DCD

MORE GHz

FIGHT NIGHT

August/September 2017
datacenterdynamics.com

EPYC 32 CORES

BATTLE OF THE

CHIPS

"THE UNDERDOG"
AMD

VS

"THE CHAMPION"
INTEL

VR TAKES A RINGSIDE SEAT

Simulations and simulacra in the world of data centers

RACE AGAINST THE CLOCK

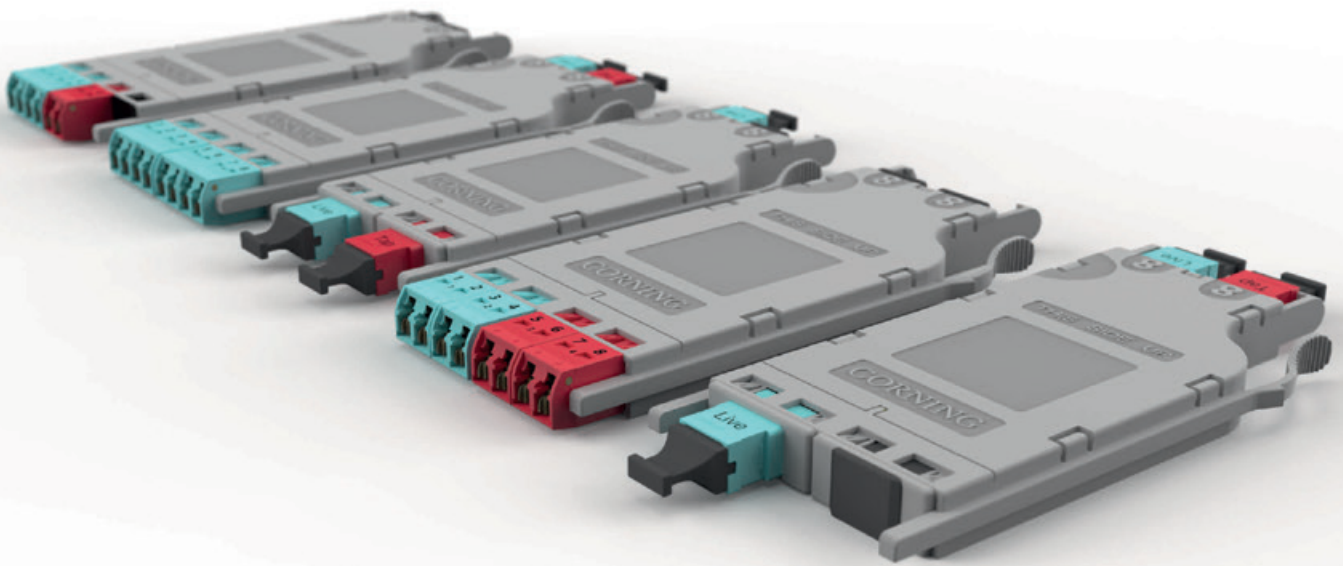Every nanosecond counts as satellites show their age

THE GRID VS DATA CENTERS

Don't be power hungry, become energy smart

★ ★ ★ ★

★ ★ ★ ★

CORNING

Connect to smarter, higher-density rack space utilization
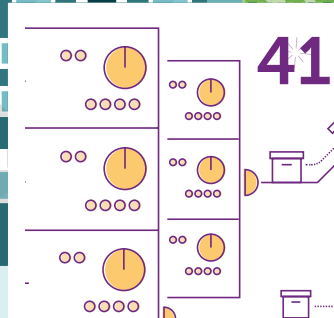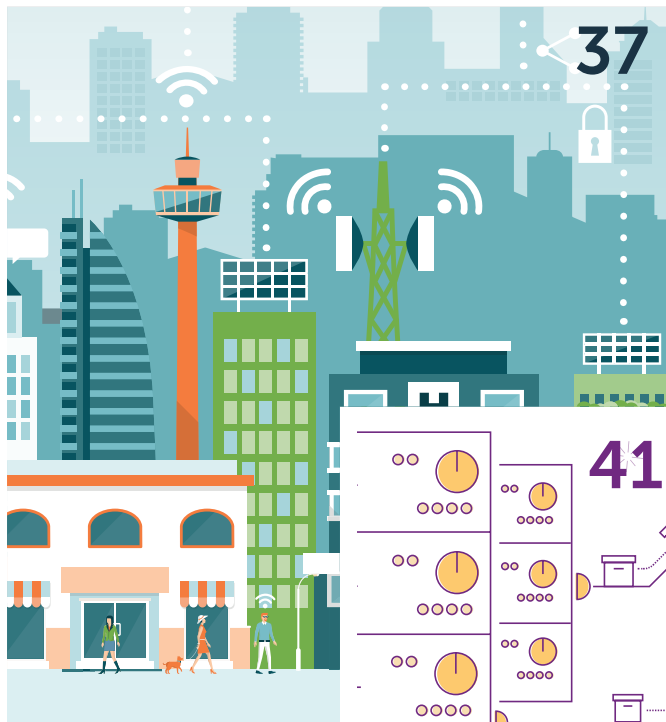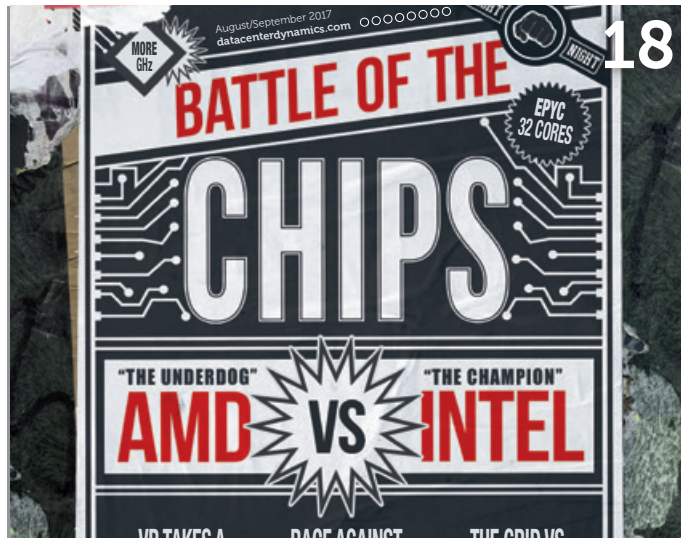
EDGE8™

Corning's award-winning EDGE8™ solution is the industry's first modular, tip-to-tip optical cabling system to feature a Base-8 design to maximize per-rack-unit density for better network scalability and improved link performance. And now our EDGE8 solution includes port tap modules: completely passive, purely optical splitters — no power or IP configuration required.

Are You Corning Connected?

Contents
August/September 2017

30

34

18

46

37

41

19

Dr Lisa Su
CEO of AMD

27

EDITOR'S PICK

A wrinkle in time
The Internet has well-developed
techniques for synchronization.
But they may not be good
enough for emerging
applications and regulations

From the Editor
White Space Fight Club

Fight metaphors aren't normally my cup of tea, but you can't avoid it when you look at the conflict in server chips, between the grizzled reigning champion Intel, and AMD, a fan-favorite contender, making a late bid for glory.

With tempers rising, I assigned my colleagues Max Smolaks and Sebastian Moss to the ringside, each covering one corner.

Max got inside information on Intel, while Seb got privileged access to AMD, including an interview with the CEO, Dr Lisa Su (p18). Between them, they also cover ARM and other players.

My only worry is - after getting roped into this bruising contest, will Max and Seb still be friends?

> *Curbing demand is a fight between humans and an inanimate global system we created*

Cooling tech is more standard fare for *DCD* readers, so Tanwen took a virtual reality look at its history (p30).

Through a combination of VR headsets and computational fluid dynamics modeling, Future Facilities gave her a magical historical tour of data center efficiency – all the stages of data center construction reproduced in virtual space.

As so many people have told us, whatever power and racks you may have, it's all about the air flow.

More adventures in time for Seb: distributed cloud services can only work when the services have an agreed and shared reference system time.

How is that time delivered? It turns out there's a whole sub-industry of time servers, using esoteric techniques like cesium fountain clocks and rubidium crystal oscillators.

That may sound abstruse, but new financial rules mean you will need to know about this (p27).

Capacity demands are growing but our Energy Smart summit in San Francisco says this is not necessarily an environmental problem (p34).

Data centers use energy steadily, which is good news for the utility grid which likes to deliver power that way.

The bad news is that both utilities and data centers have set up a system where services delivered from data centers appear to be free. Demand grows without limit and the infrastructure has so far been scarily good at meeting that demand.

It turns out that the best way to exit this cycle may be to intervene in human behavior, and create feedback which helps people to realize and limit their own environmental demands.

That sounds a lot like blaming the users. But put it another way. If curbing demand were a fight, it would be a fight between humans and an inanimate global system they created.

In dystopian sci-fi, the global system might win. In the real world, I'd say the human race could still score a knockout.
bit.ly/DCDmagazine

98%

Share of the x86 server market owned by Intel. AMD's portion "rounds up to 1%," AMD CEO Lisa Su told DCD (p19)

Peter Judge
DCD Global Editor

Most Eye-catching Datacenter in Asia

Korea, Fastest growing IT hub, becoming a hot place for major global IT businesses

Pack'em up Tight!

Unlimited higher rack power density, for lower Total Cost of Ownership

The only Carrier Neutral datacenter
operated by Tier I Network service provider

Certified as 'Tier III' by
Uptime Institute

Asia's largest capacity of
white space and power(165MW)

Green datacenter with advanced
energy efficiency(air-side economization)

TIER III DESIGN
LG Uplus Corp
Pyeong-Chon Data Center Step I
4 April 2014
UPTIME INSTITUTE CERTIFIED

ISO 27001

FROST & SULLIVAN

Uptime
Institute

Information
Security Management

Best IDC Service
Provider

Hundreds of power whips.

Or one simple solution.

Increasing power density within data centres is an ongoing challenge for facilities managers. And the continual cycle of increasing power requirements has too often translated into more cables and whips under the floor. But Starline Track Busway is self-contained, customisable and flexible. So you can avoid a jungle of wires, and enjoy power expansion in minutes, not weeks. To learn more about our maintenance-free, reliable systems, visit StarlinePower.com.

STARLINE®
TRACK BUSWAY

Vapor IO puts edge colo in cell towers

Cylindrical rack enclosure company Vapor IO announced at DCD Webscale that it will put its gear into the base of cell towers to deliver colocation services at the edge of the network, backed by wireless infrastructure provider Crown Castle.

Sri Lanka gets first Tier III certified data center

Sri Lankan telecoms company Dialog Broadband Networks has been awarded a Tier III Constructed Facility certificate for the first phase of its Malabe Data Center #2.

Asgardia to put SSD into low earth orbit

Self-proclaimed space nation Asgardia is to launch a small satellite into space with 512GB of memory in an effort to test long term data storage in space. It appears that the total storage available to end users is only just over 200MB.

Schneider extends Galaxy VX UPSs

Schneider Electric has launched new scalable 500kW, 1000kW and 1500kW UPSs in its Galaxy V-Series range. Launched at DCD Webscale in San Francisco, these models can be used to shave demand during peak periods.

New regulations force Apple to build first Chinese data center

Apple has announced plans to open its first data center in China, in the southern province of Guizhou, in partnership with data management firm Guizhou-Cloud Big Data Industry Co Ltd (GCBD).

The new facility comes as China is strengthening its controversial cybersecurity laws, mandating that foreign firms have to store user data within the country. China also said that businesses transferring over 1,000 gigabytes of data outside of the country will have to undergo yearly security reviews, in language that has been criticized as overly vague and a threat to proprietary data.

"The addition of this data center will allow us to improve the speed and reliability of our products and services while also complying with newly passed regulations," Apple told Reuters. "These regulations require cloud services be operated by Chinese companies so we're partnering with GCBD to offer iCloud."

GCBD was co-founded by the provincial government, which wants Guizhou to become a data hub. Other projects in the province include a US$23 million facility to process and store data from the world's largest radio telescope.

Apple pointed out the data would be secure. It said: "Apple has strong data privacy and security protections in place and no backdoors will be created into any of our systems."

The company added that this facility, like Apple's other sites, will use renewable energy, and that its investments in Guizhou would pass $1 billion.

Elsewhere, Apple announced plans to build a second data center in Denmark for 6 billion Danish crowns ($921 million).

In Ireland, meanwhile, the company must wait until October 12 for a judgment on its 2015 application to build a data center in Athenry.

🔗 bit.ly/littletroubleinbigchina

Vox Box

Can the pace of efficiency match the increase in demand?

For the last five years, efficiency has kept up, because of Moore's Law, virtualization and the increasing efficiency of the infrastructure measured by PUE. There are a lot of people questioning whether Moore's Law can continue, you can't get better than a PUE of 1, and you can't get more than 100 percent utilization. There could be some issues beyond 2020.

Dale Sartor
Scientist/engineer
LBNL

🔗 bit.ly/energysmartsartor

What is different about energy policy in the US?

Energy policy in the US is unique. The states are sovereign entities with certain powers under the constitution - and those states have been the primary drivers of energy policy for centuries. As well as this, virtually all the energy assets are privately owned. In most countries, natural resources are owned by the government, not individuals.

Donald Paul
Director
USC Energy Institute

🔗 bit.ly/energysmartpaul

Digital Realty and DuPont Fabros announce merger

Digital Realty – one of the largest data center operators in the world – is set to merge with another American colocation provider, DuPont Fabros (DFT), in a deal valued at approximately $7.6 billion.

This will create a global giant with 157 facilities under management, and considerably increase Digital Realty's capacity across core US metro markets, where DFT usually locates its facilities.

The all-stock transaction will see DuPont Fabros shareholders receive a fixed exchange ratio of 0.545 Digital Realty shares per DuPont Fabros share.

The transaction has been unanimously approved by the boards of directors, and is awaiting approval by the shareholders of both businesses. It is expected to close in the second half of 2017.

🔗 bit.ly/megamerger

Vertiv to sell ASCO to Schneider Electric for $1.25bn

Vertiv plans to sell ASCO, its automatic transfer switch business, to Schneider Electric for $1.25 billion

Formerly known as Emerson Network Power, Vertiv was last year rebranded and acquired by Platinum Equity for $4 billion. The company said that the sale, which is expected to close by Q4 2017, will allow Vertiv to focus on its core data center, telecoms, and commercial and industrial markets.

"This sale is a significant step forward in our evolution as the premier provider of digital critical infrastructure solutions," Vertiv CEO Rob Johnson said.

ASCO's Automatic Transfer Switches (ATSs) continually monitor the available current from two different sources and switch between them when one drops below a desirable level. In a data center, an ATS would switch to a generator in the event of a power grid failure.

"The market for ATS presents attractive long-term growth opportunities as more commercial and industrial customers move towards autonomous or multi-source power management," Schneider said in a statement.

ASCO first introduced an ATS in 1925, after operating as an elevator, compressor and generator control company for nearly 40 years. It was acquired by Emerson back in 1985.

🔗 bit.ly/switchingswitches

IBM opens four cloud data centers

IBM has opened four new data centers: two in London, one in San Jose (California) and one in Sydney.

The new facilities bring the company's total to 59 data centers in 19 countries, including four in Australia, five in the UK and 23 in the US.

In its second quarter results, net income decreased by two percent since Q2 last year, the 21st consecutive quarter of revenue decline.

But IBM's cloud revenue grew 17 percent year-on-year with adjusted currency. In 2016, its cloud revenue grew by 35 percent.

IBM has stated that it expects revenues from its so-called 'strategic imperatives' (such as AI and cloud computing) to increase in the second half of 2017 as a result of having signed new cloud contracts, including a 10 year cloud outsourcing deal with Lloyds Bank and a joint venture with China's Dalian Wanda.

🔗 bit.ly/eyebeeemm

Google launches appliance to upload data by FedEx

Google has launched a storage module designed to let users ship data to its cloud by courier, a process referred to as "Sneakernet." The product, which echoes similar moves by Amazon, holds up to 480TB, and avoids network uploads which could take years.

Google Transfer Appliance (GTA) is a 19in rackmount storage appliance, which can be installed in a customer's rack, filled with data, and then shipped to Google for upload to the Google Cloud Platform (GCP).

Two sizes are available: a 4U-high 480TB unit or a 2U-high 100TB module. The product, currently only available in the US, follows Amazon Web Service's similar 50TB/80TB Snowball device launched in 2015, and the truck-sized 100-petabyte Snowmobile launched in 2016.

