A RATING SCHEME FOR DATA CENTRES

TECHNICAL DISCUSSION PAPER

CLIMATE NEUTRAL DATA CENTER

1 Executive summary and purpose of this document

We are all familiar with A-G energy efficiency labels on domestic goods like fridges, washing machines and TVs. We are also used to energy performance certificates (EPCs) for buildings. These rating scales were invented for a good reason: to help consumers identify the most efficient devices and buildings so they can make informed buying decisions.

So developing a standardised energy rating scheme for data centres seems like a logical step to take, and pressure to take it is growing from regulators and policy makers.

However, the challenge of developing a robust rating scheme is a significant one, and has occupied industry experts for decades. The sector has developed an impressive range of standardised metrics that indicate sustainability performance in a growing number of specific areas – infrastructure efficiency, water use, renewable power adoption, and many others. There is however, no combined metric that provides an overall view. There is not even, as yet, a standardised metric to measure actual productivity in terms of data centre IT functions. This is not because of laziness or lack of ambition, but because of three broad factors: first, the intensely complex nature of the problem; second, the multiple technologies and variety of computing functions that may be housed within individual facilities; and third, the astonishing speed of change.

The overall problem has the complexity of a Gordian knot, but these are not the days of Alexander the Great. A decisive and authoritative attempt to resolve a highly complex problem with a simple and radical solution will deliver the equivalent of lots of little bits of string – along with confusion, perverse outcomes and a discontented industry. Our objectives, therefore, are to support those tasked with developing a rating scheme for data centres, and to ensure that they benefit from the knowledge and experience of the sector in developing a workable solution.

This discussion paper represents the views of the signatories to the Climate Neutral Data Centre Pact (CNDCP) on developing such a scheme. The paper:

- Sets out some of the challenges involved in designing a scheme that is universally applicable and the precautions necessary to avoid unintended consequences.
- Explains the market and operational complexities of the sector and discusses the pros and cons of alternative approaches.
- Warns that expectations must be managed around how feasible it will be to develop an approach that works, considering the complexity of the sector and the multitude of operational models and computing activities it encompasses.
- Summarises the position of Pact signatories on each of the discussion points.

Broadly speaking, Pact signatories agree that a consistent, robust and holistic approach to evaluating the sustainability performance of data centres would be a good thing. The CNDCP has been working on metrics relating to infrastructure that can inform this process. However, the IT function within data centres also needs to be scrutinised for any rating to be meaningful.

The variety of business models in the sector – including many where the infrastructure and IT functions are operated by different entities – creates reporting challenges that must be addressed. Reporting obligations placed on operators must not extend beyond the functions within their direct control. Any scheme that rates organisations on activities outside their operational remit is unfit for purpose.

In summary, Pact signatories request:

- A cautious and considered approach in view of the market and operational complexities of the sector.
- Clear objectives, which must include sustainability outcomes additional to those already being delivered by market forces.

- Scope that includes all facilities where sustainability is not market driven.
- Application of lessons learned from existing rating schemes.
- Clear differentiation of data centre metrics from those used in existing rating schemes because they reflect operational, not optimal, conditions and also include infrastructural overhead.

Furthermore, Pact signatories take the view that:

- Infrastructure metrics suitable for inclusion in a scheme include PUE, WUE and REF.
- ERF does not measure operational efficiency and cannot be a core metric.
- Timely update capability will be necessary to reflect the speed of change in IT environments
- IT metrics are not yet standardised. CSERV could be used as a trial metric in the meantime.
- Reporting must be limited to activities within direct control of the reportee.
- Allowances should be in place to accommodate constraints outside operational control but should not disguise actual performance.
- Categorisation (e.g. in terms of operational model, location etc.) will be needed to ensure a fair basis for comparison.
- Confidentiality must be protected both in commercial and security terms.

The paper is structured in a modular way for ease of reference. Pact signatories look forward to working with the team tasked with developing a ratings scheme.

2 Glossary of terms used in this document

A-G Rating: A grading system, usually using labels colour bands shaded from A (best - green) to G (worst - red) to indicate the energy efficiency of domestic products.

AI - Artificial intelligence: The simulation of human intelligence processes by machines, particularly computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions), and self-correction.

Cloud computing: The delivery of different services through the Internet, including data storage, servers, databases, networking, and software. Cloud computing offers flexible resources and economies of scale.

CNDCP / Pact - Climate Neutral Data Centre Pact: A voluntary self regulatory initiative involving 85% of the EU data centre market. Signatories commit to efficiency, renewable energy and circular economy targets and develop robust sustainability metrics.

Colo or Colocation: Data centres where businesses can rent space for servers and other computing hardware. These facilities provide the physical infrastructure, such as power, cooling, and security, while clients manage their own equipment.

Colocation, retail / MTDC (Multi-Tenant data centre): These facilities host multiple customers' servers in a single data centre facility to share costs and infrastructure resources. This was the original meaning of colocation, where servers belonging to different customers were "co-located" in a single third-party facility.

Colocation, wholesale: These facilities lease large spaces or entire buildings to clients, often large enterprises or service providers that need significant amounts of space and power. Clients typically design and manage their own IT infrastructure within the leased space.

Enterprise: Enterprise data centres are essentially "in-house" facilities that are dedicated to supporting an organisation's IT functions. Both the infrastructure and the IT are controlled by the same entity.

EPC - Energy Performance Certificate: A colour coded rating system usually deployed to grade the energy efficiency of domestic and commercial buildings. A minimum standard may be required for buildings to be let or sold.

ERF - Energy Reuse Factor: A metric to evaluate the amount of waste heat exported and used. (See Section 9)

HPC - High Performance Computing: The use of supercomputers and parallel processing techniques for solving complex computational problems. Systems can handle highly complex computations that are beyond the reach of standard computers.

Hyperscale: Large-scale data centre facilities operated by major technology companies like Amazon, Google, and Microsoft that support massive, scalable applications and services, providing vast amounts of computing power.

KPI - Key Performance Indicator: A metric that is used to measure how well a product or system performs against a given criterion or set of criteria.

LCA - Life Cycle Assessment: A systematic process that evaluates the environmental impacts of a data centre throughout its entire life cycle, from construction to decommissioning. It provides a comprehensive view of sustainability performance.

On premise data centre: A data centre that is housed within corporate offices or premises rather than in a dedicated building. By default an enterprise data centre (qv).

PUE: Power Usage Effectiveness: A metric that measures infrastructure overhead (the ratio between the power consumption of the facility and the power consumed by the IT functions it houses). Design PUE indicates how well a facility could perform under optimal conditions and operational PUE is a measure of actual performance under real-world conditions. (See Section 9)

REF- Renewable Energy Factor: A metric to evaluate the proportion of power consumed by a data centre that is low carbon. (See Section 9)

Server Utilisation: How busy a server is which is an indication of the efficiency of server usage within the data centre. Higher utilisation rates indicate better resource management and reduced energy waste.

SLA - Service Level Agreement: A contract between a data centre operator and a customer. It specifies what services the operator will provide and defines the performance standards the operator is obliged to meet.

SRI - Self Regulatory Initiative: A binding commitment made on a voluntary basis, not an externally imposed legislative requirement.

WUE - Water Usage Effectiveness: A metric to evaluate how efficiently a data centre uses water. (See Section 9)

3 What is the Climate Neutral Data Centre Pact?

The Climate Neutral Data Centre Pact was established in January 2021. It is a voluntary initiative involving data centre operators of all types: enterprise, colocation and cloud service providers, plus their representative bodies. The Pact now has over 100 signatories who collectively represent 85% of data centre market share within Europe.

The Pact seeks to support the objectives of the European Green Deal in realising ambitious greenhouse gas reductions and achieving Europe's 2050 climate-neutral goal by technological and digital means. Operators and their trade associations are committed to achieving operational climate neutrality by 2030. They wish to ensure that data centres are an integral part of the sustainable future of Europe.

In line with these commitments, Pact signatories take energy stewardship very seriously. Moreover, the electro-intensive nature of our sector means that we are highly motivated to maximise efficiency and minimise the energy consumption required to operate our facilities and support their IT load.

However, the core mission of the Pact goes beyond energy consumption to address broader sustainability challenges relating to data centres, such as water use, waste, renewables sourcing and circular economy practices. Because data centres are complex environments, an exclusive focus on one area – such as energy – could compromise performance in others, such as water use.

Signatories have been developing robust industry metrics to provide a fair basis for comparison and reporting. More information on the work of the Pact can be found at www.climateneutraldatacentre.net.

4 Broad points of agreement from Pact signatories

Pact signatories recognise that the proposal for a rating scheme comprises only one part of an ongoing regulatory process. Operators representing all business and operational models across European markets have serious reservations about the effectiveness of both the legislative process and potential rating scheme. However, they acknowledge the following:

- A consistent, holistic approach to evaluating the sustainability performance of data centres is a viable policy ambition. The data centre sector has already developed such an approach through the CNDCP's Self Regulatory Initiative (SRI). We recommend that the KPIs and targets of the SRI be adopted in any future rating scheme.
- An adequate understanding of data centre energy efficiency depends on scrutiny of both the infrastructure and the IT. However, some business models control both the infrastructure and IT, while others control only one of the two. A rating scheme should only reflect what lies within an entity's control.
- While the approach should try to level the playing field as much as possible and enable continuous improvement across diverse data centre types and climates, this will be extremely challenging to implement in practice, and no approach will be ideal. The limitations of any scheme (e.g. in terms of its inability to enable like-for-like comparisons) should be made clear so that it is not misapplied.

5 Some core principles for rating data centres

At first glance, creating a rating scheme for data centres looks both straightforward and desirable: straightforward because the energy consumption of a data centre is, in principle, easy to measure; desirable because the sector is so electro-intensive that improved efficiency could deliver significant savings. A clear metric for comparison – perhaps similar to those used to categorise the efficiency of household appliances and buildings – would seem essential to any such improvement.

The fundamental problem is that data centres operate under much more complex conditions than household appliances, and, as we shall see, they are anything but ordinary working buildings.

A key consideration is that the computing technologies that data centres host are still immature, and they are evolving very rapidly. As such, we must avoid approaches that distort the market or discourage the deployment of advanced workloads like high performance computing (HPC), Artificial Intelligence (AI) or other next-generation computing activities. New technologies tend to go through developmental stages before maturing and during these transitional phases they may not perform optimally against conventional metrics. For example, some newer, high-performance processors may need lower temperatures to operate reliably, which will impact performance against metrics like PUE (see Section 9) at least in the short term.

The exceptional complexity of data centre environments, the rapid speed of technological change and the unpredictability of technological evolution mean that the design of any rating scheme will demand very careful preliminary analysis, and should meet the following conditions:

- The objectives of a scheme must be clear and should be limited to addressing market failure.
- The scope of any proposed scheme must be set accordingly, both in terms of performance elements and the entities within the sector that would be obliged under it.
- It should not duplicate or conflict with existing schemes already widely deployed in the sector.
- The development process should identify the problems that plague existing approaches and learn from them in order to define a workable approach.
- Existing metrics used in the sector should be deployed where possible.
- A successful scheme must seek to ensure a standardised basis for comparison by:
    - Accommodating variables – for example, differences in facility age, business model, location, HPC intensity, water availability and redundancy requirements.
    - Classifying data centres by activity and operational model.

Core principles - precautionary approach: Pact position

The data centre sector has spent decades developing sustainability metrics and is best placed to understand the complexity of data centre facilities and the dynamic and changing nature of digital technology.

There is very grave concern among operators that external parties less familiar with the operational, technological and market realities of this sector will develop approaches that are inadequate in terms of addressing market failure and incentivising best practice, and that these approaches will create perverse incentives and/or market distortion. **The sector urges a cautious and considered approach.**

6 Clarifying objectives

Energy performance rating schemes are generally consumer-facing and their purpose is to help people make informed choices when they buy electrical products. Their ultimate aim is to reduce domestic energy demand. By contrast, most data centres – especially enterprise, hyperscale, or HPC facilities – are not consumer-facing.

