

IBM Comments to the California Energy Commission Efficiency Committee Product Energy Efficiency Scoping Proposal
Docket Number 11-AAER-1 2009 Rulemaking Proceeding on Appliance Efficiency Regulations Phase II

IBM appreciates the opportunity to comment on the California Energy Commission Efficiency Committee scoping proposal to develop a three year plan for the implementation of additional product energy efficiency requirements. As a designer and manufacturer of Enterprise Server systems, IBM has a strong interest in the efficiency committee's review of the appropriateness of mandatory energy efficiency requirements for server systems. IBM has a longstanding product stewardship program, established in 1991, which emphasizes a proactive and strategic approach to the company's environmental design and management of products. The program's mission is to develop, manufacture and market products that are increasingly energy efficient; can be upgraded and reused to extend product life; incorporate recycled content and environmentally preferable materials and finishes; and can be recycled and disposed of safely.

Server systems are complex products in which a single machine type will have a wide range of configurations (combinations of processor, memory, I/O, and disk storage components) with widely varying power signatures. As the number of processor sockets and cores increases, so does the types and complexity of the workload that will be managed, the number and variety of supportable components, and the range of workload capacity. The various cost/benefit ratios of all of these factors must be considered when establishing energy efficiency requirements. Thus, a purchaser of a computer server must evaluate and select a range of options when deciding what server system is best for their application – blade versus rack, virtualization capable, software configurations, and the type of peripherals – as they provide a much broader range of choices and possible configurations than are presented for other product types. In addition, the work delivered per unit of energy expended (as the sum of active and idle power) is significantly improved with virtualization and increased system utilization. Thus, server efficiency has to be measured in the context of the following server attributes: capacity of the system (the maximum workload delivered per of power consumed), the ability to keep the IT asset fully utilized, power use at idle, and the serviceability and reliability requirements of a given application.

The energy efficiency of ICT products is determined by four product attributes:

1. Power Supply Efficiency: The EPRI/ECOS 80+ program sets specific requirements for power supply efficiency over the operating range of the power supply, designating specific levels of power conversion efficiency.
2. The ability of the IT product to reduce its energy use when no workload is present: IT equipment and components are increasingly incorporating power management techniques which reduce the power use of processors, memory, and I/O components when no workload is present.
3. The Computing Capacity: This is a measure of the computing capacity which can be delivered by the equipment for each unit of power consumed.
4. System Utilization: Servers have the ability to run multiple workloads simultaneously, referred to as virtualization of workload. This enables an individual IT device to run multiple workloads and increase the utilization of the device. This

DOCKET
11-AAER-1
DATE <u>Sept 29 2011</u>
RECD. <u>Sept 30 2011</u>

allows a single IT device to support the many workloads that were previously performed on single IT devices. Use of virtualization technologies delivers more workload for each unit of energy consumed, reduces the amount of equipment required to deliver a specific workload, and reduces the required data center space to support the workload.

Establishing requirements which fairly account for all of these attributes and do not bias the requirements to low power systems is extremely difficult, as demonstrated by the time and effort required by EPA and interested stakeholders to establish criteria for the V1 and V2 of the ENERGY STAR® Server Requirements.

POWER SUPPLY CRITERIA

Power supply efficiency is a straight-forward criterion, as the EPRI/ECOS 80+ program has defined criteria against which companies are developing power supplies for server systems.

IDLE POWER CRITERIA

Idle power has been the next criteria of choice for voluntary and regulatory program proposals, but systems have widely varying power profiles dependent on their intended purpose, capabilities to support virtualization and high performance computing, their service and reliability capabilities and other factors. As an example, the two charts below show the power use of the 4 processor socket servers at idle and full power for the maximum and minimum configurations.

