Energy Proportional Servers: Where Are We in 2016?

Congfeng Jiang, Yumei Wang, Dongyang Ou
School of Compute Science and Technology
Hangzhou Dianzi University
Hangzhou, Zhejiang, China
cjiang@hdu.edu.cn,{wyumei1,oudongyang1}@gmail.com

Bing Luo, Weisong Shi
Department of Computer Science
Wayne State University
Detroit, MI, USA
{bing,weisong}@wayne.edu

Abstract—**The huge energy consumption in data centers produces not only high electricity bill but also tremendous carbon footprints. Although today's servers and data centers of leading internet companies are more energy efficient than ever before, the fluctuations in external workload and internal resource utilization calls for energy proportional computing. Insight into server energy proportionality can help improve workload placement while also reducing energy consumption. In this paper, we investigate all 477 valid published results of SPECpower_ssj benchmark from 2007 to 2016Q3 and reorganize them by hardware availability year for more accurate analysis on production servers. Through comprehensive analysis we find that: (1) The specious stagnation of energy proportionality in recent years is mainly caused by the adoption of processors of specific microarchitecture and is not the indicative trend of energy proportionality improvement. (2) Microarchitecture evolution has more influence on energy efficiency improvement than energy proportionality. (3) Today's servers' peak energy efficiencies are shifting from 100% resource utilization to 80% or 70% utilization and server energy proportionality improves with such shifting. We then conduct extensive experiments on 4 rack servers to investigate the energy efficiency variations under different hardware configurations, including memory per core installation and processor frequency scaling. Our experiments show that hardware configuration has significant impact on server's energy efficiency. Our findings presented in this paper provide useful insights and guidance to system designers, as well as data center operators for energy proportionality aware workload placement and energy savings**.

Keywords—*Energy Proportionality; Energy Efficiency; Servers; SPECpower_ssj; Data Centers*

I. MOTIVATION

The ever-growing cloud-based services, big data analytics, e-commerce, and Internet traffic make the energy consumption of data centers grow faster [1, 2]. In 2007, the U.S. Environmental Protection Agency (EPA) reported to the Congress that the energy consumption of the nation's servers and data centers in 2006 was estimated to be more than double of the electricity that of 2000 and it would reach to 107.4 billion kilowatt-hours (kWh) in 2011 if current efficiency trend of that time would continue [3]. Fortunately, the prediction did not pan out in 2011 because of the significant energy efficiency improvements in servers. The Natural Resources Defense Council (NRDC) estimated that the energy consumption of data centers in U.S. was 76.4 billion kWh in 2011, and it will increase to 138 billion kWh by 2020 [4]. NRDC's estimation is closer to EPA's prediction based on improved operation scenario, which includes "*energy-efficiency improvements beyond current trends that are essentially operational in nature and require little or no capital investment*" [3], but much higher than EPA's optimistic prediction based on *best practice* scenario or *state-of-the-art* scenario. In June 2016, Lawrence Berkeley National Laboratory (LNBL) estimated that the energy consumption of data centers in U.S. is about 70 billion kWh in 2014, and it will increase slowly to 73 billion kWh by 2020 in a *current trends* scenario [5] assuming more increase of more energy efficient hyperscale data centers than traditional data centers and more business activities shifting from localized data centers to colocations or cloud facilities.

It's conspicuous that large scale data centers operated by leading companies like Facebook, Google, Amazon, and Microsoft are highly energy efficient [6, 7]. However, there are much more enterprises, small and medium-sized organizations have lower energy efficiency in their smaller scale data centers, which account for more than 95% percentage of total installed servers [4]. Moreover, the varying workload also results in fluctuation in resource utilization. This urges the calling for energy proportional computing which provides constant energy efficiency (we use EE for short in this paper) under all levels of utilization, including low to medium utilization regions [8, 9, 10, 11]. Usually a server with high peak energy efficiency is not essentially highly energy proportional since it may have low dynamic range or superlinear power-utilization curve. In addition, different hardware configuration may also affect the energy proportionality (we use EP for short in this paper) of the server [8, 12, 13, 14]. Therefore, good knowledge of server energy proportionality is vital for better workload placement and more energy savings in data centers.

In response to the increase of energy consumption in data centers, industrial standard organizations have developed benchmarks to evaluate server's energy efficiency. SPECpower_ssj2008 [15] (we use SPECpower or

SPECpower_ssj for short in the remainder of the paper) is developed and widely adopted to characterize a system's energy efficiency at varying utilization levels. Mainstream server vendors submit their SPECpower testing results to SPEC and the results are made available online after reviewing and auditing. Therefore the published SPECpower results are ideal sources to investigate the energy proportionality and energy efficiency improvements of production servers in data centers.

Among 517 submitted SPECpower results until Sept. 30, 2016(2016Q3), there are 40 results that are non-compliant with SPECpower discloser rules and published without any energy efficiency data. Among 477 valid results, there are 74 results whose published date is different from hardware availability date, which account for 15.5% of all valid results. In these cases, the hardware availability year can date back to 1~6 years before their published year, or 1 year after their published year. For example, a vendor submitted a result in 2011 (published year) reporting the testing result of its server made in 2006 (hardware availability year). In another case, a vendor submitted a result in 2015 (published year) whose hardware availability date is 2016 (hardware availability year). Such mismatch between published year and hardware availability year significantly affects the quality of temporal analysis, including data accuracy, time scale, data distribution, and subsequent statistics. Moreover, analysis based on published year may conceal the actual energy proportionality and energy efficiency improvements with times even if it may derive similar conclusions.

Although some researchers conducted analysis on server energy proportionality and energy efficiency from published SPECpower results [16, 17, 18], comprehensive study on server energy proportionality and efficiency is currently missing. The existing work only investigates a subset of the SPECpower results and does not classify the servers according to their hardware availability date. With newer results published, the derived models and conclusions from previous work pose greater errors thus should be revised and updated. For example, with the newest SPECpower results, the coefficient of correlation of energy proportionality and overall energy efficiency score decreases to 0.741 (477 valid data points) while it was 0.83 (459 data points including non-compliant points) in [16] .

In order to make more precise analysis of server energy proportionality and energy efficiency, different from others' work, we reorganize the published results by hardware availability date to quantify energy proportionality advances more accurately. After reorganization and correction, the average and median energy proportionality in hardware availability year have -6.2%~8.7% and -8.6%~13.1% differences compared with those values in published year. Similarly, the average and median energy efficiency in hardware availability year have -2.2%~16.6% and -5.0%~20.8% differences compared with those values in published year.

Our contributions are:

(1) We reorganize the SPECpower results by the server's hardware availability date and draw the energy proportionality and energy efficiency curves of all 477 servers, namely, the *pencil head* chart and the *almond* chart, to provide a complete view of energy proportionality and efficiency of production servers. The charts show enveloping edges of energy proportionality and energy efficiency curves of a wide variety of servers.

(2) We quantify the correlation between multiple metrics, including energy proportionality, energy efficiency, microarchitecture, dynamic range, ratio of peak energy efficiency over energy efficiency at 100% utilization, and peak energy efficiency offset. Regression analysis provides better explanation and predication of the evolution of energy proportionality and energy efficiency.

(3) We conduct extensive experiments on 4 rack servers to investigate the energy efficiency variations under different hardware configurations, including memory per core installation and processor frequency scaling. Our experiments show that hardware configuration has significant impact on server's energy efficiency.

The rest of the paper is organized as follows: Section II provides brief introduction of SPECpower benchmark and metrics and terms used in this paper. In Section III we present our observations and analysis of energy proportionality and energy efficiency trends. In Section IV we investigate the shifting of peak energy efficiency and its impact on energy proportionality improvement. We present some implications and guide on server energy proportionality in Section V. In Section VI we review some related work on energy proportionality analysis. We conclude the paper with some future research directions in Section VII.

