Comments on the *Energy Star Program Requirements for Computer Servers (Draft 1)*

Bernd Schäppi and Thomas Bogner, Austrian Energy Agency

(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 first draft of program requirements for computer servers. We propose to consider the following revisions, amendments and points of discussion:

Commitment

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

- It is highly recommended that new products complying with Energy Star (which are sold as Energy Star compliant products) must be published on the US or EU database. Provision of information on complying products only once per year as indicated in the draft requirements will always result in a partly outdated inaccurate database which is then not useful as a source of information for buyers.
- To avoid this situation we strongly recommend to require a publication of new products as soon as they are put on the market together with the Energy Star label. The minimum requirement would be that new products are published on a quarterly basis.

Definitions

a) General definitions for servers including blades

- The general definition provided for computer servers is reasonable. However the specification of typical hardware characteristics is more problematic since several hardware features like for example multi socket capability, baseboard management controller etc. are also components of PC equipment. Furthermore there will be dynamic change of hardware characteristics in the future
    - It is proposed to use the general server definition together with the more robust characteristics which e.g.:
        - Reliability, Availability, Serviceability, and Manageability (RASM) features
        - Certification for use with enterprise-class server Operating Systems
        - Multiple LAN and/or WAN networking ports, such as Ethernet
    - The definition for Blade Servers may require some modification since for example not all Blade Servers use hard-drives. There are also solutions with flash drive cards and external storage.

b) Categorisation/Classification

- The proposed categorisation (3 server classes) seems problematic since on one side it is an unnecessary attempt to cover more or less the whole range of server hardware and on the other it cuts off desktop-derived servers. The three defined classes of server hardware are still very broad and not necessarily helpful for the definition of specifications. They should be reconsidered.
    - We strongly propose to limit the Energy Star specifications to volume servers (and may be the smallest class of mid-range servers) as these are the mass market. It should be feasible to define the upper boundary of this market segment of servers thereby defining what is in and what is out of scope.
    - Volume servers and small midrange-servers account for more than 80% of the energy consumption of servers worldwide (see results from EPA study and EU-study within IEE-E-Server, see www.efficient-servers.eu)
    - Consequently midrange and high-end servers are a relatively small market segment an therefore less important in terms of energy consumption. Larger mid-range servers and high-end servers a normally operated under higher load conditions at high performance but also at high availability requirements. Consequently energy efficiency often will be better than for small systems but at the same time restrictions for energy saving measures will be higher. In contrast to volume servers which is a highly dynamic market segment, the market of high-end servers and larger mid-range servers is currently not growing but rather decreasing.
    - Consequently addressing high-end servers and the majority of midrange servers will be a high effort and a comparably little gain in energy efficiency which will be difficult to justify.

- Desktop-derived servers (DDS) are currently addressed in the computer specifications. It is proposed that they should be moved to the server specifications in the future for the following reasons:
- DDS serve the same general functions as other servers and are as such different from PCs and workstations etc.. Thus it seems quite logical to cover DDS in the server specifications. The recommendation is further supported by the fact that it is not always easy to distinguish desktop-derived servers from other small volume servers and this will even become more difficult in the future due to changes of hardware components and configuration of equipment.
    - A separation of DDS from other volume servers by number of sockets may not be appropriate in the future since there will be single socket machines with multiple cores (quad core and more) on the market not belonging to the desktop-derived segment. Overall due to dynamic changes in hardware features, the distinction of equipment within the volume segment by the already defined features (definitions for desk-top derived servers in the Energy Star requirements for computers) may get difficult.
    - The DDS specifications included in the Energy Star requirements for computers currently do not seem to be supported by the market since no products have been registered by now. This may also be an indication that a revision of the definitions and specifications is necessary and a transfer to the server specifications should be supportive.

c) Blade Servers

- Specifications for the energy efficiency of blade servers will have to be defined for the complete system including blades and chassis. It has to be considered that blades are not always sold at a 100% configuration. Typical chassis sizes can host 6, 12 or 14 blades but may be sold with only 50-70% configuration (only half or two thirds of the blades installed. Efficiency of power supply and other supportive hardware located in the chassis will be dependent on the number of blades installed. Consequently energy efficiency criteria may have to be defined for different configurations.
- It is proposed that blades are covered under a separate server category and specifications are defined for different configurations (e.g. 50%, 75%, 100%).
- There may be the option of scaling chassis hardware (adding power supply modules and other hardware on demand) which could be supported by specific requirements.
- However all these aspects need to be investigated in more detail before specific requirements for blade servers are defined.

d) Definition Idle State

- The definition for idle state should be more detailed, e.g. clearer definition for „machine is not asleep“.

Efficiency requirements for qualifying products

a) Requirements for power supplies

Power supply efficiency is very much appreciated as an essential criterion. A further improvement of efficiency rates of power supplies is very likely. For this reason a two tier approach may be appropriate setting short term as well as mid term targets.

