



Ark Continuity - Continuously Delivering high integrity data centres

Exploring the concept of comparative PUE

**Is Power Usage Effectiveness a metric only suitable for
comparing a data centre with itself?**

Dr Ian F Bitterlin

PhD BSc(Hons) BA DipDesInn MCIBSE MIET MBCS MBIFM MIEEE



Ark Continuity delivers high integrity data centres for Government and Corporate occupiers. Ark is now building its data centre campus at Spring Park, Wiltshire with the first phase opening in November 2009. The facility will be one of the most secure, available and sustainable in Europe.

This document is one in a series of white papers prepared by Ark Continuity. If you would like further information, please contact Ark Continuity on 0845 389 3355 or visit www.arkcontinuity.co.uk

Copyright © 2009 Ark Continuity Limited. All rights reserved.

The information contained in this document represents the current view of Ark Continuity on the issues discussed as of the date of publication. Because Ark Continuity must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Ark, and Ark cannot guarantee the accuracy of any information presented after the date of publication.

This white paper is for information purposes only. Ark Continuity makes no warranties, express or implied, in this document.

Ark Continuity may have patents, patent applications, trademark, copyright or other intellectual property rights covering the subject matter of this document. Except as expressly provided in any written licence agreement from Ark Continuity, the furnishing of this document does not give you any licence to these patents, trademarks, copyrights or other intellectual property.

Ark and Ark Continuity are trading styles of Ark Continuity Limited, registered in England and Wales number 05656968.

Head office: Hartham Park, Corsham, Wiltshire, SN13 0RP

Exploring the concept of comparative PUE

Is Power Usage Effectiveness a metric only suitable for comparing a data centre with itself?

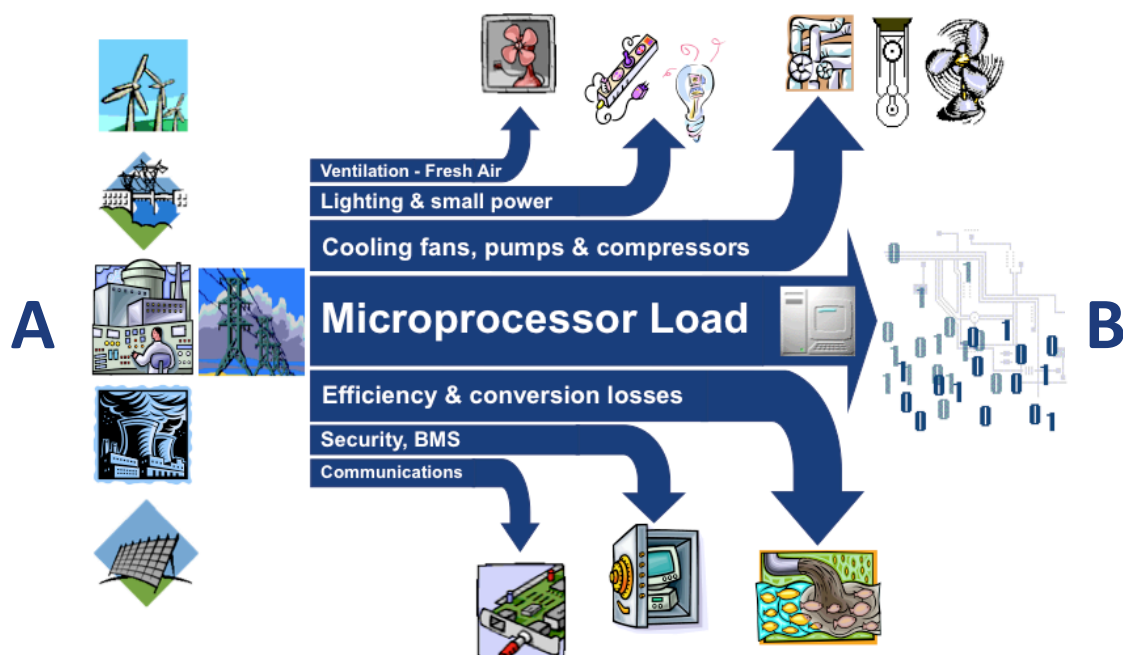
Introduction

The old adage about being able to improve only those things that you can measure has never been truer than in the data centre arena. As energy costs rise, carbon taxes become real, penalty charges loom and everybody wants to be seen to be 'green', the data centre has (rightly, as a large consumer of power) attracted the attention of the sustainability bandwagon. In a rush to be seen to be doing 'something' some trade organisations have produced metrics (measuring criteria) that can be applied to a data centre to prove that it is not wasteful of energy. If we ignore the sticky problem of 'what' the data centre actually 'does' we are left with one metric that has found almost universal favour – Power Usage Effectiveness, or PUE.

