BFEPM:Best Fit Energy Prediction Modeling Based on CPU Utilization

Xiao Zhang*, Jianjun Lu*, Xiao Qin[†]

*School of Computer Science Northwestern Polytechnical University 127 Youyi xi Road Xi'an Shaanxi China
Email: zhangxiao@nwpu.edu.cn lujianjun@mail.nwpu.edu.cn
[†]Department of Computer Science and Software Engineering Auburn University Auburn, AL USA 36849-5347
Email: xqin@auburn.edu

Abstract—**Energy cost becomes a major part of data center operational cost. Computer system consume more power when it runs under high workload. Many past studies focused on how to predict power consumption by performance counters. Some models retrieve performance counters from chips. Some models query performance counters from OS. Most of these researches were verified on several machines and claimed their models were accurate under the test. We found different servers have different energy consumption characters even with same CPU. In this paper, we present BFEPM, a best fit energy prediction model. It choose best model based on the power consumption benchmark result. We illustrate how to use benchmark result to find a best fit model. Then we validate the viability and effectiveness of model on all published results. At last, we apply the best fit model on two different machines to estimate the real-time energy consumption. The results show our model can get better results than single model.**

Keywords-**energy consumption model; green computing; data center management;**

I. INTRODUCTION

Energy cost becomes a major part of data center operational cost. Computer system consume more power when it runs under high workload. Power consumption model can be used to estimate the energy cost and do job scheduling between different servers. Current power consumption model needs many performance counters and can not be applied in heterogeneous environments. An inaccurate model will misguide the power and thermal management. In this paper, we present a flexible power estimate model BFEPM, it includes a series equations and selects best fit model for given server. We found that different servers has different power consumption characters even with same CPU. One single static model can not get accurate value for different servers. We show how to select best fit model for each server, then apply the models to analyze the published power benchmark results. We also apply the model to estimate the real-time energy consumption under complex workloads.

Four factors make our model desirable and possible:

1) The growing importance of energy saving
2) Thermal management modules make use of power models to predict cooling cost.
3) A common model is necessary for heterogeneous environments.
4) The needs for estimate the cost of each job or instance.

The power requirements of modern data centers range from 75 W/ft^2 to 150-200 W/ft^2 and will increase to 200-300 W/ft^2 in the nearest future[1]. Energy cost becomes a major part of data center operational cost. To reduce the energy cost of data center, many techniques were developed focus on saving energy consumption and cooling systems(e.g., workload consolidation, living migration, CPU throttling solutions).

Computer system consume more energy when utilization grow. While it consumes than 46% the maximum power in idle state[2]. Many techniques try to migrate jobs to make some physical machines run at high utilization and turn off idle machines. Migration policies depend on utilization and power consumption of each physical machine. A prerequisite of energy-aware scheduling is precise knowledge of any activity inside the computer system. Energy consumption model is focus on relation between resource utilization and power consumption. By Using energy consumption model, managers can predict energy consumption before and after job migration. There are many performance event counter can be used to model energy consumption(e.g., CPU, IO, Memory). Past studies show a strong correlation between performance events and power consumptions.

Different servers have different performance counters and power consumption characters. There are many developed models to estimate different servers' power consumption. Some models retrieve performance counters from chips. Some models query performance counters from OS. Most of these researches were verified on several machines and claimed their models were accurate under the test. According to our results, a single static model can not adopt to any servers. It is also inconvenient to develop models for different kinds of servers in heterogeneous data center.

In a heterogeneous data center, A job need different time and resource to finish on different servers. High-performance server can finish the job in a short time, but the power during running is high. While other server needs more time, but the power consumption is lower than high-performance server. Using power estimate model, manager can calculate the power cost of each job running on different machine. It is important to get the accurate cost in "'pay as you go'" mode. Another interesting thing is that some servers energy consumption is not linear mode. Which means cost

of running a job when server is idle is different with when the server is under 50% utilization.

