DCS
DATACENTRE SOLUTIONS
DEVELOPING DIGITAL INFRASTRUCTURE IN A HYBRID WORLD

ISSUE IX 2025

AN ANGEL BUSINESS COMMUNICATIONS PUBLICATION

HTTPS://DATACENTRE.SOLUTIONS

2026: THE YEAR THE UK'S AI INFRASTRUCTURE MOVES INTO FULL ACCELERATION


Liquid cooling proven at exascale.
Simplified for AI at scale.
We're ready for AI. Are you?
Learn more about our end-to-end AI-ready cooling technology
Scan the QR code to learn more.
se.com/datacentre
Life Is On
Schneider Electric

VIEWPOINT

BY PHIL ALSOP
EDITOR

Data Centres at the crossroads of scale, sustainability, and societal impact

The data centre industry is entering a period of profound acceleration - one defined by hyperscale expansion, soaring energy demands, tightening sustainability expectations, and an operational paradigm reshaped by ai. Across every new dataset, forecast, and report, one theme emerges with unmistakable clarity: digital infrastructure is no longer just keeping pace with global transformation; it is driving it.

Hyperscale operators epitomise this momentum. The sector's projected rise from $58.3 Billion in 2024 to nearly $591 billion by 2034 reflects an industry undergoing tenfold growth in a decade. Demand from cloud computing, ai, big data analytics, and iot continues to stretch capacity requirements at unprecedented speed. Enterprises, holding 55% of the market, are understandably prioritising hyperscale-ready infrastructure to support mission-critical workloads, modernisation, and cybersecurity. The united states, commanding over 80% of the hyperscale market share, remains the global nexus of investment and innovation.

Yet hyperscale growth is only one part of the story. Across emea, the situation could be described as a race where demand sets the pace but capital, land availability, and power constraints dictate the limits. Knight frank's latest report reveals more than £422 billion is required to meet development needs, with hyperscale cloud and ai fuelling near-total pre-leasing in cities such as dublin, milan, and london. Regions like paris and johannesburg are roaring ahead, yet vacancy rates across the continent sit below 10%, and power constraints remain a defining bottleneck.

As such pressures mount, energy efficiency has quickly shifted from a technical objective to an existential priority. The smart energy coalition's launch comes at a time when global data centre power consumption is expected to double

by 2030 - reaching the equivalent electricity usage of japan. Ai is a major contributor here: training today's frontier models demands high-density compute environments, advanced cooling, and enormous energy supply. Even as microsoft, amazon, alphabet, and meta push collective capex towards usd 360 billion by 2025, the sector increasingly recognises that investment must be coupled with efficiency innovation. Siemens, the climate group, and others argue persuasively that energy optimisation is now business-critical, not optional.

Solutions are emerging. Heat reuse, long championed in scandinavia, is finally gaining traction in the uk, with aecom's findings showing london's data centres could heat up to half a million homes. Liquid cooling is becoming the default for ai clusters. And schneider electric's electrification roadmap points to policy reforms that could reduce european energy costs and drive resilience across the continent's infrastructure.

DCS DATACENTRE SOLUTIONS

COVER STORY

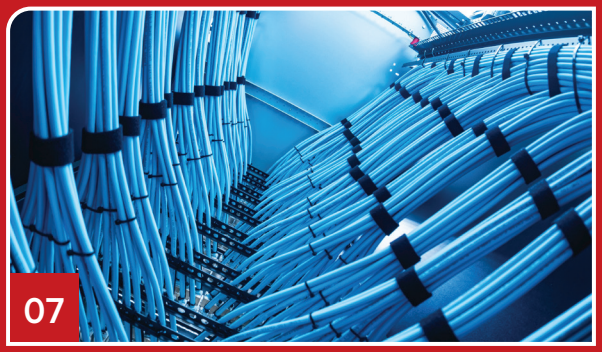
2026: The year the UK's AI infrastructure moves into full acceleration

The UK is on the brink of its most transformative infrastructure surge in decades. As AI demand soars and billions flow into new data-centre projects from Scotland to the South East, 2026 will be the year the country moves from preparation to full-scale, AI-first build-out

The data centre trade association

DCA News

NEWS



DCS DATACENTRE SOLUTIONS

Editor
Philip Alsop
+44 (0)7786 084559
philip.alsop@angelbc.com

Senior B2B Event & Media Executive
Mark Hinds
+44 (0)2476 7189701
mark.hinds@angelbc.com

Sales Manager
Peter Davies
+44 (0)2476 718970
peter.davies@angelbc.com

Director of Logistics
Sharon Cowley
+44 (0)1923 690200
sharon.cowley@angelbc.com

Design & Production Manager
Mitch Gaynor
+44 (0)1923 690214
mitch.gaynor@angelbc.com

Publisher
Jackie Cannon
+44 (0)1923 690215
jackie.cannon@angelbc.com

Circulation & Subscriptions
+44 (0)2476 718970 circ@angelbc.com

Directors
Sukhi Bhadal: Chief Executive Officer
Scott Adams: Chief Technical Officer
E: info@angelbc.com

Published by: Angel Business Communications Ltd, 6 Bow Court, Burnsall Road, Coventry CV5 6SP. UK
T: +44 (0)2476 718970 E: info@angelbc.com

Angel BUSINESS COMMUNICATIONS

Datacentre Solutions is published 10 times a year on a controlled circulation basis in Europe, Middle East and Africa only. Subscription rates on request. All information herein is believed to be correct at time of going to press. The publisher does not accept responsibility for any errors and omissions. The views expressed in this publication are not necessarily those of the publisher. Every effort has been made to obtain copyright permission for the material contained in this publication. Angel Business Communications Ltd will be happy to acknowledge any copyright oversights in a subsequent issue of the publication. Angel Business Communications Ltd. ©

ISSN 2756-1143 (Online)

Boom in hyperscale data centres: The future of digital infrastructure

The hyperscale data centre market is set for explosive growth, driven by rising demand for digital services, cloud computing, AI, and big data.

THE GLOBAL hyperscale data centre market, valued at USD 58.3 billion in 2024, is anticipated to grow at a CAGR of 26.3%, reaching approximately USD 591 billion by 2034, as per a report from Global Market Insights Inc.

This remarkable growth is largely attributed to the increasing demand for digital services, cloud computing, artificial intelligence (AI), and big data analytics. Hyperscale data centres offer massive scalability and energy efficiency, making them indispensable for technology providers, corporations, and government agencies.

The rise in data consumption from sources like social media, IoT devices, and enterprise applications further amplifies the need for these advanced systems globally.

While the COVID-19 pandemic initially disrupted construction and equipment supply in 2020, it subsequently accelerated market growth. The global shift to remote work, e-learning, and increased use of online services in 2021 prompted hyperscale providers to expand networks, edge computing, and hybrid cloud infrastructures.

Moreover, automation and remote management tools have become essential in operating efficiently amid workforce challenges, underpinning services such as AI-powered data management and workload orchestration.

In 2024, the market's solutions segment accounted for 78.8% and is projected to grow at a rate of 27.1% through to 2034. The segment's lead is due to the demand for scalable, high-performance IT infrastructure, encompassing servers, storage, networking hardware, and power and cooling solutions, needed for hyperscale settings. The burgeoning sectors of cloud computing, AI, and data analytics drive investments in robust, energy-efficient systems. Enterprises held the largest market share of 55% in 2024, forecasted to grow at a CAGR of 24.8% until 2034.

This is driven by the adoption of private and hybrid cloud infrastructures, supporting mission-critical workloads.

Key sectors, including banking, healthcare, and telecom, leverage hyperscale systems for modernisation, agility, and cybersecurity.

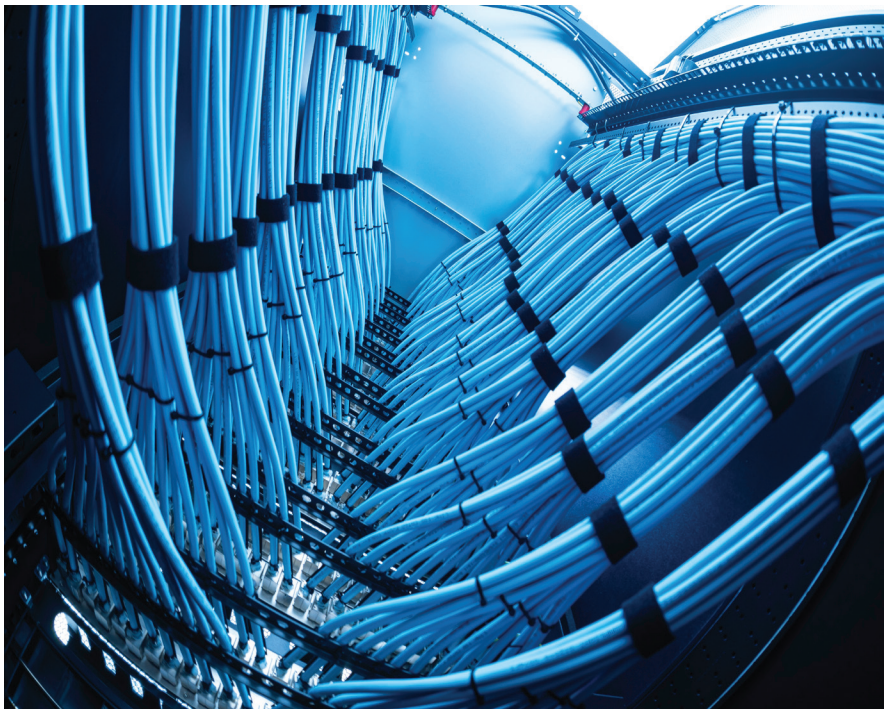
The United States dominates the hyperscale data centre market, holding an 81.6% share in 2024, generating USD 17.5 billion. The country's advantage is supported by its strong cloud computing framework, extensive digital infrastructure, and substantial investment from major technology firms. It remains the pivotal hub for hyperscale operations and cloud adoption.

Leading companies in the hyperscale data centre market include Microsoft, IBM, Amazon Web Services, and Huawei Technologies. These firms are bolstering their market presence by expanding globally, investing in innovative technologies, and diversifying services. Emphasis is placed on automation, AI-driven data management solutions, and edge computing, aiming to optimise efficiency, reduce costs, and meet the growing demand for flexible, scalable solutions.

Smart Energy Coalition rallies for energy efficiency in AI and data centres

AI and data centres demand efficiency as the Smart Energy Coalition launches global initiatives to harness energy innovation.

WORLD-LEADING companies are intensifying efforts around energy efficiency solutions which promise to shape the future of AI and data centres. Today marked the launch of the Smart Energy Coalition by the global non-profit Climate Group, focusing on energy optimisation initiatives, especially relevant during a summer of record-breaking temperatures.



The Coalition, previously known as EP100, has set its sights on enhancing energy security through smarter heating and cooling systems, presenting energy efficiency as an economic powerhouse. Comprising over 100 businesses across 200 markets and 15 industries, the Coalition's members collectively achieved a saving of US $164 million in 2024 and facilitated over 8% annual energy productivity improvements, a remarkable achievement compared to the global average of just 1%.

The demand for electricity from data centres worldwide is projected to more than double by 2030, equating to the entire electricity consumption of Japan, according to research by the IEA1. Simultaneously, as temperatures soar globally, the urgency of energy-efficient cooling systems intensifies.

The UK, along with Japan, South Korea, and parts of Europe, witnessed soaring temperatures this summer, propelling air conditioning to become the fastest-growing energy demand source for the building sector, with annual increases nearing 4% until 2035.

Sam Kimmins, Director of Energy at Climate Group, highlighted the economic and strategic advantages, stating that energy efficiency cuts costs, boosts competitiveness, and strengthens energy security. It stands as a potent solution amidst the growing demands of AI, data centres, and cooling systems.

Stacy Mahler, US Head of Sustainability at Siemens, emphasised the priority of energy efficiency amidst various technological advancements. By aligning operational and cost benefits, Siemens aims to showcase smarter ways to balance demand and supply in today's competitive landscape.

IEA's findings also underscore the broader impact of these strategic initiatives. Doubling energy efficiency by 2030 could potentially reduce global emissions by 6.5 billion tonnes while slashing global energy costs by nearly 10%.

The demand for electricity from data centres worldwide is projected to more than double by 2030, equating to the entire electricity consumption of Japan, according to research by the IEA1. Simultaneously, as temperatures soar globally, the urgency of energy-efficient cooling systems intensifies

EMEA data centre market faces investment surge and capacity challenges

Knight Frank's report reveals the EMEA data centre sector's massive growth surge, with over £422 billion in capital required to meet demand.

THE latest EMEA Data Centres Report by global property consultancy, Knight Frank, highlights the burgeoning capital needs of the region's data centre sector. More than £422 billion is required to finance the development pipeline, significantly including the £122 billion earmarked in the first half of 2025.

With EMEA's operational data centre stock now valued at £226 billion, this figure is projected to increase by 11.4% in 2025. As of the mid-year point in 2025, the region's market has accumulated a massive 50GW supply, a staggering 27.8% rise from the end of 2024. Live IT capacity hit 11.3GW, supplemented by 649.6MW worth of fresh projects.

Among the most prominent hubs, Paris leads with the fastest growth rate, needing £32.9 billion in capital to meet the 22.3% supply expansion forecast. The city has seen a 75.1% supply surge reaching 3.4GW, with live IT capacity now at 611.2MW. The rapid growth is driven partly by AI demands and partnerships with major tech firms.

In contrast, Frankfurt excels as the region's leasing activity leader, requiring £30.4 billion while marking a market valuation of £19.4 billion. An impressive take-up rate of 207MW was recorded over the previous year, bolstered by public cloud developments accounting for 93% of new leases.

Maintaining its position as the largest market, London demands over £44 billion in development capital, showcasing a valuation of £32 billion. Its capacity has expanded by 15.3%, reaching a significant 5.1GW.

Outside the leading trio, Milan ensures its place as the fastest-expanding market with expectations of a 47.9%

growth rate, while Dublin remains one of the region's highest valued at £25.7 billion despite a modest £14.8 billion capital requirement. In a notable southern turn, Johannesburg forecasts a 34.3% supply growth.

Rising demand trails the sector's growth, with an EMEA-wide colocation vacancy rate at 9.5%. Stringencies in availability highlight evident constraints, with lower vacancy in larger scale capacities.

Meanwhile, soaring pre-leasing rates exacerbate pressure with 55.2% of ongoing constructions already pre-let and 22.1% of committed capacities reserved. Cities like Dublin, Milan, and London exemplify these trends, featuring striking pre-leasing figures of 94.8%, 92.6%, and 87.8%.

Stephen Beard of Knight Frank succinctly appraises the situation: "The story of 2025 so far is one of scale colliding with scarcity. We are seeing record levels of supply delivered and planned across EMEA, yet demand from AI and cloud is growing even faster."

The report casts light on a market defined by rapid capital inflow, vigorous expansion, and relentless scarcity. With AI and cloud adoption gathering pace, the EMEA data centre sector remains poised at the vanguard of global technology development.

The story of 2025 so far is one of scale colliding with scarcity. We are seeing record levels of supply delivered and planned across EMEA, yet demand from AI and cloud is growing even faster

prysmian

BENDBRIGHT^{XS} 160 µm: A MAJOR ADVANCE IN NETWORK MINIATURISATION

The world's first 160 µm bend-insensitive fibre - higher capacity, smaller footprint

Unlock record cable density for today's space-constrained ducts, buildings, and data centres with the BendBright^{XS} 160 µm single-mode.

By shrinking the coating while maintaining a 125 µm glass diameter and full G.652/G.657.A2 compliance, it delivers faster, more cost-effective deployments that don't disrupt established field practices.

- Enables slimmer, lighter, high-count cables that install faster and travel further when blown
- Packs dramatically more fibres into the same pathway thanks to a >50% cross-sectional area reduction versus 250 µm fibres
- Ensures full backward-compatibility with legacy single-mode fibres for seamless splicing and upgrades
- Provides superior bend performance and mechanical reliability with Prysmian's **ColorLock^{XS}** coating system
- Ideal for FTTH/X access, metro densification, and hyperscale /data centre interconnects where every millimetre counts

Ready to boost capacity and shrink your network footprint?

Connect with our experts, request samples, or download the **BendBright^{XS}** 160 µm technical data today.

AI and sustainability: The twin pressures on data centres

Data centres are rapidly expanding to handle AI workloads while tackling sustainability challenges.

DATA CENTRES are swiftly evolving to meet the dual pressures of artificial intelligence (AI) workloads and sustainability demands. As AI continues to drive significant growth in capacity and investment, operators must tackle challenges like increased electricity consumption, cooling requirements, and water usage. The need for innovation in energy efficiency and green technologies has never been more critical, with experts from GlobalData highlighting these pressing concerns.

The Strategic Intelligence report by GlobalData, titled "Data Centres," underscores that AI workloads are stretching the boundaries of data centre capacity. With the majority of AI training occurring in large-scale facilities, data centres now provide essential services and stability needed for modern digital life.

Investment in data centre infrastructure is reaching unprecedented heights. Industry giants such as Microsoft, Alphabet, Amazon, and Meta have collectively reported a 2024 capital expenditure of $245 billion. This figure is projected to rise to over $360 billion by 2025, largely driven by AI-centric investments.

AI models require high-performance computing (HPC) infrastructure, specialised chips, significant memory, and advanced cooling systems. These elements position modern data centres as indispensable for powering AI innovations.

The International Energy Agency (IEA) predicts a doubling in data centres' electricity consumption to about 945 terawatt-hours (TWh) by 2030, up from 415TWh in 2024. Cooling systems and water usage also form significant sustainability challenges.

In response to these demands, tech giants and data centre operators are actively adopting low-carbon energy solutions. The development and implementation of innovative cooling technologies, particularly liquid cooling, are witnessing increased demand as operators strive to curate environmentally friendly solutions.

Bridging the AI divide

SLALOM'S latest AI Insights Survey unveils disparities between executive AI enthusiasm and workforce readiness, highlighting crucial challenges ahead.

Slalom, a leading business and technology consultancy, has shared early findings from its third AI Insights Survey, illustrating a significant difference in attitudes towards AI between executives and the workforce. The survey, which included responses from 2,000 C-Suite leaders, delved into current and anticipated AI investment trends. While AI reigns supreme in corporate strategy, gaps in skills, systems, and leadership could impede progress.

Globally, nearly all surveyed companies anticipate increasing AI spending by 2026, advancing earlier predictions. Executives view these investments as balanced, yet a substantial 93% of organisations encounter workforce challenges, with half relying on legacy platforms for core applications.

