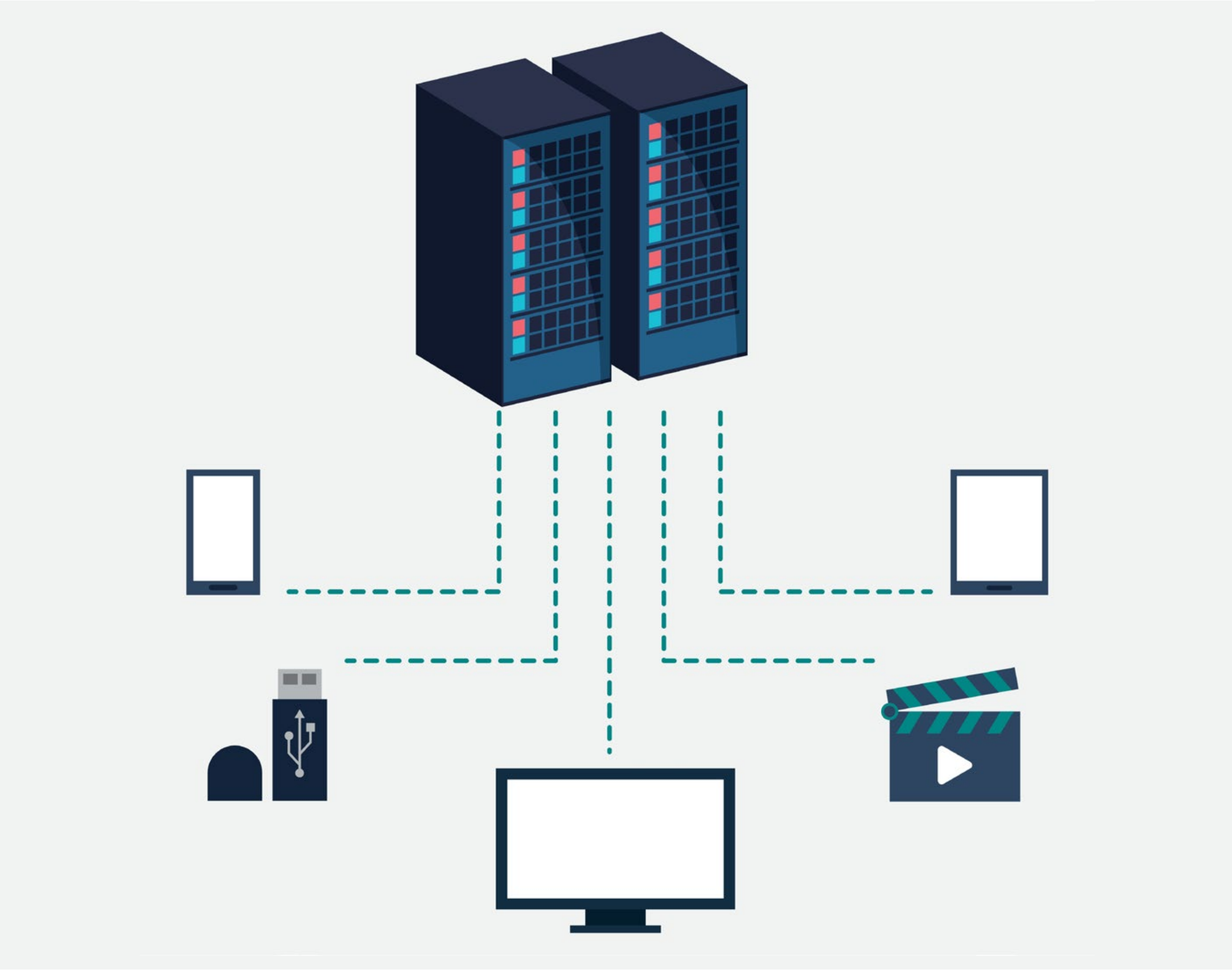
COPENHAGEN CENTRE ON ENERGY EFFICIENCY
SEforALL EE HUB

BRIEF 2

Data Centres: Digitalisation Powerhouse and Energy Efficiency Potential

KEY MESSAGES

- The digitalisation of the economy worldwide and technological innovations such as artificial intelligence (AI), internet-of-things (IoT) and blockchain are driving exponential growth in the demand for data centres' services.
- Data centres use approximately 200 TWh of electricity annually, corresponding to roughly 1% of global electricity demand[iv].
- To curb the rapid growth in the energy use of data centres, it is key that the future demand for data centres' services will be met by energy-efficient data centres and that their uptake of renewable energy is accelerated.
- Policy-makers should establish robust mechanism to collect data and publish statistics on the energy use of data centres, similarly to what has been done for other energy-intensive sectors.