In this paper, we analyzed the performance counters which energy consumption models used. We use published energy consumptions to validate the viability and effectiveness of model based on CPU utilization. The results show that the model does not always match the real energy consumption. At last, we presented results of comparison of modeled and measured energy consumption on two different machines. The results show the error of modeled energy consumption was different on different machines. The main contributions of our work are:

1) We present a best fit energy consumption model. The model use different equations to describe the character of different servers. It can get better prediction result for different server than linear model.
2) We validate the energy-consumption model by the published SPECpower results. Which includes 392 different servers. The results show our model can be used in heterogeneous environments.
3) We applied energy-consumption model on different servers and illustrate the real-time energy consumption. The results show our model can get precise result under stable state.

The rest of this paper is organized as follows. Section II introduces the existing energy consumption models and the SPECPower benchmark. Section III presents the method of how to choose best fit model for a given server. Section sec:analyze validates the accuracy through the published SPECpower results. Section V shows experimental and validation when apply the model to calculate the real-time power consumption. The last section discusses the future work of this study.

II. RELATED WORK

A. Energy consumption 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 [3]. To schedule jobs or workload consolidation in energy-efficient way, one must estimate the energy consumed by each computer node in a data center. Power meters can measure the electric power in watts of each computers. A modern digital power meter samples the voltage and current several times a second. It can get accurate real-time power consumption information.

Rather than relying on expansive power meter, many researchers find power consumption has related with performance counters. Trickle-down power modeling provides an accurate representation of complete-system power consumption[4]. Bellosa et. al use information about active hardware units (e.g., integer/floating-point unit,

cache/memory interface) gathered by event counters to establish a thread-specific energy accounting. They modified kernel data structures to hold the values of event counters of Pentium II[5]. Wattch is a framework for analyzing and optimizing microprocessor power dissipation at the architecture-level[6]. Isci developed a power model for Pentium IV using activity factors and functional unit area, similar to Wattch[7]. Singh *et al.* built some model for AMD Phenom [8].

Bricher *et al.* in University of Texas focused on modeling from CPU in recent years. They presented a simple linear model for the Pentium IV based on the number of instructions fetched/cycle in 2005[9]. They 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 in 2007[10]. In 2012, 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[2]. Different CPUs have different architecture and instruction set. The parameters and access methods are different on different CPU chips. They uses different counters even for same thing. For example, when the CPU is idle, it saves power by changing into halt state. Time under halt state can be used to calculate the energy consumption. Pentium IV processor count this by Halted Cycles. AMD CPU count this by %Halted/%Not-Halted. This is obvious in different CPU family. Pentium IV processor has 18 performance counters that can be programmed to monitor up to 59 event classesbut there are only 15 events available for Intel XScale CPU PXA255[11]. Some other studies also consider how to predict energy consumption of multi-core processors[12].

The above existing hardware performance counter solutions have the following drawbacks. First, Most of the models were tailored for special processors or computer architectures. Second it is difficult to sample counters from CPU chips. Early studies even need modify kenerl data structure. Software access to the performance counters is provided by the Linux perfctr device driver[1]. Finally, 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 system-level counters. CPU consumes most part of energy. Clock gating feature implemented in the microprocessor. I/O, disk and network subsystem consume nearly constant when in the idle and peak state. Fan *et al.* implemented models based on CPU utilization, and estimate energy consumption of each Rack,

---

[1]http://sourceforge.net/projects/perfctr/

PDU and cluster. The error stays below 1% across the usage spectrum and over a large number of PDU-level validation experiments. They focus on critical power and power usage of entire clusters with several thousand servers[3]. Mantis is a nonintrusive method for modeling full-system power consumption and providing real-time power prediction. Mantis uses CPU, memory, disk and net utilization to model power consumption. The errors range from 0% 15% on Intel CPU[13]. Heath *et al.* extended the models by using OS-reported CPU, memory and disk utilization[14].

Some other study try to find energy consumption of each process or each Virtual Machines. Snowdon *et al.* discussed approaches to monitoring power for applications [15]. 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 [16]. 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 [8].

Models based on system-level can be adopted to heterogeneous cluster systems. Most past studies show that the linear model is effective in analyzing performance counters and energy consumption. The difference of these models lies in kinds and numbers of performance counters. Most of previous models were only validated on several different kinds of server. Fan's models were validated on several thousand servers in Google, they does 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[14].