The survey refers to AI's disruption as entering its "endgame," with executives like Amy Loftus, Slalom's Chief Customer Officer, predicting a complete AI-driven transformation by 2030. The pivotal period for transition appears set between 2026 and 2028.

Sectors face misalignment between strategic AI initiatives and their practical execution. The drive towards increased AI implementation necessitates strong governance, workforce enhancement, and people-focused innovations. Amy Loftus emphasises the need for robust data foundations and frameworks to ensure trusted AI-led decisions. The survey indicates a rollout of AI confidence primarily driven by executives, while mid-tier leadership remains more cautious.

conrad energy

conradenergy.co.uk/data-centres

Tailormade energy solutions, sustainably delivered.

For most data centres, reducing energy costs is the quickest way to boost profitability. But achieving this, whilst meeting ESG targets, can often be a challenge.

As one of the UK's fastest-growing power companies, Conrad Energy develops, owns, and operates nearly 1GW of assets, with over 2.5GW in the pipeline. We also provide high-demand businesses with tailormade energy solutions, ensuring long-term reliability, cost-efficiency, and sustainability.

Our Match+ product is one of the most transparent, green supply tariffs available. We also offer Corporate PPAs and can arrange the build of on- or near-site renewable energy assets. By matching our end-to-end solutions to your unique requirements, we'll help you reduce both your costs and carbon footprint.

For further information or a free energy audit contact datacentres@conradenergy.co.uk or call 01235 427290 today, and let's power change together.

Unlocking the potential of London's data centres: A heat reuse opportunity

AECOM's report highlights the untapped potential of waste heat from London's data centres, which could warm up to half a million homes.

LONDON'S data centres have been found to emit enough waste heat to potentially warm up to half a million homes annually, an opportunity that is currently being missed. AECOM's latest report, commissioned by the Greater London Authority and in collaboration with HermeticaBlack, reveals the untapped energy residing within the capital's data centre estate.

If efficiently harnessed, it could generate around 1.6 terawatt-hours of heat per year, easily meeting the heating requirements of areas akin to Ealing.

Entitled Optimising Data Centres in London: Heat Reuse, the report suggests policy changes that could release this potential. It recommends enhancing planning guidelines, introducing infrastructural incentives, and creating a framework to enable heat offtake from data centre operators. Future data centre designs should prioritise the re-use of waste heat for potential maximum benefit.

Currently, the uptake of heat recovery in London remains limited, unlike cities such as Geneva, which utilise up to 95% of heat from data centres. However, AECOM highlights the possibility for UK cities, including London, to embrace this for sustainable heating solutions. By leveraging the data centres' heat, up to 350,000 homes could be sustained by just a portion of the excess energy being released.

In addition to addressing environmental concerns, the report points to the social benefits of exploiting such resources. With one in eight London households still grappling with fuel poverty, data centres offer a viable, local, and sustainable heat source. Situated primarily in densely populated areas of East and West London, they provide heat potentially alleviating energy poverty locally.

Crucially, data centres also align with the UK's thrust towards AI and high-performance computing. The computational demand results in higher server temperatures, subsequently producing higher-grade waste heat, which is ideal for reuse.

AECOM's associate director, Sustainability & Decarbonisation Advisory, emphasised the broader role of data centres within the UK's digital infrastructure landscape. Data centres should be viewed not only as energy consumers, but as integral contributors to the whole energy ecosystem. Recognised as one of the largest data centre hubs in EMEA, London is uniquely positioned to lead in this field. With appropriate planning and investment, these centres could significantly contribute to decarbonising residential heat, addressing housing shortages, and enhancing local energy resilience.

A precedent for such initiatives exists with the ongoing scheme in North West London, where a £36m government grant has been secured to develop a heat network, intending to reclaim heat from up to three data centres annually. As early results show promise, the optimism for widespread adaptation continues to grow.

Crucially, data centres also align with the UK's thrust towards AI and high-performance computing. The computational demand results in higher server temperatures, subsequently producing higher-grade waste heat, which is ideal for reuse

#smartconnection

www.aginode.net

LANmark™
Advanced Structured Cabling Solutions

LANsense™
Automated Infrastructure Management

Destined for density.

With more traffic within data centres and more space needed for active equipment, ultra-density is key.

Easily manage more connections and reduce errors or downtime with Aginode. Discover our single mode fibre solutions for 400G (and ready for 800G and 1,6T) or our ultra-dense patch panels with up to 144 fibres.

Plan ahead, design right and deploy agile.
Contact us: info@aginode.net

aginode

AI dominates the IT landscape

Gartner has revealed its top strategic predictions for 2026 and beyond. Gartner's top predictions span three categories: talent in the AI age, sovereignty and insidious AI.

"THE RISKS and opportunities of rapid technology change are increasingly affecting human behavior and choices," said Daryl Plummer, VP, Distinguished Analyst, Gartner Fellow and Chief of AI Research for the Gartner High Tech Leaders and Providers practice. "To properly prepare for the future, CIOs and executive leaders should prioritize behavioral changes alongside technological changes as first-order priorities."

Through 2027, GenAI and AI agent use will create the first true challenge to mainstream productivity tools in 30 years, prompting a $58 billion market shakeup.

GenAI changes will allow organizations to prioritize requirements to GenAI innovations that accelerate work completion. Legacy formats and compatibility will decline in importance, reducing barriers to entry and resulting in new competition from a wide array of vendors.

The cost and packaging of everyday GenAI is likely to change over time, with vendors moving fee-based features into a no cost tier, potentially making no cost products suitable for more users.

By 2027, 75% of hiring processes will include certifications and testing for workplace AI proficiency during recruiting

Within the next two years, expect to see many organizations implementing practical AI proficiency assessments in their hiring processes. These standardized frameworks and targeted surveys allow companies to understand candidate proficiency and close gaps in AI skills within their workforce. This trend will be especially pronounced for jobs where information capture, retention, and synthesis are major components.

As generative AI (GenAI) skills become increasingly correlated with salaries, motivated candidates will place a significantly higher premium on acquiring AI skills and need to demonstrate those abilities to solve problems, improve productivity, and make sound decisions.

Through 2026, atrophy of critical-thinking skills, due to GenAI use, will push 50% of the global organizations to require "AI-free" skills assessments
As enterprises expand their use of GenAI, hiring practices will begin to differentiate sharply between candidates who can think independently and those who lean too heavily on machine-generated output.

Recruitment will increasingly emphasize the ability to demonstrate problem-solving, evidence evaluation and judgment without AI assistance.

This shift will lengthen hiring processes and intensify competition for talent with proven cognitive capabilities. In high-stakes industries, such as finance, healthcare, and law, the scarcity of such talent will raise acquisition costs and force companies to develop new sourcing and assessment strategies.

Specialized testing methods and platforms designed to isolate human reasoning ability are likely to emerge, creating a secondary market for AI-free evaluation tools and services.

Enterprises that successfully integrate AI-free evaluation into their broader talent strategies will protect the "human edge" in decision quality and adaptability, providing an advantage that will compound as GenAI reshapes the competitive landscape.

By 2027, 35% of countries will be locked into region-specific AI platforms using proprietary contextual data

The AI landscape will fragment as technical and geopolitical factors force organizations to localize solutions, responding to strict regulations, linguistic diversity, and cultural alignment. Universal AI solutions will fade as regional differences grow.

Multinational companies will face complex challenges deploying uniform AI across global markets and will have to manage multiple platform partnerships, each with unique compliance and data governance demands. Buyers will prioritize regional platforms that offer strong performance and local compliance, while vendors will forge alliances with sovereign cloud providers and open-source models to remain competitive. Global model vendors must prove their contextual value or risk losing market share, especially in regulated or culturally sensitive sectors.

By 2028, organizations that leverage multiagent AI for 80% of customer-facing business processes will dominate

A hybrid AI model, where customer relationship management (CRM) AI handles routine tasks and humans focus on complex, emotionally charged interactions, will become the industry standard. Moreover, customers will still choose between full self-service assisted by AI interactions such as performing a transaction or learning more about a product while also choosing a human, assisted by AI to help them with things such as resolving a complex situation or billing dispute.

Organizations that fail to adopt multiagent AI for their CRM organizational processes risk losing competitive advantage as customer expectations for low effort, rapid service become the norm. Moreover, customers who find low-effort experiences often stay with the supplier/brand because of the better experience.

By 2028, 90% of B2B buying will be AI agent intermediated, pushing over $15 trillion of B2B

spend through AI agent exchanges
In this new ecosystem, verifiable operational data becomes a currency, fueling a data feed economy where digital trust frameworks and verifiability are prerequisites for participation. Products designed with composable microservices, API-first, cloud-native, headless architectures will establish a significant competitive moat. New commercial models will emerge, featuring high-frequency, frictionless sales powered by AI agents that can radically compress the sales cycle for a large range of business and technology purchases.

By the end of 2026, "death by AI" legal claims will exceed 2,000 due to insufficient AI risk guardrails

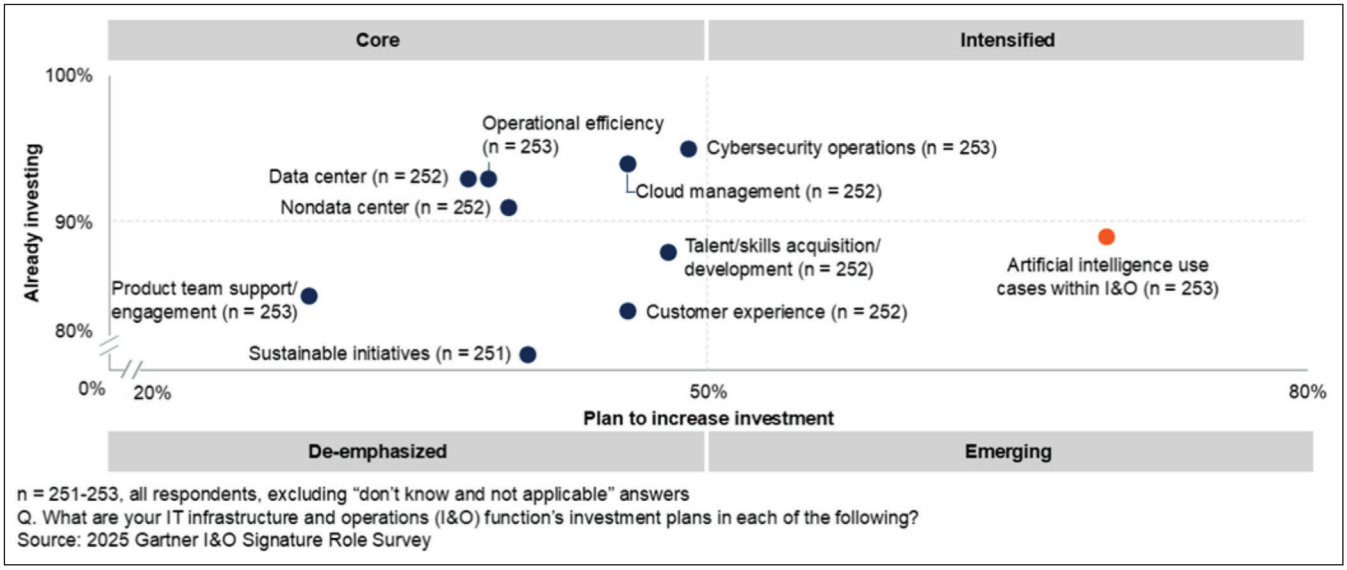
Rising wrongful death incidents of AI-related safety failures, or "death by AI," will lead to increased regulatory scrutiny and control, recalls, involvement of law enforcement agencies, and higher litigation costs.

As regulatory scrutiny intensifies, organizations will face pressure not only to meet minimum legal obligations but also to prioritize safety and transparency in their business systems through the use of AI guardrails. Somewhat paradoxically, companies will likely showcase either their AI use or lack thereof to differentiate themselves from competitors and mitigate the risk of potential litigation.

The impact of AI and decision governance failures will vary by geography due to differences in legal and regulatory systems, exposing organizations to varying risks and liabilities.

By 2030, 20% of monetary transactions will be programmable to include terms and conditions of use, to give AI agents economic agency

Programmable money is enabling new business models by allowing machine-to-machine negotiations, automated commerce, market discovery and data asset monetization, fundamentally reshaping industries such as supply chain management and financial services. Real-time programmable transactions deliver liquidity and efficiency gains by reducing friction, improving liquidity and lowering operational costs, which supports the rise of autonomous business



n = 251-253, all respondents, excluding "don't know and not applicable" answers
Q. What are your IT infrastructure and operations (I&O) function's investment plans in each of the following?
Source: 2025 Gartner I&O Signature Role Survey

➤ Figure 1. Current Investments and Planned Increases Within I&O (Percentage of Respondents).
Source: Gartner (October 2025)

operations. The rise of machine customers, such as AI agents with economic agency, will increase demand for programmable financial infrastructure, create new markets, facilitate autonomous financing and enable products that automatically adapt to changing needs. As a result, stablecoins, deposit tokens and tokenized real-world assets are evolving into mainstream financial instruments for enterprise use.

However, fragmented standards and a lack of interoperability across programmable money platforms and blockchain infrastructures will inhibit market growth and prevent AI agents and machine customers from acting as true economic actors. Security vulnerabilities in programmable money storage, access control and transaction integrity will erode trust and prompt new regulatory frameworks to govern their use.

By 2027, the cost-to-value gap for process-centric service contracts will be reduced by at least 50% due to agentic AI reinvention

AI agents will evolve to discover tacit knowledge. and interactions with them will then become the process itself. Hidden knowledge utilized by these agents will lead to new value assets. Continuous innovation-based pricing will not be limited by labor as standardized workflows are replaced by context-driven orchestration.

By 2027, fragmented AI regulation will grow to cover 50% of the world's economies, driving $5 billion in compliance investment

AI transformation is being built on AI governance. With more than 1,000 AI laws proposed last year, no two have a consistent definition of AI. AI governance can become an enabler or a barrier. While technology helps, AI literacy unlocks power. To stay safe, technology leaders will be relied on to build a perpetual "laws and regulations" mind map.

AI governance programs, with dedicated headcount and specialized software, will become the norm to manage new and evolving AI risks independent of security. These risks will be driven by both regulatory and business requirements.

AI to cut costs

Fifty-four percent of infrastructure and operations (I&O) leaders said cost optimization is their top goal for adopting AI, according to a survey by Gartner.

The survey was conducted online from May through July 2025 among 253 respondents from the U.S., U.K., India and Germany, to identify the key goals and challenges that heads of I&O face.

Expert Take:

"While AI is the top growth area for future I&O investment, integration difficulties (48% of respondents) and lack of budget (50% of respondents) were cited as top adoption challenges," said Melanie Freeze, Research Director at Gartner.

Gartner says I&O leaders need to adopt AI intentionally. "Rather than chasing big AI projects, they should start with high-value, feasible pilots and flexible upgrades," said Freeze. "For example, organizations could use GenAI for cloud cost management to automatically analyze cloud billing, resource usage and infrastructure efficiency." Cybersecurity operations and talent/skills development are the second and third most common areas, where many I&O leaders (49% and 48%, respectively) expect to increase their investments in the future (see Figure 1).

Worldwide IT spending to grow 9.8% in 2026

Worldwide IT spending is expected to total $6.08 trillion in 2026, an increase of 9.8% from 2025, according to the latest forecast by Gartner.

"The uncertainty pause that began in the second quarter of 2025 started to alleviate in the third quarter and a significant budget flush is anticipated before the end of the year," said John-David Lovelock, Distinguished VP Analyst at Gartner. "Despite being in the trough of disillusionment in 2026, GenAI features are now ubiquitous across software already owned and operated by enterprises and these features cost more money, aligning with this flush. The cost of software is going up and both the cost of features and functionality is going up as well thanks to GenAI.

"However, not all segments will experience this flush uniformly," said Lovelock. "Software and services spending growth in 2025 does not recover in the same way as devices and data center systems.

For instance, vertical-specific software spending has been slightly more affected, as vertical-industry-level software buyers are more sensitive to the policy changes and business uncertainties (see Table 1)."

Worldwide IT Spending Forecast (Millions of U.S. Dollars)

	2025 Spending	2025 Growth (%)	2026 Spending	2026 Growth (%)
Data Center Systems	489,451	46.8	582,446	19
Devices	783,157	8.4	836,275	6.8
Software	1,244,308	11.9	1,433,037	15.2
IT Services	1,719,340	6.5	1,869,269	8.7
Communications Services	1,304,165	3.8	1,363,058	4.5
Overall IT	**5,540,421**	**10**	**6,084,085**	**9.8**

Source: Gartner (October 2025)

➤ Table 1. Worldwide IT Spending Forecast (Millions of U.S. Dollars)

AI Infrastructure and Devices Drive Demand

Solid mobile phones and PCs shipment data over the first half of 2025 reported by vendors support the devices market's strong growth projection for 2025, with spending expected to reach $783 billion in 2025, an increase of over 8.4% from the previous year.

Spending on devices in 2026 is expected to be strong, but a bit slower. "This growth in 2025 is mainly driven by stronger-than-expected spending on mobile phones," said Lovelock.

"The availability of AI devices has also boosted overall spending by more than $30 billion. With the replacement cycle unchanged, the stronger performance in 2025 will result in a lower relative growth rate for 2026, as demand has been pulled forward.

"In other areas, such as data center systems, the race to build AI infrastructure has further increased demand and growth expectations for data center servers, especially AI-optimized server racks. However, server demand growth remains limited by supply constraints," continued Lovelock.

Exploring industry trends

Gartner has announced its list of top strategic technology trends that organizations need to explore in 2026. "Technology leaders face a pivotal year in 2026, where disruption, innovation, and risk are expanding at unprecedented speed," said Gene Alvarez, Distinguished VP Analyst at Gartner. "The top strategic technology trends identified for 2026 are tightly interwoven and reflect the realities of an AI-powered, hyperconnected world where organizations must drive responsible innovation, operational excellence, and digital trust."

"These trends represent more than technology shifts; they are catalysts for business transformation," said Tori Paulman, VP Analyst at Gartner. "What feels different this year is the pace. We've seen more innovations emerge in a single year than ever before.

Because the next wave of innovation isn't years away, organizations that act now will not only weather volatility but shape their industries for decades to come."

The top strategic technology trends for 2026 are:

AI Super Computing Platform

AI supercomputing platforms integrate CPUs, GPUs, AI ASICs, neuromorphic and alternative computing paradigms, enabling organizations to orchestrate complex workloads while unlocking new levels of performance, efficiency and innovation. These systems combine powerful processors, massive memory, specialized hardware, and orchestration software to tackle data-intensive workloads in areas like machine learning, simulation, and analytics.

