# Data centres and the power system: expected trends, challenges and opportunities

May 2026

entsoe

CONNECTING EUROPE,
ELECTRIFYING THE FUTURE

# Foreword

**ENTSO-E, the European Network of Transmission System Operators for Electricity, is the association of the European transmission system operators (TSOs).** The **40 member TSOs, representing 36 countries,** are responsible for the secure and coordinated operation of Europe's electricity system, the **largest interconnected electrical grid in the world.**

**Before ENTSO-E was established in 2009, there was a long history of cooperation among European transmission operators, dating back to the creation of the electrical synchronous areas and interconnections which were established in the 1950s.**

**In its present form, ENTSO-E was founded to fulfil the common mission** of the European TSO community: to power our society. At its core, European consumers rely upon a secure and efficient electricity system. Our electricity transmission grid, and its secure operation, is the backbone of the power system, thereby supporting the vitality of our society. ENTSO-E was created **to ensure the efficiency and security of the pan-European interconnected power system** across all time frames within the internal energy market and its extension to the interconnected countries.

**ENTSO-E is working to secure a carbon-neutral future.** The transition is a shared political objective through the continent and necessitates a much more electrified economy where sustainable, efficient and secure electricity becomes even more important. **Our Vision: "a power system for a carbon-neutral Europe"**[*] shows that this is within our reach, but additional work is necessary to make it a reality.

In its Strategic Roadmap presented in 2024, ENTSO-E has organised its activities around two interlinked pillars, reflecting this dual role:

- "Prepare for the future" to organise a power system for a carbon-neutral Europe; and
- "Manage the present" to ensure a secure and efficient power system for Europe.

**ENTSO-E is ready to meet the ambitions of Net Zero, the challenges of today and those of the future for the benefit of consumers, by working together with all stakeholders and policymakers.**

---

[*] **https://vision.entsoe.eu/**

# Contents

# Executive Summary

The rapid growth of artificial intelligence (AI) and digital services is fundamentally transforming the role of data centres in the European power system. Rather than niche consumers with predictable demand profiles, data centres are becoming systemically relevant electricity loads whose behaviour increasingly affects secure grid operation. Total data centre electricity demand in Europe is expected to grow by more than 50% between 2025 and 2030, driven primarily by colocation and hyperscale expansion concentrated in established metropolitan hubs.[1, 2, 3] This growth poses challenges for electricity transmission system operators (TSOs).

Data centres are not only large-scale localised loads, they also have software-driven load profiles that can introduce stability risks, some of which manifest at speeds and frequencies that conventional Supervisory Control and Data Acquisition (SCADA) systems cannot detect, requiring Phasor Measurement Units (PMUs) with high-resolution sampling for identification and analysis; operational modelling tools are also necessary in order to investigate dynamic behaviours of a large data centres and interactions among them. Their Uninterruptible Power Supply (UPS)-based power architectures can instantaneously disconnect hundreds of megawatts during minor grid disturbances, potentially amplifying problems to safe grid operation and power quality. If this risk of sudden large-load disconnection persists, transmission system operators (TSOs) may be forced to operate at a lower renewable energy sources (RES) penetration level.

Beyond stability concerns, the concentrated deployment of data centres in already constrained regions is straining transmission capacity and may, over time, also challenge generation adequacy, as already seen in the United States and Ireland. At the same time, access to power is critical for data centres, and connection lead times – ranging from a few years to more than a decade – are now the sector's most frequently cited concern.

However, some of the same technical characteristics that create these challenges also present an opportunity. The sophisticated power electronics, battery systems, cooling infrastructure, and controllable IT workloads embedded in data centres can, under the right regulatory and market frameworks, be leveraged as active grid resources. Updated connection codes can ensure grid-safe behaviour both during faults and normal operation, Safeguarding grid security while avoiding restrictions for renewable integration. Transparent hosting-capacity information and reformed grid connection capacity allocation processes can steer development towards regions where the grid can accommodate new load. Flexible connection agreements can help close the time-to-power gap. Once the control architecture for grid-safe operation is in place, data centres can progress towards providing flexibility services and participating in electricity markets as demand resources, or in some cases as virtual power plants (VPPs).[4]

This report examines these dimensions from the perspective of electricity TSOs: how data centres consume electricity (Chapters 1–2), what challenges they pose to safe grid operation and how connection requirements are responding (Chapter 3), what challenges they pose for grid planning and how these can be addressed (Chapter 4), and how their inherent flexibility can be harnessed as an opportunity for grid support and market participation (Chapter 5).

While data centres inherently possess significant potential to act as flexible grid resources, this theoretical capacity is heavily filtered in practice. A cascading series of technical limitations, business model realities, and market participation barriers progressively narrows the actual flexibility these facilities can offer to the system.

---

[1] IEA, "Energy and AI," International Energy Agency, 2025. [Online].
Available: https://www.iea.org/reports/energy-and-ai

[2] European Data Centre Association (EUDCA), "State of European Data Centres 2025," Pb7 Research, May 2025. [Online].
Available: https://www.eudca.org/state-of-european-data-centres-2025

[3] Accenture, Analysis on TWh electricity demand estimates by country from multiple sources, 2025–2026.

[4] It should be noted that the injection of energy requires additional compliancy to the Network Code on Requirements for Generators (NC RfG).

# 1 European Data Centres: Growth, Scale, and Transformation

Europe holds a strong and growing position in global data centre deployment. Driven by the expansion of cloud computing, the rise of AI, and the increasing digitalisation of economic activity, the European data centre ecosystem is not only growing rapidly but also undergoing a structural transformation with direct and growing implications for the electricity transmission system. What were once relatively small to medium-sized facilities have evolved into large-scale, operating industrial assets whose electricity demand adds to that of traditional heavy industry, all while increasing at an unprecedented pace.

## 1.1 Expected Growth of Data Centres in Europe

Current estimates indicate that there are over 10,500 data centres facilities across Europe with at least 50 kW of IT power, with capacity totalling approximately 12.7 GW in IT power supply, of which around 9.9 GW is in the EU27.[5, 6] This considerable amount of power is set to increase rapidly and will make up a substantial share of electrical energy consumption growth in the coming years. Despite a wide range of future forecasts, the consensus is that data centres will drive massive growth in electricity demand in the next five to ten years.

There are major differences in the speed of deployment across European countries, as shown in Figure 1. The total electricity demand from data centres for the region is expected to increase by over 50 % between 2025 and 2030,[4] with all major economies witnessing rapid growth. Developments such as the EU Cloud and AI Development Act, expected in the first half of 2026, aim to triple EU data centre capacity over the next five to seven years and could further accelerate these trends.[7]

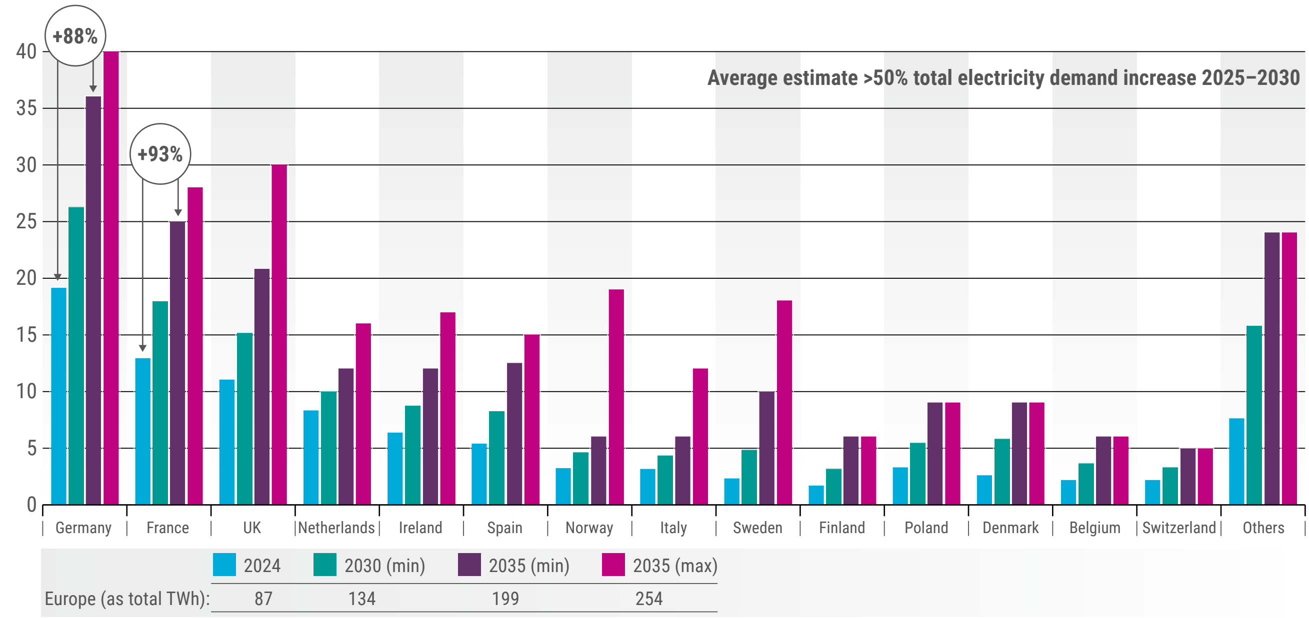


Figure 1: Estimated total data centre electricity demand, by key country (TWh per year)[8]

---

[5] Around 9.9 GW of which is in EU27 countries.

[6] European Data Centre Association (EUDCA), "State of European Data Centres 2025," Pb7 Research, May 2025. [Online]. Available: https://www.eudca.org/state-of-european-data-centres-2025

[7] European Parliament, "AI and the energy sector," Brief EPRS_BRI(2025)775859, Jul. 2025. [Online]. Available: https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2025)775859

[8] The high ranges in the figure are affected by forecasting uncertainties.

On the other hand, the growth of AI and advanced data capabilities also presents significant opportunities for the energy sector. The European Commission's Strategic Roadmap for Digitalisation and AI in the Energy Sector, also due in the first half of 2026, aims to harness these capabilities for electricity grid optimisation, demand-side flexibility, and renewable integration, recognising the power of AI and data not only as a source of demand but also as enablers of a smarter, more resilient energy system.

## 1.2 Data Centre Types

While "data centre" is commonly used as a single term, the sector encompasses fundamentally different types of facilities, whose ownership structures, commercial incentives, and operational constraints lead to very different grid behaviours. Understanding these distinctions is essential for assessing flexibility potential, defining connection requirements, and designing effective policy. Data centres can be classified into three principal categories based on their ownership and business model:

- **Colocation and service provider data centres:** These facilities lease space, power, and connectivity to customers who own and operate their IT equipment. Revenue for the colocation provider derives from leasing physical capacity rather than computing services. The critical feature from a grid perspective is the split-responsibility model: the colocation provider controls the facilities but not the workloads, creating a structural barrier to flexibility. Large "scale colocation" facilities exceeding 50 MW are increasingly leased to hyperscale tenants, blurring the boundary between data centres categories.
- **Hyperscale data centres:** These are massive facilities frequently exceeding 50 MW of IT power. These data centres are operated by major technology companies like cloud service providers, as well as large software-as-a-service (SaaS) and internet platforms companies. They use scalable, highly efficient infrastructure to support cloud services, web hosting and, increasingly, AI services.
- **Enterprise data centres:** These facilities are owned and operated by a single organisation for its own use – typically banks, government agencies, or enterprises with sensitive workloads. These data centres are cost generators rather than profit drivers, and their workloads are tightly tied to internal business processes.

When considering only data centres larger than 50 kW, the current installed IT power supply is distributed across different types, as shown in Figure 2. Colocation facilities account for nearly half of this capacity (6.8 GW), while enterprise and hyperscale segments each contribute roughly a quarter.

