Performance and Power Consumption Measurement of Java Application Servers

Hitoshi Oi and Sho Niboshi
The University of Aizu, Aizu-Wakamatsu, JAPAN
Email: {oi, nibo}@oslab.biz

***Abstract*—In this paper, we present our in-progress project of modeling performance and power consumption of Java application servers using SPECjEnterprise2010. We run the workload on two application server using two different CPUs, AMD Phenom II and Intel Atom, and investigate performance and power consumption behaviors against the increasing system sizes. We have observed that: (1) CPU utilization draws non-linear functions of the system size and their shapes are different on Phenom and Atom. However, power consumption on both servers increase proportionally. (2) Browse transaction is the source of non-linearly in the CPU utilization. (3) Estimation of the CPU utilization from that of each transaction measured separately incurs large errors (up to 65%), while the errors in the estimation of the power consumption are relatively small (up to 4%).**

I. Introduction

In this paper, we present our in-progress project of modeling performance and power consumption analysis of Java application servers using SPECjEnterprise2010 [1] (jEnt10 hereafter where appropriate) on application servers with two different CPUs. jEnt10 is an industry standard benchmark suite for Java EE application servers published by the Standard Performance Evaluation Corp. (SPEC). It is modeled after the dealers and manufacturing operations in the automobile industry and handles five transaction types within two application domains. The Dealer domain handles Browse (browsing the automobile catalog), Manage (manage the dealers' inventories) and Purchase (purchase automobiles) transactions. These transactions are received through the web interface of the system under test, which consists of the application (App) and database (DB) servers. Two other transaction types are within the Manufacturing domain: CreateVehicleEJB and CreateVehicleWS. Both transaction types emulate the periodic creation of vehicles in small quantities (14 on average). While the former transactions are received by the EJB, the latter are received via the Web Service. These transactions are issued by the load generator (included in the benchmark suite) at the rates proportional to the system size (denoted as Scaling Factor, or SF, in this paper), and the throughput of these transactions determines the performance metric of the system. For the performance metrics to be valid, there are various quality of service (QoS) metrics defined including response time of each transaction.

II. Measurement

For the App servers, we use two platforms based on AMD Phenom II [2] and Intel Atom [3] CPUs. Table I shows the specifications of these App servers and their CPUs. The DB server has a quad-core Xeon (X3210) and 6GB of memory running MySQL 5.5.13 on Oracle Linux 6.1. For the measurement of the power consumption, Watts up ? Pro 99333 power meter [4] is used.

CPU	Phenom II X6 1065T	Atom D525
Clock Freq. TDP	2.9GHz 95 W	1.8GHz 13 W
Memory Hierarchy		
L1 (/core) L2 (/core) L3 (shared) Memory	64KB (I) 64KB (D) 512KB 6MB 16GB (PC3-10600)	32KB (I) 24KB (D) 512KB None 4GB (PC3-8500)
Inst. Issue Execution Pipes # Cores/SMT	out-of-order, 3 x86 inst 3 ALU, 3 AGU 3 FP 6-Core	in-order 2 x86 inst. 2 ALU/AGU, 2 FP 2-Core/HyperThreading
Server Software OS	Glassfish v3.0.1 Oracle Linux 6.1	

TABLE I
Application Server CPUs and Software Specifications

A. Phenom Performance and Power Consumption

Fig. 1 shows the CPU utilization and power consumption of the Phenom based server. For $SF \leq 177$, all transactions met the QoS metrics. In this configuration, the DB server was the bottleneck and the CPU utilization of the App server was significantly lower than 100% at $SF = 177$. While the power consumption is almost linear to SF, the CPU utilization per SF decreases for larger SF.

