December 5, 2008

Dear EPA,

Please find enclosed Intel comments and feedback on the Energy Star for Servers Specification Draft v3.0, dated 11/4/08.

Intel remains committed and supportive of the US EPA's efforts to define energy efficiency goals and targets across the spectrum of computer products including the current proposal for Energy Star for Servers. However several significant concerns previously voiced by Intel and major manufacturers persist in the draft specification and they must be addressed before such a program can be successful in reducing energy in enterprise datacenters. Specifically for draft 3, we've summarized these concerns with additional examples.

We continue to work extensively with our industry colleagues in Standard Performance Evaluation Council (SPEC), The Green Grid (TGG), Climate Savers Computing Initiative (CSCI), IT Information Council (ITI), and Storage Network Information Association (SNIA), to target the goals of energy efficiency in addition to supporting the Energy Star for servers program.

If you have any questions please feel free to contact myself or Henry L Wong, henry.l.wong@intel.com.

Sincerely,

Lorie Wigle
General Manager
Eco-Technology Program Office

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Intel appreciates the leadership role of the EPA in driving toward greater energy efficiency in enterprises. Energy Star for Servers is an aggressive program that could be used to harmonize energy efficiency programs world wide if it is written to achieve the stated energy saving goals.

Intel welcomes the opportunity to provide the EPA with the following response to draft 3 of the ENERGY STAR Program Requirements for Computer Servers v1.0 specification. The proposed rating system, testing plans, and initial dataset were helpful in understanding the direction and decisions the EPA has taken for this program. Based upon draft 3, we believe there are high risks and potential negative energy efficiency consequences with the current EPA proposal. These should be addressed to ensure the integrity and energy savings opportunities in the program. The comments listed below are specific to draft 3 and should be considered as additional to Intel's feedback on the previous drafts.

As with our feedback on previous drafts, the response is organized per section. We've also included a general commentary reflecting the updates and changes in draft 3. We would like to continue our practice of having the opportunity to review these comments with you and the extended EPA team to answer any questions you may have.

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Overall Summary

A system idle-power-only metric, unqualified by productivity of computer servers, is a serious flaw in the specification. Intel believes this direction creates a disincentive to progress further in improving energy efficiency in the datacenter. We believe that productivity or a proxy of performance is a necessary part of an energy-efficiency metric. If the EPA insists on the current idle-power-focus, Intel highly recommends:

1. Change the categorization to be based on system capabilities (i.e. based on sockets not installed processors)
2. Change the idle limits on the 4S category to accurately accommodate power budgets for system features.
3. Change the compliance procedures to reflect how computer servers are procured and deployed to data centers. The process should specifically determine a way to accommodate blade system alternatives.
4. Revise timelines that accommodate accuracy in the dataset (and its interpretation). We believe a mid-Q2'09 specification release and an effective date of late-Q3'09 to be more realistic.

The concern is particularly significant with highly configurable 4 socket (4S) systems, which serve as the primary vehicles for consolidation and virtualization. These 4S systems represent a classification of systems that explicitly support the architectural features, RAS* attributes, expansion, and application growth ideal for many server consolidation and virtualization efforts– enabling many lower power servers to be replaced by a single system. Although these systems consume more idle power, they are ideal for maximizing energy efficiency through server consolidation (whether the systems are fully configured or partially populated by processors). As seen in the dataset provided by the EPA there is a wide range of idle power in this category. The wide range of idle power corresponds to the varied configuration afforded by these highly customizable systems. Architectural and statistical analysis confirms the need to treat these systems as a separate category for their system capability and not necessarily the number of installed processors.

*RAS: Reliability, Availability, and Serviceability

Blades remain a category which offers an energy efficient alternative in many applications compared to traditional rack based servers. Though a metric may not be ready for these systems, these systems should be made available as an alternative in RFQ (Request for Quote) processes that require Energy Star systems.

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

The data set appears sufficient in number but appears to be missing key attributes to allow for groupings and adders. The accuracy of each of the data points (systems) and the details appear to contain errors and/or inconsistent entries. We appreciate the EPA team's responsiveness in correcting the issues in the dataset brought to your attention thus far. We've also conducted a statistical review and a computer architecture confirmation on the existing EPA dataset. The preliminary statistical analysis confirms a few systematic trends that were not corrected for. For the 4S configurations, the wide distribution of idle power appears to correspond to configuration options in this category. We were unable to make an accurate determination of power budgets to correct for this, due to missing attributes on power supplies and system features. The wide range is architecturally explainable given the high level of configuration options in this group. If the EPA insists on the flawed approach of using system idle power limits even in these highly configurable systems, we highly recommend correction for these trends in the data set by means of adders or adjustment to the proposed limit formula.

