Analysis of the Influences on Server Power Consumption and Energy Efficiency for CPU-Intensive Workloads

Jóakim v. Kistowski, Hansfried Block, John Beckett, Klaus-Dieter Lange, Jeremy A. Arnold, Samuel Kounev

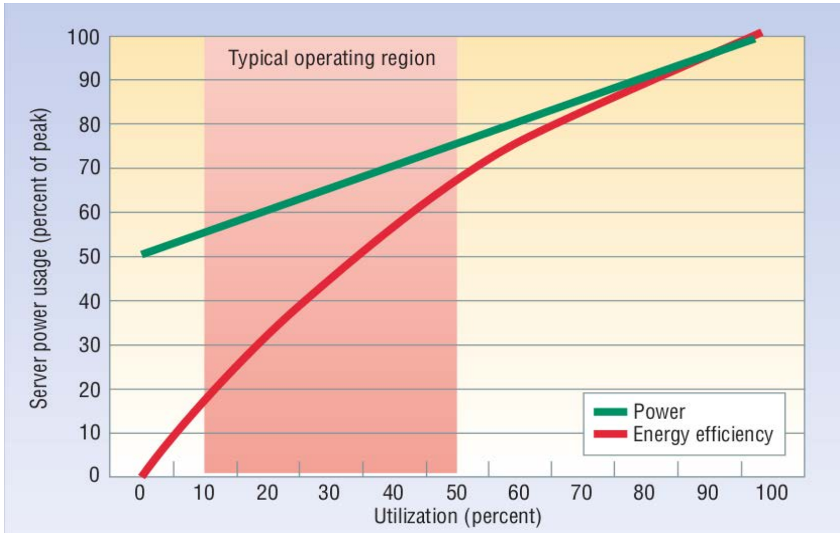
University of Würzburg, Fujitsu, Dell, HP, IBM
SPECpower Committee, SPEC

ICPE, February 3rd 2015, Austin, TX

Energy Consumption of Servers

A typical server …

- has an average utilization between 10% and 50%,
- is provisioned with additional capacity (to deal with load spikes).



Energy Efficiency and Power Consumption of Servers [1]

Power consumption depends on server utilization.

Energy Efficiency of Servers

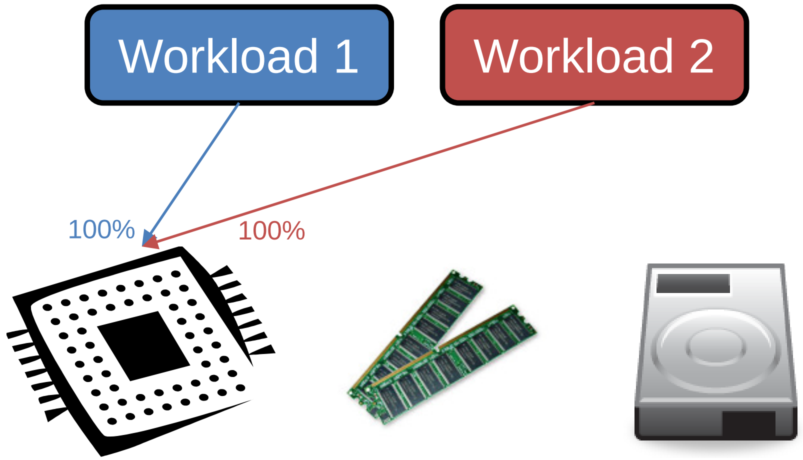
- Relationship of Performance and Power
- For transactional workloads:

$$\frac{transactions}{energy}\left[\frac{1}{J}\right] = \frac{throughput}{power}\left[\frac{1/s}{W}\right]$$

- Comparison of efficiency of different workload types is difficult
  - Different scales of transaction-counts / throughput
  - ➔ normalization

Common Power Models

- Black-box models
  - Simple
  - Fine granular models are workload-dependent [2]
- Decomposition into used hardware components [3,4]



What about different workloads targeting the same component?

Contributions

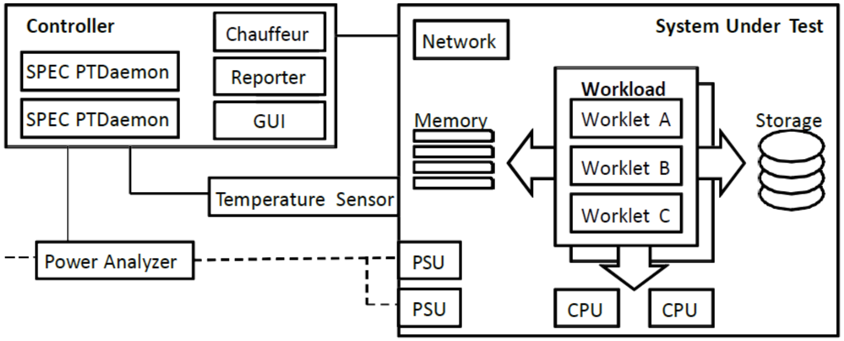
- Measure power consumption and performance for SERT's 7 CPU worklets
- Explore change of power consumption and energy efficiency depending on load level
- Demonstrate that CPU-workloads can have significantly different power consumption at the same load level
- Explore impact of different hardware and software configurations on the power/load level functions

SPEC SERT

- Server Efficiency Rating Tool
- Tool for analysis and evaluation of energy efficiency of servers
- Provides focused transactional micro-workloads (called worklets)
  - Exercise selected SUT aspects at multiple load levels
- Tests SUT at multiple load levels
- Calibrates workload intensity for target SUT load levels

SERT Architecture

- Controller System runs
  - Chauffeur: Director
  - Reporter



- PTDaemon
  - Network-capable power and temperature measurement interface
  - Can run on controller system or separate machine

