DCS DATACENTRE SOLUTIONS

DEVELOPING DIGITAL INFRASTRUCTURE IN A HYBRID WORLD

ISSUE VII 2024

HTTPS://DATACENTRE.SOLUTIONS

Cloud | Design + Optimisation | Energy Management | Infrastructure + Asset Management
Hosting + Colocation | Infrastructure + Hardware | Power & Cooling

AN ANGEL BUSINESS COMMUNICATIONS PUBLICATION

StratusPower™

Future-proof your data centre and experience tomorrow's power protection technology today

Green UPS

DARA TECH

SWISS MADE

Centiel's **StratusPower™** is the ultimate power protection solution for today's dynamic data centre environment.
With unmatched availability, reliability, and efficiency, StratusPower ensures seamless operations and business continuity, minimizing the risk of downtime.

Our innovative DARA design delivers unparalleled scalability, eliminates single points of failure, and provides a fault-tolerant architecture. From compact 10 kW modules to robust 62.5 kW options, the UPS meets a range of power requirements with the ability to scale up to an impressive 3.75 MW.

centiel
continuous power availability

www.centiel.com

VIEWPOINT

By Phil Alsop, Editor

How innovative is innovative?

THE CHALLENGES facing the data centre industry as it seeks to respond to the twin pressures of AI and sustainability (seemingly incompatible bedpersons) tend to generate a universal response: 'time to innovate'. Those of you who have watched a selection of the DCS video interviews available on our website may well have bumped into one or more where a favourite final question asks: 'Can the industry continue to adopt an evolutionary approach to change, or do we need something of a technology revolution if we are to successfully address the sustainability imperative?" I can't recall asking the exact same question when it comes to AI, but various interviews (and plenty of our magazine articles) do address head on the likely impact AI will have (is having already) on the industry – again framed in terms of the need to do things in a radically different way, or not.

No one who follows the data centre industry closely could fail to notice that the industry has a great track record for innovation, and that the pace of such innovation has picked up noticeably in recent years. Nevertheless, there are plenty of folks I talk to who bemoan the fact that data centres are still in the main built with traditional materials, and in a traditional design, using largely traditional power and cooling solutions, and often to meet 'traditional' standards and/or specifications which have hardly changed since the days when the Romans needed a place to store all that census data they were so fond of collecting...

The mission critical nature of the data centre is often cited as the reason the industry remains somewhat suspicious of unproven, leftfield innovations. And, yet, it increasingly seems that this part of the overall digital infrastructure sector moves at a pace markedly slower than others. The momentum behind AI is the

most obvious example of how the IT industry moves at considerable pace. And not so long ago the cloud was similarly game-changing. Have we had such a comparable moment, or series of moments, in the data centre sector?

There are some start-ups in the data centre sector who are looking to re-write some of the rules – plenty of our video interviews put them in the spotlight – but overall, we are waiting for that industry disruptor to arrive with something so different that the rules of the industry will be changed radically, and certainly for the better in terms of sustainability.

For how much longer can the data centre industry – the enabler of all manner of dynamic, agile digital applications – afford to be somewhat less than dynamic and agile itself? Maybe the ongoing and apparently worsening power infrastructure pending crisis (if you can have a worsening, pending crisis!) might just force the step change required?

CONTENTS

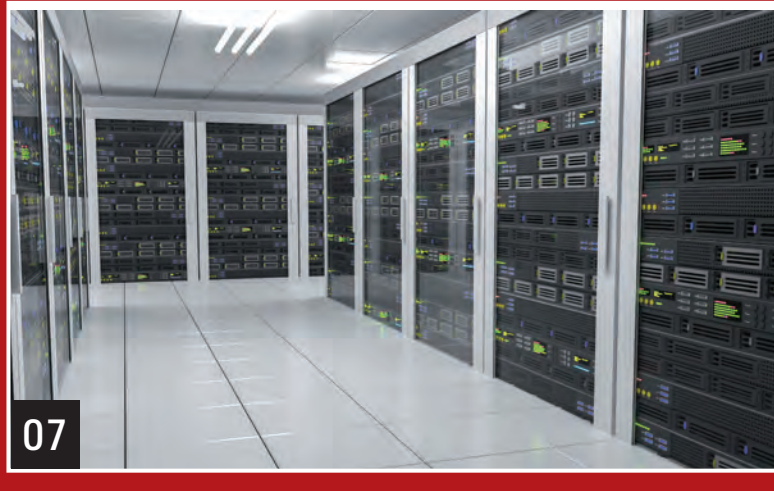
ISSUE VII 2024

DCS DATACENTRE SOLUTIONS

30

DCA News

NEWS



07

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

Marketing & Logistics Executive
Eve O'Sullivan
+44 (0)2476 823 123
eve.osullivan@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 6 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)

Leading data centre companies partner with Open Compute Project Foundation and WJE

Field Demonstration conducted on the application of low-embodied carbon concrete for data centers, targeting a significant reduction in greenhouse gas emissions.

THE OPEN Compute Project Foundation (OCP), the nonprofit organization bringing hyperscale innovations to the data center industry, announces a new collaboration to test development and deployment of low-embodied carbon concrete or “green concrete.” OCP is facilitating a collaborative effort among leading technology innovators Amazon Web Services (AWS), Google, Meta, and Microsoft to drive the adoption of low-embodied carbon concrete in data center construction.

This collective endeavor, which aligns with the objectives in a previously issued industry call to action, is conducting research through Wiss, Janney, Elstner Associates, Inc. (WJE) to test the application of low-embodied carbon concrete for data center floors, targeting a significant reduction in greenhouse gas emissions to greater than 50% per cubic yard.

“By aligning OCP Community’s ability to impact the data center building material supply chain, this demonstration project will support the creation of sustainable and scalable data center buildings. Low-embodied carbon concrete represents a significant advancement in environmentally responsible building practices. By reducing the carbon footprint associated with concrete production, we can make a tangible impact in mitigating the data center industry’s environmental impact.

This demonstration will provide valuable insights into the performance and viability of low-embodied carbon concrete, paving the way for its widespread adoption throughout the industry,” said George Tchaparian, CEO for the Open Compute Project Foundation. While numerous emerging technologies exist to achieve production of low carbon

concrete, adoption has not yet scaled. This proactive and collaborative demonstration project is an important step towards de-risking these new materials that will help decarbonize concrete. The data and “hands-on” experience provided by a demonstration project supports informed decision making for those wanting to specify these materials in the future.

During this first-of-a-kind demonstration project, the teams completed a series of slab-on-ground placements with four different concrete mixtures with decreasing global warming potential - with the lowest carbon mixture achieving a greater than 50% reduction in carbon impact as compared to today’s typical concrete. The mixture formulas use supplementary cementitious materials, and an alternative cement derived from raw materials and manufacturing processes that are commercially available today but have yet to be used broadly to this degree given perceived and technical implementation risks.

To measure and observe performance of the concrete in practice, the implementation team developed a comprehensive test plan, including extensive laboratory and field testing performed by industry experts and construction practitioners. The results of these tests will be used to better understand areas of risk, possible mitigation strategies, and ways to further optimize the mixtures to deliver concrete meeting data center structural performance requirements. These findings will be coalesced into a final whitepaper and made available to the public via OCP to inform other efforts to adopt new concrete technologies.

The demonstration event was held on August 8, 2024, at the WJE facility in Northbrook, IL, and in attendance were senior engineering representatives from AWS, Google, Meta, Microsoft, senior staff from the Open Compute Project Foundation, as well as representatives from the White House Council on Environmental Quality, White House Office of Science & Technology Policy, Rocky Mountain Institute (RMI), Natural Resources Defense Council, Urban Land Institute, US Department of Energy (DOE), US Environmental Protection Agency (EPA), as well as representatives from local and state transportation and tourism agencies, and members of the academic community.

The actions posited by the Open Letter and now this demonstration project are intended to create aggregated demand for low carbon concrete, which in turn will create a market force to drive innovation. Through this open-source approach, our organizations and others can gain confidence in new concrete technologies.

Data centre industry plans for major changes

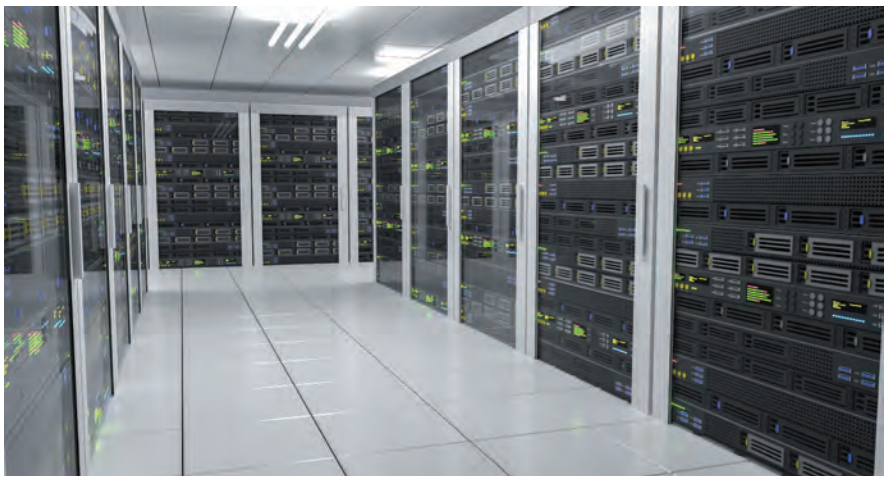
Uptime Institute has released its 14th Annual Global Data Center survey. The increased demand for digital services with both volume and compute intensity are challenging the power and cooling capabilities of much of the existing infrastructure.

TO MEET the rising demand, data center operators and their IT clients are both investing and innovating more in their IT services and facilities and using external services. The effectiveness and impact of these investments will serve to shape the industry in the years ahead.

"Our data shows operators poised for major changes ahead on multiple levels," said Andy Lawrence, executive director of research, Uptime Intelligence. "In 2024, we see the challenges of increased demand impacting power and cooling capabilities of existing facilities and the need for further investment to keep up with the demand. At the same time the industry needs to focus on continued staffing challenges to match capacity growth. And regulatory requirements are here and cannot be dismissed."

Now in its 14th year, Uptime Institute's annual survey remains the most comprehensive and longest-running study of its kind.

The findings of this report highlight the practices and experiences of data center owners and operators in the areas of resiliency, sustainability, efficiency, staffing, cloud and artificial intelligence. It provides detailed insights into the digital infrastructure landscape and a view into its future trajectory.



Key findings from the 2024 report include:

- Fewer than one half of data center owners and operators are tracking the metrics needed to assess their sustainability and/or meet pending regulatory requirements.
- Most operators recognise the benefits of AI and its potential. But despite many operators planning to host the technology, trust in AI for use in data center operations has declined for the third year in a row.
- The frequency and severity of data center outages remain mainly unchanged from 2023 or show small improvements. Operators are countering increases in complexity, density and extreme weather with investment and good management practices.
- Average server rack densities are increasing but remain below 8 kilowatts (kW). Most facilities do not have racks above 30kW, and those that do have only a few. This is expected to change in coming years.
- Enterprises continue to meet their IT needs with hybrid architectures. More than one half of workloads (55%) are now off-premises, continuing the gradual trend of recent years. Many continued to maintain their own data centers.
- Staffing challenges have neither improved nor worsened from 2023. More effort is needed to expand labor pools and skillsets to match the pace of capacity growth.
- Average PUE levels remain mostly flat for the fifth consecutive year, but this conceals advances in newer, larger facilities.

In 2024, we see the challenges of increased demand impacting power and cooling capabilities of existing facilities and the need for further investment to keep up with the demand. At the same time the industry needs to focus on continued staffing challenges to match capacity growth

Playing the hero and the villain

AI can limit the environmental damage it's responsible for causing, says investment bank.

AI'S POTENTIAL to enhance data centre sustainability has been overshadowed by the undeniable environmental impact of this technology.

Data centres, expected to account for 6% of the world's carbon footprint by 2030, are undergoing a period of transformation, driven by the rise of AI and the pressing need to combat climate change. With such rapid growth comes unforeseen environmental impacts, highlighting the significance of the application of AI technologies in optimising energy use.

It is undeniable that the data-intensive workloads generated by AI will see power consumption soar to unprecedented levels. However, the technology itself can help develop the next generation of data centres that are both high-capacity and more sustainable.

According to Julien Deconinck, Managing Director at DAI Magister, environmental concerns are driving the development of innovative AI solutions that optimise energy usage in data centres, while reducing operating costs.

Deconinck said: "Over the next five years, the amount of data generated will surpass the total produced in the past decade, necessitating a significant expansion of storage capacity in data centres worldwide. Another key factor contributing to this rising energy demand is the escalating computational power required for AI training, which is doubling every six months.

"Tech giants, recognising the scale of the problem and their significant contribution to it, are racing to mitigate the environmental impact of their operations. These companies face mounting pressure to reduce their carbon footprint and meet neutrality targets.

"Most data centres aim to operate in a 'steady state', striving to maintain consistent and predictable energy consumption over time to manage costs and ensure reliable performance. As a result, they're dependent on the local electricity grid, where outputs can fluctuate significantly. AI-driven solutions offer enormous potential to address these challenges by optimising energy usage and predicting and managing demand more effectively.

"Integrating renewable energy sources like solar and wind into the grid can improve data centre sustainability, but this presents challenges due to their variable availability. AI addresses this by forecasting renewable energy availability using weather data and predictive analytics. This enables data centres to shift non-critical workloads to peak renewable energy production periods, maximising the use of clean energy and reducing reliance on fossil fuels.

"When assessing the efficiency of a facility, the power usage effectiveness (PUE) measure serves as a crucial metric for indicating output. By monitoring and adjusting operational parameters in real time, AI sensors autonomously adjust power supply voltages, reducing consumption without compromising performance.

"AI algorithms analysing usage patterns and optimising workload distribution further reduce this energy waste associated with inadequate server management and inconsistent allocation. The optimisation of computing resources in data centres minimises the need for, and use of, excess capacity, both lowering operating costs and maximising performance capabilities."

AI can also pre-empt system issues that can lead to breakdowns or long-term disruption.

"AI sensors are facilitating predictive maintenance by analysing real-time data to detect anomalies or deviations in consumption patterns. Once identified, AI systems alert the issue to operators, preventing the activation of energy-intensive emergency cooling systems.

"Integration of AI sensors is further beneficial in thermal modelling, enabling dynamic adjustments to systems, accounting for high-intensity computing tasks and external temperature fluctuations by predicting potential hotspots within the facility, based on data collected."

Deconinck concluded: "Together, AI and green technologies are set to revolutionise data centre operations by allowing them to manage larger capacities while reducing their carbon footprint. This not only supports sustainability objectives but also safeguards the transition to low-carbon, high-capacity data centres as the demand for data storage and processing continues to surge brought about by the rise of AI."

Synergy identifies the world's Top 20 metro markets for colocation

New data from Synergy shows that just twenty metros account for 60% of the world's current colocation market.

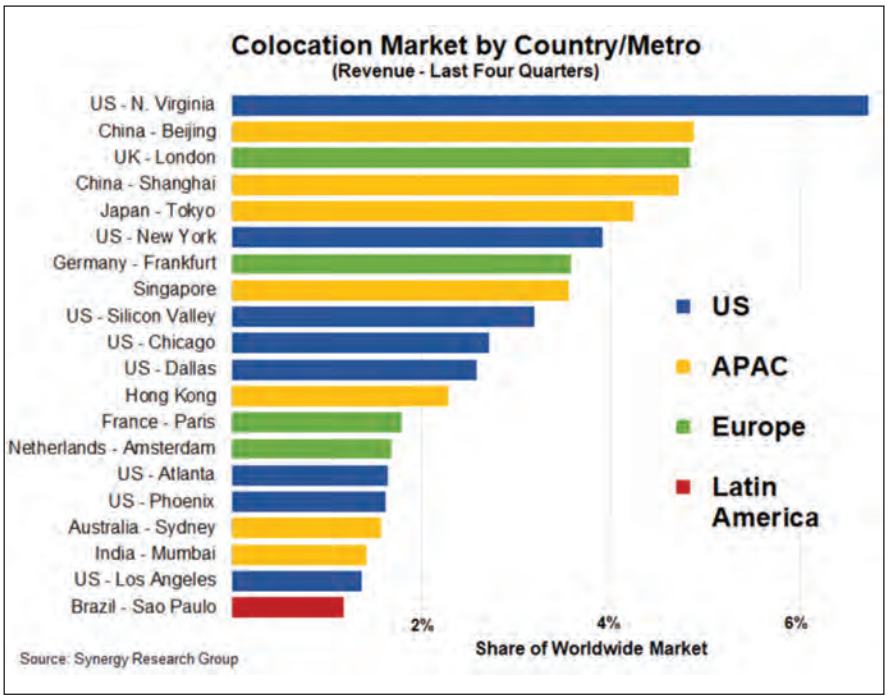
NORTHERN VIRGINIA is the single biggest market, accounting for almost 7% of the total. It is followed by Beijing, London and Shanghai, each accounting for around 5% of the total, and then Tokyo at 4%. Of the top 20 markets, eight are in the US, seven in the APAC region, four in Europe and one in Latin America.

After the top 20, the next twenty largest state or metro markets account for another 13% of the market, with the US and Europe featuring more prominently in that batch. Northern Virginia is a unique market. The area was foundational to the early years of Internet infrastructure and has remained a center of gravity for data center activity, benefitting from a rich networking ecosystem, a business-friendly environment and a good supply of low-cost power.

Beyond Northern Virginia, the other leading colocation markets mainly reflect the world's leading economic hubs. Outside of the top forty markets, some metros are growing extremely rapidly including Johor, Jakarta, Chennai, Lagos, Rio de Janeiro, Johannesburg and Queretaro.