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
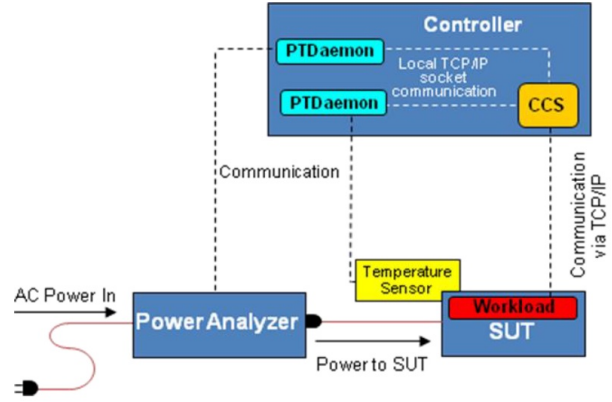

Figure 1. Typical test environment of SPECPower_ssj2008

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 1 shows testbed that makes use of the SPECPower benchmark to evaluate energy efficiency of computing systems[2].

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 [10]. CPU activities are recorded when SPECpower is running a testbed; which keeps track of CPU usage, operation per second, and power consumption.

III. CONSTRUCTION OF ENERGY MODEL

Early energy consumption model use CPU utilization as only parameter[7]. Some studies try to monitor several performance counters related with CPU to estimate the power consumption[4]. Past studies use multiple performance counters to calculate the energy consumption,including CPU, memory, disks and network[13], [14]. They extended the model with more parameters to get more accurate power consumptions. But data centers in Google also use CPU utilization to estimate the power usage[3].

There are several advantage 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. Our findings confirm that the energy consumption of the I/O devices including the monitor do not noticeably change.

[2]http://www.spec.org/power_ssj2008/

43

Second, it will cause performance slowdown to capture and handle multiple performance events, especially sampling at a small interval.

Our model calculate the power consumption only using CPU utilization. As shown in Section IV-A, the power consumption behaviors of different servers are different even they have same CPU and similar performance. We simplify the current power calculate model and use CPU utilization as the only parameter. Then we extend the model to adopt to different conditions. Finally we give the method how to choose best fit equation to calculate the power consumption.

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:

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

Different Servers have different performance characters. Power consumption can be viewed as function of CPU utilization as Equation 2. Linear model was widely used in past studies[3]. We extend the model to higher degree polynomial(up to 5th), power and Gaussian. A piecewise-defined function is used to calculate real-time energy consumption. We start with linear model equation 3a. Then we increase the order of the equation to a quadratic and cubic polynomial3. We can also extend the polynomial to higher degree.

We use R-square, Adjusted R-square and RMSE to choose which is best for certain server. We also validate the accuracy of our energy model using (Euation 5), which calculates the average error for each combination of modeled power consumption and measured consumption. There are 392 published results of SPECPower until Oct.2012. Each of published results consists 11 pairs of workload and measured power consumption. The 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. Curve fitting can be used to construct a mathematical function from given data. We illustrate how to choose best fit model for a given server and analyze all the published results.

$$P_{total} = f(R_{cpu}) \quad (2)$$

$$P_{total} = a + b \times R_{cpu} \quad (3a)$$
$$P_{total} = a + b \times R_{cpu} + c \times R_{cpu}^2 \quad (3b)$$
$$P_{total} = a + b \times R_{cpu} + c \times R_{cpu}^2 + d \times R_{cpu}^3 \quad (3c)$$
$$P_{total} = a \times R_{cpu}^b + c \quad (3d)$$
$$P_{total} = a1 \times e^{-((R_{cpu}-b1)/c1)^2} \quad (3e)$$

We also created a piecewise-defined function to describe the energy consumption of the server. Each result data set contains 11 data points. R_{cpu}, R_{mix}, and P_{total} are CPU utilization, ssj operations per second, and measured total power. We use P_i and R_i to present the total power and CPU utilization. Where i range from 1 to 11. The piecewise-defined function can not validate by published results. If none of above model can get a Adjusted R-square above 0.97, the piecewise-defined model is best for real-time energy estimation. We will validate the accuracy under real-time energy estimation.

