEEE TCAS II, vol. 60, no. 12, pp. 882-886, Dec 2013.

M. J. Walker et al., "Accurate and Stable Run-Time Power Modeling for Mobile and Embedded CPUs," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 36, no. 1, pp. 106-119, Jan. 2017.

ROBUST MODEL FORMULATION

Typical regression-based power model formulation

$$P = \text{const} + \beta_0 \text{Frequency} + \beta_1 \text{Voltage} + \beta_2 E_0 + \beta_3 E_1 + \beta_4 E_2 + \dots$$

- Relationships between power and other variables is not captured
- Too many independent variables -> instability

Our robust model formulation

$$P_{cluster} = \underbrace{\left(\sum_{n=0}^{N-1} \beta_n E_n V_{DD}^2 f_{clk} \right)}_{\text{dynamic activity}} + \underbrace{f(V_{DD}, f_{clk})}_{\text{static and BG dynamic}}$$

ROBUST MODEL FORMULATION – WHY?

Reduces the experiment time

- frequencies * core utilisations * workloads * average workload time
- By splitting model into static and dynamic, all workloads can be run at a single frequency, with just one (i.e. sleep) at all frequencies

	Avg. Error (%)	Experiment Time (hours)	Workloads
Slow	2.8	40	60
Fast	3.4	0.42 (25 min.)	30

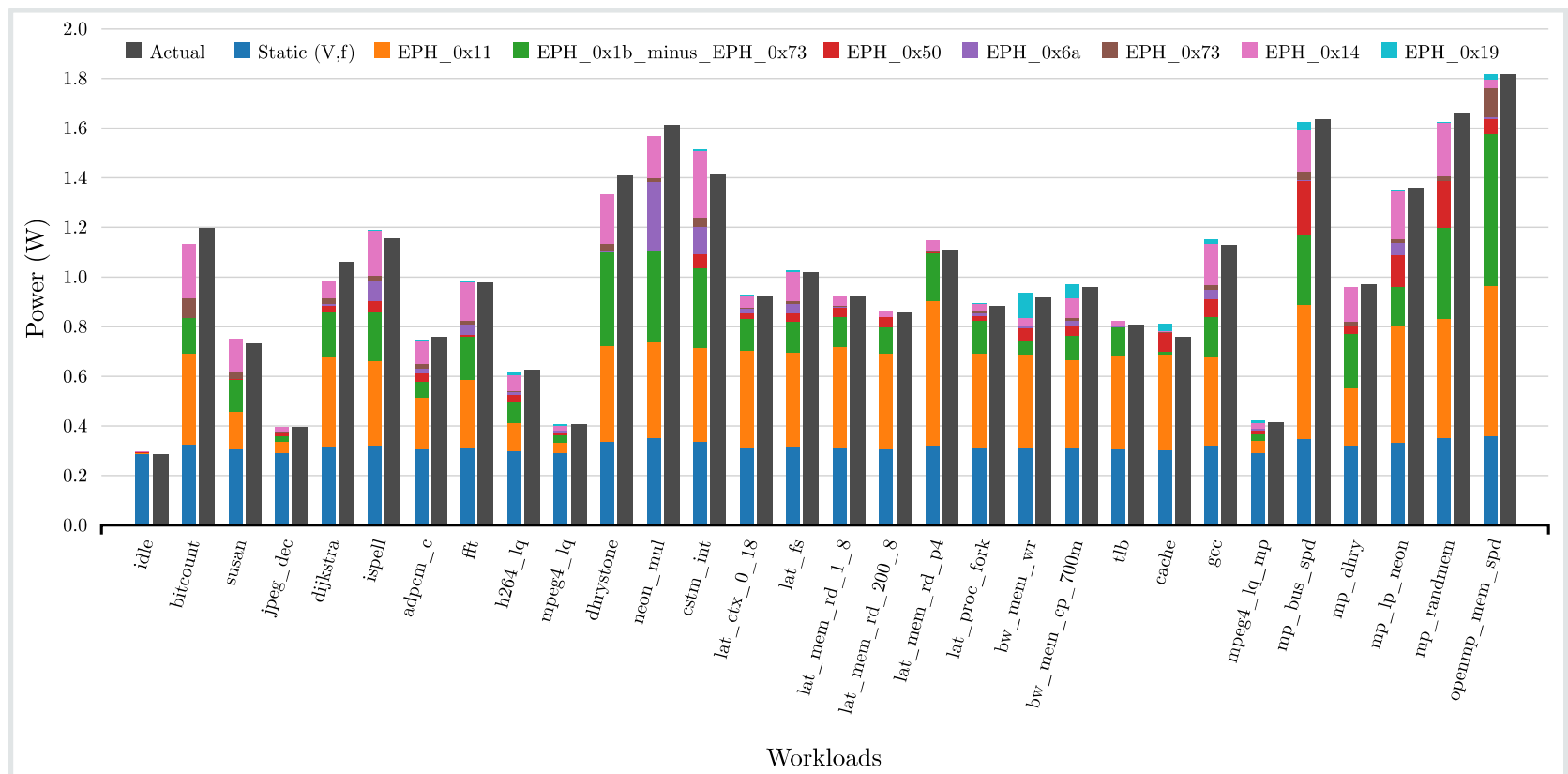
$$P_{cluster} = \underbrace{\left(\sum_{n=0}^{N-1} \beta_n E_n V_{DD}^2 f_{clk} \right)}_{\text{dynamic activity}} + \underbrace{f(V_{DD}, f_{clk})}_{\text{static and BG dynamic}}$$