The research is based on Synergy's in-depth quarterly tracking of colocation markets, including both retail and wholesale segments. Synergy provides quarterly revenue data for almost 300 colocation companies, with breakouts for over 50 countries and 85 individual metro markets. There is a strong local and national aspect to colocation, but on a worldwide basis the leading companies are Equinix, Digital Realty, NTT, China Telecom, CyrusOne and GDS.

"Proximity to customers is a key driver of the colocation market, so data centers tend to be located in metros that have a large concentration of companies and economic activity. That will not change but what we are seeing is the increasing emergence of many metros in countries that have smaller economies, but which are growing rapidly," said John Dinsdale, a Chief Analyst at Synergy Research Group.



"Countries like Malaysia, Indonesia, Nigeria and South Africa are all interesting opportunities, while in Latin America the market has developed beyond the historic focal point of Sao Paulo, and we're now seeing high growth in metros like Santiago, Rio de Janeiro and Queretaro. It is also the case that increasingly there are power or geographic constraints limiting growth in some traditional metros, pushing growth opportunities into neighboring regions. These emerging metro markets will not reach the scale of a Singapore or a Frankfurt any time soon, but they will increasingly reshape the evolving colocation market."

The research is based on Synergy's in-depth quarterly tracking of colocation markets, including both retail and wholesale segments. Synergy provides quarterly revenue data for almost 300 colocation companies, with breakouts for over 50 countries and 85 individual metro markets

Data centre market to grow by 30%

Over the past decade, the once-niche piece of IT infrastructure- data centers- has become the talk of the investment world and a massive revenue driver for tech giants.

AMAZON WEB SERVICES generate almost 20% of the ecommerce giant`s revenue today. At the same time, Nvidia, the supplier of parts for modern cloud storage and processing data centers, has grown into the world`s third-largest tech company. The surging use of AI technologies, which require significant computing power and storage, has only fuelled the data center boom, which is expected to continue in the following years.

According to data presented by Stocklytics.com, the global data center market is expected to grow by 30% and hit over $430 billion value by 2028. Network Infrastructure to Reiman the Largest and Fastest-Growing Segment The widespread adoption of cloud computing has significantly transformed the data center landscape.

While it has reduced the number of organizations operating their own data centers, it has also spurred a surge in the global data center count. This is largely due to the aggressive infrastructure expansion by major hyperscalers like Microsoft Azure, Google Cloud Platform, and Amazon Web Services to cater to their growing customer base. Other tech giants, like Nvidia, have also reaped the benefits of this trend, with its GPUs becoming an integral part of modern data centers.

The surging need for data center solutions and technology has helped the entire market grow revenues by almost 40% since 2016. However, the market projections for the following years are just as optimistic.

According to a Statista Market Insights survey, the global data centers market is expected to gross over $344 billion in revenue this year, $15 billion more than in 2023. The entire market is projected to grow by a CAGR of 6.5% in the following years, resulting in a market volume of $438 billion by 2028. Most of that value will come from network infrastructure, the market`s largest and fastest-growing segment.

Statista expects the network infrastructure to bring in $256.1 billion in revenue in 2028, almost 30% more than this year. Data center servers follow with $120 billion in revenue and a 24% growth in this period. Although far behind in revenue, the data center storage segment will also see double-digit growth, with revenue rising by 22% to $62 billion in the next four years.

The Statista data also gave an interesting insight into the biggest drivers of data center market growth and which countries have the highest share in total revenue.

Statistics show that GDP growth was the single largest market driver, with a 3.2% share in total market growth last year. The impact of the Russia-Ukraine war, which led to increased demand for data center services due to geopolitical tensions and cybersecurity concerns, followed, with 2.3% share in total market growth. Technology adoption and innovation also played a significant role, with a 1.9% share.

In global comparison, the United States will remain the single largest player in the global data center landscape, generating nearly $100 billion, or roughly 30% of total market revenue in 2024. This figure is expected to grow by 25% and hit over $125 billion by 2028.

AI - peril or promise?

Latest BCS industry survey highlights key challenges around AI.

THE CRUCIAL SPECIALIST data centres facilities needed to support the growing proliferation of artificial intelligence (AI) within global industry are significantly impacting power and water usage, presenting the sector with both challenges and opportunities for sustainability and resource management. This is according to the latest independent industry survey (number 28), commissioned by BCS, (Business Critical Solutions) the specialist services provider to the digital infrastructure industry, which captures the views of over 3000 senior data centre professionals across Europe, including owners, operators, developers, consultants and end users.

Over four-fifths of survey respondents reported experiencing an uplift in demand as a direct result of AI over the past year. However, 85% believe that the pace of widespread adoption of AI is currently being restricted by the lack of available power and facilities tailored for AI workloads.

There is however positive news as the sector continues to use AI to benefit from greater efficiencies, cost savings, driving improvements in services with the expected impact of AI on the operational side for data centres expected to be very positive. Over 65% of respondents reported that their organizations are regularly using generative AI, nearly double the percentage from their 2023 survey and around 90% of respondents expect their data centres to be more efficient as a direct result of AI applications.

James Hart, CEO at BCS, said: "There is real concern that the pace of AI adoption may be restricted by the ability of the market to deliver sufficient supply of data centres to house it. "AI data centres are power-intensive by nature, primarily due to the high computational demands of AI workloadsand the energy consumption is driven not only by the need to power the servers but also by the cooling systems required to maintain optimal operating temperatures. The environmental impact of this power usage is significant with many data centres are still reliant on non-renewable energy sources despite efforts to transition to renewables, with the pace of AI advancement often outstrips these sustainability initiatives. In addition to power usage, AI data centres also have a substantial impact on water resources used for cooling.

'The interplay between AI data centres and resource usage necessitates innovative approaches to mitigate environmental impacts. AI itself can be leveraged to enhance the efficiency of data centres with algorithms tooptimise energy use by predicting cooling needs, managing workloads more efficiently, and reducing idle times for servers. Predictive maintenance powered by AI can also prevent equipment failures, thereby reducing the need for excessive cooling.

"In conclusion, while AI data centres are indispensable for the continued advancement of artificial intelligence, their impact on power and water usage poses significant environmental challenges. Addressing these issues requires a multifaceted approach that includes transitioning to renewable energy, adopting innovative cooling technologies, and leveraging AI for operational efficiency. As the demand for AI capabilities continues to grow, so too must our efforts to ensure that this growth is sustainable and responsible, balancing technological progress with environmental stewardship."

EU electricity generation greener than ever before

The generation of clean electricity in Europe is setting records. In the first half of 2024, renewables made up more than 50% of all power generation in Europe while nuclear provided a stable share of 24% - according to latest figures from Eurelectric's electricity data platform. Demand for power, however, remains low due to sluggish growth, deindustrialisation and mild weather. Stimulating demand for electricity will be paramount to ensure continued investments in clean generation.

Europe's power generation is decarbonising at unseen pace. The latest figures from Eurelectric's Electricity Data Platform, ELDA, show that 74% of electricity produced in the EU in the first half of 2024 came from renewable and low-carbon energy sources. This is a significant increase compared to the 68% share in 2023. The main reasons behind this remarkable result were an unprecedented influx of renewables on the grid combined with the stabilisation of the nuclear fleet.

"The pace of change is impressive. These figures document that the decarbonisation efforts of electricity companies are years ahead of any other sector" - said Secretary General of Eurelectric, Kristian Ruby.

While the numbers on the supply side are promising, the same cannot be said for electricity demand. In the first half of 2023 power demand in the EU decreased by 3.4% compared to same period in 2022 and has continued to remain low in 2024 - 2.6% lower than in H1 2022. This trend is mainly due to industry relocating abroad, warmer temperatures, energy savings and slow economic growth.

"Years of stagnation in electricity demand have now turned into a regular decline. Policymakers must urgently support the uptake of electricity to provide the necessary investment signals for clean generation" – added Kristian Ruby.

"An Introduction to Data Centre Heat Reuse," a new White Paper from EUDCA

Nearly 100% of the electricity consumed by IT equipment is dissipated as waste heat, says a new white paper published by the European Data Centre Association (EUDCA) Technical Committee. Whilst the quantity of that heat can vary according to the computational load and ambient conditions, it could prove to be a useful resource from both an economic, and energy and sustainability standpoint.

"AN INTRODUCTION to Data Centre Heat Reuse," says author Alessandro Zerbetto, "is written to provide a useful and accessible overview of what heat reuse is, some of its main applications, and how heat can be recovered from a data centre. It also provides an introduction to current legislative and regulatory initiatives that support or mandate exhaust heat recovery schemes."

Making the point that generally, current data centre waste heat temperatures are not high enough for many use cases without the addition of systems to further elevate recovered heat temperature, a well-designed heat reuse system can not only improve the energy efficiency of a facility, but also reduce the overall energy consumption and greenhouse gas emissions associated with traditional heating and cooling methods.

Quoting Euroheat and Power, the white paper states the EU's total forecasted heat demand by 2025 will be 1850 TWh/ year. Waste heat from Europe's data centres could make a major contribution to meeting this demand, providing 221 TWh/ year or 12 percent of the total requirement.

According to the report there are 997 data centres in the EU-28 located within 2km of a district heating network and generating more than 75 TWh/ year of accessible excess heat.

"Reusing heat from data centres is not just an environmental imperative," says Marie Chabanon, Deputy Chair of the EUDCA Technical Committee, "it is also an economic and social opportunity, with the power to transform energy consumption into sustainable innovation."

However, one of the biggest challenges is achieving balance between the heat demand of the environment with the availability of waste heat from the data centre.

Data centres tend to produce heat at times when most cooling is required, i.e., when outdoor temperatures are hot. In winter, when European temperatures are usually at their lowest, data centres avail of free cooling availability to reduce facility power consumption and energy costs and also exhaust heat output.

Other challenges to heat reuse include the physical distance between data centres and the applications requiring the heat, and the capital and operating expenses for infrastructure including heat exchangers, piping and control systems, and in some cases equipment such as water-to-water heat pumps. It Is unclear whether these costs should be funded by the data centre operator or the waste heat consumer.

An Introduction to Data Centre Heat Reuse was written and reviewed by members of the EUDCA Technical Committee, including Alessandro Zerbetto of Vertiv (author), Tanja Gutgesell and Benoit Ploux (Vantage Data Centres), Noah Nkonge and Billy McHallum (Equinix), Marie Chabanon and Linda Lescuyer (Data4), and Steven Parker (Global Switch).

Available for sales and rental

CRESTCHIC

THE **TRUSTED** NAME FOR **LOAD TESTING**

200kW Data Centre Loadbank: Built for Integrated Systems Testing

Simulate full load operation, power outages, and cooling system failures

Verify PDU operational readiness during commissioning

Closely mimic server heat generation to test cooling systems

Control up to 20 units from one controller

- Small footprint, stackable for easy storage
- AC load bank resistive-only (200kW)
- Voltage range: 400V or 415V
- 50/60 Hz Operating frequency range
- 3 Phase
- Forklift pockets for lifting
- Sheet steel construction with IP55 rating
- Low Delta T
- Low noise
- Internal or external auxiliary supply operation

www.crestchicloadbanks.com

sales.uk@crestchic.com | hires.uk@crestchic.com | service.uk@crestchic.com

Sales & Rental Offices in : UK, North America, Singapore, Germany, Belgium, France, China and UAE

Gartner 2024 Hype Cycle highlights Developer Productivity, Total Experience, AI and Security

The 25 disruptive technologies to watch on the Gartner, Inc. Hype Cycle for Emerging Technologies, 2024 fall into four key areas: autonomous AI, developer productivity, total experience, and human-centric security and privacy programs

"GENERATIVE AI (GenAI) is over the Peak of Inflated Expectations as business focus continues to shift from excitement around foundation models to use cases that drive ROI," said Arun Chandrasekaran, Distinguished VP Analyst at Gartner. "This is accelerating autonomous AI. While the current generation of AI models lack agency, AI research labs are rapidly releasing agents which can dynamically interact with their environment to achieve goals, although this development will be a gradual process.

"Even as AI continues to grab the attention, CIOs and other IT executives must also examine other emerging technologies with transformational potential for developers, security and customer and employee experience and strategize how to exploit these technologies in line with their organizations' ability to handle unproven technologies," said Chandrasekaran.

The Hype Cycle for Emerging Technologies is unique among Gartner Hype Cycles because it distills key insights from more than 2,000 technologies and applied frameworks that Gartner profiles each year into a succinct set of "must-know" emerging technologies. These technologies have potential to deliver transformational benefits over the next two to 10 years (see Figure 1).

Four Themes of Emerging Technology Trends
Autonomous AI: AI's fast evolution is producing autonomous AI systems that can operate with minimal human oversight, improve themselves and become effective at decision-making in complex environments. These advanced AI systems that can perform any task a human can perform are beginning to move slowly from science fiction to reality. These technologies include multiagent systems, large action models, machine customers, humanoid working robots, autonomous agents, and reinforcement learning.

Boost developer Productivity
Developer productivity is about more than writing code quickly. It's influenced by developers' effective

communication and collaboration, and their feeling of energized focus, full involvement and enjoyment (being in the "flow state").

"Technology is revolutionizing the way developers design and deliver software, making them more productive than ever," said Chandrasekaran. "It's ensuring they deliver higher-quality products quickly while maximizing gains by improving developer satisfaction, collaboration, and flow."

Emerging technologies enabling developer productivity include AI-augmented software engineering, cloud-native, GitOps, internal developer portals, prompt engineering and WebAssembly.

Empower With Total Experience

Total experience is a strategy that creates superior shared experiences by intertwining customer experience, employee experience, multiexperience and user experience practices. It uses technology to address critical interactions, empowering both customers and employees, with the goal of driving greater confidence, satisfaction, loyalty and advocacy. Technologies to assess include digital twin of a customer, spatial computing, superapps and 6G.

Deliver Human-Centric Security and Privacy

Organizations will become more resilient by using security and privacy techniques that create a culture of mutual trust and awareness of shared risks between teams.

"Security practices too often rely on the premise that humans can behave in a completely safe and secure way. But when employees must make a choice between security and business delivery, they often choose business delivery, sometimes bypassing too-stringent security controls," said Chandrasekaran.

"Human-centric security and privacy weaves a tight security and privacy fabric into the organization's digital design."

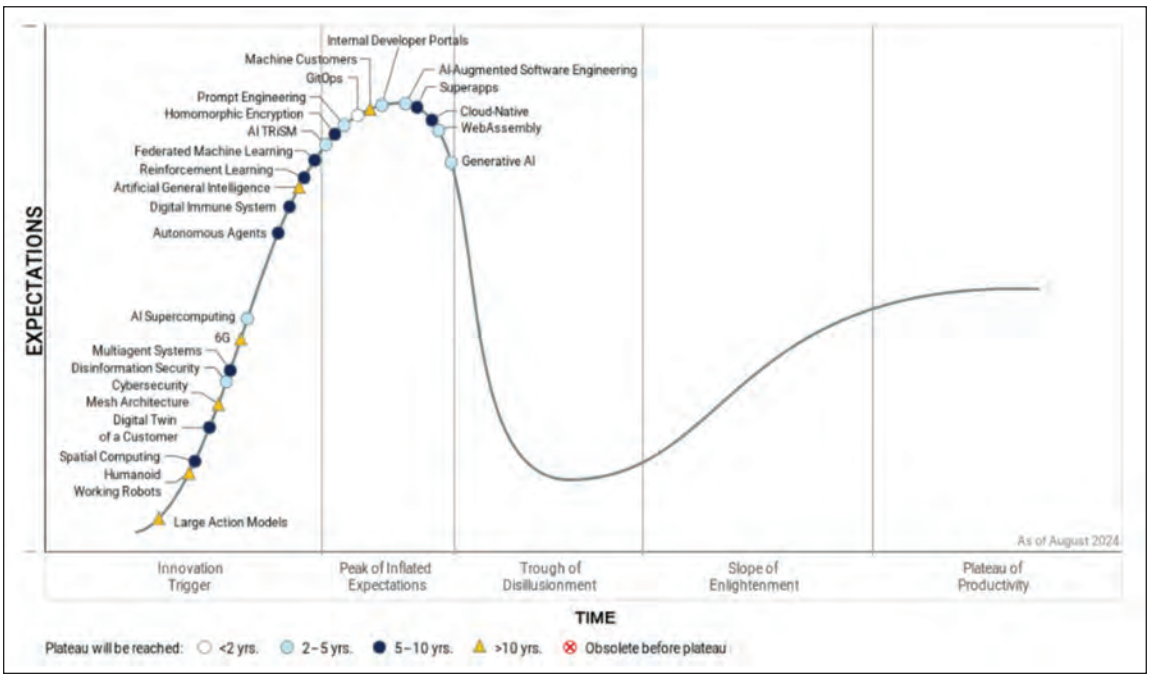
Emerging technologies supporting human-centric security and privacy include AI TRiSM, cybersecurity mesh architecture, digital immune system, disinformation security, federated machine learning and homomorphic encryption.