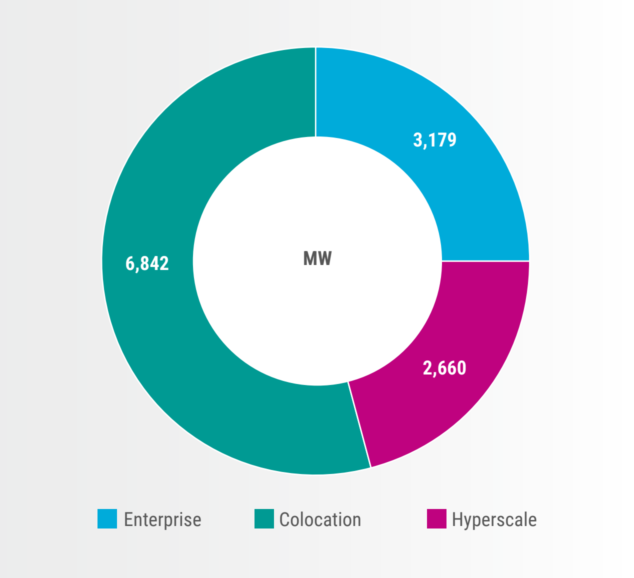
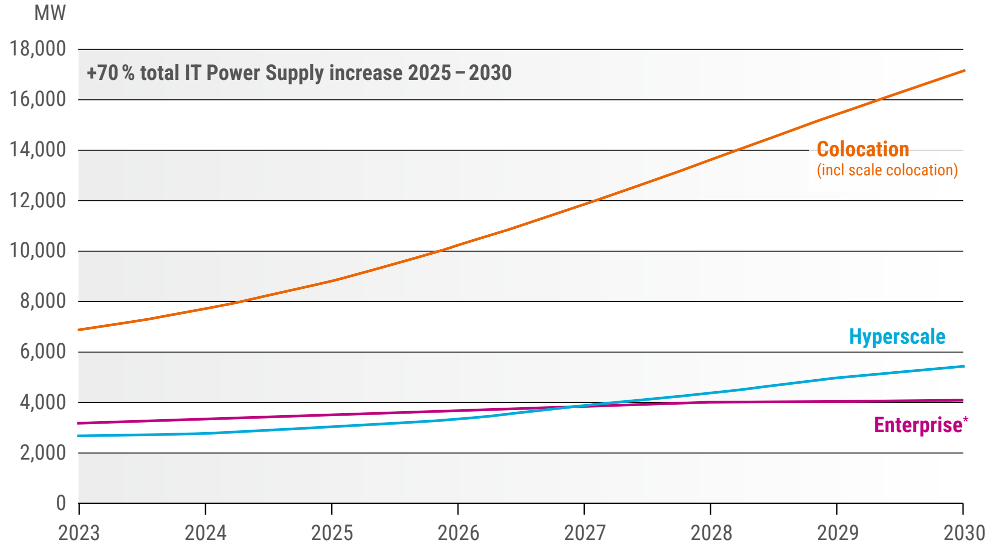


Figure 2: IT power supply (MW) in Europe by type – distribution across enterprise, colocation, and hyperscale segments

## 1.3 Structural Shift

The growth trend shown in Figure 1 will not be evenly distributed among data centre types. Instead, growth will be highly concentrated in colocation data centres, driven mainly by "scale colocation".



> Today, ca **30%** IT Power Suppy (MW) for "**colocation**" is categorised as "**scale colocation**". This category is expected to grow at 30 % yearly rates while retail & wholesale at a yearly rate of 8 %.

> Towards 2030, this trend will result in an estimated **60 –70% of colocation leased to hyperscalers**.

> The growth of **scale colocation** and **hyperscale** will lead not only to an increase in the total IT Power Supply, but also to a change in the profile of newer data centers.

Notes: * Enterprise datacenter IT load estimated and with limited longer-term forecasts available

Sources: Accenture analysis on EUDCA/Pb7, McKinsey, Synergy Research, S & P Global Market Intelligence 451 Research

Figure 3: IT power supply (MW) growth trends in Europe by data centre ownership

As shown in Figure 3, enterprise data centre capacity is largely stagnating in Europe, while colocation facilities (including scale colocation) are expected to drive more than 70 % of the European data centre IT load increase over the next five years, with hyperscale-leased capacity accounting for more than 60 % of the cumulated colocation IT load. This means that the growth is not simply more of the same; it will bring fundamental changes to the European data centre landscape.

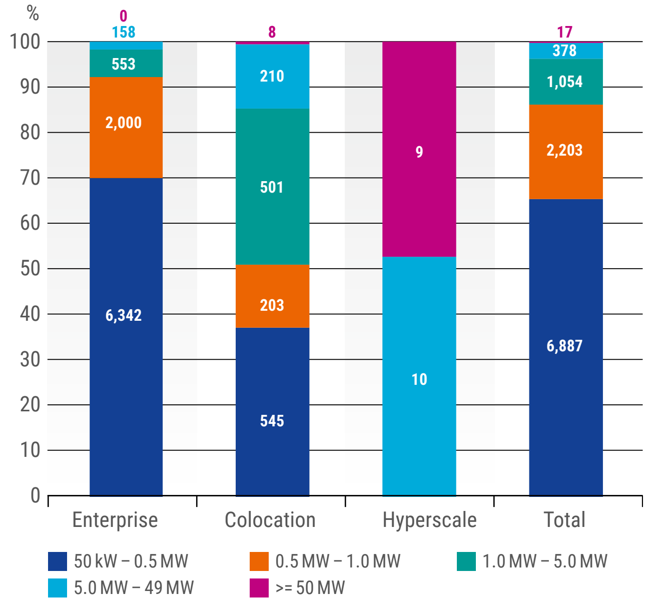


Figure 4: Number and distribution of European data centres by type and facility size

This structural shift matters for the grid because each data centre type operates at very different scales. Figure 4 illustrates the number and size distribution of European data centres by type. While enterprise facilities typically range from 1 to 5 MW, colocation and hyperscale campuses can reach tens or even hundreds of megawatts per site. These differences in individual facility size have direct consequences for the transmission system, with large-scale data centre facilities concentrated at a few points, exacerbating local transmission capacity constraints and also placing stress on the dynamic behaviour of grids.

### Key Messages – Chapter 1

- European data centre demand is growing rapidly, driven mainly by cloud computing, AI, and digitalisation.
- Growth is not uniform. It is increasingly concentrated in colocation and hyperscale facilities, while enterprise data centres are largely stagnating.
- This structural shift matters for the power system because newer facilities are much larger, often tens or hundreds of megawatts, making them comparable to major industrial loads.

# 2 Data Centre Characteristics and Behaviour

The previous chapter showed that European data centre demand is growing fast and shifting structurally towards larger colocation and hyperscale. Yet these figures describe how much electricity data centres will consume, not how they consume it. For TSOs and grid planners, this distinction matters: a 100 MW data centre does not behave like a 100 MW steel mill or chemical plant. What makes data centres electrically distinctive is a combination of the IT workloads they run and the power architecture to supply the cooling and IT loads in increasingly dense server environments. This chapter examines both dimensions and their grid impact.

## 2.1 Power Architecture

From the grid perspective, a data centre's electrical behaviour is shaped by how it supplies two fundamentally different types of loads: IT equipment and cooling systems.[9]

The IT load consists of servers, storage, and networking equipment. This equipment is financially valuable and often runs mission-critical services where even millisecond-level power interruptions can cause data loss or service failures. To protect against this, the entire IT load is supplied through UPS systems, power electronic devices that continuously condition the incoming electricity and can instantaneously switch to battery backup if the grid supply is disturbed. Centralised UPS systems are widely dominant, even in new facilities,[10] making the UPS the primary interface between the data centre's IT load and the grid. Nevertheless, novel alternative architectures are gaining traction, especially in larger hyperscale facilities. Some hyperscalers are moving towards UPS-less architectures that rely entirely on facility redundancy rather than local UPS and backup generation.[11] Alternatively, data centres are utilising rack-level UPS systems, though these decentralised models require advanced control and coordination to provide the same grid interface capabilities as their centralised counterparts.[12, 13] While these developments are not yet common, they are overrepresented in the largest facilities, creating challenges for the transmission system.

The cooling load removes the heat generated by the IT equipment. Modern AI-oriented server racks can exceed 50 kW each, and cooling systems can account for 30 % or more of a facility's total electricity consumption. Cooling and other support loads are treated as overhead, and data centre efficiency is commonly measured by the ratio of total facility power to IT power, known as power usage effectiveness (PUE). Using this metric, the average European data centre currently records a PUE of around 1.5.[14] Emerging cooling technologies, including liquid and immersion cooling, are reducing this overhead, with advanced designs achieving PUE values below 1.1 and leading hyperscalers reporting values below 1.05 in their most efficient facilities.[15]

---

[9] S. Chalise et al., "Data centre energy systems: Current technology and future direction," 2015 IEEE Power & Energy Society General Meeting, Denver, CO, 2015; Uptime Institute, "Tier Classification System," [Online]. Available: https://uptimeinstitute.com/tier-certification

[10] "Data Center UPS Market Size, Share & Trends Analysis Report," Grand View Research, 2025. Centralised segment held 65.8 % market share in 2024. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/data-center-ups-market

[11] "Infinite Scale: The Architecture Behind the Azure AI Superfactory," The Official Microsoft Blog, Nov. 12, 2025. Microsoft's Fairwater 2 data center in Atlanta forgoes on-site generation, UPS systems, and dual-corded distribution, relying on highly available grid power. [Online]. Available: https://blogs.microsoft.com/blog/2025/11/12/infinite-scale-the-architecture-behind-the-azure-ai-superfactory/

[12] IEA, "Energy and AI," International Energy Agency, 2025. [Online]. Available: https://www.iea.org/reports/energy-and-ai

[13] R. Vaidhynathan et al., "Enhancing Data Center Low-Voltage Ride-Through," arXiv preprint, arXiv:2510.03867v1, Oct. 2025. [Online]. Available: https://arxiv.org/abs/2510.03867

[14] Uptime Institute Global Data Center Survey 2025: regional view

[15] ICIS, "Europe Data Centre Power Demand." [Online]. Available: https://www.icis.com/explore/resources/data-centres-hungry-for-power/.

Unlike the IT load, cooling systems have an inherent buffer: the building, equipment, air, and coolant circuits possess thermal mass and can sustain cooling temporarily, so brief power interruptions may not immediately threaten operations. Although the high heat density of modern computing limits the practical use of this buffer, it can be increased through additional thermal storage, making the cooling load a potential source of flexibility.

Both loads are protected by layers of backup infrastructure whose extent depends on the facility's required level of reliability. The industry standard for classifying this redundancy is the Uptime Institute's Tier System (Tier I through Tier IV). At its simplest (Tier I), a data centre has a single power path with no redundant components. As the tier level increases, additional power paths, backup generators, and UPS systems are added, culminating in Tier IV, where fully independent, simultaneously maintainable power paths ensure that no single equipment failure can interrupt the IT load. At each stage, switchgear controls, isolates, and protects the electrical circuits, and where multiple sources are available, automatic transfer switches (ATS) handle the transition between them when a failure or maintenance event requires a changeover. Power is ultimately delivered to the racks through power distribution units (PDUs), which step voltage down to the level required by individual servers. The resulting architecture is shown in Figure 5.