Allows combination with ‘bottom-up’ approaches

- Once power has been divided into components, can apply theory to different parts.

Indicates where power may be being consumed

ROBUST MODEL FORMULATION – WHY?



Predicted power and modelled power for 30 different workloads

AVAILABLE TOOLS

[Home](#)[Documentation](#)[Downloads](#)[Results Viewers](#)

Run-Time CPU Power Modelling

Being able to accurately estimate CPU power consumption is a key requirement for both controlling online CPU energy-saving techniques and design-space exploration. Models built and validated using measured data from an actual device are extremely valuable as their accuracy is known and trusted.

This website makes available software tools for implementing our automated model building methodology which produces models that are both accurate and *stable*. We also provide power models for mobile CPUs (quad-core Cortex-A7 and quad-core Cortex-A15) which can be used directly in situations where an accurate reference model is required. Obtaining accurate data from mobile devices can be challenging and more time-consuming than using a simulator or desktop/server devices. For this reason, we make available our experimental platform software tools which allows workloads to be automatically run on a mobile device and Performance Monitoring Counters (PMCs), temperature, CPU utilisation, CPU power and CPU voltage to be collected. Details of our methodology can be found in the following publications:

- M. J. Walker; S. Diestelhorst; A. Hansson; A. K. Das; S. Yang; B. M. Al-Hashimi; G. V. Merrett, "[Accurate and Stable Run-Time Power Modeling for Mobile and Embedded CPUs](#)," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. PP, no. 99, pp. 1-1, doi: 10.1109/TCAD.2016.2562920
- Walker, Matthew J., Diestelhorst, Stephan, Hansson, Andreas, Balsamo, Domenico, Merrett, Geoff V. and Al-Hashimi, Bashir M., "[Thermally-aware composite run-time CPU power models](#)," In, International Workshop on Power And Timing Modeling, Optimization and Simulation (PATMOS 2016), Bremen, DE, 21 - 23 Sep 2016

Part of this work, [focussing on thermal compensation and model decomposition](#), will be presented at [PATMOS 2016](#), on Wednesday 21 September, Bremen, Germany.

This work has previously been presented at:

- [ISPASS 2016: Building Online CPU Power Models from Real Data, April 2016](#)
- [DATE 2016: RT-POWMODS Run-Time CPU Power Models from Real Data, March 2016](#)
- [MICRO-48: Building Online Power Models from Real Data, December 2015](#)

University of Southampton, UK | ARM Research

For questions or to provide feedback about the methodology, software tools or this website, email Matthew Walker (mw9g09@ecs.soton.ac.uk)

AVAILABLE TOOLS

Downloads

Here we make available our software tools. The license for our software is found in the [LICENSE.txt](#) under the top directory. If you use our software in your research, please acknowledge us by citing our paper:
M. J. Walker; S. Diestelhorst; A. Hansson; A. K. Das; S. Yang; B. M. Al-Hashimi; G. V. Merrett, "Accurate and Stable Run-Time Power Modeling for Mobile and Embedded CPUs," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems , vol.PP, no.99, pp.1-1, doi: 10.1109/TCAD.2016.2562920