Defining PUE

The definition of PUE could not be simpler, or more open to interpretation.

PUE is simply the ratio of the total power being supplied to the whole data centre facility to the power consumed by the IT load. So it is always greater than unity and the closer to unity it gets the better the data centre is meant to be. Many people regard this as the power 'overhead' but, clearly, it makes no judgement as to the 'efficiency' of the IT load itself – a sacred cow of giant proportions.



Even the diagram above clouds the issue since it implies that 'B' is the 'microprocessor load' only, whereas in practice 'B' will always include the on-board power supplies and cooling fans of the IT 'box'.

The difference between the two power measurements includes such energy consumers as:

- Fans and pumps associated with the mechanical cooling system
- Compressor motors of chilled-water and other cooling plant
- Heaters and humidifiers
- Fans & cooling compressors associated with fresh-air make-up systems
- UPS efficiency losses
- Electrical distribution losses
- NOC including Security and Building Management Systems
- Internal & external lighting and cameras

However the losses are dominated by the first three on the list – i.e. the power consumption of the mechanical cooling plant.

One of the reasons why the definition of PUE is often blurred is best demonstrated by an example:

Consider a small data centre (incl. NOC) consuming 800kW (900kVA @ 0.89 PF) located within a larger office complex that is separately consuming 720kW (900kVA @ 0.8PF). The total load for the one facility is therefore 1520kW (1800kVA @ 0.84PF). Inside the data centre the IT load consumes 445kW (445kVA @ 1.0PF). If the facility manager can only measure the IT load and the incoming building supply, the PUE appears to be 4.05 using the kVA data or 3.42 using the kW data. If the data centre load can be isolated from the rest of the building (often easier to say than to do), the PUE can now be calculated ‘properly’ – but we have two more possible answers, 2.02 by kVA and 1.80 by kW. With results varying from 1.8 to 4.05 it is clear that confusion (and sometimes deliberate ‘spin’) is easy to create.

For completeness it is worth considering which elements are not included in the PUE metric:

- The efficiency of the micro-processor e.g. in terms of MIPS/Watt
- The efficiency of the software on the hardware platform in terms of Watts used per routine
- The efficiency of the IT loads’ on-board power supply (traditionally as low as 55% but now nearer 75% average; that is 25% losses)
- The power consumed by the IT loads’ cooling fans, often 5% of the total consumption
- Transmission and transformer losses in the grid (usually only relevant when the power is locally generated)

Mechanical strategies for PUE reduction

This paper is not specifically about how to reduce the mechanical cooling load (by increasing its efficiency or eliminating it altogether) but the following actions will, step-by-step, improve the PUE:

- Manage the air-flow
 - Hot-aisle/Cold-aisle
 - Deploy vented floor tiles in only the Cold-aisles
 - Fit blanking plates
 - Contain either the Hot or Cold-aisle
- Raise the server inlet temperature, e.g. from 20-22°C up to 27°C
- Raise the chilled water temperature, e.g. from 6/12°C to 12/18°C
- Remove all humidification (allowing 20-80%RH)

- Reduce (or remove) fresh-air make-up and associated conditioning to less than 0.1 air-change per 24h – treating the space as a machine room, not an occupied space
- Apply Variable Speed Drives to all fans and pumps¹
- Install free-cooling plant (avoiding the compressor cycle when the exterior ambient permits)
- Ultimately, apply fresh-air cooling with the air-conditioning plant turned 'off', in reserve

These strategies (all contained in the 2009 EU CoC &/or ASHREA TC9.9 2008 recommendations) will reduce the energy consumption in a data centre from 5% to over 35%, reflected in PUE reductions from a typical 2.0 to as little as 1.3.

