# Mitigating the environmental impact of data centres through improved energy efficiency

## Background

Digital ecosystems enable interconnectedness and help to develop and strengthen relationships between a range of partners, including developers, suppliers of digital infrastructure, products and services, and consumer brands. Digital ecosystems have the potential to drive innovation in technology and create value in ways that would not be possible for individual stakeholders working in isolation. The benefits of this collaboration may differ between stakeholders and may include, for example, economic, social and environmental benefits (ECIS, 2019).

In this essay I am going to discuss how collaboration can create value through improved environmental sustainability. I will illustrate this through a case study of Facebook's role in driving improved energy efficiency in data centres.

## Introduction

### Energy consumption of data centres

Electricity consumption by IT industry grew continually during the last decade and the rate of increase is projected to accelerate over the next 10 years. Electricity use by data centres is a large and rapidly growing element of this consumption. Data centres are dedicated buildings that house large groups of networked computer servers, used to remotely store, process and distribute data (Jones, 2018).

Due to their electricity consumption, data centres have significant environmental impact. In 2016 it was estimated that energy use by data centres worldwide was doubling every 4 years (Bawden, 2016). In 2018 data centres used around 200 terawatt hours of electricity, that year accounting for about 1% of global electricity usage and 0.3% of total greenhouse gas emissions (Jones, 2018). With the demand for data centres continuing to increase, it has been predicted that data centres could consume over 10% of the total global energy supply by 2030 (McNerney, 2019). There is therefore a very strong imperative for investment in new technologies to contain and manage this growth.

**Social media demand for data centres**

Social media is an increasingly dominant feature of our society. As of October 2020 over 4.6 million people (59% of the world's population) have access to the internet. Of these, 4.14 billion people are active social media users (Clement, 2020a). Facebook is the most widely used social networking platform, with over 2.7 billion monthly active users worldwide (Clement, 2020b). Facebook, Inc. also owns the social network Instagram, the messaging service WhatsApp, and provides the mobile application Messenger. As of July 2020, 3.14 billion people were using at least one of these core products (Clement, 2020c).

In order to deliver these services, Facebook has a huge demand for data storage, processing and distribution. To meet this need Facebook currently operates 17 data centres worldwide, 13 in North America, 3 in Europe, and 1 in Asia. These data centres use large amounts of energy and Facebook has adopted innovative approaches to containing, mitigating and managing this demand (Facebook, 2020a).

## Analysis

### Energy efficiency of data centres

Data centres consume electricity to run: 1) Core computer operations 2) Cooling systems to prevent the hardware from overheating 3) Other operational requirements such as lighting. These are therefore key target areas for innovation and investment to improve energy efficiency.

The energy efficiency of data centres is measured by the power usage efficiency score (PUE), defined as the total energy used to run a centre divided by the amount of energy only used for computing. Ideally, data centres would have a PUEs of 1, with the only energy used that needed to run computer hardware. However, in reality energy consumed by functions such as cooling and lighting can be close to that used by core computer functions, giving PUEs approaching 2 (Jones, 2018).

In 2009 Facebook was growing exponentially and realised that it would need to radically improve the energy efficiency of its data centres in order to meet this demand, control costs, and minimise environmental impact. It therefore launched a project to design and build a highly energy efficient data centre. The result, two years later, was a facility at Prineville, Oregon with energy running costs 24% lower than their previous data centres (Open Compute Project, N.D. a). Two key elements of this improvement were cooling system developments and the move to hyperscale design.

### Data Centre Cooling

A significant amount of energy used (up to 40% in some cases) by data centres is used for cooling (Torino & Abruzzi, 2015). Several cooling technologies are available, including chilled water systems, computer room air conditioners (CRAC), and direct and indirect evaporative cooling systems.

Chilled water systems draw warm air from an IT environment through chilled water coils, with the heat ultimately rejected through water-cooled, air-cooled or glycol-cooled systems. Chilled water systems are efficient and can be used to cool multiple environments. They are also generally cheaper and have a higher heat removal capacity than CRAC units of an equivalent size (Evans, N.D.).

