DCD
April/May 2018
datacenterdynamics.com

POWERING FINANCE

Money makes the world go round, but first it requires infrastructure

Europe gets smarter

We learn about the future of energy in Stockholm

Behind the Cell

The PlayStation chip that ran a supercomputer

AI and Automation

A 10-page deep dive into AI in the data center

CORNING

Made for whatever comes next

With data demands growing every day, how can you be sure the system you install today will be up to the job not just tomorrow, but for decades to come? **Rely on our data center solutions to make the future accessible.**

Our optical solutions have led the industry since our scientists created the first low-loss fiber in 1970. And with inventions like Corning® ClearCurve® bend-insensitive fiber, which transformed data center cabling, we continued pushing the limits of connectivity. From faster data center deployments and lower operating costs, to higher data speeds and future-ready, fully upgradable solutions, we help our customers solve the problems that matter most.

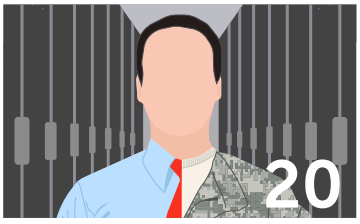
Leverage our comprehensive range of products and best-in-class services to address your data center challenges today and in the future. **Get all the details when you visit www.corning.com/edge8/dcd**

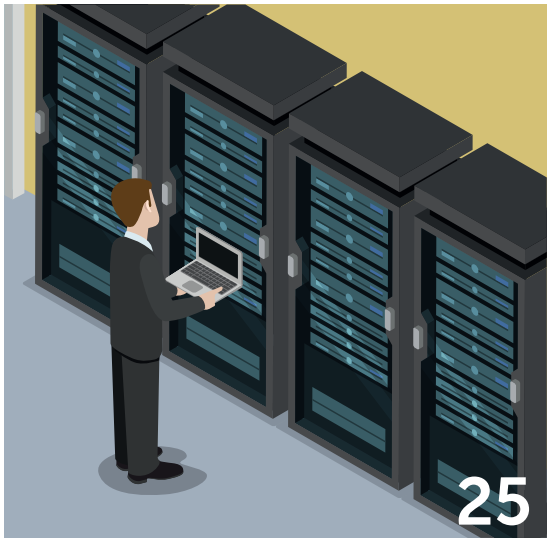
© 2018 Corning Optical Communications. LAN-2320-AEN / March 2018

Contents

April/May 2018

ISSN 2058-4946





CEO FOCUS

From the Editor

AI and finance: here comes Smart Money

Finance systems have a long partnership with digital infrastructure - well worth celebrating in this issue. Banks have made huge demands on facilities and networks, and in return tech is shaking up their business, big time.

Some of the leading names in financial data centers will be at DCD's Enterprise event in New York in May. On p14, we set out the reasons why they will be well worth listening to.

Banks are pushing technologies to their limits, which is producing interesting results in data center networks, and in databases (p18), among many other areas.

M&E kit can be hit by threats which are no longer a worry in IT

AI and automation are also simultaneously powered by data centers - and driving a revolution in how they are run. A special focus this issue rounds up several aspects of this fascinating partnership.

The move to software defined data centers is steady and inevitable... if only we understood what they are. Happily, Dan Robinson explains (p34).

Award winning ICICI Bank, from India, provides a great example of how automation can cut costs, while actually increasing flexibility (p37).

Finally, as a leader in chips for AI, Nvidia is well placed to show where things will go next: simulation (p40).

This high level approach can lead to a failure to spot dangers at a lower level - and the industry is in danger of completely failing to take into account the weakness of its physical infrastructure - generators and pumps.

Mechanical and electrical (M&E) equipment is highly likely to have connections to the Internet which are not managed, and can be attacked by simple threats, even ones which have been neutralized in the IT world.

Our article (p25) shows why this is a problem, and what you can do.

We have a skills shortage, and one answer could be military veterans according to Lee Kirby, who runs Salute, a company getting veterans into work in data centers (p20).

It might not be obvious, but think about it: Soldiers work as a team in a foreign place, handling crises, and liaising with local people.

There could be a lot of synergy with data centers.

Games consoles have often benefited from powerful silicon, whose cost is offset by the massive market these machines are designed to serve.

But what if a games console chip could be repurposed to power more general purpose systems?

That's the question answered by the Cell chip (p44), a processor which had a double life, powering Sony's PlayStation 3, and supercomputers.

It's also the starting point for what Nvidia did with its GPUs, now running world-changing AI and simulation.

bit.ly/DCDmagazine

5gms

The latency of the GTT Express submarine cable between New York and London, favored by financial institutions. The project was known as Hibernia Express, before being acquired by GTT for $590m, along with another four cables

Peter Judge
DCD Global Editor

Meet the team

Global Editor
Peter Judge
@Judgecorp

News Editor
Max Smolaks
@MaxSmolax

Reporter
Sebastian Moss
@SebMoss

Reporter
Tanwen Dawn-Hiscox
@Tanwendh

Editor LATAM
Virginia Toledo
@DCDNoticias

Assistant Editor LATAM
Celia Villarrubia
@DCDNoticias

SEA Correspondent
Paul Mah
@PaulMah

Brazil Correspondent
Tatiane Aquim
@DCDFocuspt

Head of Design
Chris Perrins

Designer
Mar Pérez

Designer
Ellie James

Head of Sales
Yash Puwar

Global Account Manager
Aiden Powell

Global Conference Director
Rebecca Davison

Conference Director
Giovanni Zappulo

Conference Manager, EMEA
Merima Dzanic

Head office

DatacenterDynamics
102–108 Clifton Street
London EC2A 4HW
+44 (0) 207 377 1907

© 2018 Data Centre Dynamics Limited All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, or be stored in any retrieval system of any nature, without prior written permission of Data Centre Dynamics Limited. Applications for written permission should be directed to the editorial team at editorial@datacenterdynamics.com. Any views or opinions expressed do not necessarily represent the views or opinions of Data Centre Dynamics Limited or its affiliates. Disclaimer of liability: Whilst every effort has been made to ensure the quality and accuracy of the information contained in this publication at the time of going to press, Data Centre Dynamics Limited and its affiliates assume no responsibility as to the accuracy or completeness of and, to the extent permitted by law, shall not be liable for any errors or omissions or any loss, damage or expense incurred by reliance on information or any statement contained in this publication. Advertisers are solely responsible for the content of the advertising material which they submit to us and for ensuring that the material complies with applicable laws. Data Centre Dynamics Limited and its affiliates are not responsible for any error, omission or material. Inclusion of any advertisement is not intended to endorse any views expressed, nor products or services offered, nor the organisations sponsoring the advertisement.

Dive deeper

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

Events | Intelligence | Debates | Training | Awards | CEEDA

Narada®

High Rate, Long Life Batteries for Critical Power - Data Centers

RELIABLE

When you need long term reliable backup power, Narada High Rate and High Operating Temperature batteries deliver. Engineered and manufactured in state of the art facilities, Narada provides solutions to meet your Critical Power needs.

ISO9001/14001/TL9000 Certified Quality
Backed by Industry Leading Warranties

Narada...Reliable Battery Solutions

www.mpinarada.com - email : ups@mpinarada.com - MPI Narada - Newton, MA
Tel: 800-982-4339

White space

A world connected: The biggest data center news stories of the last two months

Tencent launches data centers in the US, India and Hong Kong

Chinese Internet titan Tencent has opened another four cloud data centers - two in the US, one in India and one in Hong Kong.

In an effort to compete with the likes of Amazon Web Services (AWS), Microsoft Azure and Alibaba Cloud, the company has targeted key data center metro markets: Silicon Valley, Virginia, Mumbai and Hong Kong, ramping its overall data center count up to 42.

Tencent already offers cloud services hosted in the US, having made its entrance in Silicon Valley last year.

A new facility in Virginia places the company at the heart of the world's most dense data center and interconnection market.

As for the new site in Hong Kong, where Tencent brought its first data center online last year, the company says it will exclusively serve the financial sector.

The facility in Mumbai is Tencent's first in India, where there is plenty of room for growth, as the country's economy becomes increasingly reliant on digital services.

Chinese rival Alibaba Cloud launched its first local data center in Mumbai in January.

The next step for Tencent is to roll out new facilities in Thailand, Russia and Japan.

Earlier this year, the company opened a new data center in Chongqing, southwest China, to serve local manufacturing operations. China is in the midst of 'Made in China 2025,' a master plan to overhaul its factories and production processes for the modern era, building smart factories that will produce huge quantities of data.

The new Chongqing site joins an existing $1 billion data center there, which opened in 2015.

It is targeted at the public sector, as well as local businesses. In 2015, Tencent also announced that it had signed a partnership deal with the Chongqing municipal government for a smart city project.

The plan is to allow people to use the city's administrative services through Tencent's WeChat - users can apply for passports, marriage registrations, birth registrations and more in the multi-purpose social media mobile app.

bit.ly/Thirtycents

News in brief

IBM unveils 19" mainframes
To be deployed side-by-side with traditional servers.

Northstar Commercial Partners to build two data centers in Loudoun County
$100m project for an unnamed Fortune 20 company, thought to be AWS.

EdgeCore breaks ground in Mesa, expects Nevada site approval
The recently-formed infrastructure consortium has started on a $150 million, 180,000 square foot data center in Arizona, and is seeking approval for another in Tahoe Reno.

Five data centers gain CEEDA certification
Facilities in Sweden, Germany, Gibraltar and Scotland improve or update their efficiency credentials.

Tape shipments in 2017 hit new capacity records
HPE, IBM and Quantum revealed that a record 108,457 petabytes of tape capacity (compressed) was shipped last year.

Waymo self-driving trucks deliver Google Atlanta data center gear
Early tests help both Alphabet divisions.

Google breaks ground on delayed Alabama data center

Construction on the $600m facility set on the site of the defunct Widows Creek coal-fired power station was originally planned to start in early 2016, then in 2017, but the project was repeatedly delayed for unknown reasons. "They're a very secretive group," Jackson County Schools Supervisor A.J. Buckner said. "They are very, very, very, very secretive."

OCP Summit: Microsoft speeds up cloud SSDs with Project Denali

Microsoft has proposed a new interface for solid state storage, which should allow easier technology upgrades and faster integration in cloud data centers.

Solid state drives (SSDs) are normally packaged as monolithic devices which can be dropped in as direct replacements for traditional rotating hard drives. An alternative interface called Open-Channel exposes the internals of the SSD to the computer operating system, but using Open Channel means adopting proprietary interfaces specific to the device in question.

In Project Denali, Microsoft is proposing that the interface could be split up, so that there is one standardized part dealing with plain storage, and another dealing with issues like bad blocks which are specific to the particular product. The scheme would allow data centers to optimize their storage for the specific applications they run and the hardware they use, and integrate new SSDs rapidly.

The Project Denali prototype was developed with storage silicon specialist CNEX Labs, and will be offered through the Open Compute Project (OCP) as a standard when it is complete. It was demonstrated at the OCP Summit in San Jose this March by Azure hardware infrastructure manager Kushagra Vaid and senior software engineer Laura Caulfield.

"Storage paradigms have performed well on-premises, but they haven't resulted in innovation for increasing performance and cost efficiencies needed for cloud-based models," Vaid said in a blog post.

"Fundamentally, Project Denali standardizes the SSD firmware interfaces by disaggregating the functionality for software-defined data layout and media management. With Project Denali, customers can achieve greater levels of performance, while leveraging the cost-reduction economics that come at cloud scale."

Project Denali defines the roles played by the SSD and the host in a standard interface, so issues specific to the device such as media management remain on the device, while the host gets on with the business of sending and receiving data, maintaining an address map and performing garbage collection.

This means that SSD suppliers can build simpler products for data centers, and deliver them quicker, while also minimizing the disruption further up the stack when they are introduced to the facility. The project is especially useful for installations where the hosts have FPGAs or microcontrollers.

bit.ly/ScalingDenali

DataBank opens major data center in Plano, Texas; will build another in Kansas City

American cloud and colocation provider DataBank has opened a 145,000 square foot (13,470 sq m) data center in Plano, Texas.

The facility, codenamed DFW3, was acquired from Stream Data Centers in the middle of 2017. Back then, it was an empty powered shell. In the past few months, DataBank has built out the office space and two data halls (out of a potential six), and the site is now ready for customers, with up to 9MW of power capacity across two independent feeds.

It also plans to build a data center in Kansas City with 25,000 square feet (2,322 sq m) of raised floor space and 3MW of power, opening this year.

In 2016, Dallas-based DataBank was acquired by Digital Bridge, an investment fund focused on telecommunications infrastructure. The new owners proceeded to make multiple acquisitions and integrate them into the business, including the entire C7 Data Centers portfolio and two facilities from 365 Data Centers.

bit.ly/DataBanksGrandPlanos

Vox Box

Jason Waxman
Corporate VP, GM, Data Center Solutions
Intel

What's next for the data center?
There's a huge shift underway, and one of the things I plan to talk about at DCD>Enterprise in New York is the move to hyperscale in general.

My belief is that more and more of the workloads that will be run in the data center are applications or infrastructure that will be managed at scale, and that fundamentally means a shift in terms of how we think about what the data center is, does and how to architect that.

How is it that we can go take a fresh look at the data center and design it for really where the workloads and the environment is heading?

I do believe that we're going to see a shift, and those people that operate data centers on their own will be doing it at a greater scale than they've done in the past. If you look at just the size - the amount of data that's being collected, the amount of processing required for applications, such as artificial intelligence - to be able to do that effectively, it really does take a different approach, and the deployment of not just public cloud, but private cloud, is substantially on the rise.

Those are the transformations that are taking place.

Jason Waxman will be talking at DCD>Enterprise in New York on May 1-2. For more, follow the link below:

bit.ly/DCDEnterprise

Facebook settles BladeRoom lawsuit, Emerson to face trial

Facebook has agreed to settle BladeRoom's $365m trade secrets lawsuit, four days into the trial. Emerson Electric Co. will still have to convince a jury of its innocence.

According to *Law360*, Emerson requested a mistrial following Facebook's decision, on the grounds that jurors risked being influenced in their decision by the Cambridge Analytica scandal involving the social media giant - but the request was denied.

On April 11, Facebook founder and CEO Mark Zuckerberg testified before a joint hearing of the Senate commerce and judiciary committees, following revelations that the company may have allowed political strategist Cambridge Analytica to access the accounts of 87 million of its users.

On April 12, Zuckerberg testified before the House energy and commerce committee.

Emerson attorney Rudy Telscher told the court the day before Zuckerberg's appearance that, failing a mistrial, the case should be postponed, and a new jury appointed in several months' time.

But with the difficulty of enforcing jury neutrality, and in the absence of a guarantee that Facebook would be any less mired in scandal months from now, US District Judge in charge of the case, Edward J. Davila, insisted that the trial must proceed as planned.

bit.ly/BladesOut

Swedish ski slope gets longer thanks to AWS

Swedish winter sports enthusiasts will soon get more bang for their buck, thanks to a project to raise the altitude of the the Vilsta ski resort by ten meters - and cloud giant Amazon Web Services is providing most of the rock.

The resort in Eskilstuna, 120km west of Stockholm, wants longer ski runs, and the artificial geological process is being made possible by AWS, which happens to have an excess of 100,000 tons of rock, to be excavated during the building of its new data center in Stockholm.

The Tunafors Slalom ski club of Eskilstuna asked to increase the size of its ski slopes, to create the perfect race run and give skiers more options. It received the all clear from the local municipality last year, on the condition that it use rock mass free of toxic pollutants, which could be damaging to the mountain's ecosystem.

bit.ly/ADataMountain

Two petaflops in a box: Nvidia DGX-2 packs 16 GPUs on two boards

Featuring twelve of the company's newly announced interconnect fabric, NVSwitch (*see below*), the DGX-2 is comprised of 16 V100 GPUs split across two server boards, along with two Intel Xeon Platinum CPUs, 1.5 terabytes of system memory and 30TB of NVMe SSDs. The system weighs 350lb, stretches 10U and consumes 10kW.

Jim McHugh, Nvidia VP and GM, said to *DCD*: "We've had a lot of customers explain to us that they're doing bigger and bigger clusters, and one of the most amazing things about DGX-2 is that it's incredibly flexible."

Ian Buck, GM of Nvidia's data center business, added: "We have been optimizing our entire stack to scale up AI. Take the FairSeq neural network used for translating the Internet - on our DGX-1, it was measured to take about 15 days to train. We've taken that to just one and a half days.

"Just to put that in perspective - it would take about 300 Skylake servers to deliver that same level of performance."

In a keynote speech at GTC 2018, CEO Jensen Huang said: "This is 10x faster than DGX-1, it took hundreds of millions of dollars of engineering."

The system will retail for $399,000, starting in the third quarter. "This is what an engineer finds beautiful, you guys, this is sexy," Huang said.

bit.ly/HuangryForMore

Nvidia announces new interconnect fabric, NVSwitch

Nvidia has announced a new GPU interconnect fabric, NVSwitch, which allows 16 GPUs to simultaneously communicate, with multiple switches able to be connected.

The on node switch architecture has a total of 18 ports, which provide 50GB/s of bandwidth per port, giving an aggregate of 900GB/s of switching bandwidth.