- System under Test (SUT) runs
  - SERT client, executes worklets

Load Levels

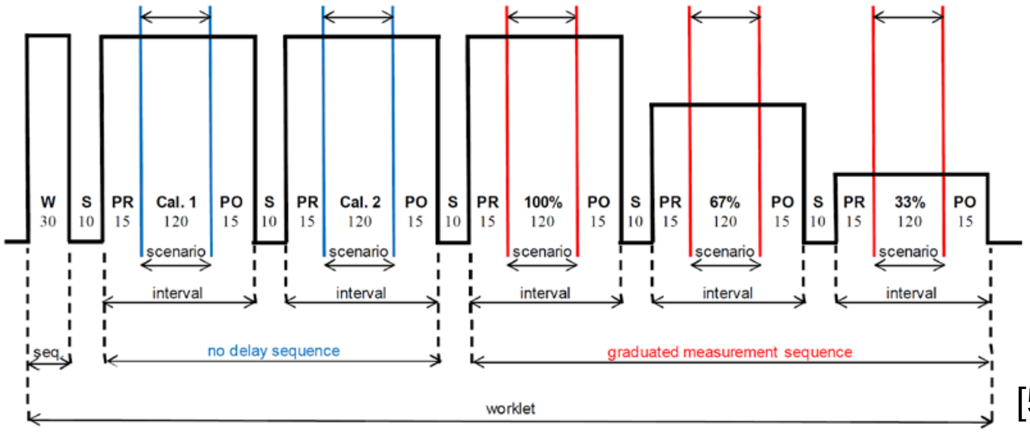
- Utilization = $\frac{t_{busy}}{t_{busy} + t_{idle}}$
- DVFS increases CPU busy time at low load
  - ➔ increases utilization
  - Power over load measurements need to compensate

    How to compare?

- SERT's solution: Machine utilization
  - 100% utilization at maximum throughput
  - Load level = $\frac{current\ throughput}{max.\ throughput}$

SERT Measurement

- Separate measurement intervals at stable states
  - 15 second pre-measurement run
  - 15 second post-measurement run
  - 120 second measurement



[5]

- Temperature analyzer for comparable ambient temperature

- Power Measurements: AC Wall Power

SERT CPU Worklets

- 7 CPU worklets:

Worklet	Description
Compress	Compresses and decompresses data
CryptoAES	Encryption and decryption
LU	Matrix factorization
SHA 256	Standard Java SHA-256 hashing and encryption/decryption
SOR	Jacobi Successive Over-Relaxation
SORT	Sorts a randomized 64-bit integer array
XMLValidate	Uses *javax.xml.validation*

- Definition CPU Worklet: 100% load level at 100% CPU utilization. CPU is the bottleneck.

Systems Under Test

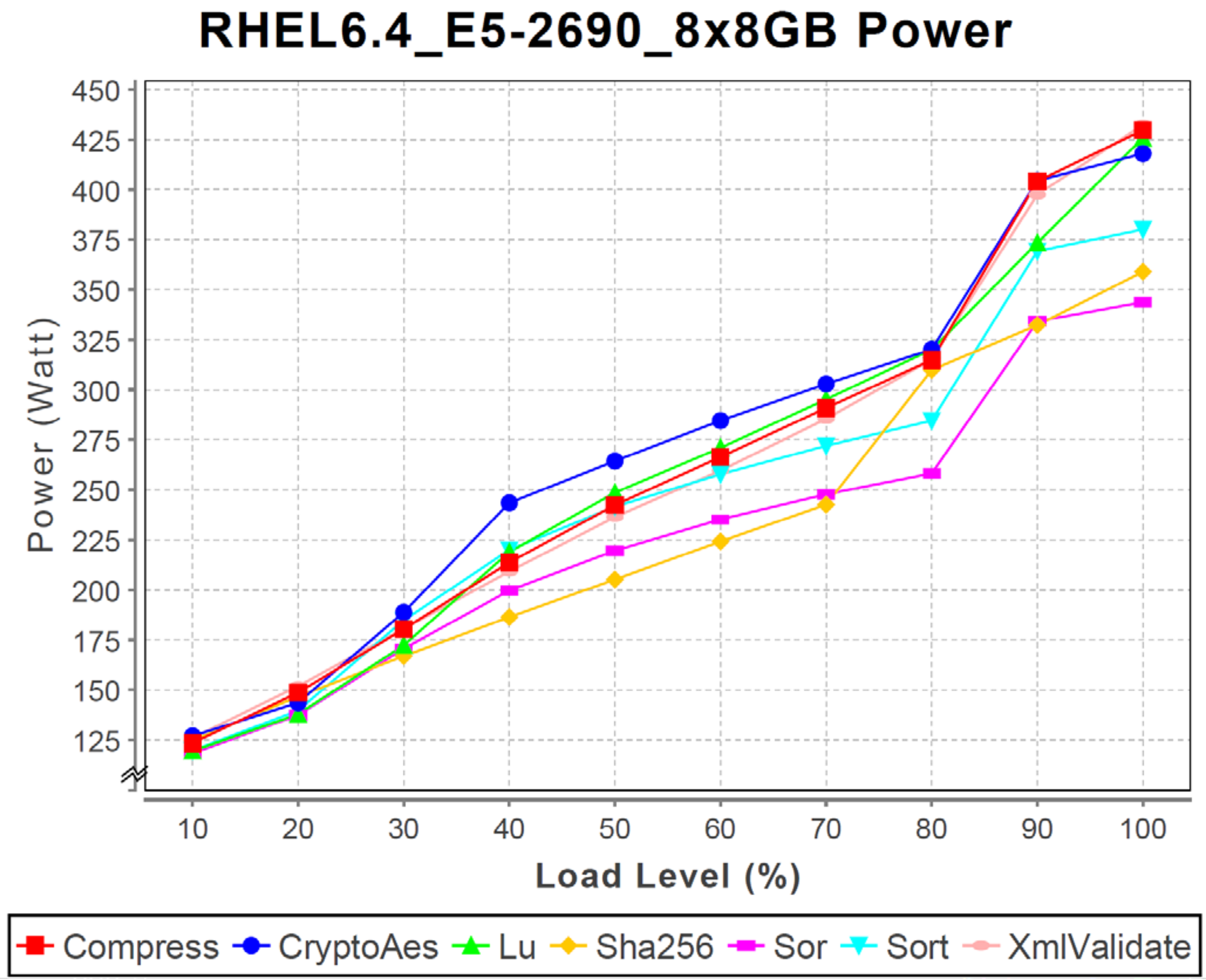
- Baseline System:
  - Tested for varying: CPUs, OS, JVM, …

