

## Energy Consumption Characteriation of Heterogeneous Servers

Xiao Zhang

*School of Computer Science  
Northwestern Polytechnical University  
127 Youyi xi Road, Xi'an Shaanxi China  
Email: zhangxiao@nwpu.edu.cn*

Jianjun Lu

*School of Computer Science  
Northwestern Polytechnical University  
127 Youyi xi Road, Xi'an Shaanxi China  
Email: lujianjun@mail.nwpu.edu.cn*

Xiao Qin

*Department of Computer Science  
and Software Engineering  
Auburn University, AL 36849-5347  
Email: xqin@auburn.edu*

**Abstract**—Energy and cooling cost is becoming the main cost of data centers. Many studies focused on how to schedule job and migrate Virtual Machine between servers to save energy. An accurate energy consumption model is the basic of energy management. Most past studies show that energy consumption has linear relation with resource utilization. We found that different servers have different energy consumption characters even with same CPU and similar workloads. Most past models were only validated on several servers. These models did not reflect the characters of different kinds of servers in a heterogeneous data center. In this paper, we verified the accuracy of the linear model on 392 servers. The data come from the results of SPECPower\_ssj2008 benchmark. The benchmark was released in 2007. There are 392 servers from 26 vendors were tested in the past six years. These servers represent most common used servers in heterogeneous data centers. We use several methods to validate the accuracy of the linear model. The results show the model does not work well for all kinds of servers. We also analyze the trend of energy consumption of last six years.

**Keywords**-Energy consumption model; Heterogeneous data center; Resource Utilization

### I. INTRODUCTION

Energy and cooling cost is becoming the main cost of data centers. To reduce the energy cost in large-scale data centers, researchers developed a wide range of energy-saving and thermal management techniques (e.g., workload consolidation, live migration, CPU throttling solutions).

Workload consolidation is one of the most effective ways of conserving power by turning off spare servers. In many cases, the workload consolidation technique is incorporated with virtual machines, which can be migrated from many physical machines into a smaller number of physical machines. This approach can save energy by reducing the number of active servers. Migration policies depend on both performance requirements and power consumption of different workloads. Power consumption models fall into two categories. Models in the first one can predict the power consumption of each physical machine; models in the second group aim to determine the impact of each virtual machine's load on energy consumption. Previous studies show a strong correlation between performance events and the power consumption. Existing models rely on various performance events (e.g., CPU, IO, memory, and cache)

to estimate energy consumption of of sub-components of a computing system; sub-components can be classified as either CPU-bound (e.g., CPU and cache) or IO-Bound (e.g., DRAM, HDD).

The existing power consumption models are practical for specific machines under certain conditions. These models are reasonably accurate for data centers equipped by homogeneous computing platforms. However, the heterogeneous and dynamic characteristics of modern data centers make these models less accurate and trustworthy. To apply the traditional models in a heterogeneous data center, system administrators must verify the models on different types of servers. When it comes to data centers containing a large number of heterogeneous computing components, it is unproductive and time consuming to adopt and validate a single model to predict energy consumption of a wide variety of servers.

Early energy consumption model use CPU utilization as only parameter[1]. Some studies try to monitor several performance counters related with CPU to estimate the power consumption[2]. Past studies use multiple performance counters to calculate the energy consumption including CPU, memory, disks and network[3], [4]. They extended the model with more parameters to get more accurate power consumptions. All of these models use linear model. Most of these models were only validated on several servers. Fan et al. validate the linear model in Google data center with thousands of servers, but they did not mention how many kinds of servers in their data center.

To address the aforementioned problem, we validate the accuracy of the linear model by comparing modeling results against available data obtained from the SPECPower\_ssj2008 benchmark (or SPECPower for short), which is a widely adopted benchmark used to evaluate the power and performance characteristics of servers. Until Oct.2012, 392 servers have been evaluated by SPECPower. These servers came from 26 different vendors through past six years; these servers can represent the heterogeneous servers in the market. To our knowledge, this is the first power usage study of so many kinds of different servers. Some of our key findings and contributions are:

- A wide validation of heterogeneous servers. We vali-

dated the linear energy model by 392 published results tested by different kinds of servers. We analyzed the accuracy through R-squared, RMSE and Average Error for different kinds of servers.

- We found that not all servers fit the linear model well. 6.5%(25 kinds of servers) of R-squared values are less than 0.95, which means CPU utilization is not significantly correlated with power consumption.
- Although the average RMSE of all servers is 14.86, there are 118 kinds of servers' RMSE values bigger than 10 and 49 kinds are bigger than 20. Which means for these servers, even though R-squared value shows the linear model fit well, there is still big difference between modeled and real power consumption.
- We also illustrate the different power consumptions of servers with same CPU. We find that different servers have different power consumption characters even with same CPU and similar workloads.

