



Issue 57 • June 2025
datacenterdynamics.com

Also inside:
Building Stargate
Quantum data centers
The cloud wars

A FRENCH REVOLUTION

The fight to become Europe's AI infrastructure capital

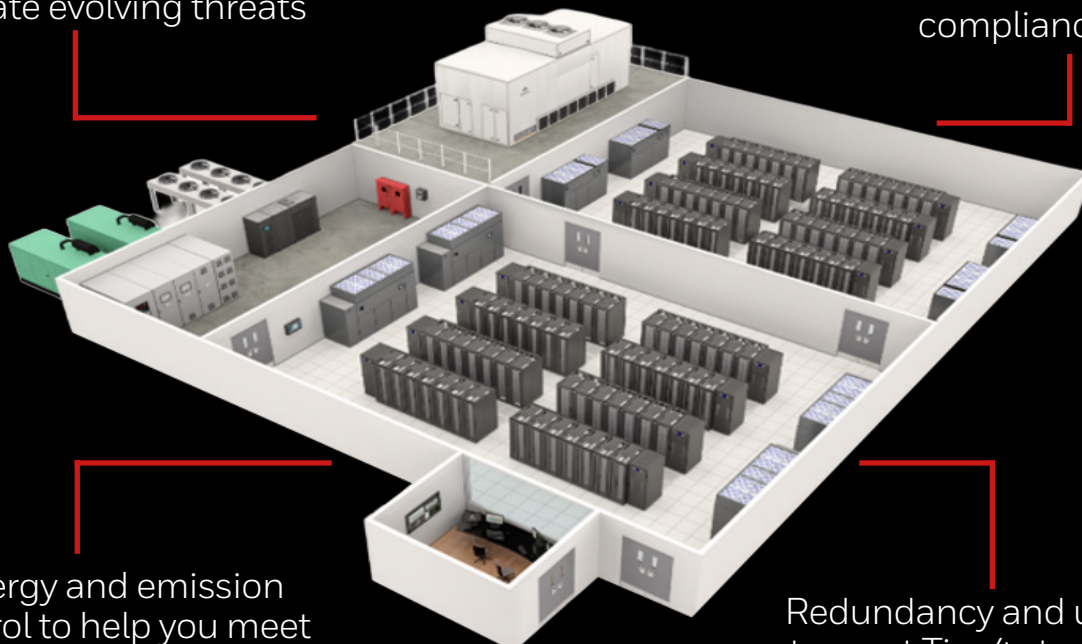
TOMORROW'S DATA CENTRES TODAY

Smart automation, integrated safety, and scalable security for resilient energy-efficient data centres

As AI, cloud, and edge computing grow, so do the demands on your data centre's resilience, efficiency, and security. Honeywell integrates building management, fire safety and physical/cyber security into one seamless system, offering:

Unified access control to mitigate evolving threats

Advanced fire detection and remote compliance tools



Energy and emission control to help you meet EED and net-zero targets

Redundancy and uptime to meet Tier 4 standards

Honeywell has deployed systems for top-tier hyperscale and colocation data centres in six continents. By leveraging advanced building intelligence, we can help you increase infrastructure and accelerate velocity at scale.

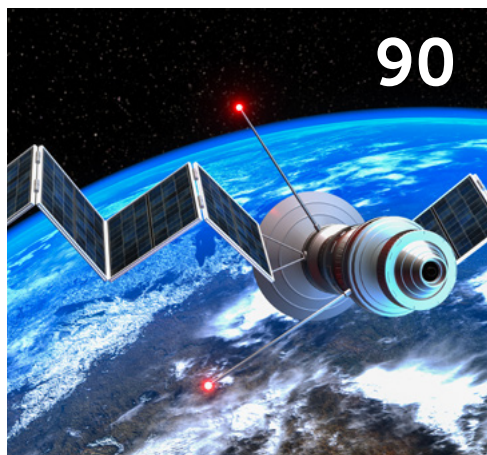
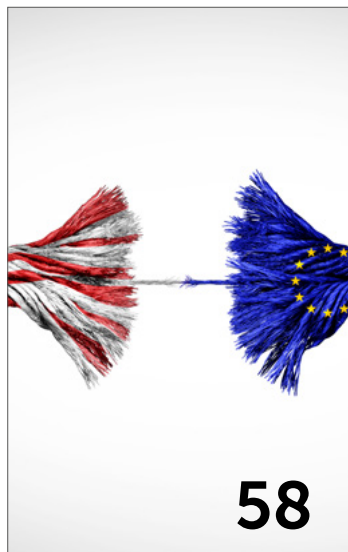
Elevate your data centre operations today

Honeywell



Contents

June 2025



6 News

15 A French revolution?

France wants to become the AI capital of Europe, but can it outpace its FLAP-D rivals?

22 Inside a quantum data center

DCD visits quantum data centers from IBM and IQM in Germany

27 The Cooling supplement

Chilling innovations

43 Geochemistry at hyperscale

Microsoft is betting on Terradot's enhanced rock weathering tech to deliver durable, verifiable carbon removals at a global scale

46 Independence Day

GDS International has rebranded as DayOne, and CEO Jamie Khoo has plans to grow across Asia and beyond



49 A global CIO 44 years in the making

How Craig Walker, former downstream VP and CIO of Shell, grew through the decades

54 A stroll down Meet-me Street

Towardex believes Meet-me-vaults can disrupt the data center industry's hold over cross-connects

58 Fighting in a cloudy arena

The battle to increase competition in the cloud computing market

65 Making Connections: The pursuit of chiplet interconnect standardization

The Universal Chiplet Interconnect Express technology and what's next for this vital piece of the chiplet puzzle

70 Building Stargate

OpenAI's director of physical infrastructure on gigawatt clusters, inferencing sites, and tariffs

76 An opportunity too good to ignore: Nokia, AI, and data centers

Better known for its mobile networking business, the vendor wants to cash in on the artificial intelligence boom

80 Watt's Next? How can batteries be best utilized in the data center sector?

A deep dive into the many use cases of Battery Energy Storage

85 Connecting the cloud

A look at the networking inside AWS data centers

90 The changing landscape of Satellite M&A

With big-LEO entrants disrupting satellite connectivity forever, consolidation is happening rapidly

94 OFC 2025: Hollow Core Fiber hype stands out amid the AI overload

The credentials of hollow core fiber were examined by the industry, as the industry grapples to power the data-hungry AI future

98 Emergency, what emergency?



Strategic data center partner

Expert delivery. Global Scale. Local precision.

Equans Data Centers is a strategic data center partner, delivering new build, fit-out, retrofit, and upgrade covering general contracting for hyperscaler and colocation requirements.

Bringing construction and engineering, global standardisation, and precision localisation, we deliver the highest quality, speed, and scalability, meeting increasing capacity needs across Europe.



Find out more

equansdatacenters.com

From the Editor

You had me at bonjour

The race for AI supremacy has governments around the world lauding multi-billion dollar investments from some of the biggest names in tech, and France is no exception.

French fancies

Our cover feature (p15) looks at the French bid to become Europe's capital of AI infrastructure. The nation's plentiful space - and access to low-carbon nuclear energy - is a potentially heady cocktail for data centers, if a host of bureaucratic hurdles can be overcome.

Making French connections

Germany's quantum leap

Over the border in Germany, they're getting serious about quantum data centers, with the likes of IBM and IQM (and possibly other companies starting with I) setting up facilities to cater for a new type of IT hardware. But what makes a quantum data center? DCD got the inside track (p22).

Storm clouds gather

Away from AI and quantum, companies across Europe are gobbling up cloud services like never before, and the hyperscale providers are cashing in. But, as revenues grow, concerns about lack of competition in the market are also on the rise, and efforts are underway to try and create a more level playing field (p58).

Day break

On the other side of the world, tensions between the US and China running higher than ever. So it's perhaps no surprise that Chinese data center operator has decided to spin off its international operations into a separate business. We spoke to CEO Jamie Khoo about her firm's plans to conquer APAC (p46).

Nok nok, who's there?

Those of you who only remember Nokia as the company behind the 3310 mobile phone may need to think again, as the firm has big ambitions in the data center, as we find out on p76.

Reach for the Stars

The biggest - in the very literal sense - data center story of the year so far has been the emergence of OpenAI's Stargate project, a grand plan to build enormous AI data centers across the US and beyond. But what's the reality behind the hype? We spoke to one of the men charged with bringing Stargate to life.

Rocking out

Efforts to decarbonize data centers have seen providers funding a vast array of novel technologies. One of the latest, which is being championed by Microsoft, among others, is enhanced rock weathering (p43), which utilizes a natural process to trap carbon in rocks. But is it a viable option to curb emissions?

Plus:

A cooling supplement, plans for universal chip interconnects, M&A in the satellite market, and much more!

€110 Billion

Investment pledged for French data centers in 2025



Sebastian Moss
Editor-in-Chief

Publisher & Editor-in-Chief
Sebastian Moss
Managing Editor
Dan Swinhoe

Senior Editor
Matthew Gooding
Telecoms Editor
Paul 'Telco Dave' Lipscombe

CSN Editor
Charlotte Trueman

C&H Senior Reporter
Georgia Butler

E&S Senior Reporter
Zachary Skidmore

Junior Reporter
Jason Ma

Head of Partner Content

Claire Fletcher
Copywriters
Farah Johnson-May
Erika Chaffey

Designer
Eleni Zevgaridou

Media Marketing
Stephen Scott

Group Commercial Director
Erica Baeta

Conference Director, Global
Rebecca Davison

Live Events
Gabriella Gillett-Perez
Tom Buckley
Audrey Pascual
Joshua Lloyd-Braiden

Channel Management Team Lead
Alex Dickins

Channel Manager
Kat Sullivan
Emma Brooks
Zoe Turner
Tam Pledger

Director of Marketing Services
Nina Bernard

CEO
Dan Loosemore

Head Office
DatacenterDynamics
32-38 Saffron Hill,
London, EC1N 8FH

© 2025 Data Centre Dynamics Limited All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, or be stored in any retrieval system of any nature, without prior written permission of Data Centre Dynamics Limited. Applications for written permission should be directed to the editorial team at editorial@datacenterdynamics.com. Any views or opinions expressed do not necessarily represent the views or opinions of Data Centre Dynamics Limited or its affiliates. Disclaimer of liability. Whilst every effort has been made to ensure the quality and accuracy of the information contained in this publication at the time of going to press, Data Centre Dynamics Limited and its affiliates assume no responsibility as to the accuracy or completeness of and, to the extent permitted by law, shall not be liable for any errors or omissions or any loss, damage or expense incurred by reliance on information or any statement contained in this publication. Advertisers are solely responsible for the content of the advertising material which they submit to us and for ensuring that the material complies with applicable laws. Data Centre Dynamics Limited and its affiliates are not responsible for any error, omission or material. Inclusion of any advertisement is not intended to endorse any views expressed, nor products or services offered, nor the organisations sponsoring the advertisement.

Dive even deeper

Follow the story and find out more about DCD products that can further expand your knowledge. Each product is represented with a different icon and color, shown below.



Events



Intelligence



Debates



Training



Awards



CEEDA

News



Virginia narrowly avoided blackout when 60 data centers dropped off grid

Sixty data centers in Northern Virginia using 1,500MW of power dropped off the grid simultaneously last summer, forcing the network operators to take drastic action to avoid widespread blackouts in the area.

The near-miss incident, revealed in regulatory filings and first reported by *Reuters*, saw the data centers in Fairfax County all switch to backup generators en masse as a result of an equipment fault on the grid.

Grid operator PJM Interconnection and local utility company Dominion were forced to quickly scale back the volume of energy going into the network from power stations. If left unattended, such a rapid increase in the amount of available power could have triggered a surge and caused systems to trip out, potentially leading to blackouts across the state.

Northern Virginia is the world's busiest data center market - home to hundreds of facilities and gigawatts of capacity - and continues to attract new developments despite constraints on the grid.

According to an incident report from the North American Electric Reliability Corporation (NERC), the trouble started at 7pm EST on July 10, 2024.

"A lightning arrester failed on a 230kV transmission line in the eastern

interconnection, resulting in a permanent fault that eventually 'locked out' the transmission line," the NERC report said. A lightning arrester is a piece of equipment designed to protect against power surges.

This led to a series of short supply disturbances, lasting a matter of milliseconds before the system corrected itself, but it was enough to cause data centers in the area to automatically switch to their backup UPS systems as a precautionary measure because "data center loads are sensitive to voltage disturbances," the report from NERC said.

"Discussions were held with data center owners to understand the specific cause of their load reductions," it continued. "It was determined that the data centers transferred their loads to their backup power systems in response to the disturbance."

NERC's investigation also discovered that if a series of faults happen in a short time period, the backup systems at many data centers do not automatically switch from UPS back to the main grid supply, and have to be changed manually.

The data center load did not return to the grid for hours, the report said.

Though voltage "did not rise to levels that posed a reliability risk," operators "did have to take action to reduce the voltage to within normal operating levels," NERC added.

NEWS IN BRIEF

QTS founder and CEO Chad Williams steps down

QTS founder and CEO Chad Williams is leaving the company after 20 years at the helm. He will be replaced by COO David Robey and chief growth officer Tag Greason, who will be co-CEOs.

Nautilus cans planned data center in Maine

Nautilus has canceled its planned 60MW data center in Millinocket, Maine. The company quietly put its floating barge data center in California up for sale last year.

CoreWeave starts trading on Nasdaq

AI cloud firm CoreWeave has gone public and is now listed on the Nasdaq. The company posted Q1 2025 revenues of \$971.63m in its first results since its float, though reported an operating loss of \$27.47 million.

DigitalBridge could be acquired by 26North

Digital infrastructure investor DigitalBridge could be acquired. A report from IJGlobal suggests alternative investment management firm 26North is in advanced talks to acquire the firm, though the companies are yet to confirm or deny the reports.

AWS data center blockaded by protesters in Canada

Activists blockaded an AWS data center in Quebec, Canada, in May in protest at the company's decision to lay off more than 4,000 staff in the province. The Alliance Ouvrière say it prevented 20 staff getting to work, but the operations of the data center were not affected.

Hyperscalers prepare for 1MW racks

Google has joined Meta and Microsoft's collaboration project on a power rack the companies hope will help them reach rack densities of 1MW. The companies are working on a new power rack side pod called Mount Diablo that will enable companies to put more compute into a single rack. Specs will be released to OCP this year.



France sets out its stall in the gigawatt era

A number of companies have outlined plans to further expand France's rapidly expanding data center market.

France saw gigawatts of data center capacity announced in February as part of an AI Action Summit in Paris.

In May, France further established its data center growth credentials as part of the Choose France Summit in Versailles.

Stargate-backer MGX, alongside French national investment bank Bpifrance, French AI firm Mistral AI, and Nvidia announced the formation of a joint venture to establish what they claim will be Europe's largest AI campus in France.

Located outside Paris, the campus is set to ultimately reach a capacity of 1.4GW. Construction is expected to begin in

the second half of 2026, with operations launching by 2028.

The same summit saw France's President Emmanuel Macron announce further investment plans, including data center developments from Prologis and others.

"The American group has decided to invest €6.4 billion to expand its logistics facilities in Marseille, Lyon, Paris, and Le Havre, and to create data centers in the Paris region," President Macron said.

Prologis reportedly aims to develop four major data center projects in the Paris region, representing a total capacity of 584MW.

"These developments are part of Prologis' broader European strategy, which already includes 435MW of secured capacity

- including 184MW in France - with an additional 400MW targeted by the third quarter of 2025," the French Government announcement said. "Full commissioning is planned for 2035."

Macron also noted that Brookfield was "accelerating its commitment to France" and had confirmed plans for a €10 billion (\$11.3bn), 1GW data center development in Cambrai - part of a €20 billion (\$20.7bn) investment pledge first announced in February and set to be channelled through its Data4 unit.

UAE firm G42 is partnering with DataOne to expand the latter's Grenoble campus to host AMD AI hardware for G42's Core42 AI subsidiary. DataOne was established last year by BSO, taking over two former DXC sites.

February also saw GPU cloud provider Fluidstack sign a Memorandum of Understanding with the French government to build an AI supercomputer in the country.

Fluidstack also signed a multi-year agreement with French modular data center operator Eclairion to host a new training cluster for Mistral in the Essonne region, south-west of Paris.

February also saw French cloud computing provider Sesterce announce it is investing €450 million (\$471.85m) into developing an AI data center in Valence.

European data center firm Evroc announced its intention to build a 96MW facility outside Cannes.

See page 15 for more on France's AI infrastructure revolution.

Saudi investors flex financial muscle



Saudi Arabia has made several large-scale data center announcements, centered on a new state-linked AI venture known as Humain.

First announced in May, Humain is an AI-focused subsidiary of Saudi Arabia's Public Investment Fund (PIF).

The company initially announced plans to develop up to 500MW of data center capacity alongside Nvidia. The GPU giant will first deliver 18,000 GB300 chips for a single supercomputer, with "several hundred thousand" more superchips in the pipeline over the next five years.

Humain then announced another 500MW deal with AMD to develop data center capacity across Saudi and the US. That deal was said to be worth up to \$10 billion.

Though it hasn't made any official announcement, Humain's website also notes it offers access to AI accelerator chips from startup Groq. Groq first deployed a cluster in Saudi Arabia in December 2024, and this February announced a \$1.5 billion deal to expand in the country with a Dammam data center.

Humain has also partnered with Amazon Web Services to develop a \$5 billion to build an 'AI Zone' in the Kingdom - reportedly separate from AWS' existing region in Saudi.

The PIF-affiliate is now reportedly seeking a US equity partner, and is targeting to build out some 6.6GW of capacity by 2034.

AWS and Microsoft pull back from some data centers, reaffirm data center capex for 2025



Microsoft and Amazon have pulled back from a number of data center leases and paused some developments.

"Over the weekend, we heard from several industry sources that AWS has paused a portion of its leasing discussions on the colocation side (particularly international ones)," Wells Fargo analysts wrote in a note published in April.

Kevin Miller, VP of data centers at AWS, downplayed the pullback in a LinkedIn post, calling the move "routine capacity management" with no fundamental changes in our expansion plans.

Amazon has previously said its total capex for 2025 - including non-data center activities - would reach \$100 billion.

The news came shortly after a series of reports suggesting Microsoft had backed off from a number of leases globally.

After previously reporting that Microsoft had walked away from some 200MW of leases, TD Cowen published a note saying the Redmond company had cancelled around 2GW of data center projects in Europe and the US.

The analyst noted Meta and Google had taken on some of the leases and capacity in Europe, though the companies are yet to comment on the matter.

In both cases, Microsoft provided the same comment: "Thanks to the significant investments we have made up to this point, we are well positioned to meet our current and increasing customer demand. Last year alone, we added more capacity than any prior year in history.

"While we may strategically pace or adjust our infrastructure in some areas, we will continue to grow strongly in all regions. This allows us to invest and allocate resources to growth areas for our future."

Microsoft has maintained that it is on track for the \$80bn spend on data centers planned for 2025.

In the company's most recent earnings call in May, Microsoft CEO Satya Nadella told analysts that, during the quarter, Microsoft "opened DCs in 10 countries across four continents" and noted that they continue to expand data center capacity. This quarter also saw the company announce plans to increase European data center capacity by 40 percent over the next two years.

Both companies noted lower capex spending for the most recent quarter, with Microsoft dropping \$1 billion and Amazon lowering spend by \$2 billion.

In its own earnings call in the wake of Amazon and Microsoft's pullback, Google CEO Sundar Pichai said the search giant's own capex for 2025 remains on track to be around \$75 billion for the year.

xAI deploys 168 Tesla Megapacks in Memphis

xAI has installed 168 Tesla Megapack battery energy storage units at its data center in Memphis, Tennessee, to power its Colossus supercomputer.

The company has integrated the Megapacks to manage outages and demand surges, which xAI claims will bolster the reliability of the data center. It's unclear whether the batteries will remain for the duration of the data center's operational life or represent a stopgap solution for the facility.

xAI has caused significant controversy over recent months, after it was revealed that the company had installed 35 onsite methane gas turbines at the Memphis supercomputer. The turbines had a combined capacity of 422MW, more than double the permitted amount.

Following remonstrations from local community and environmental groups, which argued that the presence of the turbines was in violation of the Clean Air Act, xAI removed an undisclosed amount of the turbines, after a new substation came online.

Half of the turbines are expected to remain to power the second phase of the project until a second substation, which is currently under construction, is completed. The substation is expected to be completed sometime during Fall 2025, at which time the gas turbines will be relegated to the role of backup power.

AT&T acquires Lumen's fiber business, Cox and Charter to merge

US telco AT&T is to acquire Lumen's mass market fiber business for \$5.75 billion. Announced in May, all-cash transaction is expected to close in the first half of 2026.

AT&T noted that it will acquire 95 percent of Lumen's fiber business, which covers around 1 million customers and reaches more than four million fiber locations across 11 states. The company has previously said it plans to reach more than 50 million fiber locations by 2029; following the announcement of this deal, AT&T has upped that target to approximately 60 million total fiber locations by the end of 2030.

The same month, Charter Communications and Cox Communications announced plans to merge in a deal worth \$34.5 billion. The combination of the two companies will create the biggest cable operator in the US with 69.5 million locations passed. This will be than Comcast, which ended the first quarter of this year with just shy of 64 million.

Cox Enterprises will own approximately 23 percent of the combined entity, which will change its name to Cox Communications, with Spectrum becoming the consumer-facing brand.



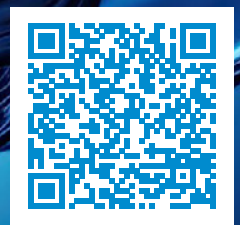
Introducing the Munters LCX: Hyperscale by design

Designed for liquid-cooled servers, our new LCX liquid-to-liquid Coolant Distribution Unit is built for efficiency, reliability, and ease of serviceability. It provides precise control of the technology fluid supply temperature and external pressure or flow, suitable for rejecting heat from cold plates, in-rack CDUs, and other liquid heat rejection devices.

What sets Munters apart is our ability to customize components, such as brazed plate heat exchangers and pumps, to optimize performance for each installation. Available in sizes from 500 kW to 1.5 MW, the LCX ensures efficient, scalable, and serviceable liquid cooling, tailored to the demands of modern high-density data centers.



Learn more at munters.com



Learn more about
Munters LCX



Equinix and DRT suffer data center fires

Several data centers suffered fire incidents in April and May.

In April, a fire at an Equinix facility in Brazil, impacting the IX.Br Brazilian Internet Exchange.

The "small" fire occurred at 12:34pm local time on Sunday March 30 at the colo firm's SP4 facility in São Paulo.

The fire was in an unnamed dark fiber provider's cage, causing a dark fiber link outage. The facility was evacuated, water was discharged, with services brought back online later that day.

Local reports suggests the fire was caused by packaging kept in turned-on equipment that had recently arrived at the data center. Equinix has not commented further.

A Digital Realty facility in Oregon also suffered a fire incident in May.

A fire broke out on Thursday, May 22, at a data center in Hillsboro leased by Elon Musk's X social media firm, causing a major global outage.

According to the Wired report, the fire was related to a room of batteries. Hillsboro Fire and Rescue spokesperson Piseth Pich said that the fire had not spread to other parts of the building, but the room was full of smoke.

Digital Realty operates two facilities across the Portland area totaling 56,500 sqm (609,000 sq ft) of floorspace.

OpenAI expands Stargate to UAE



Generative AI firm OpenAI has expanded its Stargate data center project internationally, targeting a facility in the UAE.

May saw G42, OpenAI, Oracle, Nvidia, SoftBank Group and Cisco announce Stargate UAE, a 1GW compute cluster.

The site, to be built by G42 and operated by OpenAI and Oracle, will be located within a planned 5GW AI campus in Abu Dhabi.

Set to be equipped with Nvidia's GB300 GPU systems, the first 200MW is set to go live in 2026.

The 10-square-mile UAE-U.S. AI Campus which will house Stargate UAE was announced in May by Sheikh Mohamed bin Zayed Al Nahyan, president of the United Arab Emirates, and US President Donald Trump.

OpenAI's Stargate project is a \$500bn effort to build massive data centers - originally across the US, but now globally - for the AI developer.

The likes of Oracle, SoftBank, and Abu Dhabi's MGX are named investors in the venture.

Crusoe is developing a large campus for Stargate in Texas. The Abilene campus, owned by Lancium, is expected to eventually have eight buildings and a total of 1.2GW of capacity.

Construction of the first phase, featuring two buildings and more than 200MW, began in June 2024 and is expected to be energized in the first half of 2025.

Construction of the second phase, consisting of six more buildings and another gigawatt of capacity, began in March 2025 and is expected to be energized in mid-2026. Oracle has agreed to lease the site for 15 years.

Most of the funding for the project has been provided by JPMorgan.

OpenAI is also reportedly exploring other Stargate data center options in states across the US including Arizona, California, Florida, Louisiana, Maryland, Nevada, New York, Ohio, Oregon, Pennsylvania, Utah, Texas, Virginia, Washington, Wisconsin, and West Virginia.

Beyond the US, recent reports suggested that OpenAI was looking to develop up to ten Stargate data center projects across the world. With one set to go to the UAE, OpenAI's is reportedly also looking at locations in Asia-Pacific.

As reported by Bloomberg, OpenAI is exploring locations in the APAC region, with the company's chief strategy officer, Jason Kwon, set to meet with government officials to discuss AI infrastructure and the AI provider's software offering.

Among the countries on Kwon's list are Japan, South Korea, Australia, India, and Singapore.

Dan's Data Point

Delivery backlogs for gas turbines are beginning to stretch past 2029 amid skyrocketing demand from data center customers. Three companies - GE Vernova, Siemens, and Mitsubishi Heavy Industries - currently produce the majority of gas turbines globally.



Nvidia launches new Lepton cloud service for GPU access



Nvidia has launched an AI platform, bringing together its GPUs from various global cloud providers.

Dubbed Nvidia DGX Cloud Lepton, the platform and compute marketplace connects GPUs from providers including CoreWeave, Crusoe, Firmus, Foxconn, GMI Cloud, Lambda, Nebius, Nscale, SoftBank, and Yotta Data Services.

Users of DGX Cloud Lepton can access GPU compute capacity for both on-demand and long-term computing within specific regions to support sovereign AI needs.

"Nvidia DGX Cloud Lepton

connects our network of global GPU cloud providers with AI developers," said Jensen Huang, founder and CEO of Nvidia.

"Together with our partners, we're building a planetary-scale AI factory."

Nvidia acquired server rental company Lepton AI in April 2025, with the GPU provider taking over the company that leases GPU servers from cloud providers and rents them to its own customers. Nvidia's existing DGX Cloud offering is currently offered within GCP, Microsoft, Oracle, and AWS.



DATA CENTRE SOLUTIONS

Structural Ceiling Grid For Data Halls

Engineered to your project requirements

- ✓ BIM designed and value engineered
- ✓ Higher loads: liquid cooling ready
- ✓ Faster installation: prefabricate offsite
- ✓ More sustainable: lower CO2 footprint
- ✓ The usual Hilti support!

End-to-end support

Hilti supports construction projects from start to finish through our Integrated Project Solutions: BIM design and engineering, best-in class support systems and fasteners, and trusted on-site support.



- Accommodates heavy loads above and below the grid
- Cost-saving through direct attachment to the ceiling. No
- Simple and fast assembly with the modular MT system. The grid can be prefabricated offsite.
- Seismic design available



NTT takes NTT Data private, announces DC REIT



NTT plans to take its IT services and data center subsidiary NTT Data private in a 2.37 trillion yen (\$16.4 billion) deal.

The Japanese conglomerate announced in May that it intends to purchase the entirety of Tokyo-listed NTT Data's outstanding share capital, and will pay 4,000 yen (\$27.65) per share. Currently, NTT owns 57.7 percent of NTT Data.

Operating in 20 markets around the world, NTT Data says it runs more than 150 data centers.

It is thought NTT wants to bring the firm back in-house to speed up decision-making and help it capitalize on the rapidly growing demand for AI infrastructure.

NTT Data's origins can be traced back to 1967, when Japan Telegraph and Telephone Public Corporation founded a

data division.

That company was privatized in 1985, becoming NTT, and it spun off its data business into a separate company, NTT Data, in 1988. NTT Data now claims to be the biggest ITSP in Japan.

The same month, NTT separately announced aims to establish and list a new data center real estate investment trust on the Singapore Stock Exchange, to be seeded with six NTT-owned data centers.

REITs are companies that own, and often operate, income-producing real estate such as apartments, retail outlets, offices, and data centers. They act as a fund for investors, generating revenue via leasing space and collecting rent on their properties. Several major data center firms – including Equinix, Digital Realty, Iron Mountain – operate as

REITs. Digital Realty also set up a listed data center REIT – known as Digital Core REIT – in Singapore seeded with a number of stabilized facilities from its portfolio.

NTT said it aims to transfer six NTT Limited-owned data centers to a proposed Singapore real estate investment trust NTT DC REIT. The facilities will be sold for approximately 240.7 billion yen (\$1.573bn).

The facilities set to be transferred to the REIT include data centers in Ashburn, Virginia; Sacramento, California (x3); Vienna, Austria; and Singapore. The facilities total more than 41,000 sqm (441,320 sq ft) and around 80MW. Occupancy rates vary from around 90 percent up to 97 percent.

NTT aims to list the REIT on the Singapore Exchange, but will retain a share of the new company. The company added it could sell other NTT-owned data centers to the REIT in the future.

In a busy month, May also saw NTT acquire landbanks across the world for future build-outs of totaling almost 1GW.

The company acquired land in seven global markets including Oregon, Arizona, Milan, Frankfurt, London, Tokyo, and Osaka.

"By bringing new capacity to high-growth regions, we're building the foundation enterprises need to innovate, scale, and lead confidently in an AI-driven economy," said Doug Adams, CEO and president, Global Data Centers, NTT Data.

In January 2025, NTT committed to spend more than \$10 billion on data centers globally.



Iron Mountain takes over India's Web Werks

US colo firm Iron Mountain has officially acquired all of Indian operator Web Werks.

The companies have been working together for the last four years as part of a joint venture to build data centers in India, and Iron Mountain has increased its investment to the point where it has 100 percent ownership of the business, which will now operate

under the Iron Mountain brand.

It gives Iron Mountain an Indian portfolio of six data centers located in five Indian markets - Mumbai, Bangalore, Hyderabad, Pune, and Noida - with total IT capacity of 14MW.

The company is also developing three new campuses, in Mumbai, Chennai, and Noida, which have a potential capacity of 142MW between them.

In a busy quarter, April saw Stack divest its European colocation data centers to investment fund Apollo.

The deal will see Apollo acquire seven data centers in five markets - Stockholm, Oslo, Copenhagen, Milan, and Geneva - and form a new, independent company to manage them. Financial terms have not been disclosed. Stack retains a number of hyperscale campuses across the continent.

April saw Colt Technology Services this week announced the divestment of six facilities in Germany and the Netherlands to DWS-owned NorthC, as well as two in the UK to DWS separately.

The data centers were part of the assets Colt gained with its acquisition of Lumen EMEA in 2023.

Terms of the deal were not shared.



Accelerate your data center expansion plans in Northern Europe now



Klaipėda industrial park

CAPACITY: 60MW so ready that capacity charge is incurred, upgradable to 110MW

POWER: Renewable
Reliable
N+1 redundancy
Solar plant and BESS opportunities

COOLING: DLC, RDHx and immersion cooling solutions
Free cooling
N+1 redundancy
No water use, sustainable approach

DESIGN: Modular, scalable

ATTRACTIVE PRICE & Substantial Lithuanian governmental incentives

Additional 100MW+ plots, located near high-voltage power stations, offer regional scalability.

[More information about the project](#)
or please contact egle.melinske@sentris.eu

Sentris by KV Baltic - sustainable efficiency

DESIGN
BUILD
CONSULT

Sentris has a long experience in both leading and emerging data center markets. Our solutions meet the highest reliability standards (Uptime Institute). Designed for efficiency, they deliver top performance in both high and low-power-density applications.

A man and a woman are in a data center, looking at a laptop. The man is on the left, wearing a grey sweater over a white shirt and a blue lanyard. The woman is on the right, wearing a white shirt and a blue lanyard. They are both looking at the laptop screen. The background shows server racks and a blue-tinted lighting.

PRACTICAL SOLUTIONS TO AI INFRASTRUCTURE CHALLENGES

With customizable solutions and collaborative engineering, see how Legrand's approach to AI infrastructure can help your data center address:

- Rising power supply and thermal density
- Heavier, larger rack loads
- Challenges with cable management and connectivity
- Increasingly critical management and monitoring

Click on the URL below to find out more about our solutions for resilient, agile AI data centers.

www.legrand.us/ai-data-center

#LegrandImprovingLives



A French revolution?



Matthew Gooding
Senior Editor

France wants to become the AI capital of Europe, but can it outpace its FLAP-D rivals?

"France has been a 'Cinderella' market over the last ten years,"

>>Keith Breed
CBRE

In the grounds of Château de Bruyères-le-Châtel, a 19th-century castle located southwest of France's capital city, Paris, sits a collection of buildings known as La Lisiere.

Since 2015, La Lisiere (which roughly translates to The Edge) has provided a low-cost space for exhibitions and performances, as well as a retreat for artists looking for somewhere quiet and reflective to stay while they get their creative juices flowing.

It is somewhat ironic, then, that less than 100 meters away from La Lisiere, a new data center is rising that will host one of the AI systems considered by many to present an existential threat to the creative industries.

The data center, being built on an adjacent plot by a new French operator, Eclairion, will host the first dedicated cluster of GPUs for Mistral, developer of the Le Chat chatbot and the company often lauded as Europe's best home-grown hope of challenging the big US AI labs such as OpenAI and Anthropic.

Plans for Mistral's new cluster were revealed in February, one of a string of infrastructure-related investments in France announced at the International AI Summit hosted by its President Emmanuel Macron.

Billions of Euros are set to pour into France over the next decade to fund a massive data center build-out that could create gigawatts of new capacity. And, in



theory, France, which has lagged behind the other major data center markets in Europe in recent years, is well placed to capitalize on the AI boom, with plentiful space and nuclear power, and a president in Macron who has made growing the nation's digital economy a priority since he first came to power in 2017.

However, some familiar barriers will need to be overcome if France is to outstrip its European rivals in the race for AI infrastructure supremacy.



Emmanuel Macron, friend of data centers

There's Paris, then there's the rest

As befits one of Europe's largest economies, France already boasts an active data center sector.

Paris represents one of Europe's Tier One, FLAP-D markets, alongside Frankfurt, London, Amsterdam, and Dublin, though when it comes to the volume of data centers, France as a whole is playing catch-up with some of its rivals. Data Center Map lists 265 data centers in France, compared to 427 in Germany and 425 in the UK.

Looking at data center locations within the country, the common French saying "there's Paris, then there's the rest" has never been more appropriate. The capital city accounts for 97 of the 265 facilities listed on Data Center Map, with no other hub coming close, and when it comes to the IT capacity of these facilities, the divide is even more stark. "Paris dominates to a large extent," says Keith Breed, associate director in the research division of CBRE. "About 85 percent of all capacity in France comes out of the Paris area, particularly to the south of the city, where permitting and power are easier to come by."

Elsewhere, Marseille's status as a landing point for 12 subsea Internet cables, with five more set to come into service over the next two years, means it has attracted some data center developments. The market is dominated by Digital Realty, which has been active in the city since 2014 and operates four data centers.

Breed says France has often been overlooked by investors because of perceived barriers to doing business. "France has been a 'Cinderella' market over the last ten years," he says. "It's been a slight laggard compared to the other FLAP-D locations, but over the last two years, that has started to change. And in 2024, Paris overtook Amsterdam in

"I think France has a lot of potential,"

*>>Hedi Ollivier,
Colt DCS*

terms of size and supply, so it's now the third-largest market in Europe. That was a significant moment."

