

Datacentres of tomorrow

What the future holds



In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

In this e-guide:

Ensuring uptime, resiliency and high-performance are top priorities for datacentre operators the world over, and is why it is so important for them to stay on top of the latest thinking in hardware design, server setups and facility management.

While some enterprises are undoubtedly gearing up for a wholesale, all-in move to the public cloud, others have applications and workloads that will need to remain on-premise for a long time to come. And, as a result, are investing in new servers and other assorted datacentre hardware accordingly.

In this e-guide, we take a look at some of the technologies that are being widely-tipped to become a mainstay of datacentres in the future. At the same time, it also touches upon some of the new and emerging technologies that look set to shake-up the way datacentres are managed, monitored, powered and cooled as well.

In this e-guide

- Lean, mean and green: The datacentres of the future
 - How AI and data analytics are driving composable infrastructure adoption in datacentres
 - How AI and data analytics are driving composable infrastructure adoption in datacentres
-

Lean, mean and green: The datacentres of the future

Fleur Doidge,

It's not easy being lean, although doing more with less sounds simple enough. The sustainability-seeking datacentre operator needs to pinpoint waste and reduce it, without losing value from other parts of the system. Yet the race is on to support more capacity and [high-performance computing](#) than ever.

“The industry wrestles with capacity challenges and advanced applications that are forcing significant changes to datacentres of all shapes and sizes,” says [Vertiv](#) CEO Rob Johnson. “The message to datacentre equipment providers is clear: the status quo is not acceptable.”

According to French thinktank The Shift Project's 90-page [Lean ICT report 2019](#), digital energy consumption rose globally by 9% per year between 2015 and 2020, with digital technology expected to be responsible for about 8% of greenhouse gas (GHG) emissions by 2025.

The thinktank's research pegged the average 1MW, 1000m² datacentre in 2019 as achieving a power usage effectiveness (PUE) of 2.0. Meanwhile, risk-averse

In this e-guide

- ▀ [Lean, mean and green: The datacentres of the future](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

operators typically err on the side of maximum uptime and can be over-provisioned.

[Jennifer Cooke](#), research director of the cloud-to-edge datacentre trends team at IDC, notes that although many are making sustainability announcements, there's still a lot of waste to reduce before considering entirely new solutions.

“Are you making good use of the energy you have? You can have giant facilities cooled down to the temp of a meat locker, only using 40% of the space, with 25-30% of the servers running with no one knowing what they're contributing to,” she tells ComputerWeekly.

Datacentre operators probably have more of an understanding than other businesses – and they can have a relatively standardised environment, with tighter controls and processes than mainstream enterprises.

Large hyperscalers may be able to devote a whole area to high-performance computing (HPC) – other enterprises may need to have it next to general business applications that must be kept running.

New technology will be required to tackle that disruption, confirms Cooke, but small tweaks over time can trim a lot of fat and add up to big cost savings.

“Liquid cooling has a lot of promise. It's kind of new in the datacentre, but more workloads are trying on AI [[artificial intelligence](#)], GPUs and so on, heating up the datacentre,” she says.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

The need to streamline operations

Virtualisation meant operators no longer had to buy new “stuff” for each new workload. Now, the industry needs to achieve a similar result around the physical facilities themselves.

This might include creating something like a digital twin of the datacentre to enable modelling of data volumes, power consumption, temperature controls and so on. It might also mean collaborative efforts between risk-averse datacentres protecting their own mission-critical environments and suppliers on big data projects.

“I would say we’re on the cusp of having the right technologies, with a lot down to process as well. We have cloud-based datacentre management tools, but getting people to use these consistently and getting datacentre operators OK with sharing data are big hurdles for many,” says Cooke.

Smart datacentres with sensors all around the racks can pinpoint heat emissions linked to specific workload patterns and temperature changes, disrupting other parts of the datacentre environment in situ. Here’s where AI-assisting cooling technology might come in.

Of course, energy-intensive AI adoption itself spurs [datacentre efficiency](#) requirements. Facilities can no longer remain defined by their ability to house

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

high-performance computing and power-dense equipment. But IDC says organisations are at an early stage here, with Europe potentially a little ahead and the Asia-Pacific region further behind.

“Despite greenwashing, companies that reduce waste and try to use their resources better will save money, and that makes sense,” says Cooke. “That’s going to be the tipping point for ‘I know I need to do this to get investors and to attract talent’ – students are signing pacts not to go to a company unless they have aligned sustainability goals, and so on.”