The rest of this paper is organized as follows. Section II introduces the existing energy consumption models and the SPECpower benchmark. Section III presents how to construct an energy consumption model and integrate it with other components. Section IV shows the analysis of SPECpower results, including accuracy of the linear model and the energy consumption trends in the past 6 years. The last section discusses the future work of this study.

## II. RELATED WORKS

### A. Energy Consumption Character and Model

Power management is becoming an important issue to be addressed in data centers. Managers have to reduce energy costs of servers and cooling systems in order to offer competitive services. It is straightforward to measure the energy consumption of an entire data center[5]. To schedule jobs or workload consolidation in an energy-efficient way, one has to estimate the energy consumed by each computer node in a data center.

Power meters retrieve system power usage accurately in real-time manner. This method cannot adapt to dynamic computing environments like computing clouds, because there are excessive number of computers node to be measured in real-time. The second approach is to estimate the power usage based on functional units, which sum up nominal power of each component. Values measured by this approach are constant under certain configurations, and these measured values are always larger than actual energy usage. Ludashi applied this method to estimate the energy cost of personal computers[6]. Ludashi's approach checks all components of a PC, including main board, CPU, hard disk, and memory. The measurement results are not accurate under normal usage conditions. For example, the estimated value of a laptop (i.e., ASUS U35JC) is 107 Watts, but the accurate energy consumption is 47 Watt. CPUs cost most

of the energy of computers; the power consumed by CPUs can be measured by thermal design power (TDP), which is measured under normal load[7]. Energy consumption varies sharply when workload and configuration significantly changes, thereby making a large discrepancy between estimated energy consumption and actual energy usage.

The third method is to model power usage based on performance counters. Which can be divided into two classes: chip-level performance counter queried from chips and system-level counter got from OS. Bricher *et al.* designed a system power model using hardware performance counters for vital system subcomponents. Their model relies on microprocessor performance counters to measure an entire system power consumption[8]. They used two different sets of performance counter on a quad-socket Intel CPU and AMD dual-core CPU. They selected nine events to model power consumption of Intel Quad-socket CPU and use thirteen different events to model power of AMD CPU[9]. Singh *et al.* built some model for AMD Phenom[10].

The above existing hardware performance counter solutions have the following drawbacks. First, Most of the models were tailored for special processors or computer architectures. Different CPUs have different performance counters[9], which make it impractical to apply these models to heterogeneous computing environments. Intel Pentium 4 processor has 18 performance counters that can be programmed to monitor up to 59 event classes, but there are only 15 events available for Intel XScale processors[11]. Second, there are new emerging technologies (e.g., dynamic voltage scaling or replace a SCSI disk to a SSD disk) to save energy consumption in computers. The existing static estimation methods are unable to address the dynamic features of computing environments.

Some models estimate power usage based on resource utilization. Fan *et al.* implemented models based on CPU utilization, and estimate energy consumption of each Rack, PDU and cluster. They focus on critical power and power usage of entire clusters with several thousand servers[5]. Heath *et al.* extended the models by using OS-reported CPU, memory and disk utilization[4].

Some other studies try to find energy consumption of each process or each Virtual Machine. Snowdon *et al.* discussed approaches to monitoring power for applications[12]. This method involves collecting information in real time about resources consumed by each application. Snowdon's work assumes that the energy usage of an application is directly related to the amount of CPU time. The downside of this approach is that it does not take into account loads handled by I/O devices and multi-core processing. Bohra *et al.* proposed a model called "VMeter", which predicts instantaneous power consumption of an individual virtual machine hosted on a physical node[13]. To predict energy consumption of virtual machines in cloud computing, Karan *et al.* proposed Joulemeter that uses only power models

to accurately infer the power consumption of a Virtual Machine[10].

Models based on resource utilization can be adapted to heterogeneous cluster systems. Real-life servers are characterized by different configuration, performance and workloads. Most of previous models were only validated on several different kinds of server. Fan’s models were validated on several thousand servers in Google, but they did not mention how many different kinds of servers in their data center. Heath *et al.* tested their models on 4 blade servers and 4 PCs[4].

There is a trend to include power/energy sensors in the processors. Intel presented RAPL(Running Average Power Limit) technology to estimate energy by using various hardware performance counters in recent Intel CPUs[14]. The values are also exposed to users through PAPI(Performance API). Recent NVIDIA GPUs can report power usage via the NVIDIA Management Library(NVML). The `nvmlDeviceGetPowerUsage()` function retrieves the power usage reading for the device, in milli-watts. This is the power draw for the entire board, including GPU, memory, etc. The reading is accurate to within a range of +/- 5 watts[15]. These technologies have been validated to closely follow actual energy consumption[16]. But these models are limited to special sub components like CPU and GPU. And developers have to use different interface to get energy consumption of different sub components.

