

ENERGY STAR Computer Servers: Tier 2 Stakeholder Workshop September 25, 2009

Workshop Notes

This document summarizes discussions from the ENERGY STAR Server Tier 2 Workshop held in Menlo Park, CA on September 25, 2009. Additional details can be found in the ENERGY STAR Tier 2 Server Preliminary Specification and presentations from the meeting. All of these documents are available on the ENERGY STAR Web site:

http://www.energystar.gov/index.cfm?c=revisions.computer_servers

1. Introduction

- Slides:
http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/1_ES_Servers_T2-Intro.pdf
 - Notes:
 - Overview of transition from Tier 1 to Tier 2 presented.
 - European perspective on ENERGY STAR. Tier 1 to be officially adopted in December 2009. Code of Conduct for Data Centres in place since 2008. Input and positions from the EU reported to attendees.
-

2. Program Overview

- Slides:
http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/2_ES_Servers_T2-Overview_and_T1.pdf
 - Notes:
 - Overview of guiding principles of the ENERGY STAR server program. Description of how existing requirements will “roll over” to new specification and how Active Mode requirements will be layered onto existing foundation. Review of anticipated development activities.
-

2a. Perspectives: The Green Grid

- Slides:
http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/2a_TGG_Energy_Star_Server_Workshop_v0p91.pdf

- Notes:
 - Description of The Green Grid and its mission statement. Description of need to provide common language to talk about energy efficiency – between facilities and IT, as well as needs to cover strengths and weaknesses of approaches addressing active performance-energy or idle-only metrics.
 - Goal: maximize "business output"/performance for given power envelope
 - Overview of The Green Grid's input to the ENERGY STAR Server process.
 - Continue with CSCI targets re power supply efficiency
 - Incorporate performance into evaluation tools
 - Allow room for customization
 - Prioritize general purpose systems w/ large sale volume
 - Improve taxonomy and system definition
 - Further develop power and performance datasheet format
 - Adjust reporting requirements based on datacenter monitoring
-

2b. Perspectives: eBay

- Slides:
http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/2b_EBAY_Energy_Star_presentation_v2.pdf
- Notes:
 - "Growth hides all sins." As of 2006, eBay paid more attention to energy (operating costs >> initial costs)
 - Tech refresh allows staying with Morse Law w/ same power footprint
 - PUE not the whole story – a 1.0 PUE can still be large number over large number. Focusing solely on PUE ignores the absolute power use numbers that dictate costs.
 - eBay's wish list for OEM's
 - Track lifecycle green house gas emissions – one standard across the industry
 - More generic performance-cost and efficiency metrics
 - More efficient power supplies at lower utilization
 - Warranties that cover operation at higher inlet and exhaust temperatures
 - Liquid cooling
 - Reduced components on the system, e.g. 1W LED, USB polling, stripped down video cards, SSDs vs. spindles, LEDs, etc. (1 W difference on 50k machines = huge difference)
- Discussion:
 - eBay's Primary metrics are transactions/Watt and \$/Watt – meaning of transaction is application dependent.
 - While eBay's applications are highly specialized and standardized across their datacenters, there is a need for more general efficiency evaluation tools and metrics. For a large user like eBay, this would provide "a foot in the door" to select vendors and solutions that would undergo more specialized testing under metrics closer to the expected applications.