Coming full circle

Societies are embracing the [circular economy concept](#), and datacentres must follow. Cooke says suppliers such as HP might blaze a trail, perhaps because they’ve long worked with used equipment strategies and as-a-service models.

“It’s not all about cloud and cloud providers. Organisations are trying to get cloud to run on-premise, so look at how you can do that efficiently and what to do if you’re not an expert,” she says. “[District heating](#) – taking ‘waste’ heat from one place and using it another – might be one idea.”

Datacentre operators must take a good look at their hardware and performance requirements. Rather than going out and buying the datacentre equivalent of a new Prius, ask whether the old banger still has something to offer. Even if an

In this e-guide

- ▀ [Lean, mean and green: The datacentres of the future](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

old rack of servers cannot support new requirements, can it be cascaded into another use?

“It’s really hard to recycle tech gear – it takes a lot of energy and it’s hard to use the components for something else. It makes sense to keep it in use longer or find a new home for it,” says Cooke. “How can waste from one process become inputs for another?”

“It’s really hard to recycle tech gear; it takes a lot of energy and it’s hard to use the components for something else. It makes sense to keep it in use longer or find a new home for it”

Jennifer Cooke, IDC

Knowing exactly what to do to trim energy use in specific instances is difficult as good research-based data is lacking. The Shift Project discovered some 170 papers dated 2014-2017 that simply regurgitate findings from other papers without relevant context or examination. In addition, sampling methodologies and heuristics are often obsolete.

Rigorous quantitative clarification and benchmarking of the direct environmental impacts of all digital technology. Quantitative measurement and analysis of the

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

impact of investment policies, management practices and company practices is also lacking, the thinktank concludes.

Cisco has estimated that 67ZB (zettabytes) of “useful” data will be produced by the internet of things (IoT) and industrial internet of things (IIoT) sectors in 2020 – 35 times more than the storage capacity planned in datacentres at that time. New architectures from edge to [fog computing](#) and additional storage capacity, based on SSD18 including 3D NAND19, are probably needed.

Mike Mattera, director of [sustainability at Akamai Technologies](#), says change is absolutely achievable even with older equipment. In one US datacentre, it achieved a PUE of 1.09-1.15 in 2019 using a mix of outside air and direct expansion (DX) cooling.

“After seeing the results and understanding the operating heat tolerance of our hardware, we can operate warmer than what you would traditionally see,” says Mattera.

“In addition, we reduce our need for power through software and hardware advances. Since 2015, Akamai’s platform has used 61% less energy per gigabit of network capacity while still growing by over 182%.”

In this e-guide

- ▀ [Lean, mean and green: The datacentres of the future](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

Future forward in Cornwall

Operators need to work out where they are now, where they want to be, and how to get between those two points. But without better data, it's hard to identify the best levers for improved energy efficiency, which means falling back on anecdotal evidence such as case studies.

[Chris Roberts](#), head of datacentre and cloud at [Goonhilly Earth Station](#) in Cornwall, pinpoints the difficulty. "You're trying to save the world yet need all these hugely power-intensive AI models, so you've got to look at efficient ways of doing that," he says.

"Sometimes that means not taking small steps but big adventurous steps, and at Goonhilly, we've got the closeness to the wave project and wind farm, and ultimately our ambition is to become a kind of supercomputer hub starting to address some carbon emission challenges."

Goonhilly Earth Station functions as an advanced information hub for satellite observation data streams, covering some 168 acres (with its own herd of alpacas to nibble the grass). Huge data volumes need timely processing and they consume a lot of power.

"For example, we're deploying deep-earth observation on Nvidia on a unit geared around things like optimising crop yields. That in itself would generate

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

more heat, so we use more efficient ways to cool it,” he says. “In the standard workplace server, you can use lower-power chips and so on. But with our kinds of workloads, you have to work quite hard to mitigate that – hence immersive cooling.”

As workflows rose, Goonhilly moved to liquid cooling, using [Submer](#)’s non-conductive biodegradable fluid, with a posted efficiency gain of around 65% compared with standard cooling. “We looked at firms like 3M, but they’re not biodegradable,” says Roberts.

Goonhilly’s new datacentre launched mid-2019, superseding its largely traditional historic (yet relatively efficient) comms room. It uses lower energy platforms wherever possible, but must work hard to develop the most efficient processes, looking carefully and continually at the way the infrastructure is managed, from cooling to unnecessary energy expenditure.

