KAO DATA

PUE was once just an effective tool to measure performance and efficiency. Now, during a period of high energy costs it's vital to reduce power use and realise cost savings.

Why PUE and efficiency remain essential to counter-balance rising energy costs.

Every business across the United Kingdom is feeling the crunch of rising energy costs. A survey of 200 large businesses across the UK found that 77% list energy as their number one concern across 2023/24.

Ninety percent of those businesses believe their energy costs will continue to rise and 80% believe it has become a board-level issue. Firms have already experienced a 424% rise in gas costs and 349% increase in electricity since February 2021, according to the Federation of Small Businesses (FSB), and prices are expected to continue to increase for the foreseeable future.

Now, more than ever, businesses are looking to maximise efficiencies as one strategy to keep energy costs down. Within the data centre industry metrics are already in place to help operators and their customers track power usage within the facility. While not created for the express purpose of managing an energy crisis, they can still be a very effective measurement tool.

At a time of volatility in the energy markets power usage effectiveness (PUE) can have a starring role to play. Data centre operators with a lower PUE can help customers mitigate the pain of rising energy costs. In this paper, we'll take a deeper look at what PUE is, how it's being used today and why it may be a more relevant tool than ever before.

What is PUE?

One of the most common metrics used to measure performance within the data centre environment is power usage effectiveness. PUE determines the effectiveness of a data centre by dividing the total amount of power consumed by the facility by the power used to run the IT equipment within it.

$$\text{PUE} = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

This is expressed as a ratio with the ideal target of 1.0. A PUE of 1.0 means all the power consumed by the facility was used by IT equipment. Of course, this is unrealistic as cooling, lighting, power transmission losses, and other miscellaneous services also require power which raises that figure of 1.0 northwards. Through innovative design, energy efficient infrastructure and technical operational excellence the aim is to ensure the final figure is as close to 1.0 as possible.

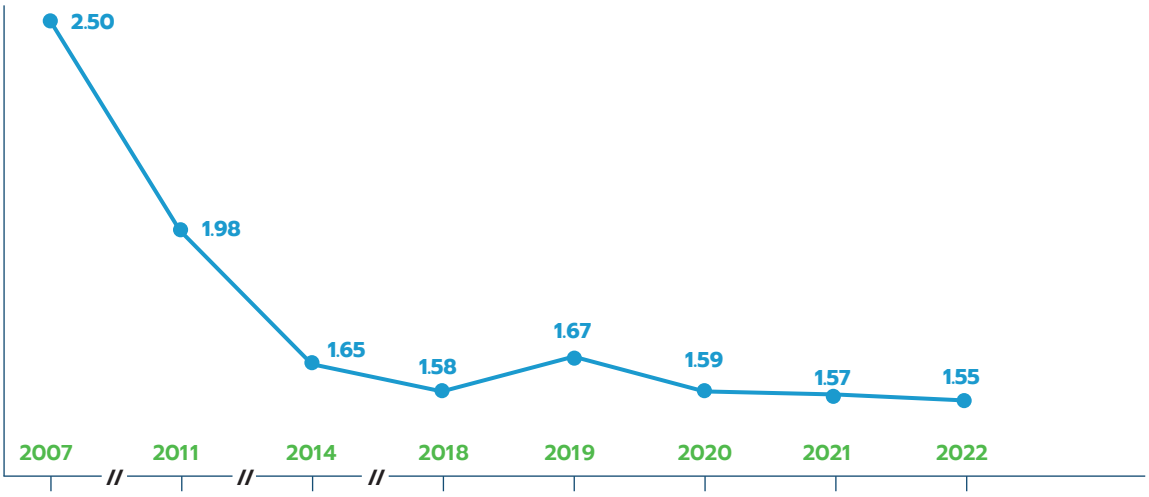
First introduced in 2006 and endorsed by The Green Grid – who later published the book 'PUE: A comprehensive Examination of the Metric through ASHRAE TC9.9, in partnership with the Green Grid' – PUE was published as an ISO/IEC 30134-2:2016 global standard in 2016. The primary benefit of PUE is its perceived simplicity. PUE is accepted by customers (and the industry) as a clear indicator of performance and efficiency. Given the drive to reduce energy usage, it is seen as one of the key ways to tackle sustainability and has become inextricably linked to green credentials. While it is considered a meaningful sustainability metric within the industry, it also has limitations.