Worldwide spending on artificial intelligence forecast to reach $632 billion in 2028

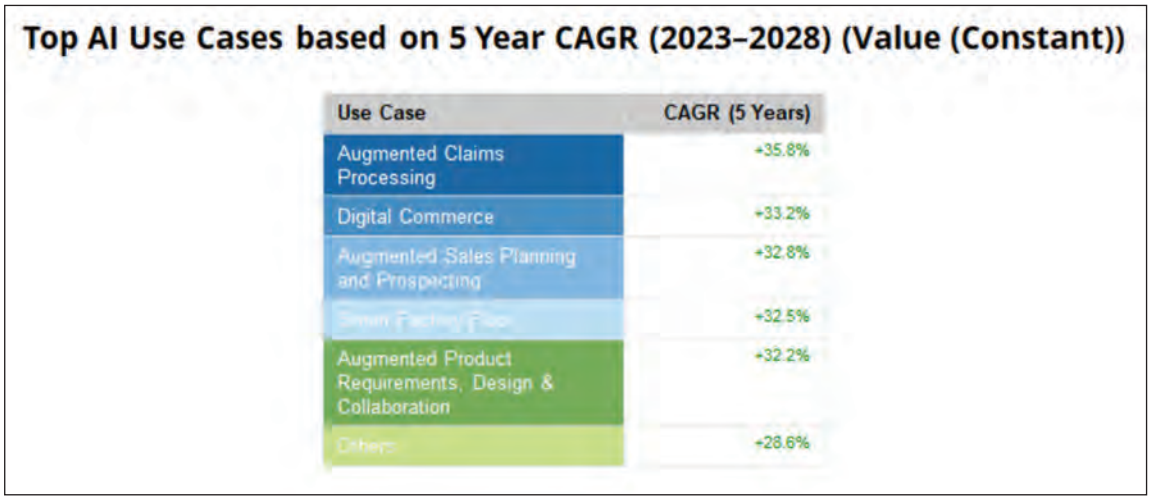
Worldwide spending on artificial intelligence (AI), including AI-enabled applications, infrastructure, and related IT and business services, will more than double by 2028 when it is expected to reach $632 billion, according to a new forecast from the International Data Corporation (IDC) Worldwide AI and Generative AI Spending Guide. The rapid incorporation of AI, and generative AI (GenAI) in particular, into a wide range of products will result in a compound annual growth rate (CAGR) of 29.0% over the 2024-2028 forecast period.

"AI-powered transformations have delivered tangible business outcomes and value for organizations worldwide and they are building their AI strategies around employee experience, customer engagement, business process, and industry innovations," said Ritu Jyoti, group vice president and general manager, AI and Data Research at IDC. "With rampant innovations in trusted AI tools and technologies and improved harmonization of human and machines interplay, barriers to AI adoption at scale will continue to diminish."

While GenAI has captured the world's attention over the past 18 months, spending on GenAI solutions will be less than the combined total of all



➤ Figure 1. Hype Cycle for Emerging Technologies, 2024. Source: Gartner (August 2024).



➤ A screenshot of a graph Description automatically generated. {Source: IDC Worldwide AI and Generative AI Spending GVuide - Forecast 2024 (August V2 2024).

other AI applications, such as machine learning, deep learning, and automatic speech recognition & natural language processing. However, the rapid growth in GenAI investments will enable the category to outpace the overall AI market with a five-year CAGR of 59.2%. By the end of the forecast, IDC expects GenAI spending to reach $202 billion, representing 32% of overall AI spending.

Software will be the largest category of technology spending, representing more than half the overall AI market for most of the forecast. Two thirds of all software spending will go to AI-enabled Applications and Artificial Intelligence Platforms while the remainder will go toward AI Application Development & Deployment and AI System Infrastructure Software. Spending on AI hardware, including servers, storage, and Infrastructure as a Service (IaaS), will be the next largest category of technology spending. IT and business services will see a slightly faster growth rate than hardware with a CAGR of 24.3%. In comparison, AI software will see a five-year CAGR of 33.9%.

The industry that is expected to spend the most on AI solutions over the 2024-2028 forecast period is financial services. With banking leading the way, the financial services industry will account for more than 20% of all AI spending. The next largest industries for AI spending are software and information services and retail. Combined, these three industries will provide roughly 45% of all AI spending over the next five years. The industries that will see the fastest AI spending growth are Business and Personal Services (32.8% CAGR) and Transportation and Leisure (31.7% CAGR). In addition, 17 of the 27 industries included in the Spending Guide are forecast to have five-year CAGRs greater than 30%.

Software will be the largest category of technology spending, representing more than half the overall AI market for most of the forecast. Two thirds of all software spending will go to AI-enabled Applications and Artificial Intelligence Platforms while the remainder will go toward AI Application Development & Deployment and AI System Infrastructure Software

AI Infrastructure Provisioning will be the leading use case for AI solutions for most of the forecast. However, with the slowest projected growth rate among the use cases included in the Spending Guide (14.7% CAGR) due to early investment by service providers, IDC expects several other use cases to catch or overtake it by 2028. These use cases include Augmented Fraud Analysis and Investigation and AI-enabled Customer Service and Self Service. The use cases that will see the fastest spending growth will be Augmented Claims Processing (35.8% CAGR) and Digital Commerce (33.2% CAGR). Thirty of the 42 AI use cases identified in the Spending Guide are forecast to have five-year CAGRs greater than 30%.

“We are thrilled to release a new version of IDC’s Worldwide AI and Generative AI Spending Guide with all new AI use cases aligned to line of business (LoB) functions and providing a GenAI/Rest of AI and industry view of each use case,” said Karen Massey, research director, Data & Analytics at IDC. “While industry-specific AI use cases approach 27% of the total spend by the end of the forecast period, the business functions that IDC expects will see accelerated AI investment are customer service, IT operations, and sales.”

AI spending in the United States will reach $336 billion in 2028, making it the largest geographic region for AI investment and accounting for more than half of all AI spending throughout the forecast period. GenAI spending in the U.S. is forecast to be $108 billion in 2028. Western Europe will be the second largest region for AI spending followed by China and Asia/Pacific (excluding Japan and China).

COOLING | HEATING | VENTILATION | CONTROLS

Ask ME about our IT Cooling solutions for Data Centres

Multi Density

Scan here

Data centres are now a crucial part of modern life and Mitsubishi Electric's **Multi Density** system makes IT Cooling easier than ever. Combining the efficiency of VRF with high-performance close-coupled air conditioning, Multi Density offers precise temperature and humidity control for high-density applications. It ensures optimal conditions for your IT equipment whilst helping increase energy efficiency and lowering emissions.

With features like advanced inverter technology, a small footprint, and pipe runs up to 165m, Multi Density is perfect for cooling high-density racks and blade servers.

Discover more about our **Data Centre Solutions** at: **les.mitsubishielectric.co.uk**

MITSUBISHI ELECTRIC

Sustainability is good for business and good for the planet, and DCIM is critical to reduce energy consumption and waste

The introduction of new model based, automated sustainability reporting features within the company's EcoStruxure IT DCIM software - discussing the importance of accurate performance metrics as data centres respond to the business imperatives of energy efficiency and performance optimisation as well as the march of environmental legislation - with the EU Energy Directive already making an impact.

BY KEVIN BROWN, SENIOR VICE PRESIDENT, ECOSTRUXURE SOLUTIONS, SECURE POWER DIVISION, AND ALISON MATTE, SUSTAINABILITY LEAD ECOSTRUXURE IT, BOTH AT SCHNEIDER ELECTRIC

DCS: *Schneider Electric has introduced some new features to the* EcoStruxure *IT DCIM. It would be good to understand the industry background that has led to the need for these new features?*

KB: It's an interesting time for the industry because it really sits at the intersection where IT and OT meet. And that is fundamentally where DCIM tools exist. One of the things that we're seeing in the industry is a bigger focus on sustainability. Now, it varies in different parts of the world. Obviously, in Europe, you have the EED and some of the reporting requirements that are coming in. But in general, when you look at the energy consumption of IT and as well as data centres, and certainly with the recent AI boom, you're seeing a greater focus coming from governments and industry on what are we doing to ensure that the energy consumption of IT is being done in a responsible and professional way. And we really feel that DCIM is in a unique position to give insight for customers on that energy consumption, because we already have a lot of the data. If somebody's deployed a DCIM solution, we have the data.

But what we realised, and it came from a project we did with our own CIO team, is the tools weren't oriented in a 'user-friendly' way - it wasn't as easy to get that data and to organise it as required, number one. Number two, sometimes the data was incomplete. As a result, we've done some work recently so that customers can meet a very simple challenge of: how much energy is my IT consuming and can I report on it so that I can manage it more effectively? We're really quite excited about what we've brought to the market, and we think we're getting some pretty good response to it as well.

DCS: *Schneider Electric rightly has a well-earned reputation for being at the forefront when it comes to sustainability. It would be good to understand what you are seeing sustainability-wise within the data centre industry, and on a global basis, as to the actions required. Is progress fast enough or are we where we need to be, for example?*

AM: Everyone realises that data centres are consuming quite a significant amount of energy and electricity, and there is a massive strain on the grid to respond to the needs of every customer and every company. Where data centre growth is exponential in certain countries, we're seeing regulations coming into those countries as well to limit the amount of data centres that are being built, but also to start controlling the energy consumption of these data centres. Europe is at the forefront right now with the Energy Efficiency Directive and with other regulations. The EU Commission has set the EED, that's coming into force in September. Additionally, Germany, for example, is adopting its own set of rules. The Netherlands is doing the same. We're seeing regulations spread throughout not only the European continent, but we're also international regulations as well.

Sustainability is a way to limit the energy consumption. It is also a way to reduce costs related to energy and to reduce waste in data centres where all these IT assets that are being managed. We're seeing DCIM play a role, and we are seeing an opportunity to help customers manage those IT assets and also to report on the regulations that are coming in.

DCS: *In terms of the legislation that's now coming in, is it because the industry hasn't moved fast enough and so it's been given a little nudge, or do you think it was going to happen anyway? However well the industry performed, legislation was inevitable?*

AM: One of the main metrics that the industry has been tracking is PUE. That's been around for decades, and it's stagnating. I do think the regulations coming in is to push customers to start actually reporting, but also to align on key metrics and ways to understand what these data centres are actually consuming, because a lot of the times there are assumptions that are being made. Regulations are creating baselines for companies to be able to actually make progress. And regulations do help push things along.

KB: Just to augment what Alison is saying, all the metrics that are the EED, they all make sense from our point of view, and they really came from the industry, and they came from things that we're doing. As Alison is saying, the EED is formalising it and putting a structure around it. Where there's going to be a lot of discussion is around how good does your data need to be in order to report against that? Because part of the reason PUE is stagnating, using Alison's term, is because sometimes you don't have all the data. You get into some of the smaller data centres, server rooms, wiring closets (which, aren't covered by the EED, but we think they will be over time), and it's not easy to get some of these metrics, and I think that's really what's going to be interesting for those in Europe, right? And the other thing for customers in Europe concerns what level of detail do I need in order to report against the EED? And this is where we've been developing tools to model it out and make it a little easier.

You know, in the other parts of the world, like in the US, I would make the argument to many CIOs, the metrics that the EED has adopted are the ones that really we would consider best practise. So, why not use those to report everywhere? Our philosophy on it is, look, even if it's model-based, it's good to start getting a benchmark. It's good to start tracking this information, even down to the smallest sites, and then you can get better over time. So, your question is quite interesting. Is it because the industry wasn't doing the right things that the regulations came in? No, I think it's because the industry was doing the right things. Now, because it's becoming so prominent in the discussions, the EED is formalising it. For those in Europe, it's going to be a real question about what's the level of detail that I need in order to comply and be auditable against the regulation? But for other people in the world where you don't have quite the same strict regulation in place, the argument is that you should be doing this anyway. We have examples of our own CIO team implementing our tools and getting some great results from a business standpoint.

DCS: *In terms of the DCIM features that have been introduced by Schneider Electric, they are providing more of the required metrics and visibility. I believe the company talks about enhanced visibility, historical data analysis, and things like fast intuitive reporting. Fundamentally, the objective is to provide more of this information and, whether you developed it specifically for the directive or not, it's a lot of the information that people will be required to report on. Is that right?*

KB: Yes. Much of the learning that we did and what we built in was as a result of watching our own CIO team do their own green IT initiative internally. They started using our own tools. They were doing

Sustainability is a way to limit the energy consumption. It is also a way to reduce costs related to energy and to reduce waste in data centres where all these IT assets that are being managed

a refresh of their IT infrastructure. And the EED was in the background because we saw the EED coming. So obviously, those two things really played together. Much of the work we did was around trying to make it easier for somebody to get those metrics. And again, as I mentioned, one of the big challenges is that people don't always have the data, and we did not limit it. I think the EED, if I remember, it's like 500 kilowatts and above, right? Certain countries might have a lower threshold. We actually think you should be implementing this at even the smallest server rooms and wiring closets, and that's what our own CIO team was doing. When you have incomplete data, what do you do? The reality was that we've spent the last 10 to 20 years working on data centres and modelling out PUE and what that looks like.

We've also spent the last 10 or 15 years, becoming a leader in sustainability as a business. You take that combined expertise - we say it's model-based, data-based modelling. Let's call it AI, because everybody has to use AI these days as a term! We looked at our knowledge of what we had internally, and we did use algorithms to help people start modelling out what these sites look like. And now you can obtain this information very easily - with the push of a button, you can get the PUE of a site in effect. Based on the model, we give people a level of confidence about what is the PUE at that site. And if we get more data, the confidence level gets better. And we're going to keep working on this because we have data coming into our cloud which we can continue to work on and analyse. These models are going to get better and better. And so in effect, it's a very easy way to get a baseline of what is your energy consumption, including the losses of the physical infrastructure. What we're most excited about is that we are enabling people to have a very fast and easy way to get started on this journey of really measuring what is your performance.

DCS: *I guess the next logical step at some stage, and it may already be happening to a certain extent, is once you've got all this information, clearly you can take actions based on it. I'm just wondering what the balance is at the moment and maybe plans for the future in terms of humans versus AI. AI has the potential to take a load of data and maybe recommend some actions based on it. What are you doing with it or what plans do you have with the decision making beyond the raw input of data that you're collecting?*

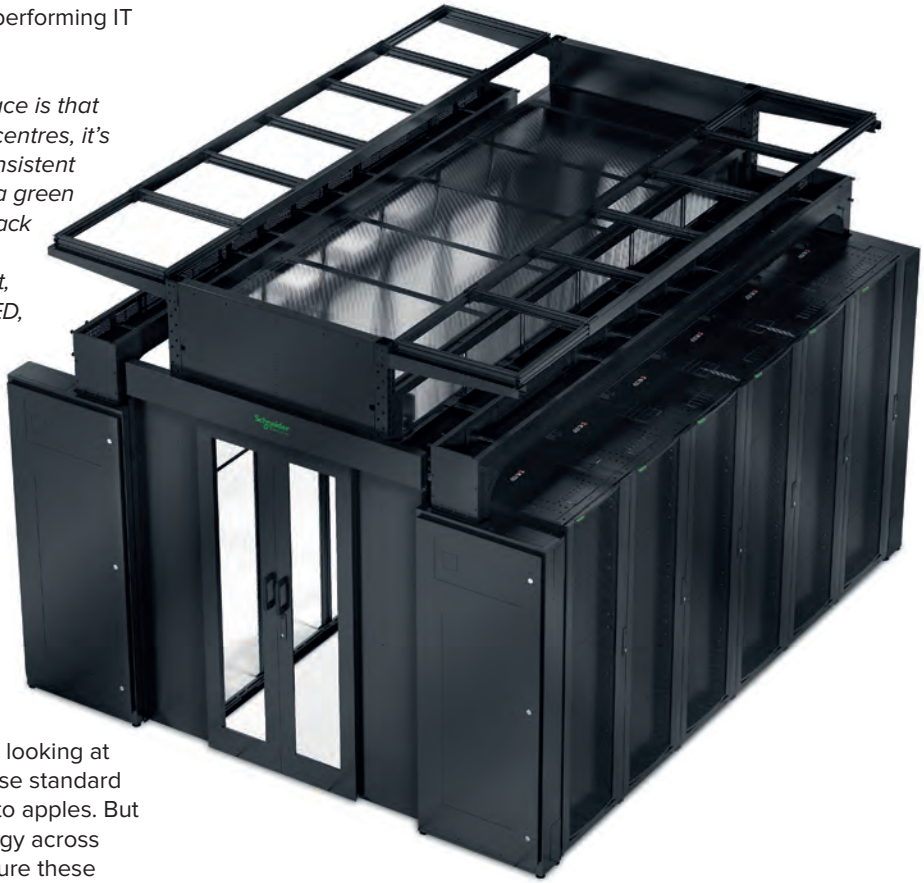
AM: I can give you an example from our own CIO team project. On our Lexington site, which had about 10 network closets, 10 IT rooms (so not a full data centre), they were able to start monitoring and tracking their energy consumption. They actually found zombie servers, and they found certain inefficiencies. With the recommendations that we were providing with our DCIM, alongside some upgrades to their assets, they were able to reduce their energy consumption by 20% in the first year. In carrying out that exercise, reducing 2 % their energy consumption, they reduced about 30% of costs. When you look at the sustainability side, they were able to reduce their carbon emissions by 17 tonnes that first year, just on that site, which is massive.

KB: The way this is going to play out is, when people go out and start measuring, it's going to point them in a direction about where's your worst performing site. And in the example Alison shared, they knew there was a lot of opportunities. Now, our software wasn't smart enough to tell them, hey, go virtualize this and go do that...you still need a human in there. But because they started running these tools and looking at the data, it pointed them in a direction to go. And what I think will happen is over time, as we start learning more from customers about what are their opportunities, you can see a day not too far in the future where we can sit there and go, based on what we see and based on benchmarks of other things that are similar, we think you may have a 30% opportunity at this site, and you should go look there. So that's really where we are in terms of the sophistication of the tool - it is giving people a very easy way to get a very broad view of their complex hybrid infrastructure and enabling them to start finding out where they should start looking and then really tracking the subsequent continuous improvement exercise.

Some of the savings that they got at our Lexington factory were because they went and did some higher-level IT stuff. They bought some new IT equipment, they virtualized things, and they were able to achieve, again, a 30% energy reduction, but that was a tremendous ROI from a business standpoint. Sustainability is good for business. I think that's an important message, and one we're trying to promote - when someone takes a sustainability initiative, a green IT initiative, they end up with a better solution that's more resilient and has a lower energy consumption. We maybe don't have don't have all the data as yet to support that assertion, but I think there's a lot of evidence that this will be the case. We're excited about this, not just because of the EED, but we think it's also

going to lead to less waste and better performing IT infrastructure.