RX300S7 RHEL6.4 E5-2690 8x8GB	
PSU Output Power	450 W
Sockets	2
CPU	Intel Xeon E5-2690
Cores per CPU	8
Threads per Core	2
Frequency	2.9 GHZ (3.8 GHz Turbo)
Memory Type	8GB 2Rx4 PC3L-12800R ECC
# DIMMs	8
Operating System	Red Hat Enterprise Linux Server 6.4
JVM	Oracle HotSpot 1.7.0 51-b13

- Other base systems:
  - Fujitsu PRIMERGY RX600S6 (4 Socket, Westmere)
  - Fujitsu PRIMERGY RX200S8 (2 Socket, Ivy Bridge)
  - Dell PowerEdge R720 (2 Socket, Sandy and Ivy Bridge)
  - HP ProLiant DL385p Gen8 (2 Socket, AMD Piledriver)

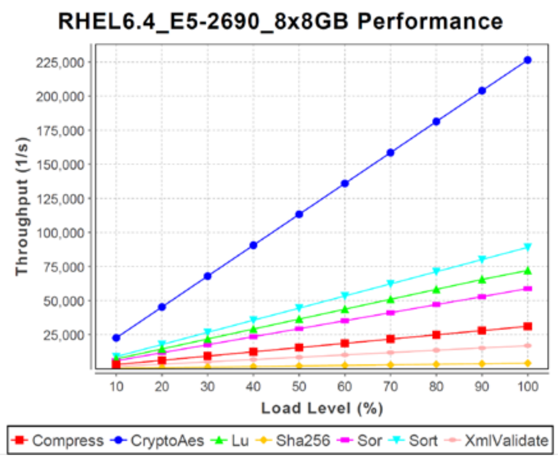
Workload Power Consumption

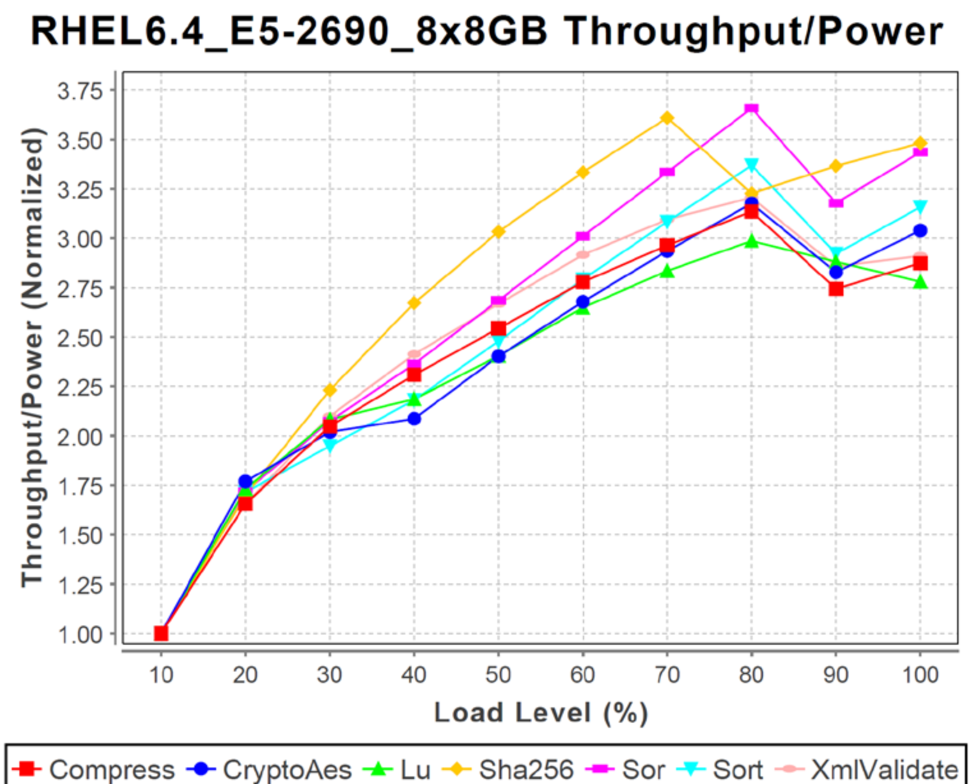
- Biggest Consumer: XMLValidate
  - 126 W @ 10%
  - 431.4 W @ 100%
- Smallest Consumer SOR
  - 118.3 W @ 10%
  - 343.3 W @100%