IBM Comments to the California Energy Commission Efficiency Committee Product Energy Efficiency Scoping Proposal
Docket Number 11-AAER-1 2009 Rulemaking Proceeding on Appliance Efficiency Regulations Phase II

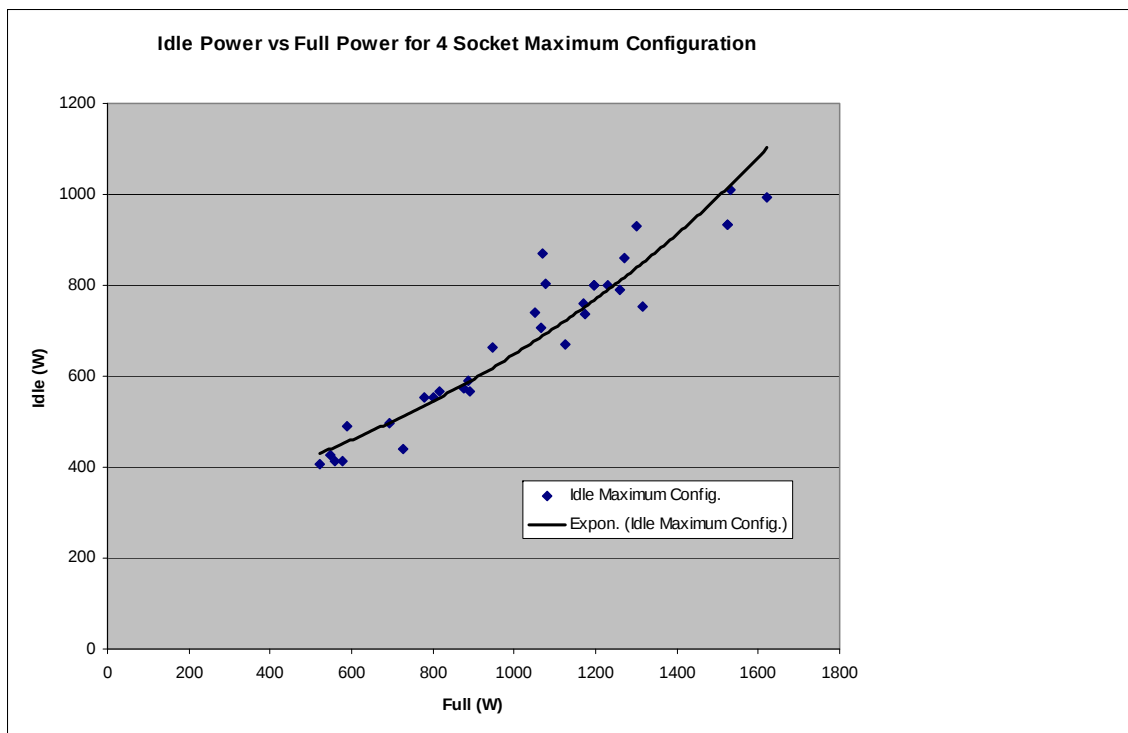


Figure 1: Distribution of Power Use for 4 Processor Socket Systems at Maximum Configuration

IBM Comments to the California Energy Commission Efficiency Committee Product Energy Efficiency Scoping Proposal
Docket Number 11-AAER-1 2009 Rulemaking Proceeding on Appliance Efficiency Regulations Phase II