Featuring some two billion transistors, the networking ASIC has been in development for the past two years. Currently it is only available in Nvidia's DGX-2 servers.

bit.ly/NVSwitchItUp

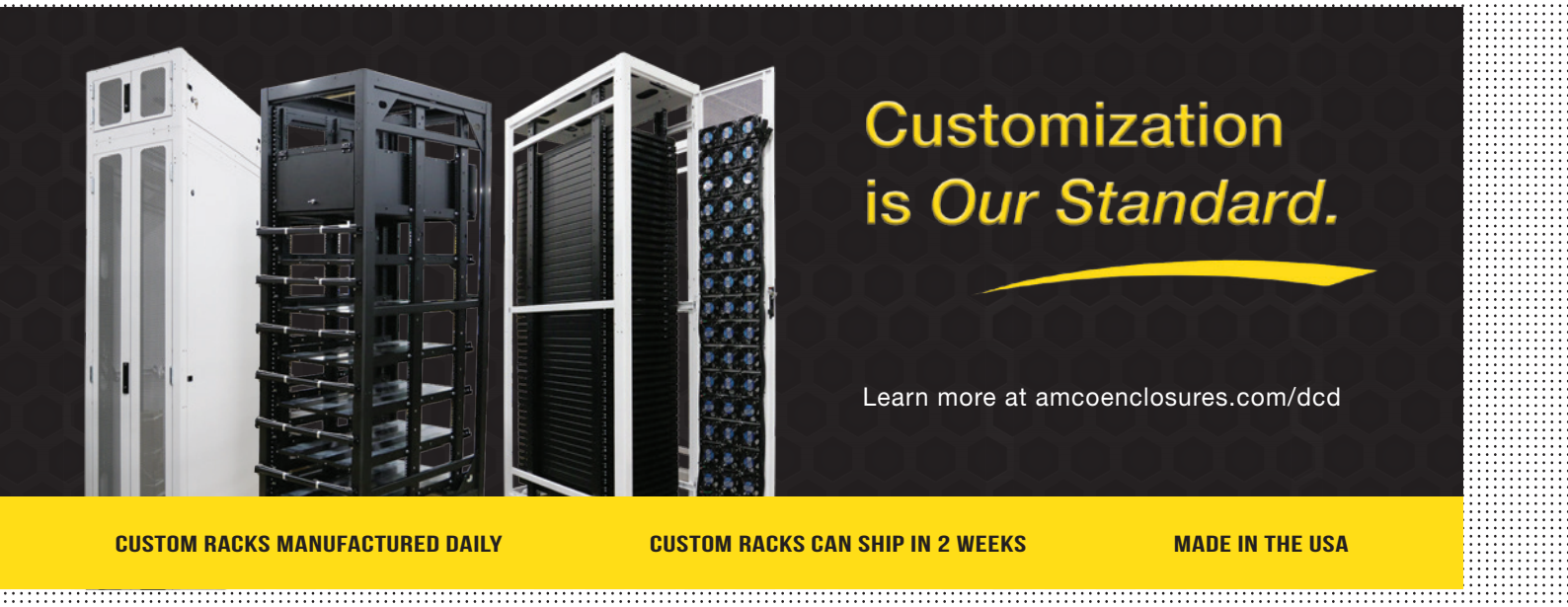
For more Nvidia, see **p40**

Vapor IO and Packet to build 5G-ready edge cloud

Edge data center company Vapor IO and bare metal cloud provider Packet plan to install '5G capable' infrastructure at edge locations, with the initiative launching in Chicago later this year as part of Project Volutus.

Operators would have to deploy the rest of the infrastructure for 5G - an emerging technology that will take years to introduce.

bit.ly/5GattheEdge



US plans to spend up to $1.8bn on two new exascale supercomputers

The US Department of Energy (DOE) has sent out a Request for Proposals (RFP) worth up to $1.8 billion for "at least two" new exascale supercomputers.

Part of the CORAL program, the new systems will follow the first US exascale supercomputer, Aurora, which comes online in 2021. The RFP also covers the possibility of Aurora upgrades and a potential successor system in 2022-2023.

The RFP asks for one system to be developed for Oak Ridge National Laboratory in Tennessee, and another for Lawrence Livermore National Laboratory in California.

Each must have a unique system design, delivered by 2021-2023 for $400-$600m. A third potential system at Argonne National Laboratory, Illinois, could have a similar or wholly different design from the first two.

The RFP states that the supercomputers can't require more than 40MW of power, with 20-30MW preferred. This power limit includes computing, storage, cooling and auxiliary equipment.

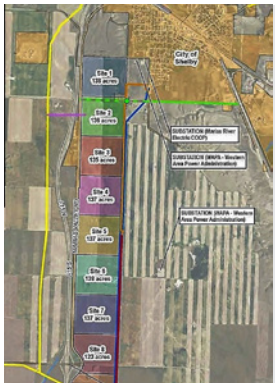
"These new systems represent the next generation in supercomputing and will be critical tools both for our nation's scientists and for US industry," Secretary Rick Perry said.

"They will help ensure America's continued leadership in the vital area of high performance computing, which is an essential element of our national security, prosperity, and competitiveness as a nation."

Perry's comments come as the US finds itself under intense competition from China, which has ramped up HPC investments, semiconductor investments and artificial intelligence investments, aiming to dominate the computing sector in the years ahead.

bit.ly/PostExascaleScalesUp

City of Shelby, Montana, earmarks 10 sites for data center development



The City of Shelby in Montana has identified 10 parcels of land that could serve as suitable locations for data centers.

All sites have a reliable power supply and access to connectivity from Tier I providers. The city highlights that any sufficiently large project would also benefit from property tax breaks approved by the State of Montana last year.

In 2017, Montana passed a package of tax incentives aimed at potential data center developers. These included a new 'class seventeen' property tax classification for qualified data centers - in this instance, facilities that comprise at least 300,000 square feet and have invested at least $150 million in land and infrastructure.

Class seventeen properties are taxed at 0.9 percent of its market value, for a period of 15 years from the start of construction.

Besides low property tax and the absence of sales tax, Montana also offers access to cheap, locally produced wind power and a cool, windy climate that should help reduce data center cooling costs.

The new initiative, launched by the Port of Northern Montana and supported by the City of Shelby, has identified and zoned 10 sites for potential data center projects in the state, at around 140 acres each.

These sites are located on the outskirts of the city and in close proximity to two electric sub-stations, served by two different utility providers. They offer a choice of network connectivity providers, water supply, access to sewers and a proposed connection to a natural gas pipeline, suggesting potential for on-site power generation.

bit.ly/LandoftheShiningDataCenters

Pentagon cuts $950m cloud computing contract with AWS reseller to $65m

The US Department of Defense has scaled back a planned cloud contract with REAN Cloud LLC, previously worth up to $950 million, to just $65m.

The news comes just a month after the major deal with the Amazon Web Services reseller was announced.

The contract was part of the DOD's overall plan to move to an enterprise commercial cloud platform, with the agreement made under an Other Transaction Authority by the Defense Innovation Unit Experimental (DIUx).

The deal was originally intended to support USTRANSCOM and other DOD organizations' migration to the cloud, but now it has been "narrowly tailored" to cover only US Transportation Command.

"After reviewing the production agreement recently awarded to REAN Cloud LLC, the Department has determined that the agreement should be more narrowly tailored to the original scope of the prototype agreement, which was limited to United States Transportation Command applications," Pentagon spokesman Col. Robert Manning said, as reported by *FCW*.

The decision comes after Oracle protested the contract with the Government Accountability Office.

After finding out that its contract would be cut by nearly $900m, REAN "announced its disappointment."

Sekhar Puli, REAN's managing partner, added: "REAN Cloud has not been made aware of the basis for the DOD's recently stated intention to reduce the contract ceiling to $65 million. However, it is clear that many DOD agencies wish to procure these services.

"Based on the threat of legal action and protest by the old guard, the only winners in this delay are those large companies that stand to lose money if the DOD proceeds with innovation. In the meantime, the cost of maintaining antiquated government infrastructure has not subsided."

bit.ly/ThatsGotToSting

Peter's JEDI factoid

The Pentagon's Joint Enterprise Defense Infrastructure (JEDI) cloud contract could be worth $10bn, and is set to be awarded to one company. With AWS the presumptive leader, Oracle, Microsoft and IBM want the deal changed

Uptime is everything—
So don't fall for the imitators.

Trust 30 years of innovation and reliability.

Originally released nearly 30 years ago, Starline Track Busway was the first busway of its kind and has been refining and expanding its offering ever since. The system was designed to be maintenance-free; avoiding bolted connections that require routine torqueing. In addition, Track Busway's patented u-shaped copper busbar design creates constant tension and ensures the most reliable connection to power in the industry—meaning continuous uptime for your operation. For more information visit **StarlinePower.com**

STARLINE®
TRACK BUSWAY

DCD Calendar

Stay up to date with the latest from DCD - as the global hub for all things data center related, we have everything from the latest news, to events, awards, training and research

Events

DCD>Enterprise | New York
Marriott Marquis, New York
May 1-2 2018

Plenary Panel: The A-Z Disrupted Data Center

10:00am - May 2 2018

This wide-ranging panel will cover any area of potential disruption to the data center - M&E, IT, operational, networking, business ops, colo/cloud demand trends, edge, interconnects, etc. It's wide open... a conversation about change, today, and in the future. If you want to help steer the conversation, ask a question!

bit.ly/DCDEnterprise

Kelly LeValley Hunt
BlockApps

Nancy Novak
Compass Datacenters

Rhonda Ascierto
451 Research

DCD>Africa | Johannesburg
Hilton Sandton, Johannesburg
Jul 24 2018

DCD returns to Africa

After a lapse of three years, DCD is excited to return to Johannesburg for an expansive debate on the future of digital infrastructure across the continent. The African data center industry is poised for growth as demand for digital services increases. Africa's leading data center businesses and thought leaders will be contributing to this year's conference program which will be held at the Sandton Convention Centre, located at the heart of South Africa's silicon capital.

bit.ly/DCDAfrica

Funke Opeke
MainOne

Dan Kwach
East Africa Data Centre

Ben Roberts
Liquid Telecom

The Spanish language magazine goes fully digital

Now our Spain-and-LATAM publication is available in an interactive format designed for tablets, laptops and phones. ¡Buen trabajo!

bit.ly/DCDspain

DCD>España | Madrid
La Nave, Madrid
May 23 2018

DCD>Shanghai
Shanghai Tower Conference Center
Jun 7 2018

DCD>Awards | Best In India

July 18 2018 // **The Sheraton Grand Bangalore Hotel**

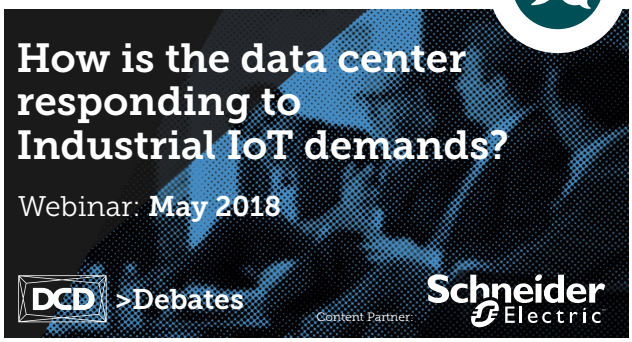
The DCD Awards are delighted to be celebrating over a decade of doing what we love – sharing the stories of innovation and cutting edge design and operation which are the hallmarks of our vitally important industry. We are proud to continue showcasing the stories which epitomize the pioneering spirit & innovative thinking behind the projects, people and teams making the industry great.

This year we are also featuring the best entries from India in a 'Best in India' ceremony on July 18 as part of the DCD>Hyperscale event in Bangalore.

The ceremony will award those entries made by Indian companies into the DCD Awards program across six categories, so there is an opportunity for everyone.

dcdasia.awardsplatform.com

DCD>Debates - A new webinar format bringing the dynamism of our live panel discussions to a global audience



For companies to store, process and act upon the vast quantities of data that a factory can create every day, they need extensive digital infrastructure. And, for optimum results, that infrastructure should be located right there, near the factory itself.

Register here: bit.ly/SchneiderIndustrialIoT

Mark Bartlett
Arup

Mark Howell
Ford Motor Company

Victor Avelar
Schneider Electric

DCD>Debate
How does Edge computing transform enterprise data center strategy?

Watch On Demand

For more information visit bit.ly/SchneiderEdgeComputing

Keep up-to-date

Don't miss a play in the data center game. Subscribe to DCD's magazine in print and online, and you can have us with you wherever you go! DCD's features will explore in-depth all the top issues in the digital infrastructure universe.

Subscriptions
datacenterdynamics.com/magazine

To email one of our team
firstname.surname@datacenterdynamics.com

Find us online
datacenterdynamics.com | dcd.events | dcdawards.global | dcpro.training

DatacenterDynamics DCDnews DatacenterDynamics

DCPRO

From industry-certified courses to customized technology training, including in-house development, DCPRO offers a complete solution for the data center industry with an integrated support infrastructure that extends across the globe, led by highly qualified, vendor-certified instructors in a classroom environment as well as online.

New Course Launched

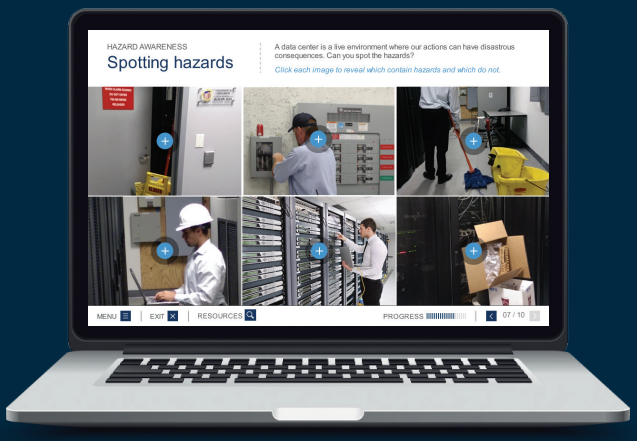
Energy Professional – 3 days
This course was designed to teach the key considerations for near and long term target measures for data center energy efficiency enablement, alongside operation and design strategies that help maximize program impacts and increase cost-effectiveness, ensuring that capital investments delivers results.

2018 Course Calendar

Data Center Design Awareness – **Shenzhen | May 10-11**
Data Center Design Awareness – **London | May 14-16**
Data Center Technician – **Singapore | May 16-18**
Energy Professional – **London | May 21-23**
Data Center Design Awareness – **Buenos Aires | May 28-30**
Data Center Design Awareness – **Madrid | June 4-6**

For more course dates visit: www.**dcpro.training**

Take our **free** Data Center Health & Safety course today!



Safety should not have a price. Take our **1 hour online** health & safety course for **free** today!

www.**dcpro.training/dc-health-safety**

POWERING FINANCE

Financial services have a close and interesting relationship with the digital infrastructure that supports them, reports *Peter Judge*

Peter Judge
Global Editor

Ever since data centers emerged as a distinct class of infrastructure, financial organizations have been their most demanding client, pushing technology vendors to develop new hardware and software. It's fair to say that financial services have had a major hand in shaping data centers - and in every development inside their doors.

When colocation data centers appeared, the major financial centers such as London, Frankfurt and New York were among the earliest locations. As the industry developed, more processing power was required to carry out the calculations powering increasingly complicated instruments and services.

Financial organizations were also among the first to make the reliability demands that have led to the expanding science of backup and disaster recovery. A financial services operation can lose millions for every minute of downtime, so a high level of investment in products that can prevent that is easily justified.

Performance was obviously important - but network speeds became particularly vital. Data had to be transmitted in vast quantities, and with low latency. It became very important to close deals quickly before the markets changed. A few milliseconds of difference in transmission times between one financial center and another can be crucial - a fact which has led to more than one new investment in submarine fiber optic cables across the Atlantic and other oceans.

And the reliability of those communications also had to be guaranteed. Cash machines (ATMs) rely on connections to core infrastructure to deliver money and advice to customers. Many of the same services have migrated to websites, which must be 100 percent reliable - and proofed against tampering.

Financial services also led the adoption of new methods of storage and processing. In a departure from their conservative public image, banks ended up spending their time and resources pursuing new technologies and turning them into a competitive advantage. Open source databases and Big Data tools like Hadoop found some of their earliest adopters amongst bankers and brokers. A system which can scan millions of transactions can potentially spot a fraud before a customer's account is compromised.

Finance has also enabled one of the most extreme developments in current data center practice: when confidence in the world's financial systems hit a new low after the financial crash of 2008, the

DCD>Enterprise | New York
New York Marriott Marquis
May 1-2 2018

Bill Beck of Credit Suisse, Brian Green of Visa, Charles Hoop of Aon, John K. Lee and Christian Pastrana of Citi and Dan Gaffney of BNY Mellon are among the scheduled speakers representing the finance industry at DCD's enterprise event in New York in May. Expect insights on how financial services are exploiting digital infrastructure.

bit.ly/DCDEnterprise

concept of Bitcoin emerged.

Supported by the blockchain distributed ledger system, in which all nodes store all the data, Bitcoin appeared to be using technology to sidestep the tainted systems of global finance. Miraculously, it appeared to provide something which could simultaneously serve as a secure and private medium of exchange for transactions, and a fast-appreciating commodity for speculators.

Looking into the future, it will be hard to square these two roles for cryptocurrency. The distributed trust mechanism has inherent difficulties, based on its fundamental requirements. Because all transactions are stored everywhere, scalability is limited, and the Bitcoin blockchain can currently only handle around seven transactions per second. In contrast, payment company Visa handles 24,000 transactions per second.

11-21 → WALL ST

Despite these drawbacks, Bitcoin has become the fastest growing application in data centers worldwide. New facilities are being built to run cryptocurrency mining, and they have to do it with manic energy and ever-increasing efficiency. New cooling techniques and faster GPUs or ASICs are thrust to the fore in the hope that these shoestring sites can steal a march on their competitors.

The irony is that this competition is self-defeating. The reward for the effort is ever-diminishing returns, and an unrestrained growth in energy consumption that could conceivably consume all the world's cheap electricity for no genuine benefit.

Against this background, we are seeing a somewhat predictable development. When technological advances seem to be at risk of becoming uncontrollable, legislation tends to emerge to rein in the excess.

While Bitcoin was conceived as a means to deliver financial transactions outside the traditional operators, the mainstream banking system has been responding to the aftermath of the financial crisis with a move towards tighter regulations designed to prevent a relapse. Service providers will have to raise their game, meet the new requirements, and deliver a new level of accountability.