3 Joint Breakout: Active Mode Rating Tool for Server Energy Efficiency

- Slides:
http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/3_ES_Servers_T2-Active_Mode_Efficiency.pdf
- Notes:
 - Discussion of barriers and challenges in implementation of active mode efficiency requirements
 - Under existing performance-oriented benchmarks, mid- and low-range systems' performance not often reported
 - Are certain existing benchmarks good first targets for use in determining efficiency rating?
 - technology-neutral/architecture-agnostic features: agnostic to CPU instruction set, CPU architectures, underlying silicon layout / topology
 - Other barriers to implementation: lots of infrastructure and cost to set up certain benchmarks and test configurations. TPC was noted in discussion as a benchmark requiring significant resources and time to operate
 - end-use scenarios: variety of workloads that stress system in diff ways
 - Software considerations: many benchmarks have numerous software settings that can be tuned to optimize performance under the benchmark.
- Discussion:
 - *Comment from stakeholder:* While it would be valuable to create a tool to display all properties, there are many conflicting factors. It is difficult to have architecture-agnostic software that covers all those areas, yet is also transparent, standardized, and complex as all application areas dictate. The SPECpower committee is balancing these factors with the need to create a tool that minimizes equipment, skill, and tuning needed to run it. Tradeoffs will be required in any specialized tool to balance all of these factors.
 - *Comment from EPA:* Ensuring fairness of the end tool can either be built into tool itself or be part of how the tool is used. Context for results important.
 - Software tuning:
 - As noted above, many benchmarks have numerous software settings that can be tuned to optimize performance under the benchmark. While this is handled through full disclosure in the benchmark world, this requires motivated and available end user attention to uncover testing variations.
 - Is this reasonable for the ENERGY STAR community? Suggestions during this discussion also noted that ENERGY STAR should focus on a HW-only benchmark to the extent possible.
 - Response/counterpoint: but gains are mostly from the whole system, not just hardware, so need to find balance for possible software settings
 - Comment from attendee: setup different groupings of software as there are groupings of hardware. Leave room for software energy efficiency.
 - In presented analogies to Miles per Gallon, MPG works because it is properly understood and applied to similar classes of products (doesn't compare across cars, planes, buses, etc, for diff scenarios). Tools can be built to fit criteria if

they're used for fair comparisons / used in fair manner, not that tool is inherently fair/agnostic.

- *Question to attendees:* is active mode energy efficiency a worthy goal to pursue?
 - *Response:* yes. Many attendees willing to accept tradeoffs to get a general sense of efficiency in operation.
- *Question to presenters:* is EPA intending to build one single test for efficiency that takes the place of existing methods?
 - *Response:* No. there are many layers of information available. EPA is seeking to provide an overview of efficiency and provide a starting point for those who wish to dig deeper (“start out with Miles per Gallon, but have access to highway vs. city results”)
- Comments on architecture-agnostic considerations: software layers play big part in efficiency, so software needs to be able to run across variety of OSs, etc. in similar ways. High-performance I/O works differently in diff software environments.
 - *Comment from attendees:* Are we after fairness or accuracy and relevancy? Rating tools should reflect different strengths (e.g. sprinter vs. distance runner); find good all-around performer instead of the best specialist.
 - *Response:* specialist uses probably not the focus of ENERGY STAR evaluation.
- Discussion of workloads that should be covered and options (matrix of existing benchmarks? Should it be up to customer to find relevant benchmarks and compare server performance with their applications)?
 - *Suggestion from attendee:* identify what class of workloads/applications has widespread use and shape efforts around improvements for this class.
 - *Response:* important task. Returning to track meet analogy, even if the tools identify the best "sprinter", maybe user wants to identify the best "decathlete"; that is, choosing the right machine for the job
- Presence of Idle Requirements should and active mode efficiency tool be included in Tier 2:
 - *Question from attendee:* will idle power requirement be maintained or replaced by active power?
 - *EPA response:* customers have the right to know the idle compute power.
- *Question to attendees:* Tier 1 was limited to a small subset of servers; should Tier 2 be applicable to anyone who wants it?
 - *Response from attendees:* Caution was suggested to measure all in industry, not just top tier of most efficient servers.
- *Question to attendees:* How does EPA get stakeholders to step forward and reveal their data (getting industry more on board)?
 - *Response from attendee:* Make sure that whatever ENERGY STAR is comparing against, that it has relevance to what users are doing or using the product to do. Utilizing SPEC seems to be the best way to take steps forward.
- *Question to attendees:* Why is it a problem if the program is only looking at the top percent of the submissions?
 - *Response from attendees:* Goal should be to compare a lot of different models. "Don't underestimate your brand." ENERGY STAR is an opportunity to open up all information on servers.
- *Comment from attendee:* What does the Power and Performance Datasheet mean? The datasheet needs to be simple and easily-understood by customers.
- *Question from attendee:* Is there discussion with Europe/Japan to standardize specs?