🔗 bit.ly/googlesneakernet

Peter's Sneakernet factoid
An Airbus full of microSD cards could achieve 5 petabytes per second between New York and Los Angeles
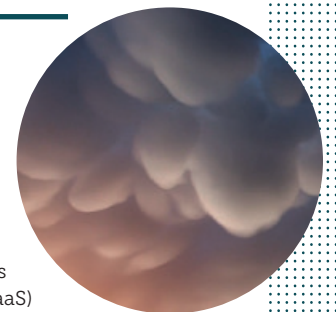
Microsoft to cut salesforce by 10 percent in major reshuffle

Microsoft is planning a significant round of job cuts, as it restructures the company around cloud computing.

The change will mostly impact its sales operations, with less than 10 percent of the sales staff – still numbering in "thousands" – set to lose their jobs.

The company told *CNBC* that this is not a cost-cutting exercise, but rather an attempt to refine its sales process: Microsoft will make changes to its sales departments, dividing their staff by industry verticals and target company sizes.

A spokesperson said: "We are taking steps to notify some employees that their jobs are under consideration or that their positions will be eliminated.

"Like all companies, we evaluate our business on a regular basis. This can result in increased investment in some places and, from time-to-time, redeployment in others."

bit.ly/cloudyfuture

Cloud and SaaS revenues to grow by a quarter every year for five years

Research collated by Synergy suggests that revenues stemming from cloud and software-as-a-service (SaaS) will grow between 23 and 29 percent every year, through to 2022.

The study predicts that the Asia-Pacific region will experience the highest growth in cloud and SaaS revenues, followed by Europe and the Middle East and North America.

Specifically, Synergy suggests that public IaaS and PaaS will grow by an average of 29 percent every year, managed and hosted private cloud by 26 percent, enterprise SaaS by 23 percent and that infrastructure sales to major cloud providers will increase by 11 percent annually for five consecutive years.

As far as Internet-based virtualized computing and managed platform growth are concerned, database and IoT-specific services are expected to feature most prominently. Enterprise resource planning (ERP) software is predicted to drive SaaS growth.

At the same time, Synergy expects that hardware and software sales to enterprises will dwindle, decreasing by two percent on average every year, reflecting the shift from privately-owned infrastructure to running workloads in the cloud.

bit.ly/DCDgrowbabygrow

Naver to build $420m data center in Yongin, South Korea

South Korean cloud giant Naver is planning to spend 480 billion won ($420 million) on a massive data center in Yongin, just south of Seoul.

According to SK publication *Platum*, the facility will start offering public cloud services in the second half of 2020.

Naver, often referred to as 'the Google of South Korea,' was established in 1999 as the first Korean web portal to develop its own search engine. The company has been successful in keeping its competitors at bay by fusing latest technologies with decisively Korean aesthetics. It also enjoys informal support from the Korean government.

Naver was the first Internet company in Korea to build and operate its own data center, the 'Gak' facility in Chuncheon, Gangwon Province. It was designed to resemble traditional rice terrace farms and is listed in the '*Top 10 beautiful data centers*' feature from our April/May issue.

Now, the company is expanding with a new facility built to power the public cloud platform it launched in April. Naver Cloud Platform already offers 30 basic infrastructure services related to computing, data, security and network, and promises to be price competitive with both AWS and Microsoft Azure.

🔗 bit.ly/anaverdatacenter
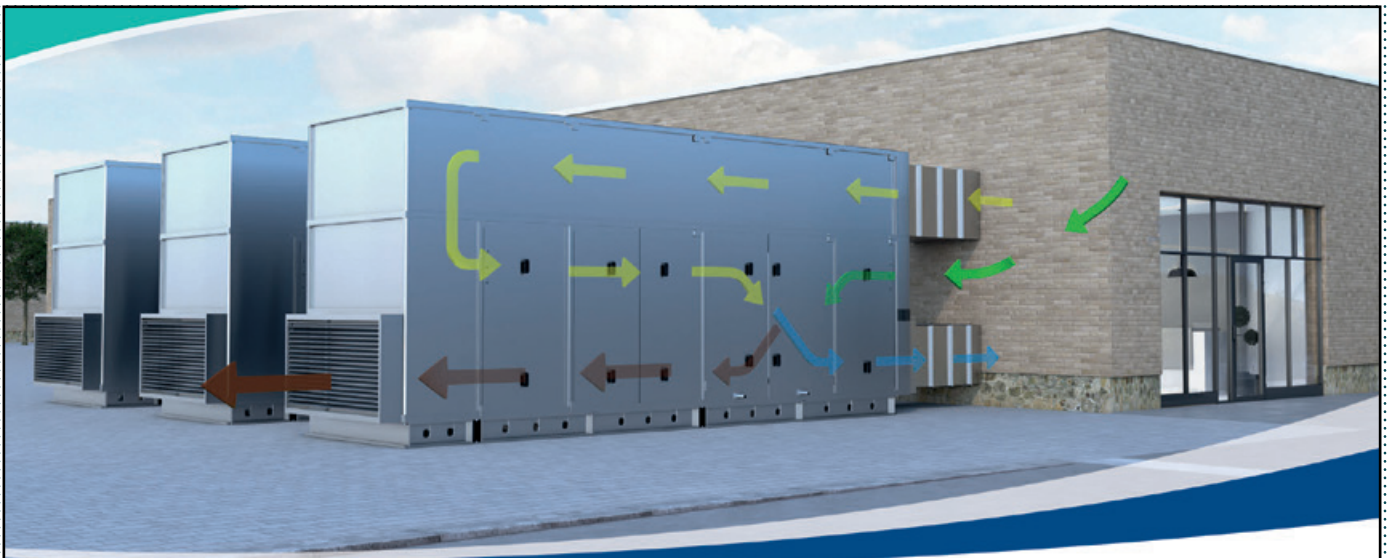
US Air Force and IBM to build 'brain-inspired' supercomputer

The US Air Force Research Laboratory (AFRL) will collaborate with IBM on a 'brain-inspired' HPC system powered by a 64-chip array of the IBM TrueNorth Neurosynaptic System - a low energy chip designed to mimic brain activity.

The company claims that the system's pattern recognition and sensory processing power will be the equivalent of 64 million neurons and 16 billion synapses, while the processor will require just 10 watts of power.

The system fits in a 4U-high (7in) space in standard server racks, with eight of the systems enabling the 512 million neurons per rack.

One processor consists of 5.4 billion transistors organized into 4,096 neural cores forming an array of one million digital neurons that communicate with each other via 256 million electrical synapses. The project was originally funded by DARPA under its SyNAPSE program.

bit.ly/planebrain

Atos set to deliver the first commercial ARM-based supercomputer

French IT services giant Atos has introduced the first ever commercial supercomputer to rely on processor cores designed by ARM.

The Bull Sequana X1310 will use the familiar Sequana X1000 architecture and Cavium's ThunderX2 chips to offer a system suitable for the most compute-intensive tasks.

"Atos designed the Bull Sequana X1000 system as an open platform, to offer HPC users a large choice of computer architectures from which to choose, and to support the future processor technologies that will make it possible to reach the exaflops level," explained Agnès Boudot, group VP and head of HPC at Atos.

"It is therefore a natural step to extend the range to include ARM processors, together with the existing CPUs, core processors and co-processors."

Processor cores designed by ARM are based on a reduced instruction set computing (RISC) architecture. This approach delivers less computational power per core, but can support more cores than a typical chip from Intel or AMD.

Such processors are cheaper, more energy efficient and require less cooling – however they have been struggling to find their niche in the data center (for more on ARM see page 23).
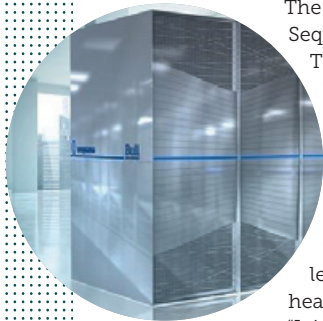
bit.ly/armingup

Estonia to create 'Data Embassy' in Luxembourg

Estonia will be backing up its government data in Luxembourg's first foreign 'data embassy,' storing copies of its most important records to protect them in case of outages, equipment failures or cyber attacks.

Last year, the country revealed that it was considering the UK to host the virtual embassy, but uncertainty surrounding Brexit negotiations led it to turn to Luxembourg instead.

The launch was announced at an annual celebration held by Digital Lëtzebuerg (Luxembourg) - an initiative set out in 2014 to promote the country's ICT sector and develop its international presence beyond the financial world - which was attended by the heads of state of Estonia and Luxembourg, Xavier Bettel and Jüri Ratas.

Peter Sodermans, senior advisor at Digital Lëtzebuerg, said that potential Russian invasion was merely a contributing factor in Estonia's decision to use Luxembourg-based infrastructure to hold a copy of its official records abroad, despite suspicions that Russia may have launched multiple cyber attacks on neighboring Baltic states in recent years.

For one, he said, Luxembourg is "one of the safest countries in the world."

bit.ly/estonesthrowaway

NEC is building a supercomputer for Czech Hydrometeorological Institute

Czech Hydrometeorological Institute (CHMI) is deploying a new supercomputer, designed by Japanese IT giant NEC.

The LX-Series machine consists of more than 300 nodes connected through a high-speed Mellanox EDR InfiniBand fabric. It will be used to improve accuracy of weather forecasts, track air pollution and provide early warning in case of extreme weather events.

The new system is scheduled to be put into operation in early 2018.

"For years, we have been successfully collaborating with meteorological institutes, and we look forward to cultivating these partnerships further," said Andreas Gottlicher, senior sales manager at NEC Deutschland.

The new supercomputer at CHMI will use 3,500 of the latest Xeon E5-2600 v4 cores to observe weather and make more accurate predictions.

Storage will be provided by an NEC LXFS-z parallel file system appliance, based on Lustre, with more than 1PB of capacity and a bandwidth of more than 30GBps.

bit.ly/rainraingoawaysimulateanotherday

Kenya opens two data centers to ensure electoral integrity

Ahead of presidential elections in Kenya in August, the country's electoral agency has coordinated the opening of two data centers. This follows an audit of the Independent Electoral and Boundaries Commission (IEBC) by KPMG that found "major security gaps" in the country's register.

The report identified more than ten security flaws in IEBC's existing data center that could be used to flout the democratic process on election day.

The integrity of the process and the ability to demonstrate it to the citizens of Kenya is all the more important given the events of the past two elections: in both instances, local and foreign observers found irregularities which cast doubt on the legitimacy of the elected candidate.

The audit of the electoral commission's current data center discovered that two active administrator accounts with access to the voter's register had default passwords that had never been changed, and preventative measures against cyber attacks were found to be lacking.

Furthermore, the data center at IEBC's headquarters had an unreliable UPS, a faulty fire alarm system that was last serviced in 2005, and its air conditioning system had failed, raising temperatures by 25 degrees Celsius (45°F).

Until now, there had been no disaster recovery site to ensure that the register was backed up and available in case of a technical failure or an attack on the IT systems. In fact, the register was backed up on tapes stored at the IEBC's head offices, where its data center was located.

bit.ly/Iwontheelectoralcollege

$3.2 million
Amazon's US lobbying spend in Q2 2017
(bit.ly/alexalobbygov)

US DOE awards $258m in HPC contracts

The US Department of Energy has awarded six American tech companies with contracts for high performance computing (HPC) research.

The $258 million will be shared between AMD, Cray, HPE, IBM, Intel and Nvidia over three years as part of the new PathForward program, itself part of DOE's Exascale Computing Project (ECP).

The companies themselves will provide additional funding amounting to at least 40 percent of the total project cost, which the DOE says will bring the total investment to at least $430 million.

The announcement comes after continuing concern over DOE funding levels under the Trump administration. The bipartisan 2017 fiscal year budget agreement slightly increased DOE funding, but Trump's own budget proposed to reduce it by up to $900 million and completely eliminate the Advanced Research Projects Agency-Energy.

Prior to this announcement, when asked about the reports over DOE funding, Cray's VP of business operations in EMEA, Dominik Ulmer, told *DCD*: "It is always a concern to us what the Department of Energy does, it's a very important partner for us - not only because they are a large customer at various places, but also because we're also developing our systems oftentimes in a co-designed fashion together with them, where we work together on technologies and try to match software and hardware technologies in a better way.

"This is a concern to us, of course."

bit.ly/racetoexascale

Other Colombian providers

GlobeNet

The carrier subsidiary of Brazilian investment bank BTG, Globenet is building a 200 sq m (2,152 sq ft) data center in Barranquilla, designed to "Tier III levels," which should be ready in August. Near Globenet's submarine cable landing station, the site supports IaaS services. The firm has a system of fiber-optic cables spanning more than 23,500 kilometers in America.

Internexa

Owned by Colombia ISA holding, Internexa has opened its newest data center in the Bogota Free Trade Zone, built by the developer of Free Zones of ZFB Group. The facility has a Tier III certification for design and construction. Internexa has other data centers in the region, including one in Medellin, along with a fiber-optic network of over 49,000km.

Helping Colombia in a time of crisis

Despite the difficult economic times, data center service providers are betting on the country, reports *Virginia Toledo*

Virginia Toledo
Editor LATAM

Monthly data is still mixed, but analysts believe 2017 has brought a significant improvement for Colombia compared to 2016. Felipe Gómez, Level 3's head of data centers in Colombia, agrees. He told *DCD* that projects which stalled in the past year could now be reactivated.

"2016 was very hard for the IT industry because of the devaluation of the peso against the dollar," he said. Investments became difficult because suppliers bought technology in dollars, but sold services in pesos.

This year there is a positive change for the Colombian economy, according to the LatinFocus forecast from *FocusEconomics*. With a strong oil sector (a key component of the economy), a low corporate tax rate, and an expansionary monetary policy, Colombia's GDP is expected to grow by 2.2 percent in 2017 and by 2.9 percent in 2018.

Gómez believes the peso's troubles are over and expects 2017 will be quiet: "The dollar will remain in the same range and organizations must accommodate the new reality and slightly increase their IT investment budgets."

Meanwhile, service providers have continued to invest in Colombia's maturing market. Cable & Wireless, for example, reached a phased agreement to buy 500 sq m (5,381 sq ft) of data center space in BT's Colombian data center over ten years. The deal is estimated at $20 million, and C&W also put

$4 million into infrastructure such as racks, servers and cleanroom equipment.

"The purchase decision was due to time to market," Marcelo Oliveira, director of data centers at C&W, told *DCD*. "With [BT's] Tier IV data center and solutions we have developed for IaaS, PaaS and DRaaS, we can provide services that the competition cannot offer."

BT's Naos data center, located in the Zona Franca de Tocancipá, opened two years ago. It was the first site in Colombia to obtain Tier IV design and construction certifications from the Uptime Institute, and received the DatacenterDynamics Latam Award for Best Service Provider Data Center in 2015.

"This data center is part of our strategic plan, not only for Colombia, but also for the region, being instrumental in the sustained growth of our company," Ricardo Imbacuan, C&W's vice president and country manager for Colombia, told *DCD*.

Another company that has recently entered the Colombian market with its own data center in Bogota is colocation giant Equinix, as part of its $3.6 billion purchase of 24 Verizon facilities.

Víctor Arnaud, Equinix's director for Latin America, said that the company sees Colombia as a strategic country, not only in of itself, but also as a bridge from South America to Central America and the United States, via Equinix's facilities in Miami.

These providers' decisions to enter the country are guided by a belief in digital transformation: they expect the slowdown in the Colombian market is temporary, and the demand for information storage and processing will continue to increase rapidly.

Level 3 has three data centers in the country, two in Bogotá and one in Cali, but it expects demand to grow there. The Cali facility was built as a backup site for customers, Felipe Gómez told us. Now, there are many customers in eastern Colombia who are migrating their data centers and outsourcing, creating new opportunities.

"At first, the idea was to build a contingency site, thinking that rules from the Financial Superintendence of Colombia would require backup data centers in Colombia to be more than 100km away, and at that time we had the two DCs in Bogota," he said. Level 3 wanted to be prepared for the regulation, but today is finding other uses for the data center, since: "we do not know whether the law will be a reality or not."

In Bogotá, Level 3 has a data center which is nearly 80 percent occupied: of four halls, three are almost full, with very little space available in the fourth room. "We believe that by 2018, we should start a new room of about 600 sq m (6,458 sq ft) in the Colombia XV data center in Bogotá," says Gómez.

Equinix already had a data center in Brazil, so Colombia was the next logical step. Thanks to its connectivity through submarine cables, Víctor Arnaud believes it has potential to become a regional hub.

Connectivity in Colombia has come a long way, says Juan Manuel González, director of data & Internet products for Colombia at Level 3: "Six years ago there were just four or five significant telecoms operators; now there are eleven or twelve companies, so the market is fairly complete and this is reflected in both coverage and cost, and even connectivity to the world through submarine cables." Level 3 implemented a Pacific cable two years ago, and there are six additional cables across the Atlantic.

