

NEWSFEED

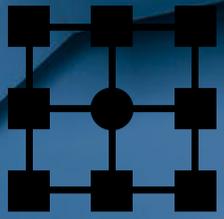
DATA CORE SHATTERS
WORLD RECORD

NEWSFEED

WEE ARCHIE, THE TINY
SUPERCOMPUTER

TECH ZONE

KEY FEATURES FOR
ENTERPRISE CLASS SSDS



HPC REVIEW

THE REFERENCE MEDIA IN HIGH-PERFORMANCE IT SOLUTIONS

2016

SUPERCOMPUTING PREDICTIONS

Industry leaders chart
the future of HPC

**ANALYTIC
BUSINESS INTELLIGENCE
CLOUD
BIG DATA
INNOVATION
INTERNET OF OBJECTS
OPEN SOURCE
STORAGE**

GIGABYTE™

A LEADING GPU SERVER, AT YOUR FINGERTIPS

G190-H44 GPU Server

- Supports 4 x double slot cards
- Supports the Intel® Xeon Phi™ cards
- 4 x Storage bays and 2 x expansion card slots
- Redundant 2000W power supply, 80 PLUS Platinum certified



Available at:

MICROTRONICA
A DIVISION OF ARROW

b2b.gigabyte.com

Supports the
Intel® Xeon® Processor E5-2600 V3
Intel Inside®. Powerful Solutions Outside.



Intel, the Intel logo, Xeon, and Xeon Inside are trademarks or registered trademarks of Intel Corporation in the U.S. and/or other countries.

HAPPY 2016 SUPERCOMPUTING YEAR!

WELCOME TO HPC REVIEW GLOBAL EDITION ISSUE 11!

This issue is chock full of predictions coming from nearly 20 industry leaders of nearly every field affecting IT, super IT and SuperComputing. As fast moving as 2015 may seem nowadays, 2016 promises to move even faster, with both emerging technologies and the pursuit of well grounded directions taken by various business segments composing the IT and HPC landscape.

USE, PERFORM AND PAY AS YOU GROW

Think global, act global is the promise technology allows businesses in every sector to keep, even and foremost for the smallest but most agile digital natives and startups. Gone are the days where several million dollars were needed to start a business, let alone be profitable. With the pay as you grow model of the Cloud extending to the storage and the compute needs, never has performance and technical availability been so accessible and scalable !

EMERGING AND ESTABLISHED TECHNOLOGIES UNITE

With emerging technologies already established around the Software Defined /X banner and even more disruptive business models remaining to be invented, 2016 is the kick off of a new era that will allow the industry to reach the next significant step by 2020: Exabyte Class Super Computing. Which will usher a new era of super use cases for academic, research, industry and enterprise. Welcome to a fascinating and never stopping world ! Happy reading.

HAPPY READING!

2016

SUPERCOMPUTING PREDICTIONS

 **NEWSFEED**

**IDC IT FORECASTS
FOR 2016 AND BEYOND**

**DATA CORE SMASHES WORLD'S
STORAGE PERFORMANCE RECORD**

**INTERVIEW: GEORGE TEIXEIRA,
CEO DATA CORE SOFTWARE**

**WEE ARCHIE, THE TINY
SUPERCOMPUTER**

 **HOW TO**

**CONTROLLING THE
TIDE OF WEATHER AND
CLIMATE DATA WITH
TIERED STORAGE**

 **VIEWPOINT**

**HOW TECHNOLOGY WILL
DRIVE THE NEXT WAVE OF
CIRCULAR AND SERVICE-
BASED ECONOMICS**

 **TECHZONE**

**THE KEY FEATURES
OF ENTERPRISE-CLASS
SSDS**

EDITOR'S NOTE
OPINION

SOURCES
THE HPC OBSERVATORY

IDC PREDICTS THE EMERGENCE OF “THE DX ECONOMY”

**FRANK GENS**SENIOR VP,
CHIEF ANALYST AT IDC

In a Critical Period of Widespread Digital Transformation and Massive Scale Up of 3rd Platform Technologies in Every Industry, International Data Corporation (IDC) announced its worldwide information technology (IT) industry predictions for 2016 and beyond.

IDENTIFYING THE 3RD PLATFORM

In 2007, IDC identified the 3rd Platform – built on the technology pillars of mobile computing, cloud services, big data and analytics, and social networking – as the foundation for the IT industry’s future innovation and growth. Since then, the rapid adoption of 3rd Platform technologies has moved into the “Innovation Stage” – an explosion of innovation and transformation on top of the 3rd Platform’s foundation. IDC predicts that this phase will accelerate over the next three to five years as enterprises commit to digital transformation (DX) on a massive scale, leading to the emergence of the “DX economy.”

“The disruptive impact of digital transformation is about to be felt in every industry as enterprises ‘flip the switch’ and massively scale up their DX initiatives to secure a leadership role in the DX economy,” said Frank Gens, Senior Vice President and Chief Analyst at IDC. “In the next two years, two-thirds of

Global 2000 CEOs will put DX at the center of their growth and profitability strategies. By the end of this decade, IDC predicts that the percentage of enterprises with advanced DX strategies and implementations will more than double.”

SCALE-UP OF DIGITAL BUSINESSES

IDC predicts that the scale-up of digital business strategies will drive more than half of enterprise IT spending within the next 24 months, rising to 60% by 2020. Mastery of 3rd Platform technologies will be table stakes for successfully executing DX business initiatives and “Cloud First” will become the new mantra for enterprise IT. Virtually none of the other 3rd Platform technologies or major DX initiatives is possible in scaled-up implementations without the Cloud as the foundation. By 2020, IDC predicts that enterprise spending on cloud services, the hardware and software to support cloud services, and the services for implementing and managing cloud services will exceed \$500 billion, more than three times what it is today.

The DX economy – operating at scale – will be driven primarily by code. Enterprises’ ability to grow and compete will increasingly depend on their digital “innovation capacity”: the size and talent of their software develop-

ment teams. In this regard, every company will increasingly be a software company. By 2018, IDC predicts that enterprises pursuing DX initiatives will more than double the size of their developer resources, focusing those developers almost entirely on DX initiatives.

Data will be equally important to the DX economy – without large volumes of quality data “fueling” innovation, the process will stall. IDC believes that success in the DX economy will depend on the ability to build robust “data pipelines” that flow both in and out of the enterprise. By 2018, IDC predicts that the flow of external data into organizations with advanced DX strategies will increase by as much as a factor of five while high-end DX performers will increase their “data out” by 500-fold or more.

