June 4, 2018

To: Ryan Fogle, EPA Manager, ENERGY STAR® for IT and Data Center Products;

John Clinger, ICF International

Re: **TGG Comments to ENERGY STAR Servers v3, Draft 3**

TGG appreciates US EPA's continuing efforts to refine the ENERGY STAR server energy efficiency requirements and the work that has gone into the development of the Draft 3 document. TGG is particularly supportive of EPA's decision to use only the SERT active efficiency score to assess and select server configurations and product families for certification to ENERGY STAR. As demonstrated by the TGG analysis of the ITI/TGG SERT dataset, the SERT metric provides the most effective measurement of a server's ability to deliver more work per unit of energy consumed and ability to differentiate more efficient servers across the range of server configurations within a product family. We think adoption of the SERT metric offers a simplified metric which will further incentivize industry efforts to offer more efficient computing platforms that meet the varied needs of our diverse group of customers.

Overall, TGG feels that the Version 3 Draft 3 specification is largely complete and ready for publication. We do have a set of recommended edits and additions to provide better clarity regarding the specific process to certify and manage a certified product family and to insure consistent application of the specification. We offer the following recommendations and justifications for edits and additions.

1.A.4.B(9) Line 96: The "and" at end of item (9) should be removed. The "and" could be understood to imply that items (9) and (10) are to be read as combined whereas these are separate items on the list of ten items. If EPA desires to have a conjunction at the end of item nine, it is recommended to use either an "or" or an "and/or".

1.A.7 Lines 129 and 135: Change the "and" between deep learning and artificial intelligence to an "or" on line 129 and line 135 (2 places). As written, the HPC definition could be implied to be required to be suitable for all three of the defined applications. HPC systems definition was revised to specifically address the new deep learning and artificial intelligence applications. The use of the conjunction "or" clarifies that a server does not have to be optimized for all three application types to be considered a HPC system.

Section 1, Definitions: TGG recommends that EPA include a definition of storage server and network server in the Draft. Because of the unique nature of the server and network servers, which tend to use lower performing processor(s) and have large number of storage devices and/or network ports, TGG continues to believe it is appropriate to define these server types and exclude them from the ENERGY STAR requirements. We have discussed this with EPA in several of our working discussions on Draft 2 and Draft 3. The definitions, taken from the ISO/IEC 21836 Draft Standard, are provided below for your consideration:

Network Server: A network product which contains more than 11 network ports with a total line rate throughput of 12 Gb/s or more in addition to the same components as a computer server, and the capability to dynamically reconfigure ports and speed and provide support for a virtualized network environment through a software defined network.

Storage Server: A storage product which contains 12 or more storage devices in addition to the same components as a computer server. A storage server can run on more than one non-vendor specific software which is designed to support storage system connectivity, capacity optimization management (COMs) deployments and virtualized storage environments arrayed in a software defined storage network.

Section 2.2.2, Line 341: Modify the storage products exclusion to read: "vi. Storage products including Blade Storage *and Storage Servers*; and"

Section 2.2.2 Line 342: Modify the network equipment exclusion to read: "vii. Network equipment, *including Network Servers*."

This is particularly important given that our analysis shows that network and storage servers are populated with the lower performing processors in a processor family, typically with lower memory capacity. These server configurations will have active efficiency score that are closest to the ENERGY STAR threshold. Not excluding the systems may inadvertently require data center operators to be unable to acquire storage or network servers if they must procure ENERGY STAR certified servers or they will need to procure servers with a higher level of processor power, and higher deployed power, than would be necessary to meet the performance requirements of the product. The idle analysis performed on the random and sequential worklets demonstrates that a storage server needs minimal processor capacity; the processor was largely idle throughout the random and sequential worklet test cycles.

1.C.1.D, Lines 183: The multi-output power supply definition indicates that the sum of any outputs that are not considered primary or secondary outputs should be greater than or equal to 20 watts. It is our recollection that this clause was written in for the single-output PSU definition so that lower voltage, lower wattage control power feeds would not be used to define a PSU as a multi-output PSU. Either it should be removed from the multi-output PSU definition, as it is not germane or it should be modified to read "…that are not primary and secondary outputs shall be no greater than 20 watts…" to match the single-output PSU requirement and reflect the desire to keep the control power feeds to less than 20 watts. A check of the previous versions (1.0, 1.1 and 2.0) of the ENERGY STAR requirements indicate the Version 3 Draft 3 definition is the same as those versions, but it appears that this may be an error.