A-G style energy rating schemes for consumer devices have been very successful in encouraging manufacturers to improve the efficiency of their products during the "in use" phase. LCA (life cycle assessment) studies indicate that the "in use" phase (rather than manufacture, transport or disposal) usually accounts for the majority of energy consumed by a given product over its lifetime. It therefore makes sense to target this phase, and doing so has led to impressive improvements in the operational efficiency of domestic appliances, to the extent that product rating schemes have had to be recalibrated to account for the scale of improvement.

In essence, energy rating schemes aim to set a standardised basis for comparison that lets consumers compare different products on the basis of energy performance. They can weigh a product's rating against its purchase price, likely use pattern and life expectancy. In contrast, it is unlikely that consumer education would be a reasonable primary objective of a rating scheme for data centres for the simple reason that they are not consumer-facing.

So if the target is not educated consumers, the objective might be to improve public sector procurement of data centre services by identifying the things that procurement teams need to look for to ensure they are procuring sustainably. However, this would duplicate existing activity because sustainable public sector procurement is already being supported through the EU Green Public Procurement (GPP) criteria for Data Centres, Server Rooms and Cloud Services initiative developed by JRC some years ago. [1]

A positive outcome of some rating schemes is that the scores are used internally by companies for trend analysis - to track how the sustainability of their product or service is improving over time. In the data centre sector, most of the existing sustainability metrics are already used, very successfully, for this purpose, so a rating scheme would add little value to this function.

Rating schemes using A-G scales have also been used to develop minimum standards of the kind we see in the EU housing market, where buildings must meet a pre-determined standard before they can be leased (the challenges of imposing minimum standards on data centres are discussed below).

The Commission must ensure that the objectives of any rating scheme address market failure while also adding value. Commercial data centres are already highly incentivised to be efficient. They are energy intensive, so their competitiveness relies heavily on good energy stewardship. As such, a data centre is a categorically different product from a consumer item like a washing machine: the manufacturer of a washing machine may have a (perverse) incentive to develop a cheap, but poorly performing, product. For the operator of a data centre, no such incentive exists[2].

---

[1] These GPP criteria are already, in a handful of Member States, incorporated into national level regulations, and there is a growing trend towards such incorporation. However, even if the objective is not to inform public sector procurement, harmonisation will be needed to minimise fragmentation and the risk of conflicting requirements. Existing regulatory and soft law frameworks (such as Taxonomy, the GPP criteria) and voluntary frameworks (like the EU Code of Conduct) should be aligned with the Energy Efficiency Directive and the upcoming Rating Scheme.

[2] The perverse incentive occurs in washing machines because the cheap and inefficient model may look much more attractive to the customer at point of sale, although it will cost more in the long term to run. For a colocation data centre customer, however, it is purely a leasing model: there is no initial purchase so the only consideration is the running cost. The colocation customer has abandoned the capex element of the purchase so there is no scope for perverse outcomes in terms of the initial purchase price

Similarly, energy intensity drives cloud service and other digital service providers to focus on efficiency, whether as customers of data centres or their owner-operators. To date, this kind of customer demand has been a far stronger driver of data centre sustainability than regulation[3]. As such, any proposed rating scheme must demonstrate how it creates additional incentives and targets market failure.

Objectives: Pact position

The objectives of a rating scheme for data centres should be to discourage and expose poor operational practices, address market failure and drive ambition among operators to achieve optimal efficiency, taking into consideration the operational model and the external conditions under which they function. A rating scheme for data centres should primarily be deployed to drive best practice through trend analysis, not for the purpose of making comparisons that will inevitably be, at best, imprecise and, at worst, misleading.

---

because the data centre is not for sale. The domestic equivalent would be for the consumer to hire the washing machine, in which case they would be very interested indeed in its efficiency, because like the colo customer, they would have to factor in an upfront cost but would pay for the energy they consume when using it.

[3] For instance, the vast majority of power procured by data centres in Europe is renewably sourced. However this is driven by internal commitments and customer demand: it is not a regulatory requirement in any country.

7 Scope

In terms of scope, we need to be clear whether any proposed rating scheme relates *only* to energy efficiency or reflects performance against a broader range of sustainability criteria. It is the view of Pact members that other factors (such as water use) that also have an energy cost should be included, even if the scope is limited to their energy impacts.

We also need to determine which data centres would be obliged under a rating scheme. While large, modern facilities are already strongly incentivised to be efficient, the scale of power being consumed means that even very marginal improvements in performance will deliver significant energy savings. At the other end of the scale, there is strong evidence of market failure: smaller, older enterprise facilities tend to be the sector's worst-performers, and the aggregate potential energy savings from many small facilities is again very considerable, so the size threshold of 500kW of IT capacity that has been used for EED reporting should be removed altogether, or set more appropriately, at no more than 50kW of IT capacity.[4]

Scope: Pact position

Scope of metrics: The scope of any metrics should be focused primarily on energy performance, but could include subsidiary or bonus elements relating to broader sustainability criteria.

Scope of obligation: In terms of those parts of the market obliged under the requirement, scope should address all data centre facilities where market failure may exist.

[4] The European Commission funded EURECA project reviewed 350 public sector data centres across Europe and discovered very poor energy stewardship, especially in the smaller facilities. **See Appendix III** for more details.

8 Avoiding existing pitfalls and perverse outcomes

Despite their apparent simplicity and intuitive-seeming, colourful layouts, A-G style rating schemes have several shortcomings.

- They tend to oversimplify complex performance factors, which can generate inaccurate and misleading results;
- They only indicate theoretical performance under ideal conditions that may never be realised; and
- While they encourage better energy and carbon productivity, they do not target overall energy consumption.

These can lead to rebound effects or perverse outcomes.

8.1 Rebound effects and perverse outcomes

Rebound effects sometimes occur with domestic appliances like fridges. It is very hard for a small fridge to obtain a good efficiency rating, because the fixed overhead imposed by elements like the compressor is amortised over a smaller volume than is the case in a larger appliance. This can make a larger fridge look like a more attractive purchase than a small one, because it performs better on the A-G chart – even though its annual energy consumption is greater, a fact that the A-G rating may not adequately reflect. The larger fridge is certainly more efficient per cubic litre (because it benefits from an economy of scale) but consumers need to consider whether that extra space is really necessary. In a situation like this, the more attractive rating is essentially incentivising consumers to buy more capacity than they need.

While A-G rating scales often include information on typical annual power use, this is not given a colour-coded rating in the same way as the efficiency data, which can be problematic (see Box 1) and footnote[5].

We can see from these examples that, in complex environments like data centres, caution is needed to ensure the approach is fit for purpose. Metrics need to be chosen with care.

[5] Among the numerous problems of the current system are: (a) property EPCs are based on estimated running costs rather than empirically-derived analysis of energy usage or carbon emissions (e.g. electric heating can result in a lower EPC rating than cheaper gas heating, even though electricity from renewable sources is greener) and furthermore, they are based on modelled rather than measured performance, and thus do not necessarily reflect how people actually use energy; (b) they disproportionately penalise older properties; and (c) there is the potential for a high degree of variance between the assessments of different assessors (owing to inconsistencies in interpretation and data input). With regard to the UK Government's ultimate ambition for all rental properties to meet a rating of C, it has been estimated that this will impose a £23.4bn cost on the sector. See: https://www.rightmove.co.uk/guides/content/uploads/2024/10/Rightmove-Greener-Homes-Report-2024.pdf

Box 1: EPCs driving perverse outcomes: In the UK, rental properties need to meet an EPC (Energy Performance Certificate) rating of E or above, and the UK Government has announced an ambition to raise this to a C. The excellent rationale is that tenants should not be footing the bill to heat poorly insulated properties. Rather, landlords should be responsible for making the necessary investments. The approach appears logical in that it takes into account the potential performance of the building and identifies room for improvement, but makes allowances where these improvements cannot be implemented in practice (for instance due to heritage listing or structural constraints).

However, while the policy addresses a genuine market failure and should improve building performance and reduce heating costs for tenants, the way that the EPC rating is calculated is deeply flawed and inadequately reflects different property types and resident lifestyles. As a result, many properties struggle to meet the EPC threshold despite having relatively low energy costs. The consequence is that many perfectly habitable properties cannot legally be let and many landlords have left the property market altogether because they fear that in time they will be left with stranded assets. As a result, the usual balance of supply and demand is distorted, especially in London. Well-meant attempts to make things better for tenants have made them unimaginably worse: finding a rental property in London now is extremely challenging. Rents have soared, and people are staying in accommodation that is unsuitable for their needs because they cannot find alternatives. Although the intention to raise the standard to C has been shelved for the moment, the shortage of rental property is still acute because landlords are not returning to the market fast enough.

8.2 Time sensitivity for data centre ratings

There is one additional shortcoming of A-G ratings for products and EPCs for buildings that makes them unsuitable for data centres: product ratings are applicable at point of sale and generally fixed. EPCs for buildings tend to be valid for long periods of time, on the basis that upgrades and refits (like double glazing or dry-lining) are one-off investments implemented for the long term. The data centre sector is much more dynamic, especially in terms of the IT assets data centres contain. Regular reassessments would be needed in order for any form of rating to reflect a continuous cycle of improvement, and compliance authorities would need to ensure they have the resources to reassess, and issue amended ratings, in an accurate and timely manner.

Avoiding pitfalls and perverse outcomes: Pact position

While efficiency rating schemes have been effective in driving manufacturer focus towards in-use efficiency, the Commission must acknowledge the perverse outcomes that have resulted from broad assumptions and over-simplified metrics. In many cases the problems have arisen from a “black box” approach to calculating outcomes that provides no transparency to reportees in terms of how the eventual score is reached. To avoid the most common pitfalls, it would be necessary that assumptions, calculation methodologies, weightings applied to contributing metrics and any allowances made for operational constraints are made fully transparent to those obliged under any scheme, while a more simplified summary is published externally.

9 Data centre sustainability indicators

The good news is that, in terms of sustainability metrics that could inform a rating scheme for data centres, we are not starting from scratch. The data centre sector has been developing efficiency metrics for over 15 years and already has an impressive array of well-developed, peer-reviewed metrics supported by publicly available methodologies developed by both European and global standards bodies (CENELEC and ISO respectively).

9.1 Metrics deployed in the data centre sector include:

- **PUE (Power Usage Effectiveness):** the ratio between the total energy consumed by the facility and the energy consumed by the IT within it. PUE is therefore a measure of infrastructure overhead. This is expressed as a numerical ratio: PUE = E_{DC}/E_{IT} where E_{DC} is the power arriving at the facility and E_{IT} is the power consumed by the IT functions within it. There are two forms of PUE: Design PUE and Operational PUE. (see 9.3 below). (ISO/IEC 30134-2)
- **WUE (Water Usage Effectiveness):** the ratio between the total water consumed by the facility and the power consumption of the IT hosted within it, measured in litres per kWh. WUE = W_{IN}/E_{IT} where W_{IN} is the water used by the facility. (ISO/IEC 30134-9)
- **REF (Renewable Energy Factor)**: the ratio between the renewable energy delivered to the facility and the power consumed by the facility. REF = $E_{RES\text{-}TOT}/E_{DC}$ where $_{RES\text{-}TOT}$ is the sum of energy produced from onsite generation, energy delivered through PPAs and energy certified under guarantees of origin. (ISO/IEC 30134-3)
- **ERF (Energy Reuse Factor):** is the ratio between the energy that is reused and the energy consumed by the facility. ERF = E_{REUSE}/E_{DC}, where E_{REUSE} is the energy reused. (ISO/IEC 30134-6)

9.2 Differentiating data centre metrics

There are several core differences between the metrics we already use in the data centre sector and the A-G energy ratings that we see on consumer devices:

- **Actual not theoretical performance:** A-G product energy ratings do not provide any indication of the operational efficiency of the device. If you have a very efficient washer-dryer machine and always run it on boil wash and then tumble dry everything in it, or if you have a very efficient fridge and leave the door open, then the actual energy consumption of this machine will bear no relation to the expected performance indicated by its score on the A-G scale.
- **Focus on trend analysis:** metrics used in the sector are used primarily for trend analysis, to incentivise a cycle of continuous year-on-year improvement and provide performance transparency to customers.

