NRDC

NATURAL RESOURCES DEFENSE COUNCIL

September 1, 2017

Mr. Ryan Fogle
United States Environmental Protection Agency
ENERGY STAR Program
1200 Pennsylvania Ave NW
Washington, DC 20460

Subject: NRDC comments regarding Draft 2 Version 3.0 Specification for Computer Servers

Dear Mr. Fogle,

On behalf of the Natural Resources Defense Council (NRDC), we respectfully submit the following comments in regards to the ENERGY STAR Draft 2 Version 3.0 specification for computer servers issued July 31, 2017.

Computer servers (hereafter referred to as servers) are the workhorses of data centers. According to a recent study by Lawrence Berkeley Labs[i], servers draw an average of 100 watts (1-socket servers) and 250 watts (2-socket servers), typically 24/7. When aggregated across the roughly 15 million servers in use in the United States, this amounts to more than 30 billion kilowatt-hours of annual energy use, with an additional 15 to 20 billion kilowatt-hours to cool these servers and keep them from overheating. This adds up to more than 1 percent of total U.S. energy consumption.

NRDC strongly supports the ENERGY STAR program for servers: it is an important tool to enable businesses of all sizes, as well as local, state and federal government agencies, to reduce their energy use through sustainable procurement policies for their data center equipment; it also enables electric utilities to implement efficiency incentive programs that reduce energy use in their service territories, and that help transform the market toward more efficient equipment.

We commend EPA for initiating the revision of the Server specification in reaction to the evolution of server technology since Version 2.0 went into effect in December 2013. We generally support EPA's Draft 1 proposal, and offer the following comments aimed at improving the draft specification. Our comments cover the following points:

1. **Active and Idle State Efficiency:** NRDC strongly supports EPA's Draft 2 proposal to set both active state and idle state efficiency criteria, to ensure that ENERGY STAR servers deliver on customer expectations of energy performance in the conditions they are typically operated in;
2. **Categories and adder framework**: NRDC supports the consolidation from 5 to 3 categories proposed in Draft 2. We also encourage EPA to revise categories to better align with performance so that ENERGY STAR is better able to differentiate on idle efficiency in a performance-based manner;
3. **Memory adders:** NRDC appreciates EPA reducing memory adders to 0.125 watt/GB in Draft 2, but this is still far higher than required by current mainstream technology. This is particularly problematic for configurations with large amounts of memory. We propose to use a hyperbolic tangent equation to better reflect the fact that power draw does not scale linearly with capacity, and to ensure that high-memory configurations are appropriately handled by the specification; We are in the process of collecting additional data to support our proposal;
4. **Power supply efficiency:** Per our Draft 1 comments, we recommend EPA tightens low-load efficiency requirements. We are in the process of collecting additional data to better support our proposal;
5. **Sleep power management:** NRDC recommends EPA defines a performance specification and provides an incentive for servers that can go into low-power sleep states when unused for prolonged periods of time while still waking up reliably and within a predictable amount of time.

Here are our detailed comments:

1. **Active and Idle State Efficiency: NRDC strongly supports EPA's Draft 2 proposal to set both active state and idle state efficiency criteria, to ensure that ENERGY STAR servers deliver on customer expectations of energy performance in the conditions they are typically operated in.**

   There are two major opportunities for reductions in server energy use in data centers:
   1. Ensuring that servers operate efficiently in the conditions they are typically operated in.
   2. Consolidating multiple lightly-loaded servers into fewer servers, such as through virtualization, as virtualized servers are typically much more efficiency per workload than dedicated servers.

   Both the Active State and Idle State metrics are important to support the first objective. Idle State remains key because SERT is more heavily focused on Active State efficiency than on

Idle State efficiency, and most servers still spend most of their time in idle or at very low load levels (55% of servers were idle or comatose in a recent study)[1].

To enable the second opportunity, it is important that the specification allows a sufficient availability of certified servers that can support high virtualization ratios, thereby replacing as many dedicated servers as possible.

Per the analysis presented by EPA at the Draft 2 webinar, EPA's Draft 2 approach of using two metrics supports both opportunities, by covering both idle and active states efficiencies and not overly restricting the availability of high-performance configurations.

NRDC is performing its own analysis of the qualified product list to further validate this. We are aiming to complete our analysis by mid-September and will share our analysis with EPA and stakeholders at that time.

2. **Categories and adder framework: NRDC supports the consolidation from 5 to 3 categories proposed in in Draft 2. We also encourage EPA to revise categories to better align with performance so that ENERGY STAR is better able to differentiate on idle efficiency in a performance-based manner.**

   Industry has indicated that some of the higher-performance servers would be affected if idle state requirements were tightened too much, because such servers have features, such as 16-core processors, that increase idle power draw relative to low performance servers.

   NRDC supports performance-based energy efficiency standards, that recognize the most energy efficient products among products of similar performance, and does not discriminate on performance. When certain features or performance characteristics substantially increase idle power draw across all products in the market, these features need to be accounted for in the specification, either through categorization or through adders. Failing to do so can leads to higher allowances across the board, just to ensure that the specification does not eliminate all products with certain features or performance levels.

