Comments on the *Energy Star Program Requirements for Computer Servers Draft 2*

Bernd Schäppi and Thomas Bogner, Austrian Energy Agency

17th September 2008

(The following position has been developed on the basis of the E-Server project currently conducted within the IEE-programme. However the position does not cover statements of the industrial members of the E-Server consortium as industry prepares its own position at an international level)

Overall we appreciate the second draft of program requirements for computer servers. We propose to consider the following recommendations and points of discussion:

0. Commitment

Declaration of Energy Star compliant products on lists on an annual basis

As we have stated in earlier positions already we would highly recommend to make more frequent reporting of data for Energy Star labelled products mandatory. In the introductory section titled “commitment” it is stated that “Energy Star partners have to provide updated lists of Energy Star complying computer servers on an annual basis”. Section 4 of the paper on testing indicates that manufacturers have to provide results from testing for all product models to EPA using the Energy Star Computer Server Qualified Product Information (QPI) form. This testing information actually should be available at the time the models are brought on the market as Energy Star complying products.

It is now unclear how the reporting of Energy Star compliant models and product data currently works in practise. While reporting data for products at the time when they are brought on the market would be highly appreciated and logical, reporting only once per year clearly would be insufficient for the highly dynamic market of IT products (not only servers).

Observations in Europe have shown considerable discrepancies between product information provided on the web-based EU Energy-Star database (www.eu-energystar.org) and information from other sources provided by manufacturers (e.g. data sheets, local websites). This problem seems to have several reasons:

First of all the EU data base (most likely also the US database) is always incomplete regarding Energy Star models available on the market which is probably is due to a lack of a specific requirement for immediate reporting of new products.

Second manufacturers sometimes provide different data on there data sheets and other information sources. Often not only Energy Star based data is provided but also information based on other testing procedures is published. This can lead to considerable discrepancies and misinformation.

To overcome these problems we would propose the following measures:

- A requirement for publication of new products/product data as soon as new models are put on the market together with the Energy Star label. Alternatively a minimum

requirement could be that new products (and the relevant product information indicated in Annex A) are published on a quarterly basis.

- If manufacturers publish data on power and energy consumption besides or instead of Energy Star based information (data based on internal or other test standards) this needs to be clearly indicated. The best and clearest solution would be to provide an indication for Energy Star in brackets behind the values published. This would allow clearly to distinguish between Energy Star and Non Energy-Star based information.

1. Definitions

A. Computer Server

We appreciate the current definition for computer servers which is broader now and does not exclude models by rather unspecific criteria (e.g. form factor). This would also correspond with our earlier recommendations. We furthermore support the focus on servers of up to 4 sockets for a first tier (see the following section on qualifying products).

It may be an option to be considered to include the type of desktop derived servers currently addressed in the computer specifications to the requirements for computer servers. This would allow a consolidation of equipment basically serving the same function (servers) and thus may be more logic.

However to date only very few models for “desktop derived servers” have been registered for Energy Star which indicates a rather low acceptance of the existing specifications. A reconsideration or revision may be advisable but does not have high priority and thus could be postponed to subsequent tiers.

I./J. High redundancy versus standard redundancy servers

The used categorisation “standard redundancy”/“high redundancy” systems seems rather uncommon. “Standard redundancy” as used here actually means “non-redundancy” specifying equipment with a single non-redundant power supply.

On the other hand servers with redundant power supplies (e.g. single redundancy) - here defined as high redundancy systems – are also found in the lower volume server product segment and are not necessarily specific for mission critical facilities. A more appropriate classification might be “non-redundancy”, standard/single redundancy and high redundancy systems.

2. Qualifying Products

We support the focus on servers of up to 4 sockets for a first tier since this product segment currently covers the majority of so called “volume servers” being responsible for 80% of energy consumption of the server market. The current proposal is in line with our earlier comments and the recommendation to exclude the mid-range product segment at a first stage.

We also support the approach by EPA to exclude server appliances for a first tier due to their more specific nature.

As a mid-term goal we would propose to address storage systems for subsequent tiers. Except for small server systems in small offices disks are often located in separate storage equipment. Especially in consolidated systems where physical servers already have been optimised in terms of number and size storage units can be responsible for a considerable part of the energy consumption. Consequently efficiency of “external” storage becomes an important aspect. Although several options of optimising energy efficiency of storage are based on management issues like rigorous storage administration, use of tapes etc., a significant part of the efficiency can also be gained by hardware features (efficient disk drives, variable disk speeds, power management for storage).

Consequently it is proposed to announce storage systems as a potential product category for subsequent tiers. Another product group to be addressed in the future is network equipment. However the recommendation would be not to delay the development of the current first server specifications but to address the additional product groups at subsequent stages.

3. Efficiency requirements for qualifying products

A) Power supply requirements

The proposed differentiation of power supplies p</> 1000 W is not relevant for the majority of rack servers which in general are in the category p<1000W. The category >1000W will mainly address blade chassis which can be powered with several kW. The proposed values seem reasonable at a first glance.

Net power loss

As indicated in the notes by EPA the definition of minimum power supply efficiency is not sufficient as a longer term goal since it does not necessarily address overall energy efficiency and in particular does not cover oversizing.

According to observations oversizing of power supplies is quite common in practise. A major argument for providing additional power capacity are the various options for hardware upgrading offered for many models and the need to meet power demand of maximum configurations. However in practise upgrading in the typical lifetime of servers primarily will be done for memory but less frequently for disks (which are often located in external storage) and rarely for CPUs. Thus options for upgrading in practise are used to a quite limited extent but increase basic power demand due to large power supplies and more complex boards.