1 Experimental Platform

1.1 Disk Image

This is the modified Ubuntu-14.04.1 image for the ODROID-XU3 and ODROID-XU4. Last updated: 7th Feb 2016. File size is 4.2 GB. Requires an SD card of 16 GB or larger.
[ubuntu-14.04.1-lts-lubuntu-odroid-xu3-20150212-power-model-20160207-r.img.xz](#)

Checksum:
[ubuntu-14.04.1-lts-lubuntu-odroid-xu3-20150212-power-model-20160207-r.img.xz.md5sum](#)

1.2 Experimental Platform Software

Below is the link to the latest experimental platform software (for the ODROID-XU3):
[experimental-platform-software-2016-02-09.tar.gz](#)

There is an also an ODROID-XU4 version (records PMCs and utilisation only):
[experimental-platform-software-xu4-2016-03-04.tar.gz](#)

2 Power Modelling Software

Below is the download link for the model building and validation software:
[power-model-analysis-2016-02-09.tar.gz](#)

3 Data and Results

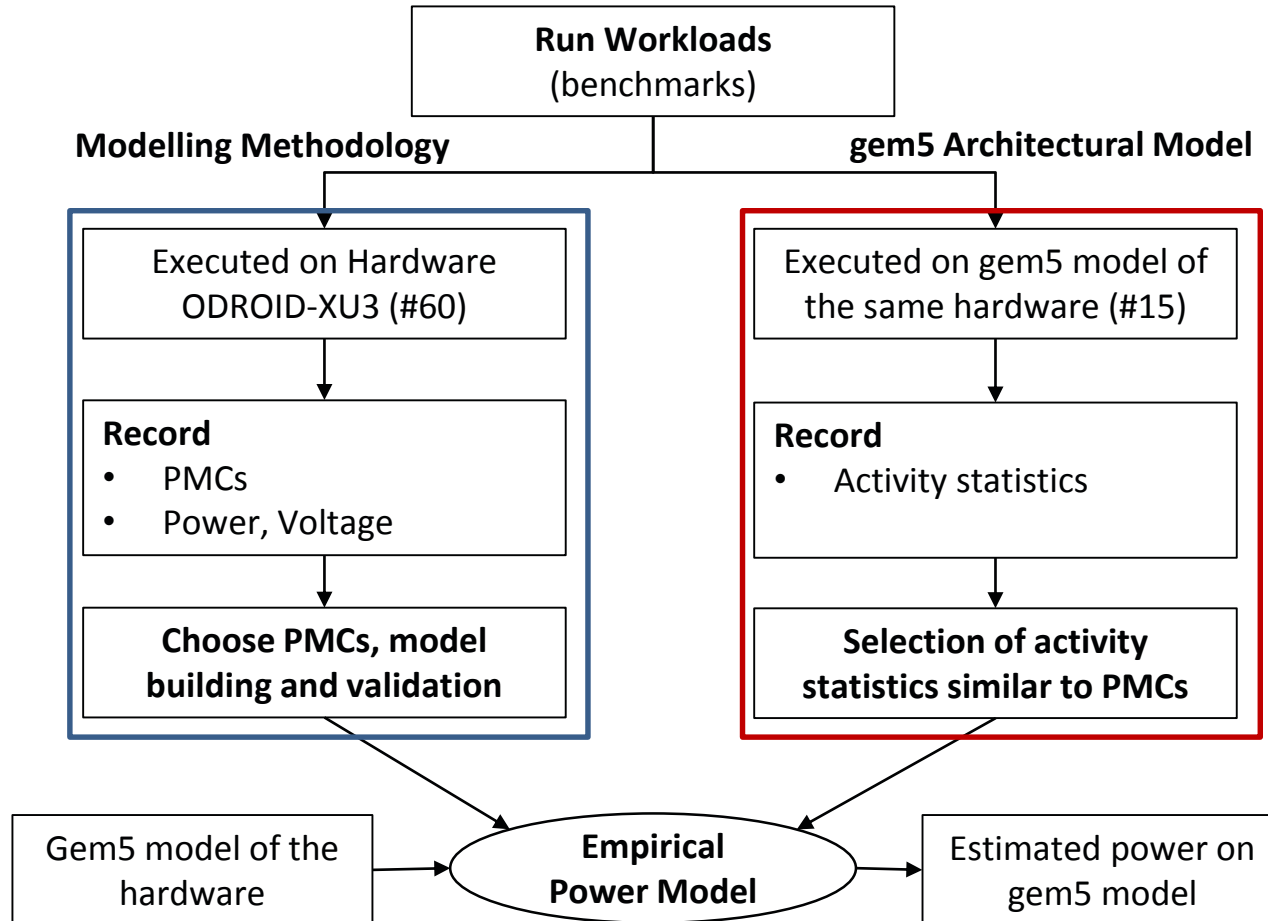
Accurate and Stable Run-Time Power Modeling for Mobile and Embedded CPUs ([link to publication](#)):