Taken together, the workload characteristics and power architecture described in this chapter define how a data centre appears to the grid at its connection point. The IT load is software-driven, capable of multi-megawatt swings within milliseconds as workloads transition between computational phases. The UPS system acts as interface between the IT load and the grid and is designed to prioritise uninterrupted power to the servers and equipment preservation, if necessary, by disconnecting from the grid entirely in response to minor disturbances. The cooling load adds significant demand that varies with server density and ambient conditions. Traditionally, cooling systems have been directly connected to the grid, often with diesel rotary uninterruptible power supply (DRUPS) systems as backup, and their motor-driven compressors have provided some inherent rotational inertia, with connection and disconnection characteristics broadly similar to those of other large loads. However, the increasing adoption of power-electronic UPS systems for cooling loads, together with variable frequency drives and other inverter-based cooling technologies, fundamentally changes this picture. These technologies decouple the mechanical load from grid frequency and thereby eliminate the inherent inertia contribution. Taken together, these characteristics differ fundamentally from those of traditional industrial loads and carry direct consequences for the grid connection.

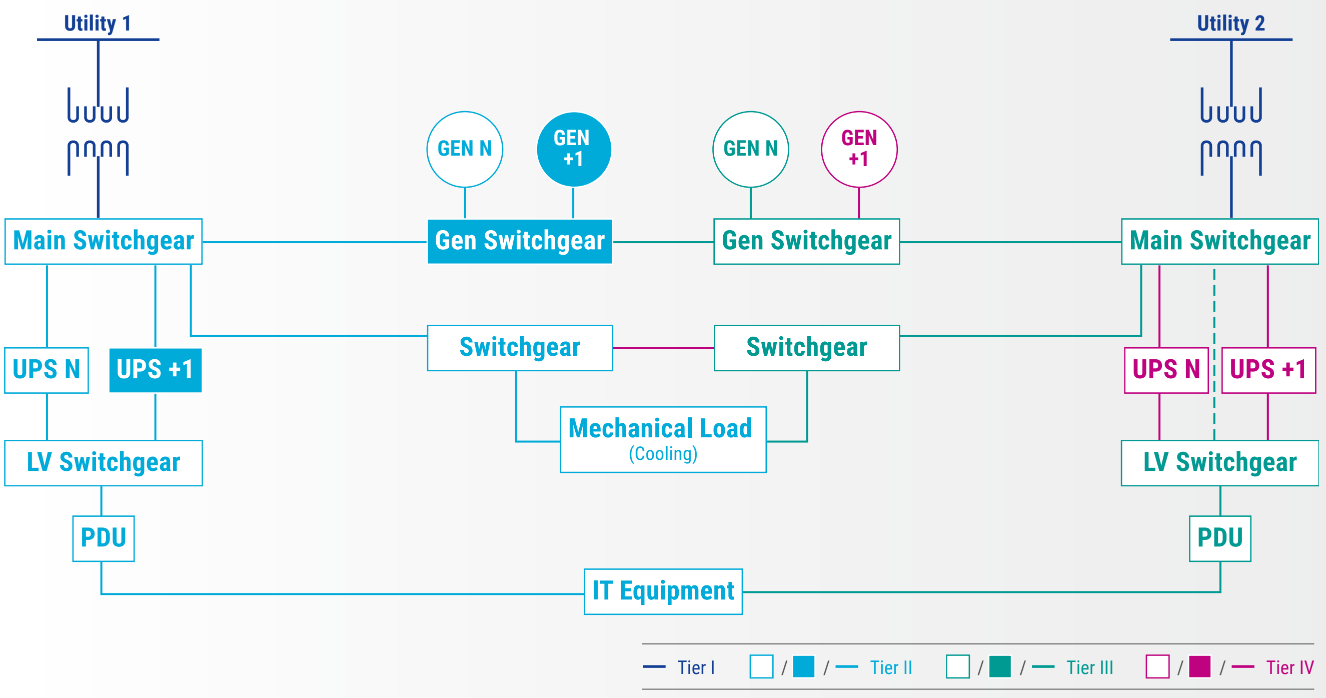


Figure 5: Illustrative power architectures by tier (redundancy level)

## 2.2 IT Workloads

The nature of the IT workload running inside a data centre is a primary determinant of its load profile, infrastructure requirements, and overall grid behaviour. Over the past two decades, the workload landscape has been fundamentally reshaped by two developments. The first is the rise of cloud computing, which shifted IT services from dedicated on-premises infrastructure to shared, virtualised platforms operated at scale, giving rise to workload categories such as web services, SaaS platforms, and content delivery networks that now account for a large share of global data centre energy consumption. The second is the emergence of AI as a major and rapidly growing category of computing demand, introducing workloads whose electrical characteristics, including extreme power density, rapid load fluctuations, and sustained high utilisation, differ from anything the data centre industry or the power system has previously encountered.

Together, these trends have diversified the types of computing occurring inside data centres and, consequently, the electrical behaviour that the grid sees at the connection point. The following paragraphs describe the principal workload categories found in modern data centres, focusing on what each workload is, what infrastructure it requires, and how it behaves electrically.

- **AI Model Training:** Runs continuously for weeks on large clusters of graphical processing units (GPUs) and tensor processing units (TPUs). Load profile shows sustained high consumption with rapid fluctuations (30–60 % within milliseconds) as clusters transition between computational phases. AI training does not require proximity to end users, enabling more location options.
- **AI Inference:** Executes trained models to produce outputs in response to user or system requests. Inference workloads are demand-driven, with aggregate demand following strong diurnal cycles tied to user activity. Latency is a primary concern, driving deployment towards metro-adjacent hyperscale campuses and urban colocation facilities rather than the remote sites suitable for training. Inference is projected to become the dominant AI workload by 2030.[16]
- **Web Services, SaaS, and Content Delivery.** This category encompasses the workloads that deliver web applications, SaaS platforms, social media, video streaming, and content delivery to end users. These are closely tied to the rise of cloud computing. The load profile is strongly diurnal and traffic-driven. The hardware is predominantly CPU-centric, but the sheer scale of deployment means these workloads have significant power consumption.
- **Big Data Analytics.** Also called big data workloads, encompasses periodic business operations (billing, reporting, analytics), data pipelines, ETL processes, and scheduled data migrations. These workloads share a common characteristic: they might be massive but operate against a completion deadline rather than a real-time latency target.
- **Backup and Disaster Recovery:** Backup and disaster recovery workloads maintain replicated copies of data and system configurations at secondary locations. Under normal conditions, they run at low utilisation, but the facility must be sized for full production loads during failover, meaning recovery sites can transition rapidly from minimal to maximum power consumption.
- **Transactional and Real-Time Workloads:** Transactional workloads process structured database operations on a continuous basis with strict consistency guarantees, while real-time workloads must complete processing within hard time constraints measured in microseconds.

A data centre may host a mix of these workload types, and the specific composition determines its overall load profile and tolerance for interruption. A facility predominantly running AI training and batch workloads will present a fundamentally different electrical profile than one serving transactional banking systems, even at a similar total power consumption. Understanding the workload composition is therefore essential for evaluating how a data centre interacts with the power system, a topic developed in Chapter 5.[17]

### Key Messages – Chapter 2

- Data centres do not behave like traditional industrial consumers – their consumption profile is largely software-defined.
- The IT load is highly sensitive and protected by UPS systems, which prioritise uninterrupted operation and can disconnect rapidly during disturbances.
- Cooling loads consume a significant share of power and have thermal inertia, representing a possible source of flexibility. Thermal storage, which is generally low-cost, could enhance potential and reduce constraints.
- Workload type strongly affects grid behaviour. AI training, inference, SaaS, batch processing, and real-time workloads all have different load shapes, flexibility levels, and infrastructure needs.

---

[16] McKinsey & Company, "The future of AI workloads," McKinsey & Company, Feb. 24, 2026. [Online]. Available: https://www.mckinsey.com/featured-insights/week-in-charts/the-future-of-ai-workloads.

[17] Workload classification and flexibility assessment based on: IEA, "Energy and AI," International Energy Agency, 2025; EPRI, "Powering Intelligence: Analysing Artificial Intelligence and Data Center Energy Consumption," 2024; Google Cloud, "Mitigating Power and Thermal Fluctuations in ML Workloads," 2024.

# 3 Risks and Solutions to Grid Security and Stability

The reliable operation of the electricity transmission power systems depends on all connected assets adhering to clearly defined technical requirements at the point of connection. In Europe, these requirements are established through the connection network codes, developed under European legislation and implemented by Member States at the national level. Three distinct EU network codes established harmonised technical requirements for grid connection: the Requirements for Generators,[18] the Demand Connection Code,[19] and the HVDC Network Code.[20]

Historically, large industrial loads were treated as predictable and relatively steady, with limited influence on system stability and less stringent requirements than generators of similar size. The existing EU regulatory framework for demand connections reflects these assumptions; however, data centres challenge them in fundamental ways.

Their scale is now system-relevant. Individual campuses can reach several hundred megawatts, and clusters in concentrated regions can exceed one gigawatt, comparable to the dimensions of European frequency reserves.[21], Their electrical demand is software-driven, governed by workload scheduling and internal control logic rather than mechanical inertia, enabling ramp rates and abrupt load steps faster than many conventional control resources can respond.[22, 23] And their response to grid disturbances is fundamentally different: UPS systems can disconnect hundreds of megawatts instantaneously when voltage or frequency thresholds are breached, then remain on battery for minutes before reconnecting at an unpredictable moment, potentially creating a second disturbance event. These behaviours introduce risks both during normal operation and during grid disturbances.

---

[18] European Commission, "Commission Regulation (EU) 2016/631 of 14 April 2016 establishing a Network Code on Requirements for Generators (NC RfG) (Text with EEA relevance)," Official Journal of the European Union, vol. L 112, pp. 1–68, Apr. 27, 2016. [Online]. Available: http://data.europa.eu/eli/reg/2016/631/oj. Accessed: Mar. 4, 2026.

[19] European Commission, "Commission Regulation (EU) 2016/1388 of 17 August 2016 establishing a Network Code on Demand Connection (NC DCC) (Text with EEA relevance)," Official Journal of the European Union, vol. L 223, pp. 10–54, Aug. 18, 2016. [Online]. Available: http://data.europa.eu/eli/reg/2016/1388/oj. Accessed: Mar. 4, 2026.

