the green grid
get connected to efficient IT

ENERGY STAR PROGRAM REQUIREMENTS FOR COMPUTER SERVERS DRAFT 3

OVERALL COMMENTARY AND KEY CONSIDERATIONS

The Green Grid (TGG) believes that the EPA's efforts to create ENERGY STAR for Servers are positive steps on the path towards making current and future data centers more energy efficient. Many of these efforts are in harmony with TGG's own efforts across the OEM and end-user community in developing standards to measure and mange efficiency across data centers and the larger computing infrastructure. TGG considers the EPA and specifically the ENERGY STAR program an industry partner in this effort. TGG supports and acknowledges the EPA and the ENERGY STAR standards in targeting the reduction of energy consumption across a range of client systems and peripherals. Cooperation and support of ENERGY STAR standards is an essential component in the ongoing effort to reduce the impact of climate change. The industry and EPA have met many times and the industry supported the development of the EPA report to congress on data centers, and it is a common point of reference in marketing communications of the entire industry.

TGG's feedback on the ENERGY STAR Program Requirements for Computer Servers remains focused on how these requirements would impact energy efficiency in the data center. In that spirit, TGG would like to see more direct input from federal data center managers to ensure that any proposed requirement would meet end-user needs. Federal data center managers will be the earliest, and most likely, consumers of ENERGY STAR for Computer Servers products. They will be an essential indicator of both the potential success of the program as well as any potential negative consequences.

After the EPA report to Congress several industry working groups and organizations formed to address the issues identified. Some are The Green Grid (TGG), Climate Savers Computing Initiative, and 80Plus. TGG has been vigorous in setting aggressive goals and definitions around energy consumption and utilization. A development effort in SPEC to enhance and potentially utilize the product and methodology of the SPECpower workgroup is already in process. These efforts created a roadmap with industry guidance of improvement opportunities, and established a baseline and direction for industry efforts to target. TGG hopes to ensure that the ENERGY STAR metrics would be consistent with the industry activities already under way.

Productivity and data have real economic value, and is a major component to energy efficiency. Data centers today sit in the center of the productivity and revenue streams of corporation and governments. As identified in the Technology CEO Council and the AeA Report for Europe, both supported in development by the ACEEE information technology; specifically data centers drive rapid and significant energy efficiency in the US and EU economies.

Due to the complexity and diversity of application, the development cycle, configuration, purchasing, and installation of servers are distinctly different from client devices such as desktops, laptops and workstations. Servers in general are targeted for business critical use in data centers, buildings or rooms specifically designed with custom environmental surroundings. Servers are building blocks and provisioned to operate a broader range of applications and generally part of a larger computing resource pool essential to the business. In many cases, the people requesting a system configuration, the purchasing agent, the installation team and the IT management are totally separate entities. In addition, the operating system image, enabled features and system settings are integrated on site during installation. Adoption of features and hardware settings such as power management tools are coordinated to intercept other changes in the data center not always coincidental to hardware changes, such as major changes in the software, application, or other IT devices.

Although typical ENERGY STAR program practices focus on retail sales, the notes listed in draft 3 calling

for "as-shipped" or reference to the "end user" is very confusing as it pertains to enterprise servers. "As-shipped", which occurs in bare hardware form, is well before the systems are configured and integrated as part of the pool of compute resources along with its configuration settings. Given that the "end user" can be any of several different entities, TGG can best assume that the system configuration requestor and potentially the purchasing agent are the influencing targets for the specification. If these are indeed the end users, TGG recommends using electronic media as part of the configuration and purchasing process to assure these users of the ENERGY STAR compliance of the product. The functionality risk and negative impact of labels, as highlighted in draft 3 is not justified by any assurance advantages, when the "end user" may not ever see this label. The "tested as-shipped" requirement implies an ability to control and predict the installed and use configuration on site, which does not exist for many system manufacturers. TGG recommends that "as-shipped" condition change to testing and selecting system model ranges, under which configuration variations would be deemed compliant. Such a method is consistent with influencing the system configuration requestor on the purchasing choice.