Sarbanes-Oxley, MiFID II and the Basel rules have made specific demands on digital financial services. Not only must there be complete clarity about how transactions happen, data on infrastructure must also be collected, and reported to the authorities when required.

Those who pay the piper have always called the tune - even more so for those who write the tune and define the means by which the piper will be paid.

Finance has driven the development of data centers, and there is no reason to believe this will change any time soon.

INNOVATION MEETS LEGACY TECH

Fintech is almost synonymous with digital innovation, but it may face its toughest challenge yet as enterprises look to embrace open banking, algorithmic trading, cryptocurrencies, machine learning, artificial intelligence and more, while remaining compliant with a raft of industry regulations

Andrew Fray / UK MD, Interxion

The UK has been a pioneer, with telephone banking established way back in 1984, as a dedicated service from Girobank. Online banking is actually older: the Bank of Scotland launched Homelink for Nottingham Building Society customers back in 1983 – years before similar ideas were tested elsewhere.

Despite this, banks and financial services find it hard to keep up with demand for constant online access to account information, and ensuring real-time payments via contactless cards, phones and even watches.

The pace of change could get faster, as innovation leadership moves from the well-funded R&D departments of traditional firms to bright young entrepreneurs and start-ups.

Traditional financial services companies are playing catch-up, as they combine digital services, including hubs that will centralize their financial functions, 'sinks' where they can store archived data, and digital superhighways that can carry huge volumes of information.

Harmonizing these elements across multiple platforms requires agility, flexibility and the right sort of infrastructure. Thirty percent of enterprises currently use a private cloud in their own data center, but fewer than one in five expect to be doing this by 2020, according to research carried out for Interxion. The proportion using public cloud is expected to increase from under a quarter (now) to nearly one-third (in 2020).

Innovation in financial services must be balanced with the 'ball and chain' of legacy systems. Financial services organizations have to maintain a lot of legacy IT, which makes migration more complex. All of this offers plenty of challenges for data center operators.

DCD>Debate
Readying for Banking & Financial Services 4.0

Watch On Demand

"It doesn't much matter which part of the financial services one touches, whether it's retail, commercial, wholesale banking, wealth management, insurance, or brokerage and trading. Each of these has its version of digital disruption," says DCD's Bruce Taylor in a DCD>Debate webinar, which is available on demand.

"Banking and insurance have been about four or five years ahead of any other industry in understanding what it means to do business electronically," says Steve Madden, senior director for vertical market development at Equinix - and that means a move to the cloud. Many financial services organizations "want to exit facilities altogether and don't want to be in the data center business anymore."

Errol Roberts, distinguished engineer at Cisco, points out that finance applications will still have strong needs for assurance and security in whatever infrastructure they are run on.

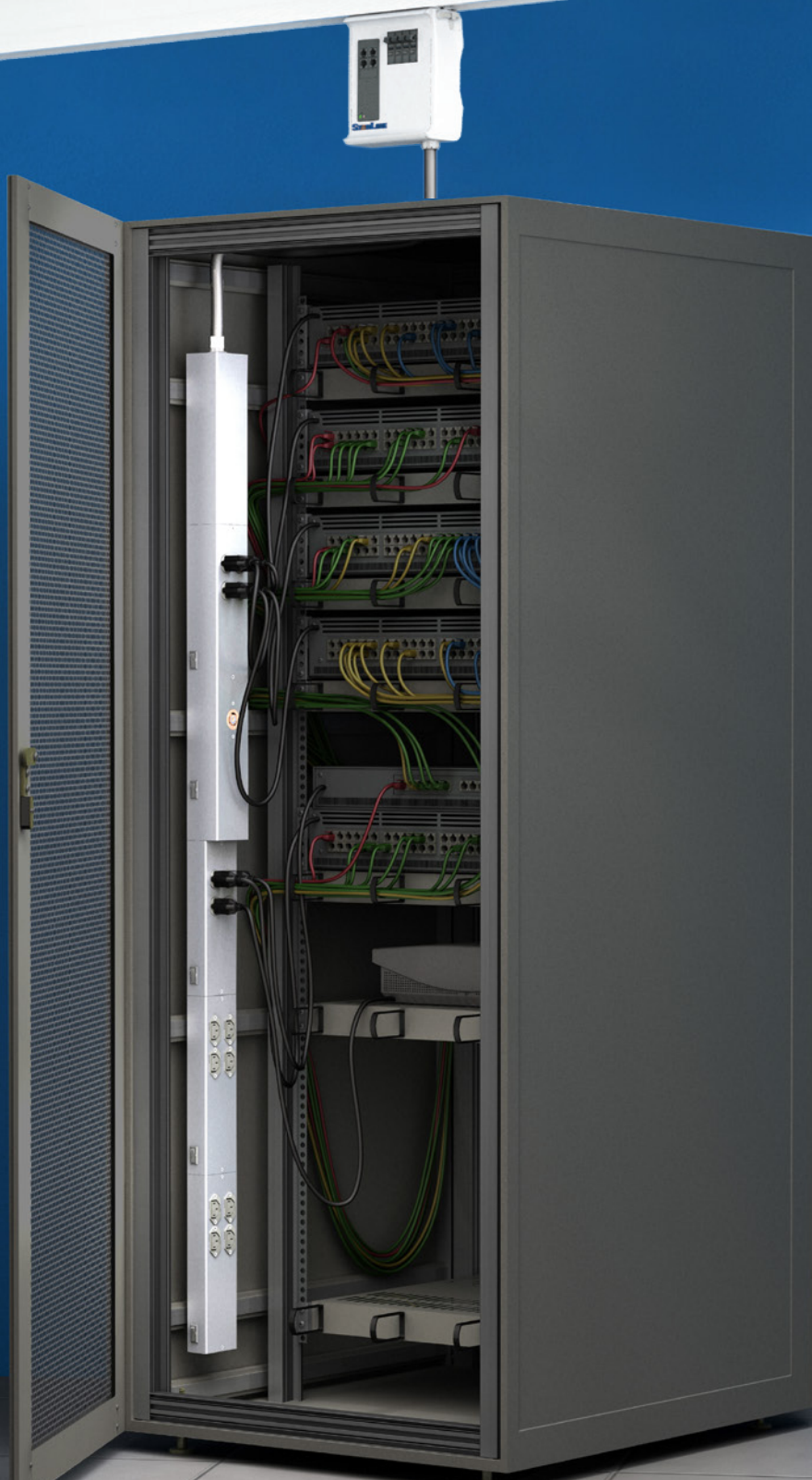
This means the networks used by finance firms have to be fast and reliable. They all have fiber to their switches, says Tom Walsh, Panduit's lead for data centers, and it could go further: "I could take fiber all the way down to the computer side."

Watch the full debate: bit.ly/PanduitFinance

Flexible Power:
From overhead to in-rack

STARLINE

When data center power requirements call for higher density, multiple types of receptacles, and future flexibility, Starline goes above and beyond. Though it may look like other power distribution systems, Starline's innovative design provides users with the flexibility to choose and use different types of receptacles on the Cabinet Busway system. To see how Starline is changing power distribution in data centers, visit **StarlinePower.com**



Dropping the legacy database

Tanwen Dawn-Hiscox
Reporter

Today's companies are using new technologies to store and manage their data, says *Tanwen Dawn-Hiscox*

For the financial sector, as well as for telecommunications, advertising and retail industries, and even government bodies, making sense of information collected from various sources and translating it into tangible results is of the utmost priority.

What can be done at the database level is but one element of the process, but its impact is quite significant, as the ways in which an organization can store and manage data vary greatly.

Firstly, it is important to differentiate between types of database management systems (DBMS).

Born in the 1970s, Structured Query Language (SQL) databases, by definition, are queried in standardized code. Most traditional DBMS are based on SQL, and follow the relational model, meaning they are very highly structured in their querying specifications.

What's more, until now, most databases have been centralized, stored and maintained in a single location, typically a desktop server or a mainframe.

One of the most popular examples of an SQL database is MySQL, an open-source relational DBMS implementation for the web, developed in 1995 by Michael 'Monty' Widenius and widely adopted among Internet giants.

More recently, we have seen the emergence of a NoSQL model (which is misleadingly short for 'Not only SQL'). These databases are non-relational - meaning they are structured in a way that supports many-to-many relations between datasets, and can be scaled without the need for hardware upgrades.

At the same time, more and more companies have shifted to a distributed approach, which allows them to run replicated databases in a number of geographically separate locations.

The modern database must be scalable, decentralized, and allow for the aggregation of disparate data types. Thus, relational databases have lost some of their appeal, and NoSQL databases have emerged from the fringes of the industry to gain a significant share of the market.

In fact, the NoSQL market is expected to be worth $4.2bn by 2020, according to Allied Market Research. In the past decade, an influx of these databases have replaced legacy engines, responding to a new operational model. Popular examples include MongoDB, Redis, HBase, MarkLogic and Cassandra.

More data is collected today than ever before, and as a consequence, speed is another characteristic on which DBMS providers are expected to improve. For this reason, another departure from tradition has been a rapid increase in the use of in-memory databases.

Whereas before, on-disk infrastructure required interaction between separate data warehouses (Teradata, Informatica, Redshift) and operational databases (Hadoop, Cloudera), with in-memory, it is possible to consolidate analytics, machine learning, AI and operations in a single place.

By relying exclusively on system memory, one eliminates the need for disk I/O to query or update data, and additional cache copies are no longer necessary.

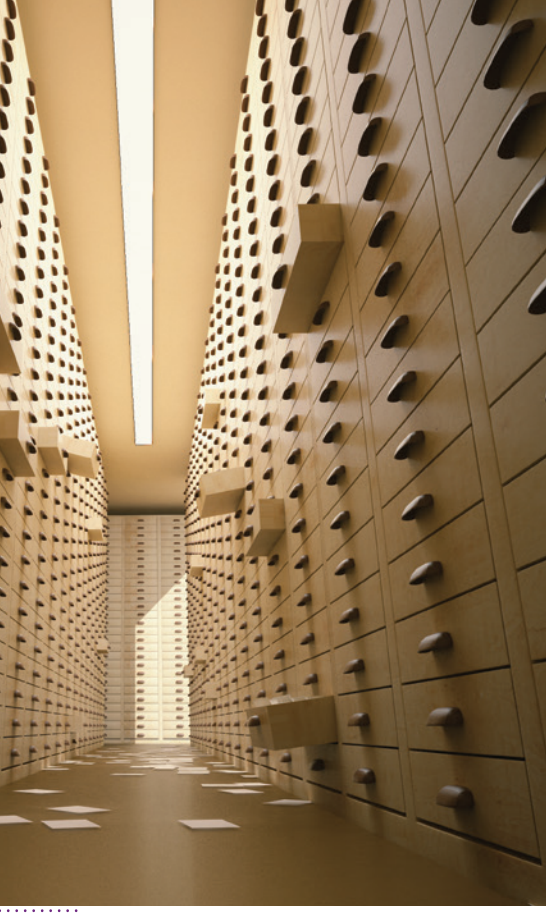
The main downside of in-memory databases is that they were historically limited by the high cost of RAM. As the price of memory has decreased, in-memory databases have grown in popularity, and the expectation of real-time analytics has gained prevalence.

Being between 10 and 100 times faster than disk-based DBMS (a difference of milliseconds), in-memory systems may bring solutions to problems posed by IoT data analysis, autonomous vehicles, modern retail platforms, transaction processing and fraud detection.

And indeed, according to research firm

"It does a lot for very agile exploration of very large datasets"

Todd Mostak, MapD CEO and co-founder



Markets and Markets, the in-memory market could reach a value of $13.23bn in 2018. Major companies traditionally working on disk have developed their own in-memory systems, including Oracle, Microsoft and SAP.

On a visit to the Silicon Valley in February as part of the IT Press Tour, *DCD* met several DBMS companies, all of which had varying approaches to database technology.

Founded in 2009, Aerospike's flash-optimized, NoSQL database connects directly to the application database, using hybrid memory architecture (combining DRAM and SSDs) and promising to deliver low latency for applications at scale. The open source key-value store system was designed to provide a solution for latency-critical applications, both in real-time transaction processing and real-time analytics - which ordinarily rely on different databases, for different reasons.

One of the main differences between the two, co-founder and CTO Brian Bulkowski explained, is that transactions require a balance of writes and reads, whereas analytics tends to be very read-intensive.

Srini Srinivasan, Aerospike's co-founder and chief development officer, added that businesses tend to use different technologies for systems of record, dealing with telco customer data, reservations and financial risk, where consistency and accuracy are key, and systems of engagement, such as real-time bidding, cyber security and fraud detection, in which performance and availability are top priorities.

By effectively replacing the cache layer, and thanks to the combined use of SSDs and RAM, Aerospike says it can ensure low latency and high throughput. What's more, its clustering capabilities can meet the needs of distributed applications too.

To improve the disk I/O, Aerospike places indexes in DRAM rather than on SSDs; it optimizes the network by colocating user and index data on the same node to avoid excessive hops between them, automates routing of client requests and balances workloads automatically. Finally, it offers multi-threading and parallel processing across multiple SSDs to facilitate scaling.

For Bulkowski, the company's mission is to bring the proprietary technologies used by Internet giants - which he is adamant are similar to Aerospike's product - to real-time, big data applications.

Emphasis on speed and scale are two priorities shared by Foster-city based GridGain, whose in-memory software can either sit between the application and the database (or data lake), or be deployed as an in-memory SQL database.

GridGain was built to support transactional and analytical applications, and is based on the Apache Ignite open source project, which the company created (and to which it is still the main contributor). Like Aerospike, it eliminates the need to separate operations from analytics and machine learning capabilities, bridging the gap between a data warehouse and an operational database.

As a consequence, Abe Kleinfeld, the company's president and CEO explained, the system is free from data integration and feedback loop interactions. All capabilities are held in a unified, in-memory datastore, and the platform has a unified API - meaning it can aggregate data in different formats and structures.

The in-memory system has proved popular with financial services, and the company doubled its revenue in the past year, boasting clients such as ING, Société Générale, Apple, Huawei and Microsoft.

GridGain's biggest project by a stretch was a 2,150 Teraflop cluster totaling 56,000 CPUs and 1536TB of memory, commissioned by Sberbank of Russia.

Both Aerospike and GridGain focus on a "scaling out" approach, increasing the node count to increase the database performance.

MapD chose to improve the performance of each individual node instead, with an open-source in-memory database which runs on anything from one to 16 GPUs per server.

It was a desire to visualize and interact with data in real-time that gave birth to MapD, and two of its products which go hand in hand: MapD Core, the GPU-accelerated SQL database, and MapD Immerse, a web-based visual analytics platform that sits atop the SQL engine and can render billions of data records in a single compressed image. The engine requires no pre-indexing or pre-aggregation: everything is done in real-time, over potentially billions of rows.

Restricted to just structured data, the company's CEO and co-founder, Todd Mostak said, MapD Core isn't as versatile as some of its NoSQL contemporaries - but instead of a replacement for data warehouses, it markets itself as "a sidecar."

"We're like a hot cache on your store record. That could be pulling data out of a data lake or a Hadoop system, that could be pulling data out of a traditional data warehouse like Teradata."

"You can hit third-party business intelligence tools, we have a nice DBI-compliant Python connector, and of course a lot of our customers leverage MapD Immerse. Even though it's not as feature complete as Tableau, it does a lot for very agile exploration of very large datasets through the SQL and rendering capabilities."

The system finds its uses in model generation for fraud, risk and anomaly detection, geo-analytics and cyber security, real-time fleet management and incentive-based insurance.

In 2014, MapD won Nvidia's $100K Early Stage Challenge, a prize awarded annually to the best start-up that utilizes GPUs, and the chip manufacturer has since participated in all three of the company's funding rounds.

MapD recently launched a software-as-a-service (SaaS) offering, which runs on Nvidia GPUs across data centers "from the leading cloud infrastructure providers," with automated provisioning, optimization, support and upgrades.

$4.2bn
Predicted value of NoSQL market by 2020
(Allied Market Research)

Together with partners Continuum Analytics and H2O.ai, MapD recently founded the GPU Open Analytics Initiative (GOAI) to integrate its platform with other GPU-based projects.

The consortium's end-game is to enable the combined use of GPU-based analytics tools, starting with the framework for GPU-native data formats and APIs.

"The idea when we all got together was that we're all running on the GPU, so let's have a zero-copy framework such that we can pass data seamlessly through these different processes without the overhead of marshaling again or going through the CPU," Mostak explained.

Military service

There's a shortage of skilled and reliable data center staff, and Salute says one answer is to employ veterans. *Peter Judge* speaks to Lee Kirby

Peter Judge
Global Editor

Data center operators have trouble recruiting staff - and the problem will get more critical in the future as an aging population comes up for retirement. It seems there aren't enough people with the skills and abilities needed to build and maintain facilities.

One answer to this could come from an unexpected talent pool. Salute Inc believes that military veterans could fill the gap.

The organization was established by Lee Kirby, who is also president of the Uptime Institute. It was inspired by his own experiences in the civilian data center sector - and in the Army. For many years, Kirby was a reservist, while also working in the technology industry. After 9/11, he got called up and served in Iraq, Afghanistan and Haiti.

In 2012, Kirby retired as a Colonel, with more than ten years of active duty.

When he retired, he noticed something: "I saw a lot of colleagues who served, lost their jobs and livelihood because of it. When I retired, young troops who had served four years, were suffering a greater than 20 percent unemployment rate - because of their service."

Military personnel have skills and attributes that ought to readily transfer to the business world, and be particularly useful in engineering projects in the data center industry, he reasoned. They know how to operate in teams, often far from home, relying on equipment they maintain themselves.

There was one obstacle: "People won't hire you without experience." And some of those veterans are at a low ebb: "Twelve percent of veterans are homeless, when we hire them," Kirby pointed out.