Figs. 1 and 2 show the CPU utilization and power consumption of each transaction type measured separately. We fitted the CPU utilization and power consumption of each transaction (except Browse) into a linear function of SF. For the CPU utilization and power consumption of the Browse transaction, we added $\sqrt{(SF)}$ and SF^2, respectively. The CPU utilization of the Browse transaction has relatively large errors (maximum 4.2% at SF = 60), but in all other cases relative errors are smaller than 1%.

Using these approximate functions of individual transaction types, we estimated the CPU utilization and power consumption of the workload as a whole (Table II). The CPU utilization has large relative errors, up to 65% at $SF = 150$. From this observation, it is not feasible to estimate the total CPU utilization from those of individual transactions measured separately. On the other hand, the total power consumption can

1526-7539/12 $26.00 © 2012 IEEE
DOI 10.1109/MASCOTS.2012.68

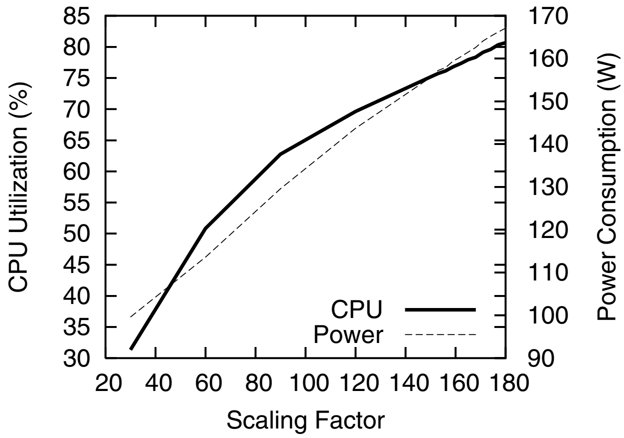

Fig. 1. Phenom Performance and Power Consumption

well be estimated from individual transactions as the errors are relatively small (up to 4% at $SF = 90$ and 120).

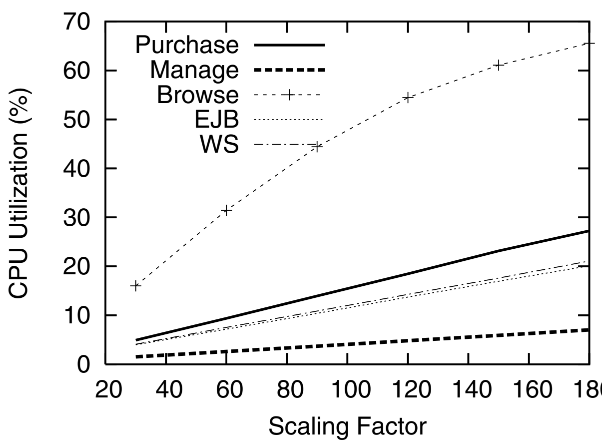

Fig. 2. Phenom CPU Utilization of Transaction

B. Atom Performance and Power Consumption

Fig. 4 shows the CPU utilization and power consumption of the Atom based server. For $SF \leq 27$, all QoS metrics are met. Unlike Fig. 1, the CPU utilization per SF increases for a larger SF.

Figs. 5 and 6 show the CPU utilization and power consumption of individual transactions, respectively. Again, the Browse transaction is the source of non-linearity in the CPU utilization, but it is not as obvious as the total CPU utilization. We fitted the CPU utilization and power consumption of each transaction to a linear function of the system size (SF), except the Browse transaction, to which SF^2 is added to compensate its higher increase rate than linear. These functions

