



**INNOVISION**  
DATA CENTRE INSIGHTS + PERSPECTIVES



# Data centre infrastructure

The foundation for digital transformation

A SPECIAL ISSUE OF DCS IN ASSOCIATION WITH





TRAINING

DEVELOPMENT

SKILLS



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

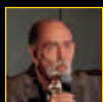


**Photonics**

**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



**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 overall experience very professional



**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



**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, Burnshall Road, Coventry, CV5 6SP. UK  
**T:** +44(0)2476 718 970 **E:** info@angelbc.com **W:** www.angelbc.com

# Editor's View

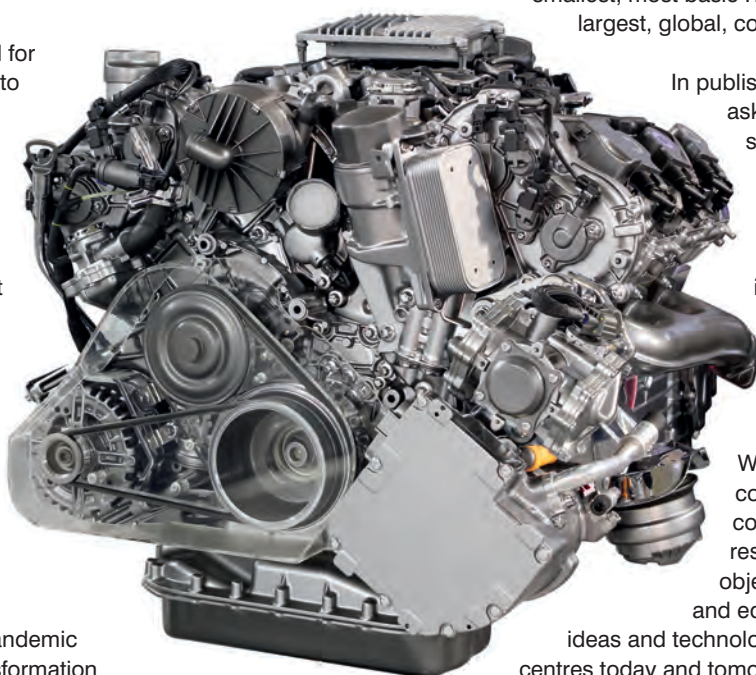
By Phil Alsop



WELCOME to the very first issue of Data Centre Solutions' annual snapshot of the data centre industry – DCS Innovision – with one eye very much on the future. As with so many new business ideas right now, it was borne out of the strange circumstances of lockdown which we have all experienced to a greater or lesser extent.

Our physical events planned for the spring and summer had to be put on hold, and our thoughts turned to both virtual events and some value-add publishing initiatives. We're organising virtual events focusing on Digital Change Management for the SME community and the emerging discipline of AIOps – both planned for late summer/early autumn. And DCS Innovision is the first of our new publication ideas, with AIOps (again) and storage networking guides planned in the near future.

Even before the Covid-19 pandemic struck, digital business transformation was on the agenda for most, if not all, organisations. And the requirements of lockdown seem to have only accelerated this focus, as companies, their employees and their customers, all look for new ways of interacting. The data centre industry is the foundation of the digital transformation drive.



There is no cloud, no edge, no AI, no 5G, no social media, no e-commerce, almost no IT whatsoever, without a data centre 'lurking' somewhere in the background. Unglamorous they might be but, as the combustion engine is to all shapes, sizes and costs of vehicles, so the data centre industry is to the smallest, most basic IT requirement, right up to the largest, global, complex infrastructure.

In publishing DCS Innovision, we asked the companies that supply the products and services which make data centres what they are to give us some thoughts as to what the future of the industry might look like, with particular reference to how they need to adapt to end user requirements.

We believe that we've compiled a fairly comprehensive and valuable resource, and met our objective of providing insight and education in terms of the ideas and technologies which are shaping data centres today and tomorrow.

Feedback is always welcome, and for those of you who missed the chance to contribute this year, we plan to repeat the Innovision exercise in 2021 – hopefully in a world where the coronavirus is, at least for the time being, a thing of the past.

## DCS DATACENTRE SOLUTIONS

<b>Editor</b> Philip Alsop	+44 (0)7786 084559	philip.alsop@angelbc.com
<b>Sales Manager</b> Peter Davies	+44 (0)2476 718970	peter.davies@angelbc.com
<b>Sales Executive</b> Jessica Harrison	+44 (0)2476 718970	jessica.harrison@angelbc.com
<b>Design &amp; Production Manager</b> Mitch Gaynor	+44 (0)1923 690214	mitch.gaynor@angelbc.com
<b>Publisher</b> Jackie Cannon	+44 (0)1923 690215	jackie.cannon@angelbc.com

**Circulation & Subscriptions** +44 (0)1923 690214

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. © Copyright 2020. All rights reserved. Contents may not be reproduced in whole or part without the written consent of the publishers. The paper used within this magazine is produced by chain of custody certified manufacturers, guaranteeing sustainable sourcing. ISSN 2396-9016 (Online)



# Data centre risk management and the next normal



Everywhere we look in our industry today, we see the impact of the pandemic and the approaches rapidly put into place to manage and cope with the crisis. That said,

the data center industry with our mission critical criteria and consistent planning is perhaps one of the industries most prepared to carry out the requirements needed for business continuity and beyond.

**BY CHRISTOPHER BROWN, CHIEF TECHNICAL OFFICER, UPTIME INSTITUTE.**

THE EFFORT CAUSED by the COVID-19 pandemic has been a herculean effort of remote and onsite collaboration, but certainly not without issues. An Uptime Institute survey of over 200 critical IT/facility infrastructure operators around the world found that a third of operators said a reduced level of IT infrastructure operations staff posed the single biggest risk to operations. This is not really surprising — earlier Uptime Institute Intelligence research has shown that insufficient staffing can be associated with increased levels of outages and incidents; and even before the pandemic, 60% of data center owners and operators reported difficulty finding or retaining skilled staff. With the onset of the pandemic, most of the concerns continued to be around existing infrastructure and operations as a direct response to staffing abilities, along with the need to make sure that additional risks were not introduced into data center operations.

When does this new criteria, like remote operation, reduced staff, deferred maintenance and social

distancing become “normal,” and as we loosen some of the lockdown criteria, how do we incorporate some of these behaviors into the “Next normal?” We must closely review what works well in our business along with how to best protect and manage staff in our practices. We have no choice in our adjusting to the pandemic, but also must look forward to what becomes normal moving forward. Our industry has always clearly understood that any operations best practices means always planning ahead, including preparation for future outcomes. And we must always operate in safe environments. A balance of planning for safe and effective operations along with an understanding of what staff requirements are required to reduce anxiety related to the pandemic and resulting employee issues is our ultimate goal.

With this in mind, we have to look forward to better determine what the “Next Normal,” has in store. The criteria, some of which is derived from the pandemic response to continue forward perhaps more permanently for best results may include:  
Some remote abilities become necessary – this is something to work into our planning and execution  
Better and more frequent collaboration between teams both on and offsite is vital, and working smart here can make a difference to success.

Continued best practices in the handling of customer access is critical both now and into the future  
Supply chain management and related planning to ensure business continuity.

It is also important to continue to review business continuity plans to make any changes or adjustments in your business to help customers work with you according to their ongoing needs. How can you help customers reduce risk? Can you bring creativity and knowledge to the table to help? With the entire world moving toward some implementations of remote business, can the data center industry rise to the occasion?

Uptime Institute has recently done just that with the addition of remote and on-site service delivery across the organization’s full digital infrastructure portfolio. New remote-compatible versions of the globally recognized Tier Standard Certifications, Accredited Educational curriculum, Data Center Risk Assessment, Management & Operations Stamp of Approval, and Facilities Management and Operations services can be utilized globally, even with non-traditional business operating plans in effect. Final Certifications and awards are finalized onsite. This can assist in reducing service delivery risk associated with digital infrastructure design and operations especially during the challenging times of the pandemic with the same overall depth and value by utilizing the remote-compatible services.

Reducing risk during these difficult times necessitates changes in our behavior professionally and personally, and in doing so, we need to review the best practices that come out of this situation and determine how to take these behaviors with us into the future.







# A new decade with new challenges calls for new data centres

By all accounts, the new decade has gotten off to a pretty tumultuous start. The COVID-19 pandemic has meant 2020 has been a year of big changes for us all – many businesses either no longer operating as business as usual or implementing working from home measures.

**BY STEVE BRAR, SENIOR DIRECTOR, PORTFOLIO MARKETING AT ARUBA**

THIS HAS MEANT now more than ever people and businesses are relying on tech to get their jobs done and provide them with some escape. With this in mind, one particular area facing major changes are our data centre networks. The proliferation of data centres in the 1990s and 2000s was subsequently followed by the centralisation of computing and its infrastructure. However, the rise of the public cloud which consolidated this centralised approach has now led to the pendulum swinging once again back to decentralisation and the edge.

In fact, according to Gartner, today 90% of data is created and processed inside centralised data centres or the cloud. By 2025 this will have massively changed, as about 75% of data will need to be processed, analysed, and acted upon at the edge.

### So, just what is the edge and why does it matter?

Simply put, with the edge or rather edge computing, IT infrastructure is moved to edge data centres that are located either directly on the premises or at locations near to end-users. It's the proximity of edge-data centres to end-users that is their strength – namely the ability to support applications that require more bandwidth, quicker response times or are latency-sensitive.

The swing back to the edge and decentralisation has largely been driven by digital transformation and the increasing need to harness data from devices using 5G, IoT or AI in order to create real-time, connected experiences. In fact, according to a study carried out by Gartner, the next generation of business innovation will be powered by the likes of 5G, IoT and AI. These

technologies will drive a massive amount of data production, with more than 7.5 billion connect devices being used by businesses in 2020 alone. Yet, these devices will not be located in centralised hubs or in the hearts of vast metropolitan cities alone, rather they will spread out throughout the populations they service. This means in order for them to reach their full potential – for example, say with 5G with its promise of low latency and high levels of interconnection – our data centre infrastructure will need to change, the fact edge data centres can be located either on or nearby to where they're needed is what makes them crucial for these technologies.

## How will data centres change?

With such a fundamental change you can undoubtedly expect changes to your data centre. Firstly, data centres as we know them will continue to shrink as workloads continue to move to the cloud but largely due to higher density from hyperconvergence. Secondly, the emergence of the aforementioned, edge data centres – these will appear more and more as workloads are moved away from the 'centres of data' to optimise performance and cost. Enterprise-owned data centres will likely come to consist of two types:

1. A blending of traditional centres and campus environments.
2. Mini data centres within IoT-enabled environments, such as large manufacturing centres.

Finally, application teams will continue to adopt more DevOps and agile practices. This will put more pressure on network teams to optimise around workload-driven operations.

The networking requirements of the future  
As infrastructure changes with the rise of edge data centres, you'll also see changes in what the requirements of our networks. This will become especially prevalent as you seek to balance the needs of edge data centres with the increasing use of cloud services and any on-prem footprint you still have. Below are a few considerations to keep in mind as data centres move towards the edge, in fact, these three requirements are applicable for any type of data centre be they private cloud, co-location, and edge.

### 1. Using automation streamline

With application teams continuing to adopt DevOps and other agile methodologies to speed up their software development, you can expect to see networking operations to become far more automated and streamlined to keep up.

Solutions needed to achieve this will need to take into account both current and future operating models as well as any existing investments. In order to simplify time-consuming yet common configuration tasks, turnkey automation could be considered. Whilst for teams with more developed DevOps, extending common automation platforms to network-related workflows will become an absolute must.

Furthermore, as DevOps and agile practices are utilised more and more in application teams, expect those practices to begin to influence how other organisations within IT function.

### 2. Getting ahead with analytics

Often one of the most time consuming and biggest drains on network operations resources is troubleshooting issues. Network visibility to incredibly important in shortening mean time to repair (MTTR), improving IT service delivery, and helping keep short-staffed teams focused on other more strategic matters.

Network-wide telemetry, captured and processed natively on each node, will be a massive leap forward once achieved. These analytics with built-in remediation will be important in providing better network stability and helping trouble-shooters proactively or even pre-empt user and/or business-impacting issues.

Predictive analytics will not just be major in dealing with issues before they arise, but also for gauging when there will be periods of high network usage – in the era of big data, cloud, streaming services and soon IoT, latency and connectivity have been the bane for many a data centre, thus being able to predict and ensure the network is right-sized to deliver the experience users demand will be vital.

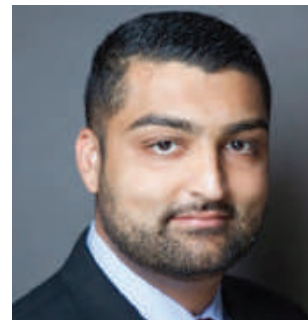
### 3. Maintaining 24-hour availability

As lockdown measures have been implemented and more of us are working at home, networks are under increasing strain as we all seek to use data sharing service. With this, the need for highly resilient networks will continue to grow – especially as even a minor hiccup can have huge ramifications for a business' productivity. Automating day-to-day operations will remove the possibility of human error, however, what networking teams will need is a far simpler and more reliable way of ensuring availability and delivering upgrades in a non-disruptive manner.

A cloud-native, microservice based operating system could be a great solution, ensuring added resiliency at a software level while also being able to utilise live software upgrades to eliminate maintenance windows will also be key.

## Conclusion

As with previous decades, we are seeing a paramount shift in how we structure data centre networks with edge data centres being the latest continuation of this trend. Now, in all likelihood you're probably already on your way toward this new era of data centre networking, however, we hope the considerations above provide you with some guidance and food for thought on how best your business can manage these changes. It might be early days for edge data centres, but it is already clear they are going to be playing a key role in shaping the future.



Steve Brar,  
Senior Director,  
Portfolio  
Marketing at  
Aruba

# AI + AUTOMATION

Artificial intelligence (AI) seems to be the buzz technology right now. In terms of the data centre industry, AI in all its guises (machine learning, IoT, robotic process automation and the like) is placing increasing strain on both data centre capacity and infrastructure, as more and more applications and workloads require more, and faster, compute, network and storage resources. The data centre infrastructure which houses this hardware cannot be the bottleneck when it comes to applications performance. AI is also finding a place within the data centre, in terms of providing a degree of intelligence and automation which helps to ensure facility optimisation.



IN ASSOCIATION WITH



**DATA CENTRE**  
**SOLUTIONS**



**DIGITALISATION**  
**WORLD**



# Deploying artificial intelligence and machine learning in MTDCs



Artificial intelligence (AI) and machine learning (ML) are the most revolutionary technologies we have seen since electricity. More powerful than the internet and mobility revolutions combined, these technologies are so impactful because they make sense of vast amounts of data quickly and efficiently. We live in a data-generating and data-driven world and without tools to make sense of the data we would drown in it.



**BY CINDY RYBORZ, MARKETING MANAGER DC EMEA, CORNING AND JUAN PENARANDA, DC MARKET DEVELOPMENT, CORNING.**

TO GIVE A QUICK EXAMPLE, this year the world will create roughly 40 zettabytes of information. That's 40 TRILLION gigabytes of information. There's no way humans can make sense of all that information – even if every human worked day and night together, it is mathematically impossible. On the flipside, AI and ML LOVE data – it is their oxygen. By using powerful and properly trained AI / ML models we can accurately process vast amounts of information, revealing the very valuable bits of data we need to act on.

## ML model training

A good example of this is MRI ML models. They are tested against known outcomes of cancer that are confirmed to be cancerous or non-cancerous (both positive and negative results). This is called the training phase. Then, a new set of MRIs are loaded into the now trained model and are analysed. This data also contains MRIs with known outcomes, but they have never been seen before by the model. These new MRIs are called a validation dataset. This data is run through the now trained model and the results are computed and displayed. The results of the validation data are then evaluated for the performance metrics chosen for that model. If the results are acceptable then the model is trained and ready for more testing/validation or live deployment. Finally, the validation phase – if the validation data fails to meet

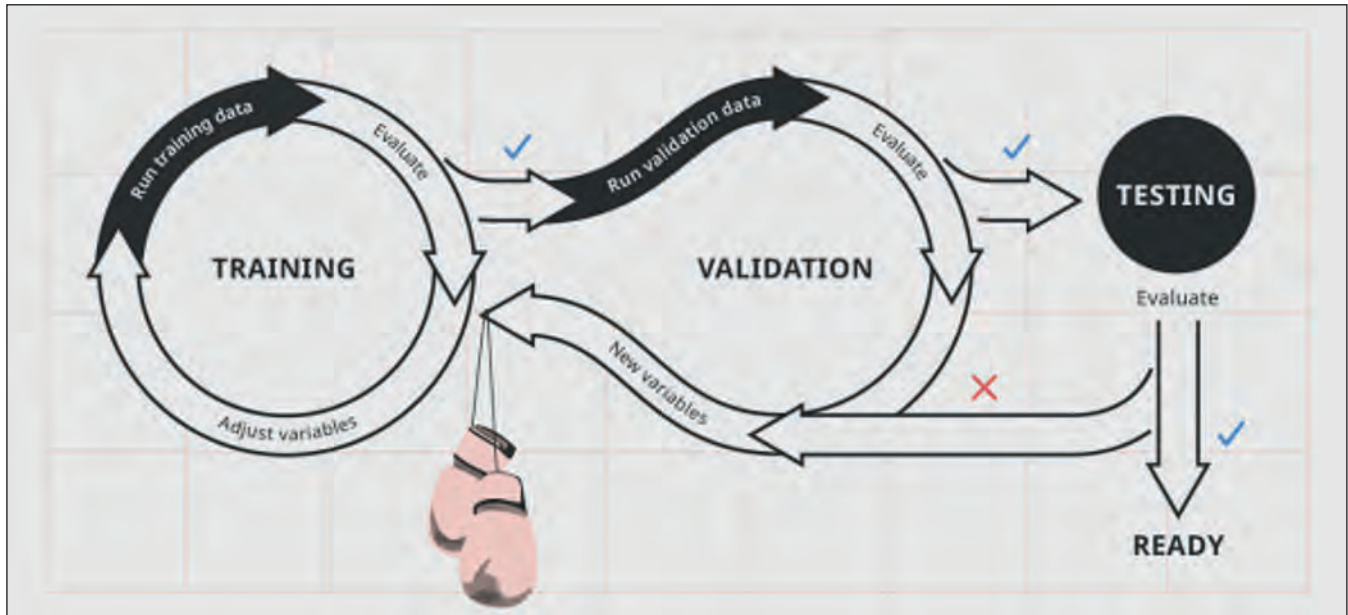
the metrics, then we circle back and either redesign the model or give it more data to be better trained for the next validation test.

## How to cable and deploy AI / ML?

The traditional problem with AI has been its enormous processing power requirements. Thankfully companies like Nvidia, Intel, AMD, and many more are closing the processing power gap. This enables companies such as BMW and countless more to deploy AI Edge capabilities. This entails installing powerful hardware on premises that will process your on-premises data using the pre-trained model. This dramatically cuts down on latency and the requirement for real-time bandwidth.

However, the problem is that no one can really do an “edge only” AI/ML deployment because while the hardware can handle the data processing of pre-trained models, it's not powerful enough to train and update the model – much more powerful hardware is needed for that.

This is where the hybrid approach comes in. In a typical hybrid design, the Edge server will process all on-premises data utilising the trained model we discussed earlier. These edge servers will select MTDCs as the best fit for their key drivers, allowing



for flexibility as the network and applications evolve. For the optical infrastructure, MTDCs typically deploy single-mode fibre to enable the end users to scale. For the company deploying AI / ML, it's important to account for both the bandwidth of today and tomorrow in their network. Citadel Analytics typically has a rule of thumb, which is to take the average bandwidth expected and multiply by four. That is how much bandwidth a company should be accounting for within their system. Bandwidth is the first thing to get eaten up in real world AI/ML deployments.

## The need for higher density

This uptick in bandwidth is also highlighting the need for high density solutions for both the MTDCs and the end users. The MTDCs maximise their revenue generating white space (space allocated for IT equipment) while the end user can efficiently utilise the white space they've invested in. This space and infrastructure can come in many shapes and sizes, but overall, it's important to look for a provider with product breadth that supports anything you want to do (single-mode or multimode, LC or MTP).

For the end user, one way to reduce the total cost of ownership is to increase the density and reduce power consumption. This can be enabled through parallel optics and utilising port break out. This application allows the use of higher speed transceivers being broken out in the cabling. It is possible to take a 40 Gigabit transceiver and operate four individual 10 Gigabit transceivers, which can allow for high density switching. To really take advantage of AI and ML, there will need to be a network of interconnected data centres providing the computing closer to where the data is being created. This creates a need to globally scale with a consistent and modular product, which offers a full solution from the edge to the central data centres.

MTDCs are the lowest risk investment and provide a faster path to revenue due to the speed of deployment through availability. Now to train and update the model, all the data is sent to the training servers located at a private data centre, primary a multitenant data centre location, or the cloud. Those training servers will use the new data to automatically train and enhance the accuracy of the model being used.

By going to less than 10km, providers can often reduce latency by 45%. This means that while there will still be a need for a central data centre, there will also be a push for smaller and more regional data centres closer to where the data is being produced. MTDCs will be the main vehicle for these smaller and regional data centres. Edge data centres will be an extension of, or hosted at, these interconnection dense MTDCs, with both services needing each other to provide a full service for the customer and network. MTDCs with the most interconnected facilities and ecosystem rich customer mixes will capitalise on the initial benefits of the edge data centre opportunity. What can we look forward to?

AI and ML applications are here to stay. There is a clear business and human case for the magnitude of revenue and productivity these applications will provide to adopters. The shift to more bandwidth at the regional level isn't optimised with the current architecture (while maintaining the lowest cost and highest quality of service).

We will see these applications running at the edge which for most companies are at MTDCs. MTDCs offer the lowest risk and fastest speed to revenue for these companies. This will result in a push for more interconnected facilities with higher density solutions in more locations, rather than in larger central hubs as we've seen before.

# Play it cool with the digital twin

Streaming movies, playing computer games or finding the latest hidden musical gems are getting easier with each passing year. Just look at the last two months of 2019. We've seen the launch of Disney Plus, Google Stadia and Spotify providing its 248 million users with custom made playlists. Powering all of this and more is the hidden force in our internet age – the data center. However, while being a booming industry, all is not quite so rosy for data center businesses. The growth in usage coupled with the 3.6 billion people who are now connected is putting pressure on the internet infrastructure like never before.

**BY JONATHAN LEPPARD, DIRECTOR AT FUTURE FACILITIES.**



THE PRESSURE is being felt too. In fact, in a recent independent report, commissioned by Future Facilities, it was found that over three quarters (77%) of decision-makers are feeling the strain of a consistent increase in demand. This increase in data center density and workloads, naturally results in more power being needed, which in turn generates more waste heat. Consequently, areas such as cooling and power are now monitored more than ever before to ensure downtimes are avoided.

Juggling the need to avoid downtime and prevent costs from dramatically increasing has seen data center operators and business decision-makers search for solutions that can help. While adding in extra cooling or power can often feel like the answer, these are expensive solutions that overlook the true state of the data center.

### Increase capacity, and capital expenditure

This is a worry for both the operations and business teams. As found in the research, the average cost of downtime is £122,000. That's a large sum of money,

and can quickly multiply if a business suffers multiple outages across a year. This figure doesn't even begin to cover the loss of loyal customers and damaged reputations. It's no wonder then that businesses are so keen to avoid outages. We need to find solutions that lower the risk of downtime, but also operations terms to increase the capacity of the data centre; and all in a way that is both cost-effective and quick to deploy.

So how are businesses currently coping? Well, as found in the research the three main areas that businesses focus on are power, cooling and networking solutions. In fact, 45% of all businesses admitted they are focused on investing in these specific areas in order to increase capacity while maintaining or decreasing the risk of an outage.

The result is though, that with all the extra cooling and power installed businesses are operating their data center far below its optimised level. The net result is that data center capacity is being left on the table, at exactly the time when it is most valuable. In fact, as an industry average, businesses are over-provisioning



to the tune of 36%! This is not just wasting power and capacity but also hitting the bottom line of businesses too.

### Digital twin on the rise

With a lack of knowledge on how data centers are performing, businesses are blind to simple, cost-effective alternatives. A key solution to this challenge is to effectively utilise any stranded capacity without having to spend on additional infrastructure. Enter the digital twin.

A Digital Twin is a 3-dimensional physics based computer model that replicates the precise behavior and performance of the real data center. The Digital Twin can be used to predict the response of the real data center to any change, but in a risk free virtual environment. By quickly trying out different setups and layouts, businesses can make small adjustments that have a huge impact on the thermal efficiency of a data center. Once a setup has been found that achieves the desired results, it can then be implemented in the physical data center itself, all safe in the knowledge that it has already been rigorously tested. This process saves money both in designing and operating a data center.

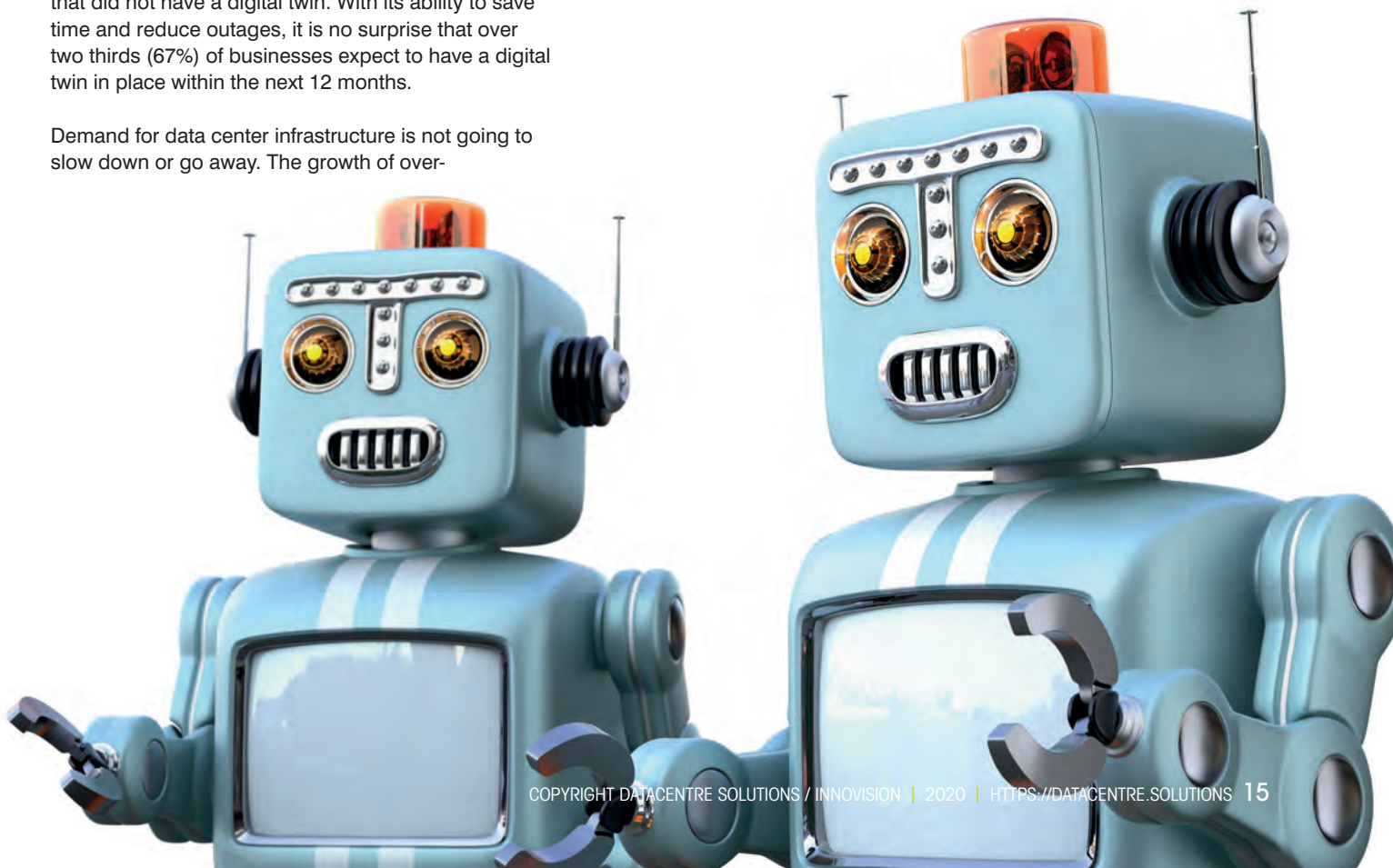
As well as helping to reclaim stranded capacity, digital twins have also proven very effective in reducing downtimes. In fact, the research found that three times as many businesses that have used a digital twin did not see an outage in a 12 month period, than those that did not have a digital twin. With its ability to save time and reduce outages, it is no surprise that over two thirds (67%) of businesses expect to have a digital twin in place within the next 12 months.

Demand for data center infrastructure is not going to slow down or go away. The growth of over-

A Digital Twin is a 3-dimensional physics based computer model that replicates the precise behavior and performance of the real data center. The Digital Twin can be used to predict the response of the real data center to any change, but in a risk free virtual environment

the-top (OTT) services like Disney Plus, as well as increased AI applications, means that data center capacity and density is going to be pushed to even greater limits. However, the good news is that there is a proven solution that is ready to help.

Already delivering benefits for data center businesses, the digital twin has become a key tool across the industry in 2020.





## AI on the edge



DCS Innovision talks to Keysource's Head of Innovation, Richard Clifford, about the launch of a new facilities maintenance service - Keysource Artificial Intelligence (KAI) - which uses machine learning and artificial intelligence technology to help organisations transform their edge operational management strategy.

**Q** Keysource has launched Keysource Artificial Intelligence (KAI) – can you give us the high level overview of this new service?

**A** KAI is a sensor and cloud based infrastructure monitoring solution that leverages AI / ML to provide our clients insight into their infrastructure behaviour globally.

Following a comprehensive audit of your estate, our critical operations team will deploy an array of the latest sensor technology to create a digital fingerprint of your infrastructure.

This digital fingerprint is then monitored 24/7 and used to outline 'normal' behaviour and highlight changes in state, through a connection to our specialist AI / ML service 'KAI'.

Leveraging our 40+ years' experience KAI will learn the relationship between changes in it's state / digital fingerprint and it's corresponding trigger. Thus ensuring the next time that incident occurs it can be automatically diagnosed and a corresponding remedial action recommended without delay.

This combination of best in breed technology, and specialist experience means you will be able to identify changes in state up to 9 months in advance, to;

- better plan and coordinate repairs or replacements around normal business in a more sustainable way (i.e. reduced travel) and at significantly reduced cost (i.e. coordinate with planned maintenance),
- facilitate repairs before they cause damage and incur additional expense (i.e. refrigerant leak)

leading to compressor failure).

- take the guess work out of life cycle replacement and prolong the life of infrastructure putting off capital spend by understand how your infrastructure is operating.
- support in providing global reach through better utilising planned maintenance windows.

**Q** It focuses on edge operational management strategy – can you explain how this works in practice? In other words what specific edge issues are being addressed?

**A** The key EDGE challenges KAI helps to overcome include;

- Ensuring experienced engineering capability can provide greater coverage over remote edge estates with greater levels of consistency

#### 1. Engineers;

- a. cant be all places at once.
- b. don't have the same experience.
- c. can become ill.
- d. will be unavailable due to

#### i. time off.

#### ii. Illness.

- e. can be distracted by family and life
- f. and the experience of the engineering teams can ebb and flow as engineers progress their careers.

#### ii. Greater visibility over the behaviour and health of infrastructure.

#### iii. Reducing the need for specialist engineers to travel to remote edge locations.

1. Reducing cost through reduced travel.
2. Increasing availability of specialist engineers for planned works.
3. Improving the sustainability of edge support through reduced travel and associated carbon produced.
4. In current pandemic, reduced risk to personnel.

#### iv. Improving the consistency of diagnostics, and remedial activities through KAI's sensory network replicating human senses, 24/7 monitoring and reduced optimised planning through early automated diagnostics.

What we typically see is facilities maintenance teams are fixing things that are broken too late and fixing things that aren't broken too early. We aim to make this a more exact science by leveraging KAI's 3 core attributes; a more human like sensory network, coupled with the latest AI / ML service and our own experienced engineering teams.

Our aim is to help our industry overcome some of its key challenges facing it. It is widely acknowledged that there is a shortage of specialist experienced engineers within our industry, which is being exasperated by the

ever expanding 'Edge' infrastructure estates required to meet societies growing need to consume data, and that's not to mention the added travel challenges associated with the current COVID-19 crisis. We asked some simple question about what challenges were being faced, which drove our creation of this service. I have outlined a sample below;

- How do we make the most of the resource we have?
- Can the most experienced engineers be in all places?
- What if you most experienced engineers leave?
- What if they are ill, on lunch or unavailable?
- Are they tired, stressed or distracted?

Experience ebbs and flows for various reasons, so to gain greater consistency and optimise diagnostics and maintenance activities we're leveraging a sensory network that mimics human senses coupled with an AI / ML service to automate the evaluation of the digital fingerprints and data and backed up by Keysource specialist and experienced engineering teams.

By its nature, Edge infrastructure is commonly geographically disparate. A great example of this is where Keysource completed Edge installations for global technology company across a number of sites that included the Arctic Circle and Dubai that were 7,000 km apart with environmental conditions ranging from -35C to +50C.

This is just one example of how big the geographical spread of edge can be, and why we should be leveraging the tools at our disposal to work smarter.

**Q** And can you talk us through how the KAI solution is implemented – from the initial infrastructure audit, through the digital fingerprint and on to the data produced?

**A** There are 5 key steps we take;

- KAI starts with an engineering audit of our clients infrastructure to determine the current condition and what key elements of infrastructure we would recommend extending the sensory network too.
- We agree the best approach to ensure security is not compromised and following this critical stage connect the sensory network to the KAI service in the cloud.
- We then allow the system to run to generate a digital fingerprint of normal behaviour.
- Once this has been generated and reviewed by our engineering team we are ready to go. Any difference to the digital fingerprint (outside of agreed tolerances) would be highlighted for review by our engineering team.
- Like an engineer the lessons learnt are shared with KAI to ensure the next time this change is seen it is able to diagnose the problem and speed up the action and remediation process.



## AI + AUTOMATION

**Q** Potential benefits of implementing the KAI solution include: better planning?

**A** Correct. With a greater understanding of the infrastructures behaviour we are able to provide early identification of potential issues and in many cases are able to remediate around a planned preventative maintenance regimes to improve engineering utilisation, reduce cost of travel and also reduce the cost to repair as we are able to identify many issues before they cause a an outage or failure requiring a replacement part (i.e. refrigerant leak leading to compressor failure).

**Q** And are KAI's operating parameters and, therefore, advice, pre-set, or are end users able to set their own policies and parameters?

**A** Typically the approach is standardised to ensure learning from the overall monitored estate can be leverage but we will work with users to tailor what plant would benefit most from this service and set custom thresholds for changes in state and therefor what triggers an alert associated with a change of state.

**Q** Before we finish, can you give us some thoughts as to how the overall edge computing/ infrastructure market is developing right now and your best guess as to the future?

**A** There is a strong consistent growth in edge requirements. This is in direct correlation with increasing consumption of data by society and the growing demands for low latency and high bandwidth connections that require close proximity to end users.

Key items driving this trend include:

- Gaming / Streaming.
- Smart Cities
- Autonomous vehicles
- Home automation.

Edge can be delivered in any number of ways including;

- On Premise regional facilities.
- Regional Colocation / Managed services.
- Containerised / Modular deployments.

We foresee a continued growth in edge. And the huge disparate spread in more and more remote areas.

**Q** Finally, the KAI solution focuses on the edge for now, are there any plans to expand its coverage to core or centralised data centres, to produce similar benefits?

**A** Absolutely. Whilst we have created this to tackle the growing challenges with managing the edge, with many of the benefits being mirrored with core or centralised data centres.



## Zack Zilakakis

Marketing Leader, Apstra

<https://apstra.com>



# Network automation helps see the way ahead

WHEN YOU ARE DRIVING A CAR, do you watch the road ahead? Or the dashboard? The answer is that you watch both – in a different way. The dashboard supplies numeric data – miles traveled, speed, revs, and fuel levels. Useful, but not immediately relevant to the driver's intention to reach the destination. (Though police hope intention includes keeping speed limits.) The road ahead is more closely aligned to intention – landmarks, road signs, and miles to go. Above all, the intention not to crash! Any obstacle suddenly in the way, and we automatically swerve, brake, or accelerate to avoid it before we even have time to think. Very different from the slower way we respond to the dashboard. Operation teams are typically notified of performance or availability issues after they've occurred. Last quarter was down 5% – damage already done. For current visibility, they use the system dashboard. After 20+ years of monitoring best practices, we are biased toward reducing outages, rather than eyes on the road and actual business objectives.