DCS: *One of the problems end users face is that when they're looking at different data centres, it's very difficult to obtain, shall we say, consistent statistics. Everyone will say, we've got a green data centre, and they might be able to back it up with some data, maybedepending on how they interpret it. Do you think that, with the combination of DCIM and the EED, at some stage we might arrive in a great place where you can compare five different data centres with the exact same set of metrics? You will know the actual performance of each data centre, as opposed to some of the marketing/ greenwashing that might still be going on?*



AM: They're using Cenelec methodology, which has been around for years, to calculate these metrics. There is a lot of standardisation. With these regulations that are coming into force, and the EED in particular, if we're looking at the same methodology to calculate these standard metrics, then you can compare apples to apples. But if you're not using the same methodology across your sites to be able to track and measure these metrics, you're not going to be able to compare. You're not going to be able to see, to have a good baseline. So that's the power of regulations and standards. It allows you to do so.

KB: Historically, we've all seen there's two things that have happened in the industry where not everyone's PUE was calculated the same way, and people would play some games with it. I think that's fair to say. And the second thing is, there's been a self-selection bias, when people go and publish 'we have a PUE of 1.15 or whatever'. Is that just for one data centre on the best day of the year? Is that an average number? What is that number? We don't know. And that's where I think, rightfully so, there's been some suspicion about some of these marketing initiatives, talking about how green data centres are. And that is a role where regulations and standards come in. It's all defined. So now the question is going to be, if you have to comply with the regulation and it's auditable, that's going to take away, let's call it, some of the game plan. And that's one of the things we did with the tools. We built in that methodology so that when you run a report, it is consistent with the standards that are in place, and it should give you a pretty good estimate.

And there's going to be a debate as to whether or not a model-based solution is good enough? Good enough to be able to comply with the regulations? But our broader point here is that, if I use these standards and I use a model-based solution for those areas where maybe I don't have all the metering in place that you need to comply with the regulations, it's still going to ensure that you're running better infrastructure, and that it's as energy-efficient as possible, it's as robust as possible. Let's not lose sight of this goal for the industry. We are looking at that broader goal as well as the narrower goal of can I report against the EED? We help you with that. But there's a broader goal about using these model-based, data-based AI engines, to help people get a handle on their whole infrastructure. Keep in mind, our best estimate is that at least half of energy consumption is outside of data centres.

If you look at energy consumption, about half of it is in server rooms, wiring closets, networking closets. And today, the EED doesn't cover that, but that's half the problem. It doesn't get the headlines that the large data centres get, but it's still there. And we would argue that over time, regulators like to stay busy, too. I mean, they're going to keep lowering this threshold. That's what I would guess. They're going to go from 500 kilowatts to 250 to 100. It's going to keep coming down. Why not get ahead of that on best practises and make sure that you are looking at your infrastructure, particularly when it's very cost-effective and very easy to do so?

This Q and A article is based on a video interview between DCS and Schneider Electric. The video can be viewed via the DCS website at: https://datacentre.solutions/videos/4660/schneider-electric-continues-to-drive-the-data-centre-sustainability-agenda

The data centre of 2030: smarter, faster and more sustainable

Data centres are poised for an era of rapid change, and by the end of this decade they will be very different, with the data centre of 2030 built for sustainability as well as computing power.

BY IAN JEFFS, UK&I COUNTRY GENERAL MANAGER AT LENOVO INFRASTRUCTURE SOLUTIONS GROUP

THE REASON for this is simple, the power demands of data centres are rising rapidly. Data centre technology already accounts for 2% of energy use worldwide, according to the International Energy Authority (IEA), and by 2026, that figure could well double, with the demands of data centres equalling the electricity consumption of Japan. In Britain, the National Grid predicts that data centre power consumption could increase six-fold over the coming decade. These predictions are set to drive a surge in demand for sustainable computing.

One of the reasons behind the exponential rise in electricity demand is artificial intelligence (AI) and in particular generative AI, alongside other innovative technologies such as quantum computing. Generative AI's appetite for energy is enormous. Research estimates that generating just one image with AI uses as much energy as fully charging a smartphone. But this must be weighed against the important role AI has to play in the battle against climate change, with AI expected to deliver breakthroughs in clean energy (for example designing nuclear fusion reactors) and in other areas such as limiting methane emissions from waste. The challenge is clear: the world must find a balance to enjoy the benefits of AI, while also containing its impact on the environment.

So how can we build data centres fit for a cleaner future? Business leaders need to take an overview of the real energy impact of data centres, taking

in everything from how servers are cooled and how that energy might be reused, to how they are shipped and the mix of energy used. Understanding the true impact of data centres is the first step towards a smarter, more sustainable future.
Powering progress

With generative AI appearing in software ranging from email apps to internet search, the energy demands of the AI industry is going to continue rocketing, with one study in the journal Cell suggesting the power demand of AI alone could match the demand of the Netherlands by 2027. The compute power required to train AI doubles every six months, and Gartner® predicts[1] that 'by 2030, AI could consume up to 3.5% of the world's electricity'. The IEA's report suggests that adding generative AI to search (as companies including Google are rushing to do at present) could multiply the energy demands of internet searches by 10.

All of this requires the technology industry to design carefully for sustainability, not just at the chip level, but at the server level and data centre level. It's also worth bearing in mind that there is a flipside to this in terms of the environmental benefits of innovations sparked by new technologies. Emerging technology such as quantum computing will be more energy-efficient and could also mean that problems are solved exponentially faster than classical computers. Both quantum and AI are expected to drive rapid innovation in everything from demand response in the electricity grid to photovoltaics to electricity generation technology. The 'smart grids' of the future will be powered by AI. In building decarbonisation, to take one example, Mckinsey estimates that AI can accelerate the process 100-fold compared to existing technologies.

Smarter systems

The data centre of the future will be designed from the bottom up with sustainability in mind. Technologies such as warm water cooling enable high performance with far less energy use, provided a host of power consumption benefits. For instance, in data centres that use warm water-cooling, there is much less need for high-speed fans to dissipate heat. Air-cooling systems can often consume vast amounts of energy in themselves.

Furthermore, any wasted heat can be reused effectively because warm water-cooling systems produce heat waste at a temperature that is more easily reused for other purposes. Not only will this improve energy efficiency across an entire data centre facility, but this energy can be successfully recycled in sustainable ways in the wider community. By 2030, recycling the heat from data centres will become the norm, from heating nearby buildings and swimming pools to piping warm liquid under roads and walkways to melt ice.

The data centres of the future will also be built around renewable energy, from renewable sources to solar panels on roofs, driving towards a future of carbon-neutral or even carbon-negative operations. Organisations will adopt 'as a service' approaches to AI to improve efficiency, and increasingly businesses will harness the power of AI to optimise electricity consumption in data centres themselves.

Just as AI will assist in demand response in the electricity grid, within the data centre, algorithms can help with optimisation, predictive maintenance and energy management. This can reduce energy consumption by improving cooling efficiency, minimising waste and optimising resource allocation.

Circular thinking

When designing and building a data centre fit for the future, it's key for business leaders to take a view across the whole lifecycle of their data centre and the servers that will work in it. Everything from how components are designed and manufactured to how they are shipped, deployed and disposed of at the end of their lives matters, and taking a holistic view is key to making real sustainability gains.

Through this decade, asset recovery services (ARS) and recycling of computer equipment will grow in importance. Other markets in Europe have already implemented legislation to reduce the environmental footprint of digital technology.

For example, France introduced a mandate that 20% of IT devices bought by organisations need to be refurbished, with a target of 40% by 2040.

Manufacturing products regionally to cut shipping miles will also be a key differentiator. The current shift towards 'as-a-service' approaches to everything from hardware to software will continue, with business leaders focusing on avoiding overprovisioning to cut carbon emissions.

Towards net zero

The challenge facing the technology sector is clear. AI's hunger for energy is set to catalyse a race to adopt smarter and more sustainable approaches in the data centre as this decade unfolds. By 2030, we will see a new kind of data centre powered by renewables, and integrated thoroughly into the community around it with excess heat efficiently re-used to heat buildings and swimming pools. They have the potential to power a new era of quantum and AI, which will help unearth breakthroughs to tackle global issues, such as climate change.

FURTHER READING / REFERENCE

- 1. Gartner Says CIOs Must Balance the Environmental Promises and Risks of AI. GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.

Could Competency Frameworks help address the digital infrastructure industry's workforce challenges?

Competence is a commonly used word that can mean different things to different people. Most frequently understood as someone's ability to carry out a task effectively, more in-depth definitions extend this to express an adequate combination of skills, knowledge and experience – perhaps even the personal aptitude – to carry out a given task.

BY SARAH PARKS, DIRECTOR OF MARKETING AND COMMUNICATIONS, CNET TRAINING

NO MATTER the intricacies of the definition, general agreement is that competence describes what most hope to achieve as a minimum in their day-to-day lives and it's undoubtedly high on the list of requirements when filling vacancies in the mission critical workforce. Recent Uptime Institute reports indicate, however, that as many as 50% of data centre owners and operators are struggling to find suitably qualified staff. So why is competence so hard to come by?

The first obstacle is the overall lack of visibility of the industry in wider society, an issue that has been discussed time and time again over many years. In our digitally reliant world that hosts an ever-increasing number of connected devices per household (2023 research found an average of 12 or 17 connected devices per household in Europe and the US respectively – and predictions of 24 billion worldwide by 2050), it's difficult to believe that still so many people lack understanding of how these devices work, or more importantly, where their data is actually being stored. This unfamiliarity means candidates who may possess the ideal aptitude to make up the industry's skills gap could be oblivious to the opportunities that are available – or might be

put off by a misunderstanding that the industry is 'too technical' for them.

The flipside of this obstacle – and another challenge in itself – is a widespread reluctance across the industry to fill job openings from 'outside'. Where a job advert specifies industry experience as essential, how can industry newbies expect to get a foot in the door? With the majority of companies demanding new hires come with industry knowledge and/or prior experience, this requirement is proving a major blockage to attracting new, competent entrants to the industry, and is severely limiting potential appointments to a diminishing pool.

If you do conquer these hurdles, you face the task of assessing competence quickly and accurately enough to be able to make an informed recruitment decision whether the individual is right for the role (or indeed, for candidates to be able to evidence it quickly enough to be able to showcase their talents). Do you use 'standard' interview questions, require applicants to complete a practical task or presentation, or rely on gut instinct – a feeling – to select the person best suited to the role?

This focus on competency is not a new one. In 1960, the "Management of Training Programs" textbook introduced the concept of the four stages of competence, which were further honed in a 1969 article published by management trainer Martin Broadwell. Naming them "the four levels of teaching", Broadwell segmented the stages that learners go through when acquiring new skills into four categories: unconscious incompetence (lacks ability and is unaware of the usefulness of having it), conscious incompetence (awareness of the value of the ability and that they do not possess it), conscious competence (ability to complete the task with heavy conscious effort), and unconscious competence (in-depth ability to complete a task, such that it becomes 'second nature').

Many have suggested furthering this hierarchy with a fifth stage, proposing names that range from chosen conscious competence to complacency, or even enlightenment. Though these four stages were introduced as a hierarchy, it's now recognised that competency is not a linear journey, with one stage leading on to the next until you reach a pinnacle. It is possible (and expected) to flit between stages as your skills and experience grow or – on the reverse – diminish, as learning is forgotten and competence declines.

Keeping this hierarchy of competence in mind while we consider the industry's workforce challenges, we gain a valuable insight, not only into the way people acquire skills and knowledge, but also how they fare over time. Acquiring a deep understanding of the way people learn and retain knowledge is beneficial in assessing competence. It can provide useful benchmarks by which to determine the level of education required to attain the expected skill level for a role. Using ongoing assessment tools that are based around a hierarchy of competence (such as Broadwell's) enables leaders to easily identify skills and knowledge gaps across teams (or in prospective candidates when used during the recruitment process) and pinpoint exactly the measures that are required to bring the individuals to a competent standard.

Beyond the digital infrastructure industry, frameworks for competency are fairly commonplace, constructed to help organisations understand the specific value-based behaviours, skill and knowledge requirements for each role in the business. Competency frameworks clearly outline expectations, enabling a fair opportunity for everyone to be assessed on a level playing field, regardless of background, age, gender, race, etc., and therefore encouraging diversity. They benefit individuals and managers by enabling transparent benchmarking which can help both parties identify opportunities for personal growth, and the clarity they provide empowers managers to place value on skills rather than qualifications.

To reap the highest rewards, competency frameworks can be tailored to reflect not only the needs of a role, but also the organisation it sits within. By using a customised approach, the framework can be adapted to include softer, behavioural skills, as well as the more technical requirements, that fit the company's unique values, ethos, goals and ambitions. Using the framework, businesses can identify any skills that they're missing across teams, and armed with this knowledge, leaders can implement skills improvement activities such as further learning, or mentoring, or recruit the people they need to meet business goals, as well as carry out a role, consequently improving business performance overall.

So, could the implementation of competency frameworks into common roles across the digital infrastructure industry help empower companies to open their eyes to the talent that lies beyond the walls of the industry and fill their vacancies with competent candidates, regardless of their background, academic achievements or prior experience?

No single change is going to fill the cracks that are forming in the industry, but there are some steps that are relatively easy to implement and could have a ripple effect across the sector. Establishing clarity of requirements for individual jobs roles, applying these frameworks to recruitment processes, and assessing skills and knowledge levels to maintain competency on a continuing basis, can only help us along the path to finding what we are looking for; competent individuals with the right mix of behaviours, skills and knowledge to fulfil the demands of our digital workforce, for today and the future.

Energising the digital economy: **the impact of IDNOs**

Independent Distribution Network Operators (IDNOs) have a vital role to play in energising data centre developments, and how IDNOs can release hyperscalers and data centre operators from the constraints that threaten to hold back delivery of their potential for the whole UK economy.

BY DAVE SWADLING, DIRECTOR OF CUSTOMER CONNECTION DEVELOPMENT AT ECLIPSE POWER

THE UK'S DIGITAL INFRASTRUCTURE plays a vital role in the country's economic growth. According to Cloudscene, the digital economy contributes around 8% of the country's GDP, generating over £65 billion gross value added (GVA).[1] No wonder that successive governments have made digital growth a core feature of the economy. As critical building blocks of the digital economy, data centres will need to keep pace with the explosion in power-intensive workloads being generated from cloud computing, IoT, machine learning and artificial intelligence.

For some time, the expansion of data centres has been subject to a 'perfect storm' of converging events that present real challenges to growth – from planning barriers to demands for greater energy efficiency and sustainability. Business leaders are increasingly focused on the role of data centres in their supply chain emissions to account for their upstream Scope 2 and 3 carbon footprints[2]. But the overriding issue for hyperscalers and data centre operators is grid capacity.

Applications made to Distribution Network Operators (DNOs) or the National Energy System operator (NESO) today are being given connection dates that are 10 to 15 years in the future. This delay puts projects – and investment – at risk. When your business plan is centred on having X-number of megawatts (MW) or a volume of data centre capacity set up and running in a certain timeframe, being unable to connect to the grid for a decade is clearly a problem. It's more than a problem for the enterprise, it's a problem for the broader economy.

The UK is an attractive location for data centres because of its political, economic and energy stability. Without connectivity at the scale and pace needed by the digital sector, investment in new projects will be redirected into countries in Europe and further afield where it's easier to find sites with suitable power and connection timescales.
Also, as we progress towards a zero-carbon economy, we are connecting more renewables to the grid, adding to the connections queue which is currently standing at 800 gigawatts. This is all taking place as data centres are becoming more prevalent and larger.

In short, data centres are waiting longer than ever for larger power connections, at a time when we've effectively topped out the grid in terms of capacity... the perfect storm. The light on the horizon, however, is that there are solutions to overcome these challenges, if you know where to look and who to work with.

The difference is in the 'I' of the beholder

Independent Distribution Network Operators (IDNOs) are perfectly equipped to unlock opportunities for data centre operators to connect to the grid. Introduced in 2004 to increase competition in electricity distribution, IDNOs, like DNOs, design, own, operate and maintain electricity networks in the UK. Also, like DNOs, they are licensed by Ofgem.

The difference lies in the 'I' of IDNO. 'Independence' means that IDNOs are not restricted to a geographical part of the UK as DNOs are. They can operate nationwide and can be more flexible about how they interpret the standards set by DNOs, which vary from region to region. Because IDNOs operate in a competitive market, they have to have more of a customer-focus than DNOs. This means that they can adapt and adjust to market challenges in a way that DNOs aren't incentivised to do. They can suggest innovative ways to overcome the challenges and simplify the complexities of getting connected to the grid.

DATA CENTRE TRANSFORMATION

22 · 10 · 24

CONFERENCE ORGANISED BY

datacentre dca alliance

22 OCTOBER 2024

THE IET BIRMINGHAM

(Institute of Engineering & Technology)

www.dca-global.org

t: +44 (0)845 873 4587 e: info@dca-global.org

MEDIA PARTNERS

DATA CENTRE WORLD · DIGITALISATION WORLD · DATACENTRE SOLUTIONS · TECHERATI · DCR Data Centre Review Magazine · WiP

Naturally, IDNOs with experience of helping data centre customers connect to the grid will understand HV connections. The effective ones will offer additional advantages, like making a capital contribution, value engineering the grid connection as well as design expertise to match the substation footprint around the needs of a specific site. They can also apply learned experience from energising other sectors.

It's often the less technical aspects of power projects that have the biggest impact on whether a project is viable or not. Successful IDNOs must have the ability to build strong relationships with stakeholders across the whole power ecosystem; the transmission and service operators, DNOs and National Grid, grid consultants, renewable developers, and more. They typically build a deep understanding of not only who to deal with among the different organisations, but importantly how to deal with them.

The really effective IDNOs do all of this in a way that is designed to help the data centre customer by providing the right advice in an open and transparent way. Ultimately, the data centre operator needs to convince investment communities of the value and viability of their project. So, the IDNO needs to distil and crystalise both the problem and solution in a language that makes it easy for the client to sell internally.

The strength of the relationships mean that they can negotiate with the DNO and National Grid to get the best solution for their customers and introduce innovative ways to solve connectivity challenges.