INTERNET OF THINGS: 22 BILLION DEVICES BY 2018

The Internet of Things (IoT) will be one of the most fertile areas for DX. By 2018, IDC predicts the number devices on the IoT will more than double to 22 billion and drive the development of 200,000 new apps and solutions that take advantage of them. These devices and solutions have the potential to redefine competitive advantage in virtually every industry. IDC predicts the most active IoT development will cluster around the manufacturing, transportation, retail, and healthcare industries.

Cognitive systems are already providing critical assistance to organizations dealing with the tsunami of data. Spending on cognitive software platforms is forecast surpass \$1 billion this year and will scale up dramatically over the next three years. IDC predicts that more 50% of developer teams – up from less than 1% today - will embed cognitive services into their apps by 2018 to leverage their data pipelines and to improve customer personalization.

In the expanding DX economy, industry cloud platform and community participation will become increasingly vital to scale up digital supply chains and digital distribu-

tion channels by as much as 100- to 1000-fold. By 2018, IDC predicts that more than 50% of large enterprises – and more than 80% of enterprises with advanced DX strategies – will create and/or partner with industry cloud platforms to scale up their digital supply and distribution networks. IDC predicts that as the number of industry clouds will reach 500 or more by 2018, up from today’s 100-plus.

SUPPORTING UP TO 10 000 MORE CUSTOMERS

As the DX economy compresses pricing in many sectors, IDC predicts that 60% of B2B and 80% of B2C enterprises will overhaul their “digital front doors” - and the customer engagement systems behind them - to support 1,000 to 10,000 times more customers and customer touchpoints than they do today. At the same time, they will have to deliver dramatically more personalized customer service. IDC predicts that Customer/Intimacy @Scale will be the biggest, most complex enterprise-wide DX initiative that organizations will have to face, requiring a fundamental cultural and operational transformation.

Finally, the DX economy will take its toll on the IT industry itself. By 2020, IDC predicts that nearly a third of today’s IT suppliers will be acquired, merged, downsized, or significantly repositioned. In this environment, enterprises will have to constantly monitor and assess the solutions offered by their suppliers and partners and be prepared to realign these relationships as needed.

An on-demand replay of IDC’s Worldwide IT Industry Predictions Web conference will also be available, along with more than 30 other FutureScape presentations covering topics including Digital Transformation (DX), the CIO Agenda, the Internet of Things (IoT), Smart Cities, Big Data and Analytics, Cloud Computing, and more. To register for any of these presentations, please go to <http://www.idc.com/idcfuturescapes2016>.



IDC'S IT FORECASTS FOR 2016 AND BEYOND

IDC Asia/Pacific releases its IT Services Predictions for 2016 and beyond, starting with Cloud First and Digital Transformation expectations



DC Asia/Pacific recently released its latest IT Services Predictions for 2016 and beyond for the Asia Pacific region (excluding Japan) highlighting the impact on services of Cloud First and Digital Transformation expectations in the next 1-3 years. The presentation titled IDC FutureScape: Worldwide IT Services Predictions 2016 – APeJ implications provides 10 predictions for services in 2016 and the implications for IT suppliers and buyers.

While 2016 is shaping up to be the banner year for Digital Transformation (DX) to take root in many Asia/Pacific organizations, IDC expects Innovation Accelerators like IoT, cognitive computing, robotics and 3D printing to accelerate this transformation in years to come.

According to IDC's latest services survey, it has shown that majority of APeJ enterprises are claiming to have DX projects in some form or another. It is becoming a competitive re-

quirement and the source of a massive wave of new investments in services to support changes to business operations and customer experiences.

“Getting on board the DX train is going to be fraught with challenges, starting with the ability to define DX. Most early stage projects involve experience transformation but little attention is being paid on the other dimensions like information governance and work-source, which often leads to disillusionment on the benefits of DX”, says Linus Lai, the Research Director with IDC's Asia/Pacific Services Research Group. Lai continues, “As part of enabling enterprises on its DX journey, service providers need to emphasize their ability to bring a full suite of services capabilities that includes integration of cloud, cognitive, mobile and analytics delivered 'as a service'. In addition, industry expertise can offer functional insights which are also needed.”

Progressive adoption of 3rd Platform technologies in the region is not only driving de-

CATHY HUANG, RESEARCH MANAGER, SERVICES AND CLOUD RESEARCH GROUP STATES, **“THIS YEAR’S SERVICES PREDICTIONS HIGHLIGHTS “CLOUD-FIRST”,** SUGGESTING 33% OF ASIA/PACIFIC ENTERPRISES WILL PURSUE A “CLOUD FIRST” STRATEGY”.

mand for related services, but is also fundamentally changing the way they interact with their broader ecosystem.

Cathy Huang, Research manager, Services and Cloud Research Group states, “This year’s Services predictions highlights “Cloud-first”, suggesting 33% of Asia/Pacific Enterprises will pursue a “Cloud first” strategy”.

“The switch to cloud-first/cloud-only points to two opportunities in the APEJ services market. Enterprises are looking for differentiated capabilities or cost competitiveness based on cloud-based offerings. In addition, it also means sourcing strategies in the past may be incompatible as organizations leverage a cloud-first strategy”, adds Ms. Huang.

In addition, this year’s Services predictions deep-dives into the impact of 3rd Platform technologies and innovation accelerators such as IoT and Next-Gen Security. As a result, more focus on governance between the IT organization and the lines of business (LOB) is demanded to ensure risk, supplier choices, fail-over, innovation, and agile development is understood between both parties, and the value of integration, management and security is brought to the front.

IDC reports that the rest of the predictions for the fast-evolving Services market for the next three to five years include the following:

#3 CLOUD FIRST: 33% of Asia/Pacific Enterprises Will Pursue a “Cloud first” strategy and the share of overall IT budgets dedicated to cloud services will increase to 26% by 2016.

#4 XAAS SUBSTITUTES FOR MANAGED SERVICES: By 2016, 50-55% of traditional outsourcing managed services will have a cloud delivery model.

#5 DX: By 2016, 65% of APEJ enterprise strategies will include a significant “digital transformation” component.