Factors such as the geographical location of the data centre, the subsequent climate, the preciseness of the facility power calculation, as well as the inability to track power consumption at the rack level, all impact the accuracy of the measurement. In addition, PUE has, in some cases, become more of a marketing gimmick rather than the technical and operational assessment tool as was originally intended. This type of unintended consequence can lead to behaviours that the metric is not well suited for.

Where is the industry today?

According to the 2022 Uptime Institute Global Data Centre Survey, the average annual PUE in the data centre industry is 1.55. This is consistent with the trends observed by Uptime in recent years that initially saw large efficiency gains, but have remained stable in the last five years. While PUE progress has stalled, Uptime has cautioned that this should not suggest the sector has reached an efficiency limit. According to the survey, "New data centre builds routinely outperform the average, achieving PUEs of 1.3 and below using facility designs and more advanced equipment that are optimised for lower energy use."

DATA CENTRE AVERAGE ANNUAL PUE



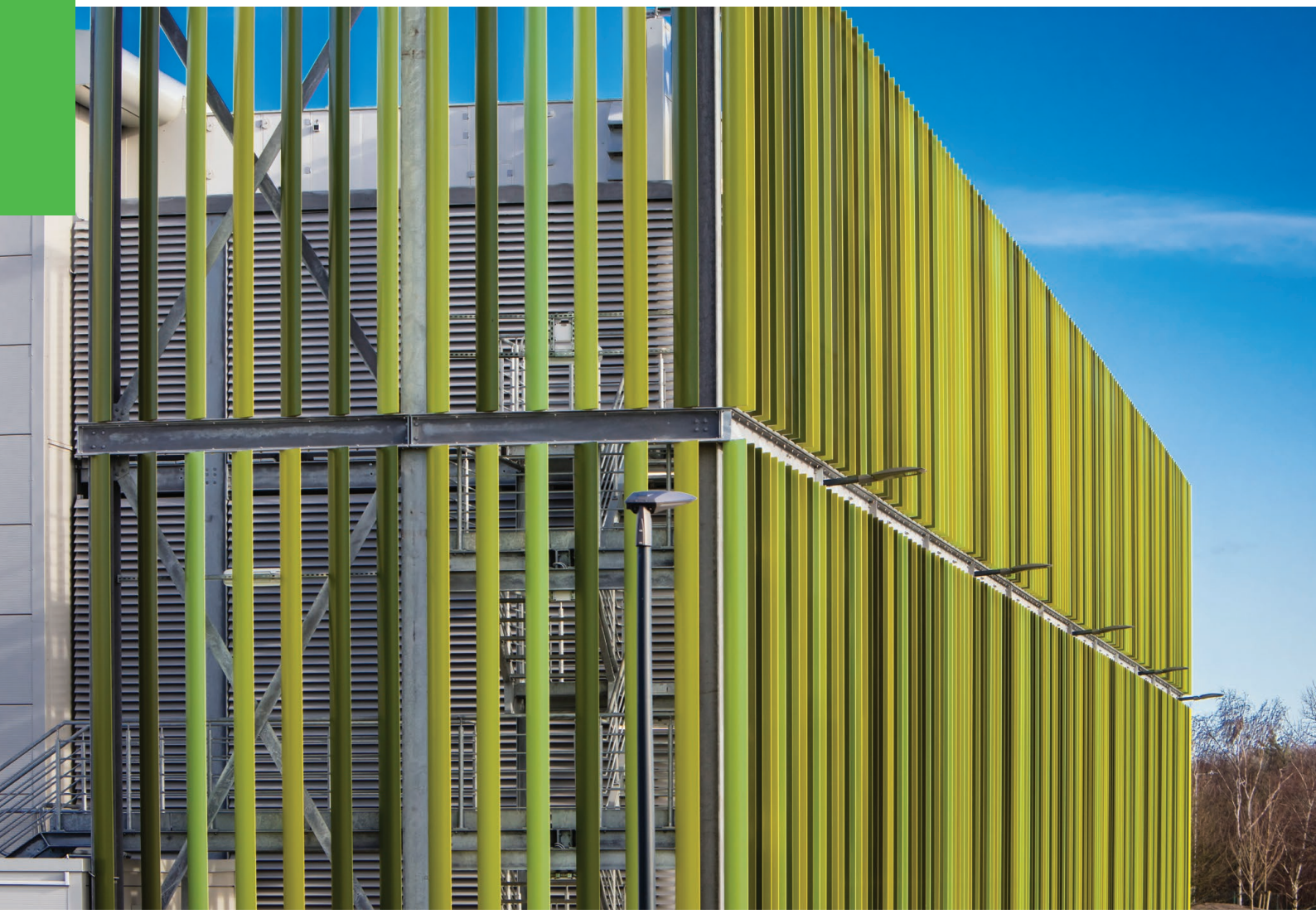
Uptime Institute Global Survey of (669) IT and Data Centre Managers 2007-2022

