Clifford Chance

March 2026

Thought leadership

Data Centres & AI Compute Infrastructure Insights 2026

Data Centres & AI Compute Infrastructure Insights 2026

Key takeaways

1. Data centres are shifting toward compute first infrastructure as AI build-out drives capital from real estate into GPUs, servers, and performance-backed capacity.
2. Regulatory, sustainability, and sovereignty pressures are intensifying, creating complex compliance requirements that increasingly influence where AI-optimised facilities can be built.
3. Contracting models are evolving toward GPU-backed capacity reservations and performance-based compute offtake agreements to manage rapid hardware refresh cycles, delivery risk, and lender bankability.

Data centres are increasingly financed less as traditional real estate and more as compute-first infrastructure. AI build-out is pulling capital up the stack, from land and power into GPUs and other specialised chips, servers and performance-backed capacity, while power availability, regulation and sovereignty requirements are becoming core underwriting issues. In this briefing we set out six trends that are shaping the market and will have a significant impact on operators and investors.

1 The rise of GPU lease finance

Industry estimates suggest, that by 2030, data centres will require around US$6.7 trillion of capex globally, with US$5.2 trillion for AI capable capacity, shifting an increasing share of digital infrastructure spend into the compute layer (GPUs/servers etc.), where refresh cycles are shorter than the underlying real estate and power. Financing GPUs has scaled in line with that shift: for example, cloud provider, CoreWeave, saw its GPU collateralised debt grow from US$2.3 billion in 2023 to a US$7.5 billion facility in 2024, showing how quickly the market has moved from early transactions to repeatable, multibillion-dollar programmes.

AI build-out is highly capital intensive, with a large portion of spend concentrated in GPUs and other short-life compute hardware. That front-loaded capex profile is pushing operators toward lease and asset-backed structures, accelerating the GPU lease finance market. AI spend occurs upfront before the full revenue stream has played out so, sponsors and operators are looking for funding that matches that procurement curve and doesn't drain equity/liquidity early on. Against that backdrop, GPU lease finance is moving beyond one-off 24-36 month equipment leases into scalable, multi-draw programmes that fund rolling acquisition of batches of GPUs over a period of time.

These financings are increasingly backed by minimum commitment levels or take-or-pay style compute contracts (and the end-user credit profile behind them) and sized to the expected contracted revenue rather than the neo-cloud's balance sheet or the theoretical resale value of the GPUs. Key risks such as delivery/installation delays, fast chip refresh cycles or the fact that the GPUs may sit in third-party data centres, are therefore addressed in part. Additionally, GPUs have an average economic life of three to five years, which can change quickly due to hardware refresh or obsolescence risk.

Therefore, long tenors can feel risky. Yet the demand for AI compute is growing fast, so the market is seeing financing structures to ensure continuous GPU deployment, while managing the three-to five-year hardware risk.

Press reports indicate that over US$11 billion has been lent to neoclouds buying Nvidia chips, suggesting a fast forming, multibillion-dollar credit market around GPU fleets. We expect 2026 to be the year that these finance structures appear more frequently as sponsors and neoclouds scale their AI capacity.

2 Increasing capital requirements

The capital needs of the data centre industry are increasing rapidly, with trillions of dollars of investment expected in data centres by 2030 and roughly 100GW of data centre capacity coming online between 2026 and 2030. Construction costs per MW are also rising as the complexity and requirements in terms of customer equipment increase. As a result, we expect continuing increases to debt requirements – for new facilities and refinancings – and the sustained uptake of other streams of financing, such as ABS and CMBS refinancings (which enable the recycling of lenders' commitments into new capex financings).

We also consider that the AI boom will lead to increased M&A and JVs in the industry as owners and operators seek scale, but face challenges with organic growth (power availability being a key constraint), including in some cases where distress arises because an owner/operator has not achieved their execution right. The development of AI gigafactories through the European Commission's InvestAI initiative is also expected to result in public-private partnership funding structures, given the scale of investment required.

It is crucial that owners and operators ensure that they are structuring operations and customer contracts with future financings and exits in mind. Ideally, to maximise flexibility, individual data centres on campuses would be ringfenced so they can be sold in different ways: individually, as an entire campus or as part of a whole portfolio. Structuring in this manner will also make the due diligence process easier with lenders.

3 Legislation and regulation: balancing oversight with the need for innovation

Governments will continue to face a dilemma over the appropriate severity and reach of legislation and regulation that, either directly or indirectly, impacts data centres and their customers (notably planning, sustainability and data). On the one hand governments are keen to ensure that they are well-placed to benefit from the AI wave and boost economic growth by avoiding over-regulation. Such over-regulation could stifle the innovation they are hoping to attract. However, this must be balanced against competing priorities around sustainability, data protection and, in some cases, data sovereignty. The primary challenge for law makers globally in 2026 is achieving the right balance, especially in a world where AI is transforming the way in which we live and work. If the regulatory burden becomes too great or the approval processes for new infrastructure become too lengthy, there is a risk that investment – particularly in certain types of data centres such as more location agnostic AI training facilities – will move to more flexible jurisdictions. The European Commission, for example, announced its Digital Omnibus proposal in November 2025 aimed at simplifying its cybersecurity, AI and data regulatory framework to address regulatory complexity. Government approaches are also being influenced by greater public awareness of and interest in data centres (particularly from a sustainability standpoint) and rising public opposition to their construction. This adds another important consideration for lawmakers.