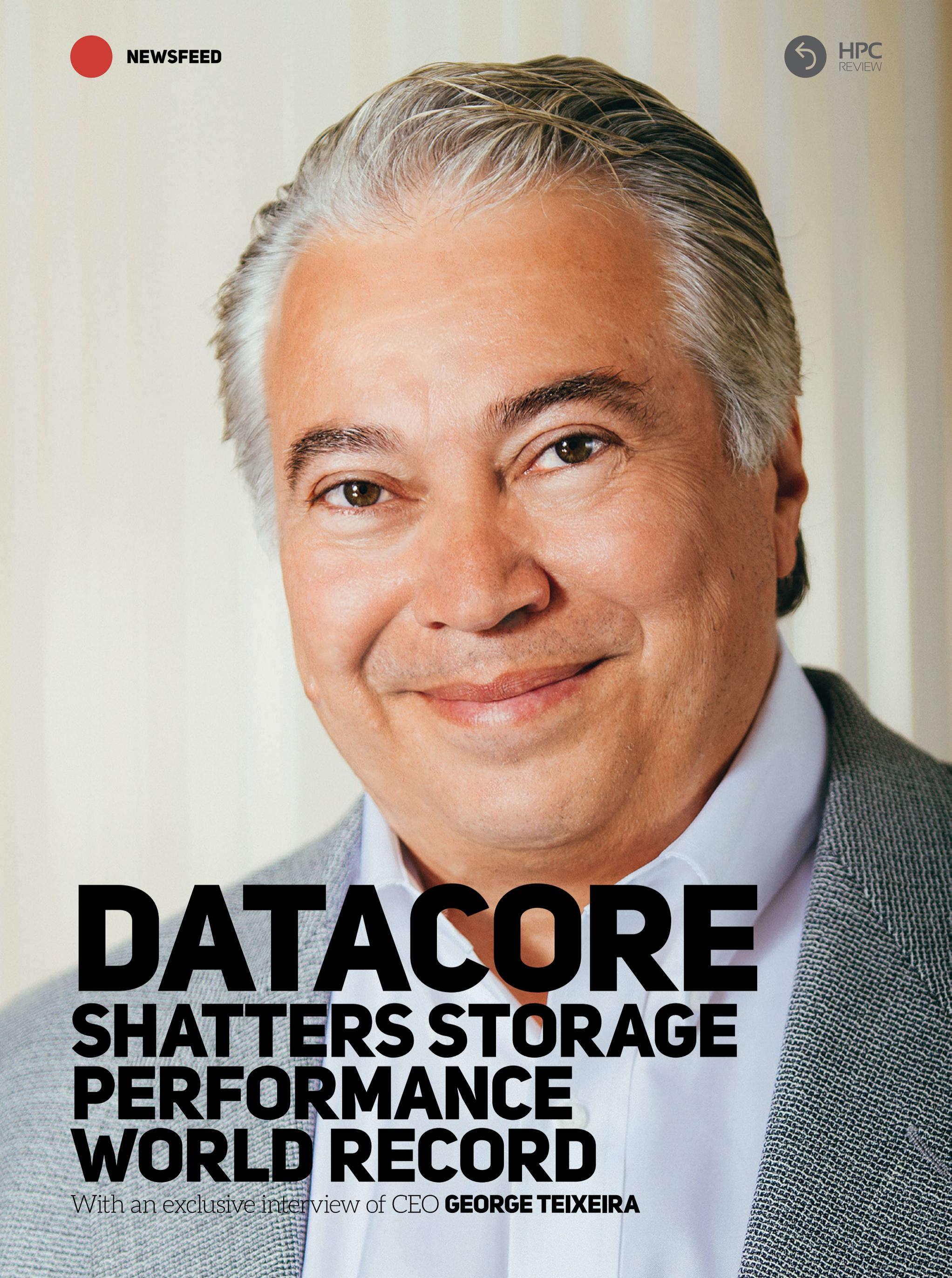
#6 LOB/IT PARTNERS: By 2018, 80% of major IT decisions will involve a partnership between the IT organization and LOBs.

#7 SOFTWARE-DEFINED IT INFRASTRUCTURE: By 2016, 30% of enterprises will implement software-defined infra in their datacenter.

#8 PLATFORMS FROM DIGITAL DISRUPTORS: By 2018, 20% of all enterprises will be leveraging digital platforms from such firms as We-Chat, Uber, GE/Predix and PTC to achieve the DX goals.

#9 ECOSYSTEMS OF ENGAGEMENT: In 2020, in excess of 75% of Ideas, Talent, and Innovation will be sourced through ecosystems of engagement.

#10 NEXT-GEN SECURITY MANAGEMENT: By the end of 2017, 55% of enterprises will rely on third-party management of their security infrastructure.



DATACORE SHATTERS STORAGE PERFORMANCE WORLD RECORD

With an exclusive interview of CEO **GEORGE TEIXEIRA**

WE ARE THE CURRENT HOLDER OF THE RECORD PERFORMANCE IN STORAGE, AS MEASURED BY THE STORAGE PERFORMANCE COUNCIL

DataCore announced a new world record for price-performance using the industry's most recognized and peer reviewed storage benchmark, the Storage Performance Council's SPC-1. Thanks largely to breakthroughs in its parallel I/O software that harnesses the untapped power of multi-core processors, this achievement places DataCore with an audited SPC-1 Price-Performance of \$0.08 per SPC-1 IOPS as the clear-cut leader in SPC-1 Price-Performance overall. DataCore certified its results on a powerful but compact 2U Lenovo System x3650 M5 multicore server featuring Intel Xeon E5-2600 v3 series processors with a mix of flash SSD and disk storage. On this same platform, DataCore also recorded the fastest response times ever attained, even compared to the many all-flash arrays and multi-million dollar name brand systems that have published SPC-1 results.

“With these first certified results, DataCore has put a stake in the ground to demonstrate our parallel I/O performance and hyper-converged capability. For us, this is just the beginning. Look for future benchmarking to incorporate multi-node high availability configurations and to demonstrate I/O originating from both inside and outside the servers - the future for all storage systems,” stated Ziya Aral, Chairman of DataCore Software. “We have only just begun to show the full potential of our inherently parallel I/O architecture.”

ONE PLATFORM FOR VARIOUS USAGES AND BEST OVERALL PERFORMANCE

Hyper-converged System Handles Compute, Parallel I/O Processing and Storage Workloads. Notably, the record-breaking price-performance results were achieved on a hyperconverged solution capable of servicing both enterprise-class storage requirements and demanding database and transaction processing application workloads – all running together on the same platform.

Hyper-converged systems must demonstrate that they can cost-efficiently handle combined enterprise-class storage and application workloads. Unlike SPC-1 results that characterize only external storage systems excluding the servers used to generate the load: DataCore's \$0.08 per SPC-1 IOPS result includes the cost to generate the workload and, therefore, encompasses the total cost and end-to-end requirements for running the enterprise application.