WHAT IS A DATA CENTRE?

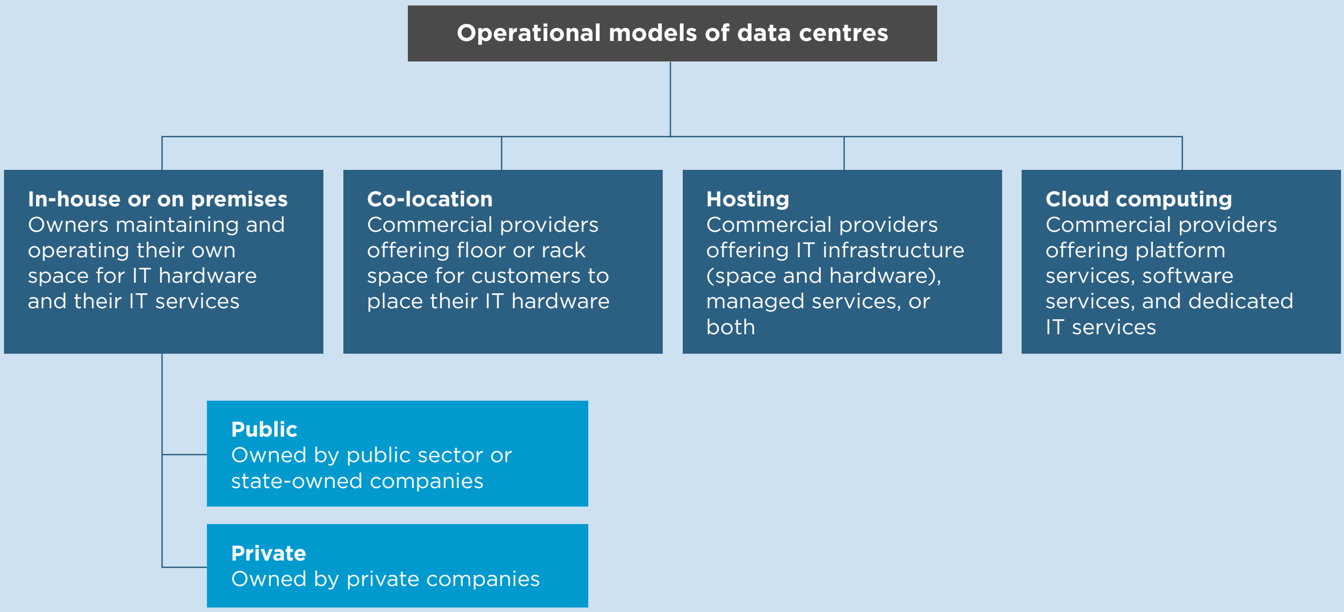
A data centre is a dedicated building or a separated room, which houses the technology for data processing, data storage and data communication of one or more organisations[i]. This space is exclusively used for the placement of information technology (IT) hardware, such as computing equipment (servers), data storage devices (hard and tape drives), all the network equipment (routers, switches, modems) for data communication, and the necessary power and cooling infrastructure. The data centre is also equipped with a resilient power supply, high bandwidth connectivity, sophisticated security systems, and controls for the management of the building, including its acclimatisation.

Some organisations have storage rooms where they keep a rack with some servers, among other equipment, but this is not considered a data centre because this is not dedicated infrastructure for data processing, data storage, and data communication.

WHY DO WE NEED DATA CENTRES?

Everyday activities such as reading an email, submitting a tax declaration, making a bank transfer, paying a bill, searching a word on the internet, and using a smartphone are all supported by data centres. The digitalised economy and networked society rely on data and connectivity, which need to be managed in a secure and efficient way. Data centres are a fundamental part of national infrastructure because data underpins many activities across the public and private sectors. It is not an exaggeration to say that data centres enable service economies just like in the past heavy industries enabled industrial economies. In an increasingly digitalised world with a growing demand for data management, processing, and storage, data centres play a key role as the supporting infrastructure to the digitalised economy. The application of AI, the deployment of IoT, and the mining of cryptocurrencies further contribute to growth in the demand for data centre services. It is estimated that in 2021 there will be 7.2 million data centres in the world, storing about 1.3 Exabytes of data. This represents nearly an eight-fold increase in the amount of stored data, compared to 2015, mostly due to the growth of big data storage[ii].