- *Load points*
  - A 100 % load requirement – either for efficiency or power factor – seems questionable, as there is no evidence of relevance in business life application. We suggest to reconsider a requirement for additional load points in the mid-range of loads (e.g. 60%, 80%), since these operating points are more realistic.
  - Also the practical relevance of a 10 % load point is arguable for many types of equipment as such low loads will not emerge in standard operating conditions. However in case of high over provisioning and a redundancy of more than a factor two this level may be relevant. This could be the case for example for blade centres with low configuration.
  - The issue of appropriate load points should be discussed further. A 10% requirement may be useful for example for blade centre power supplies.

- *DC-DC converters*
  The evaluation of DC-DC conversion efficiency is much more complex than the evaluation of power supplies. DC-DC converters as integrated parts of main boards are challenging to measure, although they are considerably influencing the power draw in general.

- *Additional Power Supply Types – Blade Center*
  As mentioned in the draft paper the measurement requirements will be based on test bed environment. Therefore we see no specific reason to differentiate between testing of PS for blade servers and other server types.

- *Power Supply Sizing and Redundancy*
  - The proposed approach to address PS efficiency is appreciated. Nevertheless the issues of appropriate provisioning as well as intelligent redundancy concepts are of high importance.
  - The Energy Star requirements should also support concepts for more intelligent redundancy. In case of redundancy one power supply should be in charge for the entire load while the others should remain in a standby mode. Energy Star may consider incentives to support intelligent redundancy concepts (e.g. setting lower efficiency requirements *per* PS, if intelligent redundancy concepts are applied).
  - It is doubted that the approach for a maximum allowable power consumption which is considered in the draft paper, will be feasible in the sense of several categories facing different levels of performance as well as loads.

- *PS cooling fans*
  In general server cooling is based on independent ventilation integrated into the chassis (explicitly separated from power supply fans). For most server equipment PS fans have little relevance for overall server cooling (probably with the exception of desktop-derived servers).

b) Requirements on idle power

- In general the idea of addressing idle power is appreciated. Discussion with data centre owners support the idea that servers (of course depending on workload) spend a significant amount of time in idle mode or under very low load conditions.
- Idle power will strongly depend on hardware configuration (e.g. number of CPUs, hard disk, RAM). Furthermore idle power is dependent on the maximum performance of the system and can not be rated without consideration of that.
    - Since server categories as defined in the draft proposal are broad and cover a spectrum of configurations, specifications for idle levels would have to be based on configuration (increasing with number of CPUs, disks, RAM etc.) or have to be related to peak-power consumption (idle power/peak power). Furthermore as stated above maximum performance has to be taken into account to arrive at a meaningful interpretation of idle levels.
    - The options for setting meaningful requirements for idle power still need further investigation. Therefore for now it is proposed to require information on idle power related to configuration or peak power but not to specify mandatory maximum levels. The specification of allowances for stand-by may be an issue for a later tier.

c) Standard information reporting requirements

- It is generally appreciated that idle as well as maximum power values will be provided based on a standardised workload (e.g. SPECpower).
- SPEC_Power_ssj at the moment only is available for one workload (SPEC_jbb_2005, JAVA). SPEC_ Power_ssj is an ideal benchmark to assess the energy efficiency for a single server dedicated to a CPU bound application scenario. However results from SPEC_power_ssj are not representative for I/O intensive scenarios and consolidated environments.
- To make SPEC_Power_ssj still more meaningful as a criterion in the context with Energy Star more workloads will have to be covered. This further development which finally may allow to use SPECpower also for mandatory requirements for power consumption may be a longer and complex process. Several issues have to be addressed as for example which workloads should be covered, what kind of aggregation of parameters can be achieved to arrive at a single Energy Star criterion etc..
- SPEC_Power_ssj primarily provides an average workload to power ratio. For real world scenarios idle power and maximum workload are relevant as well. For example servers with similar SPEC-values may differ considerably in idle power and peak performance. In practise under low performance requirements systems with low idle power (single socket) are preferable in terms of energy efficiency. High performance (multiple sockets) always comes with the trade-off of higher idle power.

d) Power management and virtualisation requirements

- Requirements for power management and virtualisation have to be separated since these are completely different issues.
- Virtualisation provides a means for high energy savings in many situations. This potential is visible and can be communicated/supported without implementing specific energy efficiency criteria. Thus options of virtualisation in cases where appropriate can be supported by general information measures. It would make little sense to support the transition from standard hardware to virtualised solutions by specific energy efficiency criteria.
- At a later stage it may become relevant to compare energy efficiency of virtualised systems themselves. However as a basis for this, new benchmarks and criteria have to be developed using a workload and performance based approach. To date a new SPEC for virtualisation is under development which may be used together with SPEC-power in the future.
- Power management for servers is certainly appropriate at level of components as for example for CPU(frequency/voltage scaling), disks etc.. On the other hand mechanisms for shut down or stand-by in extended periods of low use (e.g. weekends, nights) are relevant but their use will be strongly context/application dependent. Advanced management systems offer shut down options for whole servers. For example also in the area of virtualisation *VM-Ware VSX* allows migration of active virtual servers to one or more physical units and shut down of unused physical units.
- Dynamic high level power management e.g. power down to sleep, deep sleep and standby-modes for servers is a complex matter as it involves difficult aspects of response time and availability. These options may be addressed in the future dependent on applications but cannot be covered at present.