Parallel Computing 39 (2013) 271–279

Contents lists available at SciVerse ScienceDirect

Parallel Computing

journal homepage: www.elsevier.com/locate/parco

Analysis of Linpack and power efficiencies of the world's TOP500 supercomputers

Yuefan Deng *, Peng Zhang *, Carlos Marques, Reid Powell, Li Zhang

Department of Applied Mathematics and Statistics, State University of New York at Stony Brook, NY 11794-3600, USA

ARTICLE INFO

Article history:
Received 26 July 2010
Received in revised form 7 April 2013
Accepted 16 April 2013
Available online 24 April 2013

Keywords:
Supercomputers
Benchmark
Linpack
Power efficiency

ABSTRACT

The biannual TOP500 list of the highest performing supercomputers has chronicled, and even fostered, the development of recent supercomputing platforms. Coupled with the GREEN500 list that launched in November 2007, the TOP500 list has enabled analysis of multiple aspects of supercomputer design. In this comparative and retrospective study, we examine all of the available data contained in these two lists through November 2012 and propose a novel representation and analysis of the data, highlighting several major evolutionary trends.

Published by Elsevier B.V.

1. Introduction

While there have been many efforts to develop a comprehensive tool set to evaluate and analyze supercomputing platforms, none has achieved the popularity such as the Linpack benchmark and the TOP500 list [1]. Since 1993, `www.top500.org` has released a biannual list of the fastest 500 supercomputers that have run the Linpack benchmark. Intended as a means to facilitate comparisons between the world's most power efficient supercomputers, it also consequently fostered a sense of competition between leading vendors, driving performance improvements.

As supercomputers have reached over 1,500,000 cores and 16.3 petaflops[1] in Linpack performance in the newest release of the TOP500, November 2012 [1], energy consumption and temperature control are posing a developmental bottleneck; thus, power efficiency has garnered considerable concern in the supercomputing platform design process. In 2001, the infrastructure and energy cost of a 1U server has exceeded its purchase cost [2]. Since then, clusters like Green Destiny [3] have attempted to explore "power-aware" supercomputer designs. Inspired by the TOP500, in an effort to bring power efficiency to the forefront of supercomputer design, the GREEN500 [4] has been released as listing of the world's most powerful supercomputers, encouraging competition in the realm of power reduction. Together, the TOP500 and GREEN500 provide a rich set of data for looking into the evolution of the supercomputing landscape. Previous efforts [5–8] have noted patterns within the first 15 years of the TOP500 list. Motivation, background, and preliminary findings for the GREEN500 list appear in literature [2,4,9–11].

This paper, providing a novel presentation of the TOP500 and GREEN500 data and identifying developmental trends in supercomputer design, is organized as follows. In Section 2, we detail our techniques for analyzing the evolution of the top supercomputers. Correlating performance representations and analysis are presented in Section 3; trends are highlighted and discussed in Section 4. Finally, in Section 5, we discuss the implications of this new perspective with respect to future developments.

* Corresponding authors.
E-mail addresses: Yuefan.Deng@StonyBrook.edu (Y. Deng), Peng.Zhang@StonyBrook.edu (P. Zhang).

[1] Sequoia, BlueGene/Q, a IBM system with 1,572,864 cores.

0167-8191/$ - see front matter Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.parco.2013.04.007

2. Evaluation of supercomputers

Evaluating supercomputers is a daunting task, given the many aspects of the computers that are selectively important to individual users. We employ the most popular evaluation tools such as the TOP500 and GREEN500 while developing our own plots.

2.1. TOP500

November 2012 marks the 40th release of the TOP500 list of supercomputers. The TOP500 ranks the 500 submissions with the highest R_{max} values, a measure of maximum performance a computer (in GFLOPS) achieved when running the HPL benchmark [1]. In addition, for each submission the following responses are measured: the sites at which computers reside, computer name, year introduced, vendor, number of cores, R_{peak} for the theoretical peak performance, and power consumption in some cases. Another metric of interest, Linpack Efficiency, is determined by the unitless ratio $\frac{R_{max}}{R_{peak}}$. In summary, the web site has provided a rich set of data to further analyze and categorize the supercomputers.

2.2. GREEN500

The GREEN500 list creators [4] claim to provide "rankings of the most energy-efficient supercomputers in the world. [They] raise awareness about power consumption, promote alternative total cost of ownership performance metrics, and ensure that supercomputers only simulate climate change and not create it." Between its inception in 2007 and the November release of 2012, the GREEN500 list was a reordering of the TOP500 list in order of decreasing power efficiency measured as the maximal GFLOPS per Watt (GFLOPS/W), a metric proposed in the parallel and distributed processing community [12]. Power consumption of a system is measured by a digital power meter plugged into the system's power strip, and readings are sent to a profiling computer at a rate of 50 kHz. Newer versions of the GREEN500 list such as the Little, Open, and HPCC iterations listing different subsets of supercomputers can be found at `www.green500.org`. For the purposes of this study, we are concerned with the power efficiency of the supercomputers appearing on the TOP500 list.