Figure 1: Typical operational models of data centre operators. Authors' creation based on Hintemann (2015)



HOW CAN WE CHARACTERISE DATA CENTRES?

OPERATIONAL MODELS OF DATA CENTRES

Data centres operate based on two main types of assets: (1) the physical infrastructure, i.e. building and equipment, and (2) the service offerings, i.e. storage, management and maintenance, and security of data. These assets are owned and managed differently in each case, translating into different operation models (Figure 1). Aside from the organisations that choose to have their data centres in house, there are commercial providers that offer different layers of services to customers.

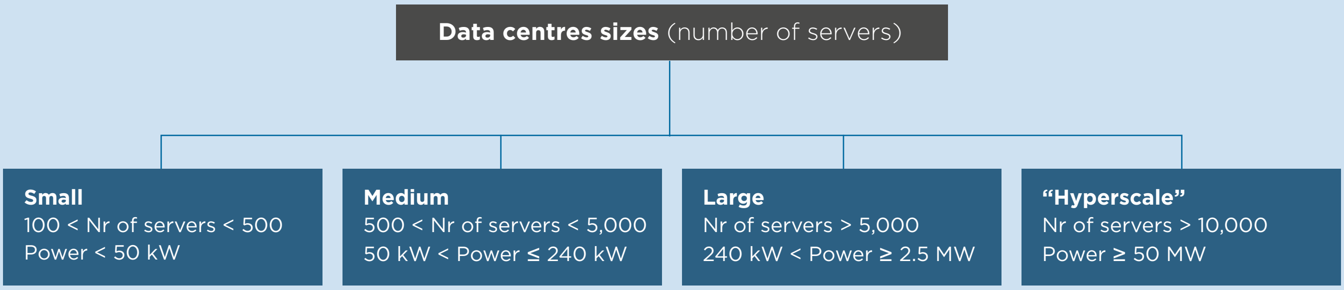
The most common operational models for data processing and storage services are:

- **In-house data centres:** also known as on premise data centres, the key principle behind this model is that the organisation, either public or private, manages its own space for the data processing and storage technology it uses. Typically, the IT department of the organisation will be responsible for all aspects of infrastructure and data centre management. Examples of companies with in-house data centres are commercial banks and telecommunications.
- **Co-location data centres:** the data centre operating company builds and owns the dedicated building and rents space in it with all the necessary power and cooling infrastructure. The customers place their own servers and data storage devices in there. Additional services may be provided, besides space, as for example IT hardware maintenance and interconnection between customers and their key partners. The co-location company guarantees security, resilience, connectivity and power. Examples of companies offering colocation services are Interxion, GlobalConnect, DigiPlex, Digital Realty, and Equinix. Under this model there are retail data and wholesale data centres. Retail co-location consists of renting space within a data centre, namely a rack, or rack space, or a caged space. Whereas wholesale co-location refers to renting entire rooms or entire buildings.
- **Hosting data centres:** the data centre operating company provides a variety of services from managed services to hosting. The range of services provided can vary: some data centre operators lease all the IT hardware for data processing, storage and communication, and the customers manage that hardware and install their own software and applications; others also manage the hardware and offer virtualised computing resources; and even others will operate the IT hardware that is property of their customers (managed hosting services). Customers pay for the services they use, and typically these are contracted under a service level agreement. These data centre operators may own the data centre building or rent space in co-location data centres. Examples of companies offering these services are Rackspace and CenturyLink.
- **Cloud data centres:** the data centre operating company builds and runs the data centre completely, including all the infrastructure and IT hardware, and sells services related to software and applications to

Table 1. Data centres' rating tiers and their characterisation. Source: Pitt et al.[iii]

Rating tier	Uptime	Downtime per year (hours)	Type of redundancy
Tier 1	99,671%	28.8	None
Tier 2	99.749%	22	Partial (N+1 components)
Tier 3	99.982%	1.6	Fault tolerant (2N+1 components)
Tier 4	99.995%	0.04	Fault tolerant (2N+1 components)

