DCS DATACENTRE SOLUTIONS

Developing digital infrastructure in a hybrid world

AUTUMN 2020

datacentre.solutions.com

AI + Automationps | ArtArchitecture + Connectivityps | ArtCloud + Managed Services
Colocation + Hostingps | ArtDesign + Infrastructureps | ArtEnergy Optimisation
Management + Maintenance | Physical Security | Power + Cooling

enel x

Energy management solutions that deliver commercial and environmental rewards

Global resources, local market knowledge and specialist technological expertise make Enel X the trusted energy management partner for data centres and large scale enterprises all over the world.

Join us as we pave the way for a cleaner, more dynamic future.

enelxuk@enel.com | www.enelx.co.uk

Editor's View

By Phil Alsop

Energy efficiency in the spotlight

WELCOME to the autumn issue of Data Centre Solutions. I'm delighted to say that this issue includes a major focus on a topic that, at some stage in the future, if not already, might just make the current chaos surrounding the coronavirus pandemic seem little more than a minor inconvenience. And that topic? The environment. An issue as divisive as Covid-19 for sure, when it comes to what action to take (or not), but one that is likely to remain centre stage for as long as humans remain on earth.

In the green corner, we have the folk who are telling us we are already too late to reverse the massive environmental changes which are damaging our planet and which will eventually lead to some kind of natural world Armageddon. I'm not sure what colour we should give to the opposite corner, but that's a space filled with the sceptics and climate change deniers, who point to the millions of years of environmental volatility experienced by our planet, suggesting that what we are experiencing is no more or less than yet one more period of extreme conditions, and we need to wait a few hundred thousand years before we can call it a significant trend.

There seems to be a general consensus that we would do well to work on the assumption that human behaviour and climate change are not unrelated topics, but beyond this, as with the coronavirus, there is massive tension between what is seen as a straight choice between health and wealth. Green the planet at a massive financial cost to us all; ignore environmental issues for the pursuit of business and wealth, and activate a ticking environmental time bomb.

Happily, there are an increasing number of occasions when the objectives of health and wealth do actually coincide.
In the data centre industry, the environmental debate has reached an interesting moment.

Many, if not all, data centre owners and operators have worked towards improving the energy efficiency and/or reducing the carbon footprint of their facilities. Initiatives such as PUE and carbon offsetting have proved valuable in this regard.

However, there is the opportunity, if not an imperative, to take energy efficiency to the next level within the data centre – what we might call Energy Efficiency 2.0. Innovative technologies and approaches offer the potential of significant environmental and accompanying financial rewards. These include:
Demand response; Waste heat re-use; Renewable energy; Decarbonisation; Microgrids/off-grid power; Liquid cooling; Data lifecycle management.

In this issue, we take a look at the topic of energy efficiency in detail, with a major focus on the potential of liquid cooling. We're also covering AI and automation, a not unrelated subject as it promises levels of operational accuracy and efficiency which can contribute to the overall energy efficiency of the data centre.

I hope you enjoy reading the magazine, and we're always grateful for feedback – whether brief comments or maybe you will be inspired to contribute content for a future issue.

DCS DATACENTRE SOLUTIONS

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

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

Sales Manager
Jessica Harrison +44 (0)2476 718970 jessica.harrison@angelbc.com

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

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

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

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

Directors
Stephen Whitehurst – CEO
Scott Adams – Joint - Managing Director
Sukhi Bhadal – Joint - Managing Director

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

Angel BUSINESS COMMUNICATIONS

recycle When you have finished with this magazine please recycle it.

Datacentre Solutions is published 4 times a year on a controlled circulation basis in Europe, Middle East and Africa only. Subscription rates on request. All information herein is believed to be correct at time of going to press. The publisher does not accept responsibility for any errors and omissions. The views expressed in this publication are not necessarily those of the publisher. Every effort has been made to obtain copyright permission for the material contained in this publication. Angel Business Communications Ltd will be happy to acknowledge any copyright oversights in a subsequent issue of the publication. Angel Business Communications Ltd.

The paper used within this magazine is produced by chain of custody certified manufacturers, guaranteeing sustainable sourcing. ISSN 2396-9016 (Online)

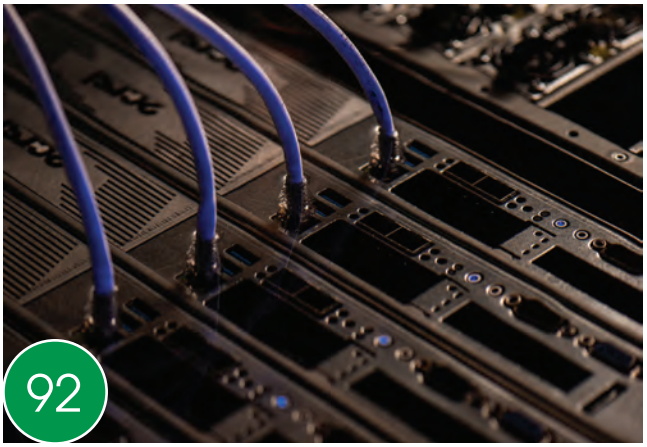
Contents

AUTUMN 2020

LIQUID COOLING NEWS

LIQUID COOLING

DCA NEWS



Cloud can reduce carbon emissions

COMPANIES are unlocking greater financial, societal and environmental benefits by moving to the cloud and developing associated circular products and services.

Migration to the public cloud can achieve significant carbon reduction in the form of a 5.9% decrease in total IT emissions or nearly 60 million tons of CO2 globally per year, which is the equivalent of taking 22 million cars off the road, according to new research from Accenture.

Informed by proprietary Accenture analysis based on its experience supporting sustainable cloud migrations for hundreds of clients globally, the report – The Green Behind the Cloud – outlines how companies can achieve the most value from environmentally-friendly use and operation of cloud services, regardless of where they are on their cloud journey.

Businesses are facing more pressure to solve large socioeconomic challenges and shift toward more responsible and sustainable practices while boosting profitability. The latest Accenture Strategy-UNGC study found 59% of CEOs say they are deploying low-carbon and renewable energy across their operations today while 44% see a net-zero, carbon neutral future for their company. Moreover, two-thirds view technologies like cloud as critical factors for accelerating change and making their commitments a reality.

"Sustainable cloud can deliver a double helix effect of shareholder and stakeholder value by simultaneously reducing costs and carbon emissions," said Peter Lacy, a senior managing director and global sustainability lead at Accenture. "Further, the magnitude of carbon reduction achieved through cloud migrations can go a long way in meeting climate change commitments and driving new levels of innovation, ultimately leading to a greener balance sheet and a greener planet."

Adding to the significant environmental impact, sustainable cloud solutions deliver key financial benefits. Accenture's analysis, based on its work with clients, shows up to 30-40% total cost of ownership savings from public cloud, driven by greater workload flexibility, better server utilisation rates and more energy-efficient infrastructure.

According to the report, the sustainability and financial benefits from cloud migration will vary based on three key factors: the cloud provider selected, the ambition level for cloud optimisation, and the level of cloud-enabled sustainability innovations.

"Companies are rapidly moving to the cloud for innovation and cost-savings, and sustainability must also be considered as a primary driver," said Paul Daugherty, group chief executive – Technology and chief technology officer for Accenture. "However, there is no one-size-fits-all approach to sustainable cloud journeys — companies must understand the migration, design, and engineering decisions that will directly determine how sustainable their solutions are and the benefits they drive."

When it comes to cloud optimisation, the report outlines three ambition levels in the cloud first sustainability journey: strategic migrations without major redesign, application of sustainable software engineering practices, and application optimisation for the "fabric of the cloud."

Accenture analysis shows initial cloud migrations alone can reduce carbon emissions by more than 84% compared with conventional infrastructure. Reductions can be pushed even higher – by up to 98% – by designing applications specifically for the cloud. To get a current cloud CO2 emission score and recommendations to reduce carbon footprint, Accenture provides a Green Cloud Advisor module as part of its myNav platform for clients interested in boosting their green index goals.

Organisations face difficulties securing budget for IT resilience

ALMOST one in four business still struggle to determine the cost of IT downtime, reports new research from Databarracks. Almost a quarter of organisations are still unable to determine what IT downtime costs their business. This is according to new research conducted by Databarracks.

Data taken from its annual Data Health Check survey has revealed almost one in four organisations (23 per cent) are not able to accurately report the financial impact of IT-related downtime on their organisation.

Given the unprecedented disruption caused by COVID-19, Peter Groucutt, managing director of Databarracks, states it is critical businesses make every effort to think more broadly about the financial impact of IT downtime.

"Being able to recognise and attribute costs related to IT downtime, helps organisations to make better-informed decisions on IT resilience, supplier management and continuity planning, which is particularly pertinent in the current climate."

The Databarracks survey also showed almost one in five (18 per cent) organisations stated as little as one hour of IT-related downtime can cost over £50,000. Groucutt continues, "Businesses can't afford unplanned

downtime. As organisations have migrated en masse to remote working, people have become accustomed to having access to data and systems from wherever they work. But when those systems are no longer available it not only means staff aren't able to do their jobs, it also comes with a severe financial cost to the business. "Costs relating to staffing, lost revenue and the financial outlay needed to fix an outage will be known, but it's important businesses go further.

For example, how do you determine the financial impact to customer perception if a vital application goes offline for more than 24hrs? These are more difficult to estimate but can have an even bigger overall impact.

"Hidden costs often take longer to materialise, meaning they can be excluded when estimating the cost of downtime. It's critical organisations do not ignore them. At a time when any outgoings will inevitably be scrutinised, it will be difficult to secure budget for IT resilience if you can't clearly show the impact of downtime. Presenting a complete downtime cost immediately gives perspective to any expenditure, which will help you gain the resources needed to make improvements.

Irish data centre industry to add 1800 new jobs

NEW SURVEY from Host in Ireland indicates 96% of data centre ecosystem companies are optimistic about business outlook despite pandemic.

A survey released from Host in Ireland finds an optimistic business outlook over the next twelve months for the Irish data centre industry. Ninety-six percent (96%) of companies surveyed are positive or very positive about business opportunities within the data centre and hosting service industry.

Eighty-four percent (84%) of these companies also expect their employee base to grow in the next 12 months, potentially creating more than 1800 new jobs. The survey was conducted with a sample of 50 companies in the data centre ecosystem in Ireland, including Host in Ireland partners and other invited participants.

"Data centres are the digital factories of today and have been one of the few completely open for business during the pandemic. As significant portions of our lives shifted to an online world, the industry has felt a great sense of purpose as essential workers being on the digital frontline," said Garry Connolly, president and founder of Host in Ireland. "The positive outlook and enthusiasm is understandable. When you take that, plus the renewable energy resources Ireland has at its disposal, there is a real opportunity to build and maintain Ireland's leading position in the global digital datasphere."

CATALYST Project releases CATALYST Green Data Centre Roadmap website

THE CATALYST project which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 768739, proudly announces the worldwide release of the "Green Data Centre Roadmap".
The website details a metro style map, an assessment tool, a handbook and a trade directory.

The Green Data Centre Roadmap
The Green Data Centre Roadmap, designed to be similar to a metro map, has 15 lines and over 130 stations, that can assist DC owners, operators, design companies and investors to consider the routes and destinations they may choose to create or retrofit their Data Centre to be as "green and sustainable" as possible. In this context "green and sustainable" means using renewable energy solutions, reusing waste heat, certified to appropriate standards and training staff.

The Green Data Centre Assessment Tool
The Green DC Assessment Tool line covers all the elements that the project considers to be required in order to be defined as a Green and Sustainable Data Centre, this includes the calculation of metrics as described in the ISO 30134 series (EN 50600 4-X series), BREEAM certification (note that this tool is designed to be used in Europe, but LEED is included on the Certification Line (but is not assessed)), participation in the EU Code of Conduct for Data Centres (Energy Efficiency) or the enhanced requirements of the Certified Energy Efficiency for Data Centres Award (CEEDA) as well as others. The result is graded Gold, Silver, Bronze, No Level and additional assistance can be obtained by dropping an email to the project team.

The Green Data Centre Handbook
The handbook chapters includes, Chapter 2 which contains a concise list of all the lines and stations contained on the Green Data Centre Roadmap.

Chapter 3 details the rationale for Green Data Centres with respect to Climate Change, Energy Security and Sustainability and includes information on the EU energy and sustainability policies.

Chapter 4 covers the Data Centre Standards Landscape

Chapter 5 details how legacy DCs can be improved by the use of the best practices contained in the EU Code of Conduct for Data Centres (Energy Efficiency).
Chapter 6 provides details on the DC new build design process, using the EU Code of Conduct for Data Centres (Energy Efficiency, the EN 50600 series of DC design, build and operate Standards in the EU, The Green Grid DC Maturity Model (DCMM) (Now EN 50600-5) and the other widely used ANSI standards, TIA 942B and BICSI 002.

Chapter 7 provides some information of integration with smart grids, smart cities, conclusions and future work on this topic.

Chapter 8 contains information on planning guidance available at EU, National, Regional and Local levels, including where appropriate enhanced requirements for potential data centre builds on campus developments.

Chapter 9 contains information on Utility Co-ordination, using information, from USEF, and Euro Heat & Power.

Chapter 10 describes information on the CATALYST business models and contractual information for relationships between the DC, Federated DCs, Utilities (Energy Flexibility, Waste Heat Reuse (District Heating/Cooling) and Smart City ECSOs.

Chapter 11 describes the "Data Centre of the Future"

NaaS growth predicted for EMEA

IN RESPONSE to COVID-19, 32% of IT leaders in EMEA plan to increase their investment in cloud-based networking, and 30% in AI-based networking, as they seek more agile, automated infrastructures for hybrid work environments.

In response to the pandemic, IT leaders in EMEA are now investing more in cloud-based and AI-powered networking technologies as business recovery plans take shape, according to research from Aruba, a Hewlett Packard Enterprise company.

The findings in a new global report ' Preparing for the post-pandemic workplace' suggested that IT leaders are responding to the challenges associated with enabling a highly distributed workforce and the emergence of the hybrid workplace – with people needing to move seamlessly between working on campus, at home and on the road – and as such are looking to evolve their network infrastructure and shift away from CapEx investments towards solutions consumed 'as a service'.

The average proportion of IT services consumed via subscription will accelerate by 41% in the next two years, from 29% of the total today to 41% in 2022, and the share of organizations that consume a majority (over 50%) of their IT solutions 'as a service' will increase by approximately 74% in that time.

The report, which surveyed 2400 ITDMs in over 20 countries and eight key industries, looked at how they have responded to IT and business demands in the wake of COVID-19, what investment decisions are being made as a result, and the consumption models now being considered. A number of key findings stood out:

Impact of Covid-19 has significant implications

ITDMs report that the impact of COVID-19 has been significant both on their employees and short-term network investments:

- 20% describing the impact on their employees as 'significant' (widespread furlough or layoffs), while 48% considered it 'moderate' (temporary reductions in some functions), and 23% 'low' (very few jobs impacted).
- In EMEA, Russia (27%), UAE (25%), Sweden and France (both 24%) ranked highest in terms of 'significant' impact with Spain (13%) and The Netherlands (15%) significantly lower.
- 74% said that investments in networking projects had been postponed or delayed since the onset of COVID-19, and 30% indicated that projects had been cancelled altogether.
- Project cancellations were highest in Sweden (59%) and lowest in Italy (11%), showing there are also significant disparities between countries within the same region, while 37% of ITDMs in education and 35% in hotels and hospitality globally said they have had to cancel network investments.

A positive outlook: Investing for emerging needs

By contrast, future plans are aggressive, with the vast majority of ITDMs planning to maintain or increase their networking investments in light of COVID-19, as they work to support the new needs of employees and customers.

- 38% of ITDM's globally plan to increase their investment in cloud-based networking, with 45% maintaining the same level and 15% scaling back. The APAC region was the global leader with 45% stating increased investment in cloud-based networking compared to 32% in EMEA nd rising to 59% among ITDMs in India. With cloud solutions allowing for remote network management at large scale, these capabilities are particularly enticing for IT teams when being on-premises is not possible or challenging.
- ITDMs are also seeking improved tools for network monitoring and insight, with 34% globally planning to ncrease their investment in analytics and assurance, 48% indicating that they will maintain their level of investment and 15% reducing it. This allows IT organizations to troubleshoot and fine-tune the network more efficiently, as demands on it are augmented by a distributed workforce.
- There is also an emphasis on innovative technologies that simplify the lives of IT teams by automating repetitive tasks. We found 35% of ITDMs globally are planning to increase their investment in AI-based networking technologies, with the APAC region leading the charge at 4% and EMEA and the Americas both on 30%.

Adoption of new consumption models is accelerating

As ITDMs shape their investment plans, they are looking at alternative modes of consumption to achieve the best balance of value and flexibility.

- 50% in EMEA say they will explore new subscription models for hardware and/or software, 51% managed services for turnkey hardware/software and 29% financial leasing – all as a result of the impact of COVID-19. This reflects the increased need for more financially flexible models in a challenging environment.
- Networking subscription models are more popular in APAC (61%) than in the Americas (52%) or EMEA (50%), and at a country level the highest demands are in Turkey (73%), India (70%) and China (65%).
- The global industries most likely to be onsidering the subscription model are hotels/hospitality (66%), IT, technology, and telecom (58%) and education (57%). The impact of COVID-19 on IT behavior has made the desire for flexibility and predictability in spending, while reducing risk from initial capital costs, greater than before.
- In stark contrast, just 8% globally plan to continue with only CapEx investments, though the proportion is higher in the Netherlands (20%), US (17%), Spain (16%) and France (15%). Across industries, 15% in retail, distribution and transport will continue to focus solely on CapEx investments, versus just 5% in education and IT, ech, and telecoms, and 2% in hotels nad hospitality.

Top 25 metros generate 65% of colo revenue

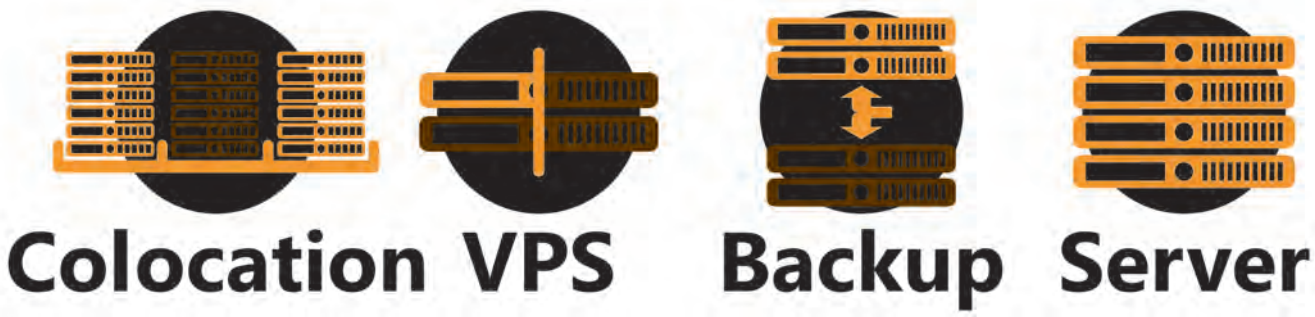
New data from Synergy Research Group shows that just 25 metro areas account for 65% of worldwide retail and wholesale colocation revenues.

RANKED BY REVENUE generated in Q2 2020, the top five metros are Washington, Tokyo, London, New York and Shanghai, which in aggregate account for 27% of the worldwide market. The next 20 largest metro markets account for another 38% of the market.

Those top 25 metros include eleven in North America, nine in the APAC region, four in EMEA and one in Latin America. The world's three largest colocation providers are Equinix, Digital Realty and NTT. One of those three is the market leader in 17 of the top 25 metros. The global footprint of Equinix is particularly notable and it is the retail colocation leader in 16 of the top 25 metros. In the wholesale segment Digital Realty is leader in seven of the metros, with NTT, Global Switch and GDS each leading in at least two metros. Other colocation operators that feature heavily in the top 25 metros include 21Vianet, @Tokyo, China Telecom, China Unicom, CoreSite, CyrusOne, Cyxtera, KDDI and QTS.

Over the last twenty quarters the top 25 metro share of the worldwide retail colocation market has been relatively constant at around the 63-65% mark, despite a push to expand data center footprints and to build out more edge locations. Among the top 25 metros, those with the highest colocation growth rates (measured in local currencies) are Sao Paulo, Beijing, Shanghai and Seoul, all of which grew by well over 20% in the last year. Other metros with growth rates well above the worldwide average include Phoenix, Frankfurt, Mumbai and Osaka. While not in the group of highest growth metros overall, growth in wholesale revenues was particularly strong in Washington DC/ Northern Virginia and London.

"We continue to see strong demand for colocation, with the standout regional growth numbers coming from APAC. Revenue growth from hyperscale operator customers remains particularly strong, demonstrating the symbiotic nature of the relationship between cloud and colocation," said John Dinsdale, a Chief Analyst at Synergy Research Group. "The major economic hubs around the world are naturally the most important colocation markets, while hyperscale operators tend to focus their own data center operations in more remote areas with much lower real estate and operating costs. These cloud providers will continue to rely on colocation firms to help better serve major clients in key cities, ensuring the large metros will maintain their share of the colocation market over the coming years."



Cloud providers 'rescue' data centre market

New Q2 data from Synergy Research Group shows that worldwide spend on data center hardware and software increased by 7% from the second quarter of 2019, thanks to a 25% jump in spending on public cloud infrastructure, which pushed it to an all-time high.

While cloud providers shrugged off the impact of the pandemic and continued to invest heavily in their data centers, enterprise spending was in the doldrums, dropping 3% from last year. In terms of market share, ODMs in aggregate account for the largest portion of the public cloud market, with Inspur now the leading individual vendor, followed by Dell, Microsoft and Huawei. The Q2 market leader in enterprise infrastructure was Microsoft, followed by Dell, HPE, Cisco and VMware.

Total data center infrastructure equipment revenues, including both cloud and non-cloud, hardware and software, were $41.4 billion in Q2, with public cloud infrastructure accounting for 41% of the total. The main hardware-oriented segments of servers, storage and networking in aggregate accounted for 75% of the data center infrastructure market. OS, virtualization software, cloud management and network security account for the balance.

By segment, Dell is the leader in server and storage revenues, while Cisco is dominant in the networking segment. Microsoft features heavily in the rankings due to its position in server OS and virtualization applications. Outside of these three, the other leading vendors in the market are HPE, Inspur, Huawei, VMware, Lenovo and IBM.

"In the middle of a global pandemic, spending on data center infrastructure was almost at an all-time high – second only to the fourth quarter of 2019. That speaks volumes about the continued robust growth in both enterprise and consumer cloud services," said John Dinsdale, a Chief Analyst at Synergy Research Group. "There was also a geographic story behind the growth. The US market grew at a good pace in the quarter, but among the larger markets it was China that was the standout performer, jumping almost 35% from the second quarter of last year."

Edge Computing spend to reach $250 billion in 2024

The concept of edge computing most often refers to intermediating infrastructure and critical services between core datacenters and intelligent end points. Proliferating enterprise and consumer devices benefit from a digitally transformed technology world through edge capabilities. However edge is defined, the compute, storage, and networking cornerstones gird data creation, analysis, and management outside of the core. A future is unfolding where extraordinary value and opportunity for essential products and services from a myriad of technology ecosystem stakeholders is being created. According to the new Worldwide Edge Spending Guide from International Data Corporation (IDC), the worldwide edge computing market will reach $250.6 billion in 2024 with a compound annual growth rate (CAGR) of 12.5% over the 2019–2024 forecast period.

"Edge products and services are powering the next wave of digital transformation," said Dave McCarthy, research director, Edge Strategies at IDC. "With the ability to place infrastructure and applications close to where data is generated and consumed, organizations of all types are looking to edge technology as a method of improving business agility and creating new customer experiences."

The stable of companies that are present and investing in edge computing continues to grow and increasingly includes a diverse set of competitors. Familiar companies in the hyperscaler space include Amazon Web Services (AWS), Equinix, Google, IBM, Microsoft, Oracle, and Switch among others. Physical infrastructure providers include companies such as AMD, Dell Technologies, Ericsson, HPE, and Intel. Meanwhile, services companies like AT&T, Lumen, and Verizon deliver critical networking capabilities to connect the thousands of planned and deployed edge datacenters.

IDC expects edge expenditures will be concentrated in the U.S. and Western Europe over the next several years. In 2020, the global regional spending shares for the Americas, EMEA, and Asia/Pacific will be 45.0%, 27.9%, and 27.2%, respectively. From an industry perspective, 11 of the 19 standard industry segments will deliver 5% or more of total worldwide spending in 2020.

The top two industries for edge spending throughout the forecast are discrete manufacturing and professional services, while retail will overtake process manufacturing to become the third largest industry by the end of the forecast. Professional services will also see the fastest growth in edge spending with a five-year CAGR of 15.4%.

From a technology perspective, services (including professional and provisioned services) will account for 46.2% of all edge spending in 2024. Hardware follows as the second largest technology category with a 32.2% share of spending, while the remaining 21.6% will go to edge-related software.

"While no technology market has been spared from the economic impact of COVID-19, edge market suppliers are poised to experience sustained growth throughout the forecast from enterprise and service provider investments," said Marcus Torchia, research director with IDC's Customer Insights & Analysis group.

How will AI advance the data centre industry?

Artificial Intelligence (AI) is touching almost every industry in one way or another. Through the power of AI, organisations can process and analyse data faster than humanly possible, enabling smarter decision-making with data-driven insights.

BY OLIVER GOODMAN, HEAD OF ENGINEERING, TELEHOUSE.

AS THE data centre industry becomes increasingly crucial to the everyday activities of consumers and organisations alike, there are growing opportunities to use AI to manage energy consumption and distribute workloads more efficiently while enhancing cybersecurity measures.

Energy efficiency is one of the most important areas of focus for data centre operators. While cost, capacity and connectivity are the three most obvious areas organisations will look for when choosing a data centre provider to work with, their environmental footprint is a growing consideration. In its digital strategy, the EU Commission notes that data centres are responsible for a significant environmental footprint, and states they “can and should become climate neutral by 2030.”

With consumers and organisations becoming increasingly concerned about their own environmental footprints, it’s likely that organisations that rely on data centres will want to see evidence of data centres operators taking positive steps toward reducing their energy consumption.

Driving energy efficiency

The infrastructure supporting a data centre is where AI can have the biggest environmental impact. Intelligentcontrol systems have been in place to drive efficiency in power and cooling infrastructure for a number of decades and the ongoing development of these systems means they are continuously getting smarter.

Cooling is essential to preventing servers from overheating. This can lead to servers shutting down and causing downtime for users which can be costly for customers. Overheating can also lead to loss of valuable data and equipment, not to mention reputational damage for both the data centre operator and the customer.

To achieve sufficient levels of cooling for the servers, datacentres are temperature-controlled environments where the temperature needs to be maintained around the clock. This requires a lot of energy and data centre operators are challenged to continually reduce energy consumption. One way around this is to build datacentres in colder climates, such as countries in the Nordic region where the ambient temperature is a lot lower. However, high bandwidth services (such as Netflix or Youtube) and the rise of edge networks (for example for Internet of Things or connected car technology) require data to be stored closer to the devices that are consuming it means there will be a growing requirement for regional data centres across even developed countries.

Since not all locations will benefit from a cold climate, this presents the challenge of running the most efficient solution for the local environment and this is where vast data collection, real time data analysis and AI will bring significant gains. Furthermore, as global extreme temperatures are on the increase and as weather patterns are becoming more erratic, Intelligent cooling systems are connected to networks of sensors throughout each data centre, measuring the performance of the customers’ hardware as well as the ambient conditions within and external to the building. The cooling is intelligently modulated to suit the conditions within the data centre and as these systems become increasingly intelligent, they are working harder and faster at achieving the best possible cooling efficiency.

As the AI capabilities of these cooling systems increase, data centre environments can be controlled with more granularity, giving them better visibility

and control over the complex systems. Predictive modelling will also come into play, using machine learning to forecast when and where temperature fluctuations will happen and dealing with those fluctuations as efficiently as possible. As the networks of sensors grow within each data centre. This will be particularly beneficial for colocation providers who aren't in control of customer equipment, the customer load, or the temperature fluctuations of those loads.

AI will also begin to play a role in managing the lifecycle of equipment and parts in data centre infrastructure. Predictive failure and maintenance will calculate when work needs to be carried out on each piece of equipment and when parts are likely to wear or fail. Not only will this safeguard the efficiency of each component in the data centre infrastructure, it will also prevent downtime caused by equipment failure.

Improving cyber security

Ensuring robust cyber security is one of the most crucial challenges for data centre operators. In 2019, 32% of businesses in the UK experienced a cyber security attack or breach. As cybercriminals continuously innovate to find new means of accessing sensitive data, cyber security will continue to be a growing challenge.

The monitoring of data centre network traffic is one way operators can ensure resilience against cyber threats. There are robust solutions that are beginning to appear on the market that will leverage AI and machine learning to monitor network traffic intelligently by learning the typical patterns of network activity and looking for anything that falls outside of that typical behaviour. These solutions look for changes in typical traffic behaviour at specific times of the day or analyse the size of data packets that are being transmitted across networks. Any anomalies that are detected will result in the automatic shutdown of ports on the networks where those anomalies are identified.

An example scenario would be an employee transmitting large amounts of data outside of office hours. The AI built into the cyber security system would detect that as an unusual pattern and flag it as an issue within a matter of seconds, potentially shutting down the ports through which the data is being transmitted. This entire process will take place at a speed that is simply impossible for a human to match. In fact, most organisations are not effectively monitoring their network activity with a view to ensure network security, rather their response is reactive to a problem that has already occurred meaning that the damage might already be done. With this functionality now appearing on the market, AI will quickly begin to add an additional layer of robustness to data centres from a cyber security perspective.

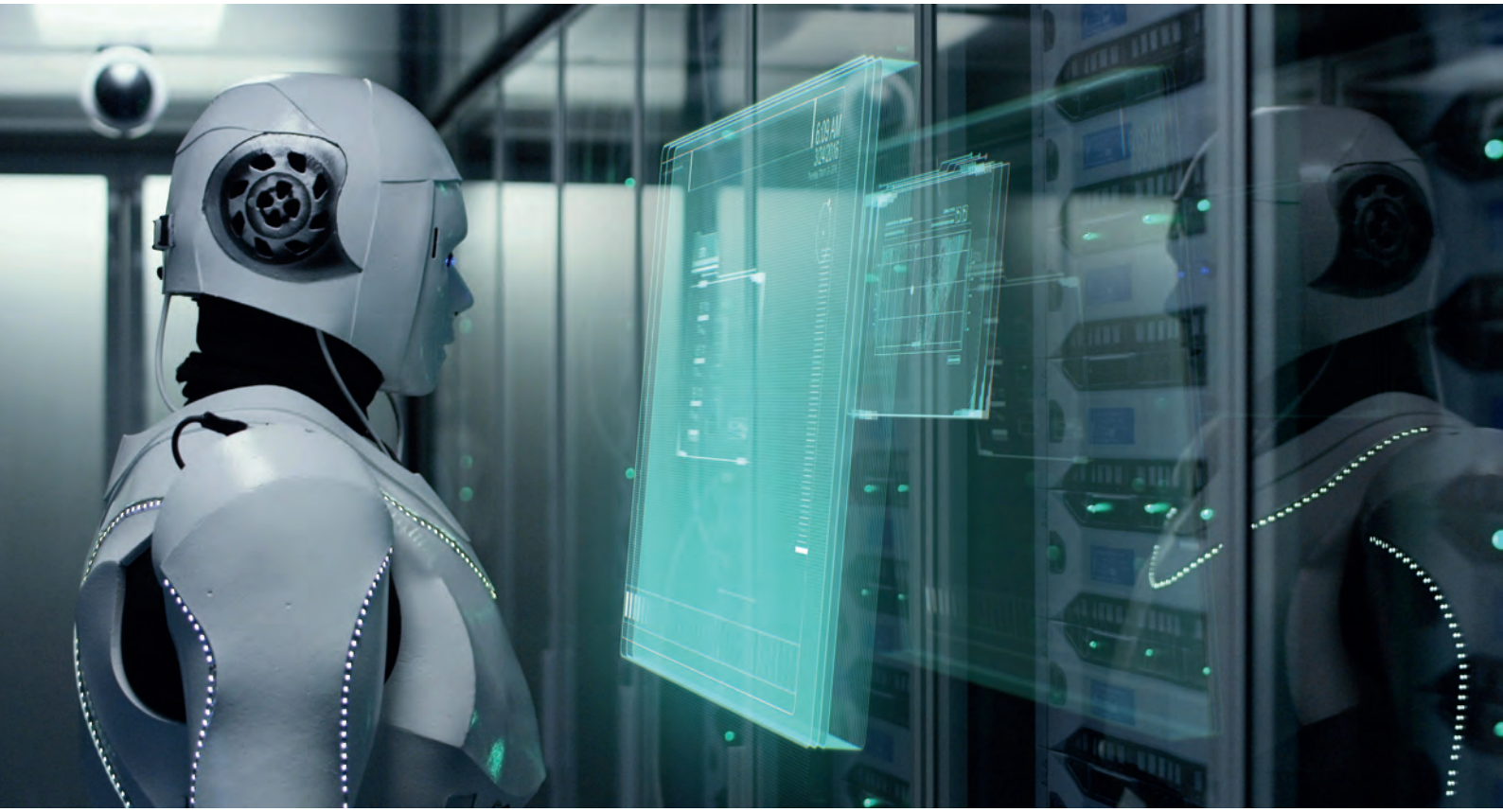
Distributing workloads more efficiently

The distribution of workloads across a data centre operator's different locations is managed by weighing

up various factors relating to business objectives. These factors can include the cost of electricity and carbon footprint of each data centre. Deciding which data centre is most suitable for each workload helps the operator to balance the whole load across its network of locations. Where the goal is to keep electricity costs as low as possible, a workload can be migrated to a specific data centre where the electricity is lowest at that exact time.