“We’ve got space for 90 racks in the new datacentre in the first data hall, and there’s capacity to expand that at least a few times. That’s why we’re looking at new technologies,” says Roberts.

Energy efficiency and sustainability are overall drivers, with the datacentre developing a denser environment over time to further enhance it, with power requirements expected to go on rising. The station uses its own on-site wood pellet biomass generator, with additional power acquired from other renewable sources.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

“The next stage is talking to the wind farm half a mile away about how we can take in 33kPa [kilopascal] rather than low voltage, and the wave project. But we’re almost waiting for it all to catch up with us, because we’re not chasing these standard workloads,” says Roberts.

Goonhilly is identifying areas on-site for more solar panels, supporting its full power requirements of 500kW, pushing energy efficiency higher still. In 12 months, it could become the world’s first carbon-neutral supercomputer and 100% solar datacentre.

“There are other solar-powered supercomputers, but in the fine print it’s only 50%. So this will be the first,” says Roberts. “We’re not as advanced in some ways as, say, Google, but we’re using the tools we’ve got and it’s absolutely about the future.”

In this e-guide

- ▀ [Lean, mean and green: The datacentres of the future](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

▀ How AI and data analytics are driving composable infrastructure adoption in datacentres

Daniel Robinson,

Enterprises are increasingly having to deal with growing data volumes, fuelling demand for advanced analytics and [machine learning](#) tools to help them make sense of it all.

This, in turn, has placed new requirements on IT infrastructure to cope with the computational demands of such techniques.

It is now a decade or so since the notion of “big data” became a hot topic among CIOs and business decision-makers in enterprise IT, but many companies are still struggling to implement a successful strategy to make full use of the data they have, and become more insight-driven.

The data springs from numerous sources, such as machine-generated or sensor data from the internet of things (IoT) and other embedded systems, transactional data from enterprise systems, or data from social media and websites.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

As a result, enterprise workloads are evolving beyond traditional ones that revolve around structured datasets and transaction processing, and are starting to incorporate analytics and other techniques, such as artificial intelligence (AI).

According to IT market watcher IDC, AI will be a core component of enterprise workloads by 2024. It believes that for three-quarters of enterprises, 20% of their workloads will be AI-based or AI-enabled, and 15% of the IT infrastructure will be accelerated by AI.

However, organisations are finding that integrating advanced analytics and AI techniques – including machine learning – into workloads can put a strain on their IT infrastructure.

Drawing parallel lines

Traditional central processing unit (CPU) architectures, in particular, have proved less than optimal for some of these techniques, which often call for a high degree of parallelism.

It quickly became clear that this was the sort of problem that graphics processing units (GPUs) could handle. Designed to offload the burden of graphics processing from the CPU in games, GPUs have a lot of relatively simple processor cores, and can handle a large number of computations in parallel.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

Other [hardware accelerators](#) have also been added to the mix, such as [field programmable gate arrays \(FPGAs\)](#) and application-specific integrated circuits (ASICs), all of which have their varying strengths when it comes to accelerating workloads.

FPGAs, for example, offer good performance for [neural networks](#) and can be reprogrammed, but are much trickier to program than developing software to run on a GPU.

These accelerators are typically plugged into a standard server and work in conjunction with the existing processor cores. This is known as heterogeneous computing, and to work best, it requires careful integration of the different types of compute engines into one system to deliver optimal performance per watt.

A good example of such a heterogeneous system is [Nvidia's DGX line](#). These products combine Intel Xeon processors with a number of Nvidia's Tesla V100 GPUs into a system aimed squarely at deep learning and other demanding AI and high-performance computing (HPC) workloads.

The first such system, the DGX-1, used Nvidia's own NVLink interconnect to join eight Tesla GPUs, but its successor required Nvidia to develop an NVLink switch in order to provide enough bandwidth for inter-GPU traffic between 16 GPUs without taking a hit on performance.

In this e-guide

- ▀ [Lean, mean and green: The datacentres of the future](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

Although the DGX line is intended as a dedicated system for deep learning and other demanding workloads, it does highlight some of the issues with deploying accelerators. The DGX-2 with its 16 Tesla GPUs cost \$399,000 at launch, making it the kind of hardware purchase you only make when you absolutely need that level of performance.