### B. The SPECpower Benchmark

SPECpower is the first industry-standard SPEC benchmark that evaluates the power and performance characteristics of volume server class and multi-node class computers[17]. The benchmark was created to compare energy efficiency among different servers. Currently, many vendors provide energy efficiency evaluations, but the vendor’s evaluation results are not comparable due to different workloads, configurations, and test environments. The benchmark helps to measure the power of computing systems under various workloads.

The newest version of the SPECpower benchmark was released on July 26, 2012. The current version exercises CPUs, caches, memory hierarchy and the scalability of shared memory processors (a.k.a., SMP) as well as the implementations of the Java Virtual Machine or JVM, the Just-In-Time compiler or JIT, garbage collections, threads and some aspects of operating systems.

The benchmark runs on a system under test (SUT) and a controller machine, which controls workloads of SUT and collects power data from a power meter connected to SUT. Figure II-B shows our testbed that makes use of the SPECPower benchmark to evaluate energy efficiency of computing systems<sup>1</sup>.

<sup>1</sup>[http://www.spec.org/power\\_ssj2008/](http://www.spec.org/power_ssj2008/)

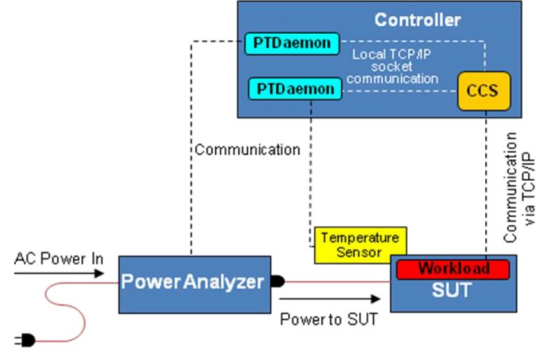


Figure 1. Typical Test Environment of SPECpower\_ssj2008

The SPECpower benchmark is comprised of processes, where CPU utilization varies from 100% to idle with an increment of 10%. SPECpower uses very little network I/O; neither does SPECpower write measured data to disks during each test. Nevertheless, SPECpower issues reads to tested disks. Previous studies show that disk subsystems have nearly a constant power consumption during the entire range of workloads[8]. CPU activities are recorded when SPECpower is running a testbed, which keeps track of CPU usage, operation per second, and power consumption.

### III. HIGH LEVEL ENERGY CONSUMPTION MODEL

We start this section by offering an overview of the energy consumption model. Then, we describe how to build the model using the measurement results of SPECPower. We also will demonstrate how to apply the proposed model in heterogeneous data centers.

#### A. Overview

The first step forward to modeling energy consumption is to break down an entire computing system into several components. Recent studies show that workload has a linear correlation with energy consumption with respect to each component[8], [18], [19]. Figure 2 illustrates the process of constructing, verifying, and deploying the model. In the initial phase, we build energy consumption model by analyzing the power consumption results of running the SPECpower benchmark. During the analysis of the measurements, we gather the workloads of SPECpower and map the test results into traditional workload metric, from which the energy consumption model is established. We verify the model by running the benchmark and other real-world applications on various servers. After completing the analysis and verification phases, we configure model parameters for each tested server. Next, we can integrate the model with performance monitor tools, including ganglia and performance co-pilot. Since the model is able to estimate energy consumption of all the tested servers, we can integrate the model into schedule management tools, which use energy consumption data as an input of schedule policies for heterogeneous data center.

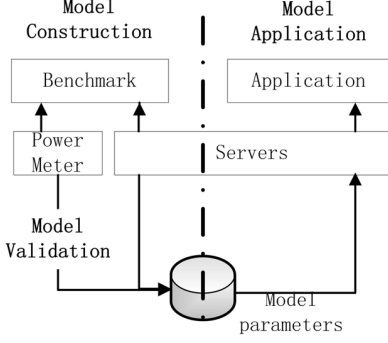


Figure 2. Energy Consumption Model Overview

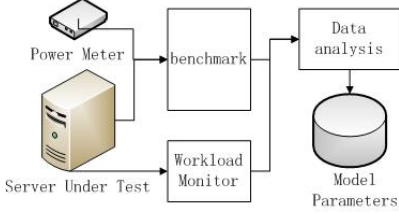


Figure 3. Energy Consumption Model Construction

### B. Model Construction

Early energy consumption model use CPU utilization as only parameter[1]. Some studies try to monitor several performance counters related with CPU to estimate the power consumption[2]. Past studies use multiple performance counters to calculate the energy consumption including CPU, memory, disks and network[3], [4]. They extended the model with more parameters to get more accurate power consumptions.