[20] European Commission, "Commission Regulation (EU) 2016/1447 of 26 August 2016 establishing a Network Code on Requirements for Grid Connection of High Voltage Direct Current Systems and Direct Current-Connected Power Park Modules (NC HVDC) (Text with EEA relevance)," Official Journal of the European Union, vol. L 241, pp. 1–65, Sept. 8, 2016. [Online]. Available: http://data.europa.eu/eli/reg/2016/1447/oj. Accessed: Mar. 4, 2026

[21] NERC, "Characteristics and Risks of Emerging Large Loads," 2025. [Online].
Available: https://www.nerc.com/globalassets/who-we-are/standing-committees/rstc/whitepaper-characteristics-and-risks-of-emerging-large-loads.pdf

[22] Google Cloud, "Mitigating Power and Thermal Fluctuations in ML Infrastructure," Feb. 2025. [Online].
Available: https://cloud.google.com/blog/topics/systems/mitigating-power-and-thermal-fluctuations-in-ml-infrastructure

[23] Kez, D. A., & Foley, A. M. (2025). "Instability Risks from Programmable AI Load Ramping in Low-Inertia Grids," SSRN. [Online].
Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5370875

## 3.1 Risks During Normal Operation

Even when the grid is operating normally, data centres can act as a source of disturbance. Because their power consumption is driven by internal software logic rather than external conditions, the resulting fluctuations are invisible to conventional monitoring and difficult to anticipate in operational planning. Two distinct mechanisms are relevant:

- **Load-driven disturbances:** Rapid power fluctuations from AI workloads, where consumption can swing by tens of megawatts within milliseconds, appear to the power system as frequent, unpredictable load-shedding or load-application events. As the aggregate data centre share of system demand grows, these effects become increasingly system-relevant, particularly in regions with a high concentration of data centres. These load-driven fluctuations can create voltage violations and frequency imbalances, among other problems.
- **Load-induced forced oscillations:** Persistent, periodic load patterns from AI training can excite electromechanical modes of the power system at frequencies that coincide with inter-area oscillation modes, creating a risk of resonance amplification that propagates across wide areas. Unlike transient disturbances that decay naturally, these oscillations are continuously sustained. Such events have already been observed in practice, including sub-synchronous oscillations not detectable by standard SCADA.[24, 25]

## 3.2 Risks During Grid Disturbances

When a grid fault occurs, data centres respond very differently from traditional loads:

- **Fault ride-through (FRT) behaviour:** Power electronics backup supply systems may transition to battery mode at the first sign of a voltage dip, instantaneously removing the entire IT load from the grid. In double-conversion mode, their constant-power characteristic can initially worsen the dip by drawing increased current as voltage falls, before ultimately disconnecting completely. DRUPS systems, while providing some natural inertia, are more sensitive to rapid frequency changes and may disconnect even earlier.[26] The issue is further compounded by the high concentration of data centres in the same area, which can exacerbate a local voltage incident into a system-wide frequency disturbance.
- **Reconnection behaviour:** After fault clearance, reconnection practices are not standardised. Power electronic UPS systems may return load to the grid within seconds, but the sudden demand shock can depress frequency back towards disconnection thresholds, risking repeated disconnect–reconnect cycles, a phenomenon known as flapping.[27] Conversely, prolonged absence from the grid may cause a sustained generation surplus that operators must manage. This is especially common in facilities with DRUPS systems and/or backup generation, which can take minutes to hours to reconnect and may require manual intervention.[26, 28] A staged reconnection approach, restoring load gradually across multiple UPS units at staggered intervals, can mitigate both extremes.[28]

Currently, there are no standards requiring grid-supportive behaviour during faults for large demand-connected facilities in most European jurisdictions. Multiple incidents across different grid systems have already demonstrated load losses ranging from several hundred megawatts to over one gigawatt following routine transmission faults[27].

---

[24] S. Dasgupta, C. Mishra, L. Vanfretti, et al., "Understanding the Inception of 14.7 Hz Oscillations Emerging from a Data Center," ResearchGate, 2025.

[25] CIGRE US National Committee, "NGN Webinar: Large Loads and Their Impact on the Grid," 2025.

[26] Eaton, "Data Center: A Good Grid Citizen," Eaton Whitepaper, 2025. [Online]. Available: https://www.eaton.com/content/dam/eaton/markets/data-center/the-impact-of-ai/whitepapers/eaton-whitepaper-technical-data-center-as-a-good-grid-citizen-en-us.pdf

[27] NERC, "Incident Review: Considering Simultaneous Voltage-Sensitive Load Reductions," Jan. 2025. [Online]. Available: https://www.nerc.com/globalassets/our-work/reports/event-reports/incident_review_large_load_loss.pdf

[28] Jimenez-Ruiz, A. (2025). Data Center Model for Transient Stability Analysis of Power Systems.

## 3.3 Connection Requirements as a Response

To address these emerging risks, connection codes are being updated where the traditional assumption of passive, predictable demand is no longer valid. Several national grid code updates are being discussed (e.g. in Ireland[29, 30, 31] and Belgium[32]), in parallel with EU-wide discussions on harmonisation also supported by ENTSO-E's recommendations.[33]

Some of the key areas under discussion include:

- **FRT:** Requiring data centres to remain connected during voltage dips rather than disconnecting at shallow thresholds, preventing cascading load losses during grid faults.
- **Rate of Change of Frequency (RoCoF) Withstand Capability:** Ensuring internal UPS and protection logic can tolerate rapid frequency changes without tripping.
- **Ramp Rate Control:** Limiting the speed of load changes to prevent software-driven steps from outpacing grid frequency and voltage control systems.
- **Voltage Control and Reactive Power:** Requiring active voltage regulation at the connection point to prevent localised voltage issues from concentrated active power consumption.
- **Oscillation Damping:** Ensuring power electronics are tuned to suppress rather than excite natural grid frequencies, addressing the forced oscillation risks described above.
- **Post-fault Active Power Recovery:** Requiring a controlled active power consumption recovery at the connection point when the network voltage resumes after a fault in the system.

The ENTSO-E position paper on national connection requirements recommends taking actions to update some of these requirements to ensure system stability. However, while the risks from data centre dynamic behaviour might be solved by updating technical connection requirements, the electricity grid is also facing challenges to system adequacy and network capacity posed by growing demand, not only from data centres, but also from other sectors such as the electrification of transport, industry, and heating.

### Key Messages – Chapter 3

- Data centres create new grid risks during normal operation because their fast, software-driven load swings can cause voltage and frequency issues.
- During faults, UPS systems may disconnect large loads almost instantly, which can worsen grid events and create subsequent reconnection shocks.
- Updated or new connection requirements, such as FRT, RoCoF withstand, ramp-rate limits, voltage control, and oscillation damping, are essential to ensure system stability.

---

[29] EirGrid, EirGrid Grid Code Version 13, Dublin, Ireland: EirGrid plc, Jan. 30, 2024. [Online]. Available: https://www.eirgrid.ie/grid/grid-codes-and-compliance/grid-code

[30] EirGrid and SONI, MPID345 Demand Facilities FRT Recommendation Paper, Apr. 1, 2026. [Online]. Available: https://cms.eirgrid.ie/sites/default/files/publications/MPID345-Demand-Facilities-FRT-Recommendation-Paper-01042026.pdf

[31] EirGrid and ESB Networks, "MPID322 Recommendation Paper: Update of Definitions for RfG and Non-RfG, HVDC and Non-HVDC, DCC and Non-DCC Units," Nov. 20, 2024. [Online]. Available: https://cms.eirgrid.ie/sites/default/files/publications/MPID322-Recommendation-Paper.pdf

[32] Elia Transmission Belgium, "Additional Technical Requirements for Grid-Forming Battery Energy Storage System and Large Inverter-Based Loads Connected to Elia Transmission System," Public consultation document, Dec. 2025.

[33] ENTSO-E, "Position on the need for national connection requirements to ensure EU power system stability," Position Paper, Dec. 2025. [Online]. Available: https://eepublicdownloads.blob.core.windows.net/public-cdn-container/clean-documents/Publications/Position%20papers%20and%20reports/2025/251211_SDC_Position_paper_on_CNC_2.0.pdf

# 4 Challenges for Grid Planning

After addressing the technical connection requirements in the previous chapter, fundamental grid planning challenges remain. The grid must have adequate generation resources to supply power and sufficient transmission capacity to transport it to the point of consumption. Both can come under pressure due to growing demand, including, but not limited to, data centre load.

System adequacy challenges arise when available generation capacity is insufficient to reliably meet peak demand, maintain reserve margins, and accommodate incremental demand from new large consumers, while also meeting climate objectives that require generation to support decarbonisation goals. In most European regions, data centres do not yet pose a generation adequacy challenge. However, their rapid projected growth is likely to make this a near-term concern, as already evidenced in Ireland and several regions of the United States.[34, 35]

Grid constraint challenges also occur when transmission lines or substation transformers cannot safely carry the additional power flows needed to serve new demand. Planning and construction of the additional infrastructure required to address both types of constraints typically takes several years, often longer than the timelines on which data centres are developed and deployed. The result is long waiting times for grid connections, as new customers must wait until reinforcements are completed before they can receive firm capacity. Cooperation with grid operators through flexible connection agreements and better information exchange can provide a temporary solution awaiting the grid reinforcement.

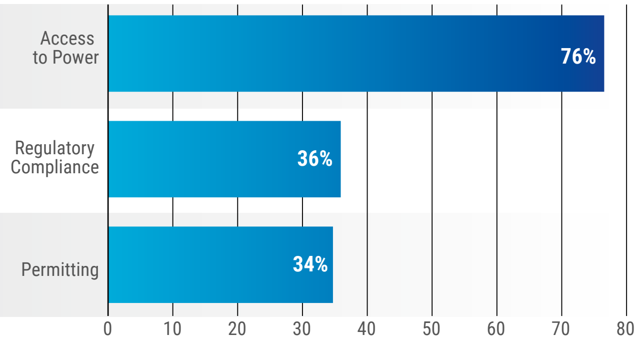


Figure 6: Biggest challenges from data centres perspective in the next three years[36]

As shown in Figure 6, access to electricity is considered the biggest challenge by almost 80 % of European data centre operators, far ahead of any other concern.[34] This reflects the fact that connection lead times in Europe can range from a few years to more than a decade in constrained regions like FLAP-D cities,[37] creating a time-to-power gap that is a central obstacle to the sector's growth[38].

On the system operator side, the task of planning and allocating connection capacity is further complicated by uncertainty over demand projections and how quickly new data centres will use the capacity they have been granted. After construction, data centres may gradually increase installed capacity depending on the type of tenants. Hyperscale data centres may ramp up utilisation relatively quickly, depending on the approved grid connection. While co-located data centres typically go through a lease-up period for multiple customers, and the full ramp-up of the maximum capacity can take multiple years:

- Retail colocation: Up to several years to reach high utilisation due to multiple smaller tenants in two to three years, the aim is to approach a target occupancy (e.g. 80–90 % of capacity sold, with some reserve for growth or redundancy).[39]
- Wholesale colocation: Potentially shorter lease-up period when a large tenant (e.g. hyperscaler) leases multiple MW upfront, the facility may reach e.g. 50–70 % utilisation quickly, followed by full commitment in about one year.

---

[34] EirGrid and SONI, "All-Island Resource Adequacy Assessment 2026–2035," Dublin, Ireland, Dec. 2025. [Online]. Available: https://cms.eirgrid.ie/sites/default/files/publications/AIRAA-2026-2035_Ireland.pdf

[35] North American Electric Reliability Corporation, "Characteristics and Risks of Emerging Large Loads," Atlanta, GA, USA, Jul. 2025. [Online]. Available: https://www.nerc.com/globalassets/who-we-are/standing-committees/rstc/whitepaper-characteristics-and-risks-of-emerging-large-loads.pdf.

[36] EUDCA, "State of European Data Centres 2025," May 2025. Available online: https://www.eudca.org/state-of-european-data-centres-2025

[37] IEA, "Overcoming Energy Constraints Is Key to Delivering on Europe's Data Centre Goals," Jun. 2025. [Online]. Available: https://www.iea.org/commentaries/overcoming-energy-constraints-is-key-to-delivering-on-europe-s-data-centre-goals
FLAP-D: Frankfurt, London, Amsterdam, Paris, and Dublin.

[38] Newmark, "2025 Data Center Site Selection Dynamics in Europe," Jul. 2025; CERRE, "Where to put data centres: A complex problem for Europe,"Nov. 2025.

[39] Umbrex, RCwireless

To align grid planning with this gradual uptake, some connection agreements use phased contractual capacities, where the contracted connection capacity increases in predefined increments over time rather than being made fully available from the start. While this reduces uncertainty considerably, the contractual capacity at any given phase remains a ceiling that may not reflect actual conditions. The installed capacity may lag behind the contractual figure, and actual consumption may, in turn, remain well below installed capacity due to workload variability and market conditions. A facility contracted for 50 MW may operate at 20–30 MW for an extended period, with limited visibility on when, or even whether, the full capacity will be required.

Further information on demand projections, expected ramp-up timelines and actual utilisation, while protecting commercially sensitive data, can support more informed system planning and reduce waiting times.