   We encourage EPA to consider revising categories to better align with performance so that ENERGY STAR is better able to differentiate on idle efficiency in a performance-based manner.

[1] Koomey J., Taylor J., "Zombie/Comatose Servers Redux," April 2017, http://anthesisgroup.com/wp-content/uploads/2017/03/Comatsoe-Servers-Redux-2017.pdf

3. **Memory adders: NRDC appreciates EPA reducing memory adders to 0.125 watt/GB in Draft 2, but this is still far higher than required by current mainstream technology. This is particularly problematic for configurations with large amounts of memory. We propose to use a hyperbolic tangent equation to better reflect the fact that power draw does not scale linearly with capacity, and to ensure that high-memory configurations are appropriately handled by the specification.**

   NRDC collected data on memory power draw by using HPE's online server configurator Power Advisor Utility[2] which calculates the increase in power resulting from adding various amounts of memory to a server. We chose a sample recent HPE server model **ProLiant DL380 Gen9**, and configured it with various amount of memory in addition to the 8 GB in the baseline configuration, with results shown in Table 1:

Table 1 – Idle Power Impact of Adding Memory in a Server

Proliant DL 380 Gen9	**Additional memory (GB)**	**Idle Power (W)**	**Incremental power (W)**	**W/GB**
Base config (8 GB)	0	43.11		
+ 8 GB	8	43.28	0.17	0.021
+ 16 GB	16	43.28	0.17	0.011
+ 32 GB	32	43.55	0.44	0.014
+ 64 GB	64	44.04	0.93	0.015
+ 128 GB	128	44.69	1.58	0.012
Average				**0.015**

While this is only for one sample server (albeit a mainstream and high-volume product), this exercise suggests that the memory allowance should be around **0.015 watt per GB, less than one tenth that currently proposed**. NRDC is in the process of collecting additional data on memory power demand, and will share our findings with EPA and stakeholders within the next few weeks.

While overly high memory allowances have limited impact for configurations with low amounts of memory, the impact becomes significant for servers with high memory capacity. The ENERGY STAR v2 Qualified Product List (QPL) shows that half of qualified models have more than 64 GB, a quarter more than 300 GB, with some as high as 6144 GB.

[2] https://www.hpe.com/us/en/integrated-systems/rack-power-cooling.html

Figures 1 & 2 – Memory Configurations in QPL

Installed Memory (GB)	Number of Configurations in QPL
0-64	388
64-128	111
128-192	105
192-256	24
256-320	65
320-384	3
384-448	51
512-576	48
704-768	1
768-832	36
1024-1088	20
1472-1536	2
1536-1600	19
2048-2112	4
3008-3072	1
3072-3136	4
6080-6144	2
Grand Total	**884**

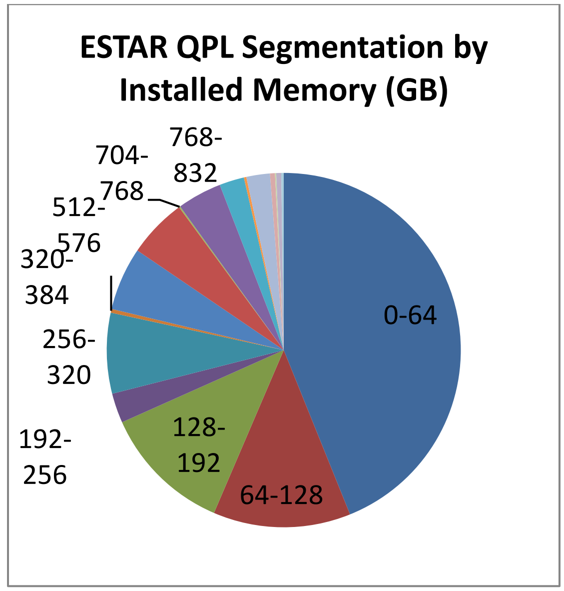

Table 2 shows that overly high allowance of 0.125 W/GB vs. 0.015 W/GB can result in several hundred watts of extra allowance for servers with large memory capacities, which dwarfs all other allowances, effectively giving those servers a free pass even if they are very inefficient.

Table 2 – Impact of overly large memory allowances on total allowance

Memory (GB)	16	32	64	128	256	512	1,024	2,048	4,096	6,144
0.125 W/GB	2.0	4.0	8	16	32	64	128	256	512	768
0.015 W/GB	0.2	0.5	1	2	4	8	15	31	61	92
Unwarranted allowance (W)	1.8	3.5	7	14	28	56	**113**	**225**	**451**	**676**

We recommend the one of the following options to address this issue:

1. Preferably, use a **hyperbolic tangent equation** to reflect the fact that power draw does not increase solely by memory capacity, but also by DIMM, and that large capacity DIMMs, while drawing high power than small capacity ones, are more efficient on a per gigabyte basis
2. To implement the same concept in a simpler manner, use a **2-step linear allowance**, with a proportional allowance of around 0.025 W/GB up to a certain limit (e.g. 64 GB, which would cover roughly half of qualified models), then a

linear allowance in the form of 0.015 * a + b where a is the memory capacity in GB for a > 64 GB.