- *Response:* EPA and The Green Grid are meeting with the EU and other international groups to discuss paths toward further collaboration.
-

3a. SPEC

- Slides:
http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/3a_20090925-SPEC_Server_Efficiency_Rating_Tool.pdf
 - Notes:
 - Reviewed SPEC's experience with SPECpower and benchmark energy efficiency evaluation, development process, and key considerations for the committee as it begins development of a candidate rating tool for ENERGY STAR.
 - Discussion:
 - *Questions from attendees:* Are benchmarks adaptive? Do they change based on cache sizes or optimize for different hardware, or is there a static, perfect model?
 - *Response:* Higher performance and performance differences stem from differences in hardware in the server. In handling cache, SPECpower does not specialize to a particular architecture.
 - *Comment from attendee:* Regarding SPECpower_ssj2008, is collected information on slide 6 covering the market as a whole, or only a subset of the most energy-efficient servers and most energy-efficient settings, or the broader market?
 - *Response:* Configuration details underlying data is important. Results are from only a subset of servers on the market.
 - *Question from attendee:* How does SPEC power capture the performance aspect of efficiency?
 - *Response:* SPECpower stresses CPU, main memory, I/O disk, and I/O network. There are different ways to slice up the performance picture.
-

3b. Themes from Morning Sessions

- Slides: n/a
- Notes:
 - Goal of Tier 2 – Should make standard available to anyone who wants to qualify as feasible.
 - "Benchmarks are beauty contests."
 - Caution to measure the entire industry, not just the beautiful
 - Why is it a problem? Want to incentivize people to be beautiful
 - Problem because people also want to figure out when to replace
 - Need to provide data for entire family line of products, not just the most presentable.
 - How to get more industry commitment and interest?
 - Make measurements relevant to end-user decision making
 - Need to get input from procurement decision makers or end-users push vendors to answer why the standard is relevant

- Have only a handful of customers asking what datasheet means – too technical
- Don't underestimate the power of your brand – ENERGY STAR label still useful
- Tools need to be simple and easily grasped – need to involve marketing etc.
- So have standardized result presentation?
- SPEC seems to have done productive work – should recycle as mechanism
 - But cannot take forever; need to raise awareness of progress to date
- Have international perspective – unifies standard with EU, Japan, China etc.
 - The Green Grid consortium can play a role?
 - Industry needs more voluntary involvement
- Two methods of communication
 - Very knowledgeable users – e.g. EBay etc.
 - Users who want very simple decisions – e.g. ENERGY STAR or not
- Directional data more useful than 100% accurate data

4a. Session 1A: Blade Servers

- Slides:
http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/BREAKOUT_1A_ES_Servers_T2-Blade_Servers.pdf
- Notes:
 - Review of Preliminary Specification. One proposal was to take measurements for aggregate blade system, divide by number of blades to get approximate individual blade info; chassis gets amortized.
- Discussion:
 - *Question to participants:* What special considerations should EPA give blades? What are best practices? Are chassis deployed filled?
 - *Suggestion:* develop cost-effective test methodology, short of using a full chassis. Use some base system (e.g. a quarter full), and then see what changes adding an additional blade makes.
 - *Suggestion:* Testing fully-configured blade chassis of 10+ blades is costly. make sure you're on "flat part" of curve in what you're measuring
 - *Suggestion:* Cost of test platform: certain electro-magnetic (EMI) compatibility tests are required anyway. Chassis gets built once, then used a lot; can add more blades at will
 - *Suggestion:* Since one gets a lower efficiency number with partially-populated system, can we extrapolate performance of a full system based off this partly filled one? Also, test procedure should take into account I/O traffic, fans, and utilization.
 - *Suggestion:* Leave it up to the vendor to fully populate or not; set the test methodology independent of this consideration (i.e. divide by number of blades in system).
 - *Response:* This could leave potential to game the system.
 - *Question to participants:* How do customers buy blades? Half/full chassis?
 - *Response:* Most purchases are populated with less than half of the blade bays full unless it is an HPC center that's buying many fully-configured blades.
 - *Question from participant:* What is "maximum configuration?" We have, for example, sixteen slots for blades, and lots of different possible types of I/O.