This trend was also noted by Olivia Ford, research analyst covering France and Benelux for DC Byte. "2024 was the year we saw a bit of catch-up in France," she says. DC Byte's research from Q1 2025 shows that there is currently data center capacity of 283MW under construction in France, with 1.8GW of early-stage projects, with the pipeline having swelled since further since then. Overall, France is "arguably less mature" than the other FLAP-D markets, Ford says. But she adds: "It's looking very healthy and compared to Amsterdam, for example, it has many larger scale developments going, particularly around AI data centers."

Indeed, France briefly became the center of the AI universe in February, when the country hosted the International AI Summit. Alongside myriad photo opportunities for world leaders - and bilateral talks between slightly baffled-looking officials rapidly learning the difference between CDUs and GPUs - came a slew of announcements about investment in French digital infrastructure. In true government style,

several of these were reheated versions of previously unveiled schemes, but plenty of new cash was also promised.

Private companies promising to invest in France during the summit included Abu Dhabi-based G42, which is planning to install AI infrastructure at a data center in Grenoble, while the United Arab Emirates government signing an agreement with their French counterparts for a €30-€50 billion (\$34-\$62bn) AI investment in France, including a new 1.4GW data center. Details of this were revealed at an investment summit in May, with French national investment bank Bpifrance, UAE investment fund MGX, Nvidia, and Mistral forming a JV to deliver the data center, which will be located on an undisclosed site outside Paris and could be operational by 2028.

Mistral had already announced it was setting up its first dedicated AI training cluster in France. The company, which has raised over \$1 billion across several funding rounds, had previously relied on infrastructure from Microsoft Azure and Google Cloud to train its models, but will now be working with Eclairion and French GPU cloud provider Scaleway to bring an initial 18,000 Nvidia GB200 GPUs online later this year.

In total, the French government claimed to have secured €110 billion (\$112bn) for digital infrastructure. This dwarfs the funds being committed to rival FLAP-D markets: By comparison, the UK government, which is also on a mission to attract data center operators, says it has brought in £25 billion (\$32bn) in investment since it took office in July 2024.

On the frontline

Ready to take a first-hand look at this French revolution, DCD arrives in Bruyères-le-Châtel, a sleepy commune of just over 3,000 people located in the Essonne department on the southern fringe of Paris, where Eclairion's first data center will be located.

The site is a hive of activity, with construction workers milling around and a large group of investors shuttling in to meet the company's CEO, Arnaud Lépinois.

The site itself is unusual in both shape and appearance. A long, narrow, plot, it is dominated by four enormous - and



Eclairion's data center pods will sit on this platform

currently empty - steel girder platforms, each spanning 2,500 sqm (26,910 sq ft), which Eclairion's Charles Huot says will eventually house the company's modular data center solution.

Huot, the company's head of development, is overseeing the Bruyères-le-Châtel build, meaning he is tasked with everything from ensuring construction crews are on task to checking on the health of the 1,500 saplings that have been delivered to the site ready for planting. He is a gracious and patient host, even as *DCD* manages to lose a pair of safety shoes halfway through the site tour.

Eclairion was founded in 2021 and is funded by HPC Group, an investment firm with its roots in the hospitality industry. The company initially had more modest ambitions for the Bruyères-le-Châtel data center, Huot explains. "When we started in 2021, 1MW was a good amount of power," he says. "We had 10MW available then, ready to deliver to ten containers, and we were ready to receive customers."

"Now, Mistral will take all the power we have, but will only use half the space."

As the AI revolution took hold, Eclairion scaled up its plans, and Mistral will occupy two of the data center's platforms, which will be supplied with 40MW of renewable energy via two dedicated substations run by grid provider Enedis. Its hardware, provided by Scaleway, will be housed in 66 large

containers, each drawing 600kW and featuring 20 racks.

Putting the containers on a platform will enable the data center to be more flexible, and make running power and cooling systems to the modules, as well as performing maintenance, a much simpler task, Huot says. The space underneath can also be used to house additional equipment or for storage, depending on a client's requirements. The data center's liquid cooling system sits at the side of the platforms, ready to be connected as and when modules arrive.

Eclairion also believes the design offers sustainability benefits, making it easy for old containers to be taken away and put to use elsewhere. The first container on the site, a demonstration unit featuring regular CPU racks, is already on its second life, having previously been deployed at

Renault as part of the automaker's crash testing system for new vehicles.

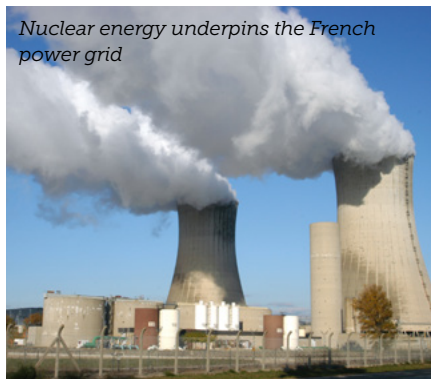
"We think our design is unique in the world," Huot says. "When we explained it to the engineers, they found it difficult to understand the concept."

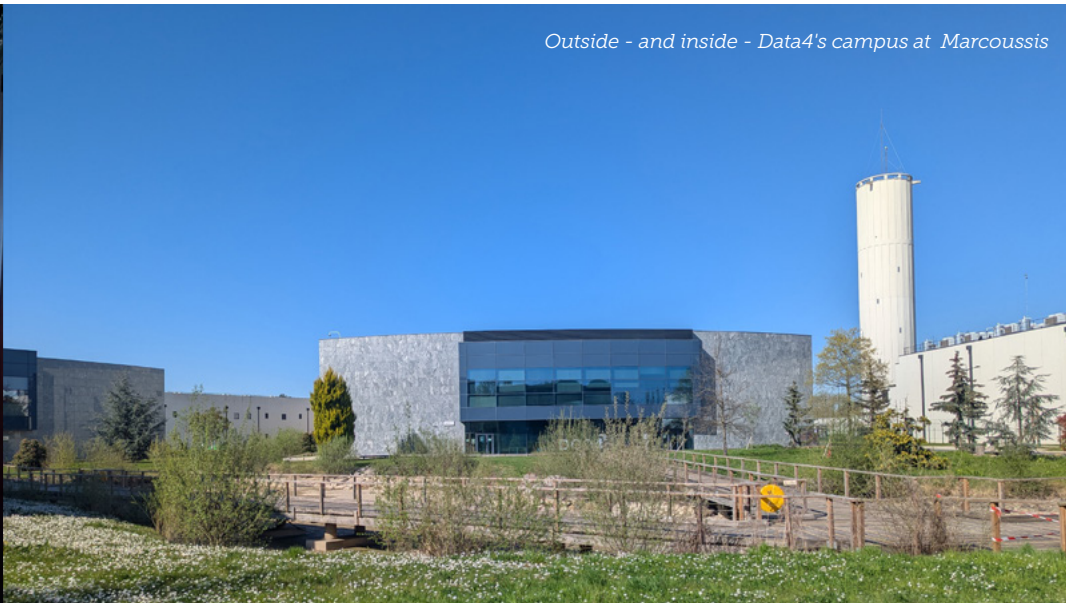
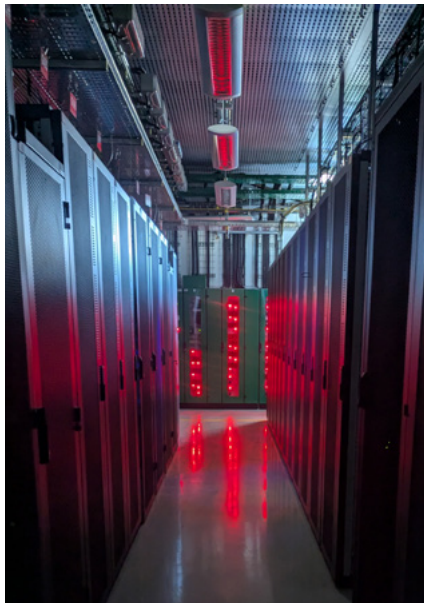
The Mistral deal was forged in summer 2024, with the French government reaching out to data center providers to help ensure that the AI lab's compute requirements could be met in France, rather than forcing it to move elsewhere in Europe. "No one else in France could find 40MW in 2025," Huot says. "We were the only option."

Accommodating Mistral has presented some logistical challenges for Eclairion; when *DCD* visits, a group of workers are busy digging out some of the tarmac under the platform where the GPUs will be housed, because the power transformers required are 30cm bigger than older generations and would not have fitted in the existing space. GPUs are scheduled to arrive on site as this magazine goes to press, and the data center is due to be operational by August.

Once Mistral is in situ, Eclairion will turn its attention to its other platforms. It has another 60MW of power guaranteed to arrive on site by 2027 via a different provider, RTE, and a queue of customers - mostly French companies deploying AI systems - ready to take space in future.

Nuclear energy underpins the French power grid





Outside - and inside - Data4's campus at Marcoussis

Twelve kilometers from Bruyères-le-Châtel, in the town of Marcoussis, another data center company has big plans to cater for the AI revolution. Data4's vast Paris campus is already home to 21 data centers, with that number set to increase to 25 by 2027. In total, the site has up to 250MW available.

Accompanying *DCD* to the roof of one of the buildings, Jérôme Totel, Data4's group strategy and innovation director, points across the fields into the distance, where several cranes are visible on the horizon. This is the site of the company's new AI campus at Nozay, 5 km from its existing site and housed at the former Nokia France headquarters, which Data4 purchased in 2023. This will also be served by 250MW of low-carbon energy.

"The first data center will be up and running in 2027," Totel says. "It will be completely dedicated to AI workloads and feature direct liquid cooling throughout." The company already runs some AI servers at Marcoussis, and Totel says the rapid pace of the technology's development means AI tasks will be performed across both campuses. The company provides colocation space for some of the biggest businesses in France, as well as catering to the needs of the hyperscalers.

"France needs a lot of data center space," Totel says. "We hope we will continue to see demand from French companies, and we want to offer our existing customers space to grow here. But we're also looking to other parts of the country."

"In the long term all the data centers in the Paris area are going to be short of power"

*>> Rogier van der Wal
Digital Realty*

Data4, which is owned by investment fund Brookfield, is headquartered in France but operates in markets across Europe. At the AI summit, Brookfield pledged €20 billion (\$20.7bn) for French data centers over the next five years, €15 billion (\$17bn) of which will be funneled into Data4. This will fund development in Paris, as well as at other locations: The company has identified a site in Cambrai, northern France, which is set to host a 1GW facility. Located on the site of the former Cambrai-Épinoy airbase, work on this data center could begin in 2026.

Totel says Data4's AI campuses are likely to be in locations like Cambrai, rather than close to the capital city, because the nature of AI workloads means they don't need to be close to the end users. "We're probably going to build the large campuses outside Paris," he says. "At the moment, we don't see a lot of AI applications that benefit from low latency."

While both Eclairion and Data4's

campuses are set in relatively rural locations, Digital Realty's Paris Digital Park, in La Courneuve, is just 7 km from Paris city center and surrounded by homes and businesses. The impressive circular construction resembles an enormous wheel of cheese divided into four pieces, but is actually four interlinked data centers catering for Digital's enterprise and hyperscale clients.

Rogier van der Wal, senior director at Digital Realty, tells *DCD* the company has developed Paris Digital Park because it is seeing "strong demand" from enterprise customers in France.

When it comes to AI, Van der Wal says that while most of the firm's French clients are consuming AI services through their cloud providers, some businesses in the country are getting more ambitious. "There is a set of customers that are mature enough to operate their own AI infrastructure," he says. "They have the skill set to run GPUs, to write training algorithms, to do the training themselves, and then to use their infrastructure for inferencing."

"These tend to be larger, more mature, enterprises, who are knocking on our door and saying 'we're in an old data center that can't run the very dense workloads we require - can you help?'"

Colt DCS is another data center firm investing heavily in France, and broke ground on its second data center in France in May 2025.

The facility, Colt Paris 2, is the first

"The government thinks it can reduce the time for planning approval from 18 to nine months"

>> Olivia Ford,
DC Byte



of three data centers planned for a 12.5-acre site in Villebon-sur-Yvette, southwest of Paris. It is part of a €2.3 billion (\$2.58bn) investment in French infrastructure, which will see five data centers constructed across two sites by 2031, bringing Colt's IT capacity in France to 170MW.

Hedi Ollivier, Colt's director of development for the EMEA region, is bullish about the market's prospects. "The power grid is very strong and in terms of network we're very well connected," he says. "And there's plenty of land available. So I think France has a lot of potential, there's a lot of capacity available, and I don't think we're likely to see the kind of problems we've seen in Ireland or the Netherlands, where it can be difficult to start building and get connected to the grid."

Plug baby, plug

Central to France's pitch to data center companies is its plentiful supply of low-carbon power, underpinned by its nuclear power plants. It has been a net exporter of electricity for many years.

"Last year [2024] we exported 90TWh, which means we can localize a lot of data centers on top of the electricity we need for our companies and households," Macron told delegates at February's AI summit, before referencing US President Donald Trump's famous pro-fossil fuel mantra. "I have a good friend across the

ocean saying 'drill baby, drill,'" Macron added. "Here, there's no need to drill, it's just plug baby, plug."

France has 17 operational nuclear plants, which are home to 57 reactors, all run by state-owned utility company EDF and providing 61GW of capacity. This is the second-largest fleet of active reactors in the world, with only the US possessing more nuclear power. Nuclear accounts for 60-75 percent of French energy generation on any given day, according to RTE's power tracker, and the grid runs almost entirely on low-carbon sources, with hydro, solar, and wind installations providing the bulk of the rest of the nation's electricity. By contrast, neighbors Germany and the UK typically source less than 20 percent of their power from nuclear plants.

This French affaire de coeur with nuclear stems from a government policy geared around energy sovereignty dating back to the 1970s. Many of the plants in service today first came online in the 1980s, and are now approaching the end of their lives, meaning a modernization program is underway. That initiative aims to increase the amount of renewables powering the French grid and reduce the reliance on nuclear so that it accounts for nearer to 50 percent of the nation's energy needs.

As the energy mix in France changes, EDF is having to adjust its operations accordingly, with the country's ARENH mechanism for nuclear power

procurement set to come to an end in 2026. ARENH was conceived in 2011, and set a fixed price of €42 per MWh (\$46.90) for purchasing nuclear power from EDF within a volume limit of 100TWh per year. It was designed to encourage competition in the market and enable French consumers to benefit from the lower prices associated with nuclear energy, while also providing EDF with a steady income that it could use to maintain the nuclear plants.

Views on whether the policy has been a success are mixed (EDF has broadly maintained its market share since it was introduced) but now that it is coming to an end, the price of buying nuclear energy from EDF is on the rise, and the state-backed utility could be left with an ARENH-shaped hole in its business model.

"When ARENH ends, there will be a move to power purchase agreements based on individual contracts," says Jonathan Hoare, who covers the French power market for analyst firm Aurora Energy Research. "So far, this has been a temperamental transition for EDF because of some managerial changes in the company and uncertainty about what the future will look like.

"Whereas previously they had 100TWh contracted, we're seeing nowhere near that kind of take-up under the new mechanism, and this is all happening in a climate where an increased amount of renewable sources are entering the

*"Mistral will take all the power we have,
but will only use half the space,"*

>> Charles Huot, Eclairion



Eclairion's first pod, acquired from Renault

system. So EDF is looking for replacement industrial off-takers for nuclear while this transition occurs."

EDF, therefore, should be a match made in heaven for data center operators and their never-ending quest for power, and the utility company is doing its best to set up the conditions for a long and happy marriage of convenience. At the AI summit, the company revealed it was offering four plots of land it owns for possible data center developments, and aims to have another two sites out to tender by 2026. It says the initial quartet of sites have a total of 3GW of power available, and that building there could cut time to market.

Colocating data centers near existing power stations means "you've got base load consumption, and you have less losses in transmission," Hoare says. "Nuclear plants are often located near large water sources because of the way the reactor technology works, so that can also potentially work well for cooling data centers," he adds.

Europe's AI capital?

In many ways, France seems like the ideal destination in Europe for data center operators, and CBRE's Breed believes the availability of power in the country could be a big draw, though he does not think the low-carbon nature of the French grid will be a factor in the decision-making of many developers.

"The biggest issue is the availability of power and the ability to scale," he says. "Clean energy becomes a secondary factor for a lot of companies. If you're a hyperscaler with ESG goals, it might be an important consideration, but I think for the other companies, such as the GPU clouds, it's less important."

Of bigger concern is the country's notorious bureaucracy, which continues to be a sticking point for potential developments, he says.

"There's a greater degree of delay involved in projects in France," Breed says. "You don't get the fast-track approvals you do in other countries. There's central government, then the

municipality, and then you might get the local mayor involved too, and suddenly you're consulting with an awful lot of people, particularly if there are ecological concerns. The whole system is less connected than somewhere like the UK."

The French government is making moves to try and speed up the planning process. Prior to last year's snap election, which saw Macron's parliamentary majority disappear and a hung parliament elected in its place, the president had been trying to push through legislation designating data centers as projects of major national interest, meaning the government could overrule local authorities to get digital infrastructure built. Now the provisions are back on the table as part of a wider economic "simplification" bill, which aims to slash red tape across the economy.

"Data centers aren't currently in the scope of projects of national interest, so this could help accelerate urban planning rulings and electricity grid connections," DC Byte's Ford. "The government thinks it can reduce the time for planning approval from 18 to nine months, and though it would likely only apply to the biggest projects, this would be seen as something positive for the data center industry, especially as it encourages developments at a time when other European markets like the Netherlands and Ireland are restricting them."

Debate on the legislation in the Assemblée Nationale began in April, but with 2,500 amendments already submitted, it is unlikely to be approved in short order.

For their part, Eclairion's Huot and Data4's Totel say their respective companies have garnered support from councils by building ongoing relationships with local politicians. Digital Realty, which announced at the AI summit it was planning to spend \$5.5 billion on new facilities in Paris and Marseille, has also worked closely with local government on the redevelopment of the area around the Paris Digital Park, which was an abandoned Airbus factory before the data center firm moved in.

Of bigger long-term concern for DRT's Van Der Wal are constraints on Paris. While the power availability situation in France is generally positive, the capital city's grid is under strain, he says. "In the



Cooling units ready for action at Eclairion

short term, things are looking good, but in the longer term, all the data centers in the extended Parisian area are going to be short of power," he says. "We're going to have to find a solution to that issue."

This highlights another potential issue for the French market, the dominance of Paris at the expense of other regions. While this reflects the nature of the French economy, which is highly concentrated on the Ile-de-France region around Paris, data center operators will need to push into other parts of the country as competition for space and power in the capital heats up.

While Data4 is looking to Cambrai, Eclairion has set its sights on a former Arjowiggins paper mill in Bessé-sur-Braye, in the Sarthe department of northwestern France. Huot says the site already has 100MW available, and Eclairion plans to put its containers inside the existing building. Work on the data center, which is expected to cost €600 million (\$675m) and has been approved by local officials, will get underway in 2026, with a view to a 2028 opening.

Ford says that, outside of Paris and Marseille, significant clusters of data centers have yet to emerge in France, but believes cities like Lille are starting to attract attention. She says the hyperscalers could change this equation if they decide to build their own facilities. Microsoft announced last year it was planning to construct a data center in Mulhouse, in the Grand Est region of eastern France. It would be the first hyperscaler-built facility in France, but the status of the project is unknown amid reports that the cloud giant has been pulling back from some of its plans for Europe. Microsoft has not commented on its plans for France.

"The hyperscalers don't have any self-builds in France yet," Ford says. "There's scope there for them to increase their presence in France, but it will be interesting to see if they maintain their colocation strategy or start building for themselves.

"That's definitely something to keep an eye on as we wait to see how many of the announcements from the AI summit actually come to fruition." ■

STORM CLOUDS

Despite the stereotypes that surround the French and their love of protests and strikes, opposition to data center developments in the country has so far been limited, apart from in Marseille.

France's second city hosted an anti-data center festival last year, organized by activist group *Le Nuage était sous nos pieds*, aka *The Cloud was beneath our feet*. The festival included a series of talks and events highlighting the effect of data centers in Marseille.

DCD spoke to 'Max' from *Le Nuage était sous nos pieds*, who says the group aims to raise awareness of the digital infrastructure in their neighborhoods. "For a long time, it felt like the data centers were hidden in our city. They don't employ a lot of staff, so they're not something people in Marseille have been aware of," she says. "We come from a techno-critical background and wanted to make people aware of the impact the industry is having on Marseille."

The activists have concerns about the impact of Marseille's data centers on the city's water supply, as well as the amount of electricity they consume. The group claims power that could have been used at the city's commercial port, or for a network of electric buses, has instead been directed to data center operators by the local authority.

Since the festival, and the announcements at the AI summit, Max says the group has been in touch with other activists around France and in neighboring countries such as Spain. "So far this has been a niche fight," she says. "But now we are becoming a point of contact for people who are concerned about these massive developments happening in their towns and cities.

"Our position is that there are already enough data centers in the world, and we need to stop and think about what the data center economy is going to look like and how it can benefit communities, rather than billionaires."

Find out more about the group's work at lenuageetaitsousnospieds.org. ■

Inside a quantum data center



Dan Swinhoe
Managing Editor

DCD visits quantum data centers from IBM and IQM in Germany

As you step into a quantum data center, the first thing that surprises you is the noise. The constant hum of fans you might be used to in a traditional server room is still there, but much quieter. Above the constant hum is the regular pumping of compressors forcing super-cooled liquids into the system to ensure atoms on chips maintain their quantum state. It's a sound that feels closer to pistons of a steam engine than electronics at the bleeding edge of computational physics.



Quantum computers might still be in their nascent stage, but even today's early-stage systems are being deployed in increasingly large numbers in a growing variety of locations. But what does a quantum data center look like, and how does it compare with what we know of data centers today?

DCD visited quantum data centers operated by IBM and IQM in Germany to see these systems out in the wild.

What does a quantum data center look like?

Much like the growing number of AI-polluted pictures of futuristic data centers you see online, the term quantum data center suggests something sleek, shiny, and perhaps slightly alien to what we're used to seeing. The reality, however, is less flashy.

Though first announced in mid-2023 and opened in October 2024, IBM's 'quantum' data center in Ehningen, just outside Stuttgart, actually dates back to the 1970s; fittingly, a time when science fiction was decidedly more tactile in aesthetic. IBM has been present in Ehningen for close to 100 years, and the company is amid an update of its presence at the site, exiting some older office buildings and developing new ones.

The building that hosts Big Blue's quantum systems previously housed client systems for what we might now consider private cloud deployments, long before the term was coined. Today, the basement continues to host traditional IT systems for IBM's own R&D teams, with the data hall featuring the company's quantum systems on the ground floor. The Ehningen site has actually been hosting a quantum system since 2021.

"It's used as a data center for our research and development teams, not primarily quantum, but also the other business units," says David Faller, IBM vice president of development, and managing director for IBM Germany Research

"For a modern data center, I don't think you have to do much to host a quantum computer"

>>David Faller, IBM

and Development. "It was a very, very good match."

IBM showed *DCD* around the quantum experience center in the facility, which included an IBM System One, hosted in an impressively sleek glass cube. At the time of our visit, the data center was home to two System Ones, each powered by IBM's 127-qubit Eagle quantum processing unit (QPU).

"We did the agreement with [research organization] Fraunhofer to set up the Quantum System One in 2021," says Faller. "The team in Yorktown [IBM's research center in New York, US] were the only ones that really worked on building quantum computers. They packaged all the boxes and sent them over here with the intent that when the boxes arrived a couple of weeks later, they would hop on a plane and get there and assemble that system. But then the Covid-19 lockdown came."

Luckily, the IBM research and development team right next door had "decades-long experience in mainframe technology, down to designing and validating the processors in our IBM Z

Systems, the cooling technology, the system packaging, all these elements," Faller says. He continues: "We recruited people from that team who know about cooling technologies and control electronics and trained them up with video conferences all day long with the US team."

The company didn't share precise specifications of the data center, nor how many quantum computers the facility could host. *DCD* is told, however, the classical IT in the basement data hall takes up "much more" power than the quantum floor; space and power is unlikely to be an issue at the site.

"There's definitely the intent to grow," Faller says. "If we have the demand over the next years, we can decide about taking other space and turn it into more quantum data center space."

"At the moment, we're using the upper floor for office space, and the quantum team has offices here as well to be close to the systems."

Despite the need to host liquid helium and nitrogen - with the helium pipes going under the raised floor - as well as accommodate taller systems than your typical racks, Faller says the site is "regular data center space."

"For a reasonably modern data center, I don't think you have to do much," he says.

The German data center of IQM, a Finnish quantum computing company spun out of Aalto University and VTT Technical Research Centre of Finland in 2018, is also not the futuristic shell ChatGPT might suggest a quantum facility might be, but a run-of-the-mill office building in Munich.

Launched in June 2024, the company is leasing space in at Georg-Brauchle-Ring 23-25, a former Telefonica office building. Acquired and redeveloped by Bayern Projekt and Europa Capital in 2017, the 39,000 sqm (420,000 sq ft) complex was rebranded as the Olympia Business Center and sold to family office Anthos in 2022.



IBM ribbon-cutting of the IBM Quantum Data Center in Europe with Chancellor Olaf Scholz



With Telefónica consolidating to the 37-story O2 Tower further down the street, the office complex is home to a number of companies; IQM's staff occupy space in the upstairs of the building, while the quantum data center resides in the basement, in white space previously used by O2. Housing two IQM quantum machines at launch, the data center is able to host up to 12 quantum computers, totaling 800kW of power capacity.

Jan Goetz, IQM co-CEO and co-founder, is German and did his doctorate on superconducting quantum circuits at TU Munich before moving to Aalto University in Finland. The company had long had a team in Munich, but decided to launch a quantum cloud-focused facility in the city.

"The goal here is really to separate the R&D and production," Goetz tells *DCD* at our visit to his facility. "So here in Munich, these systems should really be production systems, meaning high online availability; whereas the systems that we're running in Finland are mainly R&D systems, with our own engineers having access and trying out things."

He said the site combined a data center and office, which suited the company. There was little needed in the way of power and cooling, due to the relatively low power needs of quantum systems. Most quantum machines measure in the tens of kilowatts.

"There is no real plan yet to have quantum data centers in each region, but if the demand is there, why not?"

>>David Faller, IBM

"It wasn't a challenge here, and this is not a challenge in general for deploying quantum computers," says Goetz. "The only real change on the building side was we had to open up a little bit the entrance, because they're a bit bigger than the typical server rack that you have."

While Goetz seems keen to say the data center is pretty standard overall, a visit to the data floor reveals some usual characteristics. The center of the square hall features what Goetz describes as a 'separated service area' – essentially a room within the room – to host the compressors, pumps, and storage for the liquid nitrogen and helium. Each quantum system and the single accompanying rack of microwave signalling and processing equipment is placed out in the main white space area on either side of the service area in a sort of horseshoe layout.

With only half a dozen 20-qubit

quantum systems installed at the time of *DCD*'s visit, meaning only six regular air-cooled racks on the main floor, the data hall is relatively quiet and pleasant compared to some spaces. The regular pumping of compressors is the most notable – and unusual – part of the experience.

Unlike a traditional data center, Goetz says the company's data hall doesn't have a required optimal temperature to run at, but a constant temperature is more important: "We want to have a stable temperature," he says. "We don't want variations because it might affect the length of the cables if the temperature drifts too much."

Quantum form factors changing

Both IBM and IQM's systems use superconducting-based technology. These systems feature the iconic chandelier design, with the quantum chip sat within dilution refrigerators, large cryogenic cooling units that use helium-3 in closed-loop systems to supercool the entire system.

While most press pictures of IBM's quantum computers show a system hosted within a great glass cube, behind the curtain, the company operates a greater number of less presentation-worthy systems.

First revealed in 2019, the nine-foot



sealed cube, made of half-inch-thick borosilicate glass, is an impressive self-contained unit - with much of the cooling infrastructure hidden in the top and control systems behind.

The company has deployed such cubes in a number of locations for customers; Cleveland Clinic deployed such a system at its HQ in Cleveland, Ohio, instead of the company's data center on the edge of the city, while the Rensselaer Polytechnic Institute in New York also gave its own quantum system a pride-of-place display in a former chapel now housing its computing center.

Behind the scenes, however, most IBM-hosted quantum computers are less flashy and more functional for traditional white space. A simple metal frame holds a supercooled cryostat, with a traditional 19-inch rack next to it holding all the accompanying control and signalling equipment. But while many quantum companies are seeking to fit their quantum processing units (QPUs) into standard racks, IBM is pushing a new form factor that could offer more powerful systems at the cost of form factor.

IBM's System One quantum computers currently feature one QPU within one cryostat - a supercooled fridge hosting the golden chandelier commonly seen when talking about quantum computing. IBM has since launched its System Two, a much larger system that comprises three

QPUs in three cryostats hosted within one large hexagonal system.

As well as hosting three cryostats (each hosting one QPU) in one large shell, IBM has designed the system to be modular in terms of what can connect to it. While one side of the hexagon could host the control and signalling equipment, another side could feature racks of CPU or QPUs to create a hybrid system, or even more QPUs in a separate System Two hexagon.

"The System Two is the stage to be so modular that we really can build quantum computers that, in future, will bring us the quantum advantage," says IBM's Faller. "Our plan is, in the next few years, that you can connect up to seven into one virtual quantum computer so you end up with more than 1,000 qubits that you can use for your algorithms. That trio of traditional IT, GPU, and quantum processor units is something that we are anticipating for our IT solutions going forward."

A System Two has been deployed at IBM's quantum data center in New York; the Ehningen facility is set to have a System Two in future. A Bluefors hexagonal Kide cryogenic platform, developed in partnership with IBM, measures just under three meters in height and 2.5 meters in diameter, and the floor beneath it needs to be able to take about 7,000 kg of weight. IBM has also developed its own giant dilution refrigerator, known as Project Goldeneye, that can hold up to six individual dilution refrigerator units and weighs 6.7 metric tons.

IQM, however, is taking a different route, hoping to shrink its machines down to standard rack-based systems.

"We don't yet have the full quantum computer in a 19-inch rack," says IQM's Goetz. "But we are working in this direction. So from this perspective, I don't think that the kind of typical look and feel of the data centers changes so much."

Goetz notes that the liquid helium for IQM's systems is a closed-loop, and the company is working towards a quantum computer without the liquid nitrogen, meaning the system will eventually be fully closed and not require liquid refills.

"Right now, most systems still have a liquid nitrogen trap, which you need to refill," he explains, "and you need maybe one hour of training on how to deal with

the cold liquids."

What will a quantum data center look like in the future?

Despite claims from some operators, it seems there might not actually be any purpose-built quantum data centers yet. As far as DCD is aware, every space dedicated to hosting quantum computers is within an existing data center or lab, or a converted industrial building tailored to feature some white space.

That's not to say there won't be dedicated, purpose-built quantum data centers in future. And work to ensure they are standardized is already underway.

At its EMEA summit in Dublin in April 2025, officials from the Open Compute Project said they were looking to work with quantum computing companies and the operators that might such systems to develop standards around quantum computers.

Cliff Grossner, OCP chief innovation officer, said the standards group hopes to develop an OCP quantum-ready certification in the near future.

The group's existing self-assessment OCP-Ready program helps operators show their data centers meet the best practices and requirements defined by hyperscalers such as Meta and Microsoft, and are ready to host OCP-compliant equipment.

Grossner said OCP hopes to have a first draft of what a quantum-focused certification might look like out sometime in 2026; "It's better to think about it now; I believe we have enough to make a good start," he said in Dublin, noting some preliminary work with Orca and IQM had gone on during the conference.

Quantum-focused measures that might need to be considered include vibrations, electromagnetic sensitivity, and potentially even the speed of the elevators moving hardware between floors. Whether or not there would be one standard encompassing the different types of quantum computers - supercooled, rack-based, optical-tabled etc - or multiple standards to suit all comers is unclear at this stage.

The geography of quantum

While its becoming increasingly common to see quantum computing systems in supercomputing centers, its still rare

to see QPUs outside spaces hosted by quantum computing companies.

UK quantum computing firm Oxford Quantum Computing (OQC) has deployed six of its QPU systems in two colocation data centers: Centersquare's LHR3 facility in Reading, UK, and Equinix's TY11 facility in Tokyo, Japan. OVH has a system from Quandela at one of its facilities in Croix, France. Fridge supplier Oxford Instruments has installed a Rigetti quantum computer at its main factory site in Tubney Wood, Oxfordshire.

IBM does also host some dedicated quantum systems at its facilities for customers who don't want their QPUs on-site, but on-premise enterprise deployments are rare beyond the likes of IBM's deployment with Cleveland Clinic. They will likely be the exception rather than the norm for enterprises for some time to come, IQM's Goetz says.

"Corporate enterprise customers are not yet buying full systems," says Goetz. "They are usually accessing the systems through the cloud because they are still ramping up their internal capabilities with the goal to be ready once the quantum computers really have the full commercial value."

Quite what the geography of a world with commercially-useful quantum computers will look like is unclear. Will enterprises be happy with a few centralized 'quantum cloud' regions, demand in-country capacity in multiple jurisdictions, or go so far as demanding systems be placed in on-premise or colocated facilities?

"We think enterprise customers will start buying systems - larger customers that have their own computing centers or in colocation data centers," Goetz suggests. "And there will also always be some business that goes through the cloud."

IQM expanded into Germany, Goetz says, due to the strong quantum R&D ecosystem fuelled by the city's university talent, combined with lots of local industry. The CEO thinks Munich should also serve as a template for data centers the company hopes could be deployed globally should sovereignty become more of an issue.

"Sometimes it helps if you have the computer physically in a certain jurisdiction," he tells *DCD*. "If you are dealing with sensitive data or with



"Enterprise customers will start buying systems for their own computing centers or in colocation, but there will also always be some business that goes through the cloud"

>>Jans Goetz, IQM

government organizations, they might require that the computer is actually physically located in their country. So [Munich] can be a blueprint on how to build a data center in other places in the future."

What the company's future build-out will eventually look like will depend on customer demand. Goetz says he is seeing a lot of demand, and the company is preparing to have a "good number" of cloud-based systems available. He notes that the current geopolitical situation around the world is seeing budgets

growing on the defense side, which could be a "clear sign that there will be a need for sovereign quantum cloud," likely fuelling the need for in-country quantum compute.

IBM currently operates nine multi-zone cloud regions, two single-campus regions, and seven single data centers within its cloud footprint. IBM's Faller says the company is "committed" to growing its data center footprint in the US and Europe, but quite what the final picture will look like around its quantum footprint will depend on the customers.

"There is no real plan yet to have the data centers in each region," he says, "but if the demand is there, why not?"

Faller continues: "We are convinced that this will be really a fundamental element of supercomputing in the future. That's why we're talking about that quantum-centric supercomputing, bringing those worlds together."

"The hybrid model of cloud and the option on dedicated systems will make sense for quite some time. It's the same way that we have options to put mainframes or Power servers in your own data center or via the IBM Cloud. Both models exist, and I think they will be around for a number of years to come. That will not change easily."

If and when we reach quantum advantage isn't clear. Some companies suggest we're already coming against some extreme use cases where classical hardware struggles to compute a reliably accurate answer in a useful amount of time. We are at the point where some particular workloads actually might be more efficient on quantum hardware in theory; however, even if the quantum algorithm would be more effective, the quantum hardware isn't currently powerful enough for a full-scale real-world deployment.

While both companies believe we'll get there, when we reach quantum advantage isn't so clear. IQM's Goetz argues it will be a slow build-out over time, use case by use case, rather than one big bang. We might be a decade or more away from all-powerful, multi-million qubit systems delivering the ultimate quantum advantage, but Goetz notes that "doesn't mean that there's nothing on the way there," in the same way the utility of GPUs has grown over time as more companies use them for AI. ■



Sponsored by



The Cooling Supplement



Chilling innovations

INSIDE

Thinking small

> Cooling with nanofluids

As cold as ice

> Antifreeze in the data center

Joule in the crown

> Extracting water from air

Vertiv™ CoolPhase Flex

The best of both worlds.