Enter composable infrastructure

One answer to this conundrum may come in the shape of composable infrastructure, a system architecture that disaggregates some of the hardware of a traditional server into pools of resources. The thinking behind this concept is that the appropriate resources can be pulled together under software control to deliver a system that precisely matches the requirements of the workload it is intended to operate.

With existing systems, the number of processors and other resources such as memory and storage is largely fixed, or not easy to change as and when required. This typically leads to suboptimal utilisation, with systems having more resources than required by the workload.

The ability to compose a system on demand from a shared pool of resources should mean that expensive hardware such as GPUs or FPGAs can be shared across a bank of systems and allocated as and when needed, instead of being

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

permanently installed in every system that may need to run a workload that needs an accelerator.

However, as analyst Gartner detailed in a recent report ([Understand the hype, hope and reality of composable infrastructure](#)), the word “composable” has been used to describe a wide range of equipment from various suppliers.

Composable infrastructure is also currently held back by the difficulty in disaggregating DRAM from processors, as well as a lack of cross-supplier application programming interfaces (APIs). The latter means that composable kit from one supplier is unlikely to work with that from a different supplier.

What does [composable infrastructure](#) look like? As mentioned above, there is currently no practical technology for disaggregating main memory from the processor without incurring a performance penalty, so early composable platforms have compromised on using fairly standard x86 servers as the compute resources, combined with a separate pool of configurable storage.

Examples of this are HPE’s Synergy platform and Dell EMC’s PowerEdge MX portfolio. Designed around proprietary enclosures that fit a mix of compute and storage sleds, these resemble blade servers, but with a switched SAS fabric connecting the compute and storage components. This enables the SAS drives to be linked to the compute sleds in various configurations.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

Finding the right fabric

A key part of a composable system is the way the component parts interconnect so they can be used together. With Synergy and PowerEdge MX, this is provided by a switched SAS fabric linking the compute and storage sleds. DriveScale, another composable infrastructure firm, uses an Ethernet fabric.

However, if you want to compose more than just storage, you need a connectivity fabric with greater flexibility. One option is PCI-Express (PCIe), which was created as a high-speed interface for connecting devices directly to the processor, but has now been adapted as a means of connecting external hardware as well.

One company that is pushing this approach is [Liqid](#), which bases its composable infrastructure platform around a PCIe switch. This enables hardware such as GPUs, FPGAs and storage to be fitted into separate enclosures in a datacentre rack and shared among several standard x86 servers. The PCIe switch serves as the control point to compose the required configurations.

Another company, GigaIO, has taken a similar approach, creating its own interconnect technology, FabreX, based on PCIe Gen 4. The bandwidth available through PCIe Gen 4 allows the firm to target HPC deployments with its platform.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

This flexibility of composable infrastructure means it may be an attractive option for organisations seeking the best platform to support new and emerging workloads that incorporate approaches such as big data analytics and machine learning.

“I do believe that new workloads align way better with composable infrastructure, as it allows you to right-size hardware configurations as workload requirements change and evolve,” says Julia Palmer, research vice-president at Gartner.

“In addition, use of high-cost components is driving the requirements for better utilisation of, and continuous optimisation of, bare metal hardware resources, which is one of the unique benefits of composable infrastructure.”

Pros and cons

However, organisations interested in composable infrastructure should consider it carefully. According to Palmer, composable infrastructure runs the risk of creating yet another silo in the datacentre that has to be managed separately from existing systems.

This is because some composable infrastructure products do not come with built-in automation and will require additional integration with third-party

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

automation tools to enable infrastructure-as-code and intelligent infrastructure capabilities.

“Composable infrastructure is enabling technology and can only be fully leveraged within organisations that are ready to embrace fully automated server build services, which currently are suffering from lack of standards as every vendor has its own APIs,” says Palmer. “There is a hope for the Redfish ecosystem, however it is too early to tell if all of the vendors are going to fully embrace it.”

[Redfish](#) is the name for specifications drawn up by the Distributed Management Task Force for a standard API for management of software-defined IT infrastructure, including composable systems. It is supported by the Ironic bare metal provisioning service in OpenStack and the Ansible automation tool, as well as the baseboard management controller inside some servers.

Composable infrastructure can therefore be regarded as a technology still very much in its infancy. This can be seen from the fact that current platforms are unable to disaggregate CPU and memory resources and compose them independently.