By 2028, Gartner predicts that over 40% of leading enterprises will have adopted hybrid computing paradigm architectures into critical business workflows, up from the current 8%.

"This capability is already driving innovation across a diverse range of industries," said Paulman. "For example, in healthcare and biotech, companies are modeling new drugs in weeks instead of years. In financial services, organizations are simulating global markets to reduce portfolio risk, while utility providers are modeling extreme weather to optimize grid performance."

Multiagent Systems

Multiagent systems (MAS) are collections of AI agents that interact to achieve individual or shared complex goals. Agents may be delivered in a single environment or developed and deployed independently across distributed environments.

"Adopting multiagent systems gives organizations a practical way to automate complex business processes, upskill teams, and create new ways for people and AI agents to work together," said Alvarez. "Modular, specialized agents can boost efficiency, speed up delivery, and reduce risk by reusing proven solutions across workflows. This approach also makes it easier to scale operations and adapt quickly to changing needs."

Domain-Specific Language Models (DSLMs)

CIOs and CEOs are demanding more business value from AI, but generic large language models (LLMs) often fall short for specialized tasks. Domain-specific language models (DSLMs) fill this gap with higher accuracy, lower costs, and better compliance. DSLMs are language models trained or fine-tuned on specialized data for a particular industry, function, or process. Unlike general-purpose models, DSLMs deliver higher accuracy, reliability, and compliance for targeted business needs.

By 2028, Gartner predicts that over half of the GenAI models used by enterprises will be domain-specific. "Context is emerging as one of the most critical differentiators for successful agent deployments," said Paulman. "AI agents unpinned by DSLMs can interpret industry-specific context to make sound decisions even in unfamiliar scenarios, excelling in accuracy, explainability and sound decision-making."

AI Security Platforms

AI security platforms provide a unified way to secure third-party and custom-built AI applications. They centralize visibility, enforce usage policies, and protect against AI-specific risks such as prompt injection, data leakage, and rogue agent actions. These platforms help CIOs enforce use policies, monitor AI activity, and apply consistent guardrails across AI.

By 2028, Gartner predicts that over 50% of enterprises will use AI security platforms to protect their AI investments.

AI-Native Development Platforms

AI-native development platforms use GenAI to create software faster and easier than was previously possible. Software engineers embedded in the business, acting as "forward-deployed engineers," can use these platforms to work together with domain experts to develop applications. Organizations can have tiny teams of people paired with AI to create more applications with the same level of developers they have today. Leading organizations are creating tiny platform teams to allow non-technical domain experts to produce software themselves, with security and governance guardrails in place.

Gartner predicts that by 2030, AI-native development platforms will result in 80% of organizations evolving large software engineering teams into smaller, more nimble teams augmented by AI.

Confidential Computing

Confidential computing changes how organizations handle sensitive data. By isolating workloads inside hardware-based trusted execution environments (TEEs), it keeps content and workloads private even from infrastructure owners, cloud providers, or anyone with physical access to the hardware. This is especially valuable for regulated industries and global operations facing geopolitical and compliance risks and for cross-competitor collaboration.

By 2029, Gartner predicts more than 75% of operations processed in untrusted infrastructure will be secured in-use by confidential computing.

Physical AI

Physical AI brings intelligence into the real world by powering machines and devices that sense, decide, and act, such as robots, drones, and smart equipment. It brings measurable gains in industries where automation, adaptability, and safety are priorities. As adoption grows, organizations need new skills that bridge IT, operations, and engineering. This shift creates opportunities for upskilling and collaboration but may also raise job concerns and require careful change management.

Preemptive Cybersecurity

Preemptive cybersecurity is trending as organizations face an exponential rise in threats targeting networks, data, and connected systems. Gartner forecasts that by 2030, preemptive solutions will account for half of all security spending, as CIOs shift from reactive defense to proactive protection.

"Preemptive cybersecurity is about acting before attackers strike using AI-powered SecOps, programmatic denial and deception," said Paulman. "This is a world where prediction is protection."

Digital Provenance

As organizations rely more on third-party software, open-source code, and AI-generated content, verifying digital provenance has become essential. Digital provenance refers to the ability to verify the origin, ownership, and integrity of software, data, media, and processes. New tools such as software bills of materials (SBoM), attestation databases, and digital watermarking offer organizations the means to validate and track digital assets across the supply chain.

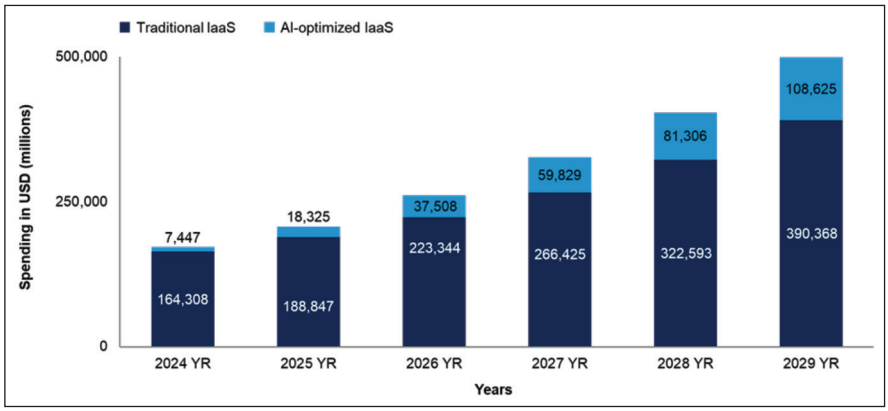
Gartner predicts that by 2029, those who failed to adequately invest in digital provenance capabilities will be open to sanction risks potentially running into the billions of dollars.

Geopatriation

Geopatriation means moving company data and applications out of global public clouds and into local options such as sovereign clouds, regional cloud providers, or an organization's own data centers due to perceived geopolitical risk. Cloud sovereignty, once limited to banks and governments, now affects a wide range of organizations as global instability increases.

"Shifting workloads to providers with an increased sovereignty posture can help CIOs gain more control over data residency, compliance and governance," said Alvarez. "This greater control may improve alignment with local regulations and build trust with customers who are concerned about data privacy or national interests." Gartner predicts that by 2030, more than 75% of European and Middle Eastern enterprises will geopatriate their virtual workloads into solutions designed to reduce geopolitical risk, up from less than 5% in 2025.

This year's top strategic technology trends highlight those trends that will drive significant disruption and opportunity for CIOs, IT and high tech leaders within the next five years. Gartner clients can read more in the Gartner Special Report "Top Strategic Technology Trends for 2026."



➤ Figure 2: Worldwide Spending and Annual Growth Rate of AI-Optimized IaaS, 2024-2029

AI-Optimised IaaS - the next growth engine

AI-optimized infrastructure as a service (IaaS) is emerging as the next disruptive growth engine for AI infrastructure. As a result, end-user spending is projected to grow 146% by the end of 2025, according to Gartner.

The AI-optimized IaaS market includes spending on high-performance computing (HPC) resources—such as graphics processing units (GPUs), application-specific integrated circuits (ASICs), and other AI accelerators—designed for large-scale AI processing.

"Traditional IaaS is maturing, however, AI-optimized IaaS spending growth projections are higher than that of traditional IaaS over the next five years," said Hardeep Singh, Principal Analyst at Gartner. "As organizations expand their use of AI and GenAI, they will need specialized infrastructure such as GPUs, tensor processing units (TPUs) or other AI ASICs, high-speed networking and optimized storage for fast parallel processing and data movement. As such, traditional central processing unit (CPU)-based IaaS will face significant challenges in meeting these demands."

Gartner estimates worldwide end-user spending on AI-optimized IaaS will total $18.3 billion by the end of 2025 and $37.5 billion in 2026 (see Figure 1). Source: Gartner (October 2025)
As AI adoption scales across industries, inferencing workloads will become a dominant force driving demand for AI-optimized IaaS. Gartner projects end-user spending on inferencing to take over that of training-intensive workloads in 2026.

aalberts

solutions made seamless

Discover our **prefabricated hydronic solutions** for data centres

Find out more

Artificial Intelligence Infrastructure spending to reach $758Bn by 2029

The global Artificial Intelligence (AI) infrastructure market is on track for unprecedented growth, poised to reach $758 billion USD in spending by 2029, according to the latest findings from the International Data Corporation (IDC) Worldwide Quarterly Artificial Intelligence Infrastructure Tracker.

ORGANIZATIONS increased spending on compute and storage hardware infrastructure for AI deployments by 166% year-over-year in the second quarter of 2025, reaching $82.0 billion. The AI infrastructure market has consistently sustained high double-digit growth for a few years, driven primarily by investment in servers for AI deployments.

Infrastructure deployed in cloud and shared environments accounts for 84.1% of the total spending in AI in 2Q25, with hyperscalers, cloud service providers and digital service providers as the largest contributors to AI spending (86.7%) in the quarter. In 2Q25, servers accounted for 98% of the total AI Centric spending, growing 173.2% compared to the same period last year.

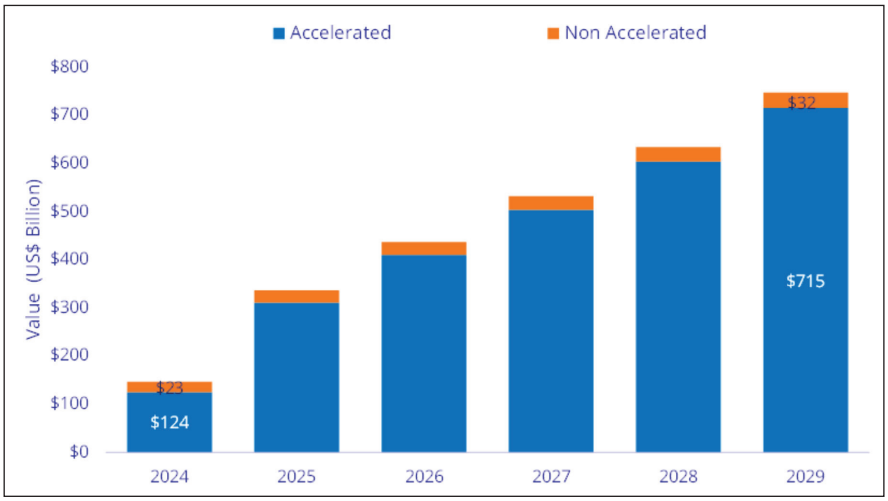
Servers with an embedded accelerator are the preferred infrastructure for AI platforms, accounting for 91.8% of the total server AI infrastructure spending – growing 207.3% in the second quarter of the year 2025. IDC projects that accelerated servers will exceed 95% of the server AI infrastructure spending by 2029, growing at a 42% 5-year CAGR.

Most of the larger changes in the AI server forecast are due to a significant reassessment of GPU and other accelerators' server demand in the US, replacing a previously expected slowdown in late 2025 and early 2026 with a new expectation that the AI investment ramp will continue through the end of this year and into 2026, based on continuously growing pipelines from major vendors and buyers.

"There is a distinct possibility that more AI-related investment will be announced in the coming years that will add to and extend the current mass deployment phase of accelerated servers well into 2026 and even beyond," said Lidice Fernandez, group vice president, Worldwide Enterprise Infrastructure Trackers. "IDC expects AI adoption to be mainly driven by hyperscalers and cloud service providers along with AI based research and education projects gaining importance by the end of the forecast period."

Storage spending in AI infrastructure has been driven by the need to manage large datasets required for training AI models, as well as storage of training, checkpoints and repositories of data for inference phases. This category reported a 20.5% year-over-year growth

in 2Q25 with 48% of the spending coming from cloud deployments. The United States leads the global AI infrastructure market, accounting for 76% of the total spending in 2Q25, followed by PRC (11.6%), APJ (6.9%), and EMEA (4.7%). Over the next five years, IDC expects the PRC region to grow at the fastest CAGR (41.5%) followed by the USA (40.5%), EMEA (17.3%) and APJ (14.3%). By 2029, IDC forecast AI Infrastructure spending to reach $758Bn with accelerated servers accounting for 94.3% of the total market spending.



➤ Worldwide Quarterly AI Infrastructure Tracker Server AI Centric VCalue US$Bn 2024-2029.

The rise of Agentic AI

IDC has released its FutureScape 2026 research, a comprehensive outlook on the next five years in enterprise technology.

Spanning more than 35 worldwide reports, IDC's FutureScape 2026 details how Agentic Artificial Intelligence (AI) will evolve from isolated pilots to enterprise-wide orchestration – transforming decision-making, operations, and competitiveness across every sector of the global economy. "Organizations today are managing through economic and geopolitical uncertainty. Amid it all, agentic AI is emerging as a strategic inflection point," said Meredith Whalen, chief product, research & delivery officer at IDC. "Our research shows that this new class of AI isn't just speeding up innovation. It's reshaping how work gets done, how people contribute, and how industries will grow in the years ahead."

Charting the agentic future

IDC's FutureScape 2026 research identifies the crosscurrents shaping enterprise strategy: economic volatility, geopolitical shifts, regulatory change, workforce disruption, changing customer engagement, and rising expectations for trust and transparency.

These pressures converge as AI accelerates, creating a pivotal moment for leadership. IDC forecasts that by 2030, 45% of organizations will orchestrate AI agents at scale, embedding them across business functions.

"Leaders can't control the geopolitical and technology crosscurrents shaping today's economy," said Rick Villars, group vice president, Worldwide Research at IDC. "But with a clear AI transformation strategy, strong data and infrastructure, and a skilled AI-ready workforce, they can turn disruption into advantage and steer their organizations toward sustainable growth in the AI era."

FutureScape 2026 delivers actionable insight to help organizations:

- Build AI-ready strategies aligned with business value.
- Develop a skilled, adaptive workforce for AI collaboration.
- Modernize technology stacks to support scalable agentic architectures.
- Embed trust, ethics, and resilience as competitive advantages.
- Understand how Agentic AI alters the business personas that interact with technology suppliers.

Top Highlights from IDC FutureScape 2026:

- **Job Role Shift:** By 2026, 40% of all G2000 job roles will involve working with AI agents, redefining long-held traditional entry, mid and senior level positions.
- **Data Readiness:** By 2027, companies that do not prioritize high-quality, AI-ready data will struggle scaling GenAI and agentic solutions, resulting in a 15% productivity loss.
- **AI Business Disruption Impact:** By 2030, up to 20% of G1000 organizations will have faced lawsuits, substantial fines, and CIO dismissals due to high-profile disruptions stemming from inadequate controls and governance of AI agents.
- **Managing Cloud Risk:** By 2028, due to geopolitical uncertainties, 60% of organizations with digital sovereignty requirements will migrate sensitive workloads to new cloud environments to reduce risk and increase autonomy.
- **Pricing:** By 2028, pure seat-based pricing will be obsolete as AI agents rapidly replace manual repetitive tasks with digital labor, forcing 70% of vendors to refactor their value proposition into new models.
- **ROI on AI Investment:** By 2026, 70% of G2000 CEOs will focus AI ROI on growth, driving C-suite efforts to boost revenue and reinvent business models without growing headcount.

> "There is a distinct possibility that more AI-related investment will be announced in the coming years that will add to and extend the current mass deployment phase of accelerated servers well into 2026 and even beyond"

2026: The year the UK's AI infrastructure moves into full acceleration

The UK is on the brink of its most transformative infrastructure surge in decades. As AI demand soars and billions flow into new data-centre projects from Scotland to the South East, 2026 will be the year the country moves from preparation to full-scale, AI-first build-out.

BY MATTHEW BAYNES, VICE PRESIDENT, STRATEGIC PARTNERS, CLOUD AND SERVICE PROVIDER

THROUGH 2026, the UK will enter the most rapid phase of digital infrastructure expansion in its history. With almost 100 new data-centres already in the pipeline representing an anticipated 20% increase by 2030, the foundations laid today will crystallise into large-scale, AI-ready campuses capable of supporting the country's next era of innovation. The scale of infrastructure now being committed makes it clear that 2026 will be the year the UK shifts decisively into high-density, AI-first digital capability.

AI demand is rewriting the UK's data-centre map

As AI adoption accelerates across every sector, 2026 will mark the moment when the UK's data-centre landscape must scale to meet the pace of demand. With an estimated 477 facilities already in operation, the UK ranks as the world's third-largest market but the trajectory ahead reflects a new level of ambition. A wave of planned sites across London, the South East, Wales, Scotland and Greater Manchester signal a shift toward not just more capacity, but a smarter, strategically distributed network of AI Factories designed for next-generation workloads.

Within this expansion, Scotland is emerging as one of the UK's fastest-growing regions for AI-oriented data-centre development. Its unparalleled access to renewable energy from wind and hydro gives it a structural advantage as operators seek sustainable ways to power high-density AI workloads.

That potential is already being realised through major projects such as Argyll Data Development's partnership with SambaNova and Schneider Electric to build the UK's first renewable-

powered sovereign AI cloud at the 184-acre Killellan. Powered by renewable energy and supported by private-wire distribution and long-duration storage, the site will scale from 100 megawatts to more than two gigawatts, demonstrating how sovereign compute and clean energy can be combined at national scale.

Alongside this, Scotland's strong innovation ecosystem spanning Edinburgh, Glasgow and Aberdeen continues to attract global investment. Developments such as DataVita's expansion, strengthened by CoreWeave's £1.5bn commitment, confirm Scotland's growing strategic relevance and its role as a natural focal point for the UK's next generation of AI infrastructure.

Further announcements by major global players will also significantly expand their UK presence. Google's £5bn data-centre programme pledge, Microsoft's $15 billion on capital expenditures, which includes the construction of the UK's largest supercomputer building in partnership with NScale at its AI Campus in Loughton and the £10bn AI campus in Blyth led by the Blackstone Group will all be at advanced construction or early operational stages sets the trajectory for 2026 as a year defined by large-scale AI infrastructure build-out.

Energy, water and resilience

As this shift accelerates, both energy and water will move to the centre of operational strategy. The National Energy System Operator estimates that data centres could drive up to 71 TWh of additional demand over the next 25 years, signalling that by 2026, energy planning will be as critical as compute planning. Regions with dense build activity, such as Slough and the Thames Valley, are likely to intensify their focus on diversified and distributed power, incorporating battery energy storage systems, behind-the-metre generation including microgrids, private wire integration and increasingly advanced liquid-cooling technologies.

This intensifying demand is directly reshaping the UK's transmission network. National Grid has begun construction on the new Uxbridge Moor substation in Buckinghamshire, which will connect more than a dozen new data centres and deliver 1.8GW of additional capacity this is the equivalent to adding a mid-sized city to the grid. As part of its £35 billon investment programme from 2026 to 2031, the substation will incorporate two SF6-free GIS substations, reducing environmental impact by around 70% and marking one of the first major deployments of SF6-free insulation technology in the country. Located beside the now-at-capacity Iver 400kV site, Uxbridge Moor exemplifies the scale of electrical reinforcement required to support the UK's AI-driven digital economy.