When we do EMC testing, manufacturers can define what the maximum configuration.

- *Comment from presenter:* Will vendors qualify for ENERGY STAR the blades and/or chassis? Do blades really warrant an ENERGY STAR rating on their own? Blades can be sold without a chassis/enclosure. Chassis meant to live through a couple generations of blades. Chassis doesn't have rating without anything in it.
 - *Suggestion:* consider each blade bay as a server (e.g. evaluate a 10-slot system as 10 servers).
 - *Suggestion:* Qualify both the chassis and blades. Use a few blades to qualify chassis.
 - *Comment/Suggestion:* There are an endless combination of blade servers and configurations. Water cooling will come into play soon. Focus on setting a couple of requirements to meet: power supply [efficiency], idle floor to meet, etc.
 - *Question from participant:* Is ENERGY STAR trying to encourage no fans/active cooling?
 - *Response:* Performance reflects amount of energy that goes into it. Want to approach this complicated environment and come up with measurements that make sense.
 - *Comment:* It is difficult to come up with a definition of "work" and this is a significantly different problem for the blade chassis configuration: it is not like specifying a bunch of rack-mount servers, rather more like specifying a network of servers. With this in mind, we almost have to come up with work for a network.
 - *Question from presenter:* how difficult is it to report out individual blade energy consumption in an active environment?
 - *Response:* it is fairly easy, depending on design. There is not currently a uniform method of doing this, but one can get to within 5% accuracy.
 - *Response:* doesn't it come down to power supply efficiency? How much overhead does chassis charge? Need work input from blades to combine with chassis to get one usable number. Power supply selection and configuration can change that overall number, but it's up to operator to make sure they're using a reasonable one.
 - *Suggestion:* To set a baseline: "not certified ENERGY STAR unless it has at least this many blades" Set a maximum configuration number?
 - *Question:* Is the chassis part of the data center (a mini rack) or is it part of the compute system? Where do you start to measure the efficiency of a data center? What is actually needed (e.g. blade, fan, and power supply)?
 - *Question from presenter:* What questions do customers ask?
 - *Comment:* Marketing: "Blade servers should just automatically get ENERGY STAR qualified because they're so much more efficient." Customers want a comparison of why they should buy 10-slot chassis with 2 blades in it instead of 2 standard servers; show them data from running some benchmark as a collective, with whole system power.
 - *Comment:* Do customers need to see stickers (indication of ENERGY STAR qualification) on chassis, extended chassis, and blades themselves?
 - *Question:* How is system sliced up? The system is very different when it requires all fans even with fewer blades. What happens when you connect a switch to the back, and how does that affect energy star rating?
 - *Suggestion:* Blades -are- servers. Treat them as such, or industry will get bogged down in the details forever; giving EPA numbers can finally put question to rest of whether or not blades are more efficient.
-