Despite the practical difficulties, Kirby believed he had potential employees, and access to training that could get people into a job: "I can take an infantryman or a military police officer, and in two years they'll be a data center technician."

Five years later, Salute has got over 200 veterans into projects. It's not a charity: it wins contracts and hires staff to fulfill them. If it operated as a charity, its employees wouldn't get competitive experience, and the organization would have to focus on getting grants.

"From the client's point of view, they have a project like installing a hot aisle or a cold aisle," he explained. "We compete and win it, then put our guys on the project."

Salute can win those deals, he said, because it's upfront that its staff may have less experience, but they have proven reliability. To win business, Salute has to win trust.

About half of Salute's business is project-based, and some involves providing visible people on-site - perhaps multifunctional staff combining network operations center (NOC) duties with security guard work: "A lot of sites don't need an engineer but someone reliable."

Unlike comparable firms, Salute does not focus on staff retention. It aims to be a stepping stone for its recruits, and welcomes their departure as they move up the employment ladder. "Salute is not a career path," Kirby said. "We expect 25 to 30 percent attrition each year. We are putting people into jobs."

Some Salute alumni remain on board as the leadership team, while others out in the industry remain friends, he said.

Service men and women with tours of duty under their belts are well suited to remote projects, where they operate semi-autonomously: "Take a typical infantry man. He maintains a weapon, a vehicle and a communications system. He must operate with local people, work around dangerous equipment, and deal with the unexpected.

"There is no civilian training or experience that can equal that," Kirby said. "They adapt and train quickly, and communication and coordination are bred into them."

He admits there may be downsides. Veterans may have post-traumatic stress disorder (PTSD), but they aren't alone in that: "Statistically there's a lot of issues in the 'normal' human race - and veterans have programs for PTSD. I was injured and I know what it is to get that trauma - but also what it is to get support."

The proportion of veterans with problematic PTSD may be lower than the proportion in the whole population, and Salute has become adept at handling those with difficulties, while keeping the organization healthy, he said.

Sign up

Salute Incorporated finds projects where highly trained American (and other nations') military veterans can work for data center owners and operators of all sizes. Its personnel have skills earnt in rigorous training programs, which can be used in data center projects that require disciplined labor - including data center cleaning, containment installation, battery maintenance, cage build-outs and general logistics support.

www.saluteinc.com

Salute operates in the US, and has expanded into the UK, Ireland and The Netherlands. Beyond that, it could operate in an advisory role in other countries: traveling in China, Kirby was asked for advice about a local chapter of Salute, but said projects in places like Russia or China should be run by local veterans: "I think every country could do this, and Salute would do best where we've been allies."

Despite the usual media cliché, 15 percent of veterans are female, so Salute can help address a missing source of talent which is under-represented in the industry. It could also cast its net wider, and work with another untapped resource: military partners.

"We have started to encourage military spouses," he said, and Salute's clients are interested. "Military spouses have project management experience. It's very hard to teach someone to multitask like a spouse has to when their partner is deployed."

Salute doesn't plan for meteoric growth. "We don't plan to do an IPO, we are trying to do a sustainable business in the industry," Kirby said. But the initiative may bring in capital partners in the future.

Kirby's main goal is to provide a steady opportunity to people who have already paid their dues. "They did their service, and have come back. All they need is a chance. They are not looking for handouts."

The data center paradox
DCD>Enterprise New York
May 1-2 2018

Rumors of the demise of the enterprise data center to be challenged at 16th DCD summit in New York. Lee Kirby will join a panel discussing: *The data center paradox – More important than ever, because of cloud adoption.*

bit.ly/DCDEnterprise

DCD New York
See us at stand **72**

DATABAR®
INTELLIGENT DATA HALL BUSWAY SYSTEM

THE MOST VERSATILE,
RAPIDLY INSTALLED,
SAFEST HOT-CONNECTABLE BUSWAY
SYSTEM ON THE MARKET.

ANORD MARDIX
> Explore our Data Center Solutions at **www.anordmardix.com**

Peter Judge
Global Editor

Europe gets Energy Smart

Thanks to its demographics, its energy grids, and its politics, Europe is getting ahead on energy efficiency, says *Peter Judge*

Europe has a large population spread across multiple countries - there are 28 in the European Union and around 50 in the continent as a whole - with a huge and growing demand for mobile and digital services. The nations have a wide array of energy sources, including nuclear and renewables, and a climate which goes from the temperate heat of Naples to the Arctic cold of Northern Sweden.

All of that makes digital infrastructure complex to build and deliver, but there's a long history of collaboration in the region on electricity grids, communications networks and on fundamental industrial and scientific research.

So when Europe gets together to consider how to make its infrastructure more Energy Smart, it's well worth listening to the ideas that emerge.

That's what happened at DCD's Energy Smart event in Stockholm earlier this year.

The overall message was integration. Data centers, like any other part of the world's infrastructure, can't be treated in isolation. Their energy use must be linked to the production of renewable power, and their reliability features can contribute to the reliability of the grid as a whole.

Waste heat is a useful resource, the conference heard. It costs money to cool the equipment in a data center, but much of this can be offset with revenue if the heat can be recovered and used.

This has been a theory for some time, but it has rarely been put into practice. At Stockholm Data Parks, run by Stockholm Exergi, money is changing hands for waste heat, according to the head of the Parks, Erik Rylander: "Cooling of the data center as a service is offered for free in exchange for the

excess heat when the data center load reaches above 10MW. In a second option, the cooling is managed by the customer and Stockholm Exergi purchases excess heat at a price reflecting its alternative heat production cost."

This works out at around $200,000 per year, per MW of heat, Rylander said.

Meanwhile, the money that is put into creating uninterruptible power supplies (UPS) is a necessary investment, which can actually produce a revenue stream, if those UPS systems are shared with the grid, according to Jussi Vihersalo, business development manager at Eaton.

A 500MW data center will have around 83MWh of energy stored in batteries, as a reserve for power failures. However, the energy could also be used more flexibly to help balance the grid in three ways. At a busy time, the data center can be run from the UPS to reduce demand; at times of higher demand, some of their energy can be fed into the local grid, and at times when too much renewable energy is being produced, the batteries can soak up some of the excess.

"Data centers can balance the power reserve, by using batteries instead of fossil fuels, and help to compensate the volatility renewable energy creates," said Vihersalo. "The next step in optimizing data center assets is the better utilization of the electrical infrastructure."

And this has a positive value for the data center, he said: 1MW can be worth €50,000 (US$61,500) to the data center - and it's paid for having it available, not for the actual amount used.

Despite the promise of money, data center operators have been skeptical, or downright hostile to this idea, fearing that other calls on their UPS will increase the risk of failure, but Vihersalo said this need not be the case. Battery sharing algorithms will include an allocation so critical backup is always allocated, and the facility is not disconnected from the mains when the UPS is powering it - so effectively the two sources switch, and the mains becomes the backup.

Other ideas are focusing on using electricity better within the data center. Replacing AC power with DC has been proposed for a long time: "There are no technical reasons to use AC power inside a data center," said Stefan Lidström, product manager, DC power at Comsys.

All the components use DC power, and AC power is responsible for up to 80 percent of the power dissipation, up to 95 percent of the weight, and up to 95 percent of the size of power distribution equipment, claimed Lidström.

Even more power distribution losses can be removed by cogeneration, which produces power on site. This can be taken to the extreme by using fuel cells to generate power within the rack - an approach being tried by automotive manufacturers and IT firms.

A lot of work on this approach is being carried out in the US, by Microsoft and the National Renewable Energy Labs (NREL), but Daimler and Mercedes-Benz are also involved. A lot of fuel cell work has come from the automotive industry.

A European project called Everywh2ere is working on a fuel-cell based replacement for diesel gensets used for backup at data centers, which will use cells from PowerCell, a Swedish Volvo spin off. Meanwhile, a French company called PowiDian has produced something similar - a 30kVA portable fuel-cell power source.

All this will need more intelligent management, which is often referred to as software-defined power (SDP). "Like the broader software-defined data center (SDDC), SDP is about creating a layer of abstraction that makes it easier to continuously match resources with changing needs," said Andy Lawrence, research director at the Uptime Institute.

SDP can make good use of microgrids, where local electricity sources and loads are synchronized with the traditional grid, in an extension of the UPS-sharing ideas mentioned above.

But the level of integration could work best at a slightly larger scale - the city, according to a number of Energy Smart presenters. "Companies and cities will lead the way," said Anders Egelrud, CEO of Stockholm Exergi, pointing out that conurbations have the flexibility and size that can make a real difference to energy use and policy.

This was underlined even further by Karin Wanngård, mayor of the City of Stockholm: "I want Stockholm to become an international beacon for how to run a truly smart and sustainable city," she said. "The Paris Climate Agreements cannot be achieved without actions by cities."

Despite a growing population, Stockholm has reduced the emissions caused by heating (see Box: *How to do CHP*). The rest of Europe, and the rest of the world could learn a lot from this.

How to do CHP

Stockholm has one of Europe's largest district heating systems, and is actively encouraging data centers to provide their excess heat alongside the output of combined heat and power (CHP) plants which burn biofuel and waste.

The district heating and cooling is provided by Stockholm Exergi, a company co-owned by Fortum and the City of Stockholm. A large part of the heat comes from Värtaverket, one of the largest a CHP plants in Europe, which burns woodchip waste from the timber industry.

The fuel is brought in by ship from the forests, and burnt in a fluidized bed system which regulates the temperature to minimize carbon monoxide (low temperature) or nitrogen oxides (high temperature). Ash is returned by ship to fertilize the forest.

The city plans to be carbon-neutral by 2040, and has an energy system which balances multiple inputs and demands, all of which depend on other factors including the weather. Renewables now make up the majority of the energy input, and sources like wind can sometimes produce so much that the effective energy price is zero.

Meanwhile, demand for hot water from the heating system is so high that CHP plants must continue to operate, sometimes using their electricity output to run heat pumps and create more heat.

When *DCD* visited, Värtaverket was producing 130MW of electricity, 400MW of heat and some 22MW of cooling.

From 2013 to 2016 the construction of a new biofueled CHP plant was underway for the new Värtaverket - one of the largest bio CHP plants in Europe.

Energy Smart Focus Day
DCD>Webscale San Francisco
June 25 2018

Microgrids and energy storage will be a major theme in the Energy Smart Focus Day which kicks off DCD's San Francisco event, opened by Robert Weisenmiller of the California Energy Commission. Sessions will also cover renewable energy and sharing backup, as well as taking the lid off new cooling technologies.

bit.ly/DCDwebscale

DCPRO Mission Critical Training

1. Mission Critical Engineering 4h
2. Reliability & Resiliency 4h
3. Electrical Systems Maintenance 4h
4. Fundamentals of Power Quality 4h

Mission Critical Awareness Online Certificate – enrol now

We've updated and redesigned the industry's most flexible series of online training courses for the mission critical sector

Visit **www.dcpro.training/mca**
Use code '**MCA33**' at checkout for **33% reduction**

DCD > **Enterprise** | New York

Come find us on the **1-2 of May** at

info@dc-professional.com

www.dcpro.training

The walls you thought you could trust

Security experts have worked hard to fix the flaws in IT systems, while industrial control systems have been left wide open, warns *Vlad-Gabriel Anghel*

Vlad-Gabriel Anghel
DCPRO Technical Consultant

Cyber security - the words are on everyone's lips as the public starts to realize the problems that come with having large amounts of personal data kept in one place. But there is more to security than protecting the private details of human beings.

The past few years have seen more and more security breaches at companies that act as trusted repositories for personal data. In 2017, a breach at Equifax leaked the Social Security numbers, addresses and dates of birth of millions of US citizens. These details are unlikely to change, and in certain instances they can be used for identity theft - for example, fraudulently obtaining credit cards and opening bank accounts.

This has happened often enough for businesses to understand the importance of cyber security, but another battleground is emerging. It also needs tight control and smart policies to keep intruders at bay. This arena is in danger of being overlooked, because instead of data, attackers are targeting machines which are normally taken for granted.

In the past, there were two separate worlds of business technology: Information Technology (IT) and Operational Technology (OT). Most people are familiar with IT - servers, switches, storage hardware and everything related to digital information - but OT is a lesser known realm.

OT is grounded in the material world; it is the technology which manages physical processes ▶

▶ through monitoring and control of devices such as pumps, valves and motors.

In the last decade, these two different worlds have begun to merge. The rapid expansion of the Internet of Things (IoT) means OT environments are now becoming digitally connected and remotely controlled. Pumps and air conditioning units are instrumented with sensors and access points and this means they are now open to new attack vectors - some of them, the same as those which have been plaguing the IT world for years.

The data center industry, which underlies so much of our daily lives, has been pioneering the integration of IT and OT - and it is therefore one of the first sectors where these new risks are being uncovered.

Most data center facilities will control power usage, airflow and cooling through a plant which is managed using the SCADA (Supervisory Control And Data Acquisition) system architecture. SCADA uses computers, networks and graphical user interfaces for process management, whilst making use of peripherals like PLCs (programmable logic controllers) and PID (proportional-integral-derivative) controllers to interface with the machinery.

In practice, large SCADA systems function in a very similar way to distributed control systems, and use several interfaces throughout the facility. These systems can work on multiple sites through large scale processes, over short and long distances.

SCADA is currently the most widespread industrial control system - and this in turn raises concerns that it could become the target of cyber terrorism.

Industrial control systems have been moving from proprietary technologies to more open and standardized solutions, like those used in IT. They have also been increasingly connected to corporate networks and the Internet. These two factors leave such systems vulnerable to attack vectors typically found within computer network security.

The risk is that data centers may have focused too much attention on their IT, while OT has become a blind spot.

"While many organizations have developed stringent security processes for IT systems, this is not the case for [industrial control systems]," Ed Ansett, chairman of I3 Solutions, told *DCD*. "MEP controllers frequently have no authentication, authorization, virus protection or security patches associated with SCADA, PLCs, RTUs, BMS and other addressable controllers often found in cooling plants, PDUs, UPS, generators, switchgear and static switches."

The OT industry will have to learn a lesson which the IT industry learnt 20 years ago: it is not enough to leave cyber security to the manufacturer. Applying updates to products requires investment and all too often, manufacturers, when left to themselves, prefer to simply cover up vulnerabilities, as it is not in their interest to make such discoveries public.

Despite this fact, certain types of vulnerabilities are widely known within the industrial control systems community, and the data center owner is not alone in dealing with them. The ICS-CERT (Industrial Control Systems Cyber Emergency Response Team) publishes known attacks and responses; in 2016, it received 305 reports involving unique vulnerabilities in control system components, the majority of them taken down through techniques like buffer overflows and DDoS (distributed denial of service) attacks - which are, again, very similar to attacks used on traditional computer networks.

Security researchers are concerned that the OT community has taken a "security through obscurity" approach to SCADA systems, trusting in physical separation of their networks.

Engineers have believed that SCADA networks are secure because they have have proprietary interfaces and are not connected to external networks.

These assumptions are no longer true - and if we dig deep, the list of potential threats to SCADA systems today is practically endless. These systems can be accessed physically by unauthorized parties, and changes can be made - accidentally or intentionally - through virus infections, malware hidden inside software upgrades, or through other threats residing on networks the system is connected to.

In many cases, the control protocol lacks any form of encryption, allowing intruders to simply modify parameters through network commands. Furthermore, the SCADA user may believe that the control system network is closed off because a VPN is applied to the IT systems involved, completely disregarding

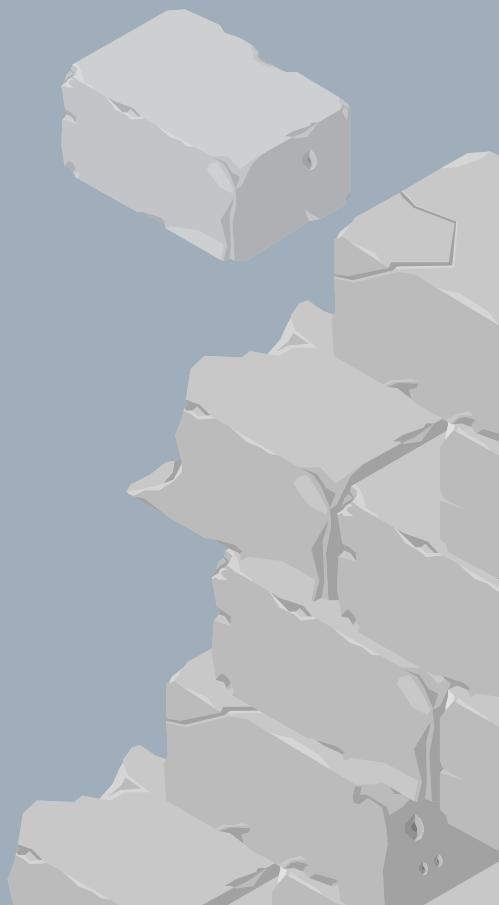
Are Russians after US utilities?

Russian government actors have been targeting US organizations in the energy, nuclear and critical manufacturing sectors, according to a March 2018 technical alert (TA18-074A) from the United States Computer Emergency Readiness Team, based on analysis by the Department of Homeland Security and the FBI.

The alert describes how parties sponsored by the Russian government carried out a multi-stage intrusion campaign, gaining unauthorized access to small commercial networks, and then conducting network reconnaissance, to access other networks and collect information about industrial control systems.

The alert explains how this is achieved, and details indicators of compromise (IoC) - evidence that penetration has occurred.

According to US-CERT, the campaign focuses on two distinct types of victims: staging targets are trusted third-party suppliers with less secure networks. Access to these networks is then used as a pivot point, to deliver malware to the intended primary target.



the possibility of physical access to the SCADA network switches.