Water usage will become an equally defining challenge. With 28 planned sites located in areas served by Thames Water, and growing scrutiny from providers such as Anglian Water, the coming years will see operators accelerate the adoption of liquid-cooling systems that significantly reduce water consumption while supporting the thermal demands of AI infrastructure at scale.

The moment to build smarter, faster and greener

This all reinforces a simple prediction that throughout 2026, the UK will start to emerge as home to a new generation of purpose-built AI factories, engineered for high-density compute, liquid cooling and integrated on-site or behind-the-metre power. As this transformation accelerates, the UK's strategic positioning will strengthen further through deepening international alliances, including the new UK–US pact on AI and life sciences, which is set to generate tens of thousands of jobs and add significant momentum heading into 2026. That momentum is already visible across the country through the rapid rollout of AI Growth Zones in the North East and Greater Manchester to the newly confirmed site in North Wales, where more than 3,400 jobs are now slated to be created and where reforms to planning and energy access will help unlock up to £100 billion in additional investment.

In North Wales, the AI Growth Zone will harness the UK's first small modular reactor at Wylfa, uniting future-facing technologies in a key clean-energy corridor. Work to secure an investment partner is already underway, with construction set to accelerate in the coming months under the government's Modern Industrial Strategy. Supported by £5 million for local skills and business adoption, the Zone will become a hub for AI development combining compute, research expertise and regional talent to drive innovations from new medical breakthroughs to cleaner, greener industry.

A new era in digital infrastructure

Taken together, these developments confirm that AI Growth Zones will be central to regenerating communities, creating thousands of jobs nationwide, and supporting the fastest-growing cycle of AI infrastructure investment the UK has ever experienced.

It is this desire to deliver that led the UK's Prime Minister Sir Keir Starmer state, "Artificial Intelligence will drive incredible change in our country... Our plan will make Britain the world leader." By 2026, that plan will move beyond intent and its impact will be visible in concrete, steel and silicon across the nation, marking a defining moment in the UK's ascent as a global AI powerhouse.

When testing becomes trust:

Redefining factory acceptance testing for the data center era

Legrand's FAT Center in Tuscany and Keor FLEX UPS deliver proven reliability, efficiency, and sustainability for next-generation data centers.

IN TODAY'S digital economy, downtime is unacceptable. Data centers power everything from AI-driven applications to global cloud infrastructure, creating a need not just for reliability but an absolute assurance that their power protection systems will perform exactly as expected, under all conditions. For operators, the real question is no longer "Will my UPS work?" but "Can I prove that it will?"

This demand for proof is where Factory Acceptance Testing (FAT) and Legrand's high-power modular UPS, Keor FLEX, come into play. Together, they not only provide a resilient infrastructure but also offer demonstrable assurance that performance, compliance with international standards, and customer-specific requirements and efficiency are fully validated before deployment.

Why proof matters more than ever

AI workloads, edge expansion, and high-density computing are putting unprecedented stress on power protection systems. In this environment, assumptions and datasheets are no longer sufficient. Operators, investors, and customers all require evidence that systems can perform under real-world conditions.

FAT provides exactly that. Conducted in a controlled environment before equipment ships, it verifies that a UPS functions flawlessly in its exact configuration, meets international standards, and aligns with customer specifications. It transforms risk into certainty.

For Legrand, FAT is not just a procedural step; it is a strategic pillar of quality assurance. Backed by over 90 years of engineering expertise through Legrand's Borri division, the company has refined factory testing into a powerful process that ensures every UPS delivered is reliable, robust, and deployment ready.

The Tuscany Test Center: Engineering confidence

Legrand's state-of-the-art FAT Center in Bibbiena, Tuscany, is designed to set a benchmark for UPS testing, delivering unparallelled scale and precision:

- Capacity to test up to three 2.4 MW systems or six 1.2 MW systems concurrently.
- Battery testing up to 3 MWh across VRLA, Zinc, and multiple Li-ion chemistries.
- Full simulation of real-world power chains, including transformers, high power busbar, switchgear, and panels.
- Energy regeneration systems that recover up to 95% of test power, minimizing environmental impact.

This process goes beyond simple equipment validation; it's an open demonstration of transparency. Customers can see their configured UPS being tested under dynamic loads,

ask questions directly to engineers, and leave with complete documentation. In a market where trust is hard-earned, this experience is invaluable.

From optional to essential: The strategic role of FAT

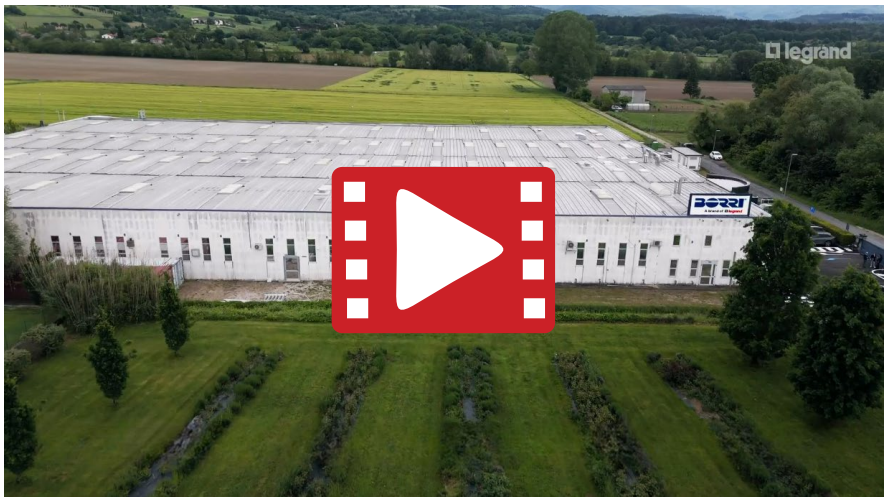
When investing in large-scale UPS systems, FATs are not optional. They represent the final checkpoint in the manufacturing journey and the first step in ensuring project success. For customers, FATs provide:

- **Risk Reduction:** Verifying that the UPS configuration matches requirements and operates flawlessly before leaving the factory.
- **Performance Validation:** Confirming efficiency, harmonic distortion, redundancy behavior, and response times under simulated real-world conditions.
- **Compliance Assurance:** Demonstrating full adherence to IEC standards and client specifications.
- **Customer Confidence:** Allowing direct observation, questions, and clarifications during the test, with complete documentation delivered.

Legrand offers a variety of FAT options tailored to meet the unique needs of each project, including Standard FAT for quality verification, Performance & Witness FAT for load testing, Reliability/Endurance FAT for long-term performance assurance, Integrated/System FAT for comprehensive power-chain simulation, and Remote FAT for live-streamed testing.This ensures that every customer receives the precise validation required, whether for individual units or large-scale redundant systems.

Keor FLEX UPS: Efficiency, modularity, and future-readiness

At the center of these tests is Keor FLEX, Legrand's flagship three-phase modular UPS system designed for data centers scaling to multi-megawatt capacities. It offers a suite of unique selling points that directly address the challenges of today's industry:



- **Unmatched Efficiency:** Achieving 98.4% efficiency in VFI mode and 99% in Eco mode, resulting in significant energy and cost savings while reducing cooling demand.
- **Scalability by Design:** Its modular architecture supports up to 1.2 MW in a single frame, expandable to 4.8 MW, enabling pay-as-you-grow scalability.
- **Smart Grid Compatibility:** The UPS can be used as a distributed energy source, supporting both reliability and sustainability.
- **Reduced Total Cost of Ownership (TCO):** Optimized energy use, minimized maintenance, and extended component lifecycles.

In summary, Keor FLEX not only protects workloads but also reduces risk, costs, and environmental impact at the same time.

Meeting the demands of AI and next-generation data centers

The surge of AI and high-density computing is forcing operators to rethink their infrastructure strategies. With quicker deployment timelines and increasing power demands, it is essential to have a reliable and efficient UPS system in place.

Legrand's expanded FAT Center directly addresses these challenges.

- **Multi-Megawatt Testing Capacity:** Supports the testing of larger and more complex systems, such as AI-driven infrastructure, simultaneously.
- **Concurrent FATs:** Running multiple FATs at the same time speeds up delivery and reduces waiting times.
- **Future-Ready Validation:** Ensures that systems are both reliable and optimized to meet AI-driven power demands.

When combined with the modular efficiency of Keor FLEX, operators gain resilience and the ability to adapt quickly as requirements change.

A European advantage

The location of Legrand's FAT Center in Italy offers customers a distinct European advantage:

- Faster scheduling and reduced shipping timelines.
- Convenient access for customer representatives to witness testing.
- Lower travel-related carbon footprint compared to transcontinental testing.

In an industry where time-to-market and sustainability goals are paramount, this proximity directly translates into a competitive advantage.

Legrand offers a variety of FAT options tailored to meet the unique needs of each project, including Standard FAT for quality verification, Performance & Witness FAT for load testing, Reliability/Endurance FAT for long-term performance assurance, Integrated/System FAT for comprehensive power-chain simulation, and Remote FAT for live-streamed testing

Beyond testing: Enhancing customer experience

Legrand has redefined FAT from a mere technical process into an enriching customer experience. Through Power Tours in Tuscany, clients can:

- Observe their UPS in operation under real-world simulations.
- Engage directly with Legrand engineers and experts.
- Enjoy a blend of technical validation and Italy's renowned hospitality.

Furthermore, the option to remotely live stream FAT provides customers with the flexibility of digital access. This ensures full visibility, assurance, and control without the need to travel, contributing to a lower carbon footprint and faster decision-making. In an environment where speed, efficiency, and transparency define success, Legrand's Live Streaming FAT provides a future-ready solution that aligns with operational and sustainability goals.

Proven. Reliable. Ready.

In the critical world of data centers, reliability cannot be left to chance.

Through its world-class FAT process and the high-performance Keor FLEX UPS, Legrand offers more than just a product; it provides proof of performance.

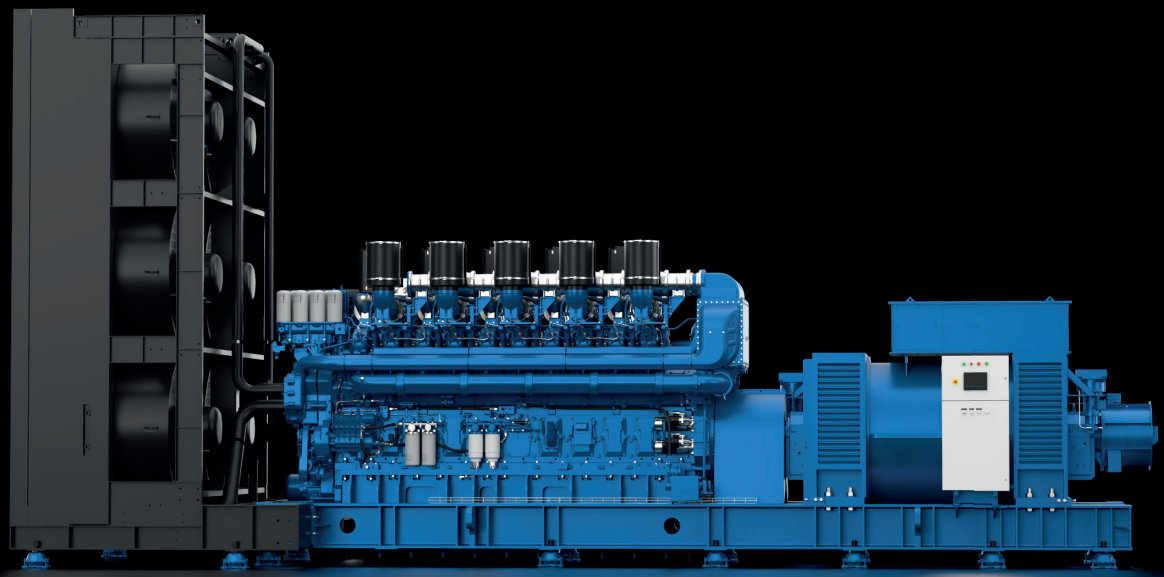
Every Keor FLEX undergoes rigorous testing from design through delivery, validating uptime, efficiency, and sustainability before it ever leaves the factory. The result is a new standard in transparency, trust, and long-term value.

For operators facing the challenges posed by AI, rising energy costs, and ambitious sustainability targets, the message is clear:

See it. Test it. Trust it.

If you would like to know more about Legrand's FAT testing or the Keor FLEX UPS, please contact us here.

BUILT TO POWER 5250 kVA



RELIABLE BACKUP POWER SOLUTIONS FOR DATA CENTRES

BEST-IN-CLASS LEAD TIME < 6 MONTHS
PRE-APPROVED UPTIME INSTITUTE RATINGS
ISO 8528 - 5 G3 LOAD ACCEPTANCE PERFORMANCE

MOTEURS Baudouin
A WEICHAI COMPANY

DURABLE. ROBUST. BUILT TO LAST. BAUDOUIN.COM

AI + modularity

Organisations that embrace modularity as an operating model, not just an infrastructure choice, will be better positioned to lead in a world where technology agility, resilience, and sustainability are non-negotiables.

BY NICK EWING, MANAGING DIRECTOR, EFFICIENCYIT

The AI infrastructure tipping point

ARTIFICIAL INTELLIGENCE has fundamentally reshaped the demands placed on data centre infrastructure.

Unlike traditional IT environments, AI workloads require unprecedented computational density, high-throughput power distribution, and advanced cooling capabilities.

The most pressing challenge for enterprises is the ability to deploy this level of infrastructure at pace, without compromising on performance, resilience or sustainability.
Traditional data centre builds often cannot keep up with the speed or velocity of AI adoption. Delays in grid access, complex construction cycles, and the high capital requirements of bespoke builds present serious bottlenecks to deployment.
Modular data centres address these challenges head-on. They allow organisations to fast-track AI infrastructure deployment without sacrificing quality or future scalability.
Prefabricated, AI-ready modules can also be deployed in as little as three to six months, compared to the eighteen to thirty-six months typical of traditional builds.
This enables enterprises to move at the speed AI demands while ensuring they remain on a path to long-term operational excellence.

Speed, scale and sovereignty

Speed-to-market has become an imperative in the race to deploy generative AI. But for many organisations, progress is hampered by long lead times, power availability constraints and an increasingly complex compliance landscape.

Modular data centres offer a strategic advantage by providing standardised, prefabricated infrastructure that can be deployed quickly and efficiently. Their controlled factory build process eliminates many of the variables and delays associated with traditional construction, and their pre-tested and pre-configured design allows for relatively straightforward or predictable

installation, even in the most challenging environments.

Prefab has also enabled a shift in the geography of compute. Instead of being limited to major metro hubs or data centre campuses, modular deployments can be positioned closer to end-users, in locations aligned with data sovereignty needs or power availability. We're now seeing organisations of all sizes take control of their infrastructure strategies by using modular solutions to scale-up operations on-premise, meet resilience and compliance objectives, and unlock new efficiencies.

AI-ready by design

To be considered truly AI-ready, a data centre must be designed from the outset to support the specific power, thermal and computational characteristics of AI workloads. This includes the ability to deliver high-density power to the rack, manage substantial heat loads through liquid or hybrid cooling, and scale capacity in line with demand.

At EfficiencyIT, we work closely with our customers to ensure every modular deployment is designed and customised to meet these exact requirements. Rather than retrofitting legacy designs or relying on general-purpose infrastructure, we provide customers with bespoke and purpose-built systems that align directly with AI workload profiles.

These are not off-the-shelf solutions. They are tailored, application-specific environments designed to support AI workloads from inception. That includes the integration of intelligent monitoring, software and control systems that deliver insight into power consumption and enable more dynamic infrastructure management as needs evolve.

Smarter infrastructure, not just bigger

Scaling AI infrastructure is not just a matter of adding more compute. It's about deploying smarter, more efficient systems that extract maximum value from every watt of power and square metre of space. Far too often, organisations rush to scale out without first optimising the infrastructure they already have, so it's important to understand what you have, how it's being used, and where further efficiencies can be made.

Modular deployments also offer a way to break this cycle. By enabling enterprises to deploy in smaller, right-sized increments, modular approaches help avoid costly overprovisioning and ensure capital is invested in line with actual demand.

Additionally, newer systems can be more conducive to integration with AI-powered infrastructure management tools, which provide real-time visibility and automation, allowing organisations to dynamically adjust power, cooling and workload distribution to maximise efficiency.

In this way, modular infrastructure can enable a more agile and intelligent approach to AI deployment, one where performance is continuously optimised rather than statically scaled.

Cooling the algorithmic arms race

New GPU systems deployed within AI infrastructure generate an extraordinary amount of heat, and traditional air-cooled environments are increasingly unable to cope. At EfficiencyIT, we're seeing liquid cooling emerge as an essential technology for next-generation deployments.

Our modular systems are designed from the ground up to support liquid and hybrid cooling methodologies, which allow for far greater thermal management and support for higher rack densities. These systems are not bolted on as an afterthought, they are integral to the design of the modular environment and allow organisations to host existing CPU and GPUs that require air-cooling, like NVIDIA DGX H100's, while futureproofing for new generations of liquid-cooled accelerated compute.

The result is improved thermal performance and far greater energy efficiency.

Our modular data centres routinely achieve excellent power usage effectiveness (PUE) ratings because they are built and tested in factory-controlled conditions and designed and modelled using VR software.
This means we can ensure optimal airflow and thermal pathways before the system is designed, configured, or it arrives on-site. As AI continues to push thermal boundaries, cooling will

be a critical differentiator, and modular design offers a clear advantage in meeting these demands sustainably and at scale.

Edge AI meets modular thinking

The proliferation of edge AI use cases—from real-time automation and analytics to autonomous operations—has brought a new urgency to deploying compute closer to the point of data generation. These applications cannot tolerate latency or rely solely on centralised infrastructure.

Modular data centres are an ideal solution for these environments. Their compact, self-contained nature means they can be deployed at the edge without the need for large-scale construction or complex integration. At the same time, they bring with them the reliability, security and performance standards of an enterprise-grade facility.
As an industry, we are also witnessing a complete rethink of the edge-core-cloud models. Modular systems are enabling enterprises to create distributed AI infrastructure that is both high-performing and tightly aligned with operational requirements.

Whether deployed in urban and industrial locations or remote environments, modular systems are supporting real-time decision-making and revealing a new value from AI at the edge.

Hyperscale has dominated the AI conversation — should it?