Dashboards designed under previous circumstances are not ideal for today's operations. Data addressing past problems rather than looking ahead to meeting business objectives. Initiatives embracing cloud, edge, and cloud-native application deployments mean that operators have to support complex, dynamically evolving projects. In the past, each new initiative added another vendor-recommended monitoring solution, building up an increasingly elaborate dashboard. Faced with constrained budgets and headcount we struggle to maintain the availability and performance of new as well as legacy services. Jeff Bezos' "Day 2" philosophy says that an expanding organization tends to rely increasingly on process. When process becomes the norm, companies overlook outcomes and focus on doing the right thing. The dashboard – not the way ahead. To optimize ongoing "day-2"

operations, we must integrate operational solutions and processes during the "day-1" planning and implementation phase. This requires pruning the dashboard and consolidating all essential monitoring functions into a single source of truth – making it easier to correlate and contextualize their information. More so with microservices-based development. Microservices are dispersed, their scale and transience make it unfeasible to discreetly track every interaction. Instead, a baseline developed during development and deployment is used to proactively identify later anomalies. A single source of truth is vital to understanding the infrastructure and how to isolate and diagnose operational events. This implementation phase offers rich data – defining a baseline for subsequent operations.

Automating this process empowers IT organizations to deploy complex solutions faster and more successfully. The network produces mountains of convoluted data, making it hard to "connect the dots" across multiple resources and isolate the issue and negative business impact. Machine learning on a single source of truth leads to automatic identification of anomalous behaviors and an advance warning allowing prompt pro-action without re-occurrence. Intent-based analytic solutions replace data from multiple disparate monitoring tools with a single source of truth, combining "rules-based" and machine learning towards awareness of the entire infrastructure. Cars too have become more complex, but their dashboards become simpler as control functions are automated – freeing our attention for the way ahead while growing faster and safer. Autonomous vehicles are next, – the same with networking. Intent-based analytic networking will automate responses and create an autonomous network that allows operators to focus on business objectives. It's what "intent-based networking" is all about.



**Michael Skurla, CTO**

BitBox USA, LLC

[www.bitboxusa.com](http://www.bitboxusa.com)



## The importance of IoT platforms

WE ARE LIVING through trying times that test operational efficiency across all industry verticals. Efficiency is no longer focused on energy or OpEx savings as a whole, but rather how critical facilities can continue operations when employees themselves are obstacles to the operation.

How do you maintain critical infrastructure operations when employees can't (or shouldn't) regularly service and manage the infrastructure?

As an industry, we have relied on building management systems (BMS) which primarily focused on the energy efficiency and operations consistency of key services in a data center environment and local staff availability to service and troubleshoot operational problems.

Interestingly, the Edge data center and telecommunications markets have tackled this problem over the last 5 years as build outs of edge infrastructure have increased due to content delivery and cellular infrastructure upgrades. Aside from the BMS, IoT Platforms have become the 'solution of choice' for these remote and unmanned operations.

Edge infrastructure has always needed to be autonomous—in traditionally wide geographic footprints, with installations typically in locations difficult or impossible to access on short notice. The IoT platforms offer a differing approach to their big brother data centers; focusing not on energy management and fault detection on specific systems but operating on the premise of a wholistic data collection from all the subsystems in a location. Harvesting that information back to a central data-lake for wholistic operational analysis through flexible BI/AI tools,

it can view a portfolio of data from many sites and make both tactical and strategic choices on the health and requirements of the entire portfolio.

IoT platforms focus on data enablement of existing infrastructure to be used for the consolidated well-being of a facility while BMS solutions focus only on their trade or simple automation. The data is the key here.

IoT platforms are designed to collect, organize, and deliver actionable data. They collect and harvest as much data from subsystems at each site as possible, back to a central data-lake. Within the data-lake, data is then organized, tagged, and stored. Delivery is where this innovative design method shines.

IoT platforms leave intact the BMS 'view' of the system from the perspective of the BMS provider but offer a highly flexible single-pane-of-glass view of the portfolio for fault detection and monitoring. The true power of an IoT platform is its standard form APIs offering to the consolidated dataset allowing refined analytics from hundreds of third parties (known as micro-service analytics/AI/BI providers), to analyze data sets for specific needs of the clients. Given this API capability, custom algorithms can be built easily in such regularly available toolkits as Google IoT, etc.

As our world faces increased levels of unpredictable disasters, pandemics, and extreme weather challenging our expanding critical facility operations, multi-site, geographically distributed critical facility operators integrating IoT platforms into their long-term corporate disaster recovery plans for remote management and operations can maintain business continuity, operational efficiency and ensure the health and safety of employees.



## Kathie Lyons

EVP and GM, Park Place Technologies

[www.parkplacetechnologies.com](http://www.parkplacetechnologies.com)



# Harnessing the power of AI

NO MATTER WHAT sector you are in, technological advances such as Artificial Intelligence (AI), Machine Learning (ML) and the Internet of Things (IoT) have revolutionised our world and thus, our perception of it. In an environment where the data deluge of the 21st century and 4th industrial revolution have provided us access to more data than ever imaginable, the role of the data centre, the hub of this valuable commodity, is more critical than ever.

With more data comes more responsibility. Those who guard the gates are responsible for ensuring that there is not a hint of downtime across the environment, otherwise huge losses are incurred by business. To quantify this risk, the average cost of an outage across businesses is \$260,000 an hour. Consequently, the upskilling of data operators is crucial in facilitating the correct handling and care of our most prized asset. With downtime resulting in colossal losses, the correct management of data centres is integral in ensuring business maintain profit margins. A degree of automation is essential here.

AI is being widely tipped as the solution to this; analysts at IDC have predicted that, by 2024, enterprises that are powered by AI will respond to customers, competitors, regulators, and partners 50% faster than those that are not using AI.

However, AI should not be solely looked at in terms of the response rate to the human entities that request information. The potential is much more than this. AI will be able to reduce

the margin for error, identify and solve potential issues before they've reared their heads. Gartner have additionally claimed that more than 30% of data centers that don't deploy AI and machine learning will not be operationally or economically viable in the very near future. Hence, the unity of AI and data operators is our next step, to minimise the risk of downtime, yet also, leverage the productive capacity for AI within this field.

Considering the average number of touchpoints in addressing data centre issues is eight, this process is both time consuming and rigorous. Therefore, the delay in finding potential faults results in business-critical down time and reputational crisis.

By harnessing the power of AI, it is now possible to predict where faults are about to arise, security measures compromised or even whether systems require patching. AI is being leveraged to instantly create a ticket once a fault is discovered, perform triage and schedule maintenance based on previous interactions with a similar fault, reducing the number of steps from eight to two and in some cases leading to a 97% first time fix rate.

While machine learning and AI can deal with the heavy lifting of analysing vast amounts of data quickly, it needs the human element to recognise nuances and sentiments that drives the value for the business. The future of AI creates the centrepiece of business analytics and monitoring, maintenance and digital transformation objectives. In pursuit of evolving with our times and maximising our capacity both artificially and naturally.

## Lee James

CTO, EMEA, Rackspace Technology

[www.rackspace.com](http://www.rackspace.com)



# How a job in antiquing has shaped the approach to customer experience

TO DESCRIBE AI as the future of data centre operations feels slightly like stating the obvious given the vital role the technology is already playing across the board. However, while the term and concept are ubiquitous, wide uptake of AI in IT management is still in progress – indeed, Gartner predicts that by the end of this year, approximately 50 per cent of enterprises will use AI for IT Operations (AIOps) technologies together with application performance management (APM), up from 10 per cent in late 2018.

The revolution, therefore, is well under way, though, with as many as half of organisations still not expected to have embraced AI to support their IT operations as 2021 begins, there is still significant room for growth. And as IT requirements get more complex, AIOps will become the only effective way for IT management in the future.

In today's world, we are always on and always connected – and this goes for businesses too. This obviously has a knock-on effect on customer expectations, not least given the ease with which opinion can now be voiced across a variety of platforms, including social media.

Acting immediately is the only real way to mitigate this – for teams to be able to understand what is happening across different services and quickly gather and analyse feedback from a number of platforms. This is where AIOps is becoming so valuable. It automatically detects, diagnoses and, in many cases, remediates service issues, all in real time, meaning far less

visible disruption of service from a customer standpoint. It can also draw upon historical knowledge to find and fix future issues quickly and accurately, again protecting businesses' revenues and reputations.

### Combining artificial and human intelligence

When we talk about AI in these terms, we should stress how far away it is from the Hollywood-fuelled headline-grabbing 'robots taking our jobs' narratives that dominate our popular culture. The philosophy of AIOps is to enhance IT operations through machine learning, analytics and big data: using multi-layered technology platforms to make operations smarter and to free up resources.

Both now and in the future, as its capabilities continue to evolve and expand, AIOps can play a pivotal role for businesses in the faster management of IT issues. Ultimately, managing cloud services is not straightforward. Things can and do go wrong, and they require time and resources to resolve – but principally they require people. However, while these people are juggling the necessary requirements to resolve whatever has gone wrong, the end product or service is not being delivered, running the risk of reputational damage and ultimately a loss of business.

With widespread adoption already well under way, AIOps will transform IT management and allow IT teams to deliver real benefits back to the business in terms of revenue, reputation management, and customer happiness.



## Mark Acton

Business Strategy and Technology Director, RiT  
Tech (Intelligence Solutions)

[www.rিতেch.com](http://www.rিতেch.com)



# It's about automated infrastructure management

Data Centres operators face challenges as the sector matures and new deployment methodologies are being adopted. The issues which need to be addressed by data centre owners and operators are driven by the following topics:

- Complex hybrid environments
- Edge adoption and remote 'dark' sites
- Increasing number of data sources
- Requirements for true integration
- Operational efficiency and cost reduction (Genuine ROI)
- Uptime, availability and time to service
- Ambition for automation, analytics, ML and ultimately AI

These challenges are leading to changes in approach, especially the tools which customers need in order to manage their day to day data centre activities. A key requirement is for a tool that provides a single source of truth. This is not simply a single pane of glass or summary dashboard. It includes full integration with business systems and processes with the sharing of data across multiple disparate systems. All assets in the business becoming accurately tracked, reported upon and viewed in the same consistent way by all departments sharing a single federated database.

As data is collected and stored in complex multi-layered and disparate systems gaining true value from this wealth of information requires new toolsets. Increased integration and analytical techniques are needed to take advantage of this key data. This data-centric approach paves the way for true automation, machine learning and ultimately AI. Automated provisioning is possible with existing levels of real-time data from facility infrastructure to IT devices and even applications. Full stack integration is available and is the key to both cost and

operational efficiencies for direct and tangible ROI.

Aligning IT, Facilities, and other business stakeholders produces competitive advantages offering the transparency necessary to manage hybrid digital architectures. These will include public and private Cloud deployments, legacy infrastructure and systems, as well as remote unmanned Edge installations. A single source of truth requires an open framework and integration ready approach that allows rapid deployment and increased speed to use. This integrated approach should not aim to be all things to all men, but rather look to seamlessly hook into existing systems, to take advantage of their unique capabilities and data.

Historically this level of integrations has been technically impossible or too financially challenging, but the shift towards distributed digital infrastructures and hybrid environments necessitates a simplified orchestration of complex digital infrastructures across multiples sites. It is widely accepted that the IT stack should be integrated into any management system as well as essential business systems. It is also recognised that Facilities, IT and 'the business' do not currently work well together. Lack of integration of information and tools across those disciplines remains a barrier to both cost and resource efficiency. Tools with the attributes described in this article are a catalyst in getting all stakeholders to work together around core business processes involving people, facilities and IT infrastructure.

RiT Tech's pioneering R&D Center in Israel is the birthplace of Automated Infrastructure Management (AIM), which is now integrated into ISO/IEC 18598. RiT's data-centric approach and big data solutions are now deployed in data centres worldwide.



**Herman Chan**  
President, Sunbird Software  
<https://www.sunbirddcim.com>



## Time for smarter capacity planning

ACCURATELY PLANNING and managing rack power capacity is a struggle for many modern data center managers, but the repercussions of failing to do so are stranded rack power and space capacity and overspending on operating expenses. Fortunately, new machine learning and AI capabilities in second-generation Data Center Infrastructure Management (DCIM) software now allow for fast, easy, and accurate power capacity planning.

Most data center managers run into the same challenges in their attempts to plan and manage power capacity. It is expensive to build out a server-ready cabinet, costing around \$15,000 per cabinet. There is immense pressure to properly monitor and understand power capacity and infrastructure health or risk costly downtime. And often, available capacity is taken for granted as overbuilding of data center capacity is a rampant and wasteful practice that leads to excessive stranded capacity.

The traditional method of capacity planning has failed data center managers. This commonplace method, where data center managers derate the nameplate value of a server device to around 60% or 70% when budgeting their power to determine how many servers can be deployed in a cabinet, is manually intensive, time-consuming, inaccurate due to being based on assumptions rather than data, and costly due to the high risk of stranding capacity and overspending on capital expenditures to add capacity.

However, with recent advances in AI and machine learning technologies now available in second-generation DCIM, the traditional method is no longer necessary. A complete solution of Sunbird's DCIM paired with intelligent rack PDUs

brings automation to data center power capacity planning and management, making it fast, easy, and accurate.

Sunbird's patent-pending Auto Power Budget feature unlocks precious rack capacity as each individual server make/model instance is assigned a unique power budget value that's automatically set from real-time outlet-measured power readings from intelligent rack PDUs – your devices' exact loads under your compute stress in your environment. This dramatically refines the traditional method of taking a percent of nameplate and improves the utilization of existing rack power and space.

Where the traditional derated budget method demands high effort and manual calculations, Auto Power Budget takes minimal effort and updates automatically. Where the traditional method provides low accuracy due to applying the same derating to all servers, Auto Power Budget provides high accuracy by using measured readings over long periods of time in its algorithms. And where the traditional approach presents high risk due to error-prone manual calculations and assumptions, Auto Power Budget's built-in algorithms and validations intelligently manage risk.

Data center managers must better leverage their data to make smarter capacity planning decisions, or risk stranding capacity and overspending. Being able to automate this process is a key differentiator between a successful data center manager and one who struggles, and now it's never been easier. Auto Power Budget takes the manual effort, inaccuracy, and risk out of power capacity planning, making it easy to deploy more compute, leveraging power in existing racks. It saves time and money, reduces risk, and eliminates stranded capacity.



# DIGITALISATION WORLD

The most comprehensive, leading information source covering  
the key technologies that underpin the digital revolution

[digitalisationworld.com](http://digitalisationworld.com)







**David Cumberworth**  
Managing Director, EMEA and APAC, Virtana  
[www.virtana.com](http://www.virtana.com)



## Monitoring and optimizing the new hybrid datacenter

IN A RECENT FORRESTER SURVEY forecasting 2020 datacenter locations around ¾ of all respondents said that hybrid cloud will be their new model. Virtana also conducted a first half of 2020 survey of enterprise customers in EMEA and US and came to the same conclusion with 74% saying Hybrid is the new strategy.

Hybrid cloud means that your infrastructure can be hosted anywhere. How do you maintain performance Service Level Agreements (SLAs) in such a diverse environment? The important solution is to centralize the monitoring, so you are getting performance-based feeds from all elements into a 'single source of truth'. It is tempting to think you can manage the new hybrid environment using your traditional tools – but you can't. Hybrid needs a hybrid management approach, not single point solutions.

For the past two years, Virtana has been working with major enterprises around the world, helping them migrate their application workloads to new environments. Understanding workload dependencies, by mapping the topology they use, is a vital starting point. Many migrations have failed to live up to their promise and it is not uncommon to find that on-premises and hosted private cloud outperforming public cloud for cost and performance.

To help you build a successful migration strategy, we have created this list of 12 topics that should first be addressed:



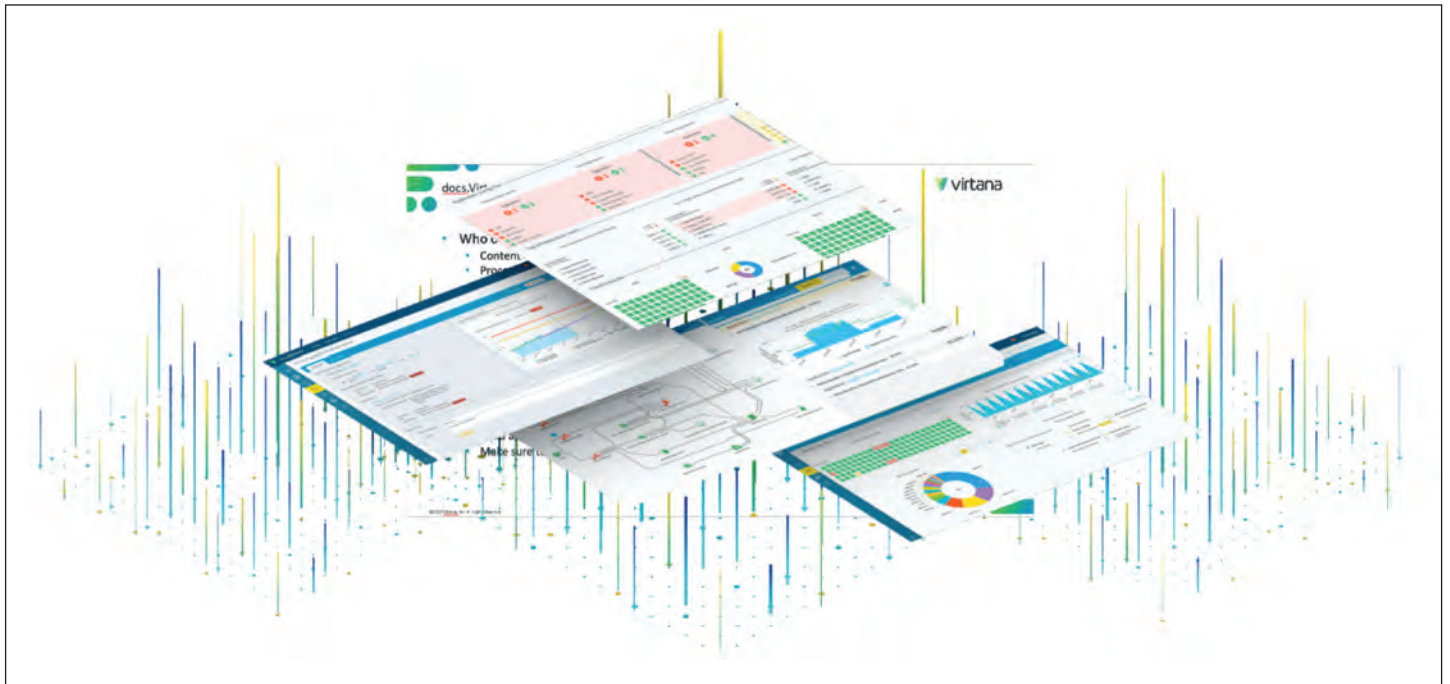
### Visualize your hybrid (or cloud) environment

In your cloud environment (hybrid or all-in-cloud) you'll need tools that allow visibility into each application from the compute stack through networking and into storage, regardless of location. Tools with automation for these capabilities are a big advantage.

### Discovering and monitoring your hybrid/cloud application infrastructure

Variability is a key characteristic of cloud environments and cloud-based infrastructure offerings such as Containers as a Service (CaaS). Monitoring and management solutions should





have a strong discovery capability for applications that run within the environment. They should also have automated dashboards for elements and metrics for alarms and capacity trending.

### The role of AI and analytics

In our recent survey, over 75% of respondents believe machine learning and AI will be needed to support their remote-working requirements. AI and analytics should permeate your solutions for monitoring and managing cloud and hybrid applications. Automatically flagging anomalies, suppressing alarm storms, performing proactive root-cause analysis, optimizing workloads, and applying intelligence to capacity management for individual elements and full application infrastructure stacks.

### Application performance

Your migrated application will have the same needs as before. Application Performance Monitoring solutions and user experience monitoring should be augmented with monitoring and management of the infrastructure that underlies the application – regardless of where that infrastructure is located.

### Problem resolution

Use application-centric views of your cloud and hybrid applications combined with AIOps-level investigations into alerts to resolve problems quickly and drive uptime.

### Performance management

Track performance trends and alerts to identify anomalies and bottlenecks before they affect application service levels and SLAs.

### Capacity management

Apply capacity monitoring by cloud element and full application stacks for compute, network and storage, and hybrid on-premises elements. Capacity management should be auto-applied on discovery and auto-alerted when needed.

### Cloud cost management for IaaS

Make sure to find a solution to the five key segments of cloud cost management – Analyzing billed usage versus resources used, identifying and eliminating idle resource, monitor for cost changes, right-size instances and reserved resources, and consider reservations for instances when you have baseline data indicating that you'll need additional capacity. Visibility across clouds and beyond is a plus.



### Network management and monitoring

Visibility into your cloud network elements, combined with your on-premises network if using a hybrid deployment model, should allow you both to understand how your applications are using cloud network resources and what settings and changes are needed to maintain performance as workloads scale.

### Automate

Apply automation using cloud-provider offerings such as AWS Elastic Load Balancing, and third-party tools that automate capacity management and workload optimization where needed to increase scalability and to prevent unplanned downtime.

requirements via real-time performance data. Although the major hosts provide cost optimization tools, these are not very accurate. Analysis of billing and how it matches capacity over time as well as in real time is what is needed. Armed with this information you can plan capacity purchases and discover wasted spend. By using a single platform for cloud management, you can monitor your infrastructure, plan capacity, and eliminate performance risks. Performance bottlenecks can be predicted before they affect clients and SLAs with multi-conditional alerting powered by advanced anomaly detection. Cloud solutions are not only publicly provided by the likes of AWS and Azure. Co-location is also a strong option where your

In order to optimize cloud resources, a third-party AIOps-based resource is needed. This will provide an independent and granular view of how applications are using capacity and if it is right-sized.

Integrating your environment with your ITSM/Service Desk. If your organization uses an IT Service Management solution such as ServiceNow or Remedy, you should link your on-premises and cloud change, monitoring, and management processes to your service desk.

### Managing application infrastructure growth

Set clear limits on cloud services' use of scalable resources and monitor for capacity well before limits are reached. Rapid escalation in resource requirements can indicate problems in application and infrastructure deployments and automation that require investigation and resolution. Use cost and performance management for cloud environments to watch trends and usage. If storage costs escalate because of swift growth in your cloud data sets, consider repatriating storage to an on-site location to control costs. In order to optimize cloud resources, a third-party AIOps-based resource is needed. This will provide an independent and granular view of how applications are using capacity and if it is right-sized. In addition, it will monitor performance of the applications in real time and provide metrics and analytics to eliminate bottlenecks. The allocated capacity can also be monitored to ensure an accurate match to workload

applications are managed on your behalf by a system integrator. This is increasingly becoming a stronger option for more business-critical applications. But to determine which is best for you, you need to start with the facts. Virtana is offering a 60-day cloud right-sizing assessment to organizations needing to better support their infrastructure as demands increase – and our cloud experts will gladly walk you through your options.

To ensure your migration is going to go well Virtana has created a Cloud Migration Readiness (CMR) service. The service looks at your entire environment and benchmarks application workflow through it. This is often the first time organizations have seen their entire infrastructure in one place. They are now ready to plan migration based on known performance, capacities, and costs. So, what is your Cloud Migration Readiness level? A good starting point is to take a free cloud migration readiness assessment. Whether you are exploring the cloud or have a current cloud initiative, find out your level of CMR.

To receive recommendations and get access to the Virtana definitive CMR white paper, visit:

<https://www.virtana.com/go-cloud-assessment>

## Niclas Sanfridsson

CEO Colt DCS

[www.coltdatacentres.net](http://www.coltdatacentres.net)



## AI will help pave the way

TODAY, we are more connected than ever before. From basic communications to mission-critical IT functions, data centres are at the forefront of ensuring smooth and efficient connectivity. The recent explosion in cloud computing and Internet of Things (IoT) uptake means that now more than ever, data centre operators play a critical role in ensuring their facilities are reliable and scalable. With ever-increasing customer demands and expectations, even the smallest degree of downtime can have disastrous consequences for both businesses and operators. This means that we must always look for new and innovative ways to manage and resolve issues, before they interrupt service.

Yet, there are some data centre providers that are silently struggling to provide reliable connectivity for businesses that require low latency and data intensive services. Often this issue is due to infrastructure not being up to scratch when it comes to today's level of data consumption.

In order to combat this issue, AI is increasingly being used by operators to streamline and enhance facility maintenance and management practices as a means to improve reliability, boost energy efficiency and meet the most demanding of service requirements.

As with dated infrastructure, some data centres may also retain dated practices from a bygone era. Traditionally, data centre operations teams are on perpetual high alert for faults in facility infrastructure. An overheating server from a malfunctioning cooling unit, for example, can have very serious consequences for customers if the problem isn't immediately solved. Since no-one can tell precisely when a component will fail, this requires

data centre engineers on constant standby, either in the data centre itself or on-call. As IT networks become larger and more data intensive due to the uptake of technologies such as the Internet of Things (IoT), data centres are growing in scale to match demand. As data centres enter the hyperscale space, these practices become unfeasible due to the sheer size of the facility and number of components.

For that reason, operators of large-scale data centres must employ sensors and AI to help staff monitor their vast and complex network infrastructure. With the increased visibility afforded by sensors, AI can monitor an entire data centre for anomalies 24/7, this includes physical and virtual infrastructure. If a sensor detects that an area of infrastructure may be at risk of failure, an alert will be sent to the operator. This gives engineers valuable time to investigate the alert before any failure can occur. If there is an issue, it can then be solved proactively minimising disruption or downtime.

In an industry where downtime is money lost, AI and machine-learning have a massive role to play in bringing down the time taken to identify and resolve infrastructure issues. It is a vital tool for identifying faults and preventing issues from compounding into service disruption. AI also helps centres to operate as efficiently as possible. This allows operators to meet challenging service requirements and bring down associated costs.

For customers, this means better value for money with more reliable services. As AI continues to be improved, we expect to see more benefits arise from AI adoption in data centre operations. When it comes to building bigger and faster data centres, AI will help pave the way.



## Nick Ewing

Managing Director, EfficiencyIT

[www.efficiencyit.com](http://www.efficiencyit.com)



# Remote monitoring - an essential tool

TODAY REMOTE MONITORING software can mitigate a number of the pain points caused by the lack of physical access to a critical IT system. By utilising next-generation Data Centre Infrastructure Management (DCIM) solutions, IT managers can gain unprecedented insight into the status of their digital infrastructure. With such visibility, organisations now have a proactive means to identify and rectify issues in real-time.

Should a team find themselves resource constrained or in need of specialist support from an external service partner, these DCIM solutions offer the ability to integrate with third-party management systems, connecting any IoT-enabled solution from any manufacturer, to ensure customer IT systems stay secure, connected and online.

With use of data analytics and machine learning algorithms, operators can minimise risk of outage through IT equipment failures. Working with such software allows them not only to work more efficiently, but provides a wealth of new opportunities for reduced costs and more efficient systems.

This recently included The Wellcome Sanger Institute, one of the world's leading research organisations focused on genomic discovery. Based at the Wellcome Genome Campus, Cambridge and established in 1993, the Institute conducts key research into improving the outcomes of human health using data derived from genomic sequencing.

At the core of Sanger's technical infrastructure are its fleet of DNA sequencing machines. This infrastructure is essential to the research operation and any downtime would present a significant loss in terms of both data and costs.

To ensure uptime each machine is protected by an Uninterruptible Power Supply (UPS) but with so many machines to protect across such a distributed campus, one of the key drivers for the project was to increase visibility and streamline operational efficiency.

To overcome the challenges presented, Sanger worked with EfficiencyIT to deploy Schneider Electric's EcoStruxure IT DCIM software. The system would connect all of the campus' infrastructure, including its on-premise data centre, to ensure the data-driven environment would be efficiently monitored and that crucial scientific research would continue without interruption.

By connecting the assets to the software and its cyber-secure smartphone application, the Institute instantly gained greater insight into the day to day operations of the IT estate on campus, allowing any issues regarding power, cooling and IT, to be rectified quickly from a remote location. In addition to providing greater insight from a single management console it also began to benchmark improvements in the data centre, ensuring that, over time, it would justify a significant Return on Investment (ROI). Today the predominant payback for the Institute is being seen through reduced operating costs, especially those of energy consumption, with 5-10 per cent expected to be saved in the first two years of deployment.

With the world of work changing around us constantly, remote monitoring has become an essential tool to IT Operators, offering a simple, fast and AI-enabled solution that delivers business continuity and bridges the gap between physical infrastructure and remote staff.

## Yair Green

CTO, GlobalDots

[www.globaldots.com](http://www.globaldots.com)



# AI becomes more and more important

AS WE SEE AI finding its way into more and more industries, it would be fair to say that many businesses are struggling to find their way with this latest technology.

What are the top challenges when your product is built on, and depends on, AI? Most of the common problems relate to people, data or business alignment, and whilst different organisations will experience the process of AI adoption in their own unique ways, there are certainly some common challenges and considerations regarding AI implementation.

**Running AI:** When you combine the advanced technology of ML systems with the deployment capabilities of containers, you are able to make machine learning systems much more useful and shareable. This ability to deploy ML applications as containers and to cluster those containers offers a number of advantages. For example, the ability to make machine learning applications self-contained. In the sense that they exist in containers, they can operate in a highly distributed environment, and so you can place those containers close to the data that the applications are analysing.

**Monitoring:** With the AI and Machine Learning (ML) explosion, it's possible for some organisations to have over a hundred models running in production every day. On this basis, making sure that everything is working well presents a huge challenge - one that many organisations are failing to address correctly. In recent years, most of the focus and open tools were focussed on exploration and developing the model stages, and also on

deployment - monitoring and tracking were seemingly neglected. So why does monitoring matter? With standard applications, it is important to keep monitoring for errors, latency and crashes etc. But when we consider the case of an AI system, these factors are not necessarily the main issue. Most importantly, organisations need to ensure that their system keeps doing what it was built for and doing so correctly and efficiently.

**Automation:** With the adoption of ML technologies continuing to rise we see a proliferation in software and tools that can potentially increase a business's efficiency and value proposition. They can help to automate the repetitive and mundane tasks within the ML or data science processes without compromising model performance and productivity. Such examples could be a platform that automatically builds predictive models that forecast the future behavior of your clients. It provides full automation of essential, yet time-consuming activities in predictive model construction.

**Retraining:** ML models deal with streams of information and they need to be updated when they are fed new data that is significantly different from the data used to train them. This is why any Machine Learning tools that you use need to be able to provide both accurate, and scalable, repeatable and ready-to-go solutions that can be easily integrated into any information system. In other words, a Machine Learning tool shouldn't necessarily look for the perfect unchangeable model, but rather for flexible solutions that can be easily retrained and brought to production in the quickest time in a shifting environment.

## Tim Doiron

Senior Director, Solution Marketing, Infinera

[www.infinera.com](http://www.infinera.com)



# How far can we take AI and ML?

“WHETHER it’s inside the data center or in the external transport network that interconnects data centers, operators are looking for ways to drive down costs, increase networking efficiency and elevate the performance of their people. While the data center and network application potential for AI, Machine Learning (ML) and data analytics is really exciting and broad, we are still fairly early in terms of broad commercial deployment of AI/ML for closed-loop automation.

The good news though is that we are on the way. In the area of security, AI/ML technology is being used to sort through a sea of networking data to identify potential threats and mitigate them rapidly - before they can cause harm. Power utilization remains one of biggest costs for data center operators.

As pioneered by Google a few years ago, an AI/ML engine can analyze internal and external parameters like the outdoor weather to dynamically and continuously adjust the cooling

system to minimize power consumption inside the data center. And in the network, we see the early application of AI/ML for network health prediction to anticipate faults and failures before they happen so that corrective action can be taken to prevent service disruptions. These three tasks could all be done without the benefit of AI/ML, but each would generally take longer, be less accurate and require more people.

This is where AI/ML and humans come together for the common good. AI/ML is good at looking at vast amounts of historical and ongoing data to establish a baseline and recognize patterns – with more time, AI/ML becomes even more precise. Humans are good at understanding a broader situational and environmental context and reacting to something completely new like a natural disaster. So how far can we take AI/ML? Over time, AI/ML will be applied to nearly every aspect of data center and network operations. However, each application will require time and trust before humans fully take their hands off the operational wheel.”

In the area of security, AI/ML technology is being used to sort through a sea of networking data to identify potential threats and mitigate them rapidly - before they can cause harm. Power utilization remains one of biggest costs for data center operators.



## Abdullah Saleh Alharbi

Data Centers Engineering Lead, Saudi Aramco

[www.saudiaramco.com](http://www.saudiaramco.com)



# The new reality, data center facilities digitisation

AS ENTERPRISES at large continue to digitally transform to gain more productivity, efficiency, and create new growth opportunities, data centers are the cornerstone of this transformation. Currently, a huge forced-surge was noticed – during the COVID-19 pandemic – in terms of enterprises who moved to digital capabilities to ensure their business continuity. Such events will accelerate the digital transformation and unleash new demand patterns. As a rule of thumb of the current digital economy, data centers will be the industry that will encounter spikes during any regional or global events, since the need for digital communication will rise accordingly.

Data centers will have to ensure the following: proper positioning to the edge, the right combination of infrastructure for agility, and facilities' digitization not become a bottleneck, with the ability to do more with less resources. The complexity of running digital facilities will boost AI, automation, and robotics development, to augment the human element and supplement any current shortages. The level of digitalization in data center facilities, especially on-premises, will normally match the enterprise digital maturity.

The use cases for disruptive technologies in data centers are massive. Planning on where to start digitizing is a critical task. The data center operator shall evaluate digitizing all resource-hungry activities and must be in line with their regulations. For example, the automated operation that all data centers have today can be deepened with the use of AI and machine learning, to maximize the energy efficiency, minimize equipment defects,

and predict future events. Additionally, all routine processes, such as physical security, will be automated with the support of facial recognition and online authorization, to minimize operation expenses, while maintaining the right level of restrictions.

Remote working for data center employees is a solution during the COVID-19 curfew period, and is expected to continue. This change will lead to centralized management of scattered data center facilities, and expansion of remote technical support for operation and maintenance activities, with the support of video conferencing and wearables.

Furthermore, face-to-face staff and vendor meetings will be reduced to maximize operational efficiency. This effect will change different concepts in the near future to remote-in-nature support, such as troubleshooting, testing (factory and site), and commissioning activities, which can be performed via live video conferencing, and recorded for quality assurance purposes.

In data center operations, it is always critical to have the right level of available staff and spare parts to ensure business continuity. Automation of spare parts management through AI and robotics will ensure accurate movement and stock level at all times, saving man-hours and the cost of management.

Thanks to technology and innovative methods that can solve many of today's problems and challenges, the sector has a number of opportunities to achieve higher reliability, efficiency, and agility."

## Adam Scraba

Director of Product Marketing, Xilinx

[www.xilinx.com](http://www.xilinx.com)



# The next wave of computing

COMPUTING INFRASTRUCTURE is undergoing a massive shift as real-time services such as virtual personal assistants, and real-time security services continue to sweep the globe. For service providers to derive value from these solutions, data and insights need to be instantly accessible and will primarily be driven through artificial intelligence (AI) enabled services.

Consequently, a shift is taking place in the data centre: architectures are being built with the ability to address varied workloads and algorithms, and compute is becoming increasingly integrated into storage and networking.

To meet the needs of the modern data centre, service providers need an infrastructure platform that can offer a huge amount of differentiation and performance. This platform should deliver the throughput and low latency, along with incredibly flexible software and hardware stack that can handle wildly different algorithms, such as RNN/LSTMs, CNNs, and Spark query acceleration.

Indeed, the next wave of computing needs to be highly adaptive, where software and hardware merge, and where both hardware and software are programmable to achieve real-time performance, maximum throughput and low latency, and power efficiency.

