



CHARACTERISTICS and ENERGY USE of VOLUME SERVERS in the UNITED STATES

**HEIDI FUCHS, ARMAN SHEHABI,
MOHAN GANESHALINGAM, and
LOUIS-BENOIT DESROCHES**

*Energy Analysis and Environmental
Impacts Division, Lawrence Berkeley
National Laboratory*

BRIAN LIM and KURT ROTH
*Fraunhofer Center for Sustainable
Energy Systems*

ALLEN TSAO
Navigant Consulting, Inc.

October 2017

LBNL-2001074

**Lawrence Berkeley
National Laboratory**

Disclaimer

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.

Acknowledgements

The authors would like to thank Steve Greenberg for his review; his comments improved the clarity and quality of this paper.

Title page background image is a derivative of "[Server room](#)" by [Torkild Retvedt](#), licensed under [CC BY-SA 2.0](#).



This work was supported by the Office of Energy Efficiency and Renewable Energy, Building Technologies Program, of the U.S. Department of Energy under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH11231.

Abstract

Servers' field energy use remains poorly understood, given heterogeneous computing loads, configurable hardware and software, and operation over a wide range of management practices. This paper explores various characteristics of 1- and 2-socket volume servers that affect energy consumption, and quantifies the difference in power demand between higher-performing SPEC and ENERGY STAR servers and our best understanding of a typical server operating today. We first establish general characteristics of the U.S. installed base of volume servers from existing IDC data and the literature, before presenting information on server hardware configurations from data collection events at a major online retail website. We then compare cumulative distribution functions of server idle power across three separate datasets and explain the differences between them via examination of the hardware characteristics to which power draw is most sensitive. We find that idle server power demand is significantly higher than ENERGY STAR benchmarks and the industry-released energy use documented in SPEC, and that SPEC server configurations—and likely the associated power-scaling trends—are atypical of volume servers. Next, we examine recent trends in server power draw among high-performing servers across their full load range to consider how representative these trends are of all volume servers before inputting weighted average idle power load values into a recently published model of national server energy use. Finally, we present results from two surveys of IT managers ($n=216$) and IT vendors ($n=178$) that illustrate the prevalence of more-efficient equipment and operational practices in server rooms and closets; these findings highlight opportunities to improve the energy efficiency of the U.S. server stock.

Table of Contents

1. Introduction	1
1.1 <i>Existing Approaches to Power Scaling in Volume Servers</i>	2
2. Data Sources and Methodology	6
2.1 <i>Characteristics of Volume Server Installed Base</i>	6
2.2 <i>Server Lifetimes</i>	7
2.3 <i>Energy Use by Server Components</i>	7
2.4 <i>Online Retail Data</i>	9
2.5 <i>ENERGY STAR Program</i>	11
2.6 <i>SPEC Benchmarking</i>	12
2.7 <i>Survey of Information Technology Managers and Vendors</i>	13
3. Results and Discussion	16
3.1 <i>General Characteristics of Volume Servers at Online Retail Website</i>	16
3.2 <i>Comparing Three Datasets of Idle Power Draw</i>	18
3.3 <i>Server Attributes Leading to Higher Power Consumption</i>	21
3.4 <i>Server Power versus Utilization</i>	25
3.5 <i>Procurement and Management Practices for Servers in Server Rooms and Closets</i>	30
4. Conclusions	43
References	45

1. Introduction

Computer servers undergird our information economy, and continue to grow in number and utility. Larger datacenters run by companies like Amazon and Google are considerably more energy-efficient than they were a decade ago, but the millions of servers located in server rooms and closets across the United States operate less efficiently than those at large datacenters (Shehabi *et al.* 2016). The conventional servers in all of these spaces are typically classified as volume servers, so named since they are sold in high volumes. International Data Corporation (IDC)'s definition of volume servers encompasses "all systems with an average selling value below \$25,000", in contrast to mid-range (\$25,000-\$249,999) or high-end (>\$250,000) servers (IDC 2014), which represent only a small portion of total servers. The field energy use of volume servers remains poorly understood, because they handle heterogeneous computing loads, are quite configurable in terms of hardware and software, and are operated over a wide range of management practices.

Currently, not many empirical data exist regarding volume servers' hardware characteristics, equipment age and efficiency, power management regimes, and their power draw across the range of utilization, from idle to maximum capacity—all of which affect energy consumption. International standards and voluntary labeling programs, such as the U.S. Environmental Protection Agency (EPA)'s ENERGY STAR, report server power draw in idle mode only, and power draw does not scale proportionately with utilization. For personal computers (PCs), several recent studies monitored real-world power draws (Desroches *et al.* 2014, Ecotope 2014, Pixley and Ross 2014), and others involved systematic testing of various configurations to derive a real-world correction factor for idle power consumption (Xergy 2016, Dewart *et al.* 2014)—but nothing similar exists for volume servers.

This paper explores various characteristics of 1-and 2-socket¹ volume servers that influence energy use to better understand how servers consume energy under real-world conditions. First, we draw upon existing IDC data and the literature to establish general characteristics of the U.S. installed base of volume servers. We then present other insights into server characteristics gained from four data collection events of server product specifications from a major online retail website.² Next, we compare cumulative distribution functions of server idle power in three separate datasets and examine the hardware characteristics to which power draw is most sensitive to explain the differences between those datasets. We subsequently examine recent trends in server power draw among high-performing servers across the full server load range and consider how likely it is that those trends apply to all volume servers, before inputting weighted average idle power load values into a recently published model of national server energy use. Finally, we observe to what extent a suite of procurement and management processes are employed in server rooms and closets, via two surveys of IT managers and IT vendors. Throughout, energy use estimates are for servers alone; power consumed by the storage and network equipment also commonly found in IT racks, as well as associated thermal management, are out of scope of this paper. Similarly, these energy use estimates apply only to compute resources, not to storage

¹ In this context, the number of sockets refers to the maximum number of physical processors than can be installed.

² We note that server purchases are also made through traditional, non-online means (*e.g.*, telephone, purchase orders), and make no claims that the range of servers sold on this website are representative of the market as a whole.

equipment or network devices. While not nationally representative given data limitations, results presented are meant to be illustrative of what kinds of servers exist, how much energy they use, and how they are operated.

1.1 Existing Approaches to Power Scaling in Volume Servers

General knowledge about server utilization has amassed in recent years. Volume servers almost never operate at maximum utilization. Usually, short bursts of high activity frequently interrupt servers running most of the time at very low loads or at idle, where efficiency is poor. According to Meisner *et al.* (2009), server utilization is below 30% in typical deployment—and as low as 10% for interactive services like file servers and transaction processing—and while “power-saving features...nearly eliminate processor power consumption in idle systems, present-day servers still dissipate about 60% as much power when idle as when fully loaded.” Koomey (2012) assumes a CPU utilization of 10% for most data centers of both the “typical existing facility” and “typical recent practice” types, and an implied CPU utilization of 40% for cloud computing and high-performance computing installations, which represent only 10% of the installed server base. More recently, NRDC (2014) asserts that average server utilization was static at 12% to 18% between 2006 and 2012. Barroso *et al.* (2013) observe that time spent in the 10% to 50% utilization range is “a perfect mismatch with the energy efficiency profile of modern servers in that they spend most of their time in the load region where they are most inefficient.” The authors argue that energy proportionality must be a computing design goal and report that simulating the energy savings of more energy-proportional servers (the percentage of utilization matching the percentage of peak power consumption) would *halve* server energy use. It is thus important—but challenging—to adequately capture power scaling in mathematical models that seek to estimate server energy use. The following reports all took different approaches to modeling how power increases with utilization, indicative of the level of complexity of this task.

The ENERGY STAR labeling program run by EPA only uses idle state power draw criteria to qualify for certification. Nevertheless, the ENERGY STAR Computers Servers version 2.1 test procedure tests servers in both idle and active states. The active state testing relies on benchmarking software known as the Server Efficiency Rating Tool (SERT), created by the Standard Performance Evaluation Corporation (SPEC). Manufacturers submitting their servers for certification must report efficiency scores for a variety of worklets, software that tests performance and power use for a particular type of computing load (ENERGY STAR 2016a). However, while manufacturers must submit the full set of SERT outputs as part of the certification process, the only power state subject to a maximum power draw criterion is idle. Consequently, evaluation of the SERT data indicates that the workload is incompatible with some servers, and it is challenging to meld worklets’ multiple outputs into a single robust active state energy consumption value.

A draft white paper for the European Union (EU)’s standards development process for enterprise servers and data storage presented eight hypothetical curves to relate arbitrary performance and power values (Intertek 2016), depicted below, which provides a first-principles perspective on the variability of power proportionality along the entire range of server utilization.

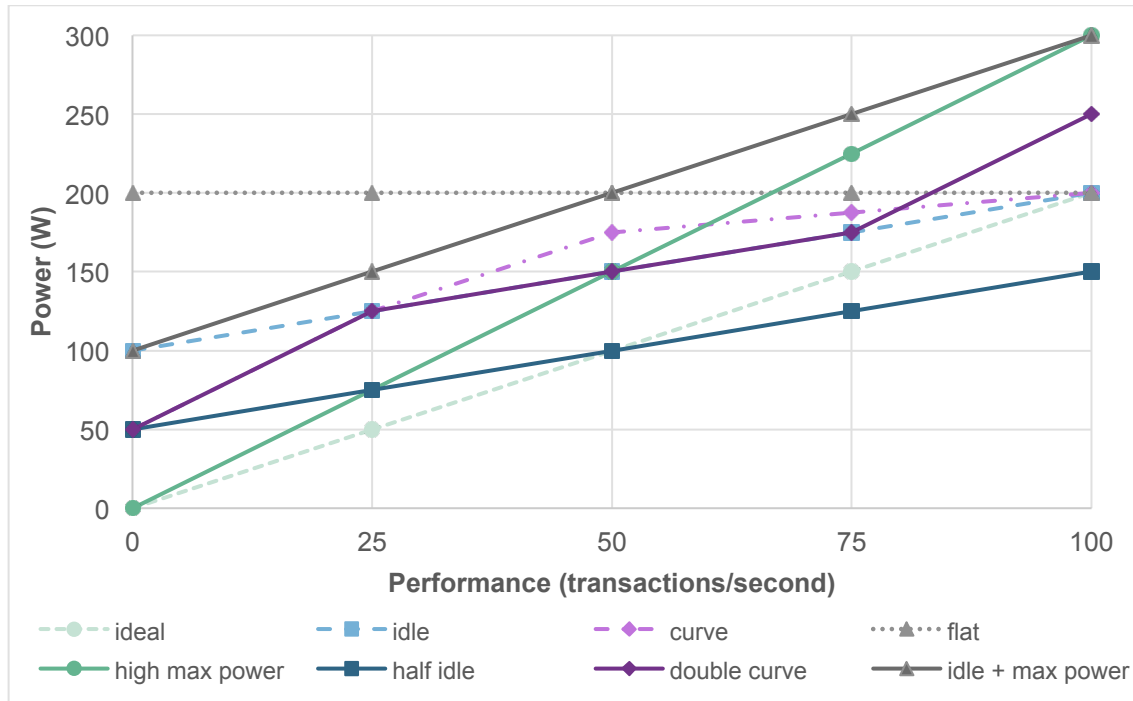


Figure 1.1: Hypothetical server efficiency curves modeled for EU Lot 9 process, adapted from Intertek (2016)

Two of these (“ideal” and “high max power”) curves embody perfectly proportional scaling between performance and power, while three increase linearly but with non-zero power draw at idle (“idle”, “idle + max power”, and “half idle”), and one involves no scaling (“flat”). Two curves depict a more complex relationship: a log curve (“curve”) and an inverse S curve (“double curve”). The least efficient of the theoretical curves are “flat” and “idle + max power.”

There are two reports that have attempted to estimate actual server energy consumption beyond idle power. The U.S. EPA submitted an authoritative report on server and data center energy efficiency to Congress in August 2007. EPA identified several technologies that enable better matching between server utilization and power use: power management through dynamic voltage and frequency scaling (DVFS), which allows microprocessor clock speed to adjust continuously to computational demand; and variable speed internal fans, which provide extra cooling only when needed (EPA 2007). Table 1.1 presents the assumptions upon which EPA relied for percent power at varying utilization levels with and without power management (*i.e.*, DVFS); authors concluded that at the time, only around ten percent of volume and mid-range servers were operated with power management enabled.

Table 1.1: Server percent power at various utilization levels with and without power management (EPA 2007)

Power management state	Server utilization					
	0%	20%	40%	60%	80%	100%
Off	63%	78%	83%	87%	95%	100%
On	51%	62%	71%	77%	93%	100%

Here we see that the energy savings from DVFS depend upon the assumed average utilization level. One practice to increase this level of utilization is virtualization, which involves replacing

several dedicated servers operating at low utilization levels with a single host server operating at higher levels—spending more time in higher-efficiency load regions while displacing the power draw of multiple servers to accomplish the same computing tasks.

A recently released report on datacenter energy usage from Lawrence Berkeley National Laboratory (LBNL), seen as the follow-up to the 2007 EPA report to Congress, updated the model of power draw as a function of utilization level. In order to determine the average power draw at typical utilization, Shehabi *et al.* (2016) used an approach based on dynamic range (DR), the ratio between minimum (at idle) and maximum power. The DR of the installed base was bounded by maximum and minimum DR trends, as seen below.

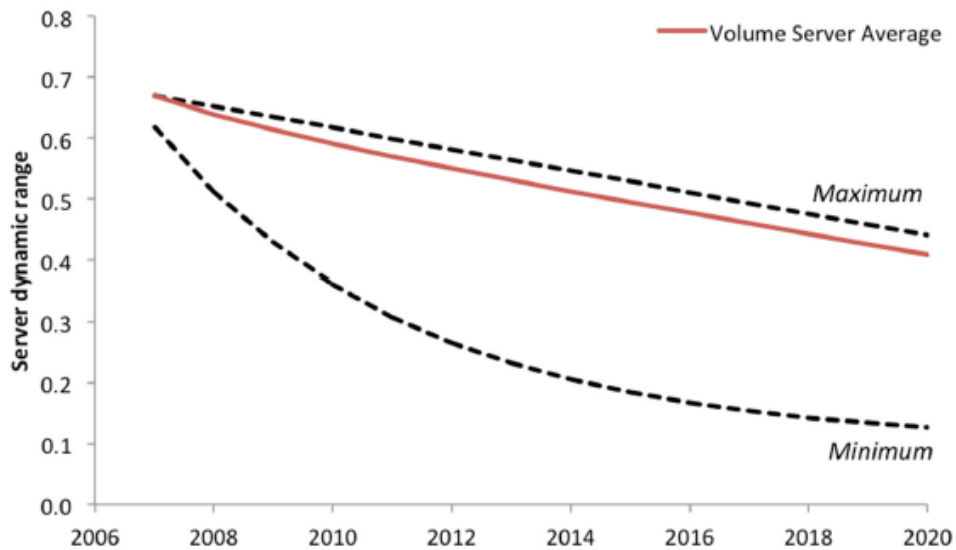


Figure 1.2: Assumed dynamic range of volume servers from Shehabi *et al.* (2016) [Fig. 8 in original; reproduced with permission]

The maximum DR trend shows all servers reaching a DR of 0.44 by 2020³, and slopes linearly from the assumed DR of 0.67 in the 2007 EPA report. The minimum DR trend was informed by a SERT dataset of SPECpower_ssj2008 worklet results. This worklet runs a hybrid workload that involves CPUs, memory, caches, and shared memory processors (SPEC 2016). Shehabi *et al.* modeled the trend from the SPECpower_ssj2008 database values from 2007–2015 as an exponential approaching a DR of 0.1. Finally, the authors determined an average scaling trend between these two bounds based upon an assessment of the current efficiency of the installed base of volume servers. From the solid red line in Figure 1.2, we see the 2016 estimate of average DR among volume servers is 0.48, dropping to 0.41 by 2020.

The limited amount of field data on server utilization has resulted in few publicly available estimates of server utilization trends, and the few that are accessible—namely Brown *et al.* 2007 and Shehabi *et al.* 2016—have only been able to provide simplified ranges of possible utilization trends. In this report we attempt to verify these previously published trends through the use of

³ Based on a SERT database of SPECpower_ssj2008 worklet results and Barroso *et al.* (2013), *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines, Second Edition*. Synthesis Lectures on Computer Architecture. Morgan Claypool Publishers.

online data collection and component-level energy analysis, and compare how national data center energy use estimates deviate when applying newly generated utilization values.

2. Data Sources and Methodology

In this section, we describe the variety of data sources we draw upon to assess various characteristics of servers, as well as their power scaling.

2.1. Characteristics of Volume Server Installed Base

IDC's market research is often cited for shipment and installed base estimates of information technology (IT) products (for example, see Koomey 2011, Yang and Williams 2009, and Bothner 2003). IDC's Worldwide Quarterly Server Tracker from December 2014 indicates that volume servers continue to make up the great majority of the U.S. market when it comes to both installed base and annual unit shipments; in every year from 2010 to 2015, more than 95% of installed servers were volume servers. Within volume servers, 2-socket servers—with the capacity for two installed processors—dominate. In 2010, they made up 79% of all volume server stock, followed by 1-socket servers at 16%. In 2016 through 2018 these shares are projected to shift to 84% and 12%, respectively. In Figure 2.1, we show the U.S. installed base of volume servers from 2010–2018 broken down by server type and socket capability, with historical quarterly sales data from 2010–Q3 2014, and forecasted sales starting in Q4 2014. For the purpose of clarity, categories where market share does not exceed 5% over this period are not displayed: ≥ 4 -socket servers of any type, as well as 1-socket density-optimized and blade servers.