CRAC systems can also be air-cooled, glycol-cooled, or water-cooled. Air-cooled systems are low cost and easy to maintain, but multiple air conditioners cannot be cooled by a single condenser.  Glycol-cooled systems can run several CRAC units from one dry cooler and pump. Water-cooled systems can also run multiple CARC units from a single cooling tower. However, both glycol-cooled and water-cooled systems have higher installation and maintenance costs compared to air-cooled systems (Evans, N.D.).

Direct and indirect air evaporative cooling systems use outside air as their main cooling source. Direct systems draw outdoor air through a filter and into the data centre, whilst indirect systems use air-to-air heat exchange and air from outside does not enter the IT environment. In both systems a wet mesh or water spray can be used to further cool the air, allowing operation at relatively higher outdoor temperatures (Evans, N.D.). Air evaporative cooling systems are more efficient in colder climates, but data protection laws in many countries now require personal data to be stored on domestically located servers, thus limiting the choice of location (Trueman, 2019).

Starting at Prineville, Facebook has developed highly energy and water efficient direct evaporative cooling technology. This system uses 100% outside air economisation which is much more efficient than recirculating the air inside a data centre. It eliminates the need for chillers or cooling towers and does not require water used in the cooling system to be chemically treated (Open Compute Project, 2013).

In 2015, alongside Nortek, Facebook began to develop a new indirect evaporative cooling system called StatePoint Liquid Cooling. This allows water to evaporate through a membrane which cools both the water and the outside air. This cooled air is then used to cool the processed air which circulates inside the centre. Although direct evaporative systems are still the most efficient solution, this indirect system will allow more efficient cooling in a wider range of environments, where direct systems may be not be appropriate due to external factors such as poor air quality (Mulay, 2018).

Facebook has also trialled other cooling technologies such as immersion cooling. This involves submerging servers into non-conductive oil baths. It can be a highly effective cooling method, but currently remains a specialist solution with difficult maintenance needs (Jones, 2018).

Further energy savings can also be made by carefully matching cooling distribution within data centres to demand at any given time. In 2016 DeepMind and Google began to develop an artificial intelligence (AI) powered system to further improve the energy efficiency of its data centres. By taking measurements from thousands of sensors across a data centre, this system can make constant adjustments to the cooling systems to minimise energy consumption. As an AI driven system, it constantly feeds the changes made by its own adjustments back into its own learning model, thus becoming more efficient over time as more data is collected. By implementing this technology across several of its own data centres, Google has seen consistent energy savings of around 30% (DeepMind, 2018).

### Hyperscale Data Centres

There are three main types of data centre, namely colocation, enterprise and hyperscale data centres. Colocation centres are owned by a third party, with space rented to individual companies and often serving hundreds of customers at a time. Enterprise data centres are owned and operated by the company they support. Hyperscale data centres are more specialised, designed and built to service the specific needs of a company and to be more easily scaled to accommodate higher demands (AFL Hyperscale, 2020).

Hyperscale centres are predominantly used by large companies including Amazon, Google, and Facebook. They are designed to be more environmentally friendly and more cost effective than conventional data centres (AFL Hyperscale, 2020). As these centres are purpose designed and built, choices can be made which reduce space and increase efficiency. For example, hardware can be stripped back to the core essentials and highly automated running and maintenance systems allow server racks to be densely packed with little or no lighting. These small savings in space and energy have a significant impact when introduced at scale. The power savings made by hyperscale centres are demonstrated by their PUEs, which are typically around 1.2, with some managing to achieve PUEs as low as 1.12 (Jones, 2018).

The Facebook data centre at Prineville is an early example of a hyperscale data centre. The benefit of this increased efficiency for such large companies is obvious. However, there are potentially even greater benefits if this technology could be accessed widely by other organisations. It has been suggested that if 80% of smaller data centres in the U.S.A. were converted to hyperscale facilities, this could result in a 25% reduction in energy consumption (Shehabi, 2016).