Figure 2. Classification of data centres according to size (number of servers). Authors' creation, based on Schomaker et al.[ix] and Hintemann[i]



customers. Service offerings include platform-as-a-service, software-as-a-service, infrastructure-as-a-service, or dedicated off-the-shelf or bespoke IT services. Customers pay only for their needed services, and this model is sometimes referred to as pay-as-you-go model. Examples of companies offering these services are Microsoft Azure, Apple's iCloud, Google Cloud Platform, Alibaba Cloud, Oracle Cloud, Amazon Web Services, IBM's private cloud services, Fujitsu, and Atos.

Whereas some data centre providers operate only under one of the models above, others offer a combination of services that combine elements of different operating models.

DATA CENTRE TIERS

Data centre tiers are an efficient way to describe the infrastructure components being utilised at a business's data centre. Although a Tier 4 data centre is more complex than a Tier 1 data centre, this does not necessarily mean it is best suited for businesses' needs.

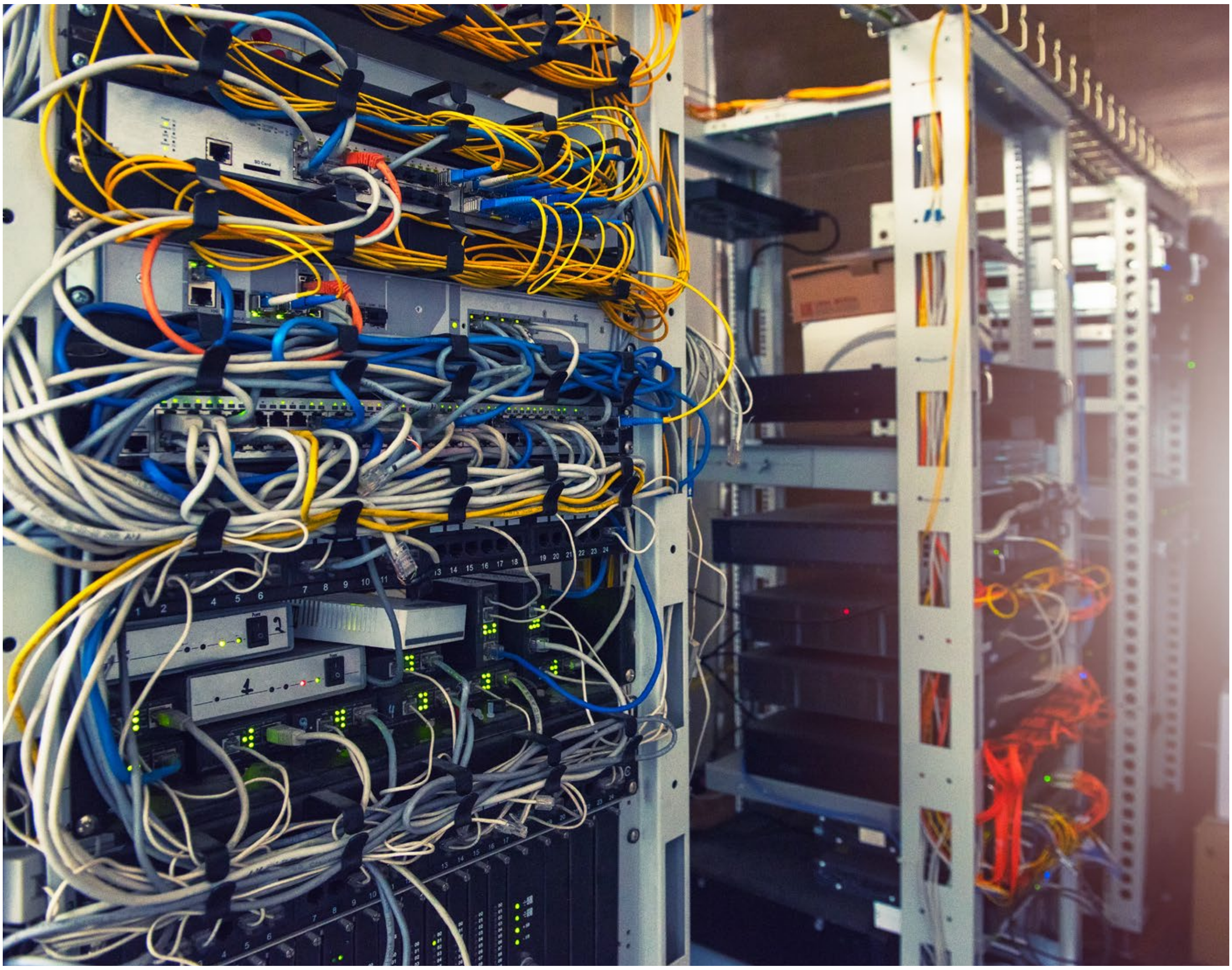
DATA CENTRE SIZES

Another way to characterise data centres is in terms of their sizes; this is considering the number of servers they contain or the space occupied by the data centre building or room. They can range from a small size, holding less than 500 servers; to a hyperscale data centre, containing more than 10,000 servers (Figure 2).

Despite the trend for data centres to become bigger, in line with the continued increase in demand for data centre services, smaller data centres will not disappear. In fact, the number of smaller data centres is expected to increase because edge data centres (or edge computing) are becoming more relevant and necessary in the context of technological developments like IoT and smart cities, which require localised information processing and very demanding latency requirements in data communication.

WHAT IS THE ENERGY USE OF A DATA CENTRE?

It is estimated that data centres account for about 1 to 1,5% of the total energy use worldwide, corresponding to approximately 200 TWh of electricity[iv]. Although the total size of data stored in data centres has witnessed rapid increase in the last decade, the global energy use of data centres has increased only modestly due to the integration of energy efficiency measures. Indeed, whereas the energy use increased by 6% in the period 2010 to 2018, the number of computing instances in data centres increased by 550%[iii]. Efficiency improvements in processors and in