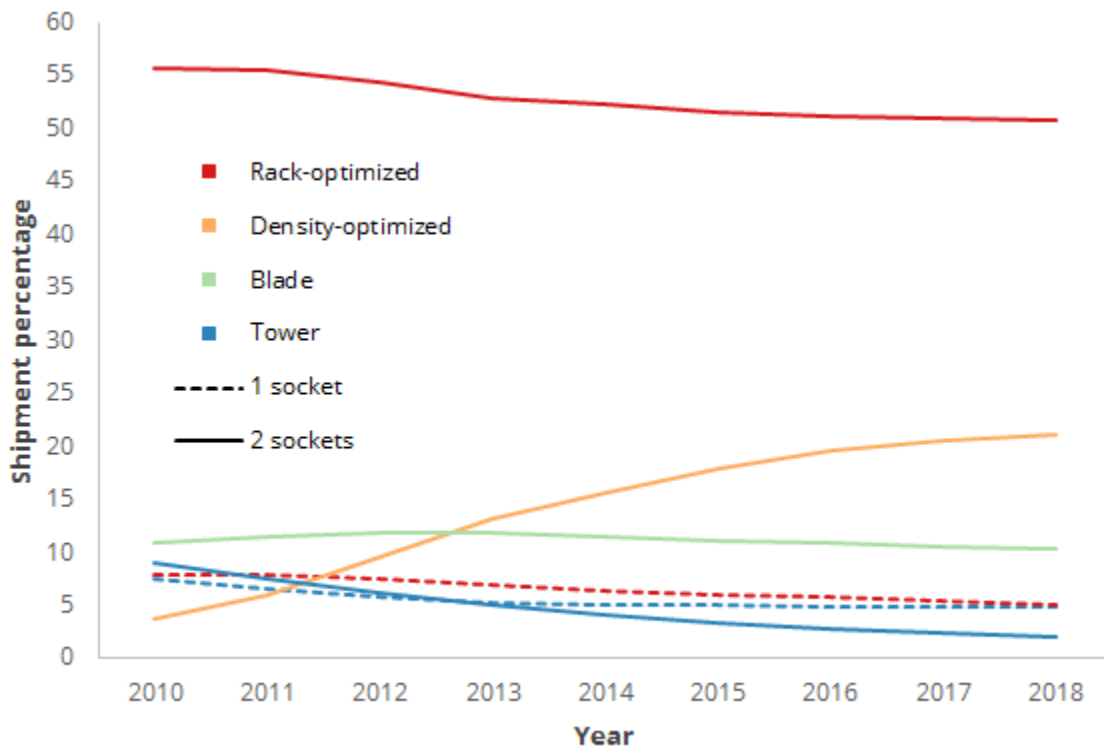


Figure 2.1: Percentage breakdown of U.S. installed base of volume servers (IDC 2014)

IDC data also project a notable growth in density-optimized servers from 2010-2018. IDC defines density-optimized servers as rack servers designed for large-scale datacenter environments with parallelized workloads, which means they differ from rack servers in having more storage. Many newer servers may thus have enough installed storage in order to handle certain applications (*e.g.*, Netflix streaming) to fall into the density-optimized category from the rack-optimized one. We also see that tower servers increasingly have only one socket, while the ratio between 2-socket and 1-socket servers among the rack-optimized and density-optimized form factors is roughly 10:1 between now and 2018.

2.2. Server Lifetimes

To determine typical server lifetimes, we searched the existing literature to establish minimum and maximum average lifetimes from an array of sources.⁴ Because many servers are replaced with new ones long before the hardware fails—if not decommissioned entirely—we chose the shorter “economic lifetime” over a lengthier “technical lifetime” where such a distinction was made. We noted the maximum and minimum values when a range for lifetime was given; if only one figure was presented, we recorded this as both the minimum and maximum value. Lifetime values from the literature are in Table 2.1.

Table 2.1: Literature values for server lifetimes

Minimum lifetime (years)	Maximum lifetime (years)	Source
5	10	Personal communication with industry experts, 2015
3	7	Bio by Deloitte and Fraunhofer IZM, 2014
3	5	Scaramella <i>et al.</i> , 2014
3	5	The Green Grid, 2012
3	10	Opinion Dynamics Corporation, 2011
6	6	Navigant Consulting, Inc., 2009

We computed servers’ average lifetime by taking the mean of all corresponding values; minimum average lifetime is 3 years, median 5 years, mean 5.5 years, and maximum 10 years.

2.3. Energy Use by Server Components

Servers are highly configurable, with many interchangeable powered components. To assess which hardware components are major determinants of server power draw and thus guide our analysis, we relied upon testing and analysis performance by Navigant Consulting on

⁴ Internal parts of servers can easily be replaced (*e.g.*, hard drives may be replaced more often than the entire server). These lifetimes refer only to the whole server, though section 3.5 presents data on required retirement age and average age of servers, power supply units, and hard drives.

approximately 20 different server configurations (personal communication with Navigant staff, May 21, 2015). The systems Navigant built and tested were representative of pre-assembled systems on the market, and their main approach was to build a full system and meter power after swapping one particular component at a time, with all other configurations and settings held constant.

Navigant’s baseline server unit consisted of one processor socket and one processor (Intel Xeon Haswell-based CPU); a single power supply unit (PSU) meeting 80 PLUS Standard qualifications; 16 GB of installed DDR3 RAM; no additional I/O interfaces beyond two 1 Gbit/s Ethernet (GbE) interfaces; and *no* graphics processing unit, dedicated hardware implementing redundant array of independent disks (RAID) level 5⁵, or integrated display. To determine the power of this base unit, Navigant measured the power of a minimal server configuration that included an HDD and installed storage, and then subtracted out the DC power of the HDD and storage. Navigant then tested another motherboard from the same manufacturer that had these features, but supported two sockets instead of one. The idle AC power use of this base unit was measured to be 46.4 W. Incremental power use associated with two sockets and one processor is 2.9W, while the incremental power of two sockets and two processors is 44.0W. This means that an additional processor requires ~95% of the power in idle mode that the base unit alone does, even for more efficient units. The number of processors is thus likely the largest single factor in energy use when it comes to individual components; accordingly, we split our analyses in the rest of this paper into 1- and 2-processor volume servers.

When it comes to power supplies, Navigant performed redundant PSU efficiency testing for the 80 PLUS Standard, 80 PLUS Bronze, 80 PLUS Gold, 80 PLUS Platinum, and 80 PLUS Titanium PSUs, assuming that a server in idle mode would draw ~10% of maximum output capacity of an appropriately sized power supply. Table 2.2 shows the additional power use associated with a relatively large PSU of 1200W DC at these certification levels.

Table 2.2: Idle power use associated with PSU certification levels and redundant PSUs

80 PLUS type	Idle AC power use for given configuration (W)		Difference in power use between redundant and non-redundant configurations (W)
	1 x 1200W PSU loaded at 10%	2 x 1200W PSU loaded at 5%	
Standard	157	184	27
Bronze	152	175	23
Gold	145	160	15
Platinum	141	154	13
Titanium	135	143	8.2

Table 2.2 shows that using a single 80 PLUS Standard power supply on a server demanding 120W requires an additional 37W of power that is lost, but replacing the Standard with a Titanium power supply would result in only 15W of additional losses. Correspondingly, the difference in

⁵ This refers to the RAID capability of the motherboard and system. Navigant tested multiple drives in some configurations, but the power for the baseline product is that with no installed storage.

power use between redundant and non-redundant configurations decreases as power supply efficiency rises.

Next, Navigant tested the impact of adding various storage drives to establish idle power use per terabyte (TB) of desktop storage as presented in the table below. These estimates, shown in Table 2.3, were taken from measurements of 2 TB drives.⁶

Table 2.3: Idle power use associated with server storage types

Storage unit type	Idle DC power use per TB (W)
3.5" HDD	3.87
2.5" HDD	0.89
Solid-state drive (SSD)	0.37

Although the component fraction of energy use in idle mode attributable to storage is small relative to the number of processors and power supply efficiency, the idle power savings on a percentage level associated with 2.5" HDDs and SSDs are notable: a 2.5" HDD uses 22% of the idle power of a 3.5" HDD, while the substitution of an SSD for a 3.5" HDD would require less than 10% of the power in idle mode.

2.4. Online Retail Data

In order to capture key characteristics of servers on the market over recent time, we relied upon four separate instances of automated web data collection over a 13-month period using customized web-crawling software (see Gerke *et al.* 2015 for an explanation and previous application of this software). Because most servers are sold to businesses, we harvested online product specifications for individual server models sold by CDW Corporation at www.cdw.com. CDW is a Fortune 500 reseller of IT products to businesses, government, and institutions, with net sales from July 1, 2015 to June 30, 2016 exceeding \$13 billion (CDW, 2016). Data collection events occurred on March 18, 2015, September 3, 2015, February 3, 2016, and May 5, 2016. The software loads a page from CDW listing the available server models, and then follows links to each individual model webpage and extracts desired data on product specifications from the HTML code that is the foundation of the page. After examining the dataset, we filtered out records where neither "server", "tower", "rack", nor "blade" was present in the short description field, in order to exclude records for non-servers, such as stand-alone processors, memory, PSUs, network management devices, racks, power cables, etc. Records for servers with more than two installed processors were also omitted, so that the dataset would be more comparable to ENERGY STAR and SPEC data. This process netted data for 1,026, 832, 634, and 792 individual server models, respectively, on each of the aforementioned dates. Note that because Dell did not partner with CDW as an authorized reseller until October 2015 (Dell 2015), and Dell's website was

⁶ These numbers can be assumed to scale linearly for 2.5" HDDs and SSDs beyond 2 TB, given the way capacity is currently added for drives of that size. For 4 TB 3.5" HDDs, the associated idle power increase is not quite double that at 2 TB.

incompatible with our web-crawling software, the first two data collection events exclude all Dell-branded servers. Consequently, we did not weight the data by manufacturer market share or other considerations.

These online retail data allow us to explore the range of certain characteristics of servers sold at CDW from March 2015–May 2016: the number of installed processors, the number of maximum installed processors, the type of installed storage, the number of installed PSUs, and the penetration of 80 PLUS certified PSUs. Because few product specification webpages at www.cdw.com are identically formatted, our general approach was to report results based on the data present in relevant fields. For instance, for the number of installed processors and maximum installed processors, we queried only where the fields “N processor” and “N processor max” held data. Neither of these fields held zeroes, meaning that some servers sold without installed processors look identical in this respect to some servers where product specifications were not in a format that allowed the web-crawling software to capture these data within these fields. Thus, reported results in section 3.1 are only for those server models with data present (and by extension, where minimum installed processors equals 1).

The number of records with data on installed processors ranged from 75–79% of total records over the four collection events, and from 96–98% for maximum installed processors. The number of records with data on drive type varied from 84–88%; moreover, while numerous servers’ product specifications mentioned SSD capabilities in the product description section⁷, drive type was only captured in terms of form factor: 3.5” and 2.5”. Next, the share of records with data on the *number* of installed PSUs spanned from 63–71%; however, the proportion of records that show some level of 80 PLUS *certification* for PSUs ranges from only 54–56%.⁸

In addition to examining available combinations of certain server characteristics with a sizeable influence on energy use, we estimated the idle mode power draw of all the servers in the pool of web-collected data by drawing upon the Navigant analysis in section 2.3. Records for 917 unique models across the four data collection events contained sufficient data on those product features; the analysis yields idle power draw associated with each of these components. Most of the 917 models (86%) were listed for sale with no installed storage, while a small fraction (1.7%) were sold with no installed memory indicating that purchasers are likely to modify the servers sold at CDW; for instance, they may install additional storage, memory, processors, or PSUs prior to use. In order to make these servers more representative of actual use cases, we assigned the median values for hard drive capacity and RAM among those servers sold *with* to those sold *without* storage and/or memory. In so doing, we determined these median values for those servers based on number of sockets and number of processors, and assigned the corresponding median value to servers missing storage and/or memory with those same socket and processor characteristics. Here we assume that the median is a more representative value than a random

⁷ For instance, “Industry-unique Lenovo AnyBay design allows multiple storage types in the same drive bay, including front-accessible PCIe SSD for ultimate performance” at <https://www.cdw.com/shop/products/Lenovo-ThinkServer-RD550-70CX-Xeon-E5-2670V3-2.3-GHz-8-GB-0-GB/3474848.aspx?pfm=srh>.

⁸ Even after parsing the “Long description” and “Miscellaneous features” fields for the next word following the phrase “80 PLUS,” as well as the phrase “xx% efficiency power supply”, and mapping the latter strings to the certification levels given for 230V internal redundant power supplies at Ecova (2016).

distribution of hard drive capacities among those servers with installed storage, given existing correlations between product features (*i.e.*, CPU, RAM, and storage).

When considering whether model counts from CDW adequately reflect the market for volume servers, we did not weight the dataset used to examine trends across time, because Dell computers were present only in the two most recent harvesting events, meaning those data were more representative of the market than those contained in the two earlier datasets. However, when estimating idle power associated with the server models sold at CDW, we applied weights—derived from the manufacturer market share among IDC 2014 shipments data—under the assumption that weighting offers an improvement over raw model counts in estimating feature distributions. The applied manufacturer market shares from IDC are shown in Table 2.4. Brands explicitly listed are those that were present in the CDW data set and had reported market share in the IDC data. “Other” brands in the CDW data (*i.e.*, those not mentioned below) were weighted by taking the sum of the market share reported by IDC for the brands not explicitly listed.

Table 2.4: Manufacturer market share of computer server sales in the U.S. (IDC, 2014)

Brand	Market Share
Acer	<1%
Cisco	7%
Dell	25%
HP	24%
Lenovo	3%
Oracle	<1%
Other	41%

When comparing unweighted to weighted mean and median results for idle power draw, the typical difference is within 5–10%. Unweighted results do not change the qualitative results, and median idle power draw values are less affected than mean values. In section 3.2 we present weighted median values as those most likely to be representative of the market.

2.5. ENERGY STAR Program

The ENERGY STAR certification program for enterprise servers has been in effect under version 2.1 since December 16, 2013. In 2016, the most recent year for which data are available, ENERGY STAR penetration among servers was estimated to be only 18% (and 15% in 2015), so the list of certified models is not representative of the wider marketplace (ENERGY STAR 2017 and 2016b).⁹ However, we can consider the distribution of power draw in active-idle mode by number of installed processors for ENERGY STAR qualified servers, in comparison to other servers on the market and industry benchmark versions. We assume that manufacturers pursue certification for

⁹ Some product specifications among the online retail data discussed in section 2.4 contained a field for ENERGY STAR certification; the observed proportion of certified servers ranged from 7–10% over the four data collection events. We excluded these data from further analysis in section 3.1 because this field was not present for every product, so some certified servers may not have been captured.

all products that would qualify, because such certification would confer advantages in the marketplace¹⁰, and thus that ENERGY STAR qualified volume servers are indeed more energy-efficient than those on the broader market.

To consider only volume servers, we examined records from the May 17, 2016 Qualified Products List dataset of ENERGY STAR enterprise servers where server type was listed as “1- or 2-socket server (neither blade nor multi-node)”, resulting in 131 models after validating data.¹¹ All but one of these models was available on the market after January 1, 2012. We considered only power draw results at typical or single configuration, which “lies between the minimum power and high-end performance configurations and is representative of a deployed product with high volume sales” (ENERGY STAR 2013). Similarly to online retail data, we weighted results for measured idle power by the relative prevalence of server brands among 2014 shipments to form a more representative picture of the market than possible through raw counts of qualified models.

2.6. SPEC Benchmarking

SPEC created the SPECpower initiative to develop a benchmark to measure the performance and power draw of servers. In December 2007, SPEC released SPECpower_ssj2008; as mentioned in section 1.1, this worklet assesses computing and power performance of a blend of diverse processor- and memory-intensive tasks, so we chose it as the most representative worklet for our purposes of examining power draw trends at various levels of server utilization (as did Hsu and Poole 2013, Subramaniam and Feng 2013). As of late April 2016, manufacturers had voluntarily published 506 SPECpower_ssj2008 performance reports on SPEC’s website (<https://www.spec.org/results.html>). 381 records were for servers with one or two processors; we analyzed this subset of data because servers with one to two processors dominate the server market in the United States (see section 2.1.). Listed form factors for these servers in our dataset are 1U, 2U, 3U, 4U, tower, or 1 of 2 blades in 1U rackmount frame (for three tested models). 1U, 2U, and tower servers make up 278 of 295 records, or 94%, for which form factor is listed. Because manufacturers self-submit test results for publication, we believe there is some self-selection bias that favors the inclusion of servers in the database that are more energy-efficient than the market as a whole (Shehabi *et al.* 2016 make similar assumptions). As such, we view these data as predictive of server performance and energy-saving potential in the near future. Correspondingly, instead of assuming these servers represent the wider market, we focus on how the broad trends in these data over time may apply to all volume servers in the future.

Unlike the online retail and ENERGY STAR data, the SPEC data allow us to look at power consumption at server utilization levels other than idle, as the worklet output reports power draw at every 10% utilization interval from idle to 100%. We calculate median percentage power vs. server utilization for all SPEC servers by the year hardware was available, from 2007-2016, to explore how power scaling has improved since the EPA 2007 report to Congress.

¹⁰ For example, all federal agencies must procure ENERGY STAR or Federal Energy Management Program (FEMP) designated products when possible, according to the Energy Policy Act (EPAct) of 2005 (Federal Acquisition Regulation Subpart 23.203 2005).

¹¹ We removed a record where idle power draw at typical configuration greatly exceeded idle power draw at maximum power configuration.

2.7. Survey of Information Technology Managers and Vendors

The manner in which servers are operated (*e.g.*, equipment age and efficiency, power management regimes, and typical degree of utilization) strongly influences their energy use, yet few data on these practices exist. We conducted two separate online surveys of IT managers ($n=216$ completed surveys) and vendors¹² ($n=178$) in March 2015 across a broad range of commercial sectors in order to understand more about procurement and management practices for volume servers and related equipment. We include analysis only of fully completed surveys. These survey results have not been published before; we present selected new findings related to server equipment requirements (*e.g.*, required retirement age; whether ENERGY STAR or 80 PLUS certification is mandatory) and practices in server rooms and closets (*e.g.*, frequency of monitoring CPU, memory, and disk utilization or power consumption; requirements for practices that promote energy efficiency).