Unlocking the queue

Given the challenges of connecting to the grid, IDNOs like Eclipse can operate independently as 'power-brokers' to unlock the queue. While there is no silver bullet or any single solution to fit all situations, a customer-centric IDNO has the commercial drive to find solutions. IDNOs are ideally suited to finding the best solution and bringing the right people together. That can involve putting the investors together with the people who need to build data centres; the grid experts with the people who have the connections; the landowners who want to do something with their land options; the grid consultants and designers; the renewable generation developers. And bring all of this together to make as many projects viable as possible by taking advantage of what is already in the grid queue. There's no sense getting in the back of the queue, at 800 gigawatts and counting, there's too much in it.

An IDNO looks for the nuggets that are already in the queue that are suitable for a data centre project. Many of the renewables projects in the connections queue are for battery storage. Ofgem, government and NESO forecasts point to the UK needing 200 gigawatts of capacity to support the energy transition. That means only a fraction of the 800 gigawatt pipeline of projects will actually connect. Some of these will be looking to switch into data centre opportunities and use their place in the queue as an incentive. Additionally, data centres have large UPS battery facilities that can be used as export capacity to support the grid – essential as more variable renewable energy sources come online. So, the right power-broker in the form of an IDNO can help unblock the connections queue and increase the penetration of renewable energy in the grid.

The common denominator

Data centres will continue to innovate to decarbonise and find new ways to work collaboratively with the grid. IDNOs are a common denominator within the highly complex equation, as they connect all of the constituent parts. It simply makes sense for any data centre developer and operator to have an IDNO inside their camp.

Working with a skilled IDNO ensures that appropriate relationships are in place so that hyperscalers and data centre operators are released from the constraints that threaten to hold back delivery of their potential for the whole UK economy.

Opening the DNO market up to competition has given the sector a fresher, faster route to solving energy connectivity challenges. The transmission network is crying out for more competition to drive regulatory and industry change. This would enable customers to connect to the network faster and more efficiently and is the natural next step that would revolutionise energy in the UK.

FURTHER READING / REFERENCE

- 1 Source: Statista
- 2 Source: Hitachi Vantara

wesco | anixter

COMMSCOPE®

Connect

Your Data Centre Priorities are Ours

Delivering value at every stage of data centre development and operations, Wesco Anixter brings reliable expertise alongside a global ecosystem of strategic partners. Together with CommScope, we provide advanced infrastructure solutions to build high-peformance, future-ready foundations for data centres.

Scan to learn more about SYSTIMAX® VisiPORT.

We build, connect, power and protect the world.

Wesco.com/datacenters-emea

Ripping up the old blueprint:
How to create data centres that are sustainable by design

Fasthosts explores how organisations that own data centres can make sure they are operating sustainably.

BY GEORGE DALY, HEAD OF DATA CENTRE FABS, FASTHOSTS

EXTREME ENERGY CONSUMPTION throughout the data centre industry is contributing to the creation of significant greenhouse gas (GHG) emissions. This is down to the growing demand for services such as cloud computing, data storage and webhosting.

While services like cloud computing are traditionally a greener alternative, this is only true if the services are managed sustainably. Worryingly, the European Commission now estimates that by 2030, EU data centre energy use will increase from 2.7 per cent to 3.2 per cent of the Union's total demand.

This would place the industry's emissions almost on par with pollution from the EU's international aviation. As these markets grow, choosing the right provider is becoming increasingly important. Businesses are now demanding that providers demonstrate genuine sustainability.

A flaw in the design

To meet these expectations, data centre owners must focus on sustainable design and operation. Appetite for change is in the air, but providers are apprehensive about taking the first step towards moving the sustainability needle.

The catch? Uncertainty.

Most businesses don't fully grasp sustainability. Not because they don't care, but because they simply don't know where to start. This often leads to them continuing as they have been, implementing half measures and turning to quick fixes like carbon offsetting, which is just the tip of the environmental impact iceberg.

True sustainability requires continuous effort, a willingness to learn and a thoughtful integration of technology. Embracing this mindset, let's explore the specific steps that data centre owners can take to build a greener future.

Step-by-step: forgetting off-the-rack design

For data centre owners looking to address harmful energy consumption in its tracks, rebuilding or relocating can hold a lot of answers. In doing so, we believe there are four steps that ensures construction keeps sustainability at its core.

First of all, providers must carefully consider location. This means finding a space that's close to a power facility. Bonus points go to those that find a renewable source. This is because distance matters. If electricity travels a long way between generation and use, a proportion is lost. Besides this, data centres located in cooler climates or underwater can cut down on the energy required for cooling.

Next, it's time to think about design and materials. Instead of completely reinventing the wheel,

providers can pull a sustainably conscious blueprint from the existing catalogue and tweak it to meet their needs. The result? Modular design.

Instead of creating a huge facility from scratch, modular design allows for gradual construction. By doing this, each new module of the centre can be purposefully built on demand with cooling and energy efficiency in mind. Nothing is produced for the sake of it.

This supports reusability as modules can be easily repurposed and relocated, as well as made near the final destination. Thereby, minimising transportation. Fewer trips in a truck, mean fewer steps on the carbon pedometer.

Once a design has been agreed, providers must focus their attention on materials. From laying sustainable concrete made with fly ash to using low-emission paint that emits fewer volatile organic compounds (VOC), a few simple switches can drastically reduce carbon output.

Now, with the building blocks beginning to fall into place, providers must consider implementing efficient and renewable technology. It offers powerful ways to enhance energy efficiency when it comes to data centre storage.

Some providers have been investing in algorithms, software and hardware designed to optimise energy usage. For example, introducing frequency scaling, which adjusts the power usage of processors based on the workload, or AI and machine learning algorithms can significantly improve how data centres manage power consumption and cooling.

Alongside this, data centre owners that are concerned about whether sustainable design presents a threat to performance should think about high efficiency uninterruptable power supplies (UPS) and power distribution units (PDU).

When operations are running smoothly, the former sits back and waits. However, when an outage occurs the UPS jumps into action and keeps the digital environment running. They are designed to generate less heat and require minimal energy.

Finally, data centre owners can't forget virtualisation and server optimisation. This means a data centre can provide high levels of service without using a copious number of servers. Virtualisation and optimisation help servers to perform well, save providers money and use minimal energy.

How does it work? Like a chest of drawers, virtualisation breaks a single computer into several different server compartments. Despite using the same hardware, this means each server can run its separate activity. In doing so, the original computers power is more efficiently used. As a result, the business can improve its resource usage, boost energy efficiency and scope for scalability.

Despite the clear benefits of greener data centres, some challenges cannot be overlooked. Primarily, the upfront costs associated with sustainable design and advanced technology can be substantial. Investing in renewable energy sources, energy-efficient hardware and innovative construction materials often requires a higher initial financial outlay. Sometimes, the newest sustainable technologies might not meet high-performance and scalability demands.

While these obstacles are significant, they are not insurmountable. By planning strategically and considering long-term savings on energy costs and potential tax incentives, companies can overcome these barriers. Addressing these challenges head-on prepares businesses for a greener future and strengthens their resilience against stricter environmental policies in the future.

Designing a more sustainable tomorrow

As it stands, the data centre industry is a key climate change culprit. And its dangerous energy consumption is slowly spiralling out of control. However, it needn't be that way. Instead, organisations can rip up the old rulebook and reinvent how data centres are designed and operated to place sustainability at the core.
True sustainability isn't a quick fix. It's not a sticking plaster. It takes time. For businesses choosing to make this transition by building or relocating their data centres, it's essential to bear location, design and materials, renewable technology and server optimisation in mind.

Ensuring data centre efficiency in an AI-driven world

With increased investment comes far-reaching implications for data centre infrastructure. As AI technologies grow more sophisticated and their applications more widespread, data centres face growing pressures to evolve at speed while ensuring security, sustainability, and scalability.

BY ALAN STEWART-BROWN, VP EMEA, OPENGEAR

IN THIS CONTEXT, scalability is no longer a luxury but a necessity. As AI solutions grow, the demand for processing power escalates, forcing data centres to expand their capacity quickly and efficiently. However, this expansion is not simply about adding more servers or storage. Effectively managing the increased load also requires advanced orchestration, automation, and network management.

Sustainability, security, and service continuity

As with any consideration in the world of data centers, the environmental impact is a key factor. More power means increased energy consumption and a larger carbon footprint, therefore data centres are compelled to not only manage their energy consumption more efficiently but also to invest in renewable energy sources and advanced cooling technologies. Ensuring the longevity of sustainable data centres amid rapid AI expansion requires a balanced approach to growth and environmental responsibility. As data centres become more integral to AI operations, the stakes for network security and operational continuity are raised even further. The complexity of AI applications and the data they process require robust protection against cyber-threats and efficient recovery systems to manage potential disruptions.

Further ramping up the challenge, data centre businesses are faced with having to do all of this with what is, in many cases, a shrinking workforce. The Uptime Institute Data Center Staffing Survey 2023 revealed that 58% of data centre operators reported that their organisation is having difficulty finding qualified candidates for open jobs.

More specifically, there is also a shortage of network engineers in many regions, with research by Opengear revealing that more than six out of ten (61%) UK CIOs expect at least 25% of their network engineers to retire in the next five years.

So, with the ongoing advance of AI continuing to place their networks under increasing amounts of pressure, many data centres have been forced to face up to a wide range of other challenges – from

APC™

40Years

Make greater

IMPACT

with an industry-leading network

Celebrate 40 years with APC.

Start leveraging four decades of uninterrupted protection, connectivity, and unparalleled reliability with the APC UPS family, a legacy marked by pioneering UPS technology and an unwavering commitment to innovation.

https://www.apc.com/uk/en/product-category/88972-uninterruptible-power-supply-ups/

Life Is On Schneider Electric

intensifying cyber-threats to growing skills shortages – that make managing these pressures effectively even more difficult.

That's where AI-driven network management solutions fit in, ensuring that data centres can expand in size and capability, even with network engineers in short supply, without impacting efficiency or performance. Alongside this core network management capability, power management and monitoring can be leveraged, helping data centres to achieve greener operations. At the same time, so as to ensure they can maintain service continuity even under adverse conditions, data centres need to be looking at solutions like Smart Out of Band Management that provide an independent overlay management network and ensure seamless connectivity for both Ethernet and serial management. In this way, they can help ensure they are achieving robust network management, whatever their evolving needs may be.

To do all this effectively, however, involves data centres upgrading their physical infrastructure and integrating AI within their operational frameworks to optimise processes and pre-emptively address potential issues.

President Biden's recent executive order in the US highlights the focus on AI›s implications for business operations, positing a future where data centres are central to AI-driven innovation. This governmental spotlight highlights the need for data centres to continue focusing on early adoption of new technology that puts them in a good position to manage the increasing demands of AI workloads.

Ensuring resilience and efficiency

The rapid growth of AI technology places immense demands on data centres, necessitating a multifaceted strategy to achieve scalability, sustainability, and security. Government investments in AI, particularly in the UK and US, emphasise the critical role of data centres in this evolving landscape. To meet these challenges, data centres must swiftly upgrade their infrastructure, incorporating advanced AI-driven technologies for network management and automation. These enhancements are essential not only for handling increasing computational loads but also for ensuring energy efficiency and robust cybersecurity.

Additionally, addressing workforce shortages through innovative solutions is crucial for maintaining operational continuity both now, and in the decades ahead. By adopting cutting-edge technologies and strategic workforce management, data centres can remain resilient and efficient. This proactive approach will enable them to support the ongoing expansion and complexity of AI applications, ensuring they continue to meet the demand in a responsible way.

StratusPower™

Future-proof your data centre and experience tomorrow's power protection technology today

Green UPS

DARA TECH

SWISS MADE

Centiel's **StratusPower™** is the ultimate power protection solution for today's dynamic data centre environment.
With unmatched availability, reliability, and efficiency, StratusPower ensures seamless operations and business continuity, minimizing the risk of downtime.

Our innovative DARA design delivers unparalleled scalability, eliminates single points of failure, and provides a fault-tolerant architecture. From compact 10 kW modules to robust 62.5 kW options, the UPS meets a range of power requirements with the ability to scale up to an impressive 3.75 MW.

centiel
continuous power availability

www.centiel.com

Don’t leave me stranded:
the expensive and environmental risk of unused assets

The potential risk of data centers becoming stranded assets due to their power-intensive nature. As data centers increasingly handle High-Performance Computing (HPC), facilities can be challenged when it comes to delivering the output demanded while balancing the economic viability with the needs of the environment.

BY GORDON JOHNSON, SENIOR CFD MANAGER AT SUBZERO ENGINEERING

DATA CENTERS, particularly those handling HPC, are at constant risk of becoming stranded assets. Stranded assets can occur when facilities do not meet their designed capacity, are no longer economically viable due to changes in technology or business needs or fail to contribute effectively to sustainability measures. In the context of data centers, this often relates to cooling challenges and the cooling capacity that the ITE can’t use.

Power, space, and cooling are the main capacity parameters in the data center. Stranded capacity is installed capacity that cannot be used to support critical load. It refers to data center resources that are not available for use. Unfortunately, most of the time, you don’t even know if you have stranded capacity or, if you do, how much there is.

While critical loads are expected to be renewed, refreshed, or replaced over the lifetime of the data center facility, older, non-energy star certified, or inefficient servers that are still turned on but no longer being used continue to use both power and cooling resources. It also includes excessive

redundancy or low utilization of the redundancy options, a lack of scalable, modular design and the use of oversized equipment or legacy lighting and controls.

While many may plan for the update and evolution of the ITE, the mismatch of power and cooling resources versus the equipment requiring the respective power and cooling inevitably results in stranded assets.

Challenging the invisible ceiling

Every data center has an invisible ceiling that limits the amount of ITE that it can cool. With the shift towards HPC data centers, facilities are even more challenged when it comes to having enough cooling capacity to match (and exceed) ITE cooling demand. As data centers become more power-intensive due to HPC requirements, the cooling infrastructure must keep pace.

Inadequate cooling can lead to increased operational costs, reduced efficiency, and environmental concerns. It's a problem since it prevents data centers from meeting design capacity and also restricts them from becoming sustainable and energy efficient. It's also an expensive problem since wasted cooling energy is not contributing to the overall cooling of the ITE.

Stranded capacity is wasted energy, cooling unnecessary equipment and lost cooling to areas that need not be cooled. Stranded cooling capacity can include bypass air (supply air from cooling units that is not contributing to cooling the ITE), too much supply air being delivered from the cooling units, lack of containment, poor rack hygiene (missing blanking panels), unsealed openings under ITE with raised floors, just to name a few.

Combating the inefficiency

Releasing stranded capacity improves energy efficiency, reduces costs, and allows for increased data center capacity without additional cooling equipment. To avoid becoming a stranded asset, organizations should conduct thorough assessments when building or upgrading data center facilities.

While the root cause for stranded assets is often down to operational issues and conflicts of management or space allocation, future scalability, energy efficiency and identifying technology obsolescence must be considered with strategic changes or updates.

A lack of metering and metrics can contribute significantly to an inefficient data center. To combat this, it's highly recommended to use data center monitoring software to help identify and eliminate wasted energy within the data center white space.

Monitoring and compliance with operational best practices can minimize the power and cooling imbalance effectively, but this needs to be an ongoing and continually updated process to reduce risk.

Realising the energy savings

Air management and control is a prerequisite to many energy efficiency measures. Wasted air is, after all, wasted money. Realising energy savings through correct air management can be done in two ways. First, to physically rearrange the space to promote the separation of hot and cold air. While this measure by itself does not save energy it does enable future savings to be made. Secondly, one of two of these actions must be taken: increase the supply air temperature and/or decrease the supply airflow rate.

Other options for energy savings include removing server waste. Idling or unused servers can consume up to 60% of full power while delivering no tangible output. Under-utilized ITE can be consolidated, saving space and optimizing the utilization of the ITE that you do need. Virtualization may be the most powerful of consolidation approaches. It consolidates applications from under-utilized ITE onto fewer items and better-utilized hardware. Some organizations opt for colocation or cloud-based solutions to minimize the risk of stranded assets by outsourcing their data center needs to providers that specialize in maintaining infrastructure.

In summary

Addressing the challenge of the stranded asset is vital for ensuring the industry remains both sustainable and economically viable while laying the financial foundation for its long-term future. It makes no sense to allow ITE to take up space that's either not being used, consumes power but gives nothing in return, or indeed utilizes energy that affects and upsets the environmental equilibrium.

A lack of metering and metrics can contribute significantly to an inefficient data center. To combat this, it's highly recommended to use data center monitoring software to help identify and eliminate wasted energy within the data center white space

Data centre cooling for AI workloads

In the rapidly evolving landscape of artificial intelligence (AI), the demand for computational power is soaring, placing unprecedented strain on data centres around the globe. The burgeoning scale and intensity of AI workloads bring with them a critical challenge: managing heat.

BY DAVID WATKINS, SOLUTIONS DIRECTOR AT VIRTUS DATA CENTRES

AS AI MODELS grow more sophisticated and data processing becomes more intensive, the thermal management of data centres has become a pivotal concern. Effective cooling strategies are not merely about maintaining optimal operating temperatures, they are integral to ensuring system reliability, maximising performance, and minimising operational costs.

Data centre cooling is no longer a peripheral consideration but a central component of data centre design and operation. Traditional cooling methods, such as air-based systems, might struggle to keep pace with the thermal demands of modern AI hardware. These systems often rely on cold air being circulated through the data centre, which is then heated by the high-density compute equipment. However, as AI workloads generate an ever-increasing amount of heat, this approach can lead to inefficiencies and higher energy consumption. To address these challenges, data centres are exploring more advanced cooling solutions methods.

The liquid alternative

Liquid cooling has emerged as a transformative addition to traditional air-based cooling systems, offering significant advantages in managing the heat produced by high-performance computing environments. By utilising liquid to cool components directly, this approach achieves more efficient heat dissipation compared to air, enhancing thermal management across data centres. This not only boosts cooling efficiency but also helps lower the overall energy footprint, making it an appealing solution for handling intensive AI workloads. Two primary approaches to liquid cooling are especially noteworthy: immersion cooling and direct-to-chip cooling.