$$P_{total} = \begin{cases} P_i : R_{cpu} = R_i \\ P_i + (P_{i-1} - P_i) \times \frac{(R_{cpu}-R_i)}{(R_{i-1}-R_i)} \\ : R_{i-1} > R_{cpu} > R_i \end{cases} \quad (4)$$

$$AverageError = \frac{\sum_{i=1}^{NumSamples} \frac{|Modeled_i - Measured_i|}{Measured_i} \times 100\%}{NumSamples} \quad (5)$$

We illustrate how to validate energy consumption model using the benchmarking results obtained from the latest study on Dell servers (i.e., PowerEdge C6145[3]). Table I shows the benchmarking data of the Dell servers. Recall that each result data set contains 11 data points. R_{cpu}, R_{mix}, and P_{total} are CPU utilization, ssj operations per second, and measured total power, respectively. We can obtain the best fit parameter of P_{const}, P_{cpu}, and P_{ssj_ops}. Using the curve fitting function, we derive the following power consumption model:

$$P_{total} = 244.1118 + 8.8858 \times R_{cpu}$$
$$P_{total} = 223.2672 + 10.2783 \times R_{cpu} + -0.0139 \times R_{cpu}^2$$
$$P_{total} = 220.1039 + 10.7824 \times R_{cpu} + -0.0272 \times R_{cpu}^2$$
$$+0.0001 \times R_{cpu}^3$$
$$P_{total} = 16.04 \times R_{cpu}^{0.8755} + 209.5 Note^4$$
$$P_{total} = 1134 \times e^{-((R_{cpu}-118.2)/97.41)^2}$$

We place the energy-consumption modeling results on the last three columns. We apply equation 5 to compute the

No	CPU	Measured Power	Modeled Power				
			1st	2nd	3rd	Power	Gaussian
1	99.7	1,105	1130.0	1109.8	1123.8	1111.2	1093.8
2	90.1	1,038	1044.7	1036.5	1043.9	1034.7	1043.5
3	80.0	962	955.0	956.6	959.8	953.1	972.4
4	69.9	883	865.2	873.8	875.0	870.3	886.8
5	60.0	790	777.3	789.9	790.7	787.6	793.6
6	50.0	690	688.4	702.4	703.7	702.3	694.6
7	39.9	595	598.7	611.2	613.4	613.9	594.3
8	29.9	510	509.8	518.2	520.9	523.7	498.6
9	20.0	437	421.8	423.3	425.7	430.4	410.4
10	10.0	362	333.0	324.7	325.3	329.9	330.2
11	0.0	196	244.1	223.3	220.1	209.5	260.1

Model	SSE	R-square	Adjusted R-square	RMSE
1 degree	18620	0.9761	0.9735	45.48
2 degree	3840	0.9951	0.9938	21.91
3 degree	3152	0.996	0.9942	21.22
4 degree	3150	0.996	0.9933	22.91
Power	2247	0.9974	0.9968	16.76
Gaussian	6267	0.9928	0.991	27.99

mean error between the modeling and measured results. The mean error of this set of tests is 4.02%(1st) 3.25%(2nd) ,3.25%(3rd), 2.59%(Power) and 4.92%(Gaussian).

To find which is best model for the server, we listed the goodness of fit by each model in II. SSE means the sum of squares due to error. This statistic measures the deviation of the responses from the fitted values of the responses. A value closer to 0 indicates a better fit. R-square refers to the coefficient of multiple determination. This statistic measures how successful the fit is in explaining the variation of the data. A value closer to 1 indicates a better fit. Adjusted R-square is the degree of freedom adjusted R-square. A value closer to 1 indicates a better fit. It is generally the best indicator of the fit quality when add additional coefficients to the fit model. RMSE is the root mean squared error. A value closer to 0 indicate a better fit. From the table, we can find that cubic polynomial is best fit model for the tested server. In the same way, we can find best model for each tested server.

Power model is the best fit for PowerEdge C6145. But it is not the best for all the servers. For example, 5th degree polynomial model is the best fit model for IBM System x3550 M4 [5](the Adjusted R-square of 0.9898(linear),0.9904(quadratic),cubic(0.9909),4th(0.9959),5th(0.9998), power(0.9877) and gaussian(0.9891)). In the next subsection, we describe the analysis of all the SPECpower published results.