### Compute acceleration

Today's modern teams are leveraging high-performance computing (HPC) as a tool for discovery and a path to solving some of the world's most impactful and complex problems. However, accelerating time to insight and deploying HPC at scale requires incredible amounts of raw compute capability,

energy efficiency, and adaptability. Fortunately, FPGAs have the unique ability to address these requirements. Applications accelerated by FPGAs enable optimised hardware and software implementation with the flexibility to adapt to changing requirements without sacrificing performance and energy efficiency.

### Networking

With the advent of virtualised computing and containerised workloads, networking has become far more complex. As these environments scale beyond a single server, they need to employ sophisticated overlay networks. Overlay networks are virtualised systems that are dynamically created and maintained using the concept of packet encapsulation.

This then adds burden on the OS or virtualisation kernel to supervise this encapsulation, when combined with traditional networking tasks together, these consume nearly 30% of a server's raw CPU cycles. One of the most commonly employed means of managing overlay networks is through the utilisation of the Open vSwitch (OvS) protocol.

FPGA-based SmartNICs have the computational capacity to offload the host CPU from the above mentioned 30% overhead. In simple terms, three servers with SmartNICs handling OvS have the computational power of four servers with standard NICs.

FPGA-based SmartNICs can also be leveraged to offload security and encryption tasks usually executed on the server CPU. Security comes in the form of deep packet inspection that then results in packets being dropped if they pose a threat. This could augment or even replace traditional firewall software



that enterprises run on their servers. Additionally, SmartNICs can easily offload various encryption and decryption tasks from IPsec to TLS.

### Storage

With the widespread adoption of high-speed storage and the increased performance requirement of data-intensive applications, traditional architectures have created CPU, memory, and storage bottlenecks. Understandably, the attention is shifting from compute to data through computational storage and its implications on improved application performance and overall infrastructure efficiency.

A viable solution is to move compute closer to the data. Samsung Electronics has adopted this approach with the launch

of SmartSSD® Computational Storage Drive (CSD), which enables high-performance accelerated computing closer to flash storage, bypassing CPU, and memory limitations. Samsung's approach increases speed and efficiency and lowers operating costs by pushing intelligence to where data lives.

### A new world order for real-time services

In the growing world of real-time services, CPUs – or multicore CPUs – simply don't meet the necessary cost, power consumption, and scale requirements. As Moore's law grinds to a halt, next-generation CPUs offer little in the way of hope. Consequently, to meet the needs of the modern data centre, the industry is already starting to turn its attention to adaptable compute accelerators as a feasible way to meet the broad compute demand and manage scale-out and operating costs.



The digital infrastructure information platform



[datacentre.solutions](https://datacentre.solutions)

IN ASSOCIATION WITH





# COLOCATION + HOSTING

In-house or to outsource? Colocation and hosting facilities offer a range of significant advantages when it comes to accessing data centre infrastructure. For many end users, these advantages are far too compelling to ignore: the very latest infrastructure, constantly updated, constantly supervised, with constant access to expert advice and support for a fixed monthly fee. Put it another way, the colocation and hosting model gives end users access to high capacity and high quality data centre resources which would be simply unaffordable if they had to build these themselves.



IN ASSOCIATION WITH



**DATA CENTRE**  
**SOLUTIONS**



**DIGITALISATION**  
**WORLD**

# A modern criteria for choosing a site for your new data centre



No longer limited by earlier technologies, many data centres can be located in distant locations where reliable services are still accessible, this reduces many of the cost associated with running a data centre. As a result of this freedom, measures should be taken to optimise the location selection process.

**BY DAVID CAPPuccio, DISTINGUISHED VP ANALYST, GARTNER**

A BEST PRACTICE for data center owners and operators is to work through an internal process that establishes the most critical selection criteria, and then assigns weights or degrees of importance to each criterion. The following four contemporary factors must be considered in a modern site selection process.

## HR issues

When looking at HR in the context of choosing a location for the data centre, leaders should be mostly focused on the workforce profile in the area. It is a very important criterion to consider if the data centre will be fully staffed because operators will then need to question the skills available in the local market, prevailing labour rates and commuting distance.

As the number of people needed to run the data centre decreases, its weighting in our decision making also falls in comparison to other concerns such as proximity to suppliers and building infrastructure. It also eventually becomes a nonfactor if you have plans to run a “lights out” data centre.

## Incentives are important

Many regions offer economic incentives to attract a “high end” operation, such as a data centre, to their communities. This is particularly true if a significant head count is residing at the data centre. Incentives might include tax abatements, energy credits, investment credits and grants.

As part of the site search, IT leaders should contact the local economic development agency in the target area to determine what incentives might apply.

## Environmental concerns

Considering the environment presents a relatively new set of criteria for site selection, because of the increased emphasis on greenhouse gas emissions and other environmental concerns. Factors such as applicability for the use of free cooling technologies –

located where the use of renewable energy solutions (such as wind and sun) is possible – and local utility companies’ use of renewable power are examples of important considerations.

Additionally, noise abatement may be an issue in some regions, when external generators are test run periodically as part of ongoing maintenance cycles. This might require either special dispensation, or enclosing the generator area to reduce external noise.

## Public service access

This is usually a criterion that is often overlooked. It concerns the degree to which our targeted site is able to be serviced (quickly) by first responders, such as fire, police and ambulance services. This could be an issue if the site is highly remote or rural, where normal public services are sparse and not readily accessible.

With fully staffed data centres this is an important factor to be considered for both the health and safety of the workforce and a protection against litigation.







## Cindy Ryborz

Marketing Manager DC EMEA, Corning OC

[www.corning.com](http://www.corning.com)



## As easy as MTDC?

DECISIONS on the optimum environment for IT applications and infrastructure can be a complicated business. Choices on data storage range from in-house, colocation, cloud, managed services or, in many cases, a combination of them all. This choice has become heightened in recent weeks, with businesses ever more reliant on data and connectivity in the midst of the COVID-19 crisis.

The impact of COVID-19 has meant an immediate pressure on data centres to facilitate rapid and seamless transition of data. However, organisations have long required that vital data is maintained in an environment that is robust in the face of major, unpredicted events. Bandwidth, access to business tools and possible network constraints are all factors that need to be considered. COVID-19 is testing organisations' data decisions in the face of catastrophe.

Multitenant data centres (MTDC) offer a viable solution in these instances, particularly for small and medium sized businesses that don't necessarily have resources to navigate the challenges of a crisis in-house. Businesses are facing huge volumes of VPN connections due to the shift to working from home. Security challenges, and changes in regular network usage patterns, are placing further load on networks.

Many businesses see an improvement to their IT team's capacity and ability to support the business when they outsource their data centre operations. Any IT team managing its own data centre must constantly be focused on maintaining, upgrading, and ensuring the data centre is working at an optimum level at all times. There must also be the capability to expedite work when demand spikes, along with addressing downtime and data loss.

Through outsourcing to an MTDC, businesses are provided with a connection to network service providers/operators and cloud providers without the need to buy new equipment or train additional IT teams to implement and maintain solutions.

Scalability is key. In the short-term, particularly in times of crisis, businesses need to be agile. An outsourced service is often a good option to achieve this. More generally, scalability within the data centre will support simple migration to higher-speed technologies and applications — a vital factor for meeting future bandwidth and transmission speeds that the customer demands. As data is continually generated and technological advancements continue to increase, the need for organisations to meet customer demands for zero latency and real-time access to data will as well.

More specifically within the COVID-19 crisis, where many industries and businesses are under pressure to shut premises and reduce on-site working, staff operating colocation centres in many instances have been given more imperative to continue their work. In many markets, data centre operators have been established as essential workers, and therefore enabled by their governments to continue in normal working environments.

For business operators both big and small, taking advantage of the capabilities of massive data is no easy task and this is amplified even further during a crisis. Migrating computing needs to an MTDC both tackles many of these challenges, while offering scalability, flexibility, modularity, and stringent SLAs to assure maximum uptime, peace of mind, and rapid deployments to high-density applications when you need it, fast.



**Richard Blanford**  
Chief Executive, Fordway  
[www.fordway.com](http://www.fordway.com)



## Why cloud may not be right for you just yet

Moving services to cloud makes complete business sense in many cases. A suitable SaaS service, which replaces an in-house (a.k.a. legacy) application and provides the capability you need, is easy to use and has comparable costs to in-house provision is a no brainer – so do it. However, that doesn't mean you should rush to a 'cloud for everything' policy.

First, any major change needs to be driven by a compelling event. This could be a move to new premises, a need to refresh existing IT infrastructure, the end of an outsourcing contract or a major organisational change. If you've recently upgraded your data centre, wholesale migration to cloud is probably not appropriate at present.

Second, most of us are not starting from scratch or working with a small number of relatively straightforward applications – basically we've all got legacy. We need to consider carefully which applications can be transferred efficiently and cost-effectively to cloud. Like any other major IT transformation, moving to cloud requires strategic planning to ensure it can be managed alongside business as usual and without unpleasant surprises or unforeseen costs. So, in my view, the future will be hybrid for a long time to come.

Your first step should be to look for a SaaS service and check that it offers the right resilience at the right cost, with appropriate contractual terms from a supplier you are confident will be there to assist and support you for at least the next few years. You should simply be able to consume it and be billed for what you use, while monitoring application and supplier performance to ensure you receive what you've signed up for. However, before you sign up, consider how it will play with your other

key applications. For example, once your Microsoft Office applications have been moved to Office 365, they will be patched and upgraded automatically. This means that if you have a legacy application which only integrates with an older version of Office or MS Exchange, at some point it will no longer talk to your Office applications. You need to understand potential interactions, which means mapping applications for the next two years and asking vendors for their roadmaps.

Many application providers are developing their own SaaS strategy, but these typically only support the latest version of software, and many can't accept customised applications or third-party add-ons. We've investigated many 'SaaS' offers for customers, and all too often the vendor will simply host and maintain a dedicated version of the user's software on a public cloud service while charging a premium price.

If SaaS isn't available, you could choose IaaS, where you move your application, as-is or with minor enhancements, to operate from a cloud provider's infrastructure. This frees you from the need to own, operate and manage the infrastructure hosting it. However, you need to retain licences and support from the application provider, and are still responsible for patching, resilience, back-up, security and application support and maintenance inside the instance – which is a full-time job.

Alternatively you could use managed IaaS, which effectively offers your legacy application back to you as SaaS. If SaaS isn't available and IaaS sounds like too much work, configure your internal IT as a private cloud or look for a specialist provider to host it for you. You can then continue to run it as is, and be ready to move it to SaaS when a suitable solution becomes available.

## Brian Roe

Commercial Director, Servecentric Ltd

[www.servecentric.com](http://www.servecentric.com)



# Infrastructure strategies – consider all your options

THE DEBATE CONTINUES regarding the best deployment scenarios for enterprise workloads. Philosophies range from cloud first/cloud only strategies to retaining local dedicated infrastructure housed within the enterprise's four walls – and everything in between.

But how should the enterprise approach the problem? What are the metrics for success? How does the CIO / CTO best blend performance, resilience, security, accessibility and support to deliver bang for their buck?

### Considerations

The cloud market has matured greatly in the past 5-6 years. Improvements in end user connectivity options and continuous introduction of business enabling PaaS services has made migration simpler. However, as the land grab by the big three continues, end users and the reseller market are finding themselves under pressure to migrate all services to cloud platforms, and sometimes, this is simply the wrong option.

I would suggest that the traditional “DC managed service” (where dedicated tin is commissioned for client use as an opex service) is now losing its appeal as an alternative.

However, the use of dedicated client owned resources in colocation or locally can absolutely be the right choice for businesses. It simply makes good business sense in so many cases.

I have seen, first-hand, organisations that have cancelled cloud migration strategies or migrated partially or wholly from hyperscale cloud back to the data centre as costs escalate

beyond anything initially projected. In truth, by design in my opinion, it is almost impossible to correctly cost some public cloud implementations. Additionally, the projected savings relating to maintenance, staff overhead, security etc. are rarely realised. For example, if you consider an always-on, compute and storage intensive application that runs 24 x 7, owning and running the infrastructure will almost certainly deliver lower overall TCOs versus hyperscale cloud equivalents.

That said, cloud is absolutely the right way forward in many or most other cases. When loads are variable, in the case of test/dev environments, when IT support resources are not available or where an organisation does not, for whatever reason, want to commit to capital outlay, cloud can be an exceptionally effective option due to its pay-as-you-go nature, straightforward set up and ready to go self-service options.

Also, for most organisations, looking beyond hyperscale to other IaaS providers can provide real alternatives that provide cost clarity and savings, improved support without compromising on service quality. They may not scale like hyperscale, but do most users really need this reach?

### Summary

We cannot underestimate the value in considering all options when creating infrastructure strategies. It is absolutely key not to be vendor-led in this, but to view all workloads from within and understand what is the best fit for each of these, without compromising on resilience or security.

For enterprise infrastructure requirements, this invariably will result in hybrid builds of colocation and cloud.

## Noel O'Grady

Sales Director, Ireland, Sungard AS

[www.sungardas.com/en-gb](http://www.sungardas.com/en-gb)



# Improving disaster recovery capabilities

Many businesses today buy into (hybrid) cloud solutions to solve cost issues, but it's simply not an option for certain workloads and data. This is driving the increase in datacentre co-location, which enables businesses to house their own private datacentres in dedicated facilities where everything from security and power to temperature and moisture control is fully managed. The approach also allows companies to access expertise and services in the most cost-effective manner and leverages the knowledge and support of a colocation provider. It also enables organisations to focus on their core business operations while being confident that new technologies can be quickly implemented as they become available.

Why is colocation as part of a hybrid-cloud strategy a better alternative to developing and delivering new technologies and services? There are some key benefits.

Firstly, for most businesses, managing an on-premise data centre is not a key component to success. As a cost centre, incorporating new technologies and business continuity is a necessary evil required to deliver core services or applications. Colocation can reduce the overall total cost of ownership (TCO) for an environment through a provider's scale where services are developed to be consumed in increments aligned to a business strategy.

Secondly, data centres are also expensive to operate and manage. New technologies have a large "up front" commitment in both human and cash capital. A DIY approach to deploying new technologies often has long evaluation, design and implementation runways which hinder an organisation's ability to keep up and get ahead of competitors. Costs include the

upfront spend on physical infrastructure, and the ongoing cost of full-time employees managing the facility and troubleshooting. There are also energy costs and network connectivity to consider alongside rent, taxes and other expenses that an organisation may be responsible for.

With colocation, capital expenses are turned into operating expenses. By renting the space in a third-party data centre, the building, infrastructure and maintenance is handed to a provider for a predictable – and lower – monthly cost. Organisations can take advantage of economies of scale to lower costs while delivering more value. As COVID-19 has demonstrated too, a third-party will ensure staff are able to get to site, minimise routes for infection and make sure there are enough staff available to manage facilities safely – something that cannot be accounted for with other setups under lockdown.

Thirdly, capacity is a legitimate concern for those focused on TCO. Often, homegrown data centres and services must be oversized initially to plan for future growth. Environments typically have a lifecycle of having too much capacity to begin with and then eventually run out of capacity as business grows. With a colocation provider, an organisation purchases and implements what best suits its current business requirements with the ability to quickly expand as those requirements change and evolve.

Making colocation a part of a hybrid strategy can greatly improve disaster recovery capabilities. Whether it is leveraging a geographically diverse colocation facility or incorporating advanced recovery solutions, working with an experienced provider can ensure organisations get the solutions.



**Vicky Glynn**

Product Manager, Brightsolid

[www.brightsolid.com](http://www.brightsolid.com)



## Choosing the optimum workload environment

CHOOSING the optimum environment for your workloads is a debate that has raged on and is one unlikely to come to a conclusion any time soon. Why? Because choosing the right environment for your workload is a decision with as many options as choosing the right holiday. You know what you want to achieve at the end, but it's about getting the mix of various elements and bringing them together to achieve that right balance.

However, rather than weigh up whether you take a slightly smaller room as a trade-off for a better location, or pay a little bit more for the hot tub you didn't know you wanted; when considering an IT environment the balance means trade-offs between elements such as performance, features, speed, costs, security, supportability, integration, migration, culture fit and future proofing.

There is sadly no real art to achieving this balance as your environment is unique to your business – so there are few easy “absolutes” as those that might be non-negotiable to you may differ to another organisation; as such the decision making process around optimum environments is complex and nuanced.

Of course, some requirements point to some more likely solutions. For example, if you want full control, to pay capex, minimal latency, and physical segregation, you are likely to be heading to the on premises/ colocation end of the spectrum. However, if you need outside skills, scalability, and global access you generally will be heading towards cloud solutions.

The debate will rage on – because it is an impossible question to achieve a general answer, there are too many variables for an answer right for all of your existing and potential workloads. Instead, the question you should be asking is around which of your environments are optimal for this particular workload. This ability to choose is what is leading over 85% of enterprises to a multi-cloud or hybrid cloud strategy (according to IBM).

One way to revolutionise the thought process on this is to lead with XaaS (anything as a service) thinking. Assume that at some point you will have workloads that may need to be placed in any environment and work from there. With this in mind, the focus will change to how do you integrate solutions into your wider environment. This could include (but not be limited to) building relationships with providers in the market and developing your security position to bring on any solution. By building up your understanding of how you would absorb a workload on any of the options, you need to be clear on what the ROI model is that appeals to you.

Make some sensible choices sure, for example you may want to just settle on one public cloud (AWS or Azure or Google rather than all three) or one private cloud but embrace knowing how to deliver. Create a secure, compliant, network environment ready for new services to be added to.

Cracking the back of this solves the complexity from the unknown and put you in upper hand as to quickly implementing the right cloud at the right time for a particular workload.



## Svein Atle Hagaseth

Chief Sales Officer, Green Mountain

<https://greenmountain.no>



# Trends that will shape the future of the industry

WE CERTAINLY LIVE in extraordinary times, as the Covid-19 crisis has affected every part of society and business. As providers of data center services, we are part of the critical infrastructure that has been truly essential during this pandemic. It has accelerated growth of demand while at the same time tested our capabilities and raised important questions on the way forward for the industry. In these uncertain times there are some trends I believe will shape this future.

### Time to market – more hyperscalers will use colo

Over the past years, I have experienced that “time-to-market” is gaining more importance as a decision criterion for clients, they want RFS as soon as possible. In Green Mountain we have implemented four different deployment models. For immediate needs we can deploy rack-by-rack or cage solutions in a few weeks’ time. In the timeframe of 3 – 6 months we can offer room-builds or even prefabricated modular data centers.

Finally, for larger and more complex needs, like from for instance hyperscalers, a 12-month time range is realistic. We can then do a custom and scalable independent data center with several MW capacity. In contrast, hyperscalers building themselves usually have a time scenario of +36 months. I predict that hyperscalers will increase their use of colocation providers to enable “smaller” builds in new locations as this will shortening their time-to-market significantly.

### The triple bottom line – the next generation of sustainable data centers

Kass and Ravagni recently presented the following call-to-action

in a recent paper: “We must rethink how we design, build, and operate next generation data centers and adopt using the sustainable cost of ownership which includes economics, resource impact, and society benefits.” I could not agree more. Lead by the big corporations like Microsoft, Google and Facebook, this new way of measuring sustainable cost of ownership is gaining widely support.

By using this new framework, we must consider not only the economic factors in building and operating a data center. We must adopt a circular economy mind-set and consider the environmental and societal impact of a data center’s input and output through the whole life cycle.

Our clients challenge us on these topics more than ever; Is your energy truly sustainable? How do you reduce waste? How do you support your local community? These client demands help accelerate speed of innovation and the decisions that are necessary to build a more sustainable industry.

### The Nordics will lead the way

With the former paragraph in mind I believe that in the European market the Nordic countries will take the lead. They have prerequisites to support both economic and environmental factors with abundant renewable energy, low power costs and cold “energy-efficient” climate. In addition, these countries are known for being politically stable societies with proactive mindset in supporting rapid scale build-out of data centers.

I am therefore confident that the center of gravity will move north when it comes to colocation growth in Europe.

## Vinny Vaghani

Co-Founder and Head of Operations & Commercials, IP House

[www.ip-house.co.uk](http://www.ip-house.co.uk)



## Keeping customers connected

THE TRADITIONAL NEED for owning and operating on-premise physical infrastructure is a perspective that is radically changing. Questions centred on ownership and access are quickly dissipating in favour of realising the benefits of moving into a colocation data centre. IP House provide a range of resilient colocation solutions to suit London-based organisations, including full hosting and migration to bespoke hybrid IT environments.

### IT Infrastructure in the optimum location

In the centre of London, where commercial real estate is at premium, IP House offer an innovative and cost-effective home for your data. The facility is designed to Tier III standards and operates at the edge of the city's busy financial district. The data centre is accredited to ISO 27001 (Information Security), ISO 50001 (Energy Management), ISO 14001 (Environmental Management) and ISO 9001 (Quality Management).

Our goal is to keep customers connected, providing peace of mind by ensuring 100% uptime for their critical IT systems. 24 hours a day, 7 days a week, the team delivers a highly responsive and customer-driven service, leveraging industry-leading technology to empower and enable senior executives to focus purely on their business objectives. All critical systems are diligently monitored and maintained, and the team of dedicated on-site engineers can be your remote hands, minimising the need for customers to travel to site.

### Connectivity

Connectivity can be a challenge for any on-premise owner if fibre routes and carriers are not already in place. As such, IP House has established partnerships with Tier I carriers, cloud platforms, managed service providers and equipment manufacturers to

ensure diverse connectivity routes. Our resilient network offers customers the fastest internet speeds and delivers ultra-low latency connectivity which improves business performance. The associated costs of downloading data from some cloud services providers can be expensive. Our ability to negotiate the best pricing with recommended partners lowers both your costs and risks. We work with CDN providers who act as a bridge between your customers and data within the cloud.

### Modular growth & flexibility

IP House adopts a bespoke approach to working with customers, advising them on how to plan and adapt to meet future requirements. Through fully migrated, hosted or hybrid IT environments, we have the flexibility to scale as demands grow. Customers can establish a core site with their own assets and full access within our data centre, whilst benefiting from substantial cost savings on commercial space, equipment and on-going maintenance. IP House also provide options to help maintain customers on-premise solutions, utilising our colocation data centre for disaster recovery and backup infrastructure.

Today we work alongside customers to develop the best solution for their critical IT requirements. Hybrid architectures utilising cloud computing, colocation and smaller, localised on-premise infrastructure, create a scalable backup solution, compliant with GDPR regulations. From a dedicated data suite to a few racks, we offer the same availability, security, compliance and connectivity to all customers. Find out more information about our resilient colocation data centre within the heart of London by visiting <https://www.ip-house.co.uk>

**Make our house, your home.**

**Alexandros Bechrakis**  
CCTO, Lamda Hellix Data Centers  
<https://lamdahellix.com>



## Getting to grips with colocation

IT IS A WELL-ACKNOWLEDGED fact that data centers play a major role in storing critical data and keeping the digital economy up and running. Today's data centers though are more than that, they also act as connectivity hubs, providing customers with incredible possibilities when it comes to deploying their networks and safeguard their critical applications.

Colocation has become a key aspect of a business's digital transformation strategy and connectivity within a data center. The COVID-19 outbreak brought many new challenges and created a new way of working remotely worldwide, reshaping the term of potential threat to the uninterrupted operation of a business while highlighting the necessity of hosting your data safely.

Choosing a colocation provider with multiple connectivity options is one of the most important decisions an organization can make given that high-speed internet and great performance are being of vital importance when referring to data center connections as it allows businesses to exhibit a strong presence in their market. The rise of high profile providers offering Colocation services along with Hybrid Cloud solutions, like Lamda Hellix, enables small & mid-sized businesses and large enterprises to move their infrastructure into Hybrid Cloud environments, leveraging all the benefits of Hybrid Cloud without the cost of investment in resources, people and skills or for the deployment of an in-house Data Center.

A neutral data center can be a place of synergy creations among its customers with minimum interconnection costs.

Colocation is certainly a differentiating factor and competitive advantage among enterprises looking to establish further growth

and increase business effectiveness. Lamda Hellix provides Colocation services that help companies safeguard their critical data and applications within a flexible and modular environment that delivers reliability and industry-leading performance.

Lamda Hellix, a leading neutral data center services provider in South-East Europe, Middle East & North Africa, provides high-quality services that aim to help customers increase the safety and availability of their mission-critical applications while maintaining the same autonomy and independence as with fully owned infrastructure.

Our tech experts help each customer determine which solutions suit best to their company's IT requirements, controlling simultaneously their expenses as well as the scalability of their infrastructure. Additionally, we offer clear, high availability services with 24x7x365 monitoring, designed to cover the needs of critical applications for data storage and disaster recovery infrastructure and the customers pay only for the resources they use.

With our Hybrid Cloud & ICT services, our customers have access to premium, high-availability Hybrid Cloud services, including HWaaS, IaaS, PaaS, Container capabilities such as AKS or Service Fabric (managed or cluster solutions), App Services or Serverless code capabilities, and DRaaS designed to support critical systems and applications.

A fully customized services package where you can select among Private or Public cloud solutions leveraging all the benefits of Hybrid Cloud without the cost of investing in resources, people, and skills.



## Eltjo Hofstee

Managing Director, Leaseweb UK

[www.leaseweb.com](http://www.leaseweb.com)



# The business case must drive the best environment

“IN THE CONTEXT of the modern business environment, where so much nowadays is run virtually, in-house can be regarded as a bit dated – it offers the least flexibility and the highest amount of upfront investment. So, while this option may be a bit of a last resort, it still might be the best choice for very specific workloads, for example a factory where engineers need to work with the equipment often.

Colocation offers more flexibility in terms of contracts and less upfront investment. A great additional advantage is that if you don't have the connections in-house, it's much easier to connect from a data centre to a cloud, than from in-house to a cloud. In this regard colocation offers wider variety, and nowadays is often used as a first step to move infrastructure from on-premises to the cloud – a springboard if you like.

Positioned between cloud and colocation services you'll find dedicated servers, which are being offered by increasingly more

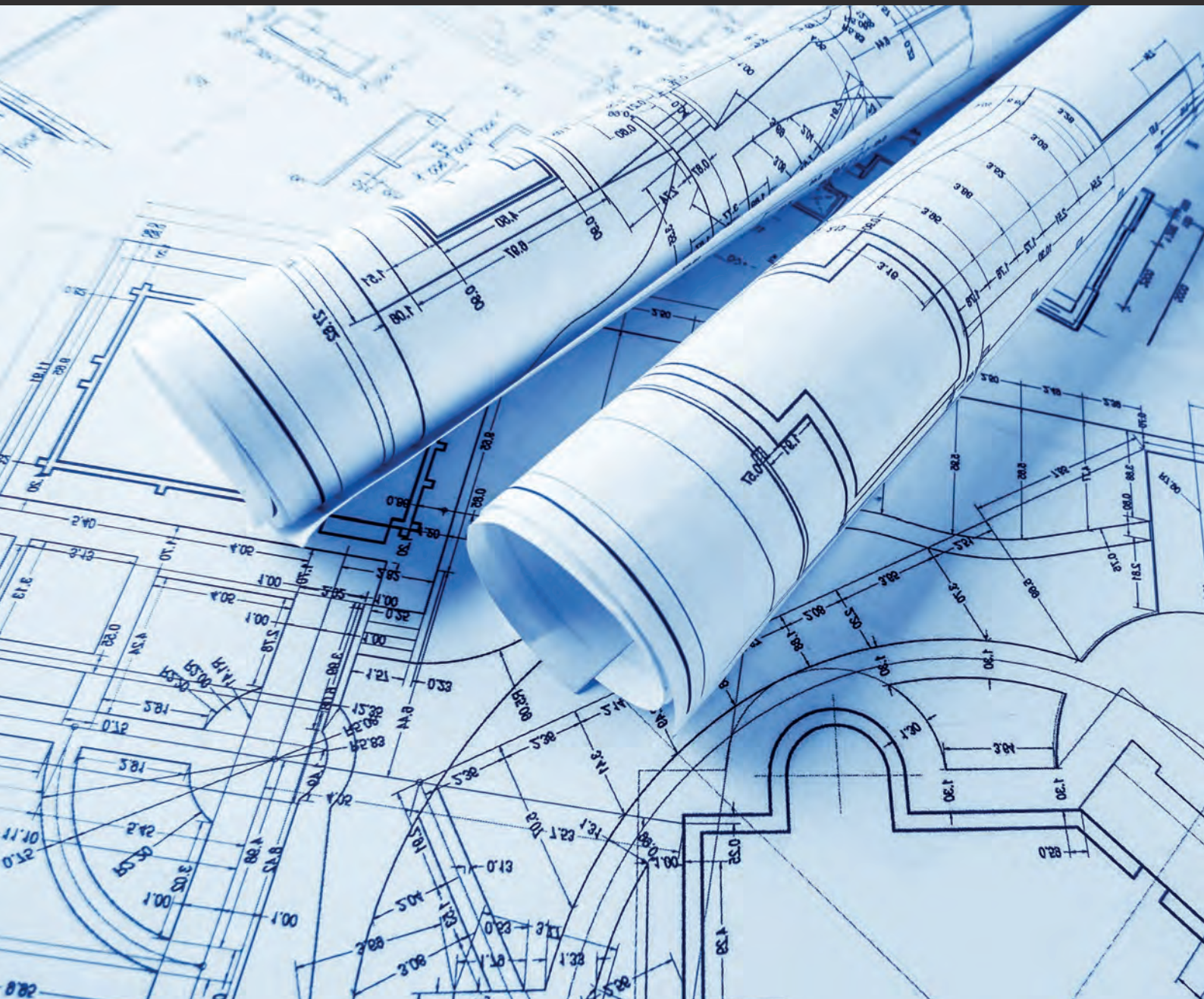
providers, almost as on-demand as cloud providers offering VPS services. Dedicated servers are managed a little like colocation, provide the advantage of leveraging more OPEX, and less CAPEX, and can either be in private or shared environments.

Data protection, regulations or very specific performance (defining or tweaking the hardware for example) are all good reasons to choose a private cloud option. Public cloud is a good choice for peak loads, or loads that are not full time.

For fulltime load, or baseload, rather than considering a hyperscaler, it may be better to source a public cloud provider who can offer services like these, but tailored to the business needs and with payment models that are not just pay-as-you-go (PAYG) to save on cost. Ultimately, the best environment for applications and IT infrastructure needs to be driven by the business case.”

Data protection, regulations or very specific performance  
(defining or tweaking the hardware for example) are all  
good reasons to choose a private cloud option

Centralised, distributed, centralised, distributed...and then along came the central, regional, local model, as edge computing came over the horizon. And that was after Cloud and managed services had already made their mark. And so we have a data centre industry requiring physical buildings and infrastructure for everything from the hyperscalers right down to the most local of edges – maybe just a small cabinet under a desk, or in a retail unit. Location, build quality and flexibility, and the facilities infrastructure required to support the IT load – the power and cooling, the cabinets and racks, the cabling and the floors – have never played such a crucial role.



IN ASSOCIATION WITH





# Offloading data centre maintenance to the cloud to support remote working



Forward-looking businesses are offloading data centre maintenance and management to the cloud to keep things running during the COVID-19 pandemic.

**BY RAJA RENGANATHAN, VICE-PRESIDENT AT COGNIZANT AND HEAD OF CLOUD SERVICES BUSINESS.**

For years, proponents have urged businesses to better enable employees to work from home, citing benefits like increased productivity, less commute time, better work-life balance and enhanced preparedness for business continuity, should a localised disaster strike, such as a tornado, hurricane, earthquake or flood.

Overnight, the COVID-19 global pandemic made the final argument for work-from-home a reality for millions of workers – ready or not. Many global enterprises must suddenly support more and more people working remotely, whether they are equipped to deliver and support workloads at scale or not. This has sent businesses scrambling to quickly embellish

digital channels and platforms, increase bandwidth, add virtual private networks (VPNs), provision more laptops, and offer thin-client applications to their employees and customers to improve operational collaboration and enforce social distancing. A proper business continuity plan followed up with precision execution can ensure that enterprises deliver such capabilities. However, what happens to the on-premises data centre where a physical presence is required? Even with workplace virtualisation technologies like remote consoles and “out-of-band networks”, which reduce the need for on-site data centre operations staff, the fact is, physical boxes in on-premises data centres still need to be managed, guarded and secured by people.

Take the February 2019 data centre meltdown of a major U.S. bank, which crippled the organisation’s online and mobile banking capabilities. The company needed to shut down one of its data centre facilities due to a smoke condition. It took two days to bring the facility back up, and only with significant effort, which required the physical presence of data centre staff. Imagine if this happened during the COVID-19 crisis. The time taken to fix the issue would increase exponentially due to a lack of people resources and hesitation to collaborate in-person. Even physical security could become compromised, which raises grave concerns.

An increasing dependence on the foundations of IT. The fact is, as our dependence on IT intensifies, data centres have become the substratum of how we live, work and play. From banking to insurance to 24x7 news, everything is supported by cloud infrastructure housed in virtual data centres. If these data centres go down, critical business functions, financial networks and in some cases our whole way of life become threatened. As a result, virtual data centres need to be continuously supervised and constantly cared for. The Uptime Institute’s 2019 Data Centre Survey puts this into context, some of the findings include: The staffing problem affecting most of the data centre sector has become a crisis. Sixty-one percent of respondents said their organisations had difficulty retaining or recruiting staff — up from 55% a year earlier.

Outages continue to cause significant problems for operators. Just over one-third (34%) of all respondents had an outage or severe IT service degradation in the past year, while half had an outage or severe IT service degradation in the past three years. Ten per cent of all respondents said that their most recent significant outage cost more than \$1 million. (The study authors note that “most recent” could have been at any time in the past.) Why transfer maintenance responsibility to the cloud? Forward-looking businesses are taking a different approach – they are offloading data centre maintenance and management to the cloud. One reason for this is scale, as cloud service providers

With many proven methods and tools, cloud migration is involved, but it is not difficult. For businesses that take a meticulous optimisation approach, the cloud can be more cost-effective than CapEx-hungry data centres

(CSPs) have mastered the art of managing scale. In addition to proactively planning for capacity, businesses can leverage auto-scaling features to rapidly meet any unplanned surge in demand. Cloud infrastructure is also highly automated and allows for the creation of scaling policies that set targets and add or remove capacity in real-time as demand changes. Thus, utilisation and costs are optimised, while the need for having more people on the ground is reduced.

Most CSPs now provide a multi-tenant architecture that allows different business units within an organisation or multiple organisations to share computing resources. This allows organisations to optimise their resources and staff vs. having their own data centres.

Lastly, the physical security in and around CSP data centres tends to be more robust and proven than what enterprises can individually afford. Most CSPs have rigorous and ongoing processes for assessment and mitigation of potential vulnerabilities, often performed by third-party auditors.

### Planning for disruption

With many proven methods and tools, cloud migration is involved, but it is not difficult. For businesses that take a meticulous optimisation approach, the cloud can be more cost-effective than CapEx-hungry data centres. Even in the context of coronavirus, digital platforms running on the cloud can unleash cost and operational advantages via centralised control while meeting bandwidth challenges that flare up during peak usage periods.

Perhaps it is our hyper-connected world, but severe disasters seem more frequent than ever. As businesses respond to the many challenges COVID-19 presents, they also need to keep their eye on the horizon to prepare their data centres to withstand any disaster that strikes in the futures.



## John Hall

Managing Director, Proximity Data Centres

[www.proximitydatacentres.com](http://www.proximitydatacentres.com)



# The rollout of 5G will dawn a new era for edge data centres – but there's work to do yet

THE ARRIVAL OF 5G has dawned a new era in connectivity and data sharing, and edge data centres will have a crucial role to play in its full rollout. Gartner's recent report highlights this, as it predicts that by 2025 edge computing will account for 75% of enterprise-generated data. The first major impact of the introduction of 5G will be the movement of massive amounts of data through faster and virtualised networking, as well as wireless infrastructure. Data centres will be at the heart of enabling 5G in all applications for the foreseeable future, and 5G itself will both drive new edge deployments, and enable new edge use cases.

There are a host of issues to be ironed out before 5G can achieve its full potential, though. The infrastructure of data centres must be strategized effectively to cope with the huge demand in data that 5G's full rollout will bring; new regulatory requirements will need to be negotiated, introduced and adhered to, and potential security vulnerabilities will have to be addressed, too.

This is by no means unachievable, however, and once 5G becomes a reality it will unleash the full capability of IoT, morphing the world into an environment that we would have never seen before. But before this becomes a reality, there are questions to be answered surrounding 5G's arrival.