9.3 PUE and energy overhead

The best known and longest established data centre sustainability metric is PUE. PUE was developed at a time when colocation was a common business model (colocation means that the facility infrastructure is provided as an outsourced service and operated by a different entity to the entity operating the IT hardware). So it made sense to be able to compare sites on the basis of infrastructure overhead. Over the years, and to improve transparency, the sector has differentiated "Design PUE" (the efficiency of the overhead under optimal conditions) from "Operational PUE" (how the facility is actually performing in the real world). Operational PUE is the version for which the industry has developed a standardised methodology and deployed as an efficiency metric for trend analysis. *Further references to PUE in this document refer to operational PUE unless stated otherwise.*

So PUE is a measure of infrastructure overhead and reflects the energy productivity of the facility's actual output. While PUE shares the characteristics of other data centre metrics in that it reflects actual rather than theoretical (or modelled) performance (and as such is deployed for trend analysis) this role as a measure of overhead is a critical difference between PUE and consumer-facing A-G labels, which do not consider overhead.

Imagine a freezer that runs 24/7 and only contains a bag of ice and a packet of sausages from 1986 that you know in your heart you will never eat. Irrespective of its efficiency rating, this freezer is imposing an energy cost in return for no useful work (as a quick glance at the running costs compared to the value of the contents would show). The freezer is being used in a way that is absolutely inefficient, and its A-G score is irrelevant. This is the essence of the problem with A-G scores: they do not reflect the effort required to deliver *specific outcomes* — in this instance, the energy required, per kilogram of usable food, to chill the freezer's contents to the correct temperature. To take a similar comparison, a fridge containing a single pat of butter will, however efficient, impose a huge overhead per kilogram of food compared to a well-stocked fridge.

9.4 Shortcomings of PUE as a regulated metric

The industry is well aware of the shortcomings of PUE. On the one hand PUE is an excellent tool for trend analysis, but it only gives us useful information about energy usage — and, moreover, only about the energy consumed by the infrastructure, not the core activity it houses. The proxy of IT energy consumption is used as an indicator of output: if the IT activity itself is inefficient, this is not identified. Nevertheless, scrutinising the overhead required to deliver a particular function is a relatively sophisticated approach. Such limitations are not always understood outside the sector, so while it is very pleasing to see industry metrics adopted as reference standards for regulations, it is disheartening when policymakers misunderstand the application of PUE, and the underlying maths. (For a detailed explanation of this issue, see Appendix I.)

The data centre sector's approach to efficiency metrics is therefore a very long way ahead of standard A-G product rating schemes and EPCs for buildings, because the real focus in terms of reporting within the sector takes into account the way the facility is run: it focuses on actual rather than theoretical performance.

(Recall that we are talking about operational PUE. None of this is to say that design PUE is unimportant – it is still needed because customers need to be sure that the data centre has the *capability* to perform optimally should optimal conditions be met.)

Differentiating sustainability indicators: Pact position

The Commission must acknowledge the fundamental distinction between the metrics deployed in the data centre sector, which measure performance under operational conditions and include infrastructure overhead, and those deployed for consumer facing products.

Policymakers should also be familiar with the limitations, and the calculation methodologies, of the metrics they choose.

10 Challenges of focusing on actual, not theoretical performance

The main disadvantage of focusing on operational rather than design performance is that operations happen under real-world conditions and are subject to external factors that may be beyond the control of the data centre operator. (The corresponding disadvantage of focusing on design performance is that it assumes that conditions are optimal). In terms of transparency and existing reporting procedures, switching to design PUE for data centres looks like a backward step, but if we decide to use operational PUE in preference, we need to consider variables that operators cannot control – otherwise we don't have a fair and equal basis for comparison. So what are these variables?

- **Facility age:** older data centres cannot benefit from technology innovations that can be implemented in new builds. Retrofitting is difficult in data centres because they have to continue running: you can't just cone off areas or impose overnight closures as you would if you were upgrading a motorway. Instead, to avoid compromising the continuity of service, you have to do the equivalent of changing the engine in a Boeing 737 in mid-flight.
- **Facility size:** smaller data centres are unlikely to be able to achieve the economies of scale of larger facilities and may suffer from proportionately higher fixed energy overheads.
- **Location (degree days):** external ambient temperatures affect PUE. More cooling is needed in hotter locations like Madrid than in cooler locations like Stockholm.
- **Redundancy:** the resilience rating of a data centre needs to be accommodated in operational metrics. To be resilient, a data centre needs to be built with duplicate systems in critical areas like power supply and cooling. Duplicating these systems is termed redundancy (on the basis that the spare kit involved generally sits around doing nothing and is therefore redundant unless there is an emergency). This adds operational overhead in terms of activities like maintenance and testing that reduces efficiency.
- **Business model and occupancy:** retail colocation accommodates multiple customers, which can make it harder to use space optimally. If you own a warehouse and use it to store things for several different customers: you can't guarantee that each customer will use their whole storage allocation, nor that their crates of stuff will pack efficiently around other customers' crates, nor that their storage needs will be the same from week to week. Occupancy in particular is an important variable that may be outside the control of retail colocation providers, either because the data centre is new and tenants are gradually taking up space, or because tenants are under-using the capacity they have leased, or because old customers are leaving and new ones are arriving ("churn"). At the other end of the scale, hyperscale customers tend to make better use of leased capacity. Hyperscale, retail, colocation and enterprise all have different usage patterns.
- **Utilisation:** neither colocation nor cloud platform providers can necessarily control the proportion of leased power that their customers actually take up. We don't downgrade the rating on a washing machine if someone puts a single sock on a full wash cycle, but a data centre's PUE will be affected if it is run below capacity – even if that is a result of its customers' choices rather than its operator's policy.
- **Activity and operating environment:** we are currently in an era of rapid change in IT. High performance and high density computing, AI and next generation workloads may depend on computing hardware with operational constraints different from those required over previous decades. For example, some new processors operate best at lower temperatures, or generate more heat, which impacts PUE.
- **Contractual terms and SLAs:** Service level agreements with customers may limit temperature and humidity ranges to smaller envelopes than necessary. Cooling is more energy-intensive than heating, so it has a disproportionate impact on efficiency.

Some data centre metrics, such as WUE, already take into account operational variables. For example, the SRI target for WUE accommodates water scarcity and potability. A robust rating scheme should acknowledge these variables to avoid penalising expansions or stifling innovation. It may be necessary to make allowances to avoid perverse outcomes, especially if a minimum standard or threshold is being applied (for example, in a situation where a new colocation data centre is installing customer load over an extended period of time, which is temporarily giving it a high PUE).

Accommodating external conditions in operational metrics: Pact position

The core sustainability metrics used in the data centre sector (operational PUE, WUE, REF) are all scored empirically. They are not based on a theoretical best case scenario, but measured under real-world conditions. Scores are influenced both by factors within the control of the operator (such as cooling technology) and outside their control (such as local climate).

This means that allowances for factors beyond the control of the operator must be accommodated in any rating scheme that sets minimum standards in order to ensure a fair basis for comparison and avoid perverse outcomes.

11 Which infrastructure metrics should form the basis of a rating scheme?

The choice of metrics should depend on whether the rating scheme solely targets energy efficiency (like typical EPCs) or reflects broader sustainability considerations (energy, water, renewable sourcing, etc). The existing data centre metrics that are candidates to form the basis of a rating scheme can perhaps be categorised as follows:

11.1 Metrics that could inform a rating scheme limited to energy efficiency only

- **Operational PUE:** Standardised operational PUE provides an indication of the actual operating efficiency of the data centre. However, as mentioned above, operational PUE is, as the name suggests, a metric that reflects performance under real-world conditions. When considering PUE we need to be aware of external factors and variables that affect the PUE rating of a facility, especially those outside of the control of the operator.

- **Design PUE:** the inclusion of design PUE as a metric, possibly a benchmark against which to measure actual performance, has been discussed in detail. There are pros and cons. For new facilities with no operational data, design PUE may be very useful as an interim indication of performance. Design PUE also gives an indication of the optimal performance of a facility. If it is unduly high, one might question whether such a facility would ever be fit for purpose. But while design PUE could provide an interim indicator, or signal inevitably poor performance, it has shortcomings as a metric. In theory, design PUE could provide a useful benchmark for data centres to aim for in their operations. However, in reality, once a facility is upgraded, its design PUE changes. Recalculating design PUE for an operational building presents a number of non-trivial technical challenges and in addition is resource-intensive and technically demanding. Setting design PUE as a benchmark could establish an unambitious target and discourage major refits. Moreover, the degree of standardisation is considered by many to be less well-developed for design PUE than for operational PUE[6]. Last but certainly not least, and as we have previously mentioned, design PUE does not account for real-world operational conditions.

- **WUE – Water Use Effectiveness:** it may seem surprising that WUE is included as an energy metric, but when data centres are under pressure to reduce PUE to very low levels, the adoption of a water-intensive cooling system can help minimise PUE. The inclusion of WUE here is to ensure that energy efficiency is not being improved at the cost of other sustainability performance elements. In this instance, including a water threshold acts purely as a disincentive to increase water intensity in order to minimise PUE. It is not necessary to be overly ambitious in terms of scrutinising or differentiating operators when water use is very low – that is only necessary if the rating scheme includes broader sustainability criteria beyond just energy. Here it is simply to signal where substantially increased water use is the "sustainability price" for exceptionally low PUE.

11.2 Additional metrics that could inform a broader sustainability rating scheme

- **REF – Renewable Energy Factor:** the degree to which renewable power is used is not, strictly speaking, efficiency-related as it focuses on the source of energy rather than how productively it is used. It is nevertheless an important consideration for customers and external stakeholders, and usually a prerequisite in colocation leases. The way that low-carbon power is sourced is also important, and different grades should apply according to the type of renewable supply. For example, the purchase of renewable credits sends the right market signals, but should be

[6] There are as yet no formal standards for deriving design PUE. By contrast, for operational PUE, ISO and CENELEC have published standardised methodologies. However, Green Grid proposals relating to design PUE did take account of external factors like local weather profile, regulatory constraints on operation and IT load.

differentiated from renewable power delivered from on-site generation or via a power purchase agreement (PPA) — both of which deliver additional, utility-scale renewable generation. That said, it may be necessary to reflect the more limited capacity of smaller operators to engage in the PPA market, which currently tends to be the preserve of only the very largest providers. Pact signatories also suggest that REF scores are not given excessive weighting: purchasing renewable power is not a substitute for other sustainability elements.

- **WUE – Water Use Effectiveness:** water scarcity is an increasing global problem, and therefore an important sustainability differentiator for data centres. Water treatment and delivery is also energy intensive, so the relationship between water and energy should be acknowledged. WUE as developed by the CNDCP differentiates potable and non-potable water, and locations that are water-scarce from those that are not. If WUE is part of a broader sustainability metric, Pact signatories consider that more stringent criteria should be applied in order to differentiate the best from the good and drive continuous improvement in water stewardship. Operators note that if the target criteria for water and renewable energy are too easily achievable to enable differentiation of the best performing facilities, then, irrespective of the weighting given to each criterion, the resultant rating would essentially be over-reliant on PUE.