Hyperscale data centres have undoubtedly played a critical role in

the growth of AI, but they are certainly not a one-size-fits-all solution. Their centralised model, while efficient for certain types of workloads, often lacks the flexibility and agility required by enterprises seeking to embed AI more deeply within their operations.

Modular data centres offer a compelling alternative. They allow businesses to retain control over their infrastructure, deploy resources where and when needed, and scale intelligently as demand evolves.

This level of responsiveness is increasingly valuable in a world where AI use cases are expanding across every sector. By focusing too narrowly on outsourcing versus owning, organisations may miss opportunities to deploy infrastructure that is more aligned with their strategic business objectives, whether that's proximity to users, compliance with local regulations, or the ability to innovate quickly.

Building fast, failing faster

The ability to deploy infrastructure quickly is one of the major benefits of modular construction, but it must be accompanied by rigorous design, high standards of security and integration, and operational planning. One of the risks we've noticed in the industry is the temptation to rush deployment at the expense of long-term reliability.

In all our deployments at EfficiencyIT, we mitigate this risk through an end-to-end design and validation process that ensures every system is configured for the customer's application-specific requirements. That includes careful workload analysis, integrated systems testing, and proactive planning for future expansion or integration.

Operationally, modular systems must be managed with the same care as any traditional data centre facility. That means implementing robust monitoring, regular maintenance schedules, and predictive management tools that anticipate and resolve issues before they affect performance.

Modular infrastructure is about speed and agility, yes, but those benefits must be delivered without compromising resilience or quality.

From CapEx to composability

The shift towards modular infrastructure also requires a shift in the procurement process. Instead of committing vast capital outlays upfront, organisations can now invest incrementally, scaling infrastructure in line with current usage and revenue capabilities. This approach is particularly attractive in today's AI-driven landscape, where demand can be unpredictable and fast-changing.

The ability to combine and reconfigure infrastructure components as needs evolve is another key advantage that modular systems provide. By decoupling power, cooling, and compute elements into modular building blocks, we can enable customers to build infrastructure that adapts to their business rather than the other way around. This changes the role of the CIO and CTO. They are no longer infrastructure caretakers, they become infrastructure strategists, actively taking a role in shaping deployment models that drive optimal business performance with a clear and tangible ROI.

The long view: AI + modularity as an operating model

Looking beyond the tech, the combination of modular infrastructure and AI represents a new paradigm for organisations. It's not just about data centres, it's about how companies approach security, innovation, resilience and energy efficiency in the face of technological disruption.

Modular infrastructure enables organisations to experiment, evolve and iterate at the speed of AI. It supports decentralised IT strategies, reduces time-to-deployment, and allows said companies to enhance sustainability initiatives via lower embodied carbon and improved energy performance. At the same time, AI is beginning to inform how data centres are managed, from predictive maintenance to real-time energy optimisation, creating a feedback loop where infrastructure and intelligence are co-evolving.

At EfficiencyIT, we see this as the future of digital operations. Organisations that embrace modularity as an operating model, not just an infrastructure choice, will be better positioned to lead in a world where technology agility, resilience, and sustainability are non-negotiables.

Integrated Modular Data Centres: Strategic solutions for pressing issues

DCs are facing growing challenges like rising power demands, labour shortages, rapid growth of AI workloads etc. Traditional approaches are often too slow, costly, and unsustainable where speed, efficiency, and scalability are required.

R&M addresses this with modular, ready-to-use solutions. These support key areas including servers and storage, computing rooms, meet-me rooms, and interconnects.

Scan to contact us.

R&M France
Parc Méliès III, Bâtiment C,
9 Rue René Clair
78390 Bois-d'Arcy, France
fra@rdm.com

www.rdm.com

R&M

Prefabricated data centres: A scalable solution to take on the AI age

Introducing scalable, prefabricated solutions delivers an essential pathway for futureproofing data centres.

BY HARQS SINGH, CTO AND FOUNDER, INFRAPARTNERS

AI REPRESENTS the most consequential technological shift in human history. For the first time, we are manufacturing intelligence itself, which is driving the largest mobilisation of capital the world has ever seen. Forecasts are suggesting that global data centre capacity could rise by as much as 200GW by 2030 – a significant increase on today's footprint.

There is no denying that the UK has stepped up to the challenge of the century with a plan to unlock heavy investment in building out AI infrastructure. The reality is we cannot scale as quickly as we need using traditional methods and we must think in different and innovative ways. Fortunately, there is great reason for optimism as new platforms are due to enter the market shortly that will build at a scale that no one has previously envisioned.

We're already seeing the proliferation of value-adding AI use cases across every sector, and this is reshaping what a data centre needs to deliver in the future. AI data centres will be AI factories, as these new workloads demand radically different infrastructure from the facilities of the past. At the same time, escalating energy demands, supply chain bottlenecks and global skills shortages are putting traditional build models under pressure.

The good news is that building better infrastructure for AI is a big priority. However, it's important we understand how the industry can design, implement and scale the data centres of the future, tackling AI-based challenges and other issues. Part of the answer is undoubtedly an industrialised and prefabricated approach to building, particularly for many large-scale deployments around the world.

Two challenges for data centre deployment

The data centre industry is currently up against two key challenges which are reshaping how capital is deployed and the timeline to deliver returns — shortage and speed. We're seeing a significant shortage of talent across both construction trades and critical infrastructure roles, worsened by the sheer demand of AI. This talent gap is inflating project costs and delaying project delivery, even in cases where capital is readily available. At the same time, supply chain volatility continues to drive price increases and equipment

Speed and schedule certainty is a well-known issue in the industry, and AI has held up a microscope to the fact that progress simply isn't going fast enough. New plans for data centres in the UK are making daily headlines right now, although these are not anywhere near nor quick enough for what is needed

shortages, shifting the bottleneck upstream.

Importantly, there's also a shortage of power to meet the explosive demand for AI compute. The demand for electricity is outstripping the available supply. As an industry, we will need to look at new and different types of power sources to bridge the gap. We need to unlock innovative ways to power these sites to get GPUs up and running as quickly as possible.

Speed and schedule certainty is a well-known issue in the industry, and AI has held up a microscope to the fact that progress simply isn't going fast enough. New plans for data centres in the UK are making daily headlines right now, although these are not anywhere near nor quick enough for what is needed. Yet, building large-scale new sites in a traditional manner will typically take two to three years, and sometimes even longer if there are labour and equipment shortages. However, the AI industry does not have time right now. We need a new way.

The potential of prefabrication

To address these key challenges, the industry must look to industrialised delivery mechanisms. Standardisation and offsite manufacturing are no longer niche; they are essential tools to de-risk capital investment and compress deployment cycles.

Prefabrication enables scalable and repeatable quality, reduces the need for on-site resources, and accelerates the time to first token and revenue generation. For institutional investors and infrastructure funds, this approach introduces greater clarity and speed into capital planning, offering a path to scaled deployment without the typical trade-offs in quality or timeline.

In addition, prefabricated solutions deliver greater schedule certainty, which is a key factor that investors and operators must consider. Manufacturing offsite in controlled environments shortens timelines, reduces risk and bypasses many local constraints, including skilled labour shortages. It allows data centres to be treated as 'products' rather than long-term conventional construction projects, reducing timelines from years to just months from conception to full deployment.

Why can't legacy sites just adapt? It's not just speed that can be tackled with prefabrication, but also power shortages. Legacy sites were never designed to handle the current power demands, and this issue is only going to increase if nothing changes.

While liquid cooling is becoming standard, this requires intricate planning and infrastructure that older sites rarely accommodate well. Retrofitting is possible but often costly, disruptive and inefficiently created. Prefabricated data centres that are purpose built for GPUs can integrate advanced cooling, optimise power and support high load densities without compromises.

There's also the opportunity to 'follow the power' if a data centre is manufactured offsite e.g., easily deploying data centres in an area where there is more grid capacity. Location flexibility enables energy-efficient deployments that can be positioned closer to available renewable energy sources, reducing both grid strain and environmental impact.

Looking ahead

It's clear the industry is at a crucial inflexion point. Amid the exponential growth of AI and soaring workloads, inference costs are falling, use cases are multiplying and demand for capacity is set to accelerate. What is built today needs to be able to overcome the challenges of tomorrow. Therefore, facilities should be designed for easy upgrades to handle shifting demands, like new cooling systems, AI workloads and shorter GPU lifecycles. And it is the responsibility of both the developers and the stakeholders to prioritise education and understanding about what these future needs will look like for their digital infrastructure.

GPU-infrastructure lies at the heart of the transformation taking place across every industry, and prefabricated data centres serve as a solution to many of the strains being put on infrastructure today. If the UK is to live up to its potential as a 'superpower' for AI technology, infrastructure cannot be compromised. Introducing scalable, prefabricated solutions delivers an essential pathway for futureproofing data centres.

The AI future we need starts with the data centres we build

By reimagining data centres as intelligent, connected systems rather than static infrastructure, we can ensure they play an active role in the broader energy ecosystem. Their impact on grid stability, sustainability and economic progress will be essential as AI adoption accelerates.

BY JUAN COLINA, EMEA DATA CENTRE & IT SEGMENT LEADER AT EATON

THE CONVERSATION around data centres is no longer novel. Beyond discussing their benefits or their role in enabling the UK's AI ambitions, the focus has shifted to industry ensuring it can deliver the data centre capacity the country needs.

Achieving this involves several priorities. One is ensuring capacity is delivered where it is needed most. Another is protecting the UK's progress on environmental goals. For data centres, this means striking a careful balance between growth and environmental impact.

One route forward is to reconsider what a data centre needs to be. The common view holds that they must be large, fixed builds with high upfront costs and long development timelines. But there is now an opportunity to move away from that format and towards modular facilities which offer more flexibility. These can adapt to customer needs and local infrastructure while placing less strain on the grid and the environment.

This approach allows providers to move at the pace of demand. It also opens up greater potential for innovation and commercial success, while encouraging the use of more responsible materials and construction methods.

Delivering capacity at pace

As the capabilities of AI increases, and understanding of its offering grows, so will demand for fast and reliable access to it. To effectively respond, operators therefore need to develop and deliver data centres when and where they are needed. Prefabricated or modular systems provide an answer to this. For these systems, components are manufactured in controlled environments off-site, allowing for high-precision production before being delivered and connected at the final location.

Unlike traditional, on-site construction methods, modular approaches avoid many of the delays that typically slow down builds. With fewer disruptions, the financial benefits of prefabrication are maximised, particularly when project delays can be so costly. They also allow operators to expand their reach quickly without causing their carbon footprint to rise at the same rate.

Through centralised production, businesses gain greater oversight of the life cycle and emissions of each component. In turn, this allows for the selection of more sustainable materials and processes, ensuring consistent improvements can be implemented across all projects. Additionally, having a single manufacturing hub reduces the need for extensive transport of people and goods, cutting logistics costs and lowering Scope 3 emissions.

Integrating flexible energy systems and storage technologies also allows for operators to better use renewable energy, which is inherently intermittent, helping to support progress towards net zero goals and reducing carbon emissions.

Becoming a grid partner

As is widely known, data centres place significant demand on the energy grid, with forecasts suggesting energy usage could increase sixfold over the next decade. This challenge cannot be ignored, and instead, infrastructure must evolve to support the increasing load imposed by new facilities.

Crucially, data centres must move towards becoming net contributors to the stability of the grid. When combined with on-site renewables and advanced technologies such as grid-interactive UPS systems, battery storage, and intelligent energy management platforms, data centres can operate as Distributed Energy Resources (DERs). This setup enables them to either draw or return energy to the grid as and when is needed, helping to balance supply and demand without compromising operations. In turn, this reduces pressure on the grid and can even open up new revenue opportunities.

Maximising your workforce

Modular data centres also help address the concerning shortage of skilled professionals in cloud computing and AI. They provide a practical response by moving work from the building site to the factory, with the added benefit of up to 35% efficiency gains from repeatable manufacturing processes.

This approach means organisations can reduce reliance on specialist teams at every site and instead deploy expert talent to where the need is greatest. Additionally, modular systems support scalability and long-term adaptability, as they can be expanded in phases. This means facilities can expand their capacity incrementally, which is critical when adapting to AI workloads that often spike unpredictably.

Delivering an AI-powered future

As AI reshapes the digital landscape, operators must be ready to support the rapid technological change. Modular data centres offer the flexibility and speed needed to scale efficiently and support evolving workloads.

Challenges remain, but the benefits of prefabrication provide a solid foundation to meet them. Faster deployment and consistent build quality help organisations grow without being held back by labour shortages or supply chain issues. At the same time, modular designs enable integration of energy-efficient systems and sustainable practices, helping data centres contribute to a more stable and lower-carbon grid.

By reimagining data centres as intelligent, connected systems rather than static infrastructure, we can ensure they play an active role in the broader energy ecosystem. Their impact on grid stability, sustainability and economic progress will be essential as AI adoption accelerates.

The data centre network's role in energy efficiency

For UK operators navigating energy uncertainty, emissions targets, and ever-increasing data demands, rethinking the switch could be the smartest move they make.

BY DARREN WATKINS, CHIEF REVENUE OFFICE AT VIRTUS DATA CENTRES

IN today's data centres, energy efficiency has become both a responsibility and a competitive advantage. Operators have already made big strides in optimising how facilities are powered and cooled. But there's a third pillar that's often overlooked - one that's now proving as critical as AI, analytics, and high-density workloads reshape infrastructure demands.

That pillar is the data centre network itself.

The hidden cost of switching

At the heart of the network lies the switch. These devices direct traffic between servers, storage arrays and compute resources, performing countless operations every second. Traditional switches, however, rely on optical-to-electrical signal conversions to route data. These conversions, while fast, burn energy and release heat, adding pressure to both power budgets and cooling systems.

In the past, this overhead was tolerable. But with newer workloads, particularly AI training and edge analytics, requiring vast volumes of data to move fast and often, the inefficiencies in the switching layer are becoming impossible to ignore.

When networking becomes a thermal issue

It's not just the power draw of the switch that's a concern, it's the heat it throws off. Every watt consumed at the switch becomes heat that needs to be extracted from the environment. And in tightly packed racks already pushing thermal boundaries, inefficient switching can tip the balance, forcing operators to limit density, increase airflow, or reengineer cooling layouts. This turns a networking issue into a mechanical one. The switch, in other words, is no longer just an IT component - it's a source of system-wide strain.

A shift to all-optical switching

In response, a new wave of technologies is challenging the status quo. All-optical switching eliminates the need for energy-intensive conversions

altogether, keeping data in its light-based form from ingress to egress.

One UK company leading the charge is Finchetto. It has developed a passive optical switching platform that enables packet-level routing without the typical power overhead. By avoiding electrical processing, it not only reduces energy use but also slashes heat generation.

Optical switching gives operators the ability to reclaim both power and thermal headroom. By keeping data in the optical layer, it not only improves efficiency but also streamlines cooling strategies and creates new flexibility in how data halls are designed - from rack layout to density planning.

Knock-on effects across the stack
The benefits of optical switching ripple through the facility. Less heat from switching means fewer hotspots, lower fan speeds, and better airflow management. That can allow for tighter rack densities or create headroom for future growth.

It also eases pressure on supporting electrical systems. With reduced draw on power distribution units (PDUs) and uninterruptible power supply (UPS) units, facilities can downsize backup infrastructure or avoid costly overprovisioning.

For sites exploring heat reuse strategies, whether through local heating networks or internal energy loops, cutting out waste heat at the source makes thermal integration simpler and more effective.

Designed for gradual adoption

Operators are right to be wary of introducing unfamiliar networking gear into live environments. That's why many next-gen optical solutions are built for modular deployment. Finchetto's system, for example, supports standard protocols and can be slotted into spine-leaf architectures without ripping and replacing legacy hardware.

This allows for a test-and-expand approach: start with one rack or row, measure the impact, then scale. It's a pragmatic route that aligns with typical refresh cycles and retrofit timelines across the UK's colocation and enterprise market.

Why it matters now

Three forces are converging to make switching efficiency a top priority:

- **Workload evolution:** AI, high performance computing (HPC), and real-time applications are pushing bandwidth and latency demands skyward.
- **Sustainability pressures:** With ESG commitments rising and energy costs volatile, every kilowatt saved makes a difference.
- **Technology readiness:** Optical switching has matured and it's no longer confined to labs and telcos and it's ready for production environments.

Thinking systemically

The days of treating the network as an isolated domain are over. From cooling dynamics to energy provisioning, decisions made in the switching layer now influence the entire facility. A holistic view is essential.

Forward-thinking operators are reframing their network fabric as both a performance driver and an energy asset. And in doing so, they're uncovering new opportunities to reduce footprint, control costs, and hit sustainability targets.

A smarter path forward

Switching may not be the most visible part of the data centre, but its impact is growing. As optical technologies mature, the opportunity to boost efficiency without disrupting everything else becomes too valuable to ignore.

For UK operators navigating energy uncertainty, emissions targets, and ever-increasing data demands, rethinking the switch could be the smartest move they make.

Optical switching gives operators the ability to reclaim both power and thermal headroom. By keeping data in the optical layer, it not only improves efficiency but also streamlines cooling strategies and creates new flexibility in how data halls are designed

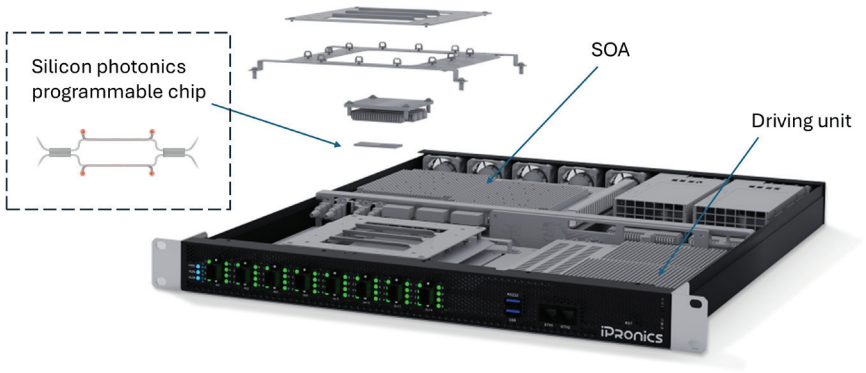
Revolutionising AI infrastructure with iPronics ONE: a new era of optical networking

As artificial intelligence continues to reshape industries, the infrastructure behind it is under pressure to keep up. Data centers are facing growing demands for speed, scalability, and energy efficiency. At the center of this transformation is iPronics, a company redefining how networks operate with its innovative technology: the first programmable silicon photonic switch, purpose-built for AI.

A Smarter Way to Move Data

UNLIKE traditional switches that rely on static routing and copper connections, iPronics ONE performs faster switch/reconfiguration times and offers lower cross-talk. It's designed to adapt in real time to changing traffic patterns, making it ideal for the unpredictable nature of AI workloads.