### The security concerns surrounding 5G

5G will vastly increase the number of devices that can be connected to core networks – which increases the opportunities

for cyberattacks. The increased number of connected devices hugely expands the volume of information that will be transmitted and shared between users and devices, making it difficult for businesses to process and maintain data securely.

One of the ways of addressing this data volume challenge is through decentralisation - bringing analytics as close as possible to the data centre where information is processed, easing the burden on businesses of having to move around large data sets. Moving data to the edge means internal IT teams can reduce the amount of data they have to store, which is crucial when addressing cybersecurity.

As more processing takes place at the edge, new analytical approaches and updates will be able to be quickly pushed and spread around the wider network once they're optimised by a central analytics engine. As the threat landscape evolves, so too must an automated approach to security. By working in a more distributed way and utilising edge data centres, organisations will be able to focus on stopping the next attack, rather than scrambling to recover from the last one.

Ultimately though, the government must take the lead on how cybersecurity will be delivered – not only for data centres to adhere to, but also for the peace of mind of both businesses and consumers. The government has a key role to play in establishing a framework that data centres, and other IoT providers adhere to, to ensure accountability, if and when an attack does occur - but also good practice to mitigate against such attacks.

Data centres will be replacing data at a fast rate as they keep pace with the increasing volume of information that is created, as more devices become connected. However, the more data has to be updated or replaced, the greater the chance of suffering a leak. Having a strict guide that provides best practice on how to do this securely will provide assurances, and defences, against data leaks taking place.

### Edge data centres' role in 5G's future

With billions of devices to be connected globally, it is inevitable that not all these devices will be located in urbanised areas throughout the UK. As a result, edge data centres must be deployed as an extension of centralised data centres throughout both urban and rural areas, as well as coming out of the Greater London conurbation too. This will ensure that connected devices throughout the country are serviced, no matter where they are operating.

Edge data centres will play a key role in ensuring the whole of the UK is reaping the rewards of 5G, not just its major cities, by overcoming the latency issues and having wide-spread data centres present. And while centralised data centres still have a crucial role to play as they are the hubs of data distribution networks, it's edge data centres that will continue to act as the local depots of data, for regions across the country.

The growth of edge data centres will also lead to a required shift in the whole operations of data centres and their deployment, not just for the rollout of 5G - but forever. Because of the ever-increasing customer demand and surface areas requiring network coverage, the business model of data centres must change entirely.

The topology and infrastructure of data centres itself will naturally evolve into a business model where data is stored and transmitted equally, through edge data centres. In fact, by 2025 Gartner predicts that 80% of enterprises will shut down their traditional data centres.

### Regulatory issues and traditional rivalries

With IoT devices connecting around the world, it is inevitable that countries across the globe will have to collaborate and agree on blanket regulations, which is no mean feat. Different countries

have different regulations when it comes to 5G installation, but new standards must be set and adhered to in order for rival countries, rival product makers and rival network operators to put aside their differences. A

s we have seen with Huawei, this is easier said than done, after the US banned the use of its equipment in their 5G networks, and is calling on the UK to do the same. Rivalries aren't so easily laid to rest, it seems. And geopolitics is only one of several key issues that need to be ironed out, if 5G is to reach half of its capability potential around the world.

### New technical standards

To make global interoperability possible, new technical standards must be set for network infrastructure equipment, smartphones and IoT sensors alike. Ongoing standards deliberations are fraught with geopolitical and commercial rivalry challenges as well as the core technical ones.

Deployment of additional network transmission capacity  
Massive infrastructure is still to be deployed and developed in many countries, and to support 5G networks, regional edge data centres on mass are still needing to be connected to cell towers with fibre optic cabling, both in the UK and the wider world. This is absolutely necessary to support 5G's implementation.

### Licensing spectrum

Each national government must allocate, and licence, sufficient spectrum for both fixed and wireless links, as well as mobile 5G connections. 5G networks need spectrum within three key frequency ranges to deliver widespread coverage and support all use cases. There remains a key focus on establishing international agreement and allocated licensed spectrums. If unlicensed, there is no legal protection against outside interference to the networks.

### A new chapter

While there are many barriers yet to overcome, 5G will eventually change the world. But for it to do so, it will need to rely somewhat on edge data centres for its comprehensive delivery. Not only is the next evolution of internet connection going to change the way business and people operate, it will certainly birth a new chapter for edge data centres, too.



## Ciarán Flanagan

Group Vice President of ABB's Global Data Centre Business

[www.abb.com](http://www.abb.com)



# Future proofing data centers for a sustainable tomorrow

TWENTY FIVE YEARS AGO, when ABB entered the data center sector, the major considerations for our customers were uptime and reliability. That quickly evolved.

Starting in 2007, following the publication of the US EPA's Data Center Energy Use Report to Congress, rising energy consumption of data centers also became a major concern. Thanks to significant efficiency gains in IT and OT, predictions of soaring energy usage of data centers in the coming decade turned out to be far from reality.

Today, the combined use of electricity by data centers in the U.S. for example amounts to just 2 percent of the country's overall energy usage. This is about the same level as data centers consumed in 2006, making this a remarkable achievement considering internet traffic has risen more than fivefold, and data center workloads have tripled with the explosive growth in smartphones and cloud infrastructure.

Running data centers at higher temperatures and using virtualization to cut down on underutilized servers have played a key role in this efficiency gain over the years. In addition, the industry's move towards larger and more efficient hyperscale centers has helped, because these centers consume much less energy for cooling as compared with their smaller counterparts.

On average, one server in a hyperscale center is said to be able to replace 3.75 servers in a conventional center with a 2016 report from the Lawrence Berkeley National Laboratory estimating that energy usage would drop by a quarter if 80 percent of servers in small U.S. data centers were moved to hyperscale facilities. This has driven the shift to colocation and multitenant data centers from a cost and sustainability standpoint.

### Moving to ultra-efficiency

Energy efficiency achieved until now, outstanding as it is, isn't enough. Data center operators feel the pressures of sustainable stewardship to continually improve. Computationally intensive applications such as video on demand, autonomous vehicles and advanced 5G technologies are growing in popularity, driving the increase in data center demand and energy consumption. To deal with this surge, data centers will need to become even more efficient. At such ultra-efficient data centers, every watt counts.

For data center operators, the most important consideration is still to maintain business continuity, which means running uninterrupted 24/7, 365 days a year. Increases in efficiency cannot jeopardize availability. Since the most common cause of outages in data centers is power supply problems, the ability to balance the need for uninterrupted, reliable and efficient power is critical.





### Resilience versus sustainability

A key solution in driving out energy losses is advanced Uninterrupted Power Supply (UPS) systems, where enhanced power measurement allows the data center operator to accurately track energy consumption. By using decentralized parallel architecture (DPA), modules that are not needed are intelligently switched by the UPS to a standby mode, ready to transfer back to active mode if demand increases.

In the case of a power failure, all modules become active within milliseconds. Along with other support infrastructure such as smart sensors, intelligent switchgear and predictive maintenance algorithms, the UPS delivers a high system resilience alongside more sustainable operations.

Another technology that can help make data centers incrementally more efficient and sustainable are gas insulated switchgears (GIS). Compared with traditional offerings, GIS offer a lower lifecycle cost. This is possible because of their compact design and reduced frequency and cost of maintenance, which in turn minimizes operating costs and downtime risk. GIS can also be commissioned more quickly and can deliver a more sustainable solution – in particular those versions which do not use SF6, a greenhouse gas popular in the industry.

### Digital solutions to realise green gains

As the industry continues to demand more sustainable operations, Data Center Automation will also be key in achieving this objective. Given the industrial scale and criticality of today's data centers it seems natural that industrial-grade automation systems should be used to monitor and manage operations.

Industrial automation systems are robust and highly hardened systems that can scale from small to extremely large and from simple to extremely complex. They are designed to talk to all kinds of equipment using many different protocols and are used in continuous operation over long periods, even while being upgraded.

By enabling inefficient power and water usage to be identified and automating reporting to track progress and compliance, digital energy management solutions can facilitate the reduction of a facilities energy usage. In the case of Lakeland Community

College, they were able to leverage data center automation to identify cooling inefficiencies and forecast the impacts of layout and capacity changes to reduce the facility's energy usage by more than 50 percent.

### Every Watt Counts

While technologies and operating practices implemented to date have successfully stabilised the energy consumption of data centers, the stakes have undoubtedly been raised. Today it is no longer enough to prevent further increases in energy usage, the aim is to achieve carbon neutral or negative data center operations, often within a very limited time frame.

Drawing on the latest innovations in power distribution and energy management will be key in delivering future proof, sustainable data center operations. No single technology can solve the challenge alone, it will take an ecosystem of fully integrated and intelligent solutions to realize this objective, evidence indeed that every watt counts.

Drawing on the latest innovations in power distribution and energy management will be key in delivering future proof, sustainable data center operations. No single technology can solve the challenge alone, it will take an ecosystem of fully integrated and intelligent solutions to realize this objective, evidence indeed that every watt counts

**Maikel Bouricius**  
Marketing Manager, Asperitas  
[www.asperitas.com](http://www.asperitas.com)



## Helping with high density

ASPERITAS is a clean and high-tech company. As leading immersion cooling specialists we provide cutting edge immersion cooling solutions for energy efficient and high density datacentres globally. We launched our first solution in 2017 after an extensive R&D phase, with an ecosystem of cutting edge partners since 2014.

### Challenges

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. This results in a situation whereby the current datacentre model does not fit with the technology it should be facilitating.

The challenge we have is to communicate as effectively as possible with Datacentre Infrastructure managers to ensure that the datacentre model is open to change, in order to be future ready. With Immersed Computing®, high density datacentres can be built on only 1 fifth of the space allowing up to a 40% increase in performance.

### Developments

Asperitas' 15" and the new 21" ready AIC24 immersion cooling solutions are designed to ensure the highest possible continuity of the IT. 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 spearhead 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 these emerging technologies are 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 can facilitate 5-10 times as much density.

### Opportunities

The AIC24 module takes up 80% less physical footprint and 95% less cooling energy for datacentres. The next generation hardware for these emerging technologies are 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, which is where the AIC24 module comes in!

Asperitas wants to live and work for reshaping and innovating the datacentre world.

## Jason Bunting

Head of Engineering, Bud Systems

<https://bud.co.uk>

# If you're building a business, build it in the cloud

BUD SYSTEMS is a management platform that supports every aspect of apprenticeship delivery. Essentially a Platform-as-a-Service, we bring together everything, from applicant management and programme creation to workforce and learner reporting and funding management, into a single package. Our focus is always on the best user experience, and creating new ways for education providers to be more cost effective, efficient and remove a lot of admin that enables them to focus their time on delivering outstanding training. Partnering with a third-party cloud provider, we built Bud in the cloud, which is a key pillar of our overall proposition and approach. This is why...

### Scalability

As a business that is rapidly growing, and has aggressive projections over the course of the next 3-5 years, trying to meet that need with on-premise infrastructure is not practical involves substantial investment and additional headcount in employing operations teams. Bud also needs the ability to ramp up its user capacity at the click of a button, and having our own infrastructure on-premise simply wouldn't be able to compete with the flexibility we get by being a full cloud delivered service.

A lot of it comes back to scale but it's driven but like any business cost efficiency has a not insignificant part to play. Any business needs to have the capacity to meet the demand of its busiest 1% of the year. Given the nature of what we do, with an on-premise solution, we'd often end up with huge amounts of capacity sitting idle for extended periods of time. The flipside of that being you may then need to go from 10% to 90% capacity instantly, so our system needs to be able to cope with that particular sort of increase in demand. When we look at our usage patterns, with users consisting primarily of learners and training providers,

we can effectively plan our provision to be both proactive and reactive, and combined with our microservice architecture be very granular in our approach. We often reduce capacity overnight. It's a pay-as-you-go model, where we're fine tuning the spend to exactly or to within just a few percent of what we need. This means that we're working with a cost-effective solution for the business in terms of operational costs, but equally we can provide an excellent service to our end users because there's always sufficient capacity to handle the load at any one time.

### Always Adding Value to the Business

I want my team focused on delivering value to the business as much as possible, and for us that's user experience and new and enhanced platform functionality.

Bud was created to fill a gap in the market. Tech solutions that were already out there were predominantly designed for 'expert users' only. Very often, really good apprenticeship trainers aren't tech-savvy – their passion lies in teaching, not administration. No system existed that offered a simple but great user experience that supported trainers in delivering training consistently across the board – that's what Bud is all about. Running an on-premise infrastructure, you can spend inordinate amounts of time on just maintenance. Being able to push the responsibility of keeping everything running correctly and securely to the cloud provider, frees my team up to focus on what we believe really counts, whether that's working on enhanced functionality or improving system performance and existing features.

A great user experience, combined with the elasticity of being able to effectively manage capacity as and when it's required through extreme peaks and troughs, is what's driving our success.

# AIRSYS

**Philip Smith**

Engineering Sales Manager, Airsys UK

[www.air-sys.uk](http://www.air-sys.uk)



## Keeping IT cool

LONG GONE are the days when the IT manager was only concerned with the safety of data and resilience of the infrastructure, with energy a secondary concern. The conventional method of cooling was, and still is, the use of downflow close control “CRAC” units (computer room air conditioners), using a raised floorvoid as a means of distributing cooled air to the racks.

Nowadays the minimisation of energy usage, is seen as the primary driver of IT infrastructure design, and this has led to the adoption of free cooling, widely available and in varied forms. However, the risk to data can be increased if the system is not designed to cater for all eventualities – periods of high ambient temperatures, especially when combined with high relative humidities; and air pollution, especially saline atmospheres, diesel fumes, and pollens.

If the end user has relaxed tolerances for temperature and humidity within the conditioned space then the use of unconditioned fresh air will provide free cooling if the ambient air is cool enough, usually a minimum of 5°C below the design space condition. The smaller the temperature difference between ambient and the space design then the more fresh air is necessary to provide sufficient cooling, which may then exceed the design capabilities of the ventilation system.

When the ambient exceeds the desired space temperature then you lose control altogether. Generally this technique allows for space temperatures around 25°C to be controlled for 90% of the year.

The use of adiabatic cooling can lower the temperature of

the incoming fresh air and enable cooling in higher ambient temperatures to be obtained, but this is also reliant on the ambient humidity levels. If the air is moist then the effectiveness of adiabatic cooling is reduced considerably, the supply air can become too humid, posing a subsequent condensation risk on IT surfaces. Generally this technique allows for space temperatures around 25°C to be controlled for 95%+ of the year.

### There are ways to improve things:

- The use of mechanical refrigeration cooling will supplement those periods when fresh air alone cannot provide the required duty. This will also allow the setting of lower space temperatures if required and will enable control for 100% of the year.
- The use of indirect free cooling will eliminate the issues of humid or contaminated fresh air. This uses a closed condenser water loop to pre-cool the air and then refrigeration is used to supplement any shortfalls.

You can even combine direct free cooling and indirect free cooling if your budgets will allow!

At Airsys, specialist cooling providers to the telecoms and data centre market, we provide all these options, from 10kW to 100kW+ and can demonstrate the pros and cons of each system so that the best solution is presented to the client. We offer a full design and build service. Airsys do it all.

To safeguard any power issues all Airsys units have the option for essential and non-essential supplies. With fans and controls via the essential supply (UPS/generator) cooling is guaranteed.



## Tim Loake

Vice President, Infrastructure Solutions Group  
UK, Dell Technologies

[www.dell.com](http://www.dell.com)



# Autonomous architecture

THE INCREDIBLE advancements in robotics, automation, and artificial intelligence (AI) over the last decade have created a material shift in our relationship with technology. While many people will see some of the visible manifestations of this shift in shiny exploratory technology, such as Starlink, real innovation is naked to the human eye: it sits, hidden within the advancements of autonomous infrastructure.

Ten years ago, if you had asked any IT professional to predict the future, they would, with confidence, have said that we would be operating in an environment where the vast majority of applications and data existed entirely in the public cloud. However, as we know now, that “nail in the coffin” of on-premises infrastructure was a touch pre-emptive. These predictions underestimated the gravitational force of on-premises IT infrastructures and the considerable cost and complexity associated with migrating applications and data to the public cloud.

Today we more clearly see that the future of on-premises IT infrastructure is less about extinction and more about evolution. Gone are the days when infrastructure must be handcrafted and carefully tended; instead, we set our sights on achieving cloud-like infrastructure and operations, developing the ability to scale compute, storage, and networking resources, autonomously, dynamically on-demand and at speed.

As such, we will continue to cross the spectrum of autonomous capabilities and see the evolution of the intelligent controller – software that takes responsibility for the operational and

maintenance processes that make up the lifespan of an infrastructure system. We can use autonomous vehicles as a parallel. Today, it is possible for a human driver to let their car know their desired destination; and, thanks to the progression of systems like lane assist, navigation, and self-parking; the computer, with an awareness of its surroundings, can take over multiple functions (e.g. braking automatically for corners, keeping a set distance from the car in front) to get there. It's the same deal for IT infrastructure - advanced automation will allow administrators to state the desired business outcome, and the system will calculate the best way to achieve it. And just as a high level of autonomy can take the pressure off the driver of a car, it can also take the strain off administrators of infrastructure.

If we agree that Utopia is an infrastructure that is entirely self-deployed, self-provisioned, self-managed and self-healing, we must prepare ourselves for the advent of the complete and absolute intelligent controller; one that can assume full operational responsibility. And when the onus of abstracting complicated management functions moves to an intelligent controller, skilled human operators will be freed up, ultimately becoming more proactive and strategic in focus.

On-premises cloud environments which include highly automated infrastructure, will eliminate previously inherent complexities, and in doing so will help businesses become more efficient and strategic. Of course, technology forecasters have been proven wrong before, but I believe that for businesses to truly transform, intelligent, autonomous infrastructure must be at the heart.



## Dave King

Product Manager, Future Facilities

[www.futurefacilities.com](http://www.futurefacilities.com)



# 2020 vision of the future data centre

AT THE END OF 2019, we saw an increase of chatter around the digital twin, liquid cooling and edge computing in the data centre industry. We predicted that 2020 would see conversations turning to how they could each be used to improve data centres. Now, and partly due to COVID-19, we're fast discovering the best ways to optimise the operations of a data centre remotely; be it managing infrastructure, application deployment in IT or power distribution.

### AI and robotics

AI and robotics are also set to change the data centre industry. We're beginning to discuss whether there will be a rise of robotic assistants in the automated data centre, completing repetitive tasks or low level maintenance. This would certainly prove extremely important during periods where remote work and reduced travel is a necessity, and the virtual workspace is not just cost effective but potentially can save lives. To achieve a reliable and resilient automated data centre in the future, there is the need for a digital representation of the physical space for the robotics to operate. Improving existing robotic abilities to judge space and depth is key to minimise the risk of accidental or prolonged downtime. This is where the digital twin steps in. A 3-dimensional physics based computer model that replicates the precise behavior and performance of the real data centre. This technology enables data centre managers to give robots a detailed perspective on spatial awareness, test setups and layouts, and practice tasks repetitively to train the digital assistants.

### Bring on more edge

Social disruption and distancing is creating demand for more mobile applications that support remote operations in

vital industries, from banking to healthcare. Thankfully, with developments in 5G networks, we have more options and stronger capabilities than ever before. This is however bringing pressure on data centres to deploy more sites at the edge to increase app performance and reliability.

### Rise in liquid cooling

The growth of edge data centres is also driving the take up of liquid cooling. The advantages of liquid cooling have been known for a while. Offering better thermal management than air cooling, liquid cooling is set to help boost rack and server density, as well as presenting the opportunity for the data centre industry to be more environmentally friendly.

The vast majority of deployments of liquid cooling direct to IT are currently still in the high-performance computing (HPC) sector. This is primarily because the high power densities are challenging, if not impossible, to cool with air cooled systems. But, the majority of these installations report that the energy consumption is much lower than it would have been using air. This also results in both reduced operational costs, and a consequent environmental benefit.

Over the next few years we expect the current resistance towards liquid cooling to disappear as more data centres realise the benefits. Already hyperscalers like Google and Alibaba are leading the way with liquid cooling and we expect more to follow. It's clear that we still have a long way to go, but the first half of 2020 has certainly changed the industry's perspective. The journey we are on towards automated, liquid cooled, efficient and environmentally friendly data centres is a long one; but one well worth covering.

**David Craig**  
CEO, Iceotope  
[www.iceotope.com](http://www.iceotope.com)



## Don't get tipped over the edge

AS ORGANISATIONS DEPLOY processing and storage capacity outside the data centre and enterprise, and into non-dedicated environments close to the point of use, the challenges multiply.

Edge applications will increasingly employ AI and edge-AI capabilities which rely on the dense heat loads generated by GPUs. Energy consumption, power availability at the site, physical security and tamper risks as well as reliability and resilience in harsh IT environments are some of the obstacles that enterprises will need to overcome to deliver real-time value at the edge.

### Traditional is best?

Various IT suppliers are tweaking traditional technology which includes air cooling and large mechanical fans and repackaging these as edge solutions. This may appear convenient in the short term, but it may not be effective in the medium-term and could prove to be an expensive and ecological mistake in the longer-term, as thousands of edge sites appear.

Traditional air cooling has served the industry well but taking it out of the climate-controlled data centre environment that supports it, and you forfeit not only control but also efficiency. Your once enviable data centre PUE goes through the roof as inefficient cooling is a major contributor to wasted energy.

Fans and moving parts are critical to cooling electronics with air but are also failure points which need regular maintenance. The in-chassis vents and holes which enable airflow to remove heat, also expose the IT load to environmental contaminants and people. Considering a high percentage of data centre and technology outages are caused by people, as we multiply the

data sites around the world we must reduce human interaction within those environments.

### Cool calm connected

To overcome these obstacles and enable edge to deliver its full potential, we need to move from legacy systems and implement new solutions that are edge smart and fit for purpose; technologies that can be deployed resiliently, securely and responsibly.

Chassis level liquid cooling uniquely allows your compute to be deployed to the edge within a sealed and safe environment which makes it impervious to heat, dust and humidity. This server format also provides a high degree of human error resilience, or what might be described as 'idiot-proofing'.

A sealed and silent chassis allows edge IT to be delivered in latency-busting proximity to the point of use with class leading energy performance from installation throughout its lifetime. Reducing single point of failure and increasing resilience, chassis-level liquid cooling technologies significantly reduce the cost and risk of servicing and maintenance at the edge.

New and emerging applications will play a major role in advancing our business and home lives. Key to this is the way the physical network infrastructure such as chassis-level liquid cooling is utilised in the up scaling of a reliable and efficient IT service for always-on users.

The rollout of edge computing must provide highly energy efficient environments in which to house the next phase in global digitisation and stop us tipping over the edge.

**Dr. Thomas Wellinger**  
Marketing Manager at R&M  
[rdm.com](http://rdm.com)



## Preparing for convergence and densification

TODAY, we can't discuss network trends in one area without looking at requirements across other areas. Developments outside the data centre are changing the way we need to think about data centre infrastructure.

First, let's look at some developments related to data centres. The number of fibers used for transport is increasing. Multimode applications of 40 and 100 Gbit/s require eight fibers in parallel pairs. Higher speeds require eight or 16 pairs. Years ago, the traditional hierarchical network topology with core, aggregation and access level was already becoming too restrictive and slow. A leaf-spine architecture, however, considerably reduces latencies and ensures data and applications are available in real-time at all times.

Migration to 100, 200 and 400 Gbit/s requires higher density and more cabling in racks. Data centres may contain hundreds of thousands of ports and patch cords. IoT-driven Smart Networks facilitate management of this complex infrastructure, unburdening humans, boosting efficiency, and reducing Mean-Time-To-Repair by around 75%. In future, AI support may be added.

However, key drivers for change in data center networks are also coming from other areas. Data centres are increasingly becoming converged with the networks and functional requirements of the wider world. For example: Statista Research figures put the global installed base of IoT devices at 75.44 billion by 2025. Technologies supporting these devices, such as 5G, and Wi-Fi 6, are changing the network landscape. 5G speeds could surpass 10 Gbps - up to 150 times faster than 4G, with 20x lower latency, down to 1ms.

However, wireless solutions' inherent limitations mean optical fiber connectivity needs to be taken to each 5G connection point. Connecting 5G base stations using existing microwave links won't suffice: new active 5G antennas produce vast data traffic - existing technology can't handle such volumes fast enough. A dense fiber backhaul network and (Ultra-High Density) edge DCs are needed to offer 5G quality of service and enable new 5G user experiences with critical low latency applications in urban areas. Centralized RAN (C-RAN)-based solutions, with multi-operator capability, will optimize management and allocation of centralized base band resources. Wavelength-Division-Multiplexing (WDM) transmits services at different wavelengths, boosting network capacity without adding extra fiber, and can cost-effectively support fast-growing requirements, offering flexible upgrade paths for different carriers and new service providers.

By extending LAN cabling with technologies such as Wireless LAN, PoE and SPE, digital building automation can be realized exclusively with Internet Protocol. The 'All over IP' approach provides high levels of standardization, availability and reliability. This is already being enabled by RJ45 and Power over Ethernet (PoE), and will be supported further by Single Pair Ethernet (SPE) based on xBASE-T1. Smart converged data networks can be extended through ceilings, making it possible to connect devices to building automation via pre-installed overhead service outlets.

Tackling new trends and arriving at the best solution for your specific application can be challenging. As there's no single 'one size fits all' solution, sharing expertise and engaging in dialogue is more important than ever.



Life Is On

**Schneider**  
Electric

## Steven Carlini

VP Innovation and Data Center, Schneider Electric

[www.se.com](http://www.se.com)



# My data center predictions for 2020: The year we set the stage for futuristic autonomous and immersive applications

CLOUD COMPUTING has been a bit of an enigma over the years but now, driven mainly by the smart phone era, almost everyone understands that their pictures live in a large cloud data center somewhere. There are certain geographical “sweet spots” where these very large data centers cluster around the world.

These sweet spots include low cost or renewable power, access to IT workforce, tax incentives, and lately, a climate that is on the cold side and conducive to “free cooling.” Unless you get lucky and live in close proximity to one of these hyperscale data centers, you most likely have experienced slower than expected performance for cloud services like Microsoft 365, Salesforce, Box, etc. Additionally, if you have moved your business-critical operations to the cloud and you are not satisfied with the performance or availability of those applications, you probably have heard that the solution could be an on-premise (your site) “cloud stack” or “tethered cloud.”

There has been a lot of speculation and promises around extending the data center cloud on-premise to the “local edge.” Smaller versions of these hyperscale data centers could increase speed, lower costs, and allow businesses to keep data within their four walls, giving them greater control over that information. The idea is to allow businesses to run similar IT infrastructure on-premise to improve the consistency of their hybrid cloud system. The ultimate goal is to enable businesses to actually use

the same tools, APIs, hardware, and functionality across both their on-premise and cloud systems to create a consistent hybrid user experience.

While Azure stack has been available for a short time, it has not exactly been affordable. The Azure solution has more cost-effective versions coming. Amazon Outpost was officially made available at the end of 2019, and Google Anthos will be right on its heels. 2020 will be more of a year of validation and site testing, but it's a move in the right direction.

**Next Gen Telco (5G) and IT Data Center Architecture Converge**  
Today's connected society and proliferation of high-bandwidth technology – video, immersion (AR/VR), haptic, etc., – require local data architecture to deliver on the promises. Yes, even 5G must deal with physics and the speed of light, and it will be deployed in small clusters with each one requiring a mobile edge computing (MEC) data center.

Think about a quantity of four new MEC data centers in the same area of a single 4G tower and base station. It's going to be a tremendous challenge first to justify the ROI, then to find locations for these MEC data centers (tops of buildings, parking garages, basements) where the cost to deploy is going to be quite expensive and access rights to these locations is questionable. At the same time, cloud providers (Amazon,

IN ASSOCIATION WITH



**DATA CENTRE**  
**SOLUTIONS**



**DIGITALISATION**  
**WORLD**



Google, Microsoft) are releasing their intention to do on-premise local cloud stacks or tethered clouds at the local edge – a very similar deployment strategy. Additionally, the application needs and drivers of the cloud going to the edge, namely latency, overlap significantly with the promises that 5G makes for mobile users (public and private) and make MEC data centers a natural and likely meeting point of cloud services and the mobile world.

Finally, 5G is a software designed technology and the intention is to operate on standard IT servers so it is quite possible and likely we will see MEC data centers providing on-premise cloud services, as well as running the 5G application.

### Energy Costs at the Edge can be Massive

I am predicting that next gen telco and tethered clouds will drive the massive build-out at the edge. If we use the local MEC needed for 5G as a proxy, we can calculate the impact on energy and carbon. In terms of scale, significant global coverage in 2/3/4G is in place with about 5 million telco tower base stations in the world with average power draw at about 6 kilowatts (kW) rising to 8-10kW at peak traffic periods.

The global footprint is 50GW at peak power! Unfortunately, most of these tower base stations were not conceived with energy efficiency in mind. They operate around a PUE of 1.5 (power in/power of the telco (IT) load), meaning that about half of the power is wasted. When deployed at scale, this power adds up quickly and waste is multiplied by the number of deployments.

Here's a real example of an initial 5G buildout: a Chinese operator recently added 100,000 5G-ready base station sites averaging 10kW each – that's 1 GW of energy! At a PUE of 1.5, this could cost 1.3 B€ (\$1.45B) and give off 9.3 million tons of CO2 annually (based on U.S. national average CO2 footprint). But if these systems were designed to be extremely energy efficient, PUE could be 1.1, and it would cost 1B€ (\$1.12B) and give off 6.8 million tons of CO2 annually.

Energy costs and controlling carbon are critical considerations for this generation of telecom. I predict a focus on standardization and innovative solutions like liquid cooling and cloud-based management systems will be leveraged.

### Drop-in Computing Racks for AI Learning

Artificial Intelligence (AI) awareness has gone mainstream,

mainly from what people view in futuristic movies. AI is chiefly seen as robots that have progressed to be able to learn on their own. We are far away from that reality, but we have begun a path where we want our computers to start learning things on their own by using a neural network model, which mimics the way the human brain works.

For example, healthcare applications are reducing the time it takes to plan treatments or diagnose ailments with the use of machines. Once the machines have developed the algorithm and the inputs needed, much less powerful computers can be used for inference AI.

The issue is that learning AI, or training AI computers, use the most powerful processors (CPU and GPU) and could produce 20, 30, or up to 50 kW of heat. These heat levels are inefficient and, in most cases, impossible to cool with any form of popular mainstream cooling.

For 2020, I see a trend where data centers of all kinds – cloud, colocation, and enterprise – will need to deploy dedicated drop-in computing racks for AI learning. The best way to cool these is some form of liquid cooling – immersion or direct to chip. These drop-in racks will not need air containment or to be lined up in rows or pods. Again, they will be limited in number and once they have developed the appropriate model, they will be dormant until a new model is needed.

### Data Center at Scale for ASiC

An application-specific integrated circuit (ASiC) is an integrated circuit (IC) customized for a particular use, rather than intended for general purpose. AI as well as digital ledger type applications (blockchain) operate far faster and much more efficiently with ASiCs vs. general purpose chips.

We will see companies that are not considered mainstream GPU and CPU developers enter the market with data center scale application specific chip sets. 2020 is an exciting time to set the stage for futuristic autonomous and immersive applications.

Computing power and storage will be packaged more effectively and operate more efficiently in the form of edge computing and ASiCs. While the industry is motivated to move fast, the stakes are high as the deployment scale, energy use, and carbon footprint are top of mind issues.



## Nicolas Roussel

Technical Manager Central Europe, Siemon

[www.siemon.com](http://www.siemon.com)



# Data centre design: the heart of IT infrastructure

THE NEVER-ENDING CHALLENGE for data centre operators is to be able to provide a reliable IT infrastructure to ensure high performing service levels without latency and to, by design, develop schemes to reduce any form of downtime. IT technology will continue to grow and change so we must be willing to adapt the infrastructure to support the emerging technologies and the strain being placed upon the data centre environment to support rapid deployment, increased performance and operational efficiency.

The first and main factor for a data centre owner and their end users is performance; data centres need to ensure high performing low latency IT infrastructure to satisfy data traffic. Especially today, when we are running many more web conferences, online videos, social media activity, using more connected IoT devices or stream music, availability must follow a growing demand. This is also why edge computing data centres are emerging, responding to these new demands, and effective design is essential for success.

A key design consideration must be to consider both scalability and rapid deployment, allowing you to easily expand computing services without redefining the physical infrastructure at each technology update. At the Siemon company, we are specialists in data centre infrastructure design and can easily help you to achieve your goal with our robust and efficient solutions.

This said, before designing any data centre environment, we all need to first consider the IT needs for today and the future, we do this by incorporating our awareness of new network application developments and then extrapolate the future needs, selecting and recommending the right scalable topology. Only

after this, can we start to think about the best reliable physical infrastructure to ensure performing low latency computing.

In order to achieve critical, high performance and low latency, we need to challenge duplex based systems with parallel optics connectivity in the same way as balancing structured cabling with our high-speed interconnect direct attach cable solutions. This will drive us as designers to the best suitable topology with the required flexibility. Some forms of IT infrastructure will perfectly fit into one topology while some others will have completely different needs.

In certain cases, top of rack (ToR) design will probably enhance energy efficiency as well as operational cost, while in some other cases, the same ToR approach could run into over subscription and be counterproductive. There is never a single answer, as every data centre environment is different.

To achieve the needs of the IT infrastructure and customer expectations/requirements, we need first to gather as much information as possible and be engaged as a team to find the best suitable solution. During this process we review a range of criteria including performance matrixes such as reliability, redundancy, scalability as well as financial factors, such as energy and cooling efficiency and operational costs.

Even if products are your major focus, we need to ensure design and installation is completed and tested to the highest quality against your performance requirements. We achieve this with the help of our unique Data Centre Design Service, but also by providing features in our products that make installation and utilisation easier to achieve and maintain.

## Herman Chan

President, Sunbird Software

<https://www.sunbirddcim.com>



# Getting clever with cable management

DATA CENTER cable management is a complex task. Often overlooked and neglected over time, improper cable management can wreak havoc across the entire enterprise and increase downtime. Poor data center cable management practices can result in “spaghetti” cabinets, difficult equipment installations, extended periods of troubleshooting and maintenance, and unsafe environments due to restricted airflow.

Tracking cabling in the data center with Excel spreadsheets and Visio diagrams no longer gets the job done today as sites grow, rack densities increase, and multi-hop connections introduce documentation complexity. Many organizations are deploying Data Center Infrastructure Management (DCIM) software to address the key challenges of data center cable management.

While the initial design of the network or structured cabling infrastructure could be documented in advance by contractors, ongoing documentation of patching as a result of moves, adds, changes could prove challenging. DCIM ensures continued accuracy of the cable plant by allowing you to document the initial design as well as the multitude of ongoing changes to enable real-time transparency of physical network/patch port capacity at every cabinet in the white space.

Poor documentation leads to difficulties planning, provisioning and maintaining equipment and troubleshooting connectivity issues. With a database of available and used patch or network switch port capacity in every rack, DCIM software makes it easy to search for capacity, plan rollouts, create work orders, provision equipment, and document patch cable installations down to the port level. A database driven tool with a web-front end allows anyone to easily view entire circuits and every hop to every device in between, down to individual physical ports on

every device. You can even visualize circuit routes in 3D between all devices in an entire room.

Accurately measuring cable lengths before installation keeps the data center free of loose cables and reduces wasted cable, but often involves a time-consuming process of walking the data center floor with a tape measure. With CAD-level accuracy, DCIM software automatically measures all routes to ensure correct cable lengths are installed, avoiding rework or downtime when changes are implemented.

The team responsible for the equipment and cable installation may be different than the team who plans infrastructure deployments, and insufficient instructions in the build sheet can lead to incorrect installations and connections.

DCIM software can radically simplify data center cable management by automatically validating the compatibility of your planned connections before allowing you to make a change. DCIM software can also provide clear work orders with cabinet elevation, equipment specifications pulled from the software’s database, and cabling instructions to ensure that the work is done accurately the first time.

With DCIM software, dashboards, reports, and what-if analysis are readily available, helping you plan for the future by seeing the impact of planned additions and decommissions on connectivity capacity.

Overcome the challenges of data center cabling with best practices and a modern DCIM solution, and you’ll be well on your way to driving sustainable behaviors that foster precise and accurate cable planning, installation, and management.