There are moves to rectify this, such as via the [Gen-Z memory fabric](#), which is designed to allow memory semantic operations over connections ranging from a simple point-to-point link to a rack-scale, switch-based topology.

In this e-guide

- ▣ [Lean, mean and green: The datacentres of the future](#)
 - ▣ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - ▣ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

In summary, the need for costly accelerator hardware such as GPUs and FPGAs could drive demand for composable infrastructure. However, as Gartner advises, companies must carefully weigh up the cost and benefits of composable infrastructure compared with traditional infrastructure alternatives.

In this e-guide

- Lean, mean and green: The datacentres of the future
 - How AI and data analytics are driving composable infrastructure adoption in datacentres
 - How AI and data analytics are driving composable infrastructure adoption in datacentres
-

▀ Datacentre liquid cooling: What needs to happen for it to become commonplace in colocation?

Caroline Donnelly, Senior Editor, UK

The deployment of [liquid cooling](#) in datacentres is far from a novel concept, but the use cases for the technology remain niche, and largely limited to exascale and high-performance computing (HPC) environments.

Fresh air cooling remains the preferred (and most prevalent) means of keeping the temperatures of datacentres across the world in check, but that may not always be the case, says Paul Finch, CEO and chief operating officer of Essex-based colocation facility Kao Data.

The company has set itself a goal of becoming the leading provider of datacentre capacity to the UK's life science community. It also claims to be the first wholesale colocation facility in the UK to make use of indirect evaporative cooling without mechanical refrigeration, whereby the hot air generated by the servers is released into the atmosphere, rather than outside air being pumped in to keep the equipment cool.

In this e-guide

- ▣ [Lean, mean and green: The datacentres of the future](#)
 - ▣ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - ▣ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

Even so, its flagship, multi-tenanted facility in Harlow is designed to accommodate a future whereby liquid cooling could become more commonplace within the [colocation space](#). The way Finch sees it, developing hybrid facilities – containing a mix of air and liquid cooling systems – is a good way of laying the groundwork for that.

“What is certain is that liquid-cooled processor technology has reached a stage where if you are in the datacentre colocation market, and not planning for your facility to implement liquid-cooled compute capability, then you are way behind the curve,” he says.

The physical space requirements and layout of a facility that is liquid cooled requires a different type of datacentre design than an air-cooled facility.

“Datacentre buildings evolve, and if liquid cooling is the next step, then there are considerations that must have been taken prior to the construction stage, such as structural floor loading, ceiling heights and logistic routes, including lift design,” adds Finch.

Colocation providers that fail to take all of the above into account when planning out their facilities will struggle to accommodate liquid cooling-requiring workloads, which – in turn – could put them at a competitive disadvantage later down the line, he says.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

“Without this flexible building design, this radical infrastructure change will mean [legacy datacentres](#) at best will struggle to remain competitive, and at worst discount themselves at the customer proposal stage,” says Finch.

This is particularly as various parties in the datacentre industry are of the view that the number of compute-heavy workloads that run in traditional colocation facilities is set to increase, which will require more efficient and targeted cooling systems than traditional air-based systems provide.

As an example, Finch cites the growing demand for denser and high-performing compute capacity in datacentres to support deep learning and [artificial intelligence](#) (AI) workloads, fuelled by advancements in processor technology.

“The technology to process the massive datasets required for real-time or near real-time applications means these chips are processing billions of instruction operations per second [IOPS], which generates heat that must be removed in the most efficient process possible,” he says.

“Liquids offer a far higher heat capacity than air, [given] water is 4,000 times more efficient than air. With rack density estimates north of 50kW, this technology offers opportunities to achieve increased [energy efficiency](#) across the datacentre compared with air-cooled sites.”

This is perhaps why the operators of hyperscale datacentres are said to be among the user groups that are most actively exploring the use of liquid cooling

In this e-guide

- ▀ [Lean, mean and green: The datacentres of the future](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - ▀ [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

in their facilities, as it could enable them to cram much more kit into a smaller amount of space.

This is also why the roll-out of [5G connectivity](#), and its reliance on edge-compute environments, is often cited as another potential use case, says David Craig, CEO of liquid cooling systems manufacturer Iceotope Technologies.

“5G can only be enabled by adopting edge compute, and with the anticipated volume of infrastructure expected to be high, we have to ensure those datacentres are designed and cooled efficiently from the start, or risk negative impacts on the environment,” he says.