II. SPECPOWER BENCHMARK AND METRICS

A. SPECpower Benchmark

SPECpower is the first industry-standard benchmark for measuring power and performance metrics of computer systems. Workload of SPECpower is designed to evaluate energy efficiency and performance of server-side Java applications for small and medium sized servers at graduated utilization levels. It is intended to test CPU, caches, memory, system architecture, JVM and some OS components without stressing the storage components. Specifically, SPECpower reports power consumption for severs at different utilization levels, from 100% utilization to idle in 10 percent intervals, over a set period of time. Detailed workload characterization of SPECpower can be found in [19].

B. Metric Notations and Terms

For consistence and convenience, we list some notations and terms used in this paper:

(1) *Utilization*. In this paper we define the server hardware utilization as the *target load* in a SPECpower result assuming that the benchmark excises all hardware components concertedly. In a SPECpower result, there are ten utilization levels from 10% to 100%. Please note that here the utilization is *not* the CPU utilization.

(2) *Peak utilization*. We refer 100% utilization as peak utilization.

(3) *Energy efficiency (EE)*. The energy efficiency is defined as performance to power ratio with unit of *ssj_ops per watt*. In

a SPECpower result, the energy efficiency values are entitled as *performance to power ratio*.

(4) *Server overall energy efficiency.* Server overall energy efficiency is overall performance to power ratio of a server in a SPECpower result, i.e., the ratio of sum of ssj_ops over sum of power for 10 utilization levels (from 10% to 100%) and active idle. The server overall energy efficiency is also referred as a server's SPECpower score.

(5) *Peak energy efficiency.* Peak energy efficiency is defined as the greatest (peak) energy efficiency of a server among all utilization levels.

(6) *Energy proportionality (EP).* In this paper we use the energy proportionality metric in [14]. Take a server from SPECpower dataset in 2016 with overall score 12212 as an example. We plot its power-utilization curve in Fig.1. Note that the power in Fig.1 is normalized to its power at 100% utilization. The solid line is the energy proportionality curve of the sample server and the dotted line is of an ideally energy proportional server. With the power-utilization curve in Fig.1, we can compute the energy proportionality of the server as following [14]:

$$\text{EP} = 1 - \frac{Area_{real} - Area_{ideal}}{Area_{ideal}} \quad (1)$$

Thus the power-utilization curve in Fig.1 is also called the energy proportionality curve. From Eq.1 we can see that EP is a value equal to or greater than zero but less than 2.0. For an ideally energy proportional server, its EP value is 1.0. For the server in Fig.1, we approximate its EP by summarizing the areas of ten trapezoids corresponding to ten utilization intervals and then get its EP value 1.02 according to Eq.1

III. EVOLUTIONARY TREND ON ENERGY PROPORTIONALITY

A. *Energy Proportionality Evolution*

During the last decade, server energy efficiency improves significantly thanks to technical breakthroughs including low power electronics and designs, dynamic voltage frequency scaling, and cross level coordinated power savings. We compute the energy proportionality value of each server according to Eq.1 and present the trend curve of all the servers' EP and EE values in Fig.2 and their statistics in Fig.3 and Fig.4. Not surprisingly, both energy proportionality and energy efficiency improve with times. The average energy proportionality increased significantly from 0.30 to 0.82 from 2005 to 2012 and seems stagnant at 0.84 in 2016. Similarly, the minimal EP in each year also improves significantly. For example, newest servers made in 2016 have minimal EP of 0.73, which is the greatest EP value in 2009. Among 477 servers, the least energy proportionality value is 0.18 (in 2008) and the highest energy proportionality value is 1.05 (in 2012). From Fig.2 and Fig.3 we observe that there are two significant improvement steps in EP, one is from 2008 to 2009, and another is from 2011 to 2012. From 2008 to 2009, the average and median EP increase from 0.37 to 0.55 (+48.65%), and 0.37 to 0.56 (+51.35%), respectively.

Similarly, from 2011 to 2012, the average and median EP increase from 0.66 to 0.82 (+24.24%), and 0.67 to 0.85 (+26.87%), respectively. From 2008 to 2009, the majority of the servers switch their processor microarchitecture from Core (Penryn) to Nehalem. From 2011 to 2012, the majority of the servers switch their processor microarchitecture from Nehalem (Westmere) to Sandy Bridge. These two switches are called *tock* in Intel's *tick-tock* chip iteration model which designates a new microarchitecture.

Although energy proportionality fluctuates with times in Fig.3, their energy efficiency increases monotonically with times in Fig.4, including its average value, maximal value and median value of each year in all three metrics. Only the minimal values of the three metrics in 2014 decrease because one server in 2014 has low EE as 1469 and low EP as 0.32. This server has a tower form factor and Intel Core i5-4570 processor usually not designed for server purpose use, which is quite different from other servers in 2014 or other years. Since there are few results for year 2004, 2005, 2006, and 2014, the statistical results in these years pose outliners and greater deviations. For example, EP decreases significantly from 2004 to 2005 and from 2012 to 2014. But the median EP in 2014 still increases compared with 2013.

We plot the CDF chart of energy proportionality in Fig. 5. Among 477 servers, 25.21% servers have EP between 0.6 and 0.7, and 17.44% servers have EP between 0.8 and 0.9. These two intervals have the most EP values. Moreover, 99.58% servers have EP less than 1.0.

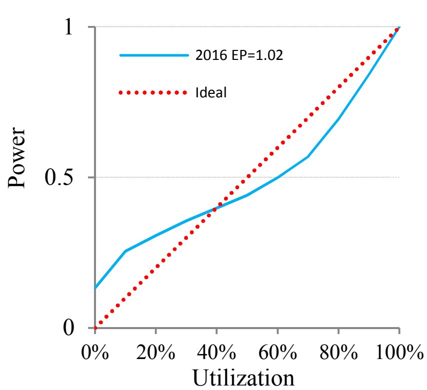

Fig. 1. Energy proportionality curve.

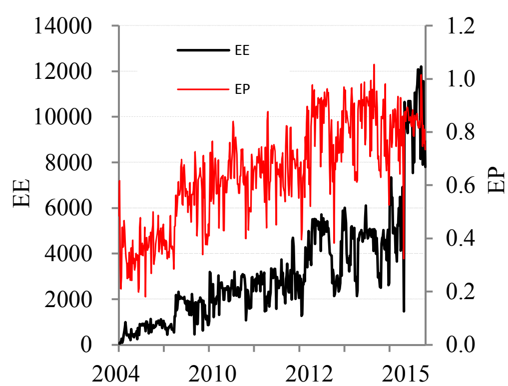

Fig.2. EP and EE evolution.

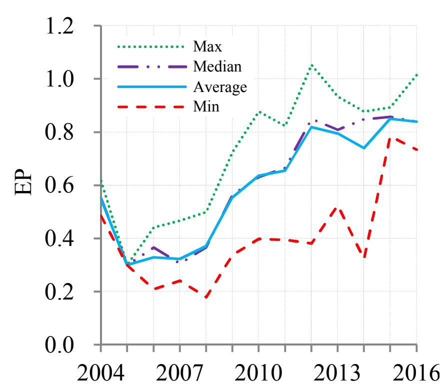

Fig.3. Stats. trend of EP.

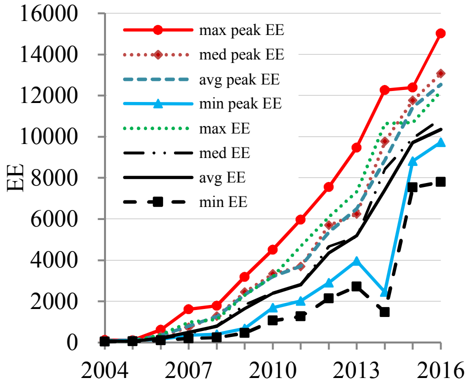

Fig.4. Stats. trend of energy efficiency.

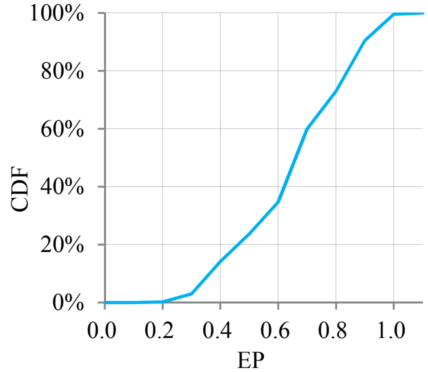

Fig.5. CDF of energy proportionality.