Survey respondents were recruited via survey panels with access to IT professionals who manage server rooms and closets, as well as the ability to effectively screen participants based on their qualifications. The survey vendor recruited people to join the panel via web advertising (at any site participating in Google AdSense at the time) or through loyalty-program partnerships. New members completed a profile; those who were employed indicated the department in which they work, their industry, and their employer size. The survey vendor invited members whose profile specified they worked in IT, screening out any who no longer worked in IT. The vendor also prioritized those who had not taken a survey in the past week, so each day of fielding, a new pool of people was available. The approach used is a purposive (non-probability) sampling methodology, subject to unknown bias and error.

Table 2.5 shows the range of experience among responding IT managers (corresponding to their position titles), while Figure 2.2 displays a breakdown of the commercial/industry classification(s) of the company at which they work. IT managers could select more than one industry type, so the percentages presented are of the number of total respondents (216).

Table 2.5: Surveyed IT managers' position titles

Position title	% of IT manager survey respondents
IT Manager	37%
IT Director	24%
IT/Systems Administrator	16%
CTO/CIO	16%
Other IT staff	7.4%

¹² Vendors were conceived of as value-added resellers for IT equipment, system integrator consultants, those offering leased services, or service providers that offer long-term maintenance, ongoing monitoring, or help desk services.

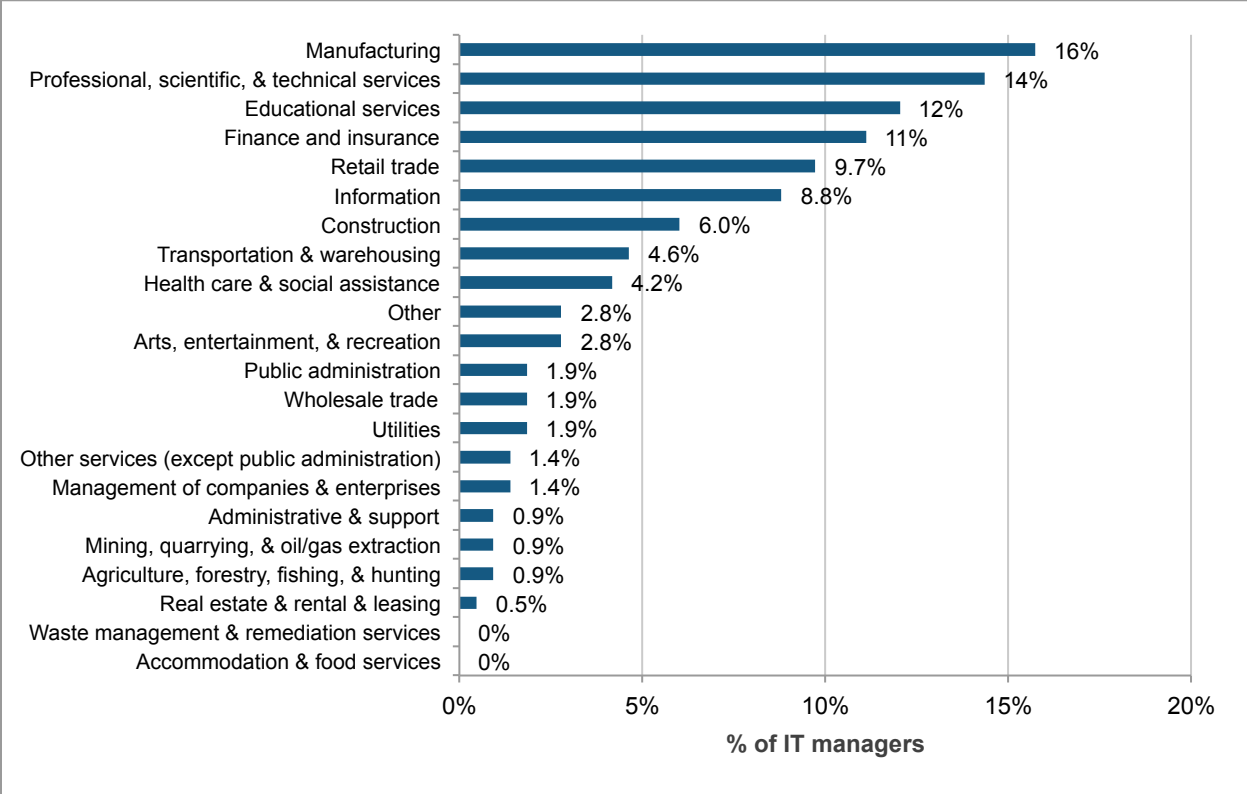


Figure 2.2: Commercial/industrial classification of surveyed IT managers' companies

Industries in the “other” category were specified by respondents as electronics, pharmaceuticals, computer software, travel, defense, and electronics (once each). Similarly, Table 2.6 and Table 2.7, respectively, present the range of experience among IT vendor respondents (corresponding to their position titles) and their companies’ roles in the server room and commercial PC market. Figure 2.3 displays which industry types to which IT vendors’ companies provide services; almost all vendors’ companies provide services to multiple types of organizations. Excepting the question regarding position titles, IT vendor respondents could select multiple answers; percentages presented are thus calculated out of the number of total respondents (178) rather than the number of total responses.

Table 2.6: Surveyed IT vendors' position titles

Position title	% of IT vendor survey respondents
IT Director	39%
IT Manager	30%
IT Engineer	10%
IT Consultant	9.0%
Vice President	4.5%
Business Analyst	4.5%
Technical Support	2.2%
Network Engineer	0.6%

Table 2.7: Surveyed IT vendors' companies' roles in server room and commercial PC market

Company's role in server room and commercial PC market	% of IT vendor survey respondents
Service provider offering long-term maintenance and/or IT support	61%
System integrator consultant	52%
Value-added reseller for IT equipment	49%
Leased services	43%
Other	1.1%

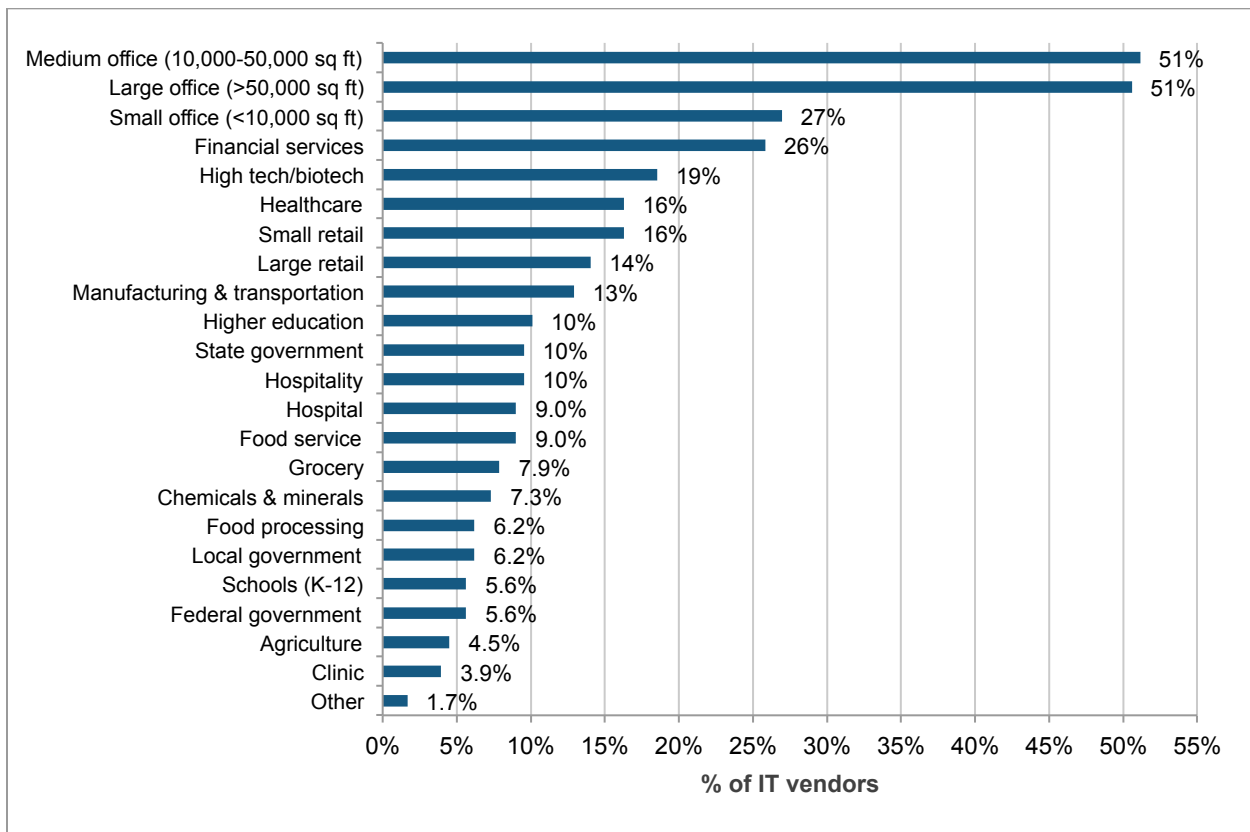


Figure 2.3 Industry types to which IT vendors' companies provide services

In the figure directly above, industries in the “other” category were specified by IT vendor respondents as travel, software, and IT (once each).

3. Results and Discussion

3.1. General Characteristics of Volume Servers at Online Retail Website

Figure 3.1 through Figure 3.4 present insights gleaned from the data on volume servers for sale at cdw.com assembled during four data collection events. Note that these data are simple unweighted model counts because the first two collection events exclude Dell-branded servers, given that CDW was not selling them at the time. This introduces a potential source of bias, unknown in magnitude and direction; however, the large n achievable from automated data harvesting should make these data reasonably robust. We see consistent results across the time span of web harvesting (March 2015 through May 2016) with a slight proportional increase in single processor servers when Dell-branded servers are included in the last two collections. Across all collections, around three quarters of volume servers sold by CDW have one installed processor, yet most have the capacity for two processors.

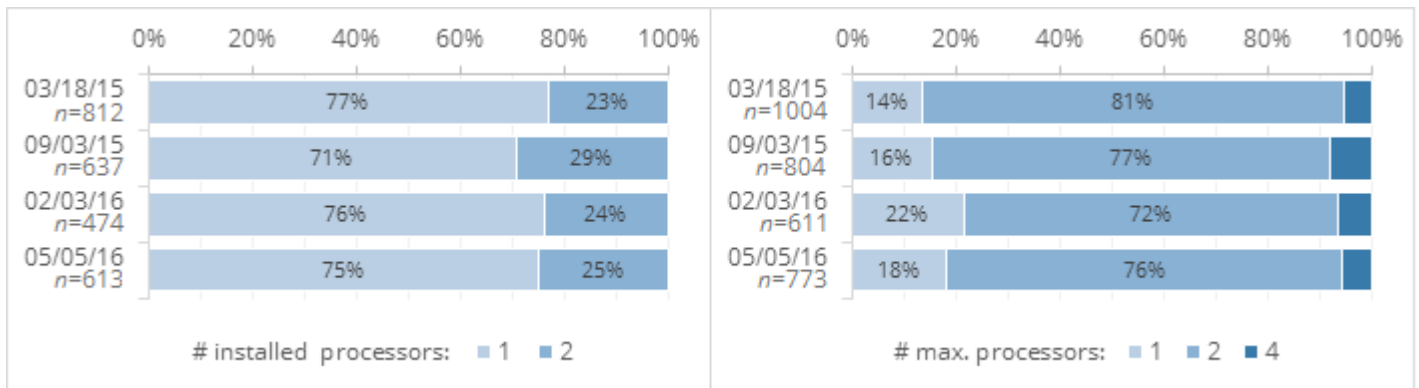


Figure 3.1: Number of installed processors (L) and maximum processors (R) in volume servers at cdw.com

Similarly consistent are results for drive sizes seen in Figure 3.2. 2.5” drive bays (which can contain a 2.5” HDD or SDD) consistently predominate, likely because of the advantages conferred by a smaller form factor, SSD capability, and reduced power requirements. 3.5” drive bays require an additional adaptor to fit an SSD.

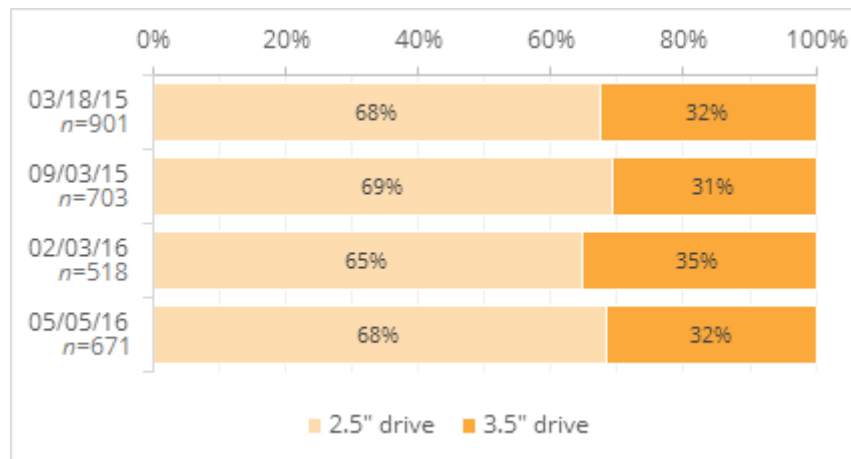


Figure 3.2: Volume server drive form factors at cdw.com

Also quite unvarying over time are the number of PSUs sold with volume servers at CDW; around 60% are sold with one PSU, while the other 40% are sold with two.

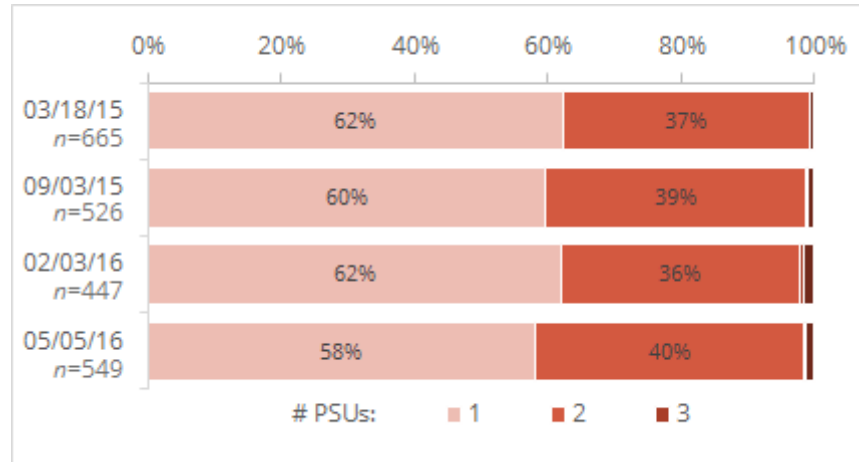


Figure 3.3: Number of PSUs sold with volume servers at cdw.com

Along with the number of installed PSUs when sold, we can examine their power conversion efficiency, as certified by the 80 PLUS program. For the characteristic of whether a PSU is 80 PLUS certified, the absence of data in the relevant field may—or may not—be meaningful; that is, the lack of data may mean that the PSU is not efficient enough to achieve 80 PLUS certification, or that the product specifications were formatted in a manner incompatible with our web-crawling software, or that the manufacturer did not disclose any information on PSU efficiency, 80 PLUS or not. Because the absence of data is not meaningless, in Figure 3.4 we present the distribution of 80 PLUS certified PSUs both with (top) and without (bottom) the records that contained no data.

The proportions and temporal trends in the top chart that includes no data records can be compared to the limited literature on 80 PLUS PSUs. A 2008 report by ACEEE claimed that 80 PLUS and ENERGY STAR unit sales should achieve no less than 8% of market penetration by the end of 2009 (Rasmussen and Wickes 2008). In 2011 the Northwest Energy Efficiency Alliance (NEEA) released a report prepared by Navigant signifying that the market share of 80 PLUS for desktops was 37% in 2010. Interviewed market actors “indicated that 80 PLUS (or equivalent) market share is likely significantly higher for servers than for desktop PCs, but there is little publicly available market data about U.S. server sales and thus Navigant was unable to estimate 80 PLUS market share for server power supplies” (NEEA 2011). However, by 2013 NEEA’s market share claim of 80 PLUS PSUs increased to 70% for desktops in the Northwest region (NEEA 2013). The bottom chart makes plain that when considering only certified power supplies, 80 PLUS Platinum PSUs are becoming markedly more prevalent among CDW’s volume servers over time, chiefly at the expense of less-efficient Gold ones.