A sign of the momentum of digital transformation is the volume of data and Level 3's network expansion plans. Level 3's Pacific cable was initially set up with 400Gbps but can reach a capacity of 4.5Tbps, and the company is planning to enable new equipment to expand beyond that. The company is involved in another project which seeks to take infrastructure directly to Quito, the capital of Ecuador.

For Level 3, Colombia, Chile and Argentina are the three major economies marking growth in the region. "If you look at the investments that operators are making, one of the most important in terms of deployment of connectivity is Colombia." ●

GDP is predicted to grow by 2.9% in 2018

Don't be held hostage

Asian experts shared their security tips for data centers at a Singapore summit. *Paul Mah* reports

The data center industry should not build security in as an afterthought, but should up its game and be alert to the latest threats to digital infrastructure, security and data center experts said at a DCD Security Summit held in Singapore as part of Interpol World in July.

"A data center can be held hostage if the digital controls to the UPS and chiller systems are compromised by hackers," said Ka Vin Wong, an independent consultant with experience helming colocation providers in Southeast Asia. With control over the mechanical and electrical systems, attackers can issue blackmail demands threatening an induced outage.

Mechanical and electrical systems can be isolated from the network, but Wong's point illustrates the need to harden modern data centers against digital threat vectors.

Everyone needs to play a part in security, and corporations can no longer be insular in their data management, said Chris Church, a senior mobile forensic specialist at Interpol (the International Police Organization). The

irony, said Church, is that executives often have irrational fears about the cloud, unaware of the extent to which they are using it.

Yet many damaging security attacks are not related to the cloud at all. One bank hack was traced to an outsourced IT support team from a neighboring country, said Church. One support staff member's laptop was infected with malware for an entire year before hackers and pulled a digital heist of $10 million within the span of an hour.

Although overstated, the threat from data leakage through the cloud is real, and hackers love the cloud as it requires just the user account and password to access the data, said Church. Almost all (90 percent) of consumers are not aware of what they are storing in the cloud; some apps save a lot more files than they expect.

Some users consider themselves unattractive targets, believing that attacks only happen to other people and large organizations, but this is simply not true, said Church. Data is a commodity like gold, and usernames, passwords and email accounts can change hands to the tune of thousands of dollars, he said.

Paul Mah
SEA Correspondent

To identify genuine threats in modern infrastructure, users may need to create a local threat intelligence base, because traditional defensive measures such as proxy servers, intrusion prevention systems (IPS) and antivirus software have repeatedly failed against targeted attacks, said Florian Lukavsky, the director of SEC Consult Singapore.

Security information and event management (SIEM) software products and services can generate a large volume of alerts, which are impossible to review manually. In the infamous Target hack of 2013, the hackers' activities were flagged, but ignored by the security team.

A honey pot – a fake environment with intentionally leaked false information can entrap attackers, said Lukavsky. The local threat intelligence gathered is unique to the data center environment, allowing security personnel to know with certainty that a security breach has occurred.

When one attacker installed ransomware and APT (advanced persistent threat) tools in a honey pot, the security team was able to collect the fingerprints of customized malware, usernames and passwords for backdoor tools, and attack patterns. These were checked against other parts of the environment to see if hackers had got in elsewhere.

Finally, a properly secured data center should incorporate capabilities including highly controllable personal access, and safeguards against actions that degrade security such as buddy punching and tailgating, says Phoon Wai Leong of ERS Group.

Physical access control protects against accidents as well

A laptop was infected with malware for a whole year, before hackers pulled a $10m digital heist

as malice, said Phoon. A systems engineer who runs out of power sockets in one rack, could reach to the next rack to find a vacant slot – and then promptly trigger a power trip as an already maxed-out circuit is pushed over the edge.

So how should modern data centers be secured? For a start, authentication systems should have the right data to identify logins from current or former employees, or blacklisted personnel.

And systems should have the ability to search, locate and track people in a near real-time manner to strengthen control for assets and personnel.

Phoon offered practical tips on securing physical security with the use of two-factor authentication (2FA). An encrypted proximity beacon could be issued upon validation and registration, which should be time synchronized to thwart duplication.

Active RFID or Bluetooth dongles can allow personnel to be tracked for incident response and auditing in the wake of security incidents.

Keypad and pin locks should be implemented at checkpoints within the facility, though facial recognition is increasingly seen as a reliable and cost-effective method of implementing a second factor control.

In fact, three-factor authentication should be considered too, suggested Phoon, and can be implemented with a mix of physical tokens, personalized passcodes and biometric authentication.

It will take some time to bring all data centers up to scratch on the security front. But as more systems within the data centers are digitized and networked, this is an area that can no longer be ignored. ●

The DCD Security Summit was held at Interpol World on 6 July in Singapore

Top tips:

- Isolate mechanical systems
- Create a local threat intelligence base
- Handle the cloud with care
- Use physical access controls
- Apply two-factor authentication, or more
- Trust **NO ONE**

BATTLE OF THE CHIPS

EPYC 32 CORES

The war for the data center will likely never end. As an old challenger returns, a new battle begins

While data centers are a complex combination of various pieces of computing equipment and industrial machinery all working in tandem like a great, living creature, at the heart of most lies the Central Processing Unit (CPU).

Servers form the center of a data center, and the chips within them the core of those servers – most of which are x86 servers, that is, servers which are backward-compatible with the instruction set which developed from Intel's original 8086 and 8088 CPUs, starting in 1978. With most of the 8086's successors ending in '86,' the name stuck, as did the architecture.

In Q3 2016, x86 servers brought in 90 percent of the server market revenue, according to IDC. Outside the x86 sector, vendors like IBM and ARM fight for scraps.

"x86 servers continue to be the predominant platform used for large-scale data center build-outs across the globe," Gartner's research vice president Jeffrey Hewitt said.

Within the x86 sector, the vast majority of processors come from Intel; AMD leads a tiny wedge of compatible manufacturers.

In 2003, AMD's Opteron line looked like it offered serious x86-compatible competition for Intel, but its position has waned.

"AMD got a little unfocused and, perhaps because of that lack of focus, had some execution issues," AMD's data center head Forrest Norrod told *DCD*. It lost its share of the market, handing Intel a near monopoly that it has only tightened its grip on.

But now things could once again change. AMD is spoiling for a comeback fight.

Under the stewardship of CEO Dr Lisa Su, the company has launched Zen – its first new microarchitecture in years.

Zen is an x86 microarchitecture, and the first processor family to implement it is Epyc. We talked to Su and profile the chip on the next two pages.

Intel, however, is keen to keep its dominant position. In an earnings call after its most recent quarterly results, CEO Brian Krzanich said: "AMD has raised up a bit with their more recent products, but you see us responding.

"This is a traditional performance battle that we're very accustomed to, and we're comfortable in reacting and competing very aggressively in. And so you see us coming out with our Xeon Scalable."

We look at the Xeon Scalable Platform and what it could mean on pages 28-29.

We may not know the outcome of this battle for some time, but one thing is for sure – the war for the heart of the data center is back on. ●

RETURN TO SPLENDOR

"We know how to win in the data center."

Dr Lisa Su, AMD's CEO, appears confident. We're backstage, sitting in a small room just after the official launch of the Epyc server chip, where journalists, analysts and potential customers were presented with a bevy of impressive stats, selling points and partnerships.

"I think we've shown there's a huge pent-up demand for different solutions. The current offerings in the CPU space, frankly, aren't very flexible for all of these different workloads that are there. It's more than just technology, it's giving the customer the choice."

Choice has indeed been lacking for some time in the x86 server market, currently dominated by Intel, which maintains a roughly 98 percent market share.

"At our peak we were at 25 percent, right now we're at... let's call it less than one percent," Su said. "It's all upside. I think a very reasonable goal for us, hopefully sooner than 2020, is to get back to double digit share. If we get back to 10 percent share plus – think about it, just the CPU market is a $16 billion market. That's a significant opportunity for us, and that's an intermediate goal, not the final goal."

That 25 percent peak, achieved over a decade ago with the Opteron processor line, feels like a distant memory. "We became distracted and we were doing many different things, server was not the first priority," Su told *DCD* in her first interview since Epyc's launch.

"As I look forward, I see the data center space as a driver of high performance technologies and a showcase of what we can do. AMD has the potential, it has always been a company of potential, and my job with this management team has been to unlock that potential with strong execution."

For Epyc, that journey started four years ago, when the company realized that pushing forward with the Bulldozer microarchitecture would not prove fruitful, and that it had to go back to the drawing board and design an entirely new microarchitecture – Zen. "That was a big decision. But the corollary with that is that it would take us years. We were okay with that."

The result is, by all accounts, a processor family that can contend with Intel's best. Depending on the workload, or the budget, each vendor offers a compelling argument – but the fact that AMD is even considered in such discussions is a major improvement.

"I feel really good about our end," Su said. "The differences that we're talking about in terms of price performance, I/O flexibility, memory capability – these are not small differences. This is not about a price game, about buying sockets, this is about enabling someone to do something that they can't do today, or something that used to be much, much more expensive."

Processors in the Epyc 7000 series range from the eight-core, 16-thread Epyc 7251 running at 2.1 to 2.9GHz in a 120W power envelope, to the 32-core, 64-thread Epyc 7601 running at 2.2 to 3.2GHz, with a 180W TDP.

But while AMD has certainly upped its game, Intel has not rested on its laurels. Since Epyc's launch in late June, the reigning champion launched the next generation of Xeon processors (for more on this, read Max Smolaks' deep dive on the following pages).

Perhaps sensing that Epyc represents a greater threat than any of AMD's recent efforts, Intel has launched an unprecedented assault on the rival processor family. Until recently, the company rarely referred to its main competitor by name; this time, it invited journalists to a special event, just before Epyc's launch, to explain that the product was inferior.

Key to its argument was the fact that Epyc processors consist of four Zen-based dies in a single package, which Intel slammed ▶
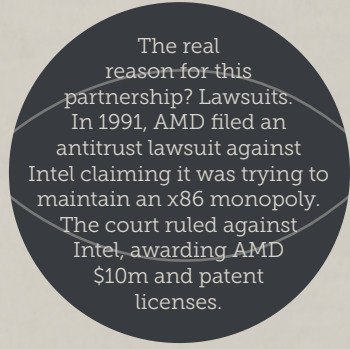
Strange bedfellows

AMD and Intel may be fighting tooth and nail over the x86 market, but they are actually strangely joined.

Since 1976, the two companies have shared numerous cross-licensing agreements to allow each partner to use the other's patented x86 innovations after a certain time period.

While the agreement has changed over the years – with the 1996 deal ensuring AMD could no longer make direct copies of Intel chips beyond the 486 processor, or use the same bus as Intel's chips – it has been fundamental in allowing both companies to work in the sector on compatible, infringement-free products.

The agreement is, however, canceled if AMD declares bankruptcy or is taken over.

The real reason for this partnership? Lawsuits. In 1991, AMD filed an antitrust lawsuit against Intel claiming it was trying to maintain an x86 monopoly. The court ruled against Intel, awarding AMD $10m and patent licenses.

Sebastian Moss
Reporter

▶ as a "repurposed desktop product for server," claiming it delivered "inconsistent performance from four glued-together desktop dies."

When asked about the decision to go with four individual dies, Su told *DCD*: "Honestly, it was one of the best decisions that we've made, I love it. The biggest issues that we have in advanced technologies is that, as you get much larger die sizes, your yields and your costs go up.

"The way we're able to offer all of this I/O and all of this capability up and down the stack is because of that decision we made to go with the multi-chip module approach. I think this idea of breaking up large die is absolutely here to stay, so I am very pleased we made that bed."

As with all commercial products, the true arbiter of whether the right decision was made will be the market. Early customers announced by AMD include Microsoft and Baidu in cloud computing, and the likes of HPE, Dell EMC and SuperMicro in the server space.

"I think you will hear about other cloud providers deploying AMD in the second half of the year," Su said. In some cases, the main challenge has been to convince customers to put their trust in AMD, a company that may know about winning in the data center, but certainly knows about losing in it. Two years ago, with large debts and no profits, its share price hovered below $2, and its future was being openly questioned.

"At the end, it always comes back to 'are we reliable, do we have a good price-performance advantage, can they expect to use us for multiple years?'

"For many of these customers, if we only showed them Epyc, it would not have been enough, they've seen the second and the third generation, and that gives them confidence in investing in AMD."

She added: "I view this as a long term investment. [Intel] are very capable, but I think we have shown that we can be very, very capable as well."

After the interview, I wandered outside to the Epyc launch party. As the sun set on another blistering Texan summer day, a few AMD employees were suddenly chanting: "Twenty dollars! Twenty dollars! Twenty dollars!"

The company's share price had rocketed over the course of the week, gaining nearly 40 percent and standing at roughly $14. The dark times of near-bankruptcy seemed behind them, but as I contemplated the rise and fall of AMD's fortunes, I knew they still had a way to go. ●
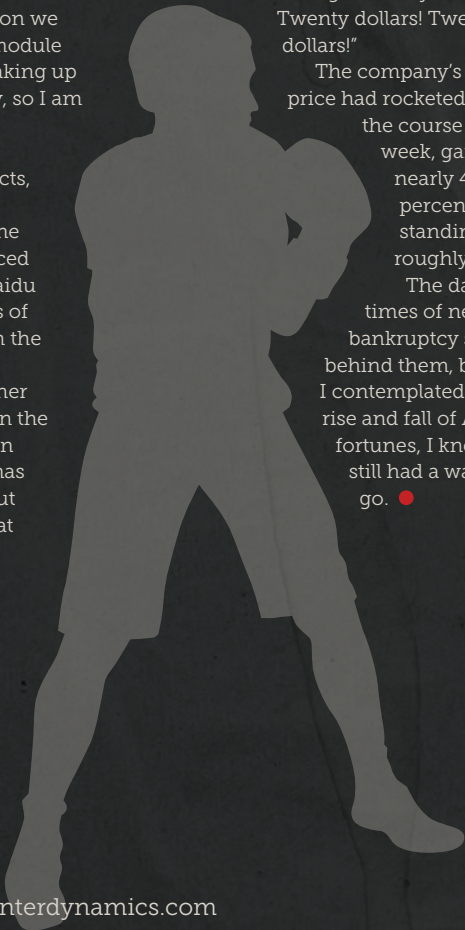
In the left corner:

Weighing in at a market cap of $12.5bn

AMD

Best known for performance per dollar

Founding member of The Green Grid

On Nov 2012, set the record for the highest frequency of a computer processor – 8.794GHz

!

Is a fabless semiconductor manufacturer, turning to for-hire foundries

Is behind Intel in CPUs and Nvidia in GPUs in market share

"THE CHAMPION"
INTEL

Max Smolaks
News Editor

In the right corner:

Weighing in at a market cap of $172bn

INTEL

Best known for performance per watt

Claims to be conflict mineral free

Acquired McAfee for $7.8bn, Altera for $16.7bn, Saffron for an unknown sum and Nervana for $400m

Accused by the European Commission of anti-competitive practices in 2007

Settled with AMD for $1.2bn after claims of unfair business practices in 2009

WE HATE SURPRISES

Intel sometimes refers to itself as the guardian of Moore's Law – an observation made by one of its founders, Gordon Moore, who stated that the number of transistors that can be packed into an integrated circuit will increase twofold approximately every two years.

This has been a double-edged sword: on one hand, it highlighted Intel's importance as the dominant supplier of CPUs, responsible for roughly 98 percent of the x86 server market today, and as a result, the direction of the industry as a whole. On the other hand, it made Intel's product launches as predictable as the rising and the setting of the sun: we know that performance of a new generation of chips will conform to the well-established tradition.

The fact that the Xeon Scalable Platform (Xeon SP) processor family, launched in June, delivers approximately 1.65 times more performance than last year's Xeon E5 v4 will surprise absolutely no one. The new CPUs feature more cores, more inter-CPU bandwidth, and support more memory channels. But they also introduce never-before-seen features that should, in theory, enable IT equipment to do new and exciting things.

According to Navin Shenoy, who assumed responsibilities for Intel's data center division in May, Xeon SP was designed to tackle three 'mega trends' that are changing the face of the industry: cloud computing, artificial intelligence and 5G connectivity.