Workload Energy Efficiency

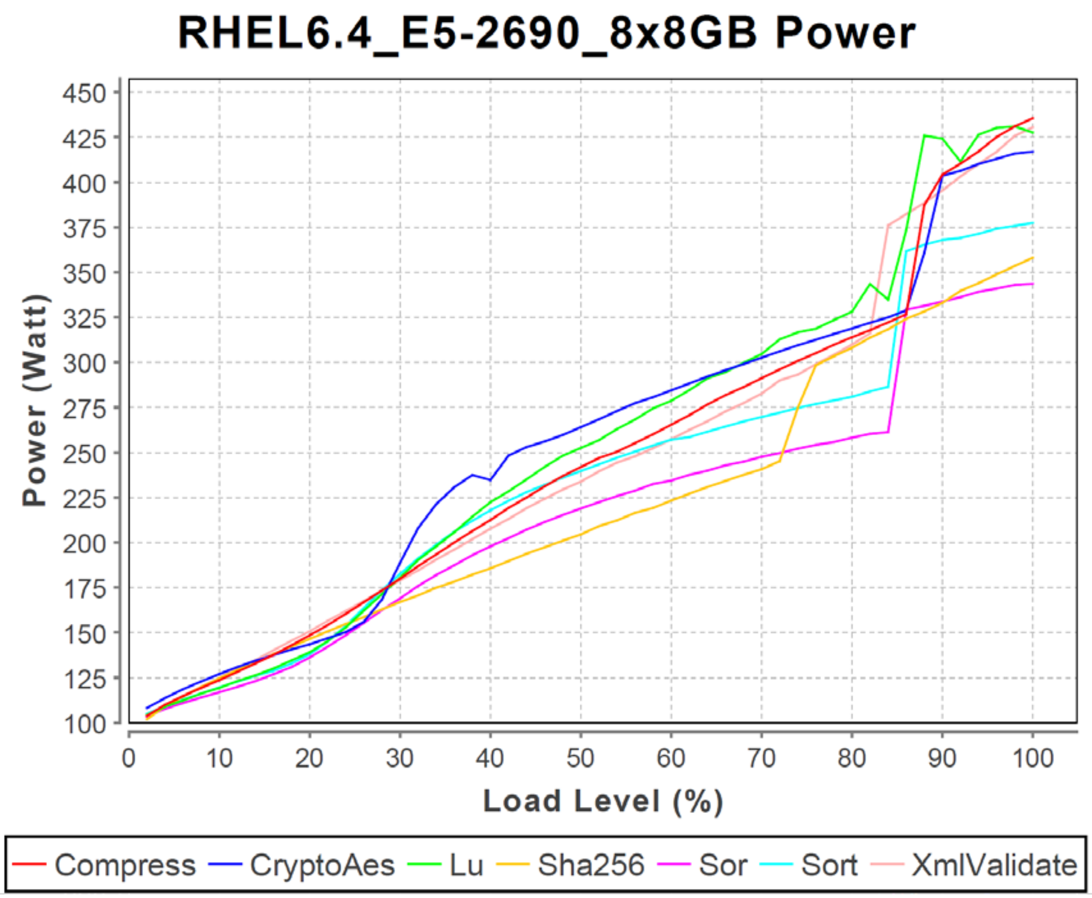
- Throughput is always linear
- Different throughput scales
  ➔ normalization
- Maximum efficiency @ 70% or 80%



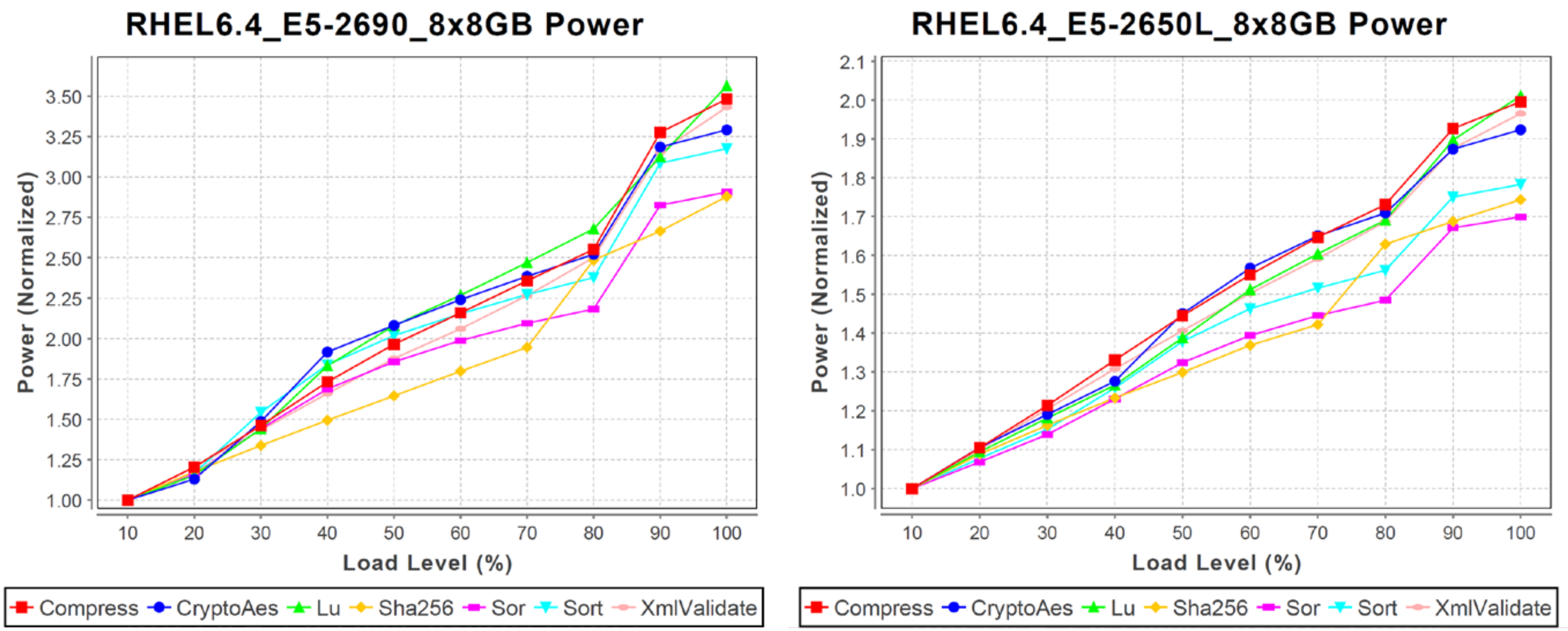


10% Measurement Intervals

- Are observations based on 10% measurement intervals accurate?
  ➔ Measurements at 2% measurement intervals

