UK Council of Data Centre Operators

Communication:

Idle Limits for Servers

February 2017

Introduction

The data that powers and enables our digital economy is processed, stored and managed by computer servers, which may be consolidated together into purpose built facilities (data centres) remote from the businesses they support, or located in-house in server rooms and cupboards ("Distributed IT").

Servers come in many shapes and sizes and the current industry trend within the data centre environment is towards fewer, larger and more powerful devices with higher processing capacity. These can deliver economies of scale because one large machine has the processing capability of multiple smaller machines. This trend has been driven by the increasing demand to compute data and by the consolidation of multiple workloads onto single servers (eg through virtualisation). In turn this has driven a growing need for machines with high performance computing capabilities. It is also driven by cost considerations because these larger machines are more energy efficient (measured by work per unit of energy consumed).

Proposed power thresholds for idle servers

Server utilisation (how busy the server is and how much of its capacity is taken up by work) varies over time and depends on the applications and services it supports. Servers are not busy all the time. A server that is not doing any work is said to be in "idle mode" but it still uses some power. The European Commission has the laudable intention of decreasing overall energy consumption used by servers. Their proposals, under the EcoDesign Directive[1], include setting maximum thresholds for power use by servers when in idle mode. Idle mode power limitations are vocally supported by NGOs who are also keen to minimise unproductive energy use.

Industry fully supports improved energy efficiency and carbon productivity but does not support idle power limits because there is, unfortunately, a problem with the Commission's legislative proposals. Setting idle power limits is not an effective way to decrease total server energy consumption and will lead to unintended consequences.

The idle limits that are being put forward are biased toward lower performing, lower power servers. Focusing solely on idle without considering performance capability will essentially preclude the sale and deployment of many high performance servers. Think of a bus which has a larger engine and higher fuel consumption than an individual car but can transport more people at once and is therefore far more efficient than using multiple cars to do the same job. Legislating idle limits is the equivalent of banning buses in favour of cars. It will also drive the market in the wrong direction, away from larger machines that can consolidate the work, and will encourage a proliferation of smaller devices. The workload remains the same regardless of whether it is performed by one high performance server or multiple low performance servers but the total idle power of multiple low performance servers will be greater than that of the individual high performance server they replace (see worked example in Table 1).

The result will be an increase in net energy and resource use because more devices take up more space, have greater collective idle and active power demand, contain more components and have higher embedded energy and material requirements.

[1] The EcoDesign Directive aims to improve the energy performance of a range of devices, grouped by product type into different lots. Servers and storage are being addressed by DG GROW through Lot 9).

techUK Data Centre Council observations

In view of these facts, techUK's Council of Data Centre Operators makes the following observations:

- Server manufacturers and operators are highly cognisant of the energy performance of servers. Energy is the biggest operational costs for data centres. Consequently there is strong market competition on server efficiency.
- The Commission proposals are likely to be counterproductive in terms of sector energy efficiency.
- The proposals are likely to create market distortion, favouring less efficient devices.
- The proposals fail to recognise that data centres are complex environments or that a system approach is necessary to achieve good energy stewardship.
- Data centre operators make extensive use of relevant standards and best practices such as the European Code of Conduct for Data Centres and various Green Grid tools, which provide systemic and auditable energy management tools more appropriate to this complex environment than one-size-fits-all requirements targeted at individual elements.
- The projected savings will not be achieved by implementing the policy proposals.

Worked example - Table 1

		SERT Active Efficiency Score (Perf/W)	Idle Power (W)	Pass or Fail Idle Limit	Number of Servers to Deliver Workload	Net Idle Power (W) of Servers Required to Complete Work
1 Processor Socket Server	High Performance Server	75.41	156	Fail	1	156
	Low Performance Server	13.23	38	Pass	22	843
2 Processor Socket Server	High Performance Server	99.5	178	Fail	1	178
	Low Performance Server	9.2	91	Pass	31	2800

Power demand (see table 1)

techUK, in conjunction with DIGITALEUROPE, has evaluated a dataset of over 100 one and two processor socket servers to understand the impact of the server idle limits proposed in the Lot 9 Server Energy Efficiency Requirements. We have applied the proposed EU idle adders and idle limits to the dataset and determined which servers would pass and which would fail. The examples we have provided in both cases compare the most inefficient server (as measured by the SERT active efficiency metric) that passes the idle limit to the most efficient server that fails the idle limit.