It seems clear from the notes regarding “as-shipped” and “labeling”, that there may be some misunderstanding on the process for requesting, configuring, purchasing, installing and maintaining these systems. These processes are significantly different for enterprise class servers as opposed to a client device, such as a personal computer (PC). We highly recommend that the procurement, configuration and deployment processes, be reviewed with system manufacturers and users to determine how best to integrate the energy star program.

We also recommend that the schedules be revisited not only for Tier 1, but, also the timing for Tier 2. The aforementioned data analysis on the existing data set and end of year schedules, already make the Feb 1, 2009 specification release target unrealistic. The market lifecycle cadence on servers is 3-4 years [1], which defines the timeframe for the creation of new efficiency measures. A full specification change 1-2 years after Tier1 is not practical as the industry would not be able to realize any changes in the design or performance to energy star criteria. We recommend that after a realistic schedule is determined for Tier 1, that a Tier2 specification be scheduled 3 years later (e.g. 2012)

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

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Review by Section
Section 1 & 2: Server Definition and Eligibility

Server Definitions
We firmly believe based on the server architecture, configuration and the July 2008 demonstration data that server categories should be separated by the possible configuration and capability of that system. The configuration options such as how many processors or other installed features determine the limits within each category. This method should not be confused with requiring the testing for every system configuration. Testing compliance, categorization and the limits within each category should be kept separate. We recommend that the following be included in the server definition:

For each server category (1S, 2S or 4S),
+ All processors must have access to shared system memory AND
+ All processors installed must be capable of being independently visible to a single OS or hypervisor

Blades
Blades need to be part of the server specification. We recommend creating both a purchasing logistics alternative and a delay in the specification release to accommodate blade systems.

Enterprise Server Purchasing and Integration
The notes section in draft 3 explicitly calls out “as-shipped” and Energy Star® identification method (labeling) that is inconsistent with how enterprise servers are defined (configuration requests), purchased, configured, installed and managed. Unlike client devices, not only are systems meant to be integrated into a pool of computing resource, entities that conduct the activities above are generally not the same “end user”. Take for example the desire to use a physical label to “assure the end user of the Energy Star compliance of the system”, which end user are we considering? If this is intended for the entity purchasing or defining what is to be purchased, the existence of a physical label will generally not be able to provide any confirmation, since those entities may never actually see the system. In fact, physical labels limit cooling design options impacting the performance of the system. A semi-permanent or removable label may actually cause maintenance issues as they may find their way into the ventilation systems. In the case of inventory assessments, the confirmation is normally done electronically referenced by a serial number or other numerical code.

Even more problematic, given this market process, is EPA’s desire to test individual systems with specific “default” settings, OS, and configurations “as-shipped”. The actual feature settings, OS image (to be consistent with the computing pool), installation, and use/management, are established well after the systems are shipped. Outside of recommendations and tools (software and hardware), system manufacturers may not have the visibility or ability to ensure “default” settings or final configurations, as opposed to simply testing systems capabilities. We believe the EPA’s goal is to emulate what the system configuration would be as installed in the data center. If that is the case, we would recommend working with an industry organization such as the Green Grid, IT Industry council, or Climate Savers Computing Initiative to change the Energy Star compliance requirements of how systems are to be tested or identified to better reflect the process indicated.

Section 3 Efficiency Requirements for Qualifying Products

Section 3a Power Supply Efficiency

Intel recommends the EPA to continue its engagement with Climate Savers Computing Initiative to resolve efficiency limit issues with low load efficiency, redundant power supplies, power measurement changes due to input line voltages, and efficiency differences between high power PSU and those less than 1200W. We will continue to apply our efforts in the harmonization around CSCI's common targets and methodology.

Section 3b Idle Power

Use of platform idle power unqualified by system compute capabilities is inconsistent with computer server energy efficiency trends for data centers. While we observe data center professionals maximizing datacenter efficiency by regularly optimizing the server compute capabilities to the energy consumed, the optimization to idle power creates a disincentive to enable the very technologies needed to achieve those efficiencies.