The International Energy Agency (IEA) concurs with this outlook. "Since 2010, the number of internet users worldwide has more than doubled, while global internet traffic has expanded 20-fold. Rapid improvements in energy efficiency have, however, helped moderate growth in energy demand from data centres and data transmission networks, which each account for 1-1.5% of global electricity use." While overall data centre energy use is likely to continue growing moderately over the next few years, IEA advises longer-term trends are more uncertain.

Part of that comes from newer, smaller data centre markets consuming more local power, particularly in locations like Ireland and Denmark. There is also the fact that while data centre facilities have become more energy efficient, IT use can still be inefficient and wasteful. Elements of the IT stack, like processors and memory, will still draw energy even while sitting idle.

As facilities become more efficient, measuring improvements becomes harder and gains become increasingly incremental. New data centre builds and technologies may see lower PUE ratings because they have the advantage of making more efficient decisions from the start – location for ambient temperature, architectural design of the building, etc.

Data centre operators can certainly retrofit existing data centres, but only to a certain degree. For example, the layout of server racks can be optimised using proven air flow management techniques. Positioning alternate rows so that all cold air intakes face one way and hot air exhausts face the other, creates a hot aisle/cold aisle configuration that helps to optimise cooling. Further supplementing with containment, results in dramatic performance both in terms of efficiency and reliability. Again, these are relatively simple and cost effective measures that can be implemented to minimise wasted energy.



Measuring energy costs with PUE

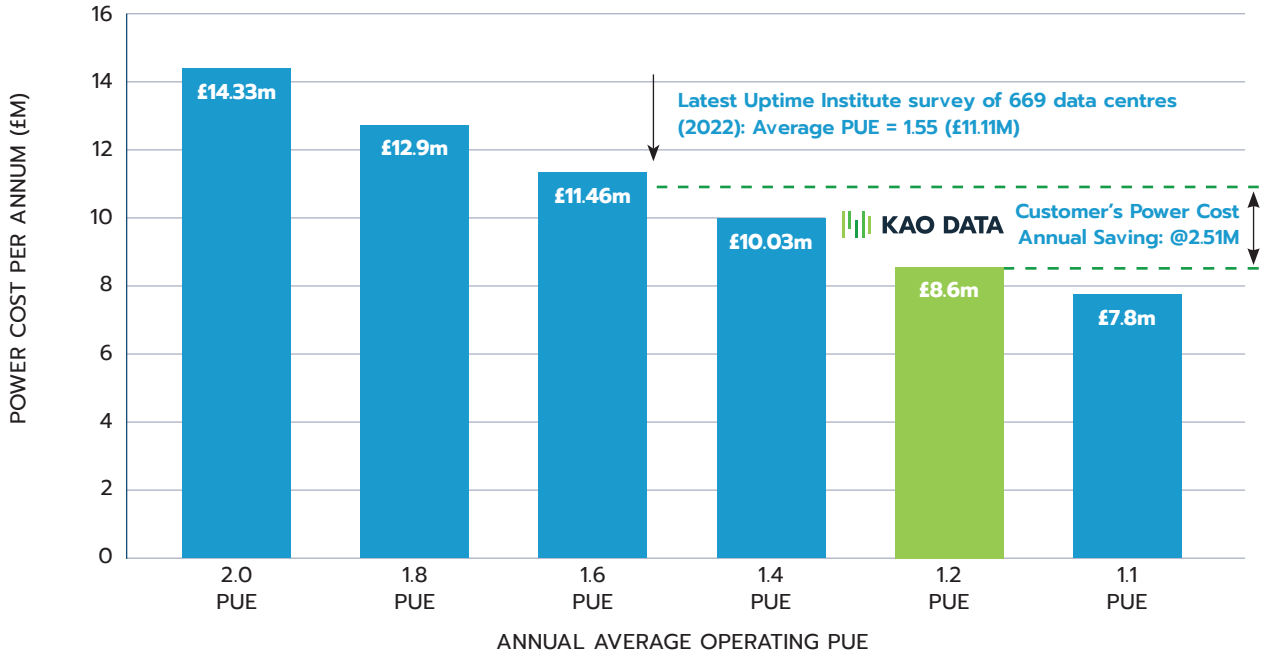
Volatility in the current energy markets creates a unique opportunity for PUE as a metric. As data centre operators and customers proportionately feel the pain of energy operating costs, those with a lower PUE will feel the impact less as they need to procure less power.

