December 3, 2008

To: Rebecca Duff
ICF International

CC: Andrew Fanara
United States Environmental Protection Agency

Arthur Howard
ICF International

Re: Hewlett-Packard Response to Preliminary Draft 3 of ENERGY STAR® Program Requirements for Computer Servers

From: Hewlett-Packard Company, Enterprise Storage and Servers Business Unit

This document may be published on the Energy Star website.

Hewlett-Packard (HP) welcomes this opportunity to provide comments on ENERGY STAR® Program Requirements for Computer Servers (hereafter in this feedback document called "Energy Star for Computer Servers") — Draft 3. HP is proud to continue its long-standing association with the Energy Star program. The comments and issues in the Draft 3 document fall into two categories and are covered in the two major sections of this review:

1. Draft 3, Energy Star Partner Commitments
2. Draft 3, Energy Star Product Eligibility Requirements

Several key issues are detailed on the pages below, but here is a brief summary of six critically important points of feedback:

1. Blades must be included in the Energy Star v1.0 specification, or else blade servers will be at a competitive disadvantage, for an unpredictable length of time, to servers that are not excluded from the Energy Star program.
2. HP objects to adding Energy Star labels to servers, when the criteria don't allow for certification of all SKUs to be based on the certification of a base model.
3. Servers cannot be tested "as-shipped." Servers are shipped with no operating system.
4. Idle power criteria (if they cannot be removed) should have realistic power budgets for each and every server feature demanded by customers, demanded by their facilities environments, and/or demanded by their applications; including scaling for computing performance (capacity), memory capacity, storage capacity, number and type of optional I/O cards, and a multitude of RASM features.
5. The dividing line between low wattage and high wattage power supplies should be 1200W and not 1000W, meaning that the efficiency and power factor requirements should be the same for all power supplies 1200W and lower.
6. Systems with small and/or "right-sized" power supply solutions should not have to meet measurement accuracy, efficiency and power factor @ 10% power supply load.

December 3, 2008

1. Draft 3, Energy Star Partner Commitments

1.1. The requirement in lines 33-37 stating that "The ENERGY STAR mark must be clearly displayed on the front or back of the product…" is in direct conflict with statements made by Andrew Fanara at the July 2008 Redmond, WA stakeholders' meeting. Hewlett-Packard and all other server manufacturers in attendance were clear in their opposition to this requirement at the stakeholders' meeting, and Andrew's statement led us to believe that this issue was closed.

Datacenter workers are the only ones that will briefly see the label before either removing them or putting the server into a lights-out data center environment. The label provides no informational benefit for data center workers, since deployment personnel do not make the decision to purchase Energy Star servers. Energy Star is a selection criterion for the CIO and for the procurement department. Energy Star information can be efficiently coded in a more useful way into a database that can leverage the unique serial number, bar code or electronic identity code that every product must already carry. Energy Star information from a database referenced by these non-label sources could still be valid and auditable once the Energy Star sticker is tossed in the trash or becomes invisible in a dark data center.

HP reiterates past objections to the cost and lack of usefulness for permanent labels. There is no room on the front or back of servers for physical Energy Star labels and their presence adds both materials cost and labor cost, with no positive benefit to either the Energy Star program or to end-users.

HP would approve of an Energy Star labeling program only if certification is based on testing base models and if (within reasonable judgment) optional features are not factored into whether an SKU is compliant or not.

1.2. With the EPA's stated goal for ~25% of shipping servers being able to meet the requirements of this Energy Star program, actual shipments will fall far short with the present pass/fail criteria. The current pass/fail criteria do not provide power allowances for the enterprise-class service levels and Reliability, Availability, Serviceability and Manageability (RASM) feature needs of the government data centers around the world that are compelled by law to either purchase Energy Star servers or to buy servers that are excluded from the Energy Star specification.

2. Draft 3, Energy Star Product Eligibility Requirements

The following is a compiled list of HP comments, referenced to the line numbers listed in the document. Comments are broken into two general lists. The Substantive Feedback section lists substantial changes to the approach documented in Draft 3, while the Editorial Feedback section details issues that are important to clarify the intent of the document.