“We'd like to see others, like Nutanix and SimpliVity, publish SPC-1 benchmark numbers to reveal how they fare against our record-breaking SPC-1 Price-Performance results. Then customers can clearly assess the cost implications of these alternatives,” challenges DataCore's CEO, George Teixeira. “There's been much speculation about how these systems perform under I/O-intensive workloads generated by mission-critical enterprise applications. Using the peer-reviewed SPC-1 full disclosure process provides an objective frame of reference for making comparisons prior to any buying decisions.”



THE RESULTS: RECORD-BREAKING PRICE- PERFORMANCE FOR BOTH STORAGE AND HYPER-CONVERGED

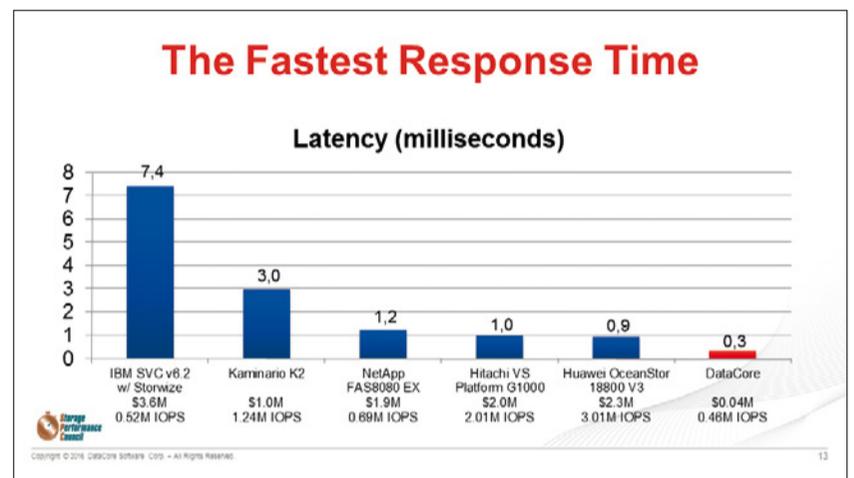
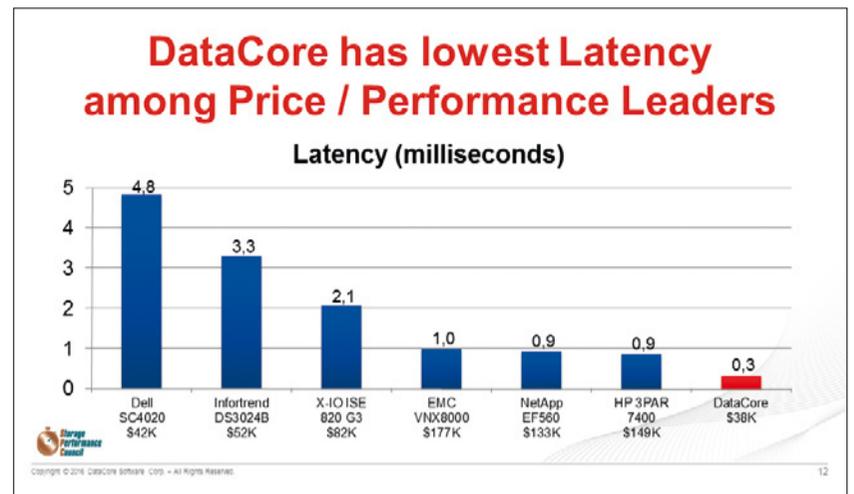
For the benchmark, DataCore used an off-the shelf, hyper-converged system targeting enterprise-class OLTP and latency-sensitive database applications rated for 459,290.87 SPC-1 IOPS, with a total cost for hardware, software and three years of support coming in at \$38,400.29, making it the top SPC-1 Price-Performance result of \$0.08 per SPC-1 IOPS™. That is a 300% improvement over the previous record of \$0.24 per SPC-1 IOPS attained by the Infortrend EonStor DS 3024B and less than 25% of the cost of popular top-of-the-line storage arrays including EMC VNX 8000, NetApp EF560 All Flash Array, Dell Storage SC4020, and HP 3PAR StoreServ 7400.

DataCore and IBM are the only companies to benchmark a hyper-converged system where the SPC-1 applications and storage workloads they generate are both serviced on the same platform. This means that DataCore’s \$38,400.29 price includes not only the storage components, but all of the host server resources and the hypervisor software needed to run the enterprise-level database/OLTP workloads generated by the benchmark.

For comparison, the only other hyper-converged system with publicly reported SPC-1 results is an IBM Power 780 Server. Their SPC-1 Price-Performance result is \$4.56 per SPC-1 IOPS . That system attained 780,081.02 SPC-1 IOPS at a total price of \$3,557,709.00, or roughly 91 times more costly than the DataCore solution.

DATA CORE ADAPTIVE PARALLEL I/O TECHNOLOGY FULLY EXPLOITS THE POWER OF MULTI-CORE CPUS

The remarkable price-performance ratings can be attributed in major part to DataCore Adaptive Parallel I/O techniques intrinsic to the design of the SANsymphony-V Software-Defined Storage Services platform. DataCore executes many independent I/O streams simultaneously across multiple CPU cores, si-



gnificantly reducing the latency to service and process I/Os by taking full advantage of cost-effective but dense, multi-core servers such as the Lenovo System X machines. Competing products serialize I/O limiting their throughput and slowing their response times.

“Lenovo initially approached us to run the demanding SPC-1 enterprise workload benchmark. They wanted proof we could fully harness the power of their multi-core servers given the abundance of unsubstantiated performance claims circulating in hyper-converged circles,” continued Teixeira. “They soon realized with parallel I/O, we had a rocket ship in our hands.”

“Lenovo is excited to partner with DataCore to disrupt the storage marketplace providing customers the best price and performance in the industry” stated Chris Frey, VP & GM, Lenovo North America. “DataCore’s industry-leading SPC-1 results on Lenovo System x demonstrate the performance, innovation and reliability that Lenovo is delivering to meet the growing storage needs to our customers.”



THE SPC-1 BENCHMARK - TESTED / PRICED CONFIGURATION

The rigorous SPC-1 performance testing is designed to demonstrate a system's performance capabilities for business-critical enterprise-level workloads typically found in database and transaction processing environments. The audited configuration that was tested and priced includes DataCore SANsymphony-V parallel I/O software on a Lenovo System x3650 M5 multi-core server featuring Intel Xeon E5-2600 v3 series processors running Microsoft Windows Server, equipped with 16 SSDs and 8 hard disk drives. The DataCore software also supports Microsoft Hyper-V, VMware ESXi, Linux KVM and other hypervisor-based solutions. It can also run directly on Windows servers when server virtualization is not appropriate.

