The Green500 List: Year One

W. Feng and T. Scogland
{feng,njustn}@cs.vt.edu
Department of Computer Science
Virginia Tech

Abstract

The latest release of the Green500 List in November 2008 marked its one-year anniversary. As such, this paper aims to provide an analysis and retrospective examination of the Green500 List in order to understand how the list has evolved and what trends have emerged. In addition, we present community feedback on the Green500 List, particularly from two Green500 birds-of-a-feather (BoF) sessions at the International Supercomputing Conference in June 2008 and SC|08 in November 2008, respectively.

1. Introduction

At the *First IEEE Workshop on High-Performance, Power-Aware Computing* in 2005, the keynote address provided the initial impetus for a Green500 List [7]. However, back then, "performance" was still only synonmous with "speed" (as measured in floating-point operations per second, i.e., FLOPS), so the mere notion of such a list was viewed as heresy by the greater high-end computing (HEC) community. Yet the authors of [18] forged ahead and made a case for a Green500 List anyway.

It was not until an invited talk in September 2006 at *Clusters and Computational Grids for Scientific Computing* [8] that the idea of the Green500 List took root and garnered significant support and encouragement by its attendees, who spanned industry, government labs, and academia. By November 2007 at ACM/IEEE SC07, the Green500 web site was established and a request for comments (RFC) issued. A year later in November 2008, the first official Green500 List debuted at ACM/IEEE SC08 with the express purpose of providing a ranking of the top supercomputers in the world by *power efficiency* rather than speed. With the latest release of the Green500 in November 2008, an entire year's worth of data is now available. This paper tracks the progress of power-efficient supercomputing on the Green500, analyzes trends in the Green500, and reflects upon the implications of these trends.

This work was supported in part by NSF grant CCF-0848670.

2. Background

For the high-end computing (HEC) community, the Green500 List serves as a complementary view to the Top500 List, where the performance metric of interest is *power efficiency* rather than *speed*. The rationale for embracing *green supercomputing* comes from three orthogonal directions: (1) a revived social and cultural conscience in the midst of climate and environmental concerns, (2) enhanced hardware reliability due to lower power consumption and lower temperature, and (3) the economic bottomline that annual infrastructure costs as well as annual energy costs each surpassed annual server purchases in 2004 and 2008, respectively, as shown in Figure 1. Collectively, the above contribute to reducing the overall total cost of ownership (TCO).

The low-power "Green Destiny" cluster [6], [9] aggressively addressed the above issues by consolidating common hardware infrastructure via blade technology and by significantly reducing per-node power consumption. The judicious use of blades by Green Destiny allowed compute nodes to be packed more densely while also *simultaneously reducing power density*, i.e., watts per square foot. Although the HEC community understood the benefits of the former, the latter was largely ignored. Consequently, while the

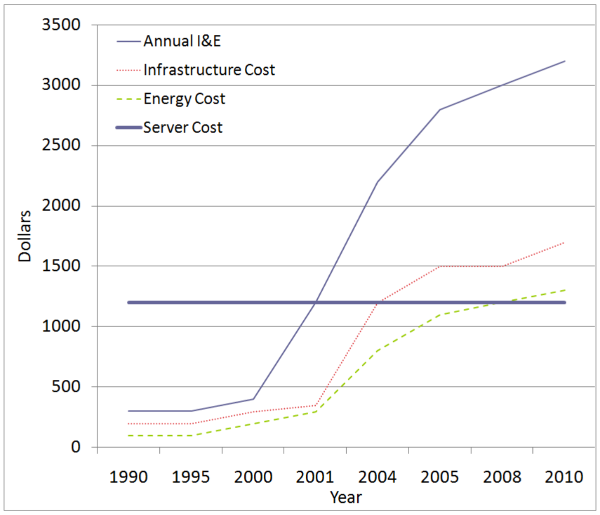

Figure 1. Annual Amortized Costs in the Data Center [4]

advent of server blades created more densely-packed components, it also resulted in a substantial and faster-than-expected increase in the power density, as shown in Figure 3.

