ASSESSING TRENDS OVER TIME IN PERFORMANCE, COSTS, AND ENERGY USE FOR SERVERS

Jonathan G. Koomey*, Christian Belady†, Michael Patterson**, Anthony Santos**,
Klaus-Dieter Lange††

* Lawrence Berkeley National Laboratory and Stanford University

†Microsoft Corporation

**Intel Corporation

††Hewlett-Packard Corporation

Contact: JGKoomey@stanford.edu, http://www.koomey.com

Final report to Microsoft Corporation and Intel Corporation

Released on the web: August 17, 2009

EXECUTIVE SUMMARY

Data centers are the heart of the global economy. In the mid-1990s, the costs of these large computing facilities were dominated by the costs of the information technology (IT) equipment that they housed, but no longer. As the electrical power used by IT equipment per dollar of equipment cost has increased, the annualized facility costs associated with powering and cooling IT equipment has in some cases grown to equal the annualized capital costs of the IT equipment itself. The trend towards ever more electricity-intensive IT equipment continues, which means that direct IT equipment acquisition costs will be a less important determinant of the economics of computing services in the future. Consider **Figure ES-1**, which shows the importance of different data center cost components as a function of power use per thousand dollars of server cost. If power per server cost continues to increase, the indirect power-related infrastructure costs will soon exceed the annualized direct cost of purchasing the IT equipment in the data center.

Ken Brill of the Uptime Institute has called these trends "the economic breakdown of Moore's Law", highlighting the growing importance of power-related indirect costs to the overall economics of information technology. The industry has in general assumed that the cost reductions and growth in computing speed related to Moore's law would continue unabated for years to come, and this may be true at the level of individual server systems. Unfortunately, far too little attention has been paid to the true total costs for data center facilities, in which the power-related indirect costs threaten to slow the cost reductions from Moore's law.

These trends have important implications for the design, construction and operation of data centers. The companies delivering so-called "cloud computing" services have been aware of these economic trends for years, though the sophistication of their responses to them has varied. Most other companies that own data centers, for which computing is not their core business, have significantly lagged behind the vertically organized large-scale computing providers in addressing these issues.

There are technical solutions for improving data center efficiency but the most important and most neglected solutions relate to institutional changes that can help companies focus on reducing the total costs of computing services. The first steps, of course, are to measure costs in a comprehensive way, eliminate institutional impediments, and reward those who successfully reduce these costs.

This article assesses trends in servers to help explain the driving forces affecting data center costs. It develops and documents detailed examples from available data, estimating costs and correcting them for inflation, and explaining the implications of the results.

Figure ES-2 summarizes some of the key technical findings of this study for the examples investigated here. With one exception, performance per server and performance per thousand dollars of purchase cost double every two years or so, which tracks the typical doubling time for transistors on a chip predicted by the most recent incarnation of Moore's law.

Power used per thousand dollars of server acquisition cost is the most important driver of power and cooling costs in data centers, because all of the costs to purchase electricity and almost all of the facility costs are directly related to the power use of IT equipment. The power-related capital costs for cooling, backup power, and power distribution are substantial (roughly $25,000 per kW of IT power use), and together with the electricity costs account for roughly half of total annualized costs in typical data centers.

Power used per thousand dollars of server cost can be broken down into two components: performance per dollar of server cost and performance per watt. Performance per dollar of server cost has in all cases examined here been increasing more rapidly than performance per watt in recent years, and this trend leads to increases in the power use per server cost.

The result is that the indirect costs for cooling and power distribution (which are directly related to the power use per dollar of server acquisition cost) start to offset the performance related benefits of Moore's law. A purchaser of servers who does not assess the total cost for purchasing new servers but instead focuses solely on performance per dollar of server acquisition cost will invariably overestimate the benefits from buying more computing power. This mismatch between costs and benefits is the primary reason why institutional changes are needed in most data center operations, which traditionally have separate budgets for the IT and facilities departments. IT departments generally don't pay the electric bill or the costs to build cooling or power distribution capacity, so they don't demand high efficiency servers, because the costs for inefficiency come out of someone else's budget. Cloud computing providers have generally been ahead of the rest of the industry in fixing these misplaced incentives, which is one economic advantage that they hold compared to in-house corporate data center operators.

There are some indications that the industry's focus on reducing power use of servers since 2006 has been paying off, although more research is needed to confirm this finding. Three of our case studies (the DL360, the DL380, and the LBNL cluster computing examples) show slowing growth in recent years for power use, resulting in longer doubling times for power use per real server cost than in the other examples.

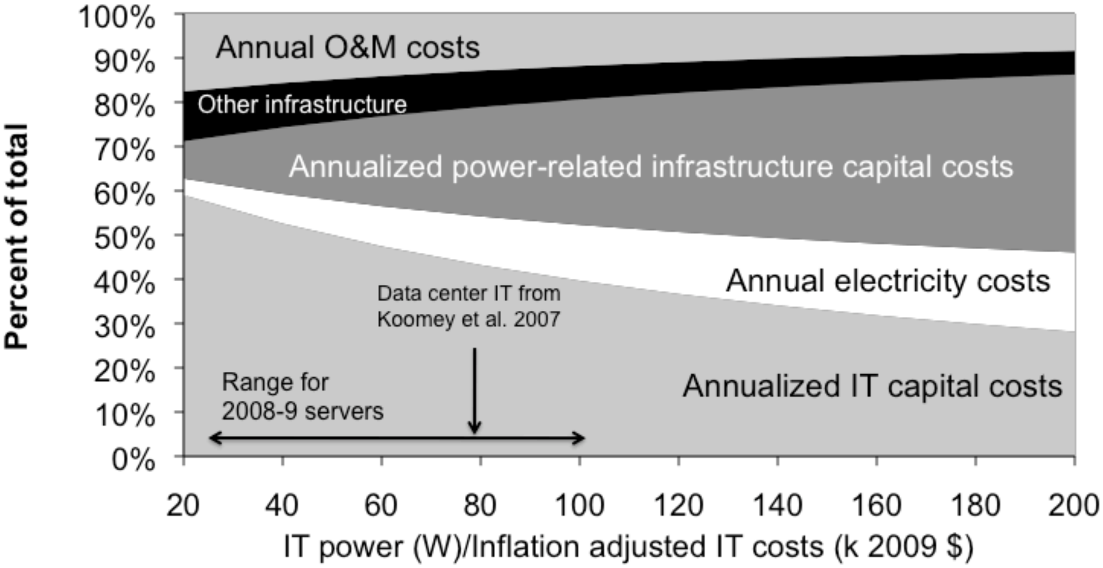


Figure ES-1: Power-related costs grow as power per server cost grows

The 2008-9 server data apply to the servers graphed in Figure 2. Capital and operating cost components derived using equations in Appendix A.

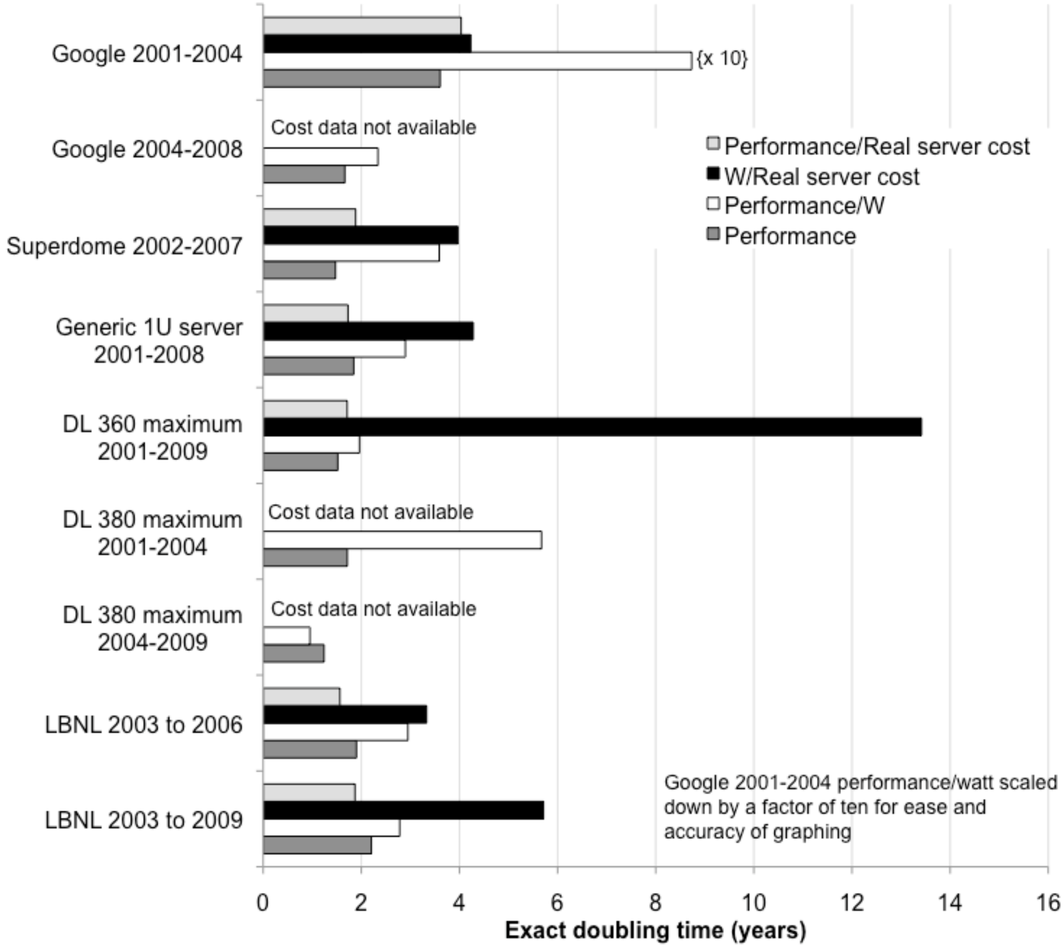


Figure ES-2: Summary of trends for servers, expressed as doubling time in years

Longer bars mean slower growth. Doubling time calculated using instantaneous exponential growth rates as described in the text.

ASSESSING TRENDS OVER TIME IN PERFORMANCE, COSTS, AND ENERGY USE FOR SERVERS

Jonathan G. Koomey, Christian Belady, Michael Patterson, Anthony Santos, Klaus-Dieter Lange

INTRODUCTION

As data centers have grown in both economic importance and cost, the need for understanding the underlying drivers of total costs in the data center has also increased. In particular, the relationships among processing power, energy use, and purchase costs of information technology (IT) equipment in these facilities strongly affect the fraction of total costs attributable to IT equipment (as distinct from facilities/infrastructure equipment like chillers and power distribution systems).

Anecdotal reports indicate that infrastructure equipment related to power and cooling may be responsible for about half of total annualized costs in typical data center facilities (Belady 2007, Brill 2007, Koomey et al. 2007) and that this fraction is growing over time as IT equipment acquisition costs decline and IT equipment energy use increases. This finding is surprising to people new to the data center arena, as they associate these facilities mainly with the IT equipment they contain.

Unfortunately, there has been little systematic, transparent, and peer-reviewed work documenting the aggregate trends in IT equipment that are driving changes in total data center costs. This lack is most keenly felt by those trying to plan for new facilities. Modeling data center costs at a high level requires abstracting from anecdotal data to generalize about trends, but the poor quality of available data and examples has prevented such generalizations from being useful to the bulk of the data center industry.

This article assesses trends in server equipment (the most important component of IT equipment in data centers) in a way that will be useful for people trying to understand data center costs at a high level. It develops and documents detailed examples from available data, estimating costs and correcting them for inflation, and explaining the reasons for differences in the results. It also describes future work to expand the number of examples and improve the usefulness of the data.

Conceptualizing the problem

One of the most important aggregate parameters affecting the cost of data centers is the amount of direct power use (watts) associated with one thousand dollars of expenditure on IT equipment hardware (in this case, servers[1]). Brill (2007) showed (using anecdotal information) that this parameter has been increasing rapidly in recent years, which has

---

[1] In principle, this parameter should be measured for all IT equipment in the data center, not just servers, but the data, sparse as they are, are most available for servers, so that's what we focus on here.

made cooling and power infrastructure costs rival the IT capital costs in some recently constructed data centers. If this trend continues, the power related infrastructure costs will significantly exceed the IT capital costs for new facilities in under a decade, a finding that has implications for how these facilities are built and how their costs are allocated within organizations (Belady 2007, Brill 2007, Koomey et al. 2007).[2]

Figure 1 shows how annualized IT capital costs compare to annual electricity and annualized infrastructure capital costs as a function of power use/server cost. The figure uses the equations in Appendix A and the infrastructure cost and electricity price assumptions from Koomey (2007) for a Tier 3 data center.[3] At one hundred watts per thousand dollars of server cost, IT capital costs represent about 40% of total costs, and at two hundred watts per thousand dollars they are responsible for less than one third of total costs, which means that for every dollar spent on IT equipment, a company would be committing to at least another two dollars for electricity use, power and cooling capital costs, and other costs. These results have implications for assessing and controlling costs in these facilities, as discussed in the future work section below.

Power per server cost (in watts per thousand dollars) can be decomposed into two component parts, as shown in equation 1:

$$\frac{\text{Power}}{\text{Server cost}} = \frac{\text{Power}}{\text{Performance}} \times \frac{\text{Performance}}{\text{Server cost}} \qquad (1)$$

or equivalently

$$\frac{\text{Power}}{\text{Server cost}} = \frac{\dfrac{\text{Performance}}{\text{Server cost}}}{\dfrac{\text{Performance}}{\text{Power}}} \qquad (1a)$$

where

---

[2] Most companies have separate budgets for IT and facilities expenditures, and if a dollar spent on IT can commit another part of the company to a dollar or more of additional expenditures in a separate budget, suboptimal behavior will generally be the result.

[3] Each data center is unique. They vary greatly depending on the reliability they deliver and the types of computing they support. This example was developed for high performance computing for financial applications. It is the best-documented published example of data center costs, which is why we rely on it for our discussion here. The conceptual points raised in our discussion are not affected by the specifics of this example, and we believe the concreteness this example lends to the discussion outweigh any potential pitfalls.

$\frac{\text{Performance}}{\text{Power}}$ = system performance divided by the measured power use for that server system to deliver that performance (i.e., performance per watt); and

$\frac{\text{Performance}}{\text{Server cost}}$ = that same performance metric divided by the server hardware capital cost as configured to achieve that performance.

This equation explains why measuring power use, performance, and server costs in a consistent fashion is so important—it allows us to understand the underlying drivers of power per server cost in an unambiguous way. It also shows that whenever performance per server cost is increasing faster than performance per watt, power use per thousand dollars of server costs will increase.

Consider **Figure 2**, which plots the two components in Equation 1a for fourteen servers selected from the available SPECpower_ssj2008 runs for recently manufactured servers, including the HP DL360 G5 machine analyzed below (see Appendix B for the underlying data and methods used to estimate purchase costs for these machines). For comparison, it shows data for the HP DL360 G1 machine assessed in Table 5, as well as two lines of constant watts per thousand dollars (one for 25 and one for 100).

The x-axis plots server performance at maximum load divided by power and the y-axis plots server performance at maximum load divided by purchase cost. The graph shows almost a factor of two variation in server performance per unit of power and a factor of four variation in server performance per server purchase cost. Combining these parameters yields a range of from 26 to 100 watts per thousand dollars of server equipment, as shown in Figure 1.

These graphs illustrate the complex and multivariate nature of the decision problem for data center design. Increasing performance per purchase cost is a sure-fire way to reduce the direct cost of delivered computing services, but if performance per watt is low, that choice will exact a penalty in infrastructure capital and electricity costs.

Consider the following stylized cost calculation for the total costs of a data center, based on the simple model developed in Koomey et al. (2007) and the equations in Appendix A. The annualized total cost (ATC) of a data center can be expressed as in Equation (2):

$$ATC = IT + INF_{nonkW} + INF_{kw} + EC_{IT} + EC_{Inf} + O\&M \tag{2}$$

Where

IT = annualized IT capital costs (which includes the acquisition costs of servers, network gear, disk arrays, and other IT equipment);

INF_{nonkW}= annualized non-kW related infrastructure capital costs (which includes building shell, office fittings, and land);

INF_{kw} = annualized kW related infrastructure capital costs (like chillers, water distribution, cooling towers, backup power systems, generators, and anything else whose sizing is dependent on the amount of power drawn by the IT equipment);

EC_{IT} = annual direct electricity costs for IT equipment;

EC_{Inf} = annual direct electricity costs for infrastructure equipment (in typical data centers, $EC_{IT} \approx EC_{Inf}$); and

O&M = annualized operations and maintenance costs (in our definitions this term includes both IT and facilities operations costs, but does not include software licenses and application development).

