Using hardware counter data to model performance and energy usage in NUMA systems

Doctoral Meeting February 2019

Miquel López Becoña

Centro Singular de Investigación en Tecnoloxías da Información





Centro Singular de Investigación

citius uscles

L Context

Motivation and goals Hardware Counters Berkeley Roofline Model Intel RAPL

- Performance analysis and optimization tool
 Introduction
 Optimization strategies
- 3 Energy usage modelling Relation with RM Other events
 - Conclusions and future work



Context Motivation and goals Hardware Counters Berkeley Roofline Model Intel RAPL

2 Performance analysis and optimization tool Introduction Optimization strategies

- 3 Energy usage modelling Relation with RM Other events
- 4 Conclusions and future work



Motivation and goals

- Memory gap: data locality is a key matter in performance.
 - Specially in NUMA systems.
- Power consumption issues.
- Goals: runtime optimization, performance and energy usage modelling.
- How? Development of a tool that:
 - Reads and gathers hardware counter information about the activity of each thread.
 - Uses of Roofline Model as the basis to characterize the performance.
 - Performs migrations to improve locality.



Hardware Counters

- Microprocessor specific registers.
- Intel PEBS: sampling.
- Low overhead and high accuracy.
- Issues while measuring FLOPS: replaced by instructions. More general.
- perf_events Linux interface to access and extend its information.
- Each sample dumps the core state.
 - Generic fields: timestamp, CPU ID, PID, TID ...
 - Hardware events: what we actually measure.



	PEBS record format
Offset	64bit/8 bytes Field
0x0	R/EFLAGS
0x8	E/REIP
0x10	R/EAX
0x18	E/EBX
0x20	R/ECX
0x28	R/EDX
0x30	R/ESI
0x38	R/EDI
0x40	W/EBP
0x48	R/ESP
0x50	R8
0x58	R9
0x60	R10
0x68	R11
0x70	R12
0x78	R13
0x80	R14
0x88	R15
0x90	IA32_PERF_GLOBAL_STATUS
0x98	Data linear Address
0xA0	Data Source encoding
0xA8	Latency (core cycles)

Berkeley Roofline Model

- Performance model.
- 2D plot: simple and easy to understand.
- Helps finding bottlenecks.
 - Clues to optimize our applications.
- X axis: operational intensity.

FLOPS bytes_read_from_memory

- Y axis: raw performance (GFLOPS/sec).
- Roof/s: hardware limits.
- Basic classification: side of the graphic.
- Extensions: Dynamic Roofline Model (DyRM), and 3DyRM (useful for NUMA)²



²Óscar García Lorenzo et al. "3DyRM: A dynamic roofline model including memory latency information" CinUS ³Samuel Williams et al. "Roofline: an insightful visual performance model for multicore architectures"

Context

Example of generated Roofline Model





4/17

Which hardware events do we currently use?

- Mainly those which are associated to the Roofline Model.
- Instruction count (performance related).
- Offcore requests: cache lines read from memory (performance related).
- Memory access latency (data locality related).
- Energy usage per node (load balancing related), with the aid of Intel RAPL.



Intel RAPL (Running Average Power Limit)

- Software interface to estimate energy usage.
- Divides consumption between logical domains:

cores

- pkg: cores + LLC + memory controller + ...
- ▷ ram
- Variable support.
- Set of buffers for each NUMA node.



Example energy usage plot





Time (s)



1 Context

Motivation and goals Hardware Counters Berkeley Roofline Model Intel RAPL

Performance analysis and optimization tool
 Introduction
 Optimization strategies

3 Energy usage modelling Relation with RM Other events

4 Conclusions and future work



Performance analysis and optimization tool

Main goal

To characterize the performance of parallel programs and performs thread and page migrations accordingly using hardware counters information.

- **Two main modes:** JUST_PROFILE and DO_MIGRATIONS.
 - CSV dumping for a posteriori analysis.
- Requires Linux OS and a Intel microprocessor.
- Not very intrusive: low overhead and easy to install and use.
- No root permissions required*.
- Automatic detection of system topology (relation of CPUs per NUMA node, etc.).
- Language-independent solution (code is not analysed).

Example of Dynamic Roofline Model plot

Additional R and Python code for analysing data and making neat plots.





Optimization strategies

- Decisions about which thread/page to migrate and where.
- Selection of a set of migrations (that might be empty) per iteration.
- Modular implementation of the strategies.
- Usually based on classic optimization strategies and search problems.
 - Simulated annealing, genetic algorithm, random, energy-balancing...
- Most of them are in an early stage of development.