Initially, managing this increased heat density simply entailed increasing fan speed or fan quantity or adopting heat pipes in order to more aggressively dissipate the heat. However, with the power density of contemporary processors exceeding that of a nuclear reactor, as shown in Figure 2, this approach quickly reached its limits. Newer high-end computers already adopt alternative cooling strategies such as chimneys [13], indoor water cooling [17], immersion in inert gas [5], and natural *open-air* cooling [3]. Alas, these new approaches do *not* come free. Again, as noted in Figure 1, the annual financial burden to operate and maintain the infrastructure that houses these high-end compute servers surpassed the annual cost of the new servers themselves in 2004.

In addition to the economic impact of excessive energy consumption, the environmental impact has also affected the HEC community — including datacenters, ranging from the Google, Microsoft, and Yahoo! to the more pedestrian industries of pharmaceuticals, insurance, and manufacturing [14]. While such datacenters and HEC facilities were previously immune to egregious energy consumption, this is no longer the case.

To this end, major players in the datacenter and HEC markets often negotiate energy deals with electricity suppliers to build or upgrade power substations, near or immediately next to, their computing facilities.

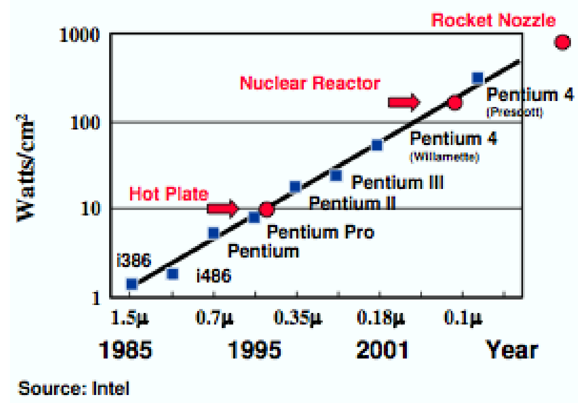

Figure 2. Moore's Law for Power Consumption [12]

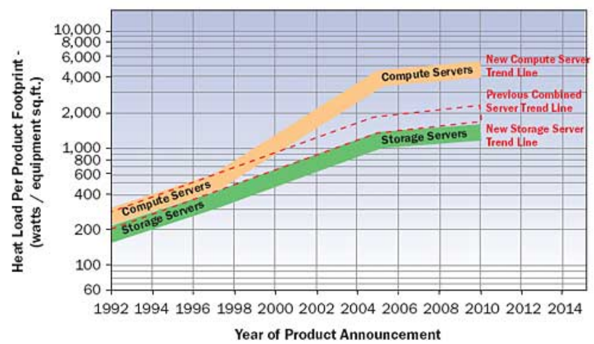

Figure 3. ASHRAE Heat-Density Projection [2], [16]

Examples include, but are certainly not limited to, the following: Los Alamos National Laboratory [23], Oak Ridge National Laboratory [24], National Center for Supercomputing Applications [15], and National Security Agency [10], [19]. Alternatively, when not enough power infrastructure can be built at or near computing facilities, many companies *move* their computing facilities to the power source, e.g., Google [1], [14] and Microsoft [27].

However, the notion of simply "feeding the beast" is not a sustainable solution as the demand for energy continues to grow rapidly. Instead, the power efficiency of these systems needs to be increased to more effectively make use of electricity, as existing centers are constrained by the physical wattage entering the building. This issue is particularly important to those centers located in megacities that are already struggling with energy demand like New York, Tokyo, and Shanghai. To this end, power-efficient computing consortiums and projects are attracting the attention of the populace and supercomputer vendors to reduce this burden and to more efficiently utilize the energy available. A sampling of the many programs that have started are Energy Star [21], SPECPower [25], Green Data Centres [22], ClimateSavers [26], 80-Plus [20], and of course, the Green500 [11].

3. An Analysis of the Green500

In this section, we present an analysis of the data that we have collected over the first year of the Green500.

3.1. Overall Power Efficiency

Since the first official release of the Green500 in November 2007, the overall power efficiency of the machines on the Green500 has improved dramatically, particularly the top half of the list. Figure 4 shows the maximum and average power efficiency of each of the Green500 releases in its first year of existence.