Both annual electricity cost and kW-related infrastructure costs are directly related to IT expenditures through the ratio of power use per server cost. In the example in Koomey et al. (2007), where the aggregate power use per purchase cost for all IT equipment in the data center is about 80 W/thousand 2009 dollars, IT capital costs account for 45% of total annualized costs, electricity use accounts for about 10%, and power-related infrastructure capital accounts for almost a quarter of the total (see **Figure 3**).

Of course, what we really care about is the cost per delivered computing cycle. Let's think about this problem in terms of the maximum number of computations possible for a given data center over the course of a year.[4] Dividing both sides of equation 2 by maximum annual computations we get:

$$\frac{\text{ATC}}{\text{Annual Computations}} = \frac{\text{IT} + \text{INF}_{\text{non-kW}} + \text{INF}_{\text{kW}} + \text{EC} + \text{O\&M}}{\text{Annual Computations}} \qquad (3)$$

Equation 3 represents in a schematic form the complete decision for data center design. The designer would like to minimize the total cost for delivering computations, but achieving this goal is not as simple as choosing the server with the maximum performance or lowest power use per dollar of equipment purchase cost. Focusing only on the ratio of IT costs per computation would result in a significantly more expensive facility than if the data center were analyzed as a whole system.

---

[4] The subtleties of measuring actual utilization and total computational output are complex ones that need not enter into our illustration here. Poorly utilized data centers can of course lower their total costs of computing substantially by increasing utilization levels. Such changes will have a large effect on computational efficiency because current server power use does not generally scale exactly with computational output, and there is a large fixed power draw when the server is idle (Barroso and Hölzle 2007).

Implications of these equations

These equations, combined with those in Appendix A, can be used to give quantitative insight about the tradeoffs among the different cost components of data centers. Let's assume Moore's law drives performance per server cost up by a factor of two over a two-year period (a doubling time of two years). The effect on the power-related components of data center costs depends on what happens to power use per server costs (and implicitly, to performance per watt).

- If performance per watt also doubles during this period, then watts per thousand dollars will remain constant (as per Equation 1) and the overall cost per computation will be exactly halved (Equation 3), because each of the numerator elements remains constant and the denominator doubles.
- If performance per watt doesn't change at all, watts per thousand dollars will double and the 50% reduction in cost per computation will become a 32% reduction in total costs (see Table A-1 for details). In this case, increased indirect power-related costs offset 36% (18%/50%) of the cost reductions resulting from increased compute performance. Another way to say this is that the total cost of building and operating a data center that delivers the increased performance would be 36% higher than in the base case.
- To make the increase in indirect costs exactly offset the benefits from increased server performance, performance per watt would have to drop to half of its initial value at the same time as performance per server cost is doubling. This change would result in power use per server cost of more than 300 W/thousand 2009 dollars, almost a factor of four increase over the base-case value (about 80 W per thousand 2009 dollars, from Koomey et al. (2007)).

These effects cut in both directions. If server manufacturers were able to triple performance per watt as performance per server cost was doubling, the total cost per computation would be 12% less than if performance per watt just kept pace with performance per server cost, because of the reduction in power-related costs. Whether investing to make this change would be economically desirable depends of course on the costs to improve server efficiency at this rate.

The focus of this study

To understand the underlying drivers for this complex situation, this article explores trends in power use, server costs, and performance. It focuses on the following questions:

- What kinds of data would be needed to accurately characterize trends in performance per watt, performance per server cost, and power use per server cost?
- Can changes in these parameters be measured in a credible, accurate, and representative way using publicly available data?

- If so, how have these parameters changed over the past ten years and what can we say about how they are likely to change in the next decade?

DATA AND METHODS

General issues

Our purpose here is to develop peer-reviewed consistent comparisons for performance, costs, and energy use over time. By peer-reviewed we mean that a broad section of knowledgeable industry observers (identified by name in the acknowledgements section to this report) have examined the assumptions, data, and analysis and found them credible. By consistent we mean that measurements of these parameters are conducted in a fashion that allows for meaningful comparisons over time.

To understand these trends for server equipment, we first need to define system boundaries. Servers can be analyzed at the CPU level, the system level, or the applications level.[5] The applications level is closest to the tasks that users are performing but data at that level are the hardest to measure and to generalize. Data are abundant at the CPU level but CPU measurements are sufficiently removed from actual computing tasks that they are of limited usefulness. System level data are in the middle in terms of both data availability and relevance to actual computing tasks. In practice, the system level data are the most likely to be both available and relevant.

It is important to ensure that any examples used be representative of IT equipment. There are at least two dimensions in which server hardware can be representative: configuration and operation. Server systems can be configured with variations in random access memory (RAM), disk drives, and network interface cards—examples chosen should be as representative of typical configurations as possible. Most business servers operate at only 5 to 15% of their maximum computing loads, but there's wide variation in compute utilization. The ideal examples would be broadly representative of the ways servers run actual applications.

To allow straightforward comparisons, we use the metric of doubling time, defined as the number of years it takes for a parameter (performance per watt, for example) to double. We first calculate the instantaneous growth rate g as in Equation 4[6]:

---

[5] The most sophisticated data center operators that have relatively homogeneous computing loads can analyze servers at the data center level, since they can shift loads between servers relatively easily. This system level analysis is not relevant for most users (who are more concerned with server level trends) so we don't discuss it further here. It is also important for improving equipment utilization, another topic we don't treat here.

[6] It is more common in most situations to use simple growth rates, calculated as g = $\left(\frac{Y_t}{Y_o}\right)^{\left(\frac{1}{t}\right)} - 1$ but this method gives erroneous answers for growth rates higher than about 10% per year. For the high

$$g = \frac{\mathrm{LN}\left(\frac{Y_t}{Y_o}\right)}{t} \qquad (4)$$

where

Y_t is some quantity at time t,

Y_o is that quantity at time 0

and t is the time over which growth occurs, measured in this case in years (from year 0 to year t).

Instantaneous growth rates assume continuous compounding, which is necessary when dealing with the rapid growth rates common in computer technology. An instantaneous growth rate of 69.3% implies a doubling every year.

We can then calculate the doubling time using Equation 5:

$$\text{Doubling time} = \frac{LN\,(2)}{g} \qquad (5)$$

Using the doubling time allows us to compare the trends in servers to another important parameter popularly reported in this fashion (Moore's law), which in its most precise form states that the number of transistors on a chip doubles roughly every two years.[7] The most widely believed incarnation of Moore's law is that performance per microprocessor doubles every 1.5 years, which happens to be true (as documented by Nordhaus (2007)) but it is unclear if this popular belief is based on real data or just a misunderstanding of what Moore actually said (Mollick 2006).

Performance

How to measure computing performance has been a source of controversy since the beginning of the computer age, and this article won't settle those issues. Each example

---

growth rates common to information technology equipment, instantaneous growth rates are more appropriate and accurate (Nordhaus 2007). The instantaneous growth formula is derived from the equation $Y_t = Y_o e^{gt}$. To convert a simple annual percentage growth rate (P) to a continuously compounded instantaneous rate, take the natural logarithm of (1+P). We are indebted to Philip Sternberg of IBM for helping to sort out the subtleties of these growth calculations.

[7] This "law" has changed in form over the years (See Mollick 2006 and http://arstechnica.com/hardware/news/2008/09/moore.ars). At first Moore (1965) referred to "components" not transistors, and correctly predicted that the number of components would double every year through at least 1975. In 1975, Moore correctly predicted that the number of transistors on a chip would double every two years in the future (Moore 1975).

we developed relies on different performance metrics, but in each case the performance metric remains consistent over time. It is the time trends that matter for this analysis, not the accuracy of one metric over another.[8]

One possible source of uncertainty in the analysis of time trends comes into play when improvements in hardware (as characterized by theoretical performance benchmarks like FLOPS or Composite Theoretical Performance (CTP)) require additional work on the software side to take full advantage of those performance improvements. This has become an issue recently as processors have moved to multiple cores on a single chip, making software redesign necessary to take full advantage of parallel processing. To the extent that theoretical benchmarks (or any other benchmarks) do not reflect real world applications (which may or may not have been optimized for improvements in hardware) then the trends we derive won't precisely reflect real world costs. Since our study is exploratory, we simply note this uncertainty as an area ripe for further investigation.

Energy use

Energy use of IT equipment has been a major focus of research for more than two decades.[9] The most common error in assessing energy use for computers is to rely on the nameplate power use printed on the computer's power supply, which is generally two to three times larger than typical power use for that device in operation.

We rely mainly on measured data for this analysis, some of which comes from SPECpower_ssj2008 <http://www.spec.org/power_ssj2008/>, available for use since late 2007. As with all benchmarks it has limitations, but for now, it's the best available option for associating power use with performance.

One source of uncertainty in the SPECpower_ssj2008 measurements relates to whether power management features are enabled in most production servers. Anecdotal information suggests that these features are often disabled in the field, which suggests that the power measurements in most SPECpower_ssj2008 runs (which assume these features are enabled) may not reflect actual field conditions.

Another source of uncertainty is that SPECpower_ssj2008 mainly stresses the processor and memory components of server systems. Disk drives, network interfaces, and other system components are also important in certain applications, and measuring power use while running other performance benchmarks may be more appropriate for those applications.

---

[8] Of course, one should always prefer benchmarks that closely approximate real-world workloads when they are available.

[9] For more details, see these references: (Baer et al. 2002, Blazek et al. 2004, Harris et al. 1988, Kawamoto et al. 2002, Koomey 2008, Koomey et al. 2002, Koomey et al. 2004, Koomey et al. 1996, Lovins and Heede 1990, Mitchell-Jackson et al. 2002, Mitchell-Jackson et al. 2003, Norford et al. 1990, Piette et al. 1991, Roth et al. 2002, Roth et al. 2006)

Costs

One of the key failings of industry assessments of cost trends in the past is that costs are almost never reported in a form that is consistent with the performance and energy use data. We treat that issue by compiling industry data on equipment prices for configurations of servers for which performance and energy use are reported.

Another issue with costs when they are reported is that they are almost never corrected for inflation. We use the annual implicit deflator for GDP from the Bureau of Economic Analysis (http://www.bea.gov/national/nipaweb/SelectTable.asp?Selected=Y) and the assumption of 2% per year inflation from 2008 to 2009 to adjust all dollar figures to constant 2009 dollars, thus eliminating inflation as a confounding variable in our time trends analysis. When cost data are available by month we use the monthly GDP deflator data from EIA (http://tonto.eia.doe.gov/cfapps/STEO_TableBuilder/index.cfm) to correct to July 2009 dollars (assuming 2% inflation month to month from 2008).

Costs depend on the characteristics of the purchaser, so absolute estimates of the power use per server cost (or other cost related ratios) are dependent on the particular context in which the servers were purchased. In general (but not always), large purchasers get more favorable pricing. In this study, we rely in part on costs produced by online stores for HP, Dell, and IBM. We do not include taxes, shipping costs, software, or service contracts. Where there is a choice we use costs for small + medium businesses (as distinct from costs for individuals or large corporations).

RESULTS

None of the following examples are perfect, but they represent a good first step towards accurately characterizing trends in server power, performance, and costs.

Generic 1U server example

Belady (2007) gives an example of server costs and electricity use over time for generic 1U “pizza box” servers (unfortunately, consistently measured performance data are not reported for this example). We extract the relevant data on costs and electricity use per server from Figure 3 in that article, as shown in **Table 1**. Nominal server cost is shown on the Figure, which we convert to real 2009 dollars as described above. The power use per server must be inferred from the Figure, using the assumptions stated in the article’s text (as summarized in the footnotes to Table 1). We add performance to this example by assuming that performance per server will track the changes documented in the DL360 example (Table 4, below), which allows us to estimate performance per watt and performance per server cost doubling times as well.

In this example, watts/server cost doubles about ever four years, power use per server doubles every five years, performance per watt doubles about every 2 years and performance/server cost doubles in 1.4 years.

HP Superdome example

Figure 2 in Belady (2007) gives performance and performance per watt trends for the HP Superdome high end server, although the graph doesn't indicate that these trends apply to that specific server. Fortunately, the Transaction Processing Performance Council (TPC) published capital cost and performance results for Superdome servers in 2002 and 2007, and we can use the performance/watt trends data from Figure 2 in Belady (2007) to add power use consistent with the measured performance and costs from TPC.[10]

Table 2 describes how we merged those two data sources. The TPC reports give extensive breakdowns of cost components, and because we are interested in trends in server hardware costs, we subtracted out the costs of software and support from the TPC cost totals. Costs were inflation adjusted as described above. We combined performance per watt trends from the Belady article with the performance data to calculate watts/server and watts/server cost, as described in the footnotes to Table 2.

The results for this example show that performance doubles ever 1.5 years, watts per server cost doubles every four years, and performance per server cost doubles every two years

Google server example

In 2005, Barroso (2005) published a graph containing cost, performance per watt, and cost per server data for three generations of Google servers. The original Figure showed data points for each generation of servers (labeled "A", "B", and "C") but didn't identify the years.

We've plotted **Figure 4** from that original graph, eliminating the intermediate generation of servers and adding the years 2001 and 2004 as the endpoints of the trends, based on an email exchange we had with Dr. Barroso in January 2009. The data from the graph are shown in **Table 3**, where we correct the price data for inflation, estimate trends data for server price and watts per server, and calculate the annual percentage change and doubling time for each parameter.

The results indicate that performance grows at a 19% per year instantaneous rate over this three year period, with inflation adjusted server price (acquisition cost) only increasing by about 2% annually. Watts per server increases at an 18% per year instantaneous rate, indicating that power use actually increased about as fast as performance did during this period (performance per watt was flat). Watts per server price increased at an instantaneous rate of 16% per year, which is a doubling time of about four years. Improvements in performance per watt are remarkably slow, with a doubling time of almost 90 years.

---

[10] TPC is currently working on an energy metric but it is not yet available as of August 2009.

This example illustrates the pitfalls of trend-based analysis over a relatively short time period and the importance of context to understanding numerical results. Google had already wrung out significant inefficiencies in their custom-designed servers by 2001, so the trends begin from a base year server that was significantly more efficient than standard industry designs. Energy use and performance are measured assuming that the application is web search, so those results may not be applicable to other kinds of server applications. Google also optimizes its servers for search so the results may not apply to "off the shelf" servers of more conventional types.

The years 2001 to 2004 preceeded a period of great innovation in chip design, where the manufacturers shifted from ramping up clock speed to increasing the number of cores on a chip. The latter approach allows substantial increases in performance for little or no increase in chip power (assuming that software was redesigned to take full advantage of parallel processing). The effects of this change are evident in **Figure 5**, which plots data supplied by Dr. Barroso in March 2009.

The figure plots an index of performance and performance per watt over time for Google servers, using a highly parallelized benchmark similar to the one used in the example above (unfortunately, consistent cost data are not available for this example). It shows modest changes in performance and performance per watt from 2001 to 2004 (comparable to the changes shown in Figure 4) with a substantial jump in the rate of change in both parameters after 2004. Whether these rapid rates of change can continue for years to come is an ongoing subject of debate in the computer industry (Bohr 2007), but doing so is dependent on significant new innovation comparable in scale to the shift from single core to multi-core computing. This innovation will also require substantial changes in software design (Asanovíc et al. 2006), which is a relatively new development for the IT industry.

HP ProLiant DL360 example

We derived the previous three examples from already published results. Our fourth example relies on new measurements of power use and performance for an older server combined with published measurements for a new version of that same server model (based on SPECpower_ssj2008, http://www.spec.org/power_ssj2008/). SPECpower_ssj2008 is one of the first attempts to associate power use with a specific performance benchmark (see also Koomey et al. (2006)). Like all benchmarks, it applies to only a subset of the actual computing loads in the real world, but associating such benchmarks with actual power use at different computing loads is the most accurate way to assess power use.