11.3 Metrics that could comprise bonus criteria for rating, but not core criteria

There are multiple metrics that could be used as bonus criteria. However, for the purposes of this document we will refer to those metrics that operators are required to report on under EED legislation. That means the only metric for consideration under this heading is ERF – Energy Reuse Factor.

Waste heat can only be reused in the presence of a heat network and/or willing offtakers. It is not a legitimate metric of operational efficiency. Heat reuse is a particularly problematic issue for data centres: current operational activities produce waste heat of low quality, in the form of warm air at around 30°C. While operators are broadly willing to design data centres to enable the collection and delivery of waste heat to the edge of sites, the productive reuse of this heat depends on the existence of third party infrastructure and demand. Moreover, ERF does not include on-campus reuse of heat.

Despite these obvious shortcomings, data centre heat reuse attracts disproportionate attention from policy makers. This suggests to the sector either a lack of understanding among regulators of the operational realities within data centres, or a misconception that data centre distribution can be dictated by the presence of heat network infrastructure. If this misconception exists, it is a fallacy – data centres have very strong locational attributes relating to power, connectivity and demographics.

There are further factors that act as strong disincentives to operators to share waste heat. These include complexities relating to responsibility for the upgrade of heat and the impact this has on operational efficiency metrics like PUE; the danger of imposing supply obligations on data centres that would oblige them to run boilers to augment waste heat if not enough was available through normal operation; and the suggestion that operators supplying free waste heat should be subject to tax on the product. There is also the potential that the carbon impact of upgrading and delivering the waste heat could erode or consume the projected carbon savings of its reuse.

Finally, there is also, in view of the rapid development of computing technology, a risk that the nature of the waste heat product associated with data centres may change radically over time, which could result in stranded assets in terms of heat network infrastructure. While this is unlikely, it should not be ignored.

It appears that data centres are being targeted by policy makers as a route to decarbonise European heat networks. If this is the case, a more objective approach may identify more carbon- and cost-effective routes to achieve that objective.

Infrastructure metrics for inclusion: Pact position

Operators consider that operational PUE, WUE and REF, as defined by the Self Regulatory Initiative, are the least-worst reference metrics for a rating scheme, but that they should be used for trend analysis – to encourage continuous improvement – and not for spot comparison. Operators do not consider ERF a legitimate measure of operational efficiency, and as such it should only be included as a bonus metric, never as a core metric.

12 IT metrics - considerations

12.1 IT metrics proposed under EED

The EED requires operators to report performance against six IT related metrics:

- ICT Capacity for Servers – CSERV
- ICT Storage Capacity – CSTOR
- Incoming traffic bandwidth, in gigabytes per second – BIN
- Outgoing traffic bandwidth, in gigabytes per second – BOUT
- Incoming data traffic in exabytes – TIN
- Outgoing data traffic in exabytes – TOUT

The majority of these metrics, in particular those related to data traffic, are not sustainability metrics, and are therefore not reliable indicators of the IT efficiency of a facility (see box 2 for explanation). Moreover, the four data traffic metrics, especially if disclosed as part of a published rating, would lead to very serious safety and security consequences.

ICT capacity, or compute capacity – CSERV – is generally considered the only data point that provides an approximate indication of the processing activity housed in a data centre that could be roughly correlated to productivity, although industry has yet to reach a formal conclusion. Its obvious limitation is that it only indicates capability, not actual activity. Measuring computing functions is fraught with difficulty, and while compute capacity may be the least-worst indicator, it is still problematic as it can change rapidly, for example, during server refresh cycles. The challenges of including IT related metrics are outlined in the next section.

Box 2: Why not use data flows / bits per Watt as a data centre efficiency KPI?

Data centres do different jobs so this metric would favour some operations and penalise others. For example, high performance computing (HPC) involves very high quantities of data processing, very high utilisation of servers, high energy intensity and high value - but low volume - output. The production of weather maps, for example, uses HPC because of the size and complexity of the models and the sheer volume of data involved. At the other end of the scale you might find an operation like Netflix, where there is storage – and enormous quantities of content are delivered - but hardly any processing takes place. So a Netflix data centre would have a high storage capacity, lower probable utilisation but very high levels of digital output as content is streamed. This operation would perform very well against a bits per Watt metric, and the weather map operation would perform very badly. The metric would not give a reliable indication of comparative efficiency because the two data centres are performing different functions.

12.2 IT variables

While data centres can internally track the energy efficiency of their IT equipment using existing metrics for trend analysis over time, significant technical limitations hinder equitable comparisons across different data centres. In summary these limitations include:

- Unknown impact of different workloads on current energy efficiency metrics;
- Lack of standardised testing methodologies (though it should be noted that some hardware is built specifically for certain workloads and may perform poorly in a standardised test);
- Uncertain effects of new cooling technologies, such as liquid cooling, on energy consumption;
- Variations in processor designs and hardware configurations affecting energy efficiency outcomes.

In more detail, examples of variability of IT activity include:

- **Diverse Equipment and usage:** data centres house a wide variety of IT equipment, each with different energy consumption patterns and efficiency levels. This diversity makes it difficult to create a one-size-fits-all standard.
- **Dynamic workloads:** the energy efficiency of IT equipment can vary significantly based on the workload. Dynamic and fluctuating workloads make it hard to establish consistent and reliable efficiency metrics.
- **Technological advancements:** rapid advancements in IT technology mean that any standard set today could quickly become outdated. Keeping standards current with technological progress is a continuous challenge.
- **Measurement and verification:** accurately measuring and verifying the energy efficiency of IT equipment requires sophisticated tools and methodologies, which may not be uniformly available or applied across all data centres.
- **Virtualisation and power management:** the use of virtualisation and power management techniques can significantly impact energy efficiency. These factors add complexity to the task of setting a standard, as they vary widely in implementation by IT equipment operators and effectiveness.
- **Resilience and utilisation:** the resilience of server workloads and the utilisation of server capacity also play crucial roles in energy efficiency. These aspects are highly variable and context-dependent, making standardisation difficult.

12.3 Standardisation of IT metrics

From a policy perspective, ensuring consistency of measurement across data centres is essential for fair competition and accurate sustainability assessments. Furthermore, standardised metrics should be comprehensive, encompassing both the workload and the methodology used to measure performance and power consumption during that workload.

Currently, there are no IT metrics available that are formally standardised, so caution is needed, since a reference IT metric must be based on industry-wide standards. Different manufacturers and operators currently use varying methods and benchmarks, leading to inconsistencies in reporting and comparison. Measuring computing functions is inherently complex, owing to the diverse range of activities and workloads that data centres handle. All this makes IT efficiency an exceptionally complex area to standardise.

Therefore, until robust, standardised, and adaptable metrics are developed to address these challenges, IT equipment energy efficiency metrics should not yet be mandated in data centre sustainability ratings that are intended for use as comparators between data centres.

However, development of standardised metrics and methodologies, including the upcoming SERT 3 standard, is well underway. This presents the opportunity for any new metric to undergo a trial period of, say, five years, during which it could be monitored without materially impacting the sustainability rating of a data centre. Such a time frame would allow industry and external stakeholders to assess the metric's effectiveness, adaptability and fairness across different technologies and data centre configurations, ensuring it can accommodate advancements like AI compute workloads and new cooling methods. After the five-year period, the feasibility of incorporating the IT capacity metric into the rating scheme could be evaluated. In the interim, the Green Grid's new CSERV calculation methodology (PerfCPU) could be adopted. This would ensure that data centres continue to report in a standardised manner under the Energy Efficiency Directive. This approach would facilitate transparency and comparability across the industry while accommodating ongoing advancements in technology and metrics development.

12.4 Accommodating refresh cycles and upgrades

Compute capacity can change quickly, especially during server refresh cycles. This makes it difficult to maintain consistent and accurate IT measurements over time. The Uptime Institute encourages the adoption of existing standards – such as the Energy Star server and storage standards – to set performance thresholds for IT equipment deployed in data centres.[7]

Reporting IT metrics: Pact position

Operators broadly accept that to understand data centre efficiency, both the infrastructure and the computing function should be scrutinised. However, the Commission needs to be conscious that IT related metrics may inadequately account for varieties of workload; that unlike infrastructure metrics they are generally not formally standardised; and that IT performance may change very quickly, resulting in rapidly outdated ratings.

Operators support the inclusion of an IT capacity metric, providing it is based on industry-wide standards, which are currently at an advanced stage of development. Until this is in place, they take the view that reporting against IT metrics should not be mandated, and propose a transitional period to test viability of approach. In the meantime, operators support the use of the Green Grid's CSERV calculation methodology: PerfCPU.

[7] Additional actions that can improve energy efficiency of products on the market include eliminating inefficient hardware models from the market.

13 Reporting obligations: principles

Rating a data centre is not like rating a washing machine. As mentioned above, in data centres the infrastructure is often operated by a different business entity from the one responsible for the IT functions hosted within it. To understand how efficiently a data centre is operating, the performance of both the infrastructure and the IT functions need to be measured and reported. In an enterprise environment in which the same business entity runs both IT and infrastructure, this is straightforward.

But it is a much more challenging proposition at a colocation data centre. In a colocation situation, the entity using the IT function ("the customer") usually has access to information about the data centre's infrastructure performance via a contractual stipulation. However, the data centre's operator, although responsible for the infrastructure – and thus able to see the total energy used by the customer's IT – will not know about the *efficiency* of the IT function, because that is the customer's confidential information. This would be a major problem if the data centre operator were to be the entity solely responsible for reporting under a rating scheme. While arguments could be made for reporting obligations to be placed on occupants rather than on infrastructure providers, this approach would not work for the multitude of multi-tenant data centre environments where there could be hundreds of different customers hosted within the same facility, and many asynchronous server refreshes happening throughout the year (see below). The operational and organisational segregation of IT from infrastructure in colocation data centres is a very significant barrier to any rating scheme that seeks to record both infrastructure and IT performance across the sector and apply a combined rating at facility level.

It is a fundamental principle of policy making that regulatory subjects should only be required to do things that are within their control (see Box 3). Colocation providers in particular will fail to see logic in a scheme that requires them to report data relating to elements of performance that they can do nothing to improve. EED already breaks this principle by placing legal obligations on colocation providers to report data that they have no legal access to and are moreover prohibited from disclosing under contractual terms. The result is reporting chaos, and significant loss of trust in the Commission's policy development and implementation capabilities (see Appendix II – CNDCP Position Statement on EED).

Box 3: Reporting obligations Think of applying this to a hotel –it's rather like requiring hoteliers to report on everything their guests chose to do whilst in their rooms. Yes, it would be feasible in a non-intrusive way for a hotel to monitor and report the power and water associated with each guest visit, but how would they force guests to disclose what they did? This requirement would immediately set guests at odds with hotel staff who would be blamed for what is in effect an entirely unacceptable intrusion on private space. The hotel is between a rock and a hard place: legally required to pry on behalf of the state. In principle this is no different for data centre operators, except that in data centres these privacy arrangements are formalised contractually.

Reporting obligations: Pact position

There is grave concern regarding the Commission's approach to EED implementation and reporting: the way that obligations are imposed on the reporting entity is a highly contentious issue that needs urgent resolution.

A regulation that requires entities or individuals to do things that are outside their control, or provide information that they do not have access to, is wholly unfit for purpose. A rating scheme should only rate entities on performance elements within their direct control.

With respect to reporting of IT metrics, Pact members were divided on one aspect of IT reporting. While all agree that reportable elements should be confined to those activities within the control of the reportee, colocation providers consider that the ratings should remain separate for IT and infrastructure, so that a well-performing colocation operator is not unfairly penalised for the poor performance of its tenants. On the other hand, for large enterprise operators who may run highly integrated facilities where it is increasingly difficult to differentiate IT from infrastructure, a single rating is needed. **We reiterate here the Pact position that reporting obligations must be limited to factors within the direct control of the reportee.**

14 Classifying infrastructure variables....

We now need to consider how to classify and accommodate the real-world, external factors, identified above, that might prevent a data centre operator from achieving optimal performance. This accommodation would need to take the form of allowances and/or caveats built into a rating scheme. So which variables should we make allowances for, and against which metrics? We also need to consider how significant any allowances should be for each variable, and when they should, and should not, be applied.