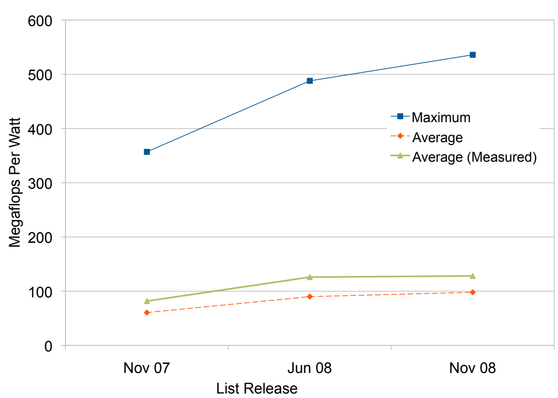

Figure 4. Maximum and Average Power Efficiency per Green500 Release

The majority of the improvement, however, has been confined to the very top of the list, where a relatively small set of machines pulls the mean power efficiency well above the median power efficiency, as shown in Figure 5. In each list, the distance between the mean and median has increased from the previous release, as shown in Figure 4. In November 2007, the difference is only 9.7 MFLOPS/W; in June 2008, it increases to 26.49 MFLOPS/W; in November 2008, the difference is 32.91 MFLOPS/W, which is also shown in Figure 5.

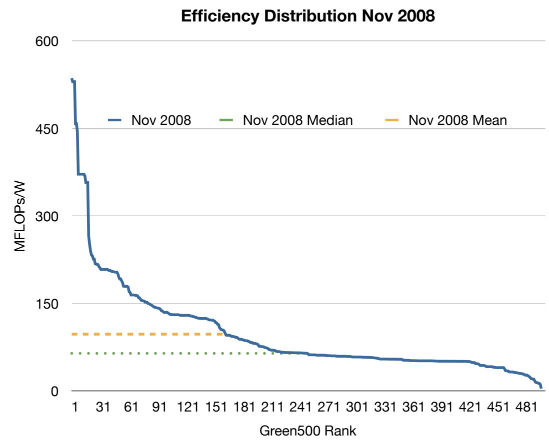

Figure 5. Power Efficiency by Green500 Rank

For the November 2007 list, the most efficient machine on the Green500 registered 357 Mflops/W. This improved to 488 Mflops/W in June 2008 and then 536 Mflops/W in November 2008. *Thus, in year one of the Green500, the power efficiency of the top-ranked machine improved by more than 50%.*

The average power efficiency of the Green500 also improved substantially from 60 Mflops/W in November 2007 to 90 Mflops/W in June 2008 to 98 Mflops/W in November 2008, which translates into a *63% improvement in the average power efficiency of the Green500.*

If we consider the aforementioned improvements in the context of Moore's Law and its corollaries (i.e., doubling of a metric every 18 months or 75% improvement per year), the Green500 List shows initial indications of a similar tracking to Moore's Law, at least with respect to average power efficiency. Perhaps the main reason why the improvement falls short of 75% is due to the fact that nearly half the machines on the list still do *not* report their actual power numbers, thus artifically suppressing actual improvements in power efficiency. If we only consider the reporting Green500 machines, the average power efficiency improved by approximately 75% in 2008.

3.2. Power Efficiency vs. Speed

Figures 6 and 7 plot power efficiency (MFLOPS/W) versus speed (Top500 Rank) for June 2008 and November 2008, respectively. The overall distribution in both graphs is quite astounding. Quite a few machines at the top of each list are at the bottom of the other. For instance, on the June 2008 list, the Fraunhofer ITWM

BladeCenter QS22 Cluster with PowerXCell 8i 3.2 Ghz achieves a whopping 488.14 MFLOPS/W to place it at #1 on the Green500 List but is a mere #464 on the Top500 List. See the upper- and right-most circle in Figure 6.