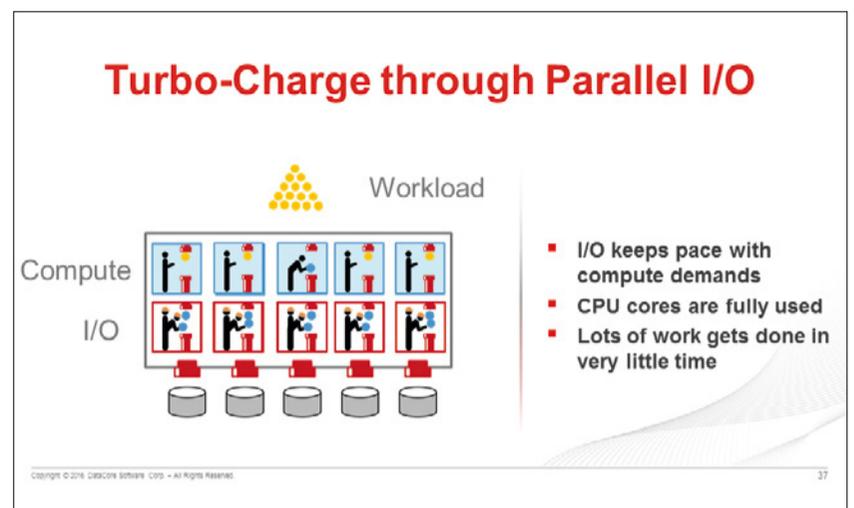
"LESS IS MORE" WITH DATACORE HYPER-CONVERGED VIRTUAL SAN AND SOFTWARE-DEFINED STORAGE

DataCore SANsymphony-V software dramatically reduces the I/O limitations and bottlenecks that restrict the number of virtual machines (VMs) and workloads that can be consolidated on server and hyper-converged platforms. The software enables industry standard x86 servers to gain the essential storage functionality needed to meet today's demanding tier-1 business application requirements. It runs on off-the-shelf servers and works infrastructure-wide across all types of storage (flash, disk and cloud) to automate and optimize performance and resource allocation. **JOSCELYN FLORES**

To understand the results achieved by DataCore, we interviewed George Teixeira, founder and CEO of the company during his recent visit to Paris.

Can you describe DataCore and your solutions?

DataCore was founded in 1998. Our vision at first was very simple : replicate in software what hardware storage controllers accomplished. Today we have 10,000 clients and



30,000 deployments worldwide. We pioneered storage virtualization and we recently we are innovating server and storage productivity with our Parallel I/O technology, which provides the best performance and response times for the storage industry. Our hyperconverged and software platform solutions are SANsymphony and DataCore Virtual SAN. Together, these solutions improve the performance and responsiveness of corporate storage infrastructure. They also radically improve server productivity.

How do you prove the results?

From our perspective, the greatest benefit we can bring to our customers is to provide them a competitive advantage by improving the performance of their storage infrastructure. We just recently won the world record performance, as measured by the Storage Performance Council benchmark, which constitutes the only independent assessment benchmark that is validated by industry peers.

What is your opinion on the main benefits for companies?

They are twofold. On the one hand DataCore enables the infrastructure to efficiently run critical business applications that require faster response times with extremely low latency. It helps companies by retaining the same infrastructure, allowing it to run more virtual machines without losing performance and improving the responsiveness of the attached



WE ARE THE CURRENT HOLDER OF THE RECORD PERFORMANCE IN STORAGE, AS MEASURED BY THE STORAGE PERFORMANCE COUNCIL

storage resources. The elimination of the traditional I/O bottlenecks allows them to run resource-intensive environments such as Oracle databases, SQL Server or SAP Hana or VDI. Our storage virtualization solutions allow companies to use any manufacturer's storage devices and systems without limitation. And frankly, we significantly enhance and improve existing storage infrastructures. Secondly, we greatly improve their performance and productivity. With performance improvements by a factor of 3-20 times faster, our solutions enable 3-20 times more enterprise applications to run. It also accelerates VMware and Microsoft's Hyper-V virtualized workload environments. Typically in Oracle, VMware, SAP, Microsoft SQL cluster based infrastructure environments it takes 8 to 10 nodes, our software solutions are used to do exactly the same thing with only two nodes !

How do you explain this dramatic performance improvement?

In terms of performance and latency, we are the current world holder in price performance in storage, as measured by the Storage Performance Council, an independent organization specializing in measuring performance on storage systems. How we got there is simple: our software is able to harness the heart and power of today's powerful multi-processor. Today, Intel servers often have 8, 12 or even more cores. The problem is that they are unused. Software developers except for DataCore have not bothered to exploit this power that is readily available.

How did you achieve this?

DataCore has worked to harness this untapped reservoir of power in order to exploit the processor cores available in today's servers and thus

accelerate the effective processing of I/O (input-output). The result is an increase in performance of 3 and 20 times better than the fastest all-flash arrays and currently available million dollar enterprise storage systems. With this significant performance improvement, this storage infrastructure can respond 3 to 20 times faster to database queries and enterprise application workload requirements. This rise in productivity also translates into more efficient consolidation, with a reduction of costs from energy consumption, floor space and manageability.

Do you have numbers to substantiate your claims?

Certainly, the results that we put forward are now established by the Storage Performance Council, the only standard industry tool for measuring performance. It is audited and documented, validated by peer industry players. We are currently the absolute leader in performance-price ratio, with a record low cost of 8 cents per SPC-1 IOPS. Measurements of the Storage Performance Council also demonstrate we achieved the lowest latency ever measured, including compared to all-flash array systems. Note, our test server included a mix of traditional disks and flash storage devices.

Even more surprising, this result was obtained with a Lenovo server costing less than 40,000 dollars, and it performed better than systems costing hundreds of thousands or even millions of dollars. I know it's hard to believe, but understand that unlike our competitors, the DataCore software fully utilizes all the available cores on server processors, whereas other software companies only use one or two of the available cores. We are very proud of this result, it truly is a revolution for productivity and the storage industry.