The industry typically represents PUE by considering data centre performance at the peak design or contracted kW load. This approach also aligns with the optimum operating efficiency of the power and cooling systems supporting the data centre. In theory, this can lead to exceptional PUE at 100% load. However, in practice, the data centre kilowatt (kW) load is dynamic, ranging between idle and maximum power. This has an adverse effect on the efficiencies of both the power and cooling systems which have typically been optimised for 100% kW load, resulting in proportionally higher PUE at part-load utilisation.

Through innovative design and applying engineering principals creatively, Kao Data is able to deliver hyperscale levels of efficiency from rack level to full occupancy. This also enables us to provide a contracted, SLA backed PUE ranging from 1.2 to 1.25. This includes partial loads as our PUE remains relatively constant from 20% to 100% of the IT load.

PUE COMPARISON & OPEX ADVANTAGE

Kao Data KLON-01 (8.8MW Data Centre Facility, Harlow, UK)



ANALYSIS BASED ON:

- Operating 8,760, hours per year
- An 8.8MW data centre facility at 85% utilisation
- £0.1094 kWh average UK power utility cost (2020)
- Average industry PUE of 1.55 based on Uptime Institute's 2022 survey of 669 data centre respondents

For a typical customer, this delivers potential energy savings against legacy solutions in the order of 23% to 25%, resulting in annual power cost savings of £2.51M. As power prices rise, savings based on a lower PUE will continue to increase. At Kao Data we provide an SLA-backed PUE, delivering certainty and guarantee, which in a volatile energy market is exactly what customers need.

With energy prices soaring, we have seen a re-emergence of PUE, as it forms an essential part of the tool set available in helping customers to make sound business decisions. Designing our data centres to ensure a lower PUE is just one of the many ways Kao Data enables total cost of ownership (TCO) savings for our customers.

"PUE has long been the cornerstone metric upon which data centre performance has been predicated during design and measured during operations. Without its creation, the environmental impact of our industry would be considerably higher," said Paul Finch, COO, Kao Data. "While more work remains to be done regarding sustainability metrics, PUE has served the industry well and provides a key method of tracking power consumption relative to compute, storage and network performance. Amid the turmoil of an unpredictable and volatile energy market, it remains one of our best weapons to counterbalance rising energy costs."

What comes next

There are a few underlying industry trends that will determine what the future holds for PUE. As Uptime Institute noted in its 2022 survey, future efficiency gains will need to focus on IT power, requiring additional metrics to supplement PUE. Measuring the IT efficiency is likely to lead to a collection of indicators for server utilisation, hardware, power management, etc. We are already starting to see this with the adoption of both water usage effectiveness (WUE) and carbon usage effectiveness (CUE) as metrics to compliment PUE.