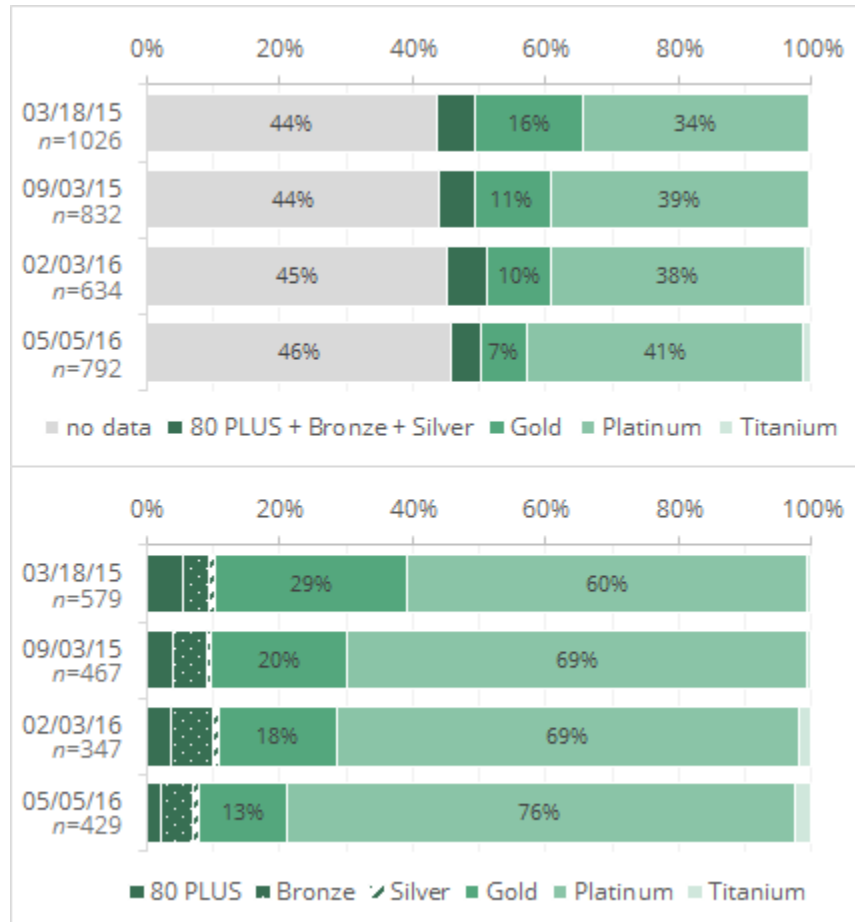


Figure 3.4: 80 PLUS PSUs of volume servers at cdw.com; top includes records with no data or certification

Except for the relative proportions of 80 PLUS certified PSUs, the results pertaining to online retail data remain fairly consistent across the four data collection events spanning fourteen months. When considered in concert with the IDC data presented in section 2.1, they support the assertion that most volume servers have two processors, but that the second processor is typically purchased separately from the server. On the other hand, approximately 40% of servers are already sold with redundant power supplies, although an additional PSU can be installed post-purchase to enable redundancy. Among the approximately 55% of power supplies that are listed as being 80 PLUS certified, 80 PLUS Platinum and Gold are in the majority, shifting over time even more towards Platinum—the second-most efficient certification of the 80 PLUS program.

3.2 Comparing Three Datasets of Idle Power Draw

Next, we present summary statistics and cumulative distribution functions (CDFs) of 1- and 2-processor servers' idle power from our three databases of individual server models. Because all but one record in the ENERGY STAR dataset were available from 2012 onward, included are SPEC records only where the date of hardware availability was later than 2011, as well as all online retail data from February and May 2016 collection events where sufficient data were present to

allow the calculation of idle power.¹³ As mentioned previously, each record is assigned a brand weight, so the CDFs and averages are also weighted. These datasets are summarized in Table 3.1.

Table 3.1: Datasets for which we compare idle power

Data source	Year	# of observations	Assumed biases
Online retail data from cdw.com	2016	917	Dataset consists of all servers for which there was sufficient information in product specifications to calculate idle power draw, with unknown biases; raw records by model then weighted by server brand market share. Servers on the market in 2016 may be more powerful than those over the whole period of 2012-2016. Servers in large data centers purchased through direct-OEM or ODM ¹⁴ contracts may differ from retail selections. Energy calculations based on Navigant (2014) component-based estimates.
ENERGY STAR qualified products list	2012–2016	131	Dataset consists of models that are weighted by server brand market share; ENERGY STAR servers are, by definition, more energy-efficient than most servers
SPECpower_ssj2008 database from 2012 on	2012–2016	156	Dataset consists of models that are weighted by server brand market share; self-selection bias, meaning servers are more power-proportional than market as a whole

Table 3.2 presents summary statistics for the idle power of volume servers in these three datasets, weighted by manufacturer market share and separated by number of processors. Figure 3.5 and Figure 3.6 display the cumulative distribution functions and weighted averages of idle power, similarly separated.

¹³ We chose data only from the latter two of these four collection events, assuming that they are more representative given that CDW did not sell Dell servers until late 2015.

¹⁴ OEM stands for original equipment manufacturer; ODM is an acronym for original design manufacturer.

Table 3.2: Summary statistics of volume server idle power in SPEC, ENERGY STAR, and online retail datasets

	Idle power draw (W) of volume servers					
	1-processor servers			2-processor servers		
	SPEC, 2012 onwards (n = 22)	ENERGY STAR (n = 34)	Online retail data (n = 651)	SPEC, 2012 onwards (n = 134)	ENERGY STAR (n = 97)	Online retail data (n = 266)
Minimum	11.6	18.1	55.5	36.0	56.8	98.4
25th Percentile	14.8	35.3	81.3	50.3	90.1	146
Median	15.8	60.1	95.8	54.1	114	166
Average	32.3	55.7	94.6	57.4	133	160
75th Percentile	19.9	69.3	109	60.7	149	174
Maximum	83.2	233	156	115	493	406

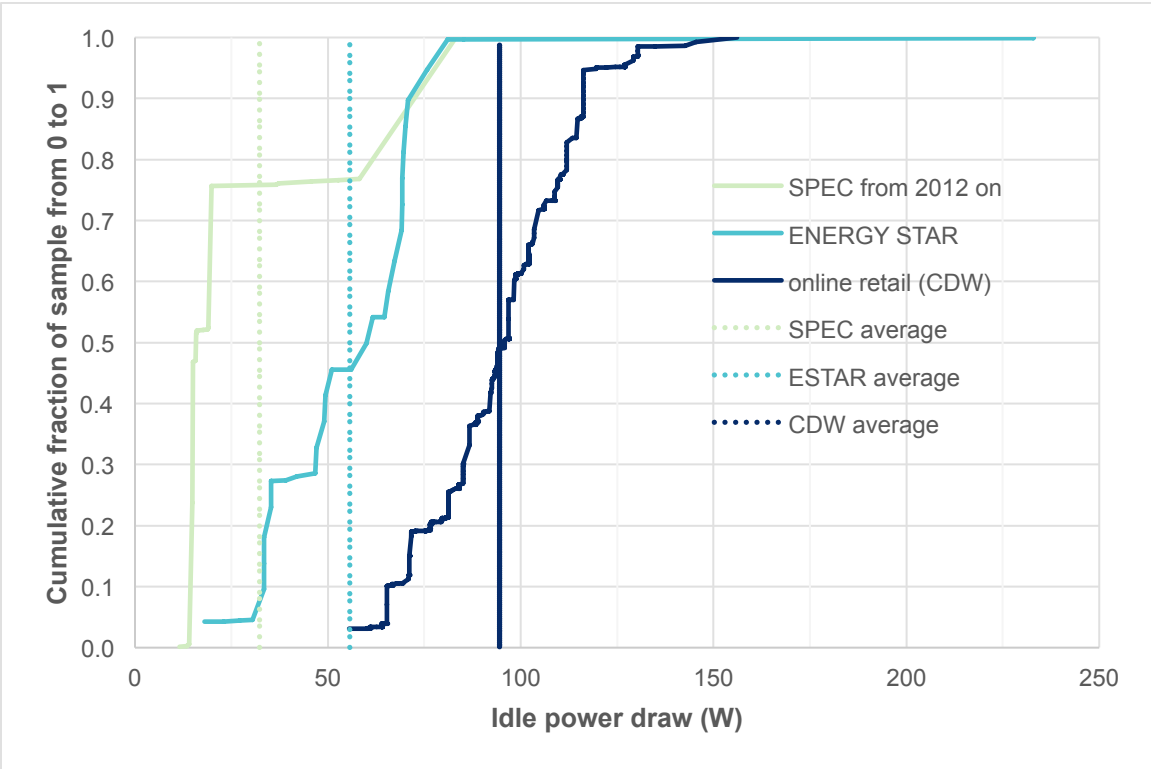


Figure 3.5: Weighted cumulative distribution function and averages of idle power draw, 1-processor servers

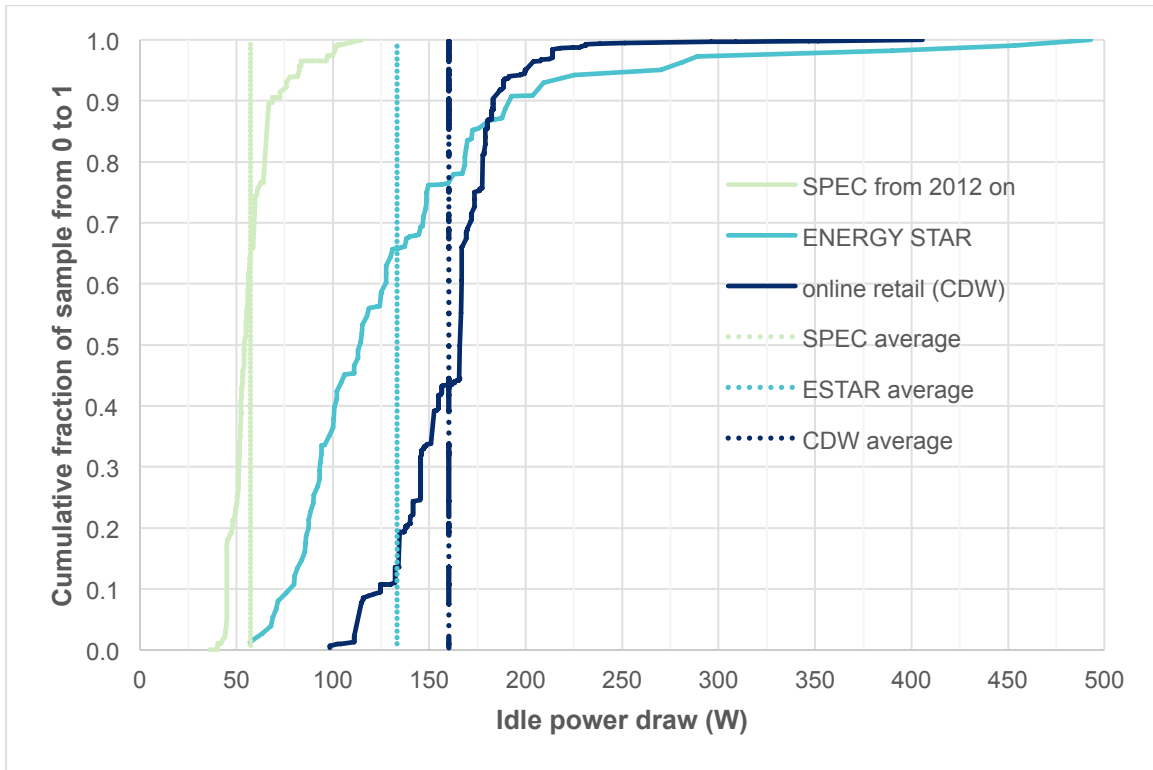


Figure 3.6: Weighted cumulative distribution function and averages of idle power draw, 2-processor servers

For both 1- and 2-processor servers, most SPEC servers use less power in idle mode than do ENERGY STAR servers, which in turn need less power than the CDW servers. In fact, the 75th percentile of idle power draw for SPEC servers only slightly exceeds the *minimum* power draw of ENERGY STAR qualified servers, for both 1 and 2-processor servers. This strongly supports our earlier assertion that servers present in the SPEC dataset are significantly more power-efficient than most, given the voluntary nature of the benchmark and bias among manufacturers. Moreover, the weighted median idle power of ENERGY STAR qualified servers is approximately two-thirds the median value of the servers sold at CDW, which is reasonable given that 15% of servers were ENERGY STAR qualified in 2015 (ENERGY STAR 2016b). In the next section, we investigate what different features of servers that constitute these datasets might be responsible for these disparities in idle power. Given the rapid evolution of server technology, the chronological differences among the data sets may be contributing to their differences in energy use, which would indicate an increase in idle power over time. However, other studies have documented a trend towards lower idle power (Shehabi *et al.* 2016). A decreasing idle power over time is also observed within the SPEC data itself (see Section 3.4).

3.3 Server Attributes Leading to Higher Power Consumption

While an underlying principle of energy efficiency labels and standards programs is that products with comparable features and performance often consume different amounts of energy, examining the data more closely might reveal whether certain server characteristics are systematically associated with higher power draw. Data outliers can sometimes reveal surprising

and useful information¹⁵, so first we turn to investigating the features of outlier servers relative to non-outlying data, both presented below. Here outliers are defined as those data points lying outside 1.5 times the interquartile range from the end of a box (*i.e.*, from the 25th or 75th percentile).

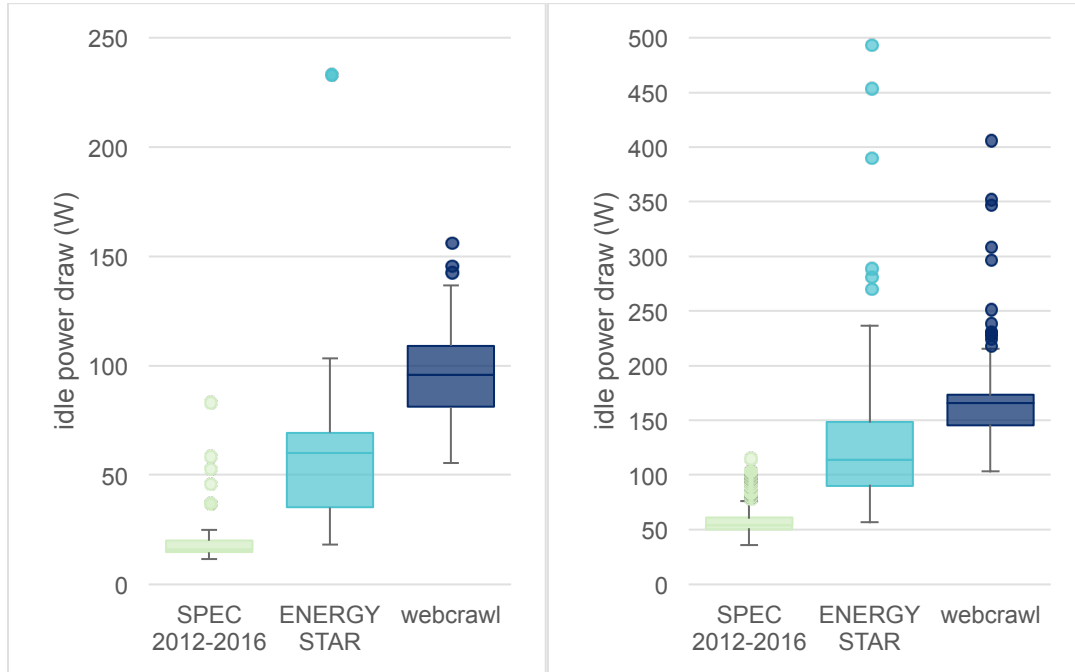


Figure 3.7: Weighted boxplots, idle power draw of 1-processor (L) and 2-processor (R) servers; outliers shown

Table 3.3 summarizes certain attributes of servers in the datasets displayed in Figure 3.7, splitting outlying data above the 75th percentile from non-outliers. Displayed server characteristics include number of cores, installed memory, drive type, number of drives, storage capacity, and number of PSUs. These were chosen for analysis because they were able to be extracted from at least two of the three rather heterogeneous datasets with a reasonable level of confidence.¹⁶ We supply median and average values for all, excepting the categorical value of drive type, which is split between HDD and SSD drives. Bold values represent the parameter for which there was the largest ratio between outlying and non-outlying values within each dataset, one way of more systematically considering what specific features of these servers entail extreme power draws.

¹⁵ Of course, outlying data points can also be due to errors in data collection, recording, or entry.

¹⁶ Graphics processing unit (GPU) acceleration holds promise for more energy-efficient high-performance computing by allowing GPUs' hardware resources to be used for non-graphical compute-intensive applications on thousands of cores instead of several CPU cores. These "general purpose GPUs" (GPGPUs) can be either discrete cards or be integrated into the CPU or motherboard (NVIDIA 2016). The only dataset with sufficient information on the type of graphics card present was scraped from cdw.com, and all records with graphics type specified showed integrated graphics chips. We assume that using GPGPUs in servers purchased through retail outlets is fairly rare as of mid-2016, that they predominantly occur in custom builds instead of volume servers, and that consumers buying a volume server would only add these units post-purchase. GPGPUs also require custom code and programming to leverage. A small percentage of CDW records did mention support for GPU acceleration, but we think it unlikely that outlying power draw values here can be attributed to GPGPUs.

More importantly, filtering out the relatively small number of outliers reveals key differences between SPEC, ENERGY STAR, and web-crawled datasets when it comes to common server characteristics. In general, we assume the ENERGY STAR attributes most closely match those of volume servers in the field, given that manufacturers were required to report a typical or single configuration that “is representative of a deployed product with high volume sales” (ENERGY STAR 2013).¹⁷ We also expect that online retail results have less memory capacity and fewer installed PSUs than do typical servers in use, since both components are typically installed after-market to meet anticipated computing needs.

From the non-outlier (unshaded) rows in Table 3.3, we see that servers are roughly similar across datasets when it comes to number of cores and processor speed. However, SPEC 2-processor servers have considerably less installed memory than both the ENERGY STAR and web-crawled servers, with both the median and average values not more than half the values for the other datasets. In addition, SPEC servers have at least two and four times fewer the number of installed drives as do ENERGY STAR servers with one and two processors, respectively; in addition, solid-state drives dominate among SPEC servers but are rarely seen (3%) among ENERGY STAR servers. Correspondingly, the storage capacity among SPEC servers is 25 to 47 times smaller than we see among online retail data, which rest upon the assumptions documented in section 2.4. Moreover, all SPEC servers have only one power supply unit, while the ENERGY STAR data—and, to a lesser extent, online retail servers—show a median installed value of two PSUs, and average values supporting the supposition that most deployed volume servers have two PSUs for redundancy purposes. These comparisons demonstrate that SPEC systems were chosen chiefly to optimize performance and efficiency, and that these manufacturer-submitted results are not characteristic of the market for volume servers. As such, most volume servers today do not achieve the low idle power draw of SPEC servers, nor are they likely to scale power with utilization to the same extent.