[5]http://www.spec.org/power_ssj2008/results/res2012q3/power_ssj2008-20120807-00523.html
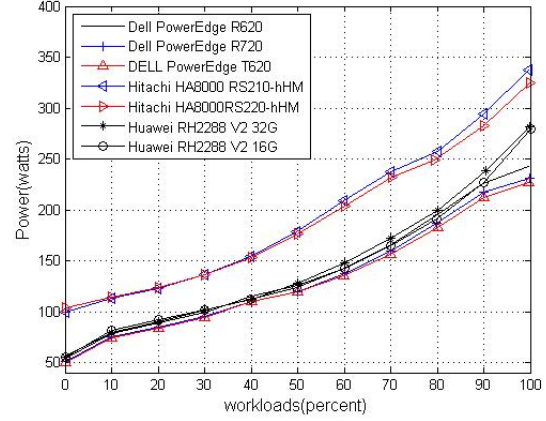

Figure 2. Power Consumption of Different Servers with same CPU

## IV. ANALYSIS OF BENCHMARK RESULTS

A. Compare 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 III 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 consumption 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 2).

| 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	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	325	190.9
Huawei	RH2288 V2 32G	32	55.6	282	146.1
Huawei	RH2288 V2 16G	16	54.8	279	142.9

B. Average error between estimated and measured data

We test 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
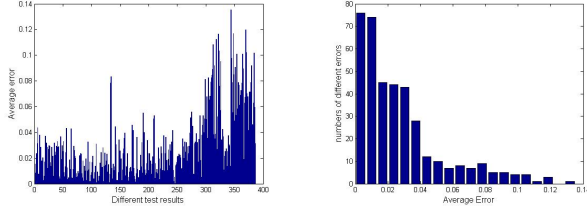
(a) Mean error rate of different re-
sults

(b) Mean error distribution

Figure 3.   Mean Error of degree 1

(a) Mean error rate of different re-
sults

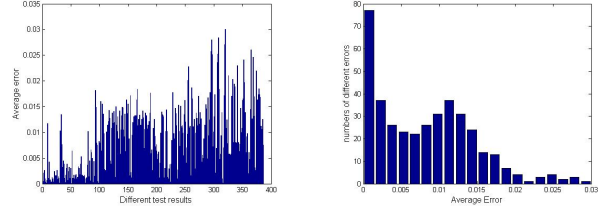(b) Mean error distribution

Figure 5.   Mean Error of degree 3
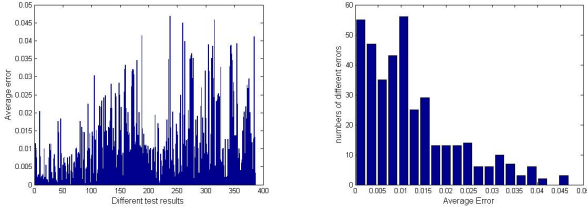
(a) Mean error rate of different re-
sults

(b) Mean error distribution

Figure 4.   Mean Error of degree 2

Figure 6.   Mean Error CDF of Different Models

degree 1,2,3 that fits the data measured power best in a least-
squares sense. The average error for each pair of measured
power consumption and modeling result is calculated by
Equation (5). We verify all the published results of different
servers.

Figure 3 shows that most of the mean errors of polynomial
of degree 1 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.

Figure 4 shows that all of the mean errors of polynomial
of degree 2 are below 5%. Most of the mean error are below
2%. While all the mean errors of polynomial of degree 3 are
below 3%(Figure 5). Most of the mean error are below 1.5%.
We use cumulative distribution function($F_x(x) = P(X \leq x)$) to show the distribution of each model in Figure 6. Our
results show that Equation 3c is best for predict the power
usage by CPU utilization. In addition to the SPECpower
benchmark, a couple of real-world applications were used
to validate the energy consumption model.