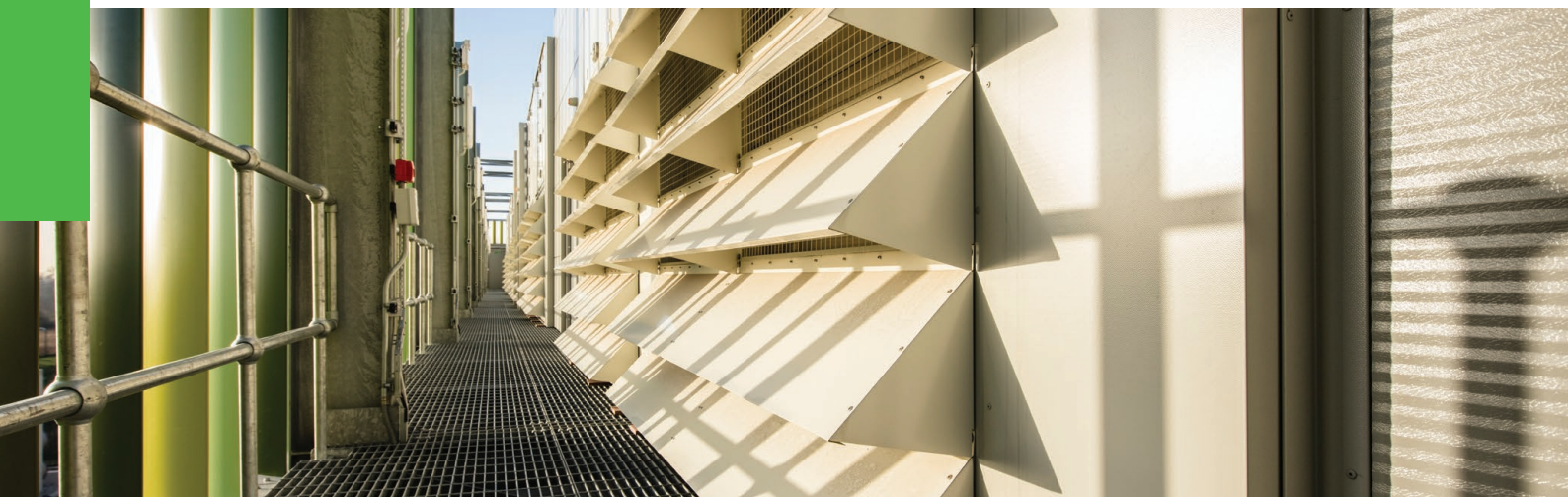
High performance computing (HPC), artificial intelligence (AI), and machine learning are also changing the way companies across all industries are doing business. These high-intensity workloads are increasing business efficiencies, improving decision-making and generating new revenue streams. They are also redefining the compute boundaries of CPUs and GPUs, leading to increased rack density and pushing the limits of traditional air-cooling technology. As a result, liquid cooling is poised for mainstream adoption in the next few years. Jacqueline Davis, research analyst for the Uptime Institute, goes further and explains that direct liquid cooling changes enough variables outside the scope of PUE that its application as an energy performance and efficiency metric may become less applicable.

What's clear is that the topic of PUE as a sustainability metric has been debated by many of the industry's engineering leaders. However, recognising the absence of a PUE metric likely causes more harm than good. Davis suggests we may also need a new metric for energy efficiency, if not as a replacement for PUE, then as a supplement.

Despite that, the energy crisis right now, is making business decisions extraordinarily complex. Any tool that is able to help guide an organisation on their power consumption still has a role to play.

With so many considerations tugging at business' attention, refocusing on PUE can still be a useful tool to simplify data centre decision-making in the short-term, while leading to considerable long-term savings through measurement and continuous operational improvements.

Switching your compute from an average PUE of 1.55 to 1.2 can save you 23-25% on your energy utilisation cost. At a deployment of 8.8MW that's a massive 2.51M!



Six practical steps to make PUE more efficient

- Dust off the calculator and work out your estate's PUE. First, determine how you are going to measure PUE by considering guidance set-out within ISO/IEC 30134-2:2016. If your legacy infrastructure doesn't have the sub-level metering required, record the metering points that are available. You might be surprised to find out how genuinely efficient – or likely inefficient – your compute and data centre facilities really are.

- Small changes can make a big difference but improvement measures should start at the IT infrastructure level, as these will not only improve energy-efficiency but also reliability. Ensure the hardware installation complements a hot and cold aisle arrangement. By installing blanking plates, brushes and seals to reduce air leakage, and utilising hot or cold aisle containment to further promote the segregation of server inlet and outlet air flows, you can improve efficiencies. Further, identify the tweaks within your architecture design to address root-cause issues, thereby reducing your PUE and driving your energy procurement lower.

- When the IT infrastructure has been addressed, the next step is to look at the engineering infrastructure supporting it. Fan speeds may now be able to be reduced and supply air set-points raised, which could then lead to increases in chilled water set-points. This will reduce the refrigeration demand across the facility and subsequent motive power driving the compressors. Greater temperature or pressure differential control can be established due to the improvements in containment, delivering a more dynamic control solution, and better match the IT demands.

- Look at increasing your cloud workloads. Due to their tremendous economies of scale and sheer size, hyperscale cloud providers run some of the most efficient data centres on earth. Take a look at your workloads to determine which need to be kept on-premises and which can be moved to the cloud, not forgetting to factor in data transit costs which all form part of the wider TCO analysis.

- Alternatively, and for compute that isn't best located in the cloud, look at migrating your compute to an industrial scale colocation facility. At Kao Data, we can provide highly optimised data centre suites, delivering SLA-backed PUEs of 1.2 via advanced design and technical excellence across our high performance infrastructure platform.

- The Green Grid, ASHRAE, and Uptime Institute all have excellent resources for further reading about PUE, with more to come from the Climate Neutral Data Centre Pact in the near future.