INTERVIEW BY **SOLANGE BELKHAYAT-FUCHS**

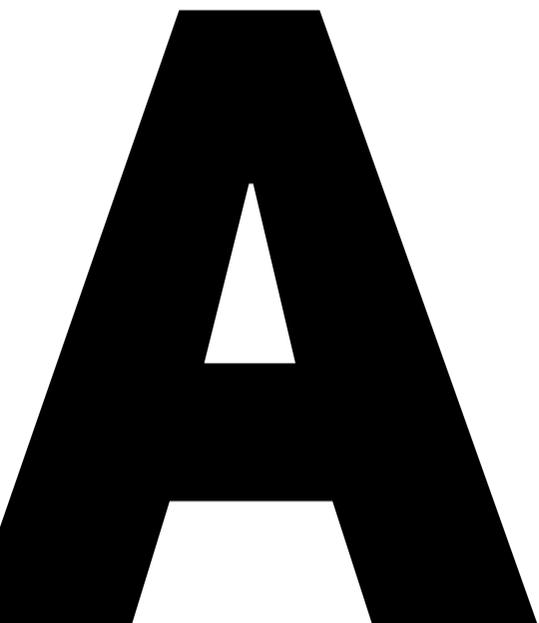


ALAN TURING INSTITUTE WILL LEAD RESEARCH IN DATA SCIENCE

The Alan Turing Institute is the new U.K. data science research organization in London, to help the U.K. as it increases research in data science to benefit research and industry.



THE TURING INSTITUTE WILL, AMONG OTHER TOPICS,
RESEARCH HOW **KNOWLEDGE AND PREDICTIONS CAN BE EXTRACTED
FROM LARGE-SCALE AND DIVERSE DIGITAL DATA.**



At the launch event, U.K. Minister for Science and Universities Jo Johnson paid tribute to Alan Turing and his work. Institute director Professor Andrew Blake told the audience that the Turing Institute is about much more than just big data – it is about data science, analysing that data and gaining a new understanding that leads to decisions and actions.

Alan Turing was a pioneering British computer scientist. He has become a household name in the U.K. following publicity surrounding his role in breaking the Enigma machine ciphers during the Second World War. This was a closely guarded secret until a few years ago, but has recently become the subject of numerous books and several films. Turing was highly influential in the development of computer science, providing a formalization of the concepts of algorithm and computation with the Turing machine. After the war, he worked at the National Physical Laboratory, where he designed ACE, one of the first stored-program computers.

The Alan Turing Institute is a joint venture between the universities of Cambridge, Edinburgh, Oxford, Warwick, University College London, and the U.K. Engineering and Physical Science Research Council (EPSRC). The Institute received initial funding in excess of £75 million (\$110 million) from the U.K. government, the university partners and other

business organizations, including the Lloyd's Register Foundation.

REALIZING THE VALUE OF BIG DATA FOR THE ECONOMY

The Turing Institute will, among other topics, research how knowledge and predictions can be extracted from large-scale and diverse digital data. It will bring together people, organizations and technologies in data science for the development of theory, methodologies and algorithms. The U.K. government is looking to this new Institute to enable the science community, commerce and industry to realize the value of big data for the U.K. economy. Cray will be working with the Turing Institute and EPSRC to provide data analytics capability to the U.K.'s data sciences community. EPSRC's ARCHER supercomputer, a Cray XC30 system based at the University of Edinburgh, has been chosen for this work. Much as we worked with NERSC to port Docker to Cray systems, we will be working with ATI to port analytics software to ARCHER and then XC systems generally.

ARCHER is currently the largest supercomputer for scientific research in the U.K. – with its recent upgrade ARCHER's 118,080 cores can access in excess of 300 TB of memory. What sort of problem might need that amount of processing power? Genomics England is collecting around 200 GB of DNA sequence data from each of 100,000 people. Finding patterns in all this information will be a mammoth task!

ATI have put together a wide ranging programme of workshops and data science summits, details of which can be found on <https://turing.ac.uk/#data-summits-workshops>



WEE ARCHIE

REPLICATING A SUPERCOMPUTER IN A TINY PACKAGE

A mini supercomputer that powers virtual dinosaur races shows how the world's most powerful computers work. What's faster: a Tyrannosaurus, Velociraptor or Dilophosaurus? Using HPC the EPCC, Edinburgh Super Computing Centre, can find out by combining paleontology with biology to form a simulation of these ancient creatures. What's more, this provides an engaging visual demonstration of how HPC can be applied to the sciences which is especially applicable to outreach events such as the British Science Festival. By allowing the public to design and race dinosaurs against each other, we can make a lasting impression about how simulation is

the third research methodology, complementing theory and experiment. The aim of this project is to further develop the prototype dinosaur-racing application. There are two aspects to this software:

1. AN EXISTING FORWARD DYNAMIC MODELLING PROGRAM KNOWN AS GAITSYM, WHICH IS THE SIMULATION WORKHORSE.

It allows the makeup of a dinosaur's skeleton, muscles and joints to be specified, and uses Newton's Laws to calculate the movements that result from these choices. GaitSym already runs on HECToR.

2. A CLIENT WHICH ALLOWS THE PUBLIC TO EASILY CONFIGURE AND SIMULATE THEIR OWN DINOSAURS, VISUALISING THE RESULTS OF THEIR CHOICES VIA REALTIME RACES. This client is currently very basic and the visualisations provided very simple.



world problems run on these machines how HPC is, without them knowing, central to their everyday lives. The hardware itself is complete, but more applications and demonstrations are needed for running on the machine in order to have the impact that we want. A common and very important application of HPC is that of weather forecasting. In conjunction with the UK Met Office we have developed a new, state of the art, weather forecasting model which the scientific community are now starting to use. This would be a fantastic code to run on Wee Archie, not least because the same simulation code is running on ARCHER for real world scientific modelling.

Wee Archie is comprised of 18 Raspberry Pis (8 control cores and 64 computation cores) housed in a custom constructed casing with all the wires and networking hardware required for it to work and providing a minimum of set

up time. The idea behind this is so it can be easily taken to outreach events where the public can gain an understanding of what a supercomputer is comprised of, illustration of parallelism and through examples of real

Ideally, dinosaurs would be designed and configured visually. This would be followed by a detailed view of their composition and how they race each other. The visualisation component is the broad context of this project. The compact machine - called Wee ARCHIE - takes its name from the £43m ARCHER supercomputer at the University's Advanced Computing Facility. Wee ARCHIE replicates in miniature high performance computing techniques to simulate races between on-screen dinosaurs.

PARALLEL COMPUTING : Wee ARCHIE and its larger namesake use parallel computing systems, which enable many calculations to be completed instantaneously on different microprocessors. It was designed and built by the University's science outreach group, FUSION, in collaboration with the Edinburgh Parallel Computing Centre. The machine has already proven popular with school pupils at outreach events designed to shed light on the complexities of supercomputing.