Immersion cooling involves submerging IT hardware - such as servers and graphics processing units (GPUs) - in a specially designed dielectric fluid, such as mineral oil or synthetic coolants. This fluid effectively absorbs and dissipates heat directly from the components, eliminating the need for traditional air-cooled systems. By directly cooling the hardware, immersion cooling greatly enhances energy efficiency and reduces operating costs. This method is particularly advantageous for AI workloads, which often generate substantial amounts of heat and benefit from the enhanced thermal management and reduced energy consumption offered by immersion cooling.

In contrast, direct-to-chip cooling, also known as microfluidic cooling, focuses on delivering coolant directly to the heat-generating components of servers, such as central processing units (CPUs) and GPUs. This approach maximises thermal conductivity by targeting heat dissipation at the source, thus improving overall performance and reliability. By addressing the cooling needs of critical components more precisely, direct-to-chip cooling minimises the risk of thermal throttling and hardware failures. This method is especially crucial for data centres managing high-density AI workloads, where maintaining peak operational efficiency and system stability is essential for optimal performance. Overall, both immersion and direct-to-chip cooling technologies provide effective solutions for the heat management challenges posed by advanced AI workloads, offering data centres the ability to enhance cooling performance while reducing energy consumption and operational costs.

Available for sales and rental

CRESTCHIC

THE **TRUSTED** NAME FOR **LOAD TESTING**

200kW Data Centre Loadbank: Built for Integrated Systems Testing

Simulate full load operation, power outages, and cooling system failures

Verify PDU operational readiness during commissioning

Closely mimic server heat generation to test cooling systems

Control up to 20 units from one controller

- Small footprint, stackable for easy storage
- AC load bank resistive-only (200kW)
- Voltage range: 400V or 415V
- 50/60 Hz Operating frequency range
- 3 Phase
- Forklift pockets for lifting
- Sheet steel construction with IP55 rating
- Low Delta T
- Low noise
- Internal or external auxiliary supply operation

www.crestchicloadbanks.com

sales.uk@crestchic.com | hires.uk@crestchic.com | service.uk@crestchic.com

Sales & Rental Offices in : UK, North America, Singapore, Germany, Belgium, France, China and UAE

A flexible approach

The versatility of liquid cooling technologies offers data centre operators a strategic advantage, enabling them to customise cooling solutions to meet specific infrastructure and AI workload needs. By employing a mix-and-match approach, data centres can enhance cooling efficiency and adapt to diverse cooling requirements. Integrating different cooling technologies - such as immersion cooling, direct-to-chip cooling, and air cooling - allows for the optimisation of heat management across various components and workload types. This flexibility ensures that each technology's strengths are maximised while mitigating its limitations.

This approach addresses the varied heat dissipation characteristics of diverse AI hardware configurations. By tailoring cooling solutions to specific workload demands, data centres can maintain system stability and performance. As AI workloads and data centre requirements evolve, a scalable and adaptable cooling infrastructure becomes crucial. Combining multiple cooling methods supports future upgrades and expansion without compromising efficiency. For instance, while air cooling remains essential for certain high-performance computing and networking components, the primary heat rejection systems, like chillers, must be appropriately sized and capable of heat reuse to manage waste heat effectively. This integrated strategy not only enhances cooling performance but also prepares data centres for ongoing advancements and growing demands. Whilst liquid cooling has emerged as the preeminent solution for addressing the thermal management challenges posed by AI workloads, it's also important to understand that air cooling systems will continue to be part of the data centre infrastructure for the foreseeable future.

Navigating challenges

Because AI workloads involve diverse hardware with varying cooling needs, it is challenging to design a one-size-fits-all cooling solution. Advanced systems must be tailored to specific components, requiring detailed engineering and testing. This need for customisation can lead to compatibility issues and increase the complexity of the cooling infrastructure. Despite the significant potential benefits, the adoption of advanced cooling technologies in data centres - particularly for AI workloads - presents additional challenges that must be overcome. Perhaps most pressingly, the initial expense of advanced cooling systems, such as liquid and immersion cooling, can be substantial. These technologies require significant investment in new equipment and infrastructure modifications. Integrating these systems into existing data centres is complex and can disrupt operations, particularly in older facilities not designed for such upgrades.

Advanced cooling solutions also demand specialised maintenance and skilled personnel. Systems like immersion cooling need careful monitoring to prevent issues such as leaks, and finding or training staff with the necessary expertise can be difficult. Ongoing maintenance is essential to prevent downtime and maintain efficiency, adding another layer of operational complexity.
As AI technologies evolve, cooling systems must continue to be scalable and adaptable. Ensuring that these systems can be upgraded or expanded without significant additional costs or disruptions is crucial for maintaining long-term operational efficiency.

What lies ahead?

As data centres grapple with the soaring thermal demands of AI workloads, the adoption of advanced cooling technologies represents both a crucial necessity and a promising opportunity. Liquid and hybrid cooling solutions offer significant advantages in managing the intense heat generated by modern AI applications, facilitating greater efficiency and performance. Realising these benefits involves navigating a complex landscape of high costs, intricate integration processes, and specialised maintenance requirements.

The shift towards these innovative cooling methods necessitates careful planning and a strategic approach. Investing in advanced cooling infrastructure not only addresses immediate thermal management issues but also positions data centres to handle future technological advancements. As AI continues to drive innovation and expand its applications, data centres must remain agile and forward-thinking, integrating cooling solutions that can evolve with emerging demands.

The effective adoption of these technologies contributes to a broader goal of sustainability. By optimising cooling performance and reducing energy consumption, data centres can minimise their environmental footprint while supporting the growing needs of AI and other high-performance computing tasks. As the industry moves forward, overcoming the challenges associated with advanced cooling will be key to achieving operational excellence and fostering a more sustainable, technologically advanced future.

COOLING | HEATING | VENTILATION | CONTROLS

Ask ME about our IT Cooling solutions for Data Centres

Multi Density

Scan here

Data centres are now a crucial part of modern life and Mitsubishi Electric's **Multi Density** system makes IT Cooling easier than ever. Combining the efficiency of VRF with high-performance close-coupled air conditioning, Multi Density offers precise temperature and humidity control for high-density applications. It ensures optimal conditions for your IT equipment whilst helping increase energy efficiency and lowering emissions.

With features like advanced inverter technology, a small footprint, and pipe runs up to 165m, Multi Density is perfect for cooling high-density racks and blade servers.

Discover more about our **Data Centre Solutions** at: **les.mitsubishielectric.co.uk**

MITSUBISHI ELECTRIC

Balancing the power and sustainability demands of AI

Artificial intelligence (AI) is dominating the digital infrastructure industry and revolutionizing how we interact with technology. From personalized recommendations on streaming platforms to autonomous vehicles navigating our roads, AI is seemingly everywhere, yet still in the early stages of its development. Which poses an important question - how are data center operators managing the impact of AI?

BY KELLEY MULLICK, VICE PRESIDENT OF TECHNOLOGY ADVANCEMENT AND ALLIANCES, ICEOTOPE

MUCH OF THE INITIAL FOCUS has been on capacity as wholesale space in most major global data center markets is limited due to a "land grab" by cloud providers to support AI workloads. What is also emerging are constraints on power infrastructure and the ability to meet sustainability objectives caused by AI. Global headlines across Europe, Asia and the US are showcasing the tension between power, sustainability, and data center growth. The International Energy Agency (IEA) predicted that global electricity demand, driven by AI growth, is set to double by 2026.

This surge in power consumption poses significant challenges for data center operators striving to maintain efficiency, sustainability, and total cost of ownership (TCO). The energy-intensive nature of AI exacerbates carbon emissions and the carbon footprint of data centers, amplifying environmental sustainability concerns. Cloud Service Providers (CSPs) are particularly concerned about TCO optimization as they grapple with the implications of AI on their operations. Similarly, telco operators in Europe and Asia prioritize improving TCO and sustainability while relying on data centers to support AI-driven services.

Data centers must also allocate a greater proportion of their resources to cooling power-hungry CPUs and GPUs to meet the computional demand of AI workloads. Nvidia made headlines with the announcement of its 1200W Blackwell GPU calling it "a new class of AI superchip". The solution is designed to build and run real-time generative AI on trillion-parameter large language models. Because of this kind of compute density required for AI, as well as the overall rising thermal design power of IT equipment and the need for sustainable solutions,

liquid cooling is rapidly emerging as the solution of choice for solving these challenges.

Liquid cooling systems offer a more efficient means of dissipating heat compared to air cooling methods. By circulating a coolant fluid directly over the hottest components, heat is rapidly transferred away, maintaining optimal operating temperatures for AI systems. As chips continue to get hotter, data center operators need to know they are future proofing their infrastructure investment for 1000W CPUs and GPUs and beyond. Choosing technologies that can meet the demands of processor and chip roadmaps and future server generations will be key.

Iceotope Labs recently conducted tests to validate how single-phase liquid cooling technology, like Precision Liquid Cooling, can go beyond the perceived 1000W limit to compete head-to-head with other cooling technologies. Initially, the testing showed that single-phase liquid cooling demonstrated a constant thermal resistance at a given flow rate as the power was increased from 250W to 1000W. More excitingly, a second round of testing found continued consistent thermal resistance up to 1500W – a threshold not yet met within the industry. It is exciting to see these results as it showcases single-phase liquid cooling technology as an indispensable solution for effectively managing the escalating thermal demands of AI workloads in data centers.

Liquid cooling is a leading solution for efficiently accommodating modern compute requirements. Embracing this technology enhances operational efficiency, lowers energy consumption, and aligns with emerging sustainability standards. While much of the market hasn't reached 1500W operation yet, it's poised to do so soon. Liquid cooling efficiently dissipates heat from high computational power and denser hardware configurations, addressing the thermal challenges of AI and optimizing performance, energy efficiency, and hardware reliability. It's indispensable for AI workloads and key to unlocking their future.

Why AI cloud computing beats on-premise infrastructure

The age-old argument of buying vs. leasing has plagued organizations for centuries. From the decision to rent an office rather than purchasing the building, to hiring seasonal workers instead of permanent staff, even to signing up for a monthly rather than annual Adobe Photoshop subscription, everyday business is flush with dilemmas regarding the permanence of places, products and services.

BY DANIEL BEERS, SENIOR VICE PRESIDENT, GLOBAL DATA CENTER OPERATIONS OF ARDENT DATA CENTERS, A PRODUCT OF NORTHERN DATA GROUP

OFTEN, there's no clear-cut answer: leaders must consider factors like payback time, storage space, control over the asset and more. But in the age of AI, during which technology is progressing at a never-before-seen rate, buying participatory tools outright is often an unwise investment. After all, they may become obsolete even before their purchase becomes profitable. Instead, many businesses are choosing to access compute power externally via the cloud.

Let's explore why cloud computing is booming – and how more businesses can charter off-site GenAI capabilities that are as powerful and accessible as on-site infrastructure.

Simple, affordable scalability

Modern AI applications require significant computational resources. But installing infrastructure can prove time-consuming and expensive, often leaving it out of the reach of SMEs. According to IBM, the physical size of an average data center varies between 20,000 to 100,000 square feet. In comparison, the average full-size football pitch equals around 64,000 sq. ft. Meanwhile, a standard GenAI data center's energy requirements range from 300 to 500 megawatts, an amount that could power as many as 500,000 homes.

Simply, operating a data center is a serious undertaking, requiring huge amounts of expensive space and resources, particularly amidst today's high energy prices. Cloud compute providers offer instant access to powerful hardware, which can then easily be scaled up or down based on demand. Then, organizations only need to pay for the resources they use, rather than the 24/7 running and ownership costs of retained infrastructure.

Advanced performance accessibility

The recent semiconductor crisis, in which car production was slashed and PS5s became akin to gold dust, offers a reminder of how supply issues can disrupt progress. Now, the rising demand for GPUs due to a widespread adoption of AI threatens to cause supply chain challenges once again.

According to Nasdaq, Nvidia, the leader in GenAI chipmaking with an estimated market share of 95%, saw huge demand for its H100 GPU. In fact, on its recent earnings conference call, the company said demand for its upcoming flagship H200 and Blackwell GPUs will extend well into 2025.

In some ways, this demand bodes well for Nvidia, and the companies that already own their chips and organizations looking for flexible access to compute power. Many cloud providers have already integrated thousands of advanced GPUs from

top manufacturers like Nvidia, which customers can instantly lease and use. Some providers even enjoy early purchase rights to manufacturers' next-gen models thanks to a longstanding, successful partnership. Customers that partner with these organizations will therefore be able to harness advanced compute power long before their competitors, helping to establish them as market leaders in an increasingly AI-first world.

Prioritization of next-gen technology

However, this AI world is also a murky, unfamiliar one. The industry has seen so much hype and so many headlines that it can be tough for everyday business owners to decipher what's important, what deserves their attention and what should be ignored or avoided. While the technology ostensibly seems to have taken over the world, "if you compare a mature market to a mature tree, we're just at the trunk," Ali Golshan, founder of an AI start-up, told The Washington Post. "We're at the genesis stage of AI."

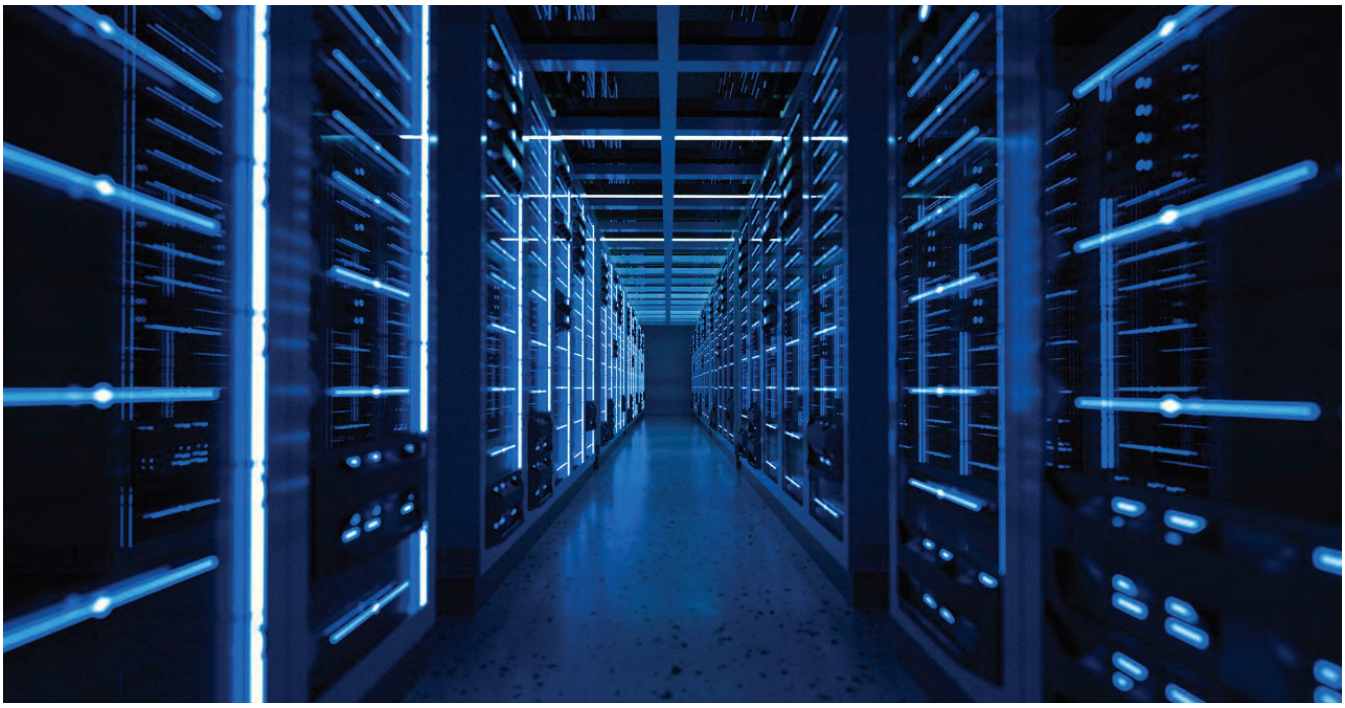
For organizations looking to capitalize on AI, it can therefore be incredibly useful to partner with a specialist provider that has inside knowledge of the industry and technology. Cloud providers regularly invest in the latest technologies first. Their experts can identify the best-in-class hardware needed for customers now and into the future, and purpose-build corresponding data center environments with proprietary performance-optimizing solutions. Similarly, cloud providers invest heavily in the latest security measures to protect data and infrastructure, while handling important maintenance tasks such as software updates to enable customers to freely focus on innovation.

Bringing the best ideas to life

AI is the future of business, but this future remains unpredictable. The technology could progress at an even faster or slower rate than foreseen,

and its impact could be felt at varying levels of consequence. Meanwhile, new laws that aim to put safety guardrails around AI's development are also set to alter its development course. The European Union's AI Act, more comprehensive than the US's light-touch compliance approach, will likely come into force in the summer of 2024. And, according to Ali Golshan of the AI start-up, one of clients' biggest concerns is that strict new AI laws will render their past investments a waste.

This unpredictability underlines the benefits of AI cloud computing. By partnering with a specialist external provider, businesses can access highly-coveted GPUs whenever and however often they like. That way, they can enjoy advanced technology support and realize their previously unachievable innovation goals – all without breaking the bank. You can too. So, why not explore cloud computing today?



How EMI shielding helps data centres manage the load

Data centres and servers are vital infrastructure components, responsible for storing, communicating and transporting massive amounts of information generated daily. They play a key role in everything from facilitating modern conveniences to ensuring mission-critical data exchanges. Maintaining data centres and servers at peak performance is a top organisational priority to avoid the major disruption and costs associated with downtime.