1.F.6.A, Line 268; APA definition: EPA chose not to change the APA definition from Draft 2 to Draft 3. In its comments to Draft 2 of the Version 3 requirements dated October 30, 2017, the SERT WG proposed the definition be expanded to indicate that an APA could include one or more GPU or FPGA chips, which the APA(s) could be direct attached or installed through an add-in card and that APA card could depend on separate, standalone switches to operate. The WG believes that the expansion of the definition is necessary to acknowledge the fact that APAs can be built with different kinds and number

of accelerator devices and dedicated switches in order to support the requirements for reporting (see comment below on Section 3.9) and to position the definition to support specific requirements in future versions of the Computer Server Product Specification.

Auxiliary Processing Accelerators (APAs): An additional compute device installed in the computer server that handles parallelized workloads. *This includes, but is not limited to, Graphical Processing Units (GPUs) or Field Programmable Gate Array chips which can be installed in a server either on Graphics or Extension add-in cards installed in general-purpose add-in expansion slots (e.g., GPGPUs, CPU accelerators, etc. installed in a PCI slot) or direct attached to a server component such as the motherboard. There are two specific types of APAs used in servers:*

A. *Expansion APA:* An APA that is on an add-in card installed in an ~~*general purpose*~~ add-in expansion slot ~~*(e.g. GPGPUs, CPU accelerators, etc. installed in a PCI slot)*~~. *An expansion APA add-in card may include one or more APAs and/or separate, dedicated removable switches.*

B. *Integrated APA:* An APA that is integrated into the motherboard or CPU package~~. or an expansion APA that has part of its subsystem, such as switches, included in the non-APA server configuration that would be used to run the energy efficiency test (SERT suite).~~

Note: The italicized text in Expansion APA definition represents text added to the Version 3 definition. The marked-up edits show which deletions and additions from the definition proposal we submitted with the October 30, 2017 comments for reference. The modifications were made based on additional understanding of the implementation of APAs and the changes in reporting/compliance requirements in section 3.9. The SERT WG has modified the definition presented in the group's version 2 comments based on additional work done since those comments were submitted.

1.G.1.C line 285: Add a sentence to this subsection which states: "A product family can be defined for a server with only partially populated sockets (e.g. 1 processor populated in a two socket processor system) as long as the configuration(s) are tested as a distinct product family, as required and meet the active efficiency limit for the number of populated sockets within a given product category."

TGG believes it should be acceptable for a manufacturer to certify servers with partially populated processor sockets as long as those servers can meet the active efficiency limit for the number of populated sockets. There is demand in the marketplace for servers with partially populated sockets to allow for future expansion or to increase the availability of memory, storage or I/O capacity to a single processor above what is available from a single socket server. If the partially populated server can meet the appropriate active efficiency limit it should be allowed to certify to ENERGY STAR.

1.G.2, lines 275-316: The SERT WG supports the Version 3 Draft 3 product family configurations.

Line 322: Add a "Section H: Qualifying a single Configuration: A server manufacturer can designate and certify a single server configuration to the ENERGY STAR requirements."

The Draft 3 Version 3 specification does not clearly indicate that a manufacturer can test and certify a single configuration. This addition clarifies that point.

3.3.3, lines 402-405: The reference to the "Power and Performance Data Sheet" needs to be changed to the "Computer Servers Qualified Product Exchange Form". The Power and Performance Data Sheet is no longer required under the Partner's Agreement (to the best of my knowledge).

3.5.3 Lines 431-432: Add a phrase at the end of the sentence: "…minimum active efficiency thresholds listed in table 3 *for each of three configurations, within a given product family, submitted for certification, and for all other server configurations represented as ENERGY STAR certified by the server manufacturer*."

Per the discussion at the EPA webinar, TGG is concerned that the requirements for certifying and representing a product are not sufficiently clear. The addition of the above phrase clearly notes the fact that all three certifying configurations must exceed the appropriate active efficiency threshold as well as all the products that the manufacturer represents as ENERGY STAR certified to its customers. The additional phrase removes any ambiguity regarding the requirements.

3.5.3, Line 432: Add 3.5.3.i, which states:

i. Where a server configuration has been tested to SERT V1.1.1, the measured worklet performance and power values can be used to calculate the SERT V2.0.1 worklet efficiency, workload efficiency and overall active efficiency scores, without re-testing. The calculation must be performed and validated by the CB certifying the SERT test data and scores.

TGG provided EPA analysis data during the Draft 2 comment period showing that performance and power measurements taken by SERT versions 1.1.1 and 2.0.1 were equivalent within the accuracy of the test and measurement processes. The differences in scores were the result of different normalization values used in V1.1.1 and V2.0.1 and the differences in calculation for the Flood and Capacity scores. Given this analysis, it is appropriate to allow manufacturers who have tested and certified a server product or product family to ENERGY STAR using SERT V1.1.1 should be able to use that test data to calculate a SERT V2.0.1 score and certify the server product or product family to ENERGY STAR Version 3. TGG can provide EPA a conversion spreadsheet that can be posted on the EPA ENERGY STAR website server products page to facilitate the conversion. This will enable manufacturers to avoid an unnecessary test to requalify their products to V3.