"Enterprises will need to think about how they handle these new data workloads, seamlessly move them between public and private cloud," he said during the launch event. "Cloud service providers will need to think about how they improve performance, security, agility, utilization across their infrastructure. ▶
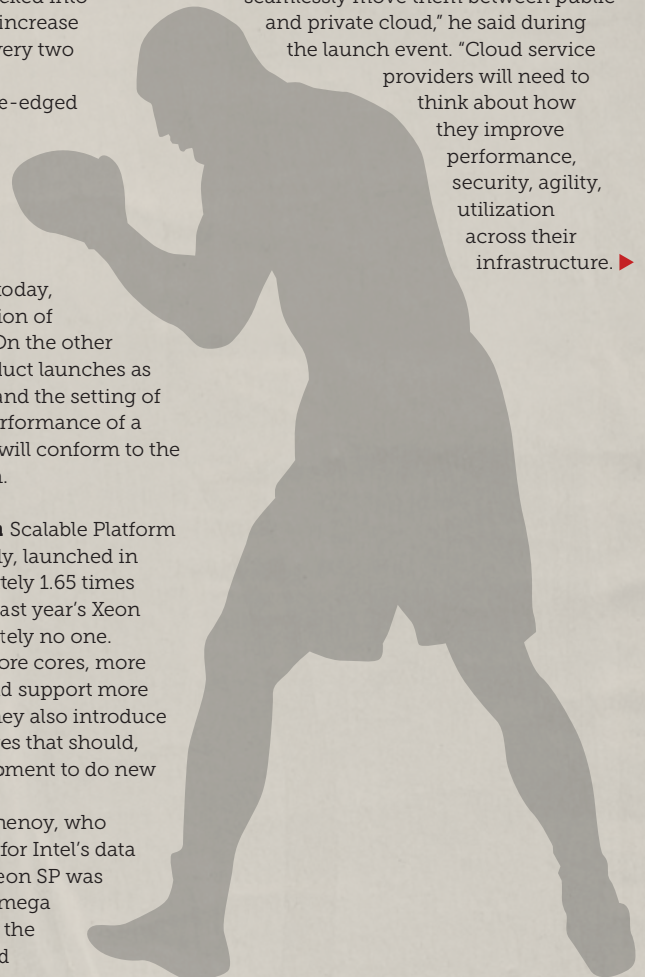
▶ Communications service providers will need to deliver a completely new level of network agility and efficiency, in the face of the new service demands as 5G is deployed. And in the world of artificial intelligence, we will see a broad range of algorithms develop, and this will require a broad range of solutions."

All new Xeon SKUs are now divided into four categories, based on their relative level of performance and target workloads: from Bronze for low-power applications to Platinum for high-performance computing. Theoretically, the same hardware could be made to serve very different purposes, depending on the choice of silicon.

The processor family is the first to replace the tried and tested ring architecture with a mesh that adds more interconnects between cores. It also supports a larger number of cores on the same die – up to 28, whereas Broadwell would max out at 22.
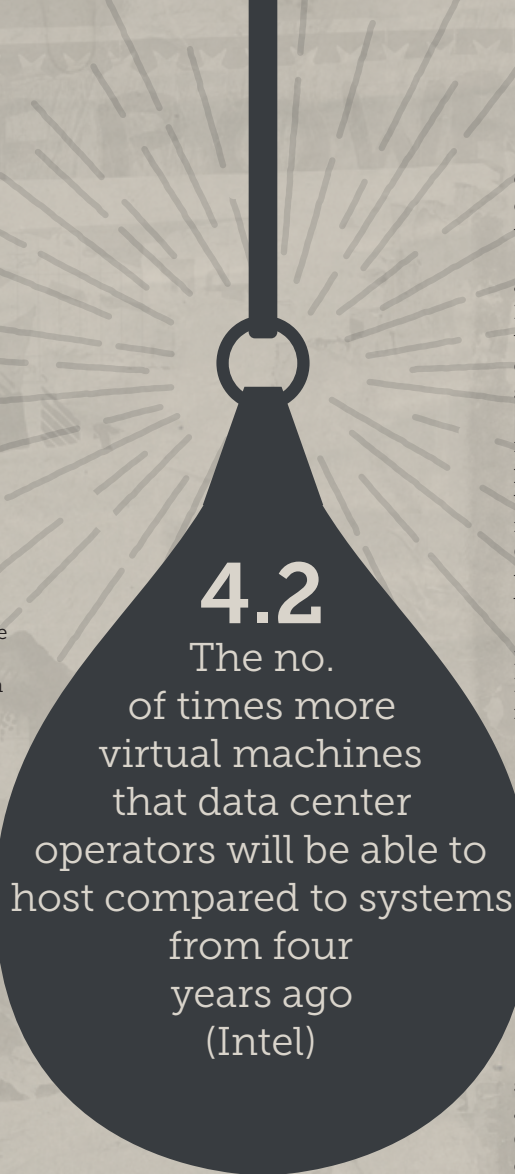
"It's a fundamental change to the processor design that allows us to increase performance by optimizing data sharing and memory access between all CPU cores," Lisa Spelman, the director of product marketing for Intel's data center group, explained.

"We realized that Xeon Scalable wouldn't actually scale the performance that we wanted to see from the product, if we didn't increase the bandwidth for the interconnect between the cores, the cache, the memory and the I/O. Without change, the interconnect, which adds a lot of value, suddenly becomes the limiter.

"That's why I'm excited about this innovation: it provides a more direct data path for traveling between cores than the previous ring architecture did."

One of the clear benefits of Xeon SP is increased server density: a single motherboard can now support up to eight sockets without the need for additional node controllers. This means system builders can squeeze more compute into a single rack unit, and for data center operators, this means an increase in demand for power and cooling per square meter. Average power densities in data centers are growing all the time, so this is nothing new, but a rapid increase could catch some providers off-guard.

"With PowerEdge FX, we can have four dual-socket Xeon SPs in 2U – that gives me 224 cores in 2U, with a large memory footprint, 1.5TB worth of RAM per sled," Mark Maclean, server product manager at Dell EMC, told *DCD*. "It's not just a speed bump – it's an architectural change. Estimates will always vary, but in certain workloads, we are seeing up to 50 percent performance increase."

4.2
The no. of times more virtual machines that data center operators will be able to host compared to systems from four years ago (Intel)

Xeon SP places more emphasis on security, with Intel's Quick Assist Technology – previously available as an optional extra, now bundled as standard. It involves a dedicated chip for tasks like compression, encryption and key management – and according to Intel, the new Xeons encrypt data twice as fast as the previous generation's processors.

Another interesting feature that aims to cash in on a change in server hardware is Intel Volume Management Device (VMD), which brings RAID-like features to your expensive, high performance NVMe solid state drives, making them hot-swappable.

"What this means is we can take advantage of some of the capabilities in the new CPUs, with enhanced bandwidth around PCIe, to have better NVMe capability," James Leach, director of platform strategy for Cisco UCS, told *DCD*.

"Because that's the real difference – when we were switching between SAS and SATA, it was very simple because we were routing the same kind of connectivity. NVMe depends on the PCIe bus, and we're just seeing the tip of the iceberg with some of the performance that it can offer.

"As the CPUs become more capable, and as we see more cores added through FPGAs and GPUs, we need to be able to feed those cores and NVMe is one way that we can really crank up the performance on the storage side."

With Xeon SP, Intel is indeed improving its HPC credentials: besides increased performance, the new Platinum chips are the first ever to offer support for AVX-512 instruction set extensions and integrated OmniPath networking. Both will be useful in the design of systems that deal with machine learning, analytics and 3D rendering.

Before the official launch, Intel supplied Lenovo with several thousand Xeon SP Platinum 8160 chips, which were then integrated into Mare Nostrum 4 – a beast of a machine hosted at the Barcelona Supercomputing Center. This immediately propelled it to the position of the 13th most powerful computer on the planet, with 11.1 petaflops of raw processing power.

But Mare Nostrum 4 is currently outperformed by supercomputers based on CPUs from IBM, Oracle and even AMD – that's got to hurt if you are positioning yourself as the world's foremost authority on chip design.

Lenovo and Intel plan to get a Xeon SP supercomputer into the Top 10 in November, and we'll see a lot more Xeons in the Top500 going forward – but whether the company can unseat the reigning champion, China's Sunway TaihuLight (based on processors from The National High Performance IC Design Center, Shanghai), is anyone's guess.

As for Intel's references to 5G – while there has been plenty of industry buzz around next generation wireless, the first networks based on this standard are not expected to appear before 2020, and will be limited to China. Meanwhile, major manufacturers still have no idea how to build this technology into handsets without having to double their size.

Intel will release several generations of processors before it needs to contend with 5G – there's simply no market for it at this stage. But there's plenty of market for servers that can run existing telecommunications networks, as more and more operators experiment with SDN and NFV in order to cut their costs.

Taking all of this into account, we can conclude that Xeon SP is much more than a marketing exercise. It offers real solutions to actual challenges in server design – with Moore's Law gradually losing its role as the force driving Intel's agenda. ●

ARM BIDES ITS TIME

While the big boys of the CPU market are slugging it out on the ring, a number of smaller, more agile chip vendors are quietly making inroads into the data center using cores designed by ARM (which was recently acquired by SoftBank). Despite the venerable (and vulnerable) Opteron branding, AMD's A1100 – launched in the beginning of 2016 – has failed to set the world on fire, and no one, including AMD itself, is even mentioning it these days – the Zen architecture looks like a much stronger contender.

Su told *DCD*: "I think ARM is a good architecture, and an important architecture. We partner with ARM, we use ARM in some of our semi-custom products. But I think relative to the data center, x86 is the dominant ecosystem and will continue to be the dominant ecosystem."

That hasn't stopped another two American companies – Qualcomm and Cavium – from having their own shot at the title of the welterweight champion. The former is an expert in mobile devices, and wants to apply its knowledge in the enterprise IT market. The latter used to specialize in networking, before deciding to try its luck with servers.

Qualcomm's Centriq 2400 is the world's first server processor based on the 10 nanometer process. For comparison, Xeon SP is still using 14nm, which was also used in last year's Xeon E5. The number of nanometers defines the resolution of the chip lithography equipment – smaller numbers mean more transistors on the same size of the die, increased speed and reduced power requirements.

Qualcomm's first foray into server chips offers up to 64 cores and will be able to run not just Linux, but Windows Server too – a testament to its lofty ambitions. It should ship in the second half of 2017.

Meanwhile, Cavium's ThunderX2 is a refined, updated version of the silicon launched in 2015, with new cores and I/O. The original chip was among the first to implement 64-bit on ARM, and the latest version continues the legacy of technical innovation, with up to 54 cores running at up to 3GHz. Each ThunderX2 supports up to six memory channels and up to 1.5TB of total memory – just like the latest Xeon SP.

The platform promises to be extremely workload-specific, with "hundreds of integrated hardware accelerators for security, storage, networking and virtualization applications," across four main varieties. Just like Centriq 2400, ThunderX2 is expected to ship in the second half of 2017.

This will be a fight to remember. ●

Why tape is still strong

Tape storage has been written off in the past, but it turns out the cloud era needs it more than ever, says *Dan Robinson*

Dan Robinson
Correspondent

Tape storage is one of those technological hangovers from the early days of computing, associated in the minds of many with rooms full of massive mainframe cabinets. Somewhat like the mainframe, tape shows no signs of going away just yet, and ironically, could even be handed a new lease of life thanks to the burgeoning volumes of data that are being accumulated in modern data centers.

Modern tape storage is a world away from the movie cliché of huge tape reels spinning backwards and forwards as the computer chomps its way through some complex computation. Today's tape drives use cartridges capable of holding up to 15TB of data, and are more often used for backup or archiving purposes than for online storage.

However, few in the data center industry can have failed to notice the rapid changes that have been taking place at the storage layer of the infrastructure. Flash-based solid state storage devices have enabled new tiers of low latency storage with higher IOPS, while hard drive makers have responded by continuing to push the storage density of rotating media, driving the cost per gigabyte ever lower.

The end result is that the cost of disk storage has fallen to a level where many businesses have begun to use disk-based backup systems where once they would have used tape drives or tape library systems. In addition, cloud-based storage services such as Amazon Glacier have proven attractive to businesses of all sizes for off-site storage of backup or archive data because of the low cost per gigabyte and the ease of transferring data over the Internet.

This does not mean that tape storage is about to vanish. For one thing, many regulated industries such as the finance or legal sectors have strict regulations which require providers to retain data and to be able to prove that its content has not been altered. Modern tape systems offer a write-once-read-many (WORM) capability that delivers this, and for this reason, tape is often mandatory for archiving data.

There are other reasons why tape is likely to be around for some time, according to Clive Longbottom, service director at analyst firm Quocirca.

"The biggest one is still investment protection: the cost of large tape libraries and robotic retrieval systems is high, and just dumping these because disks are now cheap (but failure-prone) is just not a good financial argument," he said.

"Then there is the ongoing cost. Sure, spinning disks are becoming cheaper and cheaper to acquire. However, keeping the disks spinning has a large ongoing operational cost due to the required power for spinning. A tape, once written, is close to zero cost – it holds its data until it is needed again. Hard disks can be spun down, but rarely are," he added.

Meanwhile, the shift towards cloud-based services for storage has simply moved the problem from the business to the cloud service providers. While the enterprise tape market has declined each year, cloud service providers are turning to tape as the optimal solution for backing up the ever expanding volumes of customer data they are storing, or for actually delivering archive services to customers.

"Cloud providers have a bit of a problem: they have put heavy focus on the incredible scale of their storage capabilities.

The trouble is that customers have fallen for the message. Therefore, the big players are looking at a need for zettabytes of storage capability to meet customer expectations," said Longbottom.

Fortunately, a large proportion of this data is unlikely to be accessed ever again, so if the service provider can figure out what data is likely to be accessed, that can go onto disk while the bulk of it can be written to tape with SLAs stipulating that some data may take an hour or more to be recovered.

Amazon does not say what technology its Glacier service uses, but it is widely believed that it is based on tape storage, simply because the retrieval times quoted to customers are as much as several hours.

Tape is well suited for archiving or long-term storage as it offers by far the lowest price points of any storage medium, with a raw storage cost of around $0.02 per gigabyte, and also boasts a potential longevity of several decades if stored under conditions of low temperature and humidity.

In the past, there were many competing tape formats, but most of these have largely given way to Linear Tape Open (LTO), which was developed as an open standard not controlled by any single vendor. IBM and Oracle still have their own proprietary formats while also supporting LTO.

LTO has been through multiple iterations, with LTO-7 introduced in 2015 delivering a native capacity of 6TB per cartridge, or up to 15TB with data compression. The next generation, LTO-8, is expected later this year or early next year, and is anticipated to boost native capacity to 16TB, with up to 32TB possible using compression.

IBM's 3592 series of tape systems has likewise been through multiple iterations, but the firm has recently introduced the sixth generation in the shape of the TS1155 Tape Drive, which offers a native capacity of 15TB, or up to 45TB using the 3:1 compression ratio that IBM quotes for the technology.

There is no sign yet of an end to increased tape capacities. Most recently (July 2017) IBM and Sony have pushed the record tape density to 200Gbits per square inch in an experimental process which uses sputtering, new tape heads and lubricant. This could lead to a theoretical maximum of 330TB in a single standard palm-sized tape cartridge, half the size of a 60TB SSD.

Compatibility is a key concern for technologies that will be used for long-term archival of information. For this reason, the LTO Consortium enforces strict rules to ensure that any LTO drive can read cartridges from the two preceding generations as well as its own, and can write data to cartridges from the previous generation. IBM's TS1155, for instance, supports existing JD and JC tape cartridges.

If tape vendors can continue to boost storage density, and keep the price per gigabyte of tape at rock-bottom levels, there is no reason why the old medium should not continue for several more decades for backup and archive.

"An enterprise with just less than a petabyte of data should focus on disk-based backup and archive. Greater than that, and I'd be looking at how and where tape could possibly play," said Longbottom. ●

330TB
tape cartridge promised by IBM and Sony

DCD >Awards | 2017

The Data Center Awards are Open for Entries!

Until September 29th

Category 1
Living at the Edge Award

Category 2
The Infrastructure Scale-Out Award

STULZ

Category 3
The Smart Data Center Award

Category 4
The Data Center Eco-Sustainability Award

Schneider Electric

Category 5
Energy Efficiency Improvers Award

StarLine

Category 6
Mission Critical Innovation Award

CBRE

Category 7
The Open Data Center Project Award

RITTAL

Category 8
Cloud Journey of the Year

Category 9
Data Center Operations Team of the Year – Enterprise

Category 10
Data Center Operations Team of the Year – Colo+Cloud
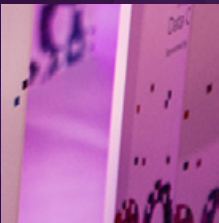
Category 11
Design Team of the Year

VERTIV

Category 12
Best Data Center Initiative of the Year

Category 13
Young Mission Critical Engineer of the Year

Category 14
Business Leader of the Year Award

Category 15
Outstanding Contribution to the Data Center Industry

MERCURY

Charity Partner

CHAIN of HOPE
mending little hearts

For more sponsorships and table booking information, please contact **Dan Scarbrough** at **global.awards@datacenterdynamics.com**

www.dcdawards.global

Open for Public Voting
October 16 2017

Category 16
World's Most Beautiful Data Center

QUALITY UPTIME SERVICES

A wrinkle in time

The Internet has ways to keep time, but they may not be good enough for a new breed of applications and regulations, reports *Sebastian Moss*

Sebastian Moss
Reporter

Time synchronization solutions

In centralized systems, the solutions for asynchronization are simple as the main server dictates time to its counterparts.