## Jackson Lee

VP Corporate Development, Colt DCS

[www.coltdatacentres.net](http://www.coltdatacentres.net)



# Hyperscale is the answer for IoT

As data centre operators, we are always striving to ensure we are a few steps ahead of the technological curve. This way, our products stay relevant for the demands of the next generation of customers. Over the past few years, we have seen a surge in large-scale, mission-critical IT and networking solutions utilise IoT technologies that require never-before-seen levels of bandwidth, connectivity and transmission speeds. Analyst firm Gartner predicts that there will be as many as 5.8 billion enterprise and automotive endpoints by the end of 2020. When we take into account the recent surge in home working, the acceleration of this trend only compounds further. With such rapid progression afoot, we can no longer afford to just think about current customer requirements. In order to avoid obsolescence, the data centre industry must build for the future.

### Hyperscale is the answer

From a facility design perspective, scalability of infrastructure is key to meeting the incredible demands posed by enterprise IoT. This means data centre facilities need to be able to rapidly increase capacity for customers, whilst also ensuring they are operating at peak efficiency at all times. As such we are already seeing an increase in demand for hyperscale data centre facilities with new builds in the 100MW+ range appearing around the globe. As our customers grow, we must grow faster. Operators who fail to incorporate scalability into their product roadmaps risk falling behind and losing enterprise customers to service providers that are better suited to their needs.

### Upscaling management and operations

Nonetheless, infrastructure is but one piece of the hyperscale puzzle. Operators must also ensure the correct management

and security protocols are in place to ensure the facility can cope with the increased load. A data centre is only as effective as its operator, so when we take into account the sheer physical and virtual size of a hyperscale data centre staffing and management expertise is crucial for ensuring the facility functions properly. This means increased staffing and enhanced management systems to proactively anticipate and diagnose infrastructure and network issues to avoid damaging downtime.

### Staying secure

It is critical that operators do not let security fall to the wayside as customer demand increases. Leveraging automation and AI-powered solutions can be key in helping ensure data centre services are protected. Physical and cybersecurity systems need to be regularly tested to ensure they can keep up with growth in facility size and data traffic.

The benefits of enterprise IoT are well known. However, we live in a world where cyberattacks are to be expected. Every connection point is a potential backdoor for hackers, so data centre providers must ensure security does not fall by the wayside to the sheer volume of connected devices on an enterprise network. In order to stay secure, customers are thinking beyond their own firewalls and expect data centre providers to maintain the most secure environment for their data as possible.

In these uncertain times, preparation now is the only way data centre operators can keep pace with customer requirements and ensure their infrastructure and services are futureproof. So, when it comes to rapid IoT adoption, hyperscale is the answer.

## Cathy Gadecki

Senior Director, Juniper Networks

[www.junipernetworks.com](http://www.junipernetworks.com)



# Application experience is the new Uptime

USERS EXPECT to experience each application interaction with fast response, without interruption, and in assured security, wherever they are and whatever the time of day, no exceptions. Application experience is therefore the new uptime. IT leaders who deliberately guide their organization's transition to this modern mindset will not only preserve business competitiveness but also increase agility in serving customers and other users. Move Away from Boundary-to-Boundary Design and Operation A first reset is the traditional architectural practice to compartmentalize infrastructure into subsets of operations running boundary to boundary. While breaking up design and operations ownership can be helpful in some ways, the method severely limits the tracking and assurance of user experience that necessarily exists across these boundaries. For experience to be the primary measure of success, IT will need to consider how the architecture affects user experience end-to-end—from access through the network to the workloads, both those in the data center as well as those running in the cloud. This change necessarily requires full consideration of operations end-to-end, including needed visibility and control.

### Embrace Automation and Programmability

Automation and orchestration platforms are essential to support administrators in operating the infrastructure end-to-end with consistent policies and workflow. Additionally, SRE/DevOps and SecOps teams require rich programmability of the infrastructure to support the required business agility in applications development and rollout. Meanwhile, architects must also demand that these new capabilities come without added NetOps complexity nor lock-in to monolithic infrastructure.

### Add AI-Driven Assistance

Also assisting operators in serving happy users is AI-driven,

decision-making tools for rapid identification, triage and resolution of issues. For example, systems are becoming available that unify network and application-level telemetry, along with AI-driven insights, to provide a knowledge graph that describes both application and network behavior in a holistic way.

These tools move far beyond today's views of network data and application data in discrete silos and coarse granularity. By describing application and network behavior together via advanced AI-driven analytics, the experience becomes the measure for dynamically making resourcing decisions and providing operational recommendations to support teams. Along with the necessary software is a requirement for switches and other devices to support microburst monitoring.

### Evolve to Connected Security Designs

Security is an essential aspect of the application experience. Here also, administrators are getting a big boost with AI-driven, cloud-based systems for the threat aware network. Enterprises with critical data to protect are adopting a connected security approach to safeguard users, applications and infrastructure by extending security to all connection points across the network and cloud.

By enabling software capabilities in existing devices, they become policy enforcement points, and the network becomes a distributed security blanket that stops threats at the point of attack and performs rapid and automated threat remediation. In the data center and cloud, application experience is the new uptime. IT leaders who successfully guide their teams in acquiring and implementing this new mindset will provide the digital fuel for their business success.



**Andrew Lavery**

Data Centre Solution Manager, Eltek

[www.eltek.com](http://www.eltek.com)



## Sustainable data centers start with new designs

21st CENTURY demands require a new breed of data centres that must be highly adaptable and exceptionally efficient. Powering the platforms we now rely on to navigate our personal and professional lives in a digital era. Solutions for critical infrastructure require seamless scalability to support the unprecedented growth models of tomorrow. Data Centres must also complement Energy usage, supporting the Global community through sustainable and strategic use of our Power resources.

In the 1880's two geniuses Tesla and Edison argued about power domination, AC vs DC and "War of the Currents" is still applicable in today's arena. Green Micro Grids include Li-ion Battery Energy Storage Systems so this debate remains very current.

Grid: Rack power conversion generates losses; seven is typical for AC distribution with only three for DC. LV directives dictate safe working thresholds of 1000Vac or 1500Vdc within their standards. Hybrid dual topology with renewable storage simplifies installations and creates enhanced functionality for multiple applications.

Moore's Law 1965, has been used in the semiconductor industry to guide long term planning and set targets, typically doubling capacity every two years. Advancement in digital electronics are strongly linked, like memory, price, even pixels in a digital camera. Often described as a driving force of technological social change even productivity and economic growth. Operator's need to pay more attention to IT energy efficiency, review analysis carefully as the slowdown in Moore's Law is creating new challenges. Some Data Centres we see today are designed as

they were 50 years ago for teletype machines and monolithic mainframe computer systems and the design of these facilities still reflect legacy computing environments. We don't need that anymore.

Legacy installation gains in mechanical and electrical efficiency have actually stalled. Studies still show that of the 65% of the power used by IT processes only generates 7% output. A consequence of aging equipment and process inefficiencies. Surely it's time for operators to pay attention to IT energy efficiency and remove idle servers. Topology becomes diverse when you review the various sectors, Telecoms, Local Edge, Collocation, Enterprise and Hyperscale. It would be wrong to define winners / losers as each market has different legacy and expansion challenges.

It's a matter of designing systems to be dynamic, able to handle different densities, variable loads as well as Hybrid options. Creative solution must achieve greater business values by reducing costly infrastructure even resources like IST. Utilize assets to their fullest extent possible, ensuring no idle systems consume electricity. System performance, availability, response time, capacity, relative efficiency form new algorithms within Machine Learning. These scalable systems maximise capability and allow for seamless transition of servers if anomalies are detected. Using these adaptive data centres we can support high, mixed, and variable density without stranding capacity or other inefficiencies. Today's expectations are to improve efficiency, rack density, utilise AI and reduce latency to 1ms. We've seen improvement in Power Usage Effectiveness and adoption of Data Centre Infrastructure Management providing real time, historical analytics, trends even site customisation functionality. Traditional racks have grown to 30kW, immersion



exceeding 100kW with others providing sealed liquid alternatives. Infrastructure is available as pre-assembled skid mounted / containerised offer, even halls are now delivered as systems. Lego block style solutions pre-built will negate site constraints that we should anticipate. Manufacturing variants to suit legacy users in multi kW / MW combinations are available for AC or DC topology. As is the ability to combine both within a single solution even transition power changes over time.

New technology within Power Conversion combines 3 functionalities, a rectifier, an inverter and a static transfer switch in one bidirectional power module. Creating new power flow architecture and new ways of designing systems. Meeting the needs of many with AC and DC outputs in a modular solution. Say goodbye to Automatic Transfers switches, load step issues with generators even system tests without isolating mains. Also

futureproofed for renewables as such wind, solar and storage are catered for.

Finally let's touch on metrics, Data Centre Infrastructure Efficiency [DCIE] is used to calculate the energy efficiency derived by dividing Information Technology equipment power by the Total facility power, isn't that more relevant.

PUE is the inverse of DCIE, describes how efficiently a data centre uses energy; specifically for computing equipment [in contrast to cooling and other overhead]. ISO/IEC 30134-2:2016. Anything that isn't considered a computing device in a data centre (i.e. lighting, cooling etc.) falls into category of facility energy consumption. I don't set standards yet know metrics, hate losses and love efficiency. Each device even conductor represent losses grey space.



The digital infrastructure information platform



[datacentre.solutions](http://datacentre.solutions)



## Eli Brovashov

Director of Communications, Huawei Technologies

<https://e.huawei.com/en>



# Two types of converged architectures

**1. Horizontal convergence** refers to full life cycle data storage management. Production storage supports transaction data processing, and can be expanded to analytics storage to support core analytics services. Backup storage also supports analytics scenarios, and primary storage provides hybrid cloud backup and tiering to migrate cold data to the cloud. In addition, the lightweight backup and management features are added to the production storage, to support more storage scenarios of both hot and cold data storage throughout the life cycle.

**2. Vertical convergence** refers to vertical optimization of data processing and storage. Transaction data processing is integrated with production storage to improve performance and reliability, whereas data analytics is combined with analytics storage to improve analytics efficiency.

- At the storage layer, the storage architecture is redefined to integrate multiple storage services, such as block, file, object, and HDFS, to eliminate data silos and facilitate diversified data storage and sharing. In addition, the multi-protocol convergence technology enables one piece of data to support analytics requirements of multiple services, such as databases, big data, and AI. This reduces unnecessary data flow and makes analytics more efficient.
- At the computing layer, multi-engine converged analytics and diversified computing are centrally scheduled to simplify mass data processing.
- At the management layer, AI is integrated into the life cycle management of thousands of nodes, implementing intelligent O&M covering resource planning, service provisioning, system optimization, risk prediction, and fault location.

## Heterogeneous computing convergence

With the large-scale use of AI and machine learning, new generations of intelligent data infrastructure must support heterogeneous computing chips such as GPU, FPGA, and ARM-chips. These chips can provide the computing power required to match data in multiple forms, and provide high-performance computing resources for the center, cloud, and edge, enabling applications to run more efficiently.

Some data is directly processed in the disk of a storage controller, and calculated, valid data is returned to the computing layer. In this way, the data volume that moves between the storage layer and the computing layer can be reduced to a maximum extent, and computing efficiency is improved.

The decoupled database and storage architecture allows database compute and storage resources to be flexibly configured and expanded based on service requirements. Resource matching is more accurate and reasonable.

Data is stored using different protocols throughout the life cycle. With the protocol convergence technology, data is logically centralized. That is, multiple applications and data sources can be accessed. Plus, the open data access framework allows the access of third-party data sources.

Standard interfaces are provided to upper-layer applications and clients to support multiple services, such as block, file, object, and big data. This eliminates application silos from multiple storage systems built with traditional data infrastructures. The unified platform uses one piece of data for the analytics of multiple services.

## **KAO DATA**

**Spencer Lamb**

Vice President, Kao Data

[www.kaodata.com](http://www.kaodata.com)



# Resilient digital infrastructure for HPC and AI

WITHIN THE DATA CENTRE, an accelerated level of processor, storage and infrastructure innovation has taken place in recent years, ushering in a new era of high performance computing (HPC) and artificial intelligence (AI).

At its core, HPC and machine learning workloads associated with intensive, algorithmic deployments such as deep neural networks, machine vision and natural language translation require compute at industrial scale. Data centre operators are increasingly having to accommodate the power and cooling demands of supercomputing clusters, additional node-to-node interconnectivity, rack densities of 30-60kW and beyond, and GPU-powered installations.

This has left many legacy data centres operating infrastructure that would have been sufficient 5 years ago, now behind-the-curve. Older, legacy data centres are not easily repurposed, and today many can of course try to replicate a previously successful design. However, even then, a facility may not necessarily be a suitable environment for HPC or optimised to support its high power and cooling demands. Here a flexible and forward thinking approach to design and build is required.

The Open Compute Project (OCP) is well known for its reimagining of Information Technology, making it more efficient, flexible and scalable. What is less known is its involvement in data centre design, working with OCP-Ready organisations to create a homogeneous shell that does for data centres what the methodology has done for IT hardware. Kao Data was proud to become the first OCP-Ready data centre in 2018 outlining our commitment to providing specialist infrastructure for HPC and AI, and offering all enterprises hyperscale levels of architecture.

For HPC and AI, location is crucial and today, the UK Innovation Corridor is a digital artery pulsing between Cambridge and London, providing tributaries that link into UK life science, pharma, fintech and research organisations – all frequent consumers of HPC. Based within its heart is the Kao Data campus, providing 40MW of IT capability, space to scale and a well-connected home for latency sensitive, demanding HPC and AI applications.

Its customisable architectures provide a flexible platform for customer's changing requirements. Whether at cell, row, pod, suite or custom-built data centre level, the Kao Data OCP-Ready facility is designed to the highest levels of technical excellence, delivering the adaptability to evolve with the next wave of HPC challenges.

And, crucially, in the environmentally aware world we now live in, all of this is powered by 100% certified green energy, and supplemented by industry-leading efficiency measures such as refrigerant-free indirect evaporative cooling resulting in an impressively low PUE of >1.2 even at partial loads. These, and other measures, enable us to reduce our carbon emissions by more than 80,000 tons per year, without reduction in capacity or capability.

Businesses utilising HPC or intensive forms of AI need a data centre partner that understands their challenges, and has the technical capabilities to cater for their workloads. Today, organisations utilising HPC and AI need look no further than Kao Data – a data centre engineered for innovation at industrial scale.

**Emily Ain**

Data Center Solution Architect, Siemens

[www.siemens.com/datacenters](http://www.siemens.com/datacenters)



## How forecasting can deliver incredible value

THERE'S A LOT that goes into managing a data center. Rising complexity due to the thousands of applications and devices plus the absence of knowledge-transfer between stakeholders makes daily management an operational challenge. Add planning for future deployments and tenants to the mix and the challenge can easily become a costly, haphazard nightmare. Data centers are critical components of modern enterprises and should be used to drive business, not hinder it. In these environments, it is not only essential to have continuous monitoring in place, but also to effectively manage capacity and discover hidden potential.

The usual answer to these challenges is the implementation of a data center infrastructure management (DCIM) platform. DCIM software generally focuses on providing real-time monitoring and a consolidated view of data with current inventory. We cannot take these functionalities for granted, but the true challenge arises when insights into future states are required. Knowing that in this moment the data center is using 80% of installed capacity is valuable, but what about a few months from now? How does knowing current capacity help when we need to know if we can support new tenants in the future, while minimizing overprovisioning?

The answer is forecasting, which uses past, current, and future operational data to monitor and predict the impact of data center deployments on capacity metrics, including but not limited to power, space, and cooling usage. Today's DCIM solutions need to do more than asset tracking – they need to support business objectives and reduce risk in decision-making for all stakeholders, whether their focus is facilities, IT, or the overall health of the enterprise. Leveraging a forecasting-enabled DCIM can help data centers manage their existing capacity

much more precisely, and address two prevalent challenges: over-provisioning and tenant management. Lack of insight into existing and future installations has made overprovisioning the only option available for data center managers who need to maintain service continuity. Overprovisioning cannot be solved with a simple asset inventory, and not only represents a heavy capital investment but also higher administration, staffing, and maintenance costs. Imagine instead being able to predict when capacity will be released, measure the impact of new tenants and projects on existing infrastructure, and plan effectively for incoming clients and installations. By leveraging forecasting to track capacity over time and see how future deployments will affect installed capacity, decisions regarding expansions can be accelerated, costly overprovisioning can be avoided, and existing infrastructure can be used efficiently and cost-effectively.

Tenant management also presents a challenge. Without insight into how tenants are utilizing available capacity over time, it's impossible to accurately predict whether you'll be affected by a capacity shortfall, and therefore be able to respect SLAs. In addition, not knowing when the biggest contracts are going to expire and their effect on the data center space blocks the use of valuable capacity. It's important to plan how to use and, more importantly, sell space that will become available. Complexity is now manageable. In a market that's projected to grow by \$284 billion by 2023, complete visibility into data center operations is critical for business continuity, and represents a large opportunity to eliminate waste, increase utilization, and save money. When it comes to capacity planning, visibility in real-time is not enough to combat the uncertainty and lack of information that prevent effective decision-making. Luckily, effective solutions such as forecasting are already available to transition us to the next phase of data center management.



## Gabriel Bonilha

EMEA Professional Services Manager

[www.vertiv.com](http://www.vertiv.com)



# Time to operate with remote services

HUGE DEMANDS for working remotely have led to important discussions around safeguarding data centres and critical infrastructure remotely. This is to ensure the physical safety of key workers and avoid nationwide downtime. The need for remote management of digital infrastructure is going to keep growing as the risk of on-site visits remains high and requires additional personal protective equipment (PPE). For now, each technician who must physically attend the site will need to be fully prepared and equipped.

The more we can do remotely, the better, considering the need to minimise on-site personnel. Therefore, changes are being made to data centre planning strategies to pave a new, more agile future for the market.

### Living in a smart world

We're moving to the Smart Machine Age; from grids, cities, metering, logistics, production and assembly lines. As more parts of our lives go smart, we'll be using more automation and artificial intelligence. This will be generating tons of data, so analytics will become an increasingly important part of data centre operations. Having actionable insights will enable continued optimisation of the data centre and allow managers to make more informed decisions regarding maintenance and remote solutions. In addition, a dramatic upsurge in demand of cloud services to ensure remote access is likely to help enterprises become more resilient to unexpected market shifts. Naturally, to be less prone to the effects of these shifts, businesses will want to be sure to partner with technology companies who offer them the best level of protection and those that can do so with remote services. With an increased demand for delivery of cloud services, our industry is also likely

to see hyperscalers and applicable colocation providers begin to distribute these services at an even faster pace. With a more distributed user base, this means a more spread traffic profile, and the need for more access and edge capacity. Being well protected in difficult times means that they will be able to keep up with the pace of user demand, which calls for concepts like prefabricated modular (PFM) data centres.

### Keep your edge with PFM

The latest Omdia report specifically addresses the role prefabricated modular data centres (PFM) will likely play in helping telecom operators deliver future edge infrastructure. It highlights how the PFM market forecast to grow from \$1.2 billion in 2018 to \$4.3 billion in 2023 thanks to the growth of edge computing, as well as overall growth by cloud service providers.

Building data centres may become complex and costly endeavours. This is where PFM comes into play, offering a streamlined approach to data centre construction. Rather than conventional – or stick builds – which rely on on-site construction, PFM involves assembling modules of data centre infrastructure in factories and then shipping them to the site. This means that they can be built fast, to the exact specifications required by telecommunication providers looking to get ahead.

### Achieve remote resiliency

Currently, not all sites are equipped for remote services, but when they are, organisations should lean heavily into those capabilities to minimise site visits and make necessary visits more efficient. With remote monitoring and services, you can collect and analyse data, and in many cases, initiate actions remotely to address a potential issue and eliminate the need for a site visit.



## Andy Mullings

EMEA Sales Manager, Enlogic

[www.enlogic.com](http://www.enlogic.com)



# Cool enough for Formula One

IN FORMULA ONE RACING, microseconds can be the deciding factor of winning and losing. The critical component is speed: from precision aerodynamics of the vehicle build to the traction control and pressure of each tyre. But what about the speed of data? Data processing time is just as much a key factor in the fast-paced world of Formula One.

During a race, every Formula One car is monitored and managed at their HQ via an onsite a Track side mobile Data Centre. The decision of that data management is of utmost importance: the speed in which data is delivered to the racecar is parallel in importance to the speed of the car.

Whilst monitoring the car during the race, if the team in the control center needs to shift the settings in the car, a data lag time of 5 seconds can mean the difference between winning and losing. Which is why a top UK Formula One racing team relies on Oper8 to provide their mobile Data Centre's at every race.

### Racing with PDUs

At the Formula One race headquarters in the UK, all of the engineers evaluate and maintain the operations of the on-track cars. It is critical that the control centre maintains the dynamic response in order to effectively manage real-time diagnostics and telemetry. This data travels to and from the prefabricated DC at the Track. Terabytes of data are transmitted during a race and hundreds of changes to the car in real time.

Losing power or communication to the mobile DC is not an option. It would mean they've given a 5-second advantage to their racing competitors. That's why Oper8 only trusts Enlogic iPDU's to power their racing operations.

### Why Enlogic

In 2013, Oper8 ran an independent evaluation of the top iPDU manufacturers to determine which iPDU could withstand the extreme conditions that were required for the success of Formula 1 Racing teams.

The test was just as extreme as Formula 1 Racing: the iPDU's were run at 100% load, in temperatures between 44-55°C, for 1-2 weeks per rack PDU. The results were unparalleled: "Raritan, Schneider Electric, Vertiv all failed the independent test. One of those brands literally started to melt. The only PDU that withstood the test, was the Enlogic iPDU."

Since that test, Oper8 has only used Enlogic iPDU's; "In the past 6 years, we've put thousands of Enlogic iPDU's into the market and have never had a single failure. That incredible success rate speaks for itself."

With the Enlogic iPDU comes the speed and reliability that is synonymous with Formula One. Enlogic has the only truly hot-swappable Network Management Card. In the world of high-performance racing, uptime is paramount. The last thing you want is to worry about is failure in a power source. This is why they only trust their power to come from Enlogic.

Using Enlogic iPDU's, anyone can harness the power proven fast enough to win with the UK's top Formula One racing team.



# DIGITAL CHANGE MANAGEMENT SUMMIT

30.09  
2020

[www.dcmsummit.com](http://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

# ENERGY OPTIMISATION

Environmental issues tend to divide like few others. The glass is either half empty or half full. The data centre industry is no exception. The optimists are confident that energy resources are infinite; the pessimists predict black outs in the not so distant future. Encouragingly, the industry as a whole has made great progress in addressing the issue of energy optimisation, with initiatives at everything from individual company level upwards. But there's more work to be done, as digital transformation places ever more demands on the IT infrastructure that is its very lifeblood, and the data centres which are the foundation layer below this.



Launching a revolution in server design  
powering a Net Zero data center.



DOWNLOAD







# Upgrading to **greener**, more sustainable data centres by using modern technology



Aligned with the United Nations 2030 Agenda for Sustainable Development, large operators are ensuring that their future data centre infrastructures are

increasingly greener and sustainable. For operators sustainability and profitability are no longer competing interests.

**BY HENRIQUE CECCI, SENIOR DIRECTOR ANALYST, GARTNER.**

DATA CENTRES are energy-intensive and energy consumption is one of their biggest environmental issues. A green data centre gleans the maximum amount of production from the minimum amount of materials and energy, without compromising performance.

In the past, this was difficult to do, however, with annual developments in technology increasing accessibility and affordability it is now a possible reality.

Operators should consider two ways they can leverage cost-effective and green energy production and storage when upgrading their existing data centres



## Modern solar systems

Electricity is the fuel that runs data centres and major operators say slashing their CO2 emissions is a primary goal - solar panels may be a way to reconcile these two facts. Solar panels were previously thought to be too inefficient for data centre usage, however as technology has developed that it is just not the case anymore.

Data centres usually take up a large footprint, which means there is plenty of roof space for solar panel installations for large amounts of passive energy generation – standard solar panels produce around 150-200 Watts per square meter in good sunlight. Multiply that with the size of your roof space and you have a good guess at what you can potentially generate.

Since panels are installed in underutilised space, whether that's on the roof of the property, or other open space such as a above a carpark, they won't disrupt the operation of the facility.

## Lithium-ion batteries

While standard data centres will most often use lead acid battery cells, major operators have already made the move to the more modern Lithium-ion batteries, like the one you'll find in your smartphone.

The cost of Lithium has dropped substantially in the past 18 months (and 87% since 2010) thanks to larger order sizes and battery electric vehicle sales making their previously prohibitive price point much more affordable.

When testing the older type of battery, Lead Acid, we see it struggles to compare with Li-ion technology on most fronts, due to their huge size, poorer efficiency and need to be replaced much more regularly – this difference in power usage effectiveness is a core reason many data centers have already made the switch.

From a sustainability perspective, Li-ion batteries don't contain hazardous chemical materials and are increasingly easier to recycle as waste-management operators become more familiar with the procedures. When coupled with greater energy efficiency, they are becoming an even more attractive choice for critical IT applications, both now and in the future.

To begin a green transformation of your data centre, establish an environmental impact benchmark for your data centre by reporting on the current environmental impacts, such as carbon emissions and water consumption, and seek to improve the system from there.



# How the data centre industry can deliver on its green ambitions



Digital services are being relied upon now more than ever. As the world deals with the impact of the Coronavirus crisis, tools for remote working, streaming, commerce, online entertainment and healthcare have become critical to keeping the country running. In some European countries, internet use has been up as much as 50 percent since the pandemic began.

**BY DAVID WATKINS, SOLUTIONS DIRECTOR FOR VIRTUS DATA CENTRES.**

FOR COLOCATION PROVIDERS this means that there is a growing demand for additional data centre and networking bandwidth. But the importance of the data centre is not just apparent in times of crisis - the data centre industry has long been considered the lynchpin of the digital economy. Anything with “smart” in front of it inevitably has a data centre behind it - providing the processing power, connectivity, networking and storage which is required to deliver digital services.

However, there is a downside. It’s no secret that the data centre industry is a contributor towards climate change and the pressure to reduce its environmental impact is high on the priority list of governments all over the world. Today, data centres generate two percent of the world’s 50 billion metric tonnes of greenhouse gasses every year and the EU Commission recently set a “green deadline”, noting that the industry “can and should become climate neutral by 2030.”

Data centre operators must take responsibility to meet their environmental targets, and at the same time, maintain a high quality service and deliver on continually growing customer demand.

### **Making cooling more efficient**

As data centres focus on building a “greener” internet, improving the energy efficiency of cooling operations will be critical. On average, as much as 40 percent of

a data centre’s electricity consumption goes toward cooling the servers – making it the most power-hungry area of data centre management.

However, the industry is already innovating at a pace, and is fast developing ways to keep equipment cool while minimising the environmental impact. Google’s Hamina, Finland data centre makes use of seawater for cooling purposes. And, Facebook adopted a cooling system at its Lulea, Sweden data centre that uses the chilly outside air to ensure its equipment is kept at the optimum temperature.

VIRTUS Data Centres is continually looking at how to optimise cooling technology. To keep its facilities as efficient as possible, the company uses a variety of innovative design elements for greater efficiency whilst actually lowering costs; this includes air flooded data halls, utilising hot aisle containment, and cooling using a variety of industry leading technologies.

### **Harnessing renewable energy**

More good news is that the ICT sector is already world-leading in its adoption of renewables. Indeed, according to Greenpeace, in the UK 76.5 percent of the electricity purchased by commercial data centre operators is 100 percent certified renewable.

Google is the world’s largest purchaser of renewable power, meeting its requirements through power





purchase agreements; direct contracts with renewable generators that create additional utility-scale renewable capacity.

Another example is a campus in the southwestern tip of Iceland, which runs almost entirely on geothermal and hydroelectric power. The Icelandic data centre owners claim theirs to be the world's first carbon-neutral data centre, and BMW has already moved a large portion of its German clusters to the campus, and more organisations look set to follow.

In the UK, VIRTUS Data Centres ensures that green credentials are in evidence. All of the energy consumed at its facilities is from 100 per cent renewable sources thanks to partnerships with companies like Bryt Energy who generates power from wind, solar and tidal sources.

VIRTUS' LONDON2 data centre also incorporates a borehole dug at the inception of the site, using natural water sources for cooling reducing demand on the mains water supply. Combined with the local climate, and efficient cooling technology this delivers low Water Usage Efficiency (WUE) for the site.

### Switch to eco mode

The idea behind "eco mode" is simple: to reduce the

enormous amount of wasted electricity consumed by uninterruptible power supplies (USPs) in data centres. Just like using the energy-saving mode on your computer or in your car to reduce energy consumption, eco mode can reduce a data centre's servers to their lowest energy-sapping level. This setting is the most efficient and can significantly improve data centre efficiency and power usage effectiveness (PUE).

This technique is already being used to significant effect. For example, VIRTUS Data Centres already uses "smart" USPs, which are able to turn on and off according to demand, and which significantly reduce power consumption.

However, being environmentally friendly is not a binary achievement - there are "shades" of green, and some providers are doing better than others. To realise environmental ambitions, sustainability must be built into every aspect of data centre construction and maintenance.

Some data centres are already working hard to mitigate the environment effects of a technology hungry society, but there is still work to be done. The industry must work together to take action and accelerate its search for new ways to minimise carbon emissions and increase energy efficiency.

## Richard Venga

Business Planning and Strategy Manager,  
Mitsubishi

[www.mitsubishielectric.com](http://www.mitsubishielectric.com)



# Increasing sustainability and operational flexibility in critical cooling

OUR MODERN WORLD is reliant on IT to an extent never seen before and, as a result, there is an ever-growing requirement to keep business-critical computing at the right temperature to ensure it functions effectively.

This comes at a time when we must also reduce energy use and look for more sustainable, IT cooling solutions. For data centres and server rooms, 24/7 cooling is a must have, not a nice to have, and ensuring that it is done in the most environmentally way possible must be the focus going forwards. And this is where the latest cooling-only VRF (Variable Refrigerant Flow) air conditioning can deliver the innovation, performance and reliability needed, according to Richard Venga, Mitsubishi Electric's Business Planning and Strategy Manager for Central Plant: "We've developed a VRF system that offers a more targeted cooling solution, especially where the cooling required per rack exceeds 15kW and you risk hot spots within your server racks."

The new Multi Density systems combine the efficiency, industry-leading quality and simplicity of Mitsubishi Electric's renowned City Multi VRF range, with high performance, close coupled cooling solutions for high density data rooms.

By moving the air conditioning closer to the computer rack, a more precise delivery of inlet air and a more immediate capture of exhaust air is possible, ensuring the most demanding IT systems are kept at optimal conditions, reducing the risk of outages, while maintaining optimal efficiency.

"These inverter-driven systems are ideal for small and medium

size IT environments with hot spots," adds Venga, "They also offer long pipe runs and are about as plug and play as it is possible to get for an IT room."

The indoor cooling package consists of multiple coolside, close coupled, in-row cooling solutions, connected to a City Multi VRF outdoor unit. The result is a full inverter multi-split system, designed according to the best quality standards and dedicated to the most reliable IT environments.

With a wide operating range and self-developed inverter compressor technology, the systems benefit from M-Net control making them suitable for application in a modern IT infrastructure that is typically characterised by high thermal loads.

With up to 8 indoor units able to connect to one outdoor unit, the system is particularly suitable for high density racks and blade server cooling in data centres with hot spots, with the in-row technology placing the air conditioning unit directly within the rows of racks to cool these localised heat sources.

An Active Redundancy function ensures that the heat loads are balanced among the units (including those units that usually remain in stand-by) according to the actual requirements of the IT infrastructure, leveraging on the multi-unit configuration of redundant systems.

"These innovative solutions really have been built to offer flexibility of both design and operation," explains Venga, "They are able to offer customers a multitude of configurations to match their desired level of reliability, in line with Tier III and IV design topologies."



## Mike West

Technical Manager, Aquacooling

[www.aquacooling.co.uk](http://www.aquacooling.co.uk)



## Invest in immersion

COMING FROM a process cooling background, for me the most exciting opportunities within the data centre industry revolve around immersion technology. Against the current COVID-19 landscape, we keep hearing that we're in unprecedented times. One thing is certain, with people working from home and the explosion of on-line activities, data volume is certainly unprecedented too.

By embracing liquid cooling, we address energy optimization which in turn impacts the theory behind the design & infrastructure of the facilities themselves.

Managing more and more data increases the need for effective and energy efficient cooling as required densities increase. Immersion technology – put simply, submerging servers or components in liquid – enables systems to run at higher temperatures. You gain higher density with less physical equipment needed – here you're talking around 100kW per rack.

This is key in situations where real estate comes with a hefty price tag – where doesn't it! - or in densely populated areas where space is restricted. A conventionally built data centre might take up several floors of a high-rise city centre building, but based on utilizing immersion tech, this can be condensed into just a few rooms on a single level of the same building.

With immersed technology, permissible temperatures within your facility also increase which opens the door to an array of highly energy efficient process cooling options based on using water. You can also look to move away from traditional chillers and move to non-mechanical "chiller less" cooling systems. This includes dry coolers and adiabatic systems, and if design allows

the former then you gain further from zero water evaporation – reducing your operational costs even further.

I think it's fair to say that across the industry the take up of immersion-based data centre facilities has been slower than initially anticipated. For me this is the biggest challenge, getting the marketplace more comfortable with, and accepting of, immersion technology.

For businesses that have taken that leap of faith the benefits have been dramatic, and their facilities are future proofed for the years ahead. I think some of the reservation from the sector has perhaps been the CapEx investment, you obviously can't just drop a water solution into a DX set-up, you need different servers and new infrastructure.

That change of physical architecture is an initial obstacle, so perhaps it's an easier concept for a new build developer to embrace, but once it's overcome the benefits are huge.

There's also the idea of using water alongside highly sensitive technological equipment, which can cause nervousness amongst some. For me that's an easy issue to resolve, if we combine a Leak Prevention System then clients have the benefit of a guarantee of 100% leak proof operation. For most when they see this technology in action, water becomes much less the enemy very quickly.

Immersion technology isn't going to be the best solution for everyone, but for those that are able to make the initial investment the result is a data centre facility designed, built and operating for the new times ahead.

**Alessandro Bruschini**  
Infrastructure Manager, Aruba  
[www.datacenter.it](http://www.datacenter.it)



## The challenges of energy optimisation

AS THE WORLD becomes increasingly digitalised, data centres are emerging as an important source of energy demand. Data centres require high levels of electricity in order to run their servers, backups, and power cooling infrastructure.

The temperature requirements for sustaining and operating a server below 26 degrees celsius can comprise of up to 40% of electricity usage in a conventional data centre, which leaves a significant carbon footprint. When you take into account that by 2025 one-fifth of all the electricity in the world will be consumed by data centres, vendors and operators alike are now realising that there is a real need to tackle the problem of reducing the environmental impact of data centres.

The discussion around energy efficiency is not only about how power is used, but also how it is generated. One view is that many different voltages are used in the transmission and distribution of energy, meaning that energy is transformed several times, with each transformation resulting in some efficiency losses. Data centres that focus on the optimisation of their energy consumption make a significant saving on costs – which can be passed on to customers. As an example, the cooling system of the Global Cloud Data Centre (IT3) campus in Milan means that we can drastically reduce the quantity of energy we use, reducing the overall cost and making better use of the energy available.

However, going 'green' is not that easy. Energy produced through green initiatives is more expensive than standard

energy. A sustainable alternative must not only be green, but also consume less energy, so that organisations are not out of pocket. One way to achieve this is by replacing traditional cooling solutions with geothermal systems that make use of the cold water found underground.

The entire air-conditioning system for the data rooms across Aruba's Global Cloud Data Centre campus is extremely energy efficient, thanks to the geothermal system that makes use of the cold water found underground. 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.

The progression of energy efficient and environmentally friendly solutions also requires industry-wide effort. Vendors and operators alike have made big strides in demonstrating that they can implement energy efficient solutions and practices, but there needs to be a concerted effort in tackling this on a global scale.

One way to do that is industry initiatives such as the 'Green Grid' and the 'European Code of Conduct' for data centres and other industry standards and certifications for energy efficient operations. Although data centre power consumption will continue to be an issue in the future, the most important aim is to consider how to achieve maximum efficiency and energy savings on a wide-industry scale, whilst reducing costs and protecting the environment.