The fragmented approach globally is also creating a challenging landscape to navigate for owners and operators who face different compliance requirements in each jurisdiction. This requires owners and operators to put in place a significant compliance function, particularly for operators of global platforms, who need to ensure that they are fully informed of both current and upcoming legislation and regulation.

4 Data centres evolve into "AI factories" (compute-first infrastructure)

We are seeing a shift from traditional colocation to AI-optimised campuses built around high-density GPU clusters (power delivery, liquid cooling, heat reuse, dedicated fibre and tighter performance monitoring). The commercial discussion increasingly starts with "time-to-compute" and "tokens/throughput under a Service Level Agreement (SLA)", rather than "space and power". These facilities are being re-imagined as industrial-scale plants that transform raw data and energy into intelligence outputs such as text, code, and video, but are not without challenges. These include:

- Evolution of technical architecture: The shift from central processing units to graphics processing units and other specialised chips requires specialised, accelerated computing systems optimised for parallel operations.
- Scaling of power requirements: Capacity is expanding toward gigawatt-scale facilities with rack densities exceeding 600 kilowatts. This scale triggers stricter regulatory obligations, as these facilities are increasingly categorised as critical national infrastructure under frameworks such as the network and information security directive (NIS2) in the EU, requiring robust incident reporting and risk assessment protocols.
- Integration of cooling technologies: The impact of such technologies includes new liability and insurance considerations regarding water leakage risks and the mandatory implementation of sustainability standards.
- Sovereignty and governance: There is active debate in several countries and in regional blocs about whether there is a need to, and how to, address technology and data sovereignty concerns. Some regulators and governments – including in the EU – are increasingly focused on who can access data (particularly where a technology provider is headquartered outside the country or bloc) and where that data should reside. Data residency requirements are already in force in a number of jurisdictions for certain categories of data. As digital sovereignty expectations continue to evolve and may proliferate in 2026 and beyond, global technology providers are responding and seeking to address concerns through a combination of contractual commitments, technological controls and operational measures designed to give governments and customers greater assurance and control over data and technology. There will be continued debate in this area throughout 2026.
- Performance-based service levels: Commercial terms are shifting from the provision of space and power to defined computational outputs, such as throughput or token generation under a service level agreement. This requires more granular technical monitoring and audit rights for the customer to verify that performance metrics are being met by the underlying GPU clusters.

5

The contracting stack will shift to GPU / capacity reservation and compute offtake agreements

Alongside lease finance, clients are asking for new forms of contracts that are closer to long-term capacity reservation or take-or-pay compute offtake: these include: (i) guaranteed capacity and performance metrics, (ii) refresh/upgrade mechanics as chips turn over, (iii) clear allocation of deployment/installation delay risk, (iv) robust step-in/termination and portability and (v) auditability and security assurances. This will increasingly drive bankability and valuation and become a key diligence item in M&A and refinancings. The main components are: guaranteed capacity and performance metrics – contracts will specify not just floor space or rack power, but also measured compute throughput (for example, GPU cycles, tokens processed, or latency guarantees) tied to service level agreements. These become the basis for pricing, performance credits and, importantly, lender underwriting.

- Refresh and upgrade mechanics: given the accelerated obsolescence cycles of GPUs, agreements will embed clear mechanisms for chip refresh, upgrades and technology substitution over multi-year terms to maintain performance targets without continual renegotiation.
- Clear allocation of deployment/installation delay risk: with long lead times for GPU supply and buildout, contracts will allocate responsibility for delays, including remedies, liquidated damages or milestone-based payment structures, to give counterparties and financiers confidence in scheduled capacity delivery. In parallel, being recognised as an NVIDIA Cloud Partner (NCP) or similar strategic GPU partner materially de-risks this execution profile for both customers and financiers. A number of companies have been positioning themselves to benefit from a more significant share of the GPU partner market.
- Robust step-in, termination and portability provisions: counterparties and lenders will require bankable rights that allow step-in or controlled termination in distress, and portability of workloads or data if the operator fails to meet obligations, reducing stranded asset risk.
- Auditability and security assurances: as compute becomes the core deliverable, contracts will contain detailed audit rights, measurement standards and security governance, so that performance and compliance can be demonstrated to both customers and financing parties.

6 Geopolitics and sovereignty will reshape data centre and AI factory projects.

Governments are increasingly using trade policies, tariffs and regulation – spanning direct investment, data, privacy, cyber, IP and AI – to assert sovereignty and channel investment and growth in their own jurisdictions. The United States and other nations will likely continue to use these measures to advance their interests amid the ongoing infrastructure boom.