Server models tend to occupy the same market niche over time, and often don't change much in nominal price over the years. This means that comparing different generations of the same server model can be an effective way to generate cost, performance, and energy trends in a consistent fashion.

The first author of this article discovered a Compaq/HP ProLiant DL360 Generation 1 server (purchased in the year 2001) that a colleague at the Midwest ISO was about to decommission. That server, which is comparable to the generic 1U server analyzed in the first example, was shipped to Intel Corporation in 2008 and Anthony Santos of Intel installed the correct software to run the SPECpower_ssj2008 benchmark on this machine. Purchase costs for this machine were derived from a manufacturer's suggested retail price (MSRP) listed on a web site that still sells old DL360 G1 servers (adjusted to reflect the RAM and hard drive configuration on the server on which Intel ran SPECpower_ssj2008).

The SPECpower_ssj2008 web site contains performance and power results for more than one hundred new servers, including a new HP ProLiant DL360 Generation 5 server using an Intel microprocessor (Xeon 5450) with advanced power saving technologies. We paired these data with the retail cost of purchasing that server directly from the HP web site, configured as described in **Table 4** and in more detail in Appendix D.

In both cases, the servers were configured with one hard drive, one power supply, and half of the maximum allowed random access memory (RAM).

Table 5 describes the results, splitting the trends into a "typical" case (10% computing utilization, similar to the processing load level found in typical business servers) and a "maximum" case (100% computing utilization, more commonly found in batch mode weather, drug, or seismic modeling computing). Measured system performance seems to track the number of transistors on each processor, doubling every 1.5 years. Power per server doubles about every seven years, while performance per watt doubles every two years. Power/server cost doubles only every 14-18 years, while performance/server cost doubles in a little under two years.

There are a couple of important caveats to these results: One is that the design of Generation 5 servers are the result of several years of focus by microprocessor and server manufacturers on improving the energy efficiency of servers, so it is likely that growth in power use slowed with this generation of server. One way to determine if this is the case would be to conduct similar measurements for the intermediate generations of the DL360, to create a real time series instead of just the two endpoints to that series.

In addition, the manufacturers have only posted data for about 100 servers on the SPECpower_ssj2008 web site as of June 2009. Nobody knows how representative those servers are. It is possible that the manufacturers have just tested their most power efficient machines—we just don't know.

Another caveat relates to the use of data like these in analyzing data center costs at a high level. Infrastructure (power delivery and cooling) capital costs are directly related to the maximum measured power use of the servers, and these capital costs are not reduced if the servers operate at a low fraction of their maximum computing load (as they often do). So the most relevant parameters related to infrastructure capital costs are based on the "maximum load" case, even though typical computing loads would be much lower based on current practice. For calculating direct energy costs the estimates called “typical” in this table are more applicable.

HP ProLiant DL380 example

Some power and performance data on several generations of the HP ProLiant DL380 server recently became available on the SPECpower_ssj2008 web site <http://www.spec.org/power_ssj2008/>. The new runs cover a Generation 4 DL380 circa 2004, two Generation 5 DL380s (2006 and 2008), and Generation 6 machine. In combination with the SPECpower_ssj2008 runs for a Generation 1 DL380 completed by Anthony Santos of Intel (see Appendices C and D), we have at least an initial indication that power-saving technologies have become a higher priority for the industry since 2004. Unfortunately, consistent cost data are not yet available for these machines.

Table 6 shows that the doubling time for performance per watt at 100% load was about 5.7 years from 2001 to 2004, but only about 1 year for the period 2004 to 2009. In that same example, performance doubled every 1.7 years from 2001 to 2004, and every 1.2 years from 2004 to 2009. These results are qualitatively similar to the increases in performance and performance per watt shown above in Figure 5, with the rate of change for these parameters substantially increasing after 2004.

Performance per server is increasing more rapidly than performance per watt in this example. If (as is often the case) real costs per server are constant or decreasing, that means that power use per thousand dollars of server cost would be increasing in this example (see Equation 1). Further analysis of actual costs is needed to verify this hypothesis.

High performance computing (HPC) example

Lawrence Berkeley National Laboratory (LBNL) is a major consumer of HPC equipment. In the last ten years most of these installations have been clusters of server computers linked together in massive parallel processing networks. Because these clusters are used for research, they often remain useful for years, which means older server equipment is available to be measured.

At the request of this report’s first author, LBNL’s computing division measured the power use for three server clusters, with servers purchased in 2003, 2006 and early 2009 (see **Table 7**). LBNL was not able to run SPECpower_ssj2008 on these machines. Instead, they measured power use at idle and at maximum load, using a software program that measured component utilization and pushed each server to its limits.

Measured performance (typically expressed using the LINPACK benchmark for cluster computing) was also not available for these clusters, but the LBNL computing division uses “Theoretical performance” as a crude measure of performance. This parameter (measured in FLOPS, or floating point operations per second) is the product of the processor speed, the number of cores, and the number of instructions completed per clock cycle.

LBNL is bound by confidentiality agreements with server vendors so it can’t reveal the exact prices of these machines, but LBNL staff was able to share the ranges of prices in each case. The earliest machine cost between \$3,000 and \$3,500, and the two later machines cost between \$2,500 and \$3,000 (all in nominal dollars). We assumed the midpoint of each range.

This example, like all the others, has both strengths and weaknesses. The same institution bought the servers in all three cases, and the servers were all for similar applications that were cutting edge at the time of purchase. The energy use is measured but performance is a synthetic benchmark and the costs are expressed in the form of a range. In addition, the server from 2003 is a no-name white box, while the 2006 and 2009 machines are from a major manufacturer (Dell). This may introduce some inconsistency in the time series, although LBNL is a sophisticated purchaser that buys a large number of servers every year using a bidding process, so it is likely that the error introduced by this shift in manufacturer is small.

Table 8 summarizes trends for this example. As in all other cases, performance per thousand dollars of server cost is increasing more rapidly than performance per watt, which drives up watts per thousand dollars of server cost. Maximum power use per server actually decreases from 2006 to 2009, which presumably is because of the strong focus of chip and server manufacturers on reducing power use during that period.

Performance per server doubles about every two years, as does performance per thousand dollars of server cost, while performance per watt doubles about every three years. The change in power use per server cost doubles every 5 years from 2003 to 2006 and every 11 years over the 2003 to 2009 period, reflecting the reduction in absolute power use per server from 2006 to 2009.

DISCUSSION

Analyzing trends accurately over time requires consistent estimates of performance, server price, and power use per server, measured over a time period sufficient to capture major step changes in chip and server system design. The data also needs to be broadly representative of major classes of server applications for lessons derived from them to be generalizable to the industry as a whole. While the examples explored here have limitations, they represent a good first step towards a deeper understanding of trends in server technology. **Figure 6** summarizes these quantitative trends in terms of doubling times.

Performance

Performance per server generally doubles every 1.5 years or so. The only outlier in the performance data is the Google example from 2001 to 2004, which doubles only every four years.

Performance/watt

Performance/watt doubles every two to four years, again with the exception of Google servers from 2001 to 2004, which show much slower growth in this parameter than do the other examples.

Performance/server cost

The Google example from 2001 to 2004 is also the outlier for this parameter. The other examples show doubling times of about two years for performance/server cost. Performance per server cost is growing more rapidly than performance per watt in all cases, which is why power use/server cost is going up in all cases (as per equation 1).

Watts/server cost

Doubling times in watts/server cost are much longer for the DL360 case than for the other examples.

Watts per server

All of the examples above include estimates of doubling times for power use/server. These can be compared to the doubling times contained in Koomey (2008) for the server market as a whole (see **Figure 7**). Power use/server appears to grow faster in the examples presented above than in the overall server market. For example, Superdome watts/server doubles in about 2.5 years, which is three times faster than the overall market for high-end servers from Koomey (2008), and the same conclusion holds for the other examples compared to the market trend for volume servers.

The market trends include both changes in the power used by consecutive generations of individual server models (like the trends we analyze in the examples above) and shifts in the market share of different server models, which may explain the lower growth rates and longer doubling times in power use/server for the aggregate market data. Further research is needed to validate this inference.

FUTURE WORK

This analysis is more detailed than any conducted previously, but it’s still largely anecdotal in nature. Each example has relevance to some part of the server market, but a more comprehensive approach would be required to accurately understand the aggregate trends. The Superdome server, for example, is only one of many high-end server systems, but high-end machines vary so greatly in their design, construction, and application that additional data is sorely needed to better characterize this market. The same lesson holds for the other examples, which fall under the category of "volume" servers. Future work should therefore include generating more and better server examples, increasing focus on collecting performance data at the applications level, encouraging wider use of energy measurements associated with performance benchmarks, assessing future trends, analyzing underlying technical trends in servers, encouraging technology demonstrations using whole system redesign, broadening data collection to cover disk drives and network equipment, and assessing the effects of these trends on total data center costs using simple models.

More server examples

Future work should include developing more examples like the DL360, where performance benchmarks of some kind are run in conjunction with power measurements, and costs are estimated based on actual system configuration details. As more SPECpower_ssj2008 runs become available such comparisons will become easier, although we suspect manufacturers have to date focused mainly on their most efficient systems (the existing data are not necessarily representative of broader server market trends).

TPC (http://www.tpc.org/) may also be a good starting point for additional comparisons of the type explored in this report. The TPC performance benchmarks are for larger server systems, and the results include both capital costs and performance. All that is needed are credible power measurements for some of the systems for which TPC benchmarks have already been run (TPC is currently developing an energy metric to be associated with future benchmark runs, but it is not yet available as of August 2009).

Better server examples

There are many ways to generate better server examples for studies like this. One way would be to create comparisons that are more consistent. So for the DL360 example above, we could in principle have compared the Generation 1 machine to a Generation 5 vintage machine that did not have significant power saving technology built in (or enabled). That would demonstrate trends for the case where manufacturers didn't do the work on system efficiency, for comparison with the case using the more advanced processor.

Another way would be to create a detailed time series, instead of focusing just on the endpoints for some period. For the DL360 example above, that would mean also compiling costs and performing SPECpower_ssj2008 runs for the DL360 G2, G3, and G4 machines, to see how the doubling times for key parameters vary from 2000 to 2008. We suspect that the curves for various parameters are not linear over this period, because the industry started to pay serious attention to server power efficiency around 2005 and because of the shift to adding multiple cores per chip (which allowed the processor manufacturers to increase CPU performance while staying within the same power envelope).

Focus on the applications level when possible

The ultimate goal for assessing costs of computing technology is to calculate the cost of delivering computing services. In practice, such analyses are usually only possible when the types of IT services within a given firm or data center are relatively homogeneous. For example, the big Internet search firms can estimate the cost for each search because search comprises the vast majority of computing tasks within those firms. The more we can estimate costs and performance at the application level, the more accurately we'll be assessing these trends for real-world installations. Server system level data are interesting and useful (and often more accessible than application level data) but software can have large effects on these trends, so application level data are strongly preferred in analyzing performance. Focusing on the application level will also highlight the importance of improving server utilization levels to reducing overall cost per computation.

This conclusion is particularly important as software design becomes a more important contributor to higher computing performance. In the past, programmers didn't have to do much related to performance, and could rely on hardware performance doubling every 1.5 years, which kept their applications running ever more rapidly. With the shift to multicore chips, the onus on software designers has become greater, and not all of them have taken up the banner of more efficient coding with equal vigor.

Develop more and better benchmarks for performance and power use

Benchmarks never exactly measure performance in real-world applications. Right now, SPECpower_ssj2008 is the only widely used protocol for associating power measurements with energy use, but it is of limited general applicability. In addition, the number of servers for which SPECpower_ssj2008 has been run is modest (somewhat more than one hundred). Manufacturers need to complete SPECpower_ssj2008 runs as a matter for course for all their new servers, and more of the SPEC performance benchmarks need to be explicitly linked to power use of servers while running those benchmarks (in April 2009 SPECweb2009 incorporated power measurements but few data are yet available from that new benchmark). SPECpower_ssj2008 and other energy/performance benchmarks should also be run on a representative sample of older servers (like we did with the DL360 G1 machine) so that the trend data can be more accurately characterized. Ideally, power measurements would be attached to real measurements of business productivity from IT investments but it will be a long time before such metrics become more common.

Assess how technology changes will affect future trends

The physicist Nils Bohr once said that "prediction is very difficult, especially if it's about the future", and nowhere is that observation more true than in predicting trends in information technology. The examples presented here provide some clues that growth in power use and power per thousand dollars of server cost have started to slow in recent years. If confirmed by additional research, this development would be an important one having implications how data centers will evolve in coming years.

Analyze the underlying technical changes affecting efficiency of servers over time

The trends analysis assesses the results of technological changes in server design and construction at the server level. Another way to gain insight into the factors driving these trends would be to measure power use and efficiency at the *component* level for servers of different vintages. Such measurements would reveal, for example, that significant efficiency improvements have already been made in microprocessor CPUs, but less progress has been evident in the other components of the server (Barroso and Hölzle 2007). This information would yield insight into which technical changes have been most productive historically, and which parts of the server system might still hold the potential for significant energy savings.

Assess the potential for whole system redesign to improve power efficiency

The data summarized in Figure 2 show a wide range of possibilities for performance per watt, and technical analysis confirms that many opportunities to improve efficiency are neglected (Eubank et al. 2004, Greenberg et al. 2006, Tschudi et al. 2006, US EPA 2007). Fixing the institutional issues that impede efficient design and procurement of servers is one important step towards improving overall efficiency of servers, but there is still an important role for technology demonstrations in showing what can be done.

One of the most important concepts in creating effective technology demonstrations is what Amory Lovins at Rocky Mountain Institute calls “clean-slate whole-system redesign”. Instead of making incremental changes in existing technology, the way to create truly superior technology is to start from the tasks people want to perform and then design devices to perform those tasks that are simply better in many ways. People generally won’t buy efficiency for its own sake, but will do so eagerly when it is combined with other desirable attributes.

This design process should ignore illusory constraints inherited from the historical development path of many technologies, which is why the term “clean slate” is so important. It should also analyze the system as a whole and not focus solely on component characteristics (except insofar as that process contributes to superior whole system performance). So for example, a focus on costs of IT equipment per unit of processing power is a component level approach. One that analyzes total costs (including power use and site infrastructure capital) is a whole systems approach.

Of course, improving components can be important to achieve short-term efficiency gains. Simply replacing inefficient power supplies with efficient ones can improve server performance per watt by 10-20% in many cases, and more efficient CPUs have already had a substantial impact on total power used by servers. But this kind of incremental approach is no substitute for looking at the entire system and optimizing it as a whole.

Collect data on storage and network equipment

Cost trends in the data center are not just affected by servers. Disk storage and network gear are also important contributors to IT power loads, so comparable data needs to be collected for this equipment to create a more complete picture. Such data collection will require collaboration with the dominant players in those markets.

Apply the trend data to modeling and controlling total data center costs

After representative trend data become available for power use per thousand dollars of server cost the next step is to assess how total costs for data centers would be affected by those trends. If companies have separate budgets, responsibility, and decision making authority for IT and facilities, perverse and suboptimal behavior is sure to be the result, but without aggregate trends like the ones we are characterizing here (and without a simple model of TCO for a company's data centers) it would be difficult for a company to assess and rationalize the incentives within the firm (Stanley et al. 2007). The schematic calculations shown in Figure 1 and Appendix A show that the IT capital costs are comparable in size to the power-related infrastructure and electricity costs in typical facilities, which demonstrates why this problem is an important one.

The schematic calculations shown in this report need to be adapted to each company's specific situation. Reliability requirements affect the power-related infrastructure costs, and these requirements vary by facility (and often within each facility). Electricity prices, land costs, and labor rates vary greatly by location. And new developments in dynamic optimization of data centers can make the cost calculation even more complicated.

CONCLUSIONS

As our economy becomes more dependent on computing networks we'll need to develop an understanding of the deep underlying trends driving their costs and capabilities. This article combines economic data (server costs) with technological information to create consistent comparisons and give insights into the key trends affecting total costs in data center facilities.

Companies that own data centers need to understand the trends affecting true total costs in their facilities. Minimizing costs of computing services requires more than maximizing computing performance per dollar of IT equipment purchased. The direct power used per dollar of IT equipment cost drives the costs of cooling and electricity, which in recent years have come to approach (in annualized terms) the cost of purchasing the IT equipment in many data center facilities.