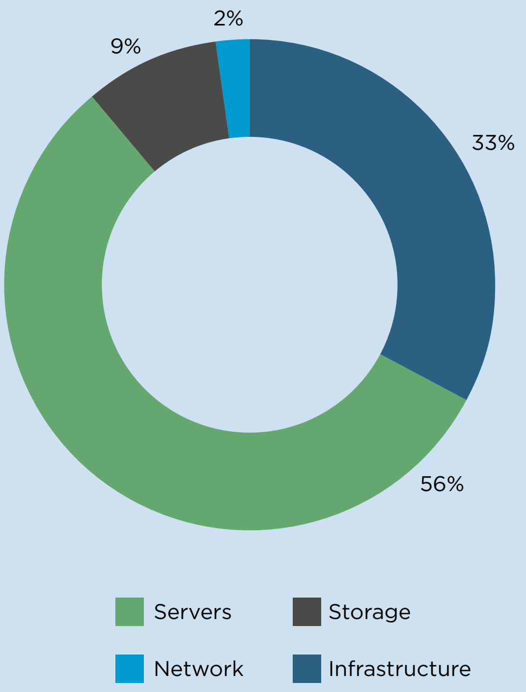


storage drives, efficient management of servers through virtualisation and consolidation, and ongoing shifts from traditional data centres to more energy-efficient cloud data centres contributed to this decoupling of growth in energy use from growth in computing instances. However, to keep the energy use of data centres in check, it is key to ensure that the future demand for data centres' services will be met by energy-efficient data centres that incorporate renewable energy sources.

Data centres are very energy-intensive buildings that require 10 to 100 times more electricity per floor space area than other commercial building types. The electricity demand of a data centre is split across four main components[v]:

- **Infrastructure:** The infrastructure is what supports the operation of IT equipment. It consists of the equipment for cooling, power distribution, lighting and security, backup batteries and generators. A metric that measures the effectiveness of the infrastructure is power usage effectiveness (PUE), which represents the total data centre energy use relative to the IT equipment energy use[iv]. A PUE value close to 1 indicates high effectiveness of the supporting infrastructure, where almost all electricity consumed at the data centre is dedicated to IT equipment. Despite the wide use of this metric across the industry, it is however inadequate to capture the efficiency of a data centre because it indicates nothing about the efficiency of the IT equipment and the performance of the data centre (e.g. number of computations) in relation to its energy use[vi].
- **Servers:** The servers are typically the IT component that uses more energy in a data centre. The power draw of servers is partially related to its utilisation. Most servers run below its maximum power capacity. Consolidation of servers, that is, substituting multiple servers running at low capacity by a single consolidated server running at higher capacity is a key measure for energy efficiency improvement[iv].
- **Storage drives:** The storage in a data centre includes storage in hard disk and in solid state technologies, which have different energy usage[iv]. Hard disk drives typically have higher energy usage, but continuous innovation is pushing their energy consumption down, in convergence with the values of solid state drives.