What we find is that the regulation, as proposed, will remove some of the most efficient servers in the market and result in an increase in data centre power use. This occurs because more, less efficient servers would be required to do the same workload in the data centre. As a result the combined power consumption and idle power demand will significantly exceed the power requirements and idle power demand of the more efficient server despite the fact that individually it has a higher idle. In the case of the two socket server, the least efficient server requires 16 times as much power as the most efficient server

In each case, the requirement for multiple servers to do the same amount of work has further repercussions on space and cooling requirements

Space: The fact that multiple smaller servers take up more space can also be demonstrated here. A small server resembles a pizza box in dimensions and is referred to as 1U. A double server (two pizza boxes stacked) is 2U. A standard rack[2] represents 42U or 42 x 1U servers or 21 x 2U servers. A workload that requires a rack of the most efficient servers would need at least an order of magnitude more racks to house the lower powered servers cited in figure 1 to do the same workload. A rack takes up about 1.5 square metres of floor space.

Cooling: In terms of cooling required, assuming the PUE[3] is 1.6, which is about average for a commercial data centre, each KW of energy to run the IT requires a further 0.6KW of power for ancilliary functions like cooling. So there will be an additional power demand to factor in, both for the aggregated demand of multiple low power servers and the demand of the single high powered device, the total power demand being 1.6 x the server power demand in each case.

About the UK Council of Data Centre Operators

techUK's Data Centre Council comprises twenty individual members who represent the full spectrum of business interests and business models across the data centre sector. Members include wholesale and retail colocation providers, cloud and hosting operators and enterprise providers and range from multinationals to SMEs. Some members specialise in the provision of professional services to data centres such as lawyers, surveyors, investors and advisors, some manufacture the IT and communications hardware that occupy these facilities and others represent the data centre supply chain. The Council is a decision-making body providing strategic direction for all techUK's activity relating to data centres. Formal Terms of Reference provide governance for the group.

The Council was established in 2009 in conjunction with the British Computer Society (BCS). Its primary objective was to provide a representative voice for the sector in policy matters, particularly those relating to energy and carbon taxation. Over the last five years the Council has been responsible for delivering a number of significant outcomes for the UK data centre sector. These include negotiating a Climate Change Agreement for Data Centres, limiting the impact of the Carbon Reduction Commitment, building a qualification framework to recognise professionalism in the sector, demonstrating the economic value of the sector to Treasury and BIS and demystifying data centres to policy makers across government. The UK has the largest data centre market in Europe by a significant margin and as a result the Council also takes a close interest in EU policy developments impacting the sector.

Comprising senior decision makers, the Council is the single most influential body representing data centres in the UK. The Council is chaired by Andrew Jay, Executive Director at CBRE and the vice chairman is Rob Coupland, MD of Digital Realty EMEA. A list of members, terms of reference, achievements and other Council communications can be found here: https://www.techuk.org/focus/programmes/data-centres/groups/data-centres-council

Contacts: Emma Fryer, Associate Director, techUK

T: +44 (0) 1609 772 137 M: +44 (0)7595 410 653 e: emma.fryer@techuk.org

About techUK

techUK is the trade association representing the digital technology sector in the UK. techUK represents the companies and technologies that are defining today the world that we will live in tomorrow. The tech industry is creating jobs and growth across the UK. In 2015 the internet economy contributed 10% of the UK's GDP. 900 companies are members of techUK. Collectively they employ more than 800,000 people, about half of all tech sector jobs in the UK. These companies range from leading FTSE 100 companies to new innovative start-ups. The majority of our members are small and medium sized businesses. www.techuk.org

[2] A rack is a purpose built structure to house servers, generally in a vertical stack where they can be slotted in and connected to power, network and other services.

[3] PUE or Power Usage Effectiveness is the ratio between total energy demand of the facility and the energy demand of the IT within it. The lower the PUE, the more efficient the facility is considered to be. PUE should be measured on the basis of cumulative energy use averaged over a time period such as a year. See the Green Grid for PUE metrics.