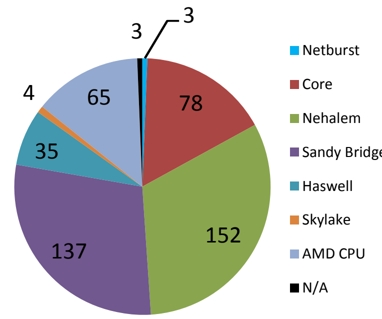

Fig. 6. CPU by microarchitecture.

B. *Is Energy Proportionality Improvement Stagnated?*

In Fig.3, we observe that the average EP decreases in 2013 and 2014. And it seems that the energy proportionality stagnated in the recent 3 years (2014 to 2016). To find out what causes this specious stagnation, we first group the servers by processor microarchitecture in Fig.6. However, we observe that even if the servers are equipped with processors of same microarchitecture, they may have different EP values since the processors are made of different codenames or generations. We divide the servers by microarchitecture codenames to narrow the EP statistics. We divide the CPU microarchitecture into subdomains with average EP in Fig.7.

From Fig.7, we observe that usually the servers with newer processor and finer manufacturing process have higher energy proportionality. This is because that generally when the die shrinks, the processor's power consumption and heat dissipation also drops. For example, in Fig.7 we observe that servers of Sandy Bridge microarchitecture, especially the 22 servers of Sandy Bridge EN microarchitecture, have the highest EP (average 0.90 and median 0.94) among 477 servers. However, its successors, the Ivy Bridge generation and Haswell microarchitecture (22nm), have lower energy proportionality compared with Sandy Bridge microarchitecture.

Furthermore, the server's energy proportionality maybe lower even if it is equipped with finer lithography process based processor. For example, most of the servers equipped with the processors of Ivy Bridge microarchitecture have lower energy proportionality than those of Sandy Bridge microarchitecture. We list the proportion of servers in 2012 and 2016 of different processor microarchitecture in Fig.8.

From Fig.8 we observe that **the energy proportionality in 2013 and 2014 decreases and maybe it's due to the adoption of processors whose microarchitecture poses lower energy proportionality. And the energy proportionality recovers in 2015 and 2016.** Therefore, the energy proportionality decrease in 2013 and 2014 is not supposedly stagnant but mainly due to specific processor microarchitecture and lack of enough SPECpower results.

C. *The Pencil Head Chart and the Almond Chart*

In order to provide a complete view of energy proportionality, we plot the energy proportionality curves of all 477 servers in Fig.9 and name it as *pencil head* chart of energy proportionality. We also plot the energy proportionality curves of 11 selected representative servers together with the ideal energy proportionality curve in Fig.10. From Fig.9 we can see that except the starting part before 10% utilization, all EP curves are located between two curves, i.e., the upper enveloping curve is the curve of a server with lowest observed EP as 0.18 and the lower enveloping curve is the curve of a server with highest observed EP as 1.05. For servers with EP greater than 1.0, their EP curves intersect with the ideal EP curve absolutely before 100% utilization. For example, the servers with EP 1.02 and 1.05 in Fig.10 are representatives of such servers.

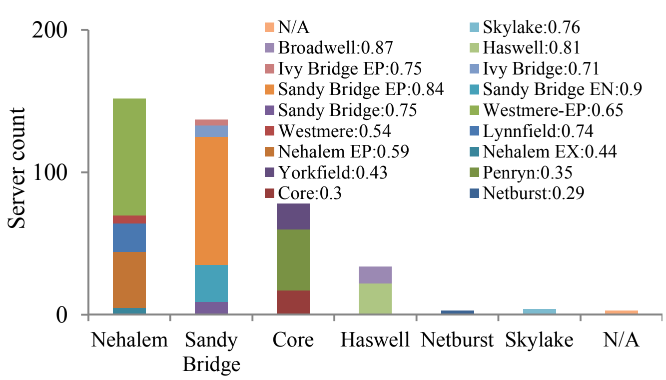

Fig. 7. Servers by microarchitecture codename (Intel).

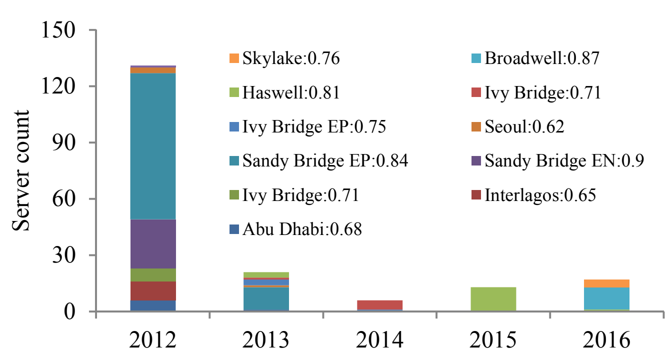

Fig. 8. Servers by microarchitecture from 2012 to 2016.

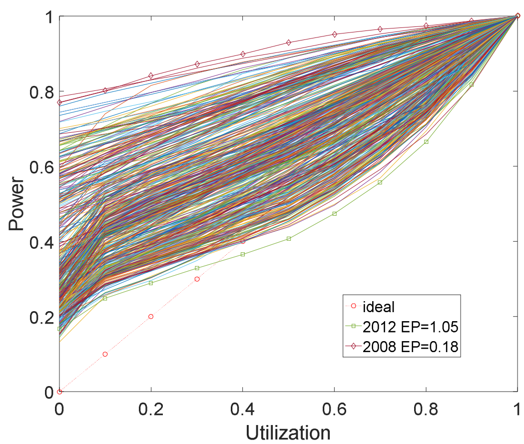

Fig. 9. The pencil head chart of energy proportionality.

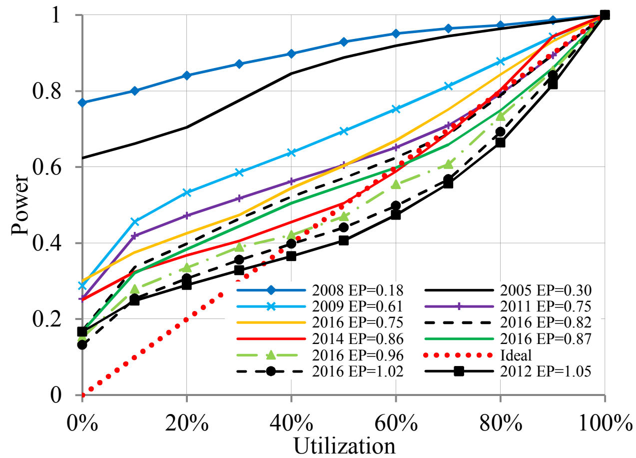

Fig. 10 Selected energy proportionality curves.

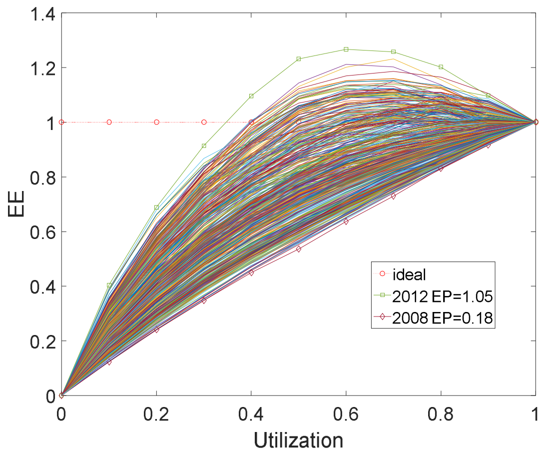

Fig. 11. The almond chart of energy efficiency.