The attacks that target these systems are similar to the ones found in IT security, but the same cannot be said about the detection and prevention methods. In a computer network, a vulnerability scanner could be used to quickly and efficiently discover hosts on a network, establish what services they are running, and which vulnerabilities are open.

However, techniques like port scanning, device fingerprinting and host probing will sometimes lock the hardware, disrupt running processes and cause erroneous readings. Industrial control systems manage machinery that must always be online. Even if they shut a device down for a couple of seconds, these techniques are not fit for SCADA.

The ICS vendors have begun addressing serious vulnerabilities, and suggest approaching SCADA security with a defense strategy based on common IT practices. Vendors have also started to address unauthorized access, by developing specialized industrial firewalls for TCP/IP based control networks and external monitoring and recording of SCADA equipment. For legacy SCADA systems, a 'bump-in-the-wire' methodology is applied - devices that offer the means of authentication and AES encryption.

As well as vendors, regulators have also taken notice of the existence of these threats. In March 2017, Cyber Security Legislation 23 (NYCRR500) was adopted by the New York State Department of Financial Services. This gave companies until 28 August 2017 to implement a cyber security program and policy - and section 500.03, clause (j), specifically referred to physical and environmental controls security.

It's not at all clear how many organizations have taken this seriously. Ansett told *DCD*: "As a minimum, data centers that host financial services will have to act urgently to undertake an audit of the vulnerabilities affecting their M&E control and monitoring systems."

It is not just the US that sees the need for regulation in this sector. The UK Government is due to set out similar regulations shortly, and other countries will follow. ●

What is being done to secure these networks?

Until recently, protection of ICS systems, particularly SCADA, was focused on reliability and elimination of random hardware faults. Currently, several industrial and government-led organizations are establishing processes and programs for securing their infrastructure.

The NERC (North American Electric Reliability Corporation), AGA (American Gas Association), API (American Petroleum Institute) and more have developed documents which recommend standards and best practices to avoid cyber incidents within ICS networks. The majority of these bodies are also authorized to enforce compliance.

Essentially, these efforts revolve around three main principles - to create awareness of security issues, to help control system operators design a security policy, and finally, to recommend basic security mechanisms for prevention.

New Online Course
Understanding Data Center M&E Cyber Security

May 2018

Whilst most organizations have a dedicated team for network security, facilities' security concerns often fall between the gaps. However, new regulations in the financial services sector have put a spotlight on the need for more education.

DCPRO is developing a new 2 hour online module - with industry experts i3 Solutions Group - that will provide students with M&E Cyber Security fundamentals. The course will be launched in May 2018.

For more information, visit: www.dcpro.training

COLOCATION NORTHWEST
AN ISOFUSION COMPANY

Is your mission critical team ready for the big one?

We're ready. Colocation NorthWest provides the safest and most reliable colocation in the Pacific Northwest.

Is your company fully prepared to scale as you grow?

It's what we do. Our South Hill Seattle 86-acre campus is built expressly for enterprise scalability as your business grows.

Is going green discussed in your strategy sessions?

We can help. Colocation NorthWest is proud of our LEED Gold rating, hydro and wind electricity sources, and green technology.

We are your single source solution provider.

LET'S TALK

1-888-210-3662
www.colonw.com/DCD

All nodes lead to Ethernet

Martin Courtney
Freelance analyst

There are ambitious plans to keep Ethernet at the forefront of data center networks. *Martin Courtney* is impressed

It might seem premature to talk about 400Gbps and 800Gbps cabling systems when the vast majority of the data center world is only just to think about laying 100Gbps wiring. Yet the sheer volume of information that hosting facilities will have to store and process over the next few years indicates that future-proofing networks against looming capacity bottlenecks should be a priority for many owners and operators.

Cisco's latest Global Cloud index predicts that global data center traffic will reach 19.5 zettabytes by 2021 for example, more than triple the six zettabytes in 2016 and representing a compound annual growth rate of 27 percent. What's more, 95 percent of that traffic is expected to be driven by the cloud as more businesses and consumers store and process the information in centrally hosted environments.

19.5 zettabytes of data center traffic by 2021 (Cisco Cloud Index)

That trend has led to a widespread expectation that hyperscale cloud hosting companies like Amazon Web Services (AWS), Facebook, Google and Microsoft will start implementing 400Gbps cabling systems as early as 2019 whilst simultaneously preparing themselves for 800Gbps upgrades at a later date. Much of that activity will be focused on the core leaf and spine parts of the network, as well as inter data center links.

Even owners and operators of more modest hosting facilities - tier two cloud providers, systems integrators and large enterprises and government departments for example - will find that the increased use of 100Mbps cabling systems at the top of rack level will inevitably create choke points within current network architecture, hastening upgrades to core backbones.

The bigger question is not so much if or when 400Gbps or 800Gbps networks will be needed, but what form they will take. And the answer to that varies according to individual data speed, cabling distance and component size/power requirements combined with planned expansion strategies and the type of network/server/storage architecture already in place.

When it comes to server interconnects, Ethernet seems to have become the specification of choice amongst large scale hosting providers and enterprise data centers focused on cloud service delivery whereas InfiniBand is favored by the high performance computing (HPC) community.

For linking external storage arrays however, Fiber Channel (FC) is widely used due to its considerable legacy install base.

Both InfiniBand and Fiber Channel have roadmaps which push existing capacity further. The T11 specifications for 64GFC and 256GFC products delivering maximum bi-directional throughput of 128Gbps and 512Gbps were completed last year, for example, with commercial products expected in 2019. Those will give data centers with existing FC infrastructure at the network edge and server interconnect level additional capacity when it comes to upgrading the 16/32/128GFC components already in situ using OM3/OM4 MMF cabling systems, especially when it comes to linking servers to external data storage devices.

The InfiniBand Trade Association (IBTA) also published a roadmap outlining 1x, 4x and 12x port widths with bandwidth reaching 600Gbps in 2017, again for server interconnects over short distance passive and active copper cables (up to 30m) or optical cables (up to 10km).

But it is Ethernet which appears to have stolen a march on both FC and InfiniBand, at least in terms of the capacity it expects to support in the nearer term. In March this

year the Ethernet Alliance, a consortium of vendors, industry experts, academics and government professionals committed to the success and expansion of Ethernet, released the latest Ethernet roadmap mapping out future iterations of the technology.

It expects to see 400 gigabit Ethernet (GbE) links deployed in hyperscale data centers by 2020, with 800GbE and 1.6 terabit Ethernet (TbE) connectivity appearing within five years or so. Of course any timescale for end user deployment depends on when individual component manufacturers can get suitable components onto market, and how affordable they are.

The 400GbE and 200GbE specifications were ratified by the IEEE 802.3 Ethernet Working Group in December 2017. Rather than making it optional, the 802.bs architecture embeds Reed Solomon forward error correction (FEC) in the physical coding sub-layer (PCS) for each rate, effectively forcing manufacturers to develop 200GbE and 400GbE Extender Sublayers to support the future development of other PCS sublayers that can utilize other types of FEC for greater efficiency at a later date.

Three 200Gbps standards - 200GBASE-DR4 (500m), 200GBASE-FR4 (2km) and 200GBASE-LR4 (10km) - all use single mode fiber (SMF) and 50Gbps per lane to achieve the desired throughput. An equivalent SMF based 400GbE standard - 400GBASE-DR4 (500m) - boosts that bandwidth to 100Gbps over four lanes, whilst 400GBASE-FR8 (2km) and 400GBASE-LR8 (10km) use eight lanes at 50Mbps. A fourth 400GbE specification - 400GBASE-SR16 - combines 16 strands of MMF fiber at 25Mbps to push 400Mbps signals over distances of 100m.

Several new optical I/O form factors to meet those standards have now emerged. These include CFP8, OSFP, QSFP-DD, and COBO, again designed for different types of MMF or SMF wiring and electrical interfaces and optimized to suit various metrics within the data center, most notably different transmission distance requirements; backwards compatibility with existing systems; and component space, heat and power consumption constraints within densely populated data center compute and network architecture.

Some manufacturers are already well advanced in their plans for these form factors. 400GbE compliant CFP8 transceivers using 50G pulse amplitude modulation (PAM4) technology having been demonstrated by various companies, including Finisair and NeoPhotonics. The transceivers have been modified for compliance with the 400GBASE-FR8 SMF standard pushing maximum transmission distance out to 2km for campus data center networks.

Mellanox too has indicated its intention to introduce ASICs supporting 400GbE at some point in 2018, whilst Huawei late last year completed tests of 400G optical network technology in partnership with China Telecom and Spirent for commercial use in access, metro and data center networks.

Name	Cable Type	Max Speed	Max Distance	Notes
100GBASE-CR10	Twinax Copper	10×10 Gbps	7 m	CXP connector, center 10 out of 12 channels
100GBASE-SR10	MMF	10×10 Gbps	100 m/125 m	MPO/MTP connector, center 10 out of 12 channels
100GBASE-SR4	MMF	4×25 Gbps	100 m/125 m	850nm wavelength over MMF
100GBASE-LR4	SMF	4×25 Gbps	10 km	Uses four lanes of SMF
100GBASE-ER4	SMF	4×25 Gbps	40 km	Uses four lanes of SMF

What no 1TbE?

Beyond 400GbE it looks like 800GbE and 1.6TbE will emerge as the default specifications at the expense of the 1TbE specifications that were initially proposed, according to John D'Ambrosia, senior principal engineer at Huawei and a leading advocate of Ethernet, having helped co-found the Ethernet Alliance 12 years ago.

"Will there ever be a Terabit Ethernet speed? I would say no," he told *DCD*. "Ethernet likes binary and it is more likely to be 800Gbps or 1.6Tbps because the ten lane problem turned out to be more difficult than people realized during the early days of 100Gbps.

800GbE will probably become a reality when single-lane 112Gbps links that double the 56Gbps Multi-Protocol SerDes (MPS) PHYs supplied by Rambus and others and compliant with PAM4 and NRZ IP cabling solutions hit the market.

Those could theoretically support either 1TbE or 1.6TbE in the future with 10 or 16 lanes running in parallel.

But whilst the 10 lane proposals came from those in the HPC community, D'Ambrosia feels that it is the hyperscale data centers and larger Internet exchange (IX) companies which are concerned that 800GbE will not be sufficient to meet their needs in the future which are most likely to push for 1.6TbE.

"We are only just getting 400GbE out and in discussions over the next speed. 800GbE seems a more plausible solution but that is only a 2x upgrade for those people already struggling with capacity," he said.

DCD>Debate | Register Now!
Hyperscale network architecture
Jun 6 2018
11.00am CST

How is HSDT tech solving hyperscale data center congestion? AI is already testing the limits of current network infrastructure as the demands of the GPU chipset are better understood, but what else is on the application roadmap and how will networking technology and the data center interconnect evolve to cope with the pressure?

bit.ly/DCDHyperscaleNA

Sebastian Moss
Reporter

Beyond the FPGA: Xilinx's new CEO talks about what's next

Victor Peng believes his company's future lies in the data center, *Sebastian Moss* reports on how he hopes to get there

In January, Victor Peng was appointed the CEO of Xilinx, the US company best known for creating field-programmable gate arrays (FPGAs) - integrated circuits that can be reconfigured using software. Now, he's ready to talk about his plans for the decades-old semiconductor firm.

"This is my first press event," Peng told *DCD* as he laid out the company's new strategy: Data Center First.

Xilinx has achieved success in its core markets, like wireless infrastructure, defense, audio, video and broadcast equipment. But Peng believes that "the data center will be our largest area for growth, and we will make it out highest priority."

The company's push into the data center comes on three fronts - compute acceleration, storage, and networking. "We can play in all three," Peng said.

"One of the challenges is that we're still primarily known as the FPGA company, and it's the classic issue of 'what got you here won't necessarily get you where you want to go.' The fact that we invented the FPGA and lead in this field is a good thing, but then there are some very outdated views on what an FPGA can do.

"Some of our competitors want to perpetuate that old mentality because they want you to think that it's too expensive, low performance and only good for prototyping. That's not true."

The other challenge is developer support, the lack of which can kill a product (see *The PlayStation Supercomputer* p44): "We need to make it significantly easier to develop for. We need to build out the ecosystem, libraries and tools and more applications. But that's exactly what we are doing.

"We've increased our hiring on the software side, but we can't do it all organically, so that's why we do have to work with partners, the ecosystem, and make investments. We're doing more things in open source, so collectively the R&D is significant."

Peng claims FPGAs are ideal for the growing number of machine learning workloads (inference, but not training), as well as video transcoding. He points towards Xilinx's deal with AWS to provide the cloud company with UltraScale Plus FPGAs for EC2 F1 instances. According to AWS, these instances can accelerate applications up to 30x when compared to servers that use CPUs alone.

Peng refuted suggestions that the instances have not been as successful as some had expected: "What tends to happen is people feel like 'it's been around a while, how come it hasn't taken off in a broader way?' But they announced in November 2016, they went to general availability in April 2017 - and at that time there were zero applications, they were just opening the doors for developers."

Today, according to Peng, there are around 20 applications using FPGAs in production for cloud customers, and more than 50 app developing companies that are targeting this field. AWS is also expanding the availability of the instances, while Alibaba, Huawei, Baidu and Tencent all plan to launch their own Xilinx FPGA-based services.

"A lot of that is part of a goal to be a software targeted platform, and FPGA-as-a-service is a fantastic vehicle for it. It gives a platform for the application developers to use this without having to design a board."

Another example of a successful, large-scale FPGA deployment that Peng likes to reference is that of Microsoft's Project Catapult - where the software giant inserted FPGAs into servers to create an 'acceleration fabric' throughout the data center. Used for Microsoft Azure, Bing and AI services, this set-up helps speed up processing and networking speeds, and is now a standard component of the company's cloud infrastructure.

"I think that model of homogeneous, distributed, but adaptable acceleration is very powerful," Peng said. "Because that gives you the best TCO overall, when you have a dynamic environment; the best TCO is not to have fixed appliances for everything you might run."

There's one problem, however - the FPGAs Microsoft uses are from Altera, the company Intel bought for $16.7 billion in 2015. "When they were an independent company, it was a duopoly, we'd be selling to the same customer, sometimes we'd win, sometimes we wouldn't. We've been talking to Microsoft, and we continue to talk to Microsoft."

While Peng expects Xilinx to continue offering classic FPGAs "for decades," the other announcement the new CEO was ready to share was the development of a brand new product called ACAP.

The 'Adaptive Compute Acceleration Platform' is made up of a new generation of FPGA fabric with distributed memory and hardware-programmable DSP blocks, a multi-core SoC, and one or more software-programmable and hardware-adaptable compute engines, connected through a network on a chip.

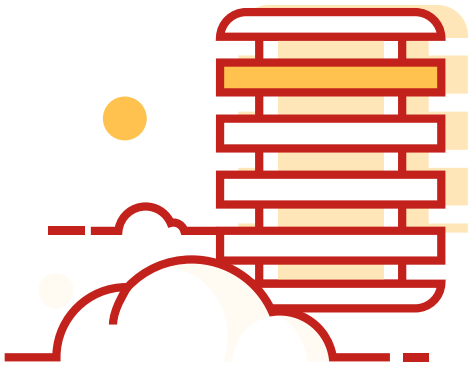
"This message of 'don't think of it as an FPGA' has got to be something we do over and over again, because after being the FPGA company, we have to say 'no we're not the FPGA company.' With ACAP, at the moment nobody even knows what that is - but they will understand over time."

The first ACAP product family, codenamed Everest, will be developed with TSMC 7nm process technology and will tape out later this year, actually shipping to customers next year.

Creating Everest took $1bn, 1,500 engineers and four years of research, and the company expects it to achieve a 20x performance improvement on deep neural networks compared to the 16nm Virtex VU9P FPGA.

Peng was coy on specifics, but confirmed that Xilinx is in advanced talks with cloud companies about ACAP.

With a renewed focus and a new product, we asked how the success of this new strategy will be judged. "It's hard for me to know how it will ramp," Peng said. "There's no roadmap, this is a whole new area, and not just for us, but for the industry. I mean who would have predicted Nvidia's ramp up? Even Nvidia didn't predict Nvidia's ramp up."



FOR SALE

50 ACRE INDUSTRIAL HYDROELECTRIC PROPERTY

With 40,000 square feet of building space. This property has several hydroelectric assets including a hydro dam, water rights, riverbed, penstocks, head gates and more.

Potential power generation is 2.2 megawatts with the potential to produce 11,000,000 kWh of electricity per year! This free electricity can dramatically lower operating costs for a data center or crypto mining operation.

- Hydro Quebec has approved the feasibility of an additional 20 megawatts electricity supply to this property.
- Videotron has approved an internet service supply with 50 gigabytes upload and 50 gigabytes download to this property.

Email: brianwmcd@hotmail.com

CALL: 418.894.0193
or Email: brianwmcd@hotmail.com

DCD

> **AI + Automation** | In-depth

INSIDE

How to get to the software-defined data center

> Look to the large Internet companies for the first attempts at building an SDDC

An intelligent bank banks on intelligence

> How ICICI Bank's smart approach to infrastructure cut costs and improved services

Nvidia dreams of simulated worlds, AI playgrounds

> The GPU king has grand plans to expand its dominion, building realities on its chips

Artificial intelligence meets DCIM

> Nlyte's CEO Doug Sabella sits down with Max Smolaks to talk smart systems

How smart is your data center?

Peter Judge
Global Editor

Artificial intelligence (AI) and robotics are coming to your facility - but the progress will be from the ground up, says *Peter Judge*

We've been watching the rise of data center automation at *DCD*. There are physical robots coming that will handle hardware upgrades much like a tape jukebox, with automated arms sliding hardware in and out of the racks. A fully lights-out data center, with no human operators, will have to work this way. But that's a long way off.

Software automation is closer to hand. Machine learning systems are taking over a lot of basic tasks. But how long until the fully automated data center arrives, and will it come from the bottom up, or from the top down?