Examples of such solutions are Cristian's algorithm and the Berkeley algorithm. The former relies on the existence of an accurate time source connected to the centralized server and variations of this algorithm exist to account for the network propagation time. The Berkeley algorithm, on the other hand, works in systems where a time measuring device is not present. As such, a time server will periodically ask clients their time measurements, average them out and send the adjusted time back.

In distributed systems, the clock asynchronization is more visible as parts of such a system might reside in different time zones and within different environments. For this, the most widely used solution is NTP (Network Time Protocol). Tried and tested over the years, this is now the norm for time synchronization across the Internet.

Because of network latency this protocol can still fail in certain cases where an offset of a few milliseconds is not acceptable, in which case PTP (IEEE 1588 Precision Time Protocol) can be used, which, when coupled with Synchronous Ethernet, can deliver sub-nanosecond synchronization.

On January 26, 2016, time went wrong. As the Space Command division for the the US Air Force began to decommission SVN 23, a satellite in the GPS constellation, things went awry – by a whole 13 microseconds.

Fifteen GPS satellites broadcast the wrong time, immediately causing issues on the Earth below. In a post-mortem, time-monitoring company Chronos detailed numerous anonymous telecoms customers who suffered 12 hours of system errors, while BBC Radio also experienced disruptions.

As those affected by the error can attest, accurate timekeeping has become increasingly important to computer networks. Distributed systems need to be synchronized, and some require traceable and accurate timestamps for everything from financial transactions, to alarm events, to the time and duration of phone calls.

Network Time Protocol (NTP) servers do some of the work, but mission-critical or large-scale applications often run their own network time servers for increased security and redundancy. For more accurate time synchronization, Precision Time Protocol (PTP) is used (see boxout for more).

With new regulations around the corner, understanding how to keep time, and how to prepare for when clocks go bad, is becoming crucial for an increasing number of businesses.

For example, the financial sector will be subject to the EU's Markets in Financial Instruments Directive II (MiFID II) from January 2018, in which subsection RTS 25 requires timestamping to be 1,000 times more accurate than is required by current legislation.

MiFID II "has been put forward in a very, very robust way," the National Physical Laboratory's strategic business development manager Dr Leon Lobo told *DCD*. ▶
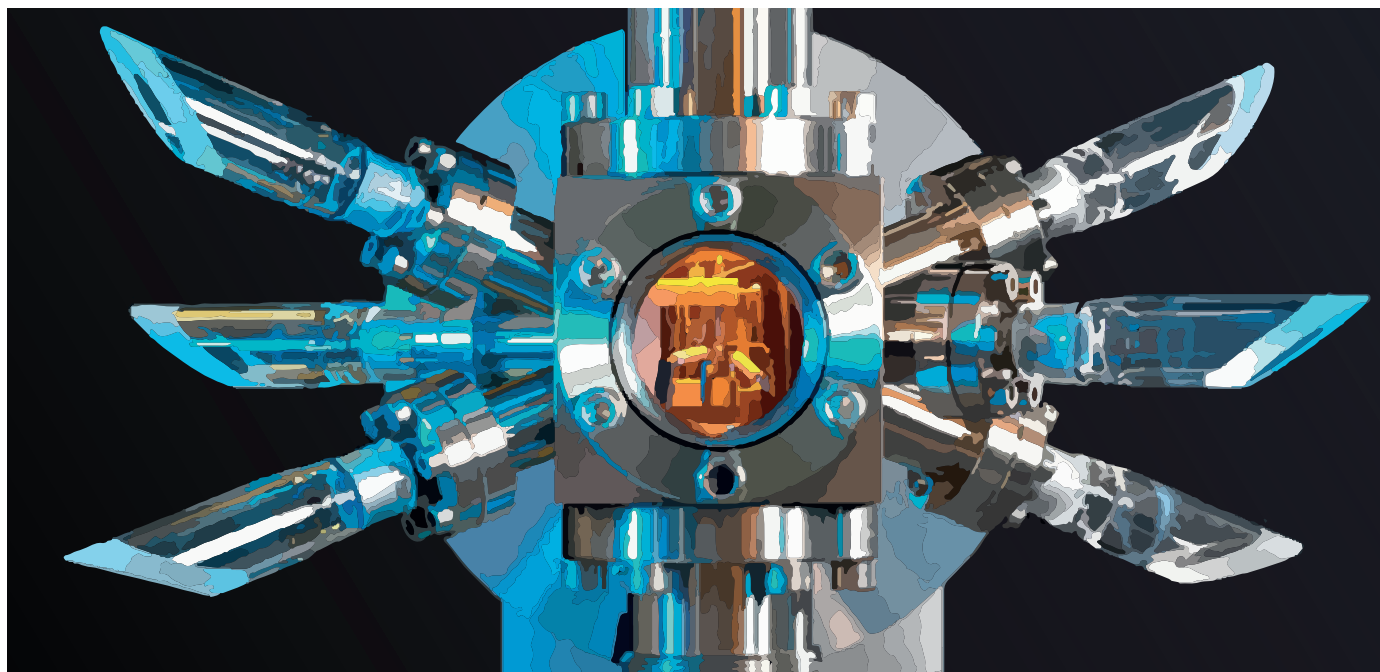
▶ "They've specified the traceability requirements, and the figures that they've put forward necessitate infrastructural upgrades. IT systems need to be upgraded."

NPL acts as the UK's National Measurement Institute, looking after Coordinated Universal Time (UTC), the world's primary clock standard.

Another regulation to watch out for is the Payment Card Industry Data Security Standard. Jeremy Onyan, director of time sensitive networks at precision time and frequency instruments manufacturer Spectracom, said: "There's a section there, 10.4, that's all about accurate and traceable time. Traceable is the key word because if you're using time off the Internet, it may seem sufficient for your application, but it's in no way traceable.

"If there is ever an event, especially in a distributed system, and you wanted to correlate how hackers got in and what they did, it's almost impossible to do because the disparity in the timestamps if you're using Internet-based time is so big that it's not even usable in a lot of cases."

Most organizations trying to track time turn to satellite receivers, but are then reliant on the satellites being accurate – predominantly via GPS, but also via other Global Navigation Satellite System (GNSS) networks such as GLONASS, Galileo and Beidou.

A lot of the receivers used for this task "are repurposed navigational receivers, that are dumb or naive – they're just taking the time from the satellite stream, and they're not qualifying or error checking it," Ron Holm, sales and marketing manager of time and frequency company EndRun, said.

"In the case of the January 2016 error, a smart receiver would have looked at that, done some integrity checks, not only on that part of the GPS data, but on other aspects of the GPS data, to reject that input."

Even with a smart receiver, however, the GPS signal itself can be susceptible to jamming or spoofing.

"Jamming is becoming more prevalent particularly because the devices are very cheap to buy and the jamming is often inadvertent," Dr Lobo said. "It's a van driver taking out the tracker in his vehicle to prevent his boss knowing where he is. So he inadvertently is taking out the city block – that sort of an event is occurring on a regular basis."

Onyan concurred: "Look at most data centers, where are they typically located? They're in warehousing districts or manufacturing districts, places where they've got lots of big buildings, and you've got trucks driving by all the time. All it takes is somebody driving by with a jammer."

Much of the conversation around GPS jamming is anecdotal, with Equinix's LD4 data center in Slough being one of the few to publicly discuss GPS interference. In that case, the issue was with a 20 year old 'rogue GPS antenna' that was reflecting a powerful GPS signal and accidentally jamming those around it.

Outside of data centers, one example frequently cited is that of Newark Airport, which in 2009 opened a new GNSS-based landing approach system. "For two years they were struggling to get it to work – sometimes it'd work great, sometimes it wouldn't, and they just couldn't figure it out," Onyan said.

"So they called in the FCC who came in with some special equipment and discovered that there was a red Ford pickup truck that one of the contractors working the site was using which had one of these jammers installed. Every time he was on site, he was taking down their landing approaches without meaning to."

In a 2014 report on such GNSS interference and jamming, Chronos looked at how widespread the issue was at the time.

"Jamming is getting worse," the paper by Professor Charles Curry said. "Some probes are now detecting five to ten events per day; over 50 websites are actively selling jammers; and the devices being seized by law enforcement agencies are now more powerful and so have considerably greater jamming ranges.

"Whereas in 2008 GPS was the only satellite [Positioning, Navigation & Timing] system under threat, and jamming targeted its L1 frequency alone, now all frequencies of all GNSS are under attack."

When a signal is jammed, or otherwise interrupted, time servers have a backup ready to take over. They go into 'holdover' and turn to what are essentially uninterruptible power supplies for time: crystal oscillators. These electronic circuits use the mechanical resonance of a vibrating crystal of piezoelectric material to create an electrical signal with a precise frequency.

Holm explained: "The base level oscillator is typically a TCXO, a temperature compensated crystal oscillator, and that'll provide acceptable performance for 24 hours. The next step up is an ovenized oscillator, that will provide acceptable time for approximately 30 days. An atomic oscillator provides time for several months."

Depending on how much one spends, these oscillators are meant to keep time while the jamming, or more likely broken cable or roof antenna, are mitigated or fixed. That still leaves spoofing, which is "particularly difficult to detect," Dr Lobo said, "because it's about somebody drifting the time rather than it being an out and out denial of service.

"The worry is that spoofing devices are becoming easier to own and to manage. They are becoming software-defined radios, as opposed to very expensive hardware solutions."

The University of Texas at Austin's Todd Humphreys looked into the possibility of spoofing data centers in the financial sector. In a research paper, he noted that "traders could use GPS spoofing as a weapon against each other. A trader could manipulate, via GPS spoofing, the timing of a competitor's trading engines, driving the competitor out of the marketplace during a crucial trading interval. After the attack, the rogue trader covers his tracks by bringing his competitor's timing back into proper alignment with true time."

The solution proposed by Spectracom is to turn to a different satellite system. Instead of, or in addition to, a GNSS signal from a satellite system 16-21,000 km (10-13,000 miles) above Earth, the company offers access to the Iridium constellation launched by Motorola in the '90s, in low earth orbit – at 800 km (500 miles).

"This is encrypted," Onyan said. Which makes it harder to spoof "because you need to know both the encryption key, which is RSA-level encryption, and you also need to know the serial number of the specific receiver in question."

The other advantage is that the signal is roughly a thousand times stronger, "which means it can work indoors in most environments."

Dr Lobo has another idea: to do away with satellites as the primary timekeeping method altogether. NPL has started to offer a precise time service, NPLTime, over fiber in the UK, with plans to expand access to Europe and beyond.

"We are delivering over fiber infrastructure in a completely managed way, so we are monitoring the latency to every endpoint on the network and effectively offsetting the time at those endpoints, as a result what's coming out of those end points is exactly the same as what's going in," Lobo said.

"We see it as the replacement where GPS could form a second priority input as a backup system, but the primary traceability requirement for MiFID II is easily satisfied by this managed service that we are providing, where we provide certification of the accuracy at that point of ingress in a customer's infrastructure."

The time measurement is provided directly by NPL's cesium fountain clock, NPL CsF2.

In it, a gas of cesium atoms is introduced into the clock's vacuum chamber. Six opposing

All the time off the world

The GPS project was started in the United States in 1973, with the first satellite launching five years later. While initially developed for the US military, it was opened up to civilian use in the 1980s - quickly becoming the predominant navigation system and way to synchronize time around the world.

But telling time from space is not an easy process, with matters of relativity coming into play.

The atomic clocks on the satellites need an accuracy of 20-30 nanoseconds, as observed by those on the ground.

However, with the satellites not in geosynchronous or geostationary orbits, they move in relation to Earth-based observers, which Special Relativity predicts makes the clocks appear to tick more slowly, from our perspective - in this case to the tune of about seven microseconds per day.

Meanwhile, as the satellites orbit at 12,550 miles (20,200 km), spacetime is less curved by Earth's mass than on the surface, which General Relativity predicts means the clocks will tick faster (from the perspective of the Earth) - this time to the tune of about 45 microseconds per day.

Together, if left unchecked, the satellites' clocks would be 38 microseconds faster than our time on our planet, quickly rendering the system useless.

Thankfully, the satellites were designed with this in mind. The atomic clocks are slightly slower than those on the ground to accommodate for General Relativistic time differences. At the same time, an onboard microcomputer computes Special Relativistic timing calculations as required.

infrared laser beams then "cool the atoms down to near absolute zero temperatures and stop them vibrating."

The two vertical lasers gently toss the ball upward, before letting it fall back down in a round trip time of about one second.

"Effectively if it was considered to be a clock, it would lose or gain a second over 158 million years," Lobo said.

But NPL is not content to stay at that level of accuracy – it is currently working on a clock where "you're looking at losing or gaining a second over the lifetime of the universe – 13.8 billion years. So there's some pretty significant capabilities there."

Questions remain over whether data centers will need that level of temporal accuracy, but one day we will find out – after all, it's just a matter of time. ●
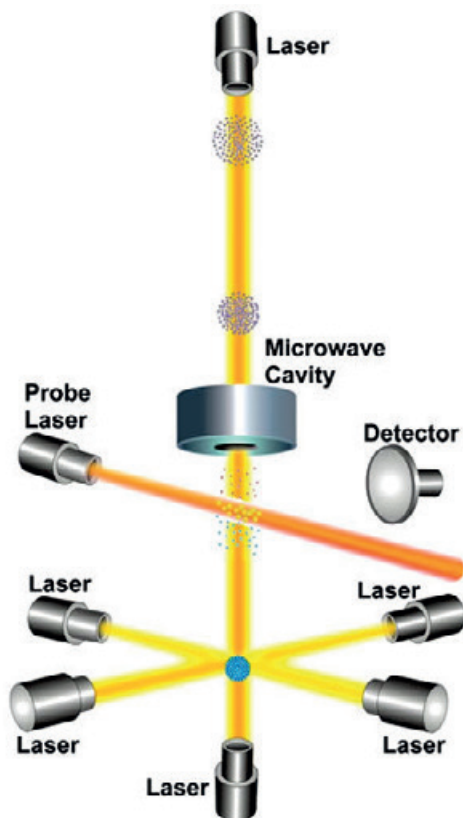
Virtually there?

Will virtual reality simulation become an indispensable tool in the data center industry? *Tanwen Dawn-Hiscox* immerses herself

Simulation is becoming integral to data centers: as architects ponder building plans, and sales teams show off a virtual facility to customers unwilling to dive into a contract without visiting the site. Even engineers and technicians increasingly use simulations to improve data center efficiency, redundancy and capacity.

Future Facilities, a London-based software company which specializes in tools for the design and operation of data centers, is testing out what it believes could be the next step in data center operations: VR simulation.

Since its inception in 2004, the company has been refining its modeling software, which is based on computational fluid dynamics (CFD) – the study of fluid and gas flows using numerical analysis and data structures – to obtain quantitative predictions and simulate the interaction between gases and fluids in the data center.

To do so, it collects live operational data from the facility's monitoring and DCIM tools; these are then informed by guidelines and standards, such as measurements for data center resource efficiency as defined by the Green Grid and ASHRAE's thermal and humidity guidelines.