Beyond this, three complementary approaches can help close the grid planning gap without compromising system security: transparency on available capacity with clear criteria to distinguish between mature and non-mature connection requests, efficiency in capacity allocation, and flexible connection agreements.

## 4.1 Transparency on Available Hosting Capacity

While many site-selection factors (market proximity, tax regimes, connectivity) are outside TSO control, access to power is not. By providing non-binding and aggregated publicly accessible information on where grid capacity is available, what connection timelines look like, and how capacity evolves under different scenarios, TSOs can steer site-selection decisions towards regions where connections are feasible. Supported by clear criteria for connection requests, this might form conditions and signals for decreasing pressure on saturated hubs and improving the utilisation of existing infrastructure.

Effective transparency goes beyond a static snapshot of current capacity. Data centre developers plan on multiyear horizons and forward-looking information based on approved grid development plans is relevant for the scope. Equally important is visibility on hosting capacity under flexibility conditions: presenting both firm and flexible capacity allows developers to assess the benefits of flexible connection arrangements and evaluate trade-offs between operational constraints and earlier access to power.

Multiple TSOs in Europe are implementing such platforms.[40, 41] One such example is Elia's hosting-capacity platform, illustrated in Figure 7, which presents current and projected capacity under different flexibility assumptions, provides a practical example of this approach, noting that multiple TSOs are implementing similar initiatives.[42]

---

[40] ENTSO-E and EU DSO Entity, "ENTSO-E TSO-DSO Joint Progress Report," Dec. 2025. Available: https://eudsoentity.eu/wp-content/uploads/2025/12/ENTSO-E_TSO_DSO_Joint_Progress_Report.pdf

[41] Terna, "Terna presents TE.R.R.A: the digital portal for the efficient planning of Italy's energy infrastructure," 2025. Available: https://www.terna.it/en/media/press-releases/detail/presentation-terra-digital-portal

[42] Elia Group, "Grid Hosting Capacity Map," Elia System Operator, Brussels, 2025. [Online]. Available: https://www.elia.be/en/customers/connection/grid-hosting-capacity

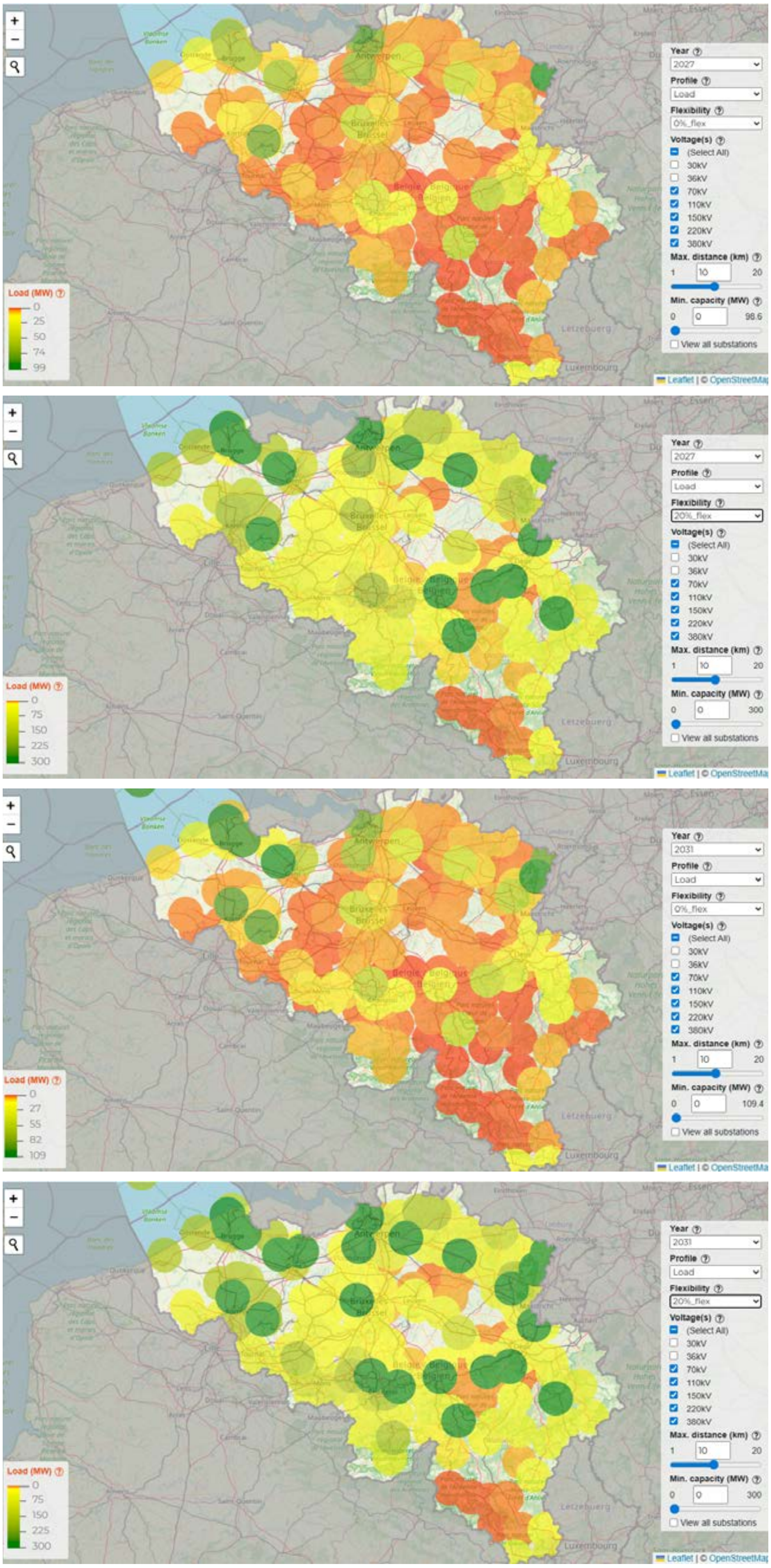


Figure 7: Hosting capacity for new large loads in Belgium (Elia) for the years 2027 and 2031, and 0 % and 20 % flexibility assumptions[43]

To serve this purpose effectively, such platforms should be updated so that users can rely on information that reflects network conditions, approved developments, and changes in queue status or available capacity. At the same time, the aggregated information published on these platforms should be clearly understood as indicative guidance rather than a binding guarantee of connection. Final connection outcomes remain subject to detailed technical assessment, formal application procedures, and the applicable regulatory and contractual framework. Keeping this distinction explicit is important to preserve the value of the platforms as planning tools while avoiding misinterpretation by project developers.

Many large-scale data centre projects are not confined to national boundaries; a consistent pan-European view is therefore needed. The EU Grid Action Plan foresees the development of a joint portal by ENTSO-E and the EU DSO Entity to collect and publish hosting-capacity information across Europe, improving comparability and supporting better-informed investment decisions at a European scale. As noted previously, transparency on capacity must be paired with effective prioritisation rules and queue management to prevent speculation and abuse; otherwise, speculative requests may concentrate in areas with available capacity to secure grid access.

---

[43] Elia Group, "Grid Hosting Capacity Map," Elia System Operator, Brussels, 2025. [Online]. Available: https://www.elia.be/en/customers/connection/grid-hosting-capacity

## 4.2 Efficiency and Fairness in Capacity Allocation

Transparency alone is insufficient if scarce capacity is not allocated to projects that are realistically able to proceed. Current allocation processes in many jurisdictions were designed for traditional industrial loads and did not anticipate the rapid surge of large, power-intensive data centre developments. A persistent challenge is the high volume of speculative or premature capacity requests: developers frequently submit applications for multiple locations simultaneously or retain queue positions beyond any feasible development horizon. While individually rational, this behaviour distorts planning signals, blocks capacity for credible projects, and increases the administrative burden for system operators.

Therefore, regulators internationally are moving away from open, low-barrier queue entry towards more structured allocation processes that reward credible progress. Recent reforms in Romania, the United Kingdom, and the United States, while differing in scope and legal context, point in a similar direction. Several policy design elements appear particularly relevant for European TSOs when considering the allocation of scarce capacity to data centres:

### Financial Guarantees & Deposits

Help deter opportunistic queue occupation and provide compensation when speculative projects consume administrative effort. These must be carefully designed to avoid unjustly favouring projects with higher budgets.

### Project Maturity Milestones & Cost Disclosure

Such as proof of land rights, permitting status, technical design completion, or financing progress, these provide objective filters that can be used to prioritise serious projects.

### Withdrawal Penalties

Discourage strategic queue holding and encourage earlier decision-making, reducing speculative congestion in connection queues.

### Capacity Release Mechanisms

Allow network operators to advance ready projects and avoid capacity being frozen by developments that do not progress.

### Social and System-Value Criteria

When foreseen by the regulatory framework, enables prioritisation of projects that support the grid, aligning queue management with broader system reliability and flexibility goals, and social value.

For TSOs, these tools offer leverage to align allocation with real system needs and maintain trust in the connection process.[44]

---

[44] Ember, "Grids for data centres in Europe," Jun. 2025. [Online]. Available: https://ember-energy.org/app/uploads/2025/06/Grids-for-data-centres-in-Europe.pdf

## 4.3 Flexible Connection Agreements

Where grid reinforcements cannot be completed in time for a firm connection and where regulation allows and TSOs find it feasible to propose such arrangements, voluntary flexible connection agreements may offer a pragmatic alternative that directly addresses the time-to-power gap. Rather than waiting years for full firm capacity, a data centre receives a reduced firm allocation complemented by conditional capacity that can be curtailed when the system is constrained. During those limited periods, facilities can rely on on-site backup generation, battery storage, or reductions in computational or cooling load.

As presented in Figure 8, these arrangements can take several forms of increasing sophistication: static capacity limitations during predefined peak windows, dynamic limitations triggered by actual or forecast grid constraints, and bring your own generation (BYOG) schemes where the data centre directly contracts or builds dispatchable generation to meet part of its demand on site (outside of the system).

Analysis of a US data centre showed that combining flexibility and BYOG could bring full operation three to five years earlier than under traditional connection processes, with curtailment limited to 40 to 70 hours per year, equivalent to more than 99 % grid availability.[45]

**Grid flexibility measures impact on lead time and controllability**

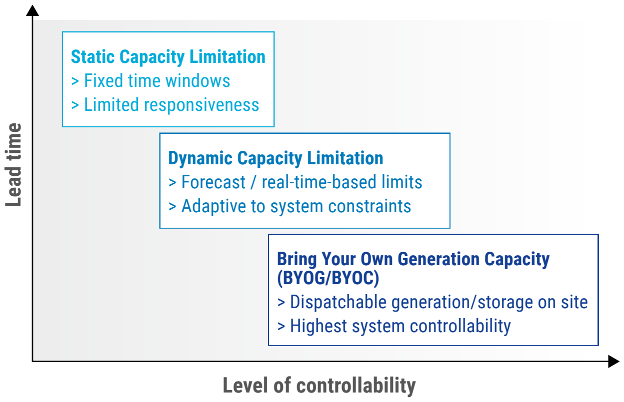


Figure 8: Different levels of sophistication from flexible connection agreements

However, it should be taken into account that on-site backup generation is still predominantly fossil fuel–based, which limits its contribution to generation adequacy under climate objectives; while several climate-friendly alternatives, such as biofuels, may help, they can still cause local problems, such as air particles. In Ireland, recent regulations require data centres to provide not only 100 % of on-site generation capacity but also additional new renewable generation to cover 80 % of its energy needs.[46] Such approaches can enable data centres to contribute to the overall achievement of energy decarbonisation goals, while also making storage-based and other innovative solutions for on-site backup generation increasingly relevant.