Potential Variable	Applicable to which metric?	Rationale	Already factored into metric?
Location (degree days)	Operational and design PUE	External temperatures have a significant impact on cooling required. Care needed in setting threshold.	No
Resilience level / Redundancy	Operational and design PUE	The additional system redundancy required for higher resilience imposes an infrastructural energy burden.	No
DC Type: Retail/ wholesale colo, Enterprise, Hyperscale etc.	Operational and design PUE	The business model impacts efficiency: retail colocation is a more challenging environment in which to achieve optimal PUEs than hyperscale and likely to suffer greater occupant churn.	No
Compute function housed	Operational and design PUE	High performance computing and AI processors may require smaller ASHRAE envelopes (temperature and humidity ranges) to operate optimally, which will affect PUE.	No
Age	Operational and design PUE	This is tricky. The age of a data centre, like the age of a house, affects its efficiency because it was built to the technological specifications relevant at the time. While it does not make sense to allow for this, because the lower efficiency is real (older houses tend to have lower ratings too), setting minimum standards that older facilities cannot meet would result in perverse outcomes because decommissioning and rebuilding would have a significantly greater energy impact.	No
Occupancy	Operational PUE	The lower the occupancy (the percentage of the data centre that is let), the higher the proportion of facility energy consumed by the fixed overhead.	No
Utilisation of power as a proportion of leased power	Operational PUE	Cloud service providers cannot influence the way their customers use the servers. If customer utilisation is low, then the impact is similar to low occupancy.	No
Operational constraints imposed by client SLAs	Operational and design PUE	Customers may impose restricted temperature and humidity envelopes when they lease space from third parties (colocation providers) presumably to ensure an operational "safety margin" in facilities where they do not have direct control of infrastructure operations. Operating high	No

Potential Variable	Applicable to which metric?	Rationale	Already factored into metric?
		performance computing equipment and high density of computing equipment means more heat is generated, and larger margins may be preferred by operators to allow them to react, should there be an issue, to avoid equipment failure (see compute function above).	
Location (water stress)	WUE	Some areas are water stressed and others not. Metric should reflect the source and type of water (mains/groundwater/ potable/industrial use but non potable, etc.) in addition to scarcity. Operators have no control over the type of water that local utilities offer access to.	Partially
Availability of heat network	ERF	Data centres have strong locational attributes and, while there is willingness to collect heat and deliver to edge of site, this metric is meaningless unless heat networks are available. Limiting data centre developments to locations with heat networks will simply constrain development and is out of line with broader policy: for instance, nuclear power stations are not required to be co-located with heat networks, despite colossal heat rejection. Instead, this heat is dissipated into large water bodies like the sea. Other factors include the heat network temperature, power availability for heat pumps, permitting, contracting requirements, etc.	No
Availability of offtakers	ERF	See above: again, while there is willingness to collect heat and deliver to edge of site, this metric is meaningless unless there are offtakers (see 11.3).	No

Making allowances for variable conditions: Pact position

Operators broadly agree that all the variables listed above are relevant and that allowances need to be made to reflect external conditions, define an equal basis for comparison and avoid perverse outcomes or incentives. Operators recognise that impacts differ, and therefore allowances will vary. The need for allowances could be reduced or even avoided if data centres were to be classed according to factors like operational model. **However, allowances should only be deployed to reduce the risk of perverse outcomes: they should never obscure actual or relative performance (i.e. a worse-performing facility must not appear better than or equal to a more efficient one).**

15 Pros and cons of a minimum efficiency rating for data centres

The introduction of a minimum efficiency or sustainability standard (or threshold) is likely to be problematic and should take account of the following considerations:

- **A minimum standard must apply to all facilities within each data centre class.** For example, within the enterprise data centre class, there should be no "carve-out" or exemption for public sector facilities, because there is strong evidence that they comprise the worst performing part of the whole data centre sector. Genuine market failure exists in public sector data centres. There is very significant scope for improvement and major savings can be realised (See Appendix III, EURECA project findings). Imposing challenging requirements on commercial operators while exempting public sector facilities would send an unfortunate message – at best making regulators appear hypocritical, at worst giving the impression that they do not wish to disclose their own sector's operational performance because it lags so far behind the private sector.
- **It should not be confused with an optimal rating.** The objective should be to set the minimum performance standard that is acceptable in the market. A minimum rating should only remove from operations those facilities that are unfit for purpose. The role of such a standard or threshold is to clean up the laggards, not to set aspirational targets.
- **A minimum standard introduces the risk that the required threshold will be viewed as an end in itself** – even when there is scope for further improvement. Aspirational improvements are more effectively achieved by a rating scale, which provides a formal pathway to performance improvement, supported by strong competition around operational efficiency at the leading end of the market.
- **It could generate hostages to fortune:** data centres may be penalised when they ramp up occupancy of new builds in stages, or have variable occupancy rates due to business churn. Transitional compliance pathways would be needed to give operators a realistic time frame to implement upgrades without derailing new builds or expansions.
- An inappropriately defined minimum standard could also disproportionately penalise HPC expansions if they briefly worsen PUE or WUE while ramping up specialised IT functions.
- **A minimum standard might also drive legacy sites to premature shutdown or hamper innovation** if strict thresholds are imposed without factoring in transitional or region-specific realities.
- **A minimum standard might be best expressed in terms of design PUE:** a high design PUE – say, greater than 1.5, or even 2 – would indicate a facility that will not run efficiently under any circumstances without a major refit. That said, this approach may be ineffective, capturing very few sites while failing to address those that are designed to achieve acceptable PUEs but are implementing poor operational practices.
- **It may add no value in the commercial data centre sector,** where poorly performing facilities will be forced out of the market anyway because their scores (either under existing metrics or under a rating scheme) will demonstrate that they are uncompetitive.
- **However, a minimum standard may be very effective for legacy or in-house enterprise facilities** that are not run as businesses, but where poor operational practices nevertheless go unscrutinised and add significant costs. This is particularly the case in small on-premises public sector data centres, as demonstrated emphatically by the EURECA project (see Appendix III), where such costs are passed to the taxpayer or ratepayer, or erode funding that should be dedicated to public services.

Minimum rating: Pact position

Operators oppose a minimum threshold for the reasons given above. However, views are divided on whether a minimum standard, based on design PUE, could help remove facilities that will never be fit for purpose. A minimum standard is likely to be more effective outside the commercial sector.

16 Weighting versus allowances– basic principles

Weighting is the relative importance we give to each of the individual metrics that contribute to a combined rating. Allowances help us to accommodate different external constraints and establish a fair basis for comparison.

16.1 Weighting

For a rating scheme that depends on combining multiple individual metrics, the relative importance given to each of those metrics needs careful consideration. Weighting should reflect the purpose of the rating scheme: if it is limited to energy, then broader sustainability criteria should attract relatively low, or zero, weighting, so that only those factors relevant to energy consumption determine the outcome. Metrics outside the control of operators should attract zero weighting.

16.2 Allowances

The objective of factoring in allowances in order to – at least partially – account for factors outside the control of the operator is to provide a fair basis for comparison to assess facility performance. It is not meant to ensure that all facilities can reach an A rating.

Legacy facilities, small facilities and those with low customer utilisation cannot achieve the extremely low PUEs achievable by large scale, modern facilities with high utilisation. Allowances enable a facility to demonstrate that it is operating optimally within the limitations imposed upon it and reduce the risk of premature decommissioning and replacement, which could impose a greater carbon and energy cost.

In principle though, allowances should not obscure actual performance. A rating scheme should reflect the performance of a facility and not disguise it. Most importantly, if a scheme is used for the purpose of comparison (an objective that is not supported by Pact members, who nevertheless anticipate that any rating scheme will inevitably lead to comparisons being made) then the rating should differentiate facilities on the basis of performance, and allowances should not mask those differences. It is inevitable that older, smaller facilities and those that require more cooling, or those that are not full, will get lower ratings. A rating scheme should present these facts, not try to re-engineer them.

This approach is broadly in line with EPCs for buildings, where allowances are only used to prevent perverse outcomes. For example, a heritage building that is unable to meet the threshold necessary to obtain an EPC (which would therefore ordinarily prohibit its lease or sale) is eligible for an exemption once the owners have taken all the improvement actions available to them. However, while the allowances recognise the special conditions that apply to such buildings, and enable them to remain within the property market, the rating – F, G or H – stands. This enables the buyer or renter to understand the energy performance of the building and factor it into their decision. Using the allowances to change the rating is neither helpful nor transparent.

While the approach chosen will to some extent define how allowances are deployed, Pact signatories take the view that allowances should generally be applied where a threshold is being proposed, to recognise special circumstances beyond operational control. This is in line with the SRI[8], which distinguishes hot from cold countries on the basis of cooling degree days, because it is imposing threshold targets. However, the allowance is expressed in the form of slightly different targets, to reflect the additional effort required to meet lower PUEs in hot countries. The way that PUE is calculated remains consistent for all facilities.

[8] The Pact SRI sets PUE targets of 1.3 in cooler countries and 1.4 in hot countries for new data centres by 1 Jan 2025, and requires existing data centres to meet these thresholds by 2030. See https://www.climateneutraldatacentre.net/working-groups/

Weighting and allowances: Pact position

Weighting: The metrics most material to operational energy performance should have highest weighting in terms of their contribution to a combined rating. Metrics outside the control of operators should have zero weighting.

Allowances: Allowances should be in place to accommodate constraints outside operational control to minimise scope for perverse outcomes. They may result in an adjustment of what is deemed acceptable to meet a threshold, but should not disguise actual performance.

17 Comparing options - broad approaches

Pact signatories considered a range of potential approaches. In this section we sketch out some possible broad approaches, with sample assessment tables for each one and short descriptions of the pros and cons identified. In summary, they are as follows:

- **Broad approach #1.** This is focused on a single metric – PUE is the most likely candidate, but with bonus metrics that can be taken into account.
- **Broad approach #2.** This approach focusses on governance and existing standards.
- **Broad approach #3.** A multi-criteria assessment derived from individual scores against a range of weighted metrics.
- **Broad approach #4.** A multi-criteria and multi-category assessment that combines core performance metrics with other criteria, such as bonus metrics, adherence to standards and governance.

These sample approaches are purely indicative, and do not represent operator views in terms of specific inclusions. Where metrics or bonus criteria are listed that are not discussed above, these are only indicative and are to be considered as being for illustrative purposes only.

17.1 Broad approach #1: single metric plus bonus categories

Core metric	Bonus metrics	Score
PUE according to ISO 30134-2	WUE REF ERF	

Feedback: Adopting this approach (and assuming PUE as the preferred metric) offers simplicity and industry-wide acceptance of terminology and standardised methodology. It is already broadly reported, and provides a simpler, quantitative framework. On the other hand, PUE is simplistic, is subject to multiple factors and only captures a minority of the energy flow through most data centres. It is too limited a measure to be helpful because it excludes other energy related performance measures, such as WUE.