With microsecond reconfiguration times, ultra-low latency, and a compact, stackable design, iPronics ONE helps data centers scale without the need for costly infrastructure overhauls. It's a smarter, faster, and more efficient way to keep up with AI's rapid evolution without compromising reliability or performance.

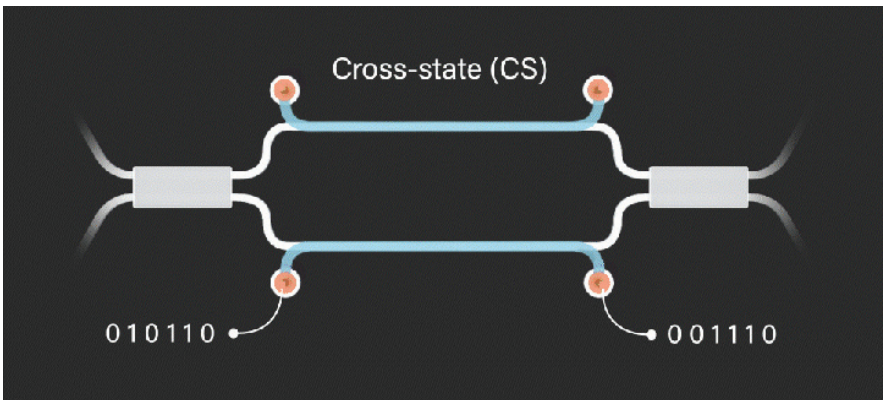


Built for Speed and Intelligence

At the heart of iPronics ONE is its ability to route data at the speed of light, eliminating the need for energy-intensive electrical-to-optical conversions. It operates across the full optical bandwidth and supports flexible scaling from 32 to 256 ports, making it suitable for a variety of AI applications. But speed isn't everything. iPronics ONE is also intelligent. It features built-in telemetry and gain control, allowing the system to monitor and adjust network conditions in real time.

This includes:

- Monitoring optical power across channels
- Tracking temperature of key components
- Alarms for specific scenarios



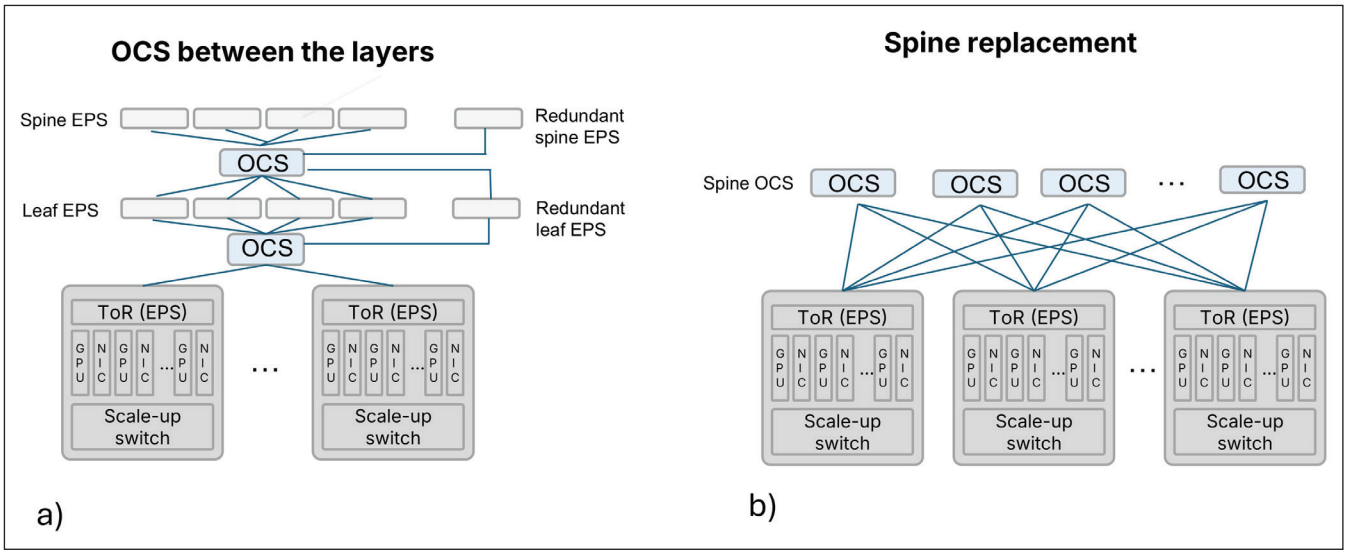
These capabilities enable dynamic reconfiguration and manual circuit setup, all managed through a software interface that integrates easily with existing data center systems using standard protocols like Yang models, Python, gNMI, and Netconf.

Ready for What's Next

iPronics ONE is designed to grow with your infrastructure. It's compatible with today's standards and ready for tomorrow's innovations, supporting:

- Pluggable transceivers
- Optical interposers
- Near package optics
- Co-packaged optics (CPOs)
- Chiplets for xPUs and switches

This flexibility makes it a future-proof investment for data centers looking to stay ahead of the curve. The company's roadmap includes models like ONE-32, ONE-64, ONE-128, and ONE-256 each doubling capacity while maintaining a compact size.



Built to handle AI's toughest challenges

Training large AI models, especially language models, requires thousands of GPUs working together. When one fails, the entire process can stall.

Traditional networks rely on checkpointing pausing to save progress which adds delays and consumes resources.
iPronics ONE changes the game. Its ability to reconfigure instantly means failures can be bypassed without interrupting the job. No pauses. No performance hits. Just uninterrupted progress.

It also supports multicast function, allowing data to be sent to multiple GPUs at once critical for efficient training during data parallelism.

Flexible topologies for real-world traffic

Most data centers are built for uniform traffic, but real-world patterns are anything but. iPronics ONE allows networks to adapt on demand, optimizing routes based on actual usage. This reduces latency, improves bandwidth efficiency, and ensures better performance for high volume data flows.

Whether scaling up within a rack or scaling out across clusters, iPronics ONE delivers the flexibility needed to support dynamic AI workloads.

Reliable, Scalable, and Ready to Deploy
iPronics has partnered with top-tier suppliers to ensure its technology is ready for large-scale deployment. From advanced chip bonding to thermo-mechanical simulations, every detail is designed for improve reliability and manufacturability.

Its solid-state design eliminates moving parts, reducing maintenance and increasing durability. And with a clear path to cost-effective scaling, iPronics ONE is not just innovative, it's practical.

The iPronics effect

iPronics isn't just about speed, it's about smarter infrastructure. It enables real-time reconfiguration, boosts GPU performance, and cuts energy use. All without the need to rebuild your network from scratch.

In a world where AI moves fast, your infrastructure should too. iPronics ONE helps you stay ahead with light, not limits.

Contact
(+886)2-2755-6111#153
hetone@hetone-group.com

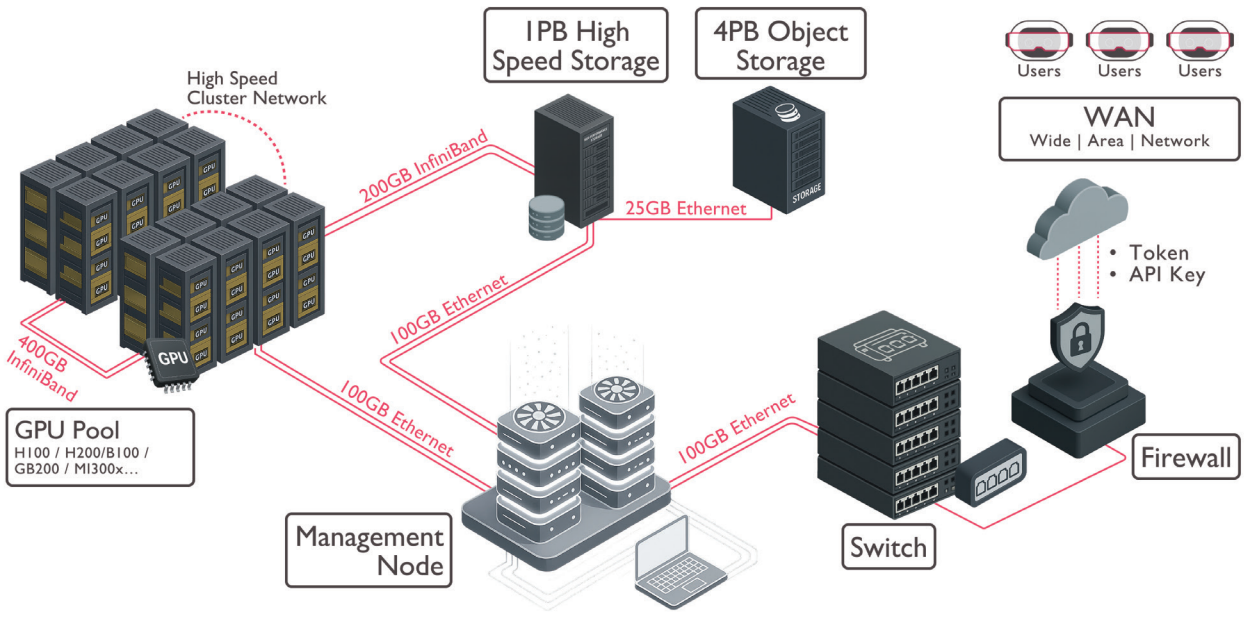
Liquid Cooling Powers the AI Era

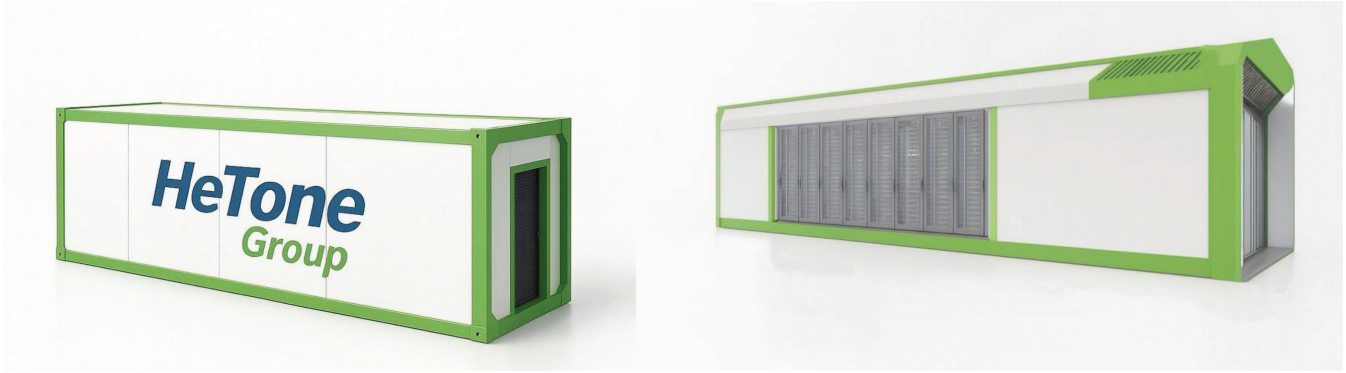
Hetone Group specializes in AI data center solutions, integrating liquid-cooling materials distribution, high-performance server sales, and AI computing platforms to help clients build high-performance, energy-efficient, and sustainable AI infrastructures. We have successfully established AI GPU compute centers in Taiwan and Thailand, and have driven the upgrade of air- cooled data centers to liquid-cooled architectures. With our global presence and localized services, Hetone is dedicated to becoming the most trusted liquid-cooling expert for our clients.

One-Stop AI Data Center Solutions

- **Company Profile**: Hetone Group specializes in next-generation AI data centers, integrating liquid cooling technology, server sales, GPU management platforms, and global connectivity services to help clients build high-performance, energy-efficient, and sustainable infrastructure.
- **Liquid-Cooled Data Center Planning & Deployment**
  Complete solutions for liquid cooling upgrades and new builds, reducing energy consumption and boosting efficiency.
- **Supermicro Server Portfolio**
  Full range of AI GPU, cloud computing, and high-density storage servers, tailored for diverse scales and applications.
- **Infinitix AI-Stack GPU Management Platform**
  Authorized software distributor, designed for GPU clusters with real-time monitoring and flexible resource management.
- **Global Data Center Connectivity & Colocation Services**
  Partnering with international operators to enable cross-border cloud and data center interconnection.
- **Value-Added Service**
- **Capital × Compute × Ecosystem**
  We connect investors with compute demand, maximizing capital value and accelerating the growth of the global AI ecosystem.

High-Performance Computing Architecture





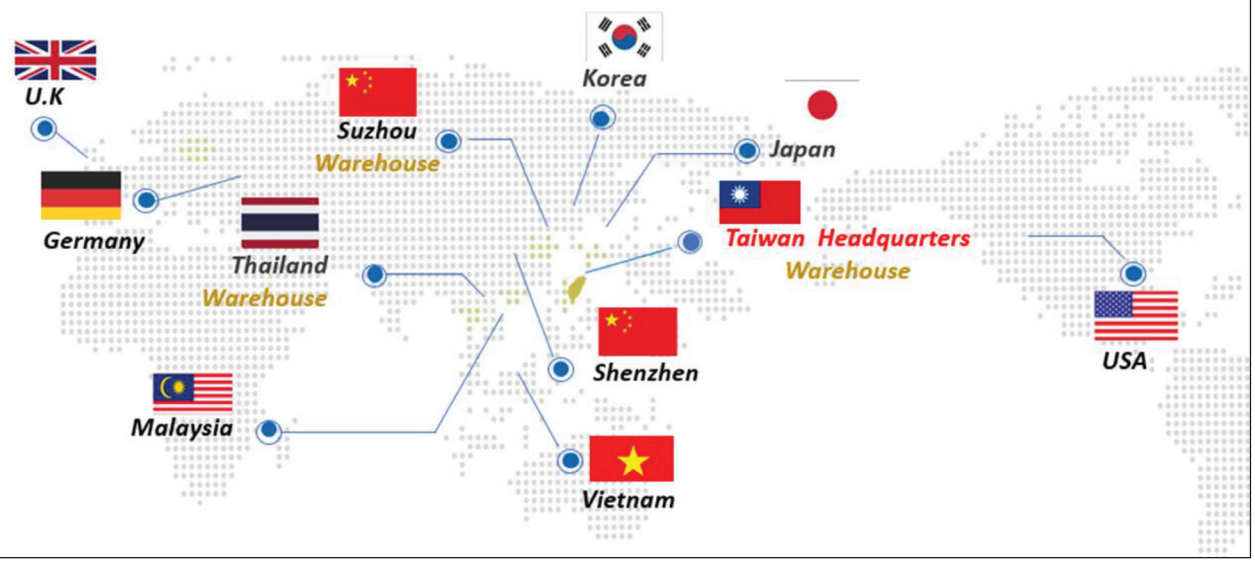
CoolPOD (POD Data Center)

1. Only USD 12.5M investment, fully operational within 7 months
2. CoolPOD Modular Liquid-Cooled Data Center – fast delivery, scalable on demand
3. Up to 64 AI GPU servers per module, built for high-performance computing
4. PUE < 1.1 with 40% energy savings, ESG & sustainability compliant
5. Compute-as-a-Service model, ensuring strong and predictable ROI

AI-Stack enhances GPU efficiency when helping enterprises implement AI

90% ↑	10x ↑	1min ↓	10x ↑
GPU Utilization	Workload execution	Development Environment Setup	Enhanced ROI
GPU Utilization	**Workload execution**	**DevEnv Setup**	**Enhanced ROI**
GPU Partitioning Increases Utilization 30% → 90%	Multiple users and tasks increase efficiency by 10x	Reduces setup time from 2 weeks → 1 min	Boosts return on investment by 10x

Global Layout of Hetone Group 合通全球佈局



Designing racks for density, airflow, and peace of mind

Modern data centres are being placed under pressure by four converging forces. AI-class compute is driving up rack power and heat densities. Exploding traffic and interconnect needs are pushing rapid upgrades to 400/800G fabrics and a far higher number of fibre terminations per rack.

RACHID AIT BEN ALI, PRODUCT & SOLUTIONS MANAGER SMART BUILDING & DATA CENTER, AGINODE

WORKLOADS are shifting toward the edge, creating vast numbers of small sites that need to be built and turned up fast with standardized, high-density connectivity. Real estate availability, sustainability requirements, and grid constraints are forcing operators to densify existing footprints rather than expand.

Net result: more ports, hotter racks, and faster rollouts – without more floor space. Meeting the demands is about getting the connectivity fundamentals right, and racks play an important part in that. In this article, we'd like to look at the rack-level requirements that make density sustainable and operations predictable.

Rack solutions and cable management

Hardware should support high port density while remaining front-serviceable so that technicians can work without disturbing adjacent cords. Cables should be routed with vertical and horizontal managers and overhead or underfloor pathways should be reserved and sized for realistic fill ratios. Don't let cable bulk intrude into airflow. Respect minimum bend radius for both fibre and copper, separate power from data where practical, and provide slack storage that doesn't block equipment intakes. It's also important to standardize labelling and U-numbering across rows, and back this up with instrumentation or automated documentation so that port-level changes are detected, guided, and reflected in DCIM and ticketing systems.

Hot–cold aisle discipline and containment

The typical cabinet is getting hotter, and the upper tail (AI/HPC rows) is getting much hotter. The latest Uptime Institute Global Data Center Survey (2024) finds that though 7–9 kW racks are gaining share. Only a minority of sites host racks above 30 kW today, but Uptime expects that to change as AI builds scale. AFCOM's 2024 State of the Data Center puts average rack density at around 12 kW – up from ~8.5 kW the previous year – and notes most operators plan further increases over the next 12–36 months. The same dataset shows many operators raising density first via better airflow and containment, with liquid cooling adoption increasing as loads grow.

Of course, air cooling only works when air goes where it should. Congested cabling blocks airflow paths at the rack face, inside cabinets, and under raised floors, increasing fan energy, hot spots, and ultimately PUE. It's vital to align equipment front-to-back and deliver supply air to cold aisles while returning exhaust to hot aisles. Use full or partial containment – end-of-row doors, roof panels, blanking panels, and brush grommets – to cut bypass and recirculation. Seal cable cutouts and side gaps and keep perforated floor tiles or supply grilles in the cold aisles only.