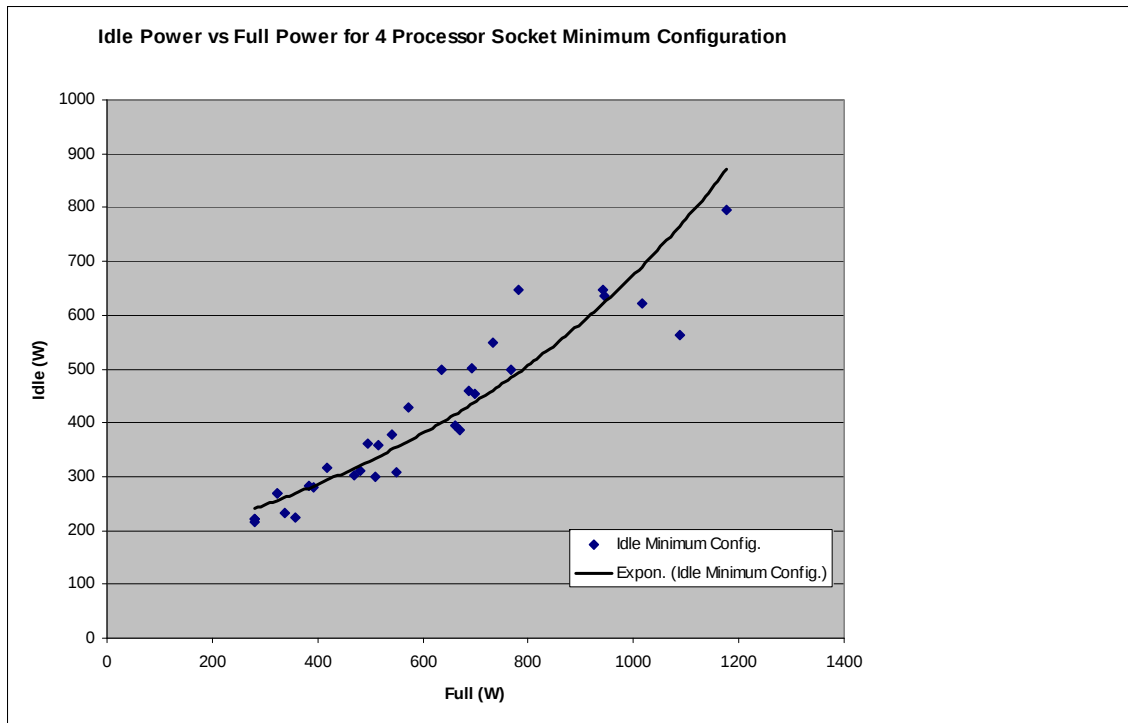


Figure 2: Distribution of Power Use for 4 Processor Socket Systems at Minimum Configuration

These two graphs illustrate the wide range in power use of 4 processor socket systems. The data for these graphs was taken from the ENERGY STAR® website in July of 2011 and represents all individual servers and server product families qualified to V1 at that time. It is expected that the same distribution would be seen for 2 processor sockets systems available on the market, but because the ENERGY STAR qualified systems are biased to the lower power systems by the current idle criteria the graph of the ENERGY STAR qualified 2 processor socket systems is not provided. The differences in power use depend on the system infrastructure and configuration. Larger, more robust systems will have a higher power use because their infrastructure is designed to provide higher reliability, availability and serviceability (RAS) as well as higher performance. More reliable servers with higher functionality and capability have these additional infrastructure power demands which differentiate them from volume or managed servers:

IBM Comments to the California Energy Commission Efficiency Committee Product Energy Efficiency Scoping Proposal
Docket Number 11-AAER-1 2009 Rulemaking Proceeding on Appliance Efficiency Regulations Phase II