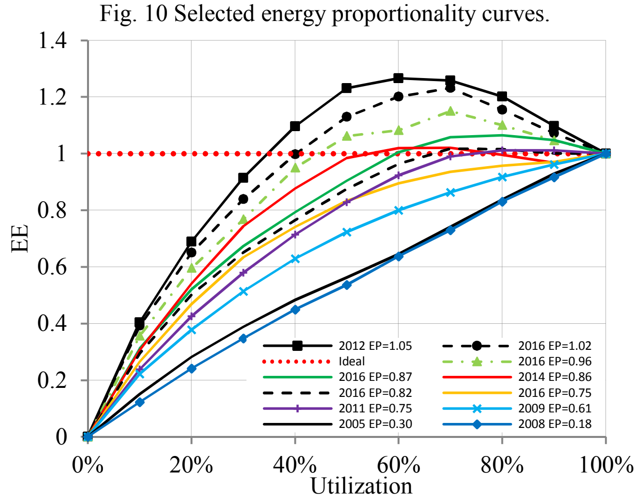

Fig. 12 Select energy efficiency curves.

However, for some servers with EP less than 1.0, their EP curves may also intersect with the ideal EP curve before 100% utilization, i.e., the EP curves of servers with EP 0.96, 0.87, 0.86, 0.82 and 0.75 (2011) in Fig.10. It's interesting that the server with EP 0.86 (red solid line in Fig.10) is a 1U server whose EP curve intersects **twice** with the ideal EP curve during utilization intervals of 50%-60% and 70%-80%, respectively. We observe that **for the servers whose EP curves intersect with the ideal energy proportionality curve before 100% utilization, the higher its EP is, the farther the intersection is away from 100% utilization.**

Moreover, in Fig.10 there are some servers whose EP values are less than 1.0 but *never* intersect with the ideal EP curves before 100% utilization, i.e., the EP curves of servers with EP 0.18, 0.30, 0.61 and 0.75 (2016). In Fig.10 we also observe that even if two servers have same EP values, i.e., two servers with EP 0.75 in 2011 and 2016, they may have quite different EP curves. Among the two curves, one intersects with the ideal curve but the other doesn't. This is because that they have different characteristics of power consumption at different utilization levels, i.e., different linear deviation (LD).

Accordingly, we plot the energy efficiency curves of all 477 servers in Fig.11 and name it as *almond* chart of energy efficiency. In Fig.11, all EE curves are located between two curves, i.e., the upper enveloping curve is the curve of a server with highest observed EP 1.05 and the lower enveloping curve is the curve of a server with lowest observed EP 0.18. We also plot the energy efficiency curves of 11 selected representative servers together with the ideal energy efficiency curve in Fig.12 (energy efficiency normalized to energy efficiency at 100% utilization). Similarly, the *almond* chart provides a complete view of the energy efficiency trends of all the mainstream production servers. More importantly, from Fig.12, we also observe that for servers with EP values greater than 1.0, i.e., 1.02 and 1.05, they reach their 0.8x energy efficiency at 100% utilization before 30% utilization and reach their 1.0x energy efficiency at 100% utilization before 40%. This also reflects that the higher a server's EP is, the earlier it intersects with the ideal EP curve. These servers reach to its high energy efficiency zone faster than servers with lower EP values. And their high energy efficiency zones above 1.0 are wider than servers with lower EP values. **These high energy efficiency zones are better places where the servers should keep working at.** From Fig.12 we observe that for some servers with EP value close to 1.0 or greater than 1.0, the higher the EP value is, the farther its peak energy efficiency is away from the ideal energy efficiency curve. We quantify these correlations in the following sections.

D. Driving Force of Energy Proportionality Improvement

According to the design goal of energy proportional computing, idle power is an important factor that may affect the server's energy proportionality. A lower idle power percentage may help the actual energy proportionality curve intersect with the ideal curve farther away from 100% utilization. We compute the idle power percentage (normalized to power at 100% utilization) from 2004 to 2016 and its correlation with EP. Among 477 servers, their coefficient of correlation between energy proportionality and idle power percentage is -0.92, which means a very significant negative correlation. We observe that the lower a server's idle power percentage, the higher its energy proportionality is. This is obvious because in energy proportionality curve the lower the server's idle power percentage is, the smaller the area under its energy proportionality curve is. Through regression analysis, we get Eq.2:

$$\mathrm{EP} = 1.2969 \cdot e^{-2.107 \cdot idle} \quad (2)$$

R^2=0.892.

where *idle* is idle power percentage normalized to power at 100% utilization.

Eq.2 indicates that server energy proportionality increases exponentially with the decrease of idle power percentage. Although it seems that the improvement of EP in recent years almost stagnated, our observation shows that if we decrease the idle power percentage further, server energy proportionality can still be improved exponentially. For example, if the idle percentage is 5%, then the energy proportionality will be 1.17. Eq.2 also suggests that the theoretical maximal EP value is 1.297 for idle percentage=0 since *idle*>=0. **Moreover, the idle power percentage decreases more significantly from 2006 to 2012 than that from 2012 to 2016. And this also helps the energy proportionality improve more from 2006 to 2012 than that from 2012 to 2016.**

E. Economies of Scale in Energy Proportionality

In order to investigate the energy proportionality on multiple-nodes servers, we plot their average and median values of energy proportionality of the servers with more than 2 counts in Fig.13. From Fig.13 we observe that the median EP values increase monotonically with the total server nodes, while the average EP value decreases a bit from 4 nodes to 8 nodes, but still increases at 16 nodes. The EP decrease at 8 nodes may suffer from the lack of enough results compared with other multiple node servers. **This suggests that the energy proportionality also benefits from economies of scale in multiple node servers**. Fig.13 also suggests that **grouping multiple identical nodes to work together on same workload is more energy proportional than letting individual identical server node work on different workloads in production environment, even without any further system tunings. This provides an opportunity to improve cluster-wide energy proportionality**.

Among 477 servers, there are 403 single node servers equipped with 1 to 8 chips. Specifically, there are 77 servers equipped with 1 chip, 284 servers equipped with 2 chips, 36 servers equipped with 4 chips, and 6 servers equipped with 8 chips. We plot the average and median EP of these 403 single node servers in Fig.14. Surprisingly, unlike the EP and energy efficiency improvements with server node increasing, all single node servers equipped with 2 chips have the highest EP and EE except the median EP. **The energy proportionality improvement beneficial from economies of scale does not hold well in single node servers when we increase the number of chips equipped on the board.** Specifically, the median EP for 1 chip and 2 chips single node server is 0.67 and 0.66, respectively. Although the energy efficiency and energy proportionality benefit from economies of scale when we increase the number of chips from 1 to 2 in a single node server, they decrease monotonically when the number of chips increases from 2 to 4 and 8. We think **the partial reason is that the increase in power density exceeds the performance gains when the number of chips increases from 2 to 4 and 8.** For the 284 single node servers equipped with 2 chips, we compare their average and median values of energy proportionality and energy efficiency with those values of all 477 servers and plot it in Fig.15.

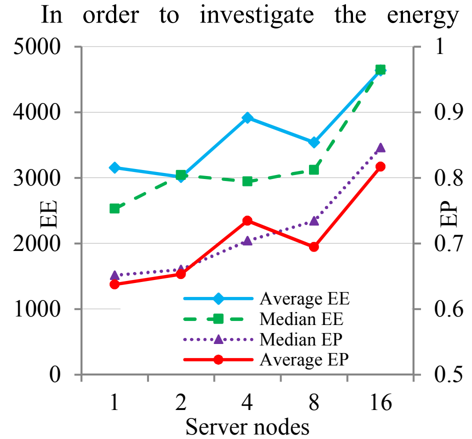

Fig. 13 EP and EE improve with server nodes.

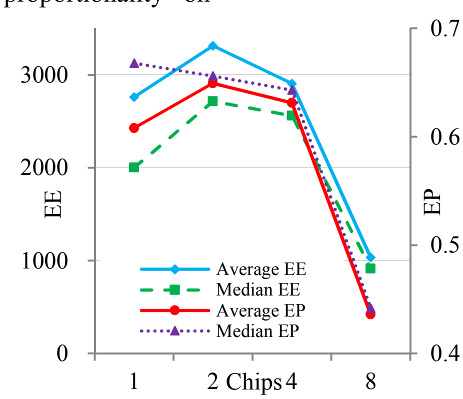

Fig.14. EP and EE of single node servers.