Large Internet operators such as Google or Facebook can build monolithic data centers, where intelligence can be applied to a small range of applications. Google, for instance, used its DeepMind AI to optimize its cooling systems, cutting the power required to run them, and reducing its PUE (power usage effectiveness) by 15 percent.

For the rest of us, *DCD* would bet on technology developing from the ground up, in the normal way. Two major movements have created potential for automating data centers; all that is required is to overlay intelligence on them to create an automated system.

The first pillar is DCIM (data center infrastructure management). It has been implemented widely in data centers, instrumenting the IT and facilities hardware, so its performance can be monitored and controlled. The obstacle it has faced is in the sheer volume of data it produces, and the complexity of the actions required to run a data center.

Artificial intelligence, and its subset machine learning, excel at precisely these tasks. Pattern recognition can pick out important events as signals from the wall of noise coming from sensors, while intelligent systems can learn the complex interrelations of the actions they might make, so as to deliver the correct response.

The second pillar is software-defined hardware management. The major resource pools in a data center - compute, storage and network - can all be delivered under software control, leading to a proposal that all resources should be available on demand, in a software-defined data center (SDDC) (see p34).

In the world outside the web-giants, the best initiatives to put intelligence in data centers have been built upwards from DCIM and software-defined resources. In 2017, DCD's Smart Data Center Award went to ICICI Bank in India, for a system which uses information from DCIM, a building management system (BMS) and virtualized IT services, and includes a special application management service built in-house to deliver data centers which can optimize their own performance and deliver more flexible services (see p37).

In fact, AI makes a natural addition to DCIM, giving it intelligence to handle levels of information that would overwhelm humans, according to Doug Sabella, CEO of Nlyte (p43).

But is automation always driven from the bottom upwards? Another approach, while still building on the data provided by low-level automation, creates an AI persona that absorbs and learns the best practices of humans in data centers.

Litbit aims to collect and apply the human expertise of data center staff, and the ambitious idea has won over partners with their feet on the ground, including real estate firm CBRE (see box).

Finally, AI works best when fast hardware is available. Nvidia's GPUs have carved out a niche in the area, and the next step is to use them to simulate everything, the company told us (p40).

Whether built from general hardware or specialized machines, and whether they comes from below or above, AI and machine learning will ease the human burden of data center operation.

REMI says hi

Commercial real estate giant CBRE is not known for following fads, but it has teamed up with AI software firm Litbit to train an automated facilities maintenance system.

The Litbit AI persona, named REMI (Risk Exposure Mitigation Intelligence), will learn how a data center works by watching how CBRE staff do their jobs. Because CBRE runs 800 data centers, with thousands of technicians and milllions of machines, REMI could amass the world's largest actionable AI repository of machine operating data, Litbit CEO Scott Noteboom said at DCD>Zettastructure.

REMI will run both in the cloud via smartphone and on the edge through permanently deployed computing devices.

Paul Saville-King, CBRE's data center solutions president, said: "The idea that expert knowledge of all the facilities, assets and equipment we manage can be in the back pocket of every CBRE technician is exciting and revolutionary."

DCD>Debate | Register Now!
AI in the data center?

May 22 2018
11.00am CST

Are AI and Machine learning ready for mainstream in the data center? DCD's Bruce Taylor talks to Enzo Greco, Chief Strategy Officer, Nlyte and an all-star cast of industry experts
bit.ly/NlyteMachineAI

SDDC is the answer - but how do we get there?

We dream of automated cloud data centers, but *Dan Robinson* wants to know how we get there from here

Dan Robinson
Correspondent

The software-defined data center (SDDC) is a trend that has been growing in importance for several years, driven in part by factors such as virtualization and cloud computing. Infrastructure has been evolving to be more dynamic and adaptable in order to serve a business requirement for greater agility. But what is an SDDC, and how do you get there from the systems in place now?

A simple definition of an SDDC is that it is a data center where the entire infrastructure is configurable under software control. This does not mean just servers, but also the networking used to interconnect them and the storage resources available to service the needs of applications.

As an example, look no further than the data centers operated by the large Internet companies such as Google, AWS and Facebook, which need to be able to respond to constantly shifting demands on resources from customers. These have set the pattern for other service providers and enterprises to follow.

These hyperscale companies built their data centers to be this way from the outset, but only need to support a limited range of services. Other organizations are likely to have to deal with a wider range of uses, and legacy infrastructure that they cannot replace all in one go. The hyperscale operators also have the resources to develop their own automation and orchestration tools, while others will have to make do with off-the-shelf commercial or open source tools.

According to Ovum principal analyst Roy Illsley, software-defined networking and software-defined storage are set to be the fastest growing areas of IT spend over the next few years, followed by cloud orchestration and management.

"This tells you that the market is heading towards the fact that everything is going to be software-defined, everything is going to be software controlled, and the internal spending of the IT department is increasingly shifting to software rather than hardware now," he said.

But there is no single definition of what software-defined means. While it is fairly straightforward for servers, in that it usually means dividing up hardware resources among multiple workloads using virtual machines or containers, things are not so simple for storage or networking.

Software-defined networking (SDN), for example, separates the network's data plane, which forwards data packets, from the control plane, which manages overall traffic flow. To support SDN, switch vendors such as Cisco have to make their hardware configurable using protocols like OpenFlow, so the way they route traffic can be dynamically managed by a central controller.

However, SDN platforms such as the OpenStack Neutron service and VMware's NSX run on the servers and manage traffic between virtual machines, using software-based switching. They also support the creation of virtual networks that overlay the physical LAN, but enable a different range of IP addresses and security policies.

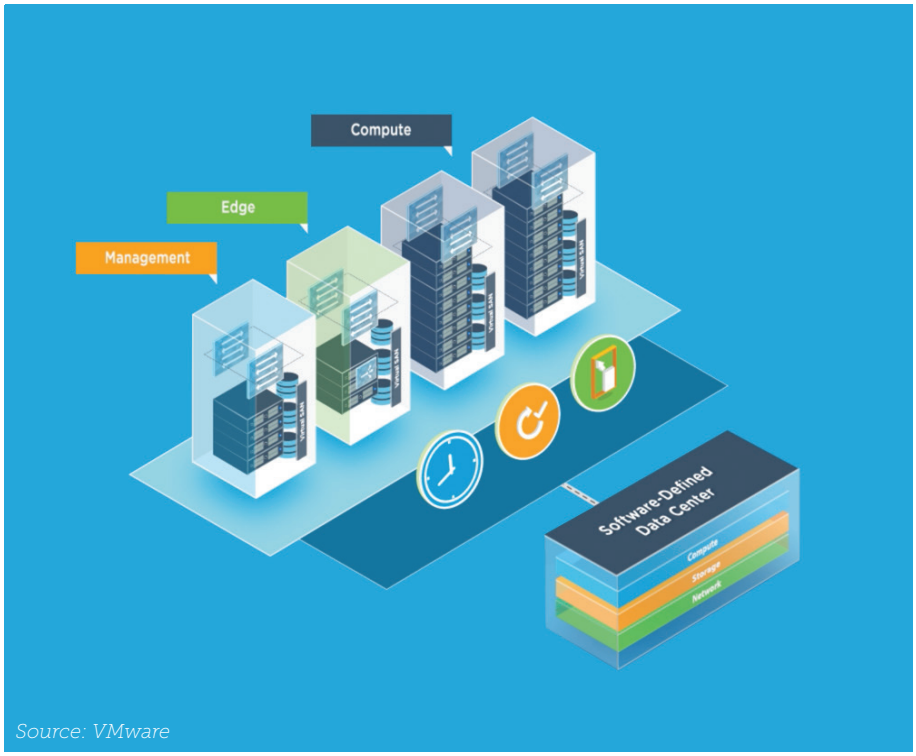
Software-defined storage (SDS) is also tricky to pin down. Perhaps the most common definition is a distributed storage service, such as Red Hat's Ceph or Gluster products. These are used to create a scalable pool of storage from the combined resources of a cluster of server nodes, and present it as block storage, object storage, a file system, or combinations of these.

> *"Everything is going to be software-defined and the budget is shifting from hardware to software"*

Meanwhile, SDS can also refer to storage virtualization such as DataCore's SANsymphony. This abstracts and pulls together existing storage infrastructure, including arrays from third party vendors into a virtual SAN. It then provides a unified set of storage services for this pool of storage, including quality of service, thin provisioning and auto-tiering.

However you characterize it, the purpose of software-defined infrastructure is to be flexible: configured, controlled and monitored by higher level management tools. This could be via configuration management software such as Puppet, or orchestration platforms such as OpenStack or Mesosphere's DC/OS.

But many organizations will have a lot of legacy kit that may not lend itself well to



Source: VMware

this model of operation. This means that they may be forced to operate a "two-speed" IT infrastructure while the refresh cycle gradually brings in modern kit that can be software-defined.

To address this, some newer platforms are described as pre-packaged SDDC solutions. A good example is VMware's platform, which is based on three pillars; vSphere for operating virtual machines; vSAN, its software defined storage product; and NSX, which provides software-defined networking.

These are combined with suites of management tools in various ways to deliver products such as VMware Cloud Foundation, which can be deployed onto hyperconverged systems hardware in a customer's own data center or on a public cloud, as with VMware Cloud on AWS.

Microsoft touts Windows Server 2016 as an SDDC platform, thanks to Hyper-V for operating virtual machines, Storage Spaces Direct and Hyper-V Virtual Switch, plus the System Center suite for management.

There are other similar offerings, and most require the customer to purchase a complete integrated platform. These can start with a few nodes and scale out to rack level, or even larger, but all essentially lock the customer into one vendor's platform.

If you prefer an open source software alternative, there is the OpenStack framework. This has a modular architecture made up of numerous separate projects, with the core modules including Nova for managing compute, Neutron for configuring networking, plus Cinder, and Swift for block storage and object storage, respectively.

OpenStack is notably used by CERN, the European particle physics laboratory, to manage tens of thousands of compute nodes forming the IT infrastructure serving the Large Hadron Collider and other experiments.

Thus far, this article has only touched on IT infrastructure, but data centers also comprise other facilities such as power distribution and cooling. Might these also be managed under software control in order to make the most efficient use of resources?

Software-defined power is starting to get some attention from vendors such as Virtual Power Systems (VPS). This firm has developed its Intelligent Control of Energy (ICE) technology to enable the use of UPS batteries to meet some of the power demand during periods of peak loads. This means that the power distribution infrastructure does not have to be over-provisioned to cope with peak loads that may only occur infrequently - and the data center owner may get a rebate from the energy utility.

In terms of cooling, data center infrastructure has become smarter, but is rarely marketed as "cooling on demand." One early approach was from HPE, which some years ago (as HP) touted a combination of sensors and computational fluid dynamics (CFD) to analyze the flow of air within the data center and route cold air to where it was most needed.

More recently, liquid cooling proponents mention their technology's ability to remove heat precisely from where it is generated, and Inertech, a subsidiary of Aligned Energy, is offering a system where the cooling units are distributed and positioned above the racks, so cooling can be delivered on demand according to requirements.

It seems clear that all aspects of data centers are becoming instrumented, and are controlled more precisely by software. The future of the data center is software-defined - but two questions remain: what exactly will that look like, and how long will it take for organizations to get there?

How hyperconverged fits with software-defined

Hyperconverged systems, otherwise known as hyperconverged infrastructure (HCI), represent a growing proportion of server shipments into the enterprise market.

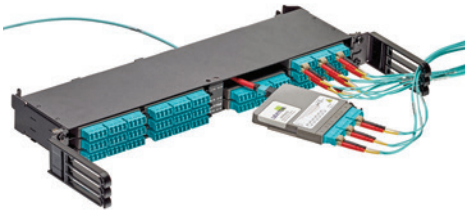
There are good reasons for this; hyperconverged systems, which combine server, storage and networking functions into a single appliance-like node, make it easier for organizations to procure and stand up new infrastructure.

In particular, IT departments have often struggled to build a private cloud out of existing deployments of servers, storage and networking from various different sources. Hyperconverged systems have emerged as a turnkey solution when such firms come to update their existing infrastructure.

Inside the box, hyperconverged systems use direct-attached storage in place of a traditional SAN. The storage resources of all the nodes in a cluster are pooled together to form a software-defined storage platform or virtual SAN that is shared between the nodes.

That sounds a lot like software-defined infrastructure, and indeed it is. But it is important to note that HCI involves compute and storage running on the same node, whereas a typical software-defined storage deployment uses clusters of nodes just operating as a shared storage system.

In this sense, HCI can be thought of as a signpost towards software-defined data centers. It takes the simpler case of a single node, while full SDDC will take a little longer to appear.



Quick, easy network deployment is possible when your cable assemblies are made-to-order.

The Leviton Data Centre Factory creates pre-terminated fibre optic cabling, delivered to you fast from the UK, for minimised network downtime and easy job site clean-up.

- ✓ **Faster deployment**
- ✓ **Guaranteed high performance**
- ✓ **Peace of mind**

Learn more about the advantages of Leviton cable and connectivity: explore our many fibre optic and copper options at **brand-rex.com/products/configurator.**

LEVITON®

7003-eu C18

Smart from the ground up

India's ICICI Bank started from fundamentals to build an intelligent data center infrastructure, reducing its energy bill and delivering better services, *Peter Judge* reports

Peter Judge
Global Editor

Making a smart data center is one thing. Making a smart data infrastructure is another. That's the job that Indian bank ICICI took on when it upgraded infrastructure which drives more than 14,000 ATMs and 4,800 branches around the world. The result won it DCD's Smart Data Center Award in 2017.

Smart data centers apply automation at all levels: they instrument the hardware to improve performance, and gather and harness data so the facility itself can learn and improve (see box).

ICICI (the Industrial Credit and Investment Corporation of India) embodied all of this in a project which collected data and automated processes at multiple data centers. The job combined IoT-based data center environmental management, centralized building management systems (BMS), adaptive capacity management and predictive analytics into a single software defined data center (SDDC) approach.

To complete their vision, the team had to build an application management tool - called App360. The end result is a more stable data center ecosystem, which operates more efficiently, requiring less energy. It also demands less management effort, as it handles requests and incidents automatically instead of manually, and can even fix problems before they happen. ▶

The project was all about the results, not technology for its own sake, ICICI managing partner Imran Shaik told *DCD* in ICICI's Award submission: "Being one of the topmost banks in India, it is imperative to ensure that the availability and performance are given the topmost priority since technology is the backbone of all banking services."

The bank has two data centers: a primary facility and a disaster recovery data center. Its IT hardware is virtualized using VMware and delivered as a service - apart from five percent of the x86 systems, which are not virtualized because of requirements for compliance or performance.

The bank has already deployed hyper-converged infrastructure using software-defined all-flash storage arrays, and software-defined networks (SDN) within the data center and across all the branches.

ICICI has unified performance management, linked to business availability, for the infrastructure and the application workloads which run on this. Across the bank, services are backed by a service level agreement (SLA), with tiered priority similar to the platinum, gold or silver service tiers offered by commercial data centers. The cost of usage is apportioned to each group within the company and recovered with a chargeback mechanism.

To enable all this, numerous tools have been implemented. IoT sensors were installed for environmental monitoring and capacity planning, managed by Vertiv's Trellis data center infrastructure management (DCIM). The Trellis tool aggregates data from temperature and humidity sensors to create real-time thermal mapping and visualization at the rack level, so the heat load can be monitored and space usage optimized.

"Implementation of DCIM enabled us to ensure online real-time measurement, making power and performance trade-offs while focusing on uptime, availability, performance and power usage," said Shaik.

Trellis helps explore the data center's total consumption, energy costs, and PUE. Combined with intelligent PDUs this helps ICICI to manage power and efficiency proactively. The DCIM operates as a closed loop system generating SNMP traps and various alerts for upstream systems such as the centralized building management system.

The BMS monitors and controls the mechanical and electrical components for five buildings across the country in one place. This includes chillers, UPS, diesel generators, precision air conditioning (PAC) units and safety equipment.

To track and manage applications, ICICI Bank developed its own customized application monitoring tool called App360, which acts as a single repository for all application details, backup policies and purging policies. App360 provides a complete mapping of all applications, virtual machines, physical servers, storage devices, backup systems and networks. It has a built-in alerting mechanism for events like SSL certificate expiry, produces reports, and tracks server activity and incidents. It also has automatic scripts and sends reminder emails: "No such consolidated tool existed in the market place," said Shaik.

October 2016
ICICI became the first Indian bank to trade using blockchain

The App360 tool ensures any unforeseen eventualities can be managed efficiently, avoiding the business impact of downtime. In an incident, the support team can check App360 for infrastructure details: when a base server goes down, it provides information about the applications which are hosted on it, so the right teams can be notified.

The bank has an incident management process with a centralized IT command center which uses multiple tools and a customized Service Manager tool from HP, to give management control over incidents and their follow-ups. The tools include Oracle Enterprise Manager, HP Operations Manager, Windows SCOM, NetApp OCI, Appnomics, Appdynamics, Dynatrace, HPOVM, SCOM, and OpsCentre - all cascaded with the in-house App360 tool.

Rack level power is handled by Sentry Power Manager (SPM), which enables socket level monitoring. Predictive analysis enables power management and capacity planning. The bank deployed modular power distribution units (PDUs), which helped manage the dynamic load, and provide a further cost saving on capex.

The monitoring systems provide historical data, and enable trend analysis of incidents, so users can identify issues and choke points, and take proactive actions before actual failures.

All this data is held in a Hadoop data lake, including structured, un-structured and semi-structured data, with data discovery, optimization and analysis. It can be accessed quickly, and searched by multiple factors, including the IP address of servers, appliances, load balancers, switches and storage, all from a single menu. All this helps staff to generate reports quickly and respond to contingencies.

All this has brought results. The bank has reduced the man-hours spent on operations by 20 percent while delivering faster responses to new requirements, enabling it to operate more flexibly.