Relying on air-based cooling systems in edge deployments within built-up, urban areas can lead to issues whereby poor air quality compromises efforts to keep facilities cool, says Craig.

“As we move forward, air cooling will be inadequate for high-density multicore-processor systems. There are a number of environmental concerns driving adoption already, including poor air quality and the overuse of water in urban environments, but overall, I expect hyperscale datacentres to be the next adoption space for liquid cooling,” he says.

“Next will come the industrial edge computing and urbanised edge applications, followed eventually – once the technology is mature, proven and easily accessible – by telcos, enterprise IT and colocation customers.”

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

Plotting the shift from air to liquid

Not everyone, however, is convinced that a wholesale shift in preference from air cooling to liquid cooling in datacentres is necessarily on the horizon, with independent datacentre consultant Mark Acton saying such a change would be decades in the making.

“My feeling is liquid cooling is a solution looking for a problem at the moment, but if we’re talking longer term, things may change. People ask me, ‘What does the datacentre of 2025 or 2030 look like?’, and my typical response is, ‘Exactly the same as it does now’.

“My feeling is liquid cooling is a solution looking for a problem at the moment, but if we’re talking longer term, things may change”

Mark Acton, independent datacentre consultant

“The datacentres we’re building today have a 20- to 30-year lifespan. So [those who think we’re going to have] completely different datacentres in 2030 are probably not thinking everything through. Yes, we might be building something slightly different by then, but we’ll certainly be using the existing stuff.”

That includes air-based cooling systems, until the density of compute within each rack increases to such an extent that [liquid cooling](#) starts to make better economic sense, he adds.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

“As we get more dense with compute, which is a continuing trend, there will come a point where we cross a threshold and liquid will become more economic and more viable, but to reach that we need to have far more dense compute environments,” he says.

“You know the number crunching-type systems where they’re operating 40-50kW a cabinet. When you’re genuinely getting to that sort of scale, it starts to make more sense.

“If you’ve just got one or two racks, you could still cool that with air, but if you’ve got a deployment of scale with that sort of density, liquid becomes a much more viable option.”

The environmental viability of liquid cooling

But what are the environmental implications if liquid cooling does become more prevalent in datacentres worldwide? It could be a positive move, claim some.

Mark Seymour, chief technology officer (CTO) at London-based datacentre engineering simulation software provider Future Facilities, says one of the reasons why liquid cooling is so popular in HPC environments is that it is far more energy efficient than traditional air-cooling setups.

“This is primarily because the high-power densities are challenging, if not impossible, to cool with air-cooled systems,” he tells Computer Weekly. “The

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

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.”

At the same time, water is much better than air for transporting waste heat away from datacentres, which opens up new opportunities for this by-product of their operations to be re-used, he says.

The warm air that is ejected from datacentres through cooling systems is, [particularly in the UK, usually released into the air because of the difficulties operators face when trying to re-use it.](#)

While many Nordic countries have district heating systems that can make use of lightly warmed, piped-in air from datacentres to heat homes and offices, the UK does not.

This warmed air is usually generated in huge amounts, which makes it costly and cumbersome to transport long distances, so datacentres effectively need a suitable receptacle for it on their doorsteps to put it to use.

Heated water, however, is relatively easier to transport, meaning liquid cooling opens up the potential for waste heat from datacentres to be re-used, while potentially cutting the amount of energy needed to heat nearby homes and offices accordingly, says Seymour.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

Making the economics of liquid cooling work

However, it is “incredibly tough” at the moment for datacentre operators that are not running compute-heavy workloads in their facilities to justify the costs and complexities involved in deploying liquid cooling systems.

This is particularly in the face of the work that the hyperscale cloud and internet giants, such as [Facebook and Google](#), are doing to [bolster the energy efficiency of their air-cooled systems](#), resulting in power usage effectiveness (PUE) scores of sub 1.1, adds Seymour.

“Taken at face value, that means the potential saving is less than 10% for a perfect system. This makes it incredibly tough for users without high-power densities to justify the perceived additional complexity and capital cost associated with liquid cooling,” he says.

“With the better transport properties of liquid and the ability to keep chips cooler even when power densities are higher, the datacentre industry has the opportunity to make a significant contribution to saving the planet by reduced impact on the environment – we just need to hurdle the barriers to liquid cooling becoming mainstream.”

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

In HPC environments, however, the environmental benefits of liquid cooling are easier to quantify, says Chris Roberts, [head of datacentre and cloud at the Goonhilly Earth Station in Cornwall](#).