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		
	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)	

Figure 1. Energy Efficiency Impact

The graph in figure 1 exhibits a comparison between idle power and the 3 primary vectors of computer server energy efficiency considerations (for an IT manager):

- Efficiency Rate – compute per watt
- Peak Capacity – Provisioning the needs of the data center
- Dynamic load balancing – ability to respond to changes in demand to lower energy requirements when needed.

Please note on the Enterprise Energy Efficiency side of the chart that information listed in the boxes are examples   of idle range and types of metrics employed. Indeed, the variety of metrics and in particular the changes from 2 socket to 4socket systems are reflective of the difficulty to settle on a single metric to cover all categories, in addition to all 3 vectors. Though the metrics differ across categories and efficiency vectors, idle power contributes little to aid in the decisions on improving efficiency.

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

The range of system idle values is approximate and the dataset already demonstrates a range wider than displayed on the chart. The range does however, highlights the risk in selecting a lower idle system rather than a more efficient one.

Optimizing for idle power will actually compete with the 3 efficiency vectors, and result in purchasing lower capable machines with as few features as possible. The increase number of machines to achieve the provisioned workload may actually slow consolidation and power management feature adoption, just to meet Energy Star requirements.

In the following example, we used a production 4 socket system, and just increased the memory to investigate the choice (per draft 3 proposals 8GB and 16GB are considered the same without adders or compensation).

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.

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

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

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

If One Must Use Non-Performance Qualified Idle-Power

If EPA insists upon using the flawed approach of system idle power, the inaccuracies and deterministic trends in data set must be addressed before setting the limits. This is particularly true in the 4S configuration where the wide range of configuration options also drives a large range in idle power. The dataset indicates an explicit need to establish a baseline set of hardware configurations for each of the categories of systems (1S, 2S, and 4S).

For the 4S server classification, redundant power supply (i.e. "high availability") adder should be included similar to the 2S category. The reason is 4P servers are even more configurable than a 2P so not having adders does not make sense. A key concern in the dataset is that not all the 4P data points submitted to the EPA have redundant power supplies, so this artificially lowers the 4P idle power limits. In addition, please note that 4P server data points submitted on the HP DL580 server have 2 power supplies as standard and 4 power supplies in a high availability configuration.

In the "high availability" category, we believe omission of a redundant fan adder is problematic, as these fans can contribute up to 15W and ensure server availability under challenging environmental conditions (external to the server). These internal cooling features become more critical as the industry advocates raising temperature limits in the datacenter for overall efficiency.

Existing Dataset

Our summary of concerns based on investigation of the existing dataset:

- 1S
    - The proposed idle power limit may be artificially low due to the inclusion of what would normally be considered "small scale servers", per Energy Star computer specification v5.0. These form factors should not be included in setting the limits for the server energy star specifications.
    - Recommend that the EPA query 1P data submissions to ensure the data points are solely enterprise class servers
- 2S
    - Clarification requested- recommend that 2S standard should be <= 16GB. Consistent with the data. Consistent with the architecture.
- 4S
    - Incorrect configurations:
        - Index 108 (142W) is a 2x 2S system
        - Index 109 (214W) OEM confirmed 1S was mis-labeled 4S
        - Index 112 (281W) has only 4GB system memory. EPA approved removal
    - Idle power targets for 4S should be reset based on corrections to the dataset.

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Statistical Observation in the Dataset
EPA's stated desire is to capture observable trends in the data set across systematic characteristics in a simple relationship. This is a challenging goal given the broad range of complexity, computing capability, and redundancy of 4S server configurations.

For 4S systems in particular, there exists a large set of configuration options in memory, redundancy, storage, and processing capability that are needed to apply to customer requirements. To find any variance in system energy consumption independent of the widely varying factors requires thorough analysis. Redundancy and high memory requirements will increase system idle power, so it especially important to capture trends correctly if undesired consequences (such as incentivizing the purchase of less energy efficient systems) are to be avoided.

Based on our analysis of EPA's Draft 3 rules we believe that Draft 3 partially but not fully identifies trends in the 4S data. We propose an additional analysis, "Rev B," which more evenly balances energy star configurations across end user requirements thought still does not compensate for all systematic trends in the 4S data.