The intelligent system predicts outages, and fixes them before they happen with proactive, preventive measures.

Rack level metrics have inspired the operations team to balance cooling and the IT heat load, alarm management, notification, and set thresholds for environmental sensors.

"Hotspots in any data center are always the biggest concern," said Shaik. "Hotspots identified through DCIM are addressed through optimizing IT equipment placement, realignment of the raised floor tiles, adding active tiles, adding baffling/blanking panels, deploying rack mount fan trays for hot air exhaust and fine tuning the cold aisle containment."

This leaves the facilities performing at optimal environmental conditions meeting ASHRAE standards. The precision air conditioning (PAC) units are using 48 percent less power, and the chiller power consumption has been cut by 13 percent - generating a direct benefit to operating expenditure (opex).

Even in the hot and semi-arid climate of Hyderabad, India, the data center has achieved a PUE of less than 1.5 - an achievement of which the bank can be justly proud.

Winner of the Smart Data Center Award
DCD>Awards | 2017

Open for Submissions

The Smart Data Center award is given to a project exemplifying software-defined automation in any aspect of the data center.

The concept of smarter data centers began with data center infrastructure management (DCIM), a move towards instrumenting and controlling the hardware inside a facility. DCIM was, if you like, the Internet of Things before its time.

The intelligence in use with DCIM has increased, and data center architects have become more ambitious. Resources within data centers can all be pooled and managed automatically, leading to the concept of the software-defined data center (SDDC), and machine learning has been applied so these facilities can improve and develop their performance.

dcdawards.global

DCD > Awards | 2018

Entry Deadline
September 21, 2018

The Data Center Awards are **Open for Entries!**

Category 1
Living at the Edge Award

Category 2
The Infrastructure Scale-Out Award

Category 3
The Smart Data Center Award

Category 4
The Data Center Eco Sustainability Award

Category 5
Energy Efficiency Improvers Award

STARLINE

Category 6
Mission Critical Innovation Award

CBRE

Category 7
The Open Data Center Project Award

Category 8
Cloud Migration 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
Data Center Manager of the Year - NEW

Category 12
Design Team of the Year

Category 13
Industry Initiative of the Year

Category 14
Young Mission Critical Engineer of the Year

Category 15
Corporate Social Responsibility Award - NEW

Category 16
Business Leader of the Year Award

Category 17
Outstanding Contribution to the Data Center Industry

MERCURY

Charity Partner

CHAIN of HOPE

For more sponsorships and table booking information, please contact:
global.awards@datacenterdynamics.com

Public voting

The World's Most Extreme Data Center

www.**dcdawards.global**

To conquer the world, Nvidia will simulate it

The GPU king dreams of virtual worlds where artificial intelligence will learn to understand our chaos. *Sebastian Moss* reports from inside the experience machine

Sebastian Moss
Reporter

Fifteen years ago, Nvidia made a decision that would eventually place it at the forefront of the artificial intelligence revolution, its chips highly sought after by developers and cloud companies alike.

It decided to change the nature of its graphics processing unit, then primarily used by the video games industry.

"We decided that a GPU, which was a graphics accelerator at the beginning, would become more and more general-purpose," Nvidia CEO and co-founder Jensen Huang said at the company's annual GPU Technology Conference.

"We did this because we felt that in order to create virtual reality, we had to simulate reality. We had to simulate light, we had to simulate physics, and that simulation has so many different algorithms - from particle physics and fluid dynamics, and ray tracing, the simulation of the physical world requires a general-purpose supercomputing architecture."

The impetus to create a general-purpose graphics processing unit (GPGPU) came when Nvidia "saw that there was a pool of academics that were doing something with these gaming GPUs that was truly extraordinary," Ian Buck, the head of Nvidia's data center division, told *DCD*.

"We realized that where graphics was going, even back then, was more and more programmable. In gaming they were basically doing massive simulations of light that had parallels in other verticals - big markets like high performance computing (HPC), where people need to use computing to solve their problems."

Huang added that, as a result, "it became a common sense computer architecture that is available everywhere. And it couldn't have come at a better time - when all of these new applications are starting to show up, where AI is writing software and needs a supercomputer to do it, where supercomputing is now the fundamental pillar of science, where people are starting to think about how to create AI and autonomous machines that collaborate with and augment us. The timing couldn't be more perfect."

Analysts and competitors sometimes question whether Nvidia could have predicted its meteoric rise, or whether it just had the right products at the right time, as the world shifted to embrace workloads like deep learning that are perfect for accelerators. But, whether by chance or foresight, the company has found itself in demand in the data center, in workstations and, of course, in video games.

"We can still operate as one company, build one architecture, build one GPU and have it 85 percent leveraged across all those markets," Buck said. "And as I'm learning new things and giving [them] to the architecture team, as the gaming guys are learning new things, etc, we're putting that into one product, one GPU architecture that can solve all those markets."

With its core GPU architecture finding success in these myriad fields, Nvidia is ready for its next major step - to push further and deeper into simulation, and help create a world it believes will be built out of what we find in those simulations.

One of the initiatives the company is working on to illustrate the potential of simulation is Project Clara, an attempt to retroactively turn the 3-5 million medical instruments installed in hospitals around the world into smart systems.

"You take this 15-year old ultrasound machine that's sitting in a hospital, you stream the ultrasound information into

Nvidia CEO Jensen Huang at the GPU Technology Conference 2018
Photography: Sebastian Moss

your data center. And your data center, running this stack on top of a GPU server, creates this miracle," Huang said, as he demonstrated an AI network taking a 2D gray scan of a heart, segmenting out the left ventricle, and presenting it in motion, in 3D and in color.

Nvidia bills Clara as a platform, and has turned to numerous partners, including hospitals and sensor manufacturers, to make it a success. When it comes to software, the company often seeks others, looking to build an ecosystem in which its hardware is key.

Sometimes, Nvidia is happy to take a backseat with software, but there is one area in which the company is clear it hopes to lead - autonomous vehicles. "We're dedicating ourselves to this. This is the ultimate HPC problem, the ultimate deep learning problem, the ultimate AI problem, and we think we can solve this," Huang said.

The company operates a fleet of self-driving cars (currently grounded after Uber's fatal crash in March), with each vehicle producing petabytes of data every week, which is then categorized and labeled by more than a thousand 'trained labelers.'

These vehicles cannot cover enough ground, however, to rival a mere fraction of the number of miles driven by humans in a year. "Simulation is the only path," Carlos Garcia-Sierra, Nvidia's business development manager of autonomous vehicles, told *DCD*.

This approach has been named Drive Constellation, a system which simulates the world in one server, then outputs it to another that hosts the autonomous driver.

Martijn Tideman, product director of TASS International, which operates its own self-driving simulation systems, told *DCD*: "You need to model the vehicle, create a digital twin. But that's not enough, you need a digital twin of the world."

"This is where Nvidia's skill can really shine," Jensen said. "We know how to build virtual reality worlds. In the future there will be thousands of these virtual reality worlds, with thousands of different scenarios running at the same time, and our AI car is navigating itself and testing itself in those worlds, and if any scenarios fail, we can jump in and figure out what is going on."

$137bn
Nvidia's market cap

Following self-driving vehicles, a similar approach is planned for the robotics industry. "I think logically it makes sense to do Drive Constellation for robotics," Deepu Talla, VP and GM of Autonomous Machines at Nvidia, told *DCD*.

It is still early days, but Nvidia's efforts in this area are known as 'The Isaac Lab,' - it aims to develop "the perception, the localization, the mapping and the planning capability that is necessary for robots and autonomous machines to navigate complex worlds," Huang added.

As Nvidia builds what it calls its 'Perception Infrastructure' to allow physical systems to be trained in virtual worlds, there are those trying to push the boundaries of simulation - for science.

"There is a definite objective to simulate the world's climate, in high fidelity, in order to simulate all the cloud layers. Imagine the Earth, gridded up in 2km grids and 3D too," Buck told *DCD*. "It's one of the things that they're looking for, for exascale."

Huang said: "We're going to go build an exascale computer, and all of these simulation times will be compressed from months down to one day. But what's going to happen at the same time is that we're going to increase the [complexity of the] simulation model by a factor of 100, and we're back to three months, and we're going to find a way to build a 10 exaflops computer. And then the size of science is going to grow again."

It is this train of thought, of ever more complex simulations on ever more powerful hardware, that has led some to propose the 'simulation hypothesis' - that we are all living in a *Matrix*-like artificial simulation, perhaps within a data center.

We asked Huang about his views on the hypothesis: "The logic is not silly in the sense that how do we really know, how do we really know? Equally not silly is that we're really machines, we're just molecular, biological machines, right?

"How do we know that we're not really a simulation ourselves? At some level you can't prove it. And so, it's not any deeper than that. But I think the deeper point is: Who gives a crap?"

DCD > Webscale | San Francisco

15th Annual

Global Content Partner
Uptime Institute

Focus Day on **ENERGY SMART** | June 25

WHAT'S NEXT FOR THE HYPERSCALE DIGITAL INFRASTRUCTURE ECOSYSTEM?

June 26 2018 // San Francisco Marriott Marquis

Limited sponsor and exhibitor opportunities - contact chris.hugall@datacenterdynamics.com

Lead Sponsors

CORNING | DELTA | RITTAL | Schneider Electric | SUB-ZERO ENGINEERING

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

@DCDConverged #DCDWebscale | Datacenter Dynamics | DCD Global Discussions

Love at first sight

AI and DCIM were made for each other, says Nlyte CEO Doug Sabella in a conversation with *Max Smolaks*

Max Smolaks
News Editor

Artificial intelligence is a tempting proposition, offering a vision of algorithms trawling through petabytes of data, making connections and assisting their human masters in their decision-making, perhaps even making some decisions on their own, without any oversight. It sounds like a perfect tool to help manage complex data center systems, since modern facilities are chock-full of probes and sensors, and already produce too much information to be processed by people. To put it simply, we could really use some help.

One of the companies working in this direction is American software vendor Nlyte, which is collaborating with an unnamed partner to integrate AI capabilities into its data center infrastructure management (DCIM) products in either 2018 or 2019. Doug Sabella, who has been CEO of Nlyte since the middle of 2012, thinks this will considerably expand the functionality of DCIM software into areas like workload management and cost analysis. Data centers of the future will require both, but there are serious technical hurdles to overcome.

"The simple things are around preventive maintenance," he told *DCD*. "But moving beyond predictive things, you're really getting into workloads, and managing workloads. Think about it in terms of application performance management: today, you select where you're going to place a workload based on a finite set of data. Do I put it in the public cloud, or in my private cloud? What are the attributes that help determine the location and infrastructure?

"There's a whole set of critical information that's not included in that determination, but from an AI standpoint, you can contribute into it to actually reduce your workloads and optimize your workloads and lower the risk of workload failure, because of where you're putting things. There's a whole set of AI play here that we see and our partner sees, that we're working with on this, that is going to have a big impact."

Other potential applications of AI in DCIM include identification of critical alerts, continuous optimization of various types of cooling equipment, and automated management of simple edge computing environments. To make this a reality, DCIM will require additional features that were historically associated with big data analytics, rather than data center software.

"We produce more and more data, but we have to become more intelligent about how we manage it," Sabella said. "Think in terms of high velocity, real-time data streaming, think [Apache] Kafka – those are the kind of things we will be introducing, so people can stream data into a Hadoop environment, for example, and create analytics they want."

In many ways, AI is a natural extension of what the DCIM vendors have been trying to achieve since the very beginning - to connect physical properties of the data center environment to the performance of the workload, and help eliminate the divide between facilities teams, responsible for things like space, power and cooling, and IT teams, responsible for servers and applications.

Sabella suggested this divide is to blame for the fact that it is currently easier to establish the cost of a workload hosted in a public cloud environment than a workload in a large in-house enterprise data center: "Because of the wall between facilities and IT, they have not invested the money or the time to actually understand [their infrastructure]."

"With DCIM deployment, there are a lot of companies that have CFOs and CIOs saying 'we don't need that, because everything is going to the cloud.' Guys, you spent billions of dollars in your data centers, and you don't even know what you have.

"The awakening that we've seen and that we've really been very successful on, is all these large enterprises realizing that they need to start understanding what's going on in their data centers if they're truly going to build a hybrid cloud.

"Without the information we provide, you can't actually do the analytics to understand what the cost of the workload is."

From a broader perspective, DCIM is changing the power dynamics within data centers, and that's not to everyone's liking. "It's not a peaceful process," Sabella told *DCD*, adding that embracing new tools like DCIM and automation through AI might require a generational change. "At least in the US, a lot of the traditional facilities folk, they are baby boomers, so they are actually moving into retirement. You have young and smart facilities people who don't think of themselves as facilities people, they are engineers. This will be driven by millennials," he said of the adoption.

In terms of business goals, the next challenge for Nlyte is to bring its software to the masses. Historically, the company had focused on very large customers – but Sabella promised a renewed effort to target SMBs, especially through DCIM as-a-service, which Nlyte launched way back in 2014. He said he hoped to grow this part of Nlyte to eventually be responsible for half of all revenue.

"At some level, it's as challenging to close a sale of 1,000+ racks as 100+ racks - it takes as much time," he said. "But, on the same token, we have a great play for small business - a lot of our SaaS customers today are small businesses."

The PlayStation Supercomputer

Sebastian Moss
Reporter

The chip that started in the PlayStation 3 ended up in the world's fastest supercomputer, and heralded a heterogeneous computing future. *Sebastian Moss* reports

In 1999, Sony was ascendant. Its video gaming business was its most profitable division, the PlayStation 2 was the world's most successful console, and hopes were high that the successor would prove even more popular.

To achieve this, the designers believed they would have to make the PlayStation 3 the most powerful console possible, with its own custom microprocessor.

Sony turned to IBM and realized that "there was a potential synergy between the consumer oriented technology that Sony worked on, and the more business and data center oriented technology that IBM worked on," Peter Hofstee, distinguished researcher at IBM, told *DCD*.

A year of discussions culminated in a grand vision to develop a radically new chip architecture that the two companies hoped would become integral to consumer electronics, edge devices and data centers.

This innovation would be known as the Cell Broadband Engine Architecture, and would help commercialize the concept of heterogeneous computing, which relies on more than one kind of processor or core.

Sony and IBM also recruited Toshiba, which previously worked on the PS2's Emotion Engine CPU, and together formed the STI Alliance, with a budget of approximately $400 million.

With the Cell chip set to launch as part of the PS3, Sony's decisions were handled by Ken Kutaragi, the 'father of the PlayStation' who was at the time widely expected to become the next Sony president.

"This was research on a development schedule," Hofstee said. "We did over 600 patents leading into this thing. There were some very significant technical challenges along the way. We would get the architecture team and the key leaders from Sony and Toshiba every morning for weeks, sometimes months on end and we would hash through these things in a very collegial way."

The principal idea of the chip was to combine a general-purpose Power Architecture core with streamlined coprocessing elements. "One proposal, which maybe does not seem as exciting today, but was pretty aggressive at the time, was a four-core Power proposal. But Kutaragi decided that that was actually not aggressive enough," Hofstee said.

"We ended up with a chip that had nine cores in total - one Power core, and eight Synergistic Processor Elements," he added. Hofstee was the chief architect of the SPEs, which consisted of a Synergistic Processing Unit (SPU) and a Memory Flow Controller (which had direct memory access, a memory management unit and a bus interface).

Initial plans called for just 64KB of memory for every SPE, but the software team realized very early on that this amount would not suffice. Memory was increased to 128KB, and still, that wasn't enough. "I said fine by me, but it is a zero sum game, so you can have eight SPEs with 128KB, or six SPEs with 256KB.

"That was when we briefly went from eight to six, then Kutaragi saw that and he said 'what is this, I didn't approve that? It has to be eight.'" Later in the meeting, when

asked why, he replied: "Eight is a beautiful number."

"I will never forget that," Hofstee told *DCD*.

The PlayStation 3 launched, after some delays, in late 2006 with a 90nm Cell chip inside it. A year later, IBM released the BladeCenter QS21, an energy-efficient server with 1.05 giga–floating point operations per second (GFlops) per watt, and a peak performance of approximately 460 GFlops. Elsewhere, the Cell chip made its way into Mercury Computer Systems servers, Toshiba TVs, Hitachi Medical scanners and Leadtek PCIe cards.

"There were applications for technology in financial services, such as high performance computing for Wall Street, as well as applications for security, oil and gas sectors, gaming and entertainment, and bioinformatics," Professor David Bader told *DCD*. "The design was very forward-looking, and the performance improvement was tremendous."

IBM's Roadrunner supercomputer

Bader directed the first Sony-Toshiba-IBM Center of Competence for the Cell Processor, at the Georgia Institute of Technology, which offered training and insights into how to use accelerators and multi-core technologies in data centers.

"This was a chip that was purpose-built for high-end computing as well as gaming systems, that spans the whole gamut of low-end to high-end. It also created a groundswell in the applications space, with new understanding that applications across many sectors, from healthcare to finance to security, could all take advantage of accelerators."

> *"We were just a tad too early," Peter Hofstee*

Indeed, IBM did target the high-end computing market: in 2008 it built the world's most powerful supercomputer to model the decay of the US nuclear arsenal for the Los Alamos National Laboratory. Roadrunner, a $100m system with 12,960 Cell chips (now 60nm with added memory) and 6,480 AMD Opteron dual-core processors, became the world's first machine to sustain 1.0 petaflops using the Linpack benchmark, roughly double the performance of the next best system.

"In retrospect, if you look at the Top500 chart [a list of the world's fastest supercomputers], we made quite a step up," Hofstee said. "You expect a certain growth rate on the list, but you can actually see that the Cell chip jumped off that quite a bit - we held the record a little bit longer than it's typically held."

The system was also incredibly energy-efficient, another benefit of the Cell. Hofstee explained: "We went with this philosophy: instead of turning something off when it's not needed, turning it on only when it is needed."