In addition, with the circles denoting specialized (custom) architectures such as Blue Gene/P, Blue Gene/L, and BladeCenter QS22 with PowerXCell, we note that specialized architectures are significantly more power efficient than their commodity counterparts. In fact, the green circles that are near 500 MFLOPS/W are the BladeCenter QS22 Clusters with PowerXCell processors. The next group of green circles that occur at roughly 375 MFLOPS/W are the Blue Gene/P machines. Finally, the last group of green circles that reside at 210 MFLOPS/W are for the Blue Gene/L machines.

Also notable in both Figures 6 and 7 is that commodity architectures have caught up to the specialized architectures, e.g., see the "x" data points between the Blue Gene/L and Blue Gene/P solutions. Additional information on this is presented in Section 3.4.

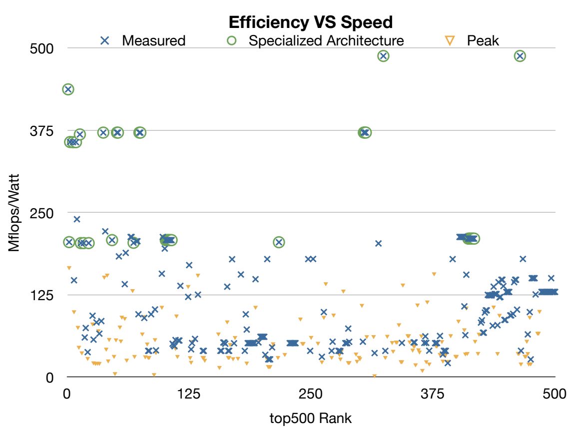

Figure 6. MFLOPS/W vs. Top500 Rank, June 2008

3.3. Power vs. MFLOPS/W

Figure 8 shows that the "Jaguar - Cray XT5" machine and the "Thunder - Intel Itanium2 Tiger4" machine consume the largest and second largest amount of total power of all the supercomputers on the Green500 List, respectively. While Jaguar makes up for its large power consumption by delivering significant enough performance to elevate its Green500 ranking to #79, Thunder does *not*, and consequently, wins the "Lanterne Rouge" for power efficiency by ranking #500

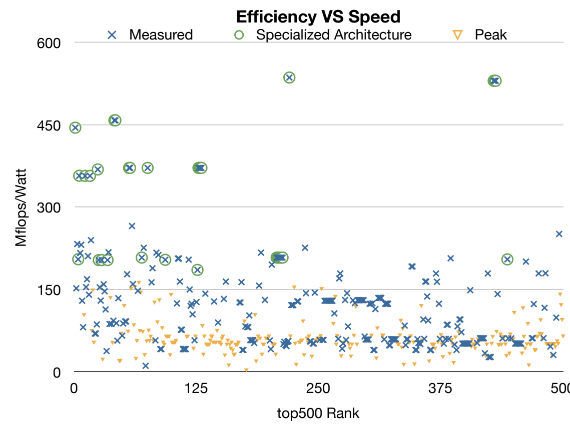

Figure 7. MFLOPS/W vs. Top500 Rank, Nov 2008

on the Green500. In contrast, the BladeCenter QS22, Blue Gene/P and Blue Gene/L machines continue to lead the way in power efficiency, as shown in Figure **??**.

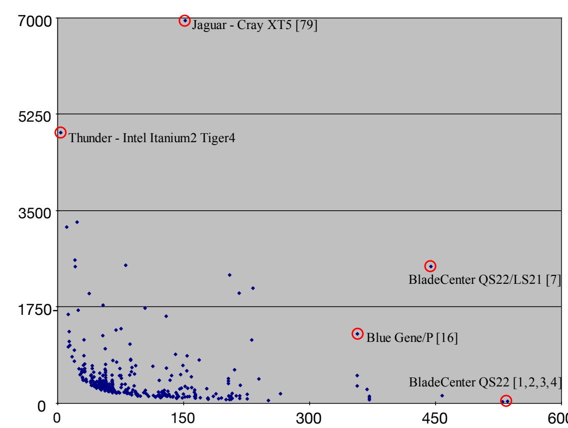

Figure 8. Total Power Consumption vs. MFLOPS/W

3.4. Commodity Catches Custom

Over the past year, IBM Blue Gene/L and Blue Gene/P and IBM BladeCenter QS22 machines have set the bar for power efficiency. With respect to the custom-built BlueGene machines, the L models deliver approximately 210 Mflops/W while the P models deliver about 370 Mflops/W.