The energy consumption of each component can be calculated as  $P = C + R \times P$ , where  $C$  is a constant power consumption when the component is idle.  $R$  is the usage ratio of the component and  $P$  is the increment of energy consumption when the usage ratio goes up. Hence, the total power consumption  $P_{total}$  can be expressed as Equation 1:

$$P_{total} = Rate \times Power$$

$$= |1, R_{cpu}, R_{mem}, R_{disk}, R_{net}| \times \begin{pmatrix} P_{misc} \\ P_{cpu} \\ P_{mem} \\ P_{disk} \\ P_{net} \end{pmatrix} \quad (1)$$

Where  $P_{misc}$  is a constant power consumption when the system is sitting idle.  $P_{misc}$  incorporates the power dissipation in all the components, including chassis, power supply, and peripheral devices.  $P_{cpu}$ ,  $P_{mem}$ ,  $P_{disk}$ ,  $P_{net}$  are power consumption of CPU, memory, disk, and network interconnect, respectively.  $R_{cpu}$ ,  $R_{mem}$ ,  $R_{disk}$ ,  $R_{net}$  are the utilization or usage ratio of the four types of resources.

Linux provides a lavish method to assess usage ratios at different levels. For example, hardware performance counters can be collected by the `perfctl` and `perfmom` driver programs[20]; and virtualized (per-process) counters also can be monitored in our testbed. The Linux servers offer utility programs (e.g., `top`, `free` and `iostat`) to measure usage ratios of memory, disks, network I/Os. SAR is applied to monitor the load of CPU load, disk, and memory at given intervals[21]. The challenge is that there are an excessive number of performance events; users have to pick a small set of representative events from the 40 detectable performance events provided by Pentium IV. Bircher and John studied the events of processor and selected representative events according to the examined architectures[8]. Their selections are determined by average error rates and a qualitative comparison of measurement and modeling results. Economou *et al.* collected data from SAR and `perfctl`, in their approach a single counter for each subsystem is used. The results show that most of average error of linear prediction models is less than 5%[18].

Data centers in Google also use CPU utilization to estimate the power usage[5]. There are several advantages to take CPU utilization as the only parameter. Firstly, Processors and memory are two major contributors to the power consumption of computing systems. Several past studies show that disk and network resources have almost constant power consumption[3], [4]. Our findings also confirm that the energy consumption of the I/O devices and network do not noticeably change. Second, it will cause performance slowdown to capture and handle multiple performance events, especially sampling at a small interval. So we use the most popular linear mode to predict the energy consumption of each servers (Eq.2).

$$P_{total} = R_{cpu} \times Power + P_{idle} \quad (2)$$

To verify the accuracy of the linear model, we use the model to fit the 392 results of SPECpower benchmark.

### C. Model Verification and Analysis

We validate the accuracy of the linear energy model using R-squared, RMSE and Mean Error, which calculates the mean error for each combination of modeled power consumption and measured consumption. The Mean Error is used in many past studies[3], [9].

The R-squared value is used to describe how well a prediction model fits a set of data (Eq.3). An R-squared near 1 indicates that the model fits the data well, while R-squared close to 0 indicates the model does not fit the data very well. Usually a value greater than 0.95 means the model is acceptable. A data set has values  $y_i$ , each of which has an associated modeled value  $f_i$ , R-squared value is the value of total sum of residuals to total sum of squares .

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (3a)$$

$$SS_{tot} = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (3b)$$

$$SS_{err} = \sum_{i=1}^n (y_i - f_i)^2 \quad (3c)$$

$$R^2 = 1 - SS_{err}/SS_{tot} \quad (3d)$$

RMSE(Root Mean Square Error) is a frequently used measure of the differences between values predicted by a model(Eq.4). The RMSE serves to aggregate the magnitudes of the errors in predictions for various times into a single measure of predictive power.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - f_i)^2}{n}} \quad (4)$$

The Mean Error is used in many past studies[3], [9]. It reflects the average error from different points(Eq.5).

$$MeanError = \frac{\sum_{i=1}^n \frac{|f_i - y_i|}{y_i} \times 100\%}{n} \quad (5)$$

The SPECpower benchmark tests the energy consumption of the servers under various workloads. Each result is comprised of eleven points, each of which is a set of energy consumption, CPU load, and operations per second. We verify all the 392 published results of different servers. Our results show that the linear model does not fit all condition well.