The collective of systems and interoperation defines a level of complexity in architecture that far exceeds client systems. As a result, industry development cycles for servers and end user changes have a cadence of three years or more. In comparison, client devices changes yearly or even more frequent. Processor, chipsets, memory, mass storage, and power supply components do share some commonality across the industry. Despite these commonalities, which are continually enhanced, the differences in architecture, feature sets and focus, at the enterprise (multi-system configurations) are expressly catered to the specific applications, such as extensive memory structures for large secured databases to extensively scaled compute structures for scientific analyses. The schedule of specification changes are far too frequent and likely to disrupt even the fastest development and replacement cycles in the server categories. Here too the categories define different development and refresh cycles. Single and two processor systems are seeing a refresh cycle approximately 3-4years.[1]

Given the level of complexity of servers in development and deployment, an initial recommended activity is to establish and proliferate power saving functions to serve as a base tool suite for energy efficiency in the datacenter. Ignoring the interactions between multiple servers and a common tool suite, would prohibit management and energy savings of these systems. With the critical business functions of these systems, changes in the datacenter will not occur unless there are persistent and reliable (power) management features that will ensure continuous and seamless operation in the datacenter (not just an individual device). With downtime or unavailable systems as serious productivity, cost and energy consequences, implementation of these features must be conducted systematically across the datacenter. Some of the energy saving features in process today include, power capping, dynamic frequency-voltage scaling, and other such tools. A focus to just platform idle instead of facilitating these tools will likely stall the implementation of power management tool adoption, and stall progress in this avenue of energy efficiency in the data center. To ensure that does not happen, TGG believes that the key measurement of the overall effectiveness of the ENERGY STAR for Servers Program, should be:

- Actual reduction or increase-avoidance in energy use by data centers

not simply

- Number of ENERGY STAR compliant servers offered by OEMs

With this focus, instituting power management feature adoption, while the industry progresses on efficiency metrics would progress energy efficiency in the data center far more effectively than trying to select lower idle

1- Quantifying the Server Lifecycle Worldwide Survey Results, IDC July 2004

systems. The Green Grid would also like to be part of a review process, to determine if the industry is indeed achieving energy efficiency targets.

While TGG members have been in discussion with the EPA on the program during the development of the possible methods, the industry has only recently been able to review the idle proposal and groupings listed in draft 3 (11/4/2008). The dataset (11/13/2008) which was instrumental in the definition and plans presented in draft 3, appears to be problematic given several errors in interpretation and submissions. For example, one system was listed as a 4 processor system, when in fact this was a single processor system. Thus far 3 errors were uncovered in the 4 processor category, reducing that category to approximately 24data points and raising the 25th percentile from 271W to between ~311W and ~339W, just based on the current information. A more statistically valid limit should be determined after the dataset and characteristics are comprehended.

Though TGG differs with the EPA's approach of using unqualified (by performance) values of system idle power, it's very clear that investigation and clarity in the dataset is necessary. Even given the EPA's desire to use unqualified system idle power, the TGG recommends that these data points be reviewed to ensure sufficient information on each has been provided. Another concern associated with the data set is the single socket platforms, in particular if any of these platforms are in reality "small scale servers" as defined in ENERGY STAR for Computers, v5.0. TGG recommends that the single processor system data be reviewed with each of the sources of those data points, such that they can be assured as part of the enterprise class of servers, and can feasibly be deployed as other systems in a rack mount configuration. A data review of the existing data set may take the remainder of December 2008 to resolve, especially given the availability of the key engineers during the US holidays. Do note that such delays may also impact the date of the final revision of the Tier 1 specification.

ENERGY STAR DRAFT COMMENT BY SECTIONS

POWER SUPPLY LIMITS: The Green Grid supports these efforts by the EPA, particularly EPA's choice not to pursue net power loss, and excluding fan power form the PSU calculations. TGG encourages the continued harmonization of these standards with the bronze level standards established by the Climate Savers Computing Initiative.