- Results can be viewed in the [Results Viewer section](#)
- [Table of raw data and results \(in tab-separated CSV format\)](#)

Thermally-aware composite run-time CPU power models ([link to publication](#)):

- [Table of raw data and results \(in tab-separated CSV format\)](#)

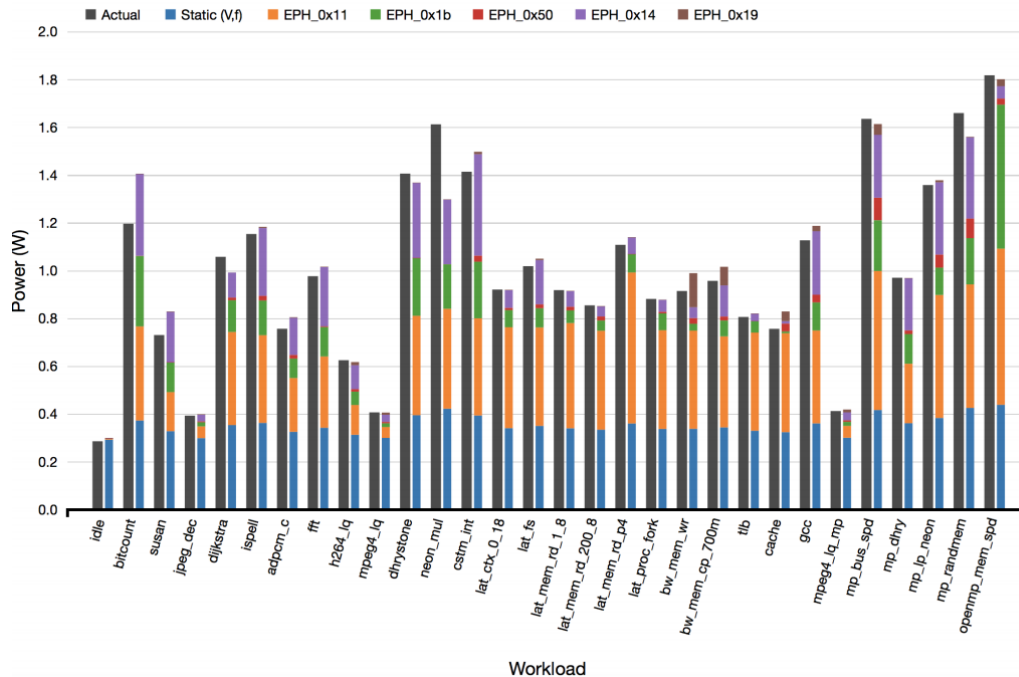
gem5 POWER ESTIMATION



PMC SELECTION

- Our Cortex-A15 power model uses the following seven PMCs:
 - **0x11 CYCLE COUNT**: active CPU cycles
 - **0x1B INST SPEC**: instructions speculatively executed
 - **0x50 L2D CACHE LD**: level 2 data cache accesses - read
 - **0x6A UNALIGNED LDST SPEC**: unaligned accesses
 - **0x73 DP SPEC**: instructions speculatively executed, int data processing
 - **0x14 L1I CACHE ACCESS**: level 1 instruction cache accesses
 - **0x19 BUS ACCESS**: bus accesses
- Suitable gem5 event counts for PMC events 0x6A and 0x73 were not available; the model was rebuilt without these