A number of countries and blocs will place growing emphasis on promoting domestic facilities, such as AI factories and data centres as well as increasing local manufacturing of chips and other critical hardware. Scrutiny of direct investment in infrastructure, AI, chips and related technologies will further intensify. The European Union's initiatives, such as the push for AI Factories and the EU AI Gigafactories programme (notably, the amendment to the EuroHPC Joint Undertaking Regulation published in January 2026), may accelerate both public and private sector demand for EU-controlled compute resources and be influenced by regulations and policies over control planes, access and data.

At the same time, customers will increasingly expect "compliance and control-by-design" to be embedded within platforms and contractual arrangements. This includes requirements for logging and traceability, incident management, cyber resilience and security governance, audit rights and subcontractor controls, and demonstrable compliance in regulated sectors. Compliance with data, privacy and cybersecurity regulations will be particularly relevant given the abundance of new cybersecurity and AI specific requirements which will come into force across the globe this year.

For more information on this topic, please contact a member of our Data Centres Team.

Contacts

Gianluca Bacchiocchi
Partner, New York
gianluca.bacchiocchi@cliffordchance.com
+1 212 878 3256

Claudio Cerabolini
Partner, Milan
claudio.cerabolini@cliffordchance.com
+39 02 8063 4248

Stella Cramer
Partner, Singapore
stella.cramer@cliffordchance.com
+65 6410 2208

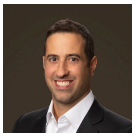
Thomas England
Partner, Singapore
thomas.england@cliffordchance.com
+65 6506 2782

Sait Eryilmaz
Partner, Istanbul
sait.eryilmaz@cepartners.com.tr
+90 212339 0 060

Jack Hardman
Partner, Dubai
jack.hardman@CliffordChance.com
+971 4503 2712

Nigel Howorth
Practice Area Leader Real Estate, London
nigel.howorth@cliffordchance.com
+44 207006 4076

Jamie McLaren
Partner, Singapore
jamie.mclaren@cliffordchance.com
+65 6661 2066



Jon Newman
Partner, New York
jon.newman@cliffordchance.com
+1 212 878 3144

Alex Pennington
Senior Associate, London
alex.pennington@cliffordchance.com
+44 207006 3583

Robert Tang
Partner, Sydney
robert.tang@cliffordchance.com
+61 2 8922 8502

Stavroula Vryna
Partner, London
stavroula.vryna@cliffordchance.com
+44 207006 4106

Carme Briera
Senior Counsel, Barcelona
carme.briera@cliffordchance.com
+34 93 344 2210

Jen Chimanga
Partner, Dubai
jennifer.chimanga@cliffordchance.com
+971 4503 2629

Matt Dunn
Partner, Dubai
matt.dunn@cliffordchance.com
+44 207006 4890

Neil Evans
Partner, London
neil.evans@cliffordchance.com
+44 207006 3388

Claire Fawcett
Partner, London
claire.fawcett@cliffordchance.com
+44 207006 8737

Marcia Hook
Partner, Washington DC
marcia.hook@cliffordchance.com
+1 202 912 5116

Emma Matebalavu
Global Head of Global Financial Markets, London
emma.matebalavu@cliffordchance.com
+44 207006 4828

Patrice Navarro
Partner, Paris
patrice.navarro@cliffordchance.com
+33 1 4405 5371

Mike Niekoop
Counsel, Amsterdam
mike.niekoop@cliffordchance.com
+31 20 711 9143

María Sabau
Partner, Madrid
maria.sabau@cliffordchance.com
+34 91 590 9445

Hein Tonnaer
Partner, Amsterdam
hein.tonnaer@cliffordchance.com
+31 20 711 9528

Charlotte Walker-Osborn
Knowledge Director – Tech Group (UK Lead), London
charlotte.walker-osborn@cliffordchance.com
+44 207006 2662

Clifford Chance

cliffordchance.com

Clifford Chance, 10 Upper Bank Street, London, E14 5JJ

Clifford Chance LLP is a limited liability partnership registered in England and Wales under number OC323571

Registered office: 10 Upper Bank Street, London, E14 5JJ

We use the word 'partner' to refer to a member of Clifford Chance LLP, or an employee or consultant with equivalent standing and qualifications

If you do not wish to receive further information from Clifford Chance about events or legal developments which we believe may be of interest to you, please either send an email to nomorecontact@cliffordchance.com or by post at Clifford Chance LLP, 10 Upper Bank Street, Canary Wharf, London E14 5JJ

Abu Dhabi • Amsterdam • Barcelona • Beijing • Brussels • Bucharest • Casablanca • Delhi • Dubai • Düsseldorf • Frankfurt • Hong Kong • Houston • Istanbul • London • Luxembourg • Madrid • Milan • Munich • Newcastle • New York • Paris • Perth • Prague • Rome • São Paulo • Shanghai • Singapore • Sydney • Tokyo • Warsaw • Washington, D.C.

*AS&H Clifford Chance, a joint venture entered into by Clifford Chance LLP.

**Clifford Chance has entered into association agreements with Clifford Chance Prague Association SRO in Prague and Clifford Chance Badea SPRL in Bucharest.

Clifford Chance has a best friends relationship with Redcliffe Partners in Ukraine.