Flexible connection agreements could benefit both sides: giving data centre operators faster access to power, while enabling the system to avoid the cost of premature or oversized network reinforcements. When combined with on-site generation and storage, these agreements might transform the data centre from a passive consumer into an active system participant, a theme developed further in the following chapter.

### Key Messages – Chapter 4

- Access to electricity is now one of the sector's biggest bottlenecks, with connection queues ranging from years to over a decade in constrained areas.
- Better information on data centre demand projections, expected ramp-up timelines, and actual utilisation can support more informed system planning and reduce waiting times.
- Better transparency on hosting capacity can steer projects towards feasible locations and reduce pressure on saturated hubs, but it must be combined with improved grid connection capacity allocation criteria to reduce speculative requests.
- Flexible connection agreements may help close the time-to-power gap by allowing earlier operation with curtailment conditions, on-site generation, or storage support.

---

[45] CAMUS Energy, encoord, and Princeton University ZERO Lab, "Flexible Data Centres: A Faster, More Affordable Path to Power," Dec. 2025.

[46] Commission for Regulation of Utilities (CRU), "Large Energy Users Connection Policy Decision Paper," CRU/2025236, Dec. 2025. Available: https://cruie-live-96ca64acab2247eca8a850a7e54b-5b34f62.divio-media.com/documents/CRU2025236_Large_Energy_User_connection_policy_decision_paper.pdf. Accessed: Jan. 5, 2026.

# 5 Opportunities for Grid Operation

The previous chapter showed that flexible connection agreements can accelerate grid access by allowing data centres to accept conditional capacity limits. The technical capabilities required to operate under such agreements, including load modulation, battery management, on-site generation control, and cooling optimisation, do not only serve the purpose of managing constrained connections. Once installed and orchestrated, the same resources can be offered as grid services, creating value for both the power system and the data centre operator. This matters economically because, although energy is not the largest component of total cost in such a capital-intensive sector, it still makes up around 50 % of operating expenditure, making flexibility especially valuable.

This chapter examines what flexibility resources exist within data centres, how meeting grid-safety requirements naturally leads to more advanced control architectures, and how the resulting capabilities map specific electricity market products.

## 5.1 Flexibility Providers Within Data Centres

Data centres contain three categories of assets that can provide flexibility to the grid, corresponding to the IT load domain, the cooling infrastructure, and on-site generation (Figure 9).[47, 48]

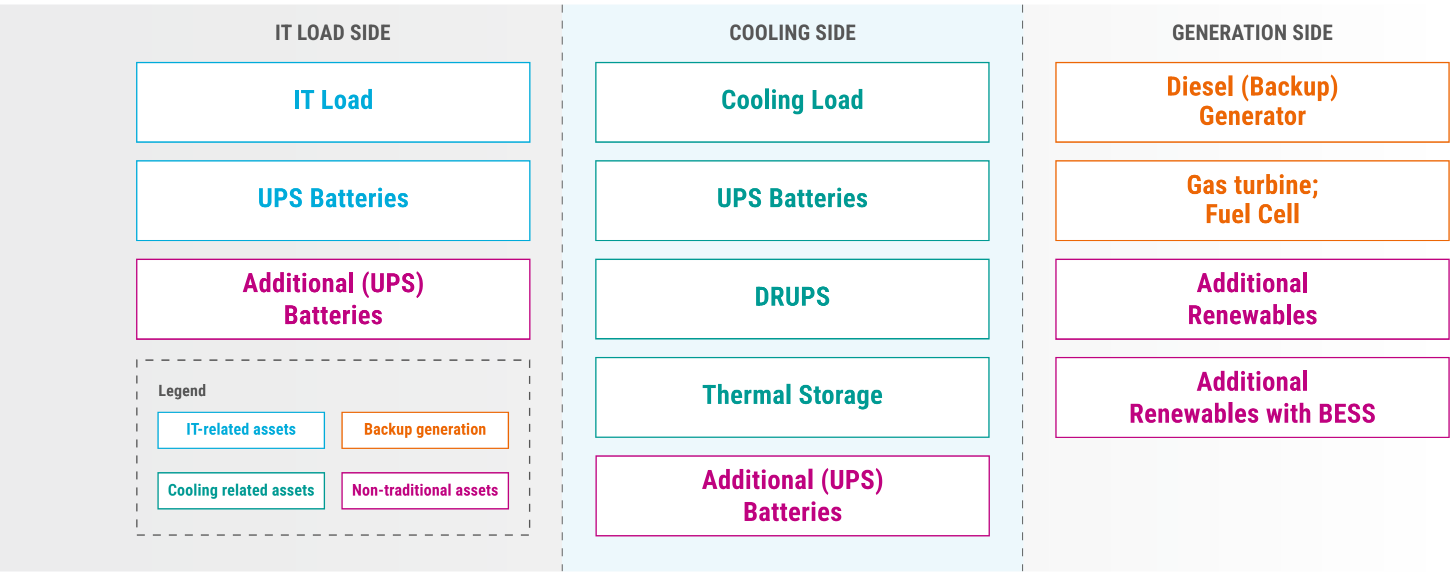
- **IT domain:** UPS power capacity is dimensioned for the maximum possible IT load under full redundancy, but in practice, the actual IT utilisation is almost always lower. The resulting surplus capacity (the backup margin not currently needed for reliability) can temporarily be used for grid services without compromising the facility's protection level. This flexibility is nevertheless limited by battery duration, since UPS systems are generally designed to support the load for only a few minutes while backup generation starts, rather than to provide sustained grid support; therefore, more extensive grid support requires additional battery storage. IT load itself can also be shifted by deferring or rescheduling batch processing, analytics, or AI training jobs to periods that are more favourable for the grid.
- **Cooling domain:** Cooling systems may account for 30 % or more of total facility consumption for a current average data centre. This load has inherent thermal inertia: buildings and coolant circuits retain cold for minutes to hours, allowing cooling load to be temporarily reduced or shifted without immediately affecting IT operations. Where cooling UPS batteries or DRUPS are installed, these provide additional short-duration flexibility. Thermal storage systems, where present, extend the available shifting duration further. Current technological advances are driving the share of cooling in the total load well below 30 % for more modern and advanced data centres, presenting a trend of decreasing flexibility potential.

---

[47] Paananen, J., & Nasr, E. (2021). "Grid-interactive data centres: enabling decarbonization and system stability."

[48] Zhang, Y., Tang, H., Li, H., & Wang, S. (2025). „Unlock the flexibilities of data centres for smart grid services: Optimal dispatch and design of energy storage systems under progressive loading," Elsevier.

- **On-site generation:** Backup diesel generators are present at almost every data centre[49] but are constrained by emissions limits and runtime restrictions. A more promising path is the integration of additional battery capacity directly into the UPS architecture: because the converter stages, protection systems, and grid-connection hardware already exist, only the battery modules need to be added, resulting in significantly lower capital costs than standalone battery storage installations.[50] Beyond incremental storage, data centres can deploy new dispatchable generation or renewable capacity, particularly where flexible connection agreements or BYOG requirements (as discussed in Chapter 4) are already in place.



| IT LOAD SIDE | COOLING SIDE | GENERATION SIDE |
|---|---|---|
| IT Load | Cooling Load | Diesel (Backup) Generator |
| UPS Batteries | UPS Batteries | Gas turbine; Fuel Cell |
| Additional (UPS) Batteries | DRUPS | Additional Renewables |
| | Thermal Storage | Additional Renewables with BESS |
| | Additional (UPS) Batteries | |

Figure 9: Data centre flexibility providers grouped by IT load, cooling systems, and on-site generation and the grid support capabilities

---

[49] J. Bañuelos, R. Patterson, M. Delwiche, B. Barnes, K. D. Lee, C. L. Houston, and J. D. Johnston, "Diesel Generators at Data Centers: Status, Impacts, and Protective Practices," Better Data Center Project, Mar. 2026. [Online]. Available: https://betterdatacenterproject.com/wp-content/uploads/2026/03/Diesel-Generators-at-Data-Centers-Status-Impacts-and-Protective-Practices.pdf

[50] ABB (2022). "PowerExchanger eases the transition to renewables and creates a revenue stream from your UPS."

## 5.2 Flexibility Potential of IT Workloads

The previous section identified IT workload shifting as one of the three categories of flexibility providers within data centres, alongside cooling systems and on-site generation. Of these three, IT workload flexibility deserves closer examination because it is genuinely specific to data centres: no other type of large electricity consumer can reshape its core productive activity in time or across geography in the way that software-defined computing can. Cooling flexibility and battery-based grid services, while valuable, rely on mechanisms available in other large industrial loads. By contrast, IT workload flexibility has few equivalents outside the data centre sector, especially in terms of the speed with which load can be shifted. At the same time, this theoretical capability should not be overstated, as the flexibility that can be activated depends on the workload mix, workload controllability and incentives.

### Workload-Level Flexibility Assessment

The IT workload types described in Chapter 2 differ substantially in their flexibility potential. Table 1 summarises the assessment along two principal dimensions. Time-shift potential captures whether computing demand can be deferred to, or brought forward into, periods that are more favourable for the power system, for example, in response to price signals, renewable availability, or grid stress. Geographical-shift potential captures whether the same computation can be executed at a different facility in a different location, enabling demand to be moved between grid areas in response to local conditions such as congestion or curtailment. Together, these two dimensions define the space within which a workload can be made flexible. The table further characterises each workload by its latency sensitivity (how tolerant it is to delays in execution), its typical load profile (steady, diurnal, bursty, or schedulable), and the business model in which it predominantly occurs. In practice, however, many facilities host mixed workloads, which reduces the share of total site load that can be shifted at a given moment. In addition, for many operators, the value of uninterrupted computing service will exceed the marginal value of grid-service revenues.

| Workload Type | Latency Sensitivity | Time-Shift Potential | Geo-Shift Potential | Load Profile | Dominant Business Model | Overall Flexibility Potential |
|---|---|---|---|---|---|---|
| AI / ML Training | **Low** | **High** | **High** | Sustained, bursty step-loads | Hyperscale | **High**[51] |
| Batch / Data Processing | **Low** | **High** | **Medium–High** | Schedulable, periodic | Hyperscale, Enterprise | **Medium–High** |
| Backup / Disaster Recovery | **Low (until failover)** | **High** | **Low** | Low baseline, surge on failover | Enterprise, Colocation | **Medium** |
| AI Inference | **Medium–High** | **Low–Medium** | **Medium** | Variable, diurnal pattern | Hyperscale, Colocation | **Medium-Low** |
| Web / SaaS / Content Delivery | **Medium–High** | **Low** | **Medium–High** | Diurnal, traffic-driven | Hyperscale, Colocation | **Low** |
| Transactional / Real-Time | **Very High** | **Very Low** | **Very Low** | Steady, mission-critical | Enterprise, Colocation | **Very Low** |

Table 1: Evaluation of flexibility potential of IT workload types

---

[51] Current limitation of GPUs and other components limit flexibility potential of frontier model training, especially for internal workloads where resource redundancy might not be available.