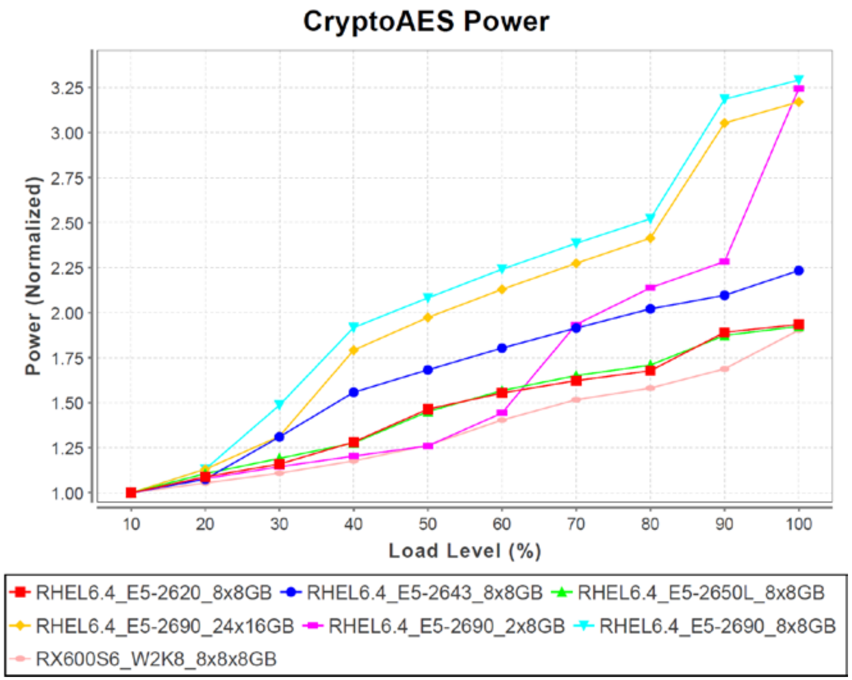


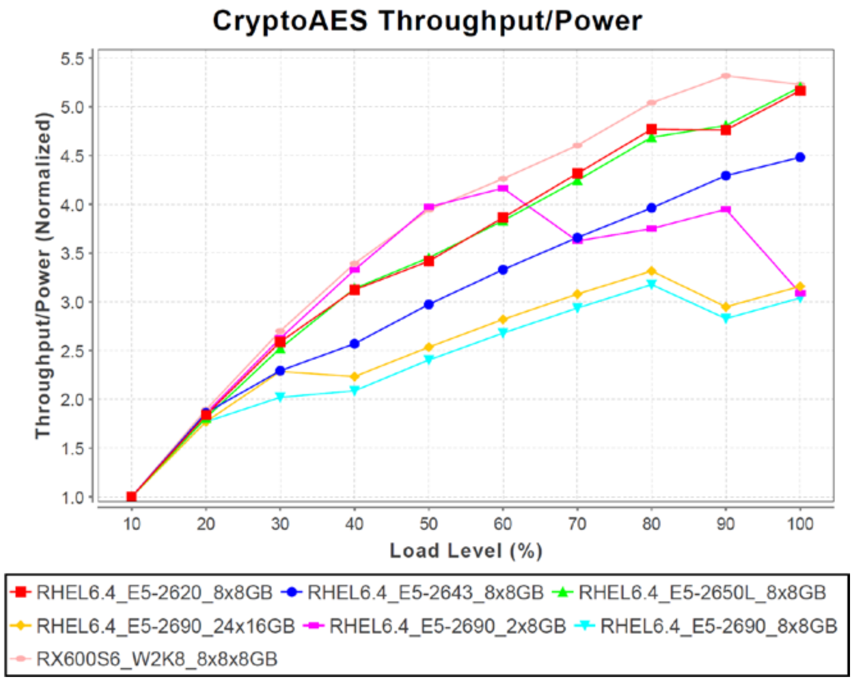
Workload Power at Lower Clock



	Xeon E5-2690	Xeon E5-2650L
#Cores	8	8
Base Frequency	2.9 GHz	1.8 GHz
Turbo Frequency	3.8 GHz	2.3 GHz
TDP	135 W	70 W