PRODUCT CATEGORIZATION: The categories listed do not comprehend the design and configurations of these systems as they apply to the deployment of servers in the datacenter. In some cases, partially configured systems are purchased and deployed prior to potential increases in capacity requirements, compute needs, software enhancements or expansion plans. The infrastructure of higher capable platforms, such as memory controllers, I/O controllers, voltage regulation, clocking devices and management logic; all serve to allow a more efficient transition to increase capacity when needed (as opposed to replacing the entire system or delaying software deployments).

The socket designation reflects some of the tools and features in the system required to perform consistency in the data and compute. For example, a 4 socket or 4S system contain the extensibility in memory, I/O, processor, and associated support functions for fast localized operations. All of these compute resources must be accessible by a single OS image (or hypervisor) in a 4S case, and that OS image would then re-provision what is actually available in the system. A 2 processor system would require that OS image to only access and manage the resources supporting that configuration. These systems also provide extensibility for expansion as the business activities expand or become more complex. Consequences of assuming adding

“installed” processor systems to achieve the scalability and application for larger capable systems ignores increased hardware (more 2S systems, more racks, repartition software, or cross bars), complexity (task managers, breaking down larger compute tasks), and resulting energy consumption of those additional machines; as opposed to simply adding capacity (e.g. 16GB to 64GB or 2P to 4P upgrades).
TGG recommends segregation systems grouped by socket capability, where as the metrics used to assess system in the group comprehend the optional configurations, such as amount of memory, I/O, processors, or processor cores actually installed within each group.

the green grid™

IDLE POWER: The presumption implicit in the specification is that “idle” mode power limits coupled with power supply efficiency limits would reduced power consumption in the data center. This may have unintended consequences in the end-user community.

Architecturally, computer servers represent significant complexity increase over desktop, laptop and even workstation systems. The Green Grid group conversations with EPA following the issuance of the draft have not lead to a clearer understanding or change in position of the TGG members with respect to the ENERGY STAR for Servers impact on server architectures. Designing servers with power management features and capabilities, designed to deliver performance per watt characteristics does not equal results from “idle” mode limits for servers.

The number of categories of servers required for ‘idle” mode limits still represents a confluence of the following characteristics.

1) Chassis (1U, 2U and Blade)
2) Operating System (No Operating System)
3) Processor/Socket Configurations
4) Memory Size
5) Disks, No Disks, Raid Configuration
6) Rotating Versus Solid State Disk Technology
7) Communications Ports (Number of Ethernet/Optical Data Interfaces)
8) PCI Express Slots

The above characteristics, summed into a standard measuring “idle power”, results in significant category complexity. TGG is concerned that the EPA is incorrectly prioritizing lower (idle) power consumption ahead of productivity improvements of varied server configurations. Though draft 3 does comprehend memory increases (beyond 32GB) and number of Hard Disk Drives (HDD), the other characteristics and associated components are not accounted for in the idle limit. The simplified motivational aspect of this metric is to minimize the integration of those items and buy more systems to accommodate the requirements of more I/O and/or features. This direction is not consistent with the energy efficiency targets the TGG advocates.

Server power-saving features represent a greater impact on power consumption than “idle” limits modes alone. Though data centers today may run for longer durations at “idle” values that are not indicative of the direction of servers and data centers, which are moving to power managed machines and increasing utilization levels of IT equipment. Power saving modes, consolidation and virtualization activities currently underway in the industry will change power consumption of data centers dramatically in the near future. Absolute “idle” limits may not deliver superior performance per watt capability in computer servers.

For this is the reason TGG strongly recommends qualifying the Idle power level both by categorization of

similarly designed systems and by some measure of compute capability. The wider the range of features and options, as highlighted by 4 processor socket systems, the greater the incentive is to purchase more low capable machines to accomplish application requirements as could be achieved with fewer, more capable machines.