## Business Model Constraints on IT Flexibility

Even where a workload type is inherently flexible, the extent to which that flexibility can be activated depends on the data centre's business model, specifically on who controls IT workloads, how contractual obligations constrain operational decisions, and what commercial incentives exist for flexible operation. Table 2 illustrates how these factors play out for time-shifting potential across the three business models.

| Type | Ownership Model | Operational Control | Contract Model | Obligations / SLAs |
|---|---|---|---|---|
| **Hyperscale** | **Cloud provider** (self-owned or hosted) | **Fully centralised** – single entity manages all infrastructure, workloades, and energy assets | **Usage-based** (IaaS, PaaS, SaaS Model) | **Strict SLAs for uptime,** latency, and performance; often >99.99 % availablity |
| **Colocation** | **Shared** (multitenant facility) | **Shared operations –** facility operator manages infrastructure; tenants manage their racks and applications | **Fixed / capacity based** (Leased Infrastructure Modell) | **Variable SLAs by tentant;** facility-wide redundancy and uptime obligations |
| **Enterprise** | **Self-owned** | **Full internal control** – single company owns and runs the facility and IT systems | **Internal IT governance;** no external service contract | **Internal performance targets,** no external penalties or credits if fault |

IT load Time-shifting potential: ● High ● Moderate ● Low ● Discretionary

Table 2: Comparison of IT load time-shifting potential based on business model characteristics [52]

Enterprise facilities possess high theoretical potential through self-ownership and full operational control, with flexibility governed by internal policies rather than standardised contractual obligations. This discretionary nature makes these facilities difficult to analyse under a general framework, as practical application is often shaped by a high density of critical workloads (such as healthcare, finance, or infrastructure) and a tendency towards conservative operational risk-management, often compounded by specific regulatory requirements.

Furthermore, the typically smaller scale of these self-owned sites can result in reduced technical sophistication and fewer specialised resources compared to hyperscale providers, which often constrains the real-world implementation of complex time-shifting strategies. Colocation environments contain large theoretical volumes of flexible computing and manage significant pools of UPS and backup infrastructure; however, the operator does not control tenant workloads. Meaningful temporal shifting depends on tenant consent, contractual flexibility, and coordinated action, which are currently not common. This structural constraint is further reinforced when energy costs are simply passed through to clients without creating incentives for active load management.

Hyperscale platforms are most structurally aligned with flexibility: their services are abstracted from physical infrastructure and designed to manage demand dynamically at the platform level. Non-latency-critical tasks can be queued, slowed, or relocated between regions. Realistic flexibility, however, varies across the cloud service stack. Infrastructure-as-a-service (IaaS), platform-as-a-service (PaaS), and SaaS form a continuum of increasing provider control: in IaaS, customers retain substantial control over workload execution; in PaaS, the provider manages the platform while the customer mainly controls the application layer; and in SaaS, the provider manages the full service stack. As provider control increases, so does the practical scope to optimise workload placement and timing, although all service models remain constrained by latency, performance, availability, and data residency requirements. [53]

---

[52] Enterprise data centres are not bound by external contracts or SLAs, rather by internal policies and operational priorities. Note that certain workloads may still be subject to regulatory availability requirements (e. g. financial services, healthcare), which can constrain flexibility regardless of internal policy.

[53] Microsoft Azure, "What are IaaS, PaaS, and SaaS?," [Online]. Available: https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-are-iaas-paas-and-saas/; Microsoft, "Shared responsibility in the cloud," [Online]. Available: https://learn.microsoft.com/en-us/azure/security/fundamentals/shared-responsibility

In practice, workloads often combine several cloud service models at once, which makes flexibility difficult to assess at either the customer or facility level. Nevertheless, hyperscale platforms retain important advantages over comparable colocation segments. Customers typically select a cloud region rather than a specific data centre, and providers operate multiple facilities within each region for redundancy and availability. This gives the provider some capability to balance workloads across facilities within the selected region, without additional infrastructure.[54]

A further source of flexibility comes from the dynamic scalability of cloud services. Customers may choose to run delay-tolerant workloads at different times and with different resource volumes when this lowers cost. Major providers increasingly support this through customer-facing carbon-aware tools and frameworks, including emissions dashboards, sustainability guidance, and software tool kits that help clients shift suitable workloads in time or location.[55] Providers also apply similar carbon-aware approaches to some of their own flexible workloads, for example by shifting compute across time and location in response to cleaner energy availability.[56]

This scalability, pricing logic, and advisory tooling further differentiates hyperscale platforms from colocation, where customers lease a specific physical structure, and gives cloud providers more scope to shape customer behaviour. A distinct case, however, concerns frontier AI training for hyperscaler models. Unlike customer-facing cloud services, this internal workload often does not benefit from the same degree of redundancy, as constrained accelerator capacity tends to keep available infrastructure at very high utilisation.[57]

The actual flexibility available to the power system therefore depends not only on the workload mix and per-site capability, but also on the relative installed capacity of each business model within a given region and on the prevailing business cycle. During periods of rapid capacity expansion, securing access to power may override flexibility incentives, while in more stable phases, cost optimisation may strengthen them.

---

[54] Amazon Web Services, "Regions and Zones," [Online]. Available: https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-regions-availability-zones.html

[55] Amazon Web Services, "Viewing your carbon footprint," [Online]. Available: https://docs.aws.amazon.com/awsaccountbilling/latest/aboutv2/what-is-ccft.html; Microsoft, "Connect to the Emissions Impact Dashboard for Azure," [Online]. Available: https://learn.microsoft.com/en-us/power-bi/connect-data/service-connect-to-emissions-impact-dashboard; Green Software Foundation, "Carbon Aware SDK," [Online]. Available: https://carbon-aware-sdk.greensoftware.foundation/

[56] Google, "We now do more computing where there's cleaner energy," [Online]. Available: https://blog.google/company-news/outreach-and-initiatives/sustainability/carbon-aware-computing-location/; Google, "Our data centers now work harder when the sun shines and wind blows," [Online]. Available: https://blog.google/innovation-and-ai/infrastructure-and-cloud/global-network/data-centers-work-harder-sun-shines-wind-blows/

[57] Microsoft, "FY25 Second Quarter Earnings Conference Call," [Online]. Available: https://www.microsoft.com/en-us/investor/events/fy-2025/earnings-fy-2025-q2; Alphabet, "2025 Q4 Earnings Call," [Online]. Available: https://abc.xyz/investor/events/event-details/2026/2025-Q4-Earnings-Call-2026-Dr_C033hS6/default.aspx

## 5.3 From Grid-Safe to Grid-Supporting

The grid-safety requirements discussed in Chapter 3 (ramp rate control, FRT, controlled reconnection) already require sophisticated coordination of multiple internal assets. Implementing effective ramp control, for example, necessitates an energy and load management system capable of orchestrating UPS batteries, cooling systems, and potentially on-site generation in real time. Once this control architecture is in place, the additional effort to extend it towards grid-support services is significantly reduced.[58]

This creates a natural progression. The first step is ensuring grid-safe operation using existing assets. The second is integrating additional battery capacity into the UPS to expand the flexibility envelope. The third is deploying on-site generation and storage with coordinated dispatch, effectively transforming the data centre into a VPP that can manage its grid interaction, provide flexibility services, and participate in electricity markets through a single control framework.

| **Flexivility Source** | **Power peak reduction** | **Energy shift** | **Operational availability** | **Business contraints** |
|---|---|---|---|---|
| **IT UPS battery** | High | Moderate | Moderate | Very low |
| **IT Load Shift** | Very high | Low | Moderate | Very low |
| **Cooling UPS Battery** | High | Moderate | High (seasonal) | High |
| **Cooling Load Shift** | Low | Moderate | Very high (seasonal) | Very high |
| **Thermal storage** | Moderate | High | Very high | Very high |

Table 3: Flexibility source capabilities

Table 3 provides a qualitative assessment of each flexibility provider across four key dimensions: power peak reduction, energy shift, operational availability, and business constraints associated with activation. The assessment illustrates that IT-domain assets offer strong power peak reduction potential but are constrained by business considerations, while cooling-related assets provide greater energy shift capability with smaller peak reduction potential, at the cost of seasonal availability limitations. These capability profiles shape how each provider maps onto specific market products, as discussed in the following section.

Importantly, much of this progression is already being driven by data centre operator needs rather than grid-support incentives. Managing on-site generation for peak reduction and connection-capacity relief, as described in the flexible connection agreements of Chapter 4, requires some of the same control capabilities, while some of the potential new connection requirements might have similar technical demands as faster ancillary services. The technical and organisational prerequisites of a VPP architecture are expected to be more widely present, creating a natural pathway towards grid-supporting operation.

---

[58] CIGRE (2025). "Leveraging Virtual Power Plants to Enhance Reliability and Sustainability in AI-Driven Data Centres."

# 5.4 Market Products and Revenue Opportunities

Once technically enabled, flexibility resources can be offered as tradable services in electricity markets, generating recurring revenue streams that partially offset operational costs.[59] The suitability of data centre flexibility for specific market products depends on the activation speed, sustained duration, and operational availability of each resource (Table 4):

- **Fast frequency response and primary reserve (FCR):** UPS batteries are well suited to these products, as their power-electronic interface can respond within milliseconds and sustain output for several minutes.[60]
- **Secondary and tertiary reserve (aFRR,[61] mFRR[62]):** Achievable through combinations of IT load shifting, cooling modulation, and thermal storage, though sustained delivery over longer periods requires careful coordination with operational constraints.
- **Energy markets (day-ahead and intraday):** Thermal storage and on-site generation are best suited for these longer-duration products, as the limited energy-shift capacity of IT and cooling flexibility constrains their contribution to energy-volume services.
- **Congestion management:** Data centre load modulation, workload shifting, and temporary operation on on-site generation can provide highly targeted, locationally valuable flexibility, particularly in grid-constrained metropolitan clusters where data centres are concentrated.

| Flexibility Provider | Inertia / Fast Frequency Response | FCR / Primary Reserve | aFRR / Secondary Reserve | mFRR / Minute Reserve | Energy Market (Day-ahead / Intraday) |
|---|---|---|---|---|---|
| **Activation time** | **< 1 s** | **< 30 s** | **30 s – 5 min** | **5 – 12.5 min** | **15 min – hours** |
| IT UPS Battery | Suitable | Suitable | Possible | Limited | Not suitable |
| IT Load Shift[63] | – | Not suitable | Possible | Possible | Possible |
| Cooling UPS Battery | Suitable | Suitable | Possible | Limited | Not suitable |
| Cooling Load Shift[64] | – | Suitable | Possible | Marginal | Limited |
| Thermal Storage[63] | – | Limited | Limited | Suitable | Suitable |

Table 4: Suitability of data centre flexibility providers for inertia, electricity market products, and ancillary services[65]

To understand the scale of the opportunity, the System Needs for the Energy Transition study projects that short-duration and ramping flexibility requirements in Europe will roughly double by 2030, reaching an effective need of approximately 15 to 30 GW.[66] One analysis of five European markets (Germany, Ireland, the Netherlands, Norway, and the United Kingdom) indicates that data centres will have a combined technical flexibility potential of around 16.9 GW by 2030, of which up to 3.8 GW could be realistically expected to be available to the grid after accounting for participation willingness and operational constraints.[67] This suggests that data centres can supply a material share of the fast-acting flexibility needed in certain regions. However, current experience indicates that this potential is not being realised, due to limited incentives and operational constraints that may be greater than anticipated.

---

[59] Le Coq, C., Bennato, A., Duma, D., & Lazarchyk, E. (2025). “Flexibility in the energy sector,” CERRE.

[60] The ability of data centres to provide fast frequency services depends strongly on their fault-ride-through capability, as disconnection during grid disturbances would prevent continued service deliver when demanded.

[61] Automatic frequency restoration reserve

[62] Manual frequency restoration reserve

[63] J. Bañuelos, R. Patterson, M. Delwiche, B. Barnes, K. D. Lee, C. L. Houston, and J. D. Johnston, “Diesel Generators at Data Centers: Status, Impacts, and Protective Practices,” Better Data Center Project, Mar. 2026. [Online]. Available: https://betterdatacenterproject.com/wp-content/uploads/2026/03/Diesel-Generators-at-Data-Centers-Status-Impacts-and-Protective-Practices.pdf

[64] Connection to Grid over a UPS assumed.