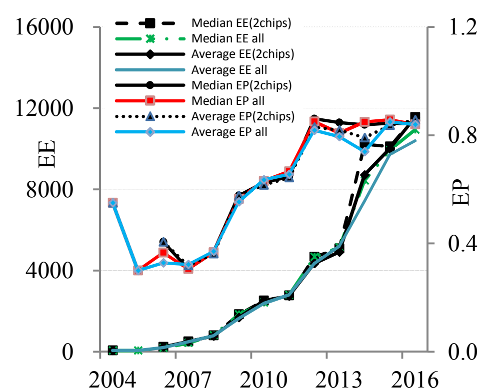

Fig.15. Comparison of EP and EE of single node servers with 2 chips to all servers.

The average energy proportionality and energy efficiency of single node servers with 2 chips are 2.94% and 4.13% higher than those of all the servers at the same hardware availability year, while the median values are 1.18% and 6.26% higher than those of all servers, respectively. This shows that the 2-chips single node server performs better than the average at the same generation or same hardware availability year.

IV. SHIFTING OF PEAK ENERGY EFFICIENCY AND ITS IMPACT ON ENERGY PROPORTIONALITY

A. *Peak Energy Efficiency Utilization Shifting*

The energy efficiency improvements benefit much from the hardware breakthroughs with times (see Fig.2 and Fig.4). However, when or where can a specific server achieve its peak energy efficiency is more important for not only data center operation but also workload placement. Keeping servers running at higher energy efficiency zones can save energy dramatically. We illustrate the chronical shifting trend of utilization spot with peak energy efficiency in Fig.16.

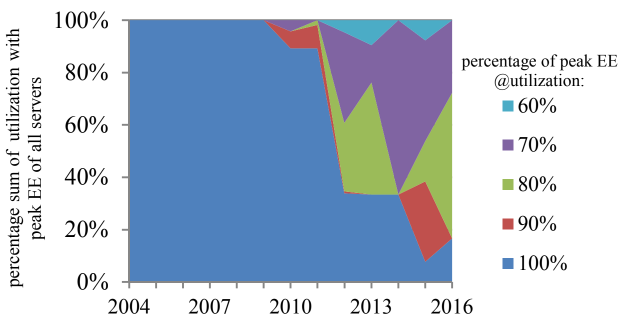

Fig.16. Chronical shifting of utilization with peak energy efficiency.

In Fig.16 we can observe that the occurrences of utilization spot with peak energy efficiency become diverse since 2010. Before 2010, all the servers achieve their peak energy efficiency at 100% resource utilization. However, among all 18 servers in 2016 (from Q1 to Q3), 3 servers achieve their peak energy efficiency at 100% utilization, 10 servers achieve their peak energy efficiency at 80% utilization, and 5 servers achieve their peak energy efficiency at 70% utilization. Among all 477 servers in SPECpower dataset, 69.25% servers achieve their peak energy efficiency at 100% utilization, 13.81% servers achieve their peak energy efficiency at 70% utilization, 11.72% servers achieve their peak energy efficiency at 80% utilization, 3.35% servers achieve their peak energy efficiency at 90% utilization, and only 1.88% servers achieve their peak energy efficiency at 60% utilization. Please note that there is a server in 2011 that achieves its peak energy efficiency at both 80% and 90% utilization. That's why we get 478 utilization spots for 477 servers.

In the former sections we already observe that the energy proportionality increases in the last decade. With the improvement of energy proportionality, the utilization spot that a server achieves its peak energy efficiency also shifts with times. We observe in Fig.16 that **the utilization spot of peak energy efficiency shifts significantly from 100% to 80% and 70% from the first time interval (2004-2012) to the second time interval (2013-2016)**. In the interval of 2004-2012, 75.71% servers achieve peak energy efficiency at 100% utilization. However, in the second interval of 2013-2016, only 23.21% servers achieve peak energy efficiency at 100% utilization while 35.71% servers achieve peak energy efficiency at 80% utilization, and 26.79% servers achieve peak energy efficiency at 70% utilization. If we investigate the latest servers released in 2015-2016 in Fig.13, we can see that this shifting of utilization spot with peak energy efficiency from 100% to 80% and 70% is still continuing since 2015. We can expect the peak energy efficiency at 50% or even 40% utilization in the near future.

B. *Asynchronization of Energy Proportionality and Energy Efficiency Evolution*

Ideally, the energy proportionality and energy efficiency of servers will improve continuously with technical breakthroughs and advancements. We can see that they do improve with times respectively, but the improvements are not concerted from Fig.2. In other words, the distributions of energy proportionality and energy efficiency with times (by hardware availability year) are not consistent. Specifically, the improvements of energy proportionality and energy efficiency are asynchronous in two folds:

(1) In terms of temporal improvement of all servers, energy proportionality does not keep pace with energy efficiency. For example, among the top 10% servers with highest energy proportionality, 91.7% of them are made in 2012, which significantly exceeds the sharing (27.4%) of servers made in 2012 in all the investigated results. In contrast, among the top 10% servers with highest energy efficiency, only 16.7% of them are made in 2012, which is less than their sharing (27.4%) of servers made in 2012 in all the investigated results, while all the servers made in 2015 and 2016 are in the top 10% servers with highest energy efficiency and significantly exceeds their sharing in all the investigated servers.

(2) In terms of energy proportionality and energy efficiency of a specific server, a server with highest energy proportionality does not essentially have the highest energy efficiency, and vice versa. Ideally, it's desired that a server has close ranking position in both energy proportionality and energy efficiency. However, our observation shows that for most of the servers with high energy proportionality, they do not have high energy efficiency values. For example, among the top 10% servers with highest energy proportionality, only 14.6% of them also have energy efficiency values fitted in top 10% highest energy efficiency range.

V. GUIDE FOR ENERGY EFFICIENT SYSTEM DESIGN AND OPERATION

A. *Energy Efficient Hardware Configuration*

Due to the difference of energy characteristics in processor, memory, and I/O systems, a server system may expose different energy efficiency with different hardware configuration and installation [26,27,28]. For example, since SPECpower does not stress storage component [19], almost all the server vendors only use single hard disk drive or solid state drive configuration to lower the whole system power consumption.

Thanks to memory technology advancement, today's servers are usually equipped with more memory modules for better performance. However, DRAM modules also consume significant power in memory intensive workloads compared with processor. Here we investigate the memory configuration of each server and compute the ratio of installed memory capacity over installed processor cores, i.e., memory per core (GB/core). We list the statistics of each ratio with more than 10 counts in Table 1(430 servers in total among 477 servers).

TABLE I. MEMORY PER CORE STATISTICS OF PUBLISHED SERVERS.

memory per core (GB/core)	0.67	1	1.33	1.5	1.78	2	4
count	15	153	32	68	13	123	26

We plot their average energy proportionality and energy efficiency of these 430 servers with different memory per core in Fig.17. Fig.17 shows that proper memory configuration is vital for both energy proportionality and energy efficiency on the majority of the servers. For energy proportionality and energy efficiency, the best memory per core is 1.5GB/core and 1.78GB/core, respectively. It also indicates that for some workload (at least SPECpower), proper memory configuration can achieve better energy efficiency and energy proportionality. In order to verify if 1.78GB/core is the best memory per core configuration, we run SPECpower(without any customization and optimization with OpenJDK1.8.0_65) on 4 different 2U rack servers with different installed memory capacities. The base configuration of these servers is listed in Table 2.

Our results on the tested 4 servers in Table 2 show that they get peak energy efficiency at peak (100%) utilization. We compare their energy efficiencies under different memory per core configurations in Fig.18 to Fig.20. From Fig.18 to Fig.20 we observed that there is a *best* memory per core value for each server where the server has the highest energy efficiency. The best memory per core value for #1, #2, and #4 server are 1.75GB, 4 GB, and 2.67 GB, respectively. Our observations presented here indicate that memory installation can affect not only the system performance but also the energy efficiency of the deployed servers. Our experiments show that when memory per core increases to 8GB or more, the energy efficiency of all the tested servers decrease significantly. For example, on server #2, energy efficiency decrease 10.6% from MCP=4 to MCP=8, while on server #4, energy efficiency decreases 4.6% and 11.1% from MCP=2.67 to MCP=8 and MCP=16，respectively. This means that server's energy efficiency does not scale well with memory capacity, especially after the *best* point.