**Maikel Bouricius**  
Marketing Manager, Asperitas  
[www.asperitas.com](http://www.asperitas.com)



## Sustainable anywhere and on any scale

ASPERITAS Immersed Computing® is responsible for delivering highly sustainable and energy producing datacentres for emerging digital technologies. To address global datacentre challenges regarding sustainability, efficiency and flexibility Immersed Computing® has been developed and goes far beyond just technology. It embraces the most efficient model for operating IT, total liquid cooling or so called immersion cooling, which reduces the energy footprint of IT by eliminating overhead energy. The technology allows for warm water cooling which provides even more energy savings on cooling installations. Moreover, the ability of the liquid to capture all the IT energy combined with warm water cooling enables solutions ready for heat reuse.

### Challenges

One of the main global datacentre industry challenges is sustainability. With Immersed Computing®, we address that challenge, as it's actually possible to reduce the energy footprint by up to 50% by using this approach. Presently, datacentres are consuming 4% of the global energy supply, using energy for IT and cooling. Each datacentre consistently churns out warm air which is too low in temperature to efficiently distribute and transport. The medium currently commonly used by datacentres to cool is air. Changing this medium for the much more efficient medium (liquid) is changing everything for datacentres. It will support addressing the global challenges: preparing datacentres for a future of high density and performance hardware, facility efficiency anywhere and sustainability.

### Developments

Total immersion cooling ensures no oxygen touches the IT components, thus preventing oxidation. Thermal shock is

greatly reduced due to the high heat capacity of liquid. The immersed environment only has minor temperature fluctuations, greatly reducing thermal expansion stress on micro electronics. These factors eliminate the root cause for most of the physical degradation of micro electronics over time. All of the Asperitas' datacentre immersion cooling solutions are driven by natural convection to circulate the dielectric liquid for heat transfer.

Consequently, pumps are not needed for this purpose, resulting in solutions whereby moving mechanical components and energy overheads are not present. All of the solutions we develop are optimised for optimal flow and heat transfer. Asperitas Immersed Computing® technology allows for climate independency and optimal heat reuse scenarios.

Similar to the previous, smaller solution launched by Asperitas, the newly introduced AIC24 21" ready system makes use of the natural convection driven fluid circulation within. This unique approach enables immersion cooling to be both climate independent (free cooling up to 48 degrees Celsius) and fully heat reuse ready (more than 98% energy transformed into hot water). Each system comes with a smart sensor-based monitoring and management control system, which is Redfish compatible for full integration in OCP standard environments.

### Opportunities

By making the energy available for reuse for heating, Asperitas 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 (100% energy reuse ready).



**Siobhan Ellis**

DPM & PM, Bamboo Systems

[www.bamboosystems.io](http://www.bamboosystems.io)



## Preventing data centers from being an existential threat

THERE is a largely unseen price for the rapid integration of technology into our lives. While experts don't always agree on the specifics, it's not uncommon to see reports of 2 to 3% of the world's power, and up to 2% of the world's greenhouse gasses are produced by data centers, collectively. The discussion is not if this will grow, but by how much. Consequently, the world is paying attention, and data centers are becoming a political issue. At Bamboo Systems, we believe that the issue of the ongoing growth in data center power usage is being approached from the wrong standpoint. We need to fundamentally rethink the problem of power consumption by taking into consideration modern software design to define the supporting hardware architectures.

### A good start

Significant steps have been taken in reducing the rate of growth in power usage. Total data center power consumption is typically a direct result of its computing power, cooling requirements and power conversion, with an ever decreasing share going towards actually driving compute. Enter the Green Grid, now absorbed into ITI, whose aim was to drive more energy efficient practices, and so developed the concept of PUE (Power Usage Effectiveness) - a ratio of data center power consumed to power used for compute with the best possible outcome of a ratio of one. What will it take to achieve a PUE of one? The closer it gets the more expensive it becomes, where the cost to achieve outweighs the financial benefit gained. Some solutions include data centers located in cold climates, where the heat they generate contributes towards warming buildings, or designing servers that can run at high ambient temperatures in an effort to eliminate air-conditioning units. These methods are either less

effective or do not work at all in hotter climates. As hyperscale data centers chase cooler climates they move further away from major centers of population, and so latency is increased.

Consolidated compute in the manufacturing process and AI/ML are all extremely sensitive to latency which is, in turn, driving edge computing. Its growth will no doubt impact data center power usage as edge data centers tend to be smaller and, as shown by a study in Germany, smaller data centers are not able to achieve all of the benefits of hyperscale data centers. We've had some accidental help along the way to combat power consumption. Moore's Law originally stated that every year we would see a doubling of transistors in the same space. The law was then revised to doubling every two years. It has been a guide to produce ever more powerful CPUs that do not increase power consumption. When combined with a focus over the last few years on cores, memory controllers and accelerators enabling increasingly faster single threads, this has enabled more dense computing for the majority of compute tasks within the same physical footprint. However, Moore's Law is generally acknowledged to be coming to an end. We are already seeing the impacts as Intel struggles with its 10nm fabrication plant, resulting in the last major change in transistor density to expand to more than two years.

### Waste not want not

Until a few years ago, these faster CPUs were dedicated to a single service. As that CPU became faster than the service required, often the CPU and memory were left underutilized. The ubiquitous rise in virtualization has taken that spare capacity





and enabled consolidation and reduced workload costs significantly. However, running those hypervisors is an overhead and consumes processing power. Hence the emergence of containers. Micro-services and containers are parallel type software architectures that are perfectly positioned to use multi-core architectures rather than a single very fastest core which is then time sliced. Yet now the software often finds that it is competing for memory through a limited number of memory controllers or is I/O bound due to sharing resources connected to a single CPU. We need to continually ask ourselves, where is the bottleneck? Is it the actual pure compute that is the bottleneck? Is it being slowed down by storage, memory controllers or even the network?

Are we expending a great deal of corporate/personal energy and power to achieve an effective and perfect processing latency when all we need are better balanced servers?  
We've had plenty of evolution, maybe it is time for a revolution?

### Time for a rethink

At Bamboo, we believe the measure of a successful data center should be the total workload delivered per kilowatt-hour; that is a true measurement of efficiency. This would support the changes needed to achieve net zero energy data centers. To do this, we decided to cure the problem, not the symptom. We have gone back to basics and redesigned the data center server from the ground up to be balanced and energy efficient, while still meeting the demands of modern highly parallelized microservice-based software. Why do servers typically include VGA and USB ports in the data center? They are there for people, and because the BIOS expects them, but when was the last time a person used

them on your server? All of those unnecessary components should be stripped away and high throughput must become the priority. Ramping up the clock speed will certainly provide more performance, but at what cost in watts consumed?

As you approach the maximum clock speed the power required, and thus heat produced, climbs logarithmically. A CPU running with overall less GHz means fewer watts and thus means less heat. That translates to less power for air-conditioning needed and a higher density of servers is possible. A number of systems running in parallel at lower power, with higher bandwidth for network, storage and memory, will provide a much higher density throughput by a factor of 10 when compared to traditional server architectures. Isn't it better to have infrastructure that delivers more in parallel in a timely manner, but uses less space and electricity and thus reduces cost while still delivering the workloads needed?

This approach would work also in hot climates, for the edge and, finally to help heat those homes in Sweden, we can just run up to 10 times the servers in the same space. A change in architecture, combined with a switch to low power processors, well balanced memory controller ratios and storage and network throughput will enable massive and revolutionary change in power consumption levels in the modern data center.

**Acknowledgements:** This work was partially funded by/in the framework of the EU project E2Data (780245)

<https://www.bamboosystems.io/you-can-have-it-all>



## Louis McGarry

Sales and Marketing Director, CENTIEL UK Ltd

[www.centiel.co.uk](http://www.centiel.co.uk)



## An agile UPS approach

AN AGILE APPROACH to UPS installation can offer datacentres a lower Total Cost of Ownership (TCO), increased flexibility and increased availability and here's how.

I often see UPS installations designed and configured for a much greater load than is actually required. However, a system which is too large wastes energy, is inefficient and costly to run. It will also cost more than necessary to maintain due to its size. Therefore, adopting constant right-sizing and a pay-as-you-grow approach, means you only pay for what you need.

Installing a true modular UPS where you have the ability to use a fully rated frame or empty carcass, provides the option to add the required number of modules to suit the actual load, plus more only when needed. All the individual modules are a UPS in their own right, all containing a rectifier, inverter, and static switch and all operating online in parallel with each other.

In this way, 'a modular approach' helps minimise the need for a large initial investment in a stand-alone solution: and datacentres can scale as they sell space. It also offers a more flexible approach.

It means datacentres have the opportunity to re-allocate resources rapidly. A colocation datacentre selling decompartmentalised space or rows or racks or kW/Sq. foot may shrink with one client and grow with another. Standard frames which can accept the same physical size and power rating of UPS modules can ensure that if one client decides to reduce

their rack space or moves on, the over and above hardware can be reutilised elsewhere.

This approach also allows the modules to be redeployed between areas or facilities based on priority, to ensure load protection and redundancy is kept at the most critical locations. It also means, if a system needs to be maintained modules can be 'hot swapped' avoiding any risk to the load (zero downtime!). This is how an agile critical power solution can support an agile infrastructure environment. Availability is maximised while at the same time costs are minimised. However, this agile modular concept relies on true modular UPS systems that are robust and of high enough quality to withstand being shipped and moved around regularly. Other brands of modular UPS are available, however, please keep in mind that not all modular UPS systems are designed to be hot swappable, with the flexibility to relocate and reutilise. As well as needing standardised frames and a robust, quality modular system, it is also necessary to manage the whole network closely too.

Implementing infrastructures that offer the ability to pay-as-you-grow and to re-deploy modules according to load requirements will require some intelligent analysis. However, with expert advice from trusted advisors who have decades of experience implementing the most efficient solutions such as our engineers at CENTIEL, it is possible to have the highest level of power protection in the most cost effective way.

For more information please see: [www.centiel.co.uk](http://www.centiel.co.uk)

**Rowland Kinch**  
CEO, Custodian Data Centres  
[www.custodiandc.com](http://www.custodiandc.com)



## Evaporative cooling and its advantages

THE TERM evaporative cooling describes a system that takes advantage of the reduction in temperature resulting from the evaporation of (in our case) water in to air.

For us as a data centre, it does several things. It gives us automatic inbuilt efficient humidity control during the winter season, which is where you get the real ESD (Electrostatic Discharge) issues. During summer you'll be reasonably humid; very rarely will you have days where you look at the temperature charts and it'll duck below 30 percent. Most of the time in the UK we sit in the 40-60 percent band and above during summer. Whereas in winter, you can get cold, dry days, where it'll plummet to next to well under 20 percent once the air has been heated up from zero to the 21 supply required within the data floors.

The big thing is efficiency. Where you are choosing to use alternatives like fresh air for cooling (compared to DX), obviously it is ambient dependent. So, for example, if it's 27 degrees outside, then that's the temperature you are going to provide internally, plus you're going to pick up a degree or two on the way in. Yes, there is a direct correlation with adiabatic cooling and ambient temperature, but for us, the big thing is the lack of requirement for electrical transformer capacity. The liquid gold for any data centre is the transformer (and backup generator) capacity, which is a massively expensive resource. Especially when you consider that it could be unused for over 95 percent of the year.

When you are designing a system, you have many considerations but one of them is peak demand. It's all very well talking about the efficiency of adiabatic vs the efficiency of a chiller, but from a systems design, you must understand what the peak requirements are.

So, for a 330kW chiller, you need somewhere in the region of 200 – 250 amps per phase reserved for its compressors and pumps. Now, unless you hit 37+ degrees ambient, you won't hit the peak electrical demand from the chillers. Typically, in this area of the world, we spend 95 percent of our time below 21 degrees according to the MET Office.

No matter how the weather turns out, you still have to have all of the available amps required for max cooling in reserve. You can't use an amp: it has to be 'held to one side'. For adiabatic, the equivalent is probably maximum electrical requirement is 15 amps. So, your additional transformer capacity becomes available for IT load and you lose none of your cooling capacity.

### Custodian Data Centres

Award-winning UK data centre provider specialising in Colocation, Cloud and Connectivity services. With 100% up-time, 24/7 technical support and state of the art facilities we are here to push the boundaries of data centre technology and provide first-class customer service.

## Marco Abdallah

Head of Engineering, Drees & Sommer UK

[www.dreso.co.uk](http://www.dreso.co.uk)



## Why we musn't lose sight of sustainability

AS AS WE COME TO TERMS with remote and home-working becoming a significant part of the 'new normal', our reliance on cloud-based technology and, ultimately, data centres is set to increase. Processing more data requires more server power which ultimately requires more energy. We are therefore obliged to ensure that data centre buildings are built as sustainably as possible — just because they are out of site for users does not mean they should be out of mind.

A data centre is a collection of thousands of smaller machines processing our data and thereby creating a lot of heat. To maintain a constant temperature in a data centre, huge amounts of cooling energy is necessary. And while it might seem to be commonsensical to simply build all of the world's data centres in the arctic tundra to keep them cool, unfortunately, the speed at which data is transferred depends on how close the centres are to clients and users; the closer they are the better, hence why so many are located in densely populated cities.

A key strategy when it comes to sustainability is to use the excess heat for heating buildings in an urban environment. This concept is already being used in cities like London, where excess heat produced by the Northern Line is today heating around 1,000 homes. The principle can be applied in precisely the same way to data centres, which can provide their excess heat to neighbouring residential buildings, thereby dramatically reducing their reliance on gas and electricity. Thereby the data centre becomes a part of an integrated urban energy concept.

This can be significantly enhanced by Smart City approaches which use data to efficiently direct the energy to where it is

needed. Freiburg City Hall, for example, is a plus-energy building that produces more electricity than it needs and feeds back excess energy into the German National Grid. Buildings like these can become local suppliers, feeding nearby high-consumption buildings — such as datacentres — with a more sustainable source of energy.

However, one interesting example of using the natural world to cool data centres without resorting to building them in Siberia is the Lefdal Mine Data Centre, a six-storey facility located in an abandoned underground salt mine in Norway. Instead of air conditioning, the mine-turned data centre uses saltwater from the depths of a fjord to cool the server racks. The water pump itself requires relatively little energy to operate and all of the facility's electricity is drawn from hydroelectric and wind power produced locally.

Expert green building management is also integral to ensuring the sustainability credentials of data centres, as was evident during Drees & Sommer consultation activities on T-Systems data centre in Biere, Germany. The facility achieved a 'Gold' certification level in accordance with the LEED system, meaning it consumes 30% less energy than similar structures.

As society becomes more and more reliant on technology and ultra-fast internet access, demand for these facilities has grown dramatically. A surge in the number of people working from home due to the COVID-19 pandemic currently sweeping the globe has only intensified that need. However, that doesn't detract from the fact that we live in an age where climate change is top of the international agenda, and there's no getting around this when it comes to energy-hungry data centres.





**Joel Kaerkkainen**  
Product Manager, Eaton  
[www.eaton.com](http://www.eaton.com)



## On the road to Net Zero

WITH THE COUNTRY WORKING FROM HOME, BT reported a 35% to 60% increase in daytime weekday fixed broadband usage. This coincides with the imminent launch of 5G which will inevitably lead to more internet-connected devices and an explosion of data.

Data centres are central to handling all of the data that facilitates our digitally-led lives, but with this comes concern about their growing impact on the environment. Data centres are notoriously touted as one of the largest consumers of energy. But what if data centres could harness this energy to increase the adoption of renewable energy on the grid?

Increased electrification means we need to balance the growing demand for energy with the urgency to reduce our reliance on fossil fuels and increase renewable adoption.

Increasingly, the UK is turning to renewable sources of energy. Sunday May 10th, 2020 marked Britain's first coal-free month since the industrial revolution. The reason this is such a huge feat is because renewable energy is volatile by nature. For example, we can't rely on the sun or wind as the only source of electricity to heat our homes in the winter. This means we still need a certain amount of fossil fuels to keep the grid stable – especially as more renewables are introduced.

This is where the data centre and its huge electrical capacity can play its part on our road to Net Zero.

Uninterrupted Power Supply (UPS) technology (UPS) can help data centres provide fast response and help to tackle issues related to low inertia by providing system (ancillary) services. This can then help to increase 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.

Beyond the environmental impact, UPS devices can also generate value for data centres. Major UPS users 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. It is this ability to reduce power failures whilst also increase renewables on the grid that puts data centres in a unique position to help. Through the support of data centre managers and the electrical utility industry, it is possible to turn data centres from energy villains, to environmental champions.

There will be backlash against the worst energy offenders, and data centres will be in the firing line. However, if the industry takes its excess stored energy to help green the grid, they're essentially enabling power companies to accelerate the adoption of green energy.

Data centres can be part of the solution and not part of the problem.



**Tony Whittle**

Head of UK & Ireland, Enel X

[www.enelx.co.uk](http://www.enelx.co.uk)



## Energy-as-a-Service: Why it makes sense

CLOUD TECHNOLOGY disrupted the old IT business models and is key to enabling X-as-a-Service delivery today. Now, the energy industry is itself embracing cloud and digital technology to tackle its own industry disruption – driven by the growth of decentralised energy resources and the urgent need to decarbonise.

Energy is high on the agenda in many data centre boardrooms. Energy forecasts suggest that the total electricity demand of ICT will accelerate in the 2020s, and that data centres will take a large slice – sometimes at the expense of public reputation. As a growing number of data centre operators look to enhance their energy profiles, adopting Energy-as-a-Service will enable them to achieve their decarbonisation goals and become better grid citizens.

Just like XaaS, Energy-as-a-Service, or EaaS, is fundamentally about giving energy consumers more control. Energy is a major cost for all data centre operators and EaaS offers a unique opportunity to gain transparency, enhance strategy and make better decisions. The choices available to transform energy portfolios present a growing list of decisions for data centre operators. Beyond simply negotiating a bespoke tariff for energy use or standard power purchase agreement (PPA), you can also choose how your power is generated; use of renewables; the pricing/payment/billing model; financing; resilience measures; distributed assets for generation or storage; the use of energy management technology... and so on.

Outside of the facility, there are other forces at play offering opportunities to improve the bottom line. The rapid uptake of distributed – but intermittent – renewable power is challenging grid system balance, which has increased the value of operational flexibility. Schemes like demand side response (DSR)

reward businesses for adjusting demand at critical times; energy users use their own assets, often supported by technology, to turn down or turn-off demand when needed. However, it's not just technology that is driving change. Regulatory frameworks are changing to support grid system needs. For example, regulators refine the rules governing DSR schemes to change market behaviour, such as delivering faster response times, longer response durations or to use low-carbon sources of generation. They also review charging schemes that are intended to support the funding of grid improvements. The transition to a new, low-carbon, technology-driven, decentralised energy landscape is more complex than it was before. For businesses looking to optimise or transform their energy use, the choices can be bewildering.

EaaS leverages outsourced strategic partnerships to develop a planned, holistic approach that puts the business's needs first. The plan may address business profitability, resiliency, and/or sustainability leveraging supply advisory; demand management; energy storage; microgrids; utility bill management; energy efficiency; and electric vehicle charging. One of the biggest barriers for any business acquiring new technology is the cost justification for capex-funded projects. The 'as-a-service' model enables services to be paid for as they are used. But EaaS's scope goes beyond cost measures. Sustainability – moving to lower carbon sources of power; quality of supply to ensure business continuity and operational simplicity are all key.

Partnering with specialists that can offer on-going advice and technology-as-a-service, enables customers to deliver higher cost savings and lower long-term risk while addressing the growing complexity – and opportunities – presented by the new energy cloud.

## Brendan Dysel

Head of Infrastructure Management, Teraco

[www.teraco.co.za](http://www.teraco.co.za)



# Energy efficiency now a prerequisite

DATA CENTRE POWER MANAGEMENT is fast becoming a key differentiator in the data centre market. Against a South African backdrop of power outages and the risk of continued load shedding, efficient power management and utilising renewable energy are now essential.

With an influx of international organisations deploying in South Africa, green and renewable energies are deemed the norm. Finding greener designs to reduce mechanical energy is an absolute prerequisite in today's South African data centre designs. With the proposed National Energy Regulator (NERSA) power increases over the next three years, these green designs will escalate the ROI of the capital equipment and provide data centre operators with better efficiencies, thus reducing the burden on clients with increased electricity prices.

He adds that issues of power uncertainty coupled with the entrance of international organisations have led to a noticeable shift in the South African data centre landscape over the past two years. Data centre designs have also moved towards accommodating hyperscale clients, but remaining flexible enough, however, to still cater for the retail colocation environment. With ongoing load shedding, voltage fluctuations are becoming more frequent and can have disastrous effects on legacy data centre builds, specifically on the refrigerant cooling systems in terms of predicted life spans.

In addition to greener and more efficient energy solutions, data centre power management provides a useful tool to understand where losses are generated within the system allowing for better

future designs and power optimisation to reduce losses linked to IT usage rather than end state design parameters. This ultimately provides vendor neutral colocation providers with a tool set to monitor and manage power usage and losses within the electrical and mechanical facility as colocation providers do not have control over clients' IT environment, which is one of several advantages of vendor neutral data centres.

Automation within data centres plays a role within energy efficiencies. Systems such as cooling, energy management, monitoring and work maintenance scheduling should be automated. The biggest consumer of power, other than the IT load, remains the cooling system and this should be the first area within a data centre to be automated as much as possible.

To effectively adopt automation in the data centre, it is best to run proof of concepts in a test environment first before deploying it in live environments.

When choosing or setting up new AI systems ensure that cybersecurity is a key selection criterion, which ensures the environment can't be access by those that are unauthorised to do so.

The advantage of a true vendor neutral data centre, is that the provider is purely focused on critical engineering infrastructure, physical security and interconnections between clients and the ecosystem. This enables dedicated skill sets to focus on the heart of the data centre, ensuring excellent service and quality in maintaining the uptime and SLAs.

## Alex Emms

Operations Director, KUP

[www.kohler-ups.co.uk](http://www.kohler-ups.co.uk)



# How to choose a UPS standby generator

ALTHOUGH POWER CUTS in the UK are unusual, they are not unknown. In November 2016, for example, London's West End was blacked out for several hours. Then, in March 2018, the Harbour Exchange data centre in London experienced a long-duration power cut. Services remained unavailable or degraded throughout the day.

Power cuts motivate data centre operators to invest in solutions to protect themselves against longer-term blackouts, in addition to mains power aberrations and failures lasting milliseconds or seconds. As such, this additional protection must come from a UPS-generator pair, rather than just a UPS.

### The generator and UPS as a pair

Correct set-up of a generator/UPS pair starts with understanding the two components' complementary roles. During normal operation, the generator is powered down, although maintained on standby and ready to start immediately on demand.

Meanwhile, the utility mains supply flows through the UPS to the critical load. The UPS handles any short-term power aberrations or dropouts. However, there is always the possibility of a power blackout that lasts longer than the battery autonomy.

Accordingly, the generator is connected to an Automatic Mains Failure (AMF) detection panel which initiates a generator start-up if a power cut becomes critically extended.

Generators need two further features to fulfil their UPS standby function. Firstly, to ensure reliable start-up, the diesel engines should be well maintained, with an adequate fuel supply and a healthy battery for starting, have sufficient coolant and oil, and

kept warm by mains-powered engine water heaters.

Secondly, the UPS should be supplied with an AC waveform that's stable in both amplitude and frequency.

### Sizing and environmental considerations

Correct generator sizing is important, with oversizing usually being advisable. As an approximate guideline, 1.5 x nominal UPS capacity should be allowed for transformerless UPSs; this rises to 2 x nominal capacity for transformer-based UPSs, and 3 x nominal capacity for air conditioning.

Generators should be selected according to their continuous rather than standby rating. The generator's ability to support a step load should be reviewed. The generator, once selected, must be installed into its environment correctly. Most standby generators use a base tank which is double-bunded to ensure capture of any spilt fuel. Acoustic housings with various noise attenuation ratings are also available to mitigate the issue of acoustic noise pollution. The usual practice is to install generators outdoors in weatherproof and acoustic enclosures, to ensure sufficient airflow for cooling and venting of exhaust fumes.

Finally, the generator must be installed on a flat, level surface, as close as possible to the AMF panel and/or the essential services board, to minimise the required power and signal cable lengths.

### Conclusion

This article has shown how generators can fulfil their role successfully, provided that they are well-prepared, carefully matched to the UPS, and installed into their environment with sufficient care.



## James Stuart

Director of EMEA Sales, Nlyte Software

[www.nlyte.com](http://www.nlyte.com)



# Energy optimization with DCIM software

ENERGY AND POWER MANAGEMENT have always been at the forefront of our management of data centres. Its importance waxes and ebbs given the political climate, be it accessibility and supply, or concerns over global warming. Recently, the EU has taken on a chivalrous charge on this subject and created the “EU Code of Conduct on Data Centre Energy Efficiency”. The aspirations of a code of conduct are admirable, however, the IT environment is always in flux and becoming ever more complex. Given the dynamic nature of the data centre, how can companies move towards the goals of increased efficiency and decreased costs?

Historically we have held PUE (Power Usage Effectiveness) as the primary metric to monitor and manage energy usage. It is a simple metric used to gauge how efficient data centres are with their energy usage. The PUE can be monitored against comparative facilities to measure improvements brought about by power and cooling improvements as well as virtualization and the subsequent increase in server density and efficiency. However, PUE is just a starting point. It was noted, albeit a bit buried, in a white paper published by The Green Grid in 2012 that while there is satisfactory value leveraging PUE metrics, it does not scale. And scale is what IT is all about these days, so we need more than simple PUE monitoring to achieve the desired nirvana of Energy Efficiency in today’s Hybrid IT environment.

Data Center Infrastructure Management software (DCIM) delivers tools to monitor and manage the data centre that can make a substantial impact on optimizing energy consumption in the data centre. Since infrastructure management solutions can centralize a data centre’s information, it can give operational analysis into what areas need change. Software such as Nlyte

Software’s DCIM solution can provide an integrated framework for metrics, analysis, and performance results across both IT and the various facilities’ capacity. Beyond just looking at the big picture, data centre management lets you look at each change and implementation on its merit, and each change’s effect on the whole environment.

The key areas that I see DCIM can specifically address are:

- **PUE:** One of DCIM’s standard features is its ability to give easily accessible information about data. DCIM readily shows PUE metrics to help you trace how effective or ineffective changes are as they are being implemented and give you direction for future action.
- **Elimination:** Some servers are inefficient. Removing such servers can reduce overall energy use significantly and have a significant impact on energy consumption. DCIM solutions can help companies quickly find and handle such inefficiencies.
- **Mapping:** Different placements and layouts can increase or decrease the energy efficiency, especially concerning the heating and cooling airflow systems. Trying out new options to deal with such temperature regulations can significantly reduce energy usage.
- **Capacity planning:** Companies can reach a sweet spot with their max capacity, best placements, and creative or innovative solutions for specific spaces. DCIM can give insight into each centre’s unique situation to enable further capacity planning.

Whether it is using DCIM’s data to eliminate inefficient servers, to enhance capacity planning, or to monitor energy consumption overall, I believe data centre infrastructure management helps optimize energy management one implementation step at a time.

## Kevin Deierling

Senior VP of Marketing, Nvidia

[www.nvidia.com](http://www.nvidia.com)



## Greener computing

“UNTIL RECENTLY, the IT industry could rely on Moore’s Law to deliver not only more powerful computers regularly but also falling power requirements — that is a new chip could have twice the transistors as the previous generation but consume the same amount of power.

“Now that Moore’s Law is dead, data centres must develop smarter and more efficient ways to run workloads, including the use of AI-aided design, domain specific processors, smart networks, containerisation and Edge Computing. The traditional compute method was to run all tasks in software and all software on general-purpose CPUs, such as the Intel and AMD X86. Software developers would simply add functions and sophistication to applications (or databases, storage, etc.) without worrying about optimising efficiency.

“Today, software must add new features and become more efficient simultaneously. Having humans go through source code line-by-line to optimise it is tedious and slow, as is switching from high-level languages like Perl to lower-level languages like C/C++ to increase efficiency. But new Artificial Intelligence (AI) tools can optimise software code automatically, quickly reducing data centre power consumption.

AI also can improve the efficiency of data centre power distribution, airflow and cooling, plus reduce the time needed to move applications to domain-specific processors such as GPUs or SmartNICs, which each can run specific tasks much more efficiently than an X86 CPU. That in turn highlights the benefits

of smarter networks that connect the distributed processing power of multiple CPUs and GPUs that also process data in the network itself using SmartNICs and intelligent switches. For example, many security, storage, virtualisation and networking tasks run 2X, 3X, or even 6X faster (and consume proportionally less power) on a properly designed SmartNIC like the Mellanox ConnectX-5 and ConnectX-6 than on an X86 CPU.

“Another way to make compute greener is to switch from virtualisation to containerisation, because containers don’t run on top of a hypervisor and don’t duplicate the entire machine/OS/kernel stack for each application instance.

However, containers can be less secure than VMs, requiring additional security in the form of next-generation firewalls and security isolation for bare metal servers. These additional security requirements would require more compute power, adding a new challenge to greening IT operations, but again can be offloaded efficiently to SmartNICs and I/O Processing Units (IPUs) – such as the Mellanox BlueField – to keep data centres moving towards green.

“Finally, Edge Computing improves efficiency by moving data collection, processing and analysis tasks closer to devices and users that both generate and consume the data. It’s greener to analyse and react to call data in the base station, retail data in the store and video in the camera than it is to ship all the data to a central data centre, process it and send the resulting actions or recommendations back to the edge of the network.”



**Francois Sterin**

Chief Industrial Officer, OVHcloud

[www.ovhcloud.com](http://www.ovhcloud.com)



## Optimising energy efficiency in the data centre

REDUCING the amount of electricity required to operate a data centre is probably the most important step in improving sustainability for the data centre industry. Using a cloud provider is of course much more environmentally friendly for most organisations than having their own data centres. Whilst I expect we'll see a reduction in energy usage from the data centre sector as more organisations shift to public cloud models, there's still lots we can learn as an industry to reduce our emissions.

Data centre cooling is one of the main demands on energy within a data centre and therefore it's an area where the industry can expect to see much innovation.

### Liquid cooling

Developing in-house water-cooling systems is one way to reduce energy usage that is generating attention. Water cooling consists of putting liquid inside a server to cool the processors. The liquid is carried by heat exchangers positioned on the processors and other components that emit a lot of heat. Most of the heat in a server is generated by the processors, either CPU or GPUs. So cooling these properly is key. On average seventy percent of the heat generated by a server is captured by the liquid, which has a much higher heat transfer capacity than air. Therefore, servers that use a water element to cool themselves, rather than electricity, inevitably reduce energy consumption and optimise air flows. Our data centres using liquid cooling need less than 10% overhead energy on top of server energy. In contrast, a data centre using traditional cooling methods needs between 40-100% more. This technique is becoming increasingly important. The newest processors are so dense with high thermal dissipation that traditional cooling techniques based on airflow don't suffice anymore.

While many data centre providers are now looking at implementing ways of liquid or immersion cooling, OVHcloud pioneered this approach over 17 years ago. Today we use it across almost 400,000 servers worldwide. It's a tried and tested method that we are continually refining.

### Autonomous racks

So beyond the processors, how do you cool the rest of the servers? Heat exchanger doors are an additional key way of reducing energy consumption in data centres. At OVHcloud we place a heat exchanger between air and water right at the back of a server rack to capture the rest of the heat generated by the other components. This heat passes through water to get cooled with the watercooling system. As such our server racks are completely self-cooled and can operate in any type of environment. Avoiding all use of air conditioning in the room, negates the need to handle large volumes of air thus reducing energy consumption in the process.

This process is operational in lower or higher temperatures. Allowing our rooms to go under 18°C or above 30°C, we fight less against external temperature which results in energy efficiency as a whole.

For organisations looking at the sustainability of their data centre supplier, scrutinising their energy efficiency is certainly a good place to start. It's important, however, to take this in the context of the provider's overall green attributes, including their values, track record in sustainability and actual policies. For data centre providers, it's time to focus on optimising your energy efficiency as well as reviewing your approach to sustainability in general. If you don't you may soon find your customers are moving somewhere greener.



**Leo Craig**

Managing Director of Riello UPS

[www.riello-ups.co.uk](http://www.riello-ups.co.uk)



## Demand side response

MENTION that phrase to a data centre operator and you're still likely to be met with a shrug of indifference. Or the well-rehearsed retort that any potential reward doesn't compare to their perceived risk: backup power systems and batteries exist for one reason and one reason only – to prevent damaging downtime when there's a power problem. Using them for anything else, it's argued, undermines that non-negotiable insurance.

But when you look at how the way we're generating power is changing, is such binary thinking still viable? Last year saw more electricity come from zero-carbon sources like renewables than fossil fuels for the first time in history.

It's almost inevitable this trend will continue in the years to come. That means smart grids and demand side response won't just be desirable to help National Grid balance supply with demand and keep a stable frequency, it'll frankly be a necessity to keep the country's lights on.

So thinking back to the data centre... They all depend on uninterruptible power supplies to minimise that underlying threat of disruption.

But here in the UK, how often do we actually experience a major power cut? Although it's a must-have insurance, UPS is often an underutilised and expensive asset. That's why data centres should view DSR as an opportunity rather than a burden. It gives them the chance to radically rethink the role of a UPS.

Firstly, there are the financial opportunities. Batteries can store cheaper off-peak electricity, then release this power at busier,

more expensive times, cutting energy bills and grid tariff charges by potentially thousands of pounds per megawatt. There's also the chance to sell stored power from batteries back into the grid via DSR mechanisms like Frequency Response, which help keep the network balanced and stable.

Think back to last August and the major blackout that affected many parts of the country during that fateful Friday rush hour. National Grid deploying nearly 500 MW of battery storage in a matter of milliseconds was one of the key reasons why it could restore frequency within minutes and keep disruption to a minimum.

But what about the risk of using batteries in this way? Unless you've installed hugely expensive battery monitoring and management systems, it's tricky to assess the true condition of traditional sealed lead-acid batteries. Hand on heart, can you be 100% sure they'll work when required?

However, to participate in DSR your premium batteries must be backed up by sophisticated monitoring. That way you know exactly when there's any cell deterioration and if you have any blocks that require replacing. Rather than undermining resilience, smart grid-ready UPS will actually enhance your reliability.

Energy bills account for up to 60% of a data centre's running costs. For bigger sites, these electricity bills run into millions of pounds. Embracing smart grids and DSR will give operators greater control of a such a fundamental part of their business, all while helping to improve the security of both their own – and wider society's – power supplies.



## Carolina Uribe

Energy & Environmental Manager, Telehouse

[www.telehouse.net](http://www.telehouse.net)

# The green future of the digital economy

DATA CENTRES (DCs) support our modern economy by transmitting, receiving, processing, storing and managing digital data. From reading emails first thing in the morning to streaming content in the evening, society's demands for any-time access to data requires DCs to operate 24/7. To maintain fully operational DCs, a reliable and constant electricity supply is required. The UK Government has set challenging carbon reduction targets, with a commitment to achieving net zero emissions by 2050. In an international context where climate change and global efforts to reduce emissions pose a great pressure in all industries, how can data centres balance continual growth and support modern needs while contributing to challenging carbon reduction targets?

### Going green

The DC industry often comes under criticism for the type and amount of power it uses. Therefore many efforts have been made to increase the use of renewable energy sources, such as wind or solar power, instead of traditional, carbon-heavy fossil fuels such as coal and oil. A recent survey of UK commercial operators revealed that 76.5% of the electricity they purchased is 100% renewable, 6.5% is between 0 and 50% renewable, 7% is between 50% and 99% renewable and 10% is purchased according to customer demand\*. When it comes to improving data centre efficiency, green design is essential from inception. By establishing proactive, tech-forward sustainability and efficiency measures from the outset, data centre companies can ensure sustainable operation, maintenance and refurbishment of their facilities, focusing on a cleaner, more circular use of materials, energy and water. However, trying to balance cost and upgrading to more efficient technologies to support the UK's low carbon future poses a challenge. DC providers are also developing their own internal environmental strategies and

policies to respond effectively to these challenges. They are also adopting appropriate international standards in Environment and Energy Management (ISO 14001:2015 & ISO 50001:2018) to manage environmental impacts and improve energy performance as well as measuring and reporting energy use and efficiency through different government and voluntary schemes to improve transparency.

### Digital activities and transparency

When it comes to helping customers to understand the energy impacts of their digital activities, there is still of work to be done. Data centres are commercial businesses rather than consumer facing organisations, therefore being transparent about energy consumption is significantly important as it enables customers to take accountability for their own emissions arising from their supply chain and outsourced activities. There is also lack of awareness about impacts of the online activities preventing many individual consumers to make conscious sustainable choices.