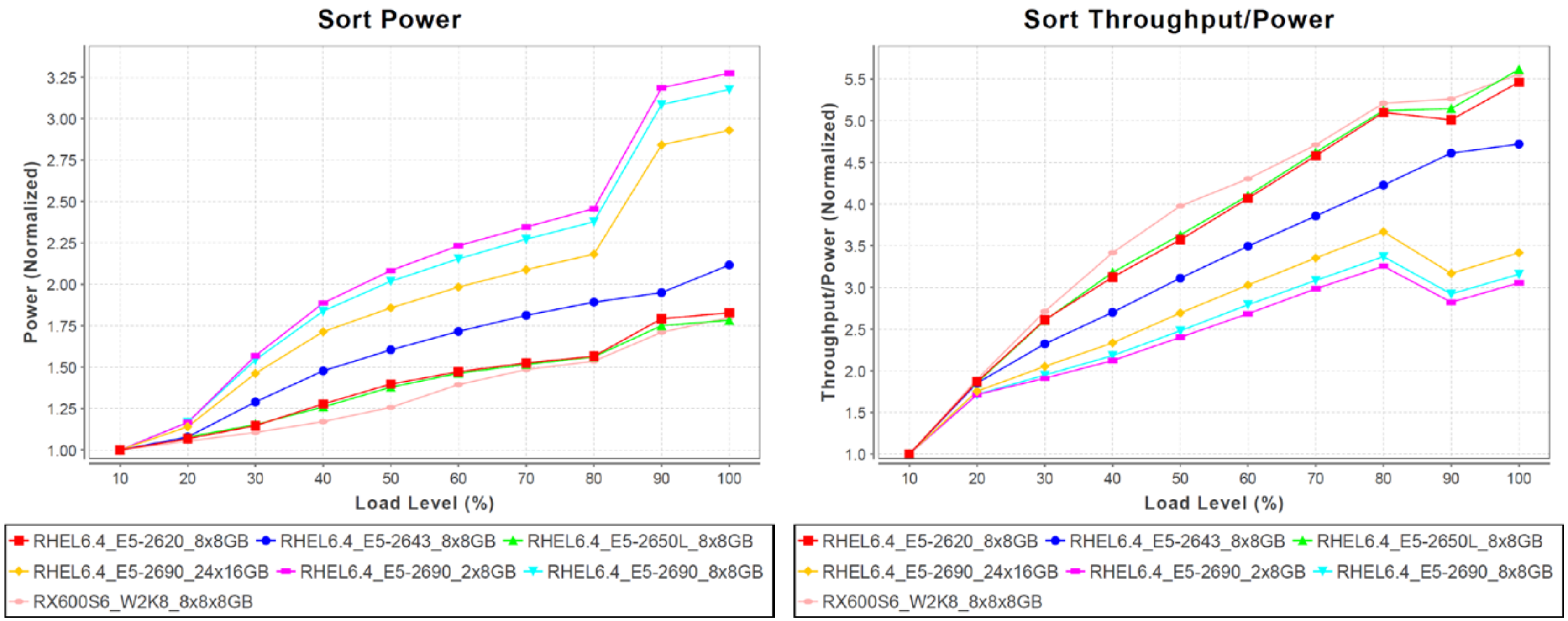
Different Configurations - CryptoAES





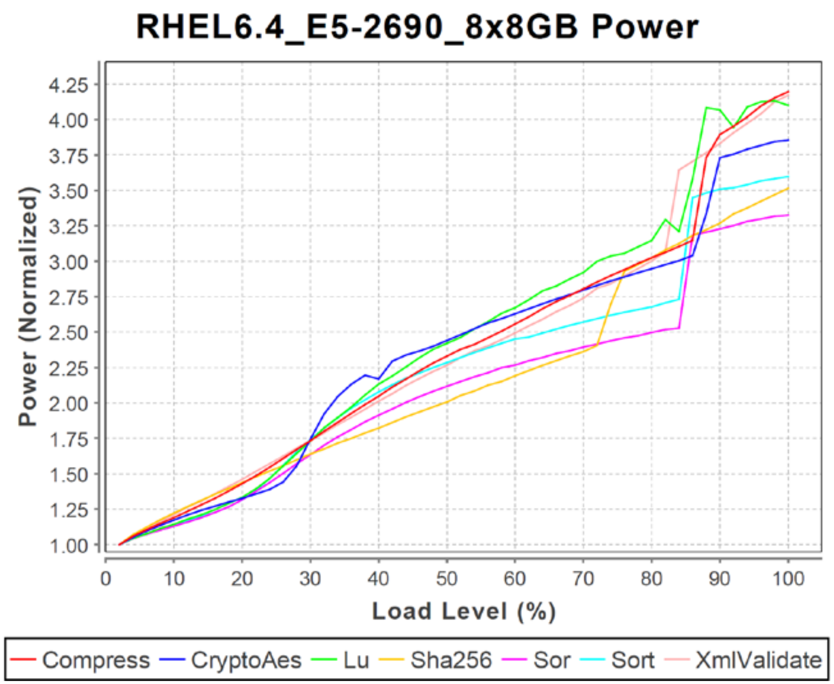
- # memory channels has a big impact.
- Big power consumption difference between min and max load is not always a sign of high energy efficiency!