Tanwen Dawn-Hiscox
Reporter
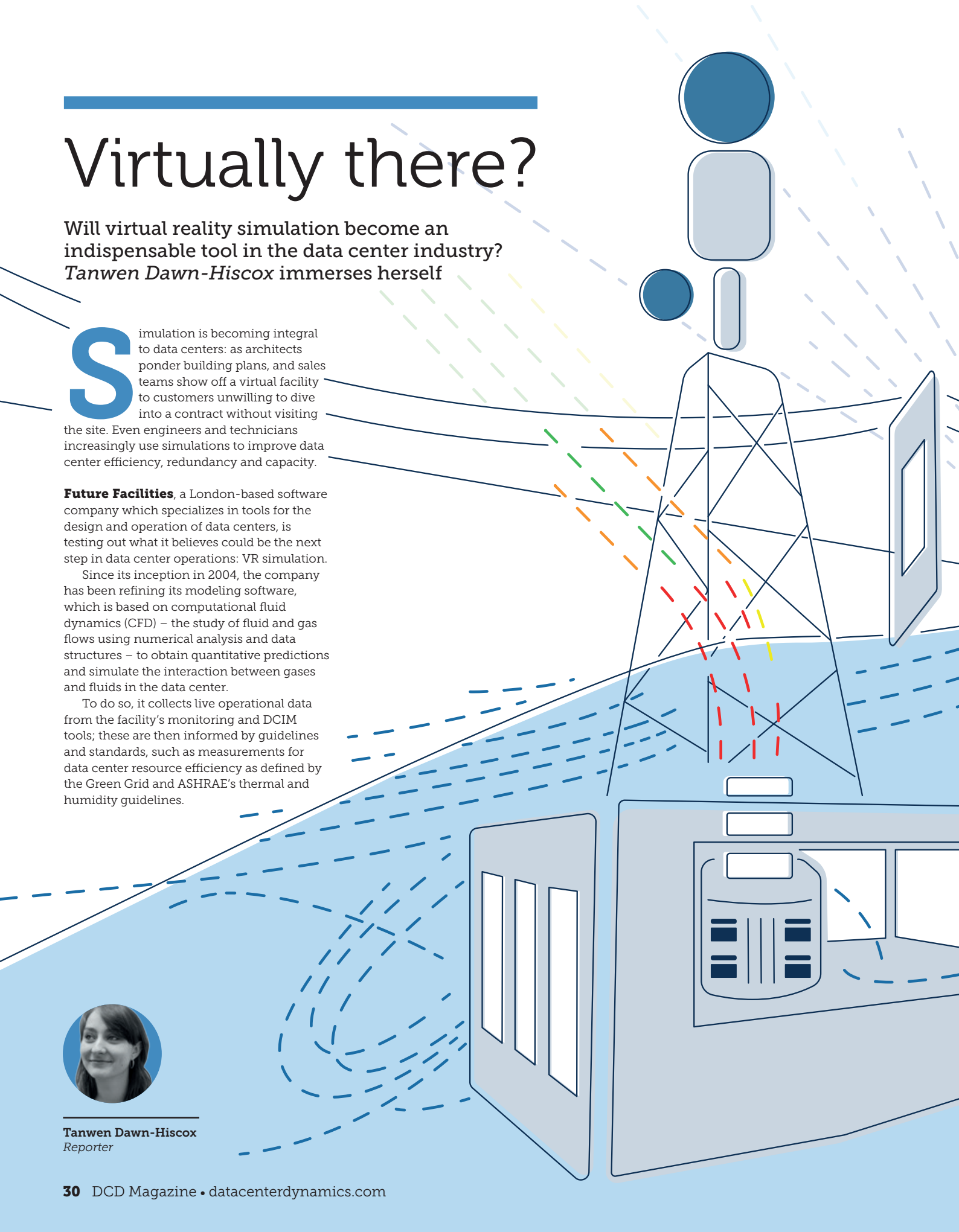
Initially, the company mostly dealt with clients seeking to understand operational failures, but eventually customers turned to it for general planning and pre-emptive purposes.

"That is the number one benefit of simulation: whatever the change, you can do it upfront, it's a kind of virtual playground. You've got a model of your room and you can do whatever you want to do, whether it is maintenance on a cooling unit, or installing new IT equipment," said Mark Fenton, chartered engineer and product manager at Future Facilities.

As well as whitespace modeling, the company models generator units, cooling plants, and its simulations take into account internal and external factors that can affect design and operations, like the weather: "You could build a beautiful whitespace and not get your cooling stuff quite right and end up with a really awful performance and actually overheating even if you've designed that whitespace well."

Future Facilities' team was initially skeptical about the importance of virtual reality, but after toying with the technology is seeing much potential.

The company has developed a demonstration using Oculus Rift which allows one to wander through a series of simulations of data centers throughout the ages. From the 1950s, when a data center was effectively just a single low powered, uncooled mainframe, through the '80s – a time of blue carpets, glass door racks, monitors on shelves and untidy cables – and the 2000s, when operators discovered the joys of raised flooring and contained aisles.

In the simulation, one can overlay all sorts of data: airflows and their temperature are represented by arrows in a gradient of colors ranging from deep blue to red. One can, for example, check the operational status of the cooling equipment and simulate its reaction to different actions (helpfully color coded in green for ideal temperature, red for overheated, and blue for overcooled), or test the effect of a new piece of IT.

The immersive demonstration gives an idea of the industry's progress in understanding how to run a data center, from zero planning to a higher level of complexity in design and execution. The development makes sense in that it follows the industry's learning curve (and by extension, the company's).

The final virtual room, an edge data center containing a Vapor IO chamber, a self-sufficient cylindrical block containing six racks and an integrated adiabatic cooling system, could bring one of the possible use cases for VR in the data center:

If, as is predicted by some with the emergence of 5G, we bring data centers closer to us to power our equally hypothetical autonomous cars, so-called smart cities and virtual reality social media, then offices, former phone boxes, rooftops and cell towers may contain a handful of racks.

In this scenario, our IT would either need to be fully self-healing, or technicians would need to have access to multiple facilities at once. So, theoretically, were everything to be software-defined, a single person could sit in a control room with a mask on, and remotely operate dozens of edge sites at once without having to make the strenuous journey to every one of them. ▶
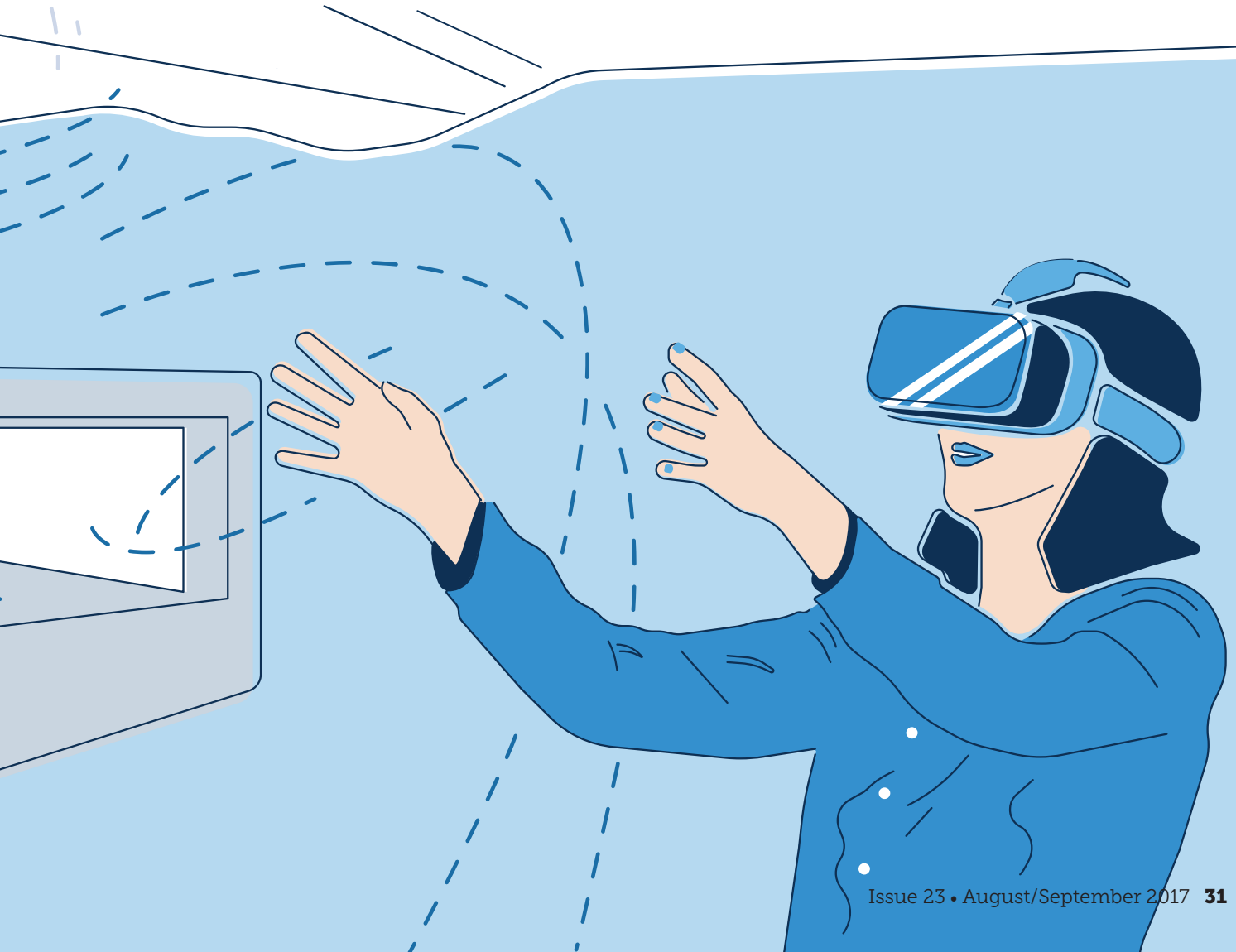
▶ But before this can happen, the next step in improving existing VR simulation technology has to be the ability to interact with the model – and the real world. "This is the early stage beta, but really the next thing we want to be able to do - because at the moment you can control where you move and where you look but you can't interact with the model in the same way that we eventually see it happening – would be that you could pick a device, get some information about it, move it, decommission it, whatever."

Such advances could prove useful for training purposes. With the data center industry growing steadily all the time, there is a massive shortage of specialized staff. An Irish university recently launched the first ever Bachelor's degree in data center engineering, an online course that will require its students to visit the north western Irish town of Sligo in Ireland or Le Mons in Belgium, for practical sessions. In theory, with interactive VR, they could do these modules remotely.

Future Facilities even sees itself taking the technology one step further, and using augmented reality in the data center. One could wear a Google Glass-like optical head-mounted display (OHMD) device whilst walking around a facility, and overlay DCIM tools and simulations to get information streamed directly in front of one's eyes.

For Mark Fenton, "that's kind of where we see the next stage for VR and then the final bit we really want to get to." But for now, the benefits of a lot of these new technologies have yet to reach their full potential.

An important factor in that process, according to Future Facilities director Jonathan Leppard, is a lack of courage (and of course resources) in the industry.

"It's almost like no-one wants to put their foot forward first. The financial sector never does it, it's going to wait until generation two or generation three. It takes someone to go "this is it, this is how it goes" and it's going to be the Google and the Facebook, the big carriers. And whoever it will be is going to take over that area."

As it stands, taking modeling into virtual reality for marketing, pre-sales and education feels like a small step up from its less immersive predecessor.

But being able to interact with a VR model of a data center would open a whole new page of potential for engineers and technicians.

Before this happens, however, DCIM and monitoring tools, network functions virtualization (NFV) and software-defined networking (SDN) will need to become ubiquitous.

"We're sitting on a plateau of technology at the moment; I don't see how data centers are moving forward quickly at this point in time. But just as soon as that next level goes and you'll see a whole new raft of data centers, big, small, in water, in space, doesn't matter. Once it becomes automated and self-healing and managed, then you'll see an explosion in the industry, that's when we'll take off again to another step of progress."

For now, though, we can only speculate. We're sitting on the verge of something that feels important, but isn't quite there yet. ●

Keeping You Ahead with Sify Hyper Connected Data Centers

Keeping your business on cloud nine with futuristic Data Center services is what we excel at. We are Sify. We design and establish Data Centers, and are one of the India's leading Data Center service providers. Our enviable track record of providing uninterrupted services extends to over 9,000 organizations across sectors and scale.

Sify Technologies provides data center services like colocation, managed hosting, DC IT services and DC migration. Through our evolved infrastructure and refreshed technology, we offer state-of-the-art ICT solutions that keep you ahead. As your technology partner, we empower you with 100% uptime and transform your business by opening innovative possibilities. Our Data Centers are ideally suited to be both the primary Data Center and DR. These are strategically located in different seismic zones across India, with highly redundant power and cooling systems that meet and even exceed the industry's highest standards.

Our impressive portfolio of over 425,000 sq.ft. of server farm is spread across 6 Tier III Data Centers, 15 Tier II Data Centers, 6 State Data Centers and several more for private clients, all built to exacting specifications and best-in-class global standards.

Hyper Connected Data Center Features

- 43 Data Centers connected by high capacity fiber access
- Internet Exchange present in Data Center in Mumbai
- Zero Latency access to AWS Direct Connect in Mumbai
- Google POP enables peering within Data Center in Noida
- Carrier neutral with presence of major telecom operators within Data Centers

With over 17 years of operational and technical expertise serving over 300 DC customers spread across BFSI, telecom, pharma, retail, manufacturing, media, etc. As a cloud and network services provider, Sify Technologies understands that reliable and affordable connectivity is key to leveraging data center and cloud investments made by organizations. Sify's cloud cover connects over 43 data centers, across the country, on a high speed network. The company provides high capacity multi-protocol low latency networks across multiple cities in India to provide solutions to address the unique requirements for data center – data center traffic as compared to traditional data center – branch traffic. Sify Technologies has marquee customers who leverage this network to connect their disaster recovery and near disaster recovery facilities to their primary data center.

Sify Technologies has evolved from a network and data center service provider to becoming a full-fledged converged ICT player with capabilities for data center transformation, application integration and transformation integration services.

Sify Technologies has demonstrated its prowess in many DC transformation projects by implementing best-of-breed solutions which enabled customers to experience the best-fit solution to address their current and future requirements and de-risk technology adoption. This transition was played on the premise that the brand had built a sound infrastructure foundation and aligned services on them in sync with the ever evolving needs of enterprises. Therefore, it believes that it is able to serve the requisite services with the same level of SLAs across the board.

Sify Technologies is also well positioned to address the increasing market opportunities on upcoming platforms of Software as a service (SaaS), Platform as a service (PaaS) and Infrastructure as a service (IaaS) as enterprises move from build to subscribe and from control to visibility.

It is positioned as one of the leading vendors in the prestigious Gartner's magic quadrant for cloud enabled managed hosting, Asia Pacific for the past three years. Its cloud platform has scale-out capabilities to compete more effectively for public cloud requirements and is also extending its support to AWS and Microsoft Azure, including offering managed services on top of third-party cloud providers.

As a full scale ICT player, Sify Technologies is in the best position to align these expectations to the market place for any enterprise.

Mumbai Data Center

sify

Contact Details
Mark Ryder
Managing Director Europe
Phone: +447429169426
Email: *mark.ryder@sifycorp.com*

Do we curb demand - or feed it?

Data centers' insatiable demand for power may actually be just what the utility grids need – and the thing that has to change is not the infrastructure but the people. *Peter Judge* reports from DCD's Energy Smart event in San Francisco

Peter Judge
Global Editor

Many people are asking if data centers use too much energy. The real issue is bigger: Can data centers and the energy supply industry work together to deliver electricity and online services which meet our needs... without costing the earth?

Ten years ago, a report from Lawrence Berkeley National Laboratory (LBNL) warned that data center power usage was out of control. Last year an updated report found that energy use by US data centers has actually leveled off. But there is a lot more to the story than that, according to senior executives from both the data center and the power industries, who gathered for *DCD*'s first Energy Smart Summit in San Francisco in June.

Online services are expanding and developing rapidly, making data centers a fundamental part of the infrastructure supporting human society. According to Jim Connaughton, a former presidential environmental adviser now working for Nautilus Data Technologies: "They are the foundation of the new economy."

But data centers are intrinsically fast moving. The IT equipment is replaced every three to five years,

Dean Nelson | Infrastructure Masons

and the cloud applications running on it can have a lifetime of months. This Protean package ironically depends on a much less flexible life support system: the electricity grid.

"Energy investments are intrinsically generational in length, or you can never make them pay for themselves," explained Don Paul of the University of Southern California Energy Institute. "But the digital world, which is wrapped around that now, like two strands of DNA, is cycling around at ten times the rate."

The two worlds are very different, said David Rinard, director of sustainability at Equinix: "You have an entrepreneurial approach [in the cloud] versus a blue-chip, older approach to the network grids." And the consequences of their failure is different too: "If you lose the Internet you go back to the 1980s. If you lose the grid, you go back to the 1880s."

But regardless of the timescales, there are big changes happening in both worlds, which have an impact on the way both services are delivered. Energy grids are moving (unevenly) towards renewable sources. And data centers are changing the way they handle reliability. Both of these changes are interlocking.

David Murray | Hydro-Québec

It seems that while we may want to curb the data center sector's energy demands, their very size could actually help the utility sector in its move to renewables, by being good, and very large – and demanding – customers.