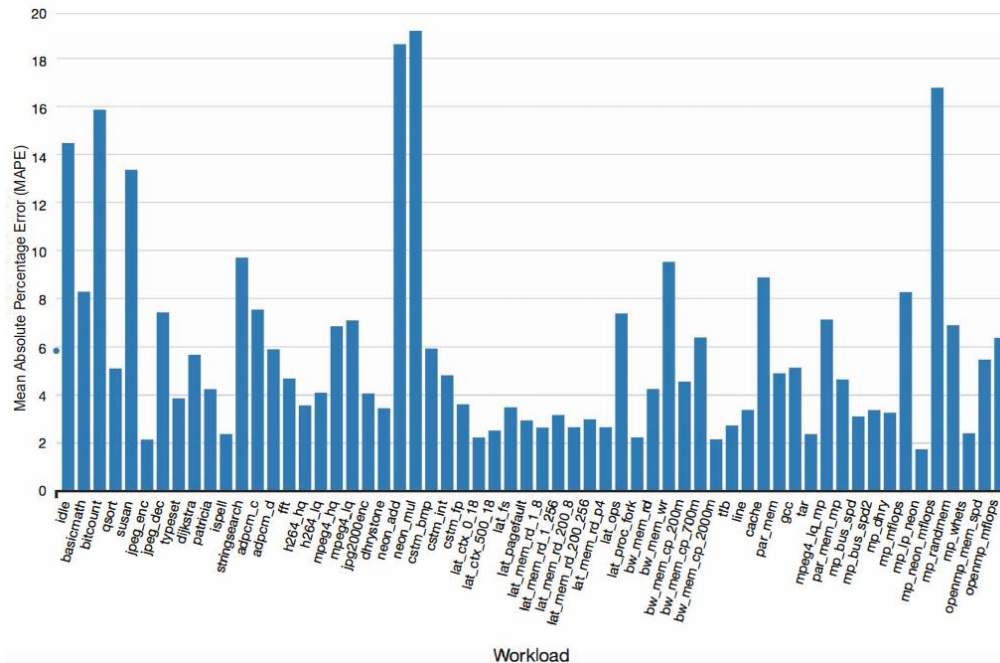
MODEL VALIDATION (vs HARDWARE)



$$P_{cluster} = \underbrace{\left(\sum_{n=0}^{N-1} \beta_n E_n V_{DD}^2 f_{clk} \right)}_{\text{dynamic activity}} + \underbrace{f(V_{DD}, f_{clk})}_{\text{static and BG dynamic}}$$

Comp.	Coefficient	Weight	p-value
Dyn. act.	$0x11 \times V^2 f$	6.198e-10	$p < 0.0001$
Dyn. act.	$0x1b \times V^2 f$	2.685e-10	$p < 0.0001$
Dyn. act.	$0x50 \times V^2 f$	3.528e-9	$p < 0.0001$
Dyn. act.	$0x14 \times V^2 f$	1.722e-9	$p < 0.0001$
Dyn. act.	$0x19 \times V^2 f$	3.553e-9	$p < 0.0001$
Static	Intercept	-1.403e+3	$p < 0.0001$
Static & B.G. Dynamic	f	2.748e-1	$p < 0.0001$
Static & B.G. Dynamic	V	4.713e+3	$p < 0.0001$
Static & B.G. Dynamic	$V f$	-1.114e+0	$p < 0.0001$
Static & B.G. Dynamic	V^2	-5.262e+3	$p < 0.0001$
Static & B.G. Dynamic	$V^2 f$	1.436e+0	$p < 0.0001$
Static & B.G. Dynamic	V^3	1.953e+3	$p < 0.0001$
Static & B.G. Dynamic	$V^3 f$	-5.979e-1	$p < 0.0001$

MODEL VALIDATION (vs HARDWARE)



ARCHITECTURAL MODEL

- A detailed OoO model of the 4-core Cortex-A15 in FS mode
- Instruction timing in execution stage configured as per (Endo et al., 2015).
- Integer instructions have latencies of 1 (ALU), 2 (x) and 12 (\div), and default latencies for FP instructions.
- Integer and floating point stages are pipelined.
- Cortex-A15 has two levels of TLB rather than one. To compensate, the ITLB and DTLB are over-dimensioned.

Parameter		Specification
Core type		Cortex-A15 (out-of-order)
Number of Cores		4
CPU clock (MHz)		200, 600, 1000, & 1600
DRAM (LPDDR3)	Size	2048 MB
	Clock	933 MHz
L2-Cache	Size	2 MB
	Associativity	16
	Latency	8 cycles
	MSHRs	11
	Write buffers	16
L1-I Cache	Size	32 kB
	Associativity	2
	Latency	1 cycle
	MSHRs	2
L1-D Cache	Size	32 kB
	Associativity	2
	Latency	1 cycle
	Write buffers	16
	MSHRs	6
ITLB/DTLB		128 each
ROB entries		128
Branch predictor type		Bi-Mode
BTB entries		4096
RAS entries		48
ROB entries		128
IQ entries		48
Front-end width		3
Back-end width		8
LSQ entries		16