A histogram of the existing dataset shows a skewed distribution and a trend analysis that is indicative of, but not conclusive evidence for, systematic trends in the data.

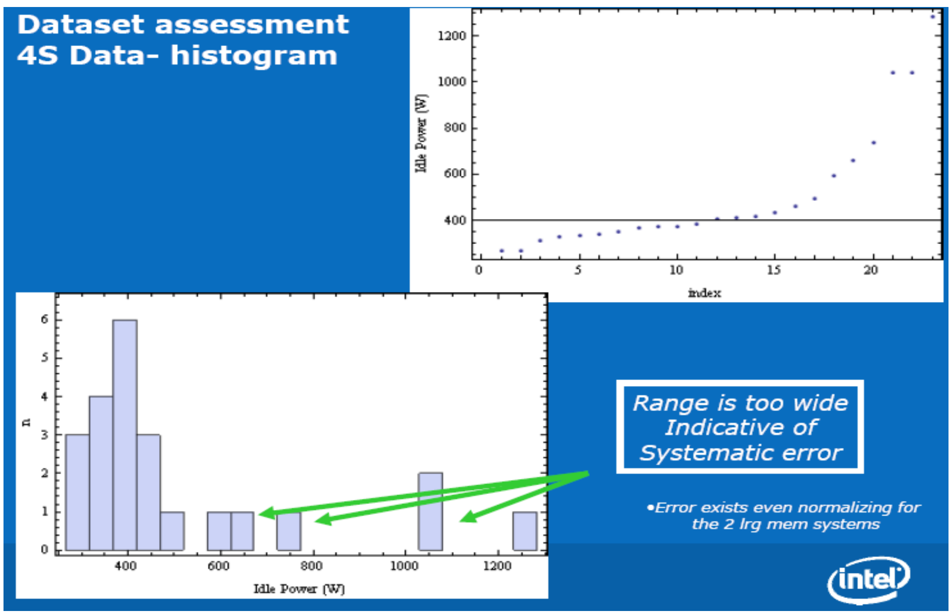
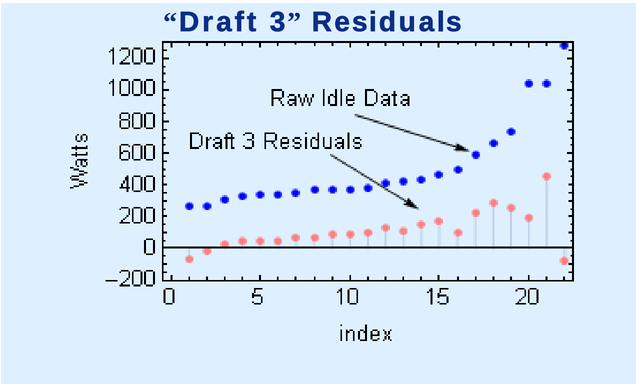
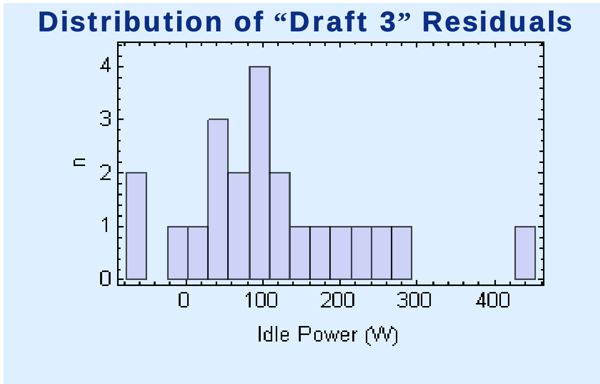


Figure 2. Histogram of EPA's 11/14/08 dataset

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Alternative Statistical View of the Dataset
EPA's Draft 3 Formula "is based on observable trends in the dataset; and represents approximately 25% of models in the dataset, across all bins and system characteristics (e.g. # of processors and hard drives, installed memory, redundancy level, etc.)."





The analysis shows the "residuals" based on the EPA Draft 3 factors for memory and HDD.

It is clear that the residuals still exhibit strong systematic trend, suggesting potential "hidden" variables. The skewed shape of the distribution (corrected) is still strongly skewed.