Distances, clearances, and service access

Cabling density is rising because modern racks are terminating far more high-speed links per RU and each link increasingly uses parallel fibre. Shipments of 400G and 800G datacom optical modules nearly quadrupled in 2024 (over 20 million units), and analysts expect another strong jump in 2025 – directly translating into more transceivers, jumpers, and trunk capacity at the rack face. On the switching side, new 51.2T/102.4TASICs are driving massive rollouts of 800G/1.6T ports (Dell'Oro projects 100 million such switch ports shipped

by 2028), which again multiplies the number of fibres landing in each cabinet. AI servers magnify this effect: compared with traditional x86 boxes, a typical AI/ML server today fans out to 12+ Ethernet switch ports (vs. ~5–6), so the cable count per rack skyrockets even before you factor in breakouts. Technology shifts are packing more terminations into the same rack space. Very-small-form-factor (VSFF) duplex connectors such as CS, SN, and MDC shrink the connector footprint, allowing panels to present more active ports in each RU. At the same time, the industry is standardizing on higher-fibre MPO formats for parallel optics – most notably MPO-16 per IEC 61754-7-3/TIA 604-18 – to support 400G SR8 and 800G SR8/DR8 applications; this increases fibre counts per link but keeps front-panel real estate efficient.

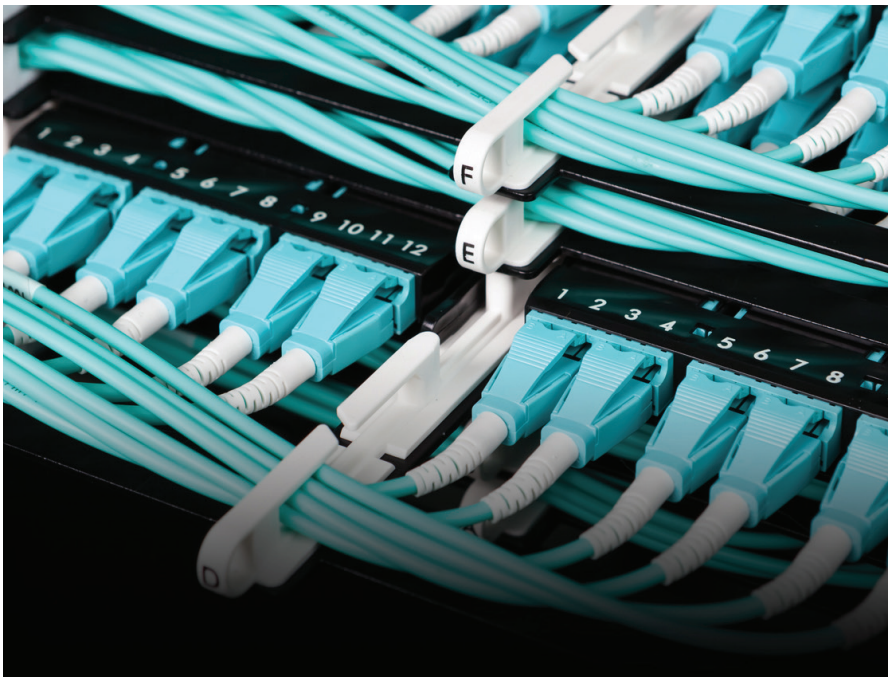
Increasing density introduces issues if you don't leave enough elbow room. If there isn't enough space to manipulate cables, small-radius bends and microbends may occur, which raise attenuation (loss) and can trigger intermittent errors, for example. Cramped conditions also slow MAC work and raise MTTR (more time tracing cords, higher risk of accidental disconnects), compounding downtime risk.

Preserve manipulation space by bringing the work surface to the technician (sliding, lockable trays with guided routing) and by moving passive patching out of the rack when density is extreme (overhead frames). Aim for roughly 1.2 m (about 4 ft) of clear space in the cold aisle to accommodate people, doors, and tools, and at least 0.6–1.0 m behind racks depending on equipment depth and door swing - this is reason why we recommend double door at the back and not a single rear door. Leave headroom for trays and ladder racks and maintain row-end clearance for egress and containment doors. Local code, seismic, and accessibility rules may push these numbers higher.

Perforation and heat dissipation

Cabinet doors need to breathe. Too little perforation in chokes airflow, raising pressure drop so server fans ramp up—driving hotter IT inlet temps, more energy use (and noise), and worse PUE. Back-pressure and resistance also cause exhaust recirculation through side gaps, creating hotspots and forcing lower practical rack densities (more cabinets for the same load). Teams often "solve" it by propping doors open, which breaks containment and security.

Specify front and rear doors with a minimum of 70/80% open area as a baseline for air-cooled rows; consider 70–80% open area where cabinet power densities are high and structural/security needs still pencil out. Populate all unused rack spaces with blanking panels to prevent bypass air. Add rear-door heat exchangers, chimneys, or tighter containment if thermal modelling shows they'll deliver measurable benefit.



Bringing it all together

To summarise: we recommend designing for density, protect airflow, and make every change auditable. When racks, pathways, containment, clearances, and door perforation are treated as a single system, you can raise ports per RU and rack power without sacrificing serviceability – or your thermal headroom. That's what keeps day-two operations calm as the environment scales. Of course, hardware/software that automates port-level documentation and change detection, exposes capacity and work orders, and integrates with DCIM/ITSM ensures cable management stays accurate after every move/add/change, not just on day one.

When specifying or upgrading DC racks, prioritize platforms that balance density with serviceability: look for 1U fibre panels that support at least 144 LC (UHD) or 96 LC (HD) and offer independent, lockable sliding trays so technicians can work section by section without disturbing adjacent cords. Ensure there's real working room and disciplined routing – overhead patch frames (≥4U) help move dense patching out of active racks, while integrated cable guides preserve minimum bend radius and provide proper strain relief.

Make the physical layer visible with Automated Infrastructure Management that detects port-level changes and ties into DCIM/ITSM via APIs for work orders, audits, and exception alerts. Where copper is required, specify 1U, 48-port Cat6A panels with integrated rear cable management, hinged shutters, and an angled front to protect airflow and handling space. Finally, choose rack and cabinet envelopes that fit your room and containment strategy – availability in 19″ and ETSI formats, widths from 600 to 900 mm, and multiple heights and door types will keep layouts flexible without compromising thermal performance.

Total engineered solutions

Following a series of complementary acquisitions, UK engineering services provider BRUSH Group has become a leader in turnkey engineering solutions for the data centre sector.

FEW COMPANIES are able to show such an enduring appetite for excellence, innovation, and adaptability as BRUSH Group. With an impressive 150-year legacy of engineering achievement, this agile solutions provider has continually reinvented itself to meet the changing demands of the UK energy market.

Recent years have seen BRUSH evolve faster than ever, growing organically and by acquisition to offer complete custom packages that bring together unrivalled expertise and technical capability to deliver tailored, future-proof solutions for the data centre sector.

From bespoke design-and-build, authorising, consultancy to sustainable systems, fully accredited ICP generation connections, custom retrofits and ongoing maintenance, the BRUSH team are on hand to provide comprehensive end-to-end solutions. Paul Ingram, Strategic Partnerships Director for BRUSH Group explained the firm's unique set-up and how it benefits customers: "We have the core capability to design, manufacture, and deliver switchgear and transformers to the data centre industry, together with the expertise of our BRUSH Power Solutions business to design, build, test, and commission data centre projects. That includes HV distribution substations and the networks required to fully connect them to the grid, only possible because our team is a NERS accredited Independent Connections Provider (ICP)."

He continued: "Whatever the requirements of our clients in the data centre supply chain, our approach allows us to combine deep technical expertise, regulatory understanding and project delivery capabilities to put together the right team at the right time, especially important with the growing expectation for service providers to be able to deploy precisely the right resources at any given moment."

Comprehensive capability

The BRUSH Power Solutions team boasts a range of core capabilities:

- HV designers, who possess the knowledge and experience to provide the full design service for data centre electrical networks and their connections, including the ICP design requirements;
- A high level of knowledge and expertise to work from 11kV to 132kV systems, including protection and design schemes, and grid interface works;
- Project managers with a proven track record in the delivery of HV distribution networks for data centres and the requirements of their controlled environments;
- Clear coordination between the client, stakeholders, and all third parties, including the management of suppliers, subcontractors, and associated risks;
- Provision of appropriate commissioning engineers, who possess a full understanding of the HV testing protocols (e.g. primary injection, relay testing etc);

- Installation teams, trained fitters, certified HV cable jointers fully experienced in GIS/AIS substations, switchgear, transformers, and management of all the permits, RAMS, and energisation protocols, with a clear focus on system integration and redundancy being considered at all times within our design and build approach;
- SHEQ and compliance specialists, fully knowledgeable in the CDM Regulations 2015, DNO / ENA standards, and any data centre specific compliance requirements.

Solutions-first approach

BRUSH has also pioneered a unique 'solutions first' approach designed to ensure the best fit to meet each project's specific requirements, leveraging the combined expertise of its family of engineering solutions providers – Eta Projects, McGowan Infrastructure, Poise and Rybka – alongside its power solutions arm, and transformer and switchgear businesses.

And with vast experience working with Distribution Network Operators (DNOs), on everything from primary and protection design through to final test, commissioning and energisation, BRUSH is well placed to deliver critical infrastructure projects. BRUSH Transformers are designed and built by the firm's engineers to handle some of the world's most demanding applications. Equally, the BRUSH Switchgear range offers many benefits, including ease of maintenance, smaller footprints, reliability, and of course, high-speed performance.

The latest product from BRUSH Switchgear is CRYNO 12kV, SF6-free fixed pattern, indoor, metal-enclosed, vacuum switchgear. CRYNO offers impressive space-saving, easier installation and enhanced safety for operators. With advanced vacuum switching technology, CRYNO boasts superior arc interruption capabilities, enhancing safety and minimising downtime. And with a Notice of Conformity Certificate from the Energy Networks Association (ENA), CRYNO is a versatile choice for data centre applications.

BRUSH delivering solutions in UK data centres

New transformers with uninterrupted operations BRUSH was commissioned to install new transformers at a Prologis data centre. The project involved replacing ageing infrastructure while maintaining uninterrupted operations. With a focus on quality and compliance, the installation required careful planning to address technical challenges and meet the client's stringent requirements. A structured approach was taken, including rigorous quality checks on all transformer components. A phased installation ensured seamless commissioning, with thorough testing of insulation, connections, and alarm systems to guarantee optimal performance. The project was completed successfully, delivering improved operational efficiency for the client. All transformers were installed and commissioned to the highest standards, with comprehensive testing and performance validation.

Replacing old DC systems to improve efficiency

The existing DC Systems at a major UK data centre had become inefficient and complex to operate by modern standards. Originally designed for dual-feed redundancy, changes over time had compromised reliability. BRUSH was approached to replace them with two efficient 240KW DC systems while maintaining uninterrupted service for all data centre operations. The BRUSH team decided on a phased approach with the first DC system installed before supplies were migrated without service disruption.

Once the old system was decommissioned, the second system was deployed in its place. An innovative cable jointing method ensured a seamless transition throughout. The installation was a success, improving efficiency and reliability across the data centre. The transition was completed without service interruptions, and all alarms and remote monitoring were fully integrated.

Innovative solution secures data centre uptime BRUSH delivered an innovative, custom-engineered 132kV transformer solution for a data centre in Hemel Hempstead, ensuring maximum uptime and redundancy in supply. Working closely with the customer, BRUSH provided the electrical balance of plant equipment within the high-voltage substation connecting the new infrastructure to the grid. Redundancy and uptime were central to the design, leading to a dual LV winding configuration in an n+1 setup. This enabled load transfer between transformers and LV connections, providing 60MVA of apparent power stepped down to 11kV. The transformers were also designed to operate under unbalanced LV scenarios, with one winding carrying up to 40MVA and the other 20MVA. Featuring a compact layout with tank-mounted radiators, the solution delivered reliability, efficiency, and civil cost savings.

Keeping up with surging density, edge, and fibre

Today, operators are facing three simultaneous pressures: denser server racks in core data centres, the rapid rollout of numerous small edge sites, and the relentless densification of fibre in central offices and points of presence (POPs).

DAVID ROBIN, COUNTRY MANAGER / MANAGING DIRECTOR FRANCE, REICHLE & DE-MASSARI

THESE TRENDS share common drivers: explosive growth in AI-class workloads, a shift of compute toward the network edge, and a structural upgrade cycle in access and transport networks.

Let's take a closer look at these drivers, starting with AI. Racks are filling with higher-TDP GPUs and CPUs so operators can add capacity without expanding floorspace or waiting for scarce grid connections. The "typical" rack is creeping from 4–6 kW toward 7–9 kW, while the number of very high-density cabinets (15–29 kW and above) is rising. Uptime Institute ties this trend directly to newer high-powered processors and the wave of dense GPU servers (often ≈1 kW per unit), and notes that many operators are retrofitting halls because power and space are constrained in major hubs. At the same time, networks are migrating to 400/800G, which increases fibre terminations per rack and pushes cabling density even further.

Growth of the Edge is driven by applications for which the lowest possible latency, uptime, and locality are key, such as audio and video streaming, and low-latency gaming. Cloud providers are meeting this demand with local/edge zones that place compute closer to users. IDC expects roughly US$261B in edge spending in 2025 and ~US$380B by 2028. Mobility forecasts likewise show mobile data and 5G usage surging through 2030, strengthening the business case for distributing computing power and content nearer to where traffic originates.

On the fibre side, Europe's FTTH/B build-out illustrates the scale: coverage reached about 74.6% of homes across the "EU39" by late 2024, with take-up passing 53%. The EU's new Gigabit Infrastructure Act is intended to accelerate fibre and 5G rollout further. In mobile, 5G architectures (C-RAN/vRAN) shift bandwidth into fronthaul/xHaul, which is far

more fibre-intensive. At the same time, the core and metro domains are upgrading to 400/800G optics at scale. All of this converges in COs/POPs: more fibres to terminate, protect, and reconfigure in the same – or smaller – footprint.

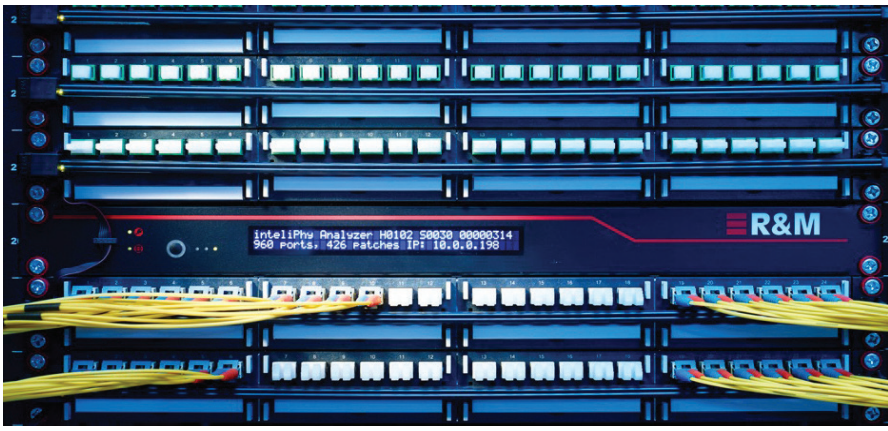
To summarize: growth of applications such as AI, streaming, and mobile traffic are raising compute and interconnect demands; power and real-estate constraints are forcing densification in existing facilities; and the economics of latency and data locality are pushing workloads outward to the edge. The result is higher-power racks in the core, a long tail of standardized micro/edge deployments, and CO/POP fibre plants that must be capable of handling thousands (or even tens of thousands) of terminations.

What this means for design and operations
Across environments, the common thread is modular density with clean operations: more ports per rack unit, faster and safer deployment, and better visibility of what is connected where. The physical connectivity platform should support very small form-factor connectors (CS/SN/MDC and MPO-16) to keep the footprint lean while remaining migration-friendly for 100/200/400/800G leaf-spine designs, and it should accommodate both Base-8 and Base-12 in a minimal footprint so teams don't need to keep swapping out hardware as optics strategies evolve. Covering the mixed-media reality of today's environments by combining fibre and copper within a single rack unit is especially useful when modernizing brownfield racks or bridging LAN/WAN domains.

An Automated Infrastructure Management (AIM) layer as a digital twin can guide each action and verify it at the port level with LEDs and RFID and sensors. Moves are recorded automatically, exceptions are obvious, and a clean API surface (SNMPv3/MQTT/REST) can feed DCIM and ticketing tools. The practical effect is fewer documentation gaps, faster troubleshooting, and better audit hygiene – without adding manual steps for technicians.

At the edge, a pre-engineered micro/edge data-centre approach accelerates rollout and reduces lifecycle cost by bundling rack, power, access control, and monitoring into repeatable blocks. Many edge rooms lack raised floors or sufficient building HVAC, so placing heat removal right next to the load and containing airflow is vital as AI-capable servers push rack densities in confined spaces. Because these sites are often unmanned, monitoring should tie into DCIM fabric so you can spot thermal excursions, or access and power anomalies before they turn into outages.

Central offices and POPs must land massive fibre counts in very little real estate. High-density aggregation frames benefit from tool-free, modular subracks with mirrored installation and patch areas,

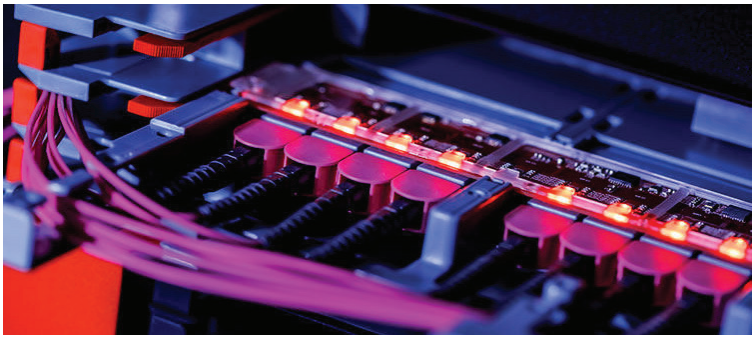


so splicing and maintenance can proceed without disturbing live circuits. Capacities may climb into the thousands of connectors per full-height rack, but physical routing must still protect bend radii and minimize handling risk. Maximizing terminations per rack unit—without locking into a single optical migration path—is ideal.

In short...

Rack power is trending upward with AI acceleration and denser leaf-spine fabrics; edge spending keeps expanding as enterprises push analytics and low-latency services closer to users and sensors; and fibre coverage—and therefore fibre terminations – continues to climb across regions. Designs that were "good enough" when racks drew 5–8 kW, edge sites were few, and ODFs hosted hundreds rather than thousands of connectors are showing their limits. A future-ready approach makes density practical to install and maintain and uses automation to ensure the physical plant matches what's on paper.