Air or liquid cooling?

Vertiv™ CoolPhase Flex delivers both in one hybrid solution. Start with air and switch to liquid anytime. Compact, efficient, and scalable on demand.

Embrace a hybrid future.



[Vertiv.com](https://www.vertiv.com)



Sponsored by



Contents

30. Making a [nano] material difference to data center cooling

How nanofluids can improve cooling efficiency, even in legacy cooling systems

36. Could plasma cooling have a place in the data center?

YPlasma and others look to swap server fans for plasma actuators

38. Joule in the crown

US firm AirJoule says its water harvesting technology could cut cooling costs at data centers

40. You're as cold as ice: Antifreeze in liquid cooling systems

The importance of stopping things from freezing to keep things moving



30



36



40

Chilling innovations

Keeping servers at a temperature where they can run efficiently in the age of artificial intelligence is a challenge for many data center operators.

While the AI revolution has heralded the dawn of the liquid cooling era, older technologies remain vital to the operation of many data centers, and vendors are busy finding ways to innovate and commercialize ideas that have been doing the rounds for years.

In this supplement, Georgia Butler finds out how one company, HTMS, is bringing nanofluids into the cooling mix.

Based on a research initiative that started more than 15 years ago at the Università del Salento in Lecce, Italy, the HTMS cooling solutions insert nanoparticles into cooling fluid.

These are ultrafine, tiny particles that measure between one and 100 nanometers in diameter. In the case of HTMS's "Maxwell" solution, the particles are made of aluminum oxide, and the company believes there are big efficiency gains to be had in pumping them around server rooms.

Meanwhile, experiments around using plasma to cool semiconductors have been carried around the world since the 1990s. Now a Spanish company, YPlasma, wants to take the idea from the theoretical to the practical, with a system that can be used to chill laptop components, and that could eventually be applied in servers too.

YPlasma's system uses plasma actuators, which it says allow for

precise control of ionic air flow. Dan Swinhoe spoke to the company about how this works, and its plans for the future, which involve finding a partner firm to help bring the system to the server hall.

Over in the US, another technology vendor, AirJoule, is seeking to help make data centers more efficient by harvesting water contained in waste heat. The company has found a way to harness metal organic frameworks, or MOFs, another technology that has thus far been confined to the lab for a variety of technical and economic reasons.

With the backing of power technology firm GE Vernova, AirJoule has built a demonstrator unit that it believes can harvest serious amounts of water. The task for the company is to convince data center operators to take the plunge and install it at their facilities.

AirJoule is targeting data centers in hot environments, where the threat of frozen pipes is probably not something that keeps data centers managers up at night.

For those running facilities in cooler surroundings, a big freeze can have big implications, which is why antifreeze remains a critical part of the data center ecosystem.

Most antifreezes on the market are based on glycol, but its high toxicity means some companies are looking to alternatives, with nitrogen emerging as a potential game-changer. With more pipework than ever filling the newest generation data centers, keeping the cooling fluid flowing has never been more important.



Georgia Butler
Senior Reporter,
Cloud & Hybrid

Making a [nano] material difference to data center cooling

*How nanofluids can improve
cooling efficiency, even in
legacy cooling systems*



At any data center conference, and you will find a plethora of new cooling solutions designed to meet the needs of the latest and greatest technologies.

At around 700W and beyond, air cooling for chips becomes increasingly difficult, and liquid-cooling becomes the only realistic option.

Nvidia's H100 GPUs scale up to a thermal design point (TDP) of 700W, and its Blackwell GPUs can currently operate at up to 1,200W. The company's recently announced Blackwell Ultra, also known as GB300, is set to operate at a 1,400W TDP, while AMD's MI355X operates at a TDP of 1,100W.

Recently, liquid cooling firm Accelsius conducted a test to show it could cool chips up to 4,500 watts. Accelsius claims it could have gone even higher but was limited by its test infrastructure, rather than the cooling system itself.

But the reality is that, while much of the excitement and drama in the sector focuses on the needs of AI hardware, the majority of workloads in data centers are not running on these powerful GPUs. The density of these racks is also creeping up, if not to the same level, and many companies are not willing to fully rip-and-replace cooling systems in costly and complex programs.

While many companies are looking to reinvent the wheel with flashy new cooling set-ups, a large cohort of data center operators are focused on refining their existing systems, looking to unlock additional efficiencies wherever possible and improve their cooling capabilities at a rate that is appropriate for their needs.

One way of doing this is to optimize the fluid used in existing cooling systems.

Nanoscale changes

A large portion of "air cooling" solutions still use liquid at some point in the system - usually water or a glycol-water mix. When it comes to the choice of fluid, water is typically considered more efficient at transferring heat, but water-glycol can be more suitable when freeze protection is needed.

But, regardless of whether it is a water or a glycol-mix base, Irish firm HT Materials Science (HTMS) says that it can



make your system more efficient with "nanoparticles."

HTMS was founded in 2018 by Tom Grizzetti, Arturo de Risi, and Rudy Holessek.

According to Grizzetti, the company's Maxwell solution stems from a research initiative that started more than 15 years ago at the Università del Salento in Lecce, Italy.

The team was exploring how nanotechnology could be used to fundamentally improve heat transfer in fluid systems. Grizzetti says of that time: "It's one of those rare startup scenarios where deep science meets a real-world market need at just the right moment."

To date, HTMS's customer base hasn't primarily been data centers, with its solution currently deployed at industrial sites and Amazon fulfillment centers, but the company is targeting the sector and talking to operators because, it says, it can see a strong use case.

Nanoparticles are ultrafine, tiny particles. In the case of HTMS's "Maxwell" solution, these are made of aluminum oxide and measure over 100 nanometers in diameter.

"Our product is a simple fluid additive that goes into any closed-loop hydraulic system, whether that be air-cooled chillers or the vapor side of water-cooling chillers," Ben Taylor, SVP of sales and business development at HTMS, explains.

Taylor makes the bold claim that Maxwell will improve efficiency at "every heat exchanger in the loop - whether it be the evaporator, barrel, chiller, the coil, or the air handler - they all see better heat transfer capabilities."

Aluminum is well known for its heat transfer properties. The metal alone has a

thermal conductivity of approximately 237 W/mK (Watts per meter per Kelvin), and the aluminum oxide compound (also known as alumina) has a relatively high thermal conductivity for a ceramic material, typically around 30 W/mK.

While having a lower thermal conductivity than pure aluminum, alumina is more suited to cooling systems as it has greater stability and durability, meaning it works well in high-temperature and high-pressure applications, and provides better corrosion resistance. All of which has been backed up by several scientific studies, including the paper *'Experimental study of cooling characteristics of water-based alumina nanofluid in a minichannel heat sink'*, published in the Case Studies in Thermal Engineering journal.

The fact is that although alumina is a well-established additive for thermal conductivity solutions, it isn't widely adopted by the data center industry. While at the Yotta conference in Vegas in 2024, DCD met with HTMS and asked about competition in the industry - only to be told that, currently, they aren't really facing any.

The nanoparticle solution can be injected into new and old cooling systems, Taylor says, "You get more efficiency in your system, and even if it's an older system, you can start to do more than you thought you could," he says.

For many data center operators, this is music to their ears. Ripping out and replacing cooling systems costs time and money, but with the growing cooling needs of hardware and increasing regulatory pressures in many markets, finding efficiencies is paramount.

There are, of course, upfront costs associated with the installation of the heat-transfer fluids, but HTMS estimates that companies see payback on this within three years, and some in as little as one year.

Efficiency and environmental impact

From the perspective of sustainability, Taylor says that Maxwell's carbon footprint is low. "A lot of customers will break even on CO2 emissions within the one to three-month mark," he says.

"We are seeing existing systems that

"You get more efficiency in your system, and even if it's an older system, you can start to do more than you thought you could"
->Ben Taylor



need more capacity, and new systems that are trying to be as efficient as possible because a lot of them are going to start taking up a lot of the power grid - and plenty are already using a lot of power.

"They are looking for a more overall energy efficient system, and one option could be a water-cooled chiller set up with open cell cooling towers. But, while they are more energy efficient, they are costly in terms of water consumption," he explains, adding that Maxwell can only be used in a closed-loop system.

He continues: "The additives are nanoparticles, so if they are open to the environment, especially in an open cell cooling tower, they could just blow away in the wind.

"The other reason is we have a pH requirement. We normally stay around 10 to 10.5, which is pretty standard in a glycol system, but it is a little higher than you would expect in a water system. And if they are open to the atmosphere, the pH will just keep dropping."

The improvements seen in cooling abilities vary depending on whether the customer uses water or a glycol mix.

When Maxwell is injected into the

system at a two percent concentration, Taylor says that a water-only system can see around a 15 percent increase in heat transfer abilities, and then with a mix of around 30 to 40 percent glycol, that 2 percent concentration can improve heat transfer capability by as much as 26 percent.

Deploying Maxwell is a simple process, Taylor says. "We have everything pre-blended and then inject it into a system," he says. "We need the pumps to be running, unless we are filling the system right out of the gate."

HTMS usually recommends also installing a "make-up and maintenance unit" (MMU) on the site - a device which monitors the system and maintains the appropriate chemistry and mixture of particles.

Jim McEnteggart, HTMS' senior vice president of applications, explains: "We usually ship at around 15 percent volume. It's a liquid, and we then pump it through the system through connections typically on the discharge side of their normal pumps.

"The concentrated products mix with their system, and then we train it

out on the suction side of the pump. There, the MMU has instrumentation that measures the pH and density. When it reaches its target, we know we have enough nanoparticles in the solution to achieve the desired outcome, and we stop injecting."

McEnteggart adds that HTMS assumes around a "five percent leaking rate per year" according to ASHRAE standards, but that once Maxwell is in a system, the company says it can last for ten years.

"In reality, everything in the loop won't degrade the aluminum oxide since it's chemically inert," McEnteggart says. "So, unless there's a leak, the product can stay in there for the life of the system."

That chemical inertia is also a positive factor in terms of the impact - or lack thereof - that the solution could have on the environment if there were a leak.

"It's not toxic," says McEnteggart, though he jokingly adds: "I still wouldn't recommend drinking it."

"If it got into the groundwater, aluminum oxide by itself is a stable element that's not really reactive with many things. But again, the protocol is that we don't want it discharged into sewer

"If we lower the pH into the acidic range - so, below seven - the particles drop out of suspension."

>>Jim McEnteggart



systems or things like that, because at the concentrations we are dealing with, it could overload those systems and cause them not to work as well as they should."

This is important to note, as additives can sometimes be PFAs - per or polyfluoroalkyl substances, also known as "forever chemicals". The nanofluids, while considered an additive, do not fall into this category.

Should a customer want to remove Maxwell from its system - which McEnteggart assures *DCD* none so far have - it's a combination of mechanical filtration and chemical separation.

"If we lower the pH into the acidic range - so, below seven - the particles drop out of suspension. They settle very quickly, and we can do that on site, just pump the system out into tanks and add an acidic compound to get the particles to collect at the bottom of the tank.

"We extract from there, and then everything else gets caught by a ceramic filter with very fine pores. Then we neutralize the solution, and if it's water, it can go down the drain, or if it's glycol, it goes to a treatment facility."

Alumina, and nanofluids by extension, have not sparked much conversation in the industry thus far.

As a result of this, it occurred to *DCD* that there may be some issues in providing the solution en masse should the industry show significant interest.

"We would definitely have to scale up very quickly," concedes Taylor. "But a good thing is that the way the product is made is quite a simple approach. Once we see interest growing, we can forecast that, and all we'd have to do is open up a building and get two or three specific pieces of

equipment.

"The only limiting factor would really be the raw material suppliers - for the alumina and, in the future, for the graphene. If we hit snags on that end of things, it would be out of our control."

HTMS has at least one data center deployment that they have shared in a case study, though the company remains anonymous beyond noting the facility is located in Italy. The likes of Stack, Aruba, TIM, Data4, Equinix, Digital Realty, OVH, and CyrusOne, as well as supercomputing labs and enterprises, operate facilities across Italy.

While unable to share identifying details

about the Italian deployment, the total power of the chiller system was 4,143 kW (1.178 RT) and consisted of three chillers and one trigeneration system. According to the case study, the data center saw a system coefficient of performance improvement of 9.76 percent.

But a big part of the challenge when entering a long-secretive industry is getting your name out there if clients won't let you share their story.

"A lot of them [data center operators] are pretty sensitive about having their names out there," admits Taylor, though the company remains hopeful that it will fully conquer the data center frontier. ●

FUTURE POSSIBILITIES

On the AI side of workloads, HTMS is currently working behind the scenes on a solution that could work with direct-to-chip cooling solutions.

Direct-to-chip (DTC) brings liquid coolants directly to the CPU and GPUs via a cold plate. The solution is becoming increasingly popular due to its ability to meet the cooling demands of the more powerful AI chips.

Details about the new product are sparse, though Taylor describes it as "more of a graphene-based solution."

He adds that the data center and chip companies HTMS is currently in discussions with have said that if they

can make a solution that improves the performance of the chip stack by between 4-5 percent, they'll be "very, very, interested," but that HTMS is targeting a 10 percent improvement.

Beyond that, the little that can be shared is that HTMS is currently talking with "some of the glycol and heat transfer fluid OEMs that are already dedicated to the DTC cooling space" to see if, down the road, they could offer some pre-blended mixtures.

At the time of talking with *DCD*, HTMS had "almost done complete iterations in R&D," but had no further updates at the time of publication. ●

Solving the puzzle of AI-driven thermal challenges

Thanks to increased power demands brought about by the rapid adoption of AI, data centers are heating up. Delivering flexible, future-ready cooling solutions is now essential

The data center landscape is undergoing a seismic shift. AI, with its immense processing requirements and dynamic workloads, is redefining thermal expectations across the industry. As rack densities climb and traditional infrastructure strains under new pressures, cooling has emerged as one of the most critical and complex challenges to solve.

But just like solving a puzzle, creating an effective data center suitable for growing densities in the AI age isn't about one single component – it's about fitting the right pieces together.

Providing operators with the tools they need to keep cool under the pressures of increased power demand, Vertiv is delivering future-ready thermal solutions designed to support AI workloads and beyond.

Jaclyn Schmidt, global service offering manager for liquid cooling and high density at Vertiv shares how the company is addressing cooling challenges of today and the future, with innovative, flexible products and services, delivered with an end-to-end, customer first approach.

The thermal puzzle that AI presents

The rise of AI has introduced a level of unpredictability previously unprecedented in data center operations. Huge workloads fluctuate rapidly, often creating thermal spikes that traditional systems struggle to manage. Schmidt compares this unpredictability to the AI market itself:

"AI workloads are just like the AI landscape – constantly shifting, evolving, and difficult to predict. Just as we can't always foresee where the next breakthrough will come, we also can't anticipate where or when thermal spikes will occur. That's why flexible and scalable cooling solutions are so essential."

This demand for agility has pushed the boundaries of what was once considered future-ready. Facilities designed with the future in mind just a few years ago are already being outpaced by the speed of technological change.

"We're seeing customers who built with scalability in mind suddenly find themselves needing to retrofit or redesign. The pace of innovation is relentless, and infrastructure must evolve just as quickly," adds Schmidt.

Beyond the box: A system-level approach

Rather than offering isolated cooling products, Vertiv takes a holistic, system-level view of thermal management – from chip to heat reuse, airflow to power infrastructure, and everything in between. It's an approach rooted in evaluating how every piece of the puzzle connects.

"This isn't just about placing one box inside another," explains Schmidt. "We examine how each component affects the entire system – controls, monitoring, airflow, and more. Everything must work in harmony with efficiency and effectiveness."

With this comprehensive lens, Vertiv has created a suite of solutions purpose-built for AI-era data centers. The Vertiv CoolPhase Flex and Vertiv CoolLoop Trim Cooler, for instance, are specifically engineered to handle the heat loads and fluctuating demands of liquid cooling deployments.

"We've taken our history with both direct expansion and chilled water systems, and evolved them to meet the needs of AI and hybrid cooling," says Schmidt. "Products like the Vertiv CoolLoop Trim Cooler

can support mass densification of the data center while also increasing system efficiency without increasing footprint, while the Vertiv CoolPhase Flex enables future-ready modularity – allowing you to switch between air and liquid cooling as your needs evolve over time.”

These options provide operators with unmatched flexibility. Whether there’s a need for direct expansion (DX), chilled water systems, or a hybrid setup that supports mixed workloads and temperatures, access to diverse solutions that address specific, personalized requirements is vital for supporting scalability across the industry.

But it’s not just about connecting indoor cooling units with outdoor heat rejection capability. Having integrated networks of products that work together and constantly communicate to provide a high level of system visibility and control is becoming vital for long-term success. Vertiv has innovated their controls and monitoring platforms to look at the data center as a whole, rather than as individual components, with products like Vertiv Unify.

Tailoring cooling solutions to unique customer demands is an essential part of effective deployments. Each element of Vertiv’s product portfolio is designed to meet customers where they are – whether they’re just beginning to explore liquid cooling or already operating dense, AI-heavy environments.

“One of the misconceptions is that if you’re deploying liquid cooling, you must use chilled water,” explains Schmidt. “But that’s not always true; it’s highly dependent on the challenges you’re trying to solve. We’ve developed solutions like the Vertiv CoolPhase Flex and Vertiv CoolPhase CDU to offer adaptable alternatives – modular, scalable tools that grow with the needs of the data center.”

By offering flexible solutions that can evolve without replacing entire systems, Vertiv helps operators plan with confidence, knowing today’s decisions won’t limit tomorrow’s potential.

Providing the right tools

Collaboration and communication between providers and operators is vital for developing innovative, resilient solutions that tackle the real day-to-day challenges. At the heart of Vertiv’s approach is a

commitment to co-creation. Rather than simply delivering products, the company works alongside customers to define the right journey, from early design consultation through to deployment and ongoing maintenance. Schmidt explains:

“We don’t believe in a one-size-fits-all approach. We ask: What are your specific challenges? Are you midway through a transition or starting from scratch? We help define a path forward based on those answers.”

That path includes expert services throughout the entire lifecycle. With approximately 4,000 service engineers and over 310 service centers worldwide, Vertiv aims to make sure that local support is never far away.

“If something goes wrong or customers need help, we’ve got boots on the ground, anywhere in the world,” she says. “That level of regional expertise allows for smoother integration, better response times, and peace of mind for our customers.”

Every data center is unique. Whether it’s a retrofit or a new build, successfully transitioning to liquid cooling, or optimizing hybrid models to manage thermal challenges, requires a strong focus on sustainability, energy efficiency, and minimizing disruption for both implementation and long-term operation.

Schmidt explains the unique challenges associated with retrofit projects:

“Retrofitting is like decoding someone else’s design puzzle. You have to figure out what can stay, what needs to change, and how to integrate new pieces without disrupting operations. We offer creative, consultative advice to make that possible.”

That includes services like recycling, repurposing, and refurbishing existing equipment – not just to reduce waste, but to extend the life of investments, lower costs, and improve operational continuity.

Research, innovation and the road ahead

Vertiv’s research and development (R&D) efforts are fueled by a dual engine: real-world customer feedback and forward-looking market analysis. This two-pronged approach enables continuous development across the liquid cooling spectrum, including direct-to-chip, immersion, and two-phase technologies.

“We’re not betting on one technology,” explains Schmidt. “We’re exploring everything from single-phase immersion to two-phase liquid cooling to meet niche and emerging needs. That diversity means we can be ready for whatever comes next.”

Crucially, this innovation is always tied to real customer pain points: “We’re constantly collecting feedback and asking: What else do our customers need? What aren’t they seeing yet? Then we build products and services to address those gaps,” she says.

This ongoing iteration makes Vertiv more than a vendor, positioning the company as a true partner in problem-solving, empowering operators to drive the vision for their own data center designs while providing the tools for success.

Although attempting to predict what the future may hold is perhaps the greatest challenge for the industry, some things are certain: workloads will continue to grow, densities will rise, and Edge deployments will proliferate. Meanwhile, regulations, chip designs, and sustainability requirements will keep evolving in response.

“We’re constantly tracking the shifts – whether it’s Edge computing, modular builds, or ultra-high-density racks,” says Schmidt. “Each one requires a different strategy, and we’re committed to staying ahead with innovation and responsive service offerings.”

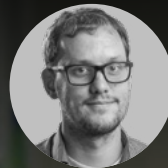
With a modular, scalable, and consultative approach, Vertiv is giving customers what they need to build, both now and in the future – meaning that as the data center puzzle grows more complex, they’ll always have the right pieces in hand.

In a world shaped by AI, flexibility and foresight are no longer optional – they’re essential. Supporting operators in designing, building, and maintaining cooling systems that aren’t just solutions for today’s thermal challenges, but can adapt to support future demand, is key to piecing it together. ●

Learn more about Vertiv’s liquid cooling solutions [here](#).



Could plasma cooling have a place in the data center?



Dan Swinhoe
Managing Editor

YPlasma and others look to swap server fans for plasma actuators

Though the data center industry is currently focused on direct-to-chip liquid cooling, most of the world is still air-cooled. Thousands upon thousands of servers remain equipped with fans blowing air away from CPUs.

And though bleeding-edge GPUs have reached the point where air cooling is no longer feasible, air cooling is still the norm for many lower-density workloads and will continue to be so for years to come.

Spanish startup YPlasma is looking to offer a replacement for fans inside servers that could give air-cooling an efficiency and reliability boost. The technology essentially uses electrostatic fields to convert electrical current into airflow.

Launched in 2024, the company was spun out of the National Institute for Aerospace Technology (Instituto Nacional de Técnica Aeroespacial, or INTA) – the Spanish space agency part of the Ministry of Defence and described by YPlasma CEO David Garcia as Spain's NASA.

The Madrid-based company's technology uses plasma actuators, which it says allow for precise control of ionic air

flow. Though interest in plasma actuators dates back to the days of the Cold War, they first left the lab in 1990s, being used to improve aerodynamics.

The company's actuator is called a Dielectric Barrier Discharge (DBD) and composed of two thin copper electrode strips, with one surrounded by a dielectric material such as Teflon. The actuators can be between just 2-4 mm thick.

When you apply high voltage, the air around the actuator ionizes, creating plasma just above the dielectric surface. This can produce a controllable laminar airflow of charged particles known as ionic wind that can be used to blow cold air across electronics.

The direction and speed can be controlled by changing the voltage going through the electrodes. Garcia says the actuator's ionic wind can reach speeds up to 40km per hour, or around 10 meters per second.

INTA's technology was originally developed for aerodynamic purposes, with YPlasma initially targeting wind turbines. The actuators can be used to push air more uniformly over the turbine

blades, producing more energy.

In the IT space, the company is looking at using its actuators to force air across chips, removing the need for fans – potentially saving space and power, offering greater reliability than fans, and removing vibrations, which could increase hardware lifespans. The company says its DBD actuators can match or exceed the heat dissipation capabilities of small fans.

This ionic wind, the company says, also removes the boundary layer air on electronics, allowing for better heat dissipation, resulting in natural heat convection combined with the forced convection of airflow over the chips. The cooling, it claims, can also be more uniform, eliminating hotspots.

YPlasma is currently working with Intel and Lenovo to look at using its plasma technology in laptops and workstations, but thinks there could also be utility within the server space.

Intel has been exploring ionic wind in chip cooling since at least 2007, previously partnering with Purdue University in Indiana; the chip giant has previously filed patents for a plasma cooling heat sink that would use plasma-driven gas flow to cool down electronic devices.

"We were curious if we were able to provide cooling capacity to the semiconductor industry," says Garcia. "So we engaged with Intel; Intel has been doing research on plasma cooling, or ionic wind, for some years, so we were speaking the same language, and they asked if we could develop an actuator for laptop applications to cool down a CPU."

"We started working with Intel in August (2024), and we developed a final prototype in January," he adds. "In May or June, we will be ready to implement them in real laptops."

Now YPlasma is looking for a project partner to begin developing and proving out server-based products. Garcia has "been in conversations" with potential collaborators but is yet to secure a project partner.

"We've been exploring this space, and realizing that every single week we are able to pull down much higher demand in applications," he continues. "We've seen that we can get higher levels of heat dissipation. And our objective is to evaluate if this technology is able to help on the data center industry. Perhaps not in the most demand applications where liquid cooling is necessary, but there's some intermediate

applications where we can replace fans."

In a talk at the OCP EMEA Summit in Dublin in April, YPlasma presented its aims in the data center space, saying it has been collaborating with "one of the main companies in the semiconductor industry" and that so far "all expectations have been exceeded," with the system showing better performance than a regular cooling fan.

In early tests, the company says its actuators have been able to match or exceed the cooling performance of an 80 mm fan under a 10W heat load. The actuators can also work below 0.05W, requiring less power than a fan to operate. A test video on the company's website shows a chip being cooled from 84°C (183.2°F) down to 49°C (120.2°F).

As well as general cooling benefits, YPlasma claims its technology comes with added bonuses for data center firms. The company says plasma coatings can provide a barrier against corrosion and oxidation, protecting components from humidity and other environmental contaminants in high humidity or industrial environments where equipment is more exposed to harsh conditions. Plasma, it adds, can effectively remove contaminants such as dust, oils, and organic residues without physical contact or harsh chemical products; dusty and dirty Edge deployments such as mining or factories could benefit from fewer fans and cleaner air running through the systems.

Beyond the servers themselves, YPlasma thinks its technology could have application within the HVAC system. By ionizing the gas inside the heat exchangers, you could enhance the heat transfer coefficient.

The system can also produce heat instead of cold air, which could be used as a type of antifreeze; Garcia says it can generate heat up to 300°C (572°F).

And another potential use, the company suggests, is to create plasma-activated water, which can have better heat transfer coefficient than regular water in liquid cooling systems. Research is still going on in this field.

In the future, the company said it will be possible to combine actuators with AI algorithms, making it possible to optimize airflow control in real-time, automatically adapting to changing conditions.

Investors in YPlasma include SOSV/ HAX

"We were curious if we were able to provide cooling capacity to the semiconductor industry"



David Garcia

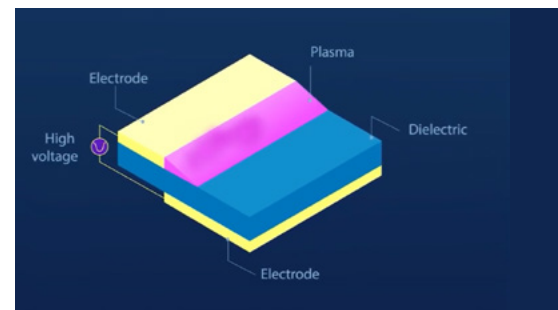
and Esade Ban, which both contributed to seed rounds totaling €1.1m (\$1.2m) in 2024. Other investors include MWC's Collider, and the European Space Agency's Commercialisation Gateway. Garcia was previously CEO of Quimera Energy Efficiency, a company that uses software to reduce energy demand in the hospitality sector; and immersive virtual reality firm Kiin. He says the company has recently closed another €2.5 million (\$2.8m) raise.

Though Intel has been exploring the space for years, it isn't the only company to have looked at plasma-based cooling for electronics over the years. Apple filed a patent for an ionic wind cooling system back in 2012, but it's unclear if the company ever took the idea further. Tessera, a chip packing company later renamed and broken out into Xperi Inc. and Adeia, demonstrated ionic wind cooling for a laptop in 2008 alongside the University of Washington.

YPlasma is not the only company that is looking at plasma for cooling today, and there could be a race to commercialize the technology. Ionic Wind Technologies, a spin-out from the Swiss Federal Laboratories for Materials Science and Technology (Empa), is also looking at developing its own ionic wind amplifier that could be used for chip cooling.

A bigger product than YPlasma's and developed as a way to dry fruit, the company claims its custom-made needle tip electrodes achieve up to twice the airflow speeds compared to conventional electrodes and use less energy. The spin-off funding program Venture Kick and the Gebert Rüt Stiftung are providing financial support to the Empa spin-off as part of the InnoBooster program to bring the product to market maturity.

"We produce the airflow amplifiers ourselves and want to sell components in the future. However, as we have patents and other ideas, a licensing model could also be conceivable," Ionic Wind founder Donato Rubinetti said in the company's early 2025 launch announcement. "I see the potential wherever air needs to be moved with a small pressure difference. In future, however, above all in the cooling of computers, servers, or data centers."

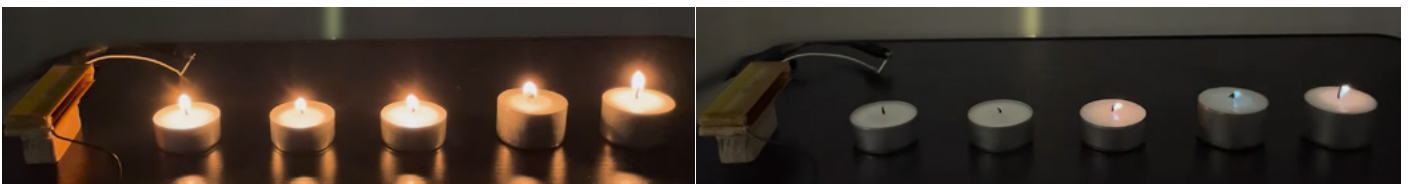


In the meantime, YPlasma's Garcia says once the company finds a server-centric partner to collaborate with, it could have a working prototype on a specific application within six months.

"We just need to understand what the requirements are that we want to achieve," he says. "It's easy to build; we've got a lot of experience testing this."

While it could be beneficial in some scenarios, the company does concede that its technology won't be applicable to all chips within a data center. It is not, for example, aiming to compete with the densities liquid cooling is designed to cool, but can help remove fans in some air-cooled applications.

"We're never going to have the capacity to remove as much heat as liquid or immersion cooling," says Garcia, "but I think there's a lot of heat we can remove without those technologies." ●



Joule in the crown



Matthew Gooding
Senior Editor

US firm AirJoule says its water harvesting technology could cut cooling costs at data centers

China has a lot of electric buses. In 2023, the nation's public transport network was home to more than 470,000 electric vehicles ferrying passengers to and from their destinations.

These buses provided an unlikely inspiration for AirJoule, a firm seeking to cut data center emissions with a system that can capture water from waste heat generated at air-cooled data centers so that it can be reused.

"One of my previous companies worked with [Chinese battery manufacturer] CATL to marry our motor technology with their batteries," says AirJoule CEO Matt Jore. "We developed an advanced power train which is operating in buses all over China.

"As part of that process we discovered that in congested cities, where 90 percent of the journey is spent in traffic stopping and starting, air conditioning is blasting away, so it's not unheard of to have seven or eight times more energy drain from the battery for air conditioning than in normal driving conditions.

"So we started looking at ways to

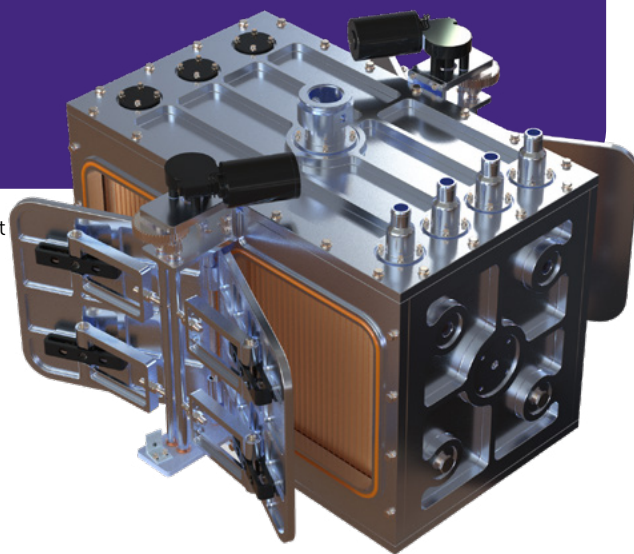
improve the efficiency of that air conditioning, and that's what got me involved in what we're doing now with AirJoule."

AirJoule's system will not be taking a ride on the bus anytime soon, but the company intends for it to become a common sight in heat-generating industrial settings such as data centers, where it says it could help cut energy usage.

Organic growth

AirJoule's machine is based on metal-organic frameworks, or MOFs, a type of porous polymer with strong binding properties and a large surface area. This, in theory, makes MOFs ideal for binding onto specific particles, such as water.

MOFs have been around for years, Jore says, but applying them in commercial settings has been difficult for two reasons; activating the useful properties of the polymer generates heat, and the material has been expensive to produce and purchase.



AirJoule solved the price issue by working with chemical company BASF to develop a cheaper MOF. By scaling the production process, Jore claims the company has managed to reduce the cost of the material from around \$5,000 per kg to "something nearer \$50." When it comes to the first problem, AirJoule's system reuses the heat generated by the MOF to turn water vapour back into liquid, creating what the company says is a "thermally neutral" process.

Explaining how the system works, Jore says warm air enters the AirJoule machine and is processed in two stages. "We divide the water uptake and the water release into two chambers that are performing their functions simultaneously," he says. "The heat generated in the absorption process is passed over to the second chamber, and it's pulled under a vacuum. We suck the air out, and now all that remains is this coated material full of water vapor molecules."

"We can reduce the amount of power needed in data centers by accessing the world's largest aquifer - the atmosphere"

>> Matt Jore



The system then initiates its vacuum swing compressor that “creates a pull on the water vapor molecules,” Jore says. “Trillions of these molecules pass through this vacuum swing compressor, and we slightly pressurize them so that they increase their temperature and condense.”

How much water comes out at the end of the process depends on how many AirJoule systems you buy. The company says one module will be able to harvest 1,000 liters of distilled water a day, either to be reused by the data center’s cooling system or supplied to the local community. “You can scale it in modular pieces,” Jore says.

Despite the claims of thermal neutrality, AirJoule does still produce a small amount of waste heat. In March, the company revealed that tests had shown that AirJoule demonstrated the ability to produce pure distilled water from air with an energy requirement of less than 160 watt-hours per liter (Wh/L) of water.

This apparently makes it more efficient than other methods of extracting water from air. “Compared to existing technologies, AirJoule is up to four times more efficient at separating water from air than refrigerant-based systems (400-700 Wh/L) and up to eight times more efficient than desiccant-based systems (more than 1,300 Wh/L),” the company said in March.

Building a partnership

The AirJoule machine is based on principles developed by Dr. Pete McGrail at the Pacific Northwest National Laboratory, a US scientific research institution in Richmond, Washington.

“The challenges the world faces around water supply are not going to be solved by any one company”

>> Matt Jore

After meeting Dr. McGrail during the Covid-19 lockdowns and finding out more about his work, Jore and his colleagues decided to try and turn it into a product. They teamed up with power technology firm GE Vernova, which was already working on its own air-to-water project with the US Department of Defense.

Recalling the early days of the collaboration, Jore says: “The two teams met in Montana in my garage. We had four PhDs from GE Vernova come out and spend the week with us and try to put our technologies together, and by the end of the week, we were clinking beers together because we knew we had something promising.”

To take the technology to market, a joint venture was formed in March 2024 between Jore’s firm, Montana Technologies, and GE Vernova. The partners each own 50 percent of the JV, though in a slightly confusing move, Montana Technologies has since changed its name to AirJoule Technologies Corporation, a decision it took to reflect its role in the JV.

The JV raised \$50 million when it went public via a special purchase acquisition company, or SPAC, last year, and in April announced it had raised an additional \$15 million from investors, including GE Vernova. So far, it does not appear to be

offering great value for its stockholders, with its share price having dropped significantly this year.

Jore and his team will be hoping to change that when it gets its machines into data centers and other facilities that produce a lot of heat.