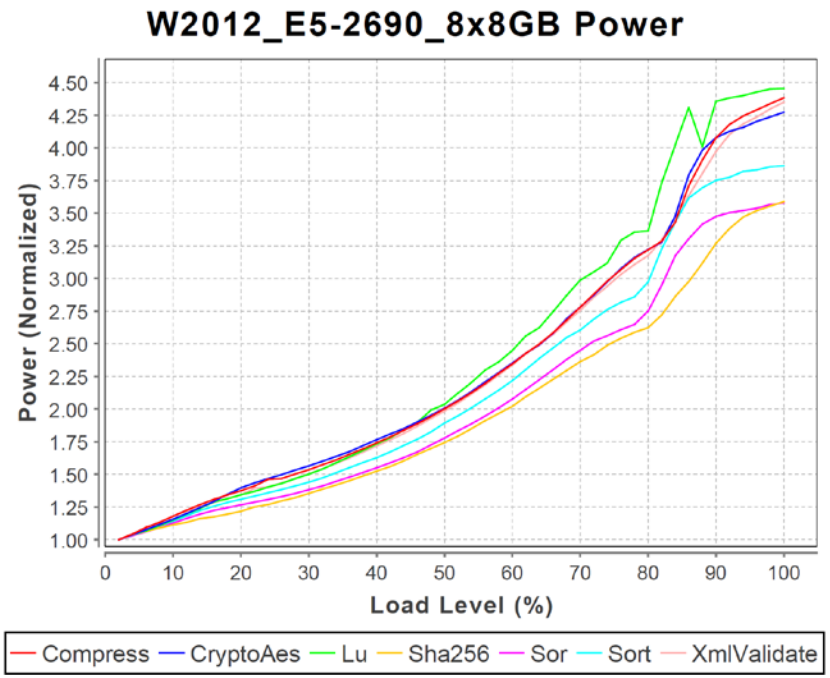
Different Configurations - SORT



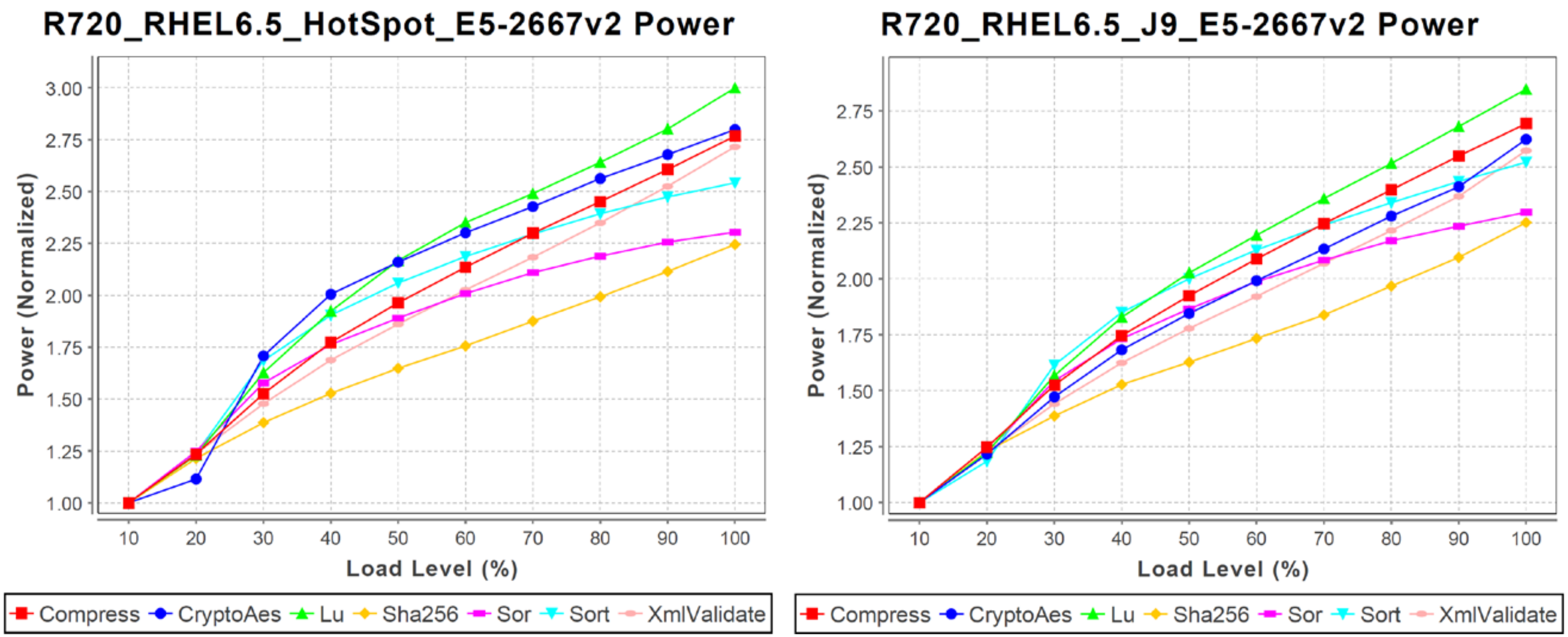
- Xeon E5-2643 system is missing the power consumption increase between 80% - 90%

Operating System



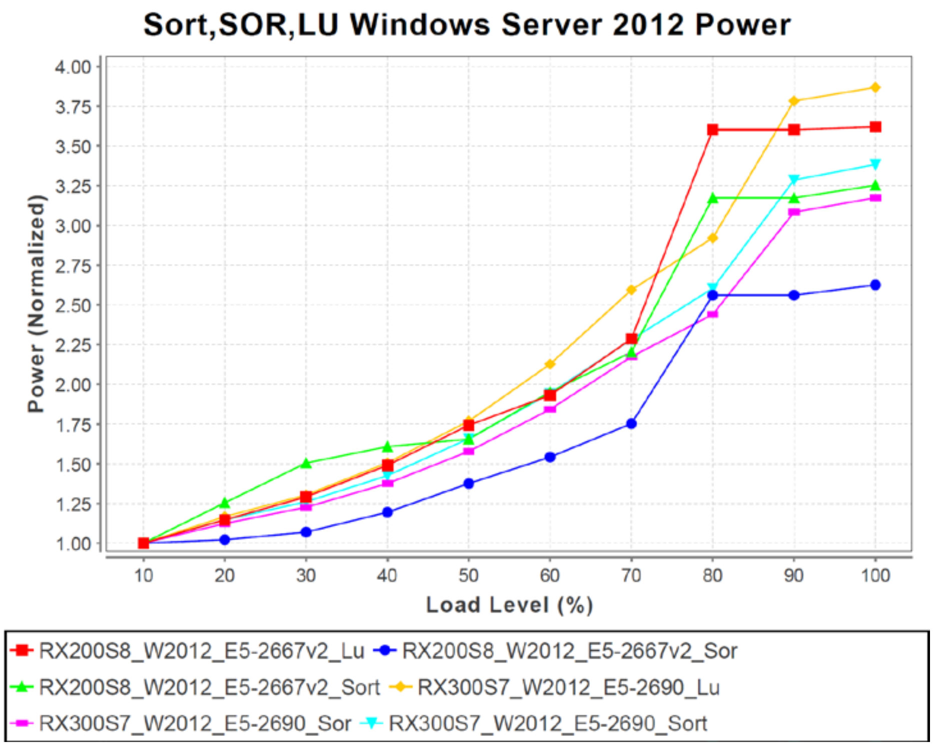


- Operating system has significant impact on power consumption per load level
  - More complex than simple constant power overhead



- JVM power impact through secondary attributes (such as instruction set support)