#### D. Applying the Model

Energy consumption model can be used independently to estimate the energy consumption of whole cluster. We use a set of parameters for servers on each "group" of same model and same enclosures. For each server in the group, we estimate the energy consumption by their utilization (different utilization get different power consumption). We can gather real-time energy consumption from each server or just collect the utilization of each server and calculate the energy consumption on another machine.

Energy consumption model can be readily integrated with any performance monitor and scheduler (see Figure 4). Ganglia is a scalable distributed monitoring system for high-performance computing systems like clusters and Grids. Ganglia - with a hierarchical design targeting at federations of clusters - can collect basic metrics (e.g., system load and CPU utilization). Ganglia also can keep track of user-defined metrics through plugins (e.g., C/Python modules)[22]. To implement our model in a heterogeneous cluster, we enable the modeling module to retrieve workloads from ganglia.

Then, the model can output energy consumption of each computing node in the cluster. Estimated power consumption data may be used by job schedulers and workload consolidation modules in the cluster.

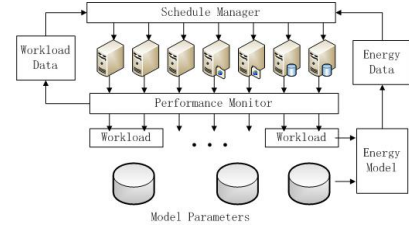


Figure 4. Energy Consumption Model Application

## IV. ANALYSIS OF BENCHMARK RESULTS

### A. Comparison Results of Same CPU

Previous studies show that CPU load is a major contributor to energy consumption. Since a CPU comes with matched North Bridge chips, we assume servers with the same CPU have a similar energy consumption. We choose seven different servers equipped with the same type of CPU (i.e., Intel Xeon E5-2670 2.60 GHz 2 chips 32 threads); All of the tested servers were single node. Table I lists the configuration of the seven servers and their power consumption. Our preliminary results show that not all the power consumption results follow the trend of a linear model. We observe that energy consumptions of the servers are very similar when the CPU load is below 50%. However, when the CPU load is increasing, the energy consumption rates of the seven tested servers are very different from each other. This trend is especially true when the CPU load is approaching 100%(see Figure 5).

Table I  
POWER CONSUMPTION OF DIFFERENT SERVERS WITH SAME CPU

Hardware Vendor	System Enclosure	Mem (GB)	Power Information		
			idle	max	avg.
Dell	PowerEdge R620	24	54.1	243	139.5
Dell	PowerEdge R720	24	51.0	231	133.3
Dell	PowerEdge T620	24	50.2	227	131.2
Hitachi	HA8000/RS210-hHM	32	99.3	337	194.3
Hitachi	HA8000/RS220-hHM	32	104.0	325	190.9
Huawei	RH2288 V2 32G	32	55.6	282	146.1
Huawei	RH2288 V2 16G	16	54.8	279	142.9

### B. Accuracy of Linear Model

We validate the energy consumption of the servers running the SPECpower benchmarks. Each SPECpower result contains 11 pairs of workload and measured power consumption. We calculate the coefficients of a polynomial P(workloads) of linear mode that fits the data measured power best in a least-squares sense.

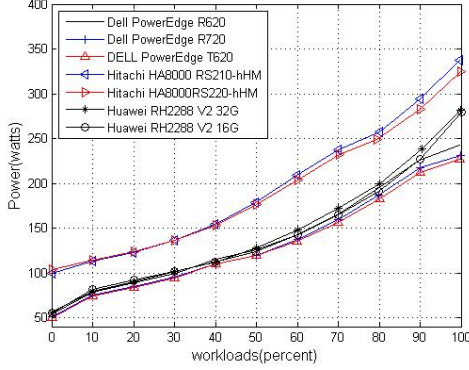


Figure 5. Power Consumption of Different Servers with same CPU

We use CDF(cumulative distribution function)  $F_x(X) = P(X \leq x)$  to show the distribution of the linear model in Figure 6(d). We find that the R-squared values of 25 kinds of servers are less than 0.95(Figure 6(a)). The minimum value is 0.8418, which comes from the server Colfax International CX2266-N2<sup>1</sup>. Energy of CX2266 can be modeled by  $P_{total} = 193.7466 + 0.9385 \times R_{cpu}$ . The mean error of this server is 3.52% and RMSE is 12.95.