COMPACT SYSTEM : The portable system displays the types of hardware found inside the world's most powerful supercomputers. It contains 18 credit card-sized processors housed in a custom-made Perspex case. LED displays on each of the processors light up when they are in use, showing how multiple parts of a parallel computing system work together to perform complex tasks.

DINOSAUR RACING : The program lets users modify the structure of dinosaurs' muscles and joints, altering their ability to run. Wee ARCHIE tests each of the configurations quickly, and presents the results as an on-screen race.

HIGH PERFORMANCE COMPUTING : Supercomputers are used for tasks that require huge amounts of processing power, such as weather forecasting and molecular modelling of biological compounds. They often occupy several thousand square feet.



books

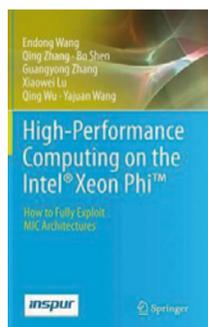
HIGH-PERFORMANCE COMPUTING ON THE INTEL XEON PHI

How to Fully Exploit MIC Architectures

Wang, E., Zhang, Q., Shen, B., Zhang, G., Lu, X., Wu, Q., Wang, Y.

Springer International Publishing, 338 pages

52,74€ / eBook 41,64€



This book explains HPC from general optimization and parallel programming concepts to the details of MIC programming and illustrates all concepts with both a standard example and extracts of real-world applications. It has

been written by a team closely involved in the development of the Intel Xeon Phi coprocessor, the backbone of the fastest supercomputer in the world (Tianhe-2)

The aim of this book is to explain to high-performance computing (HPC) developers how to

utilize the Intel Xeon Phi series products efficiently. To that end, it introduces some computing grammar, programming technology and optimization methods for using many-integrated-core (MIC) platforms and also offers tips and tricks for actual use, based on the authors' first-hand optimization experience. The material is organized in three sections. The first section, "Basics of MIC", introduces the fundamentals of MIC architecture and programming, including the specific Intel MIC programming environment. Next, the section on "Performance Optimization" explains general MIC optimization techniques, which are then illustrated step-by-step using the classical parallel programming example of matrix multiplication. Finally, "Project development" presents a set of practical and experience-driven methods for using parallel computing in application projects, including how to determine if a serial or parallel CPU program



is suitable for MIC and how to transplant a program onto MIC. This book appeals to two main audiences: First, software developers for HPC applications – it will enable them to fully exploit the MIC architecture and thus achieve the extreme performance usually required in biological genetics, medical imaging, aerospace, meteorology and other areas of HPC. Second, students and researchers engaged in parallel and high-performance computing – it will guide them on how to push the limits of system performance for HPC applications.

HPC@GREEN IT - GREEN HIGH PERFORMANCE COMPUTING METHODS

Ralf Gruber, Vincent Keller

Springer International Publishing, 221 pages

63,39€ / eBook 51,16€



The authors present methods to reduce computer energy consumption by a better use of resources and by maximizing the efficiencies of applications. The processor frequency is adjusted to the needs of the running job, leading to a power drop in servers and PCs, and increasing battery life time of laptops. It is shown how computer resources can be optimally adapted to application needs, reducing job run time. The job-related data is stored and reused to help computer managers to stop old machines and to choose new ones better adapted to the application community. Ralf Gruber won the Cray Gigaflop Performance Award in 1989 with world's fastest parallel program running at 1.7 GFlop/s sustained. He was responsible for the Swiss-Tx cluster project, a co-operation between EPFL, Compaq, and Supercomputing Systems. Since 6 years he teaches the doctoral school course on "High Performance Computing Methods". Vincent Keller received his Master degree in Computer Science from the University of Geneva (Switzerland) in 2004, and his PhD degree in 2008 from the Swiss Federal Institute of Technology (EPFL) in the HPCN and HPC

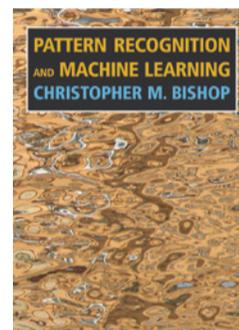
Grids fields. Since 2009, Dr. Vincent Keller holds a full-time researcher position at University of Bonn in Germany. His research interests are in HPC applications analysis, Grid and cluster computing and energy efficiency of large computing ecosystems.