Gary Demasi, Google's director of data center energy and location strategy, said: "We have always locked horns with the utilities. They are not structured to deliver us the product that we want, and we've really got to challenge that. We've been successful, but it's been a very rocky road and we've had to do a lot of what I would consider to be unnatural things."

At first, Google found utilities reluctant to sell it renewable energy, but as a champion of power purchase agreements, Google is now the world's largest renewable power customer.

In Nevada, Switch
helped change state policy so renewable power contracts are available to all, because it has purchasing muscle, according to Adam Kramer, vice president of strategy: "In 2016, in Las Vegas, Switch accounted for the entire growth in a large rate class which includes a lot of casinos. The growth rate was one percent."

"We are the utilities' best customers, because we are have a consistent baseline of demand," pointed out Dean Nelson, CTO of Uber and founder of Infrastructure Masons. "That demand is going to increase which will help stabilize the demand for the grid."

Connaughton believes "it's perfectly ok to have growth. If data centers had more demand, we could actually refurbish some of the sagging infrastructure faster, because there's more money flowing to the utilities."

Beyond that, data centers have backup power systems, which could be used to help the grid out, perhaps by using them so data centers can power themselves at peak times in so-

called "demand response" schemes. Data centers could in this way effectively "erase themselves from the grid" at critical moments, said David Murray president of distribution at Hydro-Quebec, Canada's largest utility. Alternatively, they could even sometimes contribute power from their backup sources to the grid.

There is "a dire need for energy efficiency to be married with demand response," according to Priscilla Johnson, data center water and energy strategist at utility PG&E.

The idea of sharing precious backup systems has usually been anathema to data center operators focused on delivering a reliable service, but that could change, if they deliver reliability in different ways, according to Peter Gross, vice president of mission critical systems at fuel cell vendor Bloom Energy.

"Historically there was always a buffer between the utility and the server," said Gross. "We had this very complex infrastructure support system consisting of UPSs and generators and transfer switches. It was expensive but resiliency and reliability drove pretty much every decision in the design and operation of data centers."

Things are different in 2017, where the cloud is increasingly using a "distributed resiliency" concept where the reliability of individual data centers is not quite so vital, he said.

So is everything going smoothly? Don Paul isn't so sure. The leveling of data center energy demands found by LBNL could just be a temporary hiatus caused by a period when efficiency gains canceled out the uncontrolled growth in digital demand.

"People are underestimating the digital growth rate," warned Paul. Energy demand is flat because there have been massive efficiency gains, but they have been consumed by increased demands. "We've consumed our efficiency gains, just

like we always have. History shows that as economies grow and civilizations advance, they always consume their efficiency gains."

Energy demand will rise again, he says, "unless the efficiency per byte continues to have dramatic changes," said Paul.

The demand for digital services is limitless – as long as they appear to be free, said Nelson: "We're not going to change human behavior."

Gamification might be one approach to try to change behavior, said Dr Julie Albright of USC's department of applied psychology. Utilities could feed energy usage figures back to consumers and apply social pressures so they compete with friends to use less. "We could make it fun," she said, "though the word 'fun' and the word 'utility' in the same sentence could be a problem."

David Rinard of Equinix agreed, taking us back to the culture clash between utilities and tech. The tech world makes easy to use smartphones, while the tools provided by utilities are "not of the same consumer grade quality. They are not intuitive and easy to use. They are not something you would find at Best Buy."

In the end, data center operators may have to change people's behavior, rather than themselves. ●

DCD is planning more Energy Smart events. For news and coverage, check out features and videos at bit.ly/DCDEnergySmart

> Between 1955 and 2005, computations per Watt increased by 1bn

Dr Julie Albright | USC

DCD > Webinars | 2017

ON DEMAND

Are you Hybrid Ready? It's Where IT is Going

Speakers:

Chris Ortbals, Executive Vice President of Product Management, QTS Data Centers
Andrew Boardman, Director of Sales, North East and Mid-Atlantic, QTS Data Centers

Moderator:
Stephen Worn, CTO, DCD

This on-demand webinar examines how a hybrid methodology is impacting IT management, data center and cloud deployments and the ways that providers have evolved to meet a growing demand for fully-integrated, hybrid solutions.

Watch and listen, as our subject experts address four key issues, including the challenges IT managers face when implementing a holistic hybrid strategy.

▶ **Watch Now:** bit.ly/HybridITWebinar

ON DEMAND

4 Steps to a Successful High-Speed Migration Strategy

Speakers:

John Schmidt, Vice President Global Data Centre Solutions, CommScope
Peter Judge, Global Editor, DCD

This recent one-hour webinar is packed full of insight and provides clarity on the subject of data center migration. Learn how the shift to collaborative, parallel processing is impacting infrastructure design.

▶ **Watch Now:** bit.ly/MigrationWebinar

For more information please contact **webinars@datacenterdynamics.com**

Foundations of the smart city

Chris MacKinnon
Canada
Correspondent

Data centers that support the smart grid, but are independent from it, will be the first step towards building smart cities, says *Chris MacKinnon*

The smart city, built on and around the Internet of Things (IoT), is an interconnected network of devices that gather, store, and share data while communicating with one another to improve efficiencies across a variety of functions. Smart grids, on the other hand, utilize the same IoT concepts, but in the energy grid space where two-way communications between the utility company and the end customer can improve provision of energy where and when it is needed.

Fred Tanzella, CTO of Green House Data, defines the smart concept well.

"By providing power to cities via a smart grid," Tanzella told *DCD*, "cities can communicate loads, minimize costs, restore power faster after a disruption, and integrate renewable energy sources, and provide reliable operations in multi-carrier networks."

In Tanzella's opinion, the smart grid promises to let data centers take advantage of energy price arbitrage by moving loads between geographical areas served by different different utilities. He says energy costs fluctuate across the countries and time zones. "Running

mission critical applications in time zones that are not in peak power times provides a methodology for saving cost and energy. Renewable energy sources that the Smart Grid will have access to will allow data centers to supplement peak power requirements efficiently."

Neetika Sathe, director of advanced planning at Alectra Energy Solutions, agrees that the smart grid can transform data centers. IT is changing quickly, and many data centers around the world are outgrowing their existing capacity to power and cool IT systems. When connected to a legacy distribution system, these data centers need to upgrade or ▶

▶ replace their power networks, transformers, switches, UPSs, cabling and connectors. This requires power shutdown for the data centers, which poses added cost.

Sathe elaborates: "Energy costs represent the single largest component of operating expense, and a potential barrier to future expansion. Also, while the level of power quality and reliability offered by legacy distribution systems are acceptable for the majority of customers, voltage and frequency fluctuations, harmonics and short-term power outages can have costly and disruptive effects on data centers."

Distributed energy storage and generation, and microgrids offered by smart grids, are seen as complementary or alternative solutions to meet data center energy requirements, Sathe says. "Smart grid technology can enhance reliability and also offer local, distributed energy solutions to keep a data center running even during outages. Using a smart grid allows for the diversification of energy resources that make up a data center's onsite generation. They also allow data centers to improve the ROI on assets that traditionally have been sunk costs (that is, most data centers relied on natural gas generators for resiliency). Using energy storage allows for the increased utilization of the asset for other functions like ancillary services, disaster recovery, etc."

Eric LaFance, senior trade commissioner at Hydro-Québec says: "We are slowly replacing the human brain with machine decisions, which will require far more electricity." He suggests that blockchain,

the distributed ledger technology behind the Bitcoin currency and other systems, can replace human validation in finance, "because it is more efficient and more hacker-proof."

LaFrance says self-driving cars are another example of how we are replacing human decisions by machine intelligence because, in his opinion, it is more reliable. "This new way of doing things will require a lot more energy because we will need places to store massive amounts of information and servers to make these decisions for us."

What's more, LaFrance says power utilities and data centers can meet in strategic partnership on integrated micro-grids and demand response, but mainly during peak periods. "The utility will be able to save a lot of money when, during peak periods, it can rely on customer generators, batteries, or power production facilities, instead of buying expensive energy on the spot market. Hydro-Québec is already a leader in this field and will seek nearly 1,000MW from its industrial and commercial customers during the winter peaks." LaFrance says decentralized energy production, coupled with large battery packs, will ensure data centers are online when their utilities are down.

But smart energy does not come without its challenges, as Tanzella pointed out. "There are indeed current concerns or hot topics for data centers when it comes to smart grids. Data center managers have historically been tasked to maintain power and cooling for mission critical applications regardless of

the power bill. Dynamically migrating these loads across geographic regions is simple in concept, but can be difficult in practice."

Tanzella says the more risk-averse data center managers will likely be slow to adopt, but as the technology and software becomes more prevalent, data center managers will embrace the smart grid's functionality.

LaFrance says in the future we will see more small edge computing facilities responding to the always-growing proximity demand for IoT, mobile devices and other low-latency needs. He added: "Large hyperscale facilities should be located further away from dense areas to avoid the growing demand and pressure on utilities that are located near dense areas, which minimizes a utility's investment due to the lack of space and power. Large hyperscale operators should, and hopefully will, also base site selection decisions on criteria that mainly consider the use of renewable energy."

On the power network side of data centers, Sathe says the trend could be towards meeting energy requirements of data centers in a decentralized fashion.

"The data center of the future could be fully integrated with the grid in a way that enhances the local resiliency and reliability. For those data centers where reliability is the number one priority, being connected to the bulk grid while having their own distributed energy resource (DER) would be advised. They could also participate in transactional energy markets to stack up values of the smart grid assets they own."

As *DCD* has found in its own events, data centers will be fundamental to moves towards smart cities, as well as smart grids. ●
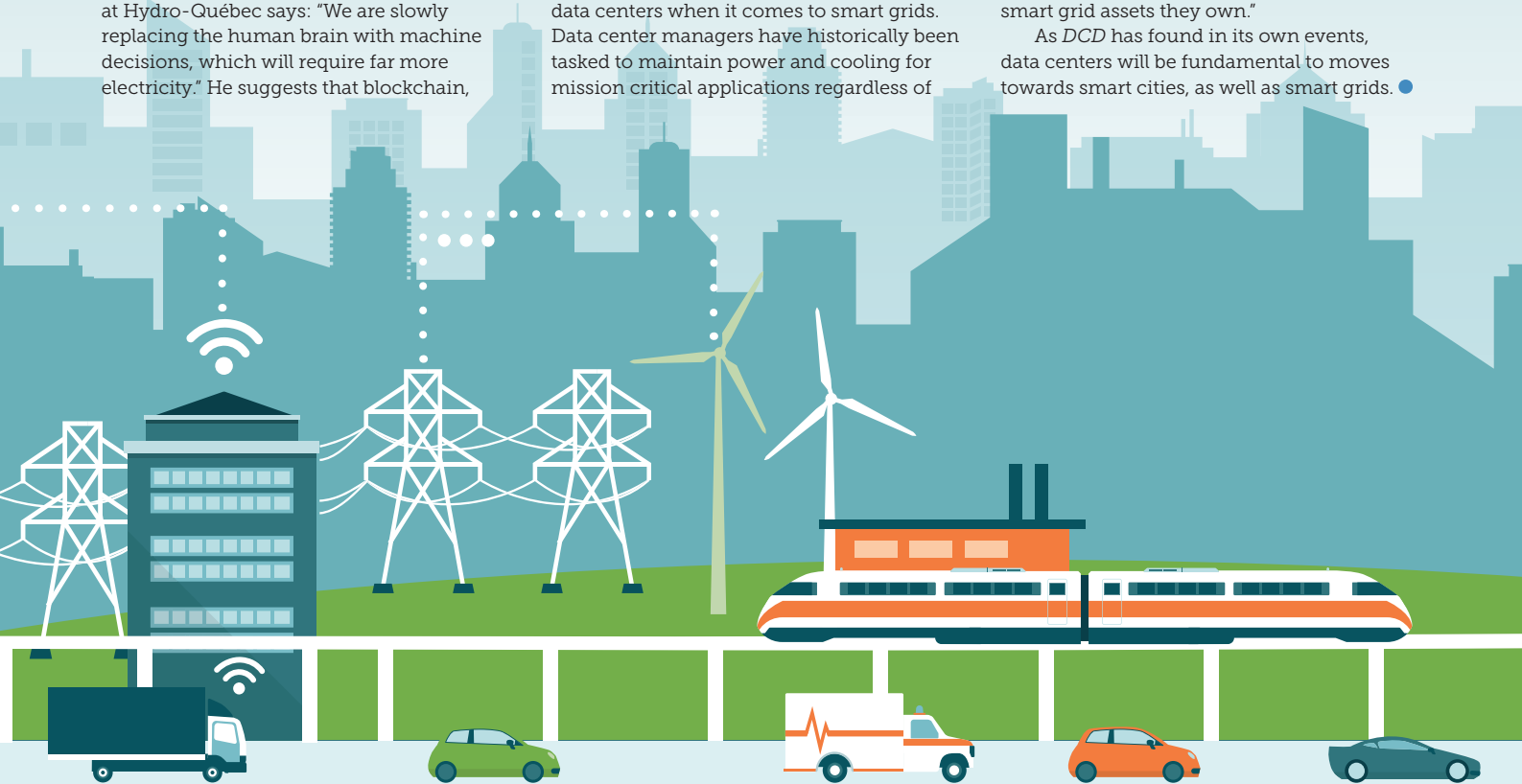
Trust Is Earned with Competence and Capability

Delta InfraSuite is the choice of the world's top smart cities

The power behind competitiveness

Delta InfraSuite – Datacenter Infrastructure Solutions empower smart cities to harness evolving mobile communications, big data and the IoT, to sense, analyze and integrate a spectrum of information, and master infrastructure challenges. The ultimate goal is to enhance government efficiency, and economic growth.

Delta InfraSuite, with its modular design, can be built fast and expanded easily. High integration and reliability ensure stable, uninterrupted system operation. Delta InfraSuite Solutions, the most robust support for smart cities, and the power behind competitiveness.

We have all the solutions.

Micro Datacenter

Modular Datacenter

Containerized Datacenter

Edge Datacenter

DCIM

Cooling

AC Power

Converged Power

More information and success stories about Delta InfraSuite

Tel: +65 6747 5155
Email: ups.singapore@deltaww.com
www.deltapowersolutions.com

Delta Energy Systems (Singapore) Pte Ltd
4 Kaki Bukit Avenue 1 #05-04
Singapore 417939

DELTA

Smarter. Greener. Together.

Containers unlock the cloud

Dan Robinson
Correspondent

They told us the cloud would allow workloads to move freely between in-house facilities and public services, but so far it hasn't happened. *Dan Robinson* believes that containers might finally get the cloud moving

Cloud computing is no longer viewed as some exotic new technology and has become an accepted part of an organization's IT toolkit for meeting business needs. At the same time, cloud is less mature than other IT sectors, and can prove to be more complex than unwary adopters may expect.

One of the promises of cloud computing has long been that users should be able to move workloads from their own data center to that of a cloud service provider and back again, if required, or even between public clouds. This might be because an application or service calls for more resources than the organization has available, or simply because it is more cost-effective to run it off-premises at that time.

Despite this, cloud adoption has broadly followed the pattern of private clouds being used to operate traditional enterprise workloads, while new-build applications and services are developed and deployed in a public cloud platform such as AWS. Even where traditional workloads have been farmed out to a service provider, this has typically been a case of colocation or of a hosted private cloud arrangement.

According to Ovum principal analyst Roy Illsley, this is because many organizations are still at an early stage of cloud adoption, and are just looking to get their foot on the first rung of the ladder.