RMSE(Root Mean Square Error) is a frequently used measure of the differences between values predicted by a model. In this case, a value greater than 20 means each estimate value has a rough error of 20 watts. One server will use about one more killowatt-hour every two days in these conditions. From Figure 6(b), we can find that 12.5%(49 servers) RMSE values are greater than 20, and several results get very large RMSE value. The average all RMSE is 14.86. The maximum value of RMSE is 330.7, which comes from a high performance server (Fujitsu PRIMERGY BX920 S3)<sup>2</sup>, which has a power consumption range from 1014 to 4965 watts. The R-squared of the linear mode is 0.9217, which is less than 0.95. It does not fit linear model. Power consumption of the server can be calculated by  $P_{total} = 762.1 + 35.92 \times R_{cpu}$ . The mean error of the server is 330.7.

Figure 6(c) shows that most of the mean errors of the linear model are below 8%. The average mean error is 2.74%. Interestingly, the model is more accurate before No.300 than after. A majority of mean errors are under 4%. The first result of 2012 is No. 292, meaning that the model offers good power-consumption estimates for servers shipped before 2012. We use CDF function to show the distribution of the linear model in Figure 6(f).

<sup>1</sup>[http://www.spec.org/power\\_ssj2008/results/res2007q4/power\\_ssj2008-20071129-00018.html](http://www.spec.org/power_ssj2008/results/res2007q4/power_ssj2008-20071129-00018.html)

<sup>2</sup>[http://www.spec.org/power\\_ssj2008/results/res2012q2/power\\_ssj2008-20120511-00459.html](http://www.spec.org/power_ssj2008/results/res2012q2/power_ssj2008-20120511-00459.html)

Table II  
AVERAGE POWER CONSUMPTION OF DIFFERENT YEARS

Year	results numbers	Power Consumption(watts)			Perf/ power
		Idle	Max	Avg.	
2007	4	173.75	236.88	277.25	627.50
2008	38	146.71	200.49	248.24	753.05
2009	37	104.42	169.05	226.00	1441.59
2010	39	71.16	167.26	235.51	2570.13
2011	24	83.82	169.17	247.71	2679.54
2012	64	64.79	149.56	257.55	4350.19

### C. Energy consumption trends

The first version of SPECpower was released six years ago. Until Oct. 2012, there have been 392 available test results. We analyze the energy consumption trends during the past six years. For example, we consider servers with eight-core processors, and compare the energy consumption of the servers that have been tested in different years. We choose 88 SPECpower results out of the available benchmarking results before 2010. Since 2010, most servers have been deployed with more than 12 cores. Only four results obtained in 2011 and 2012 are collected from servers equipped with eight-core processors. Therefore, we conduct benchmarking experiments for 12-core and 16-core processors available after 2010. In total, we choose to use 209 benchmarking results.

We listed all the idle power consumption in Figure 7(a), max power in Figure 7(b) and average power in Figure 7(c). We can find that the idle power consumption deceased over time, but the max power consumption were kept in same level. Which means although servers may use less power at different utilization ratio, the power design for data center did not change too much. Because construction must take the max power consumption into consider. Another interesting results is the trend of performance/power metric. We can find that it grows up over time form the Figure 7(d). This shows that the Efficiency of power usage improved quickly.

The above results show that three HP ProLiant DL580 servers in 2011 consume more than 400 watts even in the idle mode. These HP servers are equipped two the Intel Xeon Processor 7020, which has a typical thermal power of 135w[7]. The Intel Xeon processor consumes more than 270 watts under normal load. Because these three results are outliers (i.e., much higher than others), we do not include the results in Table II, which illustrates the average power consumption of servers in different years. Perf/Power is used to compare power-performance metric across all servers, measured by overall transactions divided by the sum of the average power consumption.

## V. CONCLUSION

In this paper, we verified the accuracy and effectiveness of the linear model in predicting energy consumption. This work was motivated by the fact that the existing models are