Figure 3. Share of energy demand by different components in data centres globally (2020). Elaborated with data from International Energy Agency[vi]



- **Networking equipment:** The energy use of networking equipment is the least significant among the IT equipment in a data centre. The trend of greater use of mobile networks might have some impact of this because the electricity intensity of mobile networks is higher than fixed-line networks. Despite 4G and 5G being less energy-intensive than 3G, the higher speed of data transmission that they allow may result in greater traffic volumes in mobile networks[vi].

Figure 3 shows the relative contribution of each component to the overall electricity demand by data centres worldwide in 2020, based on data published by the International Energy Agency (2019)[vii].

HOW DOES THE DATA CENTRE USE ENERGY IN ITS OPERATIONS?

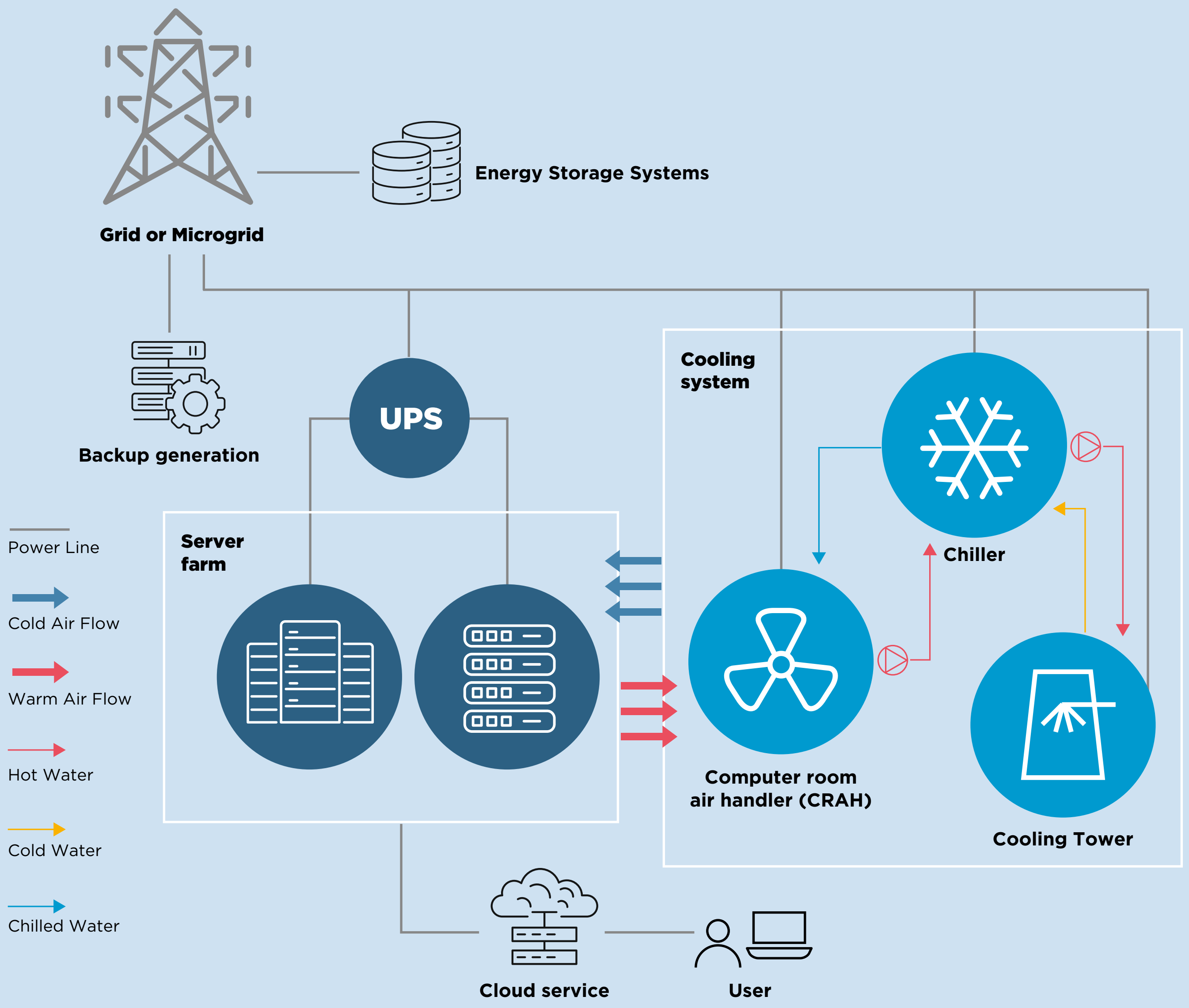
From a holistic perspective, data centres use energy in all stages of their lifecycle, namely in the construction, operation, and recycling phases. This section focusses on the energy use of data centres in the operation phase. A data centre operates different types of equipment, all of which need electricity. In general, data centres include the following equipment[viii]:

