



How to make the edge more sustainable

Data centres consume a massive amount of energy and their usage is only going to increase. If telecoms and technology companies are to meet their sustainability goals, it is imperative to make them run more efficiently. However, the rise of edge data centres could be counterproductive to this goal if their sustainability impacts are not carefully considered.

Kuba Smolorz, Consultant

Data centre sustainability has become an increasingly prevalent topic in recent years, and it's no surprise as data centres now account for around **1% of global energy consumption**. This figure is only going to increase as the demand for cloud computing (including edge) grows. In some countries, data centres will account for an estimated **5-10% of energy consumption by 2030** (24% of energy demand in Ireland by 2030). Huge efforts have already been made by a number of industry players to reduce the environmental impact of data centres. For example, the **Climate Neutral Data Centre Pact** is a movement that sets forth a vision for climate neutrality of data centres in Europe by 2030 by addressing 5 efficiency metrics including 100% carbon free energy and water conservation.

Some data centre providers seeking to minimise their footprint have already made significant headway, particularly with respect to reducing non-renewable energy consumption. Some examples of these are listed below:

- **Increased use of renewable energy:** The IT sector, including cloud data centre providers, are some of the largest pre-purchasers of **renewable energy**.
- **Cooling:** Around 40% of the energy used in data centres is for cooling. Significant investment has gone into state-of-the-art heating, ventilation air conditioning (HVAC) systems. **Google's** fans, ventilation, and other cooling equipment is controlled by AI, helping to lower energy consumption. There have also been advancements in other technologies such as water/ liquid cooling. **Microsoft** is developing an immersion cooling system, which uses a liquid with a boiling point of 50°C to submerge servers. The liquid is safe for electronics, and the heat produced by the processing boils the liquid which dissipates heat from the processor and reduces power consumption associated with cooling.
- **Other efficiencies:** Google's data centres have been using **50% less passive¹ energy** than the industry average, partly through smart lighting and temperature regulation and highly efficient custom-built servers.

Ultimately, many of these advancements require significant investments, which has meant that hyperscale cloud data centre providers (AWS, Google, Microsoft) have been able to make the necessary changes to reduce the impact of their operations in comparison to smaller providers without access to the same resources.

Not only can hyperscaler cloud data centre providers invest more capital in innovation, but the physical characteristics of large data centres, in comparison to multiple smaller data centres can also lend themselves to achieving a lower PUE. For example, smaller data centres will struggle to optimise workloads in the same way that a larger cloud data centre can. While the latter might be able to turn off some of their servers so deal with peaks and troughs in demand, smaller data centres do not have the volume of servers to be able to turn these off. This means that average utilisation of the latter are unlikely to be as high as the former.

However, smaller edge data centres *may* require less passive energy, i.e. for cooling and ventilation relative to their output and size. Having larger volume to surface ratios can help to keep servers cooler than in large data centres. This is particularly true in cooler climates that can benefit from natural cooling. However, in hot climates this benefit becomes increasingly marginal.

However, the decision between using edge data centres or cloud data centres is not predicated by sustainability considerations. Many workloads and applications will simply not be able to run in the cloud, nor would these be economically viable. Like all debates between edge and cloud, the natural answer is that edge and cloud will be used where appropriate. Therefore, the more useful question to ask is: how can edge data centres be more sustainable?

¹ Passive energy is the energy relating to the facilities operations NOT the servers

How can edge data centres be more sustainable?

The move to the edge should be seen as a new opportunity to rethink IT sustainability. Some of the suggestions for delivering on this are included below.

1. Edge data centre providers should take lessons from hyperscale cloud data centres when it comes to cooling

- Where possible, edge data centres should use renewable energy to power edge data centres including lighting, cooling and ventilation.
- Cooling (~40% of data centre energy consumption) could also be made more efficient and effective by using water cooling and HVAC systems and strategies such as separating **hot and cold aisles**. This practice separates the fresh cold air from hot air expelled by the servers. Even seemingly small impacts such as cable management can also have a positive impact on cooling by keeping cables out of critical airflow paths which cool servers.

2. Exploit the inherent technologies in modern chipsets

- For example, multi-core processors can put cores into sleep mode in microseconds (as opposed to several seconds by default). This can massively reduce energy usage, particularly in relation to more volatile workloads in smaller footprint data centres.

3. Edge compute providers should use the smaller scale of edge data centres to their advantage through innovative solutions

- Waste heat from edge servers can be used in innovative ways such as to heat buildings like homes and offices. **Heata** is a green distributed compute network that uses waste heat from compute to heat water in homes. This is done by attaching a compute server directly onto a household hot water tank. The heat generated by the processing is transferred to the cylinder, which has the added benefit of saving on household energy costs.
- New edge data centres that are located in close proximity to homes, offices and schools could explore options to use waste heat productively by engaging with innovative companies like Heata or **Meta** which is looking to provide heat to an existing district heating system in Odense, Denmark's third largest city. Most cloud data centres are located too far from towns or cities with existing district heating to make this viable.

4. Look to innovative hardware that requires less energy

- Using different servers than ones used in traditional cloud data centres can also reduce energy consumption. **G3cco** provides an alternative to traditional server racks with their EdgePods which are miniaturised servers. These smaller servers have a lower environmental footprint given their size (as therefore fewer raw resources to build and lower transporting emissions) and reduced need for cooling infrastructure relative to their size.

How to make the edge more sustainable

- Finally, employing more specialised silicon such as **GPUs (graphics processing unit)** to run edge workloads instead of CPUs (computer processing unit) can help to reduce the carbon footprint of data centres. GPUs have more computing power than CPUs but are much more energy intensive. CPUs and GPUs both contain processing cores, and each processing core can only run a single calculation at once. GPUs contain hundreds or thousands of processing cores which means they can run many calculations in parallel. CPUs on the other hand, only contain 24-48 cores (in a server environment), so can only run 24-48 instructions at once. Individual CPU cores are more intelligent, faster and more efficient than individual GPU cores. However, due to the massive parallelism of GPUs, certain applications are better suited to GPU processing due to their high computing power need, e.g. video analytics, check out our **video analytics at the edge report**. These applications can run faster using a GPU than a CPU and therefore use less overall energy. CPUs are far more versatile and intelligent than GPUs and are therefore more suited to running unspecialised workloads, like ones in a cloud data centre. Edge workloads will require more computing power so are better suited to GPU processing.

Ultimately, while cloud data centres may be able to achieve a lower PUE than edge data centres, the ability to choose where workloads are run is based on the workload requirements rather than the energy efficiency of the data centre. Therefore, edge data centres should look to leverage best practice principals to reduce their impact as the footprint of edge scales.

Kuba Smolorz is a consultant at STL Partners, specialising in 5G, edge computing and B2B strategy.

Get in touch with the author to learn more

kuba.smolorz@stlpartners.com

Or visit STL Partners' Sustainability Hub

www.stlpartners.com/telecoms-sustainability/