2.3. Development plots

Cross referencing the power efficiency data with the Linpack efficiency data, we examine supercomputers in a two-dimensional scatter plot: power efficiency vs. Linpack efficiency. Naturally, we can derive rich information from this representation. More desirable systems are those with both highest power efficiency and highest Linpack efficiency. The other three possibilities are also obvious: low power efficiency and high Linpack efficiency, high power efficiency and low Linpack efficiency, and low efficiencies in both dimensions. Fig. 1 presents the supercomputers appearing on the November 2012

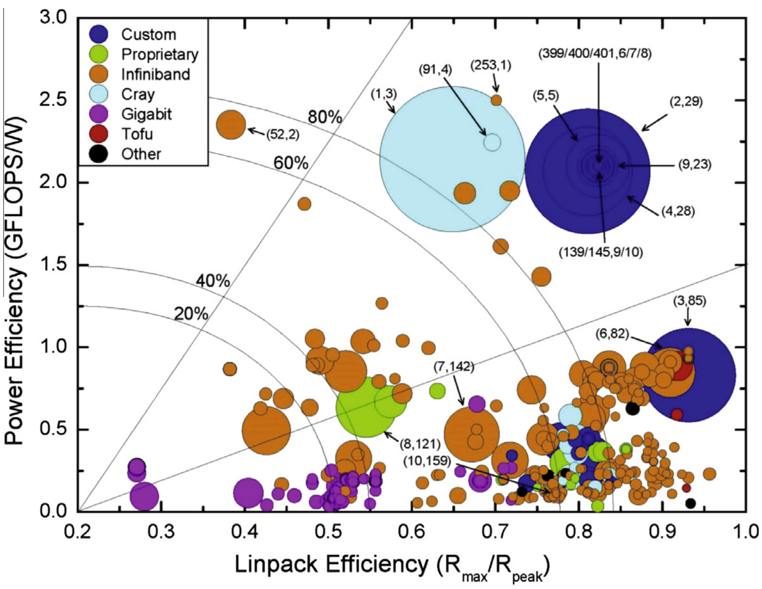

Fig. 1. Linpack Efficiency vs. Power Efficiency for the TOP500 supercomputers in 2012. The center of the circles corresponds to the respective Linpack efficiency vs power efficiency while the disc area of those circles is proportional to R_{max} performance. The color represents network type which is labeled in the legend on the top left hand corner of the plot. Top 10 supercomputers from both the TOP500 and GREEN500 are indexed with rank (i,j) respectively.

Table 1
A summary of the top ten TOP500 and GREEN500 supercomputers of November 2012, as shown in Fig. 1.

Computer name	Rank		Efficiency	
	TOP500	GREEN500	R_{Max}/R_{Peak}	GFLOPS/W
Titan	1	3	0.649	2.143
Sequoia	2	29	0.811	2.069
K computer	3	85	0.932	0.830
Mira	4	28	0.811	2.069
JUQUEEN	5	5	0.823	2.102
SuperMUC	6	82	0.910	0.846
Stampede	7	142	0.672	0.469
Tianhe-1A	8	121	0.546	0.635
Fermi	9	23	0.823	2.099
DARPA trial subset	10	159	0.779	0.424
Beacon	253	1	0.701	2.499
SANAM	52	2	0.384	2.351
Titan	1	3	0.649	2.143
Todi	91	4	0.697	2.243
JUQUEEN	5	5	0.823	2.102
BGQdev	399	6	0.823	2.101
rzuseq	400	7	0.823	2.101
BlueGene/Q	401	8	0.823	2.101
BlueGene/Q	139	9	0.823	2.101
CADMOS BG/Q	140	10	0.823	2.101

TOP500 list. The horizontal axis is the unitless Linpack efficiency, while the vertical axis is power efficiency measure in GFLOPS/W [4]; the coordinates of center of each disc indicate efficiencies while the disc area is proportional to the Linpack performance $R_{\max}$ of the corresponding supercomputer [1]. The convention of making the disc area proportional to $R_{\max}$ was so decided that we highlight the top contending supercomputers. Each label is of the form "(i,j)", where i and j are the TOP500 and GREEN500 ranks of the corresponding supercomputer, respectively. Labeled supercomputers are enumerated in Table 1 with ranking pairs in the set $\{(i,j)|i \leqslant 10 \text{ or } j \leqslant 10\}$. Each disc's color corresponds to one of the network types listed in the legend. The viewable area in the figure is a box with lower left and upper right corners corresponding to the minimal and maximal Linpack and power efficiencies, respectively. Rays from the bottom-left corner of this box divide the plot at 30° and 60° and should be viewed as guides for the eyes.