**The Open Computer Project**

A decade ago, Facebook recognised the global potential of the advances in energy efficiency it had achieved. As a result, Facebook decided to make their designs for the Prineville data centre publicly available and also to use this as an opportunity to start an open source hardware project. In 2011 this led to the launch of the Open Compute Project (OCP), designed to enable collaboration and creativity, promote innovation and help realise the benefits of technological advances on a wider scale. OCP has grown from 5 founding members into rapidly expanding global community, with well over 200 corporate members currently sharing intellectual property to help promote innovation in the IT industry (Open Compute Project, N.D. a). The origins of OCP are still identifiable through the current Data Center Facilities project and its sub projects, including one on Advanced Cooling Facilities. However, this now sits alongside 14 other main projects, covering areas including telecommunications, networking and security (Open Compute Project, N.D. b). OCP not only continues to drive technological innovation and development, but also creates opportunities for wider adoption of new technologies across the industry.

## **Discussion**

Developments in cooling technology and data centre design driven by Facebook have made direct and significant contributions to improved energy efficiency of data centres. Although creating these improvements has required significant investment, they have resulted in reduced energy costs and therefore have financial benefits for the organisation. They have also created significant additional value, through reducing environmental impact and improving sustainability.

The next step towards improving energy efficiency of data centres around the world is to encourage wider adoption of these new technologies. Although the benefits of these technologies are clear, there are also challenges associated with adoption, including upgrading older centres rather than building new. Servers in most centres are upgraded every 3 to 5 years, so over time more efficient technology, possibly employing a hyperscale model, could be installed. However, in smaller centres upgrades might necessitate taking the entire data centre offline for a period of time. A second drawback of implementing new technologies, especially for smaller centres, is cost. Over time, as with most technology, the costs involved may become more affordable for all.

The affordability of new technologies is one area where Facebook should be in a position to help. Facebook reported annual revenue of over USD 70 billion in 2019 (Facebook, 2020b). The majority (98.5%) of this income is advertising revenue, using technologies such as tracking pixels to enable more targeted advertisements (Gunnars, 2020). This financial strength gives Facebook the ability to continue to support OCP to develop affordable data centre solutions and models to facilitate adoption of new technologies in established facilities. In Technology Adoption Models (TAM), perceived ease of use and usefulness are the two main factors that influence changes in usage over time (Lai, 2017). Lowering the barrier to entry for this technology increases the ease of use, and the usefulness in terms of increased energy efficiency speaks for itself. Following the TAM, the combination of these two factors will encourage wider adoption.

## Conclusion

This analysis and discussion highlights the environmental impact of data centres. However, it also demonstrates how recent technological advances have resulted in significant improvements in their energy efficiency and how Facebook's improvements in this field have fed into the wider digital ecosystem to create additional value for Facebook and other stakeholders. Value is created through direct improvements in environmental sustainability, as well as through enabling innovation in other fields of IT. It has also allowed Facebook to accrue more social capital, which is valuable as public awareness of environmental issues caused by technology companies grows. Further improvements in energy efficiency will require continued investment in research and development of hardware and software solutions. More research into models to facilitate widescale adoption of resultant new technologies is also needed.

**Word Count:** 2199 (excluding references)

**ECIS 2020 Conference Track:** *Digital Ecosystems: The Economic, Social, and Environmental Impact of Pervasive Connectivity*

## References

AFL Hyperscale. (2020) Understanding Different Types Of Data Center. Availabe from: https://www.aflhyperscale.com/understanding-different-types-of-data-center [Accessed 10 December 2020]

Bawden, T. (January 23, 2016) Global warming: Data centres to consume three times as much energy in next decade, experts warn. *Independent.* Available from: https://www.independent.co.uk/environment/global-warming-data-centres-consume-three-times-much-energy-next-decade-experts-warn-a6830086.html [Accessed 23 October 2020]

Clement, J (2020a) Global digital population as of October 2020. Available from: https://www.statista.com/statistics/617136/digital-population-worldwide/ [Accessed 8 December 2020]

Clement, J. (2020b) Facebook: number of monthly active users worldwide 2008 – 2020. Available from: https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/ [Accessed 22 October 2020]

Clement, J. (2020c) Monthly active users of any Facebook product 2020. Available from: https://www.statista.com/statistics/947869/facebook-product-mau/ [Accessed 22 October 2020]