We will propose specific equations when we have completed our data collection.

We appreciate concerns about potentially discouraging high memory configurations which are typically used in virtualization hosts and can save energy by consolidating under-utilized servers. But this does not mean all high-memory servers should have a free pass. Both objectives can be met by defining memory allowances in a way that is more reflective of actual power draw by different memory configurations.

4. **Power supply efficiency: Per our Draft 1 comments, we recommend to tighten low-load efficiency requirements. We are in the process of collecting additional data to better support our proposal.**

We support EPA's intent to strengthen power supply unit (PSU) efficiency requirements to reflect the fact that 63 percent of configurations tested in v2 achieve Platinum level.

However, we recommend that EPA considers alternative requirements that would better identify the products that save the most energy in real-world operation. As explained in point #1 above, and in our Draft 1 comments, the average load of typical servers is in the single digits or low double digits. Power supplies are typically oversized (to ensure they can support periods of maximum load reliably without burning out), and most data center servers use redundant power supplies and some of these servers share the load between both power supplies[3], leading to the load on each power supply being half of what it would be on a single power supply. As a result, server power supplies spend most of their time in the 10 to 20 percent load range.

It would therefore make sense for PSU requirements to reflect this situation by setting more stringent requirements for 10% and 20% load than for 50% and 100%, which may never be used in real-world operation. 80-PLUS Platinum criteria are more stringent at the 50% and 100% load points than at the 10% and 20% load points, compared to the mean of all the Platinum units in the 80-PLUS database to date (530 units), as shown below:

[3] https://serverfault.com/questions/659452/how-do-servers-with-redundant-power-supplies-balance-consumption

Load:	10%	20%	50%	100%
Platinum req. (single-output)	83.00	90.00	94.00	91.00
Min	78.80	90.19	94.00	91.00
Max	94.41	95.50	96.24	95.79
Mean	88.26	92.71	94.38	92.78
STDEV	2.44	0.96	0.33	0.78
Difference mean/Platinum	5.26	2.71	0.38	1.78

It should be the opposite, to better reflect the average load of servers in real-world operation. We propose alternate requirements, based on the 80-PLUS data, in the table below.

80-PLUS Platinum is also missing a 10%-load requirement for multi-output PSUs. Because of the importance of low-load efficiency in servers, we recommend EPA sets a 10%-load requirement for multi-output PSUs, and we propose a requirement of 86%, in line with other requirements.

Lastly, EPA's proposed power factor requirement of 0.9 at 50% load isn't aligned with 80-PLUS Platinum which requires 0.95 at 50%. We recommend aligning this requirement in the interest of consistency.

We therefore propose the following updated requirements (updates in bold):

Efficiency

Load point	Rated Output power	10%	20%	50%	100%
Ac-Dc Multi-output	All output ratings	**86%**	90%	92%	89%
Ac-Dc Single-output	All output ratings	**88%**	**92%**	94%	91%

5. **Sleep power management: NRDC recommends EPA defines a performance specification and provides an incentive for servers that can go into low-power sleep states when unused for prolonged periods of time while still waking up reliably and within a predictable amount of time.**

   The biggest opportunity for energy savings in servers, possibly larger than all the other opportunities pursued in Draft 2 combined, is to get servers that are remain idle for long periods of time, to power down into very low power modes, where they can still wake up reliably and within a predictable amount of time when needed. Data centers are provisioned for peak load, with some additional safety margin, to ensure that they will be able to respond to peak load events, even if those only happen a few dozen hours per year. During the remaining 8,000+ hours in the year, these servers remain idle, or very lightly utilized, but still consuming large amounts of energy both for running idle and for keeping them cool. Many of these servers could be placed in deep sleep states, while still maintaining sufficient data center capacity to serve current workloads at all times. Unfortunately, this is far from standard practice in data centers,[4] in part for fear by data center operators that the servers may not wake up in time.

In addition to raising awareness of these efficiency best practices among data center operators, there is an opportunity for ENERGY STAR to address operator concerns and provide them with higher confidence that power management will not adversely impact their operations. This could take the form of a performance specification, including connectivity and wake capabilities, reliability and latency of wake-up from sleep, as well as sleep power levels. EPA would provide an incentive for such capabilities, rewarding manufacturers who design and test for this specification. Over time, we expect this specification would become an industry standard, it would facilitate training on, and adoption of, power management across the data center industry.

Thank you for the opportunity to participate in this specification development process and for your consideration of our comments.

Sincerely,

Pierre Delforge
Director, High Tech Sector Energy Efficiency
Natural Resources Defense Council

[i] Shehabi et al., "United States Data Center Energy Usage Report", June 2016

[4] https://www.nrdc.org/experts/pierre-delforge/new-study-americas-data-centers-consuming-and-wasting-growing-amounts-energy