There is a significant amount of workload management that takes place among larger data centre customers to make sure that they're using server space in the most efficient way possible. It's likely we'll see AI and machine learning playing a bigger role in performing the analysis, making the decisions and transferring the data instantaneously. Different data protection laws across different regions can be an issue when transferring data from country to country. AI will likely begin to factor those regulations into its decision-making, reducing yet another pain point in the distribution of workloads.

As AI-driven solutions continue to evolve, we will certainly see more intelligent decision-making taking place in various areas of data centre management. We're seeing AI give data centre operators the ability to constantly improve their service offering by analysing data in real time and deliver split-second decisions that can only be performed by computer. Energy efficiency, cyber security and workload distribution are where we can expect to see this happening soonest, but in the longer term it will be fascinating to see how AI will improve the industry in different ways.



Automation and intelligence: maximising your data centre's efficiency

The immense growth in data and the speed at which businesses operate today means manual monitoring, troubleshooting, and remediation is too slow to be effective and can put businesses at risk.

BY ANDREAS LIMPAK, DIRECTOR OF SOLUTIONS ENGINEERING, NETAPP.

DATA CENTRE AUTOMATION is central in tackling these challenges, yet implementing and successfully managing it can be tricky. While the end result is a more efficient and resilient environment, managing this level of infrastructure differs greatly from 'traditional' infrastructure, so having a well thought out, data and cloud strategy is key.

Operations and infrastructure leaders need to build and sustain dependable infrastructure; their hardware, software and their staff. It's crucial to expand infrastructure and operations skills, procedures and practices to facilitate hybrid IT operations and to develop a hybrid data placement strategy to increase the business outcome. Besides that, enterprises need to start to create an inventory of assets of the infrastructure, components, providers and processes across the whole IT landscape. This is fundamental to begin the process of automating simplifying and standardising IT operations.

So what does data centre automation actually entail? It is the process by which routine workflows and processes of a data center—scheduling, monitoring,

maintenance, application delivery, and so on—are managed and executed without human administration. This has a knock on effect of increasing both business agility and operational efficiency. Furthermore, it reduces or even eliminates the time IT staff needs to perform routine tasks, enabling them to deliver services on demand in a repeatable and automated manner.

More specifically, data center automation is immensely valuable because it:

- Delivers insight into server nodes and configurations
- Automates routine procedures like patching, updating, and reporting
- Produces and programs all data center scheduling and monitoring tasks
- Enforces data center processes and controls in agreement with standards and policies

Getting IT ready for automation

With huge dependency on a diversified IT environment, data center operation teams strive to ensure critical applications run smoothly without impacting business. Having an appropriate, established data fabric strategy allowing for the access, understanding and usage of disparate data nodes across the business' entire IT landscape is vital in taking data centre operations to the next level while keeping operations smooth. An enterprise's infrastructure is, in short, both the lock and key for automation and the roll out of intelligent services across data centres.

Once this data fabric strategy is in place, the first step in data centre automation is monitoring. Monitoring is the bedrock of almost all automation. Healthy data centres translate to better return on investment (ROI) for businesses. To achieve this ROI, we would recommend a centralized monitoring solution that can consume information from any modern-day infrastructure on the data center floor. Integrating all this data from multiple devices and multiple vendors is the key to building a centralized monitoring dashboard.

Ideally, the data center provider (vendor) would have API access to the infrastructure, enabling it to inter-operate with public clouds so that customers can migrate data or workloads from cloud to cloud. Every infrastructure operations leader will also have combining automation capabilities in cloud or hybrid cloud environments with Service Management on their agenda.

The integration of automation and infrastructure as code (IaC) is made easy by using code as machine-readable files that can be processed by ITSM frameworks. IaC reduces manual processes and optimises operational efficiency.

Rather than doing physical hardware configuration or single interactive configuration tools, all managing, provisioning, change and decommissioning in the datacenter can be done by definition files. The final, managed IT infrastructure will compromise of physical as well as virtual components and configuration resources, with a version control system maintaining all configuration changes.

Enterprises who have already begun their cloud journey will opt for server virtualization within their data centres. According to recent research, more than half of organizations plan to use storage virtualization and application virtualization by 2021. We will soon see a fully virtualized software-defined data center architecture which will play a key role in provisioning of IT resources and applications.

Keeping these services optimized in line with compute, storage and access requirements will also be key in keeping down operating expenditure. Automation of this provisioning, or at least having the capabilities to automatically assess currently workloads, will become an increasing priority as business look to scale while juggling SLAs.

Layering on the intelligence

Automation in itself is a feat to be proud of, however, the next step is moving from automation to intelligent and predictive flagging and/or remediation, and full-blown AI solutions. These tasks require unprecedented levels of computing power and large amounts of scalable storage. An explosion of data results in deep learning models that take days or weeks to train. Computing nodes, storage systems, and networks often are unable to handle the data volume, velocity, and variability of AI, ML, and DL applications at scale.

And do-it-yourself integrations are hard.

A simpler option is to look for a solutions provider who can provide white-labelled, easy to install converged infrastructure with scalable AI built in. By placing this in within the data centre racks it almost immediately enables the movement of data from its collection or storage location to the compute engines at high speed. With the right solutions, IT teams can then easily create development and test environments for sandboxes and configuration testing before rolling out system wide, at scale. Perfect for ensuring the smooth running and health of datacenters.

AI's potential in the data center is nearly limitless, and we have only scratched the surface. Increased adoption of AI will enable data centers to meet the most challenging SLAs, bring down operation costs, protect uptime – and even ensure better client servicing by reducing the number of incidents that require escalation. But it all starts with the data strategy, the birds-eye view of your infrastructure, and having the foresight to set IT teams up for future success.

Digital transformation in the hosting sector

Experts agree that, despite the current crisis, organisations will continue spending on digital transformation (DX) of business practices and products at a solid pace in the face of challenges.

BY ELTJO HOFSTEE, MANAGING DIRECTOR AT LEASEWEB UK.

OVERALL, this is good news for the tech sector, however, companies that form the basis of the digital infrastructure must now take action to anticipate new rules, and meet new customer expectations.
A recent study by Eco and Arthur D. Little outlines COVID-19's impact on the German internet sector. The study distinguishes four internet sections:
1. Network, Infrastructure and Operations,
2. Services and Applications,
3. Aggregation and Transactions and
4. Digital Business Models in the User Industries.
The study shows that companies in the B2C and B2B eCommerce, online advertising and payment platforms (part of section three) will experience the negative effects of the crisis this year. Thereafter, companies in this section will resume their annual positive growth. The specific group of companies in section 4 (Digital Business Models in the User Industries) that will suffer as a result of the crisis for a long time are in the transport, tourism and retail sectors.

Greater digitalisation

In contrast, companies in section four that offer services such as streaming and online gaming have already seen their turnover increase. A positive effect

is also being experienced in section one due to increasing investments in cybersecurity and cloud services. The report goes on to outline that through 2025, the German internet sector as a whole is growing at an average 9.5 per cent year-on-year. We can expect comparable growth for the UK: the compound annual growth rate over the past five years has been 11 per cent. If we add the pressures brought about by COVID-19, which has seen schools, hospitals and offices make massive strides forward with their digitalisation, it is logical to predict that this growth will in part be structural, resulting in further digitalisation.

This movement also means that the interests of the internet sector are growing. The functioning of entire business columns is dependent on a sound infrastructure, which at the same time is becoming increasingly complex. Service failure means en-masse disruption to the operation of entire organisations, and potentially even societies. This challenge is for hosting and cloud providers to meet head-on. Now more than ever before, services must be delivered redundantly and cloud service providers must be able to quickly fix problems, or better yet, prevent them.

Limited access to data centres

An additional problem is attempting to access data centres in our current climate. Even when a company's specialists are physically close to the data centre, they are unable to access the hardware to do recovery work. In so-called 'lights-out' data centres, there is no issue at all to enter the physical location. The consequence of this limited access to a data centre with the increased importance of digitised services is twofold. Firstly, more work will be done remotely – for example, the configuration of servers. Secondly, far more automation will have to be done within the internet service provision to guarantee the quality of the service and to be able to scale up.

Automation (the automation of processes and actions) has become increasingly popular in recent years due to the growing complexity of IT environments, in combination with the increasing dependence on the same infrastructure and underskilled IT staff. The need for automation, however, grew enormously at the beginning of this century. Server virtualisation revolutionised computing resources at the time; five to ten applications could run on one server, but they all had to be managed. Today we see that complete stacks, including network connections, can be monitored and adjusted almost completely automatically.

This situation is already a reality with the hyperscalers. We only need to look at how Netflix and Google lowered the resolution of their video streaming to use less bandwidth, or how Microsoft turned off some features in Windows 365 to create more space for Teams and other services. These interventions were only possible thanks to a very high degree of automation. Many parties in the (cloud) hosting segment will now have to take steps to make their services cost-efficient, scalable and more resilient. This is because, although the market for digital services will be much larger in 2025 than it is now, customers will be more demanding than ever, and service providers will have to quickly anticipate external disruptions.

Why 5G can't come soon enough for many operators

To reap the benefits of real-time feedback, automated orchestration – powered by AI – will become all but essential for operators and their data centres

BY SUDEEPTA RAY, VP OF TECHNOLOGY AT ALTRAN, NOW PART OF CAPGEMINI.

THE FIRST commercial launch of 4G LTE took place in 2009. At the time, it was regarded as a game changer. Speeds of 1GB/s suddenly opened the door to mobile video conferencing, on-the-move gaming and high-definition video streaming. The likes of Uber, Netflix and Airbnb were at the right place at the right time – thanks to 4G. But despite its many advantages, 4G LTE coverage has remained quite patchy.

Fast forward to the present: 5G can't come soon enough. 5G is 10 times faster than 4G. It can power use cases liker remote surgery, smart cities and automated cars, where lightning-fast speed is essential. 5G offers much more than speed, though. It's going to completely revolutionize the way services are delivered. And the sheer volume of applications and devices will require a new strategy for data centers and a need for automated orchestration coupled with AI for enhanced customer experience.

Evolving data centers and orchestration

5G applications have extremely strict requirements in terms of latency and bandwidth, so existing data center architecture must evolve. What's needed are multi-tiered data centers consisting of a centralized main data center, a regional one and one at the edge. That is why operators need automated orchestration to intelligently handle and transfer data swiftly and securely through a complex web of apps and servers. With 4G, automation was limited and services would often take two to three months to roll out. The sheer volume of applications and services that will be active on a 5G network makes this deployment timeframe practically unfeasible.

With 5G comes operator freedom

While the majority of 4G services were based on vendor-provided solutions, 5G opens things up like never before. The so-called "vendor-provided integrated stack" will become a thing of the past, and vendor lock-in will be no more thanks to the push for a cloud-native architecture and open interfaces.

Operators will be able to roll out services faster and also have a precise view of what's happening on any part of their data center. This is another reason

There are still some missing pieces of the puzzle when it comes to security and 5G, and that's largely due to the fact that we're ll still learning about the technology. Where there is connectivity, there are always security risks, so operators are going to have to keep refining their security solutions if they want to monetize 5G quickly

why automated orchestration will become all but essential for operators to reap the benefits of real-time feedback, by leveraging AI for continuous improvement for network efficiency. It a small price to pay for the freedom and reduced operating costs that will inevitably follow.

Some operators are already taking huge leaps in the right direction, evolving their architecture and infrastructure from 4G. Take Telefonica, whose end-to-end virtualization program, UNICA Next, is already maturing its operational model to take full advantage of 5G and zero-touch services that require minimal human input.

Orchestration challenges

There are still some missing pieces of the puzzle when it comes to security and 5G, and that's largely due to the fact that we're all still learning about the technology. Where there is connectivity, there are always security risks, so operators are going to have to keep refining their security solutions if they want to monetize 5G quickly.

Similarly, there is still room for improvement when it comes to the technology and processes that enable automated orchestration. For example, determining and maintaining the appropriate distance between the edge and far edge is likely to become a challenge for many operators in their efforts to monetize 5G.

More distance will mean that anomalies take longer to detect, and while this can be remedied with more controllers and servers at the edge, this route risks making the process commercially unviable.

The ultimate goal for operators and data centers in the wake of 5G will be to have open, multi-tiered architecture from multiple vendors, with full interoperability as the standard. Thankfully, this is likely to get easier as pressure mounts on vendors to adapt to a network landscape where customer lock-in is a thing of the past.

Operators should be poised to take full advantage and leverage their data centers to monetize 5G faster. No wonder 5G can't come soon enough for many operators.

Driving edge resilience with AI, machine learning and digital services

Digital transformation initiatives have become a key point of differentiation for organisations looking to gain a competitive business edge. However, this not only requires a dramatic rethink of the internal systems and operational processes, but of the critical infrastructure systems that underpin them.

BY MARC GARNER, VP, SECURE POWER DIVISION, SCHNEIDER ELECTRIC UK&I.

FROM A BANDWIDTH, latency and regulatory perspective, IoT and edge computing have become fundamental components in the success of digital transformation strategies. Recently, the technology has begun to manifest in many recognizable forms including self-service kiosks in supermarkets and fast-food restaurants, automated robots in smart manufacturing or industrial factories, and as demands increase, this sector continues to anticipate growth. 451 Research, now part of S&P Global, estimates the size of the edge computing market to be nearly $170bn, growing to $651bn by 2024.

Edge computing, or distributed IT, as it's otherwise known, streamlines the delivery of digital services by pushing the infrastructure out to the edge of the network, or the point where data is generated, processed and consumed. According to IDC, 46% of organizations that were planning edge deployments in 2019 were concerned about their ability to monitor and maintain equipment remotely.

Moreover, new research published in the 2020 Uptime Institute Data Centre Survey found that data centre outages continue to occur with disturbing frequency, and of the organisations surveyed, 75% stated that downtime was preventable with better management, processes or configuration. One could argue that security, resiliency, visibility and uptime have become ever more crucial concerns for data centre and edge computing operators, no matter the size of the system.

IoT and the edge

As the Internet of Things (IoT) accelerates the volume and type of local devices capturing data, edge computing systems are required to process and analyze that digital information more quickly, in turn helping to drive more informed business decisions. IDC predicts that by 2025, the number of devices connected to the Internet will grow to 80 billion, so while IoT facilitates greater information gathering and distribution by enabling network connections on electronic devices, this combination of edge and IoT ensures data traffic is kept in small local loops, leading to faster, more efficient and secure interactions.

For many organisations a potential Achilles Heel is the challenge of managing and maintaining these small, dispersed edge data centres, which are being deployed in geographically remote areas of the digital ecosystem.
Many will assume that edge sites play by the same rules as larger data centres, and from a resilience perspective the expectation is very much there – applications at the edge are often business critical. However, edge environments lack one thing that larger facilities have in abundance – skilled IT professionals on site.

The fact remains that at the edge, having personnel stationed at every single deployment is just not possible. It would be prohibitively expensive, and probably impractical to have IT professionals permanently placed in every location.

So how do end-user organisations overcome this challenge when, from a business and a consumer's point of view, no downtime can be tolerated? The answer lies in new developments within Artificial Intelligence (AI) and machine learning (ML) software.

The role of AI in edge data centres

Today cloud-based data centre infrastructure management (DCIM) platforms leverage IOT, AI and ML algorithms to aggregate and analyse real-time data into the health and status of distributed IT systems. They allow data centres and edge computing environments of any size and complexity to be monitored and managed from a single platform; collecting vendor-agnostic performance information from critical components including racks, servers, power distribution units (PDUs), cooling equipment and uninterruptible power supplies (UPS).

> Armed with such insights, these AI-enabled remote monitoring platforms become more proactive and in many cases predictive. Faults or capacity issues can be anticipated before they become a problem, timely maintenance or upgrades can be planned to optimise the use of resources and edge computing sites can operate more efficiently

Using the capabilities offered through AI and the cloud, edge data centres can be monitored in real-time, ensuring that proactive maintenance schedules are followed, faults are identified before they become problems, and emergency issues can be flagged - with service personnel dispatched to resolve them as a matter of priority.

Such AI capabilities allow better decisions to be made, using key learnings from recognised data patterns to identify issues that may impact performance, incur risk of malfunction and help operators to minimise downtime.

However, the sheer volume of data being collected can threaten to overwhelm those who must process and allocate resources on the basis of priority.
Such "unstructured" data is difficult to assemble, to understand and process using simple tools, so here the role of automated data analysis plays a crucial role; identifying patterns of behaviour, which can yield

deeper insights into how a system is functioning, or can be expected to perform in the future.

Armed with such insights, these AI-enabled remote monitoring platforms become more proactive and in many cases predictive. Faults or capacity issues can be anticipated before they become a problem, timely maintenance or upgrades can be planned to optimise the use of resources and edge computing sites can operate more efficiently.

Optimising performance via dispatch and servicing
From the point of view of servicing, comprehensive data analysis is driving a systematic change from monthly or annual check-ups to condition-based maintenance.

No more will service personnel upgrade a UPS battery just because a certain period of time has elapsed, rather, using a more detailed analysis of how hardware performs in the real world, based on performance and environmental conditions to implement servicing. While some components will of course need to be replaced if nearing end of life, this offers a greater appreciation of whether upgrade or repair is actually necessary, and is a more efficient use of service personnel.

No one vendor can provide all the technologies and services necessary to exploit the benefits of AI and machine learning, and the skills of expert IT professionals still play a critical role in driving resilience at the edge.

Regular communication and interaction between vendors, partners and customers is essential, and here the addition of public Application Programming Interface's (APIs) allows for greater integration between cloud-based DCIM systems and third-party management platforms.

In this way, the benefits of AI and machine learning can proliferate through all areas of a data centre, further improving the accuracy of the insights gained and the models of how a system behaves to offer greater uptime to the end-user.

As outages increase, visibility, management and greater resilience remain key issues for data centre operators. Thankfully the AI capabilities offered through vendor-agnostic, next generation data centre infrastructure management software offers a simple way to increase reliability, security and availability at the edge.

AIOPS SOLUTIONS ONLINE SUMMIT

ENABLING APPLICATION OPTIMISATION

The importance of proactive performance monitoring and analysis in an increasingly complex IT landscape.

25 November 2020

aiopssolutions.com

INNOVISION2020
DATA CENTRE INSIGHTS + PERSPECTIVES

InnoVision: A very special issue of DCS Magazine dedicated to the data centre industry's visionary leaders and technology innovators

To herald the launch of the all-new Data Centre Solutions digital publication, we have produced a very special first issue, entitled InnoVision – providing an overview of the state of the data centre industry right now.

80+ Vendors from across the supply chain have provided their viewpoint on the future and innovation.

How will the data centre industry evolve over the coming months and years, what will be the major drivers and opportunities?

Read today

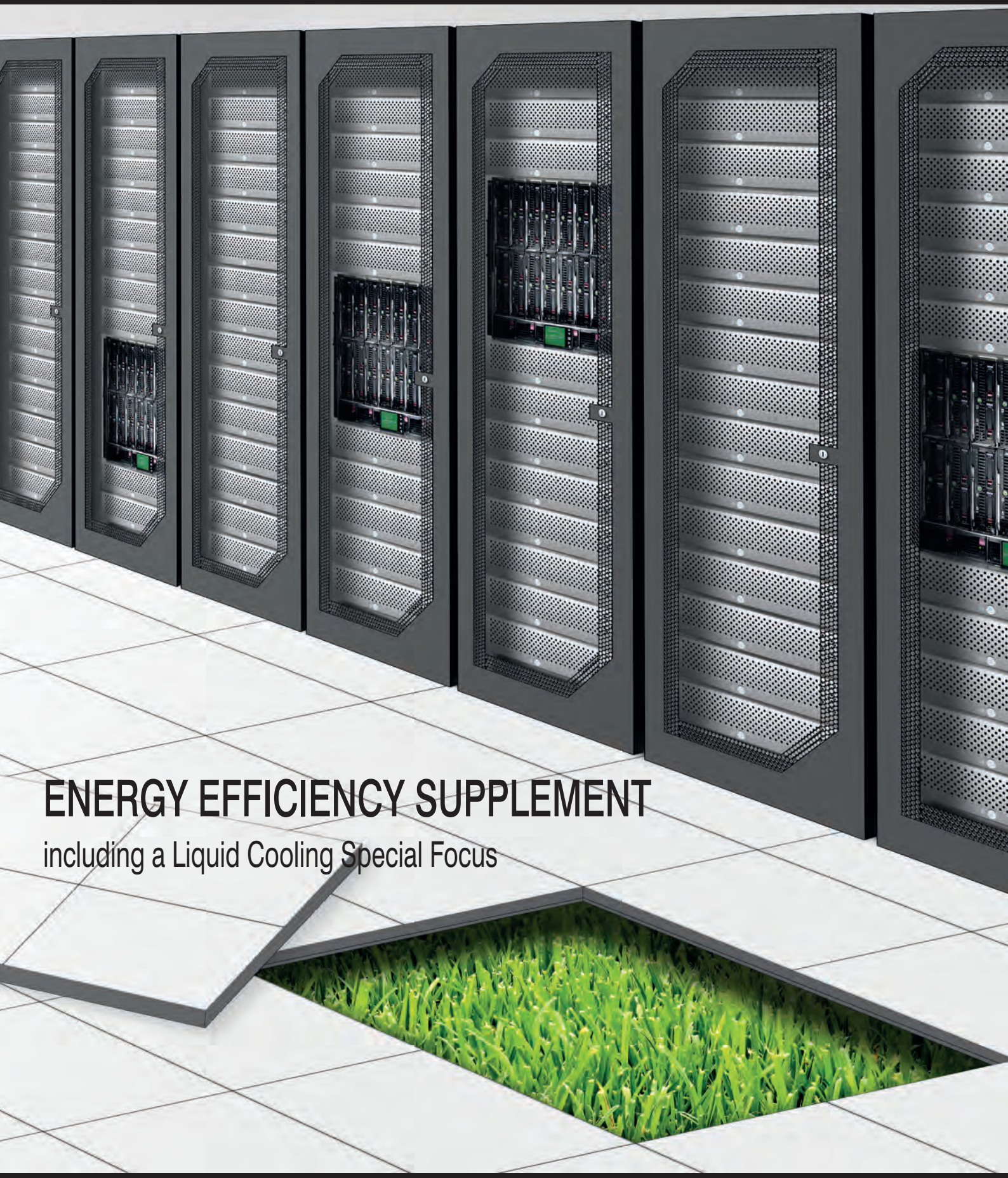
https://digitalisationworld.com/magazines

IN ASSOCIATION WITH

DATACENTRE SOLUTIONS

DIGITALISATION WORLD

DW DATACENTRE SOLUTIONS



ENERGY EFFICIENCY SUPPLEMENT

including a Liquid Cooling Special Focus

AI + Automationps | ArtArchitecture + Connectivityps | ArtCloud + Managed Services
Colocation + Hostingps | ArtDesign + Infrastructureps | ArtEnergy Optimisation
Management + Maintenance | Physical Security | Power + Cooling

Energy optimisation and the path to good grid citizenship

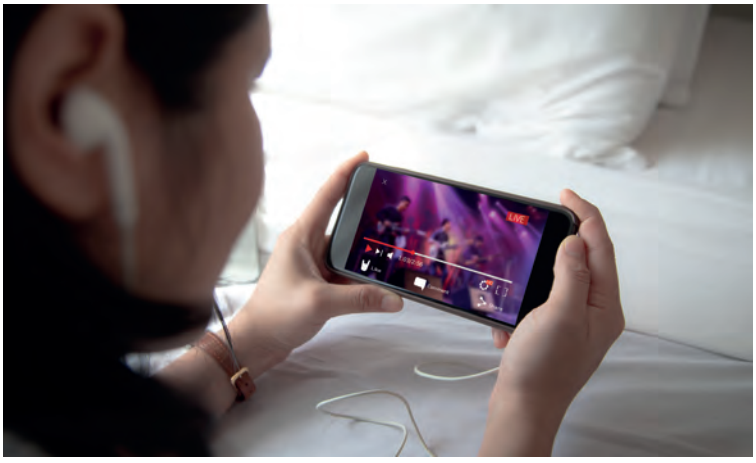
Data centre operators are looking for opportunities to reduce energy costs and are finding that their ambitions can be realised by proactively engaging with grid operators and becoming better grid citizens.

BY ANDREW TOHER, HEAD OF CUSTOMER INSIGHTS, ENEL X UK

THE GLOBAL PANDEMIC and resulting lockdown demonstrated society's growing dependence on data to enable our work and personal lives; whether due to home working or Netflix viewing, an increase in the use of video streaming caused a 40% surge in data between February and April 2020.

Data centres are on a long-term growth trajectory and the recent data surge looks like it's here to stay. Gartner predicts that 30% of the workforce will be working from home in 2024, and over two-thirds of businesses are actively working on digitisation projects to optimise office/home hybrid working models. And, according to Huawei Technologies, data centre electricity use is likely to increase about fifteen-fold by 2030 to 8% of projected total global electricity demand.

Despite this exponential growth in data and workloads, data centres have done an exceptionally good job of managing energy demand, which has increased only modestly over the past 10 years. This is partly down to better energy efficiency measures in smaller data centres, but an increase in the use of highly efficient hyperscale data centres has gone a long way to limit overall energy demand. These facilities use servers designed for purpose – they are functional and efficient, stripped of even video connectors and blinking lights as no humans walk the racks to see them.



While data centres account for 1% of the world's energy demand, the sector is responsible for only 0.25% of emissions and should be congratulated on this outcome. But grid operators should also be acknowledged for their work to maintain stability and supply to meet growing data centre needs in a rapidly changing energy landscape. Both parties continue to adapt and grow in order to maintain stability and harmony within the electricity system. Data centre operators have tackled the most achievable energy efficiency savings by optimising servers and cooling equipment. But to become even better grid citizens they must work in partnership with the grid to help keep their cooling systems running and the lights on for the rest of us.

Pathways to low-carbon energy

There are many pathways to low-carbon energy, and data centre operators have to decide what strategy is right for them. The range of options vary from simple contractual agreements (such as Guarantees of Origin) with green energy suppliers to more active involvement in the operation of the energy grid.

At the passive end of the decarbonisation spectrum, energy suppliers use Renewable Energy Guarantee

of Origin (REGO) certificates to calculate their Fuel Mix Disclosure and demonstrate how green their tariffs are. However, suppliers can trade REGOs separately from the unit of electricity that came from the renewable asset and attach them to fossil-fuel generators in order to pass them off as green. Critics of REGOs also argue that they fail to incentivise the building of new renewable sources of power.

Growth in renewable energy is set to continue as countries pursue net-zero carbon targets. According to the International Renewable Energy Agency (IRENA), over 75% of onshore wind and 80% of solar PV project capacity is due to be commissioned in 2020 without subsidies and should produce cheaper electricity than any coal, oil or natural gas option.

With renewables growing as a proportion of the overall energy mix, grid operators are finding it more difficult to modulate supply to meet demand. Instead, they are looking to increase the flexibility of the grid through a number of strategies – either storing, shifting or transporting electricity.

Demand Side Response as a resilience strategy
Using time or price triggers to shift demand away from peaks, by implementing demand side response mechanisms, has proved to be a successful flexibility strategy for grid operators for several years.

Demand Side Response (DSR) is a good fit for data centres but there are still questions around how participation works with some business models. For example, co-location data centres operate within strict customer service level agreements, and operators are understandably cautious about adopting measures that are perceived as a threat to uptime.

On the other side of the debate, some co-location data centre operators have actively embraced DSR. They see participating in DSR as part of their resilience strategy as they receive advanced notice of grid problems and can prepare for trigger events or pre-emptively move to back-up power for the duration of the grid instability. On-load testing that more accurately replicates the utility fail risk is another benefit of working with a partner to participate in flexibility schemes.

As well as being a means to more robust resiliency measures, DSR provides data centres with capacity payments simply for being on standby; valuable income to help offset energy costs.

Renewable energy leadership

The industry is showing strong leadership on corporate renewable procurement; the top four corporate off-takers of renewables in 2019 were all ICT companies. At the hyperscale end, heavy energy

A potential solution to address the funding barrier is to work with a specialist partner on an 'as-a-service' model. Energy-as-a-Service, or EaaS, helps to overcome the issue of having to find capital to fund improvements and forges a long-term relationship with a partner who can advise and deliver on PPAs, flexibility solutions, energy efficiency measures, utility bill management and so on

use is driving them closer to grid operators – often by working with partners to develop strategies that enhance their corporate reputations.

Data centres operating across all business models are looking to Power Purchase Agreements (PPAs) which enable off-takers to procure long-term contracts with operators of renewable assets. However, negotiating PPAs can be technically complex. Some key PPA parameters include the term of the agreement; whether the PPA is a corporate arrangement, includes a private wire and/or storage; how risk is allocated between procurer and generator, including the volume risk. Optimising these parameters to deliver a bespoke agreement that suits both generator and off-taker requires depth of knowledge and experience.

Data centres should consider the integration of renewables into electricity grids. Today, more than ever, grid operators need active participation from their largest energy users. This requires businesses to act as good grid citizens and embrace a degree of flexibility in their energy usage that can be used to support network operations but does not negatively impact on business operations.

Google accelerates renewable energy purchase through auctions
In 2018, Google matched 100% of its global electricity consumption with renewable energy for the second year in a row. Looking to the future, Google recognised that sustaining a 100% match would require thinking beyond its historical procurement methods. To continue meeting its users' needs in a sustainable way, it decided to streamline its renewables procurement process by running reverse auctions (where energy sellers bid for a buyer's business) for wind and solar projects. Google's goal was to find a way to source, negotiate, and sign a large wave of renewable energy deals in a single, global push.

Google used Enel X's proprietary reverse auction technology to support its industry-leading commitment to sustainability. As a result of the auctions, Google signed 10 agreements comprising more than 1.2 gigawatts of renewable energy. Running reverse auctions accelerated its procurement process and allowed it to meet its procurement cost goals.

http://bit.ly/Googlecasestudy

Broader energy initiatives

Energy management initiatives extend beyond the matter of supplying power to the data centre itself. Increasingly, businesses are seeking holistic approaches to manage their energy needs. For example, as the use of electric vehicles grows, workplaces are integrating charging infrastructure for employee and visitor use. Smart EV charging can play a role in grid balancing by integrating these more flexible, non-critical loads into an overall energy efficiency plan.

A holistic approach to managing energy typically incorporates efficiency measures, alongside other initiatives such as DSR, PPAs, EV infrastructure and even utility bill management, which provides detailed insights into energy spend that can inform planning.

While major energy users see the benefits of planning a broad approach to energy strategy, finding capital to fund the measures can be a barrier to moving forward. A potential solution to address the funding barrier is to work with a specialist partner on an 'as-a-service' model.

Energy-as-a-Service, or EaaS, helps to overcome the issue of having to find capital to fund improvements and forges a long-term relationship with a partner who can advise and deliver on PPAs, flexibility solutions, energy efficiency measures, utility bill management and so on. As well as monetising the flexibility of energy assets and reducing costs, EaaS enables profitability, improved resiliency, sustainability and better risk management – especially with respect to compliance and market exposure.

Driving grid innovation

As large energy consumers and one of the fastest growing users of power, data centres have the potential to make a tremendous impact to grid innovation. Given the data and power industries' interdependence, many now advocate that it's time for data centres to contribute to the overall stability of electricity systems by becoming better grid citizens.

Multiple strategies are available to meet different business models and energy demand profiles, enabling a business to meet its decarbonisation goals while bolstering profitability and resilience.

CELEBRATING 11 YEARS OF SUCCESS

Announcing the 11th edition of the premier IT awards: The Storage, Digitalisation + Cloud Awards 2020.

In what has been, and continues to be, extraordinary times for the business world, it seems doubly important to recognise the projects, innovations and individuals which have made such a huge difference during 2020. Almost overnight, employees switched from office working to working from home, and the new, or next, normal, means that, into the future, what might be called a 'hybrid work' model looks set to evolve, with flexible working very much the order of the day. What was already becoming a trend as part of many organisations' digital transformation programmes, has been accelerated.

The SDC Awards 2020 will celebrate the achievements of end users and the IT community as they have innovated like never before to ensure business continuity in these challenging times. This year more than any other, please do make sure that you enter our SDC Awards. There's no limit to the number of entries, all of which are free of charge, and we'll be promoting all the short-listed entries via Digitalisation World's multi-media platform over the coming months, ahead of the awards ceremony. We really do want to celebrate and recognise the many amazing achievements which have come about in response to the coronavirus.

WHY ENTER?

MAXIMISE VISIBILITY
Free PR opportunities with 5 months of marketing utilising the Digitalisation World portfolio.

POSITIONING
Position your product or service as an innovator in your field.

INCREASED CREDIBILITY
An award win, shortlisting or nomination acts as a 3rd party endorsement.

NETWORKING
Over 300 industry leaders attend the gala awards evening.