1. Use of high performance Processor Sockets: Some processors have a higher performance and functional capability which drives higher socket level power draws because of the presence of some or all of the following functionality and infrastructure.
 - 1.1 The systems are expandable to a larger socket count and include the infrastructure support for extra, built-in processor busses. The technology used in these servers provides the capability to expand systems beyond 4 sockets to provide additional capability and functionality. This requires the processors to carry additional overhead/infrastructure power.
 - 1.2 Supports a higher number of on-board, high bandwidth I/O slots and/or interfaces: The high performance I/O controller to support these slots may be embedded within the main processor socket or reside as a separate processor on the system board.
 - 1.3 They typically include chips with larger core counts or multiple (2 or more) processor chips within the same socket. This will be specific to a particular technology cycle. Under current design standards, each new generation of chips has higher core counts.
 - 1.4 They can include dual or multiple chip modules (2 or more processor chips to a socket) to drive higher core counts within the same socket. This will be specific to a particular technology cycle.
 - 1.5 The cores support Multi-Threading threading technology with 4 or more threads to enable a higher degree of consolidation.
 - 1.6 Support for an integrated hardware accelerator: The processor implements an accelerator technology either through circuitry in the processor or a bus attachment.
 - 1.7 Reliability, Availability, and Serviceability Infrastructure: The processor supports various RAS functions possibly including but not limited to: ECC on second level caches and beyond, parity on chip internal I/O paths, CRC on memory channel, processor checkpoint retry and recovery, memory RAS (ECC, chipkill, mirroring), I/O redundancy, and runtime processor de-allocation.
2. Use of memory buffers: Memory buffer can be a separate, standalone buffer chip which is integrated on the system board or on custom-built memory cards. The use of the buffer chip is required for extended DIMM support; they allow larger memory capacity due to support for larger capacity DIMMs, more DIMM slots per memory channel, higher memory bandwidth per memory channel than direct-attached DIMMs, and support of spare memory for RAS capabilities.
3. Advanced RAID support: Resilient systems have at least Level 5 RAID support in the base configuration. The systems also have dedicated infrastructure for an additional RAID controller card.
4. Greater I/O expandability: Resilient servers have larger on-board I/O infrastructure and support a higher number of I/O slots; maximum of 5 for non-resilient systems vs. 6 or more for resilient systems.
5. Requires redundant power supplies.
6. RAS Capabilities as defined by US EPA ENERGY STAR Computer Server Requirements (file "ENERGY STAR Servers Dataset Assembly Test Method.pdf") Lines 55-69.

Because of the higher power demand of the overall system based on the attributes above, a higher capacity fan and higher capacity power supplies are required. This further drives higher overall system power use and generates a power profile which is significantly different from the power profile of low power servers.

The point of this discussion is to illustrate that a single idle criteria for all servers with a given number of processor sockets does not adequately differentiate between types of servers and will result in a criteria biased to low power systems, which are likely to have a reduced computing capacity or ability to virtualize. In practice, this could require a purchaser of ENERGY STAR systems to be required to install many low power servers to execute a specific set of workloads, with a higher total net power use, when a single higher powered server could deliver the workload at a much lower net power cost. Chart 13 of the ITI presentation at the August 30th workshop provided an example of the reduced power use that can be achieved with higher power systems and several examples of consolidation projects from the IBM Silicon Valley Lab in San Jose California will be provided later in these comments.

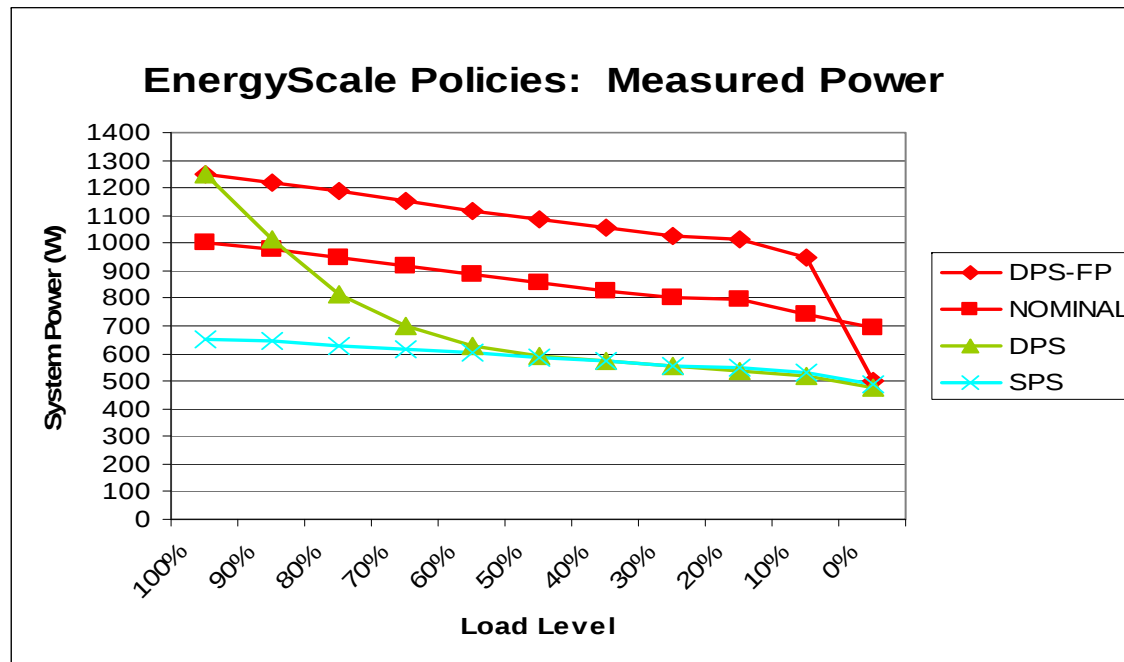
Given this information and data, IBM submits that it is important that any mandatory energy efficiency requirements for servers include the following considerations:

1. The efficiency committee needs to create at least 3 server categories as defined in the ENERGY STAR server requirements: non-managed, managed, and resilient servers. This categorization will differentiate between the base infrastructure power within the server systems as discussed above.
2. For the resilient server category (which is described by items 1 to 5 on page 4) and the 4 processor socket category, the efficiency committee needs to pursue criteria other than a straight idle criterion. IBM has the following recommendations:
 - a. Ultimately, the efficiency committee should look at utilizing the SERT metric to integrate both the idle and active energy performance into a single test and metric. This will require the CEC to put the server at the end of the 2011-2014 implementation period or the next (post-2014) implement to provide sufficient time to finalize the SERT metric and collect an initial data set to allow reasonable criteria to be set. This delay can be justified based on the current availability of energy savings attributes discussed below.
 - b. If the efficiency committee wants to set a power based metric with currently available test methods, then IBM recommends that the criteria be set based on an idle to maximum power ratio by server category. This approach normalizes for higher server infrastructure and enables the more robust systems to be included.
 - c. The efficiency committee is encouraged to consider the ENERGY STAR V1 Server Testing requirement and the V2 product family definition for classifying server products. Because of the many product configurations, testing can be extremely difficult and expensive. It is important to enable manufacturers to use an easily executable, universally recognized test method to qualify products.

IBM Comments to the California Energy Commission Efficiency Committee Product Energy Efficiency Scoping Proposal
Docket Number 11-AAER-1 2009 Rulemaking Proceeding on Appliance Efficiency Regulations Phase II

3. For blade servers, IBM recommends that the efficiency committee exclude blade servers from this implementation period. The industry is still working with the ENERGY STAR® team to develop the test method. The test method is complicated by the fact that you have to coordinate blade servers in a half full or full rack and collect power data. In addition, more test data is needed to demonstrate the impact of the differences in chassis based shared services (I/O, fans, and power supplies) between manufacturers and to validate that the test process and criteria properly measures the “fully burdened” blade server power use so that any metric or criteria is representative of the power use of a single blade server.
4. Any criteria should be designed to eliminate the bottom 20 to 40% of the available systems from the marketplace. Removal of the bottom effectively forces manufacturers to improve their system efficiency while recognizing that a power metric does not necessarily differentiate effectively between high and low power systems or systems designed for specific applications or high levels of virtualization.