“With every business I’ve worked on, you build a prototype, with cost and reliability defined, and get people to come and see it,” he says. “We’re going to have a working unit this year that we can show to our customers and partners, and from there we’ll be able to demonstrate the replicability of that unit.”

He believes the AirJoule offer will be compelling for the market, despite the plethora of water harvesting technologies that already exist.

“The companies that have been out there for the last 20 years using desiccants to absorb are not transferring the heat of absorption like we are,” he says. “We have a unique proposition in that respect.”

“But the challenges the world faces around water supply, in the global south and here in the US, in places like Arizona and California, are not going to be solved by any one company, and they’re likely to get worse over the next five years. There are some great people working in this space, and I see a lot of synergies between us. It’s going to take a common heart and a common mission to tackle this problem.”

He adds: “The conversation in the data center industry at the moment is all about power, and we can reduce the amount of power needed by accessing the world’s largest aquifer - the atmosphere. AirJoule is the first technology that makes tapping into that aquifer a truly ethical prospect.” ●

You're as cold as ice: Antifreeze in liquid cooling systems



Niva Yadav
Junior Reporter

The importance of stopping things from freezing to keep things moving

Octavian Lazar/Getty Images

Liquid cooling has emerged as the data center industry's go-to for high-density servers and advanced workloads.

The type of liquid used varies from system to system, but with a much higher heat capacity than any other liquid and impressive thermal conductivity, water makes for an excellent coolant.

However, its downfall is the result of basic physics - water freezes. Frozen pipes are a disaster in any situation, from household plumbing to a car engine or data center cooling systems. But unlike a frozen pipe in a house, when it comes to

a data center's cooling loop, you cannot simply flush the system with hot water. That's where antifreeze comes in.

What is the point of antifreeze?

Antifreeze stops the water in the primary loop of a cooling system from freezing should the outdoor temperature become too low. It also plays an important role if servers stop working, preventing water from freezing in the absence of the heat usually provided by the hardware.

"Essentially, antifreeze is glycol," says Peter Huang, global vice president for data center thermal management at oil company Castrol. "The percentage

of glycol depends on how cold the environment is. The colder the environment, the higher percentage of glycol antifreeze that you need to use."

In a very cold environment, the percentage of glycol may reach up to 50 percent. In Canada, says Huang, operators are using percentages of up to 40 percent to prevent liquid cooling systems from freezing over. Concentrations aside, there are also different varieties of antifreeze.

The first is propylene glycol (PG), and the second is ethylene glycol (EG). While EG is much more effective when it comes to cooling, it is a toxic chemical with many potentially dangerous side

effects related to exposure. PG is less toxic and more environmentally friendly, but far less efficient as a coolant. This means, Huang argues, that its green credentials are up for debate, as a less effective coolant drives up the amount of electricity used to pump the water around the loop system to cool the chips.

Though traditionally a company focused on lubricants for vehicles, Castrol has been expanding its data center business in recent years, with a range of cooling fluids for different systems, as well as antifreeze. Its products are marketed based on the type of glycol they contain and the concentration, but it is worth noting that the product is not usually shipped as a finished good, Huang explains. To save on both energy costs and shipping costs, glycol is shipped as a concentrate. Once it is received at the data center, it can be blended with water.

Keeping track of the amount of antifreeze in a system is also important, as the liquid cannot simply be left alone. Huang says Castrol recommends its clients take samples every 50 to 100 days, as antifreeze can evaporate, particularly in dry, warm climates. Most of the refills happen in the primary cooling loop, he adds.

Cold and broken

Cooling is a complicated equation, Huang explains, and depends on the chip, the density, and the system as a whole. Cold climates may be branded as a big advantage for the cooling world, says Huang, but they make liquid cooling systems susceptible to freezing.

This is where antifreeze becomes an important part of the picture.

Cold regions such as the Nordics are often hailed as the data center industry's solution to rapidly growing cooling system power needs. Operators in countries like Norway routinely rave about their sustainability, reduced capex and opex, and improved efficiency, all achieved through free cooling, or using the external air temperature to chill cooling fluid.

In 2024 Montana-based cryptomine data center firm Hydro Hash suffered an outage after temperatures dropped from -6°C to -34°C (21.2°F to -29.2°F) in a little over 24 hours. Its 1MW facility lost power,

"Cold climates may be branded as a big advantage for the cooling world, but they make liquid cooling systems susceptible to freezing"

resulting in the water block cooling system freezing solid.

Castrol's Huang explains that operators can save a substantial sum by using direct and indirect free cooling, foregoing the need for mechanical chillers. For instance, he says, in Iceland, the ambient temperature averages below 15°C (59°F) and so can be used for free cooling, although he notes that data center operators don't design facilities that purely use direct free cooling.

And while antifreeze is more necessary in cold climates, data centers operating in more temperate regions still need to consider it an important part of their cooling systems.

Huang says that the most important phrase when designing a cooling system is "supply temperature," the temperature the cooling fluid needs to be when it reaches the rack to effectively cool the hardware. This can be calculated by working backward. For example, if the customer wants the water temperature inlet reaching the racks to be 26°C (78.8°F), the water in the primary loop needs to be around 20°C (68°F), which means the ambient temperature needs to be around 15°C.

An alternative to antifreeze

As it stands, Huang believes there is no emerging alternative to antifreeze. As

"The colder the environment, the higher percentage of glycol antifreeze that you need to use"

>> Peter Huang

long as water remains the best heat-conductive fluid and liquid cooling is necessary in the data center, antifreeze is the only solution to preventing unwanted frosty surprises.

However, LRZ München, the supercomputing center for Munich's universities and the Bavarian Academy of Sciences and Humanities, has deployed an alternative to glycol-based antifreeze systems, using nitrogen. The university decided to take a new approach to antifreeze in part because of its location adjacent to the River Isar. Glycol is particularly toxic for marine life, meaning a leak into the river could have damaging consequences.

Hiren Gandhi, scientific assistant at LRZ München, explains that the system uses purified water. In comparison to a water and glycol solution, the heat capacity and thermal conductivity of pure water are higher, and the viscosity of the fluid is lower, so less energy is required to pump the liquid around the system. "It's not a secret," he says.

"What we have here is a standard chemistry lecture in schools."

If the ambient temperature reaches freezing point, the water is immediately drained out, and nitrogen is sent through the pipes. Gandhi explains: "If we use air through the pipes, oxygen will corrode them. Therefore, we use nitrogen."

With the pressure of nitrogen entering the pipes, all water is drained. Gandhi adds that the system is designed "only for an emergency situation," and is activated when the servers stop working. If the servers are not down, they are able to supply enough heat so that the water never freezes, even in freezing outdoor temperatures.

Gandhi says this has enabled LRZ München to achieve 30-40 percent energy savings. The system could be put to use in larger data centers, he says, as it is easily scalable, though obtaining and storing nitrogen in large quantities is not without its problems.

Whether viable alternatives to glycol-based antifreeze emerge for commercial operators remains to be seen, but with more and more pipe work filling the newest generations of AI data centers, keeping the water flowing has never been more important. ●

Vertiv™ CoolLoop Trim Cooler

Power-efficient cooling redefined.

Achieve up to 3 MW capacity with 70% lower power consumption. Designed for AI densification, this adaptable, liquid-ready solution delivers safe, reliable, and energy-efficient performance, meeting ASHRAE and EU Ecodesign standards effortlessly.



Geochemistry at hyperscale



Zachary Skidmore
Senior Reporter,
Energy and
Sustainability

Microsoft is betting on Terradot's enhanced rock weathering tech to deliver durable, verifiable carbon removals at a global scale

As humanity continues the arduous task of weaning itself off fossil fuels, there has been an increased onus on the development of carbon removal technologies. Unlike capture and storage, these aim to remove CO₂ permanently and durably from the atmosphere.

The data center sector has emerged as the biggest customer for carbon removal-related credits over recent years. Leading the pack is Microsoft, which in 2024 alone purchased eight million tons of durable removals. As part of its effort to become carbon negative, the cloud company has taken a scattergun approach to purchases, backing options from the technologically advanced, such as Direct Air Capture (DAC), to the holistic, such as reforestation programs.

One technology that has received strong backing from Microsoft is enhanced rock weathering (ERW). Earlier this year, the company



signed a deal with Terradot, a firm born out of Stanford University's Soil and Environmental Biogeochemistry Lab.

Terradot officially launched operations in December 2024, following the close of a \$58 million Series A fund raise, and has also secured long-term backing from Google and Frontier - a carbon buying consortium including Meta and Google's parent company, Alphabet.

Microsoft's backing of Terradot was driven by its commitment to best-in-practice science, according to Brian Marrs, senior director of Energy & Carbon Removal at Microsoft. "This deal emphasizes science and highlights our thesis that carbon removal must continue to translate best-in-class science into action," he said when reflecting on the agreement.

Terradot's CEO and co-founder, James Kanoff, sat down with *DCD* to discuss the agreement, the intricacies of ERW, and the role the data center sector can play in turning the technology into an effective and verifiable source of durable carbon removal.

What on earth is ERW?

ERW is a fairly simple concept that is "just speeding up Earth's natural carbon removal process, which has been regulating the climate for over a billion years," Kanoff says.

It removes CO₂ through a chemical reaction, called carbonation, where CO₂ trapped in rainwater reacts with carbonic acid in rocks containing silicate minerals, such as Basalt. When rainwater interacts with the rocks, the CO₂ it carries mineralizes and is stored in a solid carbonate form. The natural process currently accounts for the removal of 1 billion tons of CO₂ every year, but, according to Kanoff, with some help, it could be accelerated significantly.

To speed up the process, ERW firms like Terradot mine silicate rock from quarries, which are then crushed to a powdery consistency (Kanoff likens it to "baby powder") and spread over wide-ranging cropland. By expanding the surface area available, ERW firms claim that the potential for carbon-binding reactions increases dramatically, transforming a process that could take millennia into one that works at scale within a matter of years.

"The faster the weathering means not only higher levels of carbon removals, but also more data, quicker validation, and a faster learning curve."

>>James Kanoff

Unlike technologies such as DAC, which require massive new facilities, or reforestation programs, which can take decades to implement, ERW can leverage existing agricultural and mining infrastructure for near-term deployments. As Kanoff explains, "There are already thousands of quarries around the world - surrounded by farmland - that have all the crushers, spreaders, haulers, and loaders we need."

As a result, proponents of ERW claim that compared to alternatives, the technology has the potential to offer a low-cost, simple, and fast scaling option for durable carbon removal. However, like many other carbon removal techniques, the sector remains very much in its infancy, with concerns persisting over the efficacy of its verification procedures and ultimately how much carbon it can effectively remove. Subsequently, at present, carbon-related ERW credits remain quite pricey, ranging from \$300 to \$400 a ton, according to carbon credit marketplace SuperCritical.

Despite this, ERW firms remain bullish that if scaled, the price per credit has the potential to drop significantly, with some projections as low as \$25 per ton. Therefore, the major challenge facing the sector today is accelerating the learning curve of the ERW process to determine what the best methods are to achieve maximal carbon removal, and in turn what regions are most amenable to its widespread roll out. It is here where offtakers like Microsoft play a significant role, not only in committing to large-scale removal purchases but also backing the projects themselves.

The perfect climate

While ERW works across the world, its overall efficacy is highly dependent on the climate conditions of the region. This factor played a huge role in Terradot's decision to site its operations in Brazil. According to Kanoff, Brazil offered several key advantages that would allow for the successful scale-up of an ERW operation.

The biggest factor was the country's tropical climate, which has been found, due to its high temperatures and rain, to lead to a faster rate of silicate dissolution, a crucial step in the rock weathering process. The speed of dissolution is not only crucial in the removal process but also in the generation of data, which allows the company to develop best practices at a much faster pace.

"Tropical environments weather rock faster - so our learning process is three times as fast, [which] on a tight climate timeline, is huge," said Kanoff. "The faster the weathering means not only higher levels of carbon removals, also more data, quicker validation, and a faster learning curve."

Another major factor was Brazil's sizable agricultural sector, with ERW highly dependent on extensive farmland for maximum productivity. The country has one of the largest agricultural sectors in the world, with more than 8.51 million sqkm of agricultural land. Therefore, the country's "vast farmland and hundreds of ideal rock quarries near farms" made Brazil stand out as the obvious option, Kanoff says.

With size comes experience. The Brazilian agricultural sector has extensive involvement with "rock-based fertilizers," where rock dust is used to remineralize soil to improve crop yields.

Subsequently, Terradot has leveraged a ready-made agricultural partner network with experience and a vested interest in greater deployment of rock-based fertilizers, taking a locally minded approach as all the rock used is mined at quarries within 50km of where they are spread.

So far, the company has spread nearly 50,000 tonnes of rock across 2,000 hectares of land and expects to generate the first carbon credits tied to the project at the end of the year.

The attraction to Brazil was not solely technical, Kanoff contends, but also societal. In Brazil, soil acidification is a huge issue, with low pH levels leading to low nutrient availability for crops, adversely affecting yields. ERW helps remedy this as it provides “micronutrients that are beneficial to crops and help manage soil pH, so it ends up being a win-win.”

As a result, ERW can have a dual benefit in these communities, says Kanoff. “The idea that we can remove carbon and regenerate soils at the same time? That’s incredibly powerful.”

Microsoft's backing

Microsoft is not new to the game when it comes to ERW, having previously signed agreements to purchase removal credits from [Undo](#) to remove 15,000 tons in Canada and the UK, [Eion](#) to remove 8,000 tons in the US, and [Lithos Carbon](#) to remove 11,400 tons.

However, the deal with Terradot differs from a traditional removal purchase. “They’re not just buying credits - they’re helping build the science,” Kanoff says. Therefore, instead of merely purchasing carbon credits, Microsoft has pledged direct support for Terradot’s field trials, lab work, and the data infrastructure that underpins the company’s ERW platform.

Kanoff argues that this makes the partnership particularly powerful because Microsoft is bringing not just capital, but systems thinking and technical expertise. “Tech companies are used to thinking in terms of scale, reliability, and latency, [which are] things that really matter when you’re dealing with a global carbon removal solution,” he says.

Terradot sees backing from the data center market as a means to supercharge the process of carbon removal towards being a cost-effective and scalable solution. Much of this is down to the willingness of data center firms to commit their capital and expertise to support the growth of these projects.

“Their demand helps scale up solutions like ERW while ensuring it’s done with scientific integrity. It’s not just about making their own footprint cleaner—it’s about catalyzing a whole new ecosystem of climate solutions,” Kanoff says.

“Tech companies are used to thinking in terms of scale, reliability, and latency, [which are] things that really matter when you’re dealing with a global carbon removal solution.”

>>James Kanoff

A crucial space where Microsoft and other data center operators can support carbon removal technologies is through the validation of carbon credits. This ensures the credits purchased are of high quality and drives improvements in the overall verification process, which could have a positive impact on the verification of other projects.

Verification verification verification

One major concern that permeates the carbon removal sector at large is the accurate measurement and verification of how much CO₂ a technology removes.

This is especially true of ERW. Due to the lack of long-term field trials, it is still not certain how much CO₂ ERW captures on a large scale. In addition, there is a lack of data on how different rocks, soil, and climate have on the efficiency of the removal. Different companies are experimenting with various silicate rocks, including basalt, wollastonite, olivine, and crushed concrete, which differ in terms of weathering rates.

As a result, verification is a central priority for Terradot, with the Microsoft deal seen as a means to accelerate both validation and deployment: “[We are] measuring everything from the topsoil down to groundwater... examining both near-field and far-field zones and trying to really understand the system,” says Kanoff. “The whole idea is to demonstrate true, durable carbon removal in a way that can stand up to scientific scrutiny.”

In doing so, Terradot aims to generate as much data as possible, which it can use to inform the validation process

of its carbon removal credits, with the ultimate goal of creating a standard for true, durable carbon removal. Partnering with a cloud provider like Microsoft supercharges this effort, as they “can help us model, analyze, and improve in near real-time. That’s how this becomes scalable, repeatable, and trustworthy,” Kanoff stated.

As a result, Terradot firmly sees itself as part of a new wave of carbon removal, which focuses on permanent carbon removal, rather than sequestration and storage. In ERW, where carbon capture occurs in complex, dynamic systems, “verifying removal is the entire challenge.”

For this reason, Terradot is investing heavily in AI, synthetic data, and process-based models to build a foundation of scientific integrity. The company has also partnered with carbon registries like Puro and Isometric and has made a concerted effort to exceed standard protocols associated with carbon credit verification. In 2022, Puro became the first company to launch a carbon crediting methodology for ERW, which provides verification through the form of CO₂ removal certificates.

While ERW remains early in its development cycle, several players have already begun to emerge across a range of climates and geographies. These include InPlanet, Undo, Vesta, and Eion, who are piloting ERW from Canada to Kenya, among many others.

The show of confidence from large-scale data center firms, such as Microsoft and Google, is proving crucial in supporting the growth of these companies. While most are merely focusing on the purchasing of credits, some, like Microsoft, are “investing in actual, durable carbon removal,” says Kanoff.

Doing so, they are supporting companies like Terradot in scaling quickly without fear of financial collapse, which permits a natural learning curve where companies can scale gradually. While there may be bumps in the road going forward, in determining whether ERW can be deployed at scale and verified effectively, the technology is very much here to stay as a crucial part of hyperscalers’ push to decarbonize their value chain through carbon offsets. ■

Independence Day



Matthew Gooding
Senior Editor

GDS International has rebranded as DayOne, and CEO Jamie Khoo has plans to grow across Asia and beyond

DayOne CEO Jamie Khoo's first experience of data centers was a far cry from the multi-megawatt deployments her company is making today.

"Back in 2002, I was part of the team that set up the data center business at ST Telemedia," she says, referring to the Singapore telco that she worked at for more than a decade.

"I was on the finance side, working through some of the funding requirements, and that helped me start to build an understanding of what data centers are about. At the same time, ST was an early investor in Equinix, and I was also quite involved in that.

"At that time, we were catering to enterprise customers - you're talking about 2kW per rack with very simple cooling systems. Then the cloud came in

and gave us a spurt of growth, and now we're looking at AI giving us another spurt of growth."

Growth is the name of the game for DayOne, which started life as the international arm of Chinese data center operator GDS. Now an independent company based in Singapore, Khoo has been tasked with growing its empire of data centers across the APAC region and beyond.

A new Day

GDS was founded in 2006. Backed by STT GDC and Hillhouse, it operates dozens of data centers across Greater China. The company formed an international unit to handle its overseas data centers in 2022, and the decision to rebrand this as a standalone business, DayOne, was announced in January 2025.

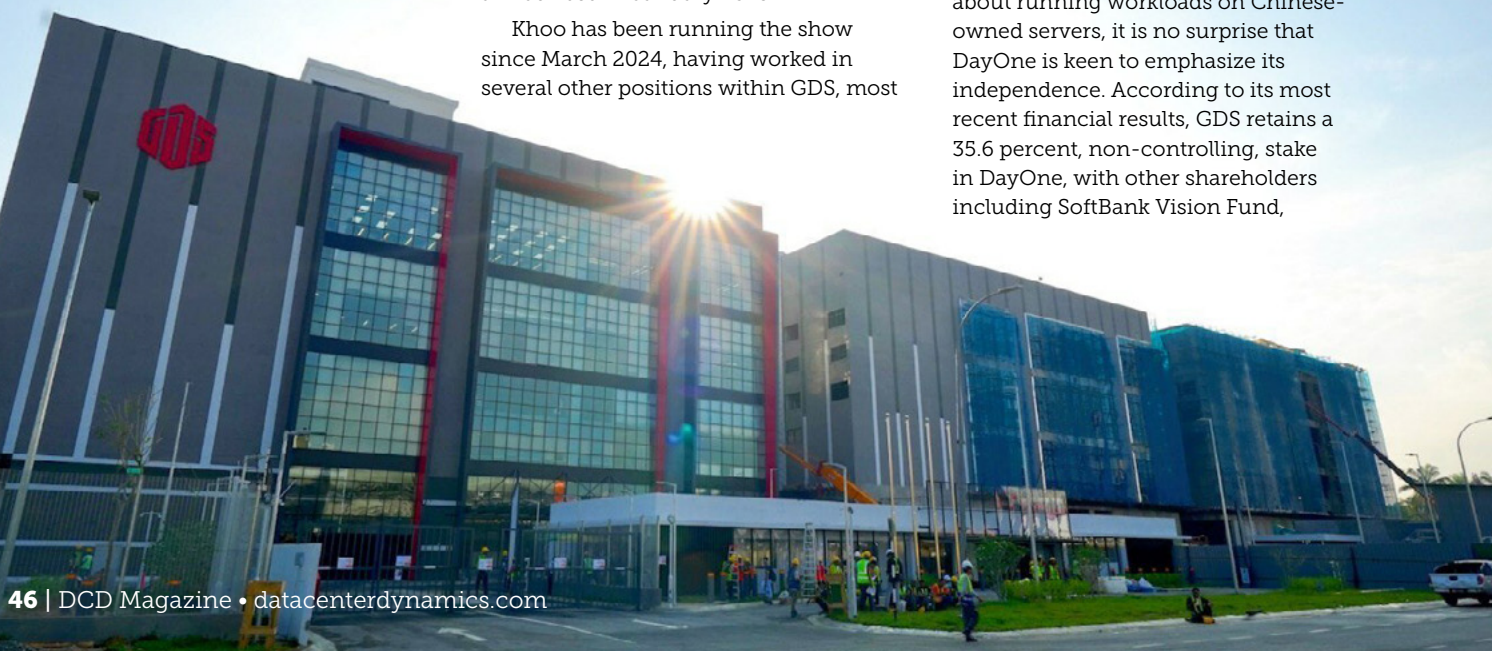
Khoo has been running the show since March 2024, having worked in several other positions within GDS, most

recently as group COO. She says it was always the company's intention to spin off its international operations into an independent entity, and that the new identity is a natural part of this.

"When we decided to go overseas, we planned to diversify our risk out of China," she says. "With the geopolitical tensions, it was identified from the beginning that the international team would need its own management team and decision-making body.

"We've been on that path ever since, but we realized we needed to take things slowly, so we raised Series A and Series B financing, then the deconsolidation from GDS, and finally the rebrand. But from the very beginning, operations and management have been totally separate."

Given that many of its international clients would probably have reservations about running workloads on Chinese-owned servers, it is no surprise that DayOne is keen to emphasize its independence. According to its most recent financial results, GDS retains a 35.6 percent, non-controlling, stake in DayOne, with other shareholders including SoftBank Vision Fund,



Citadel CEO Kenneth Griffin, Coatue Management, and Baupost Group.

On a business level, Khoo says: "We do get leads from our Chinese sales team, but those are arms-length transactions. We have our own sales and go-to market team who will then work with potential customers, both US and non-US companies, to understand their needs and work towards the kind of infrastructure they require in their data centers."

Building for the future

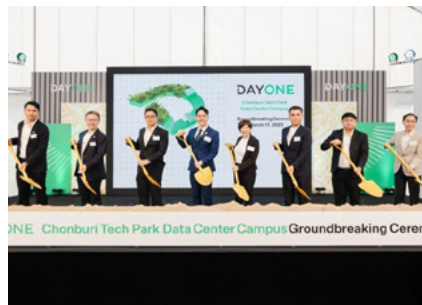
With \$1.9 billion raised across its Series A and B rounds in the last 18 months, DayOne has cash to fund a build-out across its six active markets in APAC: Indonesia, Hong Kong, Japan, Malaysia, Singapore, and Thailand.

Khoo says the company has about 480MW of data center space committed to customers, with more than 500MW under construction. It serves US and Chinese hyperscalers, as well as enterprise clients.

Data centers already operational include two facilities in Johor, Malaysia, located at the Nusajaya and Kempas Tech Parks, while in March, it broke ground on a new data center at Chonburi Tech Park in Thailand, which will offer 180MW.

"The money we raised will help us grow in our six markets," Khoo says. "Malaysia has been very successful because it was a new market at a time when Singapore was under moratorium, and we were very fortunate to be in that early spur of growth and be able to capture some of that market."

She identifies Thailand as one area with growth potential, on the back of multiple large investment announcements from the hyperscalers. AWS, Alibaba, Google, and TikTok have all announced significant projects in the country in the last 12 months. "It's in an early stage and I think we'll see growth there that's higher than the rest of the



"The cloud came in and gave us a spurt of growth, and now we're looking at AI giving us another spurt of growth"
 >> Jamie Khoo



region in the next couple of years," Khoo says.

The company is also hoping that easing restrictions on data centers in Singapore will play in its favor. Developments were banned completely in the city state between 2019-2023 due to limits on space and power, but in recent years, the government has started to relax restrictions, and DayOne, under its previous guise as GDS International, was one of four companies given permission to add 20MW of capacity in 2023. "We will be kickstarting our construction in



Singapore in July, which we hope to deliver by 2026," Khoo says.

Now up to 300MW has been made available for data centers, which shows they can build sustainable facilities. The successful bidders will be announced later this year, and Khoo is hoping DayOne will benefit. "Everyone is vying for [more capacity in Singapore] and talking to the government, and we are no different," she adds.

Beyond APAC

While Khoo expects AI to play an important part in DayOne's future, she says that "in this part of the world a lot of the demand remains cloud-based," which has an impact on the type of infrastructure offered in its data centers.

"We do have customers with AI

requirements, with different GPUs and different cooling demands," she says. "But compared to cloud and Internet content requirements, AI is on a much smaller scale."

"For us, we focus on delivery - if you can deliver fast, then I think demand will come, whether that be cloud or AI, or anything else. I believe the AI market in this part of the world is going to grow, but it will depend on the kind of technology we can get hold of."

As she moves into her second year as CEO, Khoo says growing in APAC is the big focus, though DayOne could look to other markets too. It was reported earlier this year that the company was considering going public, with Bloomberg suggesting it had contacted banks ahead of a potential IPO in the US.

"We want to grow outside Asia one day," Khoo says. "There are plans, and I hope we'll be able to do it faster than some people think."

She says her biggest challenge is around people management, and spends a lot of her time jetting between the company's various locations so she can stay connected with her teams. "Having a loyal, committed team that serves the market and our customers is the most important thing," Khoo says. "That's the thing I spend most time thinking about, because that enables a business to grow sustainably in the long term."

Khoo is a rare female CEO in an industry dominated by male executives, but neatly sidesteps DCD's question on whether she sees herself as a role model for others.

"I hope in the future there will be more female CEOs, not just in data centers but in other tech sectors too," she says. "It's a very rewarding job, and it's so fulfilling to be in a position to drive a business forward. I really hope more people can have this experience."

DATA CENTER SOLUTIONS

UNIQUE TO YOU

Designed with Your Needs in Mind



SEL is not just a supplier; we are your strategic partner in navigating the ever-evolving information and communications landscape. From network planning to design and integration, SEL offers a comprehensive suite of solutions tailored to the unique challenges posed.

SUPPORTING YOU AT EVERY STEP

- Network infrastructure expertise
- Technical consultation and support
- Customer-tailored innovations and solutions
- US & global manufacturing and supply chains
- Product training & warranties

[SumitomoElectricLightwave.com](https://www.sumitomelectriclightwave.com)

NEXT GENERATION THINKING™



A global CIO 44 years in the making



Georgia Butler
Senior Reporter,
Cloud & Hybrid

How Craig Walker, former downstream VP and CIO of Shell, grew through the decades

We often hear about the work that a CIO has done over a period of a few short years. But rarely do we get a look at several decades of experience.

Craig Walker is the definition of this. Having taken his first role as IT analyst/programmer at oil and gas firm Shell UK in 1981, fresh out of university, he went on to dedicate more than 40 years to excelling in this field. Or, in his own words, "by any yardstick, I got to the top of my chosen career."

Walker spoke to *DCD* about his lengthy career, compressing forty-odd years into forty-odd minutes. Throughout our interview, the passion he has gained along the way shone through.

He held many roles at Shell between 1981 and 2020, interspersed with a spell as a consultant at KPMG, where he advised other CIOs on decision-making in their various fields and companies. After returning to Shell, then leaving for a final time in 2020, he worked at Salesforce as a "strategic advisor to the office of the CEO, then as an advisor to the Great



"I spent three years in London, got very drunk with my Australian boss one night, and woke up the next day realizing that I had agreed to move to Saudi Arabia"



British Railways Transition Team, his current positions include co-founder and director at Veles Consulting and chair and president of CIONET UK.

Walker admits he has had a few failed attempts at retirement.

"Shell ruthlessly retires you at 60, so in 2020 I finished with them," he explains, adding that it's not a 'requirement' but a very tempting offer that is hard to refuse. He was almost immediately scouted by Salesforce, but requested a few months of rest.

2020 had different ideas, however. "I managed to retire for eight hours," says Walker, half jokingly. The Covid-19 pandemic almost immediately forced him back into work, and that led to a four-year stint at Salesforce.

"That was 43 years of corporate life now," Walker explains. "Then I decided to have a second go at retirement, but, of course, that failed as well."

Early years at Shell

Walker studied chemical engineering at university, but he tells DCD: "I would have made a terrible chemical engineer."

Explaining his transition into an IT role, he says: "Shell recognized that I loved programming. I spent three years in London, got very drunk with my

Australian boss one night, and woke up the next day realizing that I had agreed to move to Saudi Arabia."

After a stint in the Kingdom, where he was based in both Riyadh and Jeddah, Walker commenced a global tour, working in Dubai, Scotland, Colombia, Cape Town, and Texas, before heading back to London, where he finished his tenure as CIO of Shell's downstream business.

The early days of Walker's time at Shell are like a window into a different world. When he went to Saudi Arabia in 1984, there was no email or Internet. "You couldn't even phone head office, because it was too expensive and, besides, the phones didn't work half the time," he laughs.

At that point, the company used a Telex machine to communicate - an electromechanical device used for sending and receiving typed messages over telephone lines akin to a fax machine.

He recalls being told that he needed to implement some software upon his arrival, only to discover that he would have to do some of the coding in Arabic, as all the readouts and invoices in the region would also be in Arabic.

"When I started out at Shell, PCs hadn't even hit yet," he says. "You were sitting in

front of a terminal that was a hardwired box somewhere on the floor with some air conditioning. We would write some code, and put it on these massive disk platters that looked like LP records.

"You couldn't even take them on the tube because the current was too high on the rails, and it would demagnetize the disks. So we had to take them out to regional offices in the UK, load them there, then go back to head office."

It wasn't until the end of the 1990s, with what Walker describes as the "PC revolution," that Shell began to run "pretty powerful and sophisticated machines."

One pivotal shift Walker recalls is when he replaced the Telex system with email.

"I remember when I brought it in, I was summoned by the then-CEO who said to me: 'Walker, this was your idea, and I agreed to it. But how do people sign off on these emails?' and I said, 'Well, they don't. You write an email, and you send it.'"

"His response? 'I'm not too sure that's going to work.'"

But once the Internet hit, "we started to get more interoperability," Walker says.

The Apotheosis

Walker's first CIO role at Shell was heading up IT operations at Shell Markets in

Dubai, starting in 1986. He later became DS Pan-African CIO, VP, and global CIO of trading and supply with several other IT roles smattered in between, before finally settling down as the global CIO of downstream business, which covers all the B2B and B2C elements of the company - from retail, to trading, to refineries - in 2014.

"The big challenge of the downstream business was that the place was a mess," Walker says. "It took us about four years, but we managed to get [operating] costs down from around \$2.5 billion to \$1.3 billion. We virtually halved it, and saw a much better performance."

A big part of that work was aiding Shell in its migration to the cloud and out of its data centers.

According to Walker, a lot of that cost reduction was made by getting rid of global systems and renegotiating contracts. "We also recognized the legacy of our data centers in terms of facilities," he says. "The cost was too high, and we had to learn how to move to the cloud."

He says a lot of people within the firm were "skeptical" about the migration to begin with. "People worry about [the cloud] - particularly the big companies that tend to try and customize stuff a lot," Walker says, noting that it was in the early 2010s that the company finally made the "major decision" and decided it would migrate onto virtual servers.

With many moving parts in Shell, the "cloud migration" journey was a complicated one, and for different parts of the business, happened on different timelines.

The nerves on the part of senior management, however, remained consistent.

In 2014, the Shell enterprise product manager, Oskar Brink, speaking at an Amazon Web Services (AWS) Summit in London, told delegates: "It's been a four-year journey and a lot of scepticism from many of our stakeholders."

"Will it work? Is it secure? Can we get assurance that it will continue to work?", were just some of the questions stakeholders asked."

According to Brink, that particular migration effort to AWS private cloud saw between 8,000 and 10,000 applications rationalized, and took four years to complete. By the end, the company had

"By any yardstick, I got to the top of my chosen career"

reduced its load to 5,000-6,000 unique applications, which Brink said could be categorized, enabling the rationalization of the environment.

Walker estimates that by 2016-2017, it had shut down most of its data centers.

He recalls the company had three major data centers in the Netherlands, a massive one in Houston, and a big one in the UK, among others, as well as small Edge-type deployments - described as "a couple of computers in a cupboard somewhere with an AC."

The exact distribution of Shell's IT footprint today is hard to get a clear look at, though DCD has reached out to the company for information about how many - if any - data centers the company still operates, and which cloud platforms it is using for operations.

DCD reported in 2012 that Shell had renewed its data center outsourcing contract with T-Systems. The company had been outsourcing its data centers to T-Systems since 2008, with the 2012 renewal spanning five years, and included T-Systems helping Shell move customer data onto a cloud infrastructure, and was then valued at \$1.2 billion (approximately \$1.64bn today).

At that point, the company had already moved its global dynamic SAP services into the cloud. T-Systems was reported to have data centers in Europe located in



Craig Walker

Debrecen, Hungary, St. Petersburg, and the Czech Republic in 2012.

Beyond the reported relationships with AWS and T-Systems, Shell is also a known customer of Microsoft.

Ethical considerations

It's impossible to write about Shell - or any oil company - without noting both the detrimental impact that such industries have on the environment, and the political tensions that are frequently associated with them.

Cloud computing providers have also come under fire for their willingness to provide cloud services to the oil and gas industry.

Both Microsoft and AWS are keen to promote their green credentials. In 2019, AWS said it would be net-zero carbon by 2040, and added in 2023 that 100 percent of its electricity consumption was matched with renewable energy, though this claim has been disputed by some of the company's own employees.

Microsoft is looking to be carbon negative by 2050, and aims to have removed from the environment all the carbon it has emitted either directly or by electrical consumption since it was founded in 1975.

Both cloud providers are also big proponents of Power Purchase Agreements (PPAs) and carbon capture technologies, all of which seem somewhat contradictory to simultaneously getting into bed with Big Oil.

Despite they're greener pronouncements, they have not shown any inclination to turn down the lucrative contracts offered by companies like Shell.

Shell itself has a "goal" to be climate neutral by 2050, and the company's website claims it achieved a 9-12 percent reduction in the net carbon intensity of its energy products in 2024 compared to 2016, and aims to increase this to 13 percent this year.

There is some doubt, however, about the company's commitment to this effort. Its 2024 annual report notes that, of its capex for that year, \$2.2 billion went on non-energy products, \$2.4bn on low-carbon energy solutions, \$5bn on liquefied natural gas, gas and power marketing and trading, and a massive \$11.5 billion on oil, oil products, and "other."



Shell Immersion Build Week 13

The report notes that Shell defines low-carbon energy as anything with “an average carbon intensity that is lower than conventional hydrocarbon products, assessed on a life-cycle basis.” Whether this includes natural gas is unclear.

Notably, the \$2.2bn is lower than both 2023 and 2022, which saw Shell investing \$3.5bn and \$2.7bn respectively in low-carbon energy solutions.

In 2024, Shell announced that it would weaken its own efforts to reduce emissions per unit of energy, scrap some of its climate goals, scale back renewable efforts, and double down on fossil fuels.

The company laid off workers in its greener divisions, including hydrogen, and expanded oil and gas projects.