17.2 Broad approach #2: governance, compliance and commitments

Category and inventory	Check	Result
Governance (ISOs)		
▪ ISO14001		
▪ ISO 50001		
▪ ISO 9001		
▪ ISO 27001		
▪ ISO 30134 (-2, -9 etc)		
Compliance		
▪ EU Taxonomy: Technical screening completed for EU Taxonomy alignment – proof – certificate of conformity from external party		
▪ EU Taxonomy – all DNSH requirements completed		
▪ 2023 data reported as per EED and Delegated Regulation requirements		

Category and inventory	Check	Result
Commitments ▪ CNDCP (climate neutral data center pact) commitment and audit as per 2023 requirements – proof – Certificate of conformity from external party ▪ EU Code of Conduct - Participant status ▪ Official approved carbon reduction target – proof - published on SBTi website ▪ Annual reporting into CDP – Climate? Water? – proof - scores published on CDP website ▪ Ecovadis annual score – proof/score to be available annually – can be minimum requirement that the overall score is above 60 (for example)		

Feedback: This approach aligns with good practices already in place, avoids additional reporting burdens by adhering to recognised standards, many of which are already referenced in legislation, and scrutinises adherence to commitments. The standards emphasise continual improvement and allow for the use of a common data pool. The infrastructure metrics that would be reported align with those required under EED. However, this approach excludes IT. Adhering to standards does not necessarily optimise efficiency: goals may shift, smaller companies may struggle to comply with multiple standards, and the approach may not align with policy ambitions to rate and rank data centre sustainability effectively.

17.3 Broad approach #3: multi criteria assessment

This approach uses a combination of different metrics to evaluate various aspects of the data centre. This can include CNDCP categories and/or others. To arrive at an aggregate score, weighting would need to be applied, as discussed in section 16 above.

Category	PUE	WUE	REF	Compute Capacity	Materials circularity	ERF
Actual (12 month average)						
Minimum achievement						
% achievement						
Weighting						
Total scores						
Aggregate score:						

Feedback: Adopting this approach brings together existing, standardised metrics and key operational elements in a transparent way, with relative weightings made clear at point of reporting. This method includes the basics of the CNDCP SRI pillars, focuses on data already being collated, accommodates interdependencies between metrics, enables allowances to address variations such as load and resilience, and includes an IT-related metric. However, this duplicates existing frameworks, creates auditing challenges and may introduce weighting subjectivity. Administrative overhead would be significant, and the approach must be developed with care to avoid misleading comparisons.

17.4 Combined assessment: multi-criteria, bonus categories and governance

This approach includes core metrics for infrastructure and IT, bonus metrics and broader sustainability credentials, but segregates the scoring for each category and subcategory so that transparency in terms of performance against individual metrics and standards is clear at point of reporting, and any weighting or allowances are clearly stated. The segregated results could be visualised through the banded, colour-coded approach set out in Section 17 below. In addition to weighting the individual metrics, category weightings can also be applied to reflect the relative importance of, for example, governance compared to energy stewardship.

Core Metrics	Result (inc conditions)	Conversion (eg results expressed as %)	Weighting	Score	Total for Category
PUE					
WUE					
REF					
IT Related Metrics	**Result (inc conditions)**	**Conversion (eg results expressed as %)**	**Weighting**	**Score**	**Total for Category**
Compute Capacity			0 for colocation		
Bonus Metrics	**Result (inc conditions)**	**Conversion (eg results expressed as %)**	**Weighting**	**Score**	**Total for Category**
ERF					
Materials circularity					
Other…					
Governance	**Result (inc conditions)**	**Conversion (eg Yes = 1, No = 0)**	**Weighting**	**Score**	**Total for Category**
ISO Suite					
Voluntary commitments					
Other..					

Feedback: This approach is comprehensive, accommodating as it does infrastructure, IT, governance and standards – but it is complex. It is transparent at the point of reporting and can be deployed internally as a trend analysis dashboard.

All the approaches that include IT have one additional complexity, because they combine categories with very different timelines: infrastructure efficiency changes relatively slowly, but IT performance can change overnight if servers are upgraded.

Options – broad approaches: Pact position

Operators found shortcomings in all approaches. The least-worst is likely to accommodate core metrics, bonus metrics and governance, with elements beyond the control of operators given zero weighting.

18 Rating scheme visuals

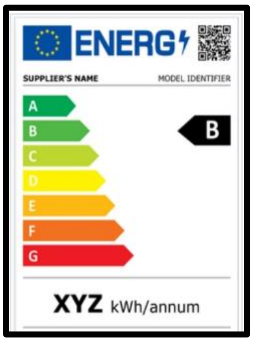
There are many options available in terms of how a rating might be represented. To a large extent these exist independently of the broad approach selected, the contributing metrics chosen and the degree of weighting applied. These options include a single rating visualisation that is the result of combining scores against a number of individual metrics, or a set of ratings that could be presented simply as a table, a graph or a spider diagram.

In general, Pact signatories prefer visuals with performance bands (rather than set points/minimum thresholds) per metric. Views varied, however, on the use of banding to accommodate allowances. Some felt that banding thresholds could vary according to operational type (and external factors, like climate zone) to set relevant baselines for each band and ensure a fair basis for comparison. Others — while agreeing that comparisons need to be meaningful within a cohort, location or category – felt that allowances should not mask actual or relative performance, and that if a worse performing facility could achieve a more favourable band than a better one, this could in effect constitute moving the goalposts (see Section 19, Categorising).

Signatories did agree that visuals should clearly display a prefix or icon to identify the category or operating model that the results refer to: eg CO for colo, EN for Enterprise, HY for Hyperscale etc. (See notes on categorisation below.)

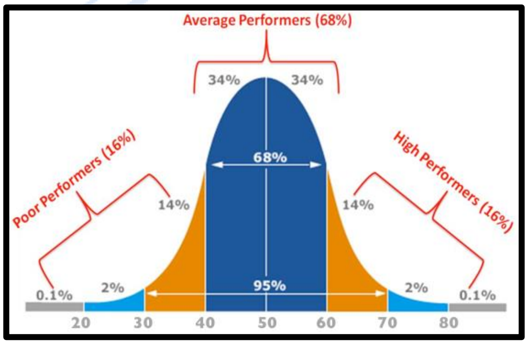
Several alternatives are considered below:

18.1 Visualisations based on a single data point



A colour-coded single performance rating (e.g. an A-G scale) displays a number or grade that usually combines scores from multiple metrics. Performance is presented as bands, or ranges, rather than specific numbers. The main advantage of this single, banded, overall rating is that it is visually simple and intuitive. However, despite being visually appealing, colour-coded ratings impose a risk of oversimplification: the information conferred by the score is less likely to be useful than other, more granular visualisations and may misinform, or under-inform, the customer because important information is consigned to the small print (see the discussion of rebound effects in Section 8).

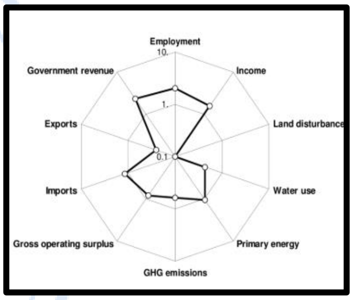
A bell curve distribution that displays relative ratings of facilities according to their distribution across a performance bell curve. This approach could work in parallel with a single colour-coded performance rating. A benefit is that it would show *relative* performance, which should encourage a continuous cycle of improvement even among the best performers (who will naturally compete to occupy the leading edge position). The bell curve can be specific to an individual operational model or to a category of data centre in order to differentiate performance within that cohort. However, a single bell curve can also be used to compare different operational models, and the way that they cluster can provide macro-level information about the relative sustainability characteristics of different facility types. The primary disadvantage is that it accepts, by default, that the rating scheme will be used for comparison.



1) Ranges and Boundaries								
Cool Climate	A++	A+	A	B	C	D	E	F
PUE	1.1	1.15	1.2	1.3	1.4	1.5	1.7	1.9
WUE	0.01	0.2	0.3	0.4	0.4	0.6	0.8	1
Renewable	100%	100%	100%	100%	100%	90%	80%	<80%
Reuse of heat	Interconnected or no possibility possible in the next 2 years	Interconnected or no possibility possible in the next 2 years	Interconnected or no possibility possible in the next 2 years	Min planned within 5 years	Min planned within 5 years	Assessment of possibilities to interconnect with district heating systems and other users of heat	Assessment of possibilities to interconnect with district heating systems and other users of heat	-

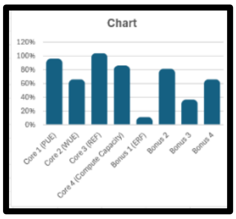
18.2 Multiple factor visualisations:

Colour-coded multiple rating displays enable a more granular representation of performance against multiple metrics, together with any weightings applied and the rationale for their application. Its disadvantages are that this may be less intuitive, and require a more sophisticated level of understanding – or at least closer attention – and as such might be better deployed in discussions with customers and regulators who have a detailed understanding of the metrics, rather than as a publicly-accessible display. It may also make reporting difficult for more integrated data centre environments (see comments below). Note that the banding in the sample graphic is for illustrative purposes only.



Spider diagrams segregate performance against individual metrics in a transparent and highly visual manner. Customers can then match performance against their individual priorities, which may vary by location (for instance, WUE would be much more critical in central Spain than perhaps in northern Sweden). Operators get clear indications of the specific areas where performance can be improved. The need for complex weighting is reduced because the results are not combined into a single figure or rating.

Similarly, if elements like resilience are included, this removes the need for the relevant allowance. Separate spider diagrams can be used for core and bonus metrics. Disadvantages are that this approach tends to rely on specific numbers. That said, banding could be overlaid in the same way that urban transport maps overlay station locations with banded charging zones.



Charts have the flexibility to provide performance indicators across as many different metrics as required, and can also include multiple bonus metrics within the same dashboard. Operators could adapt these charts to include additional information according to customer preference. Customers can match performance against their individual priorities. Operators can swiftly identify specific areas for improvement.

Visualisation: Pact position

At this stage, operators are more concerned about the broad principles of a rating system than the specific visualisation adopted. However, they propose that the publicly available visualisation should be a simple figure (single in the case of enterprise, combining IT and infrastructure, dual in the case of colocation, to reflect the different reporting entities involved) that does not include detailed performance data against each individual metric. This more detailed tier of information should be disclosed on a voluntary basis, for instance, in response to a customer's request. The spider diagram, multiple factor visualisations and charts should be reserved for such conversations.

19 Categorising facilities to ensure a level playing field

Operator views vary on the need for categorisation and how it might be achieved. Irrespective of the broad approach taken and the visuals adopted, it would be difficult to ensure a fair basis for comparison across different types of operation unless some form of classification was added (so that the rating is contextualised correctly for the type of facility it grades). Placing data centres into operating classes may also simplify, and in some cases possibly remove, the need for complex discussions about allowances – simply because there would be fewer variables. It would also set valid expectations in terms of the information that regulators and customers would see on the rating sheet (for example, the colocation class of data centres should not be expected to include information about IT performance that they do not have access to: see Section 13). There are multiple business models in the sector, but broad categories could distinguish enterprise, hyperscale, cloud operators, wholesale colocation, retail colocation and HPC.

Categorisation on the basis of location may also be considered, specifically in terms of ambient temperature. Here, industry views tend to be divided. On the one hand, categorising by location (hot/cold) and adjusting scores accordingly could disguise the fact that some locations, such as the Nordics, have climates that by their nature enable more efficient operation of data centres, and therefore put such locations at an apparent disadvantage by failing to differentiate them. There is also the issue of where the line is drawn in terms of ambient temperature.

On the other hand, when customers are seeking data centre services they need to be able to distinguish the best in class in a given location, because data centre demand, especially in Europe, is largely driven by demographics. Although some data centre workloads are location agnostic, most data centres can't be located just anywhere: in an ideal world they would all be cold countries, but they have to be geographically distributed with many individual data centres close to concentrations of population and business activity in order to deliver efficient service to areas of high demand. In other words, data centre operators in hot climates can't just move to cold ones, and should be recognised for the efforts they are making within the constraints of the external environments in which they operate. A compromise position may simply be to include very clear category labels and the retention of actual scores, even if banding thresholds change. However, this is a complex area where further discussion is needed to ensure an equitable resolution.