"We are not really seeing companies have [workload mobility] as their first concern. What they are really looking to do is 'lift and shift' to the cloud, so they might have this three-tier app or database based on legacy hardware, and they want to know how to get that to the cloud as a first step. Only once they've done that, they start looking at how to transpose that workload, change that workload, and think about moving it somewhere else," he said. ▶
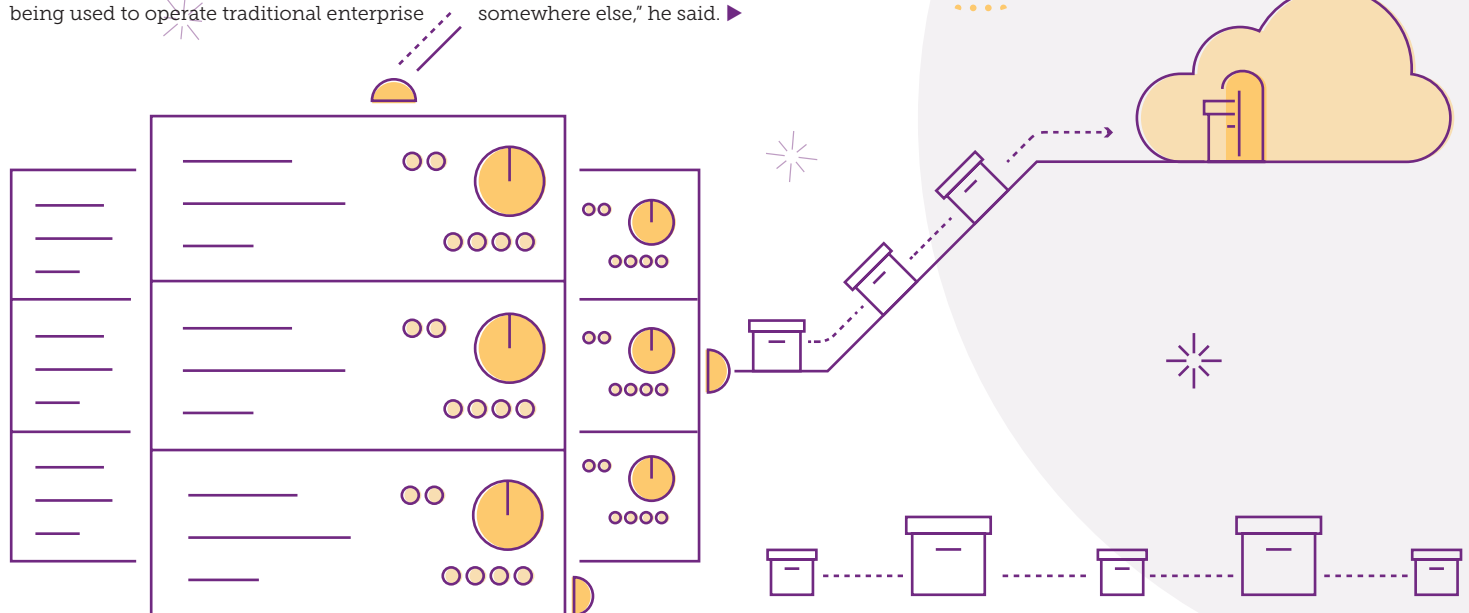
▶ There are good reasons why workload mobility has never really taken off. VMware with its vMotion feature has supported live migration of virtual machines from one host server to another for over a decade, but it is one thing to perform this inside a data center and quite another to do it from one data center to another.

Things get more complicated if you want to move a workload from your own cloud to another one that you do not control. It is not practical to move a virtual machine to a cloud based on a different platform, because of differences in the hypervisor, the APIs and the management tools used. Even if it is based on the same platform, you may not have the same level of management oversight and control as you do over workloads running on your own infrastructure.

Then there is the fact that enterprise workloads seldom operate in a vacuum; they depend on other resources to function, such as shared storage, an SQL database server or directory service. Unless such services are also replicated to the cloud, traffic between the workload and these resources will have to be routed to and fro across a WAN connection instead of across the data center itself.

Perhaps it is this reason that led VMware to change direction and ditch its vCloud Air service for Cloud Foundation, a platform that it describes as a self-contained software-defined data center (SDDC) that can be provisioned onto commonly used public clouds such as AWS and IBM SoftLayer.

However, the prospect of workload mobility could move nearer to reality thanks to containers. Most people are familiar with containers thanks to the efforts of Docker, but the ecosystem is made of various technologies and implementations.

What all container platforms share is that they enable a piece of software and its dependencies (such as code libraries) to be wrapped together into an isolated space – the container. Multiple containers can run on the same host, like virtual machines, but containers are more lightweight on resources, and a given server can operate more containers than virtual machines.

"The containers approach gives developers the opportunity of writing a new cloud-native app that, provided you've got support for containers – and if you're on Linux, you will have – you can then operate that container on a platform that runs in various other cloud environments," said Illsley.

Containers are a less mature technology than virtual machines, and the various platforms are thus still in the process of developing and perfecting associated components such as orchestration, monitoring, persistent storage support and lifecycle management tools. These are largely essential for operating container-based application frameworks at any kind of scale, as the Internet giants such as Google or Facebook do.

Docker has established its platform as the leading format for packaging and running containers, but there are several tools available for orchestration, such as the Kubernetes project, which originated at Google, or the Mesos project from the Apache Foundation, as well as Docker's Swarm tool.

Kubernetes is integrated into several enterprise platforms, such as VMware's Photon, Red Hat's OpenShift application platform and even Microsoft's Azure Container Service. Meanwhile, Mesos is used by many large web companies, such as Twitter, Airbnb, and eBay as it can scale to manage tens of thousands of nodes running containers.

The elephant in the room is that containers are tied to a particular operating system kernel. This is typically Linux, as container platforms such as Docker build on features in the Linux kernel, but Microsoft has recently begun adding support for containers into Windows Server and its Azure cloud service.

While a container image created on a Linux system should run on any other Linux system, you cannot take a container image created for Windows and run it on Linux, and vice versa. Having said that, Microsoft announced at DockerCon in April 2017 that it will support Linux container images in Hyper-V Containers, under which the container lives inside a small virtual machine.

Containers are unlikely to completely replace virtual machines, but for many workloads, especially scale-out distributed workloads, they are becoming the preferred method of deployment. This is not just because they are easier and quicker to provision, but also because they are easier to migrate from one system to another, finally beginning to fulfill more of the promise of the cloud. ●

What is Docker?

Docker is the company and platform that everyone thinks of when containers are mentioned, but it did not create the concept. Instead, it took some capabilities (known as LXC) built into the Linux kernel and developed them into a finished product that made it easy for users to create and operate a workload using one or more containers, making it popular with developers for use as a dev-and-test platform.

Part of Docker's success is that it not only created a set of APIs with which to control the Docker host, but also a package format for storing and distributing applications as container images. The Docker architecture also calls for a registry to act as an image library, and users can maintain their own registry or fetch ready-made images from a public one such as the Docker Hub.

A container under Docker's platform is therefore basically a wrapper for delivering and running an application or service along with any code libraries it depends on, all inside its own little bubble.

However, the underlying LXC capabilities it built on were more geared towards partitioning a Linux server into multiple separate user spaces, similar to Zones in Oracle's Solaris operating system or even older mainframe concepts.

DCD > Zettastructure | London

+ **OF DRIVERLESS CARS AND SERVERLESS DATA CENTERS**

Free passes for qualified end users

OPEN Compute Project

Europe Summit:
November 8

November 7-8 2017 // **Old Billingsgate Market**, London

Headline Sponsor
Life Is On | **Schneider Electric**

Principal Sponsors
Cyxtera EATON Powering Business Worldwide PANDUIT

Lead Sponsors
CORNING RITTAL VERTIV

For more information visit www.**DCD**.events

@DCDConverged #DCDZettastructure   Datacenter Dynamics   DCD Global Discussions

Calendar

DCD> Edge Roundtables
25 September, Dallas

- Data center leaders explore the impact of IoT, mobility and edge
- A Colo + Cloud pre-event special
- 200 places only

Open Compute Project workshops on servers, racks, power and more at Dallas, Toronto and London events

Events

> **México** | Mexico City
September 26-27 2017

> **Peru** | Lima
October 18 2017

> **Brasil** | São Paulo
October 30-31 2017

> **Chile** | Santiago
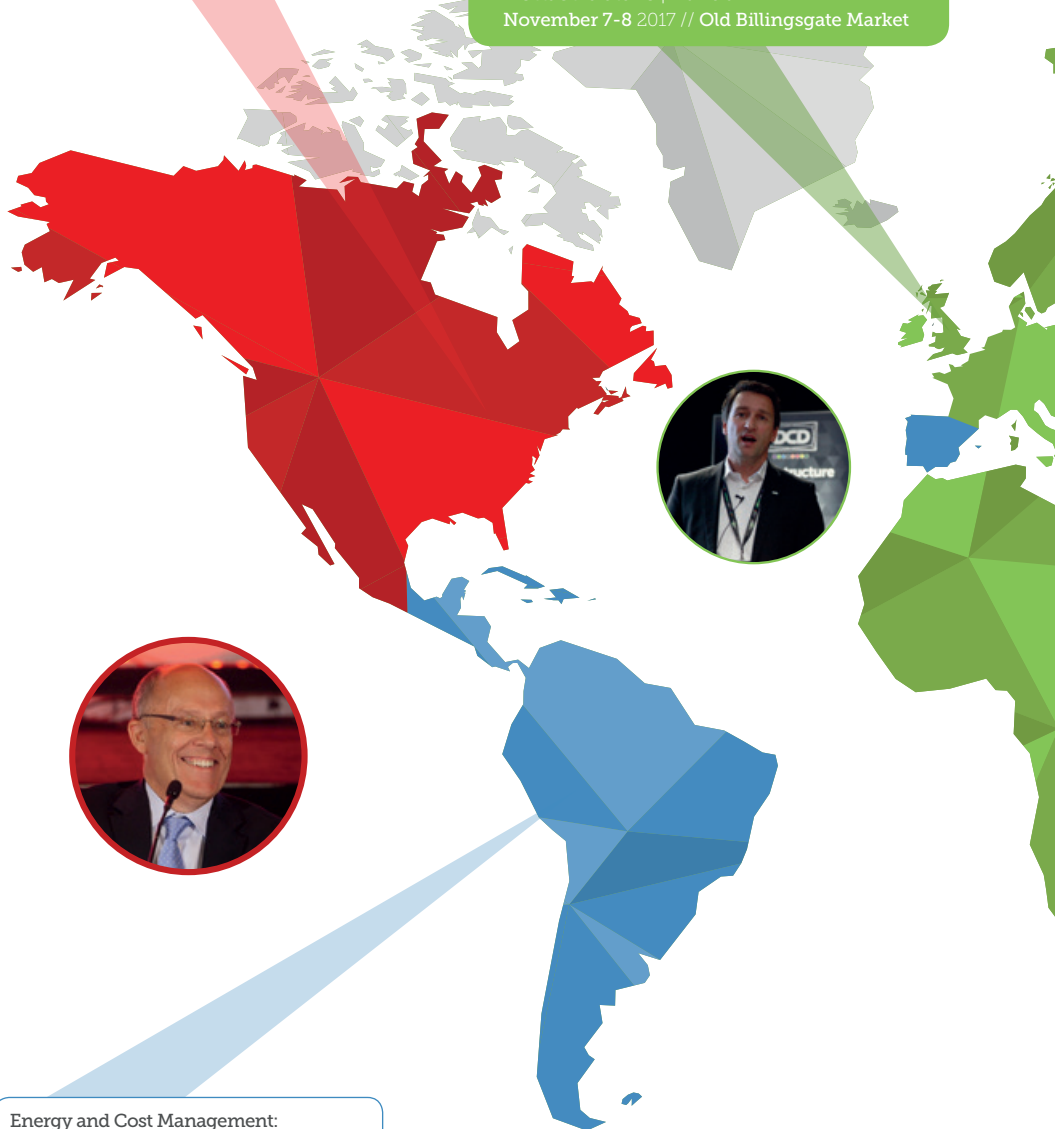November 15 2017

Training

Data Center Cooling Professional:
Lima, Peru
September 25-27 2017

Energy Efficiency Best Practice:
México, D.F., Mexico
September 28-29 2017, Hotel Novit

Data Center Cooling Professional:
Santiago de Chile, Chile
October 2-4 2017

Data Center Cooling Professional:
México, D.F., Mexico
October 9-11 2017, Hotel Novit

Events

> **Colo + Cloud** | Dallas
September 26 2017

> **Canada 4.0** | Toronto
December 14 2017

> **Zettastructure** | London
November 7-8 2017 // Old Billingsgate Market

Energy and Cost Management:
Buenos Aires, Argentina
October 16-18 2017

Critical Operations Professional: México, D.F., Mexico
October 23-25 2017

Data Center Design Awareness:
Lima, Peru
October 23-25 2017

Data Center Design Awareness:
Madrid, Spain
October 23-25 2017, Hotel Zenit Abeba

Data Center Design Awareness:
Barcelona, Spain
October 23-25 2017

Data Center Power Professional:
Porto Alegre, Brazil
October 25-27 2017

Energy Efficiency Best Practice:
Madrid, Spain
October 26-27 2017, Hotel Zenit Abeba

Critical Operations Professional: Lima, Peru
October 30-November 1 2017

Training

Data Center Design Awareness:
Stockholm, Sweden
September 11-13 2017

Energy Efficiency Best Practice:
Stockholm, Sweden
September 14-15 2017

Data Center Design Awareness:
London, United Kingdom
September 18-20 2017

Energy Efficiency Best Practice:
London, United Kingdom
September 21-22 2017

Data Center Power Professional:
London, United Kingdom
October 2-4 2017

Data Center Cooling Professional:
London, United Kingdom
October 9-11 2017

Events

> Zettastructure | Singapore
September 20-21 2017

> Hyperscale China | Beijing
November 2 2017

> Enterprise | Mumbai
November 9 2017

> Converged | Hong Kong
November 10 2017

Training

Data Center Design Awareness:
Osaka, Japan
August 7-9 2017

Data Center Design Awareness:
Beijing, China
August 29-30 2017

Data Center Design Awareness:
Sydney, Australia
August 30-1 September 2017

Data Center Design Awareness:
Hong Kong
September 4-5 2017

Critical Operations Professional:
Hong Kong
September 6-8 2017

Data Center Design Awareness:
Shanghai, China
September 8-9 2017

Data Center Design Awareness:
Singapore
September 11-13 2017

Data Center Design Awareness:
Tokyo, Japan
September 11-13 2017

Energy Efficiency Best Practice:
Singapore
September 14-15 2017

Critical Operations Professional:
Singapore
September 18-20 2017

Energy and Cost Management:
Singapore
September 25-27 2017

Events

> Türkiye | Istanbul
December 5 2017
Wyndham Grand Istanbul Levent

Data Center Week 17 Awards

SE Asia Data Center Week
15-21 September, Singapore

• DCD>Zettastructure conference & expo

• CxO and cyber security summits

• DCPRO training workshops

• Data Center Week Awards

• F1 Grand Prix dinner

DCD>Events // www.dcd.events
Sales: chris.hugall@datacenterdynamics.com
Marketing: jake.mcnulty@datacenterdynamics.com
Speaking: rebecca.davison@datacenterdynamics.com

DCD>Online // www.datacenterdynamics.com
Sales: yash.puwar@datacenterdyanmics.com
Marketing: ryan.hadden@datacenterdynamics.com

DCPRO>Training // www.dc-professional.com
Contact: liam.moore@datacenterdynamics.com

The unsung hero

People say data centers are not sexy, but they are wrong: in fact, this industry is filled to the brim with all kinds of smut. Data centers have been shaped by pornography as much as they have been shaped by developments in military technology or mobile communications.

According to conservative estimates, around 10 percent of all data transferred across the Internet features naked people. During 2016, the world's most popular adult website was responsible for streaming 99GB of data every second.

Some of the advancements in technology that were directly influenced by our penchant for X-rated material include video codecs and compression, online advertising and payments security.

They say YouTube democratized access to online video, but they are wrong: after all, no matter what content you post on YouTube, you will have to use the platform's proprietary tools. Porn gave rise to thousands of different video engines and hosting platforms, has kept an army of web developers in work and created a shadow economy worth billions. Adult websites were also some of the first to support Bitcoin at scale – as a way to avoid embarrassing items in customers' bank statements.

Carnal pleasures are a big business: the porn industry is worth around $97 billion, more money than Major League Baseball, the National Football League and the National Basketball Association combined.

Interestingly, most of the largest online porn websites – you might have heard about YouPorn or RedTube – are all owned by a Luxembourg-based company with a fairly neutral name – MindGeek (formerly Manwin).

MindGeek describes itself as "a leader in web design, IT, web development and SEO." It employs more than 1,000 people, and serves more than 115 million users every single day. It uploads 15TB of new content every 24 hours.

Initially, MindGeek didn't even produce porn – instead focusing on analytics and content delivery. But the company played its cards so well it eventually acquired some of America's largest studios, and became one of the top bandwidth consumers in the world.

Whatever you think about the subject, sex has always been a great motivator, driving technological progress forward even as it made people blush. So thank you, Internet pornography, for making our infrastructure better.

Max Smolaks
News Editor

> *During 2016, the most popular adult website streamed 99GB of data every second*

Structural Ceiling Grid by Tate

Versatile. Customized. Engineered.

- Support a wide selection of heavy data center accessories including cable trays, bus bars, and containment
- Continuous threaded slot allows for unlimited flexibility
- Decreases both installation time and cost
- Multiple extrusion types available – including M10, $3/8"$, $1/4"$, $1/4"$ hidden slot, and light structural support

Tate®
A Kingspan Group Company

Tate Is Your Single-Source Data Center Solution For:
Access Floors | Airflow Panels & Controls | Customized Containment | Structural Ceiling Grids