

From: Christian Belady [mailto:cbelady@microsoft.com]
Sent: Wednesday, March 12, 2008 5:26 PM
To: Duff, Rebecca M.; Duff, Rebecca M.
Cc: Fanara.Andrew@epamail.epa.gov
Subject: Feedback of Energy Star Specification from Microsoft

Rebecca and Andrew,

I have consolidated all of the comments and issues from Microsoft (below). Please confirm that you received these and let me know if you have any questions or clarifications. We decided to consolidate all of our comments so to minimize overlap. Hope you find this helpful.

Christian Belady

General Feedback

- 1) We approve the use of PSEfficiency though we feel we feel the following should be considered
 propose changes in table to as follow:

Table 1: Efficiency Requirements for Computer Server Power Supplies

Percentage of Rated Power Output	20%	50%	100%
Minimum Efficiency Requirement - Single voltage	TBD	TBD	TBD
Minimum Efficiency Requirement - Multi-Voltage	TBD	TBD	TBD
Power Factor	TBD	TBD	0.9

From PSEfficiency perspective, EPA proposes 10% as too low to be meaningful and 100% as too high/infrequent to achieve in the real world. From last week 2F meeting between Climate Savers and Efficient Power Supplies.Org (owner of 80plus program), we've agreed to test PS at 20%, 50% and keep 100% for power factor consideration and other legacy purposes.

- 2) Performance should only be considered in Tier 2. Performance benchmarks are tricky and the industry will need to work this in the meantime and develop consensus. For Tier 2, we like the idea that they are considering other workloads in addition to SpecPower. In case the EPA will have to specify what workloads are acceptable.
- 3) Need clarification on what configuration is tested and given the Energy Star label. How is the rating affected with thousands of configurations as a result of upgrade, etc. Need to clarify this from both hardware and software perspective.

- 4) We agree that an idle criteria is important as it dictates the lowest point to which we can get the system power. However we are concerned in how the idle criteria is determined. For example, Microsoft feels that a wattage-based idle criteria does not provide the appropriate motivation as it does not capture information on the performance of the system. However an idle criteria that is based on a percentage of the power draw at maximum load (e.g. System X uses 25% of the maximum power draw at idle) would provide motivation in a constructive direction. Also, we feel that idle is not the only criteria, it needs to be paired with other information (e.g. the number of processor packages, number of internal system drives, memory amount and type, etc.). A crisper definition of idle needs to be defined before we can provide detailed feedback and opinion. Idle may make more sense for Tier 2.
- 5) Need clarification of what is informational and what is being used for the metric and rating. We are also concerned about the binary nature of Energy Star.... would a rating system be more appropriate?
- 6) Power and Temperature Measurement Requirements - Microsoft feels that real-time data on power consumption is very useful. We also feel that this information must be provided in a standardized mechanism in-band to the operating system. This enables the OS to gather and correlate power consumption with other metrics on the system. Having this information available out-of-band (i.e. through a service processor interface) is also useful. We support working with the DMTF standards for reporting this via WS-Management. The minimum rate of power/temp data sampling should be specified. Something in the range of 10-15 seconds should be considered.
- 7) Server definition needs to be crisper. Perhaps just a few salient features should be listed without specifics number of Dimms, Sockets, etc. By listing specifics you may constrain innovation. Perhaps you should capture what servers do. Should consider looking at workloads for defining what is a server.... Some examples: OLTP, Web Services, HPC, etc.
- 8) We are assuming the Tier 1 starts in early 2009 and lasts at least through to the end of 2009 and will be obsolete when Tier 2 comes out.
- 9) We agree that Data Sheet should include the following possible information fields (bold lettering needs to be addressed):
- a. Server manufacturer, model name and number
 - b. Configuration Information
 1. Form factor (e.g., 1u, 2u, blade chassis, etc.)
 2. Available processor sockets
 3. Processor information (model number, speed, # of cores, etc.)
 4. Memory information (memory types, # Dimms, Dimm Size, etc.)
 5. Power supply information (#, size, efficiency)
 6. NIC information (#, type, integral vs add on and speed)
 7. Hard drive information (#, speed, size)
 8. Installed operating systems or those used for testing (**Need Clarification what this actually means**)
 9. Other hardware features / accessories (**optional**)

- c. Available Power Management Features (**Need Clarification what this actually means**)
- d. Virtualization Capability (**Need Clarification what this actually means**)
- e. Power and Performance Data (**This should only be Tier 2**)
 - 1. Idle power from SPECpower output
 - 2. Maximum power and throughput (using manufacturer selected benchmark)
 - 3. Estimated yearly kWh and \$ consumed (based on an agreed upon set of assumptions)

Other Feedback

- 1) Not sure what EPA means by "servers must come equipped with power management and virtualization". Why virtualization??? Ultimately Tier 2 will test this. If you can't measure the level of virtualization and power management this could be difficult to validate for Tier 1.
- 2) Why does server definition not include single socket? EPA should encourage the use of single socket server especially with low utilizations that exist today.
- 3) Agree with focus on volume servers, have large impact for small changes
- 4) Agree with focus staying with AC-DC conversion at this point and not including DC-DC conversions on motherboard
- 5) Performance benchmarks should be included for Tier 2 and SpecPower is a good starting point
- 6) Will need to specify conditions for all tests.
- 7) Need to decide what to do with fans that are part of the PS. This should not be considered part of the PS efficiency to allow apples to apples comparison
- 8) We like the Note in Section DD within the box on page 9. Please incorporate this into the SPEC.
- 9) We would like clarification on the OS permutations listed in item #2 and how that relates to measured data points as in item #5. For example, we think that idle power may differ depending on which OS is on the system when tested. Similar problems exists for maximum power and throughput
- 10) How do they define an Enterprise Class OS? What are the requirements?
- 11) What does certification for an OS mean?
- 12) Requiring hard drive on the server is the wrong approach. Consolidating storage centrally and removing disks from servers is a more power efficient direction.
- 13) **We are not sure about the relevance of the server definitions/classifications but here are some comments**
 - o I think discounting 1P systems is a big mistake. With multicore it makes sense to use 1P servers in the data center for the typical deployment of one workload per server'. 2P servers are typically very underutilized today. With 1P you save power without limiting workload performance. Server vendors are creating 1P rack mount server SKUs as a result of this market trend. They must include 1P servers.
 - o The currently defined classes are focused on today's hardware and don't take into account important directions. They are poorly defined for the Tier 1 timeframe. 4 socket servers will have 8 cores in the Tier 1 timeframe, these aren't small servers the E-4P category covers 99% of all servers shipped (maybe 8 million). Small is the most relevant category for power issues in data

centers. There will be a very small number of servers with 16 sockets shipped in 2009 (few thousand). We propose the following:

- Server classes should be 1-2 socket, 4-8 socket and 8 socket since the number of cores is increasing.

14) Page 9 – E: “Energy Star qualified computer servers must come with hardware power management and virtualization capabilities” ... Strike the word hardware.