Shell continues to explore for new sources of oil and gas, and does not expect to reduce the overall amount of fossil fuels it produces by 2030, the date by which IPCC scenarios say emissions from oil, gas, and coal will need to have substantially reduced to avoid global calamity.

DCD asked Walker if he had any qualms about the ethics of Shell's business.

“I think the company is doing all the things it can,” he says. “I do believe Shell ran to a very high standard across the world, and I can't ask any more of the company. I was pretty senior, and I was privy to the board decisions, and safety and the environment were always top of our minds.”

He also pointed to the world's current reliance on fossil fuels, beyond just energy. “If [the wider public] wanted to stop using oil tomorrow, how many things in your life do you think are still going to exist? So much is based on petrochemicals.”

Indeed, Shell's net-zero target only covers energy products, and excludes its petrochemicals business - including plastics - the production of which has a significant carbon footprint.

Walker continues: “The reason we have cheap energy and we've learned to live a certain type of lifestyle is because of oil and gas. Now, do I personally think it's sensible to burn oil and gas? No, I don't. I think that's a mistake.

“It's such a complex story that I never felt I was doing something immoral. Certainly, I was never asked to do anything immoral or ethically wrong.”

As for the company's ties to countries Walker concedes are often viewed as “corrupt and dodgy,” the former CIO says: “[Oil] has made them so much money over the decades, and that opens up the danger that money can get siphoned off in a way that we might not consider morally or ethically right.”

He adds: “We were extracting the oil and doing it in a safe way. The company gets its cut, markets it, and sells it on their behalf. Shell doesn't get a say on what that government ultimately does with the money.”

It isn't hard to find stories of corruption related to Big Oil in countries with large resources.

For its part, Shell found itself in the news earlier this year due to a cleanup operation in Southern Nigeria, when oil pollution was allegedly caused by the company's work. A BBC investigation in February 2025 found that the eight-year project is, perhaps, not going as well as Shell and the Nigerian government have previously claimed

One close observer described the clean-up project as a “con” and a “scam” that has wasted money and left the people of Ogoniland stuck living with the pollution.

A civil trial at the High Court in London was held, during which lawyers representing two Ogoniland communities of around 50,000 inhabitants argued that Shell must take responsibility for oil pollution that occurred between 1989 and 2020, which was allegedly caused by its infrastructure. A decision on the case is still pending.

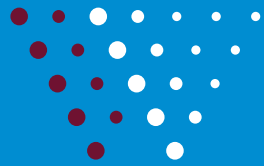
The present

Having failed at his second attempt at “proper” retirement, Walker is now co-founder and director at Veles Consulting, where he and several of his former colleagues, friends, and acquaintances pass on their extensive knowledge to other companies.

“Veles Consulting really came about because a bunch of ex-KPMG consultants got together and realized we were at a similar stage of life,” he says. “A lot of us were either retiring or thinking of retiring, and it became obvious that there was a huge amount of experience and knowledge in the room.

“Someone, who I think had had a lot less wine than I at that point, said we ought to get together more often, and that eventually became a consultancy. Since then, we've added around 35 people to Veles.”

It would seem Walker is unable to sit still. After decades of running around, he is not used to, well, not working. But he says one positive to come out of his latest endeavor is that the company now gives one percent of both its profits and time to charities and good causes. With the other 99 percent, Walker is doing what he knows best. ■



Winthrop Technologies

Turnkey Delivery of Digital Infrastructure



A dedicated data centre delivery partner,
headquartered in Dublin and delivering
throughout Europe

winthrop.ie

A stroll down Meet-me Street

Towardex believes Meet-me-vaults can disrupt the data center industry's hold over cross-connects



Depending on who you ask, the city of Boston will represent different things.

Some will think of Fenway Park and the world-famous Boston Red Sox baseball team, or the city's equally well-known basketball outfit, the Boston Celtics.

For others, it might be the TV sitcom *Cheers*, or the Boston Tea Party, an act of protest against the British government seen as a key part of the American Revolution.

Over the years, the Massachusetts city has also built up a culture of defiance that often challenges the status quo.

It's this defiance, argues James Jun, managing director and IP Core network architect at Towardex, that has now spilled over into the city's Internet infrastructure battle.

"This defiant culture has now fully propagated into Boston's Internet infrastructure, and there is now an open rebellion by telecommunications carriers against exploitative data center cross-connect practices—we've had enough, and the rent is too damn high," Jun said in a lengthy LinkedIn post last year.

Dropouts drop in on interconnectivity

Jun's firm, Towardex, is an independently owned network infrastructure provider that delivers dark fiber, blended bandwidth, and connectivity between data center facilities in the Boston area.

Founded in 2012, the company has built itself up as a connectivity hub for Internet networks in the city.

Jun and Gavin Schoch, both university dropouts, set Towardex up due to their frustrations at the inefficiencies of networking and Internet infrastructure in New England. Jun describes himself on LinkedIn as a New England interconnection specialist.

"Back in 2012, when the retail colocation was pretty strong in New England, a lot of customers were looking for a good well-peered, well-connected IP bandwidth provided for all of their colocation needs whenever they go between data centers," says Jun.

"We wanted to create a network service for our data center customers that really caters to their performance needs." The company claims to be the go-to provider in Boston, where it can help connect data centers with their tenants, supporting colos in the process.

Initially set up as a network provider, Towardex provides interconnection services for the largest telecommunications carriers, hyperscalers, enterprises, data centers, governments, and universities.

Meet me vaults

The whole premise of the company's business is to facilitate interconnectivity, and it does this through the deployment of meet-me vaults, or as Jun describes it, a meet-me-street.

These meet-me rooms tend to be more common in carrier hotels and more standard colocation facilities, which allow for the exchange of data and traffic between different networks. They aren't something that every data center will have, but are standard in almost all multi-tenant facilities.

The vaults, however, are designed to provide interconnection between different providers, such as telcos, as well as Internet and cloud service providers, with the data center but outside the facility.

Typically found at the outer edge of a data center, the vaults are buried a few feet under the ground.

Towardex's meet-me vaults are known as its Hub Express System (HEX), which operates as a common carrier, a public utility system for fiber. The company says that HEX can provide open-access underground utility for Boston's data center networks.

A typical HEX underground vault is generally a rectangular concrete box, explains Jun, who says they are 12 ft by 6 ft in length and width, and about 8 ft tall. Jun adds that a single underground vault isn't large enough to host the volume of connections between many entities.

The HEX system spans about a kilometer span of urban cityscape, underneath the streets in an entire neighborhood, with huge conduit banks (24 to 30 of 4" conduits in largest sections) and a network of over a dozen underground vaults, all interconnected together with lots of physical pathways

and high-count fibers, with no monthly recurring cross connect fees.

"Rather than doing 'Meet-Me' in a single vault (which doesn't scale), the HEX system transformed an entire street and neighborhood into a giant web of interconnected Meet-Me fabric," says Jun. "The connecting parties are often separated by several hundred feet and are located in different vaults, connected by meet-me fabric fibers and conduits."

Further explaining the importance of meet-me rooms and why they matter, Jun says they offer neutrality.

"Many carriers traditionally had difficult times working between each other, as they are often competitors," says Jun.

"Meet me rooms were usually operated by data centers or real estate owners who are not telecom carriers and therefore remained neutral (not directly competing with their carrier tenants). This neutrality of meet-me rooms is quite possibly the most important tenet of the business, allowing competing carriers to meet and exchange Internet traffic between each other at a neutral meeting ground."

The first section of the HEX system opened in 2023 to telecommunications carriers at Somerville's Inner Belt, where every telecom manhole is part of an interconnected cross-connection fabric, explains Jun.

"HEX was started because what we needed in Boston was a new and truly neutral meet-me system that is not tethered to a single private property," Jun says.

To construct the HEX system, Towardex uses precast construction (concrete) to develop durable and scalable underground network facilities. The network is made up of watertight cable vaults that interconnect.

He adds that the entire utility is an open-access system, licensing out duct space to everybody for \$1.54 per foot per year.

For context, Jun explains that this is significantly cheaper than what some of the data centers charge.

"An unnamed customer recently quipped that they're eliminating more than \$30,100 per month in potential meet-me room cross connect fees and are moving those cross-connects over to the HEX system," explains Jun. "To put



this in context, at \$350 month per each cross-connect charged by the data center, \$30,100 per month is 86 connections.

"In the HEX system, this customer would pay a total of \$1,475 per year (or ~\$123 per month) in conduit license fees (which is largely from \$1.54/ft/year price, plus additional fee for manhole space rental to place fiber equipment, etc.) to install a 288-strand fiber optic cable, which can support up to 144 connections."

Tier 2 markets

The need for these vaults, according to Jun, is due to Boston's status as a Tier 2 data center market in the US.

That said, Massachusetts is no slouch when it comes to data centers. A number of companies have data center facilities in and around the state and Boston, including Centersquare, Digital Realty, CoreSite, Equinix, Iron Mountain, TierPoint, Verizon, and tower company Crown Castle.

Meanwhile, Markley Group owns New England's only carrier hotel and provides connectivity to more than 50 domestic and international network providers.

"A lot of these types of markets across the United States have a challenge where they develop, and different data center competitors start fighting with each other," says Jun. "That creates some of the tensions and challenges for a lot of telecom providers and network tenants of those data centers."

"In the data center space, when a certain data center decides to become their own telecom provider, that data center could unilaterally refuse to renew a telecom provider license or colocation space, and that could create some issues," says Jun.

During Towardex's formative years, Jun recalls that the data center market in Boston became more competitive, to the point that it was impacting tenants.

Empower more data center companies to compete

It was at the point that the tenants were becoming affected by the situation with higher rents, that Jun remembers the need to push for something different.

"When the tenants started to get impacted, that's where we saw the need to take this meeting room dominance away from these guys and put it underground," Jun explains. "That is not to say to hurt or diminish the need for the connective data centers. That's not our goal. Our goal is to empower more competitive data centers to be able to compete in the market."

To achieve this, Jun says it was important to create an infrastructure that is based on the public rights of way that complements the existing data centers, which "forces these competing data centers to start playing nice."

It's for this reason that Jun says the company assessed how utilities are being used to drive connectivity.

In his words, the purpose of HEX is to create was to create an open utility system for everyone to use. The open element was important because Jun did not want to simply create another system "for the existing competitive telecom providers in the market that have the country on lockdown for themselves."

He explains: "We wanted to create a truly open access system where anyone, even networks that are not from Boston, could come into the manhole and do their interconnections without any concerns of dealing with competitors locking up their cross-connects, or being locked out of conduits."

Jun likens the HEX system to real-world airports, noting that, unlike data centers, the HEX system cannot discriminate against competitors, and must provide equal access to everyone on a competitively neutral and nondiscriminatory basis.

"As long as carriers are paying their fare/carriage fee to be in the system and are abiding by technical and safety rules, we cannot kick them out of the system. Conversely, data centers can, and often do, kick carriers out, for competition reasons," he adds.

The system is underground and in the public rights of ways, adds Jun, noting that it's also regulated.

Jun says that the two main principles of the meet-me vaults are to provide open access and to operate in a "pristine" manner that continues to guarantee the tenants' access over time.

In a project charter set out by the company, Towardex outlined its hopes to "operate a new multi-cable, multi-conduit fiber optic transmission system dedicated to promoting interconnection freedom, accessibility, network neutrality, digital equality, freedom of enterprise, and freedom of expression for all."

MASS Effect

HEX links in with the Massachusetts Internet Exchange, otherwise known as MASS IX, which was launched by Jun and Schoch in 2015.

The exchange is distributed, neutral, and enables public peering, cloud connectivity, and data center interconnections throughout New England. It's currently present in more than six data centers in the Boston metro.

Connections into MASS IX are available for member utilities of HEX, Towardex says, and this allows fiber providers to sell access into MASS IX natively from their network, bypassing legacy carrier hotels and without the hassle of cross-connect fees.

"The MASS IX currently captures most of the data centers in the Eastern Massachusetts around the Boston Metro," says Jun.

"There's a build happening for MASS IX to extend it to the rest of Massachusetts, but that's an ongoing project. For the most part, most of the network services that we provide are focused on Boston."

More openness

Challenges to interconnectivity come in all shapes and sizes, and Jun says one problem his company is trying to solve with HEX is access to manholes.

"Access is a real challenge in a lot of these large manholes, so we wanted to solve those challenges and maintain our manhole systems in a pristine way," he says, with the aim of making it easier for tenants to access their conduits safely and without fear of damaging other cables.

Other infrastructure, such as utility

"HEX was started because Boston needed a new, truly neutral meet-me system that is not tethered to a single private property"

>>James Jun, Towardex

poles and underground conduits, can also be difficult to get to, Jun says. "The principal problem with the utilities in general across the country is that nobody could find access to the manhole or conduit system," he adds.

"So when a new carrier comes into Boston and they want to start interconnecting onto the streets, everybody that owns conduits doesn't want them in there because they see this newcomer as a potential competitor."

Jun believes there is a better way. "Our goal was to create a large network with tons of conduits with over 24 to 34 conduits, lots of pathways, and it's regulated, low cost, affordable rent in big manholes. Let's network them all together."

Jun says Towardex has the ability to open up this entire system to everyone that comes into the market.

"The previous way utilities operate in Boston is that if you're a telecom provider trying to come in, the existing telecom guys will do everything in their power to ensure you don't get in," he says. "We're the opposite of that."

Towardex states that all tenants of the HEX system can easily license conduit space to run new cross connect cables between themselves. According to the company, HEX also allows the data center's colocation customers to become their own dark fiber provider and run their interconnections out in the streets without being subject to cross connect fees.

Taking HEX beyond Boston

Jun, a native Bostonian, is aiming for total network domination in his home city.

"Our goal is to continue our current role as a network provider in Boston, where we are essentially the facilitator for all of the major network deployments that are happening in Boston," says Jun.

"So we're going to continue that model to help our customers deploy and come to Boston, and really promote Boston as a global city for Internet traffic."

That said, long term the goal is to expand beyond Boston and Massachusetts, and into other US Tier 1 and Tier 2 markets, adds Jun.

"Long term, we'll take a look at how we succeed with this HEX model, this underground meet-me system, and think about if there is a need for a system like that in other markets, and if there is, to what extent."

Jun is confident that HEX can be implemented in other markets, but only if it's a good fit for that particular metro area.

"The reality is that not everywhere needs a system like this," he explains. "Just like Internet exchanges and meet-me rooms, there has to be a compelling business case and a market challenge that we're trying to solve. But there are definitely markets out there that would certainly benefit from a system like this."

"When a market is suffering from competitive problems and there's a demand for a lot of fiber, that's when a system like HEX could come in to ease that. It really depends on each individual market landscape," he adds.

For now, these future expansion plans are being considered, says Jun, though he does note that some additional partners would be needed to help this become a reality.

Whether the expansion into the other markets happens sooner or later isn't a concern for Towardex. Instead, the opportunity to drive interconnection is the key goal.

"We started this company as network engineers because we love the Internet," he says.

"The best part about the Internet is interconnection because the word Internet is all about interconnected networks."

"If there's no connection, there is no Internet, and that's not only challenging but the most satisfying part of being in this space." ■



COMPREHENSIVE ENGINEERING EXCELLENCE



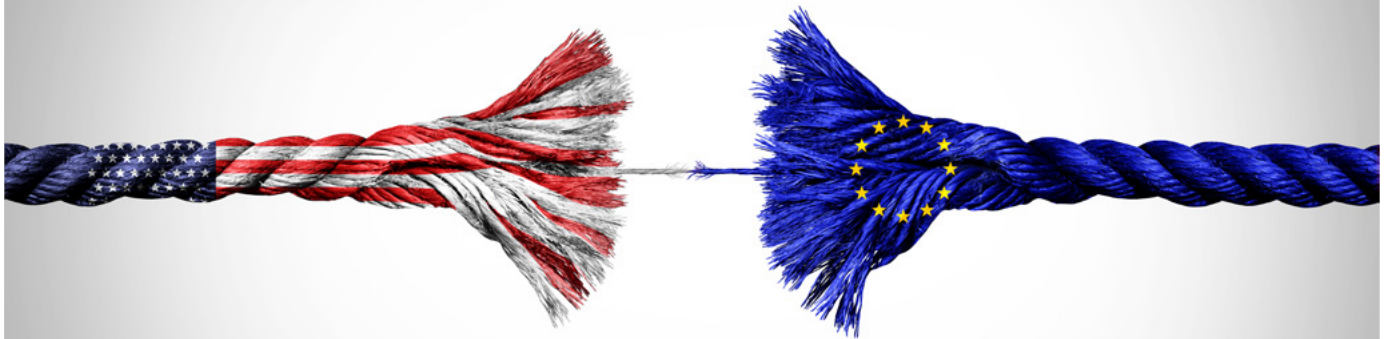
www.anordmardix.com



Fighting in a cloudy arena



Georgia Butler
*Senior Reporter,
Cloud & Hybrid*



The battle to increase
competition in the cloud
computing market

Sharing is hard. It's one of the first things you learn as a child, and for most people it's something that doesn't come naturally.

While, on a human level, most of us eventually figure it out - we let our peers have a few of our Skittles on the playground - it isn't a universal value. It also isn't business.

According to Synergy Research Group, the cloud computing market was worth \$330 billion in 2024. For comparison, Finland's GDP in 2023 was \$295 billion.

Within this enormous market, only three companies have a meaningful position: Microsoft, Google, and Amazon through its Amazon Web Services (AWS) business.

That same Synergy Research paper said that Amazon held a 30 percent market share, Microsoft 21 percent, and Google 12 percent, coming to a cumulative 63 percent.

The remaining 37 percent is split between all remaining companies - the well-established competitors like Oracle, OVHcloud, and IBM, and the Neocloud providers like CoreWeave, all the way down to the small cloud companies starting up with regional offerings.

The hyperscale trio's dominance hasn't gone unnoticed, and in the US the Federal Trade Commission has reportedly been investigating Microsoft's alleged anticompetitive practices, though this has never been confirmed by the government department.

Elsewhere, regulatory action is very much a reality.

Both in Europe, after a complaint lodged by CISPE (Cloud Infrastructure Services Providers in Europe) to the European Commission about Microsoft's licensing practises, and the UK, where competition watchdog the Competitions and Markets Authority (CMA) is looking at the overall cloud services market, tangible efforts have been made to curb the dominance of the three US hyperscalers.

Thus far, however, it seems to be to little avail.

Part of the reason the issue of competition has been taken more seriously in Europe than the US is a distinct unease with the idea of putting European data into the care and responsibility of American companies, but perhaps more importantly, the idea

*"At the end of the day,
it's like the roach motel.
You can check in, but
you can't check out,"*

>> Kevin Cochrane, Vultr

of giving money to companies that could instead be going to domestic cloud providers.

Last year, DCD spoke to Francesco Bonfiglio, former CEO of Gaia-X, an organization campaigning for federated and secure data infrastructure in Europe, about the current market share in Europe. His perspective was less than optimistic.

"We have three years," Bonfiglio told DCD at the time. "In 2017, we [European cloud providers] had 26 percent of the market share, now we are at 10 percent. We can't turn that back, so in three years it will be zero. Meanwhile, you have American providers buying data centers and gigawatts of energy."

"My question is, what's the value in having digital regulation if there is no digital market to regulate?"

The basis for concern about competition was summed up to DCD by Vultr's Kevin Cochrane.

"There are a few primary problems. Number one is that the hyperscalers leverage free credits to get digital startups to build their entire stack on their cloud services," Cochrane says, adding that as the startups grow, the technical requirements from hyperscalers leave them tied to that provider.

*"What's the value
in having digital
regulation if there is
no digital market to
regulate?"*

>> Francesco Bonfiglio

"The second thing is also in the relationship they have with enterprises. They say, 'Hey, we project you will have a \$250 million cloud bill, we are going to give you a discount.' Then, because the enterprise has a contractual vehicle, there's a mad rush to use as much of the hyperscalers compute as possible because you either lose it or use it.

"At the end of the day, it's like the roach motel. You can check in, but you can't check out," he sums up.

Mainland Europe's woes

The issue of anticompetitive practices in the cloud market has been in discussion for years at this point, but it wasn't until 2022 that two European providers - Aruba and OVHcloud, officially lodged a complaint to the European Commission, specifically focused on Microsoft.

CISPE, of which AWS was a member at the time, backed that complaint.

CISPE has been operating since 2015 but was officially registered in 2017. The trade association campaigns for a EU-wide cloud-first public procurement policy and for fairness across the industry, including software licensing practices and avoiding vendor lock-in.

Microsoft has frequently been singled out in discussions about competition in the cloud market for its software licensing practices.

CISPE itself officially filed a complaint against Microsoft in November 2023 alleging that the cloud provider had unfair licensing practices. In a press release about the filing, CISPE wrote: "Microsoft's ongoing position and behaviours are irreparably damaging the European cloud ecosystem and depriving European customers of choice in their cloud deployments. CISPE feels it has no option but to become a formal complainant and to urge the European Commission to act."

Microsoft has not responded to any requests from DCD for comment or an interview to discuss competition in the cloud market.

DCD spoke with CISPE's director of communications Ben Maynard recently at a Data Center Nation event in Milan, Italy. Against the backdrop of the conference room hum, Maynard explained a bit more about the whole process, from inception to date.

"We launched our complaints against

Microsoft with the European Commission in November 2022, and that was driven by a number of members - particularly those in mainland Europe - who felt that it was getting increasingly hard to compete with Microsoft because everyone wants to use Microsoft software, and its easier and cheaper to buy if it runs in Azure.

"We did a lot of work, and commissioned some research that showed what this was actually costing the European economy," Maynard tells *DCD*.

The research in question is likely that released in June 2023 by Professor Frédéric Jenny, which found that "additional charges levied on those choosing a non-Microsoft cloud when using the Microsoft SQL Server software sucked an additional €1.01 billion (\$1.13bn) out of the European economy in 2022."

Microsoft itself has previously conceded some wrongdoing. President and vice chair of Microsoft, Brad Smith, was reported by the Financial Times in 2021 as saying in a statement: "While not all of these claims are valid, some of them are, and we'll absolutely make changes soon to address them."

Microsoft, he added, was "committed to listening to our customers and meeting the needs of European cloud providers."

Maynard tells *DCD* that this sentiment was shared in CISPE's communications with Microsoft, with the cloud giant telling CISPE that they wanted to find some kind of solution.

"We spent a year negotiating with them on what a solution could look like. And they presented a solution to us in late spring/early summer 2024, we shared that with members who thought it was, on balance, worth giving them the benefit of the doubt."

It was in July 2024 that Microsoft and CISPE officially came to an agreement. That agreement saw Microsoft paying €20 million (\$21.7m) to CISPE members in return for the complaint to be withdrawn, as well as to develop a product



- Azure Stack HCI for European cloud providers (Hosters) - that enables CISPE's members to run Microsoft software on their platforms at equivalent prices to Microsoft's.

The company also agreed to compensate CISPE members for lost revenues related to their licensing costs for the last two years.

While Maynard explains that this was a decision that was made by the association - and thus voted for by a majority of the members - it was not universally popular.

Dissatisfied competitors

AWS spokesperson Alia Ilyas said that Microsoft was only making "limited concessions for some CISPE members that demonstrate there are no technical barriers preventing it from doing what's right for every cloud customer." Ilyas added that it would "nothing for the vast majority of Microsoft customers who are still unable to use the cloud of their choice in Europe and around the world."

AWS is and was a member of CISPE.

Amit Zavery, head of platform at Google Cloud, added: "Many regulatory bodies have opened inquiries into Microsoft's licensing practices, and we are hopeful there will be remedies to protect the cloud market from Microsoft's anti-competitive behavior.

"We are exploring our options to continue to fight against Microsoft's anti-competitive licensing in order to promote choice, innovation, and the growth of

the digital economy in Europe."

Mark Boost, CEO of UK cloud company Civo, said: "However they position it, we cannot shy away from what this deal appears to be: a global powerful company paying for the silence of a trade body, and avoiding having to make fundamental changes to their software licensing practices on a global basis."

In the months that followed this decision, things got interesting.

Reports emerged that Google offered

CISPE €14 million (\$15.3m) in cash and €455m (\$497.5m) in software licenses on the basis that CISPE continue its antitrust case. AWS reportedly contributed €6m (\$6.56m).

This has not been confirmed by either Google or CISPE, but the reports originally from Bloomberg cited "confidential documents." Around the same time, the Register reported that Google was similarly considering joining the trade association.

Sources with knowledge of Google have spoken to *DCD* about the topic. While unable to comment on the alleged "offering" to CISPE, noted that Google had asked CISPE to stick to the "Fair Software Licensing Principles" that they had agreed to uphold.

"Google heard that they were entertaining a proposal from Microsoft that would require CISPE to violate its own standards and principles, and Google made clear they would only join if CISPE would hold true to its principles," the source told *DCD*.

The *principles detailed* by CISPE that are pertinent to this include: "Customers that seek to migrate their software from on-premises to the cloud should not be required to purchase separate, duplicative licences for the same software," that "Licences that permit customers to run software on their own hardware (typically referred to as "on-premises" software) should also permit them to use that software on the cloud of the

customer's choice without additional restrictions," and also "Software vendors should not penalize or retaliate against customers who choose to use those vendors' software on other providers' cloud offerings," which includes charging more for licensing fees.

Regardless of complaint or attempted intervention, CISPE and Microsoft settled.

Part of the agreement required Microsoft to develop a product that would seemingly satisfy these software licensing issues. At the time it was called HCI Stack for Hosters, but has since been rebranded as Azure Local.

To monitor this, CISPE established the European Cloud Competition Observatory (ECCO). Interestingly, while managed by CISPE and governed independently, CISPE notes of ECCO that "technical experts contribute as needed, and Microsoft participates in the initiative."

As of February 2025, Microsoft was awarded an "Amber" status on its progress with the project in its first report.

"There were some things that they had done really well," says Maynard. They had started a beta program. They've got people onboarded. People were trying the product. They had delivered some of the easier things to do in the settlement. They had given us money to recompense for the amount of time and effort we spent, and they had given some money to the members as well.

ECCO went to Redmond, Washington, in December 2024 to view Azure Local. "It did some things, but it didn't do all the things it needed to do," admits Maynard.

As to the reason behind the delay, Maynard said: "I think it's legitimately technically challenging. The thing we really need is multi-tenancy. As a hoster, you want to be able to host as many customers on one hardware infrastructure as you can. Azure Local doesn't really do a good job of that. It's not built for that. To a



"When it came to it, it was a question of 'put your head over the parapet or your hand in your pocket,' and those in discussions would get stunningly silent,"

>> Nicky Stewart, OCC

certain extent, it's trying to fit a round peg into a square hole.

"I don't think there's any dragging of feet, I think it's just more challenging than they thought it would be."

In May 2025, CISPE ruled that the product Microsoft was proposing failed to meet its requirements, giving the cloud company two months to formally propose an alternative, effectively bringing the process back to the beginning again.

As previously noted, the compensation Microsoft paid out didn't go to the large cloud providers, despite AWS being a member of CISPE.

This is entirely logical. With an even bigger market share for cloud than Microsoft, AWS hardly needed recompense. But it raises the question of whether hyperscalers even should

be involved and have a vote in matters of competition.

DCD spoke to several companies in the industry - and members of CISPE - about this very topic.

CISPE itself actually restructured its organization to become a "fully European governance structure," which meant that only European companies could hold board positions and thus vote on decisions.

That restructuring occurred in February of this year, meaning that up till that point,

members like AWS did have a vote on decisions.

Hyve Managed Hosting, an IT Services provider and member of CISPE, director Jake Madders says to DCD: "We're trying to be very fair and open, but I mean, they're not voting members. They're not going to govern what's happening."

"But they are open to listening and being a part of these meetings, and at the end of the day, we want them to help and understand our problem. It's human beings listening to human beings."

This was reiterated, though with caveats, by Vultr's Kevin Cochrane, CMO of Vultr, a privately held cloud company and also a member of CISPE.

"It's a challenging question," agrees Cochrane. "On the one hand, we need to gather everyone together to start a community-led process, but we need to make sure the waters are not poisoned and thus maybe should exclude those market-dominating players."

"On the other hand, there are good people who work in those companies, and they are smart people. You could probably talk to someone from Microsoft; an individual may have perspectives that are very much in line with what we're talking about right now, and anytime you exclude a voice from the table, you're limiting the conversation and potentially having a worse outcome," he says.

Overall, Cochrane says he is "optimistic" and is ultimately a "big



proponent of bringing everyone together."

Finger-pointing and backstabbing

In October 2024 - around four months after CISPE and Microsoft settled - a new group entered the scene: the Open Cloud Coalition (OCC), a UK group.

The OCC describes itself as an alliance of leading cloud providers and users that aims to improve competition, transparency, and resilience within the cloud industry, and launched with ten members: Centerprise International, Civo, Gigas, Google Cloud, and domestic companies such as ControlPlane, DTP Group, Prolinx, Pulsant, Clairor, and Room 101.

At the time of writing, the OCC's website lists 21 members. Google remains the only one of the "Big Three" to join, while AWS and Microsoft remain members of CISPE, of which Google is not a member.

The OCC officially launched on October 29. The day before - October 28 - Microsoft released a blog post penned by CVP and deputy general counsel Rima Alaily said that the OCC was "an astroturf group" and that it was designed to "discredit Microsoft with competition authorities, and policymakers and mislead the public," and that Google has attempted to "obfuscate its involvement" in OCC.

Alaily suggested that Google recruited European cloud providers to serve as the public face of the organization, while Google will "present itself as a backseat

member rather than its leader. "It remains to be seen what Google offered smaller companies to join, either in terms of cash or discounts," Alaily added.

OCC's spokesperson Nicky Stewart vehemently denied these claims to *DCD*.

"Conversations about forming a coalition in the UK had been going on very loosely and informally for a couple of years," Stewart says, when asked if the group was a response to the CISPE settlement.

"Informally because when it came to it, it was a question of 'put your head over the parapet or your hand in your pocket,' and those in discussions would get stunningly silent," she adds.

"Over the parapet because most of the smaller cloud providers actually have commercial relationships with the hyperscalers. And in terms of the financial side of things, smaller companies are smaller companies, and they don't have unlimited resources to fund this kind of thing."

On the topic of whether Google was paying or giving companies incentives to join the OCC, Stewart says: "That is absolute rubbish."

Beyond Stewart's assertions that Microsoft's framing of the OCC's founding are unfounded, the key description of the group as a "shadow campaign" remains somewhat laughable.

Shadow campaign suggests secret movements, but with Google's name

listed front and center - literally in the first two sentences of the launch's press release - there is clearly no attempt to obfuscate their involvement.

At the time of Alaily's blog post, a Google spokesperson responded: "We've been very public about our concerns with Microsoft's cloud licensing. We and many others believe that Microsoft's anticompetitive practices lock in customers and create negative downstream effects that impact cybersecurity, innovation, and choice. You can read more in our many blog posts on these issues," and included links to said blogs.

Google additionally publicly filed a complaint, similar to that withdrawn by CISPE, with the European Commission in September 2024, though any progress on that effort is unknown.

The OCC has, thus far, focused on the ongoing CMA's "Cloud services market investigation" in the UK, which is evaluating the state of the sector, and whether it is currently fair.

"One of our first acts was to jump in a put a response in to the CMA," Stewart tells *DCD*. "There's a whole load of issues here, and can the OCC boil the ocean? Well, no, it can't. It's got to work out where it's best placed and where it can make the biggest difference."

The CMA investigation was launched in October 2023 following a recommendation from the UK's communications regulator, Ofcom.

Ofcom found at that time that AWS and Microsoft had a combined share of 70 to 80 percent of Britain's public cloud market in 2022, while Google was the closest competitor with five to 10 percent.

Since the launch of that investigation, all three US hyperscalers removed egress fees for those looking to fully leave the cloud in an attempt to improve competition.

After more than a year of investigation, and extending the deadline of its final outcome to August of this year, the CMA released provisional findings in January 2025.

That summary document noted AWS and Microsoft's dominant market share, their ability to invest more in data centers and servers, lower ongoing costs due to economies of scale, and dedicated a large section to Microsoft's software licensing practices, which it said had the potential to harm its rival.

Both AWS and Microsoft were recommended for further investigation to see if they should receive a "strategic market status" designation. Should they be labelled as such, they could face conduct and reporting requirements.

DCD reached out to AWS for an interview for this article, but the company said it was unable to provide an interviewee.

But, as is to be expected, both AWS and Microsoft hit back at these findings.

Microsoft complained it was being singled out, and that the CMA was using "hypothetical scenarios" not evidence for its decision. The company wrote in a response: "While AWS clearly believes it is entitled to license all of Microsoft's software for its own benefit and on favorable terms (even though AWS provides none of its own software to Microsoft or anyone else), it has not, to our knowledge, complained about an inability to compete effectively. Nor could it."

The response further draws attention to Google Cloud's successful growth, arguing that "with results like that, it bears reflecting whether the CMA must intervene to enable Google's growth in the UK market by softening competition from its competitors."

AWS disputed that its own behaviour is anticompetitive, but agreed with complaints about Microsoft's software

licensing.

Google, meanwhile, "strongly agree[s] with the CMA's finding that Microsoft's software licensing practices are giving rise to an adverse effect on competition," and said that it "broadly support[s] the package of remedies that the CMA recommends," though disagrees with proposed remedies on egress fees.

Who is really hurting?

Overall, it can't help but seem like the whole process has lost sight of what it is really about. With each response submission from the hyperscalers, one recalls the now popularized "Spider-Man pointing" meme; each hyperscaler one of three separate but identical 'Spider-Men' pointing at each other as though to shift the blame.

The real victims of a cloud market lacking in competition are not AWS and Google. In the most recent quarter, Google's cloud business pulled in \$12.3 billion globally. AWS raked in \$29.3bn.

Particularly in the UK, a good place to look to understand the gravity of the problem is in government procurement.

It is hard to actually quantify how much the UK government has spent with US hyperscalers over the years, due to the sheer quantity of contracts and extensions, but at this point in the past 10 years, it is certain in the billions.

Hyve Managed Hosting's Madders described the spend as "absolutely bonkers."

He says: "Why would the UK Government be spending billions on hosting services provided by AWS, when we have companies like Hyve in the industry and the market that are based in the UK?"

Madders later added: "We are trying to create a regulation where they need to at least give us a chance to compete, rather than just steering straight to the US hyperscalers. Even if we had one percent of the hosting contracts, it would make a huge difference to us, and very little difference to them."

OCC's Nicky Stewart is similarly frustrated by the reality, warning DCD that she would "go slightly on a hobby horse here," in particular when discussing the uncomfortable reality that just last year, the CMA itself doubled spend on AWS.

"The UK government entered into

MoUs with the big cloud providers, essentially asking for preferential pricing. The CMA has looked into this, specifically at preferential pricing that increased as volume increased.

Stewart continues: "The CMA itself had already gone into a preferential pricing deal that came to an end, and they renewed it for a further three years. Even though it was being investigated by the CMA, I suspect they couldn't afford to not increase it. The contracts were subsequently published because of transparency rules, and most of them had doubled.

"At the end of the next three years, what's going to happen then? In all probability, unless money trees suddenly sprout up, they will need to be renewed again."

To lay things out clearly, the people impacted by a government's potentially wasteful spending are the taxpayers of that country. The people who benefit from the current system are the tech oligarchs holding the contracts.

The UK's government procurement for cloud services is done through the G-Cloud framework, though according to Hyve's Madders, it is mostly ineffective.

"We've been on G-Cloud for many years. It's very hard to win any business on there. It hasn't really worked as a marketplace. They don't really use it to find you as a supplier, or at least, that's what our experience has been," he says, though he adds that "We probably haven't been leveraging it as well as we should be."

"You can find tenders in there, and there are ways of seeing what's available, but it's very difficult to compete - especially on paper. There are questions like 'What's your turnover?' and when that is versus Amazon's turnover - if they are using that as a gauge to choose you - forget it."