PATTERN RECOGNITION AND MACHINE LEARNING

Christopher Bishop

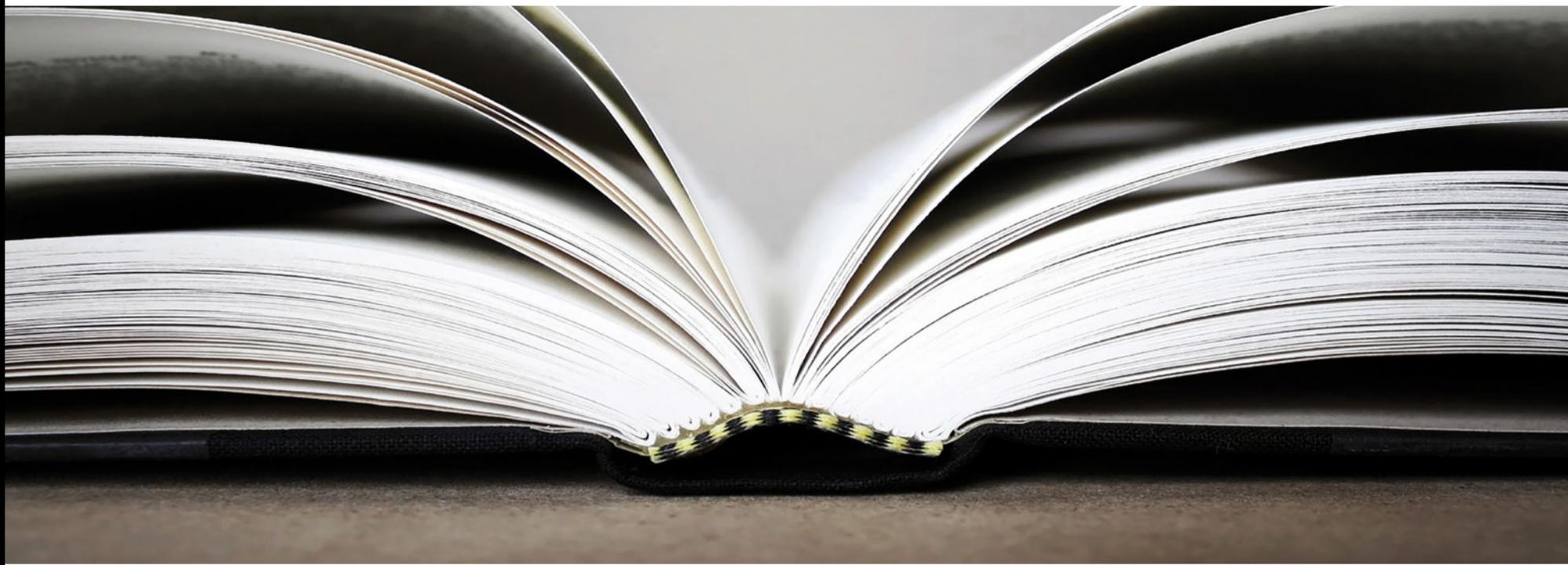
Springer International Publishing

738 pages, 73,79€



The dramatic growth in practical applications for machine learning over the last ten years has been accompanied by many important developments in the underlying algorithms and techniques. For example, Bayesian methods have grown from a specialist niche to become mainstream, while graphical models have emerged as a general framework for describing and applying probabilistic techniques. The practical applicability of Bayesian methods has been greatly enhanced by the development of a range of approximate inference algorithms such as variational Bayes and expectation propagation, while new models based on kernels have had a significant impact on both algorithms and applications.

This completely new textbook reflects these recent developments while providing a comprehensive introduction to the fields of pattern recognition and machine learning. It is aimed at advanced undergraduates or first-year PhD students, as well as researchers and practitioners. No previous knowledge of pattern recognition or machine learning concepts is assumed. Familiarity with multivariate calculus and basic linear algebra is required, and some experience in the use of probabilities would be helpful though not essential as the book includes a self-contained introduction to basic probability theory. The book is suitable for courses on machine learning, statistics, com-



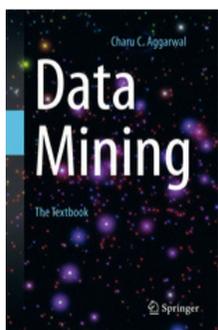
puter science, signal processing, computer vision, data mining, and bioinformatics. Extensive support is provided for course instructors, including more than 400 exercises, graded according to difficulty. Example solutions for a subset of the exercises are available from the book web site, while solutions for the remainder can be obtained by instructors from the publisher. The book is supported by a great deal of additional material, and the reader is encouraged to visit the book web site for the latest information.

DATA MINING

Aggarwal, C. Charu

Springer International Publishing, 734 pages

72,79€ / eBook 59,49€



This textbook explores the different aspects of data mining from the fundamentals to the complex data types and their applications, capturing the wide diversity of problem domains for data mining issues. It goes beyond the traditional focus on data mining problems to introduce advanced data types such as text, time series, discrete sequences, spatial data, graph data, and social networks. Until now, no single book has addressed all these topics in a comprehensive and integrated

way. The chapters of this book fall into one of three categories:

FUNDAMENTAL CHAPTERS: Data mining has four main problems, which correspond to clustering, classification, association pattern mining, and outlier analysis. These chapters comprehensively discuss a wide variety of methods for these problems.

DOMAIN CHAPTERS: These chapters discuss the specific methods used for different domains of data such as text data, time-series data, sequence data, graph data, and spatial data.

APPLICATION CHAPTERS: These chapters study important applications such as stream mining, Web mining, ranking, recommendations, social networks, and privacy preservation. The domain chapters also have an applied flavor.

Appropriate for both introductory and advanced data mining courses, *Data Mining: The Textbook* balances mathematical details and intuition. It contains the necessary mathematical details for professors and researchers, but it is presented in a simple and intuitive style to improve accessibility for students and industrial practitioners (including those with a limited mathematical background). Numerous illustrations, examples, and exercises are included, with an emphasis on semantically interpretable examples.

CHIFFRES CLÉS

44 BILLION DOLLARS

Worldwide projected HPC market value by 2020

8,3%

Yearly growth of HPC market

220 BILLION DOLLARS

Compound market value over the 2015-2020 period

Source : Market Research Media



TOP 500
TOP 3

1 TIANHE-2
National Supercomputing Center, Canton :
33863 / 54902 TFlops Manufacturer NUDT
Architecture Xeon E5-2692 + Xeon Phi 31S1P, TH Express-2

2 TITAN
Oak Ridge National Laboratory,
USA : **17590 / 27113 TFlops** Manufacturer
Cray XK7 Architecture Opteron 6274 + Nvidia Tesla K20X, Cray Gemini Interconnect

3 SEQUOIA
Lawrence Livermore National Laboratory,
USA : **17173 / 20133 TFlops** Manufacturer
IBM Blue Gene/Q Architecture PowerPC A2

The TOP500 classes every six months the 500 most powerful supercomputers in the world. The retained values, RMAX and RPEAK represent the maximum and theoretical Linpack computing power.

GREEN 500
TOP 3

1 7031,6 MFLOPS/W
RIKEN Shoubu (Japan)

2 6952,2 MFLOPS/W
Suiren Blue High Energy Accelerator
Research Organization /KEK (Japan)

3 6217 MFLOPS/W
Suiren High Energy Accelerator Research
Organization /KEK (Japan)

Green 500 list ranks the most energy efficient supercomputers in the world. Energy efficiency is assessed by measuring performance per Watt. The unit here is the MFLOPS / Watt.



COVER STORY



HPC
REVIEW

2016

PREDICTIONS FOR SUPERCOMPUTING

CHARTING
THE FUTURE
OF HPC



2016 promises to be a turnpoint year as technologies mature and their use cases expand beyond the Research and Development labs. Along with breakthrough technologies that are slated to appear in the coming months, 2016 promises to be a HPY : High Performance Year !

FIVE PREDICTIONS OF WHERE SUPERCOMPUTING IS HEADING IN 2016



From new processor technologies to quantum computing, 2016 promises to be another exciting year for supercomputing. Here are five predictions as to how the industry will push ahead in 2016

THE YEAR OF PROGRESS IN NEW PROCESSOR TECHNOLOGIES

There are a number of exciting technologies we should see in 2016, and a leader will be Intel's next-generation Xeon Phi coprocessor – a hybrid between an accelerator and general purpose processor. This new class of processors will have a large impact on the industry with its innovative design that combines a many-core architecture with general-purpose productivity. Cray, for example, will be delivering Intel Xeon Phi processors with some of our largest systems, including those going to Los Alamos National Labs (the “Trinity” supercomputer) and NERSC (the “Cori” supercomputer).

2016 should also see innovation and advances by NVIDIA with their cutting-edge Pascal GPUs, and the continuing ARM innovations will also drive a competitive environment in the processor game.

A QUANTUM LