White paper

NTT

The PUE factor

For many years bragging rights within data centers have centered on load capacity or W/m². In the mid to late 90s the norm was 250 W/m² for colocation operators, with some progressive firms building out a »massive« 500 W/m², which would clearly future proof their facilities for the next 20 years.

As a result, over the last decade, conversations between operators and designers have been dominated with W/m² claims and counter claims. However, in the last couple of years a new metric has come to prominence, Power Usage Effectiveness (PUE). Unfortunately, much like with W/m², there are plenty of ambiguities surrounding PUE, and plenty of opportunity for debate. We'll try and explore some of the issues and pitfalls in this article.

Definition of PUE

It's best to start with a formal definition of Power Usage Effectiveness. PUE is defined by the Green Grid as:

Power Usage Effectiveness

$$\text{PUE} = \frac{\text{Total facility power}}{\text{IT equipment power}}$$

The Green Grid goes on to describe the elements above:

IT equipment power: This includes the load associated with all IT equipment, including compute, storage and network equipment, along with supplemental equipment such as KVM switches, monitors, and workstations/laptops used to monitor or otherwise control the data center.

Total facility power: Everything that supports the IT equipment load, including:

- Power delivery components such as UPS, switchgear, generators, PDUs, batteries and distribution losses external to the IT equipment
- Cooling system components such as chillers, computer room air conditioning units (CRACs), direct expansion air handler (DX) units, pumps and cooling towers
- Computing, network and storage nodes
- Other miscellaneous component loads such as data center lighting

So PUE is a metric that gives us an indication of how much power is used by a data center in relation to its IT equipment load. In order to illustrate this, the best theoretically achievable number would be 1.0. Then no additional power is required for the data center other than that used by the IT equipment.

This would only be achievable in a lossless system (i.e. no transformer, transmission or UPS losses) and is impossible to achieve, but practical engineering solutions exist that enable a design PUE of close to 1.0. PUEs of legacy data centers can be as high as 5.0, while new designs can be as low as 1.21 (recently proposed by Google). In reality, most legacy data centers will fall in the 2.5 to 3.0 region, while those built more recently will fall between 1.5 and 2.5.

Why is there such a range of PUEs? No two data centers are the same, with differences in design, tier rating and location all having an effect on the PUE.

How to measure PUE effectively – new/old

Knowing what PUE is, how do we actively measure it in the field? With a new facility, this is a simple task.

In essence, if you can measure it, you should meter it. IT equipment loads can be measured at the PDUs, while total building load can be measured at the main switchboards.

The cost of installing such metering at the construction stage is significantly less than the cost of retrospectively fitting them into a live data center environment. An additional benefit of flooding the facility with services metering is that the equipment will provide invaluable management information for the lifetime of the facility.

Legacy data centers pose more of a problem. The principle issues here are that the data center environment is live, often old, perhaps inaccessible for maintenance and occasionally part of a larger building (such as an office). Here we have to work on the basis of maximum information for minimum risk. The key is to find out what the IT equipment load is. If possible fit metering to the data hall DUs, data which can be remotely accessed, otherwise a manual method of monitoring will be required. Meter all non data center loads (so as to work out what proportion of the power is not part of the measurement). These loads are often less critical and therefore less risky to interrupt. The final option is to guess. If there are shared systems (e.g. cooling plant or air handling) between data center and office, make an educated assumption and continue to use it for the purpose of measuring your PUE and therefore monitoring your data center's efficiency.

How to use PUE

So what does PUE tell us about our facility? How does PUE help us understand our facility? What can be done to reduce our PUE? Why does data center X have a better PUE than data center Y? These are all common questions that have been asked since the PUE metric rose to prominence.

PUE, once measured, is a quick 'visual' check of a data center's efficiency. Its use is much like that of miles per gallon (mpg) to a car. It is useful to the owner of the car to know if mpg/l has been getting better or worse, and similarly PUE is very useful to the facilities managers to assess the performance of their own sites. PUE allows the operator to find solutions to improve their facility and then to measure the effect of any change. However, the key to effective change management lies in the quantity of metering installed throughout a facility. A change in one area might lower the PUE of the facility but might be unduly stressing other parts of the installation.

Reducing PUE can have a significant effect on the operation of your data center, especially as it relates to running costs. A lower PUE means that your data center is running more efficiently, therefore lowering power costs. If you are building a new facility, a low PUE design will mean that less power is required for the site and may therefore make the grid connection cheaper, or even make a previously rejected site viable again. For IT managers, however, if less power is being used to run the data center, then more IT equipment power can be used.