V. REAL-TIME ENERGY CONSUMPTION

In addition to SPECpower results, real-world power es-
timation does not limit to several points. The power con-
sumption model developed in Section III can derive from
the results of SPECpower. In this section, we apply the
model to estimate the energy consumption of servers in
a real-time manner. We validate the model to estimate
the real-time energy consumption during SPECpower test
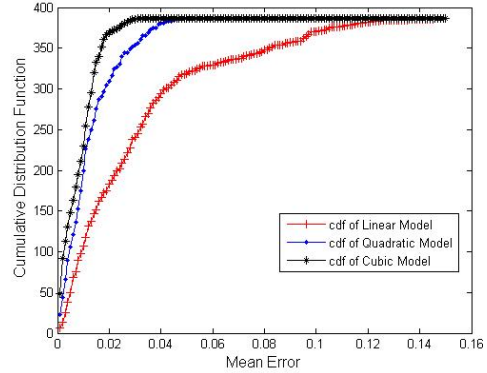
processing. SPECpower only focus on CPU and memory,
it does not generate too much IO and network workloads.
We also validate the accuracy of model under complex
workloads including IO.

A. Test Environments

In our experiments, we collect performance counters
from SAR and other utility programs in Linux. Sampling
application also needs CPU and other resource. We collect
performance counter at 10 seconds interval, it only use less
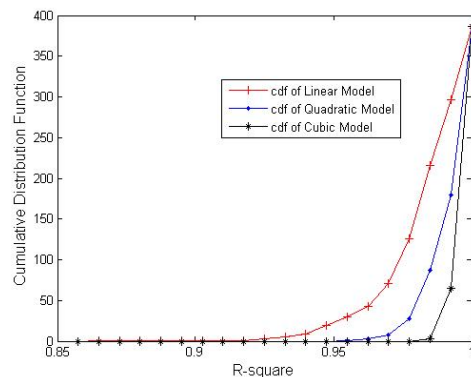than 2% CPU at the sampling time. We use a Chroma 66202

Figure 7.   R-square CDF of Different Models

Table IV
CONFIGRATUIONS OF TESTBED

Name	Hardware	Software
Server1 Inspur AS300N	1*Intel(R) Xeon(R) CPU E5502 1.87GHz(4 core) 1*16GBytes of RAM 4*160 GBytes SATA disk (Hitachi HTS545016B9A300)	Centos5.3 Linux kernel 2.6.18
Server2 Dawning A840r-G	AMD Opteron(TM) 2.3GHz 12 core 1*64GBytes of RAM 4*300GBytes SAS disk (Hitachi)	Redhat6.2 Linux kernel 2.6.32

power meter to measure the total system power. The Chroma power meter - providing readings every 0.25~2 second - can measure power ranging from 1.5w to 1000w. Power meter does not use any CPU and other resource of system under test. We collect power consumption every second. The first server system is a highly integrated blade server that includes an Intel(R) Xeon(R) processor. The second system is a high-end server containing four AMD Opteron(TM) processes. The last one is equipped with an AMD dual core processor. Table IV summarizes the testbed used in these experiments.

B. Validation of Real-time SPECpower Power

We run SPECpower on servers listed in Table IV. We use the results of SPECpower to choose the best fit model for server1 and server2. Linear model is best fit for server1 and cubic model is best for server2. We monitor the workload and power consumption during the entire testing process. We use the chosen best fit model to estimate the energy consumption during the test in a real-time manner. Figure 8 illustrates the measured and estimated power consumption of server1 and Figure 9 is for server2.

From Figure 8, we can find linear model get accurate estimate power consumption for server1. The error between modeled and measured power is between -5% and 3%. And the peak error appear when SPECpower switch to next stage, while the CPU utilization change dramatically. Most value is between +-2%. Fan *et al.* also claimed that linear model can estimate power consumption while error under 1% in Google data center[3].

Power estimate model for server2 does not fit measured power well as server1. The max error range from -50% to 179%. The peak error also occurs at switch phase. 79.13% modeled power value is between +-5%. This result also show that the model can be used to estimate energy consumption in real world.