Server manufacturers are driving reduced idle power demand through the implementation of processor, memory and I/O power management functions in their server products. IBM product design engineers have utilized power management capabilities provided by the x86 and POWER7® processors and memory and I/O components to reduce the power drawn by the server when no workload is present (idle power) by 20 to 50 percent when compared to the power used by the system at full workload for products released in 2010. Implementation of these functions on products enables increases in the ratio of idle to maximum power for a server system and allows systems to significantly reduce energy use when no workload is present. The graphic on page 8 details the 4 processor power management settings offered with the P7 processor and the reduction in system power use that occur when these settings are deployed and workload to the system varies or goes to zero. Power use is reduced over 50% when the system is at idle. X86 have similar capabilities which can deliver similar results. It must be noted that for some settings, such as DPS and SPS, the dictated processor frequency reductions results can create performance issues for customers when they are running certain workloads.



SERVER COMPUTE CAPACITY

As a rule, each generation of servers have an increased capacity per watt of power consumed because of advances in semiconductor and storage technology which underpin the ability of processor, memory, storage, and I/O components to deliver more work for each unit of energy consumed with each new product generation. In recognition of this, IBM has established a goal that each new generation of server products deliver improved performance per watt of power consumed over the previous equivalent model. Details of IBM's attainment of this goal are provided on pages 27-30 of IBM's Environmental Report¹. The chart inserted below as figure 1 provides

¹ http://www.ibm.com/ibm/environment/annual/IBMEnvReport_2010.pdf

details on performance per watt increases in IBM systems p and z products over the past 10 years.

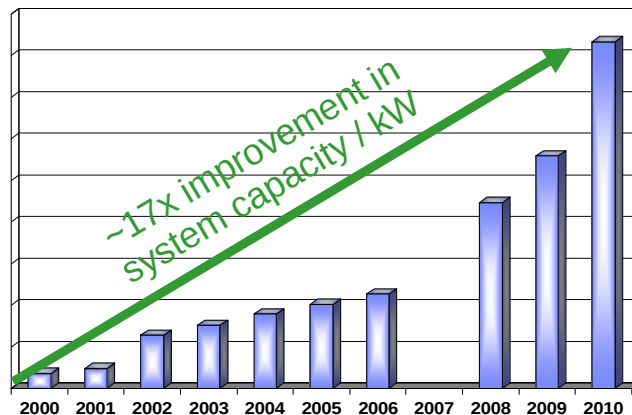
Smarter Systems for a Smarter Planet



IBM Systems : Improvements over time

POWER

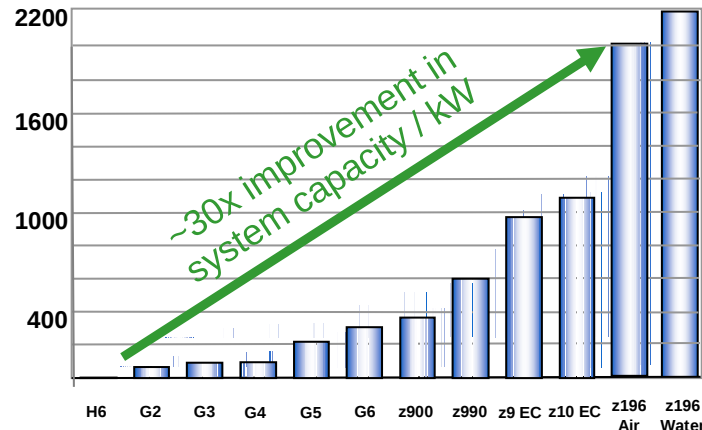
Relative Performance per Watt improvements



Net Effect: P4 to P7	
Performance increased by:	~65x
Performance / kWatt increased by:	~17x
Performance / sq ft increased by:	~65x

z196

Capacity per Watt improvements



Net Effect: G2 to z196 *	
Performance increased by:	~300x
Performance / kWatt increased by:	~30x
Performance / sq ft increased by:	~190x

Note: Based on relative performance measurements of similar system models with the different processor families.

1

© 2010 IBM Corporation

Figure 3

The Japan Energy Law sets server requirements based on capacity and necessitates that server manufacturers improve the compute capacity of their products with time to sell into the market. While this provides a useful indication of system capability, it does not provide a meaningful measure of system energy efficiency without the consideration of system utilization – what percent of the time is the system doing work.