- **Server farms:** equipment to process and store data, namely clusters of interconnected servers, stacked in racks or closed cabinets
- **Networking and communication:** equipment to maintain a high-bandwidth communication of data, including routers, switches, and cabling
- **Power distribution:** equipment to provide electricity to the data centre
- **Backup power generation:** equipment to provide backup electricity supply in case of failure of utility supply
- **Energy storage:** battery systems to provide electricity instantaneously during backup power supply start-up delay
- **Heat and airflow system:** equipment to ensure that the ambient temperature around the servers is within the appropriate range. The airflow in data centres is set up in a way that forms hot air aisles and cold air aisles. The heat and airflow system needs to extract hot air from equipment and distribute cool air. Common heat and airflow systems are the computer room air handler (CRAH) and the computer room air conditioning (CRAC)
- **Cooling system:** equipment to cool down the air being blown into the cold air aisles of the data centre. This is typically used in conjunction with computer room air handler (CRAH) units, as the computer room air conditioners use their own condensers for reducing air temperature
- **Control and security system:** equipment to control and manage power flow, making critical decisions at the time of emergency, and ensuring safety and security of the data stored

A data centre can be seen as an energy converter system that consumes electric power and converts it into heat. The data centre is a closed system in which the only connections with the outer world are electrical power and internet. Figure 4 shows an example of how data centre components are interconnected, in which the cooling system is a water chiller.

Data centres are concerned mainly with security and availability of its services[ix]. Security means that data centre operators need to prevent malicious online activity, reduce to the greatest extent the possibility of physical break-ins, and minimise the impact of natural disasters and eventual misuse by unqualified personnel. Availability means reducing to the greatest extent the impact of power failures, hardware failures and similar faults in order to prevent service downtime. To address these matters, data centres need a combination of redundant systems for all types of equipment they use. These redundant systems add up in terms of the overall energy demand for the data centre. The level of redundancy and safety of a data centre is described by its tier value: the higher the tier value, the greater level of security and availability it can provide.

Figure 4. Architecture of a typical data centre cooled by a water chiller. Source: Adapted from Rahmani et al.[vii]



HOW TO INCREASE THE ENERGY EFFICIENCY OF DATA CENTRES?

Given data centres' growing energy demand and their key role in digitalisation of the economy, efforts to ensure sustainable development of data centres are a prime consideration. These efforts focus primarily on improving its energy efficiency during operation, particularly to the IT equipment, the cooling system, and the power delivery system[iii]. Examples of such efforts include consolidation of servers, replacing chips and servers to gain operational efficiency, utilising the heat from the servers for district heating, using air and water-based free cooling, shifting to hyperscale systems, cooling servers with isolating materials, and drawing on AI for regulating the data centre's cooling system[x]. Also, addressing material shortage can include improving materials efficiency during equipment manufacturing or enhancing material reuse and recycling at the end-of-life of the equipment[xi].

Beyond energy efficiency, big players in the ICT sector, such as Google, Microsoft, and Facebook, have also started to increase the use of renewables that power their data centre operations through, for instance, photovoltaics or wind turbines. To this end, a critical factor when choosing the location of a data centre is the renewable energy content of the energy supply, the stability of the power grid and the average annual ambient temperature.