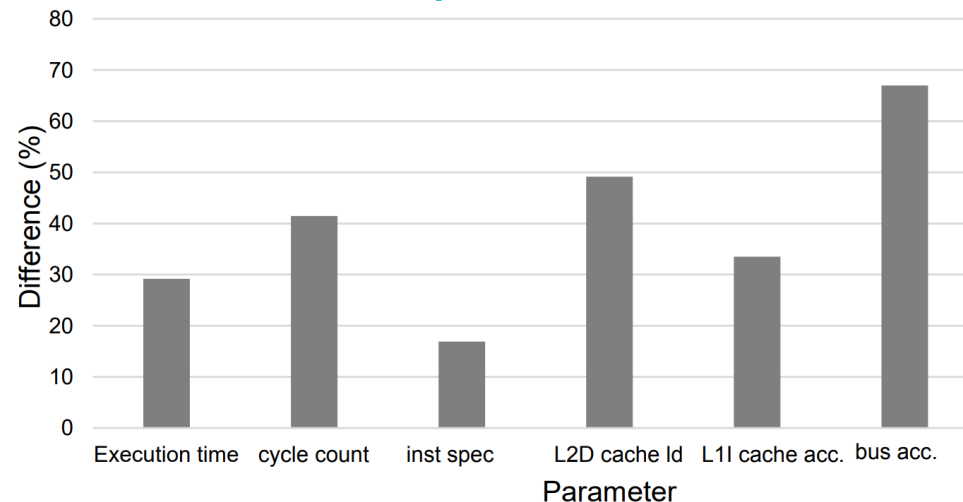
gem5 EVENTS VS HARDWARE PMCs

- 15 MiBench workloads
- 4 frequencies:
 - 200 MHz
 - 600 MHz
 - 1000 MHz
 - 1600 MHz

Hardware Event	gem5 Event
0x11 CYCLE COUNT	system.cpu.numCycles
0x1B INST SPEC	system.cpu.iew.iewExecutedInsts
0x50 L2D CACHE LD	system.l2.overall_accesses::total
0x14 L1I CACHE ACCESS	system.cpu.icache.overall_accesses::total
0x19 BUS ACCESS	system.mem_ctrls.num_writes::total + system.mem_ctrls.num_reads::total



Hardware vs gem5: execution time and PMCs/activity statistics (f=1 GHz)