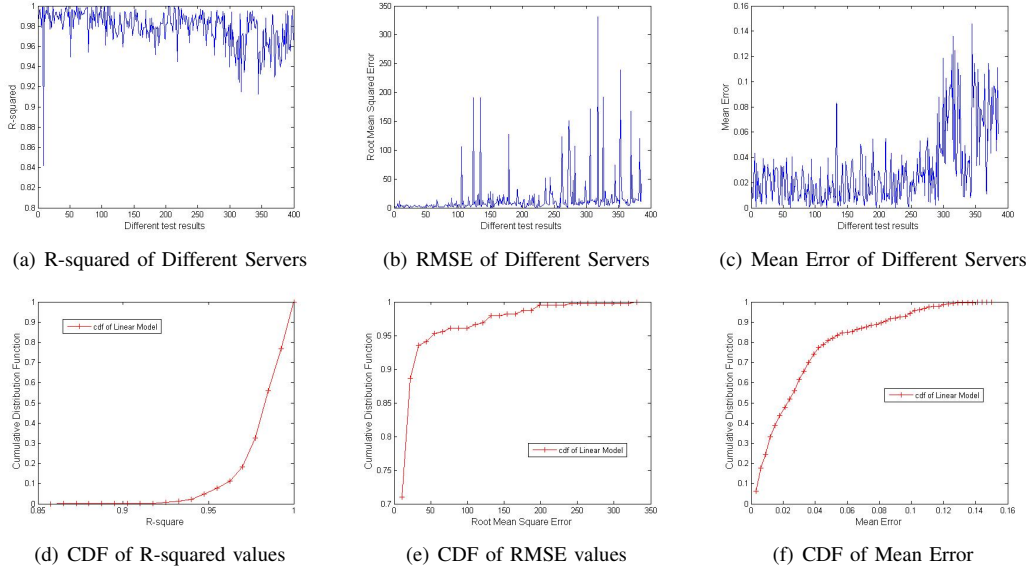


Figure 6. Validation of Different Servers

inadequate for future data centers equipped with heterogeneous servers. Most of past studies only validate their model in several machines. We use published SPECpower benchmark results to validate the linear model. The benchmark is a widely adopted benchmark used to evaluate the power and performance characteristics of servers. We found that not all servers fit the linear model well. 6.5% (25 kinds of servers) of R-squared values are less than 0.95. There are 12.5% (49 kinds of servers) of RMSE values are greater than 20. Our finding shows that the linear model cannot be used to estimate the power consumption of these servers. Energy consumption model can be integrated with performance monitoring tools deployed in data centers, thereby making job scheduling mechanisms more energy efficient.

#### ACKNOWLEDGMENT

This work was made possible thanks to the NPU Fundamental Research Foundation under Grant No. JC20110227, the National Science and Technology Ministry No.2011BAH04B05 National High-tech R&D Program of China (863) under Grant No.2013AA01A215, and the National Natural Science Foundation of China under Grant No.61033007. And this work was supported by China Scholarship Council. Xiao Qin's research was supported by the U.S. National Science Foundation under Grants CCF-0845257 (CAREER), CNS-0917137 (CSR), CNS-0757778 (CSR), CCF-0742187 (CPA), CNS-0831502 (CyberTrust), CNS-0855251 (CRI), OCI-0753305 (CI-TEAM), DUE-0837341 (CCLI), and DUE-0830831 (SFS).

#### REFERENCES

[1] C. Isci and M. Martonosi, "Runtime power monitoring in high-end processors: Methodology and empirical data," in

*Proceedings of the 36th annual IEEE/ACM International Symposium on Microarchitecture*. IEEE Computer Society, 2003, p. 93.

[2] W. L. Bircher and L. K. John, "Complete system power estimation: A trickle-down approach based on performance events," in *Performance Analysis of Systems & Software, 2007. ISPASS 2007. IEEE International Symposium on*. IEEE, 2007, pp. 158–168.

[3] D. Economou, S. Rivoire, C. Kozyrakis, and P. Ranganathan, "Full-system power analysis and modeling for server environments," in *In Proceedings of Workshop on Modeling, Benchmarking, and Simulation*, 2006, pp. 70–77.

[4] T. Heath, B. Diniz, E. Carrera, W. Meira Jr, and R. Bianchini, "Energy conservation in heterogeneous server clusters," in *Proceedings of the tenth ACM SIGPLAN symposium on Principles and practice of parallel programming*. ACM, 2005, pp. 186–195.

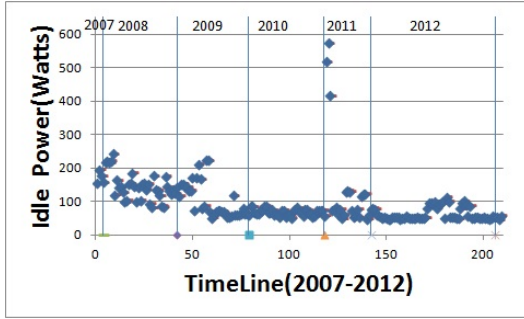
[5] X. Fan, W. Weber, and L. Barroso, "Power provisioning for a warehouse-sized computer," *ACM SIGARCH Computer Architecture News*, vol. 35, no. 2, pp. 13–23, 2007.

[6] "A software predict energy consumption," <http://www.ludashi.com/>.