Worklet Power - CPU Architectures I



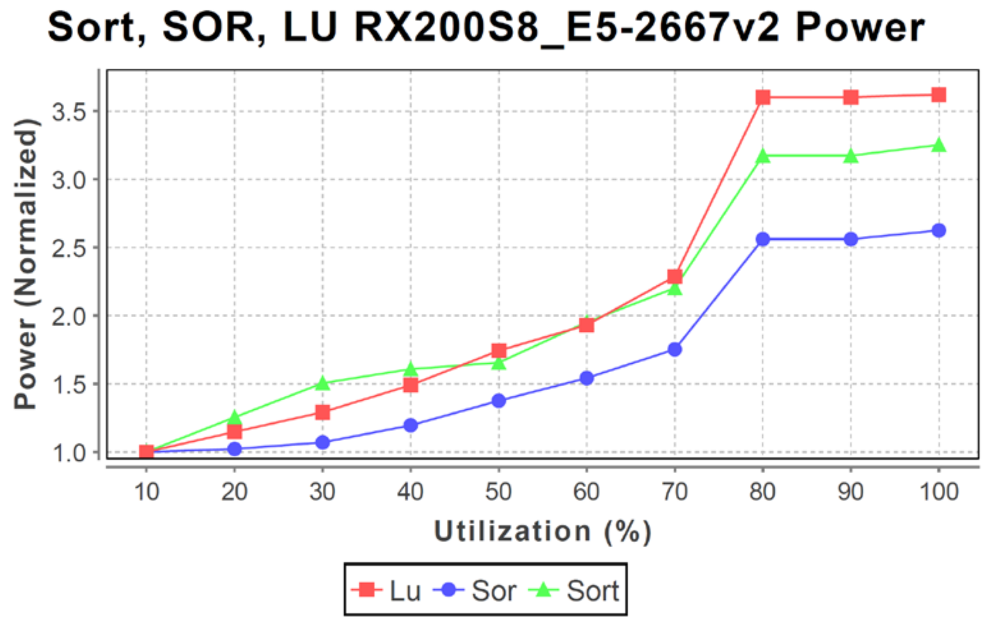
- Worklet power consumption tops out earlier on Ivy Bridge

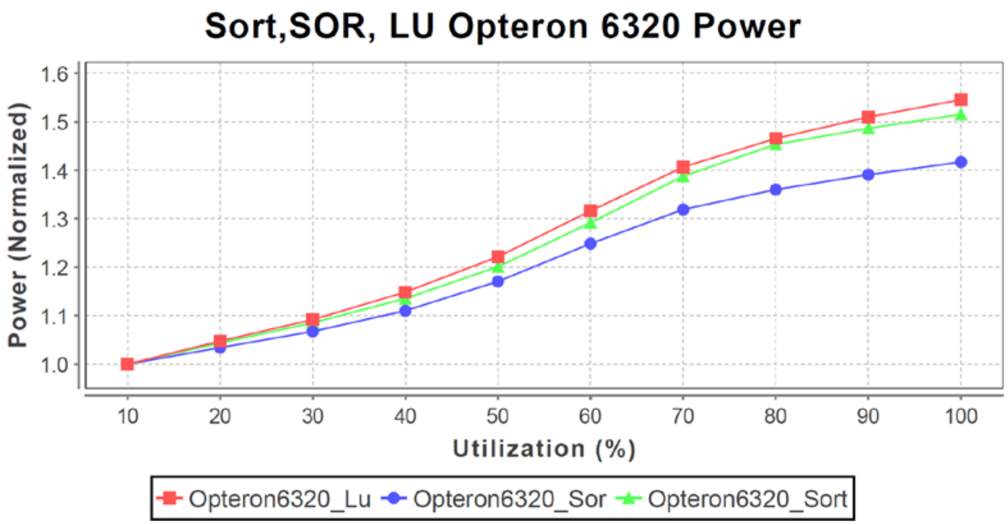
	Xeon E5-2690	Xeon E5-2657v2
Base Frequency	2.9 GHz	3.3 GHz
Turbo Frequency	3.8 GHz	4.0 GHz
TDP	135 W	130 W
Lithography	32 nm	22 nm

Worklet Power - CPU Architectures II

- Both systems run Windows Server

	Opteron 6320
# Modules	4
# Cores	8
Base Frequency	2.8 GHz
Turbo Frequency	3.3 GHz
TDP	115 W
Lithography	32 nm





Conclusions

- Power consumption and energy efficiency of SERT's CPU worklets on different systems
  - Varying operating systems, hardware components, architectures …
- Some lessons learned:
  - Power consumption varies for different CPU worklets and is affected differently by hardware/software changes
  - Operating System has significant impact on power consumption per load level
  - Load level for maximum energy efficiency depends on hardware and software configuration (usually between 70% - 100%)
  - Java Virtual Machine affects power consumption via secondary attributes

Outlook

- Power management must account for varying load levels for optimal energy efficiency

- Power models must account for
  - different workload types utilizing the same resource
  - Operating System effects

- Need to explore drops in power consumption over rising utilization

Thanks for listening!

joakim.kistowski@uni-wuerzburg.de

http://se.informatik.uni-wuerzburg.de

References

[1] L. Barroso and U. Holzle. The Case for Energy Proportional Computing. *Computer*, 40(12):33-37, Dec 2007.

[2] S. Rivoire, P. Ranganathan, and C. Kozyrakis. A Comparison of High-level Full-system Power Models. In *Proceedings of the 2008 Conference on Power Aware Computing and Systems*, HotPower'08, Berkeley, CA, USA, 2008. USENIX Association.

[3] R. Basmadjian, N. Ali, F. Niedermeier, H. de Meer, and G. Giuliani. A Methodology to Predict the Power Consumption of Servers in Data Centres. In Proceedings of the 2nd International Conference on Energy-Efficient Computing and Networking, e-Energy'11, pages 1-10, New York, NY, USA, 2011. ACM.

[4] A. Lewis, S. Ghosh, and N.-F. Tzeng. Run-time Energy Consumption Estimation Based on Workload in Server Systems. In *Proceedings of the 2008 Conference on Power Aware Computing and Systems*, HotPower'08, Berkeley, CA, USA, 2008. USENIX Association.

[5] K.-D. Lange, M. G. Tricker, J. A. Arnold, H. Block, and C. Koopmann. The Implementation of the Server Efficiency Rating Tool. In *Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering*, ICPE '12, pages 133-144, New York, NY, USA, 2012. ACM.

Trademark and Disclaimers

The SPEC logo, SPEC, and the benchmark and tool names, SPECpower_ssj, SERT, PTDaemon are registered trademarks of the Standard Performance Evaluation Corporation. Reprint with permission, see spec.org.

The opinions expressed in this tutorial are those of the author and do not represent official views of either the Standard Performance Evaluation Corporation, Transaction Processing Performance Council or author's company affiliation.