FOCUS ON IDLE POWER MAY ACTUALLY *INCREASE* ENERGY CONSUMPTION BY DATA CENTERS. If the target of ENERGY STAR is intended to incentivize reduction in energy consumption; focusing on idle for servers will result in more servers required to deliver the business output, and more energy is consumed. The outcome does not equal program goals.

To optimize for system Idle power (and ENERGY STAR compliance) servers may be required to be less robust / less capable to meet an idle power requirement, thus relegating these servers to the “low-end” of system performance. Any features, such as power management, inter-system controls, and/or increased productivity would be removed from those systems required to meet those inactive limits

A focus on Idle also drives unintended buyer / consumer behavior. Customers who buy / specify ENERGY STAR compliant servers would get lower-performing servers; and they would end up needing more servers for same datacenter provisioning – which would not result in optimal efficiency for the facility

Overall, a focus on idle power on individual systems does not encourage / motivate best data center management policies. These low idle and low productivity machines may distract data center operators from managing their servers for higher utilization, which is generally agreed as the most efficient operation for datacenters (virtualization and server features supporting virtualization is one set of technologies that aid in increased utilization). As noted by most datacenter and IT managers, idle mode is not the preferred operating mode for IT equipment.

Energy Efficiency Impact of Optimizing Idle Power

	Energy Star for Server v1.0				Enterprise Energy Efficiency			
Server Characteristics	1P	2P	4P w/16GB	4P w 128GB	1S	2Socket	4Socket; 8 DIMM	4Socket; 16 DIMM
IDLE Power	60W	151W	271W	463W (based on 8W/4GB)	Varies (60-150W)	Varies (135-400W)	Varies (250-700W)	Varies (400-900W)
Efficiency Rate	n/a	n/a	n/a	n/a	SpecPower	SpecPower	TPC/W; Linpack_Ops/W	TPC/W; Linpack_Ops/W
Peak Compute	n/a	n/a	n/a	n/a	SpecCPU-rate> xx	Linpack> yy SpecCPUrate>	TPC> xxx Linpack > yyy	TPC> xxx Linpack > yyy
Dynamic Load Balancing	n/a	n/a	n/a	n/a	xxW/20% (system)	yyW/10% (system)	VMM migration yyyW/10%load (racks; datacenter)	VMM migration yyyW/10%load (racks; datacenter)

Selection Results	Energy Star for Server v1.0	Enterprise Energy Efficiency
	Lowest performance processor	Highest Performance Processor configuration
	Smallest and slowest amount of memory	Large resident, coherent, fast memory config
	Fewest features possible	Integrated power management
	Reports data. Controls too slow to adjust.	Dynamic automated load & resource balancing
	Optimizing for low Idle results in: Low Capable system; Stalls energy efficiency (rate); Slows virtualization (smaller capacity) ; Slows down consolidation (small capacity)	

Note: Idle values and metrics listed in the Enterprise Energy Efficiency side are examples and do not encompass the full scope of variety that exists.

Figure 1. Energy Efficiency Impact of optimizing for Idle.

EXAMPLE: Comparing 2 server systems (4 Processor, same base platform processor and other hardware type. System B has 8GB more memory). Note that though both would fall under the same category in the ENERGY STAR definition, the low idle selection is the less efficient choice.

the green grid™

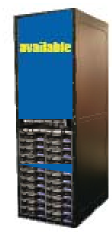
Which is more efficient? (4 Socket Comparison)

Individual Server A	
Max Power	663 W
Idle Power	336 W
Capacity	334.3 Kbops/system

Individual Server B	
Max Power	715 W
Idle Power	358 W
Capacity	411.7 Kbops/system

To meet IT Requirement of 5.0 Mbops

Rack A		Rack B	
Capacity	5.02 Mbops	Capacity	5.35 Mbops
Systems	15	Systems	13
Total Idle	5.04 kW	Total Idle	4.65kW