### Working in collaboration

As environmental obligations increasingly become intertwined within the business operations, awareness of these greener requirements are growing and organisations are looking to collaborate with those who demonstrate a clear commitment to sustainable growth and efficient use of resources. Businesses are working in partnership with industry bodies and regulators to deliver solutions that support operational efficiency, innovation, carbon reductions, resilience and value for customers. After all, almost everyone in the world is connected to a DC in some way – and by putting sustainability at the heart of the DC operations; they will continue to meet ever-growing digital needs in a progressive, sustainable and scalable way.

**Tate Cantrell**  
CTO, Verne Global  
[www.verneglobal.com](http://www.verneglobal.com)



## Optimising a data centre for machine learning

WHILE MACHINE LEARNING can take many forms, many of the recent advances in machine learning applications are due to the advance of models based on the architecture of the brain. By applying these algorithms and training the 'neural network' models against existing or manufactured data sets, applications can learn new patterns and then be deployed to gather information. The amount of compute required during the training phase of these artificial neural networks is orders of magnitude greater than earlier generations of artificial intelligence applications. The greater the network depth and the greater the data quantities for input, the greater the compute complexity, all of which requires high performance computational power and longer training times. In fact, the computations required for deep-learning research have increased 300,000-fold from 2012 to 2018.

With these types of compute intensive applications, energy consumption is enormous and keeping equipment cool consumes a substantial amount of power. The estimated carbon emissions from training a single deep learning model for language processing are nearly 40 times more than the amount generated by a flight from New York to San Francisco, per 2019 research from the University of Massachusetts. If our future is to involve more intelligent algorithms without runaway climate change, we have to optimise. MIT Labs is actively looking at concepts such as once-for-all networks that could allow for machine learning models to be fully trained once for all applications, and then optimised for individual functions. While this approach holds promise for the future, today, we must look to the data centre to provide the most immediate impact for innovation in efficiency. First and foremost, ensure the data centres that power their machine learning applications are truly

sustainable. Thanks to the proliferation of global high capacity data networks, the compute involved in training machine learning models does not need to be located near the end-user for performance or accessibility. Racks can be comfortably housed in a location like Iceland, where only green hydroelectric and geothermal power plants are connected to their truly sustainable power grid. This has a significant impact on carbon footprint, and can cause the direct carbon emissions of an algorithm to vary by a factor of 40, from 20g of CO<sub>2</sub>eq/kWh when housed in a data centre that uses renewable energy sources, versus 820 CO<sub>2</sub>eq/kWh in a location that solely relies on fossil fuels.

Data centre layout can also be modified to accommodate high-intensity AI applications. Greater emphasis is put on rack layout to maximise the bandwidth capacity between servers, while taking into consideration the cooling needs for specialised accelerated servers such as the NVIDIA DGX-2.

Data centres must also be structurally capable to handle very heavy equipment. The best data centres today allow fully populated AI computing cabinets to be rolled from the loading dock to their final computing location all on the same level and all on reinforced concrete slabs. The adoption of water cooling will also begin to emerge as we reach the limits of pure air cooling for this power intensive equipment.

The opportunities for innovation with AI and machine learning are immense. The cost of that innovation does not need to do more harm than good. Well-prepared planning and choosing the right location are just two of the positive steps companies can take to minimise their carbon footprint whilst still accessing cutting-edge supercomputing.

## Billy Durie

Global Head of Sector – Data Centres

[www.aggreko.com](http://www.aggreko.com)



# Cooling data centres into the future

AS TECHNOLOGY CONTINUES TO ADVANCE, and innovations such as 5G and AI become commonplace in a multitude of industries, it is more imperative than ever for data centre operators to avoid downtime. With data centre usage higher than ever, operators must ensure their critical infrastructure can keep up, particularly when it comes to cooling.

### Keeping their cool

A data centre's cooling infrastructure is crucial and there are a variety of methods that are more suited to particular locations or uses, whether that be air or liquid cooling. Given that the data centre industry consumes over 200TWh of electricity a year, cooling infrastructure is also required to be as efficient as possible so that the energy used by the cooling system is kept to a minimum.

An efficient means of non-mechanical temperature control, free cooling uses fresh air blended with recirculating air from within the data centre to efficiently cool the server spaces. For data centres located in hotter and more humid regions, or running at higher density, liquid cooling systems are an efficient means of cooling server racks and transmitting heat to the outside. Although often implemented in hyperscale facilities, liquid cooling is being introduced ever more in other data centres. Whichever cooling method is employed must be effective during any level of demand. As demand rises, it is crucial that cooling infrastructure in facilities globally is not affected by elevated ambient temperatures.

### Demand and temperatures rising

With the International Data Corporation (IDC) predicting that there will be 175ZB of data stored by 2025 compared to 33ZB

in 2018, it is clear that demand for data centres will continue to rise dramatically. As temperatures rise across the globe and each summer becomes ever warmer, data centre operators have to make sure that their facility can keep up with cooling requirements to avoid any costly downtime. Facilities that use free cooling technology may find that, during summer months, their systems could be limited when coping with prolonged periods of elevated ambient temperature. For data centres that use other cooling methods, high power demand from the server racks with that of the cooling system could put too much pressure on the power supply. If either failure occurs, there must be a plan in place to mitigate their impact on uptime.

### Planning pays off

As uptime is more crucial than ever in recent times, data centre professionals must be prepared for a warmer future. By having plans in place before the effects of elevated temperatures take hold, the potential for failure can be mitigated, ensuring downtime is minimal. As part of the planning for maintaining uptime during the summer and heatwaves, data centre providers could employ the use of supplementary cooling and power equipment during the periods.

Implementing the correct type of equipment is vital, as different sized facilities in distinct regions require certain solutions and amounts of equipment. When it comes to planning this, data centre professionals can work with companies such as Aggreko to ensure their facility has this equipment ready. Whether the cooling solution in place is free cooling, liquid cooling or more traditional air conditioning, with the correct planning, temporary solutions can be integrated without any delay when temperatures begin to rise in the summer months.

## John Booth

MD/Consultant, Carbon3IT

[www.carbon3it.com](http://www.carbon3it.com)



# Follow the code of conduct

## Energy Optimisation

The best way for any organisation to start their “energy optimisation” journey is to use the EU Code of Conduct for Data Centres (Energy Efficiency), aka EUCOC, a series of 150+ best practices that cover Management, IT, Cooling, Power Systems, other DC systems, Design and Monitoring.

Used correctly and with expert guidance, the EUCOC is, in my opinion, a strategic and cultural change roadmap that will transform your organisations IT into a nimble, energy efficient, and sustainable business enabler. And you don’t need to officially sign up to the endorser/participant scheme although with the potential of legislation it may be wise to get ahead of the game sooner rather than later.

## Why Strategic?

Many of the best practices really make the organisation think about how IT delivers their digital goals, whether it sticks with 20th Century thinking, or moves forward with the adoption of new concepts such as grid and cloud computing, the use of innovative solutions such as “immersed compute” and a real understanding of what is actually mission critical. It demands regular reviews of IT infrastructure, audits of services and assets and the supporting M&E infrastructure, stripping waste out of the system and enabling energy optimisation.

## Why Cultural?

Because the use of the EUCOC really makes your IT department think about what services they need to provide to enable that nimble energy efficient business enabler, in effect it forces change to be considered, do it and your organisation will prosper, don’t do it, and be left with a aging legacy environment that will cost more

and more to run, go down regularly and be a constant source of frustration for IT staff and your workforce alike.

## Is it expensive?

The EUCOC is free to download (links below) and you can go through the best practice guide and implement the best practices as necessary. However, if you’re seeking to be a full participant or endorser it may be wise to seek specialist help. In our experience, it will take between 4 and 6 weeks to get the information, or 1 day with a specialist with relevant personnel in attendance, this will cost between £3000-5000+VAT and you’ll get the reporting form completed and a comprehensive action plan.

## What next?

Well that depends on your level of maturity and the result of your review, we’ve visited many sites where the energy optimisation opportunities are restricted through SLA’s and OLA’s but it is fairly easy to get these changed although it might require some re-education. Policies, processes and procedures and temp/humidity settings are where you stand to make the biggest gains and not as you might expect, from expensive infrastructure solutions. You will need to install a monitoring solution but that’s to measure the actual energy savings as well as to calculate your PUE.

## Summary

So, in summary the EUCOC and all the relevant information can be downloaded from <https://ec.europa.eu/jrc/en/energy-efficiency/code-conduct/datacentres>

And we guarantee that the EUCOC will provide any organisation with that strategic and cultural change roadmap but seek specialist help.



# HYBRID INFRASTRUCTURE

In-house, colocation and hosting, managed services, cloud, multi-cloud...it's highly unlikely that one-size fits all an end user's requirements. Hence, the hybrid infrastructure models has been gaining traction in recent times. In an ideal world, the resources required to stand up any particular application can be specified and then a decision made, based on various factors including cost and resilience, as to how best to provision them. It maybe that the application can be accessed via a managed service; it might be best left to run in-house, or run somewhere in the cloud, or on a company's own IT hardware, housed in a colocation facility. It's rapidly becoming a hybrid world.



IN ASSOCIATION WITH



**DATA CENTRE**  
**SOLUTIONS**



**DIGITALISATION**  
**WORLD**



TRAINING

DEVELOPMENT

SKILLS



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

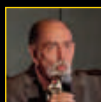


**Photonics**

**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



**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 overall experience very professional



**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



**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, Burnshall Road, Coventry, CV5 6SP. UK  
**T:** +44(0)2476 718 970 **E:** info@angelbc.com **W:** www.angelbc.com



# Data centres in a 5G era

When it comes to 5G, we're just getting started. Entire industries are simultaneously planning for a new era of connectivity, bracing themselves for what is set to be one of the most influential roll-outs in technological history. Enabling higher speeds, lower latencies, and more machine-to-machine connections, the arrival of fifth generation networking is already sparking a revolution amongst data centres. And there's more to come.

**BY ERIC LAW, VICE PRESIDENT ENTERPRISE SALES EUROPE AT COMMSCOPE.**



IN THE NEXT YEAR ALONE, Gartner has forecast worldwide 5G network infrastructure revenue will reach a staggering \$4.2 billion. That's almost double its current value. The analyst house has claimed that, despite still being in the early days, 'vendors, regulators and standards bodies' alike all have preparations in place.

Meanwhile, the UK government is pushing back against planning authorities that are attempting to halt the installation of 5G masts, with a spokesperson for the Department for Culture, Media and Sport (DCMS) expressing the government's commitment to 5G, highlighting the economic benefits of the roll-out.

Even as the first 5G services are starting to go live, however, there are still many questions that remain

unanswered. To ask how fifth generation networks will affect life inside the data centre, therefore, is similar to asking how a city would stand up to a natural disaster. It depends on the city and the storm.

### Moving to the edge

We now know enough about 5G to know that it will without doubt change how data centres are designed and, in some cases, will even alter the role that they play in the larger network. From a technical standpoint, 5G will have several defining characteristics.

The most obvious of these is the use of the 5G New Radio (NR) air interface. Exploiting modern spectrums and providing latency capabilities in just milliseconds, this enhanced performance will drive the deployment



of billions of edge-based connected devices. It will also create the need for flexible user-centric networks, pushing the compute and storage resources closer to both users and the devices.

The only way to meet these ultra-reliable, low latency requirements will be to deploy edge nodes as mesh networks, with an east-to-west flow and parallel data paths. In some cases, these nodes may be big enough to classify as pod-type data centres or micro data centres in their own right, similar to those being used by both telecom and cable providers.

## Preparing for disruption

Cloud-scale data centres – as well as larger, enterprise facilities – may be the only ones to see just some impact of this move. They are already using distributed processing and have been designed to deal with increased data flow from the edge. At the other end of the scale, retail multi-tenant data centres (MTDCs) will likely incur the most disruption, as they have traditionally grown in response to rising demand for cloud-scale services.

The biggest changes, however, will be seen among service providers. As they begin to refine their relationship between core data centres and evolving centralised RAN (CRAN) hubs, adaptation will become a choice of sink or swim. Increasing virtualisation of the core networks and the radio-access networks will be key when it comes to handling the anticipated 5G data flow. This will also enable service providers to be more flexible with compute and storage capacity, easily moving it to where is most needed.

More broadly across service provider networks, increasing virtualisation may have a more direct effect in the core data centre, with wireless and wireline networks becoming more converged. This will generate an even stronger business case for a single physical layer infrastructure, it then just depends on the degree of convergence that occurs between the core network and the RAN. Whether this occurs in a central office or in a data centre, we still don't know.

## The role of new tech

Aside from considering cloud models, we also need to focus on the impact that new technologies such as artificial intelligence (AI) and machine learning (ML) will have on data centres when 5G makes its mark. These technologies will not only require accelerated server speeds, but also higher network capacity to enable greater volume of growing edge services. When building these data models, processing vast data pools will be essential which, in most cases, are best suited to core data centre capabilities.

Most of the data that goes on to develop AI models will come from the edge. This nods to a potential shift in how more established, cloud-scale data centres will support the network. One potential use-case involves using the power of the core data centre to assemble

data from the edge to develop the AI models. These would then be pushed out to deliver localised, low-latency services. The process would then be repeated, in turn creating a feedback loop that revolutionises the operating model.

## A balancing act

As with any other aspect of digital transformation, the level and requirement of change within various data centre environments will depend on the individual application. Data streams generated by the billions of sensors and devices may well produce a steady flow of data, whereas others might be delivered intermittently or in irregular bursts. Either way, it's very much out of our control. It's how the data is then collected, processed and analysed that needs to be optimised, considering factors such as how much data should remain local to the edge device, and how much should be processed in the core centre?

Once these questions are answered, network engineers need to determine the best way to move the data through the network. Different latency and reliability requirements require the ability to prioritise data traffic at a very granular level. What can be off-loaded onto the internet via local Wi-Fi versus having to be backhauled to the cloud service provider (CSP) data centre? And remember, edge networks must fit into a financial model that makes it profitable. Infrastructure as an enabler

The widespread roll-out of 5G is still a few years away, but there is no better time to start preparing for what's to come. At the very core of this evolution, infrastructure must adapt to support higher wireless bandwidth and more universal data usage. Behind closed doors, organisations and building owners are considering more than just Wi-Fi to enable robust and dependable in-building mobile wireless with distributed antenna systems (DAS). Outdoors, it's a different story. Service providers are upgrading and expanding their fiber networks to carry wireless data back to the core of the network, or in many cases, to edge data centres.

What we really need to think about, is the applications and innovations that these changes will develop as part of a 5G era more broadly. Self-driving cars, facial recognition, smart cities and industrial automation will all be made possible and more advanced by 5G. The problem remains, however, that each of these applications all have a varied set of requirements regarding reliability, latency, and the type/volume of data traffic generated. Unless you therefore understand the parameters of the situation, it's difficult to pinpoint it's exact impact on the data centre.

Something we do know, however, is that the avalanche of new data from the network edge, and thanks to 5G, goes hand-in-hand with high compute and storage power. Exactly how much power? You guessed it: that depends!





# Hybrid cloud will be a 2020 necessity for businesses, but is it all it's cracked up to be?

2020 is set to be hybrid cloud's year, and for good reason. It's a way of working that welds together the benefits of the public cloud infrastructure with on-site data centres.

**BY SEAN ROBERTS, GENERAL MANAGER OF PUBLIC CLOUD, ENSONO.**



Hybrid, in theory, provides the pros of both, while offering a far more flexible platform: one that can scale up and down and access the latest technology features of the cloud vendors, whilst maximizing utility of the existing assets on a more predictable capex model. However, there's still much confusion around what the hybrid cloud is and isn't, which is holding

back some businesses from achieving the very best from their IT infrastructure.

### Clarifying the confusing

It's worth mentioning right off the bat that hybrid cloud is rarely something that's decided upon as a primary option – rather, it's something businesses

# HYBRID INFRASTRUCTURE

move to out of necessity. While small businesses can, with their light infrastructure and minor data footprint, jump straight to a public cloud service relatively easily, for larger organisations it's a different story. Incumbent enterprise are usually burdened with legacy IT architecture compiled over decades of operation and M&A activity. Trying to move everything to a public cloud service with the click of a button is, unfortunately, impossible which is often why a hybrid half-way house approach is adopted.

The most successful hybrid journeys begin by addressing the application portfolio rather than a straightforward lift and shift. Is there data that is required to stay on premises for contractual reasons? Are there technical, commercial or licensing requirements that demand some form of private cloud too? Where that's the case (ruling out a cloud-only option), larger organisations are finding that the hybrid cloud – the best of both worlds – is the way forward.

Still, it tends to rely on a matter of rationalizing the application portfolio. How many businesses, for instance, have separate applications that provide duplicate capabilities to separate parts of the business? This usual comes about as a result of M&A activity or via some form of decentralised IT decision making. Understandably, many have taken advantage of a migration to a hybrid cloud system to consolidate their applications and remove such duplication before moving to public cloud.

That's especially evident in applications that have a cloud-native version available. In lieu of that update

being an immediate viable path, parking existing applications on a hybrid cloud is a solid holding position - assuming, of course, that the cost benefits of shifting an application and its data to the cloud outweigh the price of doing so.

A necessity, but still, a useful place to be. What's ultimately become clear this year is that hybrid cloud is a useful interim place to be. One that's especially invaluable for organisations whose infrastructure is in a state of some flux. For companies that don't have an appropriate native cloud option in place, it can be a bridge to better. A full public cloud service is typically the end goal, and the majority of businesses now acknowledge that it's a case of 'when' not 'if'.

What's more, there are plenty of solutions currently on the market that make hybrid infrastructure a viable and long-term option. Private cloud platforms such as AWS Outposts and Azure Stack, for example, bring native AWS and Azure services to both data centres and on-site facilities. They've also made inroads into the world of hybrid cloud, which in turn demonstrates that major players in this space are working with hybrid in the medium as well as the long term.

Hybrid cloud offers businesses a way to move on from legacy systems without having to transform to a public cloud architecture overnight. It allows organisations to tread a cautious path to becoming cloud-first, offering a beneficial transitional state and the opportunity to plan the next stage of development.



**Ian Jeffs**

UK GM, Lenovo DCG

[www.lenovo.com](http://www.lenovo.com)



## Unlocking the Power of Hybrid Cloud Storage

THE INCREASING VALUE of data as a resource is clear to see and is underlined by the growth of the data centre market. In 2020, worldwide spending on data centre systems is anticipated to amount to as much as \$208bn, showing that the amount of data being collected and optimised by businesses is growing significantly. In order to keep up with this mass proliferation of data, businesses must look to optimise the power of the data available to them and adopt standardised methods of collecting, analysing and storing this data.

While legacy data systems typically use traditional databases, modern technologies are introducing new variations of data that create a need for cloud storage and high-performance computing (HPC). Whether this be through sensor data or audio or video data, there becomes a need to adopt more dynamic storage methods, which is where the capabilities of cloud storage come into their own.

Cloud data storage better enables metadata collection and enables scalable data storage. This also makes it possible for businesses to adopt globalised data storage plans, which ensures data can be accessed in real time from anyone in the organisation wherever they are in the world. As we continue to see the impact of the COVID-19 pandemic and an increase in remote work, we can see a further need for dynamic data storage that enables efficiency and speed of access. This will potentially accelerate data storage strategies for some businesses, as the increase of remote working increases the need for remote access to data.

For many businesses, however, the key is not to shift to an entirely cloud-based data storage plan, be it for reasons of security, ease of access, or budget. Instead, an increasingly popular option for small to mid-size businesses is to adopt a hybrid strategy that uses both local data storage and cloud platforms. This is because wherever fast access to large amounts of data is required, the local provision of SAN or other storage systems is still essential.

The fusion of these two approaches make it possible to guarantee fast access to data and ensure businesses can align their corporate infrastructure with the most relevant data storage solutions.

Hybrid storage is also beneficial when it comes to archiving data that is not needed on a day-to-day basis but businesses still wish to store for future use. To optimise how storage is allocated, tiering mechanisms are used which automatically decide where location data is stored.

We can be optimistic about the future of data storage, with further developments in AI and 5G meaning we will continue to see innovative data solutions and business leaders will need to stay abreast of these developments. Such innovation will empower companies to unlock their data and harness its potential. The rise in data volumes and the emergence of new storage trends will continue to fuel an already flourishing data centre market.



**Galeal Zino**  
CEO, NetFoundry  
<https://netfoundry.io>



## Orchestration key to cloud optimisation

The COVID-19 pandemic has thrown a curveball at most enterprises. Not least among the unforeseen challenges is ensuring that workers can still function effectively at a time when most are not able to go into work. Naturally, those enterprises with a robust hybrid and multi-cloud approach are faring best. The pandemic has driven up demand for cloud-based services, and further underscored the extent to which we are all reliant on the data centres where those services reside and the hyperscale cloud providers that operate them.

There is evidence that while many industry sectors are facing difficulties, the pandemic is not affecting demand among Tier 1 cloud service providers for new data centre capacity. According to research from analyst firm Dell'Oro Group, Tier 1 cloud service providers are actually expecting to increase data centre capex outlay in 2020. Dell'Oro has found that while last year most were slowing their capex spending as they consumed excess server capacity, this year they will be spending up 20% more than last year as they resume server capacity expansion in order to meet enterprise needs.

But of course there is no one uniform approach to cloud among the customers of these services. There are multiple approaches to cloud deployment, ranging from the more traditionally minded enterprise that still retains a lot of their own in-house server capacity, to the digitally focussed start-up company which might well have placed all applications and workloads in the cloud from day one, never owning infrastructure of their own. The cost of running their own compute power versus paying for it as a service makes no sense for them. For those not starting from a digital Year Zero, the best option is choosing to spread applications across a mix of public clouds, while keeping certain

functions within their own private cloud. This hybrid approach takes them closer to their ultimate goal of digital transformation. By learning to leverage the potential of cloud, and make it work alongside legacy investments, they are achieving a measure of agility and helping to free themselves up to address a world of new commercial possibilities.

Of course there are issues and challenges that can surround hybrid and multi-cloud deployments. Some organisations possess the kind of in-house resources that help them to manage the multiple connections that a multi-cloud environment inevitably entails. But traditional network management tools were not designed for this task, especially not for those who do not have huge IT resources to throw at the problem.

The issue of complexity is not getting any easier with so many organisations now developing their own applications to add to the rich mix of off-the-shelf applications they already rely on. As the Internet of Things picks up pace, we can expect to see this process accelerate. And if you are relying on multiple clouds across multiple data centre facilities, then you need to have a plan in place should one of those cloud providers go offline without warning. Data must keep flowing seamlessly in any event and workloads fail over from one cloud to another, if the whole organisation is not to suffer. An intelligent way is needed that automatically determines the best available path for data to take if one data centre, or indeed a whole region, should go offline. Relying on a mix of on-premise and public cloud services, means that organisations need some kind of orchestration between the different elements such that it all functions as one logical entity. Achieving this level of manageability and simplicity remains an elusive goal for many.



## Simon Bearne

Commercial Director, Next Generation Data

[www.ngd.co.uk](http://www.ngd.co.uk)



# Balancing cloud hosting and colocation

INCREASINGLY, the majority of IT workloads will be shared between providers of off-premise colocation and cloud hosting services. The move away from on-premise data centres is an evolution rather than revolution, a process that many enterprise organisations must transition through naturally before eventually achieving a fully optimised IT environment that's right for their specific needs; consistently delivering cost-efficient, secure, resilient, flexible future-proofed computing.

As part of the journey enterprises will inevitably need to address the implications and risks of continuing to run servers in-house for storing and processing large volumes of data and applications. Along the way they will also identify numerous workloads that just won't go into the cloud or discover that the effort or cost is a barrier. In fact some analysts are already predicting that only about half of enterprise workloads will ultimately go into cloud.

### Drivers for change

Public cloud platforms probably won't totally satisfy all enterprise application and workload needs. Some will demand extremely high performance while others will just need low-cost storage. The advantages, however, of data centres which house hyperscale cloud are the avoidance of cost and points of failure for cloud interconnect.

At the same time, private clouds can present serious engineering challenges and costs when running them on-premise. In addition - hastened by Covid-19 - users are rethinking their ongoing use of office space. Increasingly, they will want their IT stack to be independent of their office, in a lights out critical services environment that supports distributed 'pandemic' operations.

Faced with solving the rising costs, complexities and security issues involved, many will turn to finding quality colocation facilities capable of supporting their considerable requirements - from housing servers for running day to day applications, legacy IT systems, and in some cases, mission-critical systems, and for hosting private or hybrid clouds. Slowly but surely this will see most of the remaining enterprise workload share moving from on-premise to colocation - with only a minority remaining on-premise in the long term.

From what we at NGD are seeing, the signs are that colocation can look forward to a vibrant long term market. There's a growing choice of quality colocation facilities thriving in regional UK locations, some with a high level of connectedness. And in NGD's case also functioning as a regional Internet hub - as a point of presence for the London Internet Exchange's LINX Wales - alongside scalable colocation and hyperscale cloud hosting. However, CIO's shouldn't assume that colocation will always be available immediately, exactly where they need it and at low cost. As the decade wears on, some colocation providers will probably need to close or completely upgrade smaller or power strapped facilities as well as those that lack diverse connectivity. Others will build totally new ones from the ground up.

Only those that are 'at scale', with significant levels of 'forwards scale' will be capable of delivering future-proofed solutions for larger workload requirements. Few, in reality are, particularly with limitations on grid power in metro areas. Time is therefore already of the essence for CIOs currently striving to balance and optimise workloads. Commencing the evaluation of potential colocation and cloud hosting partners sooner rather than later will pay dividends in the long run.

## Rob Tribe

Senior SE Director, Nutanix

[www.nutanix.com](http://www.nutanix.com)



# Hybrid - the answer today and tomorrow

COVID-19 has underlined why we need more flexibly dispersed IT and that need is about to become more pronounced

The pattern of network traffic we see coming in and out of the datacentre is changing rapidly. Increased use of cloud computing, greater use of remote locations and home working, accelerating globalisation, the proliferation of Internet of Things devices plus more use of cellular bandwidth including 5G are all impacting what ICT leaders are seeing in their facilities.

Datacentres are changing in order to cope with these morphing use cases and the clear need is for hybrid architectures that give IT dynamic control over resources and provide the ability to handle traffic locally or at the network edge.

Now more than ever, the way that CIOs and CTOs respond to the need for more employee home and remote working is of course front of mind. The sudden shift caused by the pandemic and lockdown is driving demand for bandwidth and load balancing but also for remote management as staff work from their laptops, smartphones and tablets, often using virtual desktops and hammering server-based applications. End-user computing is reimagined in the most dramatic ways and datacentre architects are responding accordingly.

The old way of dealing with this sort of challenge would have been to throw more and more hardware capacity at the problem. Adding servers and disk arrays is an expensive approach, however.

Better today to use hybrid architecture designed with VDI, scalability and manageability in mind from the outset. It's also far

better to have a tariff that is based on usage so CIOs can rely on operating expenditure rather than upfront capital expenditure. The second road ends in the type of 'bill shock' that tends to complicate the relationship between IT and Finance!

Hybrid architecture makes it easier to dynamically control server, network and storage resources and is showing its value today, but it will come into its own when more emerging technologies enter the mainstream.

IoT will demand enormous scalability and the numbers we have seen suggest that edge computing in environments such as smart factories will ingest as much as 32 times more data than cloud, thanks in part to sensor data and image analytics. This data will often be critical to health and safety in factories, oil rigs and other mission-critical locations.

But that sort of technology will also be omnipresent in our daily lives, for example helping us to automatically pay for goods without checkout tills, as is the case with Amazon's Go stores.

Such queue-busting use cases will make everyday activities faster and less frustrating but they need intelligence at the network edge, often bolstered by AI. The same can be said of driverless cars: another paradigm that will be heavily dependent on edge architecture.

For decades now we've seen IT evolve from heavily centralised models to today's federated versions but the future will see it enter every nook and crevice of our world. For that we need optimally flexible hybrid architecture to give us the edge we need.

## Jay Gill

Sr. Director of Marketing, Pluribus Networks

[www.pluribusnetworks.com](http://www.pluribusnetworks.com)



# Distributed clouds: Enabling action at the edge

FOR EDGE COMPUTING to reach the heights its proponents have aspired to organizations will need to fundamentally rethink how to operate data center infrastructure. Meeting this challenge will require purpose-built hardware and software that is easy to deploy and manage while still meeting the stringent constraints of more, smaller and distributed data center environments.

Most vendors have tried unsuccessfully to retrofit technologies and products designed for large-scale centralized data centers and found that they just don't fit the significant space, power, environmental and cost constraints of edge locations such as telco central offices and 4G/5G base stations, let alone even smaller sites in street cabinets or harsh rural environments. And while industry organizations have published blueprints and reference architectures for edge networking, most of these are still at the stage of science experiments in the lab or, at best, in "proof of concept" field trials. This is finally starting to change, driven by new consumer applications, IoT and, at long last, by 5G. Infrastructure deployments to enable 5G are accelerating and will likewise be an accelerant of edge compute. While the action at the edge will likely grow independent of 5G, the combined momentum will bring much more pervasive access to broadband speeds and lower latency, while presenting an architecture that is inherently flatter and more distributed.

It's those applications such as VR/augmented reality, video analytics, public safety and industrial IoT that have sets of requirements that are difficult to meet with a centralized cloud architecture. The most broadly understood application, because of all the hype, is autonomous driving and the high-speed low latency networking that it requires. But the edge is moving ever closer to the most traditional industries, particularly the farming

industry, with the increasing support of precision agriculture, where edge computing and 5G public and private wireless support real-time data collection from connected IoT devices. All of this, from consumer AI to industrial IoT to the 5G future, has triggered a need for compute and storage to be deployed closer to users and things, something we call the distributed cloud.

Distributed cloud infrastructure is the fusion of edge compute and the cloud consumption model where edge locations will be networked across physical locations: between centralized data centers, colocation facilities, central offices, base stations, remote facilities and on-premise. Applications will be placed at the edge location where their requirements are met and costs are minimized. This proliferation of distributed cloud sites will require a large number of mini- and micro-data centers owned both by service providers and enterprises, all of which require network connectivity that is cost-effective, easy to manage and can provide the network slicing mandated by 5G.

The explosion in the number of sites to manage means that SDN automation capabilities become essential to basic operations. Distributed cloud infrastructure requires software-defined networking constructs that can scale cost-effectively beyond the walls of a single data center, for seamless automation and monitoring across all the various "edges" of the distributed cloud, from the central data center out to a cell tower or a factory floor. When the focus was on autonomous driving only, the edge hype was ahead of the reality. The fact is - today there are now multiple applications across consumer, agricultural, industrial and telecommunications that require high speed and low latency that it can only mean one thing: more and smaller data centers, closer to people and things.

## Simon Michie

Chief Technology Officer, Pulsant

[www.pulsant.com](http://www.pulsant.com)



## Living on the edge

ACCORDING to Gartner, edge computing will account for 75 percent of enterprise-generated data by 2025 – currently only it only accounts for 10 per cent. Some of the applications that 5G enables, such as autonomous vehicles, are very dependent on rapid processing of data and latency and responsiveness are key to both user experience and safety. Other businesses will take the opportunity to instrument their offices, factories and products at scale and integrate these sensors into their business support systems. By bringing computing power and storage capabilities closer to where the data is collected, organisations can then analyse and use data to shape decision making in near real time.

This ability to act upon data in real time, processed locally, will allow businesses to automate their operations, adapt faster to changing circumstances, and ultimately to create new and innovative services for their customers to consume.

It's not just businesses that will benefit from the adoption of 5G and edge computing, but data centres too. While the path of adoption for 5G is not yet fully mapped out, as the up-take of 5G by businesses increases and then drives growth in more complex and mission critical IoT solutions, the role of the regional data centre will become even more pivotal.

What we will see is data sent less regularly to remote centralised data centres and instead delivered across a consistent, distributed grid-like architecture spanning many data centre locations that are closer to edge devices, with less time critical processing remaining centralised. This architecture is being developed rapidly by the major Cloud technology providers like Amazon and Microsoft and engineered to work with 5G networks. Regional data centres have a key role to play in that

architecture, providing a reliable platform for delivering services to the edge, and bridging between centralised platforms and micro-edge locations, such as masts and base stations.

While the benefits are clear, to get the most out of edge networks organisations will need to ensure they have the right infrastructure in place to leverage the advantages. The physical location of the edge data centre, the design of the service architecture within it, and the quality of connectivity to the 5G network will have a dramatic effect on the success of the outcome. It will become important to be able to spin up location specific edge compute instances which are capable of supporting a modern distributed, loosely connected application set. Similarly, the data centre industry will also require a shift in mindset, particularly when it comes to the location. Data centre providers have traditionally owned or rented facilities in specific locations to serve existing business hubs, rather than developing their facilities to be closer to where their new customers will be. Low latency connectivity to 5G telecoms networks will need to be added to existing capability.

Another consideration is data security. On one hand, edge deployments increase the complexity of this challenge, with data potentially stored in multiple locations. However, equally an edge data centre can help businesses address concerns about where their data resides, with guarantees over legal jurisdiction, privacy and compliance becoming more and more important to the end customer. There are many elements to getting this right – securing the device being one area that is particularly important, along with the encryption of data passing across the networks. Regional data centres themselves will have a key part to play in addressing these challenges moving forward.





**Paul Speciale**  
Chief Product Officer, Scality  
[www.scality.com](http://www.scality.com)



## Edge and transformation of the cloud landscape

THE PROLIFERATION of connected edge devices creating and consuming data is changing the way data needs to be managed. It is estimated that by 2021 there will be between 25 Billion (Gartner) to 31 Billion (IDC) connected edge devices. At the same time, the cloud landscape is transforming: data is spread across core clouds, both within the enterprise datacentre and in public clouds, managed by hyperscalers, MSPs and Telcos.

Edge computing is impacting the way cloud providers look at the market, and in fact, edge computing seems to be a natural extension of many of the services offered in these core public cloud data centres. The large cloud providers no longer see edge deployments as threats to their dominance, and are instead investing in edge use cases and solutions that span edge-to-core. In this way, data generated at the edge can be processed locally, and easily pushed to the core cloud for management, analytics and storage.

The cloud and the edge each have their places and their advantages for applications and infrastructure deployments. The benefits to centralising data storage and processing in the cloud are well known, however, in instances where there is no tolerance for high latency, it is more beneficial to bring the processing to the edge, where the data is generated.

We are hearing about two versions of edge: the compute edge and the service edge. Today, the data collection side of the edge is massive, and growing every day: collision data from automobiles, weather data from distributed sensors, and even data generated from today's ubiquitous smart speakers, smart refrigerators, smart televisions and more. Some of this data is immediately processed at the edge, and the rest is sent to the

core data centre where it is stored for its future potential.

At the same time, service edges, which can be thought of as mini clouds or data centres, are emerging closer to where concentrations of endpoints—edge devices—collect data. These service edges are emerging and growing in locations such as hospitals, sports stadiums, exploration platforms, airports, ships, retail stores, offices and more. Even cattle ranches. They serve as waypoints for the devices that serve-up those mass quantities of data. Using the example of a hospital service edge, this mini-cloud/data centre would collect the data from patient sensors, equipment location sensors, distributed medication stations and more, to collect and process information from those nearby edge devices so that the organisation can derive the value of the data—make it actionable—more quickly. These service edge data centres are ideal for deployments of private cloud infrastructures such as Azure Stack Hub & Edge which offer local compute and storage, and can even work in a disconnected manner.

While these edge locations are physically removed from traditional core or cloud data centres, access to local infrastructure resources is vital for use cases such as local data capture or content delivery. This is the new reality, and IT services have to be more flexible—impose less friction—so that these critical decentralised “centres” at the edge, where the data is amassed, are easier to spin-up, spin-down, grow and contract as needed. Logical architectures that share services and enable the movement of data and the processing of that data freely between edge and cloud are enabled by widespread network connectivity. And, the network infrastructure is evolving quickly to support this infrastructure, with \$57 billion being spent on 5G buildout through 2022, according to IDC.

## Simon Pamplin

Technical director, Silver Peak

[www.silver-peak.com](http://www.silver-peak.com)



# SD-WAN, IoT and life on the edge

ACCORDING TO RESEARCH by Juniper, IoT connections will reach 83 billion by 2024. As such, enterprises will have a host of non-traditional IT devices connected to their networks, generating a massive amount of data.