BY JESSE HAGAR, PRODUCT LINE MANAGER, PARKER CHOMERICS

Due to the power of modern data centres and servers, high-frequency electromagnetic interference (EMI) shielding requires careful and proactive management – a consideration more important than ever with today's ultra-fast electronic systems. The latest data centres also feature an increasingly dense population of electronic sub-systems and components. These facilities are 'alive' with a myriad of signals that almost compete for attention. Interaction between signals is a primary cause of cross-talk and interference, which in turn can reduce the quality of electronic devices.

Too much noise

Interference is typically unintentional. As well as signal interaction between internal electronic components, EMI can emanate from a broad spectrum of natural and man-made sources, such as television/radio transmissions, solar magnetic storms, lightning, static electricity, power grid transmission lines, airport radar, mobile phones, factory machinery and switching-mode power supplies to list but a few.

Sometimes, however, interference is malicious. Saboteurs and other unscrupulous individuals can use directed EMI as opportunity to interfere with signals remotely. Whether the interference is intentional or not, EMI shielding is of paramount importance.

The effects of EMI can extend from inconvenient to disastrous – the latter in technologically busy areas like airports and hospitals. System administrators, network engineers and others who work directly with data centre equipment might notice anything from shaky monitor screens and communication errors, to compromised data integrity and even complete system failures where hard drives become scrambled or even wiped completely. Anyone observing a server with unexplained data errors that routine troubleshooting does not resolve

should investigate potential sources of EMI. As a general rule, high-powered devices are more likely to require a higher level of EMI shielding. They will also generate more heat, which can cause device or system failure without the right precautions.

Data centre system designers must find ways to evolve with emerging technologies and utilise products and solutions that can deliver optimal performance in these environments.

Seal of approval

To ensure adequate EMI shielding for servers and data centres, several types of high-performance electrically conductive gaskets are available. Selecting the optimal product for the application is critical.

Gaskets serve to seal any differences between surfaces, providing a complete low-impedance path for EMI signals by addressing all access, egress and ingress points. High frequencies from 100 MHz to 300 GHz are particularly sensitive to small gaps in surfaces.

Soft and flexible gaskets perform well in these instances. Anything too stiff could create gaps instead of preventing them in applications involving doors, lids or panels, for example. The enclosure will subsequently lose its ability to shield against EMI.

Material world

A common form of EMI shielding gasket is a non-conductive low-density foam surrounded by a conductive fabric wrapping (known as fabric-over-foam, or FoF). FoF provides a very economical solution for low-closure force electronic applications that require EMI shielding and electrical grounding.

By combining the conductive properties of an electrically conductive fabric and the conformability of a foam core, FoF gaskets ensure both high conductivity and high cycle reliability at an affordable price. FoF products also offer plenty of engineering flexibility and, importantly for data centres and servers, good protection against the ingress of dust.

Gaskets are available in strip form and a variety of catalogue and custom profiles comprising many foam/fabric combinations. A particular combination of benefit in data centre and server applications is a soft, open-cell urethane foam wrapped in a conductive nickel-plated copper polyester taffeta fabric. The result of this carefully engineered configuration is excellent compression set (15%) in low-closure force applications requiring high shielding effectiveness (>100 dB). Typical uses include server faceplates and blade cards.

Silver lining

Another economical EMI shielding option for data centres and servers is conductive foam. Although widely available, a certain proprietary conductive foam solution contains embedded silver-plated fibres rather than traditional conductive particles. Most conductive foams lose their integrity upon the addition of particles and subsequently fail to bounce back easily, resulting in poor compression set.

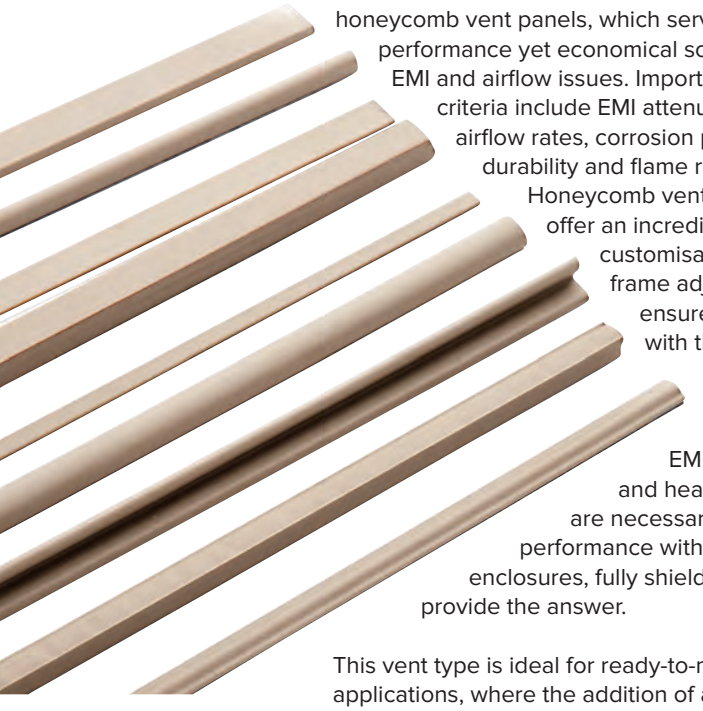
Instead of particles, conductive foams with silver fibres that run their entire length maintain contact from one side of the surface to the other. By embedding these silver-plated fibres into low-density, open-cell urethane foam, users benefit from stable electrical performance and superb conductivity with extremely short ground paths. This type of multiplanar, z-axis conductive foam also delivers great design flexibility, coming in sheets for die cutting to suit many different applications.

Electrically conductive FoF and foams, both of which are available from Parker Chomerics, provide very low compression force because they are soft and allow for a large deflection range, which means they will not damage fragile electrical components within data centres and servers.

Play it cool

Among further data centre challenges is the excessive heat generated by servers and associated devices that operate continuously. Effective heat management is vital because high temperatures can lead to reduced performance and even equipment failure. Ventilation openings are frequently used for the dissipation of heat, either by flowing cooling air into the enclosure or providing an exhaust outlet for hot air. However, despite their crucial role in preventing device overheating, these openings introduce another challenge by providing a pathway for EMI to be emitted from or into the enclosed electronic assembly.

A common pick for enclosure chassis are



honeycomb vent panels, which serve as high-performance yet economical solutions to EMI and airflow issues. Important selection criteria include EMI attenuation values, airflow rates, corrosion prevention, durability and flame resistance. Honeycomb vent panels also offer an incredible amount of customisation, including frame adjustments to ensure proper mating with the chassis.

For applications where maximum EMI protection and heat dissipation are necessary for long-term performance within electronic enclosures, fully shielded vent panels provide the answer.

This vent type is ideal for ready-to-mount applications, where the addition of a compressed gasket along the vent seam supports ultimate EMI shielding.

A further option in low-clearance or space-critical applications are slim-design vent panels constructed of expanded aluminium or Monel® (nickel-copper alloy) foil rather than conventional honeycomb.

Making the grade

High-quality EMI shielding products can only perform their roles if they have the backing of a reliable manufacturing process. Parker Chomerics takes a global approach to production that is not solely reliant on a single manufacturing plant.

With multiple production facilities across different countries, the company can offer the same high-quality products from two or more locations, enabling flexibility, fast delivery and security of supply as demand for these innovative EMI shielding solutions continues to grow.

Today, cloud deployment, the rollout of 5G and various bandwidth-hungry applications are overseeing the construction of more and more data centres. Network system manufacturers must therefore think carefully about EMI. After all, shielded data centres act as a countermeasure against EMI-impeded performance and security issues.

As always, prevention is better than cure. Design engineers need to anticipate EMI in the project's intended environment to prompt the inclusion of optimal protection. Any failure on this part can lead to extensive problems further down the line, with last minute changes and additions proving expensive, not just in financial terms, but also in time and reputation.

wesco | anixter

COMMSCOPE®

Connect

Your Data Centre Priorities are Ours

Delivering value at every stage of data centre development and operations, Wesco Anixter brings reliable expertise alongside a global ecosystem of strategic partners. Together with CommScope, we provide advanced infrastructure solutions to build high-peformance, future-ready foundations for data centres.

Scan to learn more about SYSTIMAX® VisiPORT.

We build, connect, power and protect the world.

Wesco.com/datacenters-emea

The stamp of approval

Certification standards act as key benchmarks for the planning, construction, and management of data centres, ensuring they meet specific criteria for safety, efficiency, and security. These standards are not just about compliance; they offer a roadmap to building a resilient and reliable infrastructure capable of supporting business operations even in the face of challenges. The benefits of adhering to these standards include improve reliability, enhanced security measures, and operational efficiency, all of which contribute to the bottom line of a business. Environments takes a look at data centre certification.

BY CHRIS WELLFAIR, PROJECTS DIRECTOR AT SECURE I.T.

HISTORICALLY, the evolution of data centre certification standards has mirrored the rapid advancements in technology and the growing demands of digital businesses. From basic facility requirements to advanced energy efficiency and sustainability criteria, these standards have continually adapted to support the increasingly complex needs of modern enterprises.

Tiers of joy

The Tier Classification System is a broad classification developed by the Uptime Institute to define the level of redundancy and resilience that an IT infrastructure meets. Depending on the mission critical nature of an infrastructure, it can be used as either a design guide, or a way to clearly explain to customers that it can meet their high availability needs, even when undergoing planned maintenance, or experiencing some sort of failure.

There are four tiers to the standard and each tier is designed such as that each should be able to operate at the next level down when undergoing maintenance. Tier I has is defined as no redundancy, where Tier IV has 2N+1 redundancy of all systems including power and cooling. Tier III is the most seen classification and has an N+1 redundancy arrangement of all systems, though for smaller business a Tier II N+1 arrangement which does not require power and cooling to both have

APC™

40Years

Make greater

IMPACT

with an industry-leading network

Celebrate 40 years with APC.

Start leveraging four decades of uninterrupted protection, connectivity, and unparalleled reliability with the APC UPS family, a legacy marked by pioneering UPS technology and an unwavering commitment to innovation.

https://www.apc.com/uk/en/product-category/88972-uninterruptible-power-supply-ups/

Life Is On Schneider Electric

There can be many different steps to achieving a specific data centre certification, and the requirements can vary a lot. This is something to be aware of when considering which certifications to aspire for

redundancy, may be perfectly adequate for smaller operations depending on the workloads running in that data centre.

You don't need them all

Outside the data centre tiering system there are a number of ISO certification standards that define frameworks and approaches to different aspects of data centre management. Selecting the appropriate certification requires a thorough understanding of your business needs, budget constraints, and industry-specific requirements. It's crucial to align your certification choice with your overall business strategy, ensuring that the certification supports your objectives and enhances your value proposition to customers.

There can be many different steps to achieving a specific data centre certification, and the requirements can vary a lot. This is something to be aware of when considering which certifications to aspire for. Does the time and cost investment justify the benefits. It is critical to have a solid strategy for approaching certifications, so that you are investing in those that will make a difference to your business.

It is also important to realise that maintaining standards is critical too. Once a certification is gained it is critical that operational standards are maintained to keep it. It is much easier to do this in the long run, than to approach re-certification with a poorly run ship and incomplete record keeping. Here are some of the more common certifications that apply to data centres - but remember not all will be relevant to all data centres. If you are designing or upgrading a data centre, having a clear plan on the certifications you need, and want, will be critical to the process.

Sustainability (ISO 14001) – For those that care passionately about demonstrating they take their sustainability and environmental responsibilities seriously this is an essential certification. If your business regards itself as 'green' and runs data centres, this demonstrates a commitment to minimising the environmental impact of a data centre. The standard shows that a company is running efficient environmental management systems (EMS), which can also benefit the organisation by cutting running costs an energy consumption, reducing carbon emissions and lowering waster.

Good management (ISO 9001) – This is a quality management system standard. Its emphasis is on ensuring high standards of service delivery, continuous improvement and that the data centre meets all its customer and regulatory requirements. It is mark of confidence in the way a data centre is run and highly regarded in all industries.

Security management (ISO/IEC 27001) – this standard focuses on information security management, offering a framework for managing sensitive company information. These standards are critical for businesses in industries where data security is paramount.

There are many more certifications that can apply to most data centres. These include ISO 50001, which provides a framework for organisations to improve their energy performance, by helping them implement, maintain and improve energy management systems.

Other standards can apply to specific sectors, for example PCI DSS, which is critical for any data centre that handles payment card information. Other certificates exist from organisations such as the Uptime Institute, with its Management and Operations Stamp of Approval. The key is to understand the requirements of your industry, and the expectations of your users and customers.

Looking forward

Data centre certification standards are more than just a compliance requirement; they are a strategic tool that businesses can leverage for improved reliability, efficiency, and competitive advantage. As the digital landscape evolves, understanding and adhering to these standards will be key to sustaining growth and resilience.

DATA CENTRE TRANSFORMATION

22 · 10 · 24

CONFERENCE ORGANISED BY

datacentre dca alliance

22 OCTOBER 2024

THE IET BIRMINGHAM

(Institute of Engineering & Technology)

www.dca-global.org

t: +44 (0)845 873 4587 e: info@dca-global.org

MEDIA PARTNERS

DATA CENTRE WORLD · DIGITALISATION WORLD · DATACENTRE SOLUTIONS · TECHERATI · DCR Data Centre Review Magazine · WiP

New 200kW Data Centre Load Bank ideal for Integrated Systems Testing

Crestchic has announced the launch of a new 200kW loadbank, which has been built specifically for use in data centres.

With more than 40 years of experience in the design, manufacture, sales and rental of load banks and power testing solutions, Crestchic have extensive experience in the data centre sector. Following comprehensive customer research, the team has launched a stackable 200kW loadbank, designed to make heat load testing of server halls and electrical infrastructure simpler, faster, and more efficient.

Paul Brickman, Commercial Director at Crestchic, explains, “Having served the sector for many years, we have been able to collaborate with key customers to get to the crux of what they require from a load bank. By taking that experience and insight back to our engineering team, we have been able to build a solution that fits an exacting specification. The result is a machine that provides an ideal solution for integrated system tests, is easy to operate, and easy to store. As higher data centre density becomes the norm, testing cooling systems and performing integrated systems tests is increasingly important - this product meets that need.”

At just 1159mm x 876mm x 983mm the load bank has a small enough footprint for operators to link multiple machines and strategically place them at intervals around the hall. This makes it possible to closely mimic the heat generated by servers, replicating airflow patterns and rack configuration such as hot and cold isles. Multiple load banks can be configured and controlled using just one LC10 handheld digital controller, enabling fully remote operation and simple management of incremental loads. This allows operators to use up to 20 linked load banks (4MW) to verify that the cooling installation is operating to specification in the server hall and within safe parameters, to check the redundancy of the system should a cooling unit fail, and to verify that there are no hot spots that may require reconfiguration.

In addition to heat load testing, the units are invariably connected to Power Distribution Units (PDUs), facilitating testing of the data centre’s electrical infrastructure, including power to servers, switches, bus track power units, feed tracks, tap offs, and other devices. With heavy duty casters, the new 200kW data centre loadbank has been built for easy manoeuvrability and can be positioned at each PDU in the server hall, for full system testing.

The units have one of the lowest Delta T rises on the market at below 100°C (at the exhaust face), helping to ensure a more accurate simulation of real-world conditions, and providing a more accurate test of the cooling system’s ability to handle the actual heat load it will encounter during operation. The load bank boasts a compact, stackable design, enabling the units to be safely stacked two high, reducing the storage footprint when stored. They have also been built to ensure that there are no awkward shapes or protrusions, which could be subject to knocks and bumps. Noise levels are lowest in class at 83dBA at 1m, an important consideration when operating inside

Paul adds, “With the sector now expecting 99.999% uptime as standard, it is important that data centre operators test their critical systems. Using these units for integrated systems testing pushes the systems to their limits by simulating real-world scenarios like full load operation, power outages, and cooling system failures. As well as helping to identify any weaknesses or bottlenecks before critical IT equipment is deployed, the test generates valuable data on system performance under different loads, providing a baseline for future reference that can support maintenance and troubleshooting down the line.”

Crestchic offers a range of solutions for testing backup diesel generators and UPS; heating, ventilation, and cooling (HVAC) systems; electrical bus and distribution cabling, and PDUs. Load testing at commissioning stage, and at regular maintenance intervals, can help to ensure that systems remain operational and reduce the risk of unwanted and expensive downtime.

200kW units are available for short and long-term rental. They are also available for immediate purchase, with delivery by January 2025.

For more information, visit: **crestchic.com**

The DCA Update

THIS ISSUE of Data Centre Solutions focuses on a number of key elements in the sector – DCIM / DC Management, Cloud and Liquid Cooling. DCA Members and Partners have contributed three great articles for our Update.

I'd like to personally thank Assaf Skolnik, CEO RiT Tech, Frida Cullin Persson, Msc, Global Industry Segment Manger, SWEP and Dr Jon Summers, RI.SE Sweden & DCA Advisory Board Member for taking the time to write these articles.
I do hope you will find them of interest.

Come and meet The DCA in September & October.

Platform Global 2024
10 – 12 September, Antibes
DCA representatives will be attending the event and representing DCA Members and Partners. If you would like to arrange a meeting please drop me a message – Steveh@dca-global.org

DCD Connect London
17 – 18 September 2024, Business Design Centre, London
The DCA will be in the Association Pavilion, ready to meet you and tell you more about how we support the UK DC sector, and the benefits we provide to our Partners.

The DCA's Data Centre Transformation Conference
22 October 2024, The IET, Birmingham
The DCA are delighted to be hosting this event, now in its 14th year. The focus will be on 'The Future of Data Centres' – our agenda is designed to provide strategic insight of what's on the horizon and advice on how the industry should be preparing for this. There are opportunities to get to meet other delegate, speaker and the organisers during the networking lunch, drinks and dinner.

Tickets for the event are limited so don't miss out! Click here find out more and to register for Data Centre Transformation 2024!