Categorisation: Pact position

Categorisation is needed if rating schemes are being used to compare data centres, to ensure these comparisons are meaningful. Most devices subject to A-G product rating schemes are divided into classes to help consumers compare like with like: washing machines are classified separately from washer-dryers.

Data centres should be similarly classed for the same reason. On a similar basis to allowances, ratings should be applied consistently across all classes, but comparisons only made within classes, or between classes as a whole, not between individual facilities in different classes.

20 Confidentiality

While data centre infrastructure efficiency is increasingly reported publicly – both as a regulatory requirement and as part of voluntary sustainability reporting – the collection of data relating to their IT functions is significantly more commercially sensitive. The European Commission has demonstrated a strong commitment to protecting the personal data of its citizens and should apply a similar level of stewardship to the sensitive commercial data of businesses operating within the Region.

The rating system must robustly protect sensitive data by relying on aggregated or anonymized metrics that do not reveal trade secrets. The Commission should adopt uniform data-protection protocols across Member States and identify specific situations where confidentiality constraints limit disclosure, ensuring consistent and comparable scoring. This approach maintains fairness, preserves data integrity, and safeguards operators' intellectual property.

Confidentiality: Pact position

The approach must ensure confidentiality of data considered by participants to be sensitive either from a commercial or safety and security perspective.

21 Contacts for further information

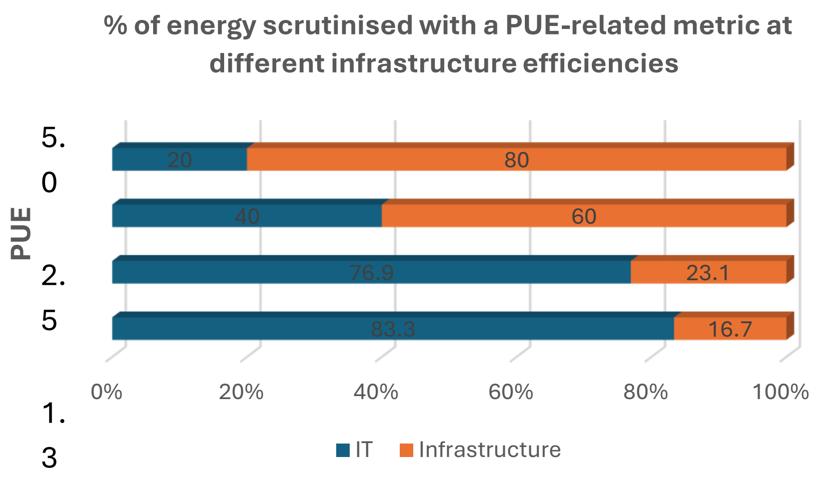
For further information please contact

- Velimira Bakalova, CNDCP Secretariat (CISPE): velimira.bakalova@cispe.cloud
- Valerio Moneta, CNDCP Secretariat (EUDCA): eudca@apcoworldwide.com
- Emma Fryer, Chair, CNDCP Working Group on Ratings for Data Centres: efryer@cyrusone.com
- Giancarlo Giacomello: Vice Chair, CNDCP Working Group on Ratings for Data Centres: giancarlo.giacomello@staff.aruba.it
- Alban Schmutz, Vice Chair, CNDCP Working Group on Ratings for Data Centres: alban.schmutz@axorya.com

Appendix I:

Effort and outcome in PUE: cost benefit Adapted from “PUE: Sith or Jedi?”

PUE only scrutinises the energy powering a data centre’s supporting infrastructure, rather than all the energy flowing through the system. In a well run facility this is a small fraction of total energy consumption. If infrastructure efficiency is poor, it is worth applying scrutiny here, but as facility performance improves, a law of diminishing returns starts to apply:



Note that at a PUE of 1.2, the energy scrutinised is 17% of the energy in the system, whereas at a PUE of 5, it is 80%. The only places that we know for certain PUEs of 5 and above exist are small enterprise data centres, notably public sector on-premise facilities, thanks to the EURECA project.

There are two simple rules to remember:

1. The effort involved in reducing PUE is INVERSELY PROPORTIONAL to the starting PUE
2. The scope for energy savings is DIRECTLY PROPORTIONAL to the starting PUE

A data centre with a high PUE can deliver substantial efficiency infrastructure improvements relatively easily. As PUE reduces, scope for efficiency improvement diminishes.

Worked Example

Let’s take for example a data centre with 500KW of IT load and a PUE of 5 (roughly the average for public sector on-premise data centres as reported by EURECA project in 2018).

This data centre actually consumes 2,500KW or 2.5MW once we have added in the infrastructure overhead (remember 5:1 – 5 total power, 1 is IT so infrastructure must be 4 of that 5).

Over a year, the total energy that this facility consumes amounts to 21,900MWh (total energy consumption is measured in Wh, and as there are 8760 hours in a year, we multiply the instantaneous demand by 8760).

Of this 21,900MWh, 17,520MWh are consumed by the supporting infrastructure compared to only 4,380MWh by the IT.

Reducing the PUE of this facility to 2 would reduce its instantaneous energy demand from 2.5MW to 1MW, which over a year would save 13,140 MWh of power.

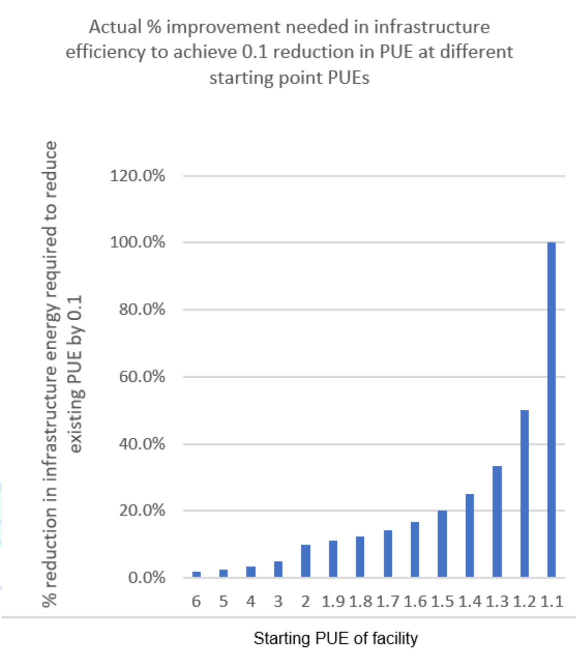
A PUE of 2 is not particularly efficient and lags well behind the commercial market average, but is relatively easy to achieve, and it’s a vast improvement on 5.

Conversely, reducing the same facility from a PUE of 1.3 to 1.2 would save just a fraction of that amount (438MWh/year), would require enormous effort and would cost far, far, more to implement.

In reality, the only way that this can be achieved is likely to be by adopting adiabatic cooling which will significantly increase water consumption. Data centres are committed to reducing water demand, so such an approach compromises these undertakings. Moreover the energy consumed in processing and delivering that water will erode the savings.

Comparative effort involved in reducing PUE

IT power consumption nKWh	Infrastructure power consumption nKWh	Existing PUE	Target PUE (reduced by 0.1)	Actual % improvement needed in infrastructure to achieve 0.1 reduction in PUE
1	5	6	5.9	2.0%
1	4	5	4.9	2.5%
1	3	4	3.9	3.3%
1	2	3	2.9	5.0%
1	1	2	1.9	10.0%
1	0.9	1.9	1.8	11.1%
1	0.8	1.8	1.7	12.5%
1	0.7	1.7	1.6	14.3%
1	0.6	1.6	1.5	16.7%
1	0.5	1.5	1.4	20.0%
1	0.4	1.4	1.3	25.0%
1	0.3	1.3	1.2	33.3%
1	0.2	1.2	1.1	50.0%
1	0.1	1.1	1	100.0%



Here you can see that reducing a PUE of 2 to 1.9 requires a 10% improvement in energy efficiency. Reducing a PUE of 1.3 by the same amount to ready 1.2 actually requires a 33.3% improvement in energy efficiency.

The key thing here is that the cost and effort involved in delivering improvements in PUE increases disproportionately as the PUE decreases.

To explain how this burden increases as PUE drops, think of a bus that takes 50 people, of whom 10 always travel free, with the other passengers divvying up their tickets between them. When there are 50 passengers, the 40 payers must add 25% each to their fare. If the bus company decides to reduce the size of the bus to 30, but keeps the ten freebies, then the cost of those ten tickets is shared between 20 paying passengers instead of 40, so each one forks out an extra 50% for their tickets. If the bus is reduced to 15, then those 5 paying passengers share those ten free tickets between them, effectively paying for 3 tickets each, an extra 200% on their fare. If the bus capacity reduces to 11, then the one payer has to fund all 11 tickets, by which point he or she has probably decided to walk instead.

Appendix II:

EXTRACTS from CNDCP Board Feedback:

Overview of practical challenges faced by data centre operators when completing mandatory reporting under the European Energy Efficiency Directive (EED)

December 2024

1 What is the climate neutral data centre pact?

The Climate Neutral Data Centre Pact was established in January 2021 and is a voluntary initiative involving data centre operators of all types: enterprise, colocation and cloud service providers, plus their representative bodies. The Pact now has over 100 signatories who collectively represent 85% of data centre operational capacity in Europe.

The Pact seeks to support the objectives of the European Green Deal in realising ambitious greenhouse gas reductions and leveraging technology and digitalization to achieve Europe's 2050 climate neutral goal. To ensure data centres are an integral part of the sustainable future of Europe, operators and their trade associations are committed to achieving operational climate neutrality by 2030.

In line with these commitments, Pact signatories take energy stewardship very seriously. Moreover, the electro-intensive nature of our business activity means that we are highly motivated to optimise efficiency and minimise the energy consumption required to operate our facilities whilst supporting increasing IT demand.

However, the core mission of the Pact goes beyond energy consumption to address broader sustainability challenges relating to data centres, such as water use, waste, renewables sourcing and circular economy practices. The Pact wishes to underline that because data centres are complex environments, exclusive focus on one area, such as energy, could compromise performance in another, such as water.

Signatories have been developing robust industry metrics to provide a level playing field for comparison and reporting. More information on the work of the Pact can be found here.

2 Data centre industry reaction to first annual reporting under EED

At the November 2024 bi-annual meeting between the Pact and the Commission, Pact members shared, at the request of the Commission, their insights and experiences of the initial round of reporting of energy data under the new Energy Efficiency Directive. At the further request of the Commission, we have summarised those comments in this document.

As an overview, we wish to reflect that the data centre sector at large, and Pact members in particular, are disappointed with a number of design and operational shortcomings of the EED as it relates to data centres.

We are disappointed because our belief and hope was that this instrument could have been deployed to expose the worst performing areas of the sector where it could have delivered very significant energy savings. We also anticipated and shared concerns in advance around some key issues, that if addressed, could have increased the effectiveness of the EED.

As requested, we have documented below the significant challenge we find with the reporting as it is practically imposed at Member State level, as well as the potential impacts of these shortcomings.

The following notes identify the most problematic issues which, if left unaddressed, could threaten the purpose of the EED: to improve energy efficiency. Concerns about exposing sensitive data, as well as issues with collecting and sharing accurate information could conspire to reduce the quality and comprehensive coverage of reporting thus undermining the Directive's core mission whilst needlessly damaging competition in this fast-changing industry. We therefore request that these measures to be revisited as a matter of urgency.

2.1 Regulation mandates data centre operators to report on things outside their control

The EED requiring colocation providers (those that 'host' IT infrastructure from one or more tenants at a single data centre facility) to report on the activity of their customers. Not only does this go against principles of good regulation by making subjects liable for things outside their control, but there are practical barriers to its effective implementation.

- Some of the information required may be known, or can be calculated, by colocation operators but they are prevented from disclosing it under the contractual terms of their client lease agreements.
- Much of the information is unknown to the data centre operator. Only the customer has this data and they may, or may not, agree to disclose it. Critically, the customers themselves are not mandated to do so by the Directive.