As of the June 2008 list, the energy efficiency of commodity machines based on Intel's 45-nm low-power quad-core Xeon reached as high as

265 Mflops/W, thus surpassing the IBM Blue Gene/L machines, which debuted in November 2004. These data points can also be seen in Figures 6 and 7.

3.5. Power Consumption by Country

As expected, the total power consumption of the machines on the Green500 list continues to increase. Since the last Green500 announcement in June 2008, the total power consumption grew 17.5% from 170.6 MW to 200.4 MW. Figure 9 shows the top 10 countries ranked by total power consumption. The range of total power consumption by the top 10 changed from [2.2, 97.6] MW in June 2008 to [2.0, 126.0] MW in November 2008.

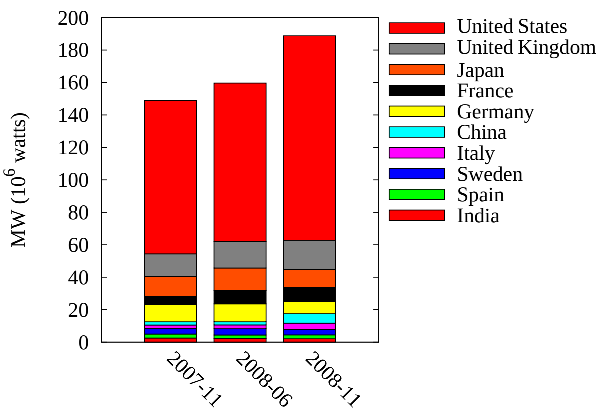

Figure 9. Top Ten Countries Ranked by Aggregate Power Consumption

The largest power consumer is the United States. This not surprising since 58.2% of all the systems listed on the Green500 are from the U.S. Nevertheless, its aggregate power consumption grew 29.1% from 97.6 MW to 126.0 MW over the last six months — the second largest increase of any country. This despite only a 12.8% increase in the number of installations. The mean power consumption per site increased by 14.5% from 378.2 KW to 433.0 KW. Thus, the increase in power consumption results from a nearly equal combination of the above two factors.

The country with the greatest relative increase in total power consumption is China. Between June 2008 and November 2008, the power consumed by the Chinese machines on the Green500 increased from just under 1.9 MW to 5.8 MW, an increase of 298%. Although the number of installation sites grew by a healthy 25%, the majority of the increase came from an dramatic increase in capacity per site with the mean power consumption per site increasing from 161.9 KW to 386.1 KW, an increase of 238.4%. Thus, for China, the main contributor to the increase in power consumption is the installation of more powerful machines with a lesser contribution from the increase in the number of sites.

Of the countries shown in Figure 9, number of installations ranging from 12% to 33%. However, the greatest increase in the number of installations over the last six months come from Poland, Mexico, Italy, and Brazil, who all doubled the number of machines.

3.6. Overall Observations

The FLOPS/watt distribution is still skewed from top to bottom. While the energy efficiency of the top-ranked Green500 machine has improved by nearly 180,000,000 FLOPS/watt since November 2007, the bottom-ranked Green500 supercomputer has only improved by approximately 400,000 FLOPS/watt.

Overall, the Mflops/watt distribution of the Green500 is skewed with only 28% of the machines achieving more than 100 Mflops/watt (up from 14% last November 2007).

Finally, 15% of the most power-hungry supercomputers on the Green500 consume more than half the total power of all the supercomputers on the Green500.

4. Community Feedback

During the first official year of the Green500 from November 2007 to November 2008, the organizers of the Green500 List were fortunate to be able to hold two Green500 birds-of-a-feather (BoF) sessions, one at the International Supercomputing Conference in June 2008 and one at SC08 in November 2008.

The two BoFs sought to solicit feedback from the high-end computing (HEC) community. Rather than have the organizers of the Green500 dictate what the Green500 List should be, the organizers sought to give a voice to the community and have the community help to drive the list.