The ultimate findings of the CMA are due for publication in August 2025, though even if it finds issues with business practices, the process of actually making marked change will be slow and painful.

We can only hope that these investigations will begin the process of working towards a fairer market, one where "sharing" extends to all, not just the elite few. ■



Grundfos Data Center Solutions

Efficient cooling that makes an impact

Optimised water solutions for efficient data flow

What if you could maximise your data center's performance while minimising the environmental impact?

Innovative Grundfos Data Center solutions make it possible. From air and liquid cooling systems to water treatment and reuse, our solutions and services give you the best chance to achieve your sustainability targets. With a 75-year history of sustainability, quality and innovation, you can rely on Grundfos to keep your servers cool, allowing you to enjoy a hassle-free workday with no worries.

Experience Grundfos Data Center solutions at
grundfos.com/datacenters

GRUNDFOS 

Possibility in every drop

Making Connections: The pursuit of chiplet interconnect standardization



Charlotte Trueman
Compute, Storage,
and Networking Editor

A closer look at Universal Chiplet Interconnect Express technology and what's next for this vital piece of the chiplet puzzle

"It may prove to be more economical to build large systems out of smaller functions, which are separately packaged and interconnected," Gordon Moore wrote in his famous article, *Cramming more components onto integrated circuits*, describing the concept that would come to be known as chiplets.

Moore's article was published in 1965. The concept of chiplets is not a new one, even if the term used to describe the technology is a more recent adoption.

On a basic level, chiplets are tiny integrated circuits with specialized functions, which can be combined to make larger integrated circuits that are then packaged and sold as a single component.

Unlike a system-on-a-chip (SoC), which is monolithic and integrates all its system components onto a single silicon die, chiplets are modular, comprising smaller interconnected dies that can be easily scaled through the addition or removal of chiplets. As a result, they provide a number of benefits over monolithic dies, including greater flexibility, faster time to market, lower costs, and improved yields, and have enjoyed a resurgence in popularity in recent years.

However, while chiplets might allow for a more efficient use of resources and improved performance, there is still one major disadvantage: their complex interconnect design. While companies had previously tried to overcome this challenge, true interoperability had been largely lacking.

Enter Universal Chiplet Interconnect Express (UCIe), set up by a consortium of chip industry giants in March 2022 to establish a standardized open specification for interconnects between chiplets.

The primary drivers for the standardization were the need for performance, power efficiency, and the ability to handle different process nodes, says Debendra Das Sharma, Intel senior fellow and chairman of the UCIe consortium since its inception.

Initially founded by 10 members, including AMD, Arm, Microsoft, Qualcomm, and TSMC, the consortium now has 12 promoter companies and lists more than 84 contributor members on its website.

"If you look into the industry landscape, everyone is doing chiplets," Das Sharma says. "If you look into server offerings, GPU offerings, CPU offerings, you name it, everyone is doing chiplets and putting together multiple chiplets and packaging them and interconnecting them in some way."

Das Sharma notes that there are a number of key drivers of this trend, the primary one being that chip manufacturers are starting to hit the limit of how much they can grow one die. There's also the ever-increasing need for performance and power efficiency."

"The need is such that you need the bigger dies, you need more functionality coming out into the packet," Das Sharma explains. "And the way to do that is by getting multiple chiplets and connecting them in some way to make it look like one big chip."

While chiplets might allow for a more efficient use of resources and improved performance, their complex interconnect design is still one major disadvantage

Another reason chiplets have grown in popularity is that they provide companies with the ability to create bespoke solutions.

"You can do mix and match," Das Sharma says. "For example, I can have an accelerator of a given cloud service provider in my offering, but someone else might want a different type of accelerator and still want it to look like it's a heterogeneous computing package."

"You can have different types of accelerators, you can have a different number of cores, different amounts of memory, all serving different people for different usages. And chiplets allow you to do that. You can put two of these, three of the other type, five of something else, you can do the mix and match and get

your own solution."

Das Sharma also explains that different process nodes, used in the production of chips, are better at different things. Memory, for example, is best produced on its own process node, separate from compute, while IO can be produced more effectively on less advanced nodes.

"There are different flavors of process technology for different types of usages," he says. "There's no reason you can't put everything on one die, but it's not going to do very well. So, that's the other reason people do chiplets."

The UCIe 1.0 specification was released in March 2022 and defined the physical layer, protocol stack, software model, and compliance test procedures. The UCIe 1.1 specification followed in August 2023 and included architectural specification enhancements, simultaneous multiprotocol support, and new bump maps in an effort to lower packaging costs.

A year later, in August 2024, the UCIe 2.0 specification was released, bringing with it support for 3D packaging, improved system-level solutions, optimized package designs, and full backward compatibility with UCIe 1.0.

Dr. Ian Cutress, chief analyst at More Than Moore, says that, at present, there's a lot of money going into chiplet-related technologies, especially through startups. However, he cautions that while they can provide great advantages, chiplets aren't perhaps quite the silver bullet the industry is searching for, and UCIe is actually only one part of a much larger puzzle.

"What we are talking about is incredibly complex and more expensive than older technologies that may already exist but may not be universal standards," Cutress says.

"Advanced packaging is expensive, leading-edge chips are expensive, putting them together is expensive, validating the whole set is expensive. The whole concept of chiplets, in and of itself, is difficult, because if you've got a design made up of chiplets, which one is the control? Which one's managing the power? How do you cool it? What's the right thermal environment? Where are the hot spots? What's the validation? These are all important questions, but UCIe only attacks the one thing, and that's how the

chips communicate with each other? All the rest of it still needs to be solved by whoever's building the chip."

Optical illusion

While UCIe was established in 2022, it's only in more recent years that hardware vendors have started to release technology based on the specifications, largely because vendors such as Synopsys or Cadence have started to provide IP offerings.

The Synopsys product, announced in September 2024, consists of a complete UCIe IP solution, operating at up to 40Gbps, faster than the industry standard.

"Our active contribution to the UCIe consortium has enabled us to deliver a robust UCIe solution that helps our customers successfully develop and optimize their multi-die designs for high-performance AI computing systems," Michael Posner, VP of IP product management at Synopsys, said at the time.

Das Sharma says that one of the tenets of the UCIe standard is that it should leverage existing packaging technologies. Because different companies have their own bump technologies, the UCIe standard accommodates these variations, with the focus being on defining the interconnect and ensuring compatibility with existing manufacturing processes.

He also described the consortium as being "forward-looking" when it introduced the specification in 2022 with a statement proclaiming that optical and co-packaged optics would be the means by which people will communicate, because it offers a lot of bandwidth within a very tight form factor.

"If you look at the amount of bandwidth we offer as UCIe, it's a lot," Das Sharma says. "Our power numbers on UCIe are fairly low, so it's a fairly good match in the sense that we are going to define things on the UCIe interface that means you can bring in your optical technology, and we are not going to constrain you about what kind of optical technology... we just want to define a common interface on UCIe that will work with any optical technology."

Ayar Labs is one such company that has already developed optical UCIe interconnect technology.

"UCIe only attacks the one thing, and that's how the chips communicate with each other? All the rest of it still needs to be solved by whoever's building the chip"

>> Dr. Ian Cutress

Unveiled at 2025's Optical Fiber Communication Conference (OFC) in San Francisco, the company describes its offering as the world's first UCIe optical interconnect chiplet.

Combining silicon photonics with CMOS manufacturing processes to support the use of optical interconnects in a chiplet form factor within multi-chip packages, the chiplet is powered by Ayar Labs' 16-wavelength SuperNova light source and is capable of achieving 8Tbps bandwidth.

The company says its compatibility with the UCIe standard helps to create a more accessible, cost-effective ecosystem, streamlining the adoption of advanced optical technologies necessary to scale AI workloads while overcoming the limitations of traditional copper interconnects.

Aside from the excitement of getting to work alongside the members of the consortium, Terry Thorn, vice president of commercial operations at Ayar Labs, says that the company wanted to develop UCIe-based technology because it was a "very natural growth path for us."

"The fact that a lot of companies were endorsing and adopting UCIe... it just made a lot of sense for us to go that direction. So that's what we did."

Ayar Labs taped out its UCIe optical interconnect chiplet in 2024, with Thorn saying that, as a result of being part of the consortium, the company has been able to work with a number of firms that have also adopted the technology in both the pre and post-silicon phases of its chip development.

In the pre-silicon phase, he said Ayar Labs was able to line up test bench

structures based on UCIe specifications to see that its chips would work together when they came back from the fabs, while post-silicon – the phase the company is in now – it has paired the chips up and see that they are connecting how they're supposed to.

"That's actually not simple to do if you're not working from a standard," he says.

"On the electric interface side, if you're going to build a co-packaged optics or you're going to build a fully optical system, there are a lot of things that have to be considered in order to do that, things like package and test.

"How do you architect and design the light source into the system? We use an external light source, which allows us to do things on a separate part of the rack, but all of that has to be considered as you start to build out fully optically enabled AI systems."

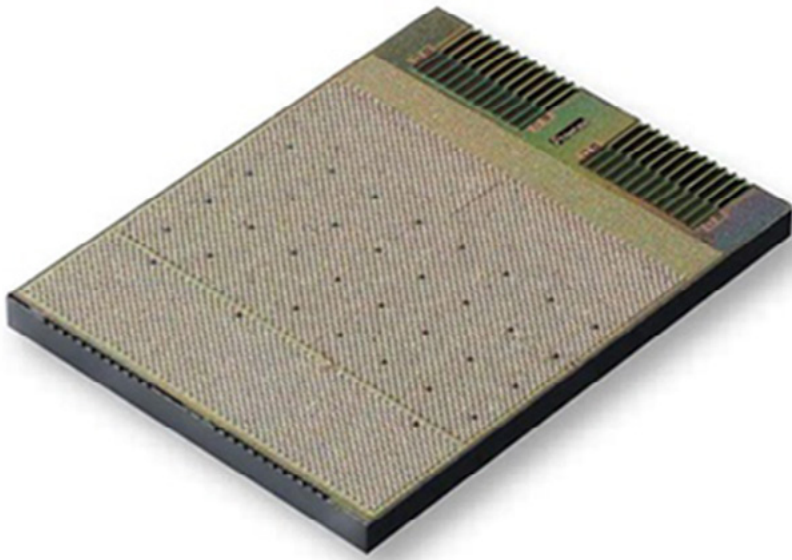
Thorn says that, as a standards-based approach, UCIe brings great benefits.

He continues: "When you look at the adoption of UCIe across the companies that have joined the consortium and are endorsing it, it really just eases the adoption from a chiplet standpoint, in a really significant way, because each customer may have some level of chiplet performance or chiplet customization they want, but you're still going back to that standard interface on the electrical side, so that that certainly helps you in addressing their needs."

Despite a number of UCIe-focused announcements made in the wake of OFC – Lightmatter also launched photonic interconnects at OCP that "use a standard interoperable UCIe die-to-die interface to facilitate scalable chiplet-based architectures" – Cutress says that the move towards optical interconnects isn't necessarily going to be a big driver of UCIe technology in the immediate future.

"Optics on its own doesn't amplify UCIe in any way, however, UCIe can amplify optics medium to long term," he says.

"The reason is that the optics industry, especially co-packaged optics, is still fairly nascent. I have a few clients in this space, and you know they're all saying 2027-2029 is more the time frame they're targeting, but for non-UCIe optics.



"What UCIE brings to the table is that with most optics solutions, whether you're going chip to chip or chip to memory, you have to connect the ASIC, your AI chip, to an optics engine, and one of the ways to connect to that optics engine is through UCIE."

Cutress says that most optics deployments that have been unveiled already, such as Nvidia's switches or Marvell's data center interconnect modules, are not using UCIE, but are instead relying on proprietary protocols.

"UCIE does help standardize that interface as we start to have more players in the optic space," he adds. "There are companies designing the optics connections and other companies designing the AI ASIC, and at some point, they have to decide what connections they're going to use if they plan to work together."

"If they both support UCIE, then it's a plus. But optics is still a bit far out, and it's not necessarily a driver of UCIE, although UCIE could help the medium to long-term adoption of optics in the industry."

The next spec

Looking to the future, Thorn says that Ayar Labs' first effort with UCIE, what is referred to as the standard package, will be its immediate focus, with an advanced package offering targeted for the following years.

"I would say in 2025 and going into

"UCIE's standards-based approach allows us to be as efficient as possible on how we and how our compute customers are designing."

>> *Terry Thorn,
Ayar Labs*

2026 we're really focused on the standard package approach and what we can do with it, and what our partners can get done with it, both pre-silicon and post-silicon," he says. "I would say the coming year, year plus, our roadmap starts to look at 'how would you implement and how do you position the value add of advanced package UCIE?'"

Meanwhile, Das Sharma, who teases that fresh UCIE news might be on the horizon – "the journey continues and we will have something exciting to be talking about pretty soon" – says the regular cadence of releases for new UCIE specifications has been in response to demand.

"There is so much pent-up demand, but in general, we always try to do a spec release based on real demand,"

he explains. "And invariably, what happens is we have more requests than we can handle. So, it's always a matter of prioritizing things. Something goes out now, something else will go out a little later, and so that's what drives the cadence. Right now, we have no shortage of ideas about what to do and what to pursue, and we all know, in this business, there is no dull moment."

However, Cutress says that while he expects to see a lot more demos regarding the technology's interoperability in the coming year, especially from startups and small to mid-sized companies, he's unsure when companies will begin to announce UCIE has been built into chips being used in actual products.

While companies like Nvidia, AMD, and Intel have all shown chiplets and prototypes that demonstrated the practical application of UCIE, Cutress says the likes of Broadcom or Marvell, which have ASIC design teams building chips for hyperscalers, never comment on their customers or talk about what technologies are used inside those chips.

"UCIE, in itself, isn't designed to be a marketing body," he says. "Would the consortium like to talk about partners who have delivered UCIE solutions? Yes. Will those partners want to tell their competitors they've enabled this, that, or the other? No."

"It's quite possible there are actually UCIE technologies, maybe in some embedded use case, or some networking appliance that's actually in a data center somewhere in the Middle East, but we don't know."

Without wanting to speculate on why companies do certain things, Thorn says the lack of transparency around the deployment of this technology likely relates to the fact that it's still relatively early days for UCIE.

"They want to make sure their designs are performant to the degree that they want them to be, and that they can do it in volume before they tend to say anything public about it. That's my suspicion," he says.

"I don't know, it's up to them, but [Ayar Labs] takes comfort in the fact that the majority of the conversations we have are rallying around the UCIE interface and what's happening in that space." ■

COMMScope®

Think Fiber

Think CommScope



Supporting applications:

- Artificial intelligence (AI/ML)
- Cloud computing
- Augmented reality (AR)
- Industry 4.0
- 5G cellular networks
- Data Center Infrastructure Management (DCIM)

Accelerate Your Data Center Fiber Connectivity for AI

Cabling considerations can help save cost, power and installation time:



Speed of
Deployment



Sustainable and
future-proof



Global reach,
capacity and scale

Scan below to learn more or visit
commscope.com/insights/unlocking-the-future-of-ai-networks



Building Stargate



Sebastian Moss
Editor-in-Chief

OpenAI's director of physical infrastructure on gigawatt clusters, inferencing sites, and tariffs



The rapid buildout of the largest data center project in history is hard to keep up with.

In the course of conversations over several months with OpenAI's director of physical infrastructure, Keith Heyde, news about the half-a-trillion-dollar Stargate project just kept on dominating the headlines, forcing new lines of inquiry.

"So much changes every day," Heyde tells *DCD* in an interview crammed between project calls. "The top-level goals still remain the same. I think the execution strategy on how to get to those goals is what is dynamic. Which partners we're leaning into more, which ones we're trying to make sure we're able to couple with tightly, is constantly in flux."

Stargate initially began as a thought experiment. The current view on training multimodal large language models (MLLMs) is that the more compute you put in (along with ever more data), the better the model. To keep up its breakneck pace in software, therefore, OpenAI needed a similarly aggressive approach to hardware.

The result of this thought process was to conceptualize an extremely large data center, akin in capacity to the entire Virginia data center market, currently the world's largest collection of data centers.

OpenAI had initially hoped to get Microsoft, at the time its exclusive cloud provider and biggest backer, on board with the \$100bn 5GW mega data center. But talks fell apart in 2024, with Microsoft increasingly cautious about how much it was spending on a growing rival in an unproven sector, and OpenAI similarly growing frustrated with the hyperscaler's comparatively slow pace.

The election of Donald Trump proved a new opportunity for Stargate to be reborn. Emboldened by a president promising to cut red tape, and keen to stay ahead of rival - and presidential confidante - Elon Musk, OpenAI announced that Stargate would be a \$500 billion project to build a number of large data center campuses across the US over four years.

The end result is a Stargate that is markedly different from the original single-facility vision. Its breadth is far broader, its scope more ambitious - but, for now, the individual data centers are smaller. And yet, they could still be larger than any currently in the market.

"The sweet spot for a training site is between 1GW and 2GW. Above 2GW, things start getting a little silly because you have back-end network distances that are a little bit too long"

At the same time, much about Stargate remains ill-defined. This is partially intentional, as the small and scrappy team looks to develop the strategy in the years to come. But it's also due to the speed of announcements, some of which are not yet backed with complete funds or contract details, but are tied to political moments - be it Trump's inauguration or his visit to the Middle East.

Much speculation has been given to how much money Stargate actually has. Soon after it was announced, Sam Altman told colleagues that OpenAI and SoftBank would both invest \$19 billion into the venture, *The Information* reported at the time. Oracle and Abu Dhabi's MGX are believed to be chipping in a combined \$7bn.

SoftBank is also in talks with debt markets about further funds, and has itself invested in OpenAI (giving it enough cash to invest in Stargate).

Individual Stargate projects are themselves an intricate web of investments. The furthest along site in

"We're looking at clusters positioned, probably more at a country level, that can coordinate across as long as you have enough bandwidth"

Abilene, Texas, has seen developer Crusoe secure \$15 billion in debt and equity, with the funds managed by Blue Owl's Real Assets platform along with Primary Digital Infrastructure. Most of that funding came from JPMorgan through two loans totaling \$9.6bn.

Oracle is currently in talks to spend as much as \$40bn on 400,000 Nvidia GB200 GPUs, per the *Financial Times*, which it would then lease to OpenAI.

Heyde, who joined last November after four years at Meta, is less concerned about the nitty gritty of the accounting. "There's sort of a money flywheel that I don't bother with," he says. "That's a related team. I try to assume that the money flywheel will provide what I need."

"I think about the power/land flywheel, which is the lowest step. Then I think about the shell flywheel. And then I think about the IT flywheel."

Abilene is set to be the first of the Stargates in the US, with OpenAI essentially piggybacking on an existing megaproject. The site is, like everything else involved with Stargate, a complex matryoshka doll of interlinked interests.

At the base is the Lancium Clean Campus, unsurprisingly owned by Blackstone-backed Lancium, a crypto miner-turned AI data center developer. The company then leased its site to Crusoe Energy, another crypto pivoter, which then began work on data centers that would be leased to Oracle.

Back in 2024, the initial plan was for Crusoe to build 200MW for Oracle. The cloud provider, which was falling ever further behind the hyperscalers in the pre-genAI era but has since spent billions on large projects to catch up, was then in talks with Musk to use it for his AI venture, xAI. When he pulled out in favor of building his own site in Memphis, OpenAI swooped in.

At the time, due to the exclusivity contract between OpenAI and Microsoft, the deal had one added layer of complexity - Oracle would lease the 200MW to Microsoft, which would then provide it to OpenAI. With Stargate, *DCD* understands that the extra step has been removed.

Also with Stargate, the ambitions have been ramped up: Crusoe is now planning on building out the entire 1.2GW capacity of the Lancium campus. However, in a

nod to the speed and chaos of the early days of Stargate, *DCD* understands that Crusoe executives at the PTC conference in Hawaii this January were just as surprised to learn about the project as the rest of the world.

"The data halls at Abilene are like 30MW," Heyde reveals. "But then the total site for us could go above a gig if we expand it as far as it could go."

Construction of the first phase, featuring two buildings and more than 200MW, began in June 2024 and is expected to be energized in the first half of 2025.

The second phase, consisting of six more buildings and another gigawatt of capacity, began construction in March 2025 and is expected to be energized in mid-2026. One substation has been built, with another on the way - power is also provided by on-site gas generators. Oracle has agreed to lease the site for 15 years.

"It's us and Oracle," Heyde explains. "And then you have your tech stuff that's going to Foxconn, and your shell stuff that's going to Crusoe. And then under Crusoe, there's [construction contractor] DPR, and there's all the subcontractors like Rosendin on the site."

With the first two buildings about to be ready, Heyde has been making regular trips to the campus. "My team was there yesterday, I was there a couple of weeks ago," he tells us. "It's more about the rack readiness than necessarily base readiness. As we get these GB200 racks off the line, that's the name of the game right now."

Roughly three hours' drive from Abilene, another OpenAI project is taking shape. Core Scientific is refurbishing its Bitcoin mining data center in the location just north of Dallas, instead pitching it as an AI data center hotspot.

That site is then leased to CoreWeave, which is working with OpenAI. "Our interactions are with CoreWeave," Heyde says. OpenAI plans to spend \$11.9-15.9 billion on services from the neocloud, although it is unclear if all of it will go through Stargate.

TD Cowen analysts say it has also signed a five year deal with Oracle for up to 5GW, while *Reuters* reports it is in talks with Google Cloud for further capacity. Some of those contracts may themselves then be farmed out to CoreWeave data centers.

With OpenAI planting its flag in Abilene as the first of its Stargates, other developers have flocked to the region to acquire land and attempt to source power, in the hope of experiencing the network effect seen in pre-genAI data center booms.

"I think that anytime you have some clustering of a certain skill set in an area, you do create a center of excellence," Heyde says, but admits he's uncertain if the same network effects will be found.

"From a tech advantage standpoint, we don't necessarily think that that is super advantageous. Because, ultimately, your back-end network distance is like the size of the cluster. So there's an argument of 'it doesn't matter if you're 15 or you're 600 [kilometers] away.'"

This network limit weighs heavily on Heyde's mind. "The sweet spot for a training site is between 1GW and 2GW," he says. "Above 2GW, things start getting a little silly because you have back-end network distances that are a little bit too long. They're too far to do what you need to do."

At that point, some racks are so far away from other racks that the data center might as well be spread out.

"Even if you have the perfect flat piece of land, the trickiest part of design is your networking distance," he says. "If you're gonna do a single story data center, at a certain level, the networking distance becomes too big to have a massive, massive cluster there."

One could move to multi-story to pack the racks closer, of course. "The challenge with multi-story is that a) There's an operation pain in the ass that comes with it; but b) there's also a little bit of logistics variability on when we start getting into these really big racks in the future," he says. "If we're talking about racks that are, like, 6,000-7,000 pounds, you start going into real funky building designs."

While a 5GW mega data center campus may not make sense, several 1-2GW campuses can still work together. "What we found is that the distance constraints that were originally considered about a year ago have somewhat relaxed, based off approaches we're using for large-scale, multi-campus training.

"So we're not necessarily looking at tight clusters associated with latency. Rather, we're looking at clusters positioned, probably more at a country level, that can coordinate across as long as

you have enough bandwidth."

Google has already publicly revealed that it trained Gemini 1 Ultra across multiple data center clusters, while OpenAI is believed to be developing synchronous and asynchronous training methods for its future models.

While some hyperscalers have dug direct fiber connections between distinct campuses, Heyde says the company isn't currently considering it. "We've had some internal conversations; I've got to make sure we get some data center campuses first."

Including Texas, OpenAI is considering developing as many as ten projects in the US, with the company saying in February it was looking at 16 states. "Candidly, I think we're a bit higher than that right now," Heyde says.

Outside of the US, the company is in the very early stages of 'Stargates for Countries,' a pledge to bring similar infrastructure to other nations, as long as they invest. The effort is in the very early stages, with the much-touted 5GW US-UAE AI campus with G42, publicly announced during Trump's Middle East tour in May, still in the Memorandum of Understanding stage.

When we asked Heyde to rattle off what he's looking for with those sites during a keynote discussion at this year's DCD>New York event, he told us: "As we're doing these multi-campus training runs, access to large bandwidth fiber is pretty critical."

As for power, OpenAI is evaluating both "the substation readiness and the associated data center side, but then also the total power of that location," Heyde says.

He continues: "I think one of the interesting things we've seen from the RFP process is that people frame their power availability very differently, depending on whether they're talking about it from the perspective of a utility, a power provider, a data center operator, or a data center builder. The time horizon that people pin to that power varies quite a bit based on assumptions baked in."

Then there's land, but "the cool thing for us is that our physical footprint is a bit smaller based on the densification," he says. "And then the last thing that I would say is really critical is that we want to go places where people want us to be there."

When it does go to those places, it



could do it via partners - be it someone like Oracle, Crusoe, CoreWeave, or Core Scientific. Or it could build its own sites completely.

"Stargate and OpenAI both have a self-build interest in mind, yeah," Heyde says. "It will definitely be a novel design, comparatively. The Abilene design was really anchored around something else, and then it got repurposed into the GB200 approach. We're working on it, and a self-build approach will be anchored around how we would go after a large training regime in the future. There are some exciting things you can do from a redundancy and risk tolerance perspective that you probably wouldn't do at a cloud site."

In a different conversation, he tells us: "We're designing around, effectively, a three-nines perspective rather than a five-nines perspective," that is building to an expectation of 99.9 percent uptime rather than 99.999 percent. "This gives us a lot of both design flexibility, but also certain trade-offs we can make in the power stack as well, as in the site selection trade-offs as well."

"Because our design is relatively unique, and because it is centered around these three-nine training workloads, we have a little bit of flexibility to play in different ways with different players that I think you might not see in more of an established reference design that gets punched out."

Here, Heyde is talking about training sites, or as he likes to call them, "research-

"We could do an inference and a consumer use case that's quite a bit smaller. In fact, this is in the market now"

enabling clusters,' principally built for pushing our capabilities on the model side forward." This, he says, is different from the training vs inference binary that was originally seen as the two distinct AI workloads of the future.

"The means and mechanisms by which we advance our models have evolved as the industry has moved in different directions. So at this point, from the reasoning models, we're able to push the capabilities of our software forward using things that are slightly different than classical training," with models always in some form of training and inference.

At the same time, however, Stargate does plan to have more inference-focused compute, a shift from the initial vision of a single mega cluster for training. This inference is alongside an ongoing extensive relationship with Microsoft Azure.

Some of the inference could take the

form of aging training clusters in the years to come. As the bleeding-edge Nvidia GPUs get older, they could shift to servicing user demand, we suggest. "I think that is not a bad approach," Heyde says. "I think there's merit to that."

The company is also looking to build specific inference clusters. "We could do an inference and a consumer use case that's quite a bit smaller [than the 1GW campuses]," he says. "In fact, this is in the market now, but my team is looking at some smaller sites on how we would deploy - I don't want to call them micro sites, they're still pretty big - an inference or customer-facing site rapidly."

He adds: "It is probably still too dynamic to truly say [what size they will be], but I can say that as we project out in the future, that part of our demand definitely increases. That is the customer-facing part of our demand. And so it's not just from increasing customer engagement, but it's also from a more and more workloads being able to move into these more efficient sites - they're not really Edge sites, but they're more like Edge site regimes. You just have more demand that gets channeled there."

That is despite the fact that latency demands are "looser than people think," he says.

We talk the night after OpenAI reveals it is set to buy Io, the company led by Apple design guru Jony Ive, with a plan to develop consumer products. The company has yet to reveal any concrete details about



The first Stargate campus takes shape at Abilene, Texas

the devices, but an announcement video references an always-on, always-listening system to allow ChatGPT to be involved at any moment.

Such a system, should it prove popular, would require even more compute and storage. Heyde isn't involved in the project, but says of the data center side: "The numbers that Sam [Altman, OpenAI CEO] shared with us earlier today, were like 'add an extra zero' type stuff. It seems like I've never had a demand break, including when DeepSeek was really shaking the market."

Another market shock has been the US President's war on the global economy, with sudden and drastic tariffs levied against virtually the entire world (tariffs are paid by the importer, not the exporter). Some of those tariffs have been walked back, some have been ramped up. At time of writing, the current state of tariffs is too fluid to make writing the individual percentages worthwhile.

TD Cowen analysts warned that tariffs could add five-15 percent costs to Stargate data center builds, an enormous cost given the scale of the joint venture. "This is an interesting one," Heyde says of the tariffs.

"A lot of the sourcing decisions that I was making were already North America-based. So the nice thing was that I didn't really see major tariff problems from that, [but] I've seen other projects that we were looking at that had major, major disruptions, though. The volatility is half of the challenge."

For now, development is continuing

unabated. Speed remains the single greatest priority, with OpenAI locked in a race for a prize it believes easily outweighs any momentary cost increases.

"In some ways, that drives us to a little bit of short-termism on our decisions," Heyde admits. "We want to really be unlocking training. I actually think we're okay with a little bit of cost inefficiency - I want to say that delicately - to enable good training moments."

A desire for speed, together with a large wallet, does come with its own challenges, and the company is inundated with development offers, most of which are unsuitable.

"Imagine you're at a restaurant, and there are just so many choices on the menu, and it's time for us to make sure we pick the ten choices we want, not the 100 choices that are at our disposal," Heyde says.

Of those ten choices, "we're trying to think about two cycles out of our data center design, you don't just get one and done," Heyde says.

"It's hard when we're talking about these power densities and it's also hard for me to not look at the lay of the land and be like, 'are racks going to be our form factor in eight years?' The whole rack form factor might change because of the power densities we're talking about."

He says he can see data centers of the future containing what is "effectively liquid, reconfigurable, compute." By this, Heyde doesn't mean immersion cooling systems, but "dynamic liquid conduits that can reconfigure as part of the data center

compute infrastructure." He adds: "This was sci-fi stuff five years ago, but not that far out now on what we're talking about."

Should his own sci-fi effort succeed, Heyde will have been partly responsible for bringing online the largest clusters of compute the world has ever seen, potentially supporting superhuman intelligences - or, at least, simulacra of intelligence.

"The mission is to create AI that solves the world's hardest problems and that, to me, is a very exciting mission. Having spent many years of my life pursuing a PhD and expanding science, to be taking a real shot at making something that can solve the world's hardest problems is pretty exciting," he says.

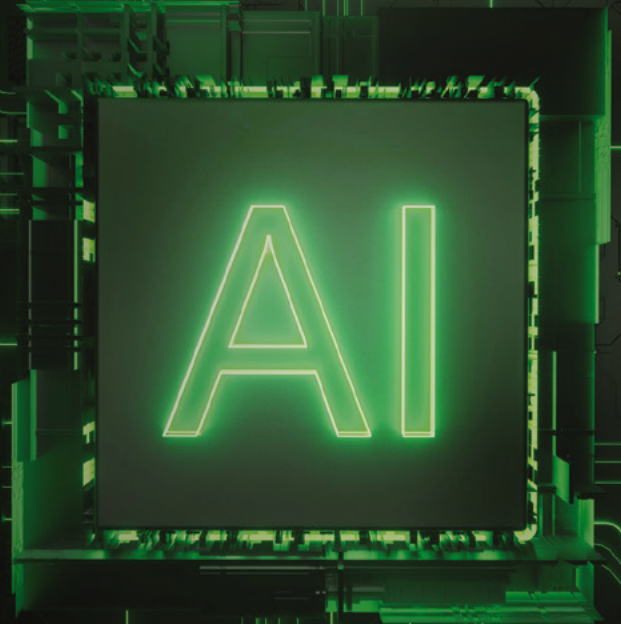
We asked Heyde how he'd use such an AGI if he could put it on a core challenge with Stargate: "What's hard for me to wrap my mind around is the sequencing of what to do to unlock the best approach. It's not so much the sequencing from a demand and supply mapping perspective, but rather which capabilities and market signals I need to unlock to create the best goal state at different training time points.

"In some ways, you unlock something by doing an Abilene, in other ways you unlock something by doing a self-build somewhere. In other ways, you unlock something by working with another partner. Those capabilities and market signals simulated out take you to different end states. Because of the variable space and the opportunity space, it's kind of hard to conceive what the perfect global solution is, even though locally, I can see a lot of positive local solutions." ■

Partnering for AI-Ready Data Centers

Transform your data center into a powerhouse of efficiency with Schneider Electric's AI-Ready solutions. Our end-to-end infrastructure services are designed to adapt to the growing demands of AI workloads, ensuring your operations remain resilient. Leverage advanced power and cooling systems tailored for high-density compute environments that maximize performance while minimizing environmental impact. Partner with us to redefine your data center's energy strategy and take the lead in the AI era.

se.com



Life Is On

Schneider
Electric

An opportunity too good to ignore: Nokia, AI, and data centers

Better known for its mobile networking business, the vendor wants to cash in on the artificial intelligence boom



Paul Lipscombe
Telecoms Editor

Nokia's Cloud Design Center

To some, the Nokia brand evokes memories of mobile phones, in particular, the 3310 of the late 1990s and 2000s with its iconic ringtone and addictive 'snake' game.

Having lost its mobile market dominance after the dawn of the smartphone era, the company has become best known for producing network hardware and software for mobile carriers across the world. It is one of the world's biggest network vendors, competing with Nordic rivals Ericsson and Chinese giant Huawei.

But in 2025, it appears that Nokia wants to be seen as something else too. The vendor wants to become a provider of data center solutions and products.

"Where the growth is happening today is on the data center side," Vinai Sirkay, head of business development for Nokia, told DCD earlier this year.

"Whether it's inside the data center

with switching, or between data centers, both at the IP layer to connect it to the Internet, or just connecting data centers to each other through IP and optical interconnectivity. We are doubling down there."

As technology such as generative AI continues to grow, there is a requirement for more data centers to support the demands placed on these technologies.

For Nokia, the data center opportunity is too good to ignore.

Data center-focused CEO

It's worth noting that Nokia's data center strategy stretches back to 2020. In July of that year, amid the turbulence of the Covid-19 pandemic, the vendor launched a new network hardware range to automate and simplify data center operations.

This year, Nokia appointed Intel data

center executive Justin Hotard as its CEO to replace long-term CEO Pekka Lundmark, who previously stated last year that Nokia sees a "significant opportunity" to expand its presence in the data center market. Hotard's appointment suggests Nokia means business on this front.

Hotard only spent around a year in the role at Intel, where he was responsible for Intel's suite of data center products spanning enterprise and cloud, including its Xeon processor family, graphics processing units (GPUs), and accelerators for servers.

Before that, Hotard spent an eight-and-a-half-year stint at HPE, where he most recently served as executive vice president and general manager of High-Performance Computing, AI, and labs for the US tech firm.

His experience and knowledge of data centers and subsequent appointment at Nokia is not a coincidence.

Infinera acquisition

Hotard joining the company in April is rather timely too, as it came around a month after the company finalized its \$2.3 billion acquisition of networking firm Infinera.

"This acquisition [Infinera] brings a number of significant benefits," Hotard told analysts during his first earnings call at the company in April.

"It gives us the scale to accelerate our product roadmaps and to drive more innovation. It also increases our access to hyperscale customers, which are a key growth driver in both cloud and AI data center investments."

During the Optical Fiber Conference (OFC) held in San Francisco in early April, Nokia wasn't quiet about its acquisition of Infinera, with its advertising on full show everywhere you looked, along with Infinera's stand, which celebrated the union.

It made sense too, given the event focuses on optical networking, and the purpose of the deal was to expand Nokia's presence in the data center interconnect space and in the US.

"Our multi-billion dollar acquisition has given us a big step forward because the Infinera team has done a really good job in building a strong customer footprint with the big web companies and web-scale companies," Manish Gulyani, head of marketing for Nokia's Network Infrastructure business, told DCD during OFC, weeks after the deal was completed.