COMPETITIVE ADVANTAGE
Gain an advantage on your competitors.

NOMINATION IS FREE OF CHARGE AND VOTING IS DONE BY THE READERSHIP OF THE DIGITALISATION WORLD STABLE OF PUBLICATIONS.

SDC AWARDS 2020

www.sdcawards.com

Underwater data centres are reliable, practical and use energy sustainably

Earlier this summer, marine specialists reeled up a shipping-container-size datacenter coated in algae, barnacles and sea anemones from the seafloor off Scotland's Orkney Islands.The retrieval launched the final phase of a year's-long effort that proved the concept of underwater datacenters is feasible, as well as logistically, environmentally and economically practical.

BY JOHN ROACH, MICROSOFT.

MICROSOFT'S Project Natick team deployed the Northern Isles datacenter 117 feet deep to the seafloor in spring 2018. For the next two years, team members tested and monitored the performance and reliability of the datacenter's servers.

The team hypothesized that a sealed container on the ocean floor could provide ways to improve the overall reliability of datacenters. On land, corrosion from oxygen and humidity, temperature fluctuations and bumps and jostles from people who replace broken components are all variables that can contribute to equipment failure.

The Northern Isles deployment confirmed their hypothesis, which could have implications for datacenters on land.

Lessons learned from Project Natick also are informing Microsoft's datacenter sustainability strategy around energy, waste and water, said Ben Cutler, a project manager in Microsoft's Special Projects research group who leads Project Natick.
What's more, he added, the proven reliability of underwater datacenters has prompted discussions with a Microsoft team in Azure that's looking to serve customers who need to deploy and operate tactical and critical datacenters anywhere in the world.
"We are populating the globe with edge devices, large and small," said William Chappell, vice president of mission systems for Azure. "To learn how to make datacenters reliable enough not to need human touch is a dream of ours."

Proof of concept

The underwater datacenter concept splashed onto the scene at Microsoft in 2014 during ThinkWeek, an event that gathers employees to share out-of-the-box ideas. The concept was considered a potential way to provide lightning-quick cloud services to coastal populations and save energy.

More than half the world's population lives within 120 miles of the coast. By putting datacenters underwater near coastal cities, data would have a short distance to travel, leading to fast and smooth web surfing, video streaming and game playing.

The consistently cool subsurface seas also allow for energy-efficient datacenter designs. For example, they can leverage heat-exchange plumbing such as that found on submarines.

Microsoft's Project Natick team proved the underwater datacenter concept was feasible during a 105-day deployment in the Pacific Ocean in 2015. Phase II of the project included contracting with marine specialists in logistics, ship building and renewable energy to show that the concept is also practical.
"We are now at the point of trying to harness what we

have done as opposed to feeling the need to go and prove out some more," Cutler said. "We have done what we need to do. Natick is a key building block for the company to use if it is appropriate."

Algae, barnacles and sea anemones

The Northern Isles underwater datacenter was manufactured by Naval Group and its subsidiary Naval Energies, experts in naval defense and marine renewable energy. Green Marine, an Orkney Island-based firm, supported Naval Group and Microsoft on the deployment, maintenance, monitoring and retrieval of the datacenter, which Microsoft's Special Projects team operated for two years.

The Northern Isles was deployed at the European Marine Energy Centre, a test site for tidal turbines and wave energy converters. Tidal currents there travel up to 9 miles per hour at peak intensity and the sea surface roils with waves that reach more than 60 feet in stormy conditions.

The deployment and retrieval of the Northern Isles underwater datacenter required atypically calm seas and a choreographed dance of robots and winches that played out between the pontoons of a gantry barge. The procedure took a full day on each end. The Northern Isles was gleaming white when deployed. Two years underwater provided time for a thin coat of algae and barnacles to form, and for sea anemones to grow to cantaloupe size in the sheltered nooks of its ballast-filled base.

"We were pretty impressed with how clean it was, actually," said Spencer Fowers, a principal member of technical staff for Microsoft's Special Projects research group. "It did not have a lot of hardened marine growth on it; it was mostly sea scum."

Power wash and data collection

Once it was hauled up from the seafloor and prior to transportation off the Orkney Islands, the Green Marine team power washed the water-tight steel tube that encased the Northern Isles' 864 servers and related cooling system infrastructure.

The researchers then inserted test tubes through a valve at the top of the vessel to collect air samples for analysis at Microsoft headquarters in Redmond, Washington.

"We left it filled with dry nitrogen, so the environment is pretty benign in there," Fowers said.
The question, he added, is how gases that are normally released from cables and other equipment may have altered the operating environment for the computers.

The cleaned and air-sampled datacenter was loaded onto a truck and driven to Global Energy Group's

Nigg Energy Park facility in the North of Scotland. There, Naval Group unbolted the endcap and slid out the server racks as Fowers and his team performed health checks and collected components to send to Redmond for analysis.

Among the components crated up and sent to Redmond are a handful of failed servers and related cables. The researchers think this hardware will help them understand why the servers in the underwater datacenter are eight times more reliable than those on land.

“We are like, ‘Hey this looks really good,’” Fowers said. “We have to figure out what exactly gives us this benefit.”

The team hypothesizes that the atmosphere of nitrogen, which is less corrosive than oxygen, and the absence of people to bump and jostle components, are the primary reasons for the difference. If the analysis proves this correct, the team may be able to translate the findings to land datacenters.
“Our failure rate in the water is one-eighth of what we see on land,” Cutler said.

“I have an economic model that says if I lose so many servers per unit of time, I’m at least at parity with land,” he added. “We are considerably better than that.”

Energy, waste and water

Other lessons learned from Project Natick are already informing conversations about how to make datacenters use energy more sustainably, according to the researchers.

For example, the Project Natick team selected the Orkney Islands for the Northern Isles deployment in part because the grid there is supplied 100% by wind and solar as well as experimental green energy technologies under development at the European Marine Energy Centre.

“We have been able to run really well on what most land-based datacenters consider an unreliable grid,” Fowers said. “We are hopeful that we can look at our findings and say maybe we don’t need to have quite as much infrastructure focused on power and reliability.”

Cutler is already thinking of scenarios such as co-locating an underwater datacenter with an offshore windfarm. Even in light winds, there would likely be enough power for the datacenter. As a last resort, a powerline from shore could be bundled with the fiber optic cabling needed to transport data.

Other sustainability related benefits may include eliminating the need to use replacement parts. In a lights-out datacenter, all servers would be swapped out about once every five years. The high reliability of the servers means that the few that fail early are simply taken offline.

In addition, Project Natick has shown that datacenters can be operated and kept cool without tapping freshwater resources that are vital to people, agriculture and wildlife, Cutler noted.

“Now Microsoft is going down the path of finding ways to do this for land datacenters,” he said.

Go anywhere

Early conversations about the potential future of Project Natick centered on how to scale up underwater datacenters to power the full suite of Microsoft Azure cloud services, which may require linking together a dozen or more vessels the size of the Northern Isles.

“As we are moving from generic cloud computing to cloud and edge computing, we are seeing more and more need to have smaller datacenters located closer to customers instead of these large warehouse datacenters out in the middle of nowhere,” Fowers said.

That’s one of the reasons Chappell’s group in Azure is keeping an eye on the progress of Project Natick, including tests of post-quantum encryption technology that could secure data from sensitive and critical sectors. The ability to protect data is core to the mission of Azure in multiple industries.

“The fact that they were very quickly able to deploy it and it has worked as long as it has and it has the level of encryption on the signals going to it combines to tell a pretty compelling vision of the future,” Chappell said.

(This article originally appeared as a blog on the Microsoft website).

DIGITAL CHANGE MANAGEMENT SUMMIT

25.11 2020

www.dcmsummit.com

How Managed Service Providers and Cloud Service Providers can help SMEs on the road to digital transformation

A unique online event to connect MSPs, VARs and System Integrators with their target market

Fuelling the green data centre revolution

Although global Covid-19 lockdowns provided a brief respite on greenhouse gas emissions, demand for fossil fuels is now climbing fast again.

BY CHRIS ROBERTS, HEAD OF DATA CENTRE & CLOUD AT GOONHILLY

THE FOREST FIRES blazing in California this summer and the devastating bushfires in Australia serve as a poignant reminder that climate change waits for no virus. Against this backdrop, the World Meteorological Organization published its assessment that the 2010s was “a decade of exceptional global heat, retreating ice and record sea levels driven by greenhouse gases from human activities”.

Despite the pandemic taking many eyes off the climate change ball, there remains a groundswell of support from industry leaders keen to do their bit to help fuel the green data centre revolution.

The race to clean power

Some data centre operators are successfully switching to renewable energy but there is a long way to go. For example, there is scope to offer incentives for improved energy stewardship to the plethora of small on-premise legacy data centres that are massively energy inefficient.

Further pressure comes from the many data centres globally that are experiencing power constraints, and the growth in demand for power-hungry HPC platforms for AI and ML computation. These HPC systems, some of which are, ironically, helping to

solve climate change, also rely on greater processor density than standard servers, making cooling another hot topic.

Our sector is increasingly finding itself in the crosshairs of climate change protestors and we need to move the green needle by switching to self-generated, clean power solutions. This involves embedding ethical practices and transparency as a matter of urgency. As the industry has not taken the bull by the horns itself, this likely means more regulation and legislation.

Legislation will help

Until now, our sector has relied on voluntary initiatives such as the Code of Conduct for Data Centres, and the work being done by trade associations such as the Data Centre Alliance in the UK to improve energy efficiency and sustainability. But these initiatives need to be augmented with regulations and legislation to drive change.

Regulations requiring data centre operators to monitor and manage users' power usage at a highly granular level is one recommendation. The argument goes that greater transparency in monitoring and reporting will usher in greater efficiencies. Incorporating these technologies into modern data centres is relatively easy, but retrofitting these systems into leaky, older data centres would require more thought.

Ethical energy

On the subject of legislation, the city government of Amsterdam made waves in 2019 when it announced it was pausing approvals for the construction of new data centres while it reviewed their impact on the local property market and on power networks.

Despite 80% of Amsterdam data centres' power requirements coming from green electricity, some parts of the city are suffering from power supply shortages. According to Marieke van Doorninck, Alderman for Sustainability and Spatial Development in Amsterdam, the city is…"going to set requirements in the area of making available residual heat free of charge for the heating of homes and the use of green energy." And the trade body Tech UK published a data centre report last November urging data centre operators in the UK to improve heat reuse, renewables adoption and energy reporting, and calling for a "coherent" regulatory regime that forces distributed, in-house and public sector IT segments to publish these reports.

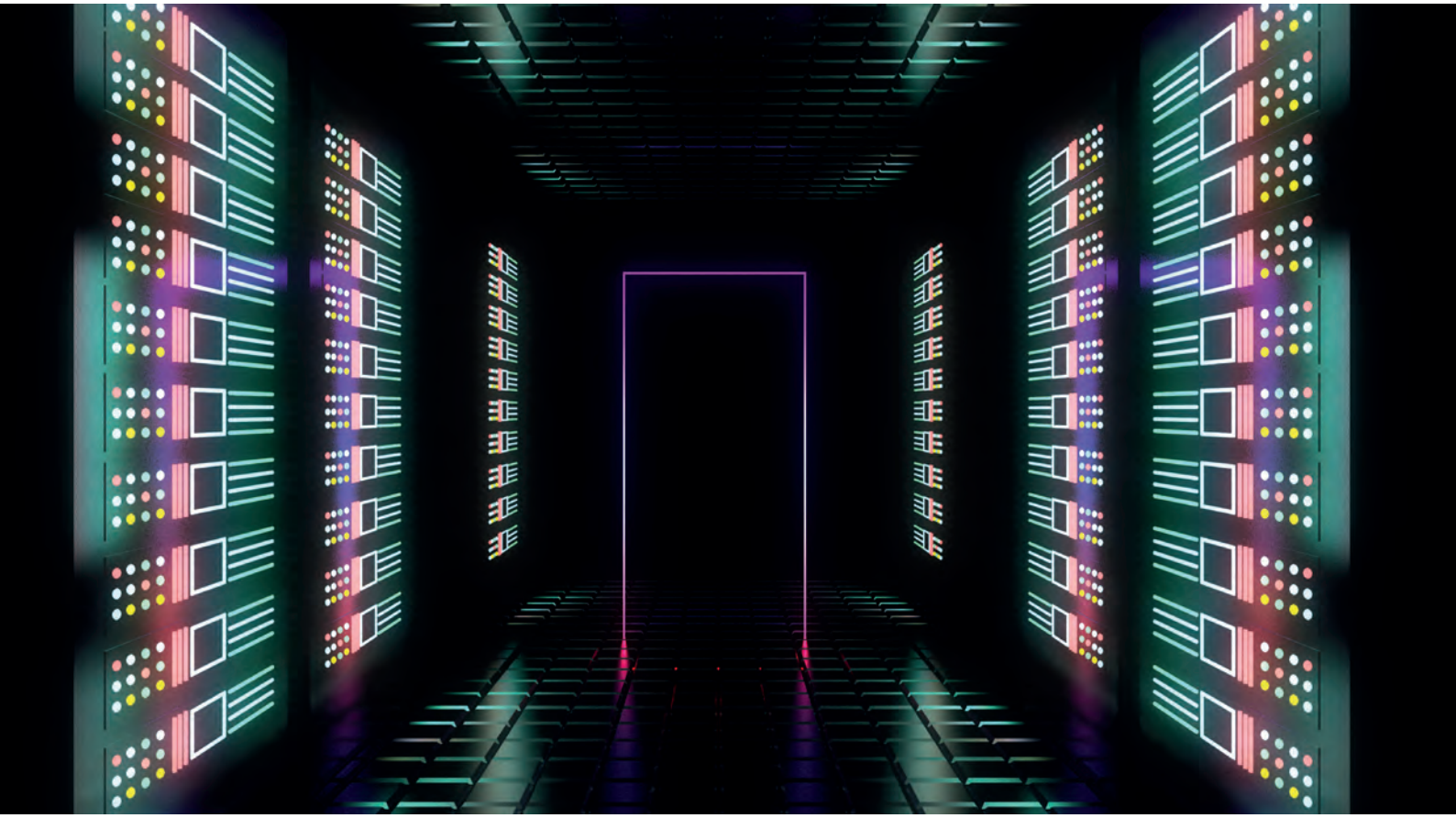
Hot buttons for a greener data centre

Fortunately, there are a number of new approaches and mind-blowing innovations that are already in motion to help reduce the environmental impact of data centres.

- For new data centres, there will be a shift to thinking 2-3 years ahead to harness the green energy technologies that will have matured by the time of construction, rather than settling for what's available at the time of planning.
- New regulations will require greater accountability and transparency on data centre energy consumption and carbon footprints, with the goal of becoming carbon-neutral sooner.
- The world's first 100% solar-powered supercomputer is now within sight. The race is on, with one focused on HPC photo-realistic rendering due to come on stream imminently. While in Texas, Hikari has already laid claim to being the world's first supercomputer to be partly powered by solar energy (up to 30%), and a 100% solar-powered machine is the goal, despite needing an acre of solar panels to generate the required 1 MW of energy.
- Robotically maintained and controlled solar panels with batteries that are 30% more efficient than those used today will gain traction, boosting the ability to generate consistent green power far more cost effectively.
- Liquid immersion cooling systems, like the one offered by Submer and deployed at Goonhilly's data centre to cool our managed HPC platform for AI and ML, will go mainstream. Submer's system is 45-50% more efficient than air cooling, cutting electricity demand in half, and allows us to use the exhaust heat elsewhere.
- We will also see AI and ML techniques being used not just to combat issues such as climate change but also to reduce the carbon footprint of compute power. Researchers in Sweden are working on a new approach to bottling solar energy using a solar thermal fuel that can reportedly store solar energy for over a decade.
- Further down the line, wave/tidal power will really start to make waves as a viable option for powering data centres in coastal areas. For example, Marine ower Systems has developed a wave energy converter with a power capacity that can match the larger offshore wind turbines, which could create a step-change in the commercial viability of wave power. And the UK's Offshore Renewable Energy (ORE) Catapult, based in Cornwall, has secured the go ahead for the €46.8m Tidal Stream Industry Energiser Project (TIGER), to deliver cost-effective tidal stream energy by installing up to 8 MW of new tidal capacity at sites in and around the Channel regions.

Summary

Controlling our environmental impact is an urgent requirement for everyone involved in the data centre industry. Fortunately progress is being made towards a carbon-neutral future. And this looks set to accelerate with the adoption of innovative green technologies and new green regulations and legislation.



Going green: powering the future of energy efficient data centres

It's well known that the data centre industry is the driving force behind the world's ever-accelerating digitalisation, managing the flow of information that is vital to many of the cloud-based platforms and services we now rely on. However, it's also recognised that data centres require huge amounts of energy and electricity to run their servers, backups, and power cooling infrastructure at the optimal levels.

BY ALESSANDRO BRUSCHINI, INFRASTRUCTURE MANAGER AT ARUBA S.P.A

WITH CLIMATE CHANGE being such a hot topic, vendors and operators are quickly realising that there is an urgent need to increase the energy efficiency of data centres and reduce their environmental impact. For example, many large data centres use enough electricity in a year to power thousands of homes – equivalent to around 30 billion kilowatt hours – and it is predicted that data centres will account for one-fifth of global electrical consumption by 2025.

In response, the EU Commission recently put pressure on data centre providers to take action to reduce their carbon footprints. In its 2020 report – 'Shaping Europe's digital future' – the EC says that the industry "can and should become climate neutral by 2030," highlighting the need to "become more energy-efficient, reuse waste energy, and use more renewable energy sources." With the data-serving requirements facing data centres only set to rise over the coming

years, the onus is on the industry to keep up with this demand, while at the same time finding ways to reduce carbon emissions.

The cooling conundrum

Cooling takes up a large proportion of a data centre's power requirements. In fact, it is the primary source of energy consumption in most conventional data centres, with up to 40% of electricity usage going towards sustaining and operating a server below 26 degrees celsius.

As such, maintaining temperature and humidity conditions at a level that enables IT hardware to operate effectively leaves a significant carbon footprint. But the importance of cooling isn't something that can be ignored. Overheating can negatively impact performance and permanently damage hardware, making it vital that businesses find the right solution – one that meets their needs while also addressing the sustainability issue.

Of course, it's not just about how power is used. Any energy efficiency discussion must also address how power is generated. The challenge is that going 'green' isn't always easy. Energy produced through green initiatives tends to be more expensive than standard energy, which means that sustainable alternatives must consume less energy so that organisations aren't left out of pocket.

The various voltages used in the transmission and distribution of energy are known to result in some efficiency losses, so data centres that focus on optimising their energy consumption have an opportunity to realise significant cost savings – which can be passed on to customers.

For example, hydroelectric plants can be used to harness the power of flowing water to generate energy in a sustainable and cost-efficient way. In addition, one way to reduce energy consumption is to replace traditional cooling solutions with geothermal systems that make use of the cold water found underground.

This is what is used to power the entire air-conditioning system for the data rooms across Aruba's Global Cloud Data Centre campus, making the system extremely energy efficient. The rack cabinets housing the servers are also equipped with an innovative cold air containment system to guarantee maximum energy efficiency and a comfortable working environment.

Striving for efficiency

The natural cooling system used in our Global Cloud Data Centre (IT3) campus in Milan provides a clear example of how data centres can drastically cut down the quantity of energy they use, reducing the overall cost and increasing sustainability by making better use of the energy available to them.

And there are plenty of innovations happening across the industry that highlight the benefit of natural cooling in reducing the carbon footprint. According to a survey published by The Green Grid, a non-profit consortium working to improve data centre energy efficiency, nearly 50% of US-based data centres are now using natural cooling to save energy and costs, with another 25% considering doing so in the near future. And this doesn't just apply to areas with extreme weather conditions, as illustrated by consulting company Capgemini. It recently opened a data centre in Swindon in the UK that has used natural cooling to reduce running costs by 80% and cut carbon emissions in half.

But, despite progress being made, the industry can't afford to stand still. The progression of energy-efficent and enviromentally-friendly solutions requires an ongoing industry-wide effort from vendors and operators alike. Although big strides have been made in demonstrating that energy-efficient solutions and practices can be implemented effectively, there needs to be a concerted effort in tackling this on a global scale.

This is where industry initiatives such as The Green Grid and the 'European Code of Conduct' for data centres, along with other industry standards and certifications for energy efficient operations, are so important. Although we can't avoid the fact that data centre power consumption will continue to be an issue in the future, we can invest in new ways to achieve maximum energy efficiency on a wide-industry scale, all while reducing costs and protecting the environment.

Connecting the world's greenest data centre

Built into the mountains of Stavanger, Norway, in what was formerly a NATO ammunition storage facility, is now one of the world's most secure and energy efficient data centres.

BY KRISTIAN GYLAND, GREEN MOUNTAIN, CEO

FOR SITE OWNERS, Green Mountain, the transition from ammunition storage facility to data centre was not without challenges. They were faced with confined spaces, existing structures and the need to install a reliable piping system to cool the server racks. The site, however, was also not without benefits – not least the cold Norwegian climate to keep data cool at lower costs and the vast supply of hydroelectric power. It has therefore proved to be a sensible investment for Green Mountain CEO, Kristian Gyland.

Together with Victaulic, a world-leader in mechanical pipe joining systems, and pipe installation contractor, Sig Halvorsen, Gyland overcame the many challenges that the site presented, constructing a truly unique data centre.

From storing weapons to storing data

The facility was constructed by NATO in 1964, during the height of the Cold War. The site initially spanned across three halls, and, in 1994, was extended to double its storage capacity to house mines and torpedoes.

Following the de-escalation of geopolitical tensions, NATO no longer required the facility and decided to sell in 2009. And what NATO no longer needed, became a golden opportunity for Green Mountain's first data centre.

"We opened this facility in 2013, with our first three customers. Since then we have continued to expand the site as we've grown. Today, we are covering 22.000 km² under the mountain and expect to reach full site capacity in 2023," said Gyland.

Secure and sustainable in Stavanger

When they learned of the location, it was obvious to Green Mountain that the site offered several opportunities that were too good to pass up. Although the construction design would be challenging, the business potential was there.

In the data centre industry, storing data in a secure environment is of upmost importance. Green Mountain had a vision that if a mountain could keep NATO's weapons secure, then it could also keep data secure; and they were right. According to Gyland, the Stavanger site is one of the most secure data centres in the world. A feature which can be largely attributed to the centre's location within the mountain.

As well as being secure, the Stavanger site is located within close proximity to the fjord, home to one of Europe's lowest-priced sources of hydroelectric power. The data centre could therefore operate at lower costs, as well as with a sustainable power source. Together with a vast supply of hydroelectric power, the site benefits from a cold Norwegian climate, providing a cost-effective means to cool data. "We are located close to a fjord and the water we are collecting for our cooling system has a constant temperature of 8 degrees, meaning that we can use the outside to cool the data centre on the inside. So, in addition to the site being the most secure data centre in the world, it is also the most energy efficient data centre in the world." stated Gyland.

Looking for versatility and durability

Designing and installing a piping system in an existing structure is never without its challenges and turning an existing structure into a data centre adds an additional level of complexity.

Faced with confined underfloor space, existing structures, and the need to have a system that is reliable and easily maintainable, Green Mountain needed an efficient pipe joining system, in addition to an installer that could do the job. Because Norway is heavily involved in the oil and gas business, welding is a common method of joining pipe in the region. However, the site needed more flexibility than welding could offer. It was important to Green Mountain that they had the ability to easily expand the facility as their business grew. Victaulic offered the ideal solution, but with one concern for Green Mountain; would it last?

"When we were presented with the mechanical pipe joining solution in 2012, it was only natural for us to use this technology for our piping systems. Using a system that is not welded provided us with the flexibility we needed and was a huge cost saver since it allowed us to build in phases. We didn't have to make assumptions on where future data racks were going to be placed, and where the cooling system should run; we were simply able to build as we grew and add customers," commented Gyland.

"The only concern we had when we installed a grooved system in our facility was whether the joints would last over time," continued Gyland. "But having operated for close to 6 years now, we're confident that this is a solution we will continue to work with, and we are happy with the way it's working here at Green Mountain Data Centre."

Beyond the flexibility and reliability of mechanical pipe joining solutions, Gyland was also pleased to partner with a manufacturer that aligned with their values on sustainability.

"As one of the most energy efficient data centres in the world, it is also evident that one of the key areas of prioritization for Green Mountain is sustainability. As the installation of the grooved system avoids toxic fumes and gases and is produced in a production facility which uses 90% recycled material, the solution fitted our company's sustainability efforts seamlessly." noted Gyland.

Overcoming installation challenges

Sig Halvorsen was the contractor tasked with installing the cooling system in the Green Mountain facility. Their employees are familiar with grooved pipe joining solutions, particularly within data centres.

Frode Horpestad, Operations Manager at Sig Halvorsen, who worked on the Green Mountain installation, stated that the pipe joining system allowed his team to overcome some of the site's installation challenges and even cited these solutions as an aid in winning additional projects.

"It is a simple system to learn and has simple check methods to ensure proper installation. Our employees also received close follow-up from the manufacturers' representatives on the construction site if required – so it has worked very well. Using a flexible and robust technology has many advantages, for example utilising this pipe joining system allows us to use the same crew on every construction site, and we do not require certified welders for joining the pipes."

Considering the internet wasn't invented until several years after the Stavanger NATO facility was built, it is safe to assume the engineers of the time never foresaw a day where the location would become a world renown data centre. But as it turned out, the safety and security of a mountain, with a nearly unlimited hydro-power supply and cool climate, made for the perfect location.

"It is a simple system to learn and has simple check methods to ensure proper installation. Our employees also received close follow-up from the manufacturers' representatives on the construction site if required – so it has worked very well. Using a flexible and robust technology has many advantages, for example utilising this pipe joining system allows us to use the same crew on every construction site, and we do not require certified welders for joining the pipes."

Powering ahead in uncertain times

Riello UPS Business Development Manager Chris Cutler explores how advances in UPS technology can help data centre operators optimise their energy efficiency as they adjust to the "new normal" of growing demand.

THESE LAST FEW MONTHS have pushed the crucial role of data centres into the public spotlight more than ever before.

Businesses turned to the cloud in extraordinary numbers, enabling their staff to work remotely. The number of Microsoft Teams users worldwide exploded from 20 million to 75 million. In April, a staggering 4.1 billion minutes of "virtual meetings" took place across the platform on just a single day.

Video streaming services, online gaming, internet shopping, and social media browsing helped alleviate lockdown boredom. Netflix gained nearly 16 million new global subscribers, while broadband providers reported up to 60% increases in weekday web usage. Hyperscalers and enterprise data centres picked up much of the slack, coping admirably with this shift in usage patterns whilst at the same time having to adapt their processes and staffing levels due to the pandemic. But while everyone's immediate attention has understandably focused on coping with coronavirus, we mustn't take our eyes off another

existential challenge – climate change. Overall electricity consumption across Britain plummeted by 20% during the height of lockdown, as thousands of offices and businesses were mothballed. The country smashed records for "coal-free" power generation as demand dropped and renewables were able to meet most of the nation's needs. But as lockdown eases, electricity use is slowly creeping back up towards pre-COVID-19 levels.

And as we emerge from the pandemic, it's anticipated many businesses will embrace the shift to remote-working as a permanent change – a survey by Gartner found 74% of companies will encourage more home working post-coronavirus.

Such a trend will inevitably lead to an increased emphasis on the cloud and quicken the migration to edge processing, with more work being done away from corporate offices. Add in the ongoing rollout of 5G and data centres will face huge pressure to keep up with demand.

With the latest Uptime Institute global data centre survey suggesting that site efficiency has flatlined in recent years, what can operators do to ensure that the growing demand for their services doesn't come at too big an environmental cost?

In amongst solutions such as increasing rack density and taking advantage of improvements such as liquid cooling, there are significant energy savings to be made from a fundamental part of a data centre's critical infrastructure – its uninterruptible power supplies. Here are a few examples of how advances in UPS technology can help operators keep control of their power consumption.

Increase efficiency with a hardware refresh

Every UPS system has a natural lifespan. Industry best-practice suggests this tends to be around year 10-12 of service life. However, with certain legacy installations it might make sense to replace your UPS even earlier in the lifecycle.

At first glance, this might seem counter-intuitive. As the benefits of Moore's Law start to plateau, many operators are starting to leave it longer between hardware refreshes. Another glance at the Uptime Institute report reveals that while back in 2015 most servers were being upgraded every three years, fast-forward to 2020 and a five-year refresh is the most common response. But with a legacy UPS, there's every chance that the upfront cost of upgrading is more than recouped in long-term efficiency and performance gains.

Compared to old-style transformer-based UPS systems, a modern transformerless solution could deliver a 5-6% increase in efficiency. Transformerless UPS also offers a significant edge when running at lower loads. They have a far flatter efficiency curve, with many models capable of high efficiency (above 95%) even at 25% load. For example, an old 400 kVA UPS would have had a 3.5% difference in efficiency depending on the load applied.

New transformerless UPS systems are smaller and lighter than a big, bulky transformer-based unit too. That means they don't generate as much heat, which could potentially lower your air conditioning requirements. With floorspace also a major consideration for data centre operators, their compact footprint is another bonus.

Rightsizing With modular UPS

If you do opt to upgrade to a new UPS, making the move to a modular solution could provide the opportunity for enhanced efficiency.
Many legacy data centre UPS installations are designed for a much higher load than is actually required. Such oversizing is inefficient, wastes energy, and costs more to run and maintain.

In essence, with a modular UPS you only need to install the "power" you need. Modular systems are made up of a frame which you populate with individual power modules that work together. Like building blocks, you just add as many as you need for your required power and redundancy.

Each individual module is a highly-efficient UPS in its own right with a rectifier, inverter, and static switch. They're hot-swappable too, which guarantees you downtime-free maintenance during service visits.
As well as eliminating wasteful oversizing, the other major benefit of the modular approach is scalability. Say you're a colocation data centre and you have a sudden spurt of clients buying more racks, increasing your power needs. With a modular UPS, you simply add in the extra modules – or install additional cabinets in parallel – to match your new load requirements.

This “pay as you grow” flexibility works the other way too. So if clients reduce their rack space, you can just as easily remove a module or two from your UPS. Such vertical and horizontal flexibility can deliver power capacity from as little as 25 kW to more than 1 MW plus redundancy in a single UPS system. This helps to future-proof your power protection needs without wasteful oversizing or using unnecessary floor space and air conditioning. This makes compact modular UPS the ideal choice to backup space-restricted edge applications.

Use Of ECO and active ECO mode

Virtually every modern uninterruptible power supply now offers some form of “economy or ECO” operational mode, which can deliver data centres significant energy savings.

In practice, these ECO modes see the UPS run in a similar way to a standby UPS. The critical load is powered by the bypass line i.e. the mains electricity, with the inverter switched off but ready to take over if there’s an issue with the utility supply. If the worst happens, there is a short break in continuous power – a few milliseconds – while the automatic bypass transfers the load back to the inverter.

Typically, running in ECO mode offers efficiency of up to 99%. Compared to the 93-97% efficiency rating of modern online UPS, that’s a difference of anywhere between 2-6%. For a large-scale facility, even a 1% increase in efficiency could add up to tens of thousands of pounds of savings a year.

Studies show that a data centre running in ECO mode can save 2.3% of their annual energy. For a 1 MW facility carrying a 50% load who pay 10p per kWh for their electricity, that could equate to £10,000 of savings a year. Of course, the main drawback with ECO mode is that your IT load is exposed to the raw mains utility without any of the vital power conditioning provided by a double-conversion online UPS. If you have a reliable, stable mains supply and loads that generate low harmonics, running in ECO mode comes with limited risk. But if your site is prone to power quality problems, or if you have sensitive equipment, then an overreliance on ECO mode could compromise your system availability. That’s why in recent years, advances in firmware control and electrical designs have seen the development of Active ECO – sometimes known as Advanced ECO – operating mode.

It’s similar to standard ECO mode, as the load is still powered via the bypass line. However, the inverter remains on at all times, running in parallel with the input without actually carrying the load current. As the inverter is always on, power transfer is far quicker than standard ECO mode in the event of a mains problem. And even though the inverter isn’t processing the load, it can absorb harmonic currents and provide power filtering similar to how full online mode does, ensuring better power quality. The efficiency savings from Active ECO are around 0.5-1% less than in pure ECO mode, due to the extra energy used to power the inverter. But they are still higher than operating in traditional online mode.
While running in either ECO or Active ECO mode will undoubtedly deliver energy savings, for many mission-critical sites such as data centres, the trade-off in resilience and power quality is often too much. That’s not to say there aren’t possibilities though. It could be an option to run economy mode when a site’s most critical loads are inactive, for example, overnight or out of hours. While another alternative involves UPS systems in an “N+X” parallel redundant configuration. Here you’d have one of the UPSs running in online load as the “master”, with the remaining UPSs operating in ECO mode until the condition of the mains supply changes and they’re required to actively support the load.

Tap into advances in battery technology

While sealed lead-acid (SLA) batteries continue to dominate the UPS market, lithium-ion (li-ion) continues to become an increasingly viable alternative. Although li-ion comes with a higher upfront price tag compared to SLA, they have up to 50 times the cycle life and a much higher power density that means they take up less than half the space and weight. In addition, they can operate safely at far higher ambient temperatures, reducing air conditioning costs. As our electricity network continues its transition towards a reliance on renewables and smart grids, li-ion also offers data centres a route into energy storage and demand side response. It’s a great way for sites to store off-peak electricity to use at busier, more expensive times of the day, cutting your energy bills and triad charges. Alternatively, li-ion batteries can store power generated onsite from renewables like solar panels, reducing your reliance on mains electricity.