SYSTEM UTILIZATION

IBM servers offer virtualization capabilities, which allow the server to perform multiple workloads on a single server device enabling system utilizations to be driven to 50% or higher. IBM's system z mainframe computers are capable of utilizations of greater than 90% with reliability and serviceability metrics of greater than 99.9%. IBM is utilizing virtualization technologies to consolidate multiple workloads from servers with low utilization onto single servers. These projects increase the utilization of the virtualized servers, deliver more workload for less energy, and reduce the number of servers and the data center floor space required to perform a given workload. The Lexington Data Center project, detailed in the paper referenced above, used virtualization technologies to consolidate 70% of the baseline workload and free up power and space for new workloads². In 2009, virtualization and consolidation projects reduced data center energy use across the IBM data center portfolio by over 30,000 MWH. Not only do these projects reduce energy use, but they free up data center space for business growth or new business opportunities. An internal project on the IBM account freed up approximately 10,000 square feet (6.6% of the specific facility's raised floor space) by consolidating workloads from individual server systems onto several system z servers.

Table 1 below details the results of virtualization projects completed at IBM SVL labs in San Jose CA. Some of these projects received energy efficiency rebates. The reductions from these projects represent a reduction of 7% in the energy use to deliver the base workload at the facility, as well as a reduction in required floor space.

² <http://public.dhe.ibm.com/common/ssi/ecm/en/sfw03001usen/SFW03001USEN.PDF>

IBM Comments to the California Energy Commission Efficiency Committee Product Energy Efficiency Scoping Proposal
Docket Number 11-AAER-1 2009 Rulemaking Proceeding on Appliance Efficiency Regulations Phase II

Table 1: Example Virtualization Projects at IBM SVL Labs: 2010/2011

Project Title	Description	KW Savings	MWH/yr Savings	Rebate (K\$)
X Series Virtualization	172 existing servers were scrapped and replaced with 25 virtual machines	78.0	892	25
Storage Refresh	34 machines were removed and replaced with 10 new	37.5	430	33
Z Series Storage	21 storage devices were removed and replaced with 3 new machines	101.1	1150	27
Z Series Processor	7 processors were scrapped and replaced with 2 new Z-10	22.0	250	28
X Series Virtualization	76 existing machines were scrapped and replaced with 17 virtual machines	32.0	360	0
	Savings include cooling energy use at a PUE of 1.3			

Unfortunately, there is no effective methodology or metric to measure or define the capability of a server system to virtualize or process multiple workloads. Several metrics groups, such as SPEC™, are investigating or developing metrics to assess utilization efficiency but those are not currently released or mature enough to serve as the basis for mandatory energy efficiency requirements for servers.

SERVER SYSTEM ENERGY EFFICIENCY CAPABILITIES

The discussion above emphasizes the fact that current server systems offer power management and virtualization capabilities to data center operators. When those function are implemented, they enable a significant reduction in the number of server systems they need to perform a defined workload and the quantity of power the servers consume when there is no workload present. This in turns optimizes/maximizes the workload delivered per unit of energy consumed in the data center. IBM encourages the efficiency commission to consider the fact that current server models offer customers significant opportunities to reduce energy use and increase the utilization of installed equipment reducing the number of servers and the power consumption required to perform a given

IBM Comments to the California Energy Commission Efficiency Committee Product Energy Efficiency Scoping Proposal
Docket Number 11-AAER-1 2009 Rulemaking Proceeding on Appliance Efficiency Regulations Phase II

workload. Installation of newer servers with power management and virtualization functions deployed was cited by Jonathon Koomey³ as an important contributing factor to the slowing of the growth in data center energy use from 2005 to 2010 and a reduction in the installed base of servers over the same period. In addition to considering energy efficiency standards for servers, the efficiency commission should work with the appropriate stakeholders, such as the efficiency utilities, to exploit these opportunities and increase the adoption of virtualization and power management technologies in California data centers. The latter effort may prove to be a more effective means to reduce data center energy use than the setting of mandatory server energy efficiency requirements.