Concentric arcs highlight several features of the data. Centered at the minimal efficiencies' values, these arcs indicate the distribution of the TOP500 supercomputers' efficiencies by displaying selected percentiles of the order statistics $\{d_1, d_2, \ldots, d_i, \ldots d_{500}\}$, where d_i represents the square of the "distance" of system i from the minimal efficiency value, i.e.

$$\left(\frac{L_i - L_{min}}{r_L}\right)^2 + \left(\frac{P_i - P_{min}}{r_P}\right)^2$$

in which r_L, r_P are the respective ranges of the TOP500 supercomputers' Linpack and power efficiencies. L_{min} and P_{min} are the minimum Linpack and power efficient supercomputer respectively. Of the first 200 systems ordered by d_i (below the 40th percentile), 99 are clusters from either IBM or Hewlett–Packard. The popularity of these clusters causes a great deal of disc-overlap in Fig. 1, identical systems having identical efficiency values; whereas the greater diversity and therefore less disc-overlap of more efficient supercomputers makes the number of these machines appear greater in comparison. To compensate for some overlap we organized supercomputers with the lowest $R_{\max}$ value to the foreground while those with the highest $R_{\max}$ to the background of Fig. 1. Fig. 1 is a rich mine for exploration. For example, we noticed that: the vast majority of supercomputers are in the bottom 30° radial sectors and only two systems are above 60° radial ray, indicating the known phenomenon: venders emphasize raw performance over power efficiency. Additionally, the Linpack efficiency distribution is bi-modal concentrating at 50% and 80% for two groups of network architectures.

3. Evolution of efficiencies

In the following subsections, we first present the Linpack and power efficiencies of the TOP500 supercomputers over the last eight iterations of the TOP500 and GREEN500 lists. The set of supercomputers from each list is partitioned into subsets based on the following characteristics: vendors, parallel architecture, network interconnect, and processor family. In each case, we highlight the developmental trends of these categories. The system averages for each list release are complimented with error bars which summarizes the spread of the subset of supercomputers within a standard deviation along their respective axes. Three-dimensional representation was considered after finding significant overlap in various subsets' evolution.

Fig. 1 summarizes the latest state of supercomputer efficiency, highlighting the best-performing ones. Such methodology can be extended to all future releases.

Figs. 2–6 present ten plots with the TOP500 systems partitioned based on particular design characteristics and highlights the evolution.

4. Analysis of the evolution of supercomputer data

In this section, we analyze the data presented in Section 3 further to draw conclusions in various other aspects.

4.1. Linpack and power efficiencies for November 2012

Fig. 1 plots the 500 most powerful supercomputers in the world as of November 2012 and Table 1 lists the labeled systems. Of the 189 supercomputers using a Gigabit Ethernet interconnect, 138 can be found at the bottom half of the TOP500 with respect to d_i. This feature is clear in Fig. 1, as the 40% arc nearly partitions the TOP500 into Gigabit Ethernet and non-Gigabit Ethernet systems. The reported 53 supercomputers using Custom interconnects include BlueGene/Q, K Computer, Cray XE6/XT6, Power 775 and Hitachi SR16000 and most can be found at the top half of the TOP500 with respect to d_i. The sporadic efficiencies of supercomputers using Infiniband interconnects are apparent in Fig. 1, where nearly half of those supercomputers can be found at the top 30% with respect to d_i.

4.2. Network families

Fig. 2 shows the evolution of both Linpack and power efficiencies for various network types. Indeed the average Linpack efficiency is rather stagnant throughout the years and in different families. The main feature in Fig. 2a is how systems of Gigabit Ethernet are, on average, considerably less efficient with respect to Linpack efficiency when compared to all others. Indeed, platforms built with Gigabit Ethernet networks reside, for the most part, in the lower left quadrant, with a maximum power efficiency of 0.150 GFLOPS/W in November 2007 to 0.654 GFLOPS/W in November 2012. As of this date, no Ethernet supercomputers have achieved top 10% Linpack efficiency performance. Only a handful of Ethernet platforms exceeding the 90th percentile in power efficiency namely, two BladeCenter HS21 Clusters using Xeon quad core processors at 2.33 GHz in

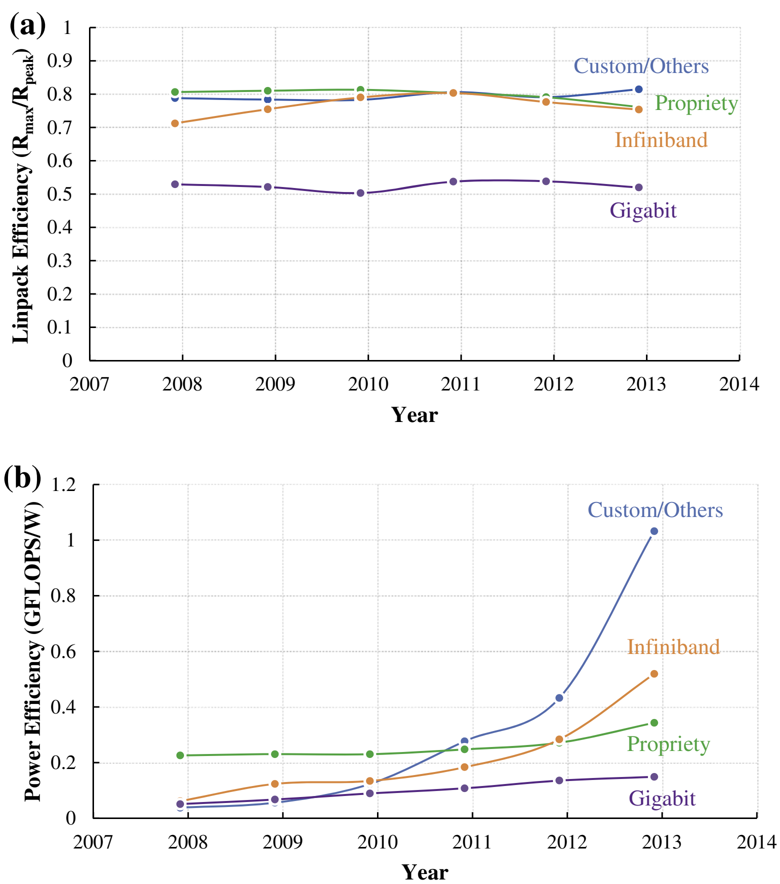

Fig. 2. (a) Linpack and (b) power efficiency evolutions for various network types.

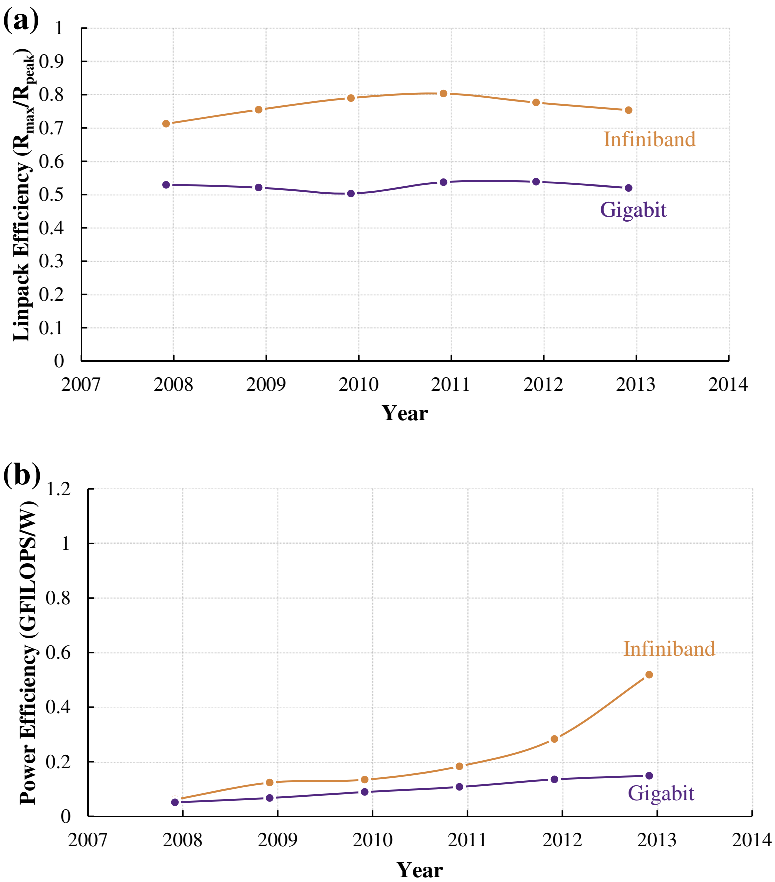

Fig. 3. (a) Linpack and (b) power efficiency evolutions for the HP Cluster Platform 3000 BL Family to compare identical supercomputers using different interconnects (Infiniband vs. Gigabit Ethernet).

2007 to 2010, an xSeries Cluster using Xeon quad core processors at 2.2 GHz in 2010, Amazon's EC2 Cluster using Xeon 8-core processors at 2.6 GHz in 2011, a HP Cluster Platform using Xeon 6-core processors at 2.66 GHz in 2012 and a HP Cluster Platform using Xeon 8-core processors at 2.6 GHz in 2012. In addition, although Gigabit Ethernet systems show consistent power efficiency as performance increases, only one supercomputer of this type exceeds 1×10^5 GFLOPS in performance.

In November 2007, no supercomputer was in the top 10% for both power and Linpack efficiencies. Since then few exceptional systems have realized this top tier, including the K Computer using SPARC64 2.0 GHz processors and a Tofu interconnect [13] in both 2010, 2011 and 2012. We have found that the only networks of supercomputers achieving top 10% performance in both dimensions were custom networks (those of BlueGene/P and K Computer), Cray networks (of Cray XK7 computer). From the November 2007 release of the TOP500 to the November 2012 release, supercomputers within the Myrinet network family gradually leave the list, being displaced by systems with non-Myrinet interconnect.

As of November 2010, systems with proprietary networks such as BlueGene/P and Q and several Cray XT systems, as well as Infiniband networks are the leaders with respect to power efficiency.

HP Supercomputer 3000BL: a Case Study of Ethernet vs. Infiniband. When examining only Hewlett–Packard's 3000BL Cluster family, we are afforded an interesting comparison of Ethernet and Infiniband networks. In Fig. 3, we plot the evolution of all HP Cluster Platform 3000BL supercomputers. The comparison shows that infiniband overshadows Linpack efficiency. Except for one cluster appearing in November 2007, each Infiniband cluster has a higher Linpack efficiency than each Ethernet cluster. On average, for any given year, HP 3000BL Cluster supercomputers using an infiniband interconnects have a higher power efficiency than identical supercomputers using Gigabit Ethernet interconnects.

4.3. Architectures

Cluster systems are improving in power efficiency, as shown in Fig. 4. Although clusters appear to have mid-range Linpack efficiency, new cluster entrants into the November 2012 TOP500 list are among the most power efficient systems on the list.

On the other hand, within the set of clusters, the subset of HP 3000BL Clusters displays the following evolutionary features. If we examine the system-center of HP Clusters, as time passes, supercomputers that are no longer powerful enough

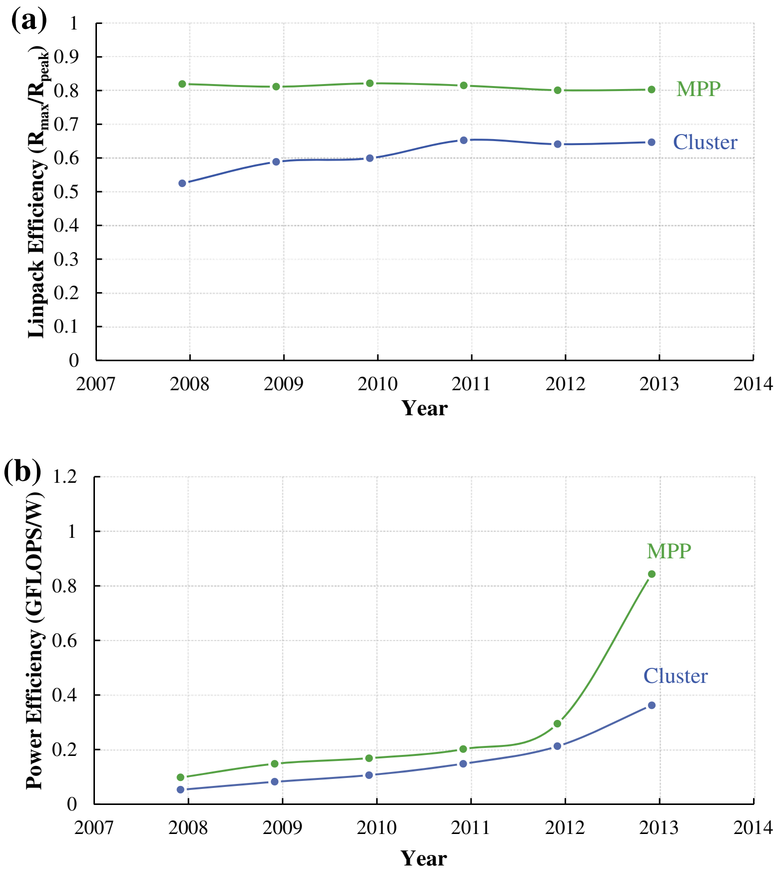

Fig. 4. (a) Linpack and (b) power efficiency evolutions for the most popular parallel architecture types.