Overall, the recommendations derived by the Green500 BoFs include the following:

1) *What to Measure?* Continue to focus on the power efficiency of the machine and do *not* consider the cooling facilities or the accompanying infrastructure.
2) *What Machines to Include?* Expand the Green500 List to be more than just a re-ordering of the

TOP500 List ranked by power efficiency in FLOPS per watt.

3) *What Benchmark(s) to Use?* Identify a different benchmark (or benchmark suite) that more appropriately stresses and captures the performance of the entire machine.
4) *What Metric to Use?* Use a different performance metric to evaluate energy efficiency.
5) *How to Present?* Enable different ways to view the Green500 data.

What to Measure? While including cooling and other infrastructure into the power measurement for the Green500 List was requested initially by several in the HEC community, further discussion made it clear that doing so would *not* be advisable. Reasons include (1) a desire to separate rating the power efficiency of the machine from the power efficiency of the cooling system, (2) *existing* difficulty in obtaining all the actual power numbers for machines on the list, difficulty that would be exacerbated by asking for cooling numbers, and (3) difficulty in identifying the cooling footprint for a single supercomputing machine as many institutions use the same cooling system to cool multiple supercomputing machines.

What Machines to Include? Two specific recommendations came out of this discussion: (1) lower the entry bar for the Green500 to a commodity supercomputer rather than a TOP500 supercomputer, and (2) create a sublist called the "Little Green500" that allows for supercomputers — which were either not fast enough to make the TOP500 list or could not abide by the arguably stricter TOP500 run rules that do not allow mixed precision floating point, e.g., GPGPU — to participate in being green.

What Benchmarks to Use? This issue was perhaps the most controversial in both BoFs. What it boils down to is "doing the right thing" versus "being pragmatic." While it has long been acknowledged that the Linpack benchmark of the TOP500 is not the "be all end all" benchmark, alternative benchmarks such as the HPC Challenge Benchmarks have arguably failed to "catch on." Why? Pragmatically, running the Linpack benchmark is relatively easy to do and tune. In contrast, the HPC Challenge Benchmarks require such a significant time investment that there are only 198 machines listed, of which many are missing numbers for certain benchmarks. In short, while the Green500 seeks to stay true to its ideals, the Green500 List must also take a pragmatic approach to ensure continued participation. For these reasons, the Green500 List will likely remain with Linpack for the foreseeable future but create an exploratory sublist, like the aforementioned "Little Green500."

What Metric to Use? For now, this metric will stay with FLOPS per watt, particularly in light of its seemingly widespread acceptance now. Two other performance metrics, however, deserve consideration as well: the energy-delay (ED) product and the Green Computing Performance Index (GCPI).

How to Present? Currently, there really exists only one way to view the list — in rank order. However, vendors and HEC researchers alike have asked for alternative views to the data, e.g., sorted by vendor, sorted by processor, and sorted by machine type.

5. Conclusion

The November 2008 release of the Green500 List marked its one-year anniversary. With a year's worth of collected data, this paper sought to study the Green500 List in more detail in order to understand how the list has been evolving and what trends that we should be paying attention to. We summarize our findings in this paper as follows:

- The overall power efficiency (on average) has improved in a manner that tracks with Moore's Law, i.e., the average power efficiency of the Green500 doubles every 18 months.
- The power efficiency of commodity machines on the Green500 now surpass the power efficiency of the custom-built IBM BlueGene/L.
- China, India, and the United States are by far the largest power consumers. In addition, China shows astronomical growth in its power appetite.

Acknowledgements

The authors would like to thank Jeremy S. Archuleta and Mark K. Gardner for their preliminary contributions to the paper as well as Kirk Cameron for his support in sustaining the Green500 List.

Thanks to Supermicro for generously sponsoring the Green500 List and for contributing to the green movement.

Finally, the authors also wish to thank the Green500 community for their continued support and feedback, particularly those participants at the Green500 BoFs in 2008.