With data centres in the frontline more than ever before, effective use of modern UPS and battery technologies will ensure operators can offer a seamless service to customers whilst also keeping their energy use under control.

CELEBRATING 10 YEARS OF SUCCESS

The 2020 DCS Awards feature 31 categories across FOUR groups.

THE DCS AWARDS are now firmly established as the data centre industry's premier annual celebration of all that is great and good. End user projects, product innovation and individual excellence are all recognised in an evening that pays more than lip service to the idea of data centre and IT convergence. So, the award categories cover both the facilities and IT aspects of the data centre, recognising the achievements of vendors, their business partners, their staff and their customers.

Getting involved with the DCS Awards couldn't be easier. Take a look at the award categories, and make sure to nominate your company, a customer, or maybe an individual – better still all three (!) – for a chance to be recognised for outstanding achievement when it comes to projects, product innovations and individual contributions within the data centre industry.

Once you've made your nominations, make sure to book a table for the Awards night. You wouldn't want to win an award and not be there to collect it! (And even if you don't win an award on the night, there's a cocktail reception, three course meal and a top comedian to entertain you – we have a track record of booking individuals on their way to the top of the comedy circuit).

To 'We look forward to welcoming you to the Awards night in December.

NOW, GET NOMINATING!

WHY ENTER?

MAXIMISE VISIBILITY
Free PR opportunities with 5 months of marketing utilising the Digitalisation World portfolio.

POSITIONING
Position your product or service as an innovator in your field.

INCREASED CREDIBILITY
An award win, shortlisting or nomination acts as a 3rd party endorsement.

NETWORKING
Over 300 industry leaders attend the gala awards evening.

COMPETITIVE ADVANTAGE
Gain an advantage on your competitors.

NOMINATION IS FREE OF CHARGE

The DCS Awards panel will validate entries and announce the final shortlist to be forwarded for voting by the readership of the Digitalisation World stable of publications in October 2020.

The winners will be announced at a gala evening at the LEONARDO ROYAL ST. PAUL'S HOTEL, London on 10 December 2020.

DCS AWARDS 2020

www.dcsawards.com

Supported by

dca The data centre trade association

Why efficiency matters in the modern data centre

For many decades, when anybody was talking about a computer's performance, it would have been taken for granted that they were referring to its speed.

BY KEVIN DEIERLING, VP MARKETING, NVIDIA NETWORKING BUSINESS UNIT

IN THE HIGH-OCTANE WORLD of today's tennis court-sized supercomputers, this is measured in FLOPS, or floating-point operations per second. Defined by speed, the world's highest performing computer right now is Japan's Fugaku, operating at 415 petaFLOPS.

But there are problems with using speed as the only metric to compare one supercomputer to another. An arms race, based around FLOPS ratings, has seen the emergence of a generation of supercomputers that burn through colossal amounts of electricity, and give out so much heat as a by-product that hugely elaborate cooling systems must be deployed at all times to keep them from melting down. An over reliance on speed in the benchmarking of computers also downplays other vital qualities, such as reliability, availability and usability. And then there's the

economics of the thing. Making speed the primary measurement of success has seen the total cost of ownership of supercomputers hit unprecedented heights, while at the same time driving up their negative impact on the environment.

The Green500 offers a different and timely approach. An alternative to the speed-focused Top500 listing, it is a ranking of the 500 most energy-efficient supercomputers in the world, and was devised to increase awareness of other performance metrics than just FLOPS. It achieves this by ranking computers according to FLOPS per watt, as well as taking into account energy efficiency and reliability. The Green500 exists also, as its name would imply, to promote the importance of environmental credentials to the various stakeholders and investors in the supercomputer sector.

The top placings in the Green500 reveal some interesting things about where supercomputing is going, and indicate why efficiency may be taking over as the number one determinant of what sorts the leaders from the followers.

Top of the Green500 is the MN-3 system, built by Japanese start-up Preferred Networks and coming in at 21.1 gigaFLOPS per watt. Measured by speed, however, MN-3 sits in 394th place in the Top500 list. It is not in commercial use, being only available for its maker's own R&D programme.

In second place is Selene, an AI supercomputer made by NVIDIA. Selene delivers 20.52 gigaFLOPS per watt, putting it in a comparable bracket to MN-3 for efficiency, but it scores 7th on the Top500 list, so it's pretty rapid too.

Selene is based around a unique type of open infrastructure, the DGX SuperPOD. Designed and built in just a couple of weeks, the DGX SuperPOD combines NVIDIA's DGX processor design with an AI networking fabric from Mellanox.

It's this configuration that gives Selene performance, efficiency and economy, as well as flexibility in terms of the variety of uses it can be put to. NVIDIA's intention with Selene was to create a supercomputer-class system powerful enough to train and run its own AI models, for use in fields such as autonomous vehicles, but flexible enough to be part of just about any academically-led deep learning research project.

Since its deployment, Selene has run thousands of jobs a week, often simultaneously. It conducts AI data analytics, traditional machine learning, and HPC applications. The power of the DGX SuperPOD is in use with companies such as Continental in the automotive sector, Lockheed Martin in aerospace and Microsoft in cloud computing.

Between them, the top machines on the Green500 list

point to a new direction in supercomputing, combining vastly superior cost of ownership compared with traditional alternatives with a design that makes them the right fit for tomorrow's top-level AI challenges. AI, powered by these machines, is transforming the planet and every aspect of life as we know it. Organizations that want to be in the vanguard of this AI-powered world understand the need for compute power that offers unprecedented scale as well as ease and speed of deployment.

The supercomputer of the future needs to be equally at home in an environmentally conscious data centre, running HPC tasks, as it is with AI research companies looking for machines that are big, fast and fit for purpose. In either case, there is no appetite any more for expensive and time-consuming custom builds with their complex interconnect trade-offs. With tomorrow's open architecture there are no more proprietary designs that take months to commission.

Modern compute needs machines that serve multiple uses and have long lifetimes, packing as much processing, memory and storage as possible into the smallest space with the least possible energy consumption. With the best of the Green500, this is now a reality.

High-quality variable-speed pumps key to improving energy and water usage in data centres

With the internet and data centres being the biggest growth 'industry' in the UK, electrical demand is continuing to drive climate change. Working alongside data centre specialist, Professor Ian Bitterlin, leading water pump manufacturer, Wilo UK, has identified how chilled water pumps are key to improving energy and water usage.

THE COVID-19 pandemic has added 20% growth to data centres in the past three months, with lockdown meaning more adults are working from home and families are streaming more content during the day. The implementation of 5G and streaming video is rapidly driving data growth and with users unlikely to reduce their usage, actions such as reducing cooling losses at data centres are important, with better pumps and more efficient controls are a key enabler in facilitating this.

There is increasing pressure from the EU for data centres to be powered by renewable energy, however the first step of this is to reduce the power consumed by the data centre infrastructure.

The Green Grid, innovator of the Power Usage Effectiveness (PUE) metric, has recently proposed the introduction of a new data-centre metric - WUF, Water Usage Factor and as PUE's drive lower, data-centres are now being judged for water consumption as well as electrical energy use.
80% of all data centres utilise chilled-water systems for cooling, 15% use direct-expansion refrigeration systems and just 5% use air based evaporative or adiabatic cooling systems.

Professor Ian Bitterlin, who carried out the research and is a Consulting Engineer & visiting Professor at Leeds University, comments: "Whilst chilled-water systems (the dominant technology used to date) have continued to evolve technically with better controls, heat-exchanger technology, variable speed compressors, fans and pumps and operationally, with flow water temperatures rising from the legacy 6°C to 18°C (and higher) enabling high percentages of free-cooling in suitable climates, an older technology has recently proved more popular – that of evaporative or adiabatic cooling.

"In pursuit of achieving an ever lower PUE, the advent of fresh-air cooling solutions brought along with it adiabatic cooling solutions, where water is used to take advantage of the wet-bulb ambient temperature and crucially humidification of high volumes of fresh-air. Perhaps for the first time in Europe, water consumption in data-centres is a growing issue."

Evaporative and Adiabatic cooling technology, known before Roman times in high-status dwellings, use water to increase the humidity of warm dry air and reduce its temperature from the dry-bulb to the wet-bulb value. For example, in the UK when the external ambient is near the record high of 35°C dry-bulb, the addition of water vapour can get the air-stream temperature down to 23°C wet-bulb and then use that to cool the data centre.

Evaporative and Adiabatic cooling systems potentially save 20-30% data centre energy (compared to chilled water systems) and do not use pumps. However, they have not proven to be universally popular as they need a lot of space and use a lot of water.

Bitterlin comments: "Despite the potential energy saving of Evaporative and Adiabatic cooling systems, chilled-water systems are, in my opinion, the way forward as they use hardly any water compared to

the ‘modern’ competition of evaporative cooling technology. The performance of chilled water systems is much improved by high-quality variable-speed pumps and chillers fitted with Evaporative or Adiabatic cooling offering total control of internal air temperature and humidity.

“The majority of enterprise and colocation data centres have partial load, typically <50% at maturity, rarely high and never 100% - this means that for energy saving reasons the chilled water pumps must be; designed for variable speed drives, optimised for operation at 40-50% - like modern power systems (UPS) .

“A pump designed for 100% flow but only having 30% load uses 100% power but the same pump running at 30% flow rate only consumes 2.7% of the energy.” David Williamson, Director of Wilo UK, comments: “Data centre cooling plants with Wilo pumps provide an opportunity to improve on past performance with partial load and variable speed pumping and offer a high level on control to meet a wide range of systems.”

“Through analysis of the hydraulic system and measurement of power performance of existing and legacy cooling systems we are able to select replacement pumps that can achieve the desired system performance whilst reducing energy use. Often such upgrades further increase data centre resilience and availability.”

The Wilo GIGA range of pumps is extensive, covering all applications within the Data Centre environment.

The Atmos GIGA Series has recently been upgraded to provide greater efficiency and the Stratos GIGA leads the industry in high efficiency providing performance greater than IE5 through EC motor technology up to 22kW.

Far from the perception that evaporative or adiabatic are modules taking over the market, the sales of chilled-water systems have proven to be increasingly resilient, aided considerably by an offshoot of the adiabatic technology; the air-cooled chiller enabled with adiabatic sprayed free-cooling coils.

With the ability to have compact cooling units in the room of 3.5m3/100kW, using water to transport heat far more effectively than air, advanced micro-channel heat exchangers, higher range chilled water temperatures, free-cooling and the opportunity in hot/dry ambient conditions to use water spray to reduce the PUE to a, location dependent, range of 1.06 in London, 1.07 in Madrid or Frankfurt and 1.18 in Dubai. If engineered correctly, chilled water cooling can no longer be considered as wasteful of energy achieving an overall PUE at full load, with all the other systems included, of 1.2 in northern European city centres.

The final piece of the puzzle is to cater for the endemic partial load and here is where the chilled water pump, allied to variable speed pumps and electronic proportional valves, come into their own. The product that, with the chiller, was most threatened by air-based direct and indirect adiabatic or evaporative cooling can look forward to a successful future in the data centre industry that we have all come to rely on so much.

How advanced UPS technologies are turning data centres from energy villains to heroes

Increased electrification will have a severe negative impact on the environment if key sectors don't find a way to balance the growing demand for energy with increased adoption of renewables.

BY JANNE PAANANEN, TECHNOLOGY MANAGER AT EATON

RENEWABLE ENERGY SOURCES now provide one-third of the world's power, and this is set to rise further. However, renewables are volatile by nature and require close management on the grid, meaning we often rely on fossil fuels to maintain system balance.

But this is where the large electrical capacity of a data centre can step in. By deploying UPS technology, large data centre operators can be compensated for immediate adjustments to power consumption that help the grid avoid power outages.

The system puts data centres in control of their energy by allowing operators to choose how much capacity to offer, when, and at what price, while helping grid operators balance sustainable energy sources.

By using its energy reserves to account for the variability of renewables, data centres can help to manage sudden disturbances in energy production that might compromise the grid. This balancing act can therefore allow for the increased adoption of renewables, and decrease our reliance on fossil fuels.

Beyond contributing to the increase of renewable energy sources by stabilising the grid, certain advanced UPS devices can be a value-generating asset. Major UPS users can be financially compensated for immediate adjustments to power consumption that help the grid avoid power outages, without compromising service levels.

Moving to a future centred on renewable energy is necessary to reduce the effects of climate change. As we expect an increase in electricity consumption, it is increasingly clear that new technologies will be the core drivers of a green energy boom.

Data centres are crucial to the running of our digital environments, creating a unique opportunity to use this role to help stabilise the grid and encourage greater renewable energy uptake whilst maintaining functionality.

Although data centres consume electrical energy, they also hold the key to overcoming decarbonisation challenges. It is their ability to leverage new technologies, like advanced UPS systems, to simultaneously reduce power failures and carbon energy usage that makes

data centres part of the solution and not part of the problem. With the support of data centre operators and the electrical utility industry, the data centre power network can pivot to boost renewable energy usage and set us on the right path to a carbon neutral future – despite the boom in human data consumption.

What approaches should organisations take to cope with the explosion of data as Digital Transformation takes hold?

BT recently reported a 35% to 60% increase in daytime weekday fixed broadband usage, caused by the widespread shift to remote working. This has coincided with the European launch of commercial 5G services which will inevitably lead to more internet-connected devices and an explosion of data across all sectors.

Data centres are central to the handling all of the data that facilitates our digitally-led lives, but with this comes concern about their growing impact on the environment. Statistics on data centre energy consumption can be staggering and it is important to question whether the insatiable desire for social media and all things digital is creating an environmental villain. They are responsible for approximately 3 per cent of global energy use and as we saw from BT, this is only going to rise as our reliance on data increases due to current movement restrictions.

Data centre owners in particular should be considering how to take a greener approach to their operations in order to counter the environmental impacts of this surge in data. Uninterrupted Power Supply (UPS) technology is one such method of handling the extra strain of higher data usage as the solution helps facilitate increased renewable energy adoption on the grid. Increasingly, the UK is turning to renewable sources of energy and last month marked Britain's first coal-free month since the industrial revolution – a phenomenal achievement given that renewable energy is volatile by nature and therefore harder to factor into the energy grid.

The fact we cannot rely on the sun or wind as the only source of electricity to heat our homes in the winter means we still need a certain amount of fossil fuels to keep the grid stable, especially as more renewables are introduced.

Eaton launches energy solution

POWER MANAGEMENT COMPANY Eaton has launched a frequency response energy solution, called Energy Aware, that enables energy-intensive industries to help stabilize the grid and contribute to more renewable power. Eaton Energy Aware UPS is based on Eaton's reserve power technology.

The first commercial application of its kind is launching in Ireland, and has been successfully piloted at Eaton's global headquarters in Dublin in co-operation with Eaton's global energy management partner Enel X. It is participating in Ireland's grid operator EirGrid's DS3 system services programme, which supports renewable integration through the management of grid frequency. Eaton Energy Aware is aimed at organizations that use large amounts of electricity, such as commercial and industrial facilities like data centers, manufacturers and airports.

Eaton's technology is a solution that industry can adopt to increase its role in the movement towards enabling a low carbon future through energy market participation. It launches at a time when there is unparalleled desire politically, in industry and amongst the public, to tackle the climate crisis and a decentralization of the energy sector.

"As individuals and as an industry we are all aware of our obligation to rethink our role as power consumers. Industry continues to struggle to make efficiency gains and to de-carbonize the energy we use. This technology turns the tables and literally means industry can give back," said Ciarán Forde, segment leader, Data Center & IT EMEA at Eaton. "This is an innovation that challenges traditional thinking, it transforms a large electrical infrastructure into a network that can contribute to the grid and not only consume power.

The primary role of the infrastructure doesn't change, but by enabling organizations to take part in EirGrid's DS3 system services, it offsets its costs and helps them hit green targets by enabling renewable energy."
Eaton Energy Aware UPS uses Eaton's UPS-as-a-Reserve (UPSaaR) technology, which has been in development since 2016 and has been successfully trialled in Norway, Sweden and the UK. The solution uses Eaton's reserve Uninterruptible Power Supplies (UPS) to enable organizations to seamlessly support grid stability while still powering critical applications.

This helps grid operators stabilize emergency fluctuations in system frequency which occur as a result of increasing variable renewable energy. Eaton has worked to bring the solution to market by partnering with Enel X, which enables organizations to take part in EirGrid's DS3 system services and earn revenue for doing so.

"The potential of the energy sector's transformation is dependent on an ecosystem of proactive energy consumers who see opportunity in using their energy differently – more intelligently – than they do right now," said Peter Connolly, Head of Sales at Enel X in Ireland.

"We are delighted to partner with Eaton to enable organizations to leverage their operational flexibility in a way that bolsters business competitiveness whilst also allowing them to become good citizens of the grid."

Ireland's electricity system is primed to test and launch the new solution because of its mix of renewables, its progressive orientation towards technology and its population of large power consumers, including many of the world's largest data center providers. Eaton's proof of concept with Enel X and EirGrid at its Dublin headquarters has demonstrated how Energy Aware manages critical loads and mission critical applications, reduces Eaton's energy consumption and improves grid stabilization.

Pilot projects have proven that the technology provides a viable and fast-acting power reserve to the grid while maintaining the UPS's primary function of securing electrical loads, further de-risking any instability from the country's increasing use of renewable generation.

UPS systems help to tackle issues related to low inertia by providing system (ancillary) services which in turn help to increase the penetration of renewables on the grid when there is more wind and sun. This means we can reduce our reliance on fossil fuels to counter the variability of renewable energy sources and therefore allow the entire grid to become greener even as the amount of data usage increases.

Aside from the green benefits of adopting this approach, UPS devices can also generate monetary value for data centres. Major deployments can be financially compensated for immediate adjustments to power consumption that help to stabilise the grid and reduce power outages, without compromising their service-levels.

Increasing the portion of renewable supply to the grid while reducing power failure puts data centres in a critical position to help organisations benefit from the uptake in data usage that is expected in the coming years.

DIGITAL CHANGE MANAGEMENT SUMMIT

25.11 2020

www.dcmsummit.com

How Managed Service Providers and Cloud Service Providers can help SMEs on the road to digital transformation

A unique online event to connect MSPs, VARs and System Integrators with their target market

A new era for the data centre has arrived

The cost to run a data center continues on an upward trajectory — and there is increasing interest in making them more energy efficient.

BY SIOBHAN ELLIS, DIRECTOR AT PRODUCT MANAGEMENT AT BAMBOO SYSTEMS

POWER USAGE EFFICIENCY (PUE) measures the ratio of power used to deliver services compared to the total power consumed by the data center, with the best ratio being 1:1. As climate change continues to be a pressing global concern, and data centers continue to be built and grow there has been a concerted focus on their carbon footprint. Data centers consume between 1 and 3% of the world's total energy and that number is only expected to climb. The big question here is, by how much? And what can be done to mitigate the impact of that energy consumption on our planet.

The technology community knows this pain point and that something needs to be done. Most major cloud providers have announced a plan to reduce carbon emissions and are taking steps to becoming carbon-neutral. 100% renewable energy is one area of concentration, although it is worth noting that no discussion has been had about how much energy is used to create a wind or solar farm. Locating data centers in cooler climates, to reduce air conditioning costs, or diverting data center heat and using it to warm homes are two of several energy conscious strategies that are becoming more common.

Over the last few years these strategies have resulted in the flattening of the increase in overall energy consumption and helped to keep the data center carbon footprint in check. Even with Moore's Law, which suggests a doubling of processing power every year, without the doubling of energy consumption, we are still seeing a noteworthy increase in the amount of energy consumed by the servers themselves.

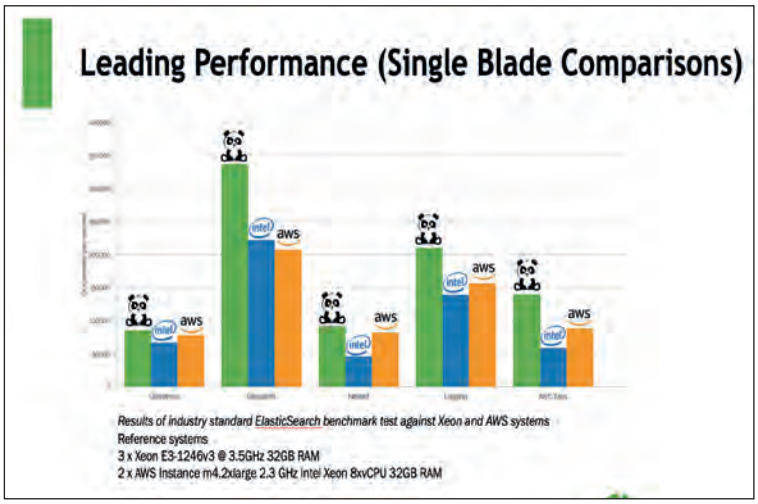
Many of the countries located in cooler climates, have benefited from their naturally lower temperatures, their ability to generate renewable energy, their low populations, tax incentives, that together enable them to deliver capabilities in this area.

Not surprisingly, these strategies do not work well in hot climates or with existing or smaller data centers. We expect to see more of the latter as edge computing becomes more common. The reduced latency requires more local computing power. Ironically, the latency can be attributed to moving data centers further away from population centers to take advantage of those cooler climates.

Additionally, new heat mitigation strategies are becoming increasingly more expensive. At some point, the upside of pursuing new technologies must make economic sense. There is recognition that PUE does not go far enough, pushing the drive towards net-zero data centers, where there is no impact on the environment and the emphasis is on energy conservation. One suggestion presented at the 2020 World Economic Forum in Davos, was put forward by the Swiss Federal Office of Energy who has been working with the IT industry. It suggests embracing new criteria such as electricity supply sustainability, recycling capabilities of waste heat, local storage utilization and how much of the server is virtualized.

Other institutions are now looking beyond simple infrastructure savings and at the basic underlying cause: the generation of heat itself. Great strides forward can be made by upgrading compute platforms. Legacy applications do come with challenges and these institutions also recognize that this may be problematical due to the sheer volume of servers that need to be upgraded and the complexity of that task. The assumption is to simply replace an existing server or group of servers, with a newer model built on the same traditional energy inefficient architecture.

All of these strategies are important, but it is necessary to question the basic assumption that servers need to have big, fast and very hot energy inefficient processors at their core. This thinking is a continuation of the traditional server architecture found in most data centers today, which is derived from the 1980's desktop PC. To meet this challenge, revolutionary High Throughput Computing (HTC) systems specifically designed for modern highly parallel workloads have been developed. They consume a



fraction of the energy of traditional server architecture, using embedded systems methodologies. Imagine the impact to the mitigation strategies discussed when applied to a computing platform such as this. The savings would be significant, even in smaller edge data centers, or in hot climates.

Now is the time to consider data center options that have been created to fundamentally rethink the foundations of server design and reduce energy consumption and its associated heat. Moore's Law is coming to an end, and while PUE is an important concept in measuring efficiency, it will only take the world so far in true data center power use reduction. Transformative server design is the next leap in both support for modern application design and delivering high throughput as well as a significant reduction in data center energy consumption at source.

As we look to the future and as the world's server and data center needs continue to grow, it is imperative that we seek out and embrace more sustainable, energy efficient and innovative technology to do the computing work. As we continue to see the need to tackle climate change grow alongside the need for more and more data center computing power, we are seeing groundbreaking technology making its way to the marketplace, built to specifically respond to these competing, urgent needs.

Bamboo Systems has created a patented server architecture called Parallel ARM Node Designed Architecture (PANDA) that delivers the highest compute throughput with 25% of the energy consumption in 10% of the physical space. Bamboo Systems servers are linearly scalable, due to dedicated resources for each processor, and achieve an industry leading high density by locating multiple servers inside a single blade, and multiple blades inside a single chassis. When combined with modern data center heat mitigation strategies and the continuing use of renewable energy, Bamboo Servers enable Net-Zero Data Centers while still delivering the needed throughput for today's compute needs.

Angel
BUSINESS COMMUNICATIONS

WEBINARS

Specialists with 30 year+ pedigree and in-depth knowledge in these overlapping sectors:

Expertise: Moderators, Markets, 30 Years + Pedigree
Reach: Specialist vertical databases
Branding: Message delivery to high level influencers via various in house established magazines, web sites, events and social media

Semiconductor (Silicon/Compound)
Publications include: Compound Semiconductor, Silicon Semiconductor, CS China, SiS China

Power Electronics
Publications include: Power Electronics World

Future Mobility
Publications include: TaaS Technology, TaaS News

Data Centres
Publications include: DCS Europe, DCS UK, SNS International

SmartSolar UK & Ireland
Publications include: Solar and Power Management, Solar UK and Ireland

Sensors
Publications include: Sensor Solutions Magazine, Sensor Solutions International

Digitalisation
Publications include: Digitalisation World, Information Security Solutions, Managed Services

Photonics
Publications include: PIC Magazine, PIC Conference

Expert Moderators

Dedicated technical and time-served experts/editors

MARK ANDREWS
Mark Andrews is technical editor of Silicon Semiconductor, PIC Magazine, Solar+Power Management, and Power Electronics World. His experience focuses on RF and photonic solutions for infrastructure, mobile device, aerospace, aviation and defence industries

PHIL ALSOP
Journalist and editor in the business to business publishing sector for more than 30 years currently focusing on intelligent automation, DevOps, Big Data and analytics, alongside the IT staples of computing, networks and storage

JACKIE CANNON
Director of Solar/IC Publishing, with over 15 years experience of Solar, Silicon and Power Electronics, Jackie can help moderate your webinar, field questions and make the overal experience very professional

DR RICHARD STEVENSON
Dr Richard Stevenson is a seasoned science and technology journalist with valuable experience in industry and academia. For almost a decade, he has been the editor of Compound Semiconductor magazine, as well as the programme manager for the CS International Conference

For more information contact:

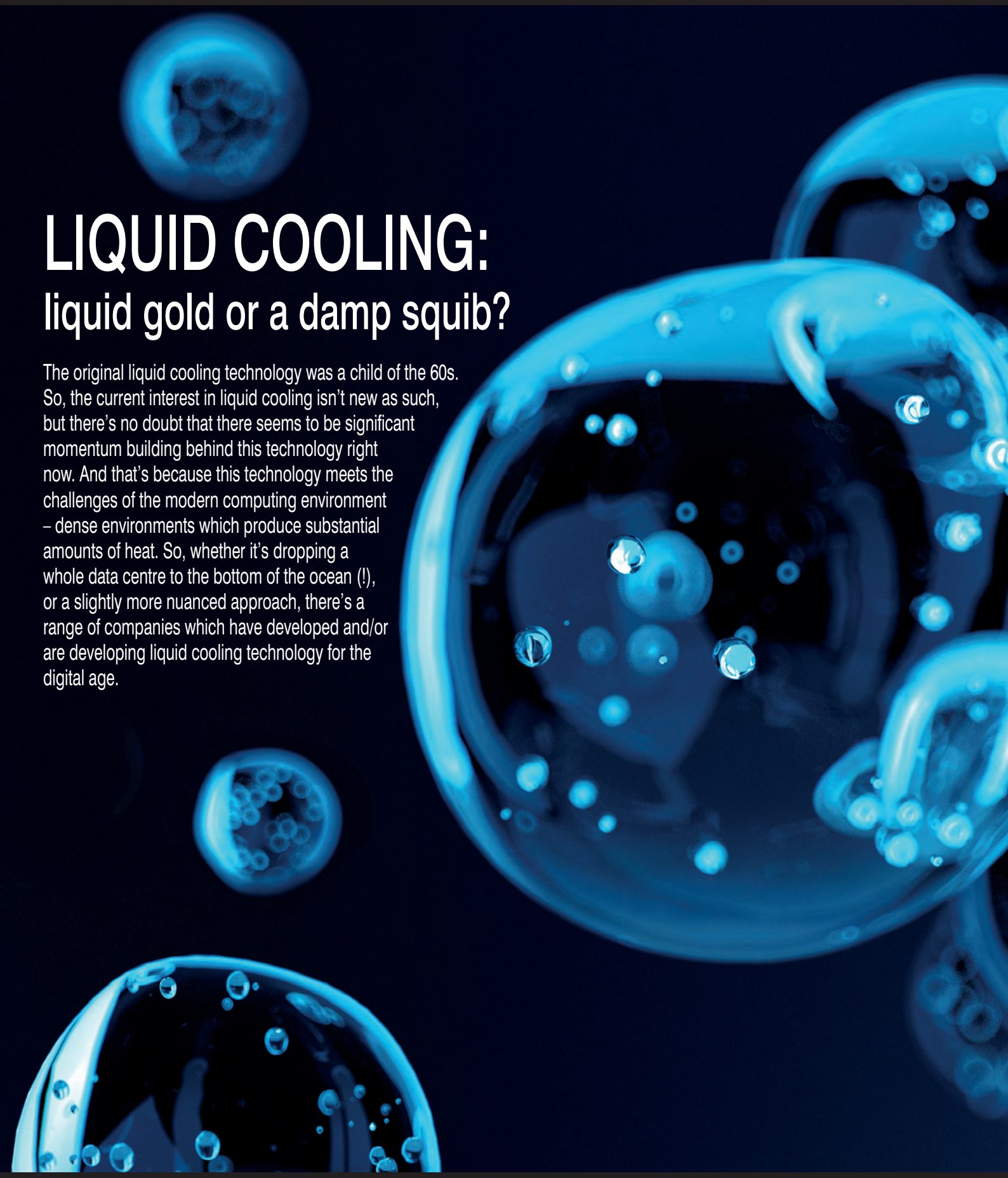
Jackie Cannon **T:** 01923 690205 **E:** jackie@angelwebinar.co.uk **W:** www.angelwebinar.co.uk
6 Bow Court, Burnsall Road, Coventry, CV5 6SP. UK
T: +44(0)2476 718 970 **E:** info@angelbc.com **W:** www.angelbc.com

DW DATACENTRE SOLUTIONS

LIQUID COOLING:

liquid gold or a damp squib?

The original liquid cooling technology was a child of the 60s. So, the current interest in liquid cooling isn't new as such, but there's no doubt that there seems to be significant momentum building behind this technology right now. And that's because this technology meets the challenges of the modern computing environment – dense environments which produce substantial amounts of heat. So, whether it's dropping a whole data centre to the bottom of the ocean (!), or a slightly more nuanced approach, there's a range of companies which have developed and/or are developing liquid cooling technology for the digital age.



Data centre liquid cooling market worth over $3bn by 2026

According to a recent study from market research firm Global Market Insights, the global data center liquid cooling market has emerged as a prominent investment ground lately for potential stakeholders, given the increasing workloads, storage densities and computing requirements of organizations worldwide. With servers across data centers becoming more sophisticated, the one major trade-off is the amount of heat they generate during their operation.

Data Center Liquid Cooling industry is set to surpass USD 3 billion by 2026, according to a new research report by Global Market Insights, Inc.

THE MATTER of overheating becomes a major point of concern when the issue results in a loss in power causing the servers inside the data centers to fail, bringing operations to a halt and costing the organization millions of dollars. This has made efficient & reliable operations of the cooling, power & support systems extremely crucial for ensuring continuous flow of data across these mission-critical facilities, a fact that has served to majorly augment data center liquid cooling market.

The increasing demand for digital services and unprecedented growth in AI and machine learning has helped push data center liquid cooling market share to new heights lately. Data centers demand efficient cooling systems to reduce power consumption as over 30% of the power is used for cooling purposes. Processor performance is also severely affected by overheating resulting in slower operations.

However, with businesses increasingly shifting towards mainframes and supercomputers, liquid cooling technology seems to be an ideal solution for modern data centers. Furthermore, the consistently changing demands for new, uninterrupted digital has been prompting data centers to adopt innovative technologies that will help boost data center liquid cooling market in the coming years.

The growing implementation of artificial intelligence (AI), big data, cloud, and machine learning in newly developed software solutions has also been driving the industry growth. With the development of more powerful chips capable of rapidly processing data for new complex applications, the amount of power consumption has exponentially increased leading to more component heat generation.

For instance, an increasing number of GPUs and FPGAs are nowadays pushing systems to their limits for smoothly running heavy applications like AI, big data analytics, HPC, media streaming, machine learning which in turn generate high amounts of heat. Reports claim that the heat profile for many GPU based servers is double than that of traditional servers.

Speaking of the regional expanse of the data center liquid cooling market, it would be prudent to state that APAC and Europe are two prominent geographical revenue pockets for the industry.

Asia Pacific data center liquid cooling industry will witness significant growth over the ensuing years, driven by the increase in data capabilities brought forth by the growing digitalization in every possible industry sector.

The expansion of the Europe data center liquid cooling industry can be majorly credited to the massive presence of numerous colocation facilities across the continent. Powered by the advent of robust digitization across the industries such as BFSI and medical care, in tandem with the rapid adoption of machine learning technology in the region, Europe data center liquid cooling industry share is set to soar high in the years to come.

Liquid-based cooling solutions have more heat removal capabilities & can reduce a data center's power consumption by about 70%. The paradigm shift toward high energy consuming applications and the deployment of AI, big data and machine learning will mandate the adoption of liquid cooling solutions for smoother operations in the years ahead.

CELEBRATING 10 YEARS OF SUCCESS

The 2020 DCS Awards feature 31 categories across FOUR groups.

THE DCS AWARDS are now firmly established as the data centre industry's premier annual celebration of all that is great and good. End user projects, product innovation and individual excellence are all recognised in an evening that pays more than lip service to the idea of data centre and IT convergence. So, the award categories cover both the facilities and IT aspects of the data centre, recognising the achievements of vendors, their business partners, their staff and their customers.

Getting involved with the DCS Awards couldn't be easier. Take a look at the award categories, and make sure to nominate your company, a customer, or maybe an individual – better still all three (!) – for a chance to be recognised for outstanding achievement when it comes to projects, product innovations and individual contributions within the data centre industry.

Once you've made your nominations, make sure to book a table for the Awards night. You wouldn't want to win an award and not be there to collect it! (And even if you don't win an award on the night, there's a cocktail reception, three course meal and a top comedian to entertain you – we have a track record of booking individuals on their way to the top of the comedy circuit).

To 'We look forward to welcoming you to the Awards night in December.

NOW, GET NOMINATING!

WHY ENTER?

MAXIMISE VISIBILITY
Free PR opportunities with 5 months of marketing utilising the Digitalisation World portfolio.

POSITIONING
Position your product or service as an innovator in your field.

INCREASED CREDIBILITY
An award win, shortlisting or nomination acts as a 3rd party endorsement.

NETWORKING
Over 300 industry leaders attend the gala awards evening.

COMPETITIVE ADVANTAGE
Gain an advantage on your competitors.

NOMINATION IS FREE OF CHARGE

The DCS Awards panel will validate entries and announce the final shortlist to be forwarded for voting by the readership of the Digitalisation World stable of publications in October 2020.

The winners will be announced at a gala evening at the LEONARDO ROYAL ST. PAUL'S HOTEL, London on 10 December 2020.

DCS AWARDS 2020

www.dcsawards.com

Supported by

dca The data centre trade association

3M - Novec profile

As the volume of data being processed continues to increase, data centre operators are seeking alternatives to traditional air cooling. Choosing the right cooling solution is a balancing act between computing power density, scalability, the environment and cost considerations.

BY ZOLTAN HANNAUER, MARKET DEVELOPMENT MANAGER AT 3M, DATA CENTRES, EMEA.

THIS IS WHY many data centre professionals are looking at single and two-phase immersion cooling to address these increasing challenges. When optimised, immersion cooling can manage much higher levels of heat generation than air cooling and a far greater density of IT hardware, while reducing overall footprint and the amount of additional equipment required to keep a data centre cool. It can also significantly reduce both energy and water consumption.

Modern immersion cooling relies on innovative cooling fluids designed for high thermal efficiency and high compatibility with delicate electronics as the foundation of its success. While hydrocarbon oils such as mineral oil are still in common use, modern fluids such as 3M™ Novec™ Engineered Fluids bring a range of benefits to data centre owners: they are non-conductive, non-ozone depleting and non-flammable, with low toxicity and low global warming potential (GWP). They also leave no residue and are non-corrosive.

Almost 70 years of cooling innovation

Novec Engineered Fluids from 3M reflect decades of research and development into fluid chemistry. Back in the 1950s, 3M introduced its first dielectric fluorochemical heat transfer fluids for direct contact cooling for transformers and military avionics, followed in the 70s and 80s by direct cooling developments for other applications like supercomputers. Then, in 1996, Novec Fluids were developed as a new class of speciality fluids designed to be replacements for ozone-depleting substances such as CFCs and greenhouse gases such as HFCs. They were quickly adopted in a range of applications, including component cleaning, fire suppression, and – of course – liquid cooling in data centres.

While widely adopted, many organisations may not even know they are using Novec Fluids from 3M, because they are just one element of the data cooling system, with a third-party specialist integrator providing the tank, expertise and anything else required. However, they can drive substantial savings. For instance, an independent study commissioned by 3M found that up to 95 per cent energy savings for cooling are achievable by using fluorochemicals for two-phase immersion cooling in data centres, while occupying 10 times less space than air cooling. Deployments have demonstrated power densities to up to 250kW per rack, with Power Usage Effectiveness (PUE) of as low as 1.02.

Different types of liquid cooling

With its 3M™ Fluorinert™ Electronic Liquids and 3M Novec Fluids ranges of products from 3M, the Company also offers solutions for both single and two-phase immersion cooling. In single-phase cooling, the fluid remains in its liquid phase.

Electronic components are immersed in a sealed tank, where the heat from them is transferred to the fluid. Pumps are most often used to flow the heated fluid to a heat exchanger, where it is cooled and cycled back

into the enclosure. In two-phase immersion cooling, the fluid is boiled and condensed to create a vapour that rises from the liquid and condenses on a heat exchanger within the tank. This reduces the need for most of the interface materials, heat sinks, fans and other equipment used in traditional cooling.

Both approaches have their own benefits, depending on the data centre concerned, and in some instances, direct-to-chip cooling with Novec Engineered Fluids from 3M may be a preferred option. The fluid never makes direct contact with electronics and can be used with both single and two-phase technologies. However, for the vast majority of current and future data centre applications, particularly at-scale in high-density environments such as hyperscale, and edge/5G, immersion cooling will likely be the most viable option.

Sustainability and safety

Of course, performance is just one element: data centre professionals are under increasing pressure to consider both environmental and human safety factors. Novec Engineered Fluids from 3M have low global warming potential and zero ozone depletion, making them a good choice for environmentally conscious data centre owners. Plus, they have gone through rigorous toxicity testing in line with global standards.

They are not classified as hazardous under the Global Harmonized System of Classification and Labelling of Chemicals (GHS) and they have been accepted for use in targeted applications by the US Environmental Protection Agency (EPA) and the REACH regulatory body of the European Chemical Agency (ECHA).

Cooling solutions for the future of data centres

As data centre owners and operators work to balance factors including safety, the environment, performance, scalability and total-cost-of-ownership in a competitive industry, cooling technology is likely to be a key differentiator. Smart, safe, sustainable fluids such as Novec Engineered Fluids from 3M can help enable that evolution from traditional air cooling to immersion cooling solutions designed for the data centres of the future.

Two-phase immersion cooling: The essential facts

With data centres quickly evolving , finding smarter ways to keep them cool now and in the long-term is vital. Many organisations are looking to liquid two-phase immersion cooling (2PIC), but there remains some confusion about how it works and its benefits in practice and at scale.

BY ZOLTAN HANNAUER, MARKET DEVELOPMENT MANAGER AT 3M, DATA CENTRES, EMEA.

What is immersion cooling?

There are two major types of liquid cooling for data centres: single-phase and two-phase cooling. These can be implemented either through technologies such as direct-to-chip cooling, where the fluid never directly contacts the electronics, or by directly immersing the electronic components into the fluid, known as immersion cooling.

In single-phase immersion cooling, electronic components are placed in a sealed tank of fluid. The heat they generate is efficiently transferred to the fluid, and pumps are often used to flow the heated fluid to a heat exchanger, where it is cooled and cycled back into the enclosure.

Single-phase immersion cooling is well-suited to many of today's data centre requirements. However, as next-generation data centres simultaneously need to provide improved performance, lower costs and higher computing density, many data centre owners are looking into the potential benefits of two-phase immersion cooling (2PIC).

In two-phase immersion cooling, components are submerged in a non-conductive liquid bath. The fluid pulls the heat away from the components until it boils and creates a vapour. The vapour in turn rises from the liquid and condenses on a heat exchanger within the tank, then flows back into the bath. Since the phase change carries the fluid straight to the heat exchanger, it does not need pumps, fans or much of the other equipment used in traditional cooling.

Two-phase benefits

2PIC provides some clear advantages to data centre owners: the ability to deal with much higher temperatures, better density and use of space, simplified design and the need for a lot less equipment. For example, an independent study commissioned by 3M of the potential impact of using 2PIC instead of air cooling in a planned data centre project found benefits including:

- A 46 per cent reduction of the space taken up by cooling: from almost 3,500 cabinets to 216 immersion tanks.
- Reduced capital expenditure for construction
- Reduced annual operating expenses
- Reduced water waste
- No need for raised floors for air cooling

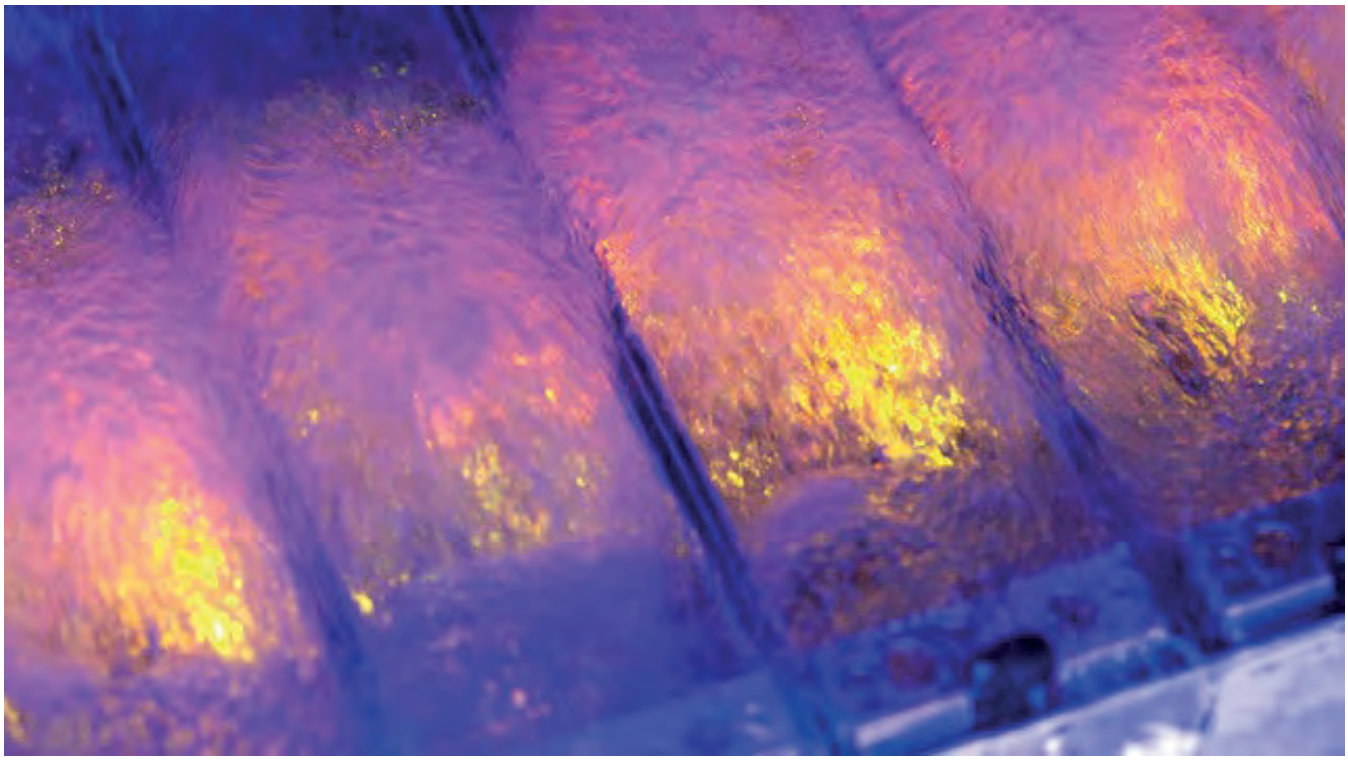
Plus, the study found that 2PIC can provide as much as 95 per cent cooling energy savings over air cooling (a 95 per cent reduction in Power Usage Effectiveness (PUE): the ratio of total energy used in a data centre to the energy used by the computing equipment). Large scale real-world implementations have borne this out, with PUEs as low as 1.02, providing highly efficient cooling to data centre equipment.

The heat extracted by 2PIC can also be reused for maximum efficiency, contributing to comfort heating for data centre operators or other useful purposes.

Factors to consider

When comparing single-phase and two-phase immersion cooling solutions, the fluid used for immersion cooling and the total cost and complexity of system ownership are also important considerations.

For example, single-phase immersion cooling typically requires simpler tank designs. Plus, both fluorochemical fluids and hydrocarbon-based fluids



like mineral oil can be used in single-phase cooling because the boiling point is not reached, offering data centre owners a range of fluid options. Hydrocarbons such as mineral oil and other oils are relatively less expensive to purchase and easier to contain due to their low vapour pressure. On the other hand, they are often combustible and can leave residue on the hardware. Fluorochemical fluids can offer benefits such as being non-combustible, chemically stable and leaving no residue behind, but their higher vapour pressure makes them harder to contain.

Two-phase immersion cooling typically has more complex tank designs and cannot generally be conducted with hydrocarbon-based oils. However, when fluorochemicals are used, this method enables much greater heat transfer efficiencies through the boiling process: for example, 500W ASICS have been successfully overclocked well over 750W. In addition, a lot less cooling infrastructure is needed to move the fluid around inside the tank, saving on equipment complexity.

Environmental considerations

As hyperscale, edge, 5G and other data centres continue to expand, their environmental impact is also important to consider. Both data centre owners and their customers are increasingly aware of environmental considerations and third party supply chains, and next-generation data centres are likely to be a key element in meeting their sustainability targets. Some factors are:

- Water usage, particularly in regions with ongoing water shortages
- Compactness, particularly for edge data centres in environments where real estate is scarce or expensive.
- Power usage, particularly as measured by Power Usage Effectiveness, which has increasingly been monitored and mandated by regulators
- Fluid choice, particularly whether the fluid has no ozone depletion potential, low toxicity, and complies with relevant regulations such as REACH in the EU

Are you ready for 2PIC?

Not all data centres are ready to adopt 2PIC, particularly if they intend to stick with legacy designs and equipment that either do not adapt well to immersion cooling, or still require a high degree of air-cooled ancillary equipment. Alternatively, single-phase immersion cooling might be an evolutionary step towards the high-efficiency 2PIC future.

However, data centres are in an unprecedented phase of their evolution. Cryptocurrency-mining is one market sector driving the current wave of 2PIC adoption. Moreover, newly built data centres have the opportunity to adopt a more modern approach from the ground-up, including modern cooling.

2PIC continues to evolve, and so do its uses. Beyond servers, other equipment can be placed in the immersion tank, such as redundant power and network switches. Even back-up batteries could be immersed in the same tank, which could both simplify the design and save space. Regardless of future developments, 2PIC has a lot to offer many of today's data centres, helping to manage density and overall costs while keeping up with escalating demand.

Asperitas – profile

After nearly 20 years in IT and a successful circumnavigation with sailing yacht Helena, Rolf Brink founded the disruptive clean and high-tech company, Asperitas, together with Markus Mandemaker. His background in product development for cloud architectures and datacentre infrastructure, combined with a passion for cleantech innovations like liquid cooling, are the foundation of the development of Immersed Computing® within Asperitas.

AS LEADING immersion cooling specialists, Asperitas provides cutting edge immersion cooling solutions for energy efficient and high density datacentres globally. The first solution was launched in 2017 after an extensive R&D phase, with an ecosystem of cutting edge partners since 2014.

In 2018, Asperitas won the international New Energy Challenge competition organised by Shell, Rockstart and YES!Delft. In May 2020, the World Economic Forum published a paper on Transformational Energy Innovations of the last 10 years, selecting Asperitas technology as the only datacentre energy efficiency solution in their ground-breaking paper.

Explanation of liquid cooling solution

The Asperitas' 15" and the recently launched 21" ready AIC24 immersion cooling solutions are designed to enable efficient, sustainable and high performing datacentres. The new AIC24 solution allows for twenty-four 1U 21" wide servers to be integrated and two additional switches, and is designed to meet the demand for high density compute both on a system and server level.

The 21" wide server space allows OEM's and integrators flexibility to design cutting edge servers utilising the full potential of the immersion cooling concept developed by Asperitas, Immersed Computing®. Both units are driven by the natural convection method, and can therefore be deployed in high ambient temperatures, even up to 45 degrees Celsius.

The liquid cooling concept is spearheaded by the fully immersed cooling of the IT within a synthetic fluid, rather than cooling it using traditional air methods. The fluid is medicinal quality GTL (gas-to-liquid) provided by Shell. The next generation hardware for emerging digital technologies is of a high performance, but also very powerful, going up to 400 watts per chip.

This amount of energy turns into heat and requires advanced cooling methods to even maintain performance. An entire new model is needed to design and operate datacentres sustainably.

Compared to an average air-cooled datacentre, the AIC24 module takes up 80% less physical footprint and 95% less cooling energy for datacentres. Asperitas' technology is trusted by cloud providers, telecom operators, global enterprises like financial institutions and high performance computing users like universities in multiple countries.

The liquid cooling concept is spearheaded by the fully immersed cooling of the IT within a synthetic fluid, rather than cooling it using traditional air methods. The fluid is medicinal quality GTL (gas-to-liquid) provided by Shell. The next generation hardware for emerging digital technologies is of a high performance, but also very powerful, going up to 400 watts per chip

Liquid cooling Q & A with Asperitas

Q: s liquid cooling the most efficient technology solution to the problem of increasing data centre heat levels and loads, as high density computing, for an increasing number of applications, becomes more prevalent?

A: Yes, but most likely in conjunction with a variety of technologies which will address the industry's needs. Each datacentre consistently churns out warm air which is too low in temperature to efficiently distribute and transport. Asperitas Immersed Computing® technology makes use of the natural convection driven fluid circulation within, allowing for warm water cooling. This unique approach enables free cooling up to 45 degrees Celsius.

Compared to an average air cooled datacentre, liquid cooling facilitates 5-10 times as much density. It's viewed as the best cooling method for holistic cooling requirements for IT kits and because of this, it can deal with relatively high densities very effectively.

Immersion cooling is offering the most efficient method as all hardware components are directly cooled and offering the most flexibility as this one method can facilitate all types of hardware.

Q: Liquid cooling has been around as a technology since the 1960s. Why is it still struggling to become a more accepted data centre cooling option?

A: In the 1960's, liquid cooling was pioneering. The real urgency for liquid cooling arrived only in the last few years with Thermal Design power (TDP) going up. Liquid cooling has impacted how we design and operate datacentres for the better, therefore it introduces a new model and not just an optimisation of current, air-cooled, facilities. With every innovation, a new model allows for more opportunities and impact, but the adaption requires time as the industry ecosystem needs to move with it. Optimised hardware for immersion is a good example, eg, Asperitas has a close collaboration with OEMs and integrators to validate and optimise hardware and future generation processors for immersion.

On another note, it may be hard for datacentre operators to see a clear path on how to actually move towards liquid cooling, while immersion cooling simplifies datacenter design with limited requirements. Ultimately, the adjustment needs to be cost-effective, which it is now in cases where density, efficiency and/or sustainability are required. As a medium for cooling liquid is a no-brainer, however it was our specific focus to develop and offer an enterprise level datacentres solution that would qualify for Tier 4 environments and other high standard datacentres. That is a big step up for immersion, where we see it now being used by such datacentre operators and users.

Q: Aside from the heat increases outlined in 1. above, what other factors will contribute to a greater uptake of liquid cooling technology over time?

A: We see a combination of three main drivers. The first one is performance requirements, higher density in the server, rack and facility plus an increase on utilisation, meaning that running processors at a higher performance can be done without any risk of overheating.

The second is efficiency of datacentre build and operations – datacentres face several constraints globally: scarcity of power, space, and water supply. The last driver is sustainability. From energy efficiency to heat reuse potential, immersion cooling is offering the highest potential when done well. Datacentres need to work towards a huge reduction in energy and physical footprint in order to keep up with the ever-growing demand.

With liquid cooling, there is a flexible deployment that in our case includes a comprehensive monitoring system and it is clean and fully contained. Downsized cooling and power requirements also reduce the amount of maintenance on these systems.

Q: How does liquid cooling work alongside 'traditional' cooling solutions within the data centre? In other words, it seems to be primarily targeted at servers, so the other IT hardware – storage and networking kit, for example – still relies on traditional cooling solutions?

A: We will see liquid cooling as part of hybrid environments for a long time. Asperitas has designed a solution in such way that it can be easily deployed in a traditional datacentre. For operators, it is the easiest way to facilitate high density and performance applications in an existing datacentre. What we are really encouraging is a datacentre set up where the right liquid cooling solutions are used for different

applications. Rear door coolers are great for storage and networking devices, while immersion cooling is optimal for compute. In 2017 we published a whitepaper called “Datacentre of the Future” where we outlined this concept for an optimised datacentre for heat reuse. In many cases immersion cooling will not stand alone, because not all equipment makes sense in immersion. For example, high connectivity devices such as network switches - having 40 switches in a single tank is not ideal.

Rear door coolers are great for that purpose. The key is to consider different technologies for different applications and connect them together effectively and efficiently. One could say that the air-cooled IT equipment itself is inefficient, as well as the cooling infrastructure to provide low temperatures with chillers.

Q: And how does a data centre have to be re-designed in order to optimise this combination of air and liquid cooling?

A: Our immersion cooling solution can be cooled with warm water, even 45 degrees Celsius, meaning it could utilise the output of the traditional datacentre space and increase the delta for the entire operation for more efficiency. Immersion cooling is not affected by high room temperatures and is also not impacting its surroundings so it can be easily used in combination. We also have seen datacentres utilise space in their datacentre originally not designed as white space. Today, datacentres are facing boundaries regarding floorspace, access to (renewable) energy, cooling costs, complexity and cost of build. This can greatly affect the performance, efficiency and expense of the whole operation. Liquid cooling takes up to 80% less physical footprint and modules can be placed on the same raised floor. High density datacentres can be built on only 1 fifth of the floor space using liquid cooling. A combination of air and liquid cooling would simply mean having space within the existing datacentre to house some modules. These can simply be placed alongside traditional air-cooling racks if required

Q: And is retrofitting liquid cooling possible and/or sensible, or does it really only work with the latest servers and in a newly designed and built data centre?

A: Retrofitting of hardware is an option. For example, we see this with pilot projects. However, as always, optimised servers offer the most potential for reliability and performance. In both cases we have a standard certification process for OEM’s, integrators and users to validate a server for immersion and identify potential for optimisation. Immersion cooling can be applied in both brown and green field datacentre projects, there is no limit, but naturally the highest level of efficiency, performance and sustainability can be achieved in projects entirely designed around liquid cooling solutions.

Q: Much is made of the waste heat reuse potential of liquid cooling solutions. How important is this as part of the overall liquid cooling package – is it an essential component, or just a 'nice to have' option?

A: In most scenarios, this is a "nice to have". Liquid cooling has a great business case without it. However, heat is in great demand in many regions and cities and it's difficult to supply it in a sustainable manner. Datacentres can be great resources for heat and in multiple regions, and in the Nordics it is strongly encouraged to make heat available for district heating.

With an open mind, one could see the potential of hyperscale and edge datacentres supplying heat to the grid and users directly. We strongly believe this integration of energy and digital infrastructure is what will need to happen to have a fully scalable and sustainable industry. The Sustainable Digital Infrastrucuture Alliance recently published an inspiring report on this development and the opportunity for utility providers.

In summary, in order to accelerate one's mission to enable sustainable datacentres on a global scale, it should be a priority. Being fully heat reuse ready (ie more than 98% energy being transformed into hot water) means that one reduces the footprint of a third-party with the heat demand, enabling the transformation of datacentres from heavy energy consumers into energy suppliers and contributors to the circular economy. In combination with the other benefits of immersion cooling, like limited required floorspace and no noise, it will allow for a great integration in an urban environment close to the user.

Q: In summary, can you outline how and where liquid cooling within the data centre makes operational and/or financial sense and how this cost/technology equation will make liquid technology more attractive over time?

A: Liquid cooling makes sense for datacentres when high density or performance is a requirement. The term "high density" can be misleading, as many people could think we are talking about 20kW or more per rack. We have designed our immersion cooling solutions in such way that most datacentres will see a positive business case from 6kW or more. At the same time, the high-end processors for high performance computing need liquid cooling for optimal utilisation according their own specifications. If these requirements are identified and the installation of multiple racks is optimum, immersion cooling is a seriously good option.

As well as this, there are several scenarios that immersion cooling can solve better than any other cooling method. Our solutions offers free cooling in ambient temperatures of up to 45 degrees, so you can imagine what that looks like for datacentres in warm climate zones. Today, all reports point to APAC as the fastest growth market for datacentres, so it is essential to apply a method allowing to standardize efficiency.

In general, high sustainability standards offer strategic value for organisations and at the moment it can still be a competitive advantage for datacentres. What we offer is datacentres that can operate close to energy neutral and even create an additional revenue stream by supplying heat. For edge computing specifically, liquid cooling makes a lot of sense, and might be actually the standard model to allow flexibility in which to deploy micro datacentres without a need for complex cooling and room design

Five delicious flavours of liquid cooling for data centres?

According to research company MarketsandMarkets, the data center liquid cooling market size is expected to grow from USD 1.2 billion in 2019 to USD 3.2 billion by 2024. Factors driving forecast growth include increasing need for energy-efficient cooling solutions and the growing demand for compact and noise-free solutions. As well as the need for lower operating costs and for better processor overclocking potential.

BY DAVID CRAIG, CEO, ICEOTOPE.

LIQUID COOLING use has until recently been primarily centred on niche applications such as high performance computing requirements (HPC) and the use of power-hungry CPUs and GPUs for artificial intelligence (AI), machine learning (ML), and advanced robotics. However, AI, for example is anticipated to become increasingly mainstream.

One of the worst kept technology secrets is the power requirements of next generation CPUs hitting the IT marketplace in the next 24 months. Coupled with the data processing needs of all AI-enabled applications, 400W – 450W microprocessors will be a considerable challenge for the cooling resources of many legacy data centres, enterprise and scientific environments. With increasingly high performance and power-hungry applications, air is not effective as a cooling medium. By contrast, liquid cooling is increasingly being looked at as the only way to ensure IT equipment can be reliably operated. Liquids have a much greater capacity to capture heat by unit volume; and remove heat more efficiently and allow chips to work harder (i.e. increased clock speed).

In addition, heat is rejected to the atmosphere either via dry coolers or, in hotter environments, cooling towers. Where the infrastructure exists, liquid cooling also enables data centre heat to be reused in other applications such as district heating schemes.

Liquid cooling comes in a selection of five flavours

The question for those new to liquid cooling (the technology itself is not new, having been around since the days of mainframe computing), is what are the different types of liquid cooling available on the market today? According a helpful white paper by Iceotope partners, Schneider Electric, there are two basic forms of liquid cooling, direct to chip (also called conductive or cold plate) and immersive. From these two categories come a total of five main liquid cooling methods:

Direct to chip single-phase liquid cooling

Liquid coolant is taken directly to the hotter components (CPUs or GPUs) using a cold plate directly on the chip within the server. Other electronic components on the server board are not in direct physical contact with the liquid coolant, although some designs include cold plates around memory modules. Fans are still required to provide airflow through the server to remove residual heat. Dedicated gamers are quite familiar with cold plate technology for direct cooling of microprocessors.

Water or a dielectric liquid can be used as the coolant to the cold plates. The use of water does introduce a downtime risk if there is leakage, but leak prevention systems (LPS) can be used to keep the water loop at slight vacuum to help mitigate this. Fluid manifolds are installed at the back of the rack to distribute fluid to the IT equipment, the interface between server and the manifold is typically achieved via a non-spill, non-drip coupling to ensure cleanliness and safety of the installation.

Direct to chip two-phase liquid cooling

This method is just like the direct to chip single-phase method, except that the fluid utilised is two-phase, it changes from one state to another – i.e. from liquid to gas in taking away the heat. Two-phase is considered more effective than single-phase (in terms of heat rejection) but requires additional system controls to ensure proper operation.

Chassis-based liquid cooling (single phase)

The server boards are immersed in a single-phase dielectric liquid, and since all the electronic components are in contact with the coolant, all sources of heat are removed. All server fans can be removed since the electronics are placed in an environment which is inherently slow to react to any external changes in temperature, enabling near silent operation. In addition, the servers are within a sealed module (compatible with standard data centre racks), mitigating the risk of leaks and rendering the boards and components immune to the ingress of corrosive humidity and pollutants.

Tub/ open bath, single-phase liquid cooling

In this method IT equipment is completely submerged in fluid. Unlike traditional IT racks, where servers are horizontally stacked from the bottom to the top, tub immersed servers are pulled out on a vertical plane. In some systems, this method incorporates centralised power supplies to provide power to all the servers within the tub. The heat within the dielectric fluid is transferred to a water loop via heat exchanger using a pump or natural convection. This method typically uses oil-based dielectric as the fluid. Note that the heat exchanger may be installed inside or outside the tub.

Tub/ open bath, two-phase liquid cooling

As with the single-phase tub method, the IT is completely submerged in the fluid. The difference here is the use of two-phase dielectric coolant which changes from one state to another – i.e. from liquid to gas in removing the heat. Because of this phase change, an engineered fluid must be used in the tubs.

Choosing which form of Data Centre Liquid Cooling is the best fit

As mentioned earlier, many applications requiring highly dense data processing are suited to liquid cooling. However, for many data centre operators there is also the growing ecological and financial requirements to deliver energy efficiency and lowered costs. Across the board, liquid cooling is perceived as the optimal cooling solution.

Direct to chip and immersive liquid cooling offer considerable benefits to data centre owners compared to air cooling. For retrofit sites, rack-based solutions like direct to chip and chassis immersive liquid cooling provide the easiest transition. For new sites, and those in harsh environments such as Edge, immersive liquid cooling is the optimum approach as it can capture all the heat as well as isolate the IT from the surrounding atmosphere.

To find out more about the different types of liquid cooling technologies, please visit the Iceotope website and download the free white paper “Liquid Cooling Technologies for Data Centers and Edge Applications.”

Liquid Cooling Q&A with Gerard Thibault, CTO, Kao Data

Q: Is liquid cooling the most efficient technology solution to the problem of increasing data centre heat levels and loads, as high-density computing, for an increasing number of applications, becomes more prevalent?

A: I believe that liquid cooling is not a response to a single situation, it is undoubtedly a solution to serve more dense computational needs, but it is also an architecture that appeals to professionals with different mindset towards more effective data centre cooling configurations; those that could result in data centres becoming more integrated into the wider communities that they belong to through waste heat re-use.
Liquid cooling of course has many advantages, from driving up the spatial efficiency of air-cooled data halls to unlocking stranded capacity in the form of air-cooled components in racks, and providing greater or more efficient cooling capabilities for high performance computing (HPC). However, it also provides an efficient way of using real estate that could be left fallow as the rack or row spacing factor often has to increase in traditional data halls, to avoid air velocities increasing sharply and adversely affecting air delivery to racks.

In raised floor environments, where floor grille air exit velocities increase, it can mean servers at the bottom of the rack become starved of sufficient air-cooling. Simply increasing the cooling air velocity is not the solution for many practical and safety reasons. Fire alarm detector heads, for example, don't operate effectively beyond air speeds of 4 m/s and water-mist sprinkler delivery systems are not reliably activated beyond 1 – 1.5m/s airflow. Therefore, the spray patterns and fire prevention systems are also likely to be affected by high velocities of air trying to reach the infrastructure more effectively.

Q: Liquid cooling has been around as a technology since the 1960s. Why is it still struggling to become a more accepted data centre cooling option?

A: Arguably, the technology came and went as the cost of air-cooled IT equipment (ITE) made it easy to deploy infrastructure within data halls using a common air delivery system to the infrastructure.
However, today we are reaching a dynamic shift with the application of new and more densely powered IT equipment, arguably, one where air can no longer cope as the sole solution, and is far less efficient. Furthermore, AI and Machine Learning are driving up processing densities, which in turn requires a more efficient way of transferring heat out of the ITE. Here liquid cooling technology and the reliability of server scale piping, chip heat-sinks and 'no-leak' flexible pipe connectors all mean that the re-introduction of liquid cooling becomes less risky and can be considered 'safe' for compute intensive applications.

Q: Aside from the heat increases outlined in 1. above, what other factors will contribute to a greater uptake of liquid cooling technology over time?

A: Liquid cooling appeals to data centre operators looking to push the boundaries of facility designs, efficiency and those working with customers leveraging power intensive processor and GPU technologies. However, with greater focus on data centre energy usage, efficiency and carbon emissions, sustainability has also become a key driver for the industry and is something we embrace at Kao Data. Sustainable operations are built into the fabric of our ultra-efficient facility, and are one of the reasons we utilise low energy consumption indirect air, adiabatic and evaporative coolers. These cooling systems do not need the power-hungry refrigerant compressors that were needed to meet the tight air temperature tolerances required by server manufacturers. As more data centres eliminate this expensive refrigerant based cooling equipment, liquid cooling technologies will

more directly correspond to sustainability objectives as well as offering a better cost benefit profile of the data centre operator. There are also other factors as mentioned earlier. With indirect water contact with on-board components such as chips and power components via efficient heat sinks or via plate heat exchangers in rack level and fully submerged cooling systems, much higher 'waste' water temperatures can be transferred to heat rejection equipment. This drives up the efficiency of cooling systems, by allowing the use of a similar reverse process – simple water to air heat rejection provided by dry air coolers.

Furthermore, using the energy dense heat rejection, water can now be more easily redirected to other external heat sinks, such as community heating projects, for homes, offices, and some industrial processes. Heat re-use is a way of 'doubling the benefit' of using the energy firstly for IT processing, then secondly, using the same energy to do other useful work in heating homes and other processes. Heat re-use via the use of ultra-efficient IEC's was far more difficult and required additional energy input, to drive up useful energy reclaim from very low temperature waste heated air. This is not the case with liquid cooling.