The data analyzed in this report point to continuing growth in power used per thousand dollars of server cost, which will only increase the importance of site infrastructure and electricity costs compared to the cost of IT equipment. This trend places a burden on most companies running internal corporate data centers, which have not yet adjusted their design, construction, and operations procedures to reflect this new reality. Incentives within many firms still do not promote minimization of the total costs of delivering computing services, oftentimes because total cost is not even analyzed in these companies.[11]

There are technical solutions that can help reduce the cost of computing services, but the problem cannot be solved without changing institutional arrangements and incentives within companies. Split incentives arise when facilities and IT departments have different budgets, or when people using the data center are charged solely per square foot, ignoring power use. Without a simple model of total costs and data assessing underlying trends, it's impossible for companies to understand the full benefits of fixing these institutional problems or to moving some of their computing demands to cloud computing providers (who have some inherent advantages in addressing these issues).

In the data compiled here, performance trends for server systems seem to track the popular interpretation of Moore's law well (it doubles in all but one case every 1.5 to 2 years). In all cases, performance per server cost increases more rapidly than does performance per watt, which drives power use per server cost up over time. While there's some evidence that the trend towards increased power use per server cost has moderated in the past few years (because of aggressive efforts by chip and server manufacturers to improve server efficiency), more research will be needed to confirm this conclusion.

---

[11] It is important to distinguish here between the large companies that supply IT services from the companies for which IT is not their core business. Most IT services companies have started down the path of fixing the misplaced incentives and structural problems that impede the minimization of total costs (with differing levels of commitment), but the latter group largely has not. In either case, understanding the trends embodied in the data presented here is critical for improving the design of these facilities.

ACKNOWLEDGMENTS

This report was produced with grants from Microsoft Corporation and Intel Corporation and independent review comments from experts throughout the industry. All errors and omissions are the responsibility of the authors alone.

The authors would like to thank Mike Manos and Rob Bernard of Microsoft Corporation and Lorie Wigle of Intel Corporation for their financial support of this project. We would also like to thank Mark Aggar (Microsoft) and Scott Shull (Intel) for their technical guidance, and Agnies Watson and Lori Blonn for their patient project management assistance.

We would like to thank James Berry of the Midwest ISO for making the HP DL360 G1 server available to Intel for testing, and to Mangesh Tamhankar of Intel Corporation for making Anthony Santos available to work on SPECpower_ssj2008 runs for this analysis and to share his insights on server power trends.

In addition, we are grateful to Gary Jung of the Lawrence Berkeley National Laboratory for making some of his servers available for testing and for assigning an excellent student, Jared Baldridge, to conduct the testing.

Finally, we would like to thank the technical reviewers for their insights and comments. The reviewers included (in alphabetical order by company):

AMD: Larry Vertal
AT&T: Michele Blazek
Cisco: Hugh Barrass, Leonid Rabinovich
Dell: Stuart Berke, David Moss, John Pflueger
Ecos Consulting: Chris Calwell
EPA: Andrew Fanara, Katharine Kaplan
Google: Luiz Barroso, Jimmy Clidaris, Bill Weihl
HP: Paul Perez
IBM: Steve Kinney, Joe Prisco, Roger Schmidt
ICF: Arthur Howard, Rebecca Duff, Cody Taylor
IDC: Lloyd Cohen, Vernon Turner
Intel: Henry Wong, Scott Shull
LBNL: Rich Brown, Bruce Nordman, Dale Sartor, Bill Tschudi, Gary Jung
Lumina Systems: Max Henrion
McKinsey: William Forrest
MegaWatt Consulting, Inc.: KC Mares
Microsoft: Mark Aggar, Stephen Berard, Richard Russell
Midwest ISO: James Berry
MIT: Walt Henry
Rocky Mountain Institute: Amory Lovins
Stanford University: Richard Mount, Phil Reese
Sun Microsystems, Greg Papadopoulos, Mark Monroe, Subodh Bapat
Team HPC: Bret Stouder
Uptime Institute: Ken Brill, John Stanley
Virginia Tech: Kirk W. Cameron, Wu-chun Feng
Yahoo: Christina Page

REFERENCES

Asanovíc, Krste, Rastislav Bodik, Bryan Catanzaro, Joseph Gebis, Parry Husbands, Kurt Keutzer, David Patterson, William Plishker, John Shalf, Samuel Williams, and Katherine Yelick. 2006. *The Landscape of Parallel Computing Research: A View from Berkeley*. December 18. <http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-183.pdf>

Barroso, Luiz André. 2005. "The Price of Performance: An Economic Case for Chip Multiprocessing." *ACM Queue, special issue on Multiprocessors*. vol. 3, no. 7. September.
<http://acmqueue.com/modules.php?name=Content&pa=showpage&pid=330>

Barroso, Luzi André, and Urs Hölzle. 2007. "The Case for Energy-Proportional Computing." *IEEE Computer*. vol. 40, no. 12. December. pp. 33-37. <http://www.barroso.org/>

Belady, Christian L. 2007. "In the data center, power and cooling costs more than the IT equipment it supports." In *ElectronicsCooling*. February. pp. 24-27.

Bohr, Mark. 2007. "A 30 Year Retrospective on Dennard's MOSFET Scaling Paper." *IEEE SSCS Newsletter*. vol. 12, no. 1. Winter. pp. 11-13.

Brill, Kenneth G. 2007. *Data Center Energy Efficiency and Productivity*. Santa Fe, NM: The Uptime Institute. <http://www.uptimeinstitute.org/>

Koomey, Jonathan. 2008. "Worldwide electricity used in data centers." *Environmental Research Letters*. vol. 3, no. 034008. September 23. <http://stacks.iop.org/1748-9326/3/034008 >

Koomey, Jonathan, Christian Belady, Henry Wong, Rob Snevely, Bruce Nordman, Ed Hunter, Klaus-Dieter Lange, Roger Tipley, Greg Darnell, Matthew Accapadi, Peter Rumsey, Brent Kelley, Bill Tschudi, David Moss, Richard Greco, and Kenneth Brill. 2006. *Server Energy Measurement Protocol*. Oakland, CA: Analytics Press. November 3. <http://www.energystar.gov/datacenters>

Koomey, Jonathan, Kenneth G. Brill, W. Pitt Turner, John R. Stanley, and Bruce Taylor. 2007. *A simple model for determining true total cost of ownership for data centers.* Santa Fe, NM: The Uptime Institute. September. <http://www.uptimeinstitute.org>

Mollick, Ethan. 2006. "Establishing Moore's Law." *IEEE Annals of the History of Computing (Published by the IEEE Computer Society)*. July-September. pp. 62-75.

Moore, Gordon E. 1965. "Cramming more components onto integrated circuits." In *Electronics*. April 19.

Moore, Gordon E. 1975. "Progress in Digital Integrated Electronics." *IEEE, IEDM Tech Digest*. pp. 11-13. <http://www.ieee.org/>

Nordhaus, William D. 2007. "Two Centuries of Productivity Growth in Computing." *The Journal of Economic History*. vol. 67, no. 1. March. pp. 128-159. <http://nordhaus.econ.yale.edu/recent_stuff.html>

Stanley, John R., Kenneth G. Brill, and Jonathan Koomey. 2007. *Four Metrics Define Data Center "Greenness" Enabling Users to Quantify "Energy Efficiency for Profit" Initiatives*. Santa Fe, NM: The Uptime Institute. <http://www.uptimeinstitute.org>

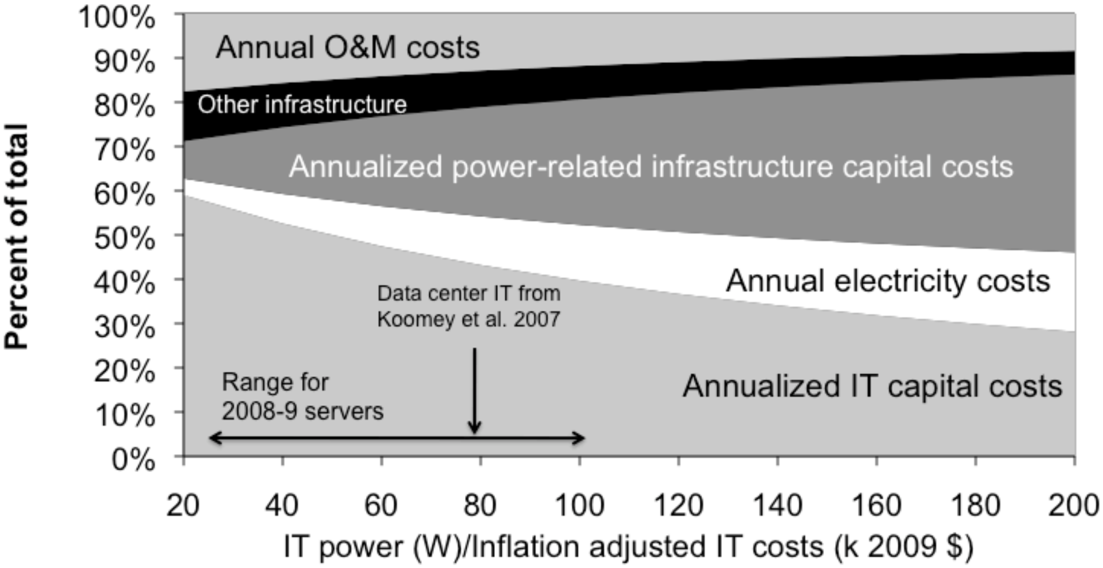


Figure 1: As power per server costs grow, power-related costs grow in importance

This graph shows annualized costs for a Tier 3 data center. The 2008-9 server data are from Fig. 2. Capital and operating costs derived using equations in Appendix A

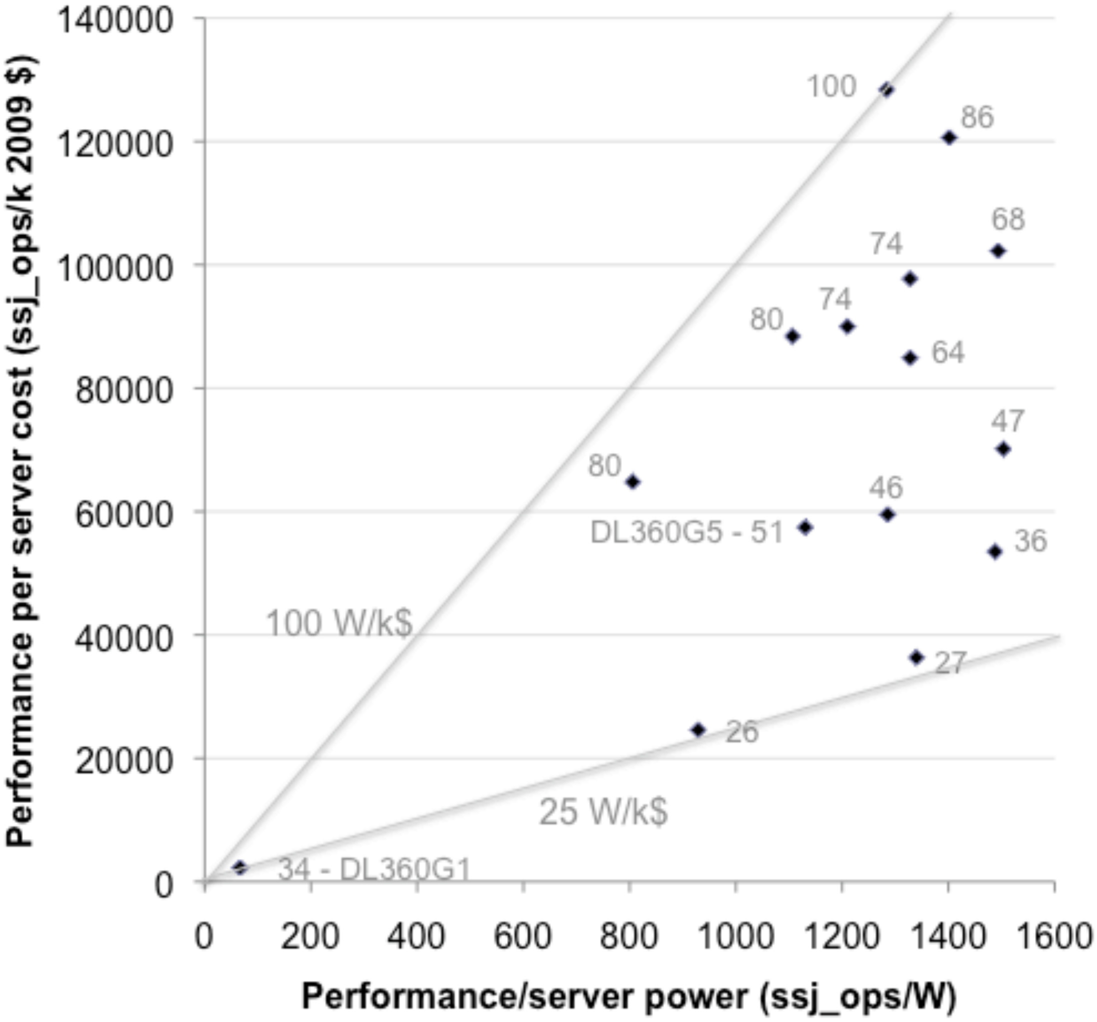


Figure 2: There are wide ranges of performance per watt and performance per server costs in currently available servers.

Performance and power based on 100% load cases from SPECpower_ssj2008 runs, as documented in Appendix B. Numbers next to each data point represent watts per thousand 2009 dollars of IT equipment expenditure for each server. DL360G1 data (circa 2001) added from Table 5 for comparison.

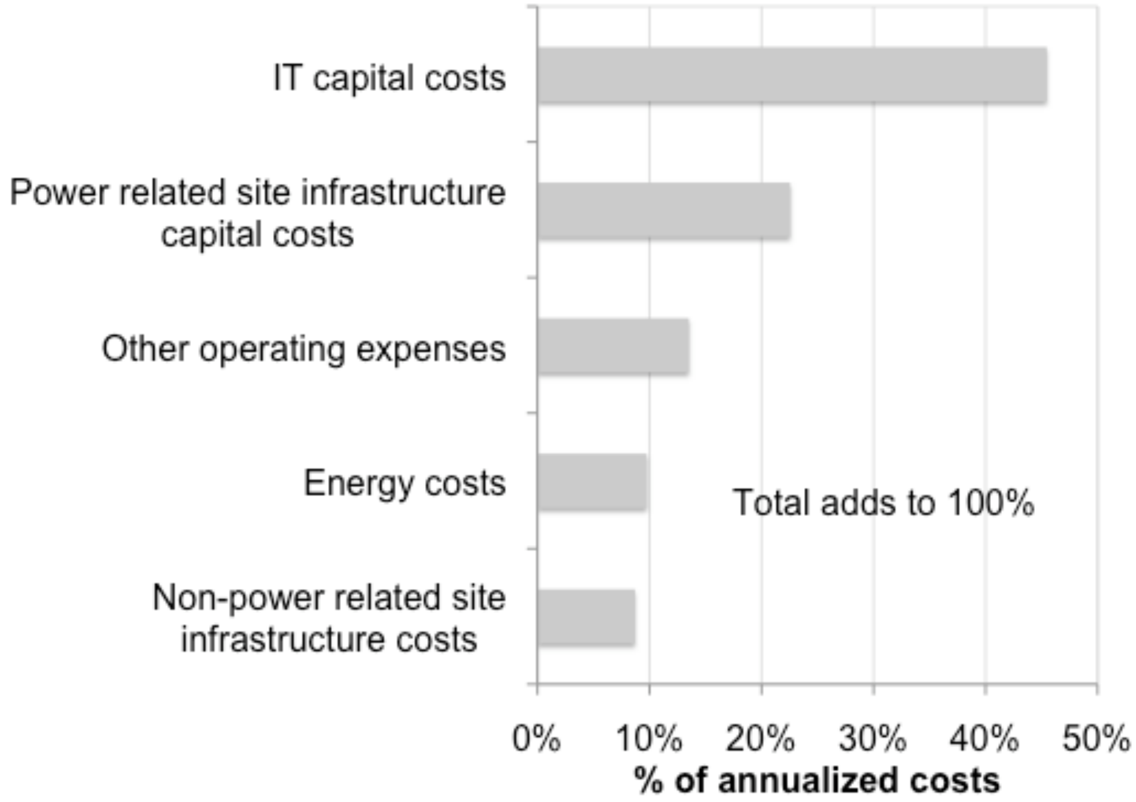


Figure 3: Infrastructure capital costs and electricity costs are substantial for Tier 3 data centers (based on Koomey (2007)).