References

[1] M. Ahmed. Google Search Finds Seafaring Solution. In *The Times*, September 2008. http://technology.timesonline.co.uk/tol/news/tech_and_web/the_web/article4753389.ece.

[2] American Society of Heating, Refrigerating and Air-Conditioning Engineers. *Datacom Equipment Power Trends and Cooling Applications*. ASHRAE, USA, January 2005.

[3] Don Atwood and John G. Miner. Reducing Data Center Cost with an Air Economizer. White Paper: Intel Corporation, August 2008.

[4] C. Belady. In the Data Center, Power and Cooling Cost More Than the IT Equipment It Supports. *Electronics Cooling Magazine*, 13(1), May 2007.

[5] Cray. Cray Unveils High-Efficiency "Green" Cooling Technology for the Petascale Computing Era. Press Release: Cray Corporation, August 2008.

[6] W. Feng. Making a Case for Efficient Supercomputing. *ACM Queue*, 1(7):54–64, October 2003.

[7] W. Feng. The Evolution of Power-Aware, High-Performance Clusters: From the Datacenter to the Desktop. In *First IEEE Workshop on High-Performance Power-Aware Computing*, April 2005.

[8] W. Feng. Global Climate Warming? Yes ... In the Machine Room. In *Clusters and Computational Grids for Scientific Computing*, September 2006.

[9] W. Feng, M. Warren, and E. Weigle. The Bladed Beowulf: A Cost-Effective Alternative to Traditional Beowulfs. In *Proc. of the IEEE Int'l Conf. on Cluster Computing*, September 2002.

[10] S. Gorman. NSA Risking Electrical Overload. In *The Baltimore Sun*, August 2006.

[11] Green500.org. The Green500 List, 2008. http://www.green500.org.

[12] C. Hsu and W. Feng. A power-aware run-time system for high-performance computing. In *Proc. of SC2005: The IEEE/ACM International Conference on High-Performance Computing, Networking, Storage, and Analytics*, November 2005.

[13] D. Kikinis et al. Method and Apparatus for Waste Heat Removal from a Computer Enclosure, Mar 1997. US Patent 5,613,906.

[14] J. Markoff and S. Hansell. Hiding in Plain Sight, Google Seeks More Power. In *The New York Times*, June 2006. http://www.nytimes.com/2006/06/14/technology/14search.html.

[15] J. Meichi. Building the Data Center of the Future. In *NCSA Newsletters*, November 2008. http://www.ncsa.uiuc.edu/News/Stories/PCF/.

[16] R. Schmidt and D. Beaty. ASHRAE Committee Formed to Establish Thermal Guidelines for Datacom Facilities. *Electronics Cooling Magazine*, Feb 2005.

[17] SGI. Altix Ice, 2008. http://www.sgi.com/products/servers/altix/ice/.

[18] S. Sharma, C.H. Hsu, and W. Feng. Making a Case for a Green500 List. In *Proc. of the Workshop on High-Performance, Power-Aware Computing, Rhodes, Greece, April*, 2006.

[19] NSA Electrical Power Upgrade, January 2008. http://cryptome.org/nsa010208.htm.

[20] 80 PLUS, 2008. http://www.80plus.org/.

[21] Energy Star, 2008. http://www.energystar.gov/.

[22] Green Data Centres, 2008. http://www.greendatacentres.co.uk/.

[23] Existing and Potential Energy Demands for Los Alamos National Laboratory, Compiled by Concerned Citizens for Nuclear Safety, September 2003. http://www.nuclearactive.org/graphix/LANLelectricity.pdf.

[24] ASCR Computing News Roundup, October 2007. http://www.sc.doe.gov/ascr/News/MonthlyNewsRoundup10-07.html.

[25] Standard Performance Evaluation Corporation, SPECPower, 2008. http://www.spec.org/power_ssj2008/.

[26] ClimateSavers, 2008. http://www.worldwildlife.org/climate/climatesavers2.html.

[27] A. Vance. Microsoft's Data Center Offensive Sounds Offensive. In *The Register*, June 2006. http://www.theregister.co.uk/2008/03/03/microsoft_data_center_blowout.