December 3, 2008

2.1. Substantive Feedback

2.1.1. Blade servers are excluded in Draft 3, but given their energy efficiency advantages, HP proposes a slight delay in releasing v1.0 until SPECpower_ssj defines the idle power test procedure for blade systems. To prepare for the addition of blades, the methodology for comprehending blade servers, blade enclosures and other options that reside in blade enclosures (switches, routers, management controllers, storage, et al.) needs to begin its definitional process in Draft 4. The imminent release of a new revision for SPECpower_ssj, that supports blades, makes a March 2009 (or later) release of Energy Star for Computer Servers v1.0 a better target date.

Since an enclosure full of blade servers often provides performance and RASM equivalent to an equal number of rack servers, it would be fair to let blade server solutions compete, on a per-server basis, directly with rack-mount server pass/fail power criteria. The power associated with blade server solutions are often lower power (per server) than their rack-mount server equivalents. This enables blade servers to compare directly to rack-mount server thresholds, but the certification of blade servers needs to be contingent on the enclosure that it installed in. E.g. an Energy Star blade server SKU must state which enclosure(s) it was tested in to achieve Energy Star compliance.

Blade enclosures also house other devices that are typically housed in racks, and typically are separate from rack-mounted servers. Network switches and routers, Fibre channel switches, I/O virtualization solutions, multi-server management devices and disk storage arrays are typically external to rack-mount servers, but may be plugged into blade enclosures. While the Energy Star for Computer Servers v1.0 specification may not want to address the pass/fail criteria for these non-server devices, the presence of these devices in a blade enclosure reduces data center power and shouldn't preclude a blade system from achieving Energy Star compliance.

HP supports having blade enclosures meet the power supply criteria defined in this specification (with changes and improvements noted in subsequent sections).

2.1.2. Lines 455-456: With the Draft 3 rules, using idle power as a pass-fail criterion is flawed. Current rules promote end-user buying behaviors that lose sight of the big picture and require more data center total energy use to deliver an equivalent computing capacity. By placing no value on the performance of servers, there is no ability for a highly capable, but higher wattage server to replace multiple low performance servers whose aggregate energy-use would actually be worse.

Pushing for artificially low idle thresholds will cause the unintended consequence that only a few de-featured, low-performance models will be able to meet the fixed idle power thresholds described in Draft 3.

2.1.3. Lines 455-456: A server with PCI or PCI Express I/O options is a key part of the definition of all rack-mount servers that are included in Draft 3. A server without optional I/O cards is a "Server Appliance", which is excluded from Energy Star for Computer Servers v1.0. Given the large power variability of PCI and PCI Express I/O functions, the power associated with those options should be ignored and optional I/O card power should be subtracted from Energy Star idle power measurements. Any configuration which meets Energy Star without the PCI or PCI Express I/O options installed should be considered compliant with one or more PCI or PCI Express options installed. This methodology will more realistically support the inclusion of all servers that require PCI, PCI Express or other types of installed option cards. Without a methodology for including richly diverse I/O cards, procurement can avoid having to buy Energy Star servers simply by requiring a PCI card that prevents all vendors' servers from passing Energy Star.

2.1.4. Lines 455-456: Idle power thresholds are meaningless if there is no correlation to a measurement of peak performance on a valid benchmark for that server type and application environment. Servers with more performance capability should get proportionally higher idle power thresholds.

2.1.5. Any single idle threshold is equally meaningless without a defined set of specific Reliability, Availability, Serviceability and Manageability (RASM) features. HP supports the general concept of having different pass/fail criteria for low redundancy and higher redundancy servers. However, lines 228-234 are inadequate to describe the rich diversity of features and power levels that are possible while delivering a wide range of RASM features.

RASM features in servers cannot be provided for zero watts of idle power, and the amount of RASM features varies widely by the applications that run on the server and by the end-user environment. RASM features may include RAID disk controllers (and the required number of hard drives to deploy the chosen RAID level), hot plug hard drives and support circuitry, system management controllers, hot plug I/O cards, ECC memory, hot-plug and/or redundant fans, hot plug and/or redundant power supplies, et al. A simple catch-all category for High Availability Server is inadequate. Power adders for specific RASM features are needed and HP would be happy to help quantify those adders.

2.1.6. Line 213: Need to define (if for no other reason than to exclude in v1.0) Computer Servers with no internal (AC-DC or DC-DC) power supplies. These servers either receive low voltage direct current (DC) power directly from an external source or provide only minor voltage regulation and conditioning of a low voltage external source. Much like blades, this is another possible way to deliver low-loss power to a rack full of servers. This type of rack-level power supply and power distribution is a valid way to efficiently deliver power to a group of servers.