While the data created by IoT devices will provide a raft of operational benefits for businesses, the increase in traffic traversing network infrastructures will be huge. Failure to rearchitect network infrastructures to address these demands will negatively affect the quality of experience to users and IT organisations. To achieve the desired results, IoT devices must communicate with each other at the 'edge,' as well as with cloud applications. Some enterprises will facilitate this by transitioning to SD-WAN, others may turn to a hybrid WAN approach. Hybrid WAN architectures bridge from dedicated data centre and MPLS resources to public shared cloud resources. While the ultimate goal may be a broadband SD-WAN, hybrid networks enable enterprises to move at their own pace.

For organisations that choose to move away from a hybrid WAN configuration and towards a broadband SD-WAN, one aspect that will need to change is the traditional router-centric approach to the WAN. This was optimised for branch to data centre communications and, while in many respects it is inefficient and handles internet traffic poorly, it did the job as most WAN traffic was from branch locations to the data centre where applications were hosted and accessed. Today, the bulk of WAN traffic is cloud destined, the limitations of this approach become clear, and a more advanced SD-WAN infrastructure is needed, especially as organisations migrate applications and services to the cloud. Since SD-WAN also leverages 5G transport, this will aid the expansion of IoT, creating a virtual network of ultra-high-

speed connections across multiple devices.

There are several key reasons that SD-WAN can enable IoT initiatives:

- 1. Visibility** – SD-WAN enables real-time visibility into the network, taking the guesswork out of problem resolution, creating process efficiencies and enabling CIOs to analyse how IoT applications are consuming bandwidth.
- 2. Security** – IoT traffic can be isolated from all other traffic on the network, enabling network managers to assign unique security policies to this traffic and reduce the attack surface.
- 3. Orchestration** – increased number of connected devices, applications and users, requires a comprehensive, intelligent and centralised orchestration approach.
- 4. Speed** – a centrally managed SD-WAN enables network managers to quickly make changes, add new locations or new applications without the overhead of configuring changes on a manual device-by-device and site-by-site basis.
- 5. Intelligence** – advanced SD-WAN solutions that apply AI to fill in network services with better context and connectivity are becoming an important extension of SD-WAN capabilities.

This last point is crucial. In this intelligent way, the WAN 'speaks' to the edge devices in branch offices. It can work out what a device is based on its IP address. When combined with other multi-level intelligence, this enables automated policy orchestration across a large number of devices and applications. IoT is a huge market and will continue to grow. However, as the number of connections increases, businesses must consider whether their networks are keeping pace. If not, they risk burdening their networks with data, slowing down connections and hampering end users. An intelligent SD-WAN can address these challenges.

## Jeff Sharpe

Director IoT & Embedded Solutions, Supermicro

[www.supermicro.com](http://www.supermicro.com)



# How our “new normal” is creating even bigger 5G, Edge and IoT opportunities

IN THE LAST FEW MONTHS, many of us have mastered new video-conferencing skills, and shared an impressive amount of data. As the world is becoming increasingly data-driven, the World Economic Forum estimates that there will be 44 ZB of data generated only in 2020.

Companies and employees will rely more and more on virtual conferencing platforms and mobile communications. Latency, and system degradation from an abundance of users can be mitigated by the inherent capabilities of 5G and the number of opportunities is huge. As businesses will require more and more real-time decision-making, streaming data becomes paramount. With the continued democratisation of machine learning, 5G networks are designed to leverage virtualisation and containers on open hardware platforms – reducing dependence on proprietary legacy hardware.

5G implementations require expertise in optimised server hardware and deep virtualisation experience. Radio Access Networks (RAN) will be centralised and virtualised into server-based Centralised Units (CUs) and Distributed Units (DUs), slimming down fixed-function hardware into Remote Radio Units (RRUs). Intelligent Edge servers will support real-time applications with computing and AI inferencing. I

In addition, 4G Evolved Packet Core (EPC) functions will be replaced by 5G Core components running as virtualised network functions (VNFs) in data centres and the cloud.

To serve users with the best efficiency and deliver the promises of the much talked about 5G revolution, networks will need to rely on dependable data centres, the cloud, virtualised RAN, and the Edge. They'll also need to be able to deploy and manage IoT device deployments from the cloud to customer premises. 5G-based architectures are being used within AI applications and bring two critical elements for all industries: substantial technology and business transformation.

The technology transformation will deliver performance of data centre technology to the edge for increased speed, open-source software and standards as well as agile software-defined infrastructure supporting simultaneous apps.

The business transformation will help to solve ecosystem dependencies, partnerships and create new business model opportunities.

These latest technologies enable AI at the edge for improved education, machine learning and autonomous intelligence for logistics, manufacturing, health, safety and energy. Needless to say, this will improve our lives boosting the development of smart cities, transportation, retail and entertainment by providing seamless computer automation and interfaces via wireless technology.

One of the latest pieces of this technology that has made a step towards this new, more data-based world on the edge is a small

box. Supermicro calls this “the outdoor edge” or “data centre on a pole”.

This recently launched platform is a compact IP65-rated server for 5G RAN, AI inferencing, and other intelligent edge-focused applications based on Intel Xeon D processors or Intel Xeon 2nd Gen Scalable processors with broad configuration options.

These new systems feature ruggedized enclosures are built for harsh outdoor environments and support the industry’s movement toward open-source software and disaggregated hardware.

Supermicro embraces the industry movement to non-proprietary hardware platforms and the growing adoption of standardised system interfaces. With our membership in the O-RAN Alliance we support its initiative to promote a cloud-native, open 5G RAN architecture for the evolution of 4G to 5G networks. Since joining the O-RAN Alliance, Supermicro has developed reference solutions with several leading telecom operators and software stack providers.

The UK is one of leading markets implementing 5G-based networks and enabling the technology infrastructure for businesses and consumers. BT, Orange, Vodafone, Deutsche Telekom, Telefonica and others are all investing in not just their networking gear, but hiring top talent to create leading edge (pun intended) technologies for software-defined networks and applications.

As x86 compute designs migrate to the telco market with the roll-out of 5G networks, operators are faced with entirely new challenges. A lot of the new 5G applications will require high-performing GPU servers for machine learning/AI training in the data centre as well as GPU card options at the edge to improve local decision-making for time-critical applications. It is therefore a two-fold investment, in the data centre and at the edge.

5G is a true game changer. It will enable entirely new applications in so many areas and it will completely transform the infrastructure that powers it in due course.



**DW DATACENTRE SOLUTIONS**  
Developing digital infrastructure in a hybrid world

IN ASSOCIATION WITH

**DCS DATACENTRE SOLUTIONS**

**DW DIGITALISATION WORLD**



## Mike Wood

CMO, Versa Networks

[www.versa-networks.com](http://www.versa-networks.com)



## Connectivity is crucial

The Coronavirus pandemic has served to reinforce the value of cloud services, and the importance of having the right hybrid cloud and multi-cloud strategy backed with appropriate connectivity. Enterprise IT departments need to know that workers can access the applications, services, and data that they need from whichever data centre it resides in, from wherever they are located, which at the moment is likely not to be from the office.

The pandemic has also led to a rise in cybercrime and DDoS attacks as bad actors seek to profit from the chaos and uncertainty. Operating a hybrid and multi-cloud strategy in this climate means more than ever that security is of paramount importance. The traditional network security that enterprises have been relying on for years to protect data and applications was not, however, designed for hybrid and multi-cloud architectures. And it certainly wasn't designed for a scenario where very few workers are operating out of a heavily guarded head office with a well-protected private data in the basement. Traditional network security relies on the virtual private network (VPN). But with working practises having moved on, and workloads more likely to be spread over a number of cloud platforms, new thinking is needed.

Those stuck in the traditional VPN era will not enjoy the seamlessness and fluidity that a hybrid cloud strategy requires. Current network architectures can also have a negative impact on the user experience and on the performance of software-as-a-service based applications, as the centralized nature of the connectivity they provide can lead to disastrous bottlenecks. The enterprise is denied the speed, agility and cost efficiency that they hoped their cloud migration would deliver.

Certain enterprises have adopted a more software-defined approach as a way to gain more control over the management of their network. SD-WAN, for example, enables more granular control over how individual applications are connected. But too many first-generation SD-WAN solutions require separate security which demands that the enterprise deploy an additional security appliance at every site where Internet connectivity is needed. What they need is a way to combine routing, SD-WAN and security all in a single solution that is fully software-defined. This would make it easy to deploy, operate and secure the network for the on-demand, multi-cloud era.

Better still, a network solution that offers contextual visibility that draws on artificial intelligence (AI) and machine learning (ML) to provide automation. This takes security to the edge of the network where the greatest dangers lie.

Successful hybrid cloud deployment is all about eliminating the complexity of managing and correlating data from a mass of individual functions. It's about a solution that supports the software-defined enterprise to extend end-to-end, across branches, data centres, clouds, campuses, 5G, edge compute, mobile, wired and wireless.

Even when the pandemic is over, companies of all sizes, and within virtually every industry, will demand less complexity. They need applications to run seamlessly, and with reliable and consistent performance. Connectivity between data centre and enterprise can no longer be a business inhibitor. There will be increased demand for the consolidation of networking and security capabilities delivered through the cloud, with deep application-awareness that understands data context.

## Juan Perez-Etchegoyen

CTO, Onapsis

[www.onapsis.com](http://www.onapsis.com)



# What is the current state of cloud security?

THE CLOUD is safe(R) in many situations, as it provides standardisation and security features across the board. It does become a safer option for organisations that cannot afford to have an IT Security team looking at everything in the organisation and I would say that is especially true for SMBs. When we talk about the large enterprise sector, embracing the cloud does provide additional attack surfaces, especially due to the sheer complexity of the business applications and the different security models. Organisations must ensure they have the right visibility and controls around all aspects of the cloud components, combined with their on-premise environments. The answer to it is unsurprisingly, "it depends", but it depends on what, right? Everyone is using the cloud, in some shape or form, unless we are talking about a company that explicitly denies using the cloud (which is more and more rare nowadays). If it is email, file sharing or any other business application, organisations are using them all. And the cloud does provide benefits in terms of security, as patches are no longer an issue for the most part, and there are additional security-relevant features.

Despite that, adopting the cloud becomes more and more complex. We are talking about IaaS, PaaS, SaaS solutions, completely integrated and with different security models. This, combined with the on-premise environment provides an ever increasing attack surface that organisations need to secure, considering all angles of the cloud, especially including the cloud applications that run the business.

### What kind of applications do people still want to keep on premise?

It depends on the cloud/risk appetite of the organisation. Some organisations prefer to keep things on premise to keep a tighter control around its data. For example, European organisations are more hesitant to adopt the cloud than US companies.

Now, in regards to applications, it depends on the business process that is supported, for example if we talk about travel and expenses, there isn't an extensive need for changing that business process, therefore a SaaS solution through its standard processes and workflows might be a good fit. Other processes such as finance, manufacturing, sales or supply chain might require additional customisations to match the organisational complexity, especially on the large enterprise sector, therefore an on premise system might work better (or a system running in AWS, Azure or GCP, to that extent it would be equivalent).

### How are companies orchestrating workflows across on-premise and public cloud environments?

Orchestration in hybrid environments is key to a successful cloud strategy. There are a number of Integration platforms that are leveraged to do this. In the case of SAP, I can mention SAP Cloud Platform Integration and SAP Process Integration, which leverage a number of technical components to successfully integrate pure cloud SaaS applications with on-premise environments. There are other vendors in the Integration platforms space that we have seen being used by organisations.

### How big of a factor is GDPR when it comes to hybrid cloud?

I think GDPR is complex in itself, regardless of the cloud service model that you run. When you are managing a hybrid environment, organisations need to go through their on-premise systems to identify personal information across all of those applications but when it comes to cloud, now companies need to involve the Cloud Service Providers, so they do their due diligence and due care to ensure the provider is doing a good job protecting that data, which is ultimately still the responsibility of the customer.

**Simon Ford**  
AVP EMEA, CyrusOne  
[www.cyrusone.com](http://www.cyrusone.com)



## The lasting legacy of Covid-19

MUCH HAS BEEN WRITTEN in recent months about the impact of the Covid-19 crisis – we're all well aware that social distancing and reduced access to workplaces has forced many organisations and enterprises to move operations and processes online for staff who are now working remotely. In some cases, naturally, the added strain on IT infrastructures has exposed areas of weaknesses or lacking flexibility within the business to adapt and cope with the extra online workload.

An earlier 2020 report by data virtualisation firm Donodo found that 75 per cent of organisations surveyed had already deployed workloads into the cloud in 2020, a ten per cent increase from 2019 (worth noting that this trend was before Covid-19). Impressively, despite the huge workloads transferred to cloud since the crisis, the infrastructure has coped, largely owing to the high levels of redundancy in the global data center network.

Cloud service providers have played a major role in supporting businesses as they continue to offer increasingly diverse services focused on individual business models. This situation has forced a rethink about how the relationships between customers and cloud provider should be structured. We've observed a couple of significant changes in the data center sector that are worth noting.

### Flexibility

Hyperscalers and other large enterprises have seen unprecedented usage of cloud services and applications during the crisis. For instance, Microsoft reported over 700% cloud usage in many regions across the world. This added pressure on the network has forced cloud service providers to exercise their lease and capacity option to increase their data storage and

processing requirements to keep pace with growing demand. However, while these large businesses have weathered the current storm, they are now quickly banking high volumes of ramp up space to anticipate long-term demand as this shift and migration to cloud environments solidifies and adoption accelerates. As the crisis deepens, the challenge to the industry is whether this further flexibility is delivered when and where it is needed the most.

### Customer Relationships

During this crisis, we have seen customers' procurement attitude shifting too. Companies are interested in working with data center providers in a new way that changes the relationship into a partnership. Rather than cherry pick from an available list of offerings, enterprises and hyperscalers are now interested in working with partners with a more comprehensive solution tailored to their requirements. We anticipate that large-scale companies will look to consolidate the number of suppliers and solutions providers as they continue to seek out bespoke cloud models that are agile and flexible enough to meet future challenges.

Modern organisations require incredibly elastic cloud environments and agile infrastructures. While it is important to note that one size does not fit all, this pandemic is also a test of an organisation's ability to swiftly, securely and systematically scale online applications and communications software, and provide remote access to all employees. This situation is still evolving and there are many unknowns regarding what ways of business will look like once the dust settles. However, we have been impressed with how flexible and efficient the data centre industry has been in responding to the crisis and feel confident that it will continue to adapt as we enter this 'new normal'.

## Rob Mather

Solutions Director, Power Control Ltd

[www.powercontrol.co.uk](http://www.powercontrol.co.uk)



### A shift in demand

THE GLOBAL PANDEMIC has dramatically changed the data centre landscape. Predictions that were made even a few months ago have been thrown into disarray and opinions are more rife than ever before.

Like many other industries, there is a lot of uncertainty in the data centre market. One thing that I do think is for certain is that we will see growth at both ends of the spectrum. The shift to the cloud has already impacted internal data centre requirements, there has been an increase in regionalised data centres and the hyperscalers are just getting bigger.

The enquiries that we are receiving from clients and consultants at design level all support this prediction. I would go as far to say that we could end up seeing 10kW on every street corner. Strong advocates of the potential 5G-era may disagree but for the moment I don't think anyone really knows what the role of 5G will be."

We have to be mindful though, that with so many available solutions and companies adopting different strategies for their ICT requirements, there is no one size fits all approach.

Despite many differing physical requirements, sustainability is high on the priority list and most definitely a hot topic in the industry. It is the step above just being able to achieve greener solutions. The industry is focussing more on the long-term operational cost viability. For example, an onsite standby diesel generator often requires planning permission, will the expected "green tax" still make this a viable option or will it become unsustainable?

Aligned to this, we also need to consider the inevitable regulations, which are sure to come into force sooner rather than later. The coronavirus pandemic has highlighted the importance of, and the role that data centres play in the wider economy, and with utility consumption set to increase, the government is likely to introduce a levy of some sort. The economic repercussions will ripple through the entire supply chain and it will be interesting to see what impact it has on data centre developments as a whole.

For the industry to remain self-regulated and avoid any government imposed regulations it will have to address sustainability and efficiency concerns that the wider public have, as well as making sure they are reliable and secure. Ultimately will the government allow an industry so crucial to nearly every aspect of the day to day operations of the country to be self regulated?

Power Control is poised for a shift in demand. We know that the industry will need knowledge, support and above all flexibility. The next chapter for data centres is going to be interesting and we will be watching the sector very closely."

When it comes to understanding power protection and the equipment available, Power Control has had over 26 years of experience in the field and has delivered countless solutions for data centre environments. It works across the entire scale of facilities. The company is highly influential within its field.

We have partnerships directly with data centre owners as well as with contractors and consultants at all levels.





## Claudio Scola

Head of Product Management EMEA

[www.centurylink.com](http://www.centurylink.com)



## In, out, shake it all about!

EDGE COMPUTING is a major trend for the data centre industry. While traditional data centres are still the core computing power for enterprises, we will increasingly see edge computing become an integral part of their data centre strategies.

Forward-looking companies seeking the advantages of next-gen technologies such as 5G, IOT and AI, and Machine Learning applications increasingly need distributed solutions to enable them to acquire, analyse and act quickly on the huge volumes of data they create. Edge takes computing out of a centralised model and places it closer to customers' digital interactions. By processing data closer to the edge of the network where it is collected and consumed, edge can greatly increase speed and responsiveness and drive innovation.

This doesn't mean that data centres will become redundant. Business leaders will continue to build and operate their own data centres and move compute resources to public data centres and specific applications to cloud service providers. However, they will look to hybrid solutions that combine a traditional data centre model and apply edge for performance-based needs.

It is clear that edge will rely on a robust level of infrastructure investment in network and compute and within the data centre, network upgrades will be needed. For new and emerging technologies to offer their full potential, they must be underpinned by low-latency, high-capacity networks that can transfer data quickly and highly reliably to where it needs to be. In addition, as data volumes rise, quickly and often unpredictably, the ability of the network to scale with these changes also becomes more essential. So, we will see significant developments in this space.

Last year, CenturyLink announced the rollout of its edge-focused strategy to help customers to advance their next-gen digital initiatives, beginning with a several hundred-million-dollar investment to build out and support edge compute services. This initiative includes creating more than 100 initial edge compute locations across the U.S., and providing a range of hybrid cloud solutions and managed services. The edge compute services blend compute, storage and networking capabilities into one package.

This expansion allows customers to leverage CenturyLink's highly diverse, global fibre network with edge facilities designed to serve their local locations within 5 milliseconds of latency. With this infrastructure, companies can complete the linkage from office location to market edge compute aggregation to public cloud and data centers with redundant and dynamically consumable network. CenturyLink's fibre network is key to the company's effort to move more compute resources to the network edge. The company is investing heavily in its infrastructure to create an extensive 4.7 million-fibre mile intercity fibre network across the US and Europe. The company is also adding on-net fibre buildings and data centres, and currently has more than 150,000 buildings and over 2,200 public and private data centre connections on-net globally.

CenturyLink's thousands of secure technical facilities combined with its network of 450,000-global route miles of fibre, expertise in high-performance cloud networking, and extensive cloud management expertise make its investment in the rapidly emerging edge compute market a natural evolution for the company.

## PHYSICAL SECURITY + SAFETY

Fencing, access control systems, biometrics, fire suppression, compliance and regulations – not, perhaps, as exciting as AI, cloud and the edge, but a crucial part of the overall data centre operational landscape. After all, there's no point in having the very best data centre halls, housing the very best IT infrastructure, running 'world-beating' applications, if anyone can wander in off the street, stroll into the computer room and play havoc with servers, network switches and storage boxes they house! Or, if fire, floods or other extreme conditions are allowed to cause havoc within the data centre.



IN ASSOCIATION WITH



**DATA CENTRE**  
**SOLUTIONS**



**DIGITALISATION**  
**WORLD**



# Why physical data centre security is as important now as ever



We live in an information age where the operations of data centres are vital to the functionality and continuity of our global economy.

**BY IAIN MORAN, DIRECTOR AT ATG ACCESS, REPRESENTING HS SECURITY.**

ACCORDING to The Data Economy Report, by 2025, in the UK alone data centres will be responsible for storing data worth circa £103 billion pounds annually. Globally the IDC report, 'Data Age 2025' estimates that the volume of data generated will have grown by 378% by 2025, compared to 2018 (33 Zettabytes to 125 Zettabytes).

The theft of information assets has serious consequences; in the UK according to IBM research the average cost of a data breach to a business has grown to nearly £2.7 million. As well as causing economic loss a data breach can also cause significant reputational damage due to loss of confidence from customers.

With such economic value at stake for businesses, the physical security of data centres must not be an afterthought but rather an integral part of how these facilities are designed, developed and built.

Often positioned in anonymous-looking buildings, these facilities are responsible for the protection,

storage and transfer of information which enables multiple critical systems to function around the world, including for example financial, healthcare and defence.

All of the above has the potential to make these facilities a potential target for malicious attack, resulting in serious and far reaching consequences.

By ensuring that the physical security element is considered as an integral part of the data centre there is opportunity to create a holistic layered approach which will effectively delay and disrupt any attempted attack.

With this in mind, below are two of the main physical security threats faced by data centres today and how these threats can be overcome.

## Restricting vehicles

Vehicles can cause extreme damage to data centres if the right security measures are not in place. Having breached the outer security perimeter, the vehicle could then proceed to damage the building itself if secondary physical protection was not installed.

In 2007, the physical impact of a travelling truck was the cause of a data breach for a large US data centre operator. The vehicle was driven into a power transformer cutting the power to the company's Dallas facility. Whilst the data centre did have a back-up generator, two of the chillers within the centre failed to re-start causing the business to take customer servers offline for a period of time.

This event demonstrated the damage posed by hostile vehicles to data centres, even if at times it is an unintentional accident. The company consequently had to restrict their services and deal with reputational damage, all because a vehicle was able to enter an area which should have been better protected.

As well causing damage externally, the movement of the vehicle within a protected area, could be a pre-cursor to physical entry into the building by unauthorised personnel. Even more worryingly, without the right security measures in place vehicles can be rammed into the building to intentionally cause damage and then launch an attack within the data centre facility.

## How can vehicles be restricted?

When considering the restriction of vehicles, a layered approach is a tried and tested methodology that is used on many different types of building. This can include measures such as HVM rated gates and bollards, perimeter fencing, associated gates, turn styles, security rated cages and external/internal doors implemented around the facility.

The first level of protection would come from the perimeter fencing and gates with the second layer

of protection being provided by bollards which can be either fixed or retractable depending on specific requirements.

Perimeter fencing and gates provide the initial layer of security to prevent unauthorised access. There are fencing solutions available that have anti-climb features incorporated within their design through mesh composition which can also offer strong visibility for CCTV and security patrols.

Introducing properly tested fencing and gate solutions around data centres will not only prevent physical damage to the building through a vehicle impact but can also stop unauthorised visitors being able to get closer to the main entrance.

Tested and certified bollards can then be used to provide the next line of defence, with retractable bollards providing a flexible security approach. When these are operated by security personnel, a warning light system can be used once access is assessed and granted but can also be immediately prevented. In this scenario, the bollards can be left up and the appropriate action can be taken.

Dynamic bollards and road blockers can also be used to enforce a 'tiger-trap' stop-and-search function for all vehicles entering a site.

These solutions provide a multi-layered security approach which when combined with various access control and detection systems monitored by the on-site security teams give the data centre a complete robust physical protection.

## Why does access need to be controlled within the facility?

In addition to perimeter security, installing measures within the entrance area of a building is essential. This helps complete the layered security approach which has commenced at the physical outer perimeter making it difficult for unauthorised personnel to gain access to restricted areas.

The importance of monitoring access to a data centre building was a problem for a national telecom's provider back in 2011 when unauthorised personnel were able to force access, stealing computer equipment and network hardware. This caused immediate disruption to the phone provider's customers who experienced the loss of SMS, internet and phone calling services.

This highlights the importance of installing comprehensive security at each and every ingress and egress point. By making it difficult for unauthorised personnel to enter a site, data centre security managers can increase the length of time that security professionals have to react to a threat, and most importantly, reduce the risk damage to the building asset and the services being provided.



There is potential for vehicles to be used to force access, as well as individuals walking into the main building if the appropriate physical measures are not in place



### How can access by foot be regulated?

In order to achieve this, a range of authentication points and 'air lock' doors should be installed. Implementing this multi-layer system can help to restrict access and also identify whether the individual trying to gain access is authorised to do so.

Having a visitor buzzer followed by an inner door can be used to grant access but also keep visitors separated from the general employee area. This allows for rigorous identification and enables security operators to establish who works for the data centre operator and who is an external, and potentially unwarranted, visitor.

Similarly, installing a floor-to-ceiling turn stile door at each entrance point can reduce the chance of an individual tail gating an authenticated user. This measure can be backed up by a 'mantrap' door. These doors are separated by an 'air lock' where one door can only be opened once the first one is closed. If an unauthenticated user does attempt to tailgate a data centre employee as they legitimately access the building, these doors allow security personnel to stop the second door being opened and can deal with the situation accordingly. For many years, the Critical National Infrastructure sectors have been building, refurbishing and hardening its infrastructure to ensure

that apparatus in data centres and access points around a facility, are protected against all levels of trespass, malicious damage and terrorism.

Whilst data centre managers will be looking to meet the legislated security standard of the Loss Prevention Certification Board LPS 1175, it is essential for a range of measures to be implemented that not only meet these standards, but help to physically protect and monitor the ongoing physical security risks at these sites.

Within data centres, a multi layered approach should be deployed to make sure that protection at each layer is robust. There is potential for vehicles to be used to force access, as well as individuals walking into the main building if the appropriate physical measures are not in place.

Most importantly, attacks can be carefully prepared, and if ill-intentioned individuals know the internal location of key assets, then a coordinated attack using a vehicle and individuals on foot could cause extreme damage to a business and its customers. That's why a holistic approach, to physical security, which considers each perimeter layer both externally and internally all the way up to the key assets, will provide the most robust, security solution.

## Dave Klein

Senior Director of Cybersecurity Engineering,  
Guardicore

[www.guardicore.com](http://www.guardicore.com)



# Vollgar Attacks SQL-Server Databases

IF YOUR ROLE involves data centre security, you might want to read April's report from cyber security researchers, Guardicore Labs. They've uncovered a widespread Botnet attack nicknamed 'Vollgar' which is targeting MS SQL-Server databases. Vollgar has been spreading fast, compromising around 3,000 machines daily for the last two years. It installs Remote Access Trojans (RATs), cryptocurrency miners and other nasties. (Full Guardicore Labs report available on request).

Dating back to at least May 2018, the campaign, which appears to stem from China, uses password brute force to breach victim machines. The attackers are especially targeting data centres in the industry sectors of healthcare, aviation, IT & telecommunications and higher education. Having first seen this campaign through several detections within Guardicore's Global Sensor Network, the Guardicore Labs team began to monitor the attackers, eventually infiltrating their systems and gaining access to all of their tools, logs and activity. At its core, the Vollgar attack group ran a network of 120 (predominantly Chinese) Command and Control machines. These, like many victim machines, were hijacked workloads owned by others.

Curiously the Vollgar attacks had a short dwell time with 60 per cent of victim machines being compromised for less than seven days with the remaining 40 per cent averaging a little longer than 14 days. With only about half a million MS-SQL workloads in the wild, reacquisition was an important part of the attack campaign and 10 per cent of the time the criminal team was able to re-compromise former victim machines. While it was relatively successful, with average dwell time in 2019 averaging a bit longer - between 58-60 days globally - one can only imagine how much larger this campaign might have been if the attackers had

been more prudent in their use of computing resources and had not given themselves away. What makes this attack especially interesting from a data centre administration perspective is to note how relatively easily attackers continue to compromise and hijack data centre resources. These MS-SQL servers were open to the Internet, had weak passwords and ran with escalated privileges allowing relatively easy, unfettered access.

We can protect ourselves from attacks like Vollgar by implementing real-time and historical visibility combined with software-defined segmentation. This allows administrators to easily understand and protect their workflows to prevent compromise, or to at least minimise impact. Such visibility and segmentation capabilities should not be disparate for individual platforms and silos. Rather, the approach should be agnostic, decoupled from the underlying operating systems and platforms to operate in a single, unified manner. Initially, even incorporating just a rudimentary I segmentation capability across an environment reduces attacker dwell time, because lateral movements found in attacks like Vollgar can then easily be detected and alerted on. Further easy things to implement are to shore up strong password enforcement and combine this with dual factor authentication, vulnerability scanning and patching. There is also a need for better elevated account control/expiration procedures, better certificate management practices and control of enterprise services like DNS, Remote Access (SSH/RDP), AD and other critical services.

Our biggest lesson from Vollgar is that the biggest threats to data centres are ironically often the easiest things to remediate. It really isn't all that hard. Owners of breached data centres are rapidly running out of excuses.

## Garry Connolly

President and Founder, Host in Ireland

[www.hostinireland.com](http://www.hostinireland.com)



## In Data We Need to Trust?

IF THERE IS ONE LESSON to be learned from a global pandemic, it is how essential data centres are as part of our global infrastructure. They have been a key piece of the jigsaw allowing us to live our lives in these disrupted times, as has the data that is enabling us to live, work, play and educate. With this move to a predominantly digital world, the question of responsibility for both the security and privacy of data becomes an important one.

Coalescing around the concept of trust, consumers, service providers and businesses all have a role to play when it comes to protecting data. Policies that put security by design and privacy by design as fundamental pillars are needed. Particularly in an environment where organisations are adapting on the fly, consequences can be catastrophic if we don't do the right thing in the right way and in the right order.

Security by design consists of the policies and procedures needed to protect the data centre campus from external threats and attacks. This includes the physical components - connectivity, infrastructure, cabinets, etc. - that need to be protected from any sort of theft or malicious attack. It is more straightforward in that many long-standing best practices have been developed across the industry to protect against these types of security concerns.

Privacy by design, on the other hand, secures the data you can't physically see. It takes into account protecting data from bad

actors with dishonourable intents, as well as human error from both individuals and within companies. It also needs to account for unintended consequences of decisions that are made for the greater good. As we move to a phase of contact tracing in effort to manage further COVID-19 outbreaks, balancing the need for data privacy with the need for containing a pandemic until a vaccine or treatment is found will continue to be a tug of war. Even with our best interests as a society at the heart of the debate.

Now more than ever, organisations will need to demonstrate they understand the seriousness of this responsibility with the actions they take. Our dependence on data is only going to increase in the short term given two-thirds (69.0%) of Irish businesses implemented remote working at the start of the lockdown.

As a result, answers to key questions - can I trust the data; can I trust how it was collected; can I trust how it was distributed; can I trust how it's stored - will become more important for consumers and businesses alike.

How organisations approach respect for data security and privacy will be a big part of the decisions on where to place data moving forward. Content generators, service providers, data centre industry, and even consumers, all have a responsibility to ensure data is safe for everyone to use. As we navigate the waters of our new normal, the concept of trust will be greater than ever before.

## Peter Jackson

Managing Director, Jacksons Fencing

[www.jacksons-security.co.uk](http://www.jacksons-security.co.uk)



# The Physical Side of Cyber Security

YOU WOULD be forgiven for thinking that, on paper, physical security for data centres is straightforward. However, in practice it's complex. On top of protecting the external perimeter, thought must also be given to myriad factors such as access control, hostile vehicle mitigation (HVM), noise reduction, protecting power infrastructure and standby generators, and localising security devices to operate independently of data centres themselves. How a site looks is more important than many may think. Getting it wrong advertises that you're protecting a valuable target. The necessary heightened security – 4m high security fences, coils of razor wire, etc. – can jar with the local aesthetic and therefore may not be appropriate in all circumstances. Getting the balance right is crucial.

A high security timber fence, for example, can be both aesthetically pleasing and disguise its security credentials. Its pleasant natural façade provides a foil to the stern steel bars and mesh often seen with high security solutions. Of course, it's still important that fencing serves its primary purposes, so make sure you refer to certifications to establish a product's security and, where required, acoustic performance.

Due to their importance for the UK's economic stability and prosperity, data centres require appropriate security investment and design. Cutting corners leads to greater long-term expense and increases the likelihood of highly disruptive attacks.

Fortunately, guidance is available through independent accreditations and certifications such as the Loss Prevention Certification Board's (LPCB) LPS 1175 ratings, the PAS 68 HVM rating, CPNI approval and the police initiative Secured by Design (SBD).

These bodies employ thorough technical evaluation work and rigorous quality audit processes to ensure products deliver proven levels of protection. With untested security measures, you will not know whether a product works until an attack occurs. Specifying products accredited by established bodies removes this concern, and will guarantee a long service life, keeping overall costs down.

However, just installing security measures will not guarantee 24/7 protection. A well-maintained perimeter ensures physical security is performing equal to its designed specification. Inspecting and maintaining the fence line needs to be a deliberate, scheduled event and from our experience that's rarely the case. According to research we conducted recently, 63% of companies admit to never testing physical security.

Look for attempted breaches and note if foliage, weather conditions or topography changes have affected security integrity. Check all fixtures and fittings, looking for damage and corrosion and clear all litter and debris.

Physical security is an important part of protecting a data centre, but should be complemented with other measures according to a site's needs. Implementing an integrated strategy helps protect against a diverse threat range. This should include security personnel and electronic security where appropriate.

Data centres are faced by an incredibly diverse range of threats including activism, sabotage, trespassing and terrorism. By working with government bodies such as the CPNI and certification boards like the LPCB, specifiers can access a vault of useful knowledge and advice. This will guide them to effective, quality products suitable for their situation.



## Alan Elder

Fellow – Engineered Systems, Johnson Controls

[www.johnsoncontrols.com/industries/data-centers](http://www.johnsoncontrols.com/industries/data-centers)



## Focusing on fire suppression

MODERN data processing centres form a vital part of the world's critical infrastructure and have changed considerably as the need for ever greater data processing speeds has evolved, leading the industry to develop storage equipment or hard disk drives (HDDs) designed to enable huge amounts of data to be processed at breath-taking speed. With that, an unforeseen consequence was the development of equipment that was much more susceptible to vibration through ever smaller clearances between the mechanical moving parts of the drives and the read-write surfaces. Several years ago, reports began to emerge of incidents occurring in data centres in connection with loss of data and/or damage to HDDs, following the release of a gaseous fire suppression system. After much research into this area it was concluded that vibration caused by the noise created by the discharge of the gaseous systems or certain alarm devices was the likely cause.

Johnson Controls (JCI) in conjunction with Michigan Technological University (MTU), researched this subject (see JCI White Paper – Effects of Sound on Computer HDD's and Risk Mitigation Measures; <https://www.johnsoncontrols.com/insights/2016/buildings/features/impact-of-sound-on-hard-disk-drives-and-risk-mitigation-measures>) and amongst many other aspects, included the analysis of HDD performance, the impact of sound on that equipment and the acoustics associated with gaseous extinguishing system discharges. JCI concluded that several factors associated with the discharge were critical in understanding how best to mitigate the effects of the discharge, to safeguard business continuity, reliability and uptime. This included the noise of the discharge itself, the vulnerability of the HDD to the effects of that noise and the general acoustics within a data centre. Initial studies focused on inert gas systems,

since the incidents reported, pointed primarily to this technology and it concluded that several factors could be responsible for creating excessive noise vibration, which included the noise from the discharge nozzle (Sound Power), the distance of the discharge nozzle from the sensitive equipment, the level of shielding enjoyed by that equipment through the server cabinets and the general acoustic in the data centre, resulting in Sound Pressure being experienced at the HDD. The frequency of the sound was also found to have an effect on the performance of the HDD.

It was clear from these studies that both the Sound Power (produced at the source) and Sound Pressure (received at the location of the HDD) needed careful attention. JCI focused on these two areas to resolve this problem. Acoustic nozzles were developed by JCI for all its Inert Gas system platforms and an acoustic calculator designed to estimate the Sound Pressure at the HDD.

The nozzles are designed to reduce the Sound Power at the point of discharge, whilst the acoustic calculator takes into account the sound performance of the nozzle at the specific system flow rate, the distance from the nozzle to the HDD, the quantity of sound absorbing equipment in the space and the structure of the data centre in order to make an estimation of the Sound Pressure and the frequency, for that very specific system. JCI was meanwhile developing a new higher pressure halocarbon system, SAPPHIRE PLUS, a 70 bar system, that gives significant benefits over lower pressure systems, allowing the use of more remote and centralized storage locations, with longer pipe runs and selector valves to protect multiple hazards from a single bank of containers.