[65] Participation is subject to the qualification requirements of the respective TSO and the specific product, in accordance with the relevant regulation. This assessment serves solely as a generic and indicative evaluation.

[66] ENTSO-E, “System Needs for the Energy Transition,” 2024. [Online]. Available: https://eepublicdownloads.entsoe.eu/clean-documents/Publications/System_Needs/entso-e_System_Needs_Energy_Transition_v10.pdf

[67] BloombergNEF and Eaton, “Data Centres and Decarbonisation: Unlocking Flexibility in European Power Grids,” 2021.

The actual participation of data centres in flexibility markets depends not only on technical capability but also on market conditions and competition from other flexibility providers. Data centres are particularly competitive in very fast flexibility products, such as FCR, because UPS systems and controllable loads can deliver highly accurate responses within milliseconds.[68] They can also be attractive for congestion management, where their large, concentrated loads provide location-specific flexibility. In these segments, however, batteries are a strong competing alternative, as they can offer a similarly fast and controllable response. Compared with dedicated standalone batteries, data centres may nevertheless benefit from lower additional investment needs, since much of the required electrical infrastructure, including converters, control systems, and backup equipment, is already installed for normal operation.[69, 70] In some cases, the main incremental investment is therefore limited to additional storage capacity.

At the same time, they may be less competitive in longer-duration or low-cost flexibility markets, where industrial demand response or flexible generation can often provide larger energy volumes at lower cost and over longer periods. In these longer-duration segments, their competitiveness is constrained by the need to prioritise service reliability, limit asset degradation, and comply with operational and contractual requirements, because flexibility provision must remain secondary to core digital service delivery.[71, 72] In addition, data centres are often partly shielded from wholesale electricity price signals through power purchase agreements (PPAs), while in colocation facilities, energy costs are commonly passed directly to customers. This can weaken the operator's incentive to actively optimise consumption in response to market prices, further reducing the attractiveness of longer-term flexibility provision beyond ancillary services. As a result, the availability and operational freedom of data centre assets may be lower than those of competing flexibility providers.

Taken together, the VPP concept provides the highest overall value by integrating grid-safe operation, flexibility provision, and core data centre operational objectives into a single coordinated control framework. However, actual participation depends on business model constraints, the predictability and controllability of IT workloads, and the evolution of regulatory frameworks and market access conditions to accommodate these new participants. As these conditions mature, data centres can transition from passive electricity consumers into active system resources that simultaneously improve system resilience and support decarbonisation while generating new revenue streams for their operators.

## Key Messages – Chapter 5

- Data centres already contain several flexibility resources: UPS batteries, cooling systems, thermal storage, controllable IT workloads, and on-site generation.
- IT workload flexibility is especially distinctive because computing tasks can sometimes be swiftly shifted in time or location, unlike most industrial processes. The potential is dependent on the IT workload.
- The same control capabilities needed for grid-safe operation form the foundation for active grid support and market participation.
- Actual flexibility also depends heavily on the data centre business model. Hyperscalers are generally best positioned to provide it, colocation faces tenant-control barriers, and enterprise sites have limited potential.
- Data centres could provide services such as fast frequency response, reserves, congestion management, and some energy market participation, creating both grid value and new revenue streams.

---

[68] Eaton, "Eaton UPS-as-a-Reserve Proven in Data Center Pilot," Nov. 7, 2018. [Online]. Available: https://www.eaton.com/us/en-us/company/news-insights/news-releases/2018/UPSasR-pilot-Norway.html. Accessed: Apr. 5, 2026.

[69] Socomec, "How energy storage boosts ROI for Data Centers," 2026. [Online]. Available: https://emea.socomec.com/en/solutions/business/data-centres/how-energy-storage-boosts-roi-data-centers. Accessed: Apr. 5, 2026.

[70] Saft, "Energy storage for data centers: how to combine resilience and ROI?," Jan. 8, 2026. [Online]. Available: https://saft.com/en/media-resources/our-stories/energy-storage-data-centers-how-combine-resilience-and-roi. Accessed: Apr. 5, 2026.

[71] impactECI, "Data Center Flexibility: Chapter 3, Stakeholder Ecosystem Analysis," Nov. 25, 2025. [Online]. Available: https://impacteci.substack.com/p/data-center-flexibility-chapter-3. Accessed: Apr. 5, 2026.

[72] J. Paananen, "Data centers as a source of fast frequency response and virtual inertia," LinkedIn, 2018. [Online]. Available: https://www.linkedin.com/pulse/data-centers-source-fast-frequency-response-virtual-inertia-paananen. Accessed: Apr. 5, 2026.

# 6 Conclusion

Data centres are no longer a niche demand category that the European power system can absorb without active management. Their scale, geographic concentration, and electrically distinctive behaviour make them systemically relevant, both as a challenge and as a potential asset.

On the challenge side, the characteristics described in this report create a new category of system concern. Software-driven load dynamics, power-electronic interfaces designed to prioritise IT resilience over grid support, and behaviours that can be invisible to conventional monitoring tools challenge the assumptions on which existing demand connection frameworks were built. **Updated or new connection requirements are an essential first step to ensure system stability**, and they should be developed in a coordinated way across Europe to aim for harmonisation and avoid fragmentation, while taking into account local specificities where needed.

**Regarding planning aspects, the central challenge is the different pace of development:** data centres can be developed and deployed far faster than the transmission and generation infrastructure needed to serve them. Addressing this mismatch requires **actions on multiple fronts**: **better information on data centres capacity needs, timelines** and **demand projections**, while preserving data confidentiality; **greater transparency** on **aggregated grid hosting capacity**, and **allocation processes** aligned with societal objectives, that direct scarce capacity towards non-speculative applications. Flexible connection arrangements can then allow earlier access to power without compromising system security where reinforcement timelines are long.

On the opportunity side, the technical assets already embedded in data centre infrastructure represent an untapped resource for the power system. The **same control capabilities needed for grid-safe operation form the foundation for active grid support and market participation**. As regulatory frameworks and market access conditions mature, data centres have the potential to transition from passive consumers into active system participants that contribute to flexibility, resilience, and decarbonisation. Yet the challenges to realising this potential should not be understated, since providing such services can conflict with the core data centre business by creating additional costs and operational constraints. Their deployment is therefore likely to depend on the policy and regulatory framework, as well as on the sustainability targets of data centre owners and operators.

**The size and impact of data centres on grid planning and grid operation require a coordinated approach, under a broader framework of European and national authorities.** The window of opportunity is defined by the speed of data centre deployment itself: the frameworks must be in place before the next wave of capacity reaches the grid, not after. In particular, the potential of acting as grid resources is illustrated in Figure 10.

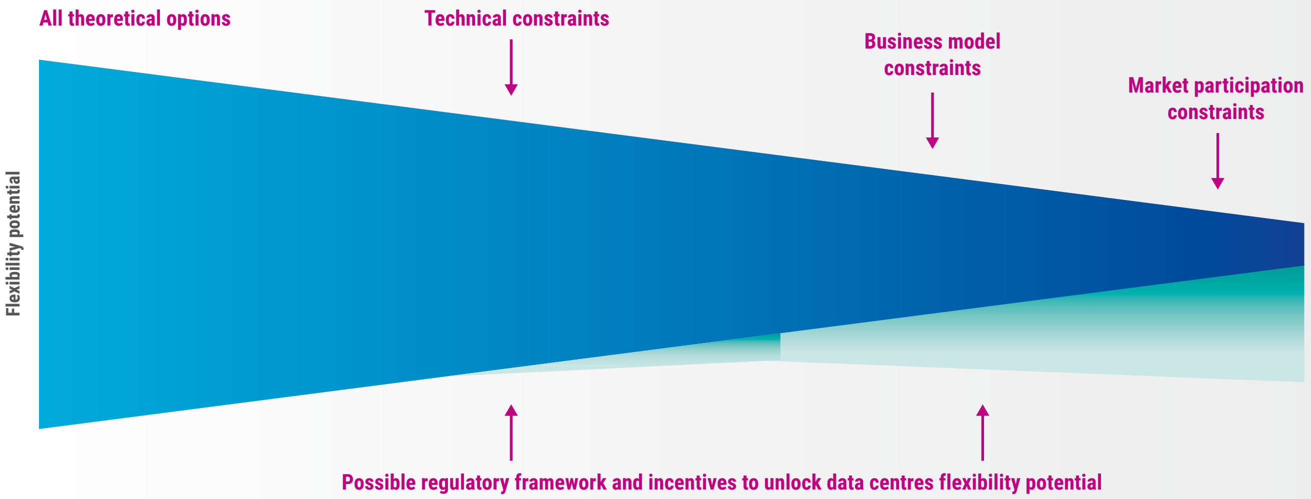


Figure 10: Unlocking data centre grid resource potential: regulatory and technical enablers can expand data centre flexibility

# Abbreviations

| | |
|---|---|
| **aFRR** | Automatic Frequency Restoration Reserves |
| **AI** | Artificial Intelligence |
| **ATS** | Automatic Transfer Switches |
| **BYOG** | Bring Your Own Generation |
| **CAPEX** | Capital Expenditure Cost |
| **DSO** | Distribution System Operator |
| **FRT** | Fault Ride-Through |
| **FCR** | Frequency Containment Reserve |
| **GPU** | Graphical Processing Units |
| **IaaS** | Infrastructure-as-a-Service |
| **EU27** | 27 members of the European Union |
| **ENTSO-E** | European Network of Transmission System Operators for Electricity |
| **mFRR** | Manual Frequency Restoration Process |
| **OPEX** | Operating Expenditure Cost |
| **PaaS** | Platform-as-a-Service |
| **PDUs** | Power Distribution Units |
| **PMUs** | Phasor Measurement Units |
| **PPA** | Power Purchase Agreement |
| **RoCoF** | Rate of Change of Frequency |
| **SaaS** | Software-as-a-Service |
| **SCADA** | Supervisory Control and Data Acquisition |
| **SNSP** | System Non-Synchronous Penetration |
| **TPUs** | Tensor Processing Units |
| **TSO** | Transmission System Operator |
| **UPS** | Uninterruptible Power Supply |
| **VPP** | Virtual Power Plant |

# Contributors

This study was conducted by ENTSO-E's Working Group Future of Energy Systems under Research, Development, and Innovation Committee with the support of Accenture.

## Review and Drafting team

Andrius Maneikis (Litgrid)

Antonio Iliceto (Terna/Co-convenor WG Future of Energy Systems)

Elena Fjørtoft (Statnett SF/Convenor WG Future of Energy Systems)

Jeanette Münderlein (TenneT TSO B.V.)

Mara Scheibenreif (APG)

Orestis Vantzos (IPTO)

Runar Moseby (Statnett SF)

Sonja Nurmiainen (Fingrid)

Ilaria Federici (ENTSO-E)

Jordanno Paiva (Accenture Industry X)

Julia Warmuz (Accenture Industry X)

Lasse Kari (Accenture Industry X)

Markus Gödde (Accenture Industry X)

The report represents the views of ENTSO-E.

## Publisher

ENTSO-E aisbl
8 Rue de Spa | 1000 Brussels | Belgium
www.entsoe.eu | info@entsoe.eu

## Images

Cover: Dragon Claws, istockphoto.com/313
page 5: Nikida, istockphoto.com/313
page 15: Gorodenkoff, istockphoto.com
page 20: Lari Bat, istockphoto.com
page 23: imaginima, istockphoto.com

## Design

DreiDreizehn GmbH, Berlin | www.313.de

## Publishing date

May 2026

European Network of
Transmission System Operators
for Electricity