However the data centre operator may have commercial constraints on what can be done in the facility that prevents many of these energy saving strategies from being implemented. For example a co-location facility may well have clients who will insist (based on fear or ignorance) upon a traditional SLA of 22°C±1°K and 50%±10%RH and that will be the end of any efficiency initiative. At the other end of the scale is the huge homogeneous server load of a search engine provider that can go straight to the end game of fresh-air cooling and have no air-conditioning at all – just fans, filters and a high rate of external air-flow.

Electrical strategies for PUE reduction

There are electrical measures to be taken (some only applicable for our 3-wire 208V/60Hz North American cousins) that can also contribute significant energy savings (with the typical efficiency percentage improvements):

- 400V_{AC} four-wire (3ph+N) power distribution negating the use of PDU transformers (2%) and using 230V_{AC} 1Ph IT hardware instead of 120/127V_{AC} (3%)
- 400V_{AC} input/output UPS negating the use of output and bypass transformers (1.5%)
- Line Interactive transformer-less UPS (4-6%)
- Battery-free UPS negating the need for plant-room air-conditioning (2%)
- Naturally ventilated switchgear

Taking all these measures can reduce the power consumption by up to 15% - not as dramatic as that obtainable in the mechanical arena but nonetheless substantial.

So, what is the problem with using PUE as a 'measure of goodness' and comparison between data centres?

PUE in the real world

The PUE can vary widely for a given facility based on four variables:

- Load
- Tier Classification
- Design & Operation
- Location, season & time-of-day

The first two are clearly linked: electrical and mechanical systems run at reducing efficiency as the load declines. They also have a no-load running power draw, often significantly high.

¹ To be precise, only for partial loads. If the load is 80-100% then avoiding the losses of the VSDs themselves is a good enough reason NOT to fit them. However, most facilities grow from 10% to 70% over long enough to warrant the investment Vs the energy saving gained.

The absolute load will (almost) never be 100% as, immediately, all systems will run at partial load and the theoretical PUE will never be reached in practice. As the Tier Classification increases (from Tier I to Tier IV) the level of system redundancy increases and the individual items of plant runs at ever lower load, with ever lower efficiency.

Partial load is the enemy of energy efficiency and the pure pursuit of energy efficiency is the enemy of increased reliability through on-line redundancy. As the prime design strategy of the vast majority of data centres is reliability we have a dichotomy.

The design can incorporate all of the energy strategies listed above but can also be executed in a modular architecture – so better matching the load to the capacity and keeping the individual plant load high to gain valuable percentage points of efficiency. In terms of operation it is a matter of house-keeping; making sure that all the energy strategies (such as blanking-plates etc) are followed to the letter.

As to the geographical location, the following points affect the potential PUE:

- Latitude & Altitude
- Ambient temperature
 - Without free-cooling
 - With free-cooling
 - City centre, rural
- Time of day, night-time temperature etc
- Heat gain of the building fabric (insolation)

The reader is encouraged to consider that a facility that reports a PUE of 20 (just for example) can be a more efficient design, and therefore ‘better’, than a facility that reports a PUE of 2 at the same location².

This highlights one of the largest limitations of PUE as a metric and why it should never be used to compare facilities; energy efficiency, despite increasing importance, is not the most important feature of a data centre design. The designer must follow the business needs of the client and, in the very earliest days of the design process, must establish the degree of redundancy and concurrent-maintainability that the client needs/wants. They may wish to use the Uptime/TIA Tier Classifications to benchmark the Availability level but, always, increasing the redundancy in plant and operational paths will increase the minimum possible PUE.

It should now be clear that it is not possible to compare the PUE:

- Of a Tier I/II facility with the PUE of a Tier III/IV facility at any load
- Of identical facilities in different latitudes (e.g. London and Edinburgh) &/or elevations (e.g. 100m and 1500mASL) at the same load
- Of a rural homogenous 50,000m² server-farm and a city centre 1,000m² Co-lo facility

So why do people, inside and outside of the industry, try to do just that?

The PUE metric is limited in its range but very useful in its application to a single facility. It will, if monitored continuously (or at least quarterly) and published as an averaged

² 20MVA Tier III with free-cooling at 10% load compared to a 2MVA Tier IV without free-cooling at 90% load. The larger facility will be heading towards <1.6 at 90% load whilst the smaller is running as well as it ever will. Note that at 10% load the small Tier IV could be up at a PUE of >20 without VSDs.