B. *DVFS Has Both Lower Power and Energy Efficiency*

DVFS[24,25] is widely used for power capping or energy aware job scheduling in data centers. Usually lower CPU frequency has lower power consumption. However, in energy efficiency wise, lower frequency does not lead to better energy efficiency.

In our experiments on 4 tested servers, the servers have lower energy efficiency at lower CPU frequency consistently on all servers at all frequency levels. Moreover, the energy efficiency does not stay even constant when frequency decreases because the completed jobs decrease more significantly. Fig.18 to Fig.20 also shows that the *ondemand* governor always almost has the highest energy efficiency and it's very close to the energy efficiency with the highest frequency. Due to space limitation, we only give the energy efficiency and peak power consumption on server#4 in Fig.21. From Fig.21 we can observe that the server consumes more power at higher CPU frequency at same memory per core configuration. When memory per core configuration increases at fixed CPU frequency, the peak power consumption also increases. And the *ondemand* almost consumes same power with the highest CPU frequency.

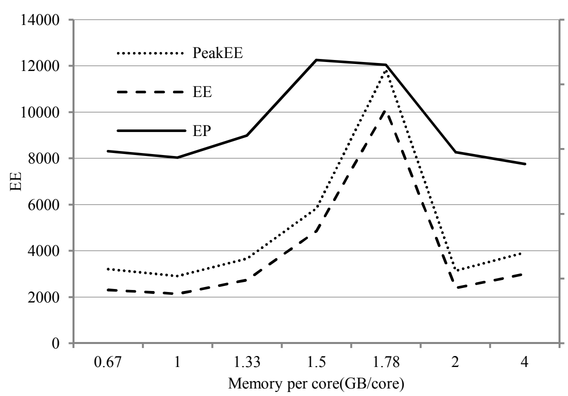

Fig. 17. EP and EE at different memory per core configuration.

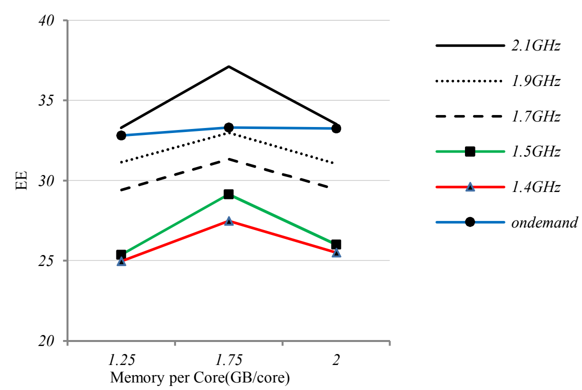

Fig.18. Energy efficiency with different memory per core and CPU Frequency on #1 server.

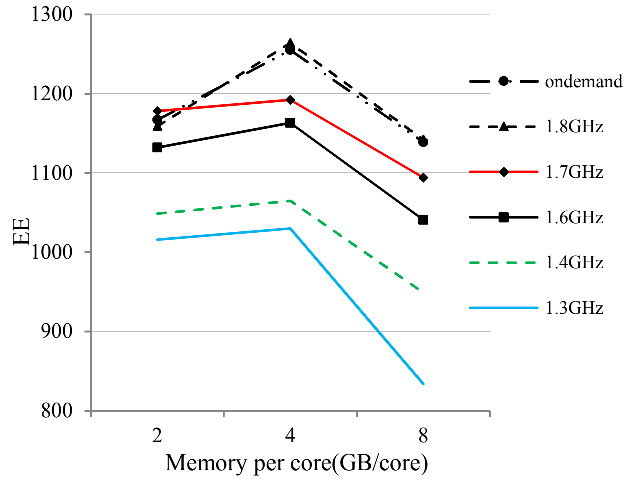

Fig.19. Energy efficiency with different memory per core and CPU Frequency on #2 server.

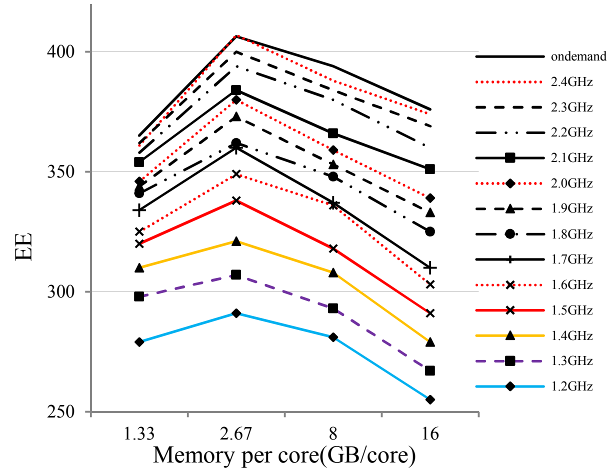

Fig.20. Energy efficiency with different memory per core and CPU Frequency on #4 server.

TABLE II. BASE CONFIGURATION OF TESTED 2U SERVERS.

No	*Name*	*Hardware Availability Year*	*CPU Model*	*Total cores*	*CPU TDP (watts)*	*Memory (GB)*	*DISK*
#1	Sugon A620r-G	2012	2*AMD Opteron 6272	32	115	64(8G*8) DDR3 1600MHz	4*SAS 300GB 10K rpm (RAID10)
#2	Sugon I620-G10	2013	1*Intel Xeon E5-2603	4	80	32(4G*8) DDR3 1600MHz	1*SAS 300GB 10K rpm
#3	ThinkServer RD640	2014	2*Intel Xeon E5-2620 v2	12	80	160(16G*10) DDR4 2133MHz	1*SSD 480GB
#4	ThinkServer RD450	2015	2*Intel Xeon E5-2620 v3	12	85	192(16G*12) DDR4 2133MHz	1*SSD 480GB

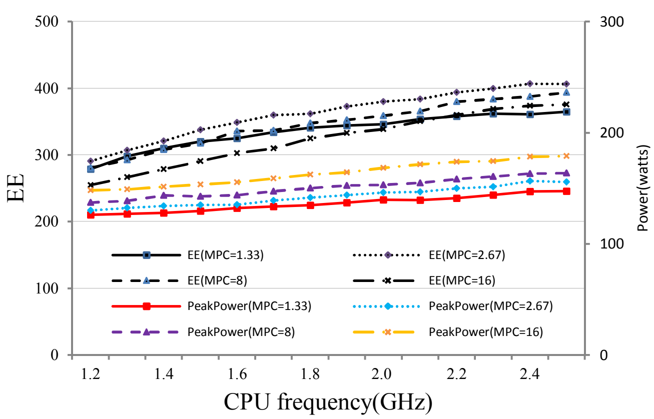

Fig.21. Energy efficiency and peak power on server#4 with different memory per core and frequencies.

C. Optimal Working Region in Operation

Except for the overall energy proportionality and energy efficiency, optimal working range for servers with high energy efficiency is more practical for data center operating and energy management. In data centers, heterogeneous servers of different generations and energy proportionality are co-working for service provisioning and processing. How to exploit the energy proportionality of these heterogeneous servers is vital for coordinated power management and energy management. For example, if a server has peak energy efficiency at 70% utilization, it may have higher energy proportionality. Our analysis shows that the 70% to 100% utilization region is better working region. Therefore, during data center capacity planning or workload placement stage, we don't need to pack as many as jobs to the server to let it fully busy. Instead, keeping the server at 70% utilization is more energy efficient. For data centers, we can first group servers by their energy proportionality values, and then subdivide the servers by their energy efficiency curves by grouping the servers with the widest working region beyond the ideal energy efficiency curve into a *logical* cluster. The optimal working region of this logical cluster is the overlapping best working region of its member servers. Then keep this logical cluster working in its optimal working region is more energy efficient. Therefore, for a fixed number of racks energy proportionality aware workload placement can maximize the throughput or do more jobs under fixed power supply.