However, much higher public data and modelling competencies are required to monitor data centre energy use in order to develop and evaluate evidence-based policies. National policy-makers should, therefore, appoint robust data collection and open data repository systems. Such efforts are particularly important given the data centre energy use growth projections[iii].

COPENHAGEN CENTRE ON ENERGY EFFICIENCY
SEforALL EE HUB

ENDORSED BY

U4E United for Efficiency — UN environment

UN environment programme

ACKNOWLEDGEMENTS

AUTHORS

Ana Cardoso
Researcher, UNEP DTU Partnership

Clara Camarasa
Post-doc Researcher, Copenhagen Centre on Energy Efficiency, UNEP DTU Partnership

CONTRIBUTORS

Geir M. Køien
Professor in Cybersecurity, University of South-Eastern, Norway

Xiao Wang
Programme Officer. Copenhagen Centre on Energy Efficiency, UNEP DTU Partnership

The Copenhagen Centre on Energy Efficiency functions as the global thematic Energy Efficiency Hub of Sustainable Energy for All (SEforALL), and accordingly works directly to support the SEforALL objective of doubling the global rate of improvement in energy efficiency by 2030.

The Copenhagen Centre fulfills its mission through:

- assisting policy change in countries and cities, with knowledge, insights and technical support
- accelerating action through innovation in project development and finance
- raising the profile of energy efficiency by communicating success stories and supporting outreach.

For more information, please visit **www.energyefficiencycentre.org** or contact us at **c2e2@dtu.dk**.

Regarding our work in Sustainable Data Centres and Smart Energy, please contact Xiao Wang at **xwang@dtu.dk**

Visit Copenhagen Centre's Knowledge Management System at **kms.energyefficiencycentre.org**

The Copenhagen Centre on Energy Efficiency is institutionally part of UNEP DTU Partnership (UDP). UDP is a UN Environment Collaborating Centre and a leading international research and advisory institution on energy, climate and sustainable development.

REFERENCES

i Hintemann, R. (2015). Consolidation, colocation, virtualisation, and cloud computing: The impact of the changing structure of data centers on total electricity demand. In L. M. Hilty & B. Aebischer (Eds.), *Advances in Intelligent Systems and Computing* (Advances i, Vol. 310, pp. 125–136). https://doi.org/10.1007/978-3-319-09228-7_7

ii Statistics from https://www.statista.com

iii Pitt, W., Iv, T., Seader, J. H., & Brill, K. G. (2001). Industry Standard Tier Classifications Define Site Infrastructure Performance. The Uptime Institute.

iv Masanet, E., Shehabi, A., Lei, N., Smith, S., & Koomey, J. (2020). Recalibrating global data center energy-use estimates. Science, 367(6481). https://doi.org/10.1007/s12053-019-09809-8

v Shehabi, A., Smith, S. J., Masanet, E., & Koomey, J. (2018). Data center growth in the United States: decoupling the demand for services from electricity use. Environ. Res. Lett, 13, 124030. https://doi.org/10.1088/1748-9326/aaec9c

vi Shehabi, A., Smith, S., Horner, N., Azevedo, I., Brown, R., Koomey, J., … Sartor, D. (2016). United States Data Center Energy Usage Report. Retrieved from https://eta.lbl.gov/publications/united-states-data-center-energy

vii IEA (2019), Tracking Buildings, IEA, Paris https://www.iea.org/reports/tracking-buildings-2019/data-centres-and-data-transmission-networks#abstract

viii Rahmani, R., Moser, I., & Seyedmahmoudian, M. (2018). A complete model for modular simulation of data centre power load. *arXiv preprint arXiv*:1804.00703

ix Schomaker, G., Janacek, S., & Schlitt, D. (2015). The Energy Demand of Data Centers. In L. M. Hilty & B. Aebischer (Eds.), *ICT Innovations for Sustainability* (Advances i, pp. 113–124). https://doi.org/10.1007/978-3-319-09228-7_6

x Huusko, J., de Meer, H., Klingert, S., & Somov, A. (Eds.). (2012). Energy Efficient Data Centers. https://doi.org/10.1007/978-3-642-33645-4

xi Laurent, A., and Dal Maso, M. (2020). Environmental sustainability of data centres: A need for a multi-impact and life cycle approach. Data Centre Brief Series. Copenhagen Centre on Energy Efficiency.