Lines 442-449; 453, 454, 459, and 460: Replace "measured" with "calculated".

The SERT worklet efficiency scores are, in all cases, calculated from the measured performance and power data.

3.5.3, line 462: The requirements need to specify the equation for calculating the worklet efficiency scores. Add the new "Equation 7: Calculation Eff_i "to the specification to clearly define how the interval measurements are combined into the worklet efficiency scores.

$$Eff_i = 1000 \frac{Perf_i}{Pwr_i}$$

Where:

$Perf_i$ is the geometric mean of the normalized interval performance measurements.

Pwr_i is the geometric mean of the measured interval power values.

Note: The numbers on subsequent equations beyond the new equation 7 need to be increased by 1.

Table 3 Line 463: TGG is very pleased that EPA has chosen to set the ENERGY STAR thresholds based on the SERT active efficiency metric. TGG strongly believes that the SERT active efficiency metric offers the best available method to differentiate servers based on the work delivered per unit energy consumed and make measurable reductions in energy use in the data center while enabling product designers to innovate to meet customer expectations for performance and capacity.

TGG is concerned about the active efficiency limit for 4 socket (greater than two installed processors) rack servers. An analysis of the 4 socket rack server configurations available in the dataset by configuration type indicate that the average active efficiency score for the low-end and minimum power configurations for the 2016 and 2017 product families are well below the 4 socket rack server active efficiency thresholds and the overall yield on the low-end and minimum performance configurations against the active efficiency threshold is 7.7%. As such, this will largely eliminate low-end performance 4 socket rack servers from ENERGY STAR certification. We don't expect this to be EPA's intent. While EPA may not see that anomaly based on its reported pass rate for 4 socket rack servers, the concern is that those pass rates were based on 2 out of 3 passing configurations. When you consider 3 out of 3 passing configurations this category pass rate dips down to unacceptable levels, not seen for other categories.

TGG recommends that the active efficiency limit for greater than two installed processor, rack servers be set at 13, the same as the 2 socket rack servers, to better enable manufacturers to offer greater than 2 socket, lower performance processor servers as ENERGY STAR certified. Given the average active efficiency values of the high-end performance and typical rack servers, this will largely serve to increase the availability of the full range of processor offerings for 4 socket, ENERGY STAR certified servers.

In addition, these systems tend to be richly configured with storage and network components, which adds additional power consumption without improving SERT performance resulting in lowered overall SERT scores making additional margin for the 4 socket category appropriate.

		# skts	4	Rack		Drft 3 limit 16		
			Average TGG Eff Sys Launch Date, All systems in data set					
	# Pass	# Fail	High End	Max Pwr	Typical	Low End	Min Pwr	
2010	0	5	10.55	8.79	7.99	6.89	10.37	
2011	0	0						
2012	0	6	4.62	3.99	4.27	4.51	3.41	
2013	0	15	7.11	7.08	5.78	5.24	6.64	
2014	5	17	13.04	9.41	12.35	12.84	6.18	
2015	1	10	11.46	4.82	9.33	3.21	4.81	
2016	2	8	17.47	9.81	13.33	10.50	11.62	
2017	4	3	30.60		25.06	13.62	13.54	
2018	0	0						
2019	0	0						
2/3	3	13	23.1%	System	1	13	7.7%	All
	Pass	Total	Yield		Pass	Total	Yield	

3.9 Line 537-545: TGG supports the revised testing approach for APAs with recommended modifications.

Add a new item ii under 3.9 which states:

ii. The idle power for the expansion APA card shall be calculated by installing the APA expansion card in the server and performing just the SERT idle test (skipping the worklet tests) and subtracting the SERT idle power measured without the APA present in the server.

a. Where a removable switch is required to support the expansion APA, the switch should be installed and SERT idle measure taken and then the APA should be installed and SERT idle measure taken, with the idle power for the removable switch and the APA calculated by subtraction of the appropriate idle power measurements.

b. Where an idle value has previously been measured and calculated for a removable switch and/or an expansion APA, those values can be reported for the expansion APA for multiple product families and/or configurations. New measurements do not have to be made for each certification submission.

This item is needed to define exactly how the idle power for the APA and associated switch are measured and calculated. TGG believes it is important to specify the process for reporting APA and switch idle power to insure consistency of the reported data.

Modify current item ii. and renumber it to iii.

iii. Manufacturers shall report the model name and model number, idle power consumption, *number of APA devices and/or switches on expansion APA card, and the total number of PCIE lanes, both input*

and output, for the supporting switch(es), for each APA device *and removable switch(es)* offered as an accessory within an ENERGY STAR product family.

In order to facilitate analysis of the idle power data, it is important have the details of the accessories. TGG found some of its SERT analysis work was occasionally limited by the lack of detail on certain configuration components.