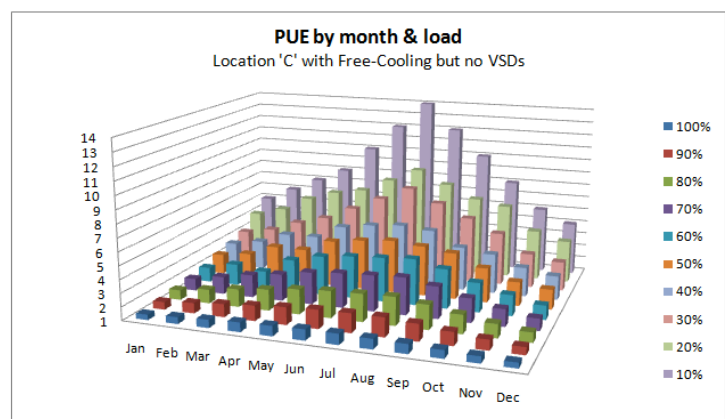
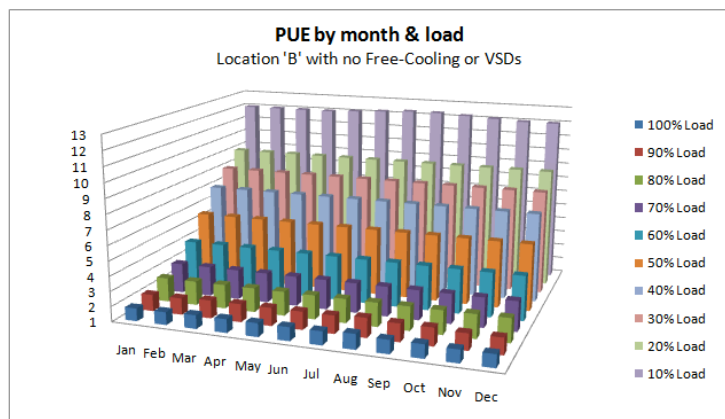
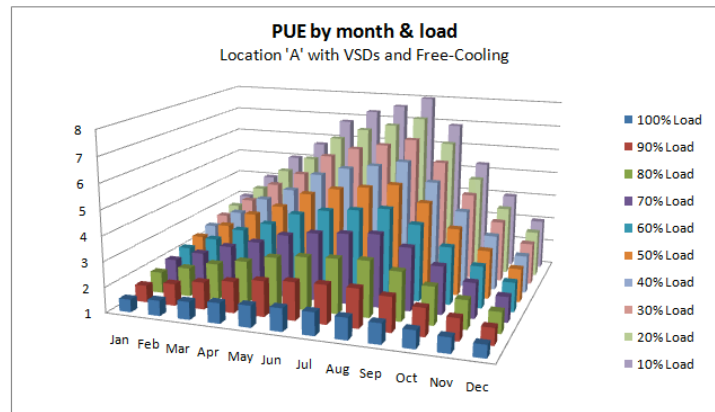
annualised figure, serve to demonstrate the improvement in the effectiveness of the M&E services. Any reported 'improvement' between adjacent quarters have to be ignored as the season will dictate the result not the facility. It does, of course, let the IT hardware efficiency off the hook completely – which is a pity, since that represents the greatest opportunity to save energy.

Illustrative Examples

The examples shown here of 'PUE by month & Load' have been calculated from basic models of Tier Classification power and cooling architectures with varying technical solutions. They each contain a modest NOC but without any additional offices or DR facilities.

To demonstrate this papers' contention that PUE should not be used for comparative purposes a PUE of 2 can be found on each facility – but at completely different set-point conditions.

Clearly, given the right climate and sufficient capital investment in the M&E infrastructure, the system that provides the best energy performance will contain variable speed drives on all pumps and fans and use free-cooling coils in the heat rejection system.



Presented by Dr. Ian Bitterlin, CTO of Prism Power Limited
at the Government ICT Goes Green 09 Conference, QEII Conference Centre, London
Tuesday 15th September 2009