leave the TOP500 list; however, these supercomputers are the more efficient ones. While it is reasonable to expect a certain overhead associated with supercomputer scale up, the set of MPP systems does not display such characteristics. Moreover, supercomputers like BlueGene/L, P and Q show near-constant efficiencies, regardless of system expansion, a clear indication of strong scaling.

4.4. Vendors

Fig. 5 shows the evolution of the top vendors on the TOP500 list. We see consistent improvements in power efficiency across all vendors. This is no mere coincidence for vendors that are now conscious of resource efficiency as a means of staying competitive [14].

IBM: In the November 2012 TOP500 list, IBM appears with 195 entries. Since June 2004, IBM supercomputers have consistently made up about 40% of the TOP500 systems, appearing with a variety of architectures, system models, processors, and networks. Fig. 5 shows consistent improvements in power efficiencies for IBM systems over the last eight list releases as well as Linpack efficiency except for the latest release.

Cray: although Cray shows little change in system averages of Linpack efficiencies between 2007 and November 2012, on average, it shows the greatest increase of any vendor in power efficiency in the same time period. Titan, Cray XK7 Opteron 6274 16C 2.200 GHz, a Cray Inc. system installed at Oak Ridge National Laboratory, is at the top of TOP500 list capable of Linpack efficiency 0.649 and GREEN500 list capable of power efficiency 2.143 GFLOPS/W. Cray supercomputers also have the smallest standard deviation with respect to Linpack efficiency which is centered around an impressive 80% Linpack efficiency.

SGI: on average, SGI is the leading vendor in Linpack efficiency and has remained competitive and improving power efficiency since 2007.

HP and Dell: HP and Dell have shown steady improvements in both Linpack and power efficiencies. Although the system average for Hewlett–Packard systems has not increased dramatically, new highly-efficient entrants cause the vertical bars for 2008 and beyond to show an increased rate of growth.

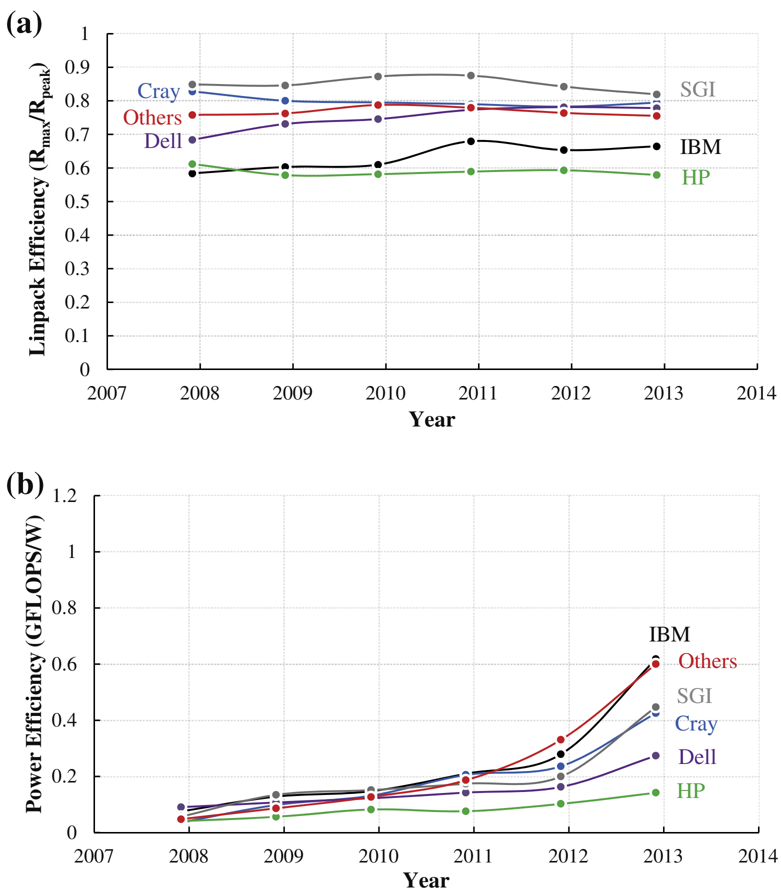

Fig. 5. (a) Linpack and (b) power efficiency evolutions for the leading vendors.