An innovative approach addressing the effectiveness and the efficiency of power supplies more comprehensively would be highly appreciated. In this regard also the net power loss approach proposed by EPA could be an important step forward. However significant research and development is necessary to move forward in this area.

10% loading level

To our knowledge in the past most servers showed relatively high power loads under idle or near idle conditions (thus low workloads but considerable power loads). Idle values were in the range of e.g. 80% of maximum (measured) power. This was due to the fact that the hardware was not designed to reduce power demand when operating under low workload. Modern equipment in the segment of volume servers shows idle levels of about 50% of maximum measured power.

However regarding the specifications for power supplies it has to be taken into account that maximum rated power (nameplate power) is not equivalent to maximum measured power which is much lower in most cases. If maximum measured power is about 70% of the rated power idle power of modern relatively efficient equipment could be about 30% of max. rated power.

Consequently power loads of 10% would be experienced only in cases with high redundancy and oversizing. However idle efficiency will further increase in the future.

The trend towards consolidation and virtualisation aims at consolidating higher work loads and power loads on single physical servers and consequently will help to avoid extended idle periods in the future.

Besides the discussion on very low power loads we propose to reconsider if it would not be appropriate to introduce an additional level between 70 and 80% of maximum rated power rather than addressing the academic 100% level. The 100% rated output power should not be required under practical operating conditions while levels in the range of 50-80% have more relevance.

B) Idle power

Idle power approach in general

In general the approach to address idle power as a first step in a first tier is supported. Many (volume) servers for different applications today are still operating near idle mode for a considerable amount of time. This fact has not been considered so far in terms of energy efficient hardware design and power management. CPUs and other hardware components for servers formerly were not designed to address low workloads in an energy efficient way and consequently power demand under idle conditions was high. Therefore it makes sense to address idle power as a relatively easy to access parameter at a first stage.

However in a second step it has to be considered that servers in the higher performance and price segment are generally not designed (resp. to costly) to be operated in or near idle for extended periods but are intended for operation at significant workloads. Furthermore workloads on single physical server units are increasing due to consolidation/virtualization strategies.

So obviously addressing idle power addresses only part of the energy saving potential for a part of the relevant applications. In addition to idle power typical workload power has to be addressed in some way in the future. SPEC power would provide a first step towards this goal although the benchmark is currently focused on one specific workload and thus not representative in general yet.

Current concept of server categorisation

The rationale behind the proposed sub-categorisation of servers for idle power provided in tables 3 and 4 is not clear. It is unclear why table 3 suggests only a division between single processor systems and multi- processor systems (2-4 processors). We would expect a significant difference between 2- and 4- processor systems and therefore would propose a further segmentation between 2- and 4- processor systems.

It is furthermore unclear why a different sub-categorisation has been chosen for the so called "high redundancy systems". Here 1-2 processor systems are lumped together in one sub-category and 4 processor systems are considered a separate (multiprocessor?) category. Also here we would propose the same segmentation as indicated above.

The proposed categorization based on memory is possible but certainly will be subject to dynamic changes since the standard of "low memory" and "high memory" will shift continually. Today 16 GB may be seen as a standard configuration for small to medium sized servers within the < 4 sockets segment. With prices for memory decreasing and technological development ongoing the range chosen now may shift quite dynamically proposing a need for dynamic upgrade of categories.

SPECpower as the proposed methodology for assessing idle power

EPA proposes SPECpower as a basis for idle measurements. In fact the benchmark itself is not used to do the idle measurement since idle is measured at no workload. Thus the idle assessment so to speak is a by-product of the benchmark tests. So in fact for the first tier addressing idle mode only, the framework and testing conditions of SPECpower would be used but not the benchmark itself.

Running the benchmark to produce on-idle measurements only, could be seen as a high effort. However it is expected that power at typical (non idle) workloads will be addressed in the future and such assessments may be based on SPECpower or its successor. Therefore it may make sense to introduce SPECpower at this stage.

Furthermore as is indicated in the section on reporting requirements (see below) EPA also expects information on maximum power for different model configurations. It is assumed (but not clear) that this information will also be required based on SPECpower?

Alternatives to using idle power based on server categorization

EPA discusses other approaches of addressing on-mode efficiency alternatively to the idle-power based on server categorization which have been discussed at a US stakeholder meeting in July. The two more prominent proposals were

- Assessing idle power as a percentage of maximum power
- A power saving checklist indicating power saving functions instead of concrete requirements

From our point of view these approaches can not be seen as an alternative to the basic idle criterion but would rather be a useful supplement. Indicating power saving functions could be a useful add-on. Idle-power as a percentage of maximum power could be a useful additional information indicating to some extent how efficiently the system power scales down from high to low workloads. However not only a percentage should be provided but both idle and max power need to be indicated and information on maximum workload then needs to be taken into account.

C. Standard information and reporting requirements

Concerning the Appendix A on required product information the sections on "System Characteristics" and "System Configurations" seem largely redundant.

From the section on "Power and Performance Data" it remains unclear what kind of information is mandatory for reporting. It seems that reporting of idle power and maximum power would be mandatory. However it is not explicitly stated how maximum power should be measured although it would be assumed that SPECpower would be applied for that purpose as it is also used for idle power assessment.

4. Test criteria

According to the first paragraph on test criteria results from testing have to be reported to EPA using the Energy Star computer server qualified product form (QPI). So in general all relevant technical information has to be provided to EPA and consequently will be available at EPA.

It remains unclear however when this type of information has to be provided by the manufacturers. Is this once a year at minimum together with the updated list of products or is it at the time an Energy Star labeled product is brought on the market? A clearer definition/information on that issue would be important (see introductory comments in this position paper).