This can prolong the life of a legacy data center by allowing additional or more powerful IT equipment to be installed while minimizing the requirement for infrastructure upgrades.

Practical example

Let's imagine two mega data centers identical in every way save their incoming power supply (Site A at 33 kV, site B at 132 kV). Which has the better PUE?

- Measured immediately downstream of the substation transformers (i.e. at the 11 kV switchboards); the PUEs are exactly the same.
- Measured at the point which the power crosses onto the site (i.e. the 33 kV breakers for site A and 132 kV breakers for site B); the extra losses for the site B transformers are likely to make its PUE marginally worse.
- Now let's take into account the transmission losses. These will be lower on 132 kV cable than 33 kV. Therefore potentially site B moves into pole position.

So we have one data center design, a singular simple difference and three ways of determining which has a better PUE. And we haven't even discussed the length of the supply cables.

Comparing PUEs

Thinking back to our car and mpg/l analogy, can a variety of cars be compared effectively using only their mpg/l?

No. If this were the case, then most manufacturers would find ways to claim an mpg/l much higher than the normal operating mpg/l.

Could this be a risk with PUE? In a word, yes. The PUE is a useful tool but needs to be considered as part of a greater list of comparators.

Here are some pitfalls to avoid when looking at PUE values:

- Design PUE versus measured PUE – the theoretical PUE that a design can achieve will not be reflected in practice for many reasons. Where possible, measured PUE should be used.
- Time of measurement – no requirement exists for the PUE to be measured at a given time or to be averaged over a seasonal cycle. Therefore, the PUE will normally reflect a facility's optimum performance and not the typical case. In this way it is possible for a data center in Iceland to have a worse PUE than a data center in the UK if the measurements are taken in summer and winter respectively.
- Fuel supply – PUE is a simple calculation that looks at the electrical energy required by the facility and the IT equipment. If a gas supply were used to provide cooling to the facility via a Combined Cooling Heat and Power (CCHP) plant, this should be accounted for in the PUE, but it isn't. In fact, is a gas supply suitable considering the dangers involved with high pressure pipelines and also the resilience of supply?
- Supply voltages – if two identical data centers were built next to each other, one with a 33 kV supply but the other a 132 kV supply, which one would be more efficient in its entirety? The PUE doesn't take power transmission into account.
- Resilience – a tier 4 data center will undoubtedly have a worse PUE than a tier 2 facility. A like-for-like comparison should be made when using PUE as a metric.
- Campus, mega or mini data center? Is a small data center with a PUE of 1.3 more sustainable than a large data center campus with a PUE of 1.5? Efficiencies of scale need to be considered with regard to the entire construction and operation of a facility. Is providing 100 x 1 MVA power supplies more sustainable than 1 x 100 MVA supply?

Conclusion

On balance, it's clearly a good thing to have PUEs and for them to be discussed. As an industry we need to strive for environmental and sustainable improvements and metrics are a useful way of bringing clarity to discussions. However, I hope we have shown that there are a number of complications and just because one data center is able to claim a lower PUE than another doesn't automatically make it superior. You need to look a little deeper, but did you expect this article to say anything different?

About the author:
***Adam Tamburini** has over 20 years of experience in the data center sector, working with end-users, telecommunication companies, IT companies and developers. He has worked on the development and construction of numerous data centers across Europe. Adam joined Global Data Centers EMEA (formerly e-shelter) to manage their developments in the UK, but for the last 6 years he has been responsible for hyperscale sales in EMEA. Now as Head of Hyperscale Sales at Global Data Centers, Adam is responsible for clients on a worldwide level. Adam is a Chartered Surveyor with a degree from Kingston University (UK).*

About the Global Data Centers division of NTT Ltd.

Global Data Centers is a division of NTT Ltd. Our global platform is one of the largest in the world. NTT is ranked as one of the top three leaders worldwide by IDC in their Colocation and Interconnection Services MarketScape, spanning more than 20 countries and regions including North America, Europe, Africa, India and APAC.

As a neutral operator, we offer access to multiple cloud providers, a large variety of internet exchanges and telecommunication network providers including our own IPv6 compliant, tier 1 global IP network. Our clients benefit from tailored infrastructure and experience consistent best practices in design and operations across all of our reliable, scalable and customizable data centers.

Visit us at our new website **datacenter.hello.global.ntt**.

NTT

Together we do great things