Figure 10 show the all measured and modeled power of different CPU utilization. Cubic model is best fit model for all data. But quadratic model does better than cubic model when CPU utilization is between 70% to 90%. There are some points show it use extreme low power under different workloads. And there are also some points show that it use high power even CPU utilization is low. We think these are caused by sampling precision. We sampling CPU utilization every 10 seconds and power consumption every 1 second.
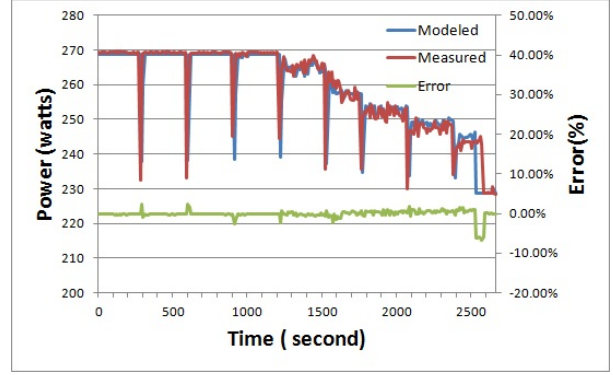

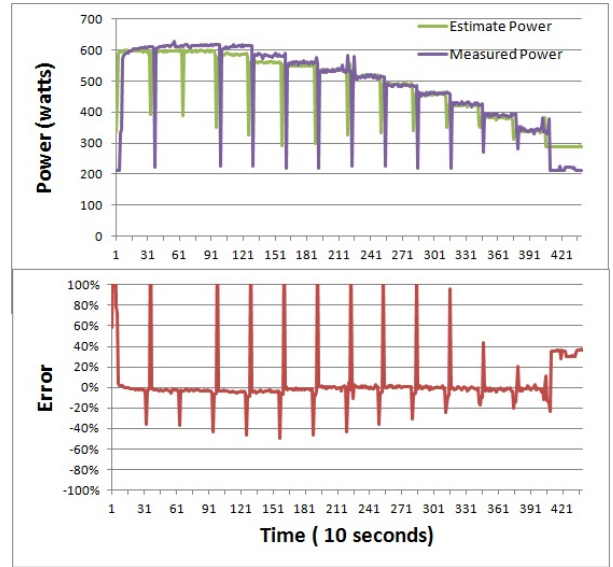Figure 8.   Estimate Power and Measured Power for Server1

Figure 9.   Estimate Power and Measured Power for Server2

If we sampling CPU and power at very small interval, it will decrease the deviation. Or if the CPU utilization keeps constant, it will get good modeled value. It is difficult to synchronize the time less than 1 second. If CPU utilization and power meter query data at different microsecond, it will cause the deviation. Another interesting thing is that none of the model can match the power consumption of idle. These cause the error of idle is about 30% after 400 in Figure 9. For this server, piecewise-defined function is better to estimate power consumption under idle state.

C. Validation of Real-time Complex Workload Power

In addition to SPECpower, real-world applications (e.g., the gcc compiler and a Linux utility program) are running on the heterogeneous servers to validate our model. Each application represents a specific workload scenario. We collect CPU, memory, disk, and network utilization every 10 seconds. The evaluated workloads are CPU-bound; for
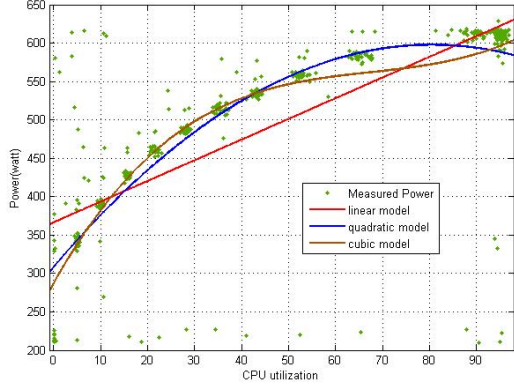
Figure 10. Curve Fitting Result of Models

(a) Compile Linux kernel

(b) find keyword in Linux kernel source codes

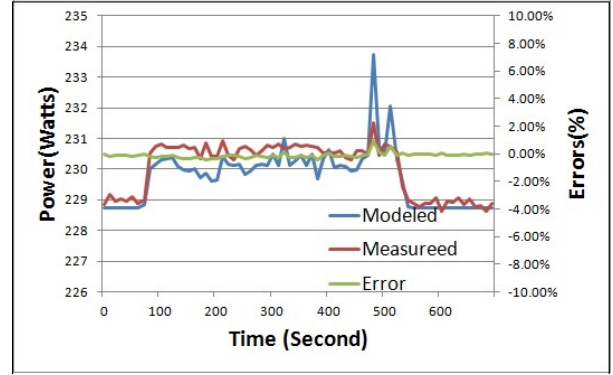Figure 11. Estimate Power and Measured Power with disk IO

example, the CPU utilization varies from 100% to 10% during a period of 40 minutes. We compile Linux kernel (3.2.32); the entire compilation time is 55 minutes. The workload of this compilation process is write-intensive (i.e., more writes than reads are issues to disks). 97% of the CPU utilization values are in the range between 23.12% and 25.52%. We use Find utility program to search keyword in the Linux kernel source codes. The utility program reads files in 12 minutes. This non-CPU-intensive workload has a high read I/O load.