Moreover, although SPECpower is a standard benchmark to mimic transactions workload, our insights presented here is also fruitful for other workloads. For specific applications, the server may exhibit energy proportionality and energy efficiency curve different from that of SPECpower workload. Energy proportionality aware workload placement and job scheduling is feasible after characterization of the server's energy efficiency curve under specific workloads.

VI. RELATED WORK

Energy consumption of large-scale data centers accounts for a significant portion of their operational cost. A large amount of recent work focused on improving the server's energy efficiency to reduce power and energy consumption [20]. Various hardware related approaches have been proposed to increase data center energy efficiency [21], including low-latency server power states [22], network traffic routing

[29], power market participation [30,31,32], transient servers power characterization [33, 34,35]. Renewable energy and novel cooling techniques are also introduced into data center wide power provisioning and carbon footprint reduction [36, 37]. Usually power and performance are tradeoff in data centers [38]. However, most of the above mentioned energy efficiency improvements and optimizations mainly focused upon full utilization or idle utilization regions. And maintaining a data center at low power mode is not a reasonable means as low power mode will make data center at low energy efficiency [25]. Barroso and Hölzle [8] pointed out that most of the time servers are running at the lowest energy-efficiency in order to save energy consumption (in 2007). This urges the energy proportional computing [8, 14] , whose goal is to improve not only energy efficiency, but also energy proportionality of severs in low, but non-zero utilization region and the reduction of overall server energy consumption [10, 16, 17].

With emergence of SPECpower, researchers also tried to characterize its workload behavior [19, 44, 45, 46] and derive trend of server energy efficiency from its published results [18, 47, 48]. Hsu and Poole[16] investigated 459 servers from SPECpower benchmark until June 2014 (including non-compliant data and is a subset of the valid published servers) to compare a wide range of metrics for measuring energy proportionality, such as ER, EP, IPR, and LD. Wong and Annavaram[17] investigated the energy proportionality trend of 291 servers from SPECpower benchmark results ranging from November 2007 to December 2011 and found that although the overall energy proportionality has been improved, not all of the utilization-levels is well energy proportional. In other words, when servers are running at low utilization there appears significant proportionality gap. Wong[41] also investigated the energy proportionality trend of 426 (a subset of the valid published servers) servers from SPECpower benchmark ranging from December 2007 to September 2015 and argued that highly energy proportional servers typically exhibit peak energy efficiency at ~60% utilization. However, we found that 69.54% servers exhibit peak energy efficiency at 100% utilization, only 2.10% servers at ~60% utilization from all the published SPECpower results. We also found that for servers with peak energy efficiency at 60% utilization, their average and median peak energy efficiency is close to those of servers in 2013.

Although there are some related work focusing on the trend analysis on server energy efficiency using published SPECpower results, up-to-date comprehensive studies about server energy efficiency and proportionality based on hardware availability date are currently missing. As we mentioned in Section I, among 477 valid SPECpower results, there are 15.5% results whose published dates are different from hardware availability date, some can date back to 1-6 year before their published year. In order to alleviate the confusion and misleading of such mismatch between published year and hardware availability year, we reorganize all the published SPECpower results by their hardware availability year. After reorganization, we conduct more comprehensive and more accurate up-to-date analysis than previous work on server energy efficiency and proportionality in the time interval of 2004 to 2016. We also identify mathematical correlations between key metrics of server energy efficiency and proportionality.

VII. CONCLUSIONS

Although estimated energy consumption in data centers is increasing steadily annually while taking advantage of technical breakthroughs on server energy efficiency, there are still ongoing energy consumption concerns on data centers from IT companies and governments. Understanding of energy proportionality and energy efficiency of servers can help data centers designers and system operators in many folds, including system capacity planning, power shifting, workload placement, migrations, and resource scheduling.

In this paper, we investigate the published results of SPECpower benchmark from 2007 to 2016 and translate the results to time scale at 2004 to 2016 based on their hardware availability date. We conduct more comprehensive and more accurate up-to-date analysis on energy proportionality and energy efficiency of typical industry servers. We find that the specious stagnation of server energy proportionality in recent years is mainly caused by the adoption of processors of specific microarchitecture and is not the indicative trend of energy proportionality improvement. For a server with multiple identical nodes, its energy proportionality increases with its count of nodes. However, for a single node server, it has the highest energy proportionality when it is equipped with 2 chips. The peak energy efficiency shifting from 100% utilization spot also helps server energy proportionality improvement. Our findings presented in this paper shed lights on energy proportionality improvements and proportionality aware workload placement and energy minimizations for system designers, as well as data center operators.

Although the SPECpower benchmark is a designed for server energy efficiency evaluation, it may not represent the different kinds of application running on servers in a datacenter. As future work, we plan to do more experiments to characterize the energy proportionality and energy efficiency variations on typical industrial servers under different workloads and hardware configurations, including processor, memory, I/O and networks. The future study may provide more insights to build servers with better than linear energy proportionality or energy proportionality reconfigurable servers.

ACKNOWLEDGMENT

This work is supported by Natural Science Foundation of China (No.61472109, 61003077, 61402140, 61572163), and Natural Science Fund of Zhejiang Province (LY14F020045, LY16F020018, Y1101092), Fund of Department of Education of Zhejiang Province (Y200907182). This work is also supported in part by National Science Foundation (NSF) grant CNS-1561216, and by the Introduction of Innovative R&D team program of Guangdong Province (No. 201001D0104726115).

REFERENCES

[1] S. Subramanya, Z. Mustafa, D. Irwin, Prashant Shenoy. Beyond energy-efficiency: evaluating green datacenter applications for energy-agility. Proceedings of the 7th ACM/SPEC on International Conference on Performance Engineering (ICPE), 2016, pp. 185-196.

[2] T. Mastelic, A. Oleksiak, H. Claussen, Ivona Brandic, Jean-Marc Pierson, Athanasios V. Vasilakos. Cloud computing: survey on energy efficiency. ACM Computing Surveys, 2014, 47(2): 1-36.

[3] Report to Congress on server and data center energy efficiency Public Law 109-431,U.S. Environmental Protection Agency ENERGY STAR Program,2007.

[4] Natural Resources Defense Council, Data center efficiency assessment, https://www.nrdc.org/sites/default/files/data-center-efficiency-assessment-IP.pdf

[5] A. Shehabi, S.J. Smith, N. Horner, I. Azevedo, R. Brown, J. Koomey, E. Masanet, D. Sartor, M. Herrlin, W. Lintner. United States data center energy usage report. Lawrence Berkeley National Laboratory, Berkeley, California. LBNL-1005775, 2016.

[6] M.A. Islam, S. Ren. A new perspective on energy accounting in multi-tenant data centers. Proceedings of the 2016 USENIX Workshop on Cool Topics in Sustainable Data Centers (CoolDC '16), 2016, pp.1-6.

[7] Q. Wu, Q. Deng, L. Ganesh, C. Hsu,Y. Jin, S. Kumar, B. Li, J. Meza, Y. Song. Dynamo: Facebook's data center-wide power management system. Proceedings of the 43th International Symposium on Computer Architecture (ISCA) , 2016, pp.469-480.

[8] L.A. Barroso, U. Hölzle. The case for energy-proportional computing.Computer, 2007, 40 (12):33-37

[9] D. Lo, L. Cheng, R. Govindaraju, L.A. Barroso, C. Kozyrakis. Towards energy proportionality for large-scale latency-critical workloads. Proceedings of the 41th International Symposium on Computer Architecture (ISCA), 2014, pp. 301-312

[10] B. Subramaniam, W. Feng.Towards energy-proportional computing for enterprise-class server workloads. Proceedings of the 4th ACM/SPEC International Conference on Performance Engineering (ICPE), 2013, pp.15-26.

[11] N. Tolia, Z. Wang, M. Marwah, C. Bash, P. Ranganathan, X. Zhu. Delivering energy proportionality with non energy-proportional Systems – optimizing the ensemble. Proceedings of the 2008 conference on Power aware computing and systems (HotPower), 2008, pp.2-2.