December 3, 2008

2.1.7. Lines 365 and 366: The 1000W choice in both Draft 2 and Draft 3, for a division between high wattage and low wattage single-voltage power supplies was arbitrarily chosen from the limited data received. This should be increased to •1200 Watts and >1200 Watts. From a technical perspective, 1200 watts is the approximate current limit for the power supply connector. Also, typically, larger power supplies than 1200W tend to be designed for 48 VDC outputs rather than low voltage 12 VDC outputs, so there is a real, scientific basis for choosing 1200W as the dividing line.

2.1.8. Line 365: Allowing multi-voltage power supplies to have an easier set of thresholds, without bounds for wattage, provides incentives for server manufacturers to move from single-voltage power supplies to less-efficient multi-voltage power supplies. An upper bound for wattage (e.g. 750W) should be added and larger power supplies should be forced to comply with the single-output power supply thresholds.

2.1.9. Lines 365-366: If the 10% load requirements are retained, then (similar to multi-voltage power supplies) the 10% pass/fail threshold for single voltage power supplies should not be required if the power supply is placed into a server that has the ability to prevent power supply loading from going below 20%. This prevention of <20% loads can be done by: 1) "right-sizing" and/or 2) via power supply technology that shifts loads and turns off unnecessary power supplies and/or 3) by shipping without N+N power supply redundancy.

2.1.10. The proposal made in lines 434-436 needs to be reflected in Tables 1 and 2 on lines 365-366. Removing the 10% load requirements for smaller power supplies is a prudent thing to do, since one of the reasons for smaller power supplies is to "right size" the systems.

2.1.11. Lines 660-678: HP suggests that it be defined how to amend the "Generalized Test Protocol for Calculating the Energy Efficiency of Internal AC-DC and DC-DC Power Supplies," without revising Energy Star v1.0 for Servers.

Consider also defining the means in the v1.0 Energy Star specification for adding testing and qualification procedures for substantially different power supply types. In the near future better power supply efficiencies and reduced total power savings may be possible with power supplies that cannot be tested using the currently defined procedures.

2.1.12. Line 677: As a specific near-term example to 2.1.11, with the growing popularity of 380-400 VDC power distribution, it makes sense to be able to include support for systems with 380-400 VDC DC-DC power supplies in v1.0.

December 3, 2008

2.1.13. More Idle Requirement Issues:

Table 3: Base Idle Power Requirements

System Type	Idle Power Limit
Single Installed Processor (1P – All systems)	60 Watts
Two or Three Installed Processors (2P & 3P)	
Standard Availability Systems	151 Watts
High Availability, Low Installed Memory (<16GB) Systems	169 Watts
High Availability, High Installed Memory (≥16GB) Systems	221 Watts
Four Installed Processors (4P)	271 Watts

Table 4: Additional Power Allowances for Extra Components

System Characteristic	Additional Idle Power Allowance
Second Hard Drive	15 Watts
Additional Hard Drives Over Two	8 Watts per Drive
Additional Memory over 32 Gigabytes	2 Watts / GB

2.1.13.1. All systems need a distinction between standard availability and high availability, not just 2P servers. HP has high-availability features in 1P servers, and a wide variety of RASM features available in 4P servers.

2.1.13.2. The EPA misinterpreted data that HP and others supplied. HP supplied data for systems with de-populated processor sockets as a means for understanding how to give power budgets for additional processors; NOT to directly compare (e.g.) a four socket system that has only two installed processors versus optimized 2-processor systems. Without a means for RASM feature power adders, this is unfair.

2.1.13.3. Installing a third processor is not a zero power adder. A power increase is required for both the processor and for the higher functionality chipset that, even with 3 processors installed, must be able to support 4 processors and higher RASM.

2.1.13.4. Eliminate the <16GB and >16GB distinctions and either specify a required DIMM density, or go solely to a Watts per Gigabyte allowance. If the distinction is not eliminated, it must be extended to all servers.

2.1.13.5. All PCI cards require power adders. The specific I/O function determines the power needed. Power allowances need to be made, OR to keep things simple, just allow PCI card power to be ignored by the Energy Star configuration and let otherwise Energy Star compliant systems be unaffected by PCI card idle power.

2.1.13.6. The hard drive and memory power allowances send a conflicting message. 2W per GB favors using very high capacity memory, where allowing only 8W per hard drive limits capacity density. The Terabyte per watt ratio is much lower on an 8W drive than on a 12W drive.