4b. Session 1B: General Requirements and Program Scope

- Slides:
http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/BREAKOUT_1B_ES_Servers_T2-Scope_and_GenReq.pdf
- Discussion:
Scope
 - *Comment to participants:* The Preliminary Specification noted that areas of the server market outside of the Tier 1 scope would be investigated for inclusion into the program.
 - *Question to participants:* What is the relevance of servers described by the new definitions for Resilient Servers and High Performance Computing Systems to the overall scope of the ENERGY STAR Computer Servers program?
 - *Suggestion from participant:* HPC servers are not comparable to the bulk of the server market
 - *Suggestion from participant:* highlight changes in different versions of drafts.
 - *Question to participants:* Are there suggested areas of the market outside of the communicated Tier 2 scope suggestions from The Green Grid that have a critical mass of products to allow effective comparison, represent a large source of energy-saving opportunity, or otherwise present an opportunity for ENERGY STAR to differentiate the market?
 - *Comment from participant:* Expanding beyond current scope is ill advised; current scope already encompasses the bulk of volume market, and adoption rate of ENERGY STAR is currently small.
 - *Comment from participant:* EPA should focus on getting the program underway, refining methods of addressing current products, and then move into other areas of the market.
 - *Comment from participant:* Specialist/niche products do not sell at high enough volume to warrant ENERGY STAR coverage priority. In addition, customers for these products likely are knowledgeable enough to implement the best practices for these server types. There is, however, nothing wrong with EPA encouraging right behavior.
 - *Comment:* Need to distinguish different kind of products – mission critical vs. others.
 - *Comment from participant:* Disagree – it would be valuable to see how different classes of servers compare against each other.
 - *Comment from participant:* Developing a workload in the active mode rating tool becomes more difficult as scope of products widens.
 - *Question to participants:* Is there data that could be shared with EPA on percentages of the server market (in units shipped, not revenue) to provide a better sense of how the program is doing?
 - *Comment from participant:* 95% of servers shipped in 2008 are 1-2 processor socket systems (currently covered by ENERGY STAR). Fewer than 1% have >4 sockets. Trends point to more cores per socket rather than more processor sockets.
 - *Comment:* Is it relevant to continue to compare AC-powered servers with DC-powered equipment? What considerations are required to ensure a fair comparison?

Taxonomy of Server Types

- *Question to participants:* one of the suggestions EPA is considering is rounding out the taxonomy of equipment presented in the definitions. Having an end to end list of definitions would help road mapping and clearly delineate what types of products are covered by the specification and which are not. Is this a valuable

step to take? Is there overlap in current definitions? What sources could EPA reference in enhancing this taxonomy?

- *Comment from participant:* The Green Grid is actively working on taxonomy, coming soon.
- *Question from participant:* Do IDC or other market research entities develop taxonomies for the server market?
 - *Response from participant:* IDC's taxonomy focus is driven by system, not technology, and may not be useful for ENERGY STAR's purposes.
- *Comment from participant:* Taxonomy in Tier 1 is incomplete. It should cover both socket and core counts, and reflect server capability.

Power Management

- *Question to participants:* What are some of the key PM features being used today – and in the near future?
 - *Suggestion:* power management will be exercised by an active mode rating tool. Perhaps specific requirements could be dropped since the effect will be seen in active mode requirements
 - *Response:* Power management needed in the absence of power-performance benchmark. Also, the rating tool may not be close to actual operation enough to show the effects of PM that an actual installation would present.

Energy Efficient Ethernet

- *Question to participants:* What are expected challenges to implementation of compliant equipment? Are vendors beginning marketing compliant components for EEE?
 - *Comment from participant:* It is unrealistic to expect pre-standard components until early 2010 (i.e. components that meet the draft standard before the standard is finalized). Standard systems using EEE should not be expected until the first half of 2011.
 - *Comment from participant:* As with most technologies, it may take a generation or two of equipment before EEE technology is used in the most optimal way to ensure energy savings. This factor should be considered when thinking about requiring EEE.

5a. Session 2A: Power Supply Requirements

- Slides:
 - http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/BREAKOUT_2A_ES_Servers_T2-Power_Supplies.pdf
 - http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/BREAKOUT_2A_CSCI_NPL_proposal_final.pdf
- Discussion:
 - Net Power Loss*
 - *Comment from presenter:* The current Tier 1 focuses on per-supply requirements. While this has helped put more efficient supplies into servers, it cannot show the effects of redundancy, sizing, on system efficiency. How does idle power specification drives right sizing?