The team would run programs and watch which pieces of the chip turned on. "If we would see a floating point unit turn on when the program wasn't supposed to have a floating point workload, we would know that something's not quite right. This is an area where people have made more progress since then, but I really think we were the first to adopt this philosophy."

Bader concurred: "Up to this point in time, when you needed memory, it would be like going to the store and buying a whole case of food and bringing it back, even if you needed just one can. With the Cell, for the first time, you could take a shopping list, go to the store, and fill up your cart just with the individual cans that you may need, and then come home. That was game-changing."

But, despite the efficiency gains and its high performance, sales of Cell-powered servers were limited. "It was early technology that lacked the set of programmers and programming tools that were readily available for the processor," Bader said

"The programming effort to use Cell required skilled and specially-trained programmers. The real challenge is to produce a radically new architecture that's more efficient and capable while at the same time having the software co-design with programmers who are ready to take advantage of it."

Hofstee added: "It was fairly difficult for ▶

$4.7bn

The amount of money Sony lost on the PlayStation 3 by Q2 2009.

Due to the Cell chip, its then-cutting edge Blu-ray drive and several other costly features, the PS3 lost Sony money every time it was sold.

It took another year for the price of manufacturing to break even.

The PS4, on the other hand, featured a PC-like architecture with traditional chips. It has been sold at a profit since launch and is currently the world's most popular games console.

commercial users to adopt Cell if you knew that you had to write your software in such a way that it would only run on Cell. People have to really take a big leap of faith if you're asking them to do that.

"We were just a tad too early. There were a number of near misses. For example, OpenCL, the open standard for parallel programming of heterogeneous systems, came just too late."

Hofstee also believes that one could have taken a compiler from supercomputer manufacturer Cray "and retargeted it to a Cell chip, basically by thinking of the local store as a large vector register file." At the time, he didn't think there were any open Cray compilers around, but later discovered the Department of Energy had open-sourced one.

"I wish I had known," he said. "That would have made it more palatable for a community that just couldn't make a decision to write their future on the architecture."

Meanwhile, as server sales struggled, Sony's core consumer product, the PS3, was also in trouble. Initially sold at a significant loss, the console was costing the company billions, its sales lagged behind Microsoft's Xbox 360, and developers found the unique architecture difficult to make games for. By 2007, Kutaragi had relinquished active management of the business he had built.

"There were plans to have a second generation Cell which was supposed to have 32 SPEs and four Power processing elements, and scale it up," Professor Gaurav Khanna told *DCD*.

"What IBM mentioned is that Sony pulled the funding and there were practical reasons for that - Sony had lost a lot of money, they made a lot of investment, the PS3 didn't quite take traction as quickly as they'd liked.

The US Air Force's Condor Cluster

83m
The number of PS3s sold worldwide

Without Sony funding the next generation, it just died. That's the hearsay I have heard."

Khanna, a black hole astrophysicist and associate director of the Center for Scientific Computing at the University of Massachusetts Dartmouth, was one of those attracted by the Cell chip's promise.

His subject of study, while fascinating, is an area "that generally is underfunded simply because there aren't necessarily very direct practical applications for the research we do, it's basic research, fundamental physics."

It does, however, require significant computing power to simulate black holes. Short on funding, Khanna had an idea - to use a cluster of PS3s. Not only did they offer a Cell chip at a low price, "Sony made it even easier by offering the possibility of putting Linux on it, and the entire scientific community uses Linux as its bread and butter. It just had to happen, it was such an obvious thing to try to do."

After contacting Sony because he "couldn't really go to the National Science Foundation and say 'give me a couple of thousand dollars for gaming,'" Khanna was given eight consoles by the Japanese corporation and created the Gravity Grid cluster.

"We were able to take those eight machines and show some nice parallel performance, the code soon produced really good results - we even published a paper or two in the first year of this system being built. Then our Dean got interested and he gave us money to double it to sixteen."

Khanna's PS3 cluster was the first to lead to published scientific results, although other researchers soon began to build their own. "There are 20 or so papers we've published using the cluster at this point," he said.

Professor Khanna's Reefer Gravity Grid cluster of PS3s

And those papers were made possible with an unusually affordable set up. "We did some nice estimates on how much it would have cost with a conventional system, it was easily tenfold. Essentially, one $400 PlayStation was equivalent in performance to 16-32 cores of your several-thousand-dollar Xeon-class x86 server."

As for energy efficiency, a "PlayStation on full load would consume about 100 watts of power; and at that time a Xeon-class dual socket type server would easily take about a kilowatt. It was quite remarkable."

Khanna's work was not just being watched by other astrophysicists - it also caught the eye of the Air Force Research Laboratory (AFRL), which was eager to explore low power consumption yet high performance chips, especially for use in autonomous drones. "They were quite impressed with the performance of a PlayStation, especially its power efficiency," he said. AFRL had a lot more resources, and set up a 300 PS3 cluster, which it soon expanded to 1,760 consoles.

The 'Condor Cluster' became the 33rd most powerful supercomputer in the world and 'the fastest interactive computer in the US Department of Defense,' capable of 500 trillion floating point operations per second.

AFRL's director of high power computing, Mark Barnell, said at the time that it cost $2 million, about 5-10 percent of an equivalent enterprise HPC system, and it consumed just 10 percent of the power that one would expect.

Khanna was called in to explain what he had learned from his cluster, and began to help AFRL with their system - in return for access to the hardware.

By late 2012, the Air Force decided to repurpose the floor space, and, with Khanna the Condor's most active user, he was offered all of the 1,760 systems. After calculating how many the campus could support, he took 400.

"One challenge we had is that they had substantial power and cooling requirements.

The bill we had to renovate a lab to house 400 units was more than $500,000," he said.

"We decided that we had to innovate again, to come up with a crazy idea again." Khanna looked to markets with large volumes, and found reefer containers - refrigerated units for shipping perishable produce over long distances.

"For $30,000 we had this relatively huge freezer unit that basically gives us cooling capacity of 80,000 British Thermal Units per hour, which is exactly what we needed," he said.

"At a tenth of the cost, we had a mini data center. That's what we've had now for a couple of years, and the PlayStations have been happily running. We've had no trouble, it's been able to maintain the temperature in a very stable way."

Khanna's decade-old system is one of the few remaining implementations of the Cell chip, along with Sony's PS3 game streaming service, PS Now, which is reported to use racks featuring eight custom console units each (Sony declined to comment). But the impact of the Cell can still be felt elsewhere.

"The heterogeneous computing aspect of that has very much survived," Hofstee said. "If you look at the supercomputers that IBM is starting to deliver now, the pre-exascale Summit and Sierra supercomputers, it's Power plus an Nvidia GPU, but I would say it's very much in the same spirit of Power plus the SPEs that we had on Cell."

He continued: "I'm very happy with the impact we've had on the overall arc of computing architecture."

As for Bader, he found that the ideas and knowledge produced in the Cell Center made an impact on the broader computing community. Meanwhile, the lessons about software tools and support "really helped provide a commercial success to emerging accelerators such as GPUs, and multi-core technologies."

He believes that the Cell was "a catalyst, an inflection point between pure CPU computing and moving to accelerators, and understanding the ecosystem required for accelerators, as well as bringing parallel computing to the mainstream.

"Up to that point, every commercial application was single-threaded and the Cell entered the market at a time where developers made the shift to understanding multi-core."

Khanna added: "Now everyone's getting back to it, a smartphone System-on-a-Chip has CPU and GPU on the same die. In many ways one could say that the legacy of the Cell is perhaps the most common processor out there.

"I feel bad, really pained that IBM and Sony had this game almost a decade before everybody else got on top of it and sort of dropped it, it's really painful to see. Maybe the Cell was just too ahead of its time, and now is the time."

The physicist, meanwhile, was left inspired by the process of turning a consumer product into a data center.

The most recent prototype he has built consists of 32 Nvidia Shield tablets. "Essentially it is a 16 teraflop system that only uses 300 watts of power, which is almost fivefold more efficient than the most power-efficient supercomputer on Green500, the list of the most efficient HPC systems in the world."

He told *DCD*: "We need to misuse more consumer electronics to do science. It's cheaper for us, it's more efficient. Why don't we just piggyback on the innovation that's happening in consumer electronics?"

The distributed supercomputer

Before the Cell chip found its way into the Roadrunner supercomputer and broke the one petaflop mark on the Linpack benchmark, it had already helped make history.

Folding@home, a Stanford University distributed computing project, became the first platform to reach a petaflop in performance on a single precision float calculation - with 74 percent of the system's power coming from PS3s.

Over its five and a half year lifetime on the games console, more than 15 million PS3s contributed over 100 million hours of computing to Folding@home, used to simulate molecular dynamics, including for protein folding and computational drug design.

"The PS3 system was a game changer for Folding@home, as it opened the door for new methods and new processors, eventually also leading to the use of GPUs," research lead Vijay Pande said. "We have had numerous successes in recent years, [with simulations leading] to a new strategy to fight Alzheimer's Disease."

While it no longer runs on the PS3, with its early help Folding@home has grown to become one of the world's fastest computing systems, hitting around 135 petaflops in January 2018.

A new Australian giant

Nick Parfitt
Lead Analyst
Asia Pacific

AirTrunk has brought hyperscale data centers to Australia. *Nick Parfitt* visited the Sydney facility and spoke to AirTrunk's CEO, Robin Khuda

AirTrunk's data center in Western Sydney, once fully built out, will become the largest data center in Australia and one of the largest in the Asia Pacific. It is accompanied by another facility in Melbourne, opened in late 2017, with both projects representing the recent 'hyperscale' wave in Australia, and all that means for standards of design and operation, business strategy and branding.

The two sites will offer total power capacities of more than 80MW and 50MW respectively, and on that basis are the largest facilities in their states.

Approaching AirTrunk Sydney down the shallow incline from Huntingwood Drive, our first view is of the AirTrunk company colors, blue and white, contrasting with the black and grey of power modules and cooling equipment located down the side and on the roof of the building. Just beyond, construction machinery rakes the surface of the ground where the second phase will be built.

An immense on-site power setup sits at the front of the building and connects the data center to the grid.

This is AirTrunk Sydney, six months on, and the company's founder and CEO, Robin Khuda, is here to tell us how satisfied he is with the work: "So far we are very happy with our progress. We delivered super-fast to match customer demand."

The Sydney and Melbourne facilities embody the speed and size of the hyperscale data center. The review of the first phase of construction accelerated the process by which they were delivered to market: the initial 20MW phase at Huntingwood took 44 weeks, by Melbourne this had been reduced to 42 weeks, while the next 10MW at Huntingwood, on the land towards Sydney's Western Motorway, will take 35 weeks.

US$400m
The amount AirTrunk raised in 2017

And then there is scale. AirTrunk has invested around AU$50m in its own power infrastructure, so each site can scale up and grow with customers. The design is modular, with standardized components for power, cooling and security, so the company's deployment will remain consistent.

"We will have fully autonomous data halls

Ensure your Invaluable is backed up by our Reliable

WE DESIGN • WE BUILD • WE DELIVER

here," Khuda explained. "They are like Lego structures with Lego blocks for the utilities on the outside."

The environmental footprint of hyperscale facilities is important, but managing this in Australia is complicated. For now, AirTrunk is addressing it onsite through efficiencies in design and operation, rather than through purchasing renewable energy.

"Sustainability is important for us and for our customers, but we have some challenges in Australia, both in the cost of renewables, and that they still have to go through the grid," Khuda said. "You can't buy them direct. We have a sustainability roadmap over time. Not just 1MW or 2MW for marketing purposes, we are looking at the whole thing."

The facility has removed a number of inefficiencies through the use of automation and AI. Electricity is delivered to the rack from the sub-station in just three stages - a fraction of the number of stages in a legacy system.

Thanks to these measures, AirTrunk Sydney can claim a PUE as low as 1.15 at 70 percent load, while Melbourne reaches 1.13 on an annualized basis. The modular design of the data center means that PUE can be measured individually for each customer. Modularization also means the space can be customized to individual customers' requirements. Once completed, there will be over 30 data halls within AirTrunk Sydney.

The CRAC units and power infrastructure are located outside the data hall where they can be maintained and repaired away from the server racks. As Khuda points out, this increases reliability by minimizing human presence: "90 percent of downtime is human error."

Following the modular principle, the power train units and UPS are factory-fitted before they arrive on site, so they are ready to be connected into the data hall late in the fit-out program. "The first years of the business focused on the R&D phase so we can build cheaper, faster and scalable," Khuda said.

It may sound obvious, but to build based on demand, you have to understand and follow client needs, and AirTrunk has put considerable efforts into developing partnerships and supply chains.

The company chose Australia at the behest of an unnamed anchor tenant, but Khuda expects continued growth to come from other sources, including consumer services and the migration of legacy IT: "I can see strong growth in IT services continuing to move to the public cloud - the content providers, the telcos as they launch 5G, Netflix – which accounts for more than 40 percent of the data traffic in Australia. Then there is AI and IoT so I think we are in a very exciting time at an early time of growth."

AirTrunk has plans to expand, based on demand, into Tier 1 markets where requirements are strong, with Japan, Singapore, Hong Kong and India mentioned as possible locations: "China is a different market and there are a lot of regulatory requirements. Then there are the emerging markets - Indonesia, Philippines, Vietnam - but we are not seeing enough demand for hyperscale there at this stage."

Our visit to AirTrunk Sydney suggests the future of the service data center in Australia is bright. More than that, the model for design and operation based on efficiency, scalability and modularity is convincing, so long as it is supported by a service model which can deliver facilities quickly, and which balances the need for standardization with the ability to meet individual requirements. ●

THOMSON™
POWER SYSTEMS

SMART SWITCHGEAR &
DATACENTER SOLUTIONS

WWW.THOMSONPS.COM 1.888.888.0110

Now that's what I call television

Photo by Alessio Jacona

> *"Move fast and break things"*
> *- probably not the best motto for a business*

Watching Mark Zuckerberg being grilled in front of the House Energy and Commerce Committee was a strangely satisfying experience. Here was a man who, mostly by chance, established one of the world's largest Internet companies. A man who, by age 25, got a film about his life, based on a book about his life. A man who is worth more than $60 billion, despite earning a $1 salary.

He was visibly uncomfortable. He was sweating. And sitting on a special four-inch pillow to make himself appear taller. Mark Zuckerberg was there, on the screen of my TV, giving meaningless answers like a living embodiment of Facebook's 30,000-word Terms and Conditions. How did this happen?

Some people like to note that Facebook is not in the business of connecting you with your friends, it's in the business of advertising. On closer examination, that's not very accurate. Facebook is in the business of profiling - reducing a person to a list of categories (as many as 29,000 categories, according to *ProPublica*) and selling access to this database. Your likes and dislikes, your relationship status, your religious beliefs, and yes, your political opinions are its business.

Along the way towards reaching two billion users, Facebook continued to experiment with this formula. "Move fast and break things." Not tied down by regulations, or even morals - remember the one where they made people sad on purpose? In case of Cambridge Analytica, one of these experiments overpowered the guards, escaped the laboratory and went on a rampage.

During the hearing, most commentary on Twitter seemed to revolve around the generation gap: look at all these grandparents, struggling to understand technology. But really, what surprised me was how many of them really knew their stuff. My personal heroes are Congresswoman Debbie Dingell, who highlighted the scale of data collection Facebook does on third-party websites through the ubiquitous 'Like' button, and Senator Dick Durbin, who asked some tough questions about the new 'Messenger Kids' app, aimed at children as young as six.

Even the extent of Facebook's lobbying in Washington was touched upon: the company's lobbying budget totaled $11,510,000 last year, and I've got a hunch it will see a considerable increase in 2018.

Zuckerberg's responses ranged from vague to misleading, to downright inane. Being a sort of libertarian saint, he would never accept the idea that any level of government oversight was necessary. He endlessly referred to the many ways the users can control their information - the whole debacle was on us, you see. He would also repeatedly mention a non-existent AI that could somehow miraculously identify "bad content," even though he had difficulties explaining what kind of "bad" he really meant.

At the end of the day, Zuck's showdown against the House Committee is unlikely to change anything: the only thing that scares Facebook is a mass exodus of users. But the hearing managed to last ten excruciating hours across two days. Making one of the richest people in the world this uncomfortable, for this long, is an achievement in itself. ●

Max Smolaks
News Editor

STULZ

CLIMATE. CUSTOMIZED.

Best server chilled

CyberAir 3PRO from STULZ stands for maximum cooling capacity with minimum footprint. Besides ultimate reliability and large savings potential CyberAir 3PRO offers the highest level of adaptability due to a wide range of systems, variants and options.

www.stulz.de/en/cyberair-3-dx

Data Center Infrastructure Total Solutions

Micro Data Center | Modular Data Center | Containerized Data Center | Power Container

The power behind competitiveness

Delta InfraSuite - Data Center Infrastructure Solutions

With current trends in big data, the IP traffic of enterprises is expected to grow rapidly over the next few years. To meet the changes in the IT business environment, Delta's InfraSuite - Data Center Infrastructure Solutions - offers a comprehensive, modular and highly integrated portfolio to support the creation of high-performance datacenters. As a global leader in thermal and power management solutions, Delta has further strengthened its leading position in data center infrastructure with a complete offering of AC or DC power, cooling systems, monitoring platforms, from micro and modular to containerized solutions.

More information and success stories about Delta InfraSuite
www.deltapowersolutions.com

Contact our local consultant for knowing more on data center infrastructure solutions.

EMEA: dc.et@deltaww.com | dc.mea@deltaww.com | dc.russia@deltaww.com
APAC: dc.china@deltaww.com | dc.hongkong@deltaww.com | dc.india@deltaww.com
dc.sea@deltaww.com | dc.south.korea@deltaww.com | dc.taiwan@deltaww.com

DELTA
Smarter. Greener. Together.