Choosing Server B is MORE Efficient
Saves 4.54 MWhrs to 3.73 MWhrs per yr.**

Performance is Key to Server Energy Efficiency

* bops= business Operations Per Second

** savings based on 50% utilization to 15% utilization. Does NOT include HVAC or environmental controls

Source: Intel Test Labs

intel

COMPARISON NOTES:

System A is

HP DL580 G5, 4 x X7350 (Tigerton 2.93GHz 130W)

4 x 2G FBD Micron MT18HTF25672FDY-667E1E4

HDD: SAS2.5’ 36G x1

System Idle Power: 336 W

System Workload Power: 663 W

SPECjbb2005 bops performance: 334.35 k

System B is

HP DL580 G5, 4 x X7350 (Tigerton 2.93GHz 130W)

8 x 2G FBD Micron MT18HTF25672FDY-667E1E4

HDD: SAS2.5’ 36G x1

System Idle Power: 358W

System Workload Power: 715W

SPECjbb2005 bops performance: 411.67 k

CALCULATIONS (NOT INCLUDING ENVIRONMENTAL)

Annual 50% utilization (est. using 50% peak and 50% idle):

A) ((50% * 9945 W) + (50% * 5040 W)) * (356 *24) = 65634.3 kWhrs

B) ((50% * 9295 W) + (50% * 4654 W)) * (356 *24) = 61096.6 kWhrs

yrly savings = (65634.3 – 61096.6) = 5537.7 kWhrs

Annual 15% utilization (est. using 50% peak and 50% idle):

A) ((15% * 9945 W) + (85% * 5040 W)) * (356 *24) = 50595.6 kWhrs

B) ((15% * 9295 W) + (85% * 4654 W)) * (356 *24) = 46867.3 kWhrs

yrly savings = (50595.6-46867,3) = 3728.3 kWhrs

the green grid™

DATA SET ACCURACY IS NECESSARY TO DETERMINE NUMERICAL LIMITS (EVEN IF IDLE IS USED)

The actual idle limits appear to be based on a simple cutoff of a list of values (Idle) submitted. As indicated previously, review of this data is necessary to understand whether there are systemic issues that augment the distribution. The consequence of ignoring these systemic changes in the metric may fail to recognize contributions such as power management or availability features which would actually save energy in the datacenter. This becomes a particular concern in the case of a 4 socket system, which have the largest configuration expansion options in the group of products being considered. The distribution distinctly shows systemic issues not being addressed (figure1). TGG recommends normalizing the systemic issues in the distribution, such that any limit or cut off would be based on a normal distribution.

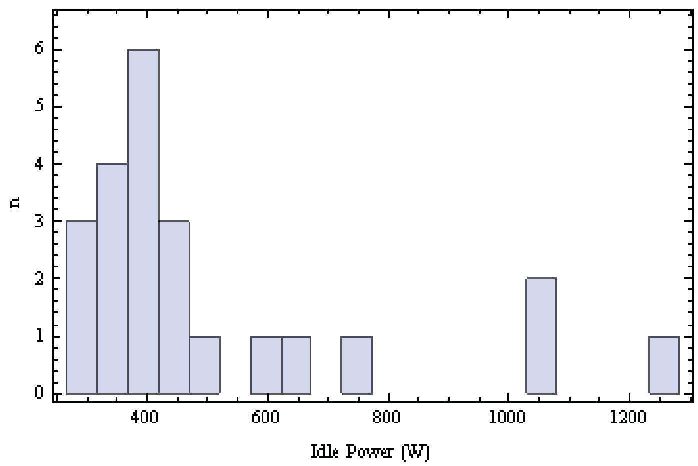


Figure 1. 4Processor Systems in EPA's dataset 11/14/08.

POWER AND TEMPERATURE MEASUREMENT REQUIREMENTS: All servers must have the ability to provide real time data on AC power consumption, inlet air temperature, and processor utilization during server operation.