- *Suggestion from participant:* Could system-level efficiency requirements (idle, active) reduce the need to focus on power supplies?
 - *Response from presenter:* One of the important side benefits of the power supply requirements is that they exist to help shape market dynamics. On one hand they improve the supplies that are in qualified servers, on the other they help create the demand to make power supply efficiency the norm and reduce the cost impact through economies of scale.
- *Comment from participant:* EPA should take care to not over-police power supplies. The industry is nearing the point of diminishing returns with improved power supplies.
- *Comment from participant:* Net power loss is configuration-specific and very burdensome to measure and report. It may lead to conflicts with a family qualification system.

Power Management of the Power Supply

- *Comment from presenter:* There is little widespread information on “power management of the power supplies.” Some anecdotal comment provided that in certain instances of duplicate or multiple power supply servers, the system is able to shut off unnecessary power supplies based on loading requirements and the amount of power needed.
 - *Response from participant:* Manufacturers have the capability to do this now. It is likely unnecessary for EPA to push.
 - *Response:* Could this be handled as an additional feature presented in the Power and Performance Datasheet, w/o pass/fail criteria?
 - *Response:* Even though there may be the capability out there, there still is a need to promote the capability.

5b. Session 2B: Reporting Criteria and Data Measurement/Output Requirements

- Slides:
http://www.energystar.gov/ia/partners/prod_development/revisions/downloads/computer_servers/BREAKOUT_2B_ES_Servers_T2-Reporting_and_Data_Meas.pdf
- Discussion:
Power & Performance Datasheet
 - *Comment:* Put data on different sheets in the same layout; calculations need to take into account the fact that there are min and max values and the effects different configurations will have on the two different extremes. Allow qualification of "families" of configurations: different amounts of memory, different processors, etc. "Typical" configurations are harder to determine.
 - *Comment:* In terms of why the Datasheet is important in addition to the QPI form, the QPI is intended to make sure the server(s) meet requirements, but Datasheet provides 2nd/3rd level of information that manufacturers can use as an indicator to the user of "it is better or worse for me?". The Datasheet gives information that is more applicable to them and operations. QPI and Datasheet shouldn't just duplicate information.
 - *Question:* Are there additional items for inclusion in the Tier 2 Datasheet? Do current items on datasheet fit customers' needs?

- *Suggestion:* Discrepancies between test and real machines, effects of power management on performance.

Power Management, Fans, Temperature

- *Question:* How much impact on performance does power management have? Is there data on this question? It would be good to review power-saving features to see that they do not impact performance.
 - *Response:* EPA wants to encourage PM, but it needs to promote full disclosure of what features are enabled on shipment and during testing so that user can make the right choices.
 - *Comment:* Fans and pressure – There is lots of interaction between fan speeds, data centers, and racks, yet there is no "as tested", "as shipped" for fan energy; ENERGY STAR ratings should be used for data centers as direction for fans.
 - *Comment:* See fan power as it corresponds to temperature.
- *Question to attendees:* How can the program deal with perception that power management is antithetical for achieving max performance (and how prevalent is this attitude)? (e.g. data center is scared to shut down over-provisioned portions of data center).
 - *Suggestion:* Some people must be using it and think it's good because the original specification was considering using it.
 - *Comment:* Where are "light-up" times long, or where does it require lots of energy to change states? Perhaps ENERGY STAR could first shine light on areas that are quick and low-energy to change states (e.g. changing processor states).
 - *Response:* A very large part of the market mistakenly thinks it has performance-sensitive workloads. We need to actually demonstrate scenarios in which we use power management and still maintain good performance.
- *Question:* Thermal side: are there additional factors that should come into play for Tier 2? (i.e., both inlet and exhaust air temperature). If so, we need to give this info to the manufacturers.
 - *Response from participant:* Staff at participant's organization say exhaust temperature is unnecessary.
 - *Response from participant:* Disagree. Data center and system operators want to know exhaust temperature at different points.