2.1.13.7. It is not clear if the high availability servers actually have redundant power supplies installed or if they are just capable of having redundant power supplies.

There are many possible combinations of availability features to comprehend:

- Redundant power supplies (N+1 or higher)
- Redundant power grids (2N)
- Redundant fans (N+1, 2N, …)
- Error detection and correction (ECC, CRC, chipkill memory…)
- RAID disks
- NonStop error tolerant server solutions (100% availability)

2.1.13.8. It is never stated if the "Idle Power Limit" in Table 3 and the "Additional Idle Power Allowance" in Table 4 is DC power out of the power supply to the server, or input power into the server's power supply. Presumably these are input power measurements (AC or DC depending on the power supply type).

2.1.13.9. Nomenclature in Table 3 is inconsistent with the nomenclature on lines 228-234. One says "High Availability Systems" and "Standard Availability Systems", while the other uses the terms "High Availability Servers" and "Standard Availability Servers".

2.1.14. Idle power level implementation suggestions:

- Establish a minimal feature set and assign a DC power requirement; then add the DC power of optional devices that are installed. Sum the power and divide by the established Energy Star power supply efficiency level @ 20% load to estimate the AC power at idle for that platform.
- LFF HDD (3.5" hard disk drive form factor) – 12W DC
- SFF HDD (2.5" hard disk drive form factor) – 8W DC
- DRAM Memory – 2.5W DC per GB
- System Overhead – 20W DC per socket available
- Installed processors – 15W DC
- Management processor – 10W DC
- GB NIC – 4W DC per NIC
- 10GB NIC – 12W DC per NIC

2.1.15. Lines 724-727: Servers cannot be tested "as-shipped." 1) Servers are most-often shipped without operating systems and software needed to provide the lowest power usage. 2) Unless reasonable power budgets are assigned, optional equipment like PCI Express I/O cards should be exempted from the testing and certification.

2.1.16. Input Power Measurement Issues:

Standardized Data Measurement: All Computer Servers must have the ability to provide data on input power consumption in Watts, inlet air temperature, and utilization of all processor cores during normal operation.

Required Accuracy: All measurements should meet the following accuracy levels:

Input power measurements: +/-10% accuracy

Utilization measurements: +/-5% accuracy

Input air temperature measurements: +/- 2° C

2.1.16.1. Power supply internal circuitry can measure input voltage and input current accurately within 10%. Only apparent power measured in volt-amperes (VA) can be reported within 10% at a reasonable input VA. For real power measurement (VA*PowerFactor), power supply should be operating at 50% or higher capacity to achieve >0.9 power factor. Recommend input power accuracy specification be changed to either "10% error when greater than 100VA, 20% when between 50 and 100VA and no requirement below 50VA" or if real input power is required to be reported, it needs to be "±10% accuracy only between 50% and 100% loads."

2.1.16.2. The power meter does not necessarily need to be inside the individual server. There are other devices in the market which can be used to monitor input power accurately. Please remove the input power meter requirement.

2.1.16.3. For any server to report input power within 10% accuracy at a low system load condition, separate power measurement circuitry is required, adding cost and impacting efficiency. If this requirement is not removed entirely, then suggest that it should move out to Tier II since incorporating this into a server takes a design cycle. No current generation systems can accurately report this parameter.

2.1.17. Sampling Requirement issue:

Sampling Requirements: Hardware polling rates of the embedded sensors must meet a minimum of one sample per second. Data should be averaged on a rolling basis of 15, 30 and 60 seconds. All three rolling averages shall be made available for collection.

Measurements should not be required to have a 15, 30 and 60 second average. A running average, which is reset-able by a monitoring program, should be used instead, to allow a monitoring program to determine what polling rate it wants.

2.1.18. Lines 779-829: HP does not want Idle power to be part of a Tier 2 Energy Star pass/fail criteria. Once Energy Star is based on procedures for fairly measuring performance per Watt on a wide range of applications, the folly of measuring idle on a product that isn't supposed to be idle must go away.

December 3, 2008

2.2. Editorial Feedback

2.2.1. Lines 671-676: HP approves of the Draft 3 improvements to enable multi-output power supply testing at 230VAC and to allow both 60Hz and 50Hz testing of all power supplies at 230VAC.

2.2.2. HP approves of the decision to exclude fan power for the purposes of efficiency testing. This is the only fair way to compare power supplies that may or may not have power supply fans.