Revealing standardized data measurements to data center operators is an important feature for ENERGY STAR. Many servers reveal this information today. There are however, some concerns with the ability to report power consumption on several of the lower power range power supplies and those systems without separate system controls. TGG recommends reviewing the availability of this feature, resolution, sampling rate and accuracy across all power supply ranges, before implementing the real time power reporting requirement. Implementation without availability or capability may inadvertently force upsizing or over-sizing the power supply to those which can meet this requirement. Over-sizing is recognized as not a very energy efficient behavior with the current power supply technologies.

TESTING "AS-SHIPPED" AND PHYSICAL LABELING ENERGY STAR FOR END "USER" IDENTIFICATION IS INCONSISTENT WITH THE PURCHASING AND DEPLOYMENT OF ENTERPRISE CLASS SERVERS.

Testing as-shipped is contrary to how enterprise systems are configured, purchased, shipped and installed. Indeed, areas such as O/S images, memory installation, and other items are conducted on site, and may be done by groups outside of the system manufacturer. Consequences of forcing as-shipped methods such as those employed on client, self-enclosed systems, would constitute a significant change to how systems are purchased today. The as-shipped requirement fails to comprehend the significant issues of enterprise integration that is necessary in the datacenter. There are in fact datacenters where the manufacturer is NOT allowed to install their recommended O/S images or hardware into the production facility. The test methods need to reflect the configuration, purchasing, and installation methods used by the consumers of these enterprise systems. In fact, unlike client devices, the IT requestor defining the configuration, the purchaser, the installer for the systems, and IT manager are separate entities. To accommodate an ENERGY STAR-like rating system, the TGG recommends addressing the configuration and purchasing areas, by testing and validating a overall model configuration that would represent the model of system being defined and purchased. The designation of an ENERGY STAR system should also be consistent with entity targeted by the term "user", i.e. which entity. Specifically, use of a label on the system, though it may negatively impact the actual thermal or mechanical fit of the system, provides no influencing advantage to the requestor or purchasing agent listed above. One has to question the use of a label which would both negatively impact efficiency and fail to influence the person purchasing the system.

CONCLUSION: TGG focus has been on the efficiency of the data center. To that purpose the delivery of useful work over energy in has been a focus. TGG supports and endorses activities and standards that lead the industry to a delivery of improved data center efficiency, such as the EPA ENERGY STAR program. Our recommendations drive to the core of a program that delivers optimized energy consumption of data centers without compromising the economic value or productivity data centers deliver.

FUTURE DRAFTS AND SPECIFICATIONS:

Clarify Server metrics and features to achieve Energy Efficiency in the Data Center

- Replace Idle with a focus on the primary metrics of energy efficiency for the datacenter.
    - Capacity assessment such as peak compute and peak power
    - Efficiency rate metric such as productivity for incremental energy
    - Load balancing such as ability to adjust energy to load changes.

- Establish minimum feature set bar for compliance. Server power management features are far more likely to help reduce data center power consumption

- Consider listing power management or power saving feature sets desirable for an ENERGY STAR Server, such as:
    - Capable of Power Saving Function
    - Capable of Running Virtualization Hypervisor
    - Capable of Power Capping
    - Capable of Power Management

Provide a vision of the Data Center Efficiency for the Future

- Power Performance Scalability
- Integration and increase capacity model
- Allow IT vendors to choose from a collection of benchmark tests, so as to properly reflect the appropriate workload vs. system architecture in their submission efforts

Use the ENERGY STAR for Servers Program to recognize / reward server products that:

A) Consume power proportional to the useful work they produce (A.K.A. “power scaling”)

- Not solely low idle power
- Autonomous (hardware-based) or Operating System directed power management are mechanisms to achieve this scaling

B) Cooperate / actively participate in the higher level energy conservation solution

Support remote manageability capabilities:

- Report power consumption
- Report inlet temperature
- Respond to directives to cap power consumption

the green grid™