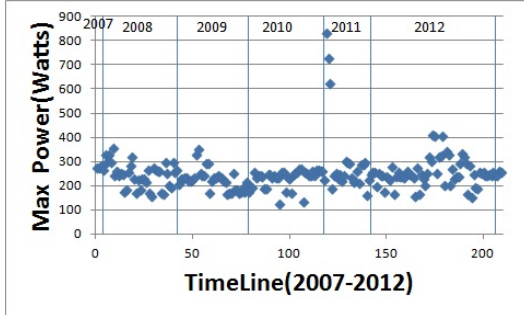
[7] "Thermal design power list of intel xeon processors," [http://en.wikipedia.org/wiki/List\\_of\\_Intel\\_Xeon\\_microprocessors](http://en.wikipedia.org/wiki/List_of_Intel_Xeon_microprocessors).

[8] W. Bircher and L. John, "Complete system power estimation: A trickle-down approach based on performance events," in *IEEE International Symposium on Performance Analysis of Systems & Software, 2007. ISPASS 2007*. IEEE, 2007, pp. 158–168.

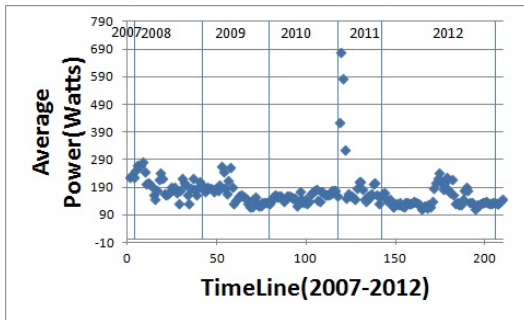
[9] Bircher, W.L. and John, L.K., "Complete system power estimation using processor performance events," *IEEE Transactions on Computers*, vol. 61, no. 4, pp. 563–577, 2012.



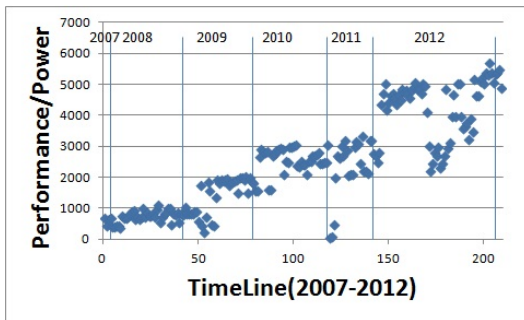
(a) idle power at different years



(b) max power at different years



(c) average power at different years



(d) performance/power at different years

Figure 7. Server Power Consumption in Different Years

- [10] K. Singh, M. Bhaduria, and S. McKee, "Real time power estimation and thread scheduling via performance counters," *ACM SIGARCH Computer Architecture News*, vol. 37, no. 2, pp. 46–55, 2009.
- [11] G. Contreras and M. Martonosi, "Power prediction for intel xscale® processors using performance monitoring unit events," in *Low Power Electronics and Design, 2005. ISLPED'05. Proceedings of the 2005 International Symposium on*. IEEE, 2005, pp. 221–226.
- [12] D. Snowdon12, E. Le Sueur, S. Petters, and G. Heiser, "A platform for os-level power management," in *The European Professional Society on Computer Systems 2009*. ACM, 2009.
- [13] A. Bohra and V. Chaudhary, "Vmeter: Power modelling for virtualized clouds," in *Workshops and Phd Forum (IPDPSW), 2010 IEEE International Symposium on Parallel & Distributed Processing*. Ieee, 2010, pp. 1–8.
- [14] I. Intel, "Intel 64 and ia-32 architectures software developer's manual volume 3b: System programming guide," *Part*, vol. 2, pp. 14–19, 2013.
- [15] N. NVIDIA, "Nvml api reference manual," vol. Version 3.295.45, p. 46, 2012.
- [16] E. Rotem, A. Naveh, D. Rajwan, A. Ananthkrishnan, and E. Weissmann, "Power-management architecture of the intel microarchitecture code-named sandy bridge," *Micro, IEEE*, vol. 32, no. 2, pp. 20–27, 2012.
- [17] "Standard performance evaluation corporation," <http://www.spec.org>.
- [18] D. Economou, S. Rivoire, C. Kozyrakis, and P. Ranganathan, "Full-system power analysis and modeling for server environments," in *Proceedings of Workshop on Modeling, Benchmarking, and Simulation*. IEEE, 2006, pp. 70–77.
- [19] D. Tsirogiannis, S. Harizopoulos, and M. Shah, "Analyzing the energy efficiency of a database server," in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*. ACM, 2010, pp. 231–242.
- [20] "Linux hardware performance counter collection," <http://sourceforge.net/projects/perfctr/>.
- [21] "Linux performance counter collection-sar," <http://www.linuxjournal.com/content/sysadmins-toolbox-sar>.
- [22] "Ganglia monitoring system," <http://ganglia.sourceforge.net/>.