Others: vendors not part of the five most frequently appearing are entitled "Others." Companies such as Fujitsu, Hitachi, NEC, as well as self-made platforms are part of the entries represented here. While this group has shown consistent yet slight improvement in Linpack efficiency, November 2010 marks the first decrease in Linpack efficiency which continues to 2012. However, this group also has the greatest power efficiency change over the dominating vendors and the standard deviation suggests an improvement of power efficiency at an accelerated rate.

Among the "Others" in Fig. 5, Fuijitsu has become very competitive with their release of the K computer in 2012. Fujitsu systems appear on the first TOP500 list with single processor computers. By list two, a transition to MPP begins. Between November 1993 and November 2001, Fujitsu's performance share is dominated by their VPP family of systems, until another transition to SMP, constellation, and cluster architectures occurs.

Hitachi platforms begin with a trajectory similar to that of Fujitsu, starting in November 1993 with single processor vector systems and making a transition to MPP systems with crossbar interconnect.

NEC follows this progression as well with the notable exception of several platforms with constellation architecture. Additionally, the then immediate impact of NEC's Earth Simulator.

4.5. Processor family

The major processor families represented on the TOP500 list are AMD's x86-64, Intel's EM64T and IA-64, and Power; the evolution of the efficiencies for systems built with these, as well as those using other processor families, are plotted in Fig. 6.

The "Others" frame shows varied distributions over the dates considered. In addition to the relatively small size of the set, elements of "Others" occupied two extremes. In June 2009, only three systems used processors not from the four major processor families. The GRAPE-DR Cluster also had the minimal Linpack efficiency value (0.26) for that date, yet ranked fifth overall on the GREEN500 List in power efficiency (428.91 GFLOPS/W). The remaining two systems, Earth Simulator – an NEC Vector system, and a Fujitsu Cluster, ranked 22 and 28 on the June 2009 TOP500 list, having Linpack efficiency over 0.91; however, they performed at 51.00 and 66.85 GFLOPS/W, respectively, on the GREEN500 list. In the latest release, the GRAPE cluster no longer appears, leaving the set "Others" in the upper-half of the Linpack efficiency and the lower-half of the power efficiency distributions.

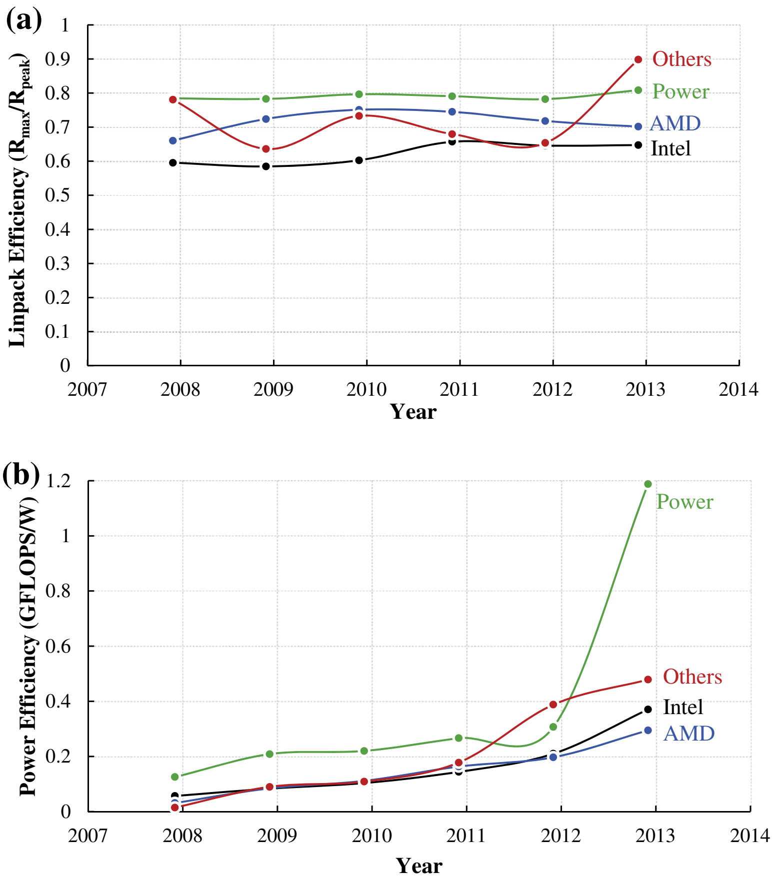

Fig. 6. (a) Linpack and (b) power efficiency evolutions for leading processor families.