1 Context

Motivation and goals Hardware Counters Berkeley Roofline Model Intel RAPL

- Performance analysis and optimization tool
 Introduction
 Optimization strategies
- 3 Energy usage modelling Relation with RM Other events

4 Conclusions and future work



Energy usage modelling with Roofline Model

- Let's model how energy usage is affected by performance metrics.
- RAPL domains!
- First approach: Roofline Model.
- pkg related to CPU activity (ginsts/s), ram related to memory operations (OI).
- First approach: energetic Roofline.



Local ops per second vs ram



Energetic Roofline



Operational Intensity (Insts/Byte traffic)



Relation with DyRM

Good results in some cases, but not in all of them...





Other hardware events

- Maybe we should take into account other/more metrics?
- Search for more correlations...

	Event-energy us													rgy usage correlations for benchm										nark bt																									
-g 0.002									0541	0.395		0.584		0.742	0.0556	0.00279			0 908	0.167								-0.K27	0.636	0.541	0.667	0.627	0.566		0.507	0.99	0.678	4.969			0.0371	0.981						0.552	0.425
0.045	0.787	0.89	0.662	0.925	0.853	0.808	0.425	0.428	0.642	0.373	0.444	0.699	0.362	0.748	0.16	0.471	0.855	0.855	0.643	0.916	0.813	0.717	0.734	-0.824	0.916	-0.524	0.924	0.245	0.240	0.395	0.241	0.57	0511	0.373	0.721	0.44	4 2 1 5	0.120	0.82	0.35	0.0153	0.351	0.341	0.963	0.953	0.909	0.924	0.653	0.413
B NR DICHTH CODUCH	ARTINE ON ACTUR	OVOLE, MCTINEY STRUES, LOUD, PENDING	באבוע אבאאונא באווער אין אבאנאר	THEOR STIMUS AURILLE VICES	כוכר? אכווועג גדוונג אווינא	DREMITE_SMITCHES/FEWUTY_CYCLES	NUM V SESINO SSINSESSIN OVOT ELLO	GT13. LOAD_MISSESSIMLE COMPLETED	30 JH STUSSION OPOT STUD	NOTE, LORD, MISSESWICK, DURATION	OT 5 LOAD WESTSWALK CONFLICTED AK	DILA, STORE, MESRES STLA, HT	TLB_STORE_MISSES MISS_CAUSES_A_MACK	2018_57046_M85855118_H1_MT_M	CTLA_STORE_WISSESSMUC_COMPLETED_AK	ULL STORE MERCINNER DURING	R. Jamescauel, Dorke, P.	PLARMAN PROVIDE	16900	THES' VERESSON	CLOR MISSS	5401 200 00	Solon' Fistrous' BIST Inv bo	STELC'SADIT BSD Da	ngaann	STUD'SADI'ILMOD	SOUL, THE D	arcool and savbo	5400 ⁻¹ 800 ⁻¹ 500 ⁻¹ 50	Secon 1100 Secon	serouo"saon"aasi'aakba	suon'iseba	sacure' seba	DOLUPIS-NOT DELABREDCORE	strad'saon'saoba	T_DELMERED CYCLES_0_UGHS_DELM_CORE	NOTE ANY DELIVERED CYCLES, PE, MAS, DK	DELMEREDICYCLER_LE_2_UOP_DELM_CORE	pervesebickcresine 1 nov pervicese	BRINGREDICHCIER JE. J. 109- DRIN, CORE	JOTTINS OF	INSTRUCTION, SETTRON	4, 100, 061963 ANY 9	NT JASC SAT STALL CYLLES	MI_MISCRECOMERY_CYCLES	NU MACKEDOVERN CHILES	M_JM8C98C0MBH_STMLLS_COUNT	1012 202	FLIL) INSIGES INSIS, CAUGHS, A., UNLOC
-0.8 -0.4										0.0										0.4 8						2,006	2,00%	0.8																					



Current stage: machine learning

- Join data for lots of events: too many columns in CSV.
- Dimensionality Reduction: PCA (Principal Component Analysis).
- Machine learning model with the most N relevant events.



1 Context

Motivation and goals Hardware Counters Berkeley Roofline Model Intel RAPL

- Performance analysis and optimization tool
 Introduction
 Optimization strategies
- 3 Energy usage modelling Relation with RM Other events
 - Conclusions and future work



Conclusions

Development of a new tool

- To characterize the performance of parallel applications in NUMA systems.
- To optimize memory accesses through page and thread migrations.
- To lower power consumption.

Energy usage research work

To explain it based on hardware counter information.



Future work

- Progressing in the energy study.
- Algorithm refinement for current strategies.
- Implement new strategies.



miguel.becona@usc.es





Cofinanciado pola Unión Europea

Fondo Europeo de Desenvolvemento Rexional "Unha maneira de facer Europa"