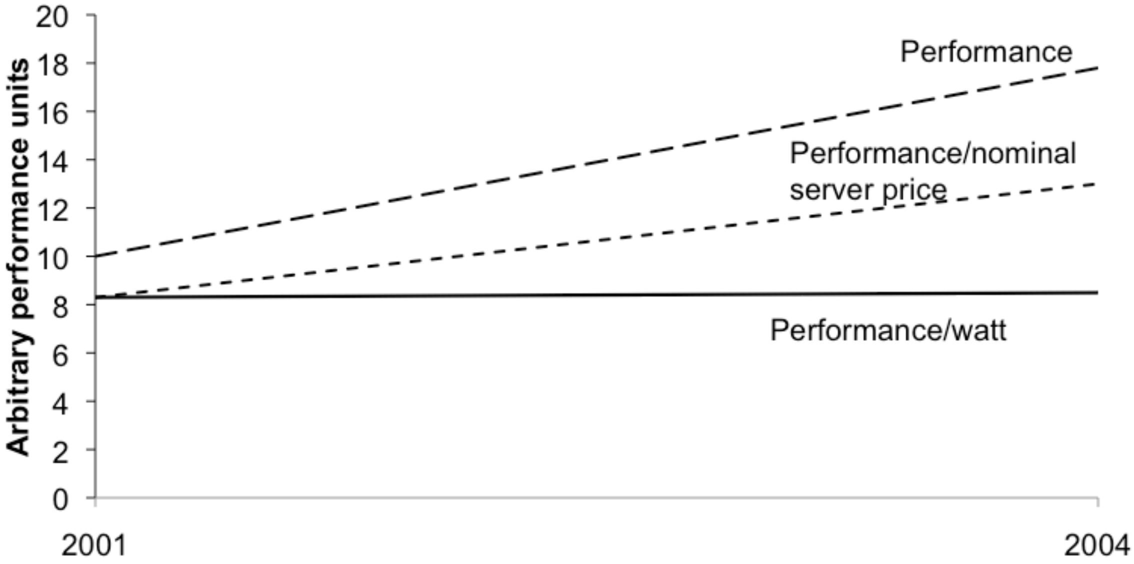


Figure 4: Trends for Google servers, 2001 to 2004 (from Barroso 2005)

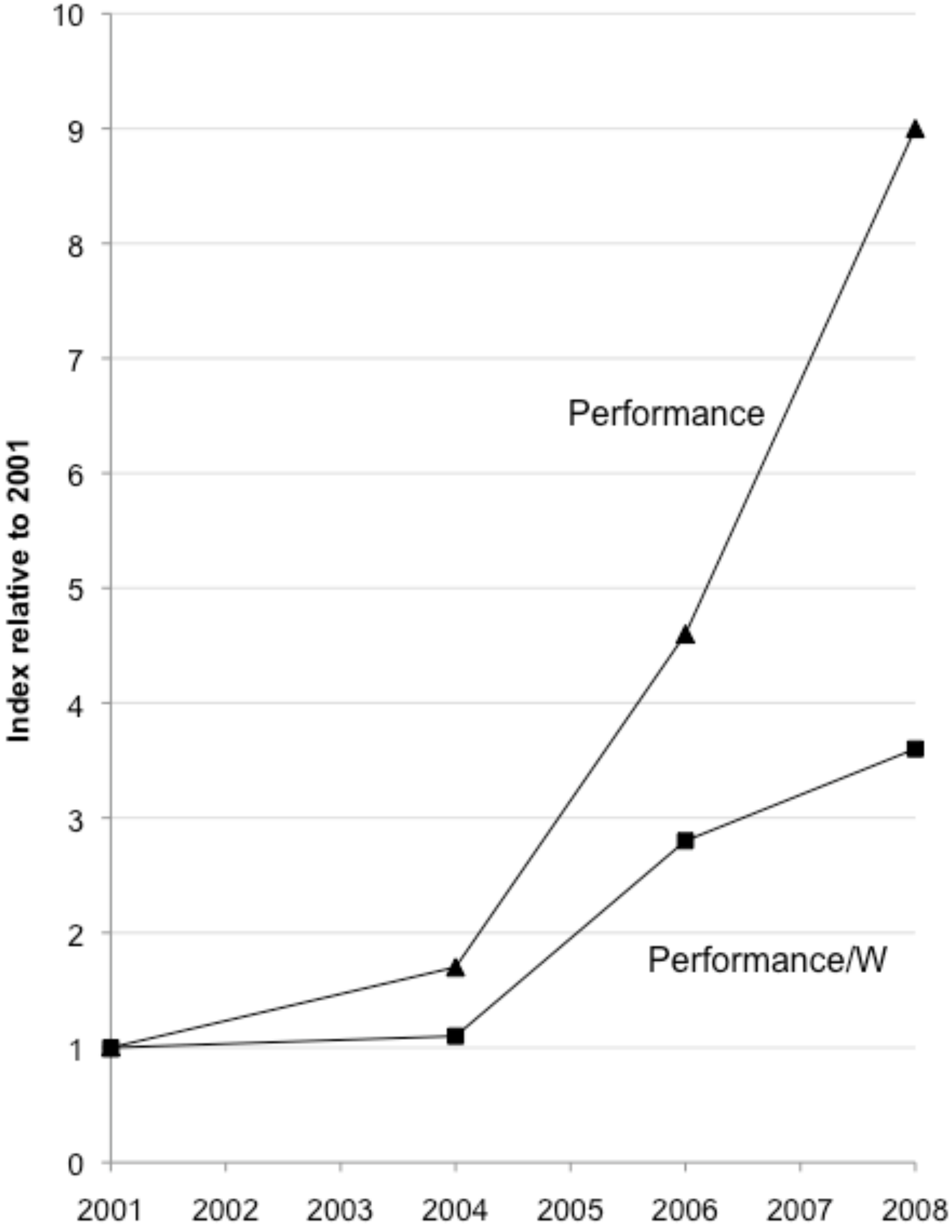


Figure 5: Performance and performance per watt trends for Google servers, 2001 to 2008

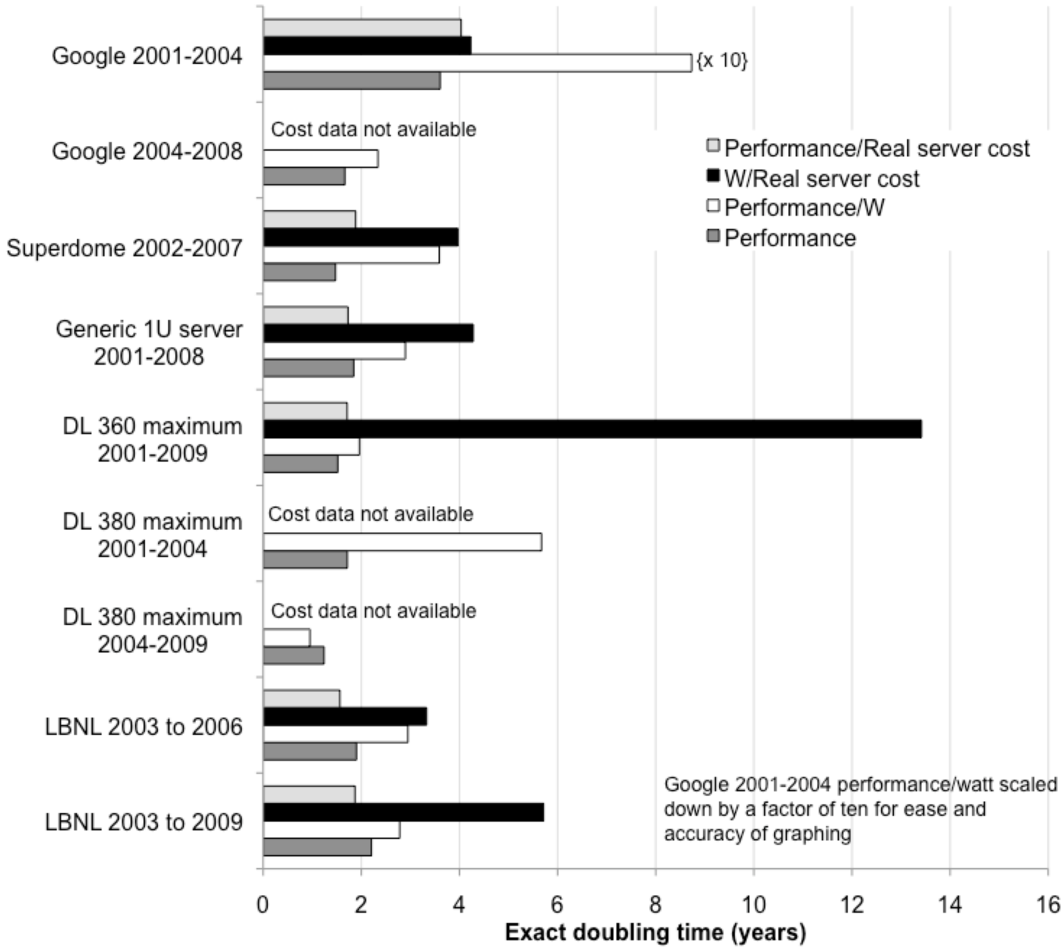


Figure 6: Summary of trends for servers, expressed as doubling time in years

Longer bars mean slower growth. Doubling time calculated using instantaneous exponential growth rates as described in the text.

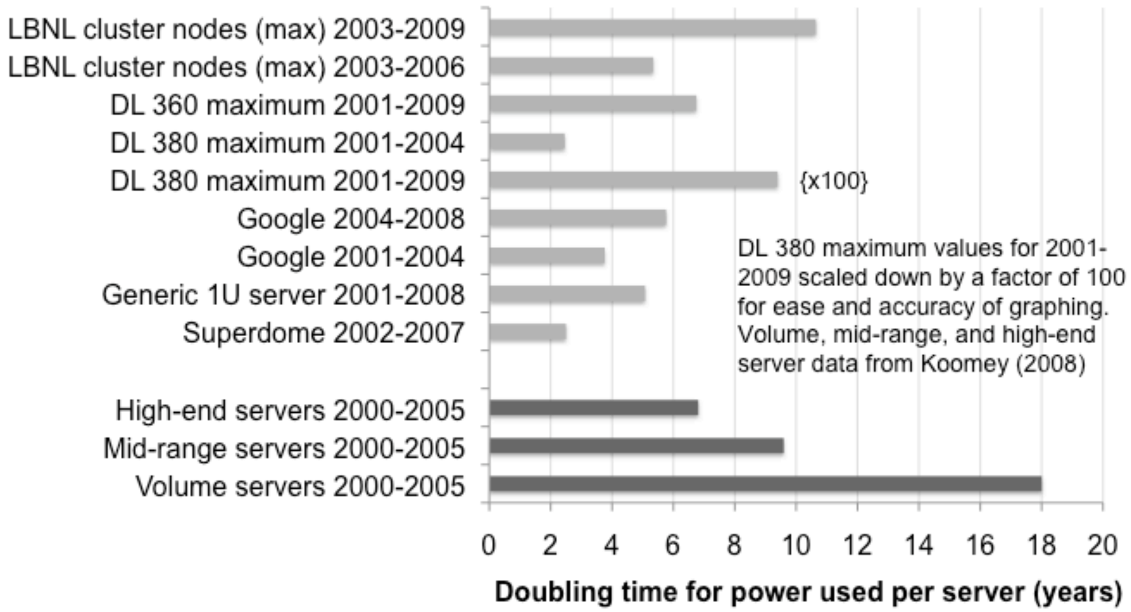


Figure 7: Doubling times for power used per server (years)

For the DL 380 servers, power use per server actually declined from 2004 to 2009, bringing it back to about the 2001 level by 2009 (see Table 6).

Table 1: Trends in generic 1U servers

	Units	*2001*	*2004*	*2008*	*Ratio 2008/2001*	*Instantaneous Annual % Δ*	*Doubling time Years (1)*	*Notes*
Nominal server cost	$	1300	1300	1300				2
Real server costs	2009 $/unit	1584	1482	1326	0.84	-3%	-27.2	3
Three year energy cost	$	500	1000	1300	2.60	14%	5.1	2
Price of electricity	$/kWh	0.1	0.1	0.1				2
Implied electricity use	kWh/year	1667	3333	4333	2.60	14%	5.1	4
Implied power	Watts/server	95	190	247	2.60	14%	5.1	5
Watts/server cost	W/k 2009$	60	128	186	3.11	16%	4.3	6
Performance					13.88	38%	1.8	7
Performance/watt					5.34	24%	2.9	8
Performance/server cost					16.58	40%	1.7	9

Assumptions

Hours per year	8766		
Server lifetime	3	Years	2
PUE	2		2
Load factor	100%		

(1) Instantaneous annual growth rate and exact doubling times calculated using Equations 4 and 5.

(2) Nominal server cost, three year energy cost, price of electricity, server lifetime, and Power Utilization Effectiveness (PUE) taken from Belady, Christian L. 2007. "In the data center, power and cooling costs more than the IT equipment it supports." In *ElectronicsCooling*. February. vol. 13, no. 1. pp. 24-27.

(3) Nominal server cost adjusted for inflation using annual GDP deflator data from BEA (http://www.bea.gov/national/nipaweb/SelectTable.asp?Selected=Y) and assuming 2% inflation from 2008 to 2009.

(4) Implied electricity use for server plus infrastructure (cooling and power delivery) calculated by dividing the three-year energy cost by the average electricity price.

(5) Implied power use per server (not including cooling and power delivery) calculated using the PUE of 2.0, a load factor of 100%, and the length of the year (8766 hours on average, including leap years).

(6) Watts per thousand 2009 $ calculated by dividing the implied power per server by real server costs.

(7) Performance taken from annual % changes in performance of the HP DL360 1U server (Table 4), applied over the period 2001 to 2008.

(8) Performance/watt ratio calculated by dividing the ratio for performance by the ratio for watts/server.

(9) Performance/server cost ratio calculated by dividing the ratio for performance by the ratio for server cost.

Table 2: Trends in HP Superdome high end server, 2002-2007

	Units	2002	2007	Ratio 2007/2002	Instantaneous Annual % Δ	Doubling time Years (1)	Notes
Total system costs	M $	6.39	11.98	1.87	13%	5.5	2
Software costs	M $	1.07	2.02	1.89	13%	5.5	2
Support costs	M $	0.64	0.97	1.52	8%	8.3	2
System cost less software + support	M $	4.68	8.99	1.92	13%	5.3	2
Inflation adjusted system costs	2008 $	5.49	9.18	1.67	10%	6.8	3
Performance (TPC throughput)		389,434	4,092,799	10.51	47%	1.5	2
Performance/watt				2.63	19%	3.6	4
Watts/system				4.00	28%	2.5	5
Watts/server cost	W/k2008$			2.40	17%	4.0	6
Performance/server cost	Perf/k2008$			6.29	37%	1.9	7

(1) Instantaneous annual growth rate and exact doubling times calculated using Equations 4 and 5.

(2) Total system costs, software costs, support costs, and performance taken from runs by the Transaction Processing Performance Council (TPC) for Superdome servers, posted at http://www.tpc.org/tpcc/results/tpcc_results.asp?orderby=hardware. The 2002 runs are for the HP Superdome PA-RISC/750 MHz-64p/64c while the 2007 runs are for the HP Integrity Superdome-Itanium2/1.6GHz/24MB iL3.

(3) Nominal system costs adjusted for inflation using annual GDP deflator data from BEA (http://www.bea.gov/national/nipaweb/SelectTable.asp?Selected=Y) and assuming 2% inflation from 2008 to 2009.

(4) Performance per watt calculated from Figure 1 in Belady, Christian L. 2007. "In the data center, power and cooling costs more than the IT equipment it supports." In *ElectronicsCooling*. February. vol. 13, no. 1. pp. 24-27 (which represents trends for Superdome servers). That graph shows increases of a factor of 75 for performance and a factor of 16 for power use, implying a factor of 4.7 increase in performance per watt over the eight year period 1999 to 2007. That represents a 19% instantaneous annual rate of change in performance per watt. When that rate of change is compounded continuously and extended over the five year period in this example, it results in a factor of 2.63 increase in performance per watt.