2.2.3. Miscellaneous Issues from Draft 3:

E. Direct Current (DC) Server: A Computer Server designed to operate with a DC-DC power supply.

2.2.3.1. The Direct Current (DC) Server definition is ambiguous. Most products are designed to primarily operate with AC-DC power supplies, with DC-DC power supplies as options. Direct Current Servers with non-isolated DC-DC power supplies should be excluded. No mention was made where the DC-DC power supply is located. HP plans for servers to be able to accept either AC-DC or DC-DC power supplies.

G. High Availability Server: Computer Server with more than one power supply for the purpose of redundant power delivery (2n or n+1), ECC memory, and a dedicated management controller (e.g., service processor). High availability systems are designed for mission-critical applications and very low downtime

H. Standard Availability Server: Basic Computer Server with a single power supply and few other Reliability, Accessibility, Serviceability and Management (RASM) features.

2.2.3.2. The definitions of "standard availability" and "high availability" servers do not cover the set of all servers. E.g. a server with redundant power supplies and no dedicated management controller is an example of a server that does not fall into one of these two definitions.

N. DC-DC Power Supply: A power supply which converts a DC voltage input to one or more different DC voltage output(s) for the purpose of powering the Computer Server. Any DC-to-DC converters (also known as voltage regulators) internal to the product and used to convert low DC voltage (e.g. 12 V DC) into other DC voltages for use by the individual Computer Server components are not considered DC-DC power supplies under this specification.

2.2.3.3. In the DC-DC power supply definition, the specification needs to specify the meaning of "low DC voltage." One example is given, but it isn't clear what else is included or excluded by this undefined term.

The National Electrical Code, Article 551, section 551.2 defines Low Voltage to be "an electromotive force rated 24 volts, nominal, or less, supplied from a transformer, converter, or battery." As a DC voltage, if

this definition is used, please clarify that the voltage should be between -24V and +24V to qualify as low DC voltage, since "less than 24 volts" could be interpreted to include large negative voltages in the definition.

Is an AC output with a DC offset included? A DC output with a high ripple? Should 48V be included? Should it exclude all but 12 VDC?

Q. Idle: An operational state in which the operating system and other software have completed loading and the Computer Server is capable of completing workload transactions, but no workload transactions are requested or performed by the system (i.e. the Computer Server is operational, but not processing any useful work).

Note: Based on several stakeholder requests, a definition for Idle is now included to support the proposed Idle requirements outlined in Section 3, below. This new definition combines language from the ENERGY STAR Version 5.0 Computer Idle definition and the SPECpower_ssj2008™ Run and Reporting Rules. Manufacturers are required to use this SPEC benchmark when measuring power consumption in Idle. Stakeholders are encouraged to provided feedback on this newly proposed definition. In addition, a definition for the SPEC organization was added for reference.

2.2.3.4. The definition of Idle on lines 298-301 should be where Idle is defined completely, with other places referencing this definition. Lines 716-717 add detail that should be in the definition section. Idle should be defined as the active idle value in SPECpower_ssj benchmarks.

However, EPA does recognize that smaller power supplies are at a disadvantage at low loads because of the impact that fixed losses have on the efficiency of these units at low power draw. EPA shares a common goal with manufacturers of ensuring that high efficiency, properly sized power supplies are incorporated into all Computer Servers. Using smaller, more efficient power supplies will further help to reduce the total energy used by Computer Servers, which EPA supports. For this reason, EPA is interested in proposals on setting requirements that address energy consumption at low loads (i.e., ≤ 20%) without penalizing smaller power supplies. For example, EPA received the following proposal:

- PSUs rated at 750W or lower are exempt from efficiency and power factor levels at 10% of rated load. Furthermore, PSUs rated at 375W or lower must meet 80% efficiency and 0.6 Power Factor at 20% of rated load instead of the normal requirements in Tables 1 and 2, above.

EPA is interested in developing a similar proposal in order to set baseline levels for smaller power supplies to ensure: (1) that these units can qualify for ENERGY STAR and (2) manufacturers continue to consider energy consumption at these lower load points in future Computer Server design.

2.2.3.5. Draft 3 asked for ideas to not burden lower power supplies with the 10% standard. HP wants to propose that any power supply under 750W be exempt from the 10% load requirements.

2.2.4. Draft 3, with a few exceptions previously noted, is well-written and generally without typographic errors. HP is looking forward to a Draft 4 with fewer "Note" sections, more-scalable definitions for pass/fail criteria, and present HP concerns addressed.