Of the 61 entries using AMD x86 processors, over half are Cray XT and XE MPP supercomputers, one of which is the number one ranked Titan, while 18 others are in the top 100. The 34 non-Cray supercomputers using AMD x86 chips are all clusters with a similar, though less skewed, distribution over the TOP500 rankings. The power efficiency of the AMD-based supercomputers shows consistent increase over the last eight releases.

PowerPC systems have made the largest increase in power efficiency as of the latest release, whereas Intel IA-64 systems have reached a plateau and are disappearing from the list.

Among the largest competitors, PowerPC, Intel's EM64T and IA-64, AMD x86-64, Intel's EM64T is dominant with nearly 80% system share in latest release, with consistent increase in both power and Linpack efficiency over the last three releases.

5. Conclusions

This analysis has highlighted many performance properties of the TOP500 systems. It also points out other areas to explore, such as increasing the number of characteristics noted for each system. For example, as the TOP500 has begun to record number of cores per processor, properties like network material (copper, fiber, etc.) would offer additional comparisons of the TOP500 systems by exploring to what extent each design parameter affects Linpack performance and power efficiency.

We have presented and analyzed the correlations of the Linpack and power efficiencies of the past ten listings of 500 supercomputers. We group such supercomputers according to their architectures and original design manufacturers and analyze conveniently the time evolutions. Our analysis has for the first time revealed, in many cases reaffirmed, the comparative properties of such supercomputers. First, with Ethernet interconnect systems are low in both power and Linpack efficiencies and they have made slight improvements over the nine-listing periods, while systems not using Infiniband, Myrinet or Ethernet the best in both efficiencies. Myrinet and Infiniband are the middle-level performers. Second, Cray and IBM are leaders of power and Linpack efficiencies and both made significant progress in power efficiency over the ten listing periods. HP has performed the poorest with low power efficiency low Linpack efficiency, but has shown examples of improvement with the latest release. Third, PowerPC processors perform the best with high efficiencies in both while Intel IA-64 has the lowest power efficiency (and decent Linpack efficiency). Finally, overall improvements to power and Linpack efficiency have been observed throughout several listings. These improvements spark competition and overall quality for supercomputers. Among the trends observed, new and emerging technologies have been a key component for displacing ranks in the TOP500 and GREEN500 lists. IBM has been a leader in both power and Linpack efficiency throughout our analysis and hence their presence is felt in the TOP500 and GREEN500 lists. To stay competitive companies will need to focus on developing and implementing new technologies of processors, networks and overall architectures.

References

[1] TOP500 Supercomputer Sites, 2012. <http://www.top500.org>.
[2] W.-C. Feng, T. Scogland, The GREEN500 list: year one, in: IPDPS' 09: Proceedings of the 2009 IEEE International Symposium on Parallel & Distributed Processing, IEEE Computer Society, Washington, DC, USA, 2009, pp. 1–7.
[3] W.-C. Feng, Global climate warming? Yes in the machine room, in: Clusters and Computational Grids for Scientific, Computing, 2006.
[4] GREEN500 Supercomputers, 2012. <http://www.green500.org>.
[5] D.G. Feitelson, On the interpretation of TOP500 data, Int. J. High Perform. Comput. Appl. 13 (1999) 146–153.
[6] J. Dongarra, H. Meuer, H. Simon, E. Strohmaier, Biannual Top-500 computer lists track changing environments for scientific computing, SIAM News 34, Number 9. (2001)
[7] D.G. Feitelson, The super computer industry in light of the TOP500 data, Comput. Sci. Eng. 7 (2005) 42–47.
[8] H.W. Meuer, The TOP500 project: looking back over 15 years of supercomputing experience, Informatik-Spektrum 31 (2008) 203–222.
[9] S. Sharma, C.-H. Hsu, W.-C. Feng, Making a case for a GREEN500 list, in: Proc. of the Workshop on High-Performance, Power-Aware Computing, 2006.
[10] W.-C. Feng, K. Cameron, The GREEN500 list: encouraging sustainable supercomputing, Computer 40 (2007) 50–55.
[11] W.-C. Feng, X. Feng, R. Ce, Green supercomputing comes of age, IT Professional 10 (2008) 17–23.
[12] C.-H. Hsu, W.-C. Feng, J.S. Archuleta, Towards efficient supercomputing: a quest for the right metric, in: International Parallel and Distributed Processing Symposium, vol. 12, 2005, p. 230a.
[13] Y. Ajima, T. Inoue, S. Hiramoto, T Shimizu, The Tofu interconnect, in: 19th Annual IEEE Symposium on High Performance Interconnects, 2011.
[14] The Green Grid Site, 2012. <http://www.thegreengrid.org>.