For the full year of 2024, Infinera posted record revenue with webscalers, noting that its total revenue exposure (direct and indirect) was greater than 50 percent of its full-year revenue.

"That gives a very good footprint for people who are building these big data centers for cloud and for AI workloads," added Gulyani, referring to Infinera's financial performance.

"So that's already there, but we'll build on that our technology. That technology really gives us all the elements we need to really pursue it from an optical perspective, and on the IP side."

According to Gulyani, Nokia plans to invest in driving more data center switching and routing solutions into the data center space, plus increasing investment on the IP side.

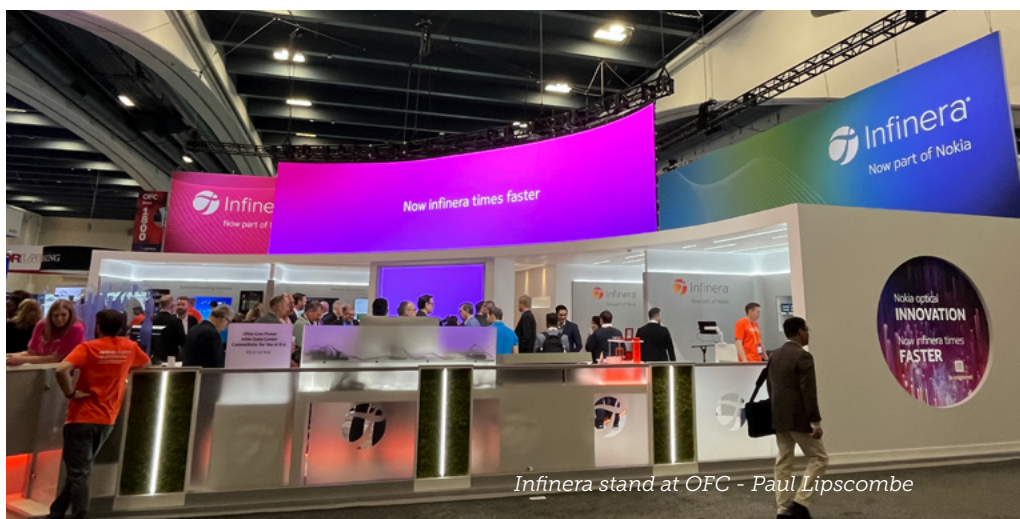
"Where the growth is happening today is on the data center side."

*>> Vinai Sirkay,
Head of business development,
Nokia*

notable switching deals with Microsoft and Apple, with the iPhone maker being one of the first to deploy the switching platform back in 2020.

In the past year, Nokia has struck further data center-focused partnerships. Most recently, it was selected by American Tower-owned CoreSite to provide routing upgrades for CoreSite's data center footprint across the US.

In November 2024, Nokia extended its existing agreement to supply Microsoft Azure with data center routers and



Nokia's data center offering

Nokia is "not in the business of building data centers," Sirkay says, but rather providing the equipment that goes inside them.

It offers networking infrastructure, routers, and switching products as part of its data center suite, and has penned

switches. A month later, Nokia, along with ITSP Kyndryl, announced plans to offer advanced data center networking solutions and services to global enterprises.

Last year, Nokia also launched a data automation platform, which it describes as an "event-driven automation" (EDA) platform.

"Some will buy directly from us like the hyperscalers, while some want us to be part of a solution," added Sirkay.

All roads lead to AI

The growth around Artificial Intelligence (AI) has been a key reason for Nokia's excitement around data center interconnectivity.

AI is demanding that data center infrastructure is in place to support the workloads that are being used and the networks that are carrying these workloads.

Federico Guillen, president of network infrastructure at Nokia, remarked at OFC



that "AI is the lifeblood of the network."

While Subho Mukherjee, vice president & global head of sustainability, Nokia, told DCD earlier this year that "AI has been a once-in-a-generation catalyst for data center boom."

Nokia sees AI as an opportunity to support data centers with its own products, such as switches.

"The way I think about the market in AI, particularly with optical is, if you look at the build-out of the data center, what's happening with AI is it's driving, as we all know, significant new data center build, but it's also driving new connectivity demands between data centers, both whether it's for training or inference or some of the convergence we're seeing with AI reasoning models," Hotard stated during the company's Q1 earnings call.

Networking for AI

As more data centers are needed to keep up with the demands for AI use cases and workloads, networking is essential to support this.

Companies such as GPU giant Nvidia have dominated the discussion around the build-out of these data centers, as the industry focuses on the compute hardware required to deal with the AI boom.

However, as Gulyani points out, the networking aspect of these build-outs is every bit as important and plays into Nokia's expertise.

"Our investment to build infrastructure to house compute for AI is just massive, and as driving not just larger data centers, but more data centers everywhere, because that's kind of driven by both where you can get the infrastructure," says Gulyani.

"AI has been a once-in-a-generation catalyst for data center boom."

*>> Subho Mukherjee,
Vice president & global head of
sustainability, Nokia*



Nokia stand at OFC - Paul Lipscombe

"You're getting more and more distribution of data centers, and we don't spend much time talking about the network piece of it, because people focus on GPUs and the other things you need to build those monster supercomputers. But you need networking inside the data center too."

When asked about why the data center opportunities excite Nokia, Gulyani speaks of the chance to support businesses meet the demands of AI.

"Switching, routing, interconnect. These are the three legs of our solution," he says. "All of them have massive addressable markets with new, innovative solutions required to meet the needs."

"What we had till the last generation was good enough for non-AI workloads, but with AI, it's creating another sort of requirement around higher performance, more resilience, more security."

Beyond its bread and butter

For Nokia, don't expect the data center

focus to usurp its bread and butter offering, which is its radio access network (RAN) products.

The vendor will be hoping that its Infinera deal will boost its opportunities in the US, where it's found life difficult in recent years, notably missing out on lucrative 5G RAN deals with AT&T and Verizon to Ericsson and Samsung Networks, respectively. That said, the company did strike gold with the country's third major operator, T-Mobile, recently. Emerging technologies such

as 6G are also central to the company's future plans, with R&D efforts well underway.

However, in the short term, its data center play will provide an opportunity for Nokia to diversify its revenue streams beyond telecom carriers.

"Nokia has two sides, the service provider business and

now the data center business," explains Gulyani. "The telecom business is strong. We are number one or number two in almost all markets."

"But now there is the cloud networking business, essentially connecting all clouds to each other and eventually to the end users. So that's where we are adding increased focus within Nokia is regarding the connectivity to anybody who plays."

He points out that this could include cloud companies, data center hosting firms, colocation providers, and other companies in the ecosystem.

"There's a lot of players in the ecosystem that are providing connectivity that need connectivity solutions. We want to be working with anybody in that ecosystem that can take advantage of the same solutions," says Gulyani.

RAN might be the bread and butter, but Nokia certainly has an appetite bigger than what the telecoms industry is serving up. The vendor sees the demand for data centers and wants a slice of that growth. ■



**Power
Generation**

ON

We understand the pressures data center and mission-critical operators face – where reliability is non-negotiable and sustainability is no longer optional. That's why we support solutions that help reduce environmental impact without sacrificing performance. For your most critical operations, switch on sustainability. And keep it on.



Scan the QR code
to learn more about
Cummins Data Centers'
sustainability initiatives.

Watt's Next? How can batteries be best utilized in the data center sector?

A deep dive into the many use cases of Battery Energy Storage



Zachary Skidmore
Senior Reporter, Energy
and Sustainability

The Battery Energy Storage (BESS) market is going through a coming-of-age moment, having grown exponentially over recent years. According to Wood Mackenzie, it has seen a 44 percent expansion in 2024, with more than 69GW of new BESS capacity installed globally.

Despite the growth, the role of BESS within data center architecture remains in the nascent stage, with debate raging on how it can be best utilized within the sector. For some, BESS offers a potential clean energy replacement for diesel generators, which remain a crucial backup failsafe for the vast majority of data centers in the event of outages. For others, BESS at scale is seen as a potential primary power source for data centers and a crucial component in changing the perception of data centers from net consumers to net contributors to the grid.

Against this backdrop, data center

operators are beginning to explore the use of BESS as a core component of data center energy architecture, with several interesting test cases already underway.

What is BESS?

Batteries already play an integral role in data center architecture, in the form of uninterruptible power supply (UPS) systems. Most UPSs have an average capacity of 50 to 300kW, providing around 20-30 minutes of backup power in case of sudden outages.

Comparatively, BESS units are, on average, much larger than UPS systems, capable of scaling into the hundreds of megawatts. Lithium-ion batteries are the dominant player, holding around a 90 percent market share in the utility-scale market. They offer an average storage duration of between two to six hours, which has mainly led them to be used in grid balancing roles, especially when tied to intermittent renewable assets.

Despite the market's growth, data center operators have been reluctant to integrate the technology within their architecture. This is due to concerns over short storage capacity, high costs, and fire risks, with lithium-ion being more prone to combustion due to overheating.

However, in recent years, several companies have taken the plunge and announced deployments of BESS at their data center sites, with each example providing an interesting test case on how the future of BESS in the data center could look.

BESS as a backup option

One of the most notable deployments of BESS within the data center sector is at Microsoft's Stackbo data center in Gävleborg, Sweden.

In a first-of-its-kind pilot initiated in 2022, Microsoft partnered with Saft, the battery subsidiary of TotalEnergies, to

deploy four independent containerized lithium-ion BESS units, each with 4.6MWh of storage and a peak output of 3MW at the Stackbo data center.

The deployment's main aim was to test the applicability of replacing diesel backup generators with BESS.

"The question was, can we install a system that is 100 percent capable of replacing the diesel generator, from a technical, safety, and operational point of view, with full compliance to regulations and grid codes," Michael Lippert, director of innovations and solutions for energy at Saft, says.

To achieve this, Microsoft approached the project with an open mind, going against the prevailing consensus that data centers should have at least 48 hours' worth of backup capacity available at all times.

Instead, at Stackbo, the BESS backup units were designed for only 80 minutes of backup capacity. The decision was based on extensive studies of the local grid, with Microsoft concluding that due to its high resilience, the Swedish grid was highly unlikely to face prolonged outages. Therefore, Microsoft and Saft calculated that an 80-minute backup would be sufficient to mitigate the risk of service disruption, with an acceptably low probability of a longer outage occurring.

In approaching the question of backup on a case-by-case basis, Lippert says that the project proved that the requirement for 48 hours of backup is not always a technical necessity. "With backup ranging from two to six hours, you would most likely be able to cover a very high level of reliability," he contends.

According to Lippert, this approach not only replaces diesel generators from a safety and operational standpoint but also enables advanced functionalities like black start capability - an ability to re-energize the system from a complete power-down state without external grid support - and grid-forming operation.

"Technically speaking, that means you can operate the data center on a microgrid," Lippert says. "The battery forms the grid, and then you can add generators or other devices on top."

As a result, at Stackbo, BESS is positioned not only as a backup option but also as an active participant in the energy ecosystem, providing voltage



B-Nest's hyperscale energy storage

"Each location has its constraints... but the core principles remain: replace diesel with batteries, use the battery as a grid-forming resource, and design the system to scale."

>>Michael Lippard

control, frequency regulation, and even supporting greater integration of renewable assets.

For Saft, the Stackbo model is considered extremely replicable globally, with slight tweaks to the duration of the BESS system depending on regional variations in grid codes and renewable access. "What we've done in Sweden can be applied elsewhere," says Lippert. "Each location has its constraints... but the core principles remain: replace diesel with batteries, use the battery as a grid-forming resource, and design the system to scale."

Therefore, use of BESS, especially in locations with resilient grid infrastructure, is only likely to become more widespread if it continues to prove its reliability across multiple deployments. This could shift industry standards and regulatory expectations toward more tailored, data-driven approaches rather than blanket duration requirements.



A Saft BESS system

However, given the recent high-profile mass outages across the Iberian Peninsula, which lasted upwards of 18 hours in some places, the sole reliance on BESS without further redundancy could place operators under significant risk, stifling the repeatability of this model.

Dual value

In contrast to Microsoft's approach, Keppel DC REIT's deployment at two of its Dublin data centers represents a distinctly different model.

Led by GridBeyond, a grid enhancement firm, the project involves two separate installations: a 2 x 2MW/2.2MWh system at Keppel's Citywest data center and a 4MW/6.1MWh system at its Ballycoolin facility.

The BESS system is designed to complement the existing UPS and standby diesel generators. To mitigate fire risk, it sits outside the resilient pack ring, which GridBeyond CEO, Michael Phelan, says allows Keppel to explore interesting energy use cases that previously could have conflicted with customer resilience requirements.

"By placing the battery outside the ring, it can enable both carbon reduction and support for grid flexibility — things customers increasingly value," Phelan says.

Again, the nature of the market made a significant impact on the deployment, with Dublin currently subject to a data center moratorium due to its outsized impact on the grid, consuming up to 20 percent of the country's power output.

A point made clear by Phelan, who states: "[The BESS units] allow data centers to store energy during non-peak hours, reducing reliance on the grid during peak demand. This not only reduces carbon emissions but also contributes to grid stability."

The Keppel deployment aims to

illustrate how BESS can provide dual value to data center operators in locations where the grid is facing challenges, not only enhancing sustainability but also accelerating development timelines for data centers.

To achieve this, Keppel has connected the BESS to an AI-powered energy management platform that enables dynamic demand response, helping to stabilize the grid during peak periods.

"By integrating battery storage, data centers can discharge during peak hours, allowing utilities to allocate energy elsewhere. This flexibility makes it possible to build data centers more quickly while ensuring grid reliability," says Phelan.

Therefore, unlike Stackbo, the batteries do not entirely replace diesel generators but supplement them, creating a hybrid system that is more resilient and efficient, according to Phelan. As a result, the model has the potential to grow in popularity, especially within areas where securing a grid connection is a much more difficult process.

BESS as a primary power option

One company pushing the boundaries of BESS use in data centers is Energy Vault. Earlier this year, the company partnered with RackScale Data Centers (RSDC) to deploy its novel B-Nest system across multiple RSDC campuses in the US.

Unlike at Stackbo and Keppel, Energy Vault's BESS will not be a supplementary or even backup solution, but rather a primary power source for the facilities. Energy Vault aims to achieve this through a vertical, modular design capable of storing up to 1.6GWh per acre. According to the company, this multi-story configuration allows data centers to conserve horizontal space for core computing infrastructure while maintaining energy autonomy.

"The innovation here is in the form factor and how we stack the batteries vertically over multiple stories to maximize energy density and reduce land footprint. That's what's new, and that's what's tailored to the data center world," says Marco Terruzzin, Energy Vault's chief commercial and product officer.

Scheduled for deployment in 2026, the B-Nest system aims to deliver 2GW/20GWh of storage across multiple campuses, which the company claims

"By integrating battery storage, data centers can discharge during peak hours, allowing utilities to allocate energy elsewhere."

>> Michael Phelan

could provide more than ten hours of power at full load operation. Energy Vault argues that by implementing BESS infrastructure up front, RSDC can ensure reliable power from day one, independent of the pace of utility build-outs.

Terruzzin emphasizes that in today's market, energy storage is as much about accelerating project timelines and maximizing land use as it is about decarbonization. Batteries are usually prioritized for connection over other power generation projects, due to their fast-to-connect interconnection and the volatility of most electrical markets, increasing the attractiveness of the technology for utilities as a balancing tool.

This, the company claims, will allow data center operators the ability to secure a grid connection at a much faster rate. The focus on the US market reflects this, with the US grid facing significant constraints amidst the surge of AI-driven demand, which has led to severe delays for data center operators. Reframing the utility of BESS makes particular sense in the US, where anything considered green or low carbon has been thrown to the sidelines.

Therefore, by framing the use of BESS as a way for data centers to achieve their operations in an expedited time frame, Energy Vault hopes to attract greater numbers of operators to trial its system.

However, the deployment, efficacy, and scalability of the model remain to be seen. Concerns include its high capital costs, large space requirements, and questions over grid integration. In addition, it seems like the model may rely on some form of onsite baseload power, which ultimately could be in the form of thermal gas generation, and would harm its sustainability credentials.

Giving back

A company taking a small-scale,

sustainability-minded approach to BESS deployment is Scandinavian Data Centers (SDC), which this year installed a BESS unit at its ScandiDC I, in Eskilstuna, Sweden.

The deployment was predicated on a "responsibility to give back," says Svante Horn, CEO of SDC. "As a company, we view power as a privilege, not a right. Since we were granted a 10MW [grid] connection, set to be ramped up to 15MW, installing batteries allowed us to support and stabilize the grid while making use of a capacity we pay for regardless of utilization."

At present, the BESS isn't being used as a backup solution. However, going forward, the company intends to use the batteries intermittently to prevent the need to run its diesel generators, unless there is an extensive fault in the system.

Therefore, the company envisions BESS as not only a potential backup solution but also as a net positive for the grid, with Horn explaining: "We want our UPS and our batteries to help stabilize the grid... [and] have proper discussions with grid operators about our backup power being auxiliary resources for the grid."

For SDC, the utilization of BESS is part of a broader concept of "data center 3.0, where everything's behind the meter: renewable energy production, battery storage, and heat reuse, creating a fully integrated, flexible energy system."

Looking further afield, Horn sees Europe as leading the charge for BESS deployments within the data center market, especially in countries with high renewable integration, such as Spain. However, the onus is very much on the hyperscalers to take the lead in building the batteries, says Horn. "When the economics start making sense... large sites can have industrial battery deployments and data center deployments at the same time. But we are not there yet."

So Watts next?

Lithium-ion BESS systems have so far made up the bulk of deployments at data centers, as seen with the examples mentioned.

However, while effective as a short-term option for storage, several concerns are notable in their applicability to the sector. Major concerns include the aforementioned fire risk, as well as a heavily constrained supply chain, with

the Chinese market representing more than 70 percent of global production.

As a result, long-duration storage technology (LDES) is increasingly being touted as a potential solution to lithium-ion's shortcomings. Two prominent examples are thermal and flow batteries, whose proponents argue offer greater flexibility in deployment and significantly longer duration of storage, in addition to no risk of combustion.

One LDES company explicitly targeting the data center sector is Exowatt, a US-based firm developing a thermal self-contained battery. The system dubbed P3 uses optical collectors to capture solar energy, stores it in a solid block of material as "sensible heat," and then converts it to electricity via a modular heat engine. Unlike most other batteries, the thermal storage medium does not degrade over time and is highly modular, allowing for significant scalability.

For its CEO, Hannan Happi, the solution offers a much more efficient form of energy storage for data centers due to its simplicity. "There's no fire risk, no degradation, no active cooling — it's a simple system designed for a 30-year service life," he says.

Exowatt is positioning itself as a prime or hybrid power source for the data center market, supporting new builds by avoiding the costly delays associated with waiting for a grid connection.

"We can energize a greenfield site immediately, which often creates the incentive for utilities to follow and build substations."

Exowatt has already begun commercial deployment in the US and is working to scale manufacturing toward 10–100GWh of annual capacity. With no rare earths and a US-based supply chain, Happi believes the company is well-positioned.

"This isn't natural gas or nuclear. There are no major regulatory hurdles — just land, sunlight, and a willingness to rethink how we store energy," Happi states.

Given its reliance on solar power, the solution is more constrained in terms of effective deployment, with Exowatt initially targeting the Sun Belt states of the US and the Middle East due to their high solar potential. Happi admits, "You could run it in the UK, but only during a few summer months — and at much higher cost."

"The innovation here is in the form factor and how we stack the batteries vertically over multiple stories to maximize energy density and reduce land footprint."

>> Marco Terruzzin

In addition, due to the significant order backlog, widespread deployment will require substantial capital and supply chain coordination.

"We're not looking internationally yet — we're very backlogged in the US," says Happi. "But as we expand, we'll need decentralized manufacturing to avoid shipping heavy modules globally."

Despite these constraints, the solution offers an interesting alternative. And while it might not be a universal fix, could be extremely effective in locations with high solar potential.

Go with the flow

Another LDES player to emerge is XL Batteries, an organic flow battery developer. In May, the company partnered with Prometheus Hyperscale to deploy a 333kW demonstration project at one of its data centers by 2027. Following this, Prometheus plans to acquire a 12.5MW/125MWh commercial-scale system in 2028, with another identical system to follow in 2029.

XL's solution differs from traditional flow batteries, which rely on vanadium, a rare earth metal, and instead uses non-flammable, pH-neutral organic chemistry built around globally available petrochemical feedstocks. By sidestepping vanadium, XL says it avoids the rare earth metals supply chain and, as a result, can offer a safer, cheaper, and more scalable energy storage solution for the sector.

For Thomas Sisto, CEO of XL, flow batteries have several advantages over lithium-ion when it comes to their applicability to the data center market. He argues that while lithium is compact and energy dense, making it ideal for short bursts or mobile applications, it is much

less suited for long-duration scenarios, which are emerging as crucial, especially for AI data centers.

"We view lithium-ion as optimized for zero to six-hour applications," Sisto explains. "It starts to fall apart beyond that. Our system can provide power from minutes to multiple days."

This, Sisto contends, makes the battery much more flexible in its use. Unlike traditional storage systems, which excel in fast response or sustained output, XL's system can act as a "shock absorber," capable of smoothing unexpected load surges while also supporting overnight or multi-day operations when renewables aren't generating. This is especially useful for inference workloads where sudden surges are much more common.

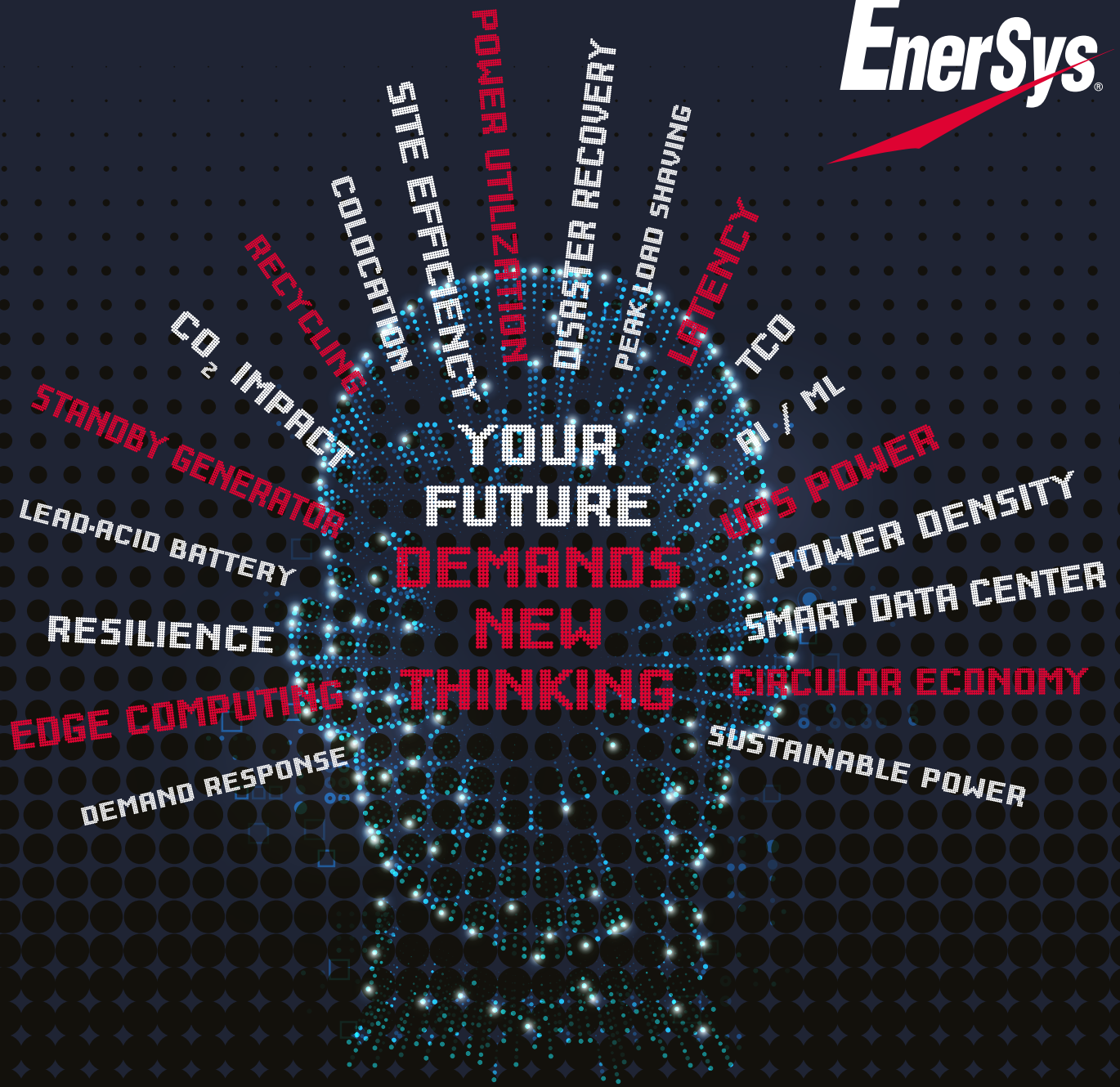
In addition, the battery has the potential to act as a long-term backup solution, potentially displacing diesel generators, as well as being a key part of a hybrid power solution for data centers, acting as a connective actor between distributed energy sources.

The technology does have some drawbacks, with a much larger physical footprint required and a supply chain that remains somewhat in its infancy. However, with data centers increasingly targeting greenfield development, space is often available, which would allow for the deployment of a larger number of batteries if necessary.

The Prometheus deal is only the starting point for XL's solution, with the company saying it is already in talks with a range of data center customers. With deployment expected before the end of the decade, we will soon see whether flow batteries could be the perfect fit for data centers.

As Sisto argues, there is no one-size-fits-all solution for battery storage in the data center sector, with a range of battery chemistries and solutions all with a potential role to play. Going forward, it's likely that we will see increased diversity in deployments, as companies continue to explore how batteries can best complement both the sector and the grid at large.

"We're not trying to be the one battery to rule them all, but we do believe there's a massive opportunity in the six to 100 hour range — and right now, almost no one is addressing that space well." ■



There is **so much** to consider.

Requirements are **changing**.

There is **another** way.

Improve your
backup efficiency:



www.enersys.com

Connecting the cloud



Georgia Butler
Senior Reporter,
Cloud & Hybrid

A look at the networking inside AWS data centers

By any measure - quantity, square footage, industry percentage - Amazon Web Services (AWS) has one of the largest data center footprints in the world.

In total, the AWS Cloud spans 114 Availability Zones in 36 regions, and each availability zone may be served by one or several data centers. This number is constantly growing.

Each data center features myriad internal and external connections, and with that comes a wildly complex networking system, one that is designed and mostly developed by AWS.

"We started developing our own hardware around 15 years ago," Matt Rehder, VP of core networking at AWS, tells *DCD*.

The reason behind that move, which becomes clear as Rehder goes on to detail the networking in AWS data centers, is the need for simplicity and scale.

"Commonly, companies will use slightly different hardware in each of those roles - connecting the servers, or connecting different data centers," Rehder says. "AWS is a bit different in that way because, many years ago, we decided to basically have the same thing



"The largest AI data center we have has more than 100,000 links or fiber connections within one building"



everywhere to make our lives simpler."

Of course, "simpler" is a relative term. Connecting massive data centers with hundreds or thousands of servers can never be quite considered "simple."

But, in a bid for an easier life, AWS developed what it calls a "Brick" - a rack of AWS network switches.

"It's a building block, and we can use Bricks wherever we need to - and then we make some of them functionally unique with some software changes," Rehder explains.

Each AWS data center has lots of Bricks, which can be either "connecting a bunch of servers to the network, connecting other AWS data centers together in the local region, connecting to third-party networks, or even connecting AWS Regions together across long distances."

Slightly different software gets loaded onto a Brick depending on the task, but "under the covers, they look almost identical," Rehder says.

This applies even when looking at AWS Local Zones, or to an extent, Outposts.

Local Zones are one of AWS' "Edge" offerings, and see the AWS cloud brought closer to end users in a data center where the company does not have a cloud region.

In some cases, these are third-party data centers, but according to Rehder, they are increasingly Amazon-owned facilities that are not part of a geographic cloud region.

"If you walk into a Local Zones facility, it looks much the same as one of our other data centers," he says. "There are some minor variations, for example, we add some extra layers of security and



resilience because it's further proximate from the region itself."

Outposts, on the other hand, are AWS' on-premises offering in which AWS hardware goes to a customer's data center.

Though Rehder says Outposts are "different" by their nature, he adds: "We try and make it as much of a pure AWS experience as possible, so it's the same servers we use, and the same effective network concepts, but oftentimes there are some extra layers on top of that for integrating into the customer's network in the building."

Brick by brick

AWS' Bricks are of their own design, and within the Brick itself, the company also provides networking hardware and switches.

The hardware, Rehder notes, has a single switching ASIC within it - a strategy that differs from most vendors, which will have multiple chips for this task.

"This is very good for certain places, but because you have a lot of these chips inside the system that are connected together, there's a lot of internal complexity in the switch, in the router, and that's hard to see or manage," Rehder argues. "It's much easier if you just break it down to that primitive unit, like the atomic unit, of just giving me the one chip in one device."

AWS doesn't actually make the ASIC - nor specify what company does - that goes into the network switches, but it does lay out the specifications for the components and the behaviour of the switch, as well as iteratively looking to improve its software and hardware design.

"We still have vendor devices, and in our experience, our switches are much better than theirs."

Another area of hardware that AWS has "gone very deep into" is the optical transceivers, which shine the laser down the fiber in cables.

Rehder explains that AWS network switches can have 32 or 64 ports in them, creating a really "fascinating complexity."

"Back in the days of 10G or 25G, optical transceivers were generally reliable, but it gets more complicated to push more speed through the fiber, and as we've gotten to say 400G or 800G, the

reliability has gone down," he says.

"We started to make material investments and went deeper into specifying how these optical module designs work, and trying to understand why links were failing.

"That led to a lot of investments in our fiber plant, and in how we redesign the modules to make them simpler and more reliable."

The biggest investment, Rehder says, was in the software that runs on these modules.

"Two or three years back, we decided to start investing in that, and we now own all the software that runs on those models, and we actively patch them all the time," he explains. "Now, our 400G generation is actually more reliable than our 100G. This is a huge advantage for us, especially with GenAI, where there are more links and the workloads in general are just more sensitive to failures."

Oodles of noodles

Connecting all of this hardware is, to put it politely, rather a lot of cabling.

The network switches are stacked up and connected together in the rack, and are then connected to AWS' server racks via fiber optic cables. According to Rehder, currently the cables within the racks use copper (he notes that the copper is "much fancier" than it used to be), but this isn't possible outside the rack due to distance limitations.

When asked how much cabling AWS has, Rehder laughs. "We have many, many, many miles of cables. I don't know the exact number, but it's definitely one of the harder parts of data center networking.

"The largest AI data center we have has more than 100,000 links or fiber connections within one building."

Rehder estimates that, given the "huge density of fiber connections in all AWS data centers," a single data center could have hundreds or thousands of miles of fiber cables within it, and they all have to be very carefully organized. This, apparently, is not as simple as just a color-coordination method.

"In the past year of my job, I've probably spent more time on the fiber cables - the building, how we design and

"In the past year of my job, I've probably spent more time on the fiber cables than any other part of the technology stack"



install them - than I have on any other part of the technology stack," Rehder emphasizes.

When establishing a new data center, Rehder says the company wants to turn up capacity as quickly as possible, and this means bringing in all the network racks and cables and plugging them in. "That is a very real work, physical job, and the sheer volume of cables is very high. The more connections you need to install, the longer lead time you'll have."

He also reiterates that it isn't like plugging in a power outlet - when working with fiber, any disturbance can

cause a degradation of the signal.

To help with this, AWS uses "structured cabling," which Rehder illustrates with a road analogy.

"You can route or bundle lots of the little pairs or threads of glass into creative physical structures. You basically have a freeway of these cables, and they can be dense - with hundreds or thousands of the fibers bundled in them, that go between rooms or rows, and then break off into smaller outlets and off-ramps into smaller 'streets' that flow to the racks."

Then, by using higher-density connectors, AWS can reduce the number of things that need to be installed or touched over time.

This saves time in the actual deployment, but adds swathes of complexity during the planning stage.

Machine learnings

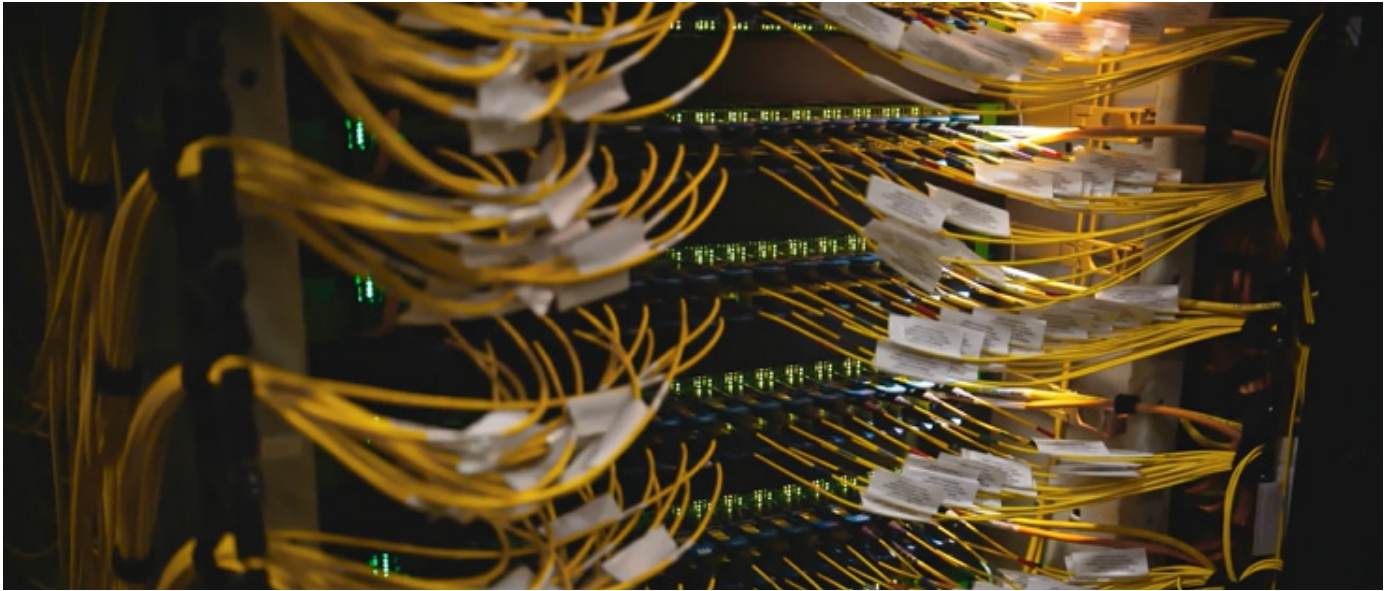
The cloud giant applies a strategy it calls "over subscription" to its networking in data centers.

Rehder explains this as having "more capacity facing the servers than you might have for them to talk to the Internet or to talk between data centers." The company can balance this "dynamically" when deploying capacity, so it can "turn the dial up or down depending on the workloads of the servers we have in place," he adds.

This, however, does not apply as effectively with AI or machine learning hardware. While the oversubscription model does enable AWS to "dial up" more capacity, Rehder notes that a machine learning or generative AI server will often need "two to three times the bandwidth of what other servers would have."

"A lot of the ML networks aren't oversubscribed, which means you build the capacity so all the servers can indeed talk to each other all at once, but that's something our fundamental architecture and building blocks allow us to do relatively easily," he says. "It's the same hardware, same switches, same concept, same operating system."

Across the industry, AI hardware is refreshing at a rapid rate, with Nvidia having moved to a roadmap featuring yearly updates for its GPUs. This is not the case with networking, with Rehder



noting that new generations emerge every three to four years.