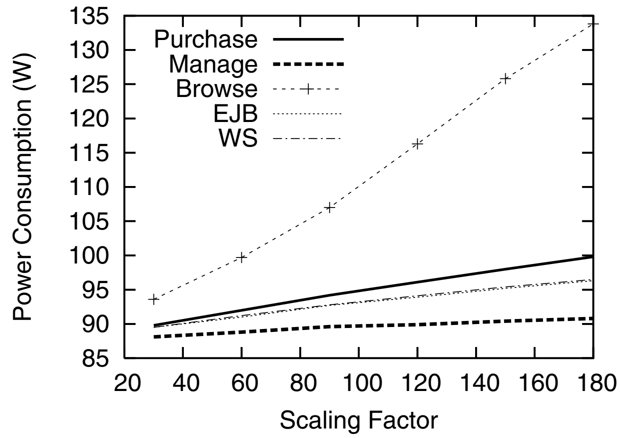

Fig. 3. Phenom Power Consumption of Transaction

SF	30	60	90	120	150
CPU Utilization (%)					
Mix.	31.4	50.9	62.8	69.6	75.2
Est.	30.0	59.5	83.5	104.7	124.0
Err. (%)	-4.3	17.0	33.0	50.3	64.9
Power Consumption (W)					
Mix.	99.6	113.6	129.6	143.7	155.6
Est.	99.4	111.7	124.6	138.1	152.3
Err. (%)	-0.2	-1.7	-3.9	-3.9	-2.1

TABLE II
ESTIMATION OF PHENOM APP SERVER CPU UTILIZATION AND POWER CONSUMPTION FROM INDIVIDUAL TRANSACTION. MIX. : NORMAL MEASUREMENT, EST. : ESTIMATED FROM INDIVIDUAL TRANSACTION TYPE MEASUREMENT, ERR. : RELATIVE ERROR.

approximate individual transactions well with the maximum error of 1.8% for Manage at $SF = 18$.

Next, as in the previous section, the total CPU and power consumption are estimated from these approximations of individual transactions. As presented in Table III, while the CPU utilization of each transaction can be approximated with a simple function well, that of total workload shows significantly large relative errors with the maximum of -25% at $SF = 27$. The difference with the case of Phenom is, this approximation underestimates the total CPU utilization for larger system sizes, while it was overestimated in the case of Phenom. On the contrary, the total power consumption, as in the case of the Phenom-based system, can be approximated with relatively small errors (up to 0.5% at $SF = 18$).

Table IV shows the relative performance of the Atom-based server against the Phenom using the CPU utilization normalized by the SF (at $SF = 150$ for Phenom and 27 for Atom). We expected that each transaction would behave on two CPUs differently and thus the relative performance of the Atom varies among transactions. However, the relative performance of Atom is almost constant among transaction types, around 35% of the Phenom. However, the relative performance for the mixed executions is significantly lower