A recent study shows that disk I/Os have little impact on the total energy consumption [10]. Figure 11(a) and 11(b) plot the real-time results of power consumption of server1. We observe that the estimated energy consumption is slightly higher than the measured data. The error between estimated and modeling results is less than 3.5% in the case of the gcc compiler and less than 1% in the case of the Find utility program.
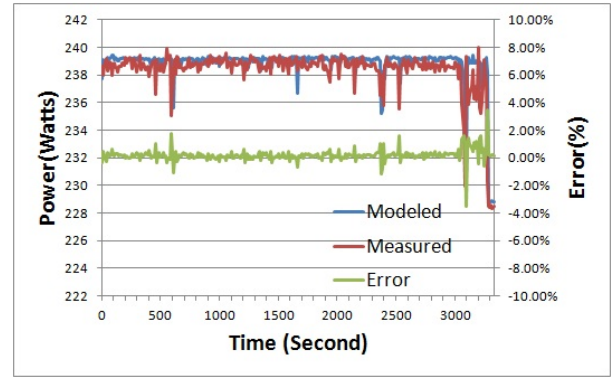
VI. Conclusions

In this paper, we present a method of modeling power consumption based on SPECpower benchmark result. Different servers have different power consumption characters even they have same CPU and chipset. We use several model to fit the SPECpower results to choose which fit the character best. We also validate the method on the results of all published SPECpower. The results show cubic polynomial fit better on most servers than linear model. Which is the most popular used model in past studies. We also validate the model in real-time manner. The results show that it can get a precise estimate results under stable state. This method improved the energy modeling method, it can get better results than current energy prediction model.

Some subsystems in a server automatically reduces their power usage after sitting idle for a period of time. Energy consumption depends on both workloads as well as previous and current power states(e.g. Fan). In our experiments, all cores in CPU has similar CPU utilization, it needs new model for multi-core systems. Our model can be used to schedule tasks among different servers, a job may need different time in different servers and generate different CPU utilization. Which means the cost of running job on different server is different. We will continue to present energy-aware job schedule based on our model.

REFERENCES

[1] "Data center needs for a sustainable hpc center: Issues facing the ohio supercomputer center," http://www.cac.cornell.edu/~lifka/Downloads/SRCC/24. 100315NSFSustainableCenters.pdf.

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

[3] 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.

[4] 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.

[5] F. Bellosa, "The benefits of event: driven energy accounting in power-sensitive systems," in *Proceedings of the 9th workshop on ACM SIGOPS European workshop: beyond the PC: new challenges for the operating system*. ACM, 2000, pp. 37–42.

[6] D. Brooks, V. Tiwari, and M. Martonosi, "Wattch: a framework for architectural-level power analysis and optimizations," in *ACM SIGARCH Computer Architecture News*, vol. 28, no. 2. ACM, 2000, pp. 83–94.

[7] 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.

[8] K. Singh, M. Bhadauria, 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.

[9] W. L. Bircher, M. Valluri, J. Law, and L. K. John, "Runtime identification of microprocessor energy saving opportunities," in *Low Power Electronics and Design, 2005. ISLPED'05. Proceedings of the 2005 International Symposium on*. IEEE, 2005, pp. 275–280.

[10] 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.

[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] I. Ahmad, S. Ranka, and S. U. Khan, "Using game theory for scheduling tasks on multi-core processors for simultaneous optimization of performance and energy," in *Parallel and Distributed Processing, 2008. IPDPS 2008. IEEE International Symposium on*. IEEE, 2008, pp. 1–6.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] "Standard performance evaluation corporation," http://www. spec.org.