There are sound commercial and contractual reasons why information exchange between operator and customer is kept to a minimum, including compliance with existing European regulations. Making the colocation operator legally responsible for the disclosure of information that they have no access to not only creates tension with its customers, but creates the risk that estimated, inaccurate or simply 'made up' data is substituted to meet reporting requirements.

2.2 Failure to implement adequate safeguards regarding data collection

The failure to safeguard data collected under EED is a particular concern. As noted above, much of the data mandated to be reported not only has commercial significance but remains the property of data centre customers, not the operators themselves. The Pact has advocated for strong confidentiality guarantees for reported data from its earliest submissions on the proposed EED.

Unfortunately, the final text has not provided the clarity and direction that we would have liked on this fundamental issue. Furthermore, we believe that the Commission has not done enough to present and prepare Member States with clear guidelines on how to treat the data they receive from operators through the EED reporting.

As a result, individual countries can impose wildly different interpretations of how data should be treated. For instance, the Netherlands, one of the first countries to implement EED reporting, has published all data received, irrespective of its commercial sensitivity or the representations made by those who submitted it in good faith.

Without adequate safeguards regarding confidentiality or publication of data many parties will rightly be concerned about what they report. Many will simply refrain from submitting data requested under EED citing evidence that it will not be protected.

2.3 Gathering data with no clear relationship to energy efficiency

The Pact has advocated from the outset that only information that has a bearing on energy efficiency should be included in EED reporting. Specifically, we have argued that data flows and data storage information is not relevant and should not be collected.

In line with previous points, the inclusion of this data in mandated reporting creates unnecessary barriers to accurate reporting because often the information required is not owned by the regulatory subjects and is confidential and commercially sensitive.

With concerns over confidentiality already high and with widespread understanding n the industry that many elements have no bearing on sustainability, energy consumption or efficiency there is significant risk that this information will be omitted or fictionalised.

2.4 Practical reporting compliance issues

The fast pace at which the EED is being implemented to significant variations in transposition, both in terms of timelines for reporting and legislative approach. We fear that regulators in some Member States are confused and have misinterpreted elements of the Directive. As noted above, this has already led to some widely differing interpretations on key aspects of reporting such as data aggregation and confidentiality. A firmer drive to greater harmonisation is necessary with clear direction over the elements that are causing confusion.

Article 12 of the EED already places, in our view, a disproportionate compliance burden on European colocation providers who need to report their data in different Member States. The reporting procedures for operators only exacerbates this. They are impractical, cumbersome and time intensive. An individual account must be established for each country meaning that operators may have to report through 27 separate accounts if they operate in all EU Member States. The process of registration is protracted with operators reporting many snagging issues that have to be resolved individually. Each account must have a unique email address and log-in which means that a European portfolio of data centre assets cannot be managed or overseen collectively by a single responsible owner within a company.

In some states, we believe that the EED has been repurposed to try to help governments meet domestic targets that are outside its remit. In Germany, for instance, it appears that it is being deployed to help deliver broader carbon reductions rather than incentivise efficiency. The UK made this mistake in 2011 with the Carbon Reduction Commitment, which was abolished in 2016 having delivered no tangible policy outcomes other than imposing cost and encouraging businesses to offshore their carbon intensive activity.

Implementation issues have also emerged because Member States have not adequately familiarised themselves with the core metric that the Commission has adopted, PUE (Power Usage Effectiveness). While we are delighted that the Commission has adopted a peer reviewed industry metric underpinned by a standardised calculation methodology, it is clear that not all Member States understand either the limitations of this metric, or the maths. (see annexe I).

3 Immediate implications for colocation providers

The consequences of the shortcomings of EED as currently implemented are significant:

Colocation providers who cannot provide the client information required because they do not have access to it and the customer is not willing to supply it are breaking the law. We are not aware of any colocation providers who are in a position to provide the client information requested. Customers have either stated that they will not provide the information or have simply not provided it.

Colocation providers who supply client information that they do have access to will be breaking their contractual obligations with their clients. Clients have in many cases formally requested that colocation providers refrain from providing this information on the basis that, under common business models where a large colocation facility may have very few, or only a single client, this amounts to the disclosure of commercially sensitive information. Operators are complying with client requirements in these instances.

No regulation should place an entire industry cohort in such a position.

4 Strategic implications for the data centre sector in Europe

The immediate impacts on operators are evident but longer term, less obvious, implications of the current shortcomings of EED are also likely to be significant.

4.1 Eroding EU competitiveness as a place to do business

Firstly, the current approach to reporting is eroding trust in the policy process. If large technology companies cannot be confident that the EU will enforce adequate stewardship of their commercially sensitive data, then the EU will cease to be a destination of choice for these entities. The sector is currently enjoying a buoyant phase of expansion and at this moment data centre demand is growing at a CAGR of 75%. To ensure that Europe has a fair share of the infrastructure that will underpin the growth of the digital economy, policy makers need to ensure that legislation like EED does not curtail our collective digital future. We worry that issues with EED reporting may ultimately compromise the EU's competitiveness as a place to do business.

4.2 Creating schism within the industry

Secondly, the EED is creating division in the market by setting those with different business models against one another. This places colocation providers and their customers at odds and it is undermining the positive collaboration that characterised the industry over the last decade.

4.3 Generating hostages to fortune by failing to accommodate technology trends

Thirdly, new trends in computing depend on GPUs, the next generation of [servers] which run hotter and thus require lower ambient temperatures to operate, making the attainment of very low PUEs extremely challenging. If the EU implements laws that operators cannot comply with, they will seek alternative locations to invest.

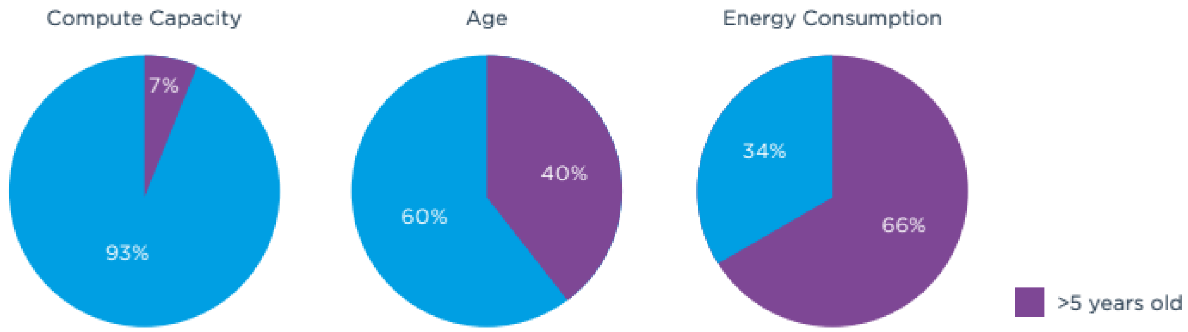
Appendix III:

EURECA project – summary findings

The Commission-funded EURECA Project reported over six years ago, in 2018, that energy stewardship in small on-premises data centres (in this case from a survey of 350 public sector facilities in Europe) was shockingly poor. The study reported average PUE at around 5 which means the supporting infrastructure consumed four times the power of the IT. See image extracted from the EURECA presentation slides and report.

Range of PUEs. Pointer indicates average (around 5)

It also reported very low server utilisation: 40% of the servers were over 5 years old, and these servers consumed 66% of power while only delivering 7% of compute. So both the infrastructure and the IT functions were hopelessly inefficient.



Server age and activity

Not a single one of these 350 data centres is currently in scope of EED – they all fall beneath the threshold of 500KW IT capacity. There is a sense that their aggregate energy consumption will not add up to much compared to the commercial sector, so they don't really matter. There are two problems here:

- Firstly we simply don't know how many data centres there are that sit below the 500KW threshold – estimates suggest that 90% of European data centres by number don't reach this threshold, but until the data is collected systematically, we cannot even estimate.
- Secondly, the combined impact of poor infrastructure and old, inefficient servers has a multiplier effect on energy consumption that is so significant that the size of market share of this activity does not have to be very large for it to have a major energy impact – and in parallel the potential for even modest efficiency improvements to deliver huge energy savings. See table below, again extracted from the EURECA project report.

This table, reproduced from the EURECA project findings, demonstrates clearly how a combination of high PUE and low server performance has a multiplier effect on energy consumption, which means that there can be orders of magnitude in the difference between the energy required to perform computing functions in different data centres.

The x-axis of the table includes a range of hardware ages, from old to current. Modern servers are much more energy efficient than old ones, as the EURECA results illustrated.

The y-axis compares operating environments from on premise traditional to highly virtualised public cloud. For each operating environment a range of PUE values and server utilisation rates are given. Note that, while the highest PUE given in the table is 3, the actual PUE of the 350 study data centres average around 5, so the energy use estimates are very conservative for this cohort.

The results in the table essentially provide comparative figures for the energy required for running a given workload of computing activity in each environment. The comparisons are striking. In an on-premise non-virtualised environment using 7 year old servers, over 50,000,000 KWh is needed compared to 189,000 KWh in a fully virtualised public cloud environment. This is 285 times as much energy. While these figures are very out of date they show the multiplier effect plainly.

	Scenario	PUE	β	Annual Use Phase Energy in KWh (for running workload ω)					
				Hardware 1 (7.5Y old)	Hardware 2 (6Y old)	Hardware 3 (4.5Y old)	Hardware 4 (3Y old)	Hardware 5 (1.5Y old)	Hardware 6 (Current)*
On-Premise (non-virtualised)	Worst	3	5%	51,372,685	15,414,061	12,840,312	6,257,229	2,453,698	2,093,779
	Average	2	10%	17,708,754	5,533,001	4,617,433	2,356,780	952,302	820,422
	Best	1.5	25%	5,838,699	2,015,383	1,688,826	950,967	406,652	356,373
Colocation (non-virtualised)	Worst	2.5	5%	42,810,571	12,845,052	10,700,260	5,214,358	2,044,749	1,744,816
	Average	1.8	10%	15,937,879	4,979,702	4,155,690	2,121,102	857,072	738,380
	Best	1.3	25%	5,060,206	1,746,666	1,463,650	824,172	352,433	308,857
On-Premise (virtualised)	Worst	3	6%	43,102,834	13,042,542	10,868,925	5,349,876	2,111,950	1,806,064
	Average	2	30%	6,682,286	2,370,976	1,988,917	1,146,976	496,637	436,802
	Best	1.5	60%	2,944,252	1,185,352	998,841	633,394	287,041	255,673
Private Cloud	Worst	2.5	7%	30,996,498	9,457,166	7,883,993	3,918,139	1,556,537	1,333,795
	Average	1.8	30%	6,014,058	2,133,878	1,790,026	1,032,279	446,974	393,122
	Best	1.3	60%	2,551,685	1,027,305	865,662	548,941	248,769	221,583
Public Cloud	Worst	2	7%	24,797,198	7,565,733	6,307,194	3,134,511	1,245,229	1,067,036
	Average	1.5	40%	3,977,983	1,481,792	1,245,265	746,813	329,759	291,637
	Best	1.1	70%	1,942,527	807,147	680,852	440,725	201,546	179,958

* *As of 2016*

So a non virtualized, on-premise data centre with old servers and a PUE of 3 would consume nearly **300 times more energy to run the same workload** as an optimized environment with new servers and a PUE of 1.1. Bear in mind that the actual average public sector data centre PUE was recorded as 5, not 3. The multiplier effect of outdated IT and inefficient infrastructure is very powerful. But this isn't just about outsourcing - reducing the PUE of an on-premise data centre to 1.5 by improving infrastructure performance and energy stewardship and installing new servers would also be transformational, improving efficiency nearly 150-fold. This is broadly true for any data centre environment, but if commercial drivers are absent, other incentives may be needed to drive that change.