OBSERVATIONS AND RECOMMENDATIONS

The discussion above details the progress that IBM server systems, and more generally server systems designed and sold by all manufacturers, have made in improving energy efficiency performance in four key areas delineated in this discussion, resulting in the delivery of more workload for each watt of power consumed by those servers where data center operators have chosen to implement these capabilities. Several examples of the actual energy use reduction benefits of these capabilities were provided. These improvements and progress in improved energy efficiency and functionality which enables improved utilization of the hardware assets has occurred in the absence of regulation.

Given the progress that companies have made the fact that current server systems offer power management and virtualization capabilities, which if fully deployed, can significantly reduce data center energy consumption, the energy efficiency commission should consider programs to promote the adoption of these systems in California data center operations in lieu of setting energy efficiency requirements for servers, as this approach may deliver a higher rate of energy use reductions.

IBM strongly recommends that if the efficiency committee intends to move ahead with energy efficiency requirements for servers that it carefully consider the categories of servers that are covered by the requirements and the nature of the requirements themselves.

1. Limit the coverage of the requirements to high volume, one and two processor socket server systems which lend themselves to specific requirements. IDC current data and projections indicate that approximately 98% of servers sold have a value of \$25,000 and the majority of these systems are expected to be one and two processor socket systems.
2. Set any power supply efficiency standards against the 80Plus Criteria, as this criterion is recognized by server manufacturers and is in use in several voluntary and mandatory programs.

³ Koomey, Jonathon G.; *Growth in Data Center Electricity Use, 2005 to 2010*; August 1, 2011; A Report by Analytics Press completed at the request of the New York Times.

IBM Comments to the California Energy Commission Efficiency Committee Product Energy Efficiency Scoping Proposal
Docket Number 11-AAER-1 2009 Rulemaking Proceeding on Appliance Efficiency Regulations Phase II

3. The efficiency committee should defer inclusion of blade systems in the requirements at this time, due to the fact that an effective test method has not been finalized and the current proposed ENERGY STAR test method for blade systems has not been validated to demonstrate that it provides a representative distribution of chassis power to a single blade.
4. The efficiency committee should carefully consider requirements for 4 processor socket, as these servers are more complex, integrated systems with larger power footprints, costs, and testing complexities than 1 and 2 socket systems. Ideally, the first set of mandatory requirements should be targeted to just one and 2 processor socket systems. If the commission chooses to proceed with standards for 4 processor systems, it is recommended that at least two categories, managed and resilient, be created and that the power metric is based on an idle to maximum power ratio. More work and data collection needs to be done to find the right algorithm to analyze an idle to max ratio and consider the different base load power use for a minimum and maximum configured server. At a minimum, IBM recommends that these requirements for 4 socket systems should be deferred to the next scoping period to provide adequate time for development of appropriate metrics and collection of supporting data to implement requirements.
5. Because of their complexity, wide range of configurations and power profiles, and limited market share, systems with more than 4 processor sockets should be excluded. These systems are typically designed for high RAS applications and workloads and because of their higher costs and capabilities are run at utilization rates which can exceed 50%.
6. If the commission chooses to select specific idle or active power criteria, it needs to carefully consider the energy efficiency criteria to insure that the standard does not unreasonably exclude systems from the market, specifically high performance, and higher power servers.
7. Servers are complex products, particularly 4 processor socket and blade servers, with a wide range of configurations. Any metric should be set to exclude the bottom 10 to 40% of the server systems, as systems with higher power use profiles may in fact offer the best delivered workload to power consumption ratio when considering power management capability, computing capacity, and system utilization for a given set of applications.

For questions or additional information on these comments, please contact Jay Dietrich, jdietric@us.ibm.com.