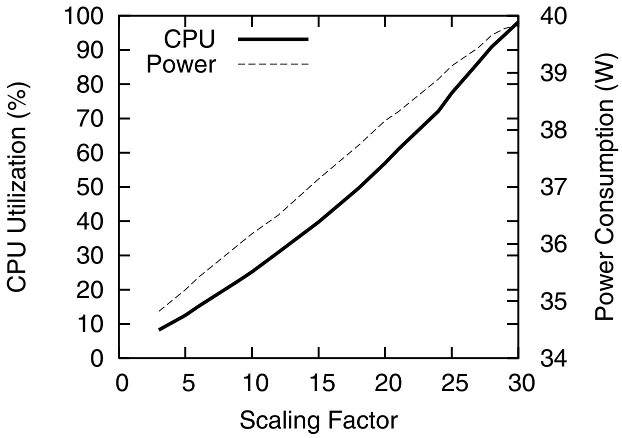

Fig. 4. Atom Performance and Power Consumption

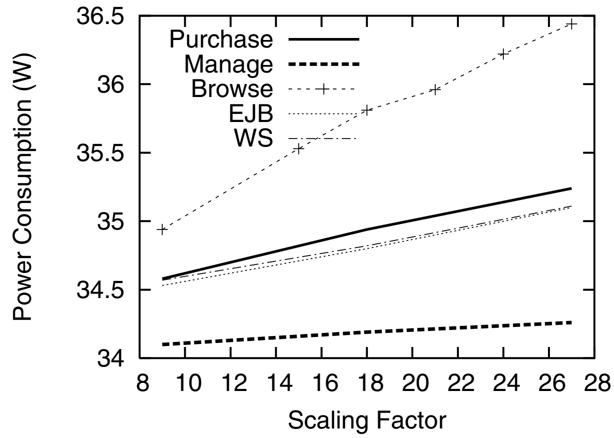

Fig. 6. Atom Power Consumption of Transaction

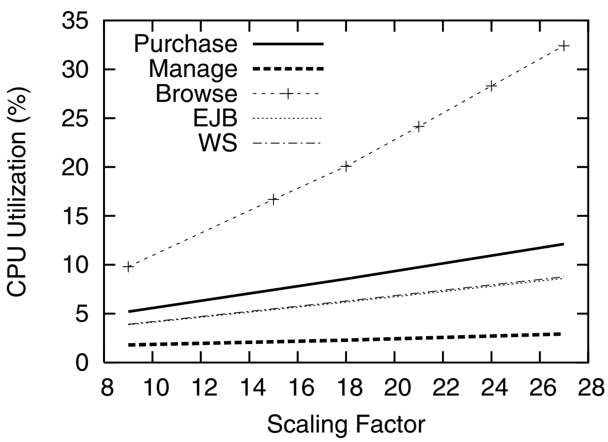

Fig. 5. Atom CPU Utilization of Transaction

SF	9	18	27
CPU Utilization (%)			
Mix.	22.6	49.6	86.3
Est.	24.5	43.7	64.8
Err. (%)	8.3	-11.9	-25.0
Power Consumption (W)			
Mix.	36.0	37.7	39.4
Est.	36.1	37.9	39.5
Err. (%)	.4	.5	.2

TABLE III
ESTIMATION OF ATOM APP SERVER CPU UTILIZATION AND POWER CONSUMPTION FROM INDIVIDUAL TRANSACTION.

than those of individual transactions. This should be due to the difference of their CPU utilization behaviors; when the system size is increased, CPU/SF of Phenom decreases but that of Atom increases. As seen in Tables II and III, separately measuring the CPU utilization of each transaction overestimates its contribution to the total CPU utilization when all transactions are executed together. This overestimate should also be another cause of the differences in the relative performance between the total and each transaction.

III. CONCLUSION AND FUTURE WORK

In this paper, we reported the results of performance and power consumption measurements of Java application servers using two different CPUs. It was found that CPU utilization exhibited non-linear scaling against the system size and the Browse transaction was the source of the non-linearity. Further investigations, especially in the non-linear scaling and the overhead of the CPU utilization in each transaction when executed separately, are required. Analysis of these issues should also help us identify various overhead of Java runtime systems [5].

[6] proposed that the operating system codes should be dispatched to simple cores on a heterogeneous multi-core system. We are aiming to develop a similar approach, based on the transaction types of the jEnt10, for the power-efficient execution of Java application workload.

REFERENCES

[1] SPECjEnterprise2010, http://www.spec.org/jEnterprise2010/ .

[2] "Software Optimization Guide for AMD Family 10h and 12h Processors," Rev. 3.13, Feb. 2011.

[3] "Intel ®64 and IA-32 Architectures Optimization Reference Manual,"

[4] Electronic Educational Devices, https://www.wattsupmeters.com/

[5] Nick Mitchell et. al, "Four Trends Leading to Java Runtime Bloat," in IEEE Software, January/February 2010, Vol. 27, No. 1. pp56–63.

[6] Rakesh Kumar et. al, "Single-ISA Heterogeneous Multi-Core Architectures for Multithreaded Workload Performance, " in *Proceedings of the 31st International Symposium on Computer Architecture*, pp64–75, June, 2004

Mix	Purchase	Manage	Browse	EJB	WS
15.7	34.4	36.6	33.9	35.5	36.1

TABLE IV
RELATIVE PERFORMANCE OF ATOM (%).