Liquid cooling appeals to data centre operators looking to push the boundaries of facility designs, efficiency and those working with customers leveraging power intensive processor and GPU technologies. However, with greater focus on data centre energy usage, efficiency and carbon emissions, sustainability has also become a key driver for the industry and is something we embrace at Kao Data

Q: How does liquid cooling work alongside 'traditional' cooling solutions within the data centre? In other words, it seems to be primarily targeted at servers, so the other IT hardware – storage and networking kit, for example – still relies on traditional cooling solutions?

A: Future developments may well see water cooled versions of ancillary devices and storage. Today we are seeing rack level immersion cooled power supplies, which traditionally, as the lower cost intensive part of the system, have been often left as air-cooled devices. As the power densities of server deployments increase, power supplies will either need to take up more space if left as air cooled devices, or transform to become direct liquid cooled as well. Until all the equipment within a facility is direct liquid cooled, a hybrid-mix of air and liquid cooling technology spaces will exist, with air accounting for maybe as much as 30% of the total, and perhaps reducing to 5% of the data hall load in time. Also, history informs us that the take up of equipment cannot be rushed and the market can be slow to embrace new technologies or architectures! ITE refresh cycles can often mean that for the majority of the market, there is a pause in new technology adoption, to ensure that ROI in existing air-cooled server equipment is fully realised.

Q: And how does a data centre have to be re-designed in order to optimise this combination of air and liquid cooling?

A: Often this will depend on the mix and relative densities within a facility! What may happen is that whilst the percentage of air-cooled equipment reduces, the actual absolute value may not. Or put another way, the density of a technology hall can be so greatly increased that in room air delivery might remain quite similar. For instance, an average data hall may deliver 2MW of air-cooled capacity, and in addition, deliver a further 8MW of liquid cooled capacity in the same technology space. That is only 20% air cooled (as opposed to the original situation of 100%. However, to accommodate that volume, the liquid cooled racks will require leak and condensation free liquid distribution systems sharing the same real estate.

Perhaps more importantly, the plant space, for both heat rejection and critical UPS power, will see significant increases due to the increase in server numbers. UPS rooms and generators will have to increase 5-fold and heat rejection equipment will be required to remove an additional 4 times the original heat capacity, but due to the temperature and ambient delta temperature available, this may not be a 4 fold additional space take.

Consequently, the overall data centre real estate model will change, driving down the traditional useful 'Net White Space' to gross floor area ratio. What is more significant is the overall IT Application process or IT Power Usage to gross floor space ratio. Hypothetically, a 2 MW data centre, based on 5kW racks, drives 400 racks at a rough average of say 2msq of data hall space, yet the overall space taken is 800msq for 2MW. Consider now a 10MW capacity in 800msq. Say 20% of the IT load is 5kW racks and

hence 80 racks or 160msq. If water-cooled racks at 44kW are deployed (driven by 63Amp connectors for IT power), then only another 180 racks are required, or approx. 360msq. In total the occupied space will hit a limit of 520msq, leaving about 280msq or about 1/3. However, this space will likely be eaten up by increased circulation space required to accommodate pipework for the liquid cooling.

Q: And is retrofitting liquid cooling possible and/or sensible, or does it really only work with the latest servers and in a newly designed and built data centre?

A: Clearly, the latest servers with the liquid cooling facility can be used in this developing architecture; however, we are seeing 'retro-fit' applications where standard servers are 'stripped' and rehoused in specifically designed enclosures allowing them to be adapted to immersion cooling individually at rack level. In this case, retrofit within the data centre is possible and practical. However, to redesign and retrofit a legacy air-cooled colocation facility for use with liquid cooling is likely impractical, costly and could cause too much disruption to tenants. As processors, chips and GPU power increases, legacy facilities may find themselves in a position where they cannot offer the power or cooling required. Metrics for the efficiency of retrofitting liquid cooling in a technology hall would need to be assessed on a case-by-case basis. As discussed the density will outstrip the space required for racks, so additional space will be available to pipework distribution systems. There is obviously a risk factor when working in live facilities, so again, the extent to which liquid cooling can be accommodated may depend on the layout and flexibility of the Technology Suites.

Q: Much is made of the waste heat reuse potential of liquid cooling solutions. How important is this as part of the overall liquid cooling package – is it an essential component, or just a 'nice to have' option?

A: I believe that heat re-use ability is an essential part of the equation. As data usage increases and more data centres are developed to meet demands, thereby using a greater quantity of energy, this needs to be reused with local communities; impacting more positively on the surrounding environment.
What is needed is for the wider community is greater governmental and environmental regulations or developer incentives, to deploy systems that will be able to consume the reused energy. This will help transform the digital revolution in to a society that benefits from an ecosystem with a common 'force for good', those like the smart cities that have been promoted in recent years!

This development also needs to be backed by market regulation around sustainable pricing for heat re-use. This is not 'free energy' on the back of the data centre sector, the system must encourage a healthy demand or price for reused energy, providing benefits for the community, and allowing the data centre sector to invest in the best technologies and still make a profit. In this way it allows the market to provide reliable and redundant heat rejection solution, for those times when the heat cannot be reused off-site!

Q: In summary, can you outline how and where liquid cooling within the data centre makes operational and/or financial sense and how this cost/technology equation will make liquid technology more attractive over time?

A: There are mny factors that will drive how economically viable liquid cooling will become. At the lowest level, the time taken for IT equipment refreshes and the introduction of reliable liquid to ITE interfaces alongside a way of safely removing the heat risk free, will become primary factors.

Government and regulatory changes, allowing heat re-use to be driven harder can provide clear environmental and economic advantages alongside sustainability gains. This could be the most-lengthy part of the cycle, taking at least a decade to achieve. Finally, the fastest up take will be driven by the processing requirements, the customer applications and the impact of energy on bottom lines. The ability to sell applications that use more dense processing, driving up the power density and hence energy consumption will become fundamental factors. Furthermore, the use of localised heat rejection equipment and rejection of 'waste heat' in to the atmosphere will be in front of the heat re-use cycle, but in itself, like the demand for renewable energy now being driven through the utilisation of schemes like 100% Renewable Energy Guarantees Origin (REGO) power, it will stimulate the pressure and demand for heat re-use, but the engine has yet to build momentum.

AIOPS SOLUTIONS ONLINE SUMMIT

ENABLING APPLICATION OPTIMISATION

The importance of proactive performance monitoring and analysis in an increasingly complex IT landscape.

25 November 2020

aiopssolutions.com

Midas Green Technology – profile

Midas Green Technologies may be the best company you never heard of or know the least about in the immersion cooling industry!

THAT IS BY DESIGN. When we formed our company and started in this industry in March 2011, we made the decision not to go to market until we had the best product on the market! We spent the next several years fulfilling that goal.

We were the second company to enter this industry. We thought we had a better idea and, as it turns out, we did!

Our Accomplishments:

We built and operated the worlds first totally immersion cooled data center in 2011.
Our patents have been granted in 23 Countries.
We were the first to put 2 cooling modules on every tank so each one could be fully redundant.
We were the first and still the only ones to allow for "concurrent maintainability" by allowing the cooling modules to be replaced in under 5 minutes without tools. Therefore, maintaining redundancy while a module is rolled away for scheduled or unscheduled maintenance.

We were the first to allow a continuous slow flow, so we do not have heat layering and stratification.
We were the first (and own the technology) to build in an inner wall for an accumulator to aid in cooling and cut down on external space.

We were the first to use square piping for return flow and make it reinforcement for the frame of the tank.
We were the first to build a tank specially designed for Crypto miners that could also handle standard servers.
We were the first to build a system made entirely of a metal shell with no PVC connections and to keep the dielectric within the footprint of the tank. Almost eliminating the risk of leaks during a fire and most daily operations.

We were the first to allow for additional rack units of space in a tank for "other IT" that DOES NOT take away from the compute available space.
We were NOT first to build a containerized solution, but we ARE the first to build a containerized solution from 400 kW's to 2 MW's using a hybrid evaporative cooling and dry cooling system all in one and containerized.

We can help you build a new construction data center, with identical processing capacity as air, for 60% to 70% of the capital, in 40% to 60% of the space and it will cut the operating cost by 50% and save you as much as 90% on cooling cost! We will also remove any problems caused by dust, tin whiskers, zinc whiskers, noise, changes in humidity and temperature hot spots.

We have been told we build the best quality in the industry and we are proud of this fact. We were given 7-800 words for this document, but we do not need that many to say that "we build the sky for you to put your clouds in"! Brick and mortar buildings or containerized solutions, if we build it, it works!

We built and operated the worlds first totally immersion cooled data center in 2011. Our patents have been granted in 23 Countries. We were the first to put 2 cooling modules on every tank so each one could be fully redundant. We were the first and still the only ones to allow for "concurrent maintainability" by allowing the cooling modules to be replaced in under 5 minutes without tools

Midas Green Technology Q and A

Q: Is liquid cooling the most efficient technology solution to the problem of increasing data centre heat levels and loads, as high density computing, for an increasing number of applications, becomes more prevalent?

A: Absolutely! Not only does immersion allow you to cut cooling costs by 90%, it also allows you to remove the fans from the servers thereby reducing critical power consumption. Eliminating the need for CRACS, raised flooring, hot/cold aisles and reducing the size of the power infrastructure not only allows you to significantly reduce your electric bill but also the capital investment.

Q: Liquid cooling has been around as a technology since the 1960s. Why is it still struggling to become a more accepted data centre cooling option?

A: It is widely misunderstood. It is a resume updating event. Get it right and you can update your resume with a star for what you did; get it wrong and you will need to update your resume! People are afraid to "mess with their core business and "take a chance"! In ten years, we have gone from "You do WHAT?" to "I've heard of that." to "we are looking at it as an option" as responses. Since we started, servers have become more dense (and continue to do so) and power savings more important.

Q: Aside from the heat increases outlined in 1. above, what other factors will contribute to a greater uptake of liquid cooling technology over time?

A: Longer meantime to failure, no dust, no zinc whiskers, no tin whiskers, no noise, much more efficient use of space (typically about 60% of space for air with same compute load). In addition, if you are so inclined, you can overclock processors in immersion more than is possible in air. This fact is very, very important to Cryptocurrency miners.

Liquid cooling needs solid state or helium sealed hard drives. If you have servers with spinning platters you have to change out the drives. Also, every server in air has debris inside. Dust, contaminants from the air and microscopic residuals from packing

Q: How does liquid cooling work alongside 'traditional' cooling solutions within the data centre? In other words, it seems to be primarily targeted at servers, so the other IT hardware – storage and networking kit, for example – still relies on traditional cooling solutions?

A: It can work just fine alongside "traditional cooling solutions" but it is a waste of money and space when immersion does not need CRAC's, raised floors, special fire suppression etc. to run efficiently. This is an economics decision. Immersion cooling offers its highest value at higher densities. While it is possible to immerse other IT hardware, it is not as critical because of the lower power draws and heat generation. In our design we have room for 20-23U of "Other IT" in the back plane of the tank or in created space in the tank that does not take away from the available space for compute servers.

Q: And how does a data centre have to be re-designed in order to optimise this combination of air and liquid cooling?

A: Even though an immersion cooling system does not need to be segregated from traditional air cooled equipment, it does not need special AC, raised floors or special fire suppression. It could include the immersion in the same space as the air cooled but, why add the additional expense?

Q: And is retrofitting liquid cooling possible and/or sensible, or does it really only work with the latest servers and in a newly designed and built data centre?

A: Liquid cooling needs solid state or helium sealed hard drives. If you have servers with spinning platters you have to change out the drives. Also, every server in air has debris inside. Dust, contaminants from the air and microscopic residuals from packing. That debris will all washout and settle to the bottom of a tank and have no long-term effect. However, if you have had a server running in the air for say 3 years and you shut it down, replace the hard drive, remove the thermal paste from the processor and replace it with a small Indium patch (looks like duct tape and very inexpensive) and put it in the dielectric, it may not reboot (simply because it is has been running for 3 years) or it may boot just fine and in 3 months, you lose, say, a raid controller, you natural instinct will be to blame the immersion when the raid controller was going to fail anyway. As long as you understand this issue, retro fit should not be a problem. Or first servers were DELL 1905's that were 3 years old when we started but they ran fine.

Q: Much is made of the waste heat reuse potential of liquid cooling solutions. How important is this as part of the overall liquid cooling package – is it an essential component, or just a 'nice to have' option?

A: We think it is a nice to have option. In hot climates there is very little need for waste heat. In cold climates it is a great benefit that makes a great solution even better.

Q: In summary, can you outline how and where liquid cooling within the data centre makes operational and/or financial sense and how this cost/technology equation will make liquid technology more attractive over time?

A: Current compute trends suggest that server density will continue to increase over the years. Additionally, power costs and efficiency incentives or carbon taxes will continue to force a strong push towards higher efficiencies. In the short term and for existing installations we recommend identifying those pieces of IT equipment that have the highest density. Then migrate those systems to immersion cooling. Moving the most dense equipment to a system that allows you to use almost half of the power and allows you to add more equipment to a previously capped installation. As technology refreshes comes and new denser equipment is required, then this new equipment should be installed into liquid cooling. This process should continue until a large majority or the totality of the equipment is cooled using immersion cooling.

In the long term and for new installations it is our strong belief that building Data Centers using at least 75% of immersion cooling is the optimal way to go. A few years ago, it was necessary to have at least 25kW/tank in order to have a short ROI on an immersion installation. Today this number is closer to 18kW/tank. In a 50U tank, this is a density of 360W per U, a very achievable threshold for most servers.

Rittal - profile

Rittal's history of cooling solutions

BY CLIVE PARTRIDGE, RITTAL'S TECHNICAL MANAGER FOR IT INFRASTRUCTURE

RITTAL had been manufacturing 19" racks for the cabling industry for many years before a fortuitous alliance with a computer manufacturer in the early 1990s catapulted the company into the server rack age.

The computer manufacturer in question was Compaq.

At the time, data centres were rapidly expanding and Compaq had had the foresight to recognise the benefits that could be realised through vertical racking of server and storage equipment, and how this arrangement could save valuable space. This strategy proved hugely successful; so much so that in 2002, Compaq was acquired by HP.

During those early years most data centres were using perimeter cooling systems and most server racks were cooled by air which circulated under the floor and then emerged from perforated floor tiles immediately in front of the racks. The racks had glazed doors with perimeter apertures and perforated rear doors, all of which was perfectly adequate for the relatively low loads of the racks (ca. 2-3kW) of the time.

Fast forward a few years and it became rapidly apparent that system miniaturisation and higher processing capacities was leading to increasingly higher rack densities. This meant that the electronic components within the racks were getting hotter and legacy perimeter cooling systems were beginning to struggle. It led to the adoption of fully vented front and rear doors, and stricter room layouts, using cold and hot aisles.

Rittal was at the forefront of these developments by virtue of its ongoing relationships with vendors such as Compaq – later HP. These have helped Rittal develop its current, substantial portfolio of IT cooling products that now range from water-cooled products specifically aimed at large enterprises wanting High Performance Computing (HPC) within limited spaces (using closely coupled air flow into conventionally aspirated IT 19" equipment), right through to smaller DX system associated with local edge computing.

If needed, external chiller plant can be supplied by Rittal to water-cooled solutions.

Rittal's LCP (Liquid Cooling Package) cooling units are enclosed in a compact format, typically only 300mm wide, and are physically attached in a 'side-car' fashion to the racks ensuring an even flow of cooled air over the entire height of the racks.

The LCP-CW range is available in two major cooling capacities, 30kW or 55kW, and as one of two systems: InRack (or closed loop) and InRow, usually deployed in conjunction with aisle containment systems. The main advantage of these products is the low space requirement for the high density, high power computing we are all used to today.

The recent, rapid adoption of edge computing in traditional IT and industrial environments has led to the introduction of similar products based on DX. As a result, Rittal's portfolio has been increased with the introduction of new variants for 12kW and 20Kw, both of which are 300 mm wide, plus two new versions for 35kW, both with 600mm-wide housings. One of these variants is hybrid, with DX and a closed-water loop to take much greater advantage of free cooling, making this particular innovation highly desirable.

Finally, a mention of Rittal's latest partnership with ZutaCore, a company that creates cooling solutions that get right to the 'core' of the subject, namely 'on-chip cooling' to remove the heat at the source. In a strategic alliance, the two companies have developed the first "Rittal HPC Cooled-by-ZutaCore" portfolio.

This has expanded Rittal's comprehensive range of IT solutions, using innovative, direct-chip cooling for high-power densities and hot-spot situations. As a result, customers will be able to obtain highly efficient, scalable, cooling solutions for demanding cooling requirements - from edge, hyperscale and colocation through to high-performance computing (HPC) – and all from a single source.

IT Cooling:

Look beyond the norm for true innovation...

For an outsider, the term "IT" conjures up images of cavernous futuristic facilities, myriads of twinkling LEDs, and technologies at the "bleeding edge" of the latest developments, being implemented at a blistering pace in response to the demands of our "connected life".

BY KARL LYCETT, RITTAL UK'S PRODUCT MANAGER FOR CLIMATE CONTROL.

Don't get me wrong, there are facilities like this; but if you look closer to home, most installations are operating on technology that has been in place for years, and in my opinion this is not always effective.

The Legacy

From a climate control perspective, we're talking CRAC units...

Computer Room Air Conditioning Units are the mainstay of the 'White Space', supplying the cavity below the raised floor with cold air, which is then directed up to the equipment via perforated floor tiles in front of the racks.

For today's IT professional, there are some possible downsides:

Upheaval of the business

First, you must have the right location, and employ a professional installer to fit the raised floor. Many small and medium businesses don't have the funds, space or management backing to make large office moves or extensions a reality.

A CRAC unit requires planning and pre-work to be a success. Businesses are becoming increasingly dependent on agile solutions to meet their ever-changing needs, and CRACs don't fit the bill.

Temperature fluctuations

As the air is delivered from below the floor, it rises and increases in temperature. Equipment placed in the highest Us of a rack receives air which can be degrees warmer than its lower counterparts. The discrepancy

between required and actual air temperatures can be detrimental to the performance, lifespan and reliability of the equipment.
In some cases, the distance of the CRAC to the equipment results in a delay in response to any increase in heat load. This delay, however slight, means that equipment will be exposed to higher temperatures. This can have similar effects as above, degrading the overall life or performance of the equipment, and reducing installation efficiency.

Lack of scalability

Clearly, most businesses can't know for certain how fast their IT installations need to grow to keep pace with change. A large project, or influx of new staff, may mean new racks have to be installed, thus increasing the heat load. If a business hasn't also budgeted for new cooling equipment, then problems may quickly arise.

A CRAC unit has a maximum cooling capacity. If that is exceeded, there are two options:
- Remove the existing unit and replace with a larger and more expensive alternative.
- Supplement the existing unit by buying another smaller CRAC to support it.

Both options have downsides. Replacing a unit that has, for example, only been in place for a short time is likely to be an unpopular – if not financially unviable - option.

Supplementing the unit with something smaller is possible, but as mentioned, CRACs take up significant space.

CRACs may be an option for some businesses, but for those wanting to take their install to the next level, there are other (better) alternatives...

The present

In-row & in-rack cooling

These two solutions have a lot in common; what sets them apart is how they deliver the air to equipment. In-Row requires your racks to have perforated doors and should be used in conjunction with Aisle Containment to prevent hot and cold mixing. Cold air is pushed directly out of the front of the cooler into the cold aisle and then passed across the IT equipment. In-Rack, by contrast, requires the use of glazed doors and doesn't require Aisle Containment. Instead it creates a "mini cold aisle" by supplying the cooled air sideways, directly in front of the equipment.

Using in-row to cool the aisle means that if one cooler was to break down, then the other, within what is a shared aisle, could increase output to support. Meanwhile, for businesses with only limited space, or where the equipment is placed within the office, a glazed in-rack option is all-enclosed, quiet solution.

> If you have always done it that way, it is probably wrong.
>
> Charles Kettering

There is also no requirement for a raised floor, so general office space could be converted over just a few days.

If the heat load increases, they can have their output increased rapidly by inserting more fan units into the front of the cooler. Added to which, with cooling capacities of 53kW in a unit that occupies just 0.36m^2, it's highly scalable.

Rear door coolers

Simply put, this means replacing the existing rear door of the rack and installing a replacement that houses a large Air to Water Heat Exchanger.

Ambient air in the room is drawn through the rack equipment using internal fans. Hot air is exhausted and passes over the rear door cooler. Chilled water running through the rear door removes heat from the air and it passes back into the room for reuse at the set temperature.

This solution is particularly suited to lower cooling requirements (<20kW per rack) and if you already have a chilled water supply on site that they can be connected to, this can be a serious contender.

The future

If the business needs and set-up suit the following options, it may be worth investigating more "cutting edge" solutions

On-Chip cooling

"On-Chip" directs cooling inside servers to small heat sinks which are attached to chips and other important hardware within.

This method has been around for some time, using air as the medium for removing heat; however, as processing power continues to climb, and miniaturisation is still the focus, the amount of heat generated is increasing. Liquid On-Chip increases the efficacy of this method and continues to push the limits of cooling within increasingly smaller footprints.

Full Immersion Cooling

Full Immersion Cooling takes things to the next level. Instead of passing liquid by the heatsink to wick away excess heat, the whole server is placed into a reservoir of 'thermally conductive dielectric coolant' which comes into to contact with the equipment from all angles allowing maximum removal of waste heat.

It's a method which is becoming increasingly popular with data centres focusing on "green" credentials and mindful of how they are utilising the waste heat.

Combining this with other methods of cooling is a big positive as Full Immersion can operate at higher water temperatures than other techniques. Even after all the cooling has been performed, users can use the gathered waste heat to warm offices, raising the efficiency levels up further!

Evolving spaces

The rise of Automation, Industry 4.0, IIoT etc. means that, whether a facility is focused on producing tins of baked beans or jet aircraft, the amount of data that can now be generated is exponential. My own experience in factories is that businesses are moving away from the traditional "White Space" and placing their IT equipment closer to its source, which means the demands placed on the cooling systems are also changing.

Robust design, improved IP ratings, and increased functionality into building management systems are part of the new reality. However, nothing will function if your computing equipment doesn't run smoothly. That expensive new robot arm will become, instead, a very expensive sculpture...

Make the break

The reality is that, despite the newer (potentially better) options, CRAC units will be around for some time. Legacy equipment will continue to be replaced and businesses will simply perform like-for-like swaps on their old raised floor installs.

However, I implore you to look beyond the familiar and embrace the new; you won't regret it!

CELEBRATING 11 YEARS OF SUCCESS

Announcing the 11th edition of the premier IT awards: The Storage, Digitalisation + Cloud Awards 2020.

In what has been, and continues to be, extraordinary times for the business world, it seems doubly important to recognise the projects, innovations and individuals which have made such a huge difference during 2020. Almost overnight, employees switched from office working to working from home, and the new, or next, normal, means that, into the future, what might be called a 'hybrid work' model looks set to evolve, with flexible working very much the order of the day. What was already becoming a trend as part of many organisations' digital transformation programmes, has been accelerated.

The SDC Awards 2020 will celebrate the achievements of end users and the IT community as they have innovated like never before to ensure business continuity in these challenging times. This year more than any other, please do make sure that you enter our SDC Awards. There's no limit to the number of entries, all of which are free of charge, and we'll be promoting all the short-listed entries via Digitalisation World's multi-media platform over the coming months, ahead of the awards ceremony. We really do want to celebrate and recognise the many amazing achievements which have come about in response to the coronavirus.

WHY ENTER?

MAXIMISE VISIBILITY
Free PR opportunities with 5 months of marketing utilising the Digitalisation World portfolio.

POSITIONING
Position your product or service as an innovator in your field.

INCREASED CREDIBILITY
An award win, shortlisting or nomination acts as a 3rd party endorsement.

NETWORKING
Over 300 industry leaders attend the gala awards evening.

COMPETITIVE ADVANTAGE
Gain an advantage on your competitors.

NOMINATION IS FREE OF CHARGE AND VOTING IS DONE BY THE READERSHIP OF THE DIGITALISATION WORLD STABLE OF PUBLICATIONS.

SDC AWARDS 2020

www.sdcawards.com

Schneider Eectric Q&A

Q & A with Robert Bunger, Liquid Cooling Program Director, CTO Office, Schneider Electric Secure Power Division

Q: Is liquid cooling the most efficient technology solution to the problem of increasing data centre heat levels and loads, as high-density computing for an increasing number of applications becomes more prevalent?

A: As both chip and rack densities increase, liquid cooling becomes one of the most efficient and effective ways of supplying cooling to the compute. Today there are new processor chips being developed, and some already available like Intel's latest 400W Xeon chips, that can only be deployed in liquid-cooled systems.

Away from high-powered chips, rack densities are increasing to the point where, typically above 20kW/rack, air cooling, although still feasible, becomes less effective, less efficient, complicated and inevitably more expensive.

That's not to say, however, that organisations are unable to use air cooling effectively. There are still many applications where traditional air-cooled infrastructure makes sense, for example, in legacy colocation and on-premise enterprise facilities. However, when you're looking to cool racks above 20kW effectively, especially when looking to get as high as 60kW, liquid cooling becomes the superior choice.

Q: Liquid cooling has been around as a technology since the 1960s. Why is it still struggling to become a more accepted data centre cooling option?

A: Long ago, chips were becoming increasingly efficient with low voltages, low power consumption while delivering more compute performance, so there was no need for organisations to consider liquid cooling. Air-cooling was easy to deploy, familiar and widely understood, so people were reluctant to consider liquid cooling as a solution. However, today, in order to get more computing performance, the power level has had to come up significantly and we have reached a tipping point in the last few years where traditional air cooling is no longer suitable for some applications – especially with increased power requirements for GPUs in high performance computing (HPC).

Liquid cooling was also once considered a solution only for niche HPC applications, but now the tide is turning. Manufacturers are partnering together with systems integrators to create solutions for users in hyperscale, colocation and edge computing environments, and today Schneider Electric are working with Iceotope and Avnet to deliver commercially available chassis-immersion liquid cooling solutions.

This shift towards liquid cooling has been happening for a while and nowadays HPC is not just the preserve of National Laboratories or secret military applications. Universities have implemented compute intensive workloads and emerging applications like Bitcoin mining will need chip and rack densities that require liquid cooling.

The hyperscale community has also been testing and implementing their own direct-to-chip cooling systems for many years, and with the adoption of 5G I believe another big driver will be the edge - especially in applications where the advantages of efficiency, high-density and inherent protection from harsh environments will be beneficial. Overall, I think it's fair to say that we are on the cusp of liquid cooling becoming mainstream.

Q: Aside from the heat increases and chip density increases what other factors do you think will contribute to greater uptake of liquid cooling technology?

A: There are a number of reasons to consider liquid cooling outside of 'niche' applications, and working together with our Data Centre Science Centre, Schneider Electric has produced a series of white papers on the technology.

White paper 279 discusses five reasons to deploy liquid cooling and in summary these are;

- Rising chip and rack densities.
- The pressure to reduce energy consumption.

The desirability of reducing physical data centre real estate, especially in urbanised areas. The need to reduce water consumption, because many highly efficient air-cooled data centres use evaporative cooling and in some areas, water can be an expensive or precious commodity. The inherent advantage of using immersed cooling in harsh environments, where cooling systems are protected against dust and humidity.

Major benefits from an environmental and sustainability perspective also include reduced water and energy consumption, lower noise pollution and greater efficiency, which means lower emissions. So there are a number of factors to consider.

Q: How does liquid cooling work alongside traditional air-cooling solutions within a data centre? It seems to be primarily targeted at servers, so the other IT hardware--storage and networking kit – still relies on traditional cooling solutions?

A: You can absolutely cool storage systems with liquid cooling. In fact, one study at an Open Compute conference showed that traditional hard drives, with spinning disks as opposed to solid-state drives, had 50% fewer failures when cooled in immersed liquid systems because of the very stable temperature environment they produced! For networking systems, cold-plate technology can be used for high-end systems, but we expect that there will be a mixture of air and liquid cooling within data centres for quite some time yet. I don't think anybody really knows exactly how or when a stable balance will be achieved between the two, but this is likely more achievable in new build data centres or optimized, OCP-Ready™ facilites like Kao Data in the UK.

Q: How does the date centre have to be redesigned in order to optimise this combination of air and liquid cooling?

A: If you're starting from scratch and want to take full advantage of liquid cooling, the data centre should have a warm-water loop from the outset for liquid cooling. This allows end-users to get rid of their mechanical cooling systems, such as chillers and computer room air-conditioners (CRAC), and reap the advantages of cheaper infrastructure.

There is also a challenge for people planning data centres to try and calculate the trade-offs that will produce optimum price performance by the time the centres become operational.

Q: Is retrofitting liquid cooling possible and or sensible and does it really work with the latest servers in a new, and in a newly designed and built data centre?

For many organisations that are using air-cooled IT equipment right now, converting all that to liquid

any developments along the lines of heat re-use may also be driven by local or national government initiatives and motivated by environmental concerns, especially in Europe with a number of organizations making greater commitments towards sustainable business practices

cooling will not make sense. We see this in older legacy data centres where the cost and complications of retrofit far outweigh the benefits. However, that doesn't mean that liquid cooling isn't a more energy efficient or cost effective option.

In another one of our research papers, #282, "Capital Cost Analysis of Immersive Liquid-Cooled vs. Air-Cooled Large Data Centres", we provide a detailed comparison of the capital costs of a 2MW data centre using traditional air-cooling, versus a second utilising racks where each server chassis is partially filled with a dielectric fluid. Overall, the cost analysis showed that the capital expenditure (CapEx) of each approach is comparable at similar power densities of 10kW/rack. However, applications utilising greater compaction and higher power densities of up to 40kW per rack, possible only with liquid cooling, resulted in far greater cost savings. At 20kW/rack the space savings result in an overall reduction of 10%, whereas at 40kW/rack the CapEx savings are 14%.

For now, we believe that liquid cooling will likely apply to new servers coming into the data centre and we foresee brand new racks being dropped on to the data centre floor as closed-loop systems or within scalable pod architectures. As more racks come on stream, there will be a staged progression with people first segmenting a small section of the data centre for liquid cooling and finally renovating a larger section where they go directly to dry cooling. It will largely depend on how much liquid cooling needs to get deployed within an existing data centre.

Q: Much is made of the waste heat reuse potential of liquid cooling solutions. How important is this as part of the overall liquid cooling package? Is it an essential component or just a 'nice to have' option?

A: Currently I put heat re-use in a "nice to have" option. For liquid cooling to make financial sense, you don't have to take into account extra benefits like heat re-use. However, there are a number of organizations using this as a business case and deploying the technology as early adopters because of the benefits to the surrounding community. You can see this as a primary factor in our work with EcoDataCenter in the Nordics. However, any developments along the lines of heat re-use may also be driven by local or national government initiatives and motivated by environmental concerns, especially in Europe with a number of organizations making greater commitments towards sustainable business practices.

Q: In summary, can you outline how and where liquid cooling within the data centre makes operational and or financial sense, and how this cost/technology equation will make liquid technology more attractive over time?

A: Some applications will use liquid cooling because it's the only feasible option. However, we are quickly moving towards the point where from the point of view of both capital expenditure (CapEx) and total cost of ownership (TCO), liquid cooling will offer clear and obvious advantages. A purpose-built data centre using liquid cooling, for example, will produce savings within the facility, in terms of reduced chillers, less switchgear or infrastructure and space savings - as well as greater cost savings and sustainability benefits from lower energy usage.

Inevitably, as vendors become more aligned with liquid cooling they will produce more standardised liquid cooled servers and infrastructure to drive costs down further. We expect to see this evolution happen over the next several years and at least within the next five years, you will be able to go and buy liquid cooled IT equipment, and the supporting infrastructure as standardised products.

INNOVISION2020
DATA CENTRE INSIGHTS + PERSPECTIVES

InnoVision: A very special issue of DCS Magazine dedicated to the data centre industry's visionary leaders and technology innovators

To herald the launch of the all-new Data Centre Solutions digital publication, we have produced a very special first issue, entitled InnoVision – providing an overview of the state of the data centre industry right now.

80+ Vendors from across the supply chain have provided their viewpoint on the future and innovation.

How will the data centre industry evolve over the coming months and years, what will be the major drivers and opportunities?