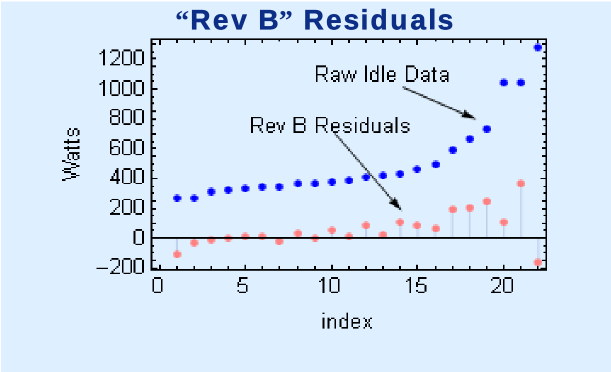
We conclude that Draft 3 (formula) does not fully reflect observed trends.

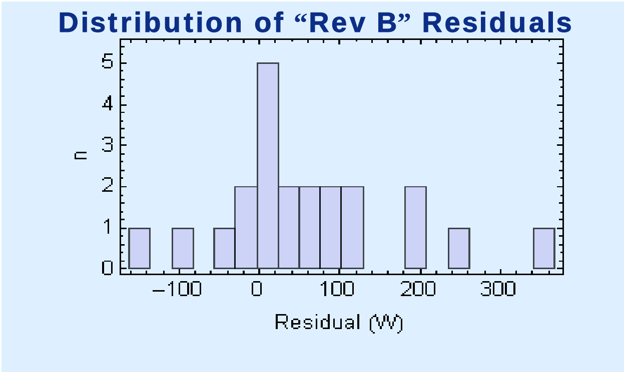
Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Alternative Formula (Rev B)

We have undertaken an analysis for additional systematic trends based on the factors available in the EPA data set (with the "outlier" systems mentioned previously removed). Note that analysis could not be performed for some factors known to be important in system configuration choices, such as storage controllers, I/O device type and number, etc.

The method takes into account EPA's allocation for HDD and Memory, but uses a "design of experiment" approach to understand the significance of additional variables.





(Considered preliminary due to limited information on the data points in the dataset, see recommendation on additional information needed on the data points)

The Rev B formula, whose results are shown in the figure above, use EPA's methodology and retains the proposed allocations for HDD and Memory, but includes an allocation of 22W per PSU.

The PSU parameter falls out of the data quite naturally. The fit parameter of 22.1 W has a calculated error of less than 2% This indicates at least one additional trend the data which should be corrected.

Based on the residual analysis above, we believe other systematic factors are present in the data. Unfortunately they could not be adequately analyzed because the data on system configurations was not available. We believe a corrected idle power spec for 4S systems would account for these additional factors.

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Section 4 Test Criteria

As noted in previous comments, many computer servers are shipped without an operating system. In fact, the final configuration, OS image, system settings, and management tools are installed well after the hardware has been shipped from the system manufacturer. We recommend testing models based on their full scale capabilities, as a means by which to rank systems. The procedures should be reviewed and evaluated with the industry organizations previously defined (TGG, ITI, or CSCI) as part of the process of modulating the Energy Star practices to accommodate the procurement and integration methods for enterprise servers.

Testing system power levels are very dependent on efficiency and conversion that occurs at the power supply. As observed with the idle power limits, the number and type of conversion has a direct affect on the value obtained at idle. The input line voltage has a direct affect on these values and the resulting difference in platform power can not be controlled to provide a similar rating. We recommend a solution be derived with CSCI, such as settling on a fixed, worst-case line level to run the compliance testing.

48Vdc systems

The bulk of the 48Vdc systems are actually -48Vdc systems. Despite its name, -48Vdc systems almost NEVER (<1% of the time) operate at -48Vdc. The ANSI standard for -48Vdc is ATIS-0600315.2007, and this states a nominal voltage to be -53Vdc (this is a compromise between the two major battery types: VRLA-based systems typically operate at -54Vdc, while flooded (wet cell) battery systems typically operate at -52Vdc). There is an appreciable difference in operating efficiency of some equipment at -48Vdc vs. -53Vdc, and we encourage the industry to optimize their power converters for the voltage that is typically used rather than at some arbitrarily different voltage used for comparison. For the purpose Energy Star testing, we recommend a test voltage of -53Vdc +/- 1Vdc for "-48Vdc" systems.

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124