¹⁷ While ENERGY STAR certified servers have only 15% market share in 2015 (ENERGY STAR 2016b), we refer here to the hardware attributes of the typical or single configurations.

Table 3.3: Summary of server characteristics within and among datasets, including outliers

Dataset	Outliers?	<i>n</i>	# cores		Processor speed, GHz		Installed memory, GB		Drive type		# drives		Storage capacity, TB		# PSUs	
			Med.	Avg.	Med.	Avg.	Med.	Avg.	% SSD	% HDD	Med.	Avg.	Med.	Avg.	Med.	Avg.
SPEC (2012 on) 1-processor	Y	6	4	5.3	3.2	3.0	12	12.0	50	50	1	1.0	0.37	0.32	1	1.0
	N	16	4	4.0	2.6	2.7	16	12.5	69	31	1	1.0	0.064	0.11	1	1.0
SPEC (2012 on) 2-processor	Y	25	16	18	2.4	2.4	64	71.4	32	68	1	1.4	0.30	0.39	1	1.1
	N	109	16	20	2.2	2.3	24	35.2	81	19	1	1.0	0.16	0.19	1	1.0
ENERGY STAR 1-processor	Y	2	2	2.0	3.3	3.3	256	256	0	100	2	2.0	-	-	2	2.0
	N	32	4	3.6	3.3	3.2	16	18.4	3	97	2	2.4	-	-	2	1.5
ENERGY STAR 2-processor	Y	6	14	14	4.2	3.9	192	181	0	100	3	3.7	-	-	2	2.2
	N	91	16	17	2.4	2.4	64	104	3	97	4	4.0	-	-	2	1.8
Web-crawl 1-processor	Y	3	6	6.0	2.1	2.2	8	48.0	-	-	-	-	5.0	4.7	2	1.7
	N	648	6	7.0	2.4	2.4	8	10.9	-	-	-	-	3.0	2.9	1	1.1
Web-crawl 2-processor	Y	13	16	16	2.4	2.4	144	181	-	-	-	-	4.8	14	2	2.0
	N	253	16	17	2.4	2.5	64	70.0	-	-	-	-	4.8	4.9	2	1.9

3.4 Server Power versus Utilization

In this subsection, we present how the median power at each utilization level changed over time for 1- and 2-processor servers via SPECpowerssj_2008 results, compared to the single median data points for ENERGY STAR and web-scraped data at idle. The results presented in this section are not weighted by server brand, because we were more interested in exploring power scaling potential from the SPEC database with all records given equal weight, in case certain manufacturers with limited American market share have achieved better scaling. Comparison of median idle power values for CDW and ENERGY STAR with Table 3.2 shows that weighted versus unweighted medians differ by a few watts in most cases, with bigger differences for ENERGY STAR than CDW values, likely because the latter dataset contains many more records. In Figure 3.8 and Figure 3.9 we only consider the last five years of SPEC records (2012–2016), under the assumption that they are more directly comparable—in terms of date of hardware availability—with ENERGY STAR and online retail results for idle power draw.

For 1-processor servers, median power draw values post-2012 cluster closely together until approximately 70% power and up. 2-processor servers exhibit tight grouping amongst all years at lower utilization levels, but gradually diverge with increased utilization. Unweighted median power draw values for ENERGY STAR and online retail data are again illustrative of differences between datasets in terms of efficiency. The gap between SPEC and ENERGY STAR at 0% (idle), in particular, may reinforce that SPEC systems that optimize power proportionality are not representative of volume servers. Shehabi *et al.* 2016 determined an average maximum power of 118 W for 1-socket servers and 365W for servers with 2+ sockets, significantly higher than the median maximum power of SPEC servers in the figures below. These manufacturer-submitted servers may maximize energy efficiency at the cost of not being able to run typical computing loads.

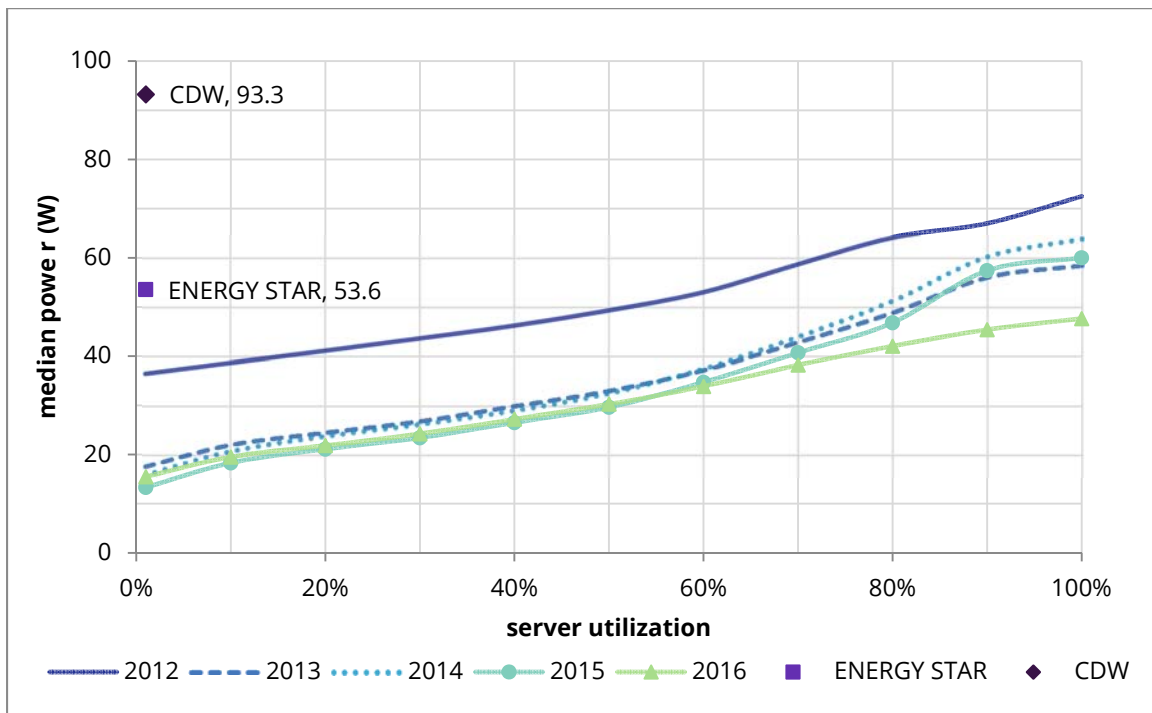


Figure 3.8: Median power vs. utilization for 1-processor SPEC servers from 2012-2016

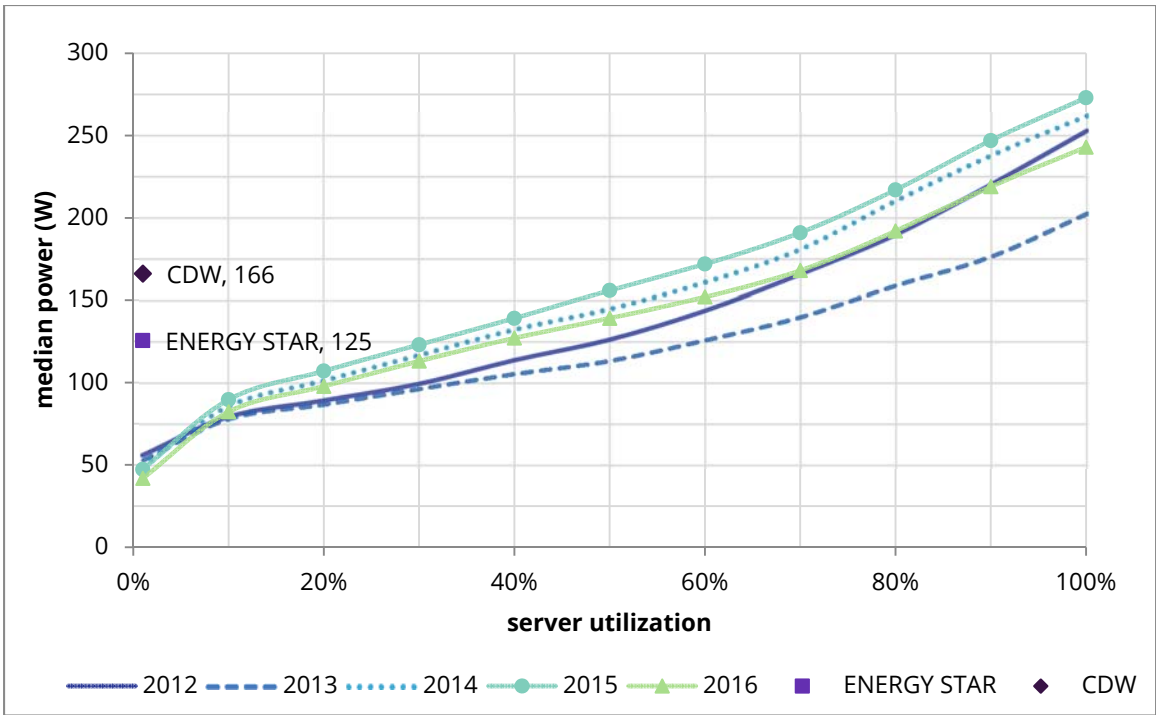


Figure 3.9: Median power vs. utilization for 2-processor SPEC servers from 2012-2016

Next, to visualize how power vs. utilization has changed over a decade of SPEC results, Figure 3.10 shows median values for percentage—not absolute—power yearly from 2007–2016.

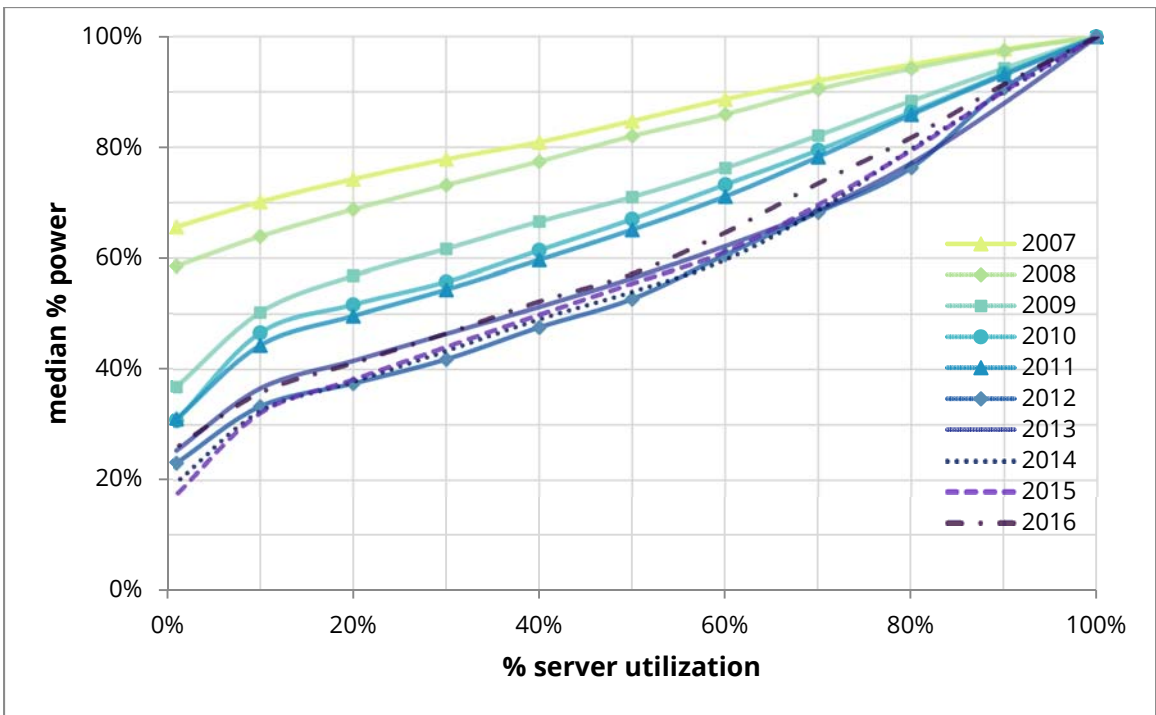


Figure 3.10: Median % power vs. utilization for 1- and 2-processor SPEC servers, by year hardware available

This chart displays the evolution of lower power at low utilization over the past decade. A sharp curve from idle to 10% utilization of 2-processor servers in more recent years is demonstrative of technological progress in decreasing power draws when servers are idle, which may be driven by efforts to certify servers for ENERGY STAR or other programs relying on a test procedure that measures only idle power.

Next, in Figure 3.11 boxplots at several key utilization levels—100%, 20%, and idle—illustrate the scatter in the SPEC results grouped together in two five-year generations, 2007–2011 and 2012–2016. The boxes do not overlap between generations at each utilization level displayed here. We see a greater spread between 25th and 75th percentiles at all three utilization levels for the earlier generation of SPEC servers, while several less-power proportional outliers appear at 20% utilization and idle for the later generation.

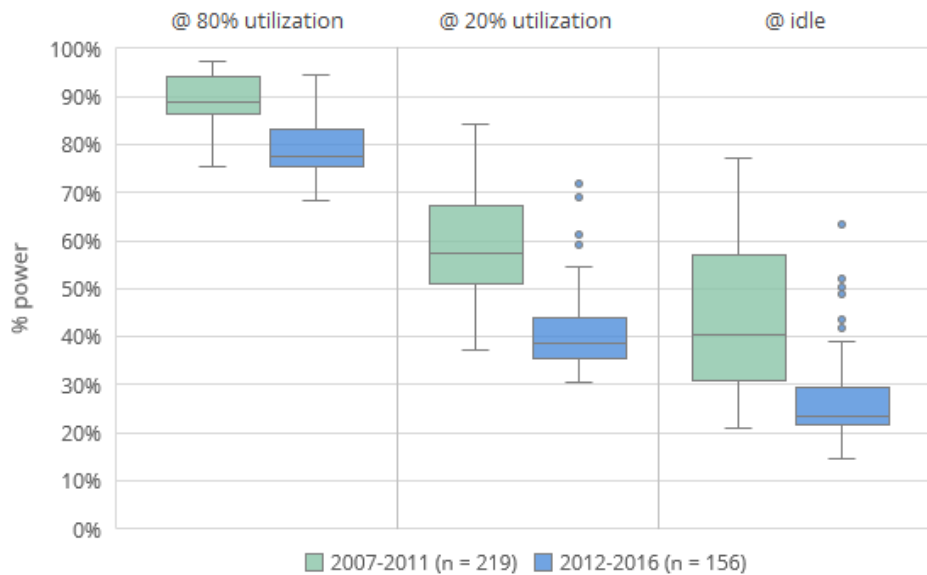


Figure 3.11: Percent power at three utilization levels for all SPEC servers, by generation

Next, Figure 3.12 is a different look at the temporal trend in average idle power as a percent of maximum power (plotted in navy). Given the self-selection bias present for SPEC data, this trend can be seen as representative of the technological potential for achieving lower power levels while servers are idle. The rate of improvement in power scaling over time slows as average power as a percent of maximum power approaches zero. We expect this is because over time there are fewer unnecessary components and functions that can be turned off to improve power proportionality. The data are consistent with a logarithmic function, and echo the shape of the lower bound trend seen in Figure 1.2, whose authors used a different methodology with the same source dataset (Shehabi *et al.* 2016).

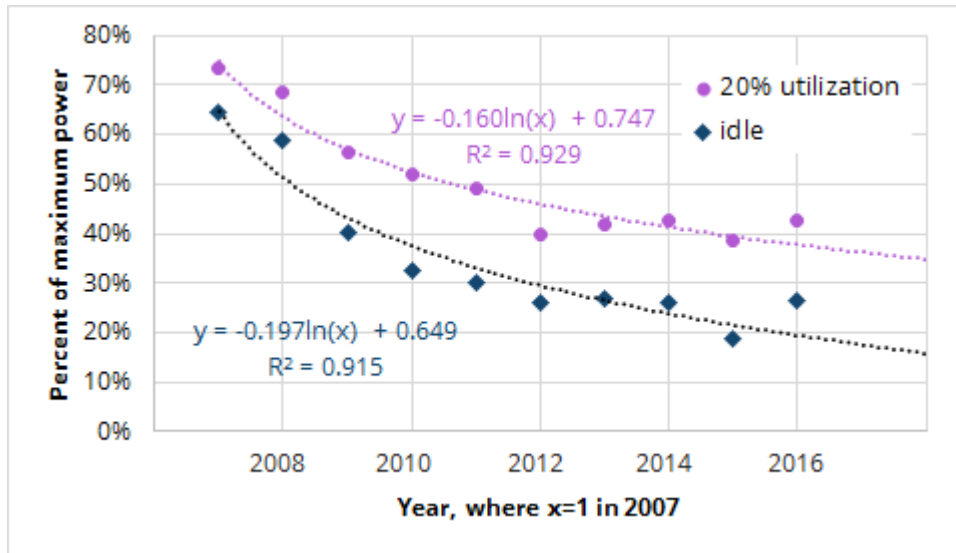


Figure 3.12: Temporal trend in percent of maximum power at 20% utilization and idle for SPEC servers

When considering the trend at 20% power (plotted in purple)—a reasonable approximation of real-world average utilization for most volume servers—we note a smaller improvement in power proportionality via a flatter curve. Thus, simply using the temporal trend for idle power observed in SPEC servers can overstate the technological improvement potential for volume servers’ power scaling.

Finally, we input the weighted average idle power for each of the three datasets seen in Table 3.2 into the model used by Shehabi *et al.* (2016), separately for 1- and 2-processor servers. Their model accounts for varying utilization levels across different data center space types, plus projections of future trends in power scaling, maximum server power, and market shares of 1- and 2-processor servers. The model also includes other data center energy demands beyond servers, such as infrastructure cooling, and relies on a set of assumptions validated by industry experts. Inputting these different server values¹⁸ allows us to compare how close the server power draw we calculated for online retail data—meant to reflect real-world use—would be to Shehabi *et al.*’s results, as well as visualize two alternative scenarios, where the fleet of installed volume servers runs at the average ENERGY STAR and SPEC powers. In Figure 3.13 and Figure 3.14, which display predicted U.S. volume server electricity use and total data center electricity consumption (including non-server infrastructure), respectively, the line marked by LBNL denotes the Shehabi *et al.* values.