The site is home to a datacentre powered by renewable energy. It is marketed as a collaboration and data processing hub for enterprise and academic organisations focused on creating machine learning and AI services.

The facility predominately makes use of free air cooling, but also has a relatively small-scale Submer liquid immersion cooling deployment on-site to support some of its clients' HPC workloads.

This setup involves submerging the datacentre servers in a thermally, but not electrically, conductive dielectric coolant, and any heat generated is removed through the deployment of cool heat exchangers.

"It reduces the cost of cooling, and because we've only been ready to run it a few months now, the data suggests it is making our cooling 45% more efficient," he says.

"It's reducing the amount of power we consume by increasing the efficiency of the cooling, which is obviously still a significant factor in the electricity costs."

This is important, as the energy consumption habits of the datacentre sector are increasingly a front-of-mind issue for environmental lobbyists, as well as IT

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

sustainability researchers and campaign groups, including the United Nations-backed International Telecommunications Union (ITU).

The datacentre sector's energy usage habits come under scrutiny in the ITU's [Frontier technologies to protect the environment and tackle climate change](#) report, which also features input from several other UN-related groups as well.

“Datacentres are consuming vast amounts of electricity and are leaving significant environmental footprints. Unless these datacentres begin to move towards utilising renewable energy sources in their operations, they will remain responsible for a significant portion of global emissions”

ITU report

“Datacentres are consuming vast amounts of electricity and are leaving significant environmental footprints,” the report states. “Unless these datacentres begin to move towards utilising renewable energy sources in their operations, they will remain responsible for a significant portion of global emissions.”

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

And if the deployment of liquid cooling technologies brings about energy efficiency improvements in datacentres, if its use becomes more prevalent, that can only be a good thing.

The ITU is currently in the throes of weighing up the [environmental pros and cons of liquid cooling](#), as part of its wider commitment to ensuring there are guidelines in place to ensure sustainability is a key consideration in the adoption and development of any new and emerging technologies.

In the case of liquid cooling, the organisation is working on the creation of standards to ensure that, if adoption of the technology were to take off, the IT industry would put it to use in an environmentally sustainable way, ITU counsellor Cristina Bueti tells Computer Weekly.

“The problem we face is that with traditional datacentres, how the datacentre is built, the way the datacentre is installed, and most importantly, how it is maintained is very clear, but the same cannot be says for [datacentres builds featuring] liquid cooling,” she says.

“What we’re trying to do is to put in place a set of specifications, which proves that that there is a possibility for liquid cooling to be considered as an alternative to the traditional [air-cooled] datacentre.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

“Water re-use could be an additional value that we can put on the table for customers that consider the environmental impacts to be a key requirement [of any datacentre build].”

These standards and specifications are still some way off, and it is difficult to say with any degree of accuracy now what they will contain, but Bueti is hopeful that a first draft will be available by the end of the year.

Even so, the general consensus across the datacentre industry is that widespread adoption of liquid cooling is still some way off, but – in the meantime – Kao Data’s Finch is of the view that many facilities will probably be operating facilities that are cooled with a mix of air and liquid for some time to come.

This is particularly as there new air-cooled server products coming to market now with 40-to-50kW load capabilities, which mean the technology is unlikely to be swapped out in favour of liquid-based alternatives in a hurry.

“For most enterprise customers of datacentres, there is an opportunity to investigate liquid-cooling processing. But when considering the broader information technology equipment architectures – comprising of compute, memory, storage and networking – we are likely to see the deployment of far more complex mechanical engineering infrastructures adopting a hybrid strategy capable of supporting both air-cooled and water-cooled information technology equipment,” he says.

In this e-guide

- [Lean, mean and green: The datacentres of the future](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
 - [How AI and data analytics are driving composable infrastructure adoption in datacentres](#)
-

■ Getting more CW+ exclusive content

As a CW+ member, you have access to TechTarget's entire portfolio of 140+ websites. CW+ access directs you to previously unavailable "platinum members-only resources" that are guaranteed to save you the time and effort of having to track such premium content down on your own, ultimately helping you to solve your toughest IT challenges more effectively—and faster—than ever before.

Take full advantage of your membership by visiting www.computerweekly.com/eproducts

Images; stock.adobe.com

© 2019 TechTarget. No part of this publication may be transmitted or reproduced in any form or by any means without written permission from the publisher.