[12] K.T. Malladi, B.C. Lee, F. A. Nothaft, C. Kozyrakis, K. Periyathambi, M. Horowitz. Towards energy-proportional datacenter memory with mobile DRAM. Proceedings of the 39th Annual International Symposium on Computer Architecture (ISCA), 2012, pp.37-48.

[13] K.T. Malladi, I. Shaeffer, L. Gopalakrishnan, D. Lo, B. C. Lee, M. Horowitz. Rethinking DRAM power modes for energy proportionality. Proceedings of the 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO-45), 2012, pp.131-142.

[14] F. Ryckbosch, S. Polfliet, and L. Eeckhout. Trends in server energy proportionality. Computer, 2011, 44(9):69-72.

[15] https://www.spec.org/power_ssj2008/

[16] C. Hsu, S.W. Poole. Measuring server energy proportionality. Proceedings of the 6th ACM/SPEC International Conference on Performance Engineering (ICPE), 2015, pp.235-240.

[17] D. Wong, M. Annavaram. KnightShift: scaling the energy proportionality wall through server-level heterogeneity. Proceedings of the 45th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO-45), 2012, pp.119-130.

[18] D. Wong, J. Chen, M. Annavaram. A retrospective look back on the road towards energy proportionality. Proceedings of the IEEE International Symposium on Workload Characterization (IISWC), 2015, pp.110-111.

[19] L. Gray, A. Kumar, H. Li. Workload characterization of the SPECpower_ssj2008 benchmark. Proceedings of 2008 SPEC International Performance Evaluation Workshop (SIPEW), 2008, pp 262-282.

[20] Y. Zhu, V.J. Reddi. High-performance and energy-efficient mobile web browsing on big/little systems. Proceedings of the 2013 IEEE 19th International Symposium on High Performance Computer Architecture (HPCA), 2013, pp.13-24.

[21] A. Mamun, I. Narayanan, D. Wang, A. Sivasubramaniam, H.K. Fathy. Battery health-conscious online power management for stochastic datacenter demand response. Proceedings of the 2016 American Control Conference (ACC), 2016.

[22] C. Isci, S. McIntosh, J. Kephart, R. Das, et al. Agile, efficient virtualization power management with low-latency server power states. Proceedings of the 40th International Symposium on Computer Architecture (ISCA) , 2013, pp.96-107.

[23] D. Fan, Z. Tang, H. Huang, H. Huang, G. R. Gao. An energy efficient TLB design methodology. Proceedings of the 2005 international symposium on Low power electronics and design (ISLPED), 2005, pp.351-356.

[24] R. Rodrigues, A. Annamalai, I. Koren, S. Kundu. Reducing energy per instruction via dynamic resource allocation and voltage and frequency adaptation in asymmetric multicores, Proceedings of the IEEE Computer Society Annual Symposium on VLSI (ISVLSI), 2014, pp.436-441.

[25] D. Meisner, T.F. Wenisch. Does low-power design imply energy efficiency for data centers? Proceedings of the 17th IEEE/ACM international symposium on Low-power electronics and design (ISLPED), 2011, pp.109-114.

[26] M. Bi, S. Chandrasekharan, C. Gniady. IAMEM: interaction-aware memory energy management. Proceedings of the 2013 USENIX Annual Technical Conference (ATC), 2013, pp.267-278.

[27] Q. Deng, D. Meisner, A. Bhattacharjee, T.F. Wenisch; R. Bianchini. CoScale: coordinating CPU and memory system DVFS in server systems. Proceedings of the 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO-45), 2012, pp.143-154.

[28] Y. Zhang, J. Liu, M. Kandemir. Software-directed data access scheduling for reducing disk energy consumption. Proceedings of the 2012 IEEE 32nd International Conference on Distributed Computing Systems (ICDCS), 2012, pp.596-605.

[29] C. James, N. Carlsson. Green domino incentives: impact of energy-aware adaptive link rate policies in routers. Proceedings of the 6th ACM/SPEC International Conference on Performance Engineering(ICPE), 2015, pp.211-221.

[30] H. Chen, C. Hankendi, M.C. Caramanis, A.K. Coskun. Dynamic server power capping for enabling data center participation in power markets.Proceedings of the International Conference on Computer-Aided Design(ICCAD), 2013, pp.122-129.

[31] R. Singh, D. Irwin , P. Shenoy , et al. Yank: enabling green data centers to pull the plug. Proceedings of the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13), 2013, pp.143-156.

[32] Y. Guo, M. Pan. Coordinated energy management for colocation data centers in smart grids. Proceedings of 2015 IEEE International Conference on Smart Grid Communications (SmartGridComm),2015.

[33] R. Singh, P. Sharma, D. Irwin, P. Shenoy, K.K. Ramakrishnan. Here today, gone tomorrow: exploiting transient servers in data centers. IEEE Internet Computing, 2014, 18(4):22-29.

[34] D. Wang , C. Ren , S. Govindan, A. Sivasubramaniam ,B. Urgaonkar , A. Kansal , K. Vaid. ACE: abstracting, characterizing and exploiting peaks and valleys in datacenter power consumption. Proceedings of the IEEE International Symposiam on Workload Characteriazation(IISWC), 2013, pp.44-55.

[35] J. Kelley, C. Stewart, D. Tiwari, S. Gupta. Adaptive power profiling for many-core HPC architectures. Proceedings of the 2016 IEEE International Conference on Autonomic Computing (ICAC), 2016.

[36] W. Deng, F. Liu, H. Jin, B. Li, D. Li. Harnessing renewable energy in cloud datacenters: opportunities and challenges. IEEE Network, 2014, 28(1): 48-55.

[37] Baris Aksanli, Jagannathan Venkatesh, Inder Monga and Tajana Simunic Rosing. Renewable energy prediction for improved utilization and

efficiency in data centers and backbone networks. Computational Sustainability, 2016, pp.47-74.

[38] Z. Zhou, F. Liu, H. Jin, B Li, B. Li, H. Jiang. On arbitrating the power-performance tradeoff in SaaS clouds. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(10): 2648 - 2658.

[39] J. Guerra, W. Belluomini, J. Glider, K. Gupta, H. Pucha. Energy proportionality for storage: impact and feasibility. ACM SIGOPS Operating Systems Review, 2010, 44(1): pp. 35-39.

[40] D. Wong, M. Annavaram. Scaling the energy proportionality wall with KnightShift. IEEE Micro, 2013, 33(3): 28-37

[41] D. Wong. Peak efficiency aware scheduling for highly energy proportional servers. Proceedings of the 43th International Symposium on Computer Architecture (ISCA) , 2016, pp.481-492.

[42] A. Fanara, E. Haines, A. Howard. The state of energy and performance benchmarking for enterprise servers. Proceedings of the 1st TPC Technology Conference (TPCTC), 2009, pp.52-66.

[43] M. Poess, R.O. Nambiar, K. Vaid, J.M. Stephens, Jr., K. Huppler, E. Haines. Energy benchmarks: a detailed analysis. Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking (e-Energy), 2010, pp.131-140

[44] K.D. Lange. Identifying shades of green: The SPECpower benchmarks. Computer, 2009, 42(3):95-97.

[45] K.D. Lange, M.G. Tricker, J.A. Arnold, H. Block, S. Sharma. SPECpower_ssj2008: driving server energy efficiency. Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering (ICPE), 2012, pp.253-254.

[46] C. Hsu, S.W. Poole. Power signature analysis of the SPECpower_ssj2008 benchmark. Proceedings of the IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), 2011, pp.227-236.

[47] G. Varsamopoulos, Z. Abbasi, S.K.Gupta. Trends and effects of energy proportionality on server provisioning in data centers. Proceedings of the 2010 International Conference on High Performance Computing (HiPC), 2010, pp.1-11.

[48] D. Wong, M. Annavaram. Implications of high energy proportional servers on cluster-wide energy proportionality. Proceedings of the IEEE 20th International Symposium on High Performance Computer Architecture (HPCA), 2014, pp.142-153.

[49] B. Subramaniam, W. Saunders, T. Scogland, W. Feng. Trends in energy-efficient computing a perspective from the Green500, Proceedings of the 2013 International Green Computing Conference (IGCC), 2013, pp.1-8.