(5) To estimate the ratio of power per server (2007 over 2002), we divided the performance ratio by the performance per watt ratio for that same period.

(6) To estimate the ratio Watts per thousand 2008 $ (2007 over 2002), we divided the watts per system ratio by the inflation-adjusted system costs ratio for that same period.

(7) To estimate the ratio of performance per thousand 2008 $ (2007 over 2002), we divided the performance ratio by the inflation-adjusted system costs ratio for that same period.

Table 3: Trends in Google server performance, energy use and costs over time

	2001	*2004*	*Ratio 2004/2001*	*Instantaneous Annual % Δ*	*Doubling time Years (1)*	*Notes*
Performance	10	17.8	1.780	19.2%	3.6	2
Performance/server cost	8.3	13	1.566	15.0%	4.6	2
Performance/watt	8.3	8.5	1.024	0.8%	87.3	2
Nominal server price			1.136	4.3%	16.3	3
Watts/server			1.738	18.4%	3.8	4
Inflation	102.4	109.5	1.069	2.2%	31.2	5
Server price adjusted for inflation			1.063	2.0%	34.0	6
Performance/inflation adjusted server cost			1.674	17.2%	4.0	
Watts/inflation adjusted server cost			1.635	16.4%	4.2	7

(1) Instantaneous annual growth rate and exact doubling times calculated using Equations 4 and 5.

(2) Performance, performance/server price, and performance/watt in arbitrary performance units taken from Barroso, Luiz André. 2005. "The Price of Performance: An Economic Case for Chip Multiprocessing." ACM Queue, special issue on Multiprocessors. vol. 3, no. 7. September. Barroso informed us in an email January 22, 2009 that the three generations of Google servers described in that article covered the period 2001 through 2004, before multicore processing became commonplace.

(3) 2004/2001 ratio for nominal server price derived by dividing the Performance ratio by the Performance/server price ratio above.

(4) 2004/2001 ratio for watts per server derived by dividing the Performance ratio by the Performance/Watt ratio above.

(5) Inflation index based on GDP deflator data from BEA (http://www.bea.gov/national/nipaweb/SelectTable.asp?Selected=Y) and assuming 2% inflation from 2008 to 2009.

(6) Server price ratio 2004/2001 divided by the inflation ratio 2004/2001 to get ratio for server price adjusted for inflation.

(7) 2004/2001 ratio for watts per dollar of inflation adjusted server price derived by dividing the Watts per server ratio by the server price adjusted for inflation ratio above.

Table 4: Characteristics of DL360 servers

	Units	DL360 G1	DL360 G5	G5/G1
Date of hardware availability		Feb-00	Feb-09	
CPU		Intel Pentium III (introduced Oct 99)	Intel Xeon E5450	
Number of Transistors	Millions	56	1640	29.3
CTP CPU performance metric	MTOPS	3734	182000	48.7
GFLOPS performance metric	GFLOPS	3.2	96	30.0
Bus speed	MHz	133	1333	10.0
Operating System (OS)		MS Windows Server 2003, Enterprise Edition, 32 bit	MS Windows Server 2003 x64 Enterprise Edition	
OS Version		5.2.3790 Service Pack 2 build 3790	R2	
JVM version		SPEC Java VM 5.0 (build 1.2.3.4-tricore 20071111)	BEA JRockit(R) (build P27.5.0-5-97156-1.6.0_03-20080403-1524-windows-x86_64, compiled mode)	
Benchmark version		SPECpower_ssj2008 1.1	SPECpower_ssj2008, 1.1	
Power management		Enabled	Unknown	
Power supply input power rating	Watts	292	NA	
Power supply output power rating	Watts	180	700	
Power supply details		LiteOn Model PS-6191	HP part # 399542-B21	
Clock speed	MHz	800	3000	3.75
# of CPUs	#	2	2	1
# of cores/CPU	#	1	4	4
Total # of cores	#	2	8	4
RAM	GB	2	16	
Size of RAM		4 x 512 MB	4 x 4096 MB	
# of Ram slots		4, all populated	PC2-5300F CL5 LP; slots 1A,3A,5B,7B populated	
# of hard disk drives	#	2	1	
Size of each hard disk drive	GB	18	120	
Total hard drive space	GB	36	120	
Hard disk notes			HP, 1.5G, 5.4K, 2.5" SFF SATA HDD	
# and type of Network Interface cards		1, integrated	2 x NC373i	
Network speed	Mbit	10	1000	

Table 5: Comparison of performance, power use, and costs for DL360 servers

	Units	DL360 G1	DL360 G5	G5/G1	Instantaneous Annual % Δ	Doubling time Years (1)	Notes
Year of purchase		2001	2009				
Number of CPU transistors	Millions	56	1,640	29.3	42%	1.6	2
CTP CPU performance metric	MTOPS	3,734	182,000	48.7	49%	1.4	3
Number of processors	#	2	2	1.0	0%	NA	
# of cores	#	2	8	4.0	17%	4.0	
Clock speed	MHz	800	3,000	3.8	17%	4.2	
Performance typical	ssj_ops	879	31,730	36.1	45%	1.5	4
Performance maximum	ssj_ops	8,297	318,769	38.4	46%	1.5	5
Power typical	Watts	86	179	2.1	9%	7.5	6
Power maximum	Watts	124	282	2.3	10%	6.7	7
Performance per watt typical	ssj_ops/W	10.3	178	17.3	36%	1.9	8
Performance per watt maximum	ssj_ops/W	66.8	1,129	16.9	35%	2.0	9
Cost as configured	2001 $	3,000					10, 12
	2009 $	3,656	5,500	1.5	5%	13.6	11, 12
Typical watts/server cost	W/k 2009$	23.4	32.5	1.4	4%	16.8	13
Max watts/server cost	W/k 2009$	33.9	51.3	1.5	5%	13.4	14
Typical performance/server cost	Perf./k 2009$	240	5,769	24.0	40%	1.7	15
Max performance/server cost	Perf./k 2009$	2,269	57,958	25.5	41%	1.7	16

(1) Instantaneous annual growth rate and exact doubling times calculated using Equations 4 and 5.

(2) Number of CPU transistors represents the sum total of transistors on the CPU for all cores.

(3) Composite theoretical performance (CTP) is a synthetic performance benchmark used by Intel Corp. to assess compliance with Federal government export requirements going back to the 80386 processor in the mid to late 1980s; http://www.intel.com/support/processors/sb/CS-017346.htm

(4) Performance typical represents ssj operations per second from the SPEC power benchmark (http://www.spec.org/power_ssj2008/) at 10% computing load (the utilization level of typical business server).

(5) Performance maximum represents ssj operations per second from the SPEC power benchmark (http://www.spec.org/power_ssj2008/) at 100% computing load (which is the utilization level of typical servers in high performance computing applications).

(6) Power typical represents average Wconsumed from the SPEC power benchmark at 10% computing load.

(7) Power maximum represents W consumed from the SPEC power benchmark at 100% computing load.

(8) Performance per watt typical calculated from typical performance and power use.

(9) Performance per watt maximum calculated from maximum performance and power use.

(10) Cost for G1 server estimated from manufacturers suggested retail price on a web site for a refurbisher of DL360 G1 servers: <http://www.networkliquidators.com> and confirmed by original purchase cost of server from purchaser.

(11) Cost for G5 server taken from HP online store, assuming 16 GB of RAM, 120 GB HDD. We needed to subtract the cost of a redundant power supply not included in the SPEC power run ($199, from another HP server on the same site). The SPEC power run used the 80W version of the 5450 (personal communication from Klaus-Dieter Lange at HP to Koomey on March 16, 2009) so that's the one we chose for cost purposes.

(12) Dollars adusted to 2009 using annual GDP deflator data from BEA (http://www.bea.gov/national/nipaweb/SelectTable.asp?Selected=Y) and assuming 2% per year inflation from 2008 to 2009.

(13) Typical watts/thousand dollars typical from server cost and and typical power use.

(14) Typical watts/thousand dollars typical from server cost and and typical power use.

(15) Typical performance/thousand dollars from server cost and and typical performance.

(16) Maximum performance/thousand dollars from server cost and and maximum power use.

Table 6: Power and performance data for several generations of the HP DL380

Values from SPEC power runs	*Year*	*Performance ssj_ops*	*Power W*	*Performance/ Watt*
DL380 G1	2001	7,991	169	47.3
DL380 G4	2004	26,880	394	68.2
DL380 G5 (Intel 5160)	2006	159,151	258	616.9
DL380 G5 (Intel 5430)	2008	306,620	253	1211.9
DL380 G6	2009	439,831	170	2587.2
Indices relative to DL380 G1				
DL380 G1	2001	1.0	1.0	1.0
DL380 G4	2004	3.4	2.3	1.4
DL380 G5 (Intel 5160)	2006	19.9	1.5	13.0
DL380 G5 (Intel 5430)	2008	38.4	1.5	25.6
DL380 G6	2009	55.0	1.0	54.7
Instantaneous growth rates (%)				
DL380 G1 to DL380 G4	2001 to 2004	40%	28%	12%
DL380 G4 to DL380 G5 (Intel 5160)	2004 to 2006	89%	-21%	110%
DL380 G5 (Intel 5160) to DL380 G5 (Intel 5430)	2006 to 2008	33%	-1%	34%
DL380 G5 (Intel 5430) to DL380 G6	2008 to 2009	36%	-40%	76%
DL380 G4 to DL380 G6	2004 to 2009	56%	-17%	73%
DL380 G1 to DL380 G6	2001 to 2009	50%	0%	50%
Doubling times (years)				
DL380 G1 to DL380 G4	2001 to 2004	1.7	2.5	5.7
DL380 G4 to DL380 G5 (Intel 5160)	2004 to 2006	0.8	-3.3	0.6
DL380 G5 (Intel 5160) to DL380 G5 (Intel 5430)	2006 to 2008	2.1	-70.8	2.1
DL380 G5 (Intel 5430) to DL380 G6	2008 to 2009	1.9	-1.7	0.9
DL380 G4 to DL380 G6	2004 to 2009	1.2	-4.1	1.0
DL380 G1 to DL380 G6	2001 to 2009	1.4	939.9	1.4

(1) Instantaneous annual growth rate and exact doubling times calculated using Equations 4 and 5.

(2) Power and performance data for DL380 Generation 1 (G1) taken from the 100% load case for servers with 2 GB of RAM in Table C-2.

(3) Power and performance data for DL380 G4 through G6 machines taken from the 100% load case for each machine from the SPEC power web site: <http://www.spec.org/power_ssj2008/>

Table 7: Characteristics of LBNL cluster computing nodes

	Units	*Whitebox 1 U compute node*	*Dell PowerEdge 1950 Woodcrest*	*Dell PowerEdge 1950 III Harpertown*	*Notes*
Purchase date		May-03	Oct-06	Feb-09	
Processor		Intel Pentium 4 Xeon	Dual Core Xeon 5150 4MB Cache	Quad Core Xeon E5410, 2x6MB Cache	
Processor speed	GHz	3.06	2.66	2.33	
Bus speed		512K/533MHz-FSB	1333MHz FSB	1333MHz FSB	
Theoretical peak performance	Gflops	12.24	42.56	74.56	1
Number of CPU transistors	Millions	110	1350	1640	2
CTP CPU performance metric	MTOPS	24,988	83,346	141,354	3
# of processors		2	2	2	
Cores/processor		1	2	4	
Total cores		2	4	8	
RAM size	GB	2	8	16	
RAM configuration		4x512MB	4x2GB	8x2GB	
# of hard drives		1	1	2	
Hard drive size	GB	40	73	750	
Hard drive characteristics		7200RPM WD400JB	SAS, 3.5-inch 10K RPM	7.2K RPM Un. SATA 3Gbps 3.5-in HotPlug	
Network interface cards		On Board Dual GigE Intel 10/100/1000 Lan	Dual Embedded Broadcom NetXtreme II 5708 Gigabit Ethernet NIC		
Power supply output rating	W	350	Unknown	Dell model number D670P-S1, 670W	
# of power supplies		1	1	1	
Idle power	Watts	113	217	217	4
100% load power	Watts	255	398	371	4
Purchase costs	Nominal $	3250	2750	2750	5
	2009 $	3840	2936	2779	6
Performance/purchase cost	Gflops/k 2009 $	3.2	14.5	26.8	
Performance/power (W max)	Mflops/W	47.9	107.0	200.7	
Power (W max)/purchase cost	W/k 2009$	66.5	135.5	133.7	

(1) Theoretical peak performance estmated as clock speed in GHz times # of cores times # of instructions per clock cycle (2 for the 2003 server and 4 for the other two servers) for each processor.

(2) Number of transistors per processor taken from <http://www.intel.com/pressroom/kits/quickreffam.htm>, multiplied by 2 processors per server.

(3) CTP performance metric taken from <http://www.intel.com/support/processors/sb/CS-028241.htm> and <http://www.intel.com/support/processors/xeon/sb/CS-020863.htm>. May 03 machine assumed to use Intel 519 processor instead of the 519K processor (web site isn't clear which is appropriate). Per processor CTP multipled by number of processors (2).

(4) Power measured by Jared Baldridge (JRBaldridge@lbl.gov) using a PLM meter to track watt-hours over a 30 minute period and then multiplying by 2 to get watt-hours/hour.
For full load measurement LBNL ran Advanced Clustering's software called "Break in" that builds and rebuilds a Linux kernel and tracks utilization of all parts of the server architecture.
<http://www.advancedclustering.com/software/breakin.html>.

(5) Nominal purchase costs (not including tax or shipping) given as a range of $3 to 3.5k for the 2003 server and $2.5k to 3k for the other two. We chose the midpoints of the ranges.

(6) Real server costs calculated using monthly GDP deflators from EIA: http://tonto.eia.doe.gov/cfapps/STEO_TableBuilder/index.cfm and assuming 2% inflation 2008-2009.

Table 8: Calculation of doubling times for key parameters from LBNL computing cluster node data

	Ratio			*Annual % Δ*			*Doubling time (years)*		
	2006/2003	*2009/2006*	*2009/2003*	*2003-2006*	*2006-2009*	*2003-2009*	*2003-2006*	*2006-2009*	*2003-2009*
Theoretical peak performance	3.48	1.75	6.09	36%	24%	31%	1.90	2.9	2.21
Number of CPU transistors	12.27	1.21	14.91	73%	8%	47%	0.94	8.3	1.48
CTP CPU performance metric	3.34	1.70	5.66	35%	23%	30%	1.97	3.1	2.30
Idle power per server	1.93	1.00	1.93	19%	0%	11%	3.61	805	6.06
100% load power per server	1.56	0.93	1.45	13%	-3%	7%	5.35	-23.6	10.65
Inflation adjusted purchase costs	0.76	0.95	0.72	-8%	-2%	-6%	-8.83	-29.4	-12.33
Performance/real purchase cost	4.55	1.85	8.42	44%	26%	37%	1.56	2.6	1.87
Performance/power (W max)	2.23	1.88	4.19	24%	27%	25%	2.95	2.6	2.78
Power (W max)/real purchase cost	2.04	0.99	2.01	21%	-1%	12%	3.33	-118	5.71
# of months	41	28	69						

(1) Data for ratios taken from Table 7.

(1) Instantaneous annual growth rate and exact doubling times calculated using Equations 4 and 5, with elapsed time expressed as months divided by 12.

APPENDIX A: SIMPLE COST MODEL FOR DATA CENTERS

The annualized total cost (ATC) of a data center can be expressed as in Equation (B-1):

$$\mathrm{ATC} = \mathrm{IT} + \mathrm{INF}_{\mathrm{kw}} + \mathrm{INF}_{\mathrm{nonkW}} + \mathrm{EC} + \mathrm{O\&M} \tag{B-1}$$

Where

IT = annualized IT capital

$\mathrm{INF}_{\mathrm{kw}}$ = annualized kW related infrastructure capital;

$\mathrm{INF}_{\mathrm{nonkW}}$= annualized non-kW related infrastructure capital;

EC = annual electricity costs, typically about half for infrastructure and half for direct IT electricity use, and

O&M = annualized operations and maintenance costs.

To annualize capital costs we use the capital recovery factor, defined as

$$\mathrm{CRF} = \frac{d\left(1+d\right)^{L}}{\left(\left(1+d\right)^{L}-1\right)} \tag{B-2}$$

where d is the discount rate (7% real) and L is the lifetime of the equipment (3 years for IT equipment and 15 years for infrastructure equipment).

Of course, what we really care about is the cost per delivered computing cycle. For simplicity, let's assume 100% equipment utilization. This means that the maximum number of computations possible for a given data center over the course of a year is the maximum number of operations per second times the number of seconds per year.[12] Dividing both sides of equation 2 by maximum annual computations we get

$$\frac{\mathrm{ATC}}{\text{Annual Computations}} = \frac{\mathrm{IT} + \mathrm{INF}_{\text{non-kW}} + \mathrm{INF}_{\mathrm{kW}} + \mathrm{EC} + \mathrm{O\,\&\,M}}{\text{Annual Computations}} \tag{3}$$

---

[12] Measuring actual utilization and total computational output is complicated. Most data centers produce more than one type of computing, and the costs and value of that computing varies by time of day and sometimes by geography. For purposes of this simple example we need not worry about these complexities, but companies wrestling with assessing total costs surely must.

Let's assume we'll spend \$1,000 on IT equipment. That yields annualized costs of \$381/year for IT (CRF calculated using a 7% discount rate over 3 years).

Both power related terms (INF_{kw} and EC) can be expressed as a function of the power use per server cost.

The kW related infrastructure capital costs can be expressed as

$$INF_{kw} = \$1{,}000 \text{ x } \frac{\$1\text{k}}{\$1{,}000} \text{ x } \frac{\text{Watts}}{\text{k } 2009 \text{ } \$} \text{ x } \frac{\text{kW}}{1{,}000 \text{ W}} \text{ x } \frac{\$24{,}800}{\text{kW}} \text{ x CRF (7\%, 15yrs)} \quad \text{(B-4)}$$

Where

Watts/k 2009 \$ = 79.6 in the base case, based on the data for all IT equipment in a data center found in Koomey (2007), and

\$24,800/kW = the capital cost of Tier 3 infrastructure in 2009\$ from Koomey (2007), based on Uptime institute data.

The energy costs can be expressed as

$$EC = \$1{,}000 \text{ x } \frac{\$1\text{k}}{\$1{,}000} \text{ x } \frac{\text{Watts}}{\text{k } 2009 \text{ } \$} \text{ x } \frac{\text{kW}}{1000 \text{ W}} \text{ x } \frac{8766 \text{ hours}}{\text{year}} \text{ x LF x PUE x EP} \quad \text{(B-5)}$$

Where

LF = load factor, defined as average electricity load divided by peak load (typically close to 100% for data centers, though climate variations and other factors can reduce this number to 85-90% in some cases),

PUE = Power Utilization Effectiveness, also known as the Site Infrastructure Energy Overhead Multiplier. This term characterizes the ratio of total data center electricity use to the IT electricity use, and it is typically about 2.0, and

EP = Electricity price, which is around \$0.07/kWh for large industrial users in the U.S.

The other two terms can be expressed as a fraction of the annualized IT costs, based on the data in Koomey (2007).

$$INF_{nonkW} = IT \text{ x } 0.19 \quad \text{(B-6)}$$

$$O\&M = IT \text{ x } 0.30 \quad \text{(B-7)}$$

The equations above are used to make Figures 1 and ES-1.

These equations, combined with Equation 1 in the main text, can be used to give quantitative insight about the tradeoffs among the different cost components of data centers. Let's assume Moore's law drives performance per server cost up by a factor of two over a two-year period (a doubling time of two years). The effect on the power-related components of data center costs depends on what happens to power use per server costs (and implicitly, to performance per watt).

Table A-1 shows several scenarios for data center costs, to illustrate the interactions among key parameters. Case 1 corresponds to the costs reported in Koomey et al. (2007), which we treat as the base case. Data center facilities vary a lot, but this source is the most well-documented published data on total costs for data centers currently known to the authors, and it is sufficiently well grounded in current industry practice that relying on it for this schematic example will not lead us too far astray.

The table examines four other cases. For each of these cases the ratio of performance to IT costs doubles compared to the base case. We assume that we always spend $1,000 for IT equipment, which implies that total performance will go up by a factor of two (in this example we rely on arbitrary performance units to simplify the calculations).

For scenario 2, we assume that performance per watt also doubles during this period, implying that watts per thousand dollars will remain constant (as per Equation 1). For scenario 3, we assume that performance per watt will remain the same as in the base case, implying that power use per unit of server cost will double. Scenario 4 assumes that performance per watt triples over two years, implying that watts per thousand dollars of IT cost will reach 2/3 of its value in the base case. Finally, in Scenario 5 we calculate the performance per watt (relative to the base case) that would result in the same total cost per computation as in the base case.

Table A-1: Schematic cost calculation for data centers (based on $1,000 of IT costs)

CASE	*1 (Base)*	*2*	*3*	*4*	*5*
Performance/real server cost (Base = 1.0)	1.00	2.00	2.00	2.00	2.00
Performance/watt (Base = 1.0)	1.00	2.00	1.00	3.00	0.53
Watts/k 2009$	79.6	79.6	159.2	53.1	302.8
Arbitrary number of computations	1000	2000	2000	2000	2000
Annualized costs (2009$)					
IT	$381	$381	$381	$381	$381
kW related infrastructure	$217	$217	$433	$144	$824
Electricity costs	$98	$98	$195	$65	$372
Non-kW related infrastructure	$73	$73	$73	$73	$73
O&M costs	$113	$113	$113	$113	$113
Total	$882	$882	$1,196	$777	$1,763
Index compared to base case	*1.000*	*1.000*	*1.357*	*0.881*	*2.000*
Annualized cost per computation	$0.88	$0.44	$0.60	$0.39	$0.88
Index compared to base case	*1.000*	*0.500*	*0.678*	*0.441*	*1.000*

(1) Case 1: Base case, circa 2007 for a data center delivering high performance computing for financial applications (based on model in Koomey 2007).
Case 2: Computations per $ of IT cost doubles as does performance per watt, keeping Watts/k$ constant.
Case 3: Computations per $ of IT cost doubles and performance per watt doesn't change, making Watts/k$ double.
Case 4: Computations per $ of IT cost doubles and performance per watt triples, reducing Watts/k$ to 2/3 of its base case value.
Case 5: Computations per $ of IT cost double and performance per watt declines almost 50% (enough for increased indirect power related costs to completely offset the IT related reduction in costs per computation).
(2) IT capital expenditures assumed to remain constant at $1,000. Electricity price = $0.07/kWh. Load factor = 100%. PUE = 2.0. Discount rate = 7% real. Lifetime of IT equipment = 3 years, lifetime of infrastructure equipment = 15 years.

APPENDIX B: COSTS, POWER USE, AND PERFORMANCE FOR SELECTED SERVERS

Table B-1 shows selected data from the runs posted on the SPECpower_ssj2008 website <http://www.spec.org/power_ssj2008/>, which we used to produce Figure 2. These runs give power use and performance data for 100+ servers as of June 2009—we focus here on the power levels for maximum performance. We chose 14 of these servers for which to estimate purchase costs using the online stores for Dell, HP, and IBM on March 5 and 6, 2009. No sales taxes, software, support, or shipping costs are included. When small differences in configuration arose (like the online store only giving a price for a server with a redundant power supply when the SPECpower_ssj2008 run for that model included only one power supply) we used data from other parts of the same site to estimate component costs and correct for those differences. Dollar costs adjusted to July 2009 dollars using the monthly GDP deflators from the EIA Short Term Energy outlook custom table builder: <http://tonto.eia.doe.gov/cfapps/STEO_TableBuilder/index.cfm>.

Table B-1: Purchase costs, peak performance, and peak power for selected servers circa March 2009

Company	Model	Peak watts W	Performance ssj_ops	Purchase cost 2009 $	Performance/ power ssj_ops/W	Power/purchase costs W/k 2009 $	Performance/ purchase costs ssj_ops/k 2009 $
Dell	R300	117	155,342	1,589	1,328	74	97,747
Dell	2950III	276	305,413	3,454	1,107	80	88,425
Dell	2970 (2.60 GHz, 2382)	258	331,257	2,580	1,284	100	128,390
HP	DL120G5	136	190,630	1,580	1,402	86	120,641
HP	DL160G5	233	281,914	3,133	1,210	74	89,981
HP	DL180G5	189	282,281	2,761	1,494	68	102,249
HP	DL360G5	282	318,769	5,550	1,130	51	57,439
HP	DL385G5 (2.30 GHz, 2356)	299	240,914	3,717	806	80	64,809
HP	DL385G5p (2.70 GHz, 2384)	257	341,306	4,020	1,328	64	84,902
HP	DL580G5	387	359,523	14,616	929	26	24,598
HP	DL785G5	796	1,066,480	29,359	1,340	27	36,326
IBM	x3250	127	188,975	3,532	1,488	36	53,510
IBM	x3350	125	187,946	2,679	1,504	47	70,156
IBM	x3450	252	323,998	5,443	1,286	46	59,529
HP	DL360G1	124	8,297	3,656	67	34	2,269

(1) Performance and power measurements are posted results from SPEC power for the 100% load case for each server, downloaded from http://www.spec.org/power_ssj2008/.

(2) Cost data estimated using online cost configurators for IBM, HP, and Dell by Jonathan Koomey on 5-6 March 2009. When given an option, we choose the prices for small/medium businesses. Prices do not include operating system or application software, maintenance contracts, taxes, or delivery charges. Prices adjusted to July 2009 $ assuming 2% inflation from July 2008 to July 2009.

(3) DL360G1 machine characteristics from Table 5.

APPENDIX C: SPECPOWER_SSJ2008 RESULTS HP DL360 + DL380 G1 SERVERS

Table C-1: SPECpower_ssj2008 results for HP ProLiant DL360 G1 server

Generation 1 DL360, 2 GB RAM
Run in research mode (no temperature sensor) 24 July 08

Target load	Actual load	ssj_ops	Average Power (W)	Performance to Power Ratio
100%	98.00%	8,297	124	66.8
90%	91.90%	7,780	122	63.7
80%	79.20%	6,709	117	57.2
70%	68.40%	5,789	113	51.2
60%	61.40%	5,199	110	47.3
50%	49.00%	4,151	105	39.7
40%	39.00%	3,301	100	33.0
30%	30.10%	2,549	95.6	26.7
20%	20.10%	1,698	91.1	18.6
10%	10.40%	879	85.6	10.3
Active Idle		0	80.1	0.0
			Σssj_ops/Σpower=	40.5

Generation 1 DL360, 4 GB RAM
Run in research mode (no temperature sensor) 18 Nov 2008

Target load	Actual load	ssj_ops	Average Power (W)	Performance to Power Ratio
100%	99.20%	6,837	127	53.6
90%	88.10%	6,068	124	49.0
80%	80.50%	5,545	121	45.8
70%	70.90%	4,883	118	41.4
60%	61.50%	4,240	114	37.3
50%	48.80%	3,366	108	41.1
40%	40.30%	2,774	104	26.6
30%	28.90%	1,989	99.8	19.9
20%	20.40%	1,404	96	14.6
10%	10.60%	734	90	8.2
Active Idle		0	84.4	0.0
			Σssj_ops / Σpower =	31.9

Table C-2: SPECpower_ssj2008 results for HP ProLiant DL380 G1 server (2 power supplies but only 1 hooked up to the AC analyzer, the other disconnected)

Generation 1 DL380, 2 GB RAM
Run in research mode (no temperature sensor) 24 July 08

Target load	Actual load	ssj_ops	Average Power (W)	Performance to Power Ratio
100%	99.0%	7,991	169	47.2
90%	92.1%	7,438	167	44.6
80%	78.5%	6,340	161	39.4
70%	70.9%	5,728	157	36.5
60%	59.2%	4,781	151	31.7
50%	48.8%	3,936	145	27.1
40%	39.5%	3,187	139	22.9
30%	30.3%	2,445	134	18.3
20%	21.1%	1,707	127	13.4
10%	10.0%	804	120	6.7
Active Idle		0	113	0.0
			Σssj_ops / Σpower =	28.0

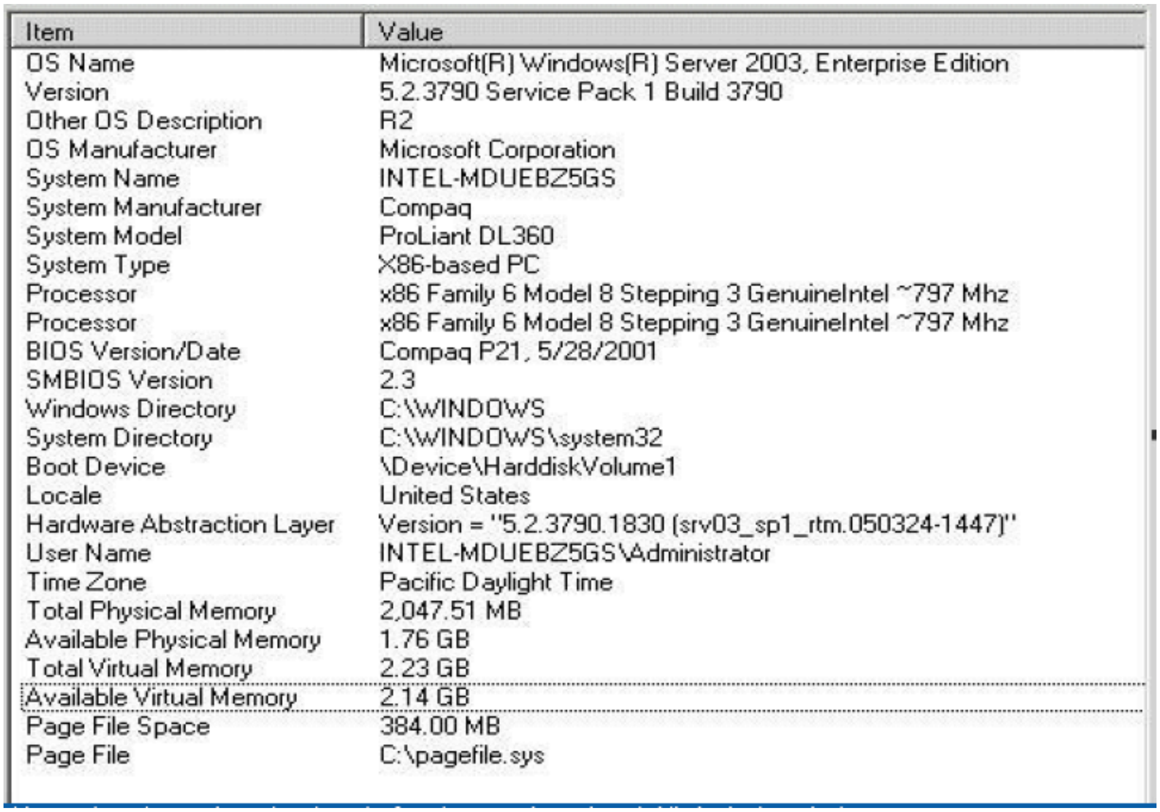
Generation 1 DL380, 4 GB RAM
Run in research mode (no temperature sensor) 18 Nov 2008

Target load	Actual load	ssj_ops	Average Power (W)	Performance to Power Ratio
100%	99.5%	8,419	170	49.5
90%	87.8%	7,432	166	44.8
80%	83.8%	7,089	165	43.1
70%	69.7%	5,894	157	37.5
60%	51.8%	5,230	153	34.1
50%	51.1%	4,326	148	29.2
40%	40.4%	3,416	142	24.1
30%	28.8%	2,438	135	18.0
20%	20.8%	1,765	130	13.6
10%	9.8%	829	123	6.8
Active Idle		0	116	0.0
			Σssj_ops / Σpower =	29.2

APPENDIX D: CHARACTERISTICS OF THE HP PROLIANT DL360 + DL380 G1 SERVERS

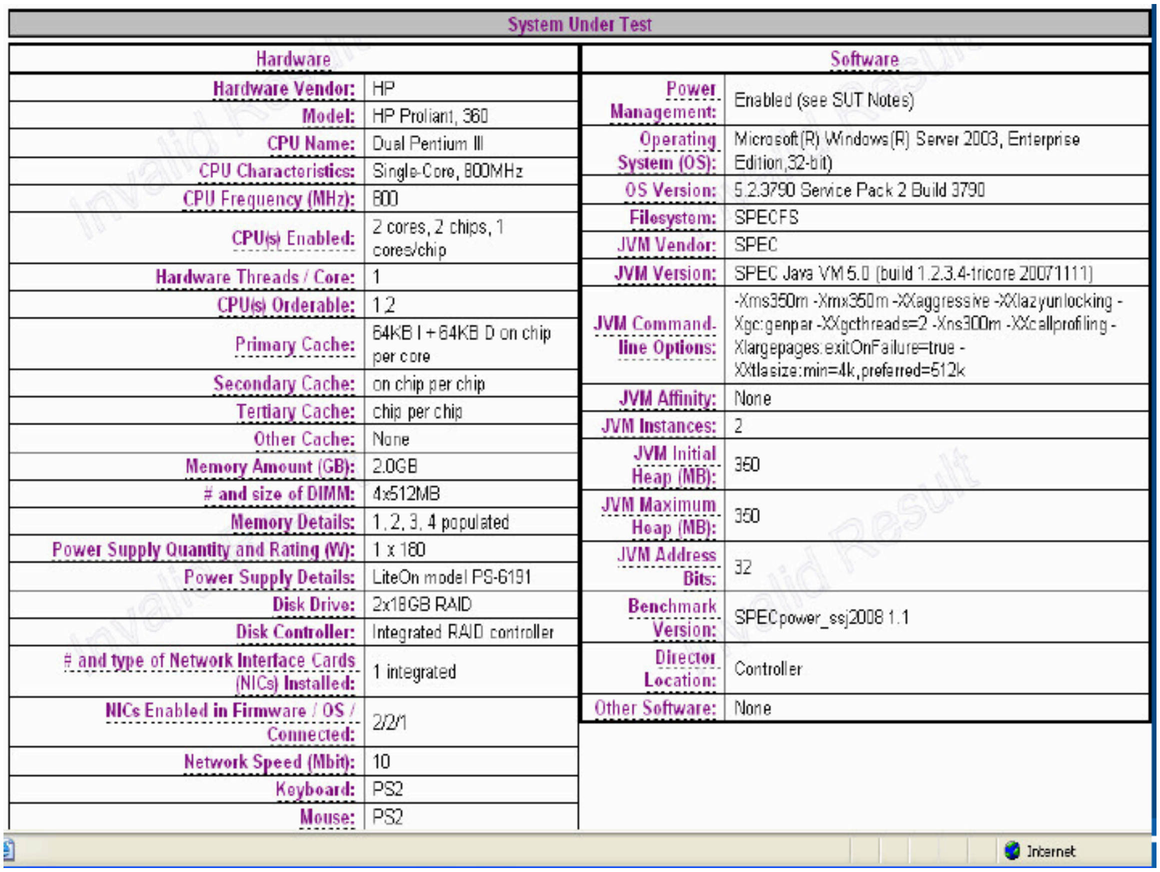
The following pages summarize the characteristics of the HP ProLiant DL360 and DL380 G1 servers upon which Anthony Santos of Intel ran the SPECpower_ssj2008 benchmark in July and November 2008. In the July runs, the servers included 2 GB of RAM, which is half the maximum allowed for this machine. Santos also ran SPECpower_ssj2008 in November 2008 for a 4 GB configuration of the DL360 + DL380 G1 machines and experimented with different power supply configurations for the DL380 machine.

SYSTEM INFORMATION FOR HP PROLIANT DL360, DUAL PIII, 800 MHZ



Item	Value
OS Name	Microsoft(R) Windows(R) Server 2003, Enterprise Edition
Version	5.2.3790 Service Pack 1 Build 3790
Other OS Description	R2
OS Manufacturer	Microsoft Corporation
System Name	INTEL-MDUEBZ5GS
System Manufacturer	Compaq
System Model	ProLiant DL360
System Type	X86-based PC
Processor	x86 Family 6 Model 8 Stepping 3 GenuineIntel ~797 Mhz
Processor	x86 Family 6 Model 8 Stepping 3 GenuineIntel ~797 Mhz
BIOS Version/Date	Compaq P21, 5/28/2001
SMBIOS Version	2.3
Windows Directory	C:\WINDOWS
System Directory	C:\WINDOWS\system32
Boot Device	\Device\HarddiskVolume1
Locale	United States
Hardware Abstraction Layer	Version = "5.2.3790.1830 (srv03_sp1_rtm.050324-1447)"
User Name	INTEL-MDUEBZ5GS\Administrator
Time Zone	Pacific Daylight Time
Total Physical Memory	2,047.51 MB
Available Physical Memory	1.76 GB
Total Virtual Memory	2.23 GB
Available Virtual Memory	2.14 GB
Page File Space	384.00 MB
Page File	C:\pagefile.sys

SYSTEM UNDER TEST CONFIGURATION FOR HP PROLIANT DL360, DUAL PIII, 800 MHZ

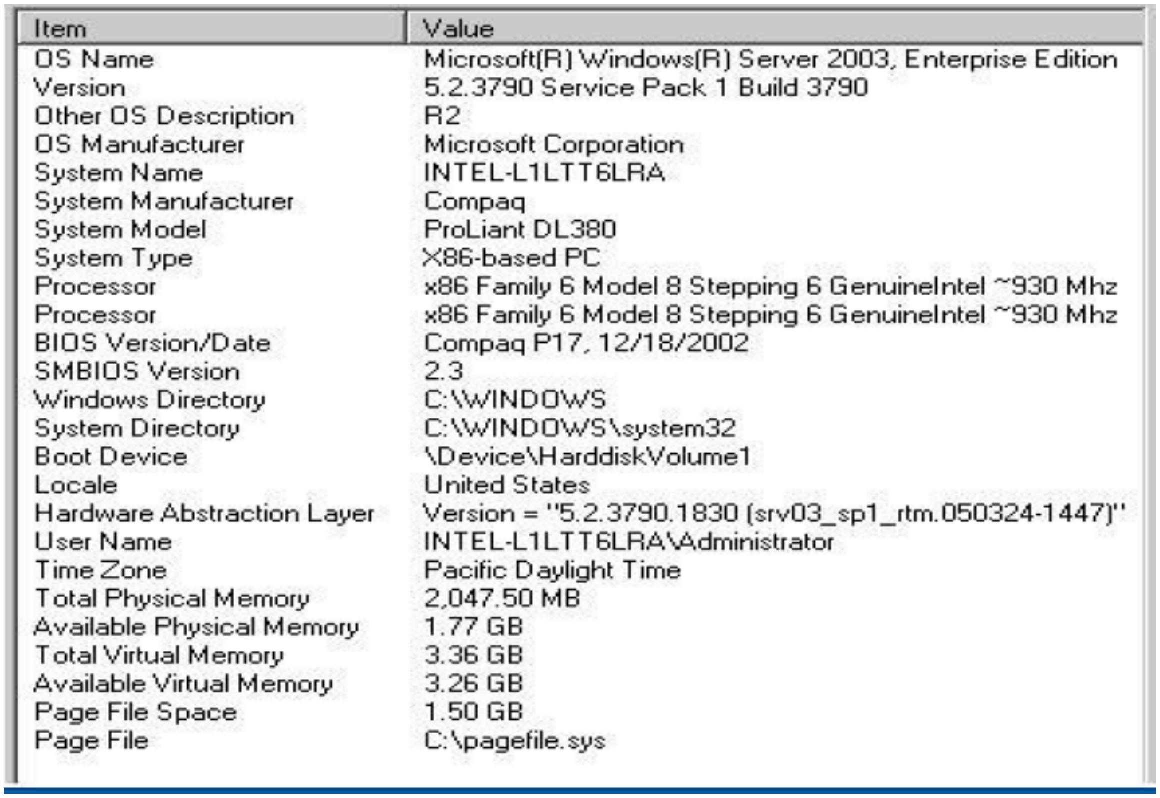


System Under Test

Hardware	
Hardware Vendor:	HP
Model:	HP Proliant, 360
CPU Name:	Dual Pentium III
CPU Characteristics:	Single-Core, 800MHz
CPU Frequency (MHz):	800
CPU(s) Enabled:	2 cores, 2 chips, 1 cores/chip
Hardware Threads / Core:	1
CPU(s) Orderable:	1,2
Primary Cache:	64KB I + 64KB D on chip per core
Secondary Cache:	on chip per chip
Tertiary Cache:	chip per chip
Other Cache:	None
Memory Amount (GB):	2.0GB
# and size of DIMM:	4x512MB
Memory Details:	1, 2, 3, 4 populated
Power Supply Quantity and Rating (W):	1 x 180
Power Supply Details:	LiteOn model PS-6191
Disk Drive:	2x18GB RAID
Disk Controller:	Integrated RAID controller
# and type of Network Interface Cards (NICs) Installed:	1 integrated
NICs Enabled in Firmware / OS / Connected:	2/2/1
Network Speed (Mbit):	10
Keyboard:	PS2
Mouse:	PS2

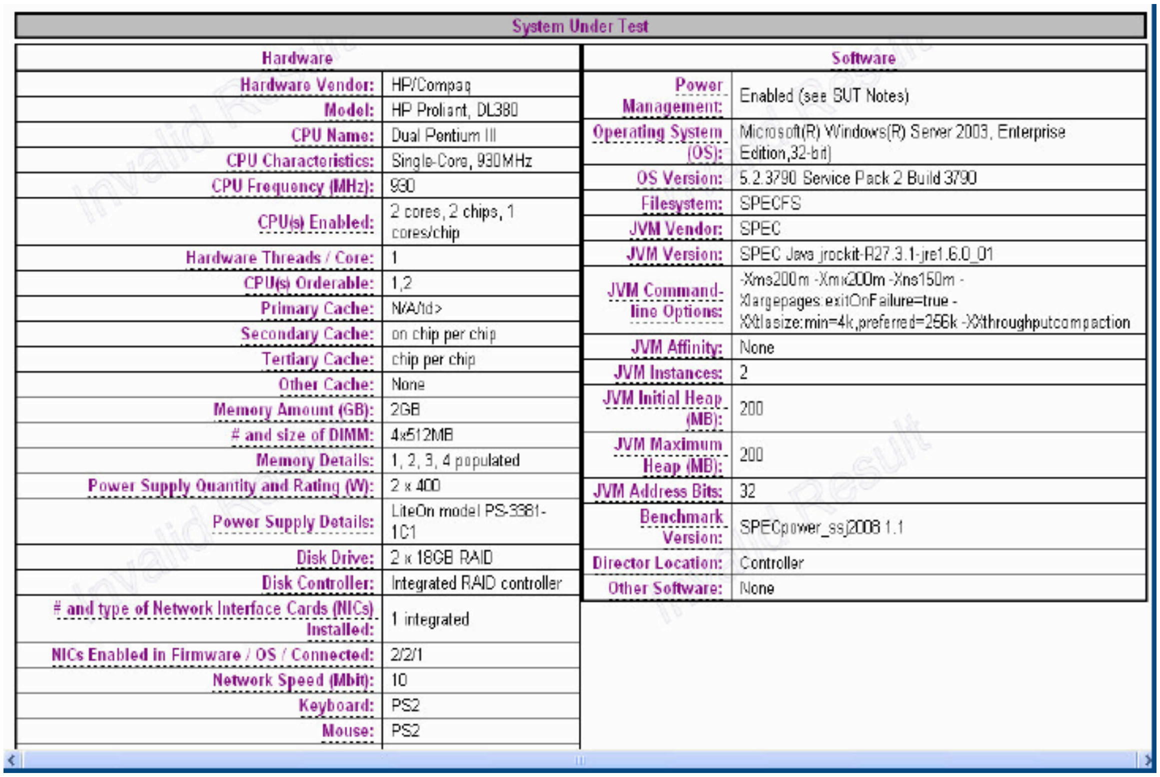
Software	
Power Management:	Enabled (see SUT Notes)
Operating System (OS):	Microsoft(R) Windows(R) Server 2003, Enterprise Edition,32-bit)
OS Version:	5.2.3790 Service Pack 2 Build 3790
Filesystem:	SPECFS
JVM Vendor:	SPEC
JVM Version:	SPEC Java VM 5.0 (build 1.2.3.4-tricore 20071111)
JVM Command-line Options:	-Xms350m -Xmx350m -XXaggressive -XXlazyunlocking -Xgc:genpar -XXgcthreads=2 -Xns300m -XXcallprofiling -Xlargepages:exitOnFailure=true -XXtlasize:min=4k,preferred=512k
JVM Affinity:	None
JVM Instances:	2
JVM Initial Heap (MB):	350
JVM Maximum Heap (MB):	350
JVM Address Bits:	32
Benchmark Version:	SPECpower_ssj2008 1.1
Director Location:	Controller
Other Software:	None

SYSTEM INFORMATION FOR HP PROLIANT DL380, DUAL PIII, 900 MHZ



Item	Value
OS Name	Microsoft(R) Windows(R) Server 2003, Enterprise Edition
Version	5.2.3790 Service Pack 1 Build 3790
Other OS Description	R2
OS Manufacturer	Microsoft Corporation
System Name	INTEL-L1LTT6LRA
System Manufacturer	Compaq
System Model	ProLiant DL380
System Type	X86-based PC
Processor	x86 Family 6 Model 8 Stepping 6 GenuineIntel ~930 Mhz
Processor	x86 Family 6 Model 8 Stepping 6 GenuineIntel ~930 Mhz
BIOS Version/Date	Compaq P17, 12/18/2002
SMBIOS Version	2.3
Windows Directory	C:\WINDOWS
System Directory	C:\WINDOWS\system32
Boot Device	\Device\HarddiskVolume1
Locale	United States
Hardware Abstraction Layer	Version = "5.2.3790.1830 (srv03_sp1_rtm.050324-1447)"
User Name	INTEL-L1LTT6LRA\Administrator
Time Zone	Pacific Daylight Time
Total Physical Memory	2,047.50 MB
Available Physical Memory	1.77 GB
Total Virtual Memory	3.36 GB
Available Virtual Memory	3.26 GB
Page File Space	1.50 GB
Page File	C:\pagefile.sys

SYSTEM UNDER TEST CONFIGURATION FOR HP PROLIANT DL380, DUAL PIII, 900 MHZ



System Under Test

Hardware	
Hardware Vendor:	HP/Compaq
Model:	HP Proliant, DL380
CPU Name:	Dual Pentium III
CPU Characteristics:	Single-Core, 930MHz
CPU Frequency (MHz):	930
CPU(s) Enabled:	2 cores, 2 chips, 1 cores/chip
Hardware Threads / Core:	1
CPU(s) Orderable:	1,2
Primary Cache:	N/A/td>
Secondary Cache:	on chip per chip
Tertiary Cache:	chip per chip
Other Cache:	None
Memory Amount (GB):	2GB
# and size of DIMM:	4x512MB
Memory Details:	1, 2, 3, 4 populated
Power Supply Quantity and Rating (W):	2 x 400
Power Supply Details:	LiteOn model PS-3381-1C1
Disk Drive:	2 x 18GB RAID
Disk Controller:	Integrated RAID controller
# and type of Network Interface Cards (NICs) Installed:	1 integrated
NICs Enabled in Firmware / OS / Connected:	2/2/1
Network Speed (Mbit):	10
Keyboard:	PS2
Mouse:	PS2

Software	
Power Management:	Enabled (see SUT Notes)
Operating System (OS):	Microsoft(R) Windows(R) Server 2003, Enterprise Edition,32-bit)
OS Version:	5.2.3790 Service Pack 2 Build 3790
Filesystem:	SPECFS
JVM Vendor:	SPEC
JVM Version:	SPEC Java jrockit-R27.3.1-jre1.6.0_01
JVM Command-line Options:	-Xms200m -Xmx200m -Xns150m -XXlargepages:exitOnFailure=true -XXtlasize:min=4k,preferred=256k -XXthroughputcompaction
JVM Affinity:	None
JVM Instances:	2
JVM Initial Heap (MB):	200
JVM Maximum Heap (MB):	200
JVM Address Bits:	32
Benchmark Version:	SPECpower_ssj2008 1.1
Director Location:	Controller
Other Software:	None