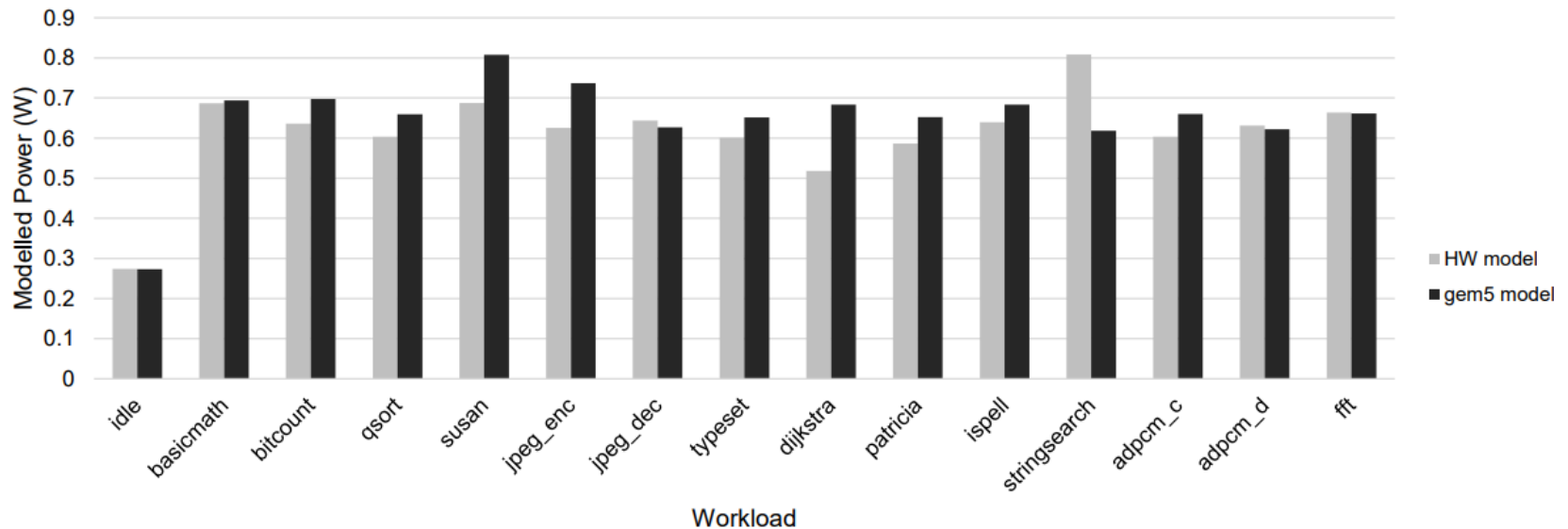
gem5 EVENTS VS HARDWARE PMCs

This difference is likely due to factors including:

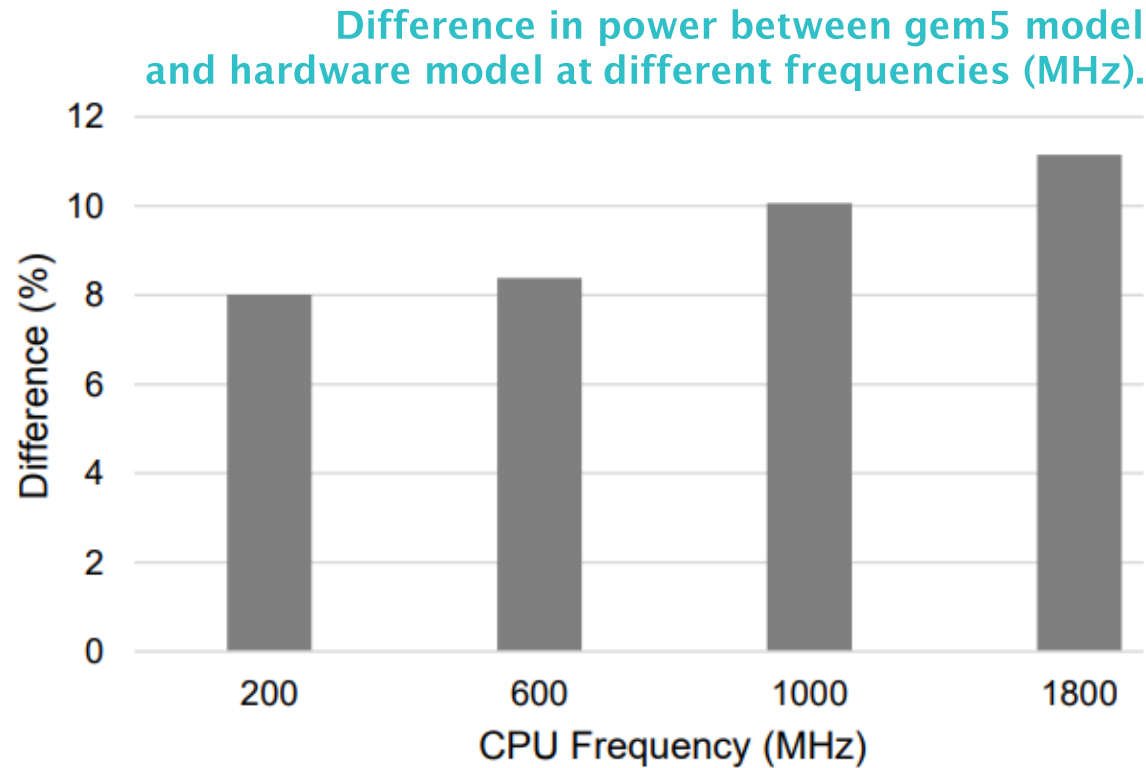
- Specification error in the simulator:
 - in the fetch stage contributes to the I-cache miss error.
 - in the TLB models contributes to the reported error in execution time and activity statistics.
- LPDDR3 DRAM in gem5 corresponds to 800 MHz, vs 933 MHz in the hardware.

MODEL VALIDATION (gem5 vs HARDWARE)

Modelled power for each workload @1000 MHz.