"It depends on where industry hardware is going," says Rehder. "By staying on generations a little longer, you do get to a level of maturity with working the kinks out where it's easier and more reliable for customers if you aren't actually making changes all the time."

"It is moving a little faster right now, but it's still early days in terms of generative AI demand. We haven't really pulled in our hardware refresh cycles yet, we are more keeping an eye on when the next generation is coming, and if we want to do it when we normally would, or try and move a little bit earlier."

A technology shift that Rehder sees as potentially being interesting in the context of ML is co-packaged optics. This is an advanced integration of optics and silicon on a single packaged substrate, and aims to help increase bandwidth while reducing power consumption. However, the technology remains stubbornly on the horizon, perpetually ready to come to fruition "next year."

But Rehder believes we are getting closer to that "next year." He says: "It's real and it's happening, but there's a trade off."

"There are advantages in less power usage, and the short distances things are travelling.

But if you bake all the optics into the switch, then you've eaten all the costs associated with it, and if you aren't going to actually use every port on the switch

to plug something in, there's an extra cost associated."

Other emerging technologies are also looked at by AWS, but aren't necessarily in place as yet.

On the topic of Free Space Optics, Rehder responds: "Every couple of years, someone has the April Fool's Day project of Free Space Optics with the disco ball. I don't think we are there yet in terms of where it's a technology that is interesting."

He is, however, more optimistic about the potential of hollow-core fiber, though he notes that this is likely to play a bigger role outside of the data center than in.

AWS put hollow-core fiber into production for the first time in 2024. While it can reduce latency, within a data center itself, this is negligible.

"The latency inside the data center is already relatively low from the fiber in the building - they are all short runs, and there's a marginal advantage to reducing that further.

"Everything from testing and talking to customers suggests a couple percent reduction in latency doesn't really move the needle much in terms of performance," Rehder shrugs.

Reliability is king

While new and experimental technology is always exciting, for a business the size of AWS, reliability and uptime are the most important things for customers.

While AWS looks at new solutions, Rehder explains that there's always the question of "can we do something fundamentally different with new hardware?" If the answer is no, then "we don't see an advantage in that," he says.

He adds: "Moving to new hardware unlocks other risks, and it's nothing against any of the providers; there are always more kinks and bugs when you are learning.

"The new generations are always going to see complexities - and when you are looking to make things faster, bigger, and better, and pushing lots of limits across lots of different processes, the jump to the next complexity level is going to be even harder than the previous one.

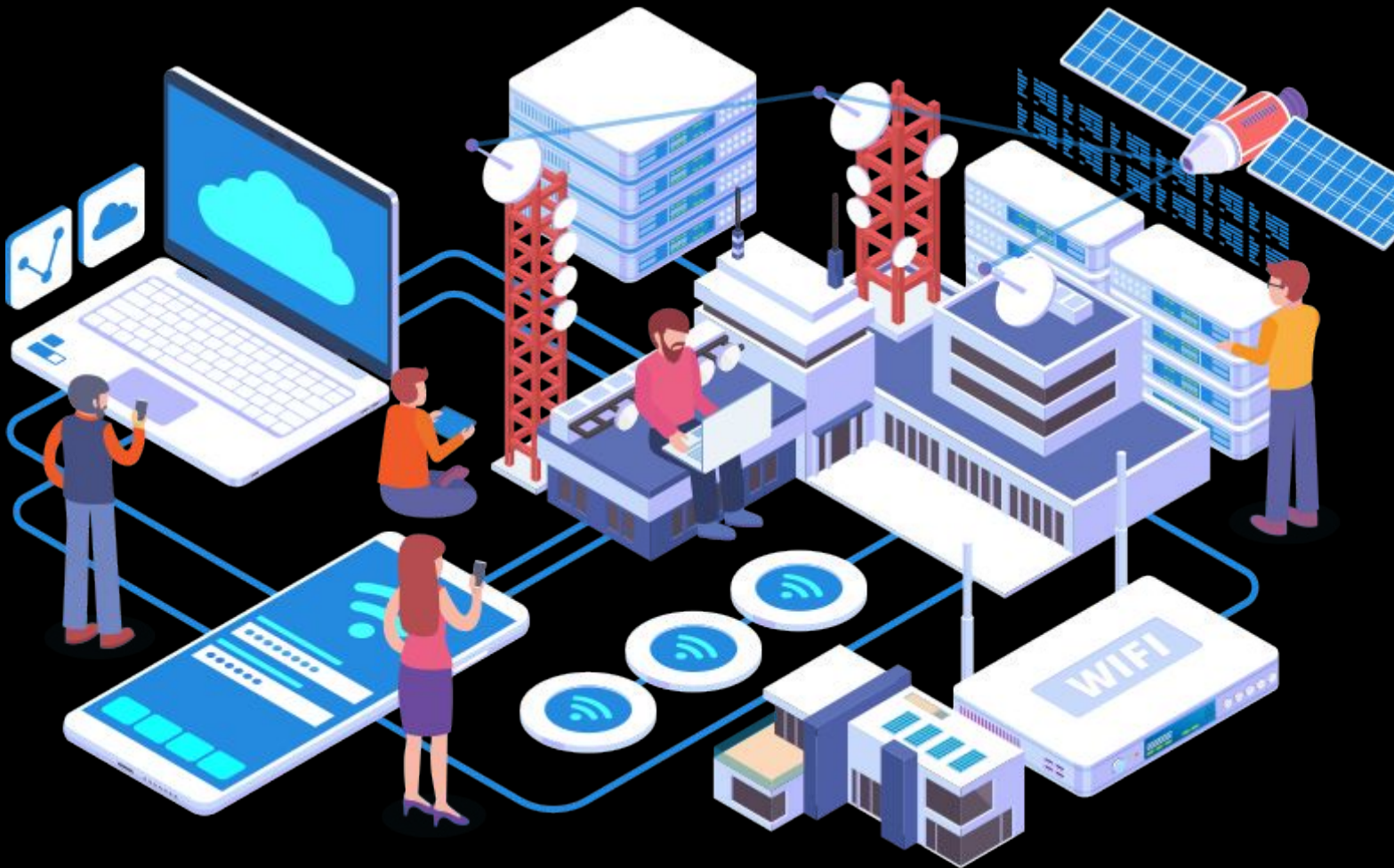
"There's going to be a lot of learnings, for example, in manufacturing and design processes. We try not to bring our customers through that. We'd rather already have the kinks worked out, and know it's going to work at scale."

With that backdrop, full-scale outages caused by the networks are very rare, but part of preventing those is always being prepared.

As Rehder tells *DCD*: "Everything fails. Everything will fail more than you expect it to, and it will fail in unique and exciting and creative ways."

By opting for a "simpler" networking solution, Rehder and AWS will be hoping to avoid too much of this kind of excitement in the future. ■

The Business of Data Centers



A new training experience
created by

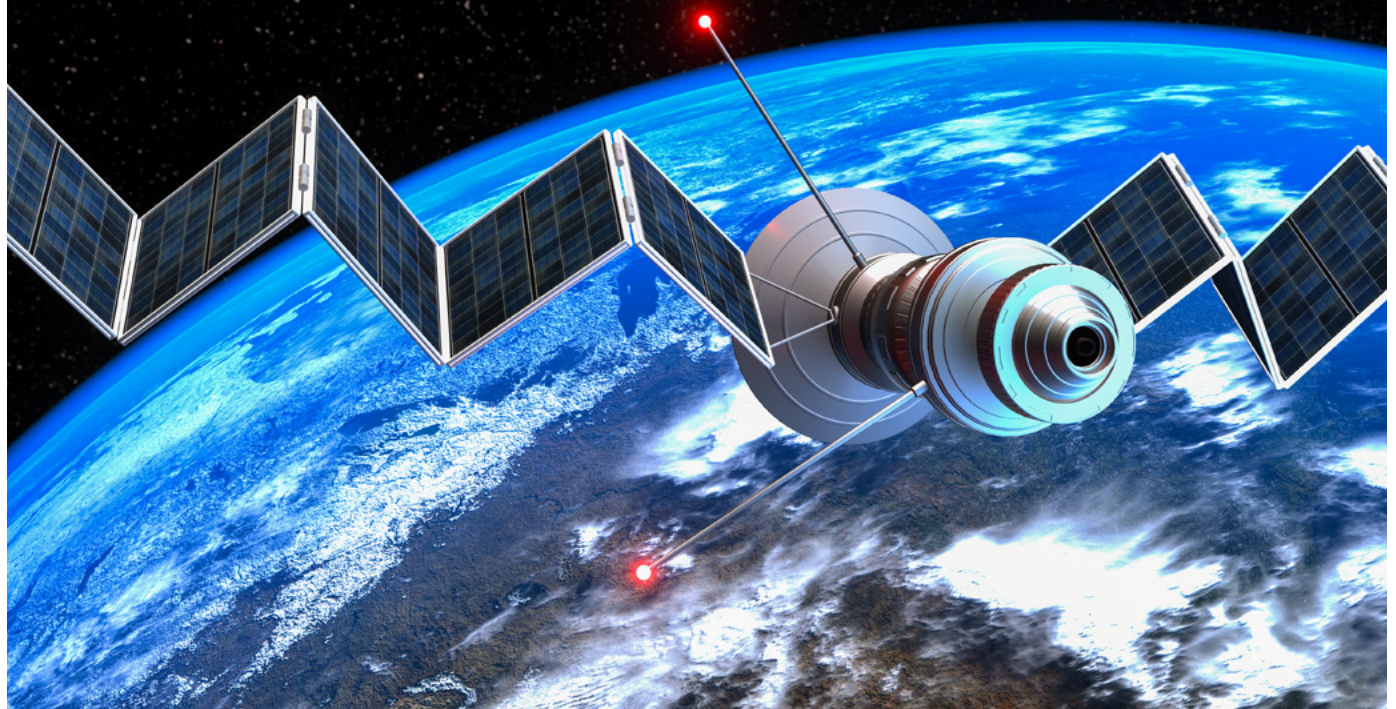


The changing landscape of Satellite M&A



Laurence Russell
Contributor

With big-LEO entrants disrupting satellite connectivity forever, consolidation is happening rapidly



The last three years have seen an unprecedented wave of merger and acquisition activity sweep across the satellite market, bringing a new diversity of capabilities to the connectivity market as operators gain access to one another's satellite fleets, forming multi-orbit offerings.

After years of excitement from some about the potential of low-earth orbit (LEO) satellite technologies, accompanied by the anxieties of others about its potential for disruption, Viasat acquired Inmarsat in a \$7.3 billion deal in June 2023. That same year, in September, Eutelsat completed its merger with OneWeb for \$3.4bn.

Most recently, SES, a satellite industry giant founded by the Luxembourg government, said it intends to acquire Intelsat for \$3.1 billion.

The impetus for these moves could be the emergence of big LEO offerings from Elon Musk's Starlink and Jeff Bezos's Kuiper; billionaire-backed all-sector platforms for consumers, enterprises, and governments that run the gamut of connectivity needs with an unmatched appetite for capex and competitive pricing, headed by enigmatic tycoons with a knack for always being in the headlines.

A less sensational explanation could be that economic circumstances have rendered consolidations organically inevitable.

"The efficiency of merged entities is the way they have to do it these days," a specialist in satellite M&A tells *DCD*. "The operational synergies are needed in the new market."

Satellite markets of yesterday

Charting the changing climate of this market means taking a look at how it used to work. Satellite commercialization was initially modeled upon the scientific achievements that pioneered the technology.

"Since the late 1990s, and until 2020, most commercial satellite infrastructure was located in geostationary orbit (GEO) at 36,000km," Pierre Lionnet, research and managing director at ASD-Eurospace, says. "This favored the development of large, high-capacity satellites able to provide global coverage of the Earth. They provided high relays

in the sky, complementing deep-sea cables to connect any point on Earth, even in remote areas lacking terrestrial infrastructure, including oceans and deserts."

This GEO dominance was mostly B2B, involving a mix of TV broadcast and data transmission for enterprise and government.

"It was a great run because of growth in direct-to-home," adds a specialist in satellite M&A. "[Broadcast satellites brought] long-term contracts which were very financeable for a satellite with a 10-15 year life. There was a huge tailwind in GEO growth because of the expansion of television channels. That came to an end with the concept of over-the-top content delivery that went through the internet. The business became about data, not video streams."

This need for the transmission of data was something GEO satellites always struggled with, never being able to compete with the speeds of terrestrial connections. Since the 1990s, LEO has been a preferable alternative.

"Commercial LEO communications systems from the late 1990s served niche market segments (mobile voice and data communications, narrowband and IoT), but only two operators eventually survived the first LEO wave, Iridium and Globalstar," Lionnet explains.

First experimented with in the 1980s, LEO satellites fly far closer to the Earth, allowing much higher throughput and lower latency, though this altitude means they cannot maintain geosynchronous positions, zooming over continents and oceans in minutes. They could only be useful in large constellations of satellites that would ring the planet, with at least one platform always in range of Earth's inhabited continents.

The principle was once as outlandish as moon bases and space elevators, until demonstration of low-cost heavy rideshare rockets brought the price of launch per satellite crashing down, and dozens of powerful proliferated satellites could be inserted at once, enabling new highways of data transmission.

Despite this new vector of disruption, the balance has yet to be upended. While LEO is growing fast, geosynchronous satellites continue to represent the bulk of the market.

"[The LEO] segment is currently worth less than a billion," Lionnet says. "The GEO segment represents \$15-20bn in annual revenues."

In addition to GEO's legacy status, it also possesses core advantages over LEO that will never quite be matched, namely its power to cover whole continents with a single satellite.

Multi-orbit necessity

The fast rise of Starlink combined with the verticalization strategy of SpaceX, which Kuiper may well seek to emulate, have drawn suggestions of monopolizing forces in the satellite market, one which suggests these networks may adopt a GEO capability to establish a truly comprehensive multi-orbit offering like the newly consolidated entities of Eutelsat-OneWeb or SES-Intelsat (though the latter uses medium-Earth orbit through the O3b mPOWER constellation).

"GEO systems are still the most competitive of all large-area broadcasting solutions (in terms of cost/bit transmitted), and there is still demand for broadcast," Lionnet says. "They are also inherently more efficient to serve areas with higher density of customers, because there is no wasted capacity. Also, the government segment in GEO is still very strong, and GEO satellites remain a very efficient solution for all applications where some latency is not a critical issue (or for those governments that can't afford a global LEO system)."

While SpaceX has opted to develop many impressive technologies in-house, it isn't above growing by acquisition. In 2021, it purchased startup Swarm Technologies for an undisclosed sum, gaining access to 30 smallsat specialists and a network of 120 tiny satellites and equally small antenna designs.

While heralded as a rarity for the spacefaring giant, the news isn't the first case of SpaceX shelling out for market-beating tech, given the company took a 10 percent stake in Surrey Satellite Technology Ltd in 2005, whose founder Sir Martin Sweeting is thought to be the father of the modern small satellite.

Satellite orbital realms aren't the only decisive, differentiating advantage in satellite markets that influence consolidation pressures.



Are some applications consolidation-proof?

Some of Starlink's contemporary success has come from working with existing satellite resellers to market capacity, much of which goes unused as satellites orbit above nations and oceans with which the SpaceX marketing teams are not as well acquainted. Since the second half of 2022, Starlink has been working partners like Speedcast, SES, KVH Industries, netTALK maritime, Singtel, Tototheo Maritime, and Marlink in maritime shipping alone.

That relationship suggests that as Starlink further establishes itself, the services it provides will become essential enough that it has the leverage to cut out or acquire companies acting as its middlemen.

Our expert in M&A calls this a definite option for resellers on the consumer side, and maybe even in-flight connectivity in the aviation market, but these industries behave very differently.

"It's easy to install these terminals at your house, but it doesn't seem like Starlink wants to be responsible for guaranteeing the performance of antennas on cruise ships and oil rigs," they say. "You're talking about chartering helicopters for offshore maintenance. I don't see SpaceX committing to delivering something like that. Guys like Speedcast have been honing the ability to do this for years. I would say that the more complex the installation and maintenance,

the better the business case of an independent value-added reseller."

Acquire to compete?

When asked if he foresaw the possibility of Starlink and Kuiper directly buying their competitors, in LEO or otherwise, Lionnet has a simple answer.

"No. But I do think that with all the plans announced, the competition in the LEO B2B space - IFC, maritime, private networks - will drive prices down and put some players out of business," he says. "The B2G space will create some opportunities for medium-sized constellations, such as IRIS² and PWSA, but the global B2C segment will probably not be able to keep in business more than one or two very big, Starlink-like, operators with thousands of satellites."

One such competitor, Rivada Space Networks, which is building the Outernet constellation for enterprise and government, recognizes the build-or-buy ultimatum for unique satellite technologies.

"Some may try to build [unique technologies like Outernet] but I think it's extremely likely that those operators - and perhaps the cloud service providers themselves - will seek to buy those capabilities to offer a full portfolio of connectivity services to their end customers, just as the GEO operators have been buying LEO assets to fill out their portfolio," speculates Brian Carney, senior

vice president at Rivada Space Networks.

Our expert in M&A predicts more satellite manufacturers becoming affiliated with an operator, perhaps further consolidating some newly merged satellite players.

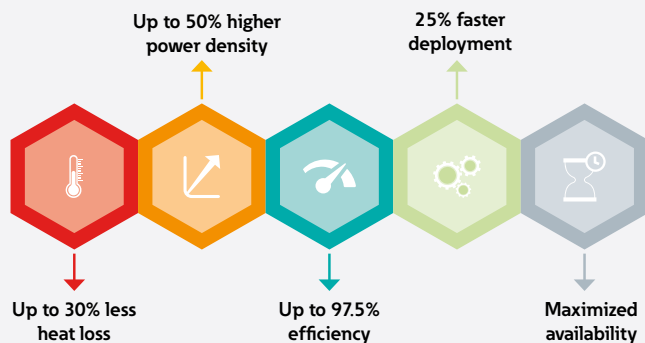
"Building satellites is a lumpy business," they say. "Operators guarantee annuity which is good from a finance perspective."

Their understanding of the market was one of compounding joint ventures, collaboration, and mergers, something they believe has now slowed.

"The last 120 days have put a spanner in the works on that," they admit. "Within respective geographies like Europe, you'll see a continuation of demand for in-house technology, which is going to drive M&A. In the US, you'll see consolidation and verticalization between primes and subcontractors, from manufacturing to operation. Developing nations will be seeing more use for government applications as part of a trend for satellite to become a core requirement of modern militaries. All things considered, I'm more optimistic about the essentiality of the satellite industry as a solid slice of global GDP than I was ten years ago."

As markets normalize following recent shocks, consolidation activity could well resume the pace experts had been tracking, but the reality of monopolization will depend on Starlink and Kuiper's more volatile drivers finding equilibrium. ■

Eaton 9395X – the neXt generation UPS



The Eaton 9395X is a new addition to our large UPS range. It builds on a legacy of proven power protection of Eaton's 9395 family, providing a market-leading footprint with the best power density, leaving more space for your revenue generating IT equipment.

This next generation Eaton 9395X UPS offers more power, with ratings from 1.0 to 1.7 MVA, in the same compact footprint and brings you even more reliable, cost-effective power, which is cleaner thanks to our grid-interactive technology. With world-class manufacturing processes and a design optimized for easy commissioning, the 9395X offers the shortest lead-time from order entry to activation of your critical load protection.



Ease of deployment

- Faster manufacturing due to ready-built sub-assemblies
- Simplified installation with inter-cabinet busbar design
- Plug-in self-configuring power modules
- One-click configuration of large systems



Compact footprint

- Occupies up to 30% less floorspace, leaving more room for revenue-generating IT equipment
- Installation can be against a wall or back-to-back
- Integration with switchgear saves space, as well as the cost of installation and cabling



Cost efficient & flexible

- Save on your energy bill with improved efficiency of 97.5% and reduced need for cooling due to up to 30% less heat loss
- Choose the correct size capacity for your immediate needs, and easily scale up later in 340 kW steps
- Optimized battery sizing with wide battery DC voltage range



Easy maintenance

- More reliable self-monitoring system
- Less need for scheduled maintenance checks
- Safe maintenance while keeping loads protected
- System status information provided



More resilient

- Builds on the capabilities of the proven Power Xpert 9395P UPS
- Improved environmental tolerance for modern datacenters
- Component condition monitoring
- HotSync patented load-sharing technology



EnergyAware

- Native grid-interactive capabilities
- Reduce facility operating costs or earn revenue through energy market participation



Powering Business Worldwide

Scan QR code
to know more



OFC 2025: Hollow Core Fiber hype stands out amid the AI overload



Paul Lipscombe
Telecoms Editor

The credentials of hollow core fiber were examined by the industry, as the industry grapples to power the data-hungry AI future

This year marked a special milestone for the Optical Fiber Communication Conference (OFC). It was the 50th edition of OFC, an event that dives deep into the world of optical networking and communications. Held in San Francisco, California, this year's OFC attracted 16,700 attendees from 83 countries.

DCD was among those that flocked to the Moscone Center for the event, and found that AI seemed to pin together all the key themes, as the optical communication industry gathered to discuss the opportunities that the technology presents, and how the sector can build it into the networks of the future.

A key focus was on how AI will influence the demands on data center infrastructure as hyperscalers push to adapt and build future-proof networking.

Developments around hollow

core fiber (HCF), subsea connectivity, pluggables, and Generative AI were all discussed, plus silicon photonics and much more.

High scores for HCF

It was evident from the get-go that HCF was going to be a key theme at OFC, given the number of keynotes dedicated to the topic. HCF features a hollow space in the middle through which light is transmitted, rather than the glass core found in traditional fiber.

This could offer considerable speed benefits, but is the industry ready for it, and what sectors will it serve?

"We don't change our fiber very often," Andrew Lord, optics and quantum R&D lead, BT, noted in a keynote, where seating was at a premium.

BT began deploying fiber back in 1982. Some of that fiber was recently tested by the telco, explains Lord, who stated that it

is still performing as well as it did 40 years ago.

"The fiber still works, it hasn't changed," he said. "Fiber is really good. I mean, the fiber guys here [referring to an image from the 1980s] have done a good job, maybe too good. But maybe since 1982, it's time to start to refresh that infrastructure," he adds.

Lord told *DCD* earlier this year that HCF won't replace the fiber that is being used to roll out networks across the world at present, as telcos pivot away from legacy copper infrastructure. He reiterated this point during OFC, explaining that it's too expensive to install at that sort of scale. The use cases HCF has been earmarked for include high-frequency financial trading.

"Does anybody here think that the use case here is a complete wholesale swap out of existing fiber for a new hollow core fiber? I don't think so," said Lord.

Instead, he envisions that HCF could play a role in the future in the development of the Radio Access Network (RAN), ultimately helping to reduce latency. He suggested that HCF could go that bit further, to areas standard fiber can't reach.

"It's a niche, but that's what we want," said Lord. "We want examples where it's not a wholesale replacement, but where it steps in to help in situations where standard fiber doesn't do the job."

Lord added that HCF could be utilized for resiliency, and possibly even for quantum networks, and could play a role in the future of subsea cables.

A rare opportunity for fiber

The discussion around HCF and its potential is only likely to grow, according to Jason Eichenholz, co-founder, Relativity Networks. The company specializes in providing fiber, including hollow-core fiber, for its clients.

"If somebody is not talking about it now [HCF] they're asleep at the switch," Eichenholz told *DCD*.

"The challenge is that in the market there's an insatiable demand for hollow core fiber. You can't go through a technical session at this conference talking about the future without bumping into hollow core.

"Quantum is another hot topic, but that's a bit further out. Hollow core has real-world applications today, especially in the data center space. We have an opportunity that doesn't happen very often. Every 50 years or so, you get a big paradigm shift. We're changing latency and we're changing the speed of light and fiber."

Much like Lord, Eichenholz doesn't see a need for HCF to go directly to the home, but does think there are opportunities around the passive optical network (PON). PON is where a single fiber optic cable is used to deliver data to multiple users, usually for services such as broadband.

"I don't think it's a world where, in the next 20 years, you'll see hollow core fiber going to the home. That being said, if you look at PON networks, I absolutely think there are proven opportunities for hollow core fiber, because of the power-handling capability in that network. It won't be that last mile to the home, but getting that PON signal down to the

splitters? Absolutely."

Network vendor Nokia also outlined opportunities around HCF. During a media briefing, Edward Englehart, VP of engineering and head of subsystems at Nokia Optical Networks, stated that HCF can drive the open line system (OLS), which is a modular architecture for optical networks.

"Hollow core can really bring some key application advantages and some really interesting technology advantages that will drive the OLS and where we're going with the OLS," says Englehart.

"So first of all, with hollow core, there's significant latency improvements.

Second, there is much higher power you can transmit over this, which gives you some better reach, and more importantly, there is significantly lower electro-optics or impairments based on the fiber, which again gives you better performance."

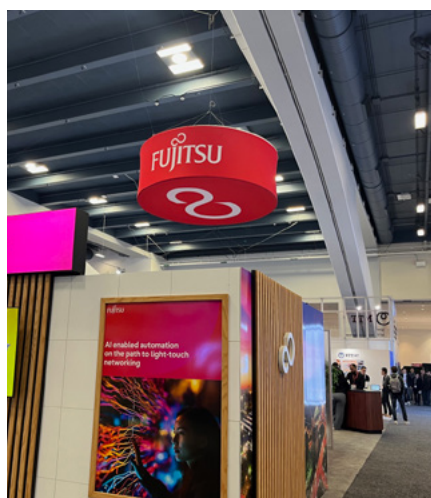
The subsea opportunities of the future

Relativity Networks' Eichenholz noted that the potential of HCF was also touched upon in other keynotes covering



"It's a niche, but that's what we want,"

>>Andrew Lord, BT



areas beyond telecoms and data center opportunities.

One such market that noted opportunities in this field was the subsea cable industry. More than 370 cable routes are active across the globe. During a keynote on the future these systems, Alexis Carbo Meseguer, system design, Alcatel Submarine Networks (ASN), explained that HCF could be a potential alternative for future submarine cables.

"Under power constraints, it could be interesting, but also the fact that this fiber has no linearities, so it can help to adopt multi-band systems," said Meseguer. But, he warned, "it's not a mature technology at all."

Discussing potential trends around subsea cable demand in the future, Pascal Pecci, submarine cable systems engineer at Meta, noted that AI is the "driver for the need for additional capacity."

He said that the current technology

maximum of 24 fiber pairs is nowhere near enough to meet the capacity demands. On this basis, only half a Petabit of capacity is possible.

"For transporting distance at a capacity of half petabits, we are using the maximum number of fiber pairs that subsea cables can manage, which is 24 fiber pairs," said Pecci.

He explains that 1Pb repeatered cable builds will launch soon, possibly within the next two to three years. "We believe that in the coming years a 1Pb cable will be built, because we all know that for subsea, if we include unlimited distance on repeated cable, one petabit is already reached," says Pecci.

However, he doesn't stop there, noting the next step will be for 2Pb, though this will come somewhere between five and 10 years down the line.

Nvidia's CPO push

Another hot topic on the menu at this year's show was co-packaged optics (CPO), which is the integration of optical and electrical components into a switch ASIC package, rather than using separate pluggable modules.

CPOs have been identified as an important way of addressing the growing bandwidth and power challenges in data centers.

Just weeks before OFC, Jensen Huang, founder and CEO of Nvidia, showcased a strong appetite for silicon photonics during Nvidia's GTC event.

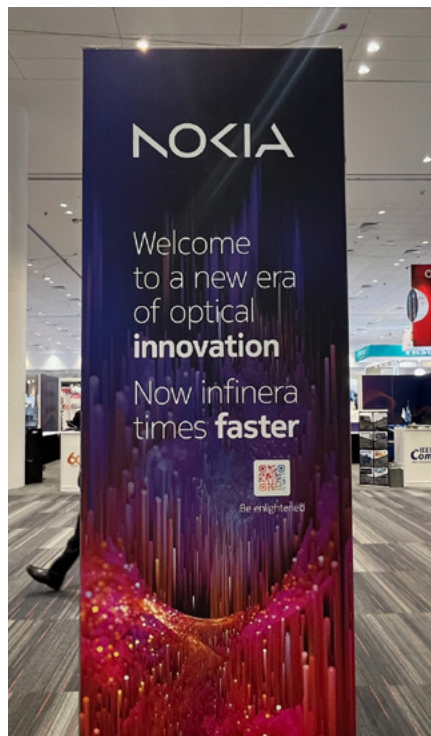
The chipset giant unveiled its silicon photonics networking switches which the company claims can lower power consumption and improve deployment speeds. Compared to undisclosed 'traditional methods,' the company promises 3.5x more power efficiency, 63x greater signal integrity, 10x better network resiliency 'at scale,' and 1.3x faster deployment.

A key area of photonics technology is CPO, and this is something that Nvidia is putting a lot of effort into, explained Craig Thompson, VP LinkX products, networking group, Nvidia.

"We made an announcement at GTC around our CPO investment. This is a major initiative inside the company, and we're putting a lot of attention and resources behind it," Thompson told a media briefing at OFC, which delved into

"If somebody is not talking about hollow core fiber now they're asleep at the switch,"

>> Jason Eichenholz, Relativity Networks.



the need to scale networks to support the next generation of AI workloads.

"We've put our weight behind CPO for a number of reasons, but I would say the two most pressing are power and reliability, or cluster uptime."

He added that CPO is a way for Nvidia to "simplify the high-speed interfaces from the core switches to the optical interface."

Best of the rest - Gen AI, chatbots, Infinera, and the road to 1.6Tb

It wasn't just data center folk at OFC, the big US network carriers were also present. Unsurprisingly, the topic centered around everyone's favorite buzzword, AI.

Generative AI's role for telcos was

touched upon during the event. Larry Zhou, lead member of technical staff, AT&T, highlighted some of the earlier generative AI use cases being adopted by carriers, and unsurprisingly, it featured use cases such as chatbots for customer service support.

"We believe agentic AI workflow is going to rewrite the whole industry," explains Zhou, who says this will go beyond chatbots and is pivotal towards creating truly autonomous systems.

Zhou also added that it aims to use generative AI to "predict problems before they escalate."

Meanwhile, network vendor Nokia had a big presence at this year's event, pushing its data center ambitions.

This included the company showcasing its recent \$2.3 billion acquisition of Infinera, notably through advertising across the showfloor. [Read more on page 76 about Nokia's data center play]

It was also difficult to ignore the 1.6Tb hype on show in North California. As networks move away from 200G to 400G and 800G, the industry is already pushing the idea of 1.6Tb optical speeds.

This, of course, matters because the demands of AI mean that the computing power needed is greater than ever, and will only continue to grow. There were many 1.6Tb optical modules on show at OFC. Even if these modules are some way off deployment, it shows that the industry is already looking in the future, being proactive, not reactive.

Thompson added that he expects the bandwidth required for the optical networks, to continue growing.

"GPU bandwidth is doubling every two years," he noted. "We've seen bandwidth across a single lane double in roughly the same period for the last few generations and I don't really see it slowing down yet."

Thompson expects this to continue in the coming years. Although 1.6Tb is yet to hit deployment, it probably won't be too long before 3.2Tb is plastered around the OFC showfloor.

"The ride will continue for at least the next few years, and it will be a really exciting time," he says.

"We need the ecosystem to come along with us. We need innovation, startups, and investment. We need bigger networks." ■



LIQUID LEAK DETECTION SYSTEMS

LIQUID COOLING SOLUTIONS

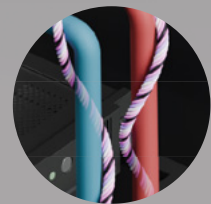
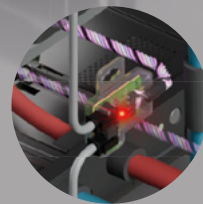
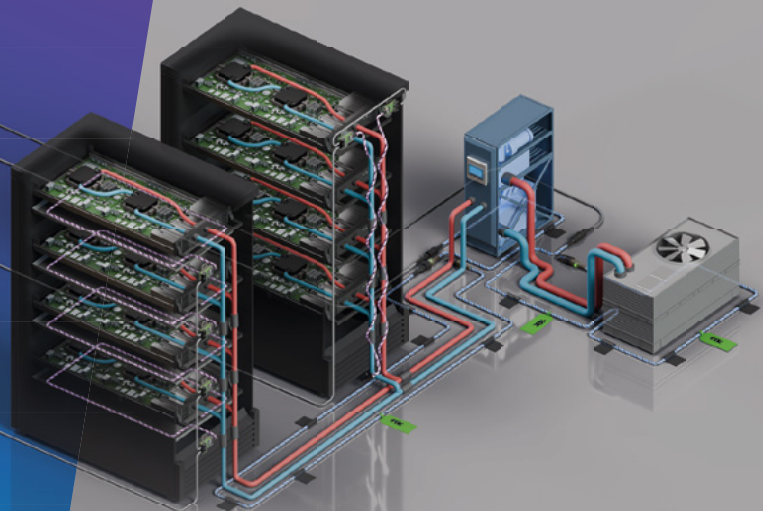
LIQUID COOLING READY !

TTK Leak Detection Solutions FG-DLC for Direct-to-Chip System:

**Compact. Precise.
Purpose-built**

- ✓ Compact Design
for Efficient
Server-level
Monitoring
- ✓ Detects Conductive
Coolants Including
PG25
- ✓ Instant, Accurate
Leak Detection with
Zero Compromise

**Protect
your servers
with confidence**



Learn more at
www.ttkusa.com

Act like it



During the flurry of announcements and executive orders that marked US President Donald Trump's first week, one of the most promising ones was the declaration of a "national energy emergency."

The returning president argued that a "reliable, diversified, and affordable supply of energy" was critical for manufacturing, transportation, agriculture, and defense industries.

He added that the situation was set to "dramatically deteriorate" if left unfixed, thanks to the high demand for power from AI data centers.

The only way to address this, he argued, was a dramatic effort to improve power generation and the grid. Data center operators, unsurprisingly, cheered on his comments, as they struggle to find power amid an unprecedented boom.

A few months later, the reality has been less than inspiring. Instead of a genuine attempt to improve the grid, ideological attacks on renewable energy and government research have left the nation weaker than ever.

Offshore wind has been killed entirely, thwarting plans for gigawatts of power across the US. Wind and solar projects on federal land have been stopped.

DOGE cuts have crippled the DOE, including non-renewable grid utility efforts, and led to a mass staff exodus. Further cuts are expected, with budget funding for wind, solar, and battery storage projects cut to zero for its 2026 financial year.

Critical research is being abruptly ended, potentially impacting grid improvements for decades to come.

Utilities and suppliers have scaled back plans in the US thanks to the loss of government support and tariff confusion.

This is not just about reducing emissions. Active efforts to block renewable energy projects and fire researchers are bad for the grid, bad for the economy, and bad for the data center sector - whether you believe in climate change or not.

China knows this, deploying transformative quantities of solar and wind, even as it continues to use fossil fuel sources. It is building the infrastructure to support gigawatt clusters and beyond.

As for the US, if the administration truly believes that there's a national energy emergency, then it's time to act like it. ■

- Sebastian Moss, Editor-in-Chief



ABB Data Centre Solutions.

Trusted. Reliable. Proven.

ABB's complete electrical portfolio provides the energy efficiency and energy insights to monitor and reduce power usage as well as technological advancements to lower carbon and GHG emissions. ABB is your premiere sustainability partner, providing 100+ years of our domain expertise in electrical solutions to solve some of your most difficult sustainability issues.

Let's write the future. Together.





> Zero Downtime

Brought to
you by DCD's
award winning
editorial team

Listen to the
DCD **podcast**,
new episodes
available every
two weeks.

ON AIR



Hear from Vint
Cerf, Synopsys,
Microsoft, Google,
Digital Realty, and
more!

Tune in for free:
bit.ly/ZeroDowntime