Find out more about The DCA or email us mss@dca-global.org

Best regards
Steve
CEO, DCA

The resurgence of DCIM: Navigating the future of data centre management

BY Assaf Skolnik – CEO, RiT Tech

DATA CENTRE INFRASTRUCTURE MANAGEMENT (DCIM) was first introduced almost 15 years ago and due to a range of early adoption challenges, it has remained largely overlooked until now.

DCIM has re-emerged as a critical tool for organisations seeking to optimise their data centre operations, enhance efficiency, and stay competitive in a fast-evolving industry where data centre models are shaped by forces other than just power and capacity, for example, demand for sustainable practices and reporting in line with EED and CSRG legislations and the rapid rise of new technologies like Artificial Intelligence (AI).

The early days of DCIM

Early DCIM solutions focused on purely monitoring - power usage, cooling and troubleshooting of physical assets within the data centre. Its initial role was to help data centre managers manage resources more effectively. However, these early tools were often limited in scope, siloed into different departments, and lacked integration with other critical IT management systems.

As a result, the adoption of DCIM was slow and many organisations struggled to realise its full potential.

In recent years, however, DCIM has undergone a significant transformation and continues to evolve, to the level where Gartner now have DCIM back on the agenda. Today DCIM solutions are far more comprehensive, seamlessly integrating with other operational management tools such as BMS, CMDB, ITSM CRM or BPM - Business Process Management tools to provide a holistic view of data centre operations. Standing at the forefront of this transformation are innovations such as Universal Intelligent Infrastructure Management (UIIM), a concept initially created by RiT Tech, a provider of DCIM solutions. UIIM represents the next generation of DCIM, pushing the boundaries of what is possible for infrastructure management.

Unlike earlier iterations, UIIM integrates advanced DCIM capabilities, analytics, AI-driven automation, and real-time monitoring into a unified platform. For example, in the instance of financial institutions,

the move has been to adopt UIIM to ensure their data centre operate with maximum efficiency, security, and compliance. These institutions handle vast amounts of sensitive data, where uptime is critical, and regulatory requirements are stringent. UIIM's ability to provide predictive analytics and proactive monitoring means that potential issues can be identified and addressed before they lead to downtime or security breaches.

Driving forces behind the DCIM resurgence

Increasing Data Centre Workload Requirements
A significant factor behind the resurgence of DCIM is the exponential growth in data generation and the requirement for more infrastructure capacity. Businesses, consumers and devices are producing data at unprecedented rates, driven by trends such as cloud computing, digital transformation and the Internet of Things (IoT).

This influx of data has created a critical demand for advanced tools that can offer comprehensive visibility into resources and infrastructure. Organisations are increasingly seeking DCIM solutions that enable them to efficiently scale their data centres to handle this growth while maintaining optimal performance. These tools are essential for managing capacity, ensuring uptime, and optimising the allocation of resources across both physical and virtual environments.

The rise of AI

The rise of AI is not just adding to the volume of workloads in data centre but is also changing their nature, requiring more sophisticated infrastructure, real-time processing capabilities, and advanced storage solutions to keep pace with the evolving demands. Modern DCIM solutions have evolved to meet this challenge by integrating with AI and providing these capabilities. As a result, DCIM has become indispensable for organisations looking to leverage these new technologies while maintaining control over their data centre operations.

Growing sustainability and environmental requirements

Sustainability has emerged as a central concern for organisations, particularly those operating large data centres across the UK and Europe. The introduction of stringent regulations such as the Corporate Sustainability Reporting Directive (CSRD) and the Energy Efficiency Directive (EED) are driving this shift.

These regulations are forcing organisations to reduce their carbon footprint, improve energy efficiency, and adopt sustainable practices. Modern DCIM solutions, like RiT Tech's, XpedITe, are equipped with features designed to track, analyse, and reduce energy consumption in alignment with standards such as ISO/IEC 30134. These capabilities not only help organisations meet regulatory requirements but also contribute to broader corporate sustainability goals, making DCIM a vital tool in the push toward greener data centre operations.

The role of DCIM in data centre management today

One of the most significant advancements in modern DCIM solutions is the ability to provide real-time monitoring and analytics. This allows data centre operators to monitor critical metrics such as power usage, temperature and resource utilisation in real-time, resulting in the ability to identify and address any issues before escalation into a major problem.

Therefore, reducing costly downtime and maintaining optimal data centre performance. Modern DCIM solutions, such as RiT Tech's XpedITe, also leverage AI and machine learning to provide predictive maintenance capabilities. By analysing historical data and identifying patterns, it will predict when equipment is likely to fail and automatically schedule maintenance ahead of any failure as well as providing automation of routine tasks such as resource allocations.

As data centres continue to grow in size and complexity, effective capacity planning becomes increasingly important. DCIM solutions provide the tools needed to plan and optimise capacity, ensuring that data centre resources are used efficiently and that there is sufficient capacity to meet future demand. By providing detailed insights into resource utilisation, DCIM helps organisations avoid overprovisioning or underutilising their data centre infrastructure.

The role of industry bodies in DCIM

The resurgence of DCIM has also been supported by initiatives from industry bodies and publications such as the Data Centre Alliance (DCA). The DCA has played a crucial role in promoting best practices, setting standards, and fostering innovation within the data centre industry.

By advocating for the adoption of advanced DCIM solutions and emphasising the importance of sustainability and efficiency, the DCA has helped accelerate the adoption of modern DCIM tools across the industry. This collaboration has been instrumental in driving the evolution of DCIM from a niche tool to a core component of data centre management.

The future of DCIM

The resurgence of DCIM marks a pivotal moment in the evolution of data centre operations. It is not just a response to the current challenges facing data centres; it also reflects the growing importance of data centre infrastructure in the digital economy. As organisations continue to generate and process vast amounts of data, the need for efficient, reliable, and scalable data centre infrastructure will only increase.

The Artificial Intelligence energy paradox to data centers

By Frida Cullin Persson, Msc, Global Industry Segment Manger, SWEP

ENERGY CONSUMPTION from data centers continues to rise, and with artificial intelligence (AI), energy demand will increase even more. However, AI also holds the potential to monitor and manage energy usage much more efficiently. As such, the role of AI in data centers therefore presents a paradox: energy consumption is significantly increased, AI also offers solutions for better energy management in data centers.

To fully comprehend this paradox with drawbacks and opportunities, here are some key elements to consider:

- Machine Learning, High-Performance Computing and similar AI based algorithms are data and energy intensive and require high computational power capability.
- Data storage and data management across systems becomes paramount to AI and is also energy intensive.
- Data centers require reliable operations 24/7, and many AI applications require continuous and real time processing, which also leads to continuous energy usage.
- With AI, it will also be possible to foresee equipment failure and take action for preventive maintenance.
- Utilising AI, it is possible to analyze energy consumption patterns in real-time and predict peak-time usage so that inefficiencies and potential areas for improvement can be identified and energy distribution can be optimized.

Subsequently, as AI advances, it poses both challenges and opportunities. One inevitable consequence that will impact datacenters is that heat generated by higher power density to perform AI related task will change cooling needs and require advanced and efficient cooling systems.1

Liquid cooling for the future

Data centers are estimated to account for 3-4% of global power consumption by the end of this decade.2 While AI can help predict heat generation patterns and adjust cooling mechanisms in real-time, their high-TDP chips generate more heat and require more efficient cooling systems. TDP (Thermal design power) refers to the maximum amount of heat a CPU or GPU is expected to generate under general usage and traditional air-cooled systems are becoming inadequate as the power increases. The data center liquid cooling market size is expected to see exponential growth in the next few years. It will grow to $10.61 billion in 2028 at a compound annual growth rate (CAGR) of 25.8%.3

Liquid cooling holds a promise for the future because of its superior advantage in cooling capacity. Water has more than four times the cooling capacity when compared to air and outperforms air at heat transfer conductivity by 25 times. Removing excess heat with water requires 50 times less energy than with air. It's clear that liquid cooling becomes part of the answer as we move towards more power dense data centers.

Liquid cooling holds a promise for the future because of its superior advantage in cooling capacity (Source SWEP)

One of the main components of a liquid cooled system is the brazed plate heat exchanger, also known as BPHE. BPHEs are the most energy efficient heat exchangers on the market and are used in these systems to quickly and effectively transfer heat energy away from the processors.

Pledge for climate neutrality accelerates shift
Besides the inherent challenge that AI will increase the energy usage of Data centers there are also other challenges. New energy goals and regulations are driving the push towards climate neutral data centers. One of the main goals in the EU is to make Europe climate neutral by 2050 and data centers climate neutral by 2030.4 This goal is shared by the CNDCP, CISPE and EUDCA. This ambitious goal is forcing data center owners to explore new ways of cooling data centers and capturing and reusing excess energy whenever possible.

The CNDCP is the Climate Neutral Data Centre Pact, a consortium of over 100 data center operators and trade associations, including Google, AWS, IBM, Microsoft, Meta, among others, that are accelerating the shift to climate neutral data centers.

The five pledges of the Climate Neutral Data Centre Pact include:

- Purchasing 100% carbon-free energy
- Prioritizing water consumption and conservation
- Reusing and repairing servers
- Proving energy efficiency with measurable targets
- Looking for ways to recycle heat.

While AI inevitably will drive-up energy consumption, there is no doubt that AI will also accelerate a shift for solutions that can contribute to fulfill the ambitious climate neutral goals for data centers. With superior cooling capacity and heat transfer conductivity, liquid cooling will certainly play its part as we move towards more power dense data centers. The goal for carbon neutral data centers by 2030, while ambitious, seems within reach.

REFERENCE

- https://www.goldmansachs.com/intelligence/pages/AI-poised-to-drive-160-increase-in-power-demand.html
- https://time.com/6987773/ai-data-centers-energy-usage-climate-change/
- https://www.globenewswire.com/en/news-release/2024/02/14/Data-Center-Liquid-Cooling-Market-Forecasted.html
- https://www.climateneutraldatacentre.net

- **About SWEP**
  SWEP has been designing and developing brazed plate heat exchangers (BPHEs) for more than 40 years to deliver effective and efficient heating and cooling in countless business- and vital liquid cooling systems for HVAC, industrial and residential applications. SWEP BPHEs are installed in data centers around the world and used in various systems such as the coolant distribution unit, for immersion cooling but also for applications outside the white space such as free cooling, chillers and as heat recovery.
  To learn more, visit **swep.net**

Liquid cooling of the immersion kind

Dr Jon Summers, Scientific Lead in Data Centres at the ICE datacenter of RISE, Sweden & DCA Advisory Board Member

dca INDUSTRY EXPERT

HAVING BEEN on numerous panels and given presentations on liquid cooling at various events it is time to write an informed opinion piece about immersion liquid cooling.

Immersion cooling uses dielectric liquid coolants in which servers are submerged in a contained enclosure. The coolants have a high dielectric strength that protects all immersed microelectronics from short circuits and whose prime function is to transfer the heat away by convection. Variations include natural, forced or phase change convection. Some configurations pump the heated dielectric out of the enclosure (usually a tank) to an external heat exchanger. Other arrangements use targeted cooling, pushing the dielectric coolant to where it is required and certain targeted solutions use low volumes of coolant and are not in tanks, but rack mounted enclosures.

There is another solution which uses engineered dielectric coolants that can change phase at low temperatures and boil off the hot components in the immersion tank, notedly two-phase immersion that operates by the liquid becoming vapour bubbles which move rapidly to the tanks surface by buoyancy. This approach has engineering complexity but offers greater heat transfer rates due to the latent heat of vaporisation – such systems require the dielectric vapour to condense and return to the tank in liquid form, therefore the tank must be well sealed.

The adoption of immersion cooling so far depends on the sector. The pure scientific high-performance computing (HPC) systems have adopted liquid cooling, but rarely immersion as far as I have seen. Not saying that HPC has not tried immersion cooling, they have. One market segment that has seen the adoption of immersion cooling is in the cryptocurrency data centres and as such exemplifies undoubtedly one of the benefits of lower total cost of ownership, less cooling infrastructure and where network and storage is low volume.

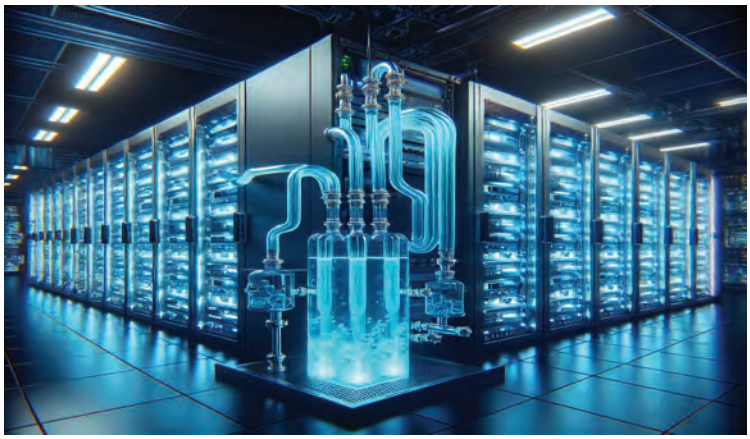
Apparent growing markets include those that require dense compute systems in hot and dusty climates and some anecdotal evidence that edge data centre deployments could make use of immersion cooling. All digital infrastructure is potentially poised to deploy immersion cooling that needs to contend with denser compute and as a result could fly the flag of operational efficiency and improved sustainability. However, these require explaining in more detail. First it is important to note that the power required by the servers is dictated by its digital workload, but also it can draw more power at higher operational temperatures. There is the so-called microprocessor case temperature that needs to be monitored, but there is an interesting system efficiency dilemma – operating servers hotter leads to leaner cooling systems that can operate all year round, but the servers draw more power due to current leakage. The initiated reader will know that this has a marked effect on the so-called ISO PUE

(Power Usage Effectiveness, ISO 30134-2), evaluated as the annualised measure of consumed energy by the entire facility divided by the annualised measure of the energy used by the servers.

Operating immersion cooling with higher temperatures helps this energy ratio value get closer to the ideal 1.0, but it may mean that the infrastructure consumes more energy in total. Easy to see that the ratio 100/80 versus 95/75 is smaller, but the numerator 100 is greater than 95 despite the ratio being lower. In our recent open access journal paper in Applied Thermal Engineering entitled "Experimental and numerical analysis of the thermal behaviour of a single-phase immersion-cooled data centre" we highlighted experimentally an increase in energy requirements of 6% when the dielectric coolant operating setpoint is increased from 30oC to 50oC, however it should be noted that different models of microprocessors have different current leakage behaviour.

The deployment of immersion cooling has two different viewpoints: one from the data centre facility perspective and the other from the IT systems configuration. From the facility point of view, they could be largely plug and play if there is already a chilled water circuit in the whitespace. The arrangement of the immersion tanks in the whitespace would depend on where the connection points to the chilled water loop are located. If the whitespace has used rear door or in-row cooling equipment, then these connection points could be in the middle of the whitespace, although pipework sizes need to match expected heat loads and chilled water flow rates for immersion system's coolant distribution unit (CDU). On the other hand, if the whitespace has used downflow computer toom air handling (CRAH) units at the edge of the room then the tanks could be placed against the wall of the whitespace. All of this presupposes a raised floor and a lot of data centre build out has opted for a slab floor – an arrangement that challenges all variations of liquid cooling not just immersion.

With slab floors the connecting pipework for the liquid cooling path would have to be at the walls of the whitespace or overhead (a difficult option for the risk averse). If the deployment is with a couple of immersion tanks these could go into a whitespace that is geared up for air cooling and an adjacent air-assisted liquid cooling unit could be deployed next to the tanks which essentially transfers the thermal energy into the air stream, and this would then be cooled by the existing air-cooling infrastructure. From the server angle things are fast changing. It is recognised that most standard air-cooled IT equipment can be immersed into the tanks, but one should consider the material compatibility and how the dielectric coolant interacts with all the components. For example, standard cables usually contain polyvinyl chloride (PVC), such as network and power cables, which leach over time into the dielectric coolant and therefore it is



advisable to use non-PVC cables. There is also the issue with the thermal interface material (TIM), squeezed under pressure to make a good thermal connection between the microprocessors and their heatsinks or heat spreaders, which for air-cooled servers is usually a thermally conductive paste. For immersion cooling the thermal paste would need to be replaced with a thermal pad of a compatible material, either thin sheets of indium or graphene have been used successfully. Equally the heatsink or heat spreader should be replaced with those that are optimised for immersion cooling.

In addition, there is the fact that the air-cooled servers have fans (air blowers) which must be removed, but the servers need to be configured to boot without fans - not usually a problem with modern servers. Servers have many labels on all components, which will come loose in the tank unless the provider of the server has used an adhesive that is compatible with the dielectric coolant. Many of the dielectric coolant suppliers have done extensive material compatibility analyses to help the immersion cooling technology advance. After all of this, the question of the server warranty arises. Is this covered by the tank provider or will the provider of the server? Naturally this leads to the alternative which is to approach providers that have immersion ready servers, in fact they should be, and in some cases are, in partnership with the immersion tank providers to offer the full system via partnerships, that indeed removes certain headaches.

Lastly, the largest hurdle for the deployment of immersion cooling systems is operator perception and perhaps resistance to change. Training and education would help, but it sometimes boils down to "religion", personal beliefs, but the full facts really do need to be presented. I would say that the open compute project (OCP) initiative is one of the major driving forces behind standards and drivers for the adoption of immersion cooling technologies. I urge the interested reader to visit opencompute.org and under "Projects" go to the "Cooling Environments" project where there is a sub-project called "Immersion". Here one can participate and contribute to the open meetings as well as access the sub-project wiki to get access a whole host of materials that support the adoption of immersion cooling.

APC™

40Years

Make greater

IMPACT

with an industry-leading network

Celebrate 40 years with APC.

Start leveraging four decades of uninterrupted protection, connectivity, and unparalleled reliability with the APC UPS family, a legacy marked by pioneering UPS technology and an unwavering commitment to innovation.

https://www.apc.com/uk/en/product-category/88972-uninterruptible-power-supply-ups/

Life Is On Schneider Electric