DeepMind. (2018) Safety-first AI for autonomous data centre cooling and industrial control. Available from: https://deepmind.com/blog/article/safety-first-ai-autonomous-data-centre-cooling-and-industrial-control [Accessed 10 December 2020]

ECIS. (2019) Digital Ecosystems: The Economic, Social, and Environmental Impact of Pervasive Connectivity. Available from: http://ecis2019.eu/programme/research-tracks/digital-ecosystems-the-economic-social-and-environmental-impact-of-pervasive-connectivity [Accessed 13 December 2020]

European Union. (N.D.) The economy. Available from: https://europa.eu/european-union/about-eu/figures/economy_en [Accessed 21 October 2020]

Evans, T. (N.D.) The Different Technologies for Cooling Data Centres. Available from: https://download.schneider-electric.com/files?p_Doc_Ref=SPD_VAVR-5UDTU5_EN [Accessed 9 December 2020]

Facebook. (2020a) Sustainable data centers. Available from: https://sustainability.fb.com/innovation-for-our-world/sustainable-data-centers/ [Accessed 20 October 2020]

Facebook. (2020b) Facebook Reports Fourth Quarter and Full Year 2019 Results. Available from: https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-Fourth-Quarter-and-Full-Year-2019-Results/default.aspx [Accessed 20 October 2020]

Facebook. (N.D.) Specifications for Facebook pixel standard events. Available from: https://business.facebook.com/business/help/402791146561655?id=1205376682832142&recommended_by=952192354843755 [Accessed 21 October 2020]

Gunnars, K. (2020) How Does Facebook Make Money? 7 Main Revenue Sources. Available from: https://stockanalysis.com/how-facebook-makes-money/ [Accessed 21 October 2020]

Jones, N. (2018) How to stop data centres from gobbling up the world's electricity. *Nature* 561:163 – 166. DOI: https://doi.org/10.1038/d41586-018-06610-y

Lai, P. (2017) The Literature Review of Technology Adoption Models and Theories For The Novelty Technology. *Journal of Information Systems and Technology Management* 14(1): 21-38 Available from: https://www.researchgate.net/publication/317412296_THE_LITERATURE_REVIEW_OF_TECHNOLOGY_ADOPTION_MODELS_AND_THEORIES_FOR_THE_NOVELTY_TECHNOLOGY [Accessed 12 December 2020]

Liu, D. (2019) Simplifying Payments with Facebook Pay. Available from: https://about.fb.com/news/2019/11/simplifying-payments-with-facebook-pay/ [Accessed 23 October 2020]

McNerney, M. (2019) The Data Center Dilemma: Is Our Data Destroying the Environment? Available from: https://www.datacenterknowledge.com/industry-perspectives/data-center-dilemma-our-data-destroying-environment [Accessed 23 October 2020]

Mulay, V. (2018) StatePoint Liquid Cooling system: A new, more efficient way to cool a data center. Available from: https://engineering.fb.com/2018/06/05/data-center-engineering/statepoint-liquid-cooling/ [Accessed 13 December 2020]

Open Compute Project. (N.D. a) About. Available from: https://www.opencompute.org/about [Accessed 12 December 2020]

Open Compute Project. (N.D. b) Projects. Available from: https://www.opencompute.org/projects [Accessed 12 December 2020]

Open Compute Project. (2013) Water Efficiency at Facebook's Data Centre. Available from: https://www.opencompute.org/blog/water-efficiency-at-facebooks-prineville-data-center [Accessed 13 December 2020]

Shehabi, A. et al. (2016) *United States Data Center Energy Usage Report.* Available from: https://eta.lbl.gov/publications/united-states-data-center-energy [Accessed 13 December 2020]

Torino, P. & Abruzzi, C. (2015) Cooling Systems in Data Centers: State of Art and Emerging Technologies. *Energy Procedia* 83: 484 – 493. DOI: https://doi.org/10.1016/j.egypro.2015.12.168

Trueman, C. (August 9, 2019) Why data centres are the new frontier in the fight against climate change. *Computerworld.* Available from: https://www.computerworld.com/article/3431148/why-data-centres-are-the-new-frontier-in-the-fight-against-climate-change.html [Accessed 23 October 2020]