If you're evaluating systems, we'd recommend focusing your checklist on five areas: the density and optics roadmap (today's ports per RU and your 12–36-month plan, including VSFF and MPO-16 readiness); operations and integration (port-level guidance and analyser scale, and how events flow into your tools); thermals and power (expected rack kW and whether closed-loop or in-cabinet cooling is justified); deployment logistics (the value of pre-assembled panels and reduced on-site work); and fibre-plant evolution (ODF capacity per room and reconfiguration cadence). Address these, and you'll have a network fabric that's denser, easier to operate, and ready for the next wave of AI, edge, and fibre growth.



DCA Update – Issue 09 DCS

Steve Hone, CEO The DCA – Data Centre Alliance

THANK YOU to all of you that took the time to attend The DCA's annual conference – Data Centre Transformation 2025! The IET lecture theatre was cram packed, and we had to resort to providing additional seating!

The DCA team worked extremely hard to compile a varied and interesting conference programme which included an opening keynote from the Department of Science and Technology and presentations from Prospore AI Coaching, TreQ – Quantum Computing and Copenhagen Atomics! We were absolutely delighted with the finalised agenda for the day! It seems you all enjoyed the programme as the feedback has been tremendous!

So, here we are again already trying to work out how we can surpass this year and make 2026 an even better event! Watch this space and we'll come back with details soon! In the meantime, we have set the date next year please hold Tuesday 20 October 2026 in your diaries!

EVENTS FOR THE REMAINDER OF 2025

Data Centres Ireland, 19-20 November 2025 – Dublin

The DCA team are on stand 721 in the Conference area come and meet us. We are also hosting a 10X10 session within the conference programme towards the end of Day 1, this is followed by networking.

PowerEx Live London, 11 December 2025 – DCA Co-hosting

Park Plaza Hotel, Westminster Bridge, London.
The DCA will again be hosting a Data Centre Track within the conference programme.

The DCA team also available in the exhibition area to meet up and talk to you!

Data Centre Solutions – Issue 09

The DCA feature is comprised of articles from DCA Partners and Industry Experts.

Thank you to all the authors for providing their contributions.

- **Copper Cabling in Data Centres and Beyond**
  Hans Obermillacher, Business Development EMEA, Panduit provides us with an interesting insight into why copper cabling remains a vital component in enterprise and data center networks.

- **Why Liquid Cooling Must Be Driven by the IT Stack**
  You can't just bolt on liquid cooling and call it a day. To stop AI servers from melting down, cooling needs to be as smart and fast as the workloads themselves.
  Mark Acton & Venessa Moffat, The DCA Executive Board provide us with their thoughts.

- **The four techniques you need to know to cool Ai Data Centres –** Thank you to Alan Farrimond, VP Worldwide Accounts at Wesco – Alan believes the rise of AI is driving high-density demands that traditional air-cooling alone can't handle – as he explains the four critical methods teams should consider.

If you'd like to find out more about The DCA and how we support the sector and those working in it drop me an email, steveh@dcauk.org

Best regards,
Steve

The four techniques you need to know to cool AI data centres

The rise of AI is driving high-density demands that traditional air-cooling alone can't handle – as he explains the four critical methods teams should consider.

BY ALAN FARRIMOND, VICE PRESIDENT OF THE WORLDWIDE ACCOUNTS AT WESCO

THE GROWING need for high-density data centres to support AI workloads is driving more data centre teams to adopt liquid cooling rather than traditional air-cooling methods. But outside of large-scale hyperscalers, many data centre teams are still learning the different liquid-cooling methods that exist and how they can make sure deployment goes smoothly.

There are four base design options for liquid cooling to consider: traditional hot/cold aisle containment, rear-door heat exchangers, direct-to-chip cooling and immersion cooling. The latter three options outperform traditional air-cooling systems, which may be insufficient for cooling the power-intensive racks in high-density data centres.

Each option has its own unique performance ranges, deployment and maintenance demands that need to be considered.

Let's explore each of the liquid cooling options and what data centre teams should be thinking about so they can choose the right option for their project.

1. Traditional hot and cold aisle cooling
Traditional hot and cold aisle cooling is a widely used method in data centres. This technique involves arranging server racks in alternating rows with cold air intakes facing one cold aisle and hot air exhausts facing the opposite hot aisle.

The cold aisle is supplied with cool air from computer-room-air-conditioner (CRAC) and computer-room-air-handler (CRAH) units, while the hot aisle collects the heated air and returns it to the cooling units for re-cooling. This separation helps to prevent the mixing of hot and cold air, thereby improving cooling efficiency and reducing energy consumption. However, this technique is generally restricted to circa 15-20kW per cabinet.

2. Rear-door heat exchangers
This technology is commonly used by data centres that are switching to liquid cooling because it provides an efficient and complete cooling solution.

Rear-door heat exchangers sit on the back of a cabinet and capture hot air from IT equipment before it enters the white space. The captured heat is transferred via a coil to a chilled water source, and the heat is then discharged out of the back of the cabinet.

A benefit of rear-door heat exchangers is that they often can remove 100% of the heat generated by the server without requiring other heat-dissipating technologies. They're also room neutral. The temperature of the air discharged from the cabinet will be the same as the room's ambient temperature.

Rear-door heat exchangers generally support power densities up to 85kW to 90kW per rack. Some manufacturers state their technology can support up to 200kW, but these offerings are typically designed for specialised use cases or cabinets. When considering rear-door heat exchangers that claim to support above 90kW, verify that they're appropriate for the application and cabinet where they'll be deployed.

3. Direct-to-chip cooling
This cooling method, also known as direct liquid cooling, puts water directly on a heat sink or cold plate inside the equipment to remove heat at its source, before it is discharged as hot air.

There are two options for direct-to-chip cooling. The most common is single-phase cooling. Here, the coolant or water that has absorbed the IT equipment's heat is moved to a coolant distribution unit (CDU), which transfers the heat to a larger loop. The cooler fluid is then pumped back to the hardware as part of a continuous cycle.
In the two-phase option, the liquid that absorbs the heat is boiled off. The resulting vapor then condenses back into a liquid and cycles back through the system.

Direct-to-chip cooling typically supports up to 100kW per rack, although in some cases it can go as high as 120kW. A trade-off of this cooling method is that it only cools the chip, not the rest of the cabinet. This means that another cooling solution like rear-door heat exchangers needs to be used. As it is not a full-cabinet cooling solution, it's typically not the first choice as a cooling solution.

4. Immersion cooling
Fully submerging servers or racks in liquid is the final closed-loop liquid-cooling option.

This process uses an inert and non-conductive dielectric fluid in an immersion enclosure to absorb heat generated by servers, GPUs and other associated hardware. The heated fluid is then circulated from the enclosure to a cooling system where the heat is extracted, such as with heat exchangers or direct liquid-to-liquid cooling. Next, the cooled fluid is transferred back to the enclosure.

A drawback of immersion is that it hinders maintenance. If equipment needs to be repaired, for instance, workers need to lift it from the immersion liquid, often using a gantry, and then wait for it to dry. This has largely limited the use of immersion cooling to applications like bitcoin mining where the hardware doesn't experience regular change and maintaining equipment is not a high priority.

The maintenance pain point can be addressed with a two-phase option for immersion cooling. This approach uses a dialectic liquid that boils off as it captures the heat generated by the equipment. The vapor that is produced is then captured, condensed and cycled back through the system (or not, if maintenance needs to be done). However, the two-phase immersion option is still early in its adoption phase.

Other considerations before deploying
After a data centre team decides on a liquid cooling approach, keeping a few things top of mind during planning and implementation can help them avoid surprises.

First among them is to plan for the entire system, not just the technology.

Sometimes data centre teams charge ahead with building and acquiring a bill of materials for a liquid-cooling project before they understand what will be needed to make those materials work together. This can create roadblocks when they discover the installation contractor they hired isn't familiar with creating a secondary water or fluid loop, which is required to remove heat loads and is separate from the primary loop that cools the entire facility.

Addressing the day-two considerations for a liquid-cooling system is also important. Equipment like the CDUs used in direct-to-chip cooling systems are far less tolerant of contaminants than the CRAC and CRAH units that have been used in data centres for decades. Data centre teams now need to think about contamination that can happen at the microscopic level, such as by using additives like inhibitors that help prevent bacterial growth.

Teams should also be thinking about how they can standardize their liquid-cooling system if their scope extends beyond one data centre. This is where a supply chain partner with a wide reach and liquid-cooling know-how can be especially helpful. They can help create a basis for design that includes not only stamp-ready drawings but also installation and commissioning packages, essentially creating a playbook for every deployment. This can save significant time, effort and money compared to treating each liquid-cooling deployment as a bespoke project.

Keep cool and carry on

Growing power demands are making liquid cooling a necessity in today's data centres. By understanding the nuances of the different liquid-cooling options, data centre teams can choose the best cooling strategy for their facility and manage the risks that come with bringing liquids into white space.

Why liquid cooling must be driven by the IT Stack

Data centres have become the physical engines of AI, HPC, and real-time analytics.

WITH increasing power densities being deployment to support these platforms, thermal management has evolved from an important element to manage to a primary operational constraint. Power density increases mean that traditional air systems can no longer reliably or economically handle some of these higher loads; the industry consensus is clear that some form of liquid cooling is inevitable.

Yet what's less often discussed is how the delivery of liquid cooling should be controlled. Liquid cooling cannot just be bolted on to replace air or managed in the same way. To deliver its full potential reliably, efficiently, and sustainably, it must be driven by the requirements of the IT stack itself.

From facility-led to workload-led cooling

Historically, cooling has been facility-led: mechanical systems respond to environmental sensors, and the IT load operates totally independently to the capabilities or capacities of the thermal conditions that are being maintained. The environment is typically managed according to customer SLAs rather than the IT workload demand.

This works tolerably well for steady-state workloads, but AI training clusters, inference nodes, and high-performance compute create unpredictable, spiky loads that demand agility, and at scale can challenge cooling systems that are managed manually and purely by humans.

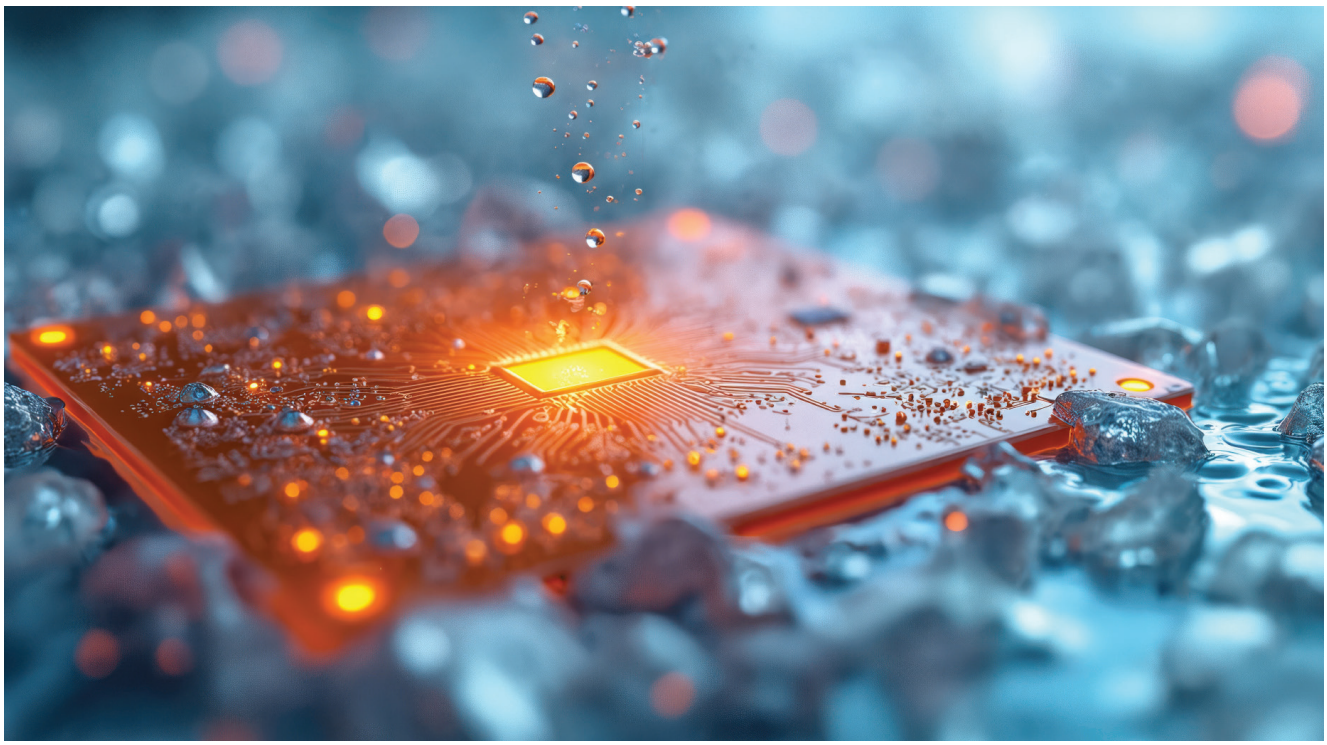
When a GPU cluster spins up from a quiescent state, temperatures can rise by tens of degrees in seconds. Waiting for facility sensors and cooling systems to react is too slow. Air systems have thermal inertia that allows a degree of flexibility, but liquid systems are more unforgiving. If the liquid isn't flowing fast enough when the heat arrives, it doesn't remove energy - it absorbs it, warms, and it stops being a coolant.

That's not a minor inefficiency; it's a design-breaking risk. In a tightly coupled liquid loop, thermal lag can cause localised hotspots, trigger component throttling, or even force system shutdowns. In extreme cases, the coolant itself can become the heat source.

The conclusion is unavoidable: pre-emptive cooling is required to start before the workload arrives.
Pre-Emptive Cooling: A Shift in Control Logic
To achieve this, control must move up the stack. Instead of the cooling system waiting to detect heat, it should anticipate it. The IT stack knows what's coming: job schedulers, orchestration layers, and AI workload managers all understand when and where compute demand will increase. That knowledge is the key to enabling pre-emptive thermal readiness.

Imagine a system where:

- The workload manager signals upcoming GPU activity to the cooling controller.



- The coolant flow rate and pump pressure ramp up before the heat is generated.
- Return-loop temperatures are monitored dynamically, allowing micro-adjustments in near real time.
- The 'DCIM' or orchestration layer tracks workload distribution and automatically throttles or redistributes tasks if a loop is nearing thermal saturation.

This is not science fiction; it's the logical next step in the integration of IT and facilities. The technical enablers already exist in modern APIs, telemetry systems, and edge-based control logic.

The physics of "too slow"

Liquid cooling is extraordinarily efficient, but only within its operational envelope. The rate of heat transfer depends on several factors: temperature differential, flow velocity, surface area, and specific heat capacity. If the flow slows down, the coolant's ability to absorb heat without excessive temperature rise diminishes rapidly.

In an ideal steady state, coolant enters cold, absorbs heat evenly, and exits warm enough to release energy through a heat exchanger. In a transient state (for example, when a new AI job starts and hundreds of kW of GPU heat load comes online within milliseconds) the coolant in contact with those hot components can reach equilibrium extremely quickly. Without pumps working to provide sufficient flow rate, that local section of liquid can become a stagnant thermal mass.

At that point, it's no longer transferring heat; it's retaining it. The liquid becomes part of the thermal problem.

This is why liquid cooling must be dynamic — responsive not just to average temperatures, but to predicted workload patterns.

Integration is everything

The future of data-centre cooling lies in control integration between IT operations and facility systems. Through APIs, the IT stack can publish telemetry on CPU/GPU utilisation, thermal forecasts based on machine learning, or AI training schedules capable of being pushed to the building management system (BMS). The BMS, in turn, can adjust valves, pumps, and chillers in real time. This is "chip to chiller" cooling optimisation.

This bi-directional communication creates a closed-loop control system where both IT and mechanical systems operate as one. It's not about facilities chasing IT heat signatures — it's about IT leading the dance.

The data-centric advantage

AI and HPC clusters already generate rich telemetry on utilisation, power draw, and temperature at motherboard and component level. If harnessed correctly, this data allows predictive cooling models to be built using machine learning. Over time, the system learns the thermal fingerprint of each workload type and adjusts coolant dynamics accordingly.

For example, a data-centre AI scheduler could "know" that a certain training job will hit 95% GPU load for four hours at a specified time. The control layer could pre-chill that loop, ramp up flow, and coordinate secondary systems like heat-recovery or free-cooling modules — achieving stability and energy efficiency simultaneously.

Currently the only way to manage the issues created by fluctuating workloads is to run the pumps moving the coolant around at full speed 100% of the time – which is hugely inefficient and will increase energy consumption on the cooling side significantly.

Because large pumps are increasingly being located in Cooling Distribution Units (CDUs) which are located in the data halls it might even be the case that those already guilty of gaming PUE figures may start to claim this energy spent on CDU pumps as 'IT load' to mask this issue.

Another New Paradigm

Liquid cooling represents a fundamental shift in data-centre physics, but unless it's also matched by a shift in control philosophy, it risks underperforming. The next generation of digital infrastructure must treat cooling as a workload-aware process, not a static utility.

Used properly this form of control could even reduce capital costs for data centre builds. If the system could anticipate when workloads should not be run due to high ambient data centre could be designed not to have to cope with extreme conditions which might only happen on a few days a year. These workloads could be shifted to a data centre with capacity and good energy pricing temporarily, and if the DC does not have to handle the extreme conditions, it would reduce the heat rejection requirements for ambient temperature extremes. Therefore it would reduce the capital costs of installed mechanical equipment, as well as allowing more of the incoming power to be dedicated to the IT load by sacrificing a limited number of hours per year of productive data centre time.

In short:

- Cooling must be predictive, not reactive.
- Control must be IT-driven, not facility-led.
- Liquid cooling must flow at the speed of data.
- Used properly energy efficiency could be improved, and both capital and operating costs reduced.

Because when the liquid stops moving fast enough, it stops being a coolant. In the age of AI, that's not a margin for error anyone can afford.

Critical Operations Rely on Dependable Power Control

DSE335 MKII
Auto Transfer Switch Control Module

Reliable power control is essential in critical operations where downtime is not an option. Whether in data centres, hospitals, or industrial environments, even the slightest power disruption can lead to serious safety, financial, or operational consequences. Robust, intelligent control systems ensure continuous performance, system protection, and complete confidence in essential infrastructure performance.

The DSE**335 MKII** is an advanced Auto Transfer Switch (ATS) controller designed for seamless power switching between two different power sources (S1 / S2). Ideal for mission-critical applications, it offers configurable inputs/outputs, comprehensive monitoring, and a close transition feature for no-break power transfer. Its robust design and flexible programming make it a trusted choice for reliable power continuity.

To find out more please visit **www.deepseaelectronics.com**

Made in **UK**