Read today

https://digitalisationworld.com/magazines

IN ASSOCIATION WITH

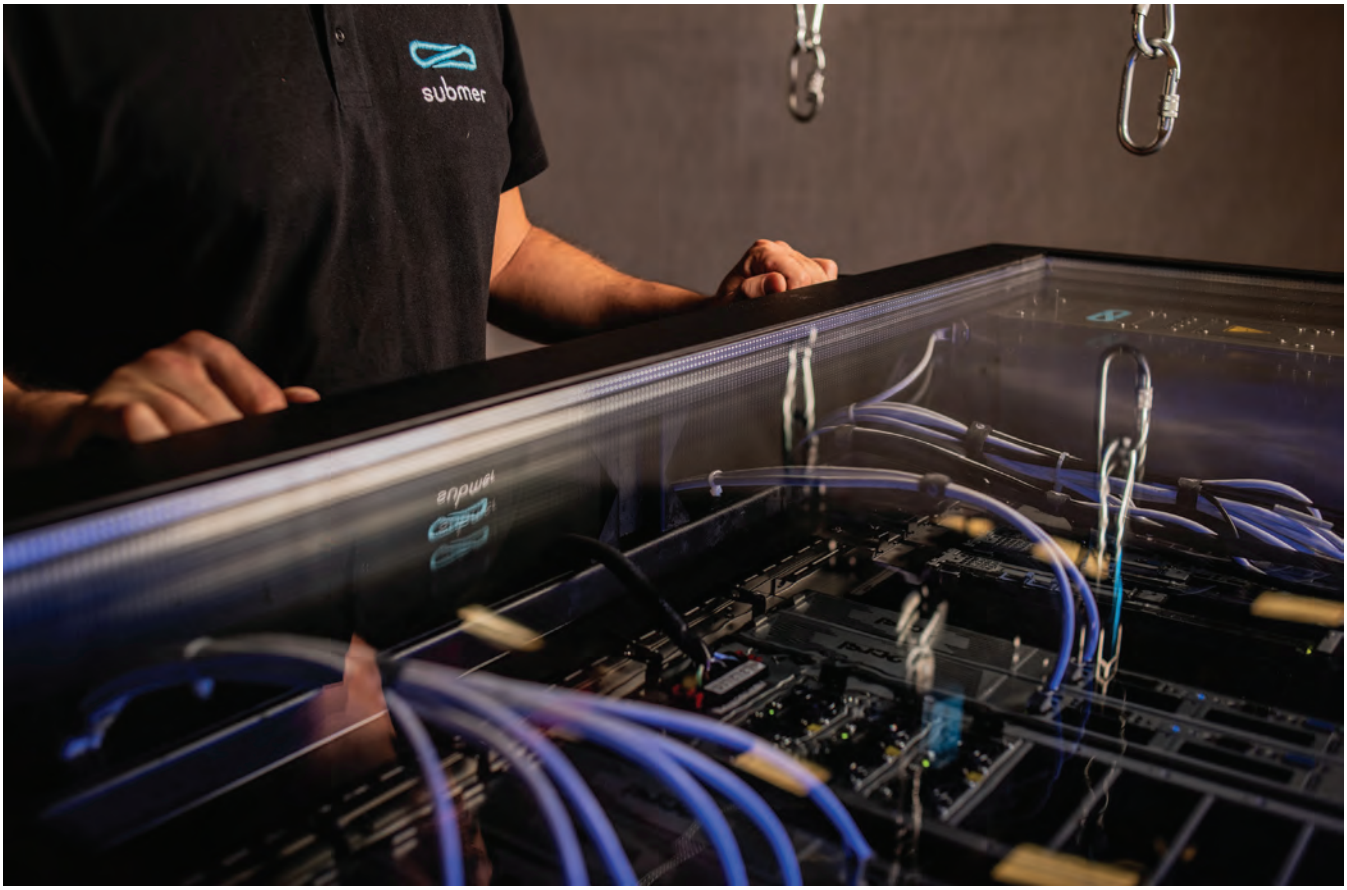
DATACENTRE SOLUTIONS

DIGITALISATION WORLD

Submer profile:
Smart and sustainable solutions for data centres that make sense

THE “Data Centres and The Environment” annual report by Supermicro shows a discouraging reality. According to a survey of more than 5,000 IT experts, 88% of data centres are not eco-friendly, and the environmental impact is not considered on the same level of other aspects like data centres costs, infrastructure expansion, performance efficiency, etc. The fast-paced growth of the new digital trends is responsible for the rapid expansion of data centres and supercomputers. These new trends that are transforming the social, urban and economic landscape, require to process large quantities of data. And this translates into a necessity for greater computational capacity and, consequently, more efficient ways to cool those IT loads with an obvious greater consumption of electricity.

In the last few years, the data centre industry has witnessed a major change in the global compute capability, with an increasing shift of workloads from on-premise infrastructure to the cloud. A new study revealed that data centres computing output jumped six-fold from 2010 to 2018, with a general energy consumption rise of about 6%. In 2018, data centres accounted for about 205 terawatt-hours of electricity usage, i.e. roughly 1% of all electricity consumption



worldwide, according to a report from Science. And by 2030, in some countries, the data centres electricity consumption may grow to between 15–30%.

Water is another essential resource for any data centres operations as important as electricity. Data centres can consume up to millions of gallons of water per day, with clear and understandable concerns regarding the impact of the industry on the environment.

Finally, data centres emit roughly as much CO2 as the airline industry (the global data centre share of carbon emission is estimated to rise from 307 million tons in 2007 to 358 million tons in 2020). In the US alone, data centres are set to consume 139 terawatt-hours in 2020 (139 billion kilowatt-hours) – that's about 3.5% of all the electricity produced there.
In a market that is expected to grow globally at a CAGR of 6.9% from 2019 to 2025, according to experts' forecast, there is simply no more room for innovation without sustainability.

To make the next digital revolution (in fact, already undergoing), sustainable, we as an industry must find smarter solutions and push the envelope of green innovation. That is what Submer is doing.

Submer has developed a series of technologies for a sustainable future. With Submer's solutions, data centres can start already now preparing themselves for the next generation of smart, eco-friendly, "post-human" data centres. Submer's Liquid Immersion Cooling solutions help HPC, hyperscaler, data centres, Edge, AI, deep learning and blockchain applications, increase efficiency by significantly lowering cooling (in a conventional data centre, about 40% of the electricity is used for cooling), and space costs while achieving unrivalled compute densities (>100 kW) and consequent computational capacity.

Submer's SmartPodX Immersion Cooling technology, for example, thanks to its modular and compact design and the physical-chemical properties of the SmartCoolant, allows to improve efficiency and have a smaller (or even positive) impact on the environment than a traditional air-cooled data centre.
Thanks to our SmartPodX technology - the first commercially available Immersion Cooling system that conforms to both standard server formats and Open Compute Project specifications - data centres can achieve:

- Up to 95% saving on cooling costs (corresponding to about 50% of the electricity consumption).
- Up to 85% saving on physical space (with the added benefit of a completely silent machine hall due to the absence of fans).
- Up to 50% saving on OPEX.
- 25%-40% saving on CAPEX (thanks to the higher IT ardware density and the easy retrofitting process of the IT hardware equipment).
- -60% in IT hardware failure rate (the SmartCoolant, apart from whisking heat away from the cores, protects the servers and their components from dust, particles, abrupt changes of temperature and moisture).
- +30% in IT hardware lifespan.
- An mPUE of 1.02 (and, consequently, a better DCiE) that translates into more efficiency and better performance, consuming less energy than air ooling.
- The use of a water-closed loop system free from any evaporation (water to interface with the heat exchanger of the SmartPodX is only required during the installation phase, to fill the closed secondary cooling loop). This allows your facility to avoid any water waste.
- A WUE (Water Usage Effectiveness) of 180.
- Non-hazardous biodegradable proprietary coolant.
- The ability to re-use the waste heat to heat the data centre's host building or surrounding urban and industrial areas.

Submer was founded in 2015 by Daniel Pope and Pol Valls, to make operating and constructing data centres more sustainable and efficient. Submer helps enable data centres that make sense through products, platforms, APIs, processes and installations that will move hyperscalers, colocation and huge industries to new levels of efficiency and innovation.

Q & A with Submer

Q: Is liquid cooling the most efficient technology solution to the problem of increasing data centre heat levels and loads, as high-density computing, for an increasing number of applications, becomes more prevalent?

A: Liquid Cooling definitely represents a much more efficient solution than air-cooling. If we just limit to the physical properties of air and water, the latter has a 4.23 times more specific heat capacity than the former and a better thermal conductivity. Now, at this point, it is necessary to specify that, there are different types of Liquid Cooling solutions, each of which has its own precise characteristics. Also, when we talk about Liquid Cooling, not always the liquid used is water. There are a number of coolant fluids specifically designed for two-phase and single-phase Immersion Cooling, that increase efficiency, security (electronics can safely coexist with liquid) and improve the dissipation of the heat produced.

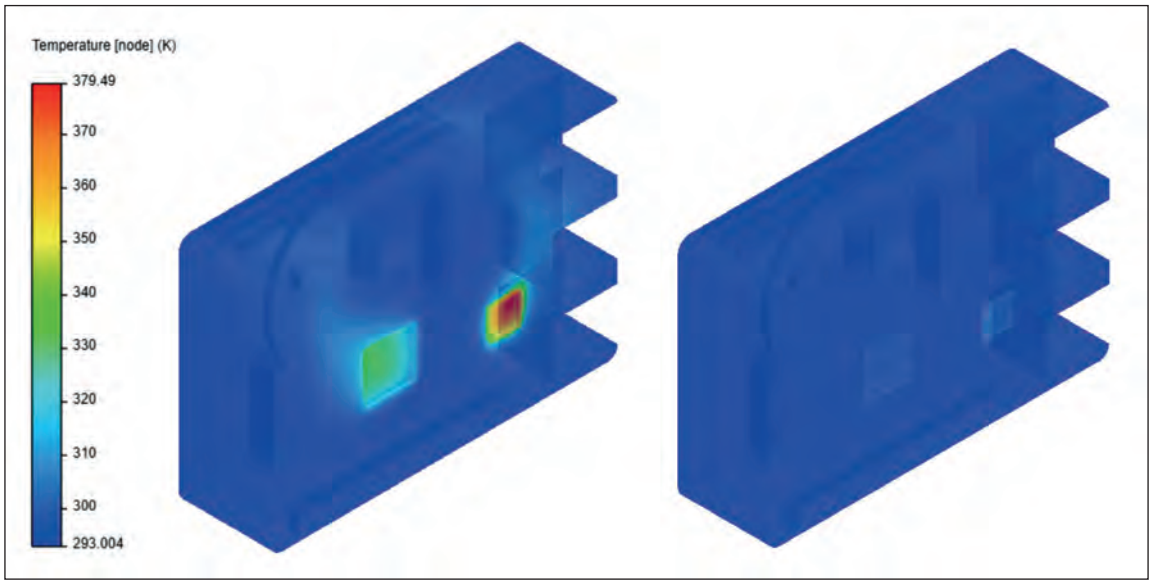
Submer's Immersion Cooling solutions, for example, thanks to their design and the physical and chemical properties of our proprietary, dielectric, synthetic SmartCoolant, allow to reach >100 kW of dissipation (in a SmartPodXL+ configuration). This clearly translates into huge savings from an economic and spatial perspective.

Given that, according to recent studies, in 2020 data centres accounted for about 1% of all electricity consumption worldwide, and that the rack average power density is skyrocketing, it appears clear the necessity of the industry to find new, smarter cooling solutions, because, simply put, we have reached a point where air cooling cannot cope anymore with the huge quantity of data and high computational density. Liquid Immersion Cooling represents a change of paradigm, and at Submer, we understand the importance to create awareness and educate around the benefits of Immersion Cooling for customer's business and the planet.

Q: Liquid cooling has been around as a technology since the 1960s. Why is it still struggling to become a more accepted data centre cooling option?

A: Yes, as a technology, Liquid Cooling is not new, as we all know (IBM developed the very first direct liquid cooling system in the 1960s).

One of the most common critiques that you hear from sceptical about Liquid Cooling is: "If in 60 years, the industry hasn't adopted this technology, it means that it is not worthwhile". Today, air cooling is still the most diffused cooling solution in datacenters (about



Thermal radiation comparison of the same heat-sink in air (on the left) and in Submer's SmartCoolant (on the left).

99% of the market): until now, air cooling worked perfectly fine, and it still does. But what are the effects in terms of electricity consumption, and, consequently of the impact of the budget of companies and on the environment?

Today, cooling systems are definitely more complex and high-tech than 4-5 years ago, as cooling has had to catch up with the demands for denser computing and in part, the beginnings of Moore's law breaking down.

As we were saying in the previous question, we are rapidly approaching a point where air cooling will not be able to cope with the high density required by the new digital trends (AI, Big Data, Machine Learning, etc). Datacenter industry is known to be very slow to implement changes. So, the real question would be: what are you going to do, when air cooling is not a viable option anymore? Do we have a plan B ready to kick in? We need to start developing smarter solutions already now, because unfortunately, even if there is much of a debate around the efficiency of data centres and their environmental impact, the truth is that data centres that are being built today, are still designed and planned according to an old model.

Jaime Pita, Submer's Thermal Engineer, says (in our next-to-be-published white paper about convection, the principle behind Submer's technology): "For electronics, historically fans and heatsinks were good enough, but in the present times, more computation density is needed. The chips are required to operate at their maximum, and one of the known problems is that temperature affects their performance. Those chips can withstand high temperatures, but would it be great to have them operating in a very comfortable environment? [...] The problem is how to improve the energy extraction from a starting point that is already pushed to its limits. [...] The SmartCoolant is the liquid that Submer has developed in order to solve the problem."

Liquid Cooling is still seen as a niche technology, and not really practical, and probably in the past, it was just like that. But in the last 5 years, the market has seen more and more players of the datacenter industry starting to consider Liquid Cooling as a possible solution. Think about for example hyperscalers like Facebook, Microsoft or Google. If the "big guns" are studying how to integrate Liquid Cooling solutions in their facilities, it means something. We are talking about companies that are constantly at the forefront of innovation. At Submer, we have designed solutions that overcome the problems and pain points of Liquid Cooling. Our Immersion Liquid Cooling technology is the result of years of experience in the data centre field.

As Daniel Pope, Submer's CEO, explained during a roundtable at last virtual DCD Madrid 2020, our duty is to keep on educating the audience about the practical, economical and operational benefits of our Immersion Cooling solutions.

Q: Aside from the heat increases outlined in 1. above, what other factors will contribute to a greater uptake of liquid cooling technology over time?

A: Liquid Cooling has all the numbers to become the mainstream cooling solution in the future. Again, a great job in terms of education is absolutely necessary to prepare the ground for its acceptance. Also, we should always have in mind that what we are doing now will have a disrupting effect on the next-generation data centre industry (design, planning, maintenance, personnel training, etc.). So, this is not a process with a limited temporal scope.
We are inevitably moving towards a new business model where sustainability is not a fancy "nice-to-have" option anymore (and this is evident also in other industries). It is a necessity for the planet and it has a clearly profound impact on the social image of companies.

Submer addresses data centres' problems offering a portfolio of customised solutions that are clearly aimed at improving the business of our customers, and that, at the same time, are also respectful of the environment (savings in energy and water, use of a non-hazardous biodegradable proprietary coolant, the ability to re-use the waste heat to heat the datacenter's host building or surrounding urban and industrial areas). In this way, our customers are given the opportunity to reach unprecedented levels of efficiency in a sustainable way.

There are also other aspects that are relevant to determine the uptake of Immersion Cooling. With Submer's technology, any data centre can save:

- Up to 95% of cooling costs
- Up to 85% of physical space (with the added benefit of a completely silent technology)

- Up to 50% OPEX saving
- 25-40% CAPEX saving (thanks to the higher hardware density and easy retrofitting)

Also, the SmartCoolant protects the servers and their components from dust, particles, abrupt changes of temperature and moisture, guaranteeing a 60% lower hardware failure rate and a 30% longer hardware life-span.

Last but not least, our solutions have been designed to be eco-friendly and have a minimum impact on the environment. Apart from guaranteeing an mPUE of 1.02, our technology allows for zero waste of water.

Q: **How does liquid cooling work alongside 'traditional' cooling solutions within the data centre? In other words, it seems to be primarily targeted at servers, so the other IT hardware – storage and networking kit, for example – still relies on traditional cooling Solutions?**

A: Our Immersion Cooling solutions can easily be integrated into an already existing facility or in a completely new one. The modular, easily scalable design allows a fast deployment, also in less traditional datacenter structures, such as an office building(with the added benefit of a completely silent machine hall due to the absence of fans). If our customers would like to install our solutions in an already existing facility, the retrofitting process required is really minimum (our solutions, for example, do not need a raised floor, and can be integrated with any secondary cooling system, such as an adiabatic cooling tower or a dry cooler). Also, being a "datacenter-in-a-box", each of our Immersion Cooling unit can be easily deployed side by side to traditionally air-cooled racks.

In terms of which components Immersion Cooling targets, it is obvious that servers are the most "logic" elements to be submerged. This does not mean that other IT hardware cannot be submerged in our solutions, but it would make sense to deploy a LIC solution to cool those components that clearly produce much heat, hence servers, miners, etc.

Q: **And how does a data centre have to be re-designed in order to optimise this combination of air and liquid cooling?**

A: Again, the retrofitting process required to deploy our technology is minimum. Submer's solutions just need to be connected to the power system, to the Internet connection and to the secondary cooling system (it might be an adiabatic cooling tower or a dry cooler). Our Immersion Cooling systems can be connected to existing chilled water loops, but the biggest benefits will come by implementing higher temperature water loops over time. Also, no raised floor is required, so our SmartPodX can be installed practically anywhere (also in an office building, being a very silent machine thanks to the lack of fans).

So, to sum up, when a customer wants to deploy our technology, we first share a set of basic facility requirements and then we can start right away the deployment process. And, being an all-in-one solution, the installation of our SmartPodX Immersion Cooling units, is perfectly compatible with traditional air-cooled racks.

Q: **And is retrofitting liquid cooling possible and/or sensible, or does it really only work with the latest servers and in a newly designed and built data centre?**

A: As said, there are different types of liquid cooling, and each of them has clearly different characteristics and retrofitting requirements.

The solutions offered by Submer allow you to easily take your air-cooled servers and submerge them into our tanks with a minimum retrofitting process. We have a video on our YouTube Channel where we show how to adapt an air-cooled server for our Immersion Cooling solutions. So, to reply to the question, yes, retrofitting is possible. Submer is also actively working with OEM's and ODM's to develop IT hardware specifically designed for Immersion Cooling. As far as the retrofitting of the facility is concerned, as already explained, our solutions can be easily integrated into an already existing data centre.

Q: Much is made of the waste heat reuse potential of liquid cooling solutions. How important is this as part of the overall liquid cooling package – is it an essential component, or just a 'nice to have' option?

A: We have to change our perception and idea of datacenters since the planning stage. Datacenters need to change their role, switching from big, ugly, energy-guzzling buildings to smaller, more agile and smarter structures organically and positively integrated into their surrounding environment.

Submer's technology has been designed with a clear environmentally friendly approach, limiting the waste of water and offering the possibility of re-using the waste heat to heat the datacenter's host building or surrounding urban and industrial areas

8. In summary, can you outline how and where liquid cooling within the data centre makes operational and/or financial sense and how this cost/technology equation will make liquid technology more attractive over time?

As summarized by Daniel Pope, CEO of Submer, in one of our eBooks:

- From an operations perspective, Immersion Cooling simplifies datacenter commissioning and maintenance, thanks to the use of scalable building blocks on the compute and cooling side. The combination of our SmartPod platform together with dry-cooler or adiabatic cooling infrastructure for the secondary cooling loops, enables infrastructures that have much simpler requirements on the water loop and consume less energy and resources than any other type of cooling infrastructure. It's also important to realize that an immersion solution should be considered infrastructure with an amortization schedule greater than 15 years in comparison to technologies like Direct to the Chip Cooling where most of the costly components are embedded into the servers and thrown away once server life-time is met. This means an unprecedented PUE throughout the lifetime of the SmartPod technology.
- From a practical perspective, deployment is quicker and simpler. Thanks to its density, Immersion Cooling gives the datacenter the possibility of deploying in prime locations at city centres and solving latency barriers without imposing restrictions on high chip densities of today's compute requirements.
- The economic benefits of Immersion Cooling go from numbers like 99% saving of electricity cooling osts, a reduction of 50% of the datacenter TCO (Total Cost of Ownership), zero water consumption nd, for the first time open the possibility to reuse data centre heat in a closed and perfectly efficient-system.

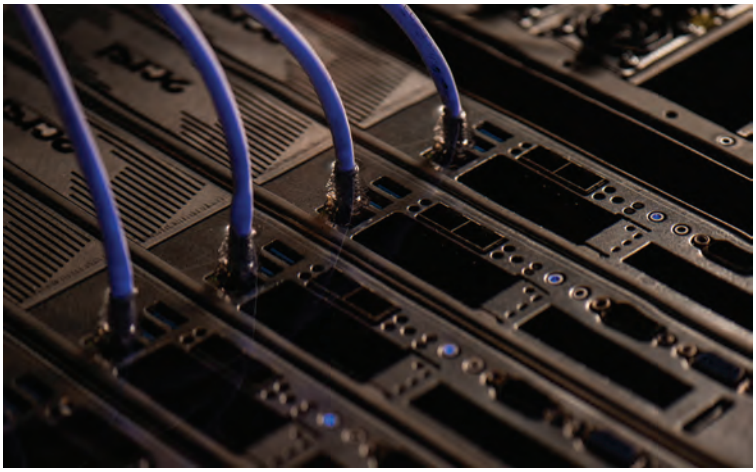
What is Liquid cooling?

A brief history of datacenter cooling.

BY MATTEO MEZZANOTTE, PR, COMMUNICATION & CONTENT MANAGER AT SUBMER.

COOLING is a fundamental aspect in datacenters. Practically, cooling technologies have the task to maintain certain environmental conditions suitable for the IT equipment to work properly avoiding overheating. Today, cooling represents about 40% of the electricity consumption in a datacenter[1]. It is then clearly understandable why the datacenter industry needs to constantly find new and better ways to cool the IT equipment.

What we refer to as datacenters today, started out as computer rooms in the 60's to house the first mainframe computers. While computers and servers evolved quickly over the years, an evolution of the datacenter itself per se and more specifically the cooling was not really necessary as the evolution in computing also brought significant innovations that allowed (apart from a few notable exceptions such as the CRAY 2 in 1985) denser computing and continued cooling with air and fans.
On the contrary of what many might think, liquid cooling is not a new technology, as it has been widely used since the 1940s to cool high-voltage transformers.
?In the 1960s, IBM developed the very first direct liquid cooling system.



In the 1980s, with the advent of metal oxide semiconductors, the concept of a liquid cooling solution became less of a priority as this new method of packing multiple transistors did not generate the same proportionate heat as the size of the transistors and therefore their voltages were drastically reduced.
The Need for a Better Cooling
Liquid Cooling is a much more efficient solution for datacenter cooling (just think that water has 4.23 times more specific heat capacity than air, and the coolants used in liquid cooling have even better specific heat capacity than water).
There are different reasons behind the necessity of datacenters to find better cooling solutions than just air-cooling: practical, economic and environmental.

Average power density has gone from 4-5kW (about 10 years ago) to 7-10kW, according to AFCOM[2]. According to Uptime Institute[3], "back in 2012, the highest density racks the organization could find were consuming 26kW. Last year, ten percent of respondents to its data centre survey reported that they were running some of their racks at above 40kW".[4] Packing more computing power into IT equipment means boosting the power density of a datacenter, but it also means increasing the heat produced by racks and cabinets.

Today, cooling systems are definitely more complex and high-tech than 4-5 years ago, as cooling has had to catch up with the demands for denser computing and in part, the beginnings of Moore's law breaking down.

From 2010 to 2018, the datacenters computing output jumped six-fold, with a general energy consumption rise of about 6%[5]. In 2018, datacenters accounted for about 205 terawatt-hours of electricity usage, i.e. roughly 1% of all electricity consumption worldwide, according to a report from Science[6]. And by 2030, in some countries, the datacenters electricity consumption may grow to between 15–30%[7].

The debate around energy consumption in datacenters too often does not consider another fundamental resource apart from electricity: water. Datacenters can consume up to millions of litres of water per day, with clear and understandable concerns regarding the impact[8] of the industry on the environment ("In 2019 alone, Google requested, or was granted, more than 2.3 billion gallons of water for data centers in three different states, according to public records posted online and legal filings"[9]).

Finally, datacenters emit roughly as much CO_2 as the airline industry (the global datacenter share of carbon emission is estimated to rise from 307 million tons in 2007 to 358 million tons in 2020). In the US alone, datacenters are set to consume 139 terawatt-hours in 2020 (139 billion kilowatt-hours) – that's about 3.5% of all the electricity produced there.

Different Methods of Liquid Cooling for Datacenters
Liquid Cooling has come back into the spotlight and there is evidence (see the testing of Liquid Cooling solutions by Microsoft and Google) that it will represent the most efficient cooling solution in the years to come and become mainstream in the not so distant future. Traditionally, there are three different ways to cool a datacenter: air-based cooling, liquid-based cooling and a hybrid of the two, and within these categories are a couple different methods. Let's focus here on the liquid-based cooling systems. First of all, it is necessary to specify the kind of liquid used in each system. It can be water, a synthetic fluid or a mineral coolant.

Water-cooled racks (otherwise known as rear-door chillers) have water flowing alongside the racks, but never actually touching the servers. This solution works well, but there is still the legitimate concern of water leaking onto the servers and components and compromising the IT hardware integrity, not to mention the fact that the chilling is done with compressors and uses quite some energy to cool. Direct Liquid Cooling mostly refers to water or other liquids that are delivered directly to a hot plate which is only on CPUs and GPUs.

With rear door cooling and direct liquid cooling, it is normally advised to have a secondary form of cooling such as a traditional CRAC unit to handle the excess heat in the room. Liquid Immersion Cooling means that the entire servers (switches, etc.) are completely submerged in a dielectric (i.e. it does not conduct electricity) synthetic or mineral fluid (single phase and two-phase Immersion Cooling use different types of coolants), designed to deal with 100% of the heat and preserving the IT hardware (no particles being shot) and is the most efficient method to save energy if combined with dry coolers (with or without adiabatic functionality) using only ambient temperatures to lower the warm water of the secondary cooling loop. According to Data Center Frontier, it is 1400 times more effective than traditional air-cooling. With air-based systems, fans need to be blasting 24 hours a day even when it isn't needed (the old, unsolved issue of "idle" datacenters).

Finally, Liquid Immersion Cooling can be single phase or two-phase:

- **Single Phase:** In single phase Immersion Cooling, the dielectric fluid (it can be either synthetic or mineral), is circulated within the tank by pumps or by natural convection flow. In the single phase Immersion Cooling, the coolant always maintains is liquid state while operating (i.e. it never boils or freezes). The dielectric coolant is either pumped through an external heat exchanger where it is cooled with any facility coolant, or the facility coolant is pumped through an immersed heat xchanger, which facilitates heat transfer within the electric liquid.[10]
- **Two-phase:** In the two-phase Immersion Cooling system, the heat removal is performed - as the name itself suggests - in two phases: the coolant (fluorocarbons in this case), literally boils when it gets in contact with hot components, due to its low boiling point. The system takes advantage of a concept known as "latent heat" which is the heat (thermal energy) required to change the phase of a fluid, this occurs when the two-phase coolant comes in contact with the heated electronics in the bath that are above the coolants boiling point. Once the two-phase coolant enters its gas phase it must be cooled or condensed, typically through the use of water-cooled coils placed in the top of the tank. Once condensed the two-phase coolant drips back into the primary cooling tank.[11]

1 https://www.sciencedirect.com/science/article/pii/S1876610215009467#:~:text=In%20data%20centers%2C%20about%2040,bill%20of%20large%20data%20centers.
2 https://datacenterfrontier.com/rack-density-trends-higher-but-ai-could-boost-a-high-density-horizon/
3 https://datacenter.com/wp-content/uploads/2018/11/2018-data-center-industry-survey.pdf
4 https://virtusdatacentres.com/item/389-power-density-the-real-benchmark-of-a-data-centre
5 https://www.nytimes.com/2020/02/27/technology/cloud-computing-energy-usage.html
6 https://science.sciencemag.org/content/367/6481/984
7 https://www.nature.com/articles/s41558-020-0837-6?utm_content=137771421&utm_medium=social&utm_source=linkedin&hss_channel=lcp-35620104
8 https://www.bloomberg.com/news/features/2020-04-01/how-much-water-do-google-data-centers-use-billions-of-gallons
9 https://time.com/5814276/google-data-centers-water/
10 https://en.wikipedia.org/wiki/Immersion_cooling
11 https://en.wikipedia.org/wiki/Immersion_cooling

Vertiv Q & A

with Nigel Gore, Global Offerings, High Density and Liquid Cooling at Vertiv.

Is liquid cooling the most efficient technology solution to the problem of increasing data centre heat levels and loads, as high density computing, for an increasing number of applications, becomes more prevalent?

There are a variety of cooling technologies out there – established and emerging – which could be used to manage increasing thermal loads in the data center. The most efficient option often depends on the customer and the characteristics of their data center in question.

However, liquid is well known as an effective and efficient heat capture and transport medium due to its specific heat properties compared to air which has a specific heat value of 1000J/kgK dielectric specific heat ranges from 1300J/kgK to 2300 J/kgK. Pure water has a value of 4179 K/KgK which is of interest to maximise heat transfer. In determining the heat transfer coefficients, the heat flux and the variance between the temperature difference are used to determine thermodynamic flow forces. The heat transfer coefficient is measured in SI units as watts per squared meter kelvin: W/(m2K). Comparison of the specific heat and heat transfer coefficient are fundamental considerations when considering how to best manage high heat loads and the energy stack to transport heat from heat generating electronic components through the facility to heat rejection devices.

Liquid cooling has been around as a technology since the 1960s. Why is it still struggling to become a more accepted data centre cooling option?

It's not that liquid cooling is struggling to become accepted, the issue is that until relatively recently the business case wasn't there outside of specific use cases such as high-performance compute. Technically, and theoretically, liquid cooling should be more efficient but there are other factors at play when it comes to deciding the most effective thermal technology for a given site. That said, liquid cooling has been gaining traction over the last five to ten years. For example, at Vertiv we have extensive liquid cooling deployments across our global customers.

We offer infrastructure solutions across a variety of cooling methods for the data centre and liquid cooling is one of them. Our XDU direct to chip liquid cooling distribution unit is specifically designed to support IT equipment that requires a controlled water circuit. There is an extensive range of rear door heat exchangers that capture the exhaust heat from racks and with the specific heat effectiveness of water can transport the heat for reuse or rejection based on the facility design. To support these rear door units there are numerous options of liquid distribution infrastructure that assist in efficient transport of warm liquid from the racks and return cool liquid through the heat exchange to rejection process.

Aside from the heat increases outlined in 1. above, what other factors will contribute to a greater uptake of liquid cooling technology over time?

We are witnessing an increased desire to manage the environmental impact of deploying digital infrastructure. This can be achieved through reduced energy use and the flow on benefit of reducing water use and water use reduction in the cooling of data centres. Carbon effects of building, operating and decommission are also under the umbrella of environmental considerations either by reduced time to build or ensuring materials used are sustainable and reduce carbon emissions. Vertiv work closely with industry bodies, partners and suppliers to continually apply improvements to best practices.

How does liquid cooling work alongside 'traditional' cooling solutions within the data centre? In other words, it seems to be primarily targeted at servers, so the other IT hardware – storage and networking kit, for example – still relies on traditional cooling solutions?

Firstly, it's important to distinguish between different forms of liquid cooling. So-called cold-plate approaches primarily capture heat from the processors with new designs extending to other components of the server. There remains some form of additional air-cooling in most cases to support ancillary equipment. Immersion liquid cooling captures the vast majority of the server heat and as such reduces the demand on air-based cooling, including climate control and comfort cooling combined with efficient external heat rejection.

However, we have seen interest in mapping out the

components contributing the highest heat outputs and candidates include networking equipment, telecommunications infrastructure, power delivery and data storage. There is a wide net of devices that fall into this category already installed in data centres globally. The life cycle and specification of the data centre will determine suitability of which devices can be considered to deploy liquid cooling. There is not a one size fits all approach to deployment with a hybrid approach seen as the logical next step. Geographical needs will determine priorities based on climatic conditions and economic factors.

And how does a data centre have to be re-designed in order to optimise this combination of air and liquid cooling?

Vertiv works with industry associations such as, The Green Grid, ASHRAE and the Open Compute Project who provide guidance and best practices for considerations of liquid cooling infrastructure. There are fundamental considerations including the heat load devices chosen to be liquid cooled which will determine the infrastructure required to transport liquids through the facility. There are construction materials to be considered included wetted materials that are compliant for use based on the chosen liquid medium. The transport of fluid will require pipework layouts designed to the chosen liquid cooling format. Most data centre designers could be consulted for the optimal methods following industry best practices and guidelines outlined. Including regional legislation requirements relating the liquid properties and mechanical considerations to deploy them.

And is retrofitting liquid cooling possible and/or sensible, or does it really only work with the latest servers and in a newly designed and built data centre?

The retrofit debate related to IT equipment depends on the age and scale of IT deployed. In most cases IT equipment already deployed that are designed for air cooling are not economically viable for retrofit to liquid cooling mainly due to the labour overhead and level of conversion required. Where there is a need to manage heat densities the use of a rear door that uses liquid to remove the exhaust heat is recommended. This can be used for rack densities up to 60kW. Where there is a need to deploy specific liquid cooled IT a first step before retrofitting the data centre an air to liquid CDU is recommended. This will allow for a liquid cooled rack of servers utilizing direct to chip cold plates to be deployed without a dedicated liquid loop taken through the data centre. There is an added benefit of the volume of liquid being controlled and in small volumes. The development of modular data centres is another consideration where a liquid cooled deployment could be added to a facility without significant reconstruction and the benefit of being a contained system. The speed of deployment capable makes this a viable option to be considered.

Much is made of the waste heat reuse potential of liquid cooling solutions. How important is this as part of the overall liquid cooling package – is it an essential component, or just a 'nice to have' option?