¹⁸ Shehabi *et al.*’s 2016 report modeled volume servers run at different utilization levels depending on data center space type, but the hardware of volume servers—with respect to energy use—was assumed to stay the same. Therefore, the values mentioned here affect all volume servers modeled in Shehabi *et al.*, regardless of space type.

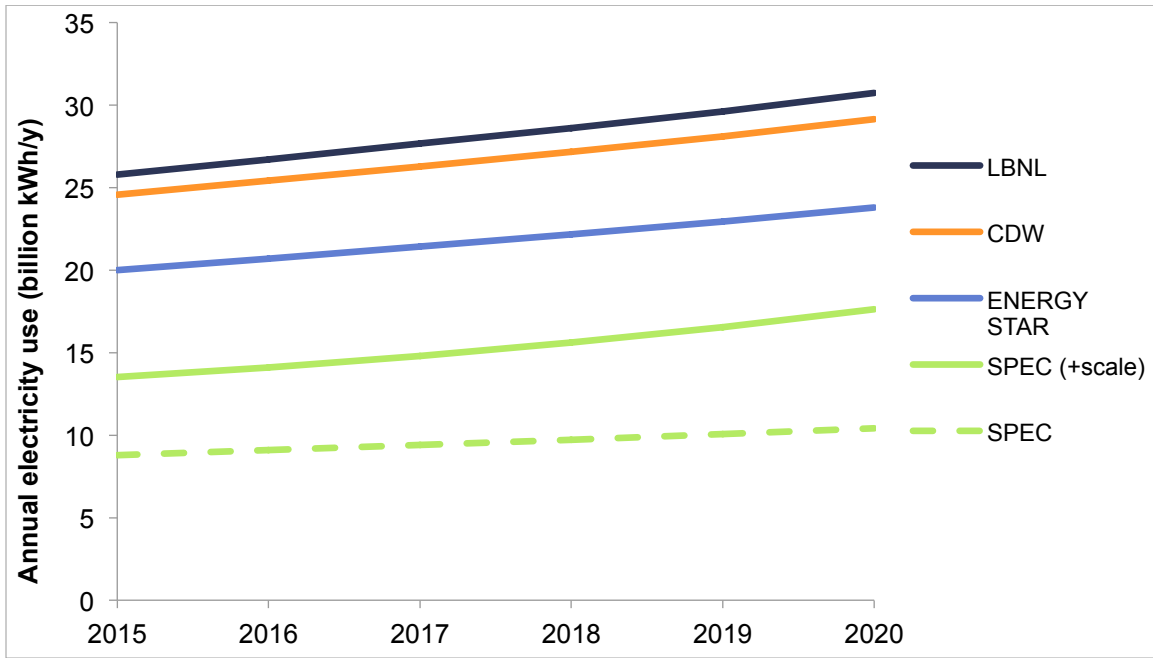


Figure 3.13: U.S. volume server electricity use 2015–2020, with varying inputs to Shehabi *et al.* (2016) model

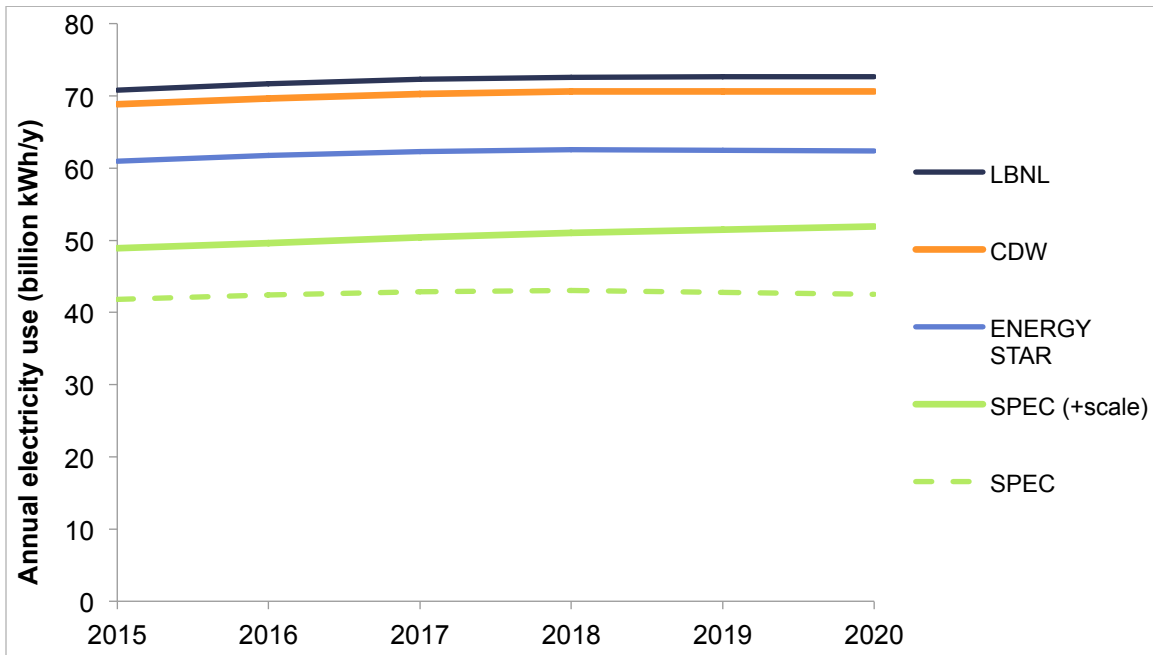


Figure 3.14: U.S. data center electricity use 2015–2020, with varying inputs to Shehabi *et al.* (2016) model

These figures show that the calculated weighted average idle power assumptions from idle retail data are very close to the assumptions in the original model, with the higher LBNL values possibly due to the inclusion of self-assembled, direct-ODM, and other non-retail servers that may be configured to be more powerful than typical off-the-shelf servers, and thus draw more power. These charts also demonstrate the potential for considerable energy savings with wider uptake of ENERGY STAR-qualified servers.

3.5 Procurement and Management Practices for Servers in Server Rooms and Closets

This section reports descriptive statistics from surveys of IT managers ($n=216$ completed surveys) and vendors ($n=178$) conducted in March 2015 and characterized in section 2.7 of this paper. All results in this section refer to procurement and management practices specific to servers in server rooms and closets (<1000 ft²)—a space type estimated to contain at least half of the servers in the United States (Bard *et al.* 2014). These findings must also be viewed in light of a sampling uncertainty of $\sim\pm 7$ percent.¹⁹ All values also were calculated inclusive of the answer choices “don’t know” and “not familiar”, even where not made explicit in the figure or table. IT managers may have a more accurate sense or deeper awareness of server management practices within their own organizations, but IT vendors are exposed to a greater range of client companies. We present comparable or complementary results for both surveys when possible, because “don’t know” and “not familiar” response rates are generally higher for the IT manager survey than the IT vendor one. IT managers may not be aware of procurement practices due to the structure of their companies. This falls into the realm of the split-incentive problem, where those departments making decisions about how servers are run are neither concerned with nor responsible for paying for the energy implications of those choices; NRDC estimated in 2014 that only one-fifth of organizations’ IT departments pay the data center power bill (NRDC 2014). In some cases, the “not familiar” response rate increases with certification stringency, so perhaps some IT managers are less educated with respect to new developments concerning energy efficiency than are IT vendors. Generally, these results give insight into procurement and management practices at a sample of U.S. organizations with different locations, industry classifications, and sizes.

First, we queried respondents about required retirement and average ages of servers, server PSUs, and server hard drives. Table 3.4 shows the distribution of IT managers’ responses for required retirement and average age of the equipment in server rooms and closets.

Table 3.4: Required retirement and average age of server equipment in IT managers’ server rooms & closets

	Required retirement age			Average age		
	Servers	PSUs	Hard drives	Servers	PSUs	Hard drives
1 year	2%	4%	7%	4%	9%	14%
2 years	8%	12%	13%	25%	23%	22%
3 years	17%	20%	14%	31%	31%	20%
4 years	14%	9%	9%	15%	11%	12%
5 years	23%	20%	16%	11%	11%	11%
5+ years	13%	8%	12%	5%	4%	6%
No requirement	18%	21%	23%	-	-	-
Don’t know	5%	6%	5%	9%	12%	14%
Average	4.2 years	3.8 years	3.7 years	3.2 years	3.0 years	3.0 years

¹⁹ Assuming a 95% confidence level; the Bureau of Labor Statistics reports 348,500 computer and information systems manager jobs nationwide in 2014 at <https://www.bls.gov/ooh/management/computer-and-information-systems-managers.htm>.

We observe that the most commonly required retirement age for all considered equipment is five years or no requirement, while the most common average age is three years (excepting hard drives, where two years predominates slightly more than three years as average age). The last row displays average required retirement age and average equipment age weighted by the number of responses, calculated excluding the “no requirement” and “don’t know” responses and assuming that “5+ years” is best approximated by six years. IT vendors were similarly surveyed about required retirement/refresh age of servers, as well as the average age for servers and hard drives, in their customers’ server rooms and closets. Using the same methodology as for IT managers, the weighted average required retirement age of IT vendors’ customers’ servers is 3.9 years, and the average age of servers and hard drives 3.6 years. As a whole, these results may somewhat understate equipment lifetime, especially for required retirement age, since a significant share of IT manager respondents indicate that their organizations have no requirement, and one quarter to one third of servers in data centers are thought to be older devices overlooked by managers that are still drawing power while providing little or no use (“comatose” servers or “zombies”) (Koomey and Taylor 2017). When contrasted to IT manager and IT vendor results, lifetime values derived from the literature (in section 2.2) bolster this assertion of underestimation, with minimum average lifetime of 3 years, median 5 years, mean 5.5 years, and maximum of 10 years.

Second, we asked IT managers whether their organization currently requires ENERGY STAR and Electronic Product Environmental Assessment Tool (EPEAT)²⁰ certification for servers specifically located in server rooms and closets. We see from Table 3.5 that in 2015, more than four in ten IT managers’ employers required ENERGY STAR version 2.0, released in 2013, compared to only 30% for version 1.0 (2009), and that the former, newer specification will be required by about half of organizations by 2018. ENERGY STAR looks to be more widespread than EPEAT, which is a more stringent specification: all EPEAT servers are also ENERGY STAR certified, but not vice versa. However, we also see that EPEAT Gold servers will be required more commonly within the next three years than will the other less-stringent EPEAT certifications.

Table 3.5: ENERGY STAR and EPEAT certification requirements in IT managers’ server rooms & closets

Certification	Currently required			Required within next 3 years		
	Yes	No	Unknown/ unfamiliar	Yes	No	Unknown
ESTAR v1.0 (2009)	30%	43%	27%	23%	40%	37%
ESTAR v2.0 (2013)	44%	31%	25%	49%	18%	33%
EPEAT Bronze	14%	50%	36%	17%	39%	44%
EPEAT Silver	23%	43%	34%	25%	32%	43%
EPEAT Gold	23%	43%	35%	31%	25%	45%

²⁰ The non-profit Green Electronics Council manages EPEAT certification, which is based on a group of environmental performance criteria. Federal regulations generally mandate that 95% of federal agency electronics purchases must meet EPEAT criteria if such standards exist for each product type purchased (Federal Acquisition Regulation Subpart 23.704). EPEAT certification did not appear in the product specifications of online retail data presented in section 3.1.

Turning to IT vendors, they report the fraction of their clients that will require ENERGY STAR version 1.0 within three years is 41% in the “more than now” and “almost all” response categories together, 40% in the “about the same portion” category, and 12% in the “fewer than now” and “almost none” bins. For version 2.0, these numbers shift to 46%, 42%, and 8%, respectively. The remainder (4–7%) vendor responses were “don’t know” or “not familiar”. For all EPEAT levels, 37–40% of their clients fall within the “more than now” and “almost all” categories, 40–42% in the “about the same” bin, and 12–15% in the “fewer than now” and “almost none” categories, with 6–7% of IT vendors “not familiar”.

We assume these requirements for certification refer to procurement practices, so Figure 3.15 below displays to what extent currently installed servers are ENERGY STAR- and EPEAT-certified. IT managers indicated what portion of their organizations’ servers in on-site server rooms and closets met these certifications, whereas IT vendors specified the average percentage of ENERGY STAR and EPEAT-certified servers in customers’ server rooms and closets; individual vendor responses were averaged for each percentage bin. The purple bars represent the portion of respondents for which 0% of servers are certified; light green bars represent 1–50% certification; dark green bars represent 51–100% certification; and gray bars are the portion of respondents who do not know how many of their servers are certified, or are unfamiliar with the certification (classified together as “unknown”).

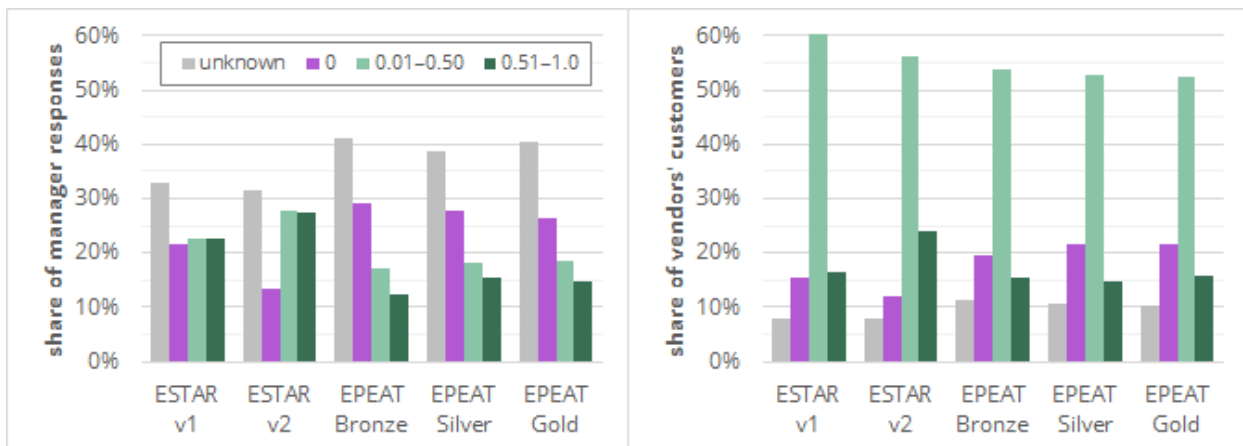


Figure 3.15: (L) shows ENERGY STAR and EPEAT certification in IT managers’ servers; (R) shows the same for IT vendors’ customers’ servers

Respondents were asked to indicate what portion of their servers met these certifications in decile ranges as well as discrete values at 0%, 50%, and 100% (i.e., 0%, 1–10%, 11–20%,...,91–99%, 100%). For another look at the data shown above, Table 3.6 presents median ranges and averages for these grouped data, calculated excluding unknown responses and, for the average, using the midpoint of each decile range.

Table 3.6: Median range and average proportion of ENERGY STAR and EPEAT certification

Certification	IT managers		IT vendors	
	Average	Median	Average	Median
ESTAR v1.0 (2009)	39%	31–40%	32%	21–30%
ESTAR v2.0 (2013)	45%	41–49%	38%	31–40%
EPEAT Bronze	24%	1–10%	28%	21–30%
EPEAT Silver	29%	11–20%	27%	21–30%
EPEAT Gold	29%	1–10%	28%	21–30%

Several insights emerge from these data. The penetration of ENERGY STAR-certified servers, particularly the newest specification, is higher than that of EPEAT qualification, which aligns with the prevalence of requirements in Table 3.5. While the share of managers’ organizations with more than half of their servers qualified under each certification (13–27%, depending on certification) is roughly similar to the share among vendors’ customers (15–24%), there is a striking difference in the proportion of “don’t know” and “not familiar” responses among managers (31–41%) in contrast to vendors (8–11%). At the same time, the shares of organizations where half or fewer servers are certified is far higher among IT vendors’ customers (68–76%) than among IT managers’ firms (41–46%). Consequently, we assume that at most of the organizations whose IT managers are unaware of the prevalence of these environmental certifications, half or fewer of the servers are certified—in accordance with the estimated 15% marketplace penetration of ENERGY STAR servers in 2015 (ENERGY STAR 2016b).

Next, we can similarly examine the penetration of 80 PLUS power supply certification, also relevant to server energy efficiency. Questions about PSUs were only asked of IT managers. Figure 3.16 is a slopegraph in which the darker the line is, the higher the efficiency of the power supply. 80 PLUS Silver and Gold certifications are the most common at the time of the survey and three years in the future, possibly because 80 PLUS Silver is required for ENERGY STAR certification (ENERGY STAR 2013). However, the three most efficient certifications—Gold, Platinum, and Titanium—exhibit the fastest forecasted uptake.

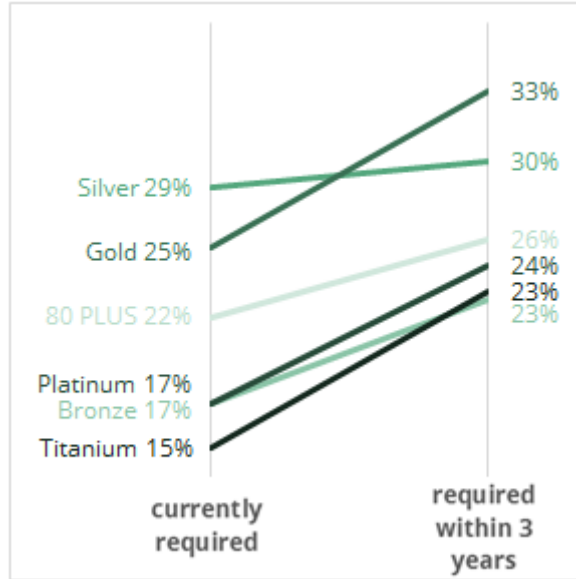


Figure 3.16: Share of IT managers' organizations requiring 80 PLUS PSU certification in server rooms/closets

Given the large number of “don’t know” and “not familiar” responses that is also quite variable, we include Table 3.7; here we observe that 80 PLUS Silver and Gold certifications are the only ones that IT managers' organizations will require within three years as often or more often than not.

Table 3.7: 80 PLUS PSU certification requirements in IT managers' server rooms/closets

80 PLUS Certification	Efficiency @ 50% load	Currently required			Required within next 3 years		
		Yes	No	Unknown/unfamiliar	Yes	No	Unknown
80 PLUS	0.80	22%	47%	31%	26%	37%	38%
Bronze	0.85	17%	51%	32%	23%	37%	41%
Silver	0.88	29%	39%	32%	31%	29%	40%
Gold	0.90	25%	41%	33%	33%	26%	41%
Platinum	0.92	17%	48%	35%	24%	33%	43%
Titanium	0.94	15%	48%	37%	24%	33%	44%

These values presumably represent purchasing requirements, so Figure 3.17 illustrates the penetration of 80 PLUS certification among installed PSUs.

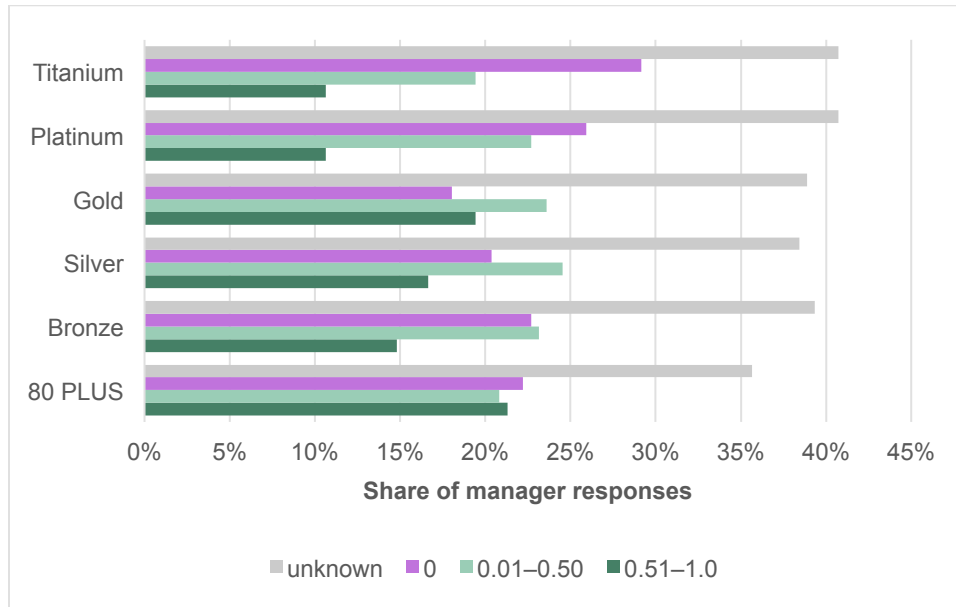


Figure 3.17: Proportion of 80 PLUS PSU certification in servers in IT managers' server rooms/closets

Respondents were asked to indicate what portion of their servers met these certifications in decile ranges, as well as discrete values at 0%, 50%, and 100% (i.e., 0%, 1-10%, 11-20%,...,91-99%, 100%). For another look at the data shown above, Table 3.8 presents median ranges and weighted averages for these grouped data, calculated excluding “unknown” responses and, for the average, using the midpoint of each decile range.

Table 3.8: Average and median range proportion of 80 PLUS PSU certification of IT managers' servers

	80 PLUS	Bronze	Silver	Gold	Platinum	Titanium
Average	39%	32%	35%	37%	26%	24%
Median	21-30%	11-20%	21-30%	41-49%	1-10%	1-10%

Similarly to Table 3.7, these data reveal a significant lack of knowledge about or familiarity with 80 PLUS certifications and whether the servers they run are 80 PLUS certified. In addition, 80 PLUS standard and Gold certifications are the most prevalent among the installed base of servers, followed by Silver and Bronze certifications. Titanium and Platinum certifications are rare.

Next, we switch our focus to operational practices. From the IT manager survey, Table 3.9 presents monitoring requirements and measurement frequency for the following server performance metrics: CPU utilization, memory utilization, disk utilization, and power consumption. IT managers indicated whether these performance metrics were currently required, will be required within the next three years, how often these metrics are measured, and how often these metrics are reviewed and evaluated. The least-monitored server performance metric is disk utilization, while all performance metrics are most commonly measured and reviewed on a monthly basis.

Table 3.9: Server performance metrics and recurring practices according to IT managers

	Performance metrics				Recurring practices	
	CPU utilization	Memory utilization	Disk utilization	Power use	Server inventory	Application mapping
	Currently required					
Yes	75%	76%	65%	70%	71%	52%
No	14%	15%	24%	18%	19%	32%
Unknown	10%	9%	12%	12%	10%	16%
	Required within next 3 years					
Yes	73%	72%	66%	67%	73%	57%
No	13%	15%	19%	15%	14%	21%
Unknown	13%	13%	15%	18%	14%	23%
	Measurement frequency				Practice frequency	
At least daily	15%	14%	14%	10%	6%	4%
Weekly	17%	19%	21%	20%	12%	15%
Monthly	38%	38%	32%	33%	36%	28%
Annually	17%	15%	15%	17%	28%	19%
Never	6%	6%	9%	11%	6%	13%
Unknown	7%	7%	8%	9%	12%	21%
	Review frequency					
At least daily	6%	7%	6%	6%		
Weekly	20%	20%	20%	18%		
Monthly	36%	38%	36%	33%		
Annually	22%	21%	20%	24%		
Never	6%	6%	9%	11%		
Unknown	9%	8%	9%	9%		

On the right side of the table above, IT managers reported on two recurring practices: server inventory/capacity planning (to track and review which servers have high or low utilization), and application mapping (to track and review which server applications have high or low utilization). Server inventory happens more widely than application mapping, and both are most typically practiced on a monthly basis. The results in Table 3.9 indicate that a large portion of IT managers are required to review server inventory on a monthly or annual basis. However, the efficacy of these required review practices in maximizing equipment use, removing unnecessary server loads, and avoiding unnecessary build-out is not clear from the survey alone. What exactly these practices comprise can vary extensively, and IT managers likely conduct these reviews for a range of reasons. One opportunity for improvement relates to the nearly 20% of respondents having no recurring evaluation practices (and an additional 10% not knowing if such practices are required). With nearly the same portion of IT managers with current requirements and those expected in the next three years, this opportunity may not be sufficiently addressed by industry.

Finally, we turn to the suite of operational practices that enhance energy efficiency outlined in Table 3.10, surveyed among both IT managers and vendors. Unfortunately, a rigorous assessment of which practices are most significant with respect to energy conservation is difficult,

because for most of the practices surveyed, energy savings depend on to what extent each measure is implemented.

Table 3.10: Operational practices that decrease power consumption of servers

Operational practice (A-Z)	Definition/example
Application consolidation	To bundle multiple functions or servers into one application (<i>e.g.</i> , consolidating database applications from different departments into one database)
Power capping	To keep each server within a pre-allocated power draw limit
Power cycling	To enable servers to automatically sleep after a period of being idle
Remote wake/power on	<i>E.g.</i> , Wake-on-LAN (WoL)
Running only as needed	<i>E.g.</i> , shutting servers down at night and on the weekends
Server consolidation	To bundle multiple applications into one server to reduce number of servers
Server decommissioning	To switch off and store servers that are no longer used
Server virtualization	To partition a physical server into multiple virtual instances
Thin provisioning	To make use of virtualization technology to allow storage resources to be easily allotted on a just-enough and just-in-time basis
Using solid-state storage	Enable ultra-high performance input/output operations per second, low latency, and lower power consumption and operating temperatures

Table 3.11 displays whether these operational practices are required for servers in IT managers' server rooms and closets currently (as of 2015) and three years on. Currently, server virtualization, consolidation, decommissioning, power cycling, and remote wake/power on are required practices in more than half of respondents' organizations, while running only as-needed, power capping, and thin provisioning are least commonly found among survey respondents. We highlight the significant proportion of "unknown/unfamiliar" responses—13 to 25% for currently required practices, and 21–30% three years from now. Organizations with IT managers unaware of or unfamiliar with approaches they are responsible for implementing likely do and will not require application of these practices.

Table 3.11: Operational practices requirements in IT managers' server rooms/closets

Operational practice	Currently required			Required within next 3 years		
	Yes	No	Unknown/unfamiliar	Yes	No	Unknown
Server virtualization	59%	28%	13%	60%	19%	21%
Server consolidation	56%	28%	17%	62%	18%	21%
Server decommissioning	55%	27%	18%	56%	22%	23%
Power cycling	54%	27%	19%	54%	21%	25%
Remote wake/power on	53%	30%	18%	51%	27%	22%
Application consolidation	46%	37%	17%	53%	25%	23%
Using solid-state storage	43%	38%	19%	44%	28%	28%
Running only as needed	37%	46%	18%	44%	31%	25%
Power capping	34%	43%	23%	40%	31%	29%
Thin provisioning	32%	43%	25%	38%	32%	30%

In comparison, IT vendor responses suggest that taken as a whole, these operational practices will be implemented to a greater extent within the next three years than they are currently. IT vendors were queried regarding the portion of their clients that will require each practice within the next three years, with possible response choices of “almost all”, “more than now”, “about the same portion”, “fewer than now”, “almost none”, and “don’t know.” The bin of “almost all” and “more than now” holds 44% of responses, on average, across all practices, while that of “about the same portion” is another 44%, that for “fewer than now” and “almost none” is 9%, and “don’t know” was 4%. The practice with the highest response rate for “almost all” and “more than now” was server virtualization (48%), while that with the lowest was application consolidation (38%).

To assess whether some share of respondents work at organizations that require several of these power-saving operational practices, we present a matrix of Spearman’s rank correlation coefficients (ρ or rho) for each pair of variables in Table 3.12. We assigned “yes” responses a value of 1 and “no” responses a value of 0; “don’t know” and “not familiar” were coded as NaN (not a number) and excluded from the calculation. Server practices are ordered according to the share of IT managers’ organizations currently requiring these practices (the first numerical column in Table 3.11) to facilitate comparison between these tables. In addition, rho values are color-mapped to reflect the degree of correlation. The tightest correlation exists between server consolidation and application consolidation, possibly owing to similar consolidation approaches available for both practices, followed by correlations among the less prevalent practices—thin provisioning

and power capping, then running only as needed and power capping—which may be implemented by the same, more knowledgeable IT managers.

Table 3.12: Spearman’s rank correlation coefficients (ρ) of practices in IT managers’ server rooms/closets

	server virtualization	server consolidation	server decommissioning	power cycling	remote wake / power on	application consolidation	solid-state storage	running as needed	power capping	thin provisioning
server virtualization	1	-	-	-	-	-	-	-	-	-
server consolidation	0.44	1	-	-	-	-	-	-	-	-
server decommissioning	0.46	0.49	1	-	-	-	-	-	-	-
power cycling	0.22	0.37	0.31	1	-	-	-	-	-	-
remote wake/ power on	0.22	0.37	0.36	0.47	1	-	-	-	-	-
application consolidation	0.39	0.64	0.47	0.23	0.43	1	-	-	-	-
solid-state storage	0.37	0.39	0.44	0.32	0.38	0.52	1	-	-	-
running as needed	0.24	0.34	0.32	0.41	0.33	0.44	0.4	1	-	-
power capping	0.24	0.4	0.35	0.36	0.43	0.45	0.5	0.58	1	-
thin provisioning	0.35	0.36	0.48	0.26	0.29	0.45	0.51	0.47	0.59	1

Finally, although these operational practices are not as closely linked to the procurement process as the ENERGY STAR and 80 PLUS requirements explored earlier, it is likely that they are implemented in an incremental fashion, as certain servers reach the end of their lifetime and managers re-configure how they operate their remaining or replacement fleet. We asked IT managers to indicate the portion of their servers in on-site server rooms and closets for which each practice was implemented in the past year, while IT vendors were similarly queried for the installed base prevalence of these practices for customers that they service. Figure 3.18 and Figure 3.19 display the extent to which these practices are implemented in the installed base of servers for IT managers’ organizations and IT vendors’ customers, respectively.

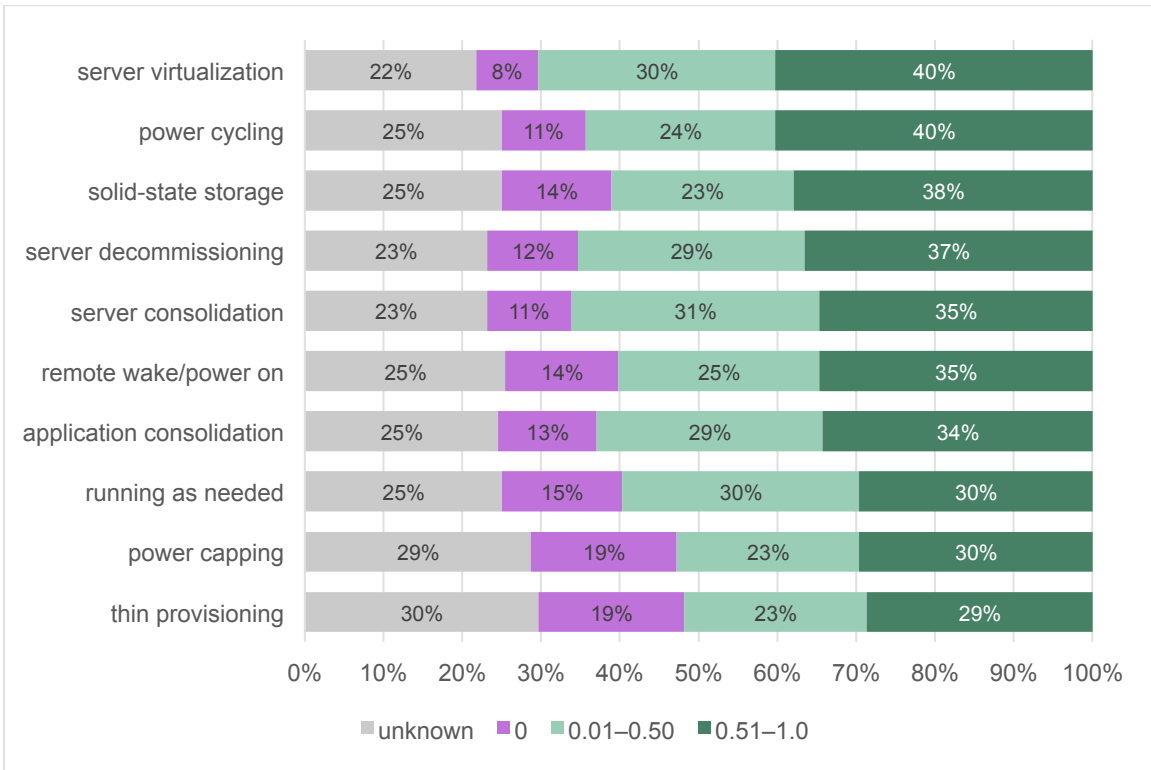


Figure 3.18: Prevalence of certain operational practices in IT managers' server rooms and closets

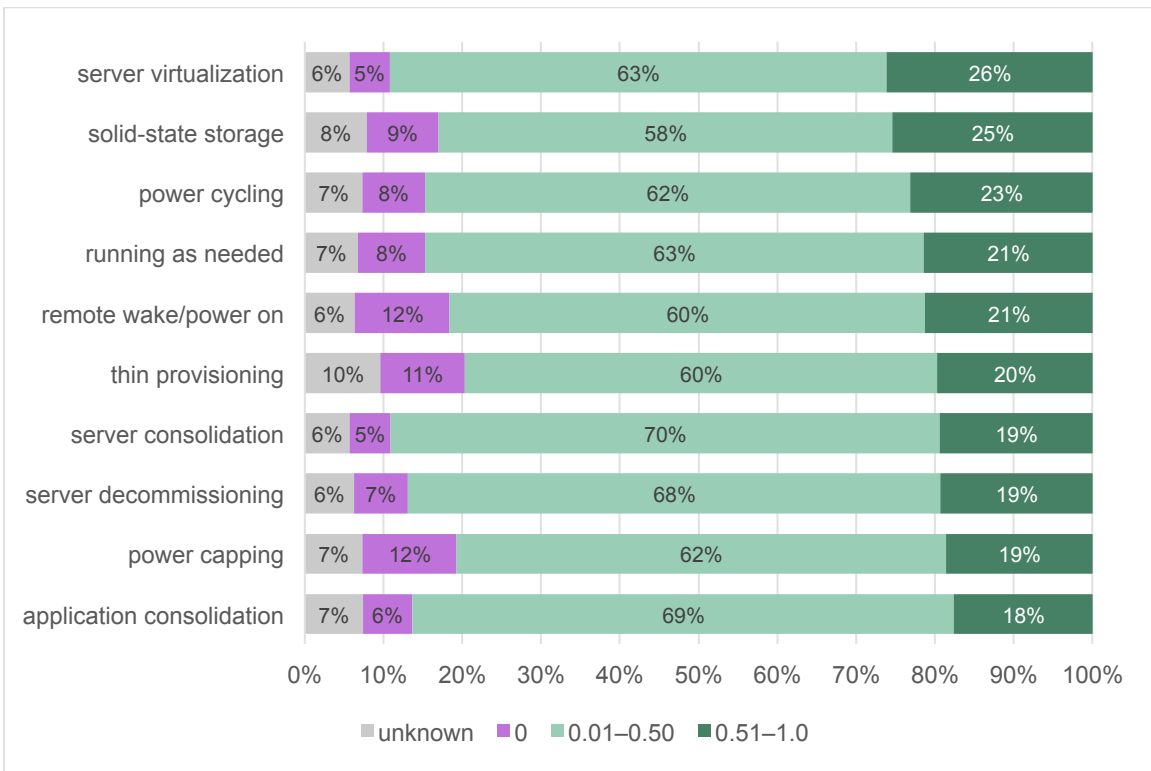


Figure 3.19: Prevalence of certain operational practices in IT vendors' customers' server rooms and closets

Each chart is organized in descending order of 50% or greater prevalence. We first note the greater share of IT managers who don't know or are unfamiliar (aggregated into "unknown") with these practices, relative to IT vendors. It is unlikely that these respondents carry out these procedures to any significant extent. Second, according to these figures, the three most common practices across IT manager and vendors are server virtualization, using solid-state storage, and power cycling, with power capping one of the least frequent methods used to limit server energy use. Only one-fifth to one-quarter of IT vendors see these energy-saving practices employed on more than half of their customers' servers.

Because respondents chose answers among decile ranges and discrete values at 0%, 50%, and 100% (i.e., 0%, 1–10%, 11–20%, ..., 91–99%, 100%), we show another view of the same data via weighted averages and median ranges in Table 3.13. These values were determined via excluding unknown responses and using the midpoint of each decile range to calculate averages.

Table 3.13: Average and median prevalence of operational practices in server rooms and closets

Operational practice	IT managers		IT vendors	
	Average	Median	Average	Median
Server virtualization	55%	51–60%	40%	31–40%
Power cycling	55%	51–60%	36%	31–40%
Server consolidation	52%	50%	36%	21–30%
Application consolidation	51%	50%	35%	31–40%
Using solid-state storage	50%	51–60%	37%	31–40%
Server decommissioning	49%	50%	35%	21–30%
Remote wake/power on	49%	50%	34%	21–30%
Running only as needed	45%	41–49%	35%	21–30%
Power capping	44%	41–49%	33%	21–30%
Thin provisioning	44%	50%	34%	21–30%

With “unknowns” excluded, the three most widespread practices among IT managers are server virtualization, power cycling, and server consolidation (according to average values) or, considering median ranges, server virtualization, power cycling, and using solid-stage storage. For IT vendors' customers, server virtualization, using solid-stage storage, server consolidation, and power cycling also predominate. Power capping and thin provisioning fall among the three least common in both groups. However, given the large share of “unknowns” among IT managers that were excluded from the analysis—which probably means insignificant uptake of these methods—their results in Table 3.13 likely overstate how common these energy-saving measures are. IT vendor results thus probably better reflect operational measures in the field.

The results presented in this section from the two surveys are fairly discrete in nature, but shed light on various aspects of how servers are operated in server rooms and closets, from procurement requirements and prevalence of energy-efficiency certifications like ENERGY STAR and 80 PLUS to operational practices like virtualization and consolidation. They also demonstrate a lack of familiarity with certain technologies or practices among a significant portion of surveyed

IT managers compared to IT vendors. This is perhaps illustrative of split-incentive problems arising from company structures, and/or from IT vendors' greater familiarity with the range of technological solutions and services they sell to customer companies. IT vendors typically sell to customers with a range of sophistication, so they need to be familiar with a greater range of practices and technologies than the average IT manager does. Regardless, promoting awareness of these practices promises improved energy efficiency in small data centers.

4. Conclusions

This paper characterizes volume servers by compiling and comparing available data on server hardware features and electricity use. Various literature sources provide an incomplete understanding of typical server operation and component configurations, and the impact of these characteristics on electricity demand. We observe a prevalence of 2-socket servers that are often sold with a single processor, indicating that a second processor may or may not be added after market. For both single- and double-processor servers, component-based power use testing and analysis by Navigant Consulting indicates an idle server power demand significantly higher than ENERGY STAR benchmarks and the industry-released energy use documented in SPEC. A closer look at the server configurations that are used to establish power-scaling trends in SPEC shows a deviation from typical servers, indicating that the power-scaling trends themselves may be atypical as well. A comparison of SPEC-reported power-scaling to the power scaling assumption in literature (*e.g.*, Shehabi *et al.* 2016) further support the assertion that the U.S. operating stock of servers does not scale power to utilization at levels observed in SPEC. While the SPEC and ENERGY STAR databases are explicitly designed not to represent all servers, the lack of server information outside of these two data sets can inadvertently cause them to be overrepresented in data center energy use estimates. This report places the energy use of SPEC and ENERGY STAR servers in the context of other available data on servers and quantifies the difference in power demand between servers in those two higher-performing data sets and our best understanding of a typical server operating in the U.S. Finally, surveys of IT managers and vendors help establish the prevalence of more efficient IT equipment and operation practices in server rooms and closets, which highlights opportunities to improve the energy efficiency of the U.S. server stock.

While these results provide insight into the characteristics of servers in data centers, it is important to note that our understanding of typical server power use is restricted by the limitations in the data and methods employed in this report. Server configurations, which drive energy use, are based on the prevalence of different configurations available in online retail data. Without access to sales data for these different configurations, servers available for sale online but rarely purchased will inevitably be overrepresented in our estimates. Our use of online retail configurations also limits our understanding of servers operating in smaller data centers, since very large enterprise or hyperscale data centers are likely to purchase servers directly from manufacturers rather than through consumer channels like the CDW website. Our understanding of procurement and management practices in data centers from the IT manager and IT vendor surveys also excludes larger data centers. Future research could help ameliorate these limitations by collecting sales data for server configurations that includes business-to-business purchases between server manufacturers and large data center entities.

Ultimately, the increased digitization and interconnectivity of information has resulted in the growth of data centers upon which economic activity increasingly depends. Computer servers are at the heart of the services provided by data centers, as well the key driver in their energy consumption. Significant energy efficiency gains have been made in data centers as the energy demand of these buildings has become better documented and publicized, but the lack of empirical data on server electricity use and the proprietary nature of server operation can limit the efficiency opportunities available to data center IT equipment. This report reviews available

literature to outline the server components and trends that influence electricity demand. While our understanding of data center servers is limited, we have attempted to establish what is known and where additional information is needed. Server components and operations will continue to change as core technologies evolve and new services emerge that require novel forms of server computation. Maintaining an understanding of how these changes affect server electricity demand will be necessary to insure that energy efficiency continues to remain a relevant factor as IT equipment evolves and integrates throughout society.

References

- Bard, A., R. Huang, and R. Friedmann. 2014. "From Our Closet to Yours: Fashioning Energy Efficiency Programs for Small Data Centers." ACEEE Summer Study on Energy Efficiency in Buildings.
- Barroso, L.A., J. Clidaras, and U. Hölzle. 2013. *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines*. Second Edition. Synthesis Lectures on Computer Architecture, Mark D. Hill, Editor. Morgan & Claypool Publishers.
<http://www.morganclaypool.com/doi/pdf/10.2200/S00516ED2V01Y201306CAC024>.
- Bio by Deloitte and Fraunhofer IZM. 2014. *Preparatory Study for Implementing Measures of the Ecodesign Directive 2009/125/EC, DG ENTR Lot 9 – Enterprise Servers and Data Equipment, Task 2: Markets*. Draft Final Report prepared for European Commission.
- Bothner, M.S. 2003. "Competition and Social Influence: The Diffusion of the Sixth-Generation Processor in the Global Computer Industry." *American Journal of Sociology* 108(6): 1175–1210.
<http://www.jstor.org/stable/10.1086/375200>.
- CDW. 2016. "Who We Are: CDW Overview." <http://www.aboutcdw.com/>.
- Dell. 2015. "Dell and CDW Expand Partnership to Make All Dell Solutions Available in North America, Europe and Asia." <http://www.dell.com/learn/us/en/id/press-releases/2015-10-12-dell-and-cdw-expand-partnership>.
- Desroches, L.B., H. Fuchs, J.B. Greenblatt, S. Pratt, H. Willem, E. Claybaugh, B. Beraki, M. Nagaraju, S.K. Price, and S.J. Young. 2014. *Computer Usage and National Energy Consumption: Results from a Field-Metering Study*. Lawrence Berkeley National Laboratory Report No. 6876E.
http://eetd.lbl.gov/sites/all/files/computers_lbnl_report_v4.pdf.
- Dewart, N., P. Delforge, P. May-Ostendorp, V. Zivojnovic, and D. Mista. 2014. *Codes and Standards Enhancement (CASE) Initiative for PY 2013: Title 20 Standards Development – Analysis of Standards Proposal for Computers – Addendum to Submittal on August 2013*. California Energy Commission Docket No. 12-AAER-2A. Prepared for Pacific Gas & Electric Company, Southern California Edison, San Diego Gas & Electric, Southern California Gas Company, and Natural Resources Defense Council.
http://www.energy.ca.gov/appliances/2014-AAER-01/prerulemaking/documents/comments_12-AAER-2A/California_IOUs_Standards_Proposal_Addendum_Computers_2014-10-27_TN-73899.pdf.
- Ecotope, Inc. 2014. *Residential Building Stock Assessment: Metering Study*. Prepared for Northwest Energy Efficiency Alliance. <https://neea.org/docs/default-source/reports/residential-building-stock-assessment--metering-study.pdf?sfvrsn=6>.
- Ecova. 2016. "80 PLUS Certified Power Supplies and Manufacturers." Plug Load Solutions, implemented by Ecova. <http://www.plugloadsolutions.com/80pluspowersupplies.aspx#>.
- ENERGY STAR. 2017. *ENERGY STAR Unit Shipment and Market Penetration Report – Calendar Year 2016 Summary*. U.S. Environmental Protection Agency and U.S. Department of Energy.
https://www.energystar.gov/ia/partners/downloads/unit_shipment_data/2016_USD_Summary_Report.pdf?12be-ead3.
- ENERGY STAR. 2016a. "ENERGY STAR Enterprise Servers Template V2."
https://www.energystar.gov/index.cfm?fuseaction=third_party_certification.xmlDataRequirements&qxt_id=68.

- ENERGY STAR. 2016b. *ENERGY STAR Unit Shipment and Market Penetration Report – Calendar Year 2015 Summary*. U.S. Environmental Protection Agency and U.S. Department of Energy.
https://www.energystar.gov/ia/partners/downloads/unit_shipment_data/2015_USD_Summary_Report.pdf.
- ENERGY STAR. 2013. *ENERGY STAR Program Requirements: Product Specification for Computer Servers— Eligibility Criteria, Version 2.1*. U.S. Environmental Protection Agency and U.S. Department of Energy.
<https://www.energystar.gov/sites/default/files/asset/document/Version%202.1%20Computer%20Servers%20Program%20Requirements.pdf>.
- Federal Acquisition Regulation (FAR) Subpart 23.203. 2005. Energy Policy Act of 2005.
https://www.acquisition.gov/sites/default/files/current/far/html/Subpart%2023_2.html.
- Federal Acquisition Regulation (FAR) Subpart 23.704. 2007.
https://www.acquisition.gov/far/html/Subpart%2023_7.html.
- Gerke, B.F., A.T. Ngo, and K.S. Fisseha. 2015. *Recent Price Trends and Learning Curves for Household Led Lamps from a Regression Analysis of Internet Retail Data*. Lawrence Berkeley National Laboratory Report No. 184705. http://eetd.lbl.gov/sites/all/files/lbnl-184075_o.pdf.
- The Green Grid. 2012. *Data Centre Life Cycle Assessment Guidelines*. White Paper #45, v2.
<http://www.thegreengrid.org/~media/WhitePapers/WP45v2DataCentreLifeCycleAssessmentGuidelines.pdf>.
- Hsu, C-H. and S.W. Poole. 2013. “Revisiting Server Energy Proportionality.” *Proceedings of the 42nd International Conference on Parallel Processing*. Lyon, France.
<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6687423>.
- International Data Corporation (IDC). 2014. *Worldwide Quarterly Server Tracker. Installed Base, 2006-2018*. Framingham, MA: IDC. December.
- Intertek. 2016. *White Paper: Investigation of Potential Approaches to Energy Efficiency Metrics for Enterprise Servers, Based Upon the SERT Rating Tool – Draft Final Report*. Ecodesign Technical Assistance Study for Lot 9: Enterprise Servers and Enterprise Data Storage. Produced for European Commission, Directorate General Growth. <http://www.server-standards.eu/wp-content/uploads/2016/04/Investigation-SERT-Metrics-draft-Paperv1-5.pdf>.
- Koomey, J. 2011. “Growth in Data Center Electricity Use 2005 to 2010.” Analytics Press, at the request of *The New York Times*.
<http://www.mediafire.com/file/zzqna34282frr2f/koomeydatacenterelectuse2011finalversion.pdf>.
- Koomey, J. 2012. “The Economics of Green DRAM in Servers.” Analytics Press, at the request of Samsung Electronics Co., Ltd.
http://www.mediafire.com/view/uj8j4ibos8cd9j3/Full_report_for_econ_of_green_RAM-v7.pdf.
- Koomey, J. and J. Taylor. 2017. *Zombie/Comatose Servers Redux*. A report by Koomey Analytics and Anthesis.
<http://anthesisgroup.com/wp-content/uploads/2017/03/Comatsoe-Servers-Redux-2017.pdf>.
- Meisner, D., B.T. Gold, and T.F. Wenisch. 2009. “PowerNap: Eliminating Server Idle Power.” *Proceedings of the 14th International Conference on Architectural Support for Programming Languages and Operating Systems*, Washington, DC. Association for Computing Machinery. P. 205-216.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.152.483&rep=rep1&type=pdf>.

Opinion Dynamics Corporation. 2011. *California Statewide Business and Consumer Electronics Program New Products Baseline*. Prepared for Pacific Gas & Electric Company, Southern California Edison, and San Diego Gas & Electric.

http://www.calmac.org/publications/BCEP_New_Products_Market_Baseline_Report_Jan_2011_V2.pdf.

Natural Resources Defense Council. 2014. *Data Center Efficiency Assessment – Scaling up Energy Efficiency across the Data Center Industry: Evaluating Key Drivers and Barriers*. Issue Paper 14-o8-A.

<https://www.nrdc.org/sites/default/files/data-center-efficiency-assessment-IP.pdf>.

Navigant Consulting, Inc. 2009. *Energy Savings Potential and RD&D Opportunities for Commercial Building Appliances: Final Report*. Prepared for U.S. Department of Energy—Office of Energy Efficiency & Renewable Energy—Building Technologies Program.

http://apps1.eere.energy.gov/buildings/publications/pdfs/corporate/commercial_appliances_report_12-09.pdf.

Northwest Energy Efficiency Alliance. 2013. *80 PLUS Power Supplies, Innovation to Action, Success Story*. Northwest Energy Efficiency Alliance.

<https://neea.org/docs/default-source/previously-funded-initiatives/neea-previously-funded-initiative-80-plus.pdf?sfvrsn=16>

Northwest Energy Efficiency Alliance. 2011. *NEEA Market Progress Evaluation Report #3: 80 PLUS*. Prepared by Navigant Consulting.

<http://neea.org/docs/reports/neea-market-progress-evaluation-report-3-80-plus--e11-226.pdf?sfvrsn=6>.

NVIDIA. 2016. “What is GPU-Accelerated Computing?” <http://www.nvidia.com/object/what-is-gpu-computing.html>.

Pixley, J.E. and S.A. Ross. 2014. *Monitoring Computer Power Modes Usage in a University Population*. California Plug Load Research Center, University of California, Irvine. Prepared for California Energy Commission. <http://www.energy.ca.gov/2014publications/CEC-500-2014-092/CEC-500-2014-092.pdf>.

Rasmussen, R. and G. Wickes. 2008. “80 PLUS: Market Impacts and Lessons Learned.” *2008 ACEEE Summer Study on Energy Efficiency in Buildings*. Pacific Grove, CA.

http://aceee.org/files/proceedings/2008/data/papers/6_477.pdf

Scaramella, J., J. Daly, M. Marden, and R. Perry. 2014. *The Cost of Retaining Aging IT Infrastructure*. IDC White Paper sponsored by Lenovo.

<http://www.lenovo.com/images/products/server/pdfs/whitepapers/IDC%20Whitepaper%20246755.pdf>.

Shehabi, A., Smith, S.J., Horner, N., Azevedo, I., Brown, R., Koomey, J., Masanet, E., Sartor, D., Herrlin, M., Lintner, W. 2016. *United States Data Center Energy Usage Report*. Lawrence Berkeley National Laboratory Report No. 1005775. <http://eta.lbl.gov/sites/all/files/lbnl-1005775.pdf>.

Standard Performance Evaluation Corporation (SPEC). 2016. *Server Efficiency Rating Tool (SERT) Design Document 1.1.1*. https://www.spec.org/sert/docs/SERT-Design_Document.pdf.

Subramaniam, B. and W. Feng. 2013. “Towards Energy-Proportional Computing for Enterprise-Class Server Workloads.” *Proceedings of the 4th International Conference on Performance Engineering*. Prague, Czech Republic. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.294.9573&rep=rep1&type=pdf>.

U.S. Department of Energy—U.S. Energy Information Administration. 2012. *Commercial Buildings Energy Consumption Survey (CBECS)*. <http://www.eia.gov/consumption/commercial>.

U.S. Environmental Protection Agency—ENERGY STAR Program. 2007. *Report to Congress on Server and Data Center Energy Efficiency – Public Law 109-431*.
https://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final.pdf

Xergy Consulting. 2016. *Determining a Real-World Adjustment Factor for Computer Energy Use: Laboratory Testing the Impact of Real-World Idle, Active Mode and Peripherals*. California Energy Commission Docket No. 14-AAER-02. Prepared for Pacific Gas & Electric Company, Southern California Edison, San Diego Gas & Electric, and Southern California Gas Company.

Yang, Y. and E. Williams. 2009. “Logistic Model-based Forecast of Sales and Generation of Obsolete Computers in the U.S.” *Technological Forecasting and Social Change* 76(8): 1105–1114.
<http://www.sciencedirect.com/science/article/pii/S0040162509000390>.