MODEL VALIDATION (gem5 vs HARDWARE)



CONCLUSIONS

Robust and Stable Power Modelling

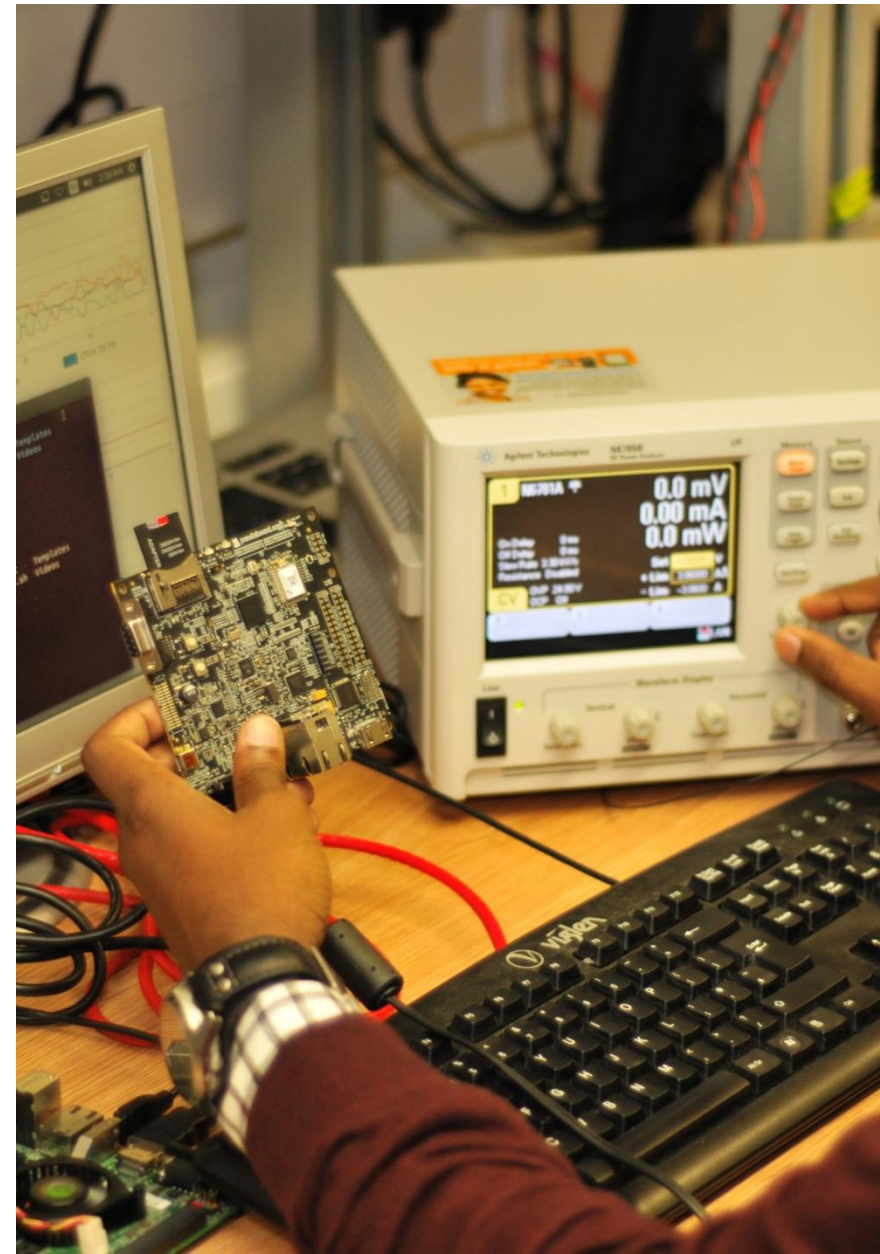
- Appropriate workload selection
- Stable PMC selection
- Robust model formulation

Applying Models to gem5

- Real hardware vs modelled architecture
- PMCs vs gem5 event stats/exec. time
- 10% error in gem5 vs hardware model

Tools Available!

- www.powmon.ecs.soton.ac.uk



ACKNOWLEDGEMENTS



Matthew Walker
Uni. Southampton (PhD)



Prof Bashir Al-Hashimi
Uni. Southampton



Dr Domenico Balsamo
Uni. Southampton (Postdoc)



Stephan Diestelhorst
Arm Research



Karunakar Basireddy
Uni. Southampton (PhD)



Andreas Hansson
(previously) Arm Research



Any Questions?

UNIVERSITY OF
Southampton

Dr Geoff V Merrett
Associate Professor

Electronics and Computer Science

Tel: +44 (0)23 8059 2775

Email: gvm@ecs.soton.ac.uk | www.geoffmerrett.co.uk

Highfield Campus, Southampton, SO17 1BJ UK