There are industry metrics for energy use to compare facility efficiency and recently heat reuse featured as a contribution to improve efficiency. There are a growing number of industry references where, with government support, these initiatives are seeing adoption particularly in Nordic regions. The logical next step is how to effectively utilise the warm liquid and supporting the heating system where seasonal demand requires this can be very efficient. The deployment options will come down to regional requirements and proximity to facilities where heated liquid can be most beneficial. Industry concepts have seen the use of 5G network infrastructure deployed with liquid cooling within housing estates with waste heat repurposed for water heating requiring less energy to top up the heat for central heating within the building.

In summary, can you outline how and where liquid cooling within the data centre makes operational and/or financial sense and how this cost/technology equation will make liquid technology more attractive over time?

The economic view considers Total Cost of Ownership (TCO) to design, build and operate a data centre. These include capital expenditure on constructing the facility, time taken and removing complexity from the installation of equipment to power, cool, secure and operate the facility. The operational side of the equation relates to the IT deployment and related energy costs which are the largest cost contributors. The number of IT refreshes and the economic life of the facility are to be included. Liquid cooling has benefits to energy use by reducing the need for fans within the server and energy to move air around the facility compared to the heat transport effectiveness of liquids. Where there is a high energy cost, the economic payback sees a faster rate of return. Where there is an environmental need, either by climatic conditions or constraints on natural resources the TCO equation will be deemed attractive through the support of legislative and geographic constraints.

Conquering the next challenge for colocation providers: Speed

Ashish Moondra – Senior Product Manager, Power, Electronics & Software at Chatsworth Products (CPI).

AS THE TRANSITION from the Information Age to the Age of Artificial Intelligence gives way to heightened significance of connectivity, cloud service providers and the IT industry work around the clock to ensure the life most of us know today, high-speed internet, mobile connectivity, self-driving cars and machine-to-machine (M2M) learning. A recent Cisco Annual Internet Report confirms this reality.

By 2023, for example, nearly a third of the global population is expected to have Internet access – that is about 5.3 billion users. Meanwhile, the number of IP networks is projected to be more than three times that number.

Within the data centre space, the colocation market may see the most growth, with an estimated CAGR of almost 11% from 2020 to 2025. Faster time to market- in lieu of undertaking an on-premise data centre project that may take months to complete – is the primary reason for the attention toward this segment. Needless to say, delays in bringing up a new customer within a multitenant environment directly translates into lost revenue. Therefore, it is no surprise that colocation providers are challenged to scale up with solutions that are quick to deploy, manage and service.

The following are two key points for colocation vendors to consider when looking to quickly get new customers up and running.

Vendor selection

Within colocation environments, end customer requirements generally vary based on budgets, functionality required and the IT equipment that will be housed within the cabinets. Service-level agreements (SLAs) require colocation facilities to be able to quickly provide the infrastructure equipment that meets the needs of their end customer. Partnering with equipment vendors that have local manufacturing capabilities and a build-to-order model provides colocation vendors with the ability to quickly procure products aligned with end customer requirements. In-region manufacturers typically have a wide breadth of standard solutions and the ability to create and deliver custom solutions in a short timeframe.

While evaluating equipment vendors for their ability to deliver products in short lead times, it is critical that data centre professionals ask questions related to location of the supply chain as well as their risk mitigation plans. With the booming demand for more things to be connected to the Internet, some electronic components as well as populated, printed circuit board assemblies can have lead times spanning several months.

Equipment manufacturers in North America that rely on in-region sources for long lead time components will have a better ability to scale quickly to meet demands of larger projects.

The common denominator within the date centre white space is the equipment cabinet. Dealing with vendors who can preinstall all infrastructure solutions within the cabinet, including power distribution equipment, cable management solutions, access control and environmental monitoring per the end customer's needs will save colocation vendors significant time, effort, and money. Additionally, preconfigured solutions that are tested together before they leave the factory minimizes any surprises that could otherwise delay schedules when multivendor equipment is received separately. Finally, consider that preinstalled solutions require minimal packaging, helping reduce waste and the tie required to deal with it.

Product considerations

To allow remote manageability of the off-premises equipment, colocation vendors provide intelligent hardware solutions that allow monitoring and control of power and environmental parameters within the cabinet. Growing regulatory and security demands also require end customers to control physical access to the cabinet and maintain an audit log of all success attempts.

While these solutions provide significant advantages to the end customer, the challenge is to deploy them speedily over the network and quickly configure them to be fully operational. Intelligent power distribution units (PDUs) that also integrate environmental monitoring and access control provide a unified solution that require just one single network connection. The speed of deployment can be further enhanced by utilizing intelligent power distribution units with Secure Array IP Consolidation that allow up to 48 intelligent PDUs to share one primary IP address and an alternate one for failover capability.

This setup allows the white space infrastructure for complete rows of cabinets to be managed by one or two ports on a network switch. The alternate and inefficient solution would have been to first install, wire and configure extra network switches purely for infrastructure monitoring, connecting them to every monitored device and then taking a crash cart to each device to perform their IP setup.

Once the PDUs are deployed on the network the next step that could take a considerable amount of time is the configuration of every monitored device that includes network access, threshold, and notification settings. In this scenario, choose PDUs with bulk configuration capabilities over the network. However, the preferences of end customers for mass configurations can differ.

For example, while a data centre operations group may prefer bulk configuration through a data centre infrastructure management (DCIM) software solution, network professionals or developers may prefer automated configuration using a Command Line Interface (CLI) or Application Programming Interface (API), This means colocation vendors that deal with multitude of end customers will be ahead of the competition if they provide a solution that supports most types of bulk configuration methods. All these capabilities not only make initial deployment and configuration easier, but also simplifies ongoing management.

Another important and usually overlooked aspect to consider is the serviceability of the products. The most common maintenance to be performed on Intelligent PDUs is timely firmware upgrades. The products chosen should allow for these upgrades to be easily performed over the network or through UB ports on the equipment. A field-replaceable controller on the unit also allows for seamless serviceability and upgradability. These upgrades should be capable of being performed while the units continue to provide basic power distribution to connected equipment. Finally, consider that intelligent products such as PDUs should include warranties with an advanced replacement coverage as a norm rather than exception.

With data consumption growing faster than ever, speed of deployment and delivery is the most pressing challenge for colocation providers. The ones who consider the two recommendations above will be able to have a competitive edge that will ultimately allow them to grow their top line revenue faster and be ahead in the race.

Ashish Moondra

Ashish Moondra has a total of 20 years of experience developing, managing and selling rack power distribution, uninterruptible power supply (UPS), energy storage and Data Centre Infrastructure Management (DCIM) solutions. Ashish has previously worked with American Power Conversion, Emerson Network Power and Active Power, and has been an expert speaker at various data centre forums.

The importance of choosing the right battery for Uninterruptible Power Supplies

Mark Coughlin - Applications Manager for Reserve Power at EnerSys®

TODAY'S data centres depend on uninterruptible power supplies (UPS) to provide clean, continuous power throughout the facility's entire operational life. While the mains supply is available, the UPS protects the data centre's sensitive information and communication technology (ICT) equipment from electrical noise and any other power aberrations that may appear on the incoming power feed. If the supply fails, then the UPS battery must take over instantly and use its stored energy to support the load until either the mains is restored, or a generator can be started, or the data centre systems can be shut down safely.

A UPS battery's ability to store energy reliably and efficiently during normal operation, making it immediately available to the load during a mains failure, is critical to data centre security. Accordingly, in this article, Mark Coughlin, Applications Manager for Reserve Power at EnerSys® reviews the key battery technologies currently available, to inform commercial and technical specifiers responsible for battery selection.

The article starts by looking at how data centre evolution is affecting the demands placed on batteries. It then compares Lead-Acid, the pre-dominant battery chemistry used within data centres, with alternative technologies, in particular Lithium-ion (Li-ion), which has been generated rising interest in recent years. Finally, it examines why data centre operators should consider advanced Thin Plate Pure Lead (TPPL) technology to optimise the performance of their UPS and power systems.

The evolving data centre landscape

Data centres today experience a rising incidence of power outages and grid fluctuations caused by increased urbanisation and demand. Meanwhile, their workload is expanding, with a move to multi-user hosting services and larger data storage capacity requirements. These factors increase pressure for 'best in class' technologies and reliable power.

UPS batteries are also directly impacted by reduced autonomy times, now typically between 30 seconds and 5 minutes, compared with historical averages of around 15 minutes. This is because of the shorter times needed to start up generators and switch loads. Fast recharge times are also desirable, allowing batteries to be recharged quickly in order to be able to support further power outages.

Energy efficiency has become an overarching concern for all data centres, not just because of the financial impact of large-scale operation and rising energy costs, but also due to pressure from stakeholders – and legislation – to pursue effective carbon footprint reduction policies.

Concerns about energy costs and grid power availability are driving growing interest in using UPS battery assets for energy storage applications, as a way to generate further revenue. In Firm Frequency Response applications, for example, UK-based data centres could provide battery energy back to the National Grid on demand.. Alternatively, the batteries could be used for peak shaving, reducing data centre energy costs by supporting loads when electricity cost is high and then recharge the battery when low-cost electricity supply is available.

Such strategies can bring significant cost savings, and generate money when supplying energy back to the grid. However, they demand longer battery autonomies than the five minutes typically needed for UPS backup. Currently, there are relatively few active sites deploying this strategy. Nevertheless, manufacturers such as EnerSys® have conducted trials with batteries that can support these applications.

Battery technologies and trends

Battery chemistries currently available for UPS backup include Lead-Acid, Lithium-ion (Li-ion) and Nickel-Cadmium. There are also non-battery technologies like flywheels and Super-Capacitors. However, this article focuses on the two types that currently dominate the data centre industry: Lead-Acid, which represent over 90 per cent of the UPS market share, and Li-ion, which is attracting increasing interest due to its purported performance benefits and high visibility through its use in electric vehicles.

Li-ion is attracting interest through being attributed with performance features superior to traditional Lead-Acid VRLA batteries, which are typically either gel or absorbent glass mat (AGM) designs. Compared to traditional VRLA equivalents, Li-ion offers a high cycle life, together with a significant size and weight reduction. Li-ion batteries also have high charge efficiency, with excellent partial state of charge tolerance – in fact, partial charge is preferred for long cycle life and operation in float conditions at full state of charge is avoided. The self-discharge rate of Li-ion is also low, which results in prolonged shelf life when in storage. Finally, it has good high and low temperature performance, and no gas emissions.

However, Li-ion's comparison with traditional VRLA reveals some challenges along with its benefits. Accordingly, we show how TPPL technology, as an advanced form of Lead-Acid chemistry, offers a number of advantages over traditional VRLA batteries. Despite historical cost reductions, Li-ion pricing remains a barrier for many users. With pricing depending on many factors including purchase volumes and the exact chemistry used, Li-ion is currently significantly more expensive than Lead-Acid. Furthermore, although space-saving may be important within data centres, weight reduction, which Li-ion batteries offer, is seldom critical. Similarly, the high cycling capability of Li-ion isn't a driving factor for selection within UPS applications, where batteries are mostly floating at near full state of charge.

While considered a safe technology, any Li-ion solution, unlike Lead-Acid, must include a battery management system (BMS) to ensure safe charging and discharging. This increases complexity, and requires users to have a thorough understanding of Li-ion technology. However, the BMS provides built-in diagnostics, which identify most problems and allow minimal maintenance.

Additionally, consideration must be given to the MTBF (Mean Time Before Failure) of the electronic components factored into Li-ion calendar lifetime calculations. Lifetimes of 15 years are claimed, but service life is not proven in the field. By comparison, advanced TPPL, with 12+ years design life, provides eight to 10 years' service life, while traditional VRLA 10-year design life batteries typically provide five to six years' service life.

Charging is another important consideration. Firstly, to fast-charge Li-ion, higher charging capacity, with increased cost, may be required. Also, in many cases the charging architecture would need to be replaced or changed to support different Li-ion battery charger voltages, so two different UPS rectifier types would be required across a data centre attempting to deploy both Li-ion and Lead-Acid batteries.

Other factors, while not immediately specific to the data centre environment, should also be considered when selecting a battery technology. During transportation, Li-ion faces legislative shipping restrictions, while Lead-Acid batteries, including AGM and TPPL, is classified as non-hazardous for all transportation modes. Then, at end-of-life, Lead-Acid has an inherent value, and is about 95 per cent recyclable by a very well-established network of smelters; this possibility, however, is not mature for Li-ion.

TPPL: Optimised performance without the drawbacks

Above, we have seen why Li-ion, while attracting increasing attention, has been slow to penetrate to

data centre market. On-going development driven by the powerful automotive sector may change this, but advanced TPPL technology offers data centre managers the best of both chemistries.

As a Lead-Acid based battery technology, TPPL is reliable, well-proven, and easy to transport, handle and recycle. Crucially, advanced TPPL technology significantly improves energy efficiency, by providing up to 43 per cent energy reduction compared with traditional VRLA batteries through reducing float current requirements. Further energy savings accrue as it can operate, within warranty, at elevated temperatures, reducing air-conditioning requirements.

Meanwhile, advanced TPPL battery technology reduces data centre vulnerability to multiple mains blackouts, through very short recharge times and time to repeat duty. For example, with 0.4C10 A charging current using fast charge methodology, TPPL can be fully recharged, following a one-minute discharge to 1.6 Vpc, in 2½ hours, and ready to repeat duty in 22 minutes.

Battery replacement costs are also reduced through low internal corrosion rates, yielding a service life 25 per cent longer than for traditional VRLA. Additionally, storage life is increased from six to 24 months due to low self-discharge rates.

Advanced TPPL technology is used today in many demanding critical applications. Data centre users can access TPPL through DataSafe® XE batteries, which are specifically designed for UPS applications. They support autonomies of under five minutes, while offering all the above TPPL features.

What of the future?

Lead-Acid technology is expected to dominate the market for at least the next few years, although enquiries and niche projects suitable for Li-ion will continue to grow. In particular, applications requiring high cycling will be seeking advanced TPPL or Li-ion solutions.

Depending on the application, Li-ion could be the preferred battery type. Nevertheless, before opting for Li-ion as the technology for a particular application, a full consideration of the requirements should be undertaken. The assessment should reflect the Total Cost of Ownership, with the benefits and challenges of Li-ion compared against other available technologies, including TPPL.

Irrespective of the technology chosen, battery monitoring systems will become increasingly popular, due to the battery condition visibility and opportunities for predictive maintenance that they provide. This will also bring UPS applications into the increasingly pervasive Internet of Things (IoT) environment, making them visible as components of the larger data centre infrastructure.

For further information on UPS batteries, specialists can be contacted via the website https://www.enersysdatacentres.com

Angel
BUSINESS COMMUNICATIONS

WEBINARS

Specialists with 30 year+ pedigree and in-depth knowledge in these overlapping sectors:

Expertise: Moderators, Markets, 30 Years + Pedigree
Reach: Specialist vertical databases
Branding: Message delivery to high level influencers via various in house established magazines, web sites, events and social media

Semiconductor (Silicon/Compound)
Publications include: Compound Semiconductor, Silicon Semiconductor, CS China, SiS China

Power Electronics
Publications include: Power Electronics World

Future Mobility
Publications include: TaaS Technology, TaaS News

Data Centres
Publications include: DCS Europe, DCS UK, SNS International

SmartSolar UK & Ireland
Publications include: Solar and Power Management, Solar UK and Ireland

Sensors
Publications include: Sensor Solutions Magazine, Sensor Solutions International

Digitalisation
Publications include: Digitalisation World, Information Security Solutions, Managed Services

Photonics
Publications include: PIC Magazine, PIC Conference

Expert Moderators

Dedicated technical and time-served experts/editors

MARK ANDREWS
Mark Andrews is technical editor of Silicon Semiconductor, PIC Magazine, Solar+Power Management, and Power Electronics World. His experience focuses on RF and photonic solutions for infrastructure, mobile device, aerospace, aviation and defence industries

PHIL ALSOP
Journalist and editor in the business to business publishing sector for more than 30 years currently focusing on intelligent automation, DevOps, Big Data and analytics, alongside the IT staples of computing, networks and storage

JACKIE CANNON
Director of Solar/IC Publishing, with over 15 years experience of Solar, Silicon and Power Electronics, Jackie can help moderate your webinar, field questions and make the overal experience very professional

DR RICHARD STEVENSON
Dr Richard Stevenson is a seasoned science and technology journalist with valuable experience in industry and academia. For almost a decade, he has been the editor of Compound Semiconductor magazine, as well as the programme manager for the CS International Conference

For more information contact:

Jackie Cannon **T:** 01923 690205 **E:** jackie@angelwebinar.co.uk **W:** www.angelwebinar.co.uk
6 Bow Court, Burnsall Road, Coventry, CV5 6SP. UK
T: +44(0)2476 718 970 **E:** info@angelbc.com **W:** www.angelbc.com

Smart cities and urban data centres – a digital marriage for future connectivity

Andrew Fray, UK Managing Director at Interxion: A Digital Realty Company

THINK ABOUT SOME of the busiest cities in the world. What sort of picture springs to mind? Are you thinking about skyscrapers reaching into the atmosphere, competing for space and attention? Perhaps you're thinking about bright lights, neon signs and the hustle and bustle of daily life. Chances are that whichever city you're thinking about – London, New York, Tokyo, Singapore – is actually a smart city.

Smart cities are so-called based on their performance against certain criteria – human capital (developing, attracting and nurturing talent), social cohesion, economy, environment, urban planning and, very importantly, technology. Speaking specifically to this last point, real-life smart cities are less about flying cars and more about how sensors and real-time data can bring innovation to citizens.

And the key to being 'smart' is connectivity, to ensure that whatever citizens are trying to do – stream content from Netflix, drive an autonomous vehicle or save money through technology use in the home – they can do so with ease, speed and without disruption. Crucial to this, and the often-overlooked piece in the connectivity puzzle – the urban data centre. Urban data centres are the beating heart of all modern-day smart cities, and we'll explore why here.

Becoming smart

Just as Rome wasn't built in a day, neither was a smart city. According to the United Nations, there will be 9.8 billion of us on Earth by 2050. Now, consider the number of devices you use on a daily basis, or within your home – smartphones, laptops, wearables, smart TVs or even smart home devices. Now multiply that by the amount of people there will be in 2050, and you get a staggering number of devices all competing for the digital economy's most precious commodity – internet access.

In fact, Gartner predicts that by the end of this year, there will be 20.4 billion connected devices in circulation. At a smart city level, this means being able to translate this escalating demand for access to the fastest connection speeds to unparalleled supply.

Connectivity is king, and without it, cities around the world will come screeching to a halt. To keep up with the pace of innovation, we need a connectivity hub that will keep our virtual wheels turning – the data centre.

Enter the urban data centre

In medieval times, cities protected their most prized assets, people and territories by building strongholds to bolster defenses and ward off enemies. These fortresses were interconnected by roads and paths that would enable the exchange of people and goods from neighbouring towns and cities.

In today's digital age, these strongholds are data centres; as we generate an eyewatering amount of data from our internet-enabled devices, data centres are crucial for holding businesses' critical information, as well as enabling the flow of data and connectivity between like-minded organisations, devices, clouds and networks. As we build more applications for technology – such as those you might find in a typical smart city – this flow needs to be smoother and quicker than ever.

Consider this – according to a report by SmartCitiesWorld, cities consume 70% of the world's energy and by 2050 urban areas are set to be home to 6.5 billion people worldwide, 2.5 billion more than today. With this is mind, it's important that we address areas, such as technology, communications, data security and energy usage within our cities.

This is why urban data centres play a key role in the growth of smart cities. As organisations increasingly evolve towards and invest in digital business models, it becomes ever more vital that they house their data in a secure, high-performance environment. Many urban data centres today offer a diverse range of connectivity and cloud deployment services that enable smart cities to flourish.

Carrier-neutral data centres even offer access to a community of fellow enterprises, carriers, content delivery networks, connectivity services and cloud

gateways, helping businesses transform the quality of their services and extend their reach into new markets.

The ever-increasing need for speed and resilience is driving demand for data centres located in urban areas, so that there is no disruption or downtime to services. City-based data centres offer businesses close proximity to critical infrastructure, a richness of liquidity and round-the-clock maintenance and security. Taking London as an example, the city is home to almost one million businesses, including three-quarters of the Fortune 500 and one of the world's largest clusters of technology start-ups. An urban data centre is the perfect solution for these competing businesses to access connectivity and share services, to the benefit of the city's inhabitants and the wider economy.

The future's smart

London mayor Sadiq Khan revealed last year his aspirations for London to become the world's smartest city by the end of 2020. While an ambitious goal, particularly with the covid-19 pandemic disrupting many industries, London's infrastructure can more than keep pace. Urban data centres will play a significant role in helping the city to not only meet the Mayor's challenge, but become a magnet for 'smart tech' businesses to position themselves at the heart of the action. The data centre is already playing a critical role – not just in London, but globally – in helping businesses to innovate and achieve growth. As cities become more innovative with technological deployments, there's no denying that smart cities and urban data centres are a digital marriage made in heaven.

The ever-increasing need for speed and resilience is driving demand for data centres located in urban areas, so that there is no disruption or downtime to services. City-based data centres offer businesses close proximity to critical infrastructure, a richness of liquidity and round-the-clock maintenance and security

Maintaining the Edge....

Mark Acton, Critical Support Director, Future-tech SCi Limited & DCA Advisory Board

What is the Edge?

Despite much discussion about Edge, the concept should really be considered a deployment strategy rather than a new technology as some are inclined to suggest. As a concept Edge is not new and has been around for multiple decides. Consider Akamai's Content Distribution Network (CDN), which was being built in the late 1990's and is a clear example of an Edge deployment in all but name.

In reality 'Edge' is something of a marketing term used to describe a strategy using technology that is not necessarily new. It is also evident that 'Edge' means very different things to different people and organisations, depending upon their business requirements and technology base. Equipment deployments ranging from 20 Watts (a so-called Nano Data Centre), to over 5 MW have been described as 'Edge' sites.

It could be argued that the term "Edge Data Centre" does not really relate to a new phenomenon, but is rather a collective name for several different types of data centres which serve a number of different purposes, none of which are new, rather that these have been relatively unrecognised or "unclassified", to date. These are likely to be relatively small remote sites which might include responses to increasing data sovereignty requirements, expectations for reductions in latency and proximal workload processing. Others would argue that the 'Edge' is entirely new and relates to very specific and separate elements within the 'Cloud', which offer new functions and features, particularly in relation to IoT and mobile services. Either way there is a clear perception that Edge sites are different to the more traditional 'Core' data centre and therefore need to be maintained and managed in a different way. This will vary depending on the technology deployed and an organisation's business objectives and requirements.

A better and more explanatory term for Edge might be 'Proximal Processing and Storage", with the intention being to place services closer to the end user for a variety of reasons which might include one of more of the following: application and network latency, network transmission costs, data sovereignty requirements, data protection requirements, advantageous taxation regimes and more widespread data gathering.

The growth of the Edge

One driving factor is the fact that customers have become accustomed to (and expect) instant results or low latency connection to services no matter where they are located. The 'Edge' in this sense is computing facilitated by smaller satellite data centres often themselves supported by one or more larger Core Data Centres which might aggregate data and perform the deep analytics that businesses now expect and increasingly require. The increasing deployment of IoT devices will also drive a degree migration to the 'Edge' but this is by no means the only driver. The Edge as a concept will continue to grow shaped by the demand for Proximal Processing and Data Storage including data sovereignty and regulatory compliance as well as latency etc.

For many businesses there is an increasing expectation and need to have processing power and data held locally to wherever their customers happen to be. This is potentially a major consideration in delivering increased service quality and ultimately increasing both customer satisfaction and providing new or significantly enhanced products. In this sense we could reference the terms 'Competitive Edge' or 'Gaining an Edge'. This does not mean the death of the traditional core data centre though as some have predicted. Core Data Centres will still be required for core functions including the storage and processing of large volumes of data accumulated from Edge sites. This is being offered in different ways by Hyperscale, Large Enterprise, Colocation and Managed Service Providers.

What are the challenges?

Whatever the definition or type of Edge is being deployed there is one aspect that is somewhat being overlooked. We increasingly rely upon, and expect, high reliability of data centre site operations and high (24×365), availability of the services hosted within those sites. This high level of service is often delivered by sites with multiple levels of equipment redundancy, a very comprehensive preventative maintenance plan and dedicated data centre engineers based on site in shifts 24×7. The high (possibly unreasonable), expectations around uptime / availability cannot automatically apply to smaller, remote dark sites which are the standard edge model. How do we achieve this expected level of reliability without staff on site running traditional Planned preventative Management operations? Since Edge sites are typically remote and unmanned there needs to be a new industry approach to providing support for these sites in order to deliver the levels of reliability and availability that customers now expect. This can be achieved if all applications and data hosted within a site have multi-site resilience such as site failover or genuine Synchronous

Replication, however this is not the case in all Edge sites. This therefore raises a challenge. How do we maintain the continued availability of Edge sites when they are remote and unmanned?

One of the biggest overheads for data centre operators is its 'on premise' engineering teams and maintenance services. In the UK the typical support model is to employ a team of onsite engineers that work on shift patterns at a large cost to the business. Where a data centre is designed and built with a high degree of site resilience and correctly maintained it should be recognised that the risk to uptime is very low and a permanent team of engineers based on site is not necessary.

A better way to operate

The typical Edge data centre will be unstaffed and follow a dark site operating model typically with a smaller footprint than traditional core data centres. In this scenario, without full time engineers on site, new operating models and predictive maintenance techniques need to be applied to ensure operational reliability in addition to the standard preventative maintenance plans.

Edge sites should be seen as typically 'dim' or 'dark sites' where engineering support is performed remotely under contract by a third-party operator. This also has the great advantage of significantly reducing OPEX costs. Companies like Future-tech, who are already operating Remote Services for data centres on behalf of multiple customers, are therefore particularly well suited to those wishing to move to a 'Dark Site' Operating model in order to reduce the extremely high costs of onsite personnel, yet maintain site reliability and availability. We need to make more use of Predictive Maintenance, effective remote monitoring tools, intelligent management systems using Machine Learning and Analytics, as well as engineers and engineering teams with multiple skillsets including IT related skills.
Routine site visits do need to be performed, yes, but with greater purpose and range of planned activities / upgrades / installations, depth of knowledge and insight rather than mere adherence to a rigid PPM Schedule.

This new approach along with specially trained engineers who can perform tasks across multiple disciplines rather than the more traditional siloed approach to engineering delivery will ensure that the site infrastructure is kept in the optimal operating condition without the requirement for site-based staff. Additional predictive maintenance techniques ensure that any potential issues are highlighted and dealt with according to an agreed and scheduled plan before they develop into problems in order to maintain site operational effectiveness.

Finally

Maintaining the Edge means looking after Edge sites properly in order to maintain the operational availability and integrity that customers expect. This requires a different set of skills and new approaches to the current traditional Facilities Management way of managing data centres. At Future-tech we take this on board to ensure that we are not only maintaining the Edge, we are also here to help you gain and sustain your competitive Edge.

I am confused, virtually...

Ian Bitterlin, Consulting Engineer & formerly Visiting Professor, University of Leeds

THEY SAY that the best way to learn is to teach, well, for me and Virtual Power Purchase Agreement's (VPPAs) for data centres, that has not proven to be entirely accurate. I cover the VPPA subject in a 3-day data centre Energy training course and, although switching from classroom to live on-line delivery during this pandemic has made it tougher, I find it easy to explain the principle of PPAs but harder to explain VPPAs to the students.

The background is simple. Data centres have come under pressure (from internal PR demands as well as external forces) to be renewably powered and PPAs/VPPAs are the only way to achieve that in power grids have some element of fossil fuel or nuclear generation. For example, the UK, with 55% gas, 20% nuclear, 10% biomass and 15% renewable (mainly wind, low solar) – is constantly varying those proportions during these times of the lowest demand for decades, and notably with almost no coal. The fact that this pressure goes against the principles of sustainability engineering, where the three steps are to 'reduce the demand', then 'improve the process' and, only finally power from renewable sources.

If we followed the three steps in order we would minimise energy and not waste any valuable, finite, often intermittent, renewable energy. Our data centre industry has never done the first step, blindly accepting that the ICT load is exponentially rising and doing everything to encourage that through increased bandwidth and speed, has been self-obsessed with the second step, in the pursuit of ever-lower PUE, and, most recently, we have been acting as if ICT with the data centres at the core has to be renewably powered and 'green'. Why this should be applied to data centres (as opposed to hospitals, care homes, schools, universities or government tax-offices) I cannot justify – except to say that it goes some way to excuse a large growing energy demand from ICT whilst demand from domestic and industrial loads decline.

So, if you want or need to declare your data centre to be 'green' and sustainable, or any other shade of grey using low-carbon nuclear, then you don't need to move it to a location where the grid provides what you want but simply pay for a specific PPA and claim the resulting carbon emissions as your own. It is worth noting here that no energy source is carbon-free, and you can search out (IPCC 2014) data for the internationally agreed grams of CO2/kWh carbon footprint, e.g. for wind 11g and solar-PV 45g.

To demonstrate the national variability, as I submitted this at 7am BST on 22nd June, the instantaneous grid power emissions for England/Wales/Scotland were 193g/kWh, NI 589g, Eire 397g, West Norway 36g (hydro), Germany 392g and Estonia 880g (coal) etc. Go see www.energymap.org for data refreshed every 10 minutes.

You may pay more, or less, than the average industrial price per kWh, depending upon how much you consume and how desperate the supplier is to sign you up, but usually there is a small premium when the renewable capacity is limited. Note that, unlike hydro, wind and solar are highly intermittent and getting above 50% in a national grid brings increasing risks of instability and blackout events such as the UK experienced on August 10th 2019 – where a lack of spinning reserve could not help in maintaining a safe frequency when a large off-shore wind resource tripped off-line. So getting to 'zero' is going to be near impossible in the next 30 years when you add in electric transport and heat-pump domestic heating in a county like the UK – the target is OK but the expectation should be moderated.

All this is straightforward in a single national system, e.g. pay for 100% nuclear and claim to be 'low-carbon' (like Eurostar), or 100% wind and declare to be '100% renewable' and simply ignore the 11g CO2/kWh reality. Of course, the grid is the grid and the weather is the weather, so there will be many periods when there is no wind and its dark, but over one year you will be able to assume that you have used what you have paid for – that is a PPA.

Now comes the bit I struggle with - what is a virtual PPA? Buying energy from a 'remote' supplier, often in another grid (region or even country) and sometimes even from a grid that is not actually connected to the one you are connected to. In some cases, your local supplier will be paid to deliver energy 'sleeved' through their distribution system from another supplier, or sometimes you will jut pay them to deliver what they produce (the local grids fuel mix, e.g. the UK's 193g) and you will wave and shout about your renewable energy credits/certificates at anyone who will listen. The supplier who has sold you the certificates (which can be less than, or all, or more than, your actual demand) will simply pump out their production into the local grid but not attribute any carbon content to any local consumers.

If you get the principle and are OK with it, then well done. It is perfectly legal, and it appears to be commercially acceptable in our, and other, industries – but, to me, perhaps in isolation, it appears to be a slight-of-hand, even slightly deceptive? Consider the extremes: A hyperscale facility in the USA that is located in an area where the power is cheap (reportedly a 25-year deal for US$3.5c/kWh based on brown-coal production) that consumes 40MW from the local grid but buys/claims all the certification output from a 100MW wind facility in a remote regional grid that is only interconnected via another regional grid.

The energy does not originate in one place and be consumed in another, it does not 'flow', and, somewhat worse is the fact that when the wind blows hard in the remote region it produces 100MW at 11gCO2/kWh and when the air is still it produces nothing but the average output over one full seasonal year is 40MW - yet the data centre ploughs forward burning energy from the highest carbon source of 890gCO2/kWh continuously at 40MW demand.Or the London data centre that is powered by 100% Norwegian hydro, again perfectly legal, bought and paid-for, but the energy does not specifically flow from Norway to the UK, only the paperwork?

If you are as confused as me then consider the case highlighted be the BBC video from the 5th March 2020 on BBC Three and iPlayer called 'Dirty Streaming: The internet's big secret' – a video originally intended for schools to highlight the impact of the internet on energy and climate change. It covers the huge data centre infrastructure (>120 data centres, >1million m2 of white-space, >10GW of electrical demand) of Loudoun County's Ashburn Data Centre Alley, Virginia, currently 100% coal fed by the monopoly grid supplier, Dominion Energy.

I almost apologise for the self-promotion but I was pleased to be included in the video, so beware, but, if you can stand it, get to the last couple of minutes and compare the statements from Dominion and only one of the many data centre businesses, AWS, that has a very large footprint – not to say massive by European standards.

Dominion state that their plan is to be '100% carbon and methane free by 2050' but AWS are more ambitious by 20 years. And how is that gap possible? A VPPA that AWS will, no doubt, achieve – while their actual consumption will be 100% coal-based probably until well after 2030. 10 years is a short time in which to re-fuel an entire grid system that is larger than many small countries.

Let me know what you think about the push for 100% renewables in data centres before other industries and the VPPAs that will get them there and the full video can be streamed at https://www.bbc.co.uk/programmes/p083tb16.

enel x

Energy management solutions that deliver commercial and environmental rewards

Global resources, local market knowledge and specialist technological expertise make Enel X the trusted energy management partner for data centres and large scale enterprises all over the world.

Join us as we pave the way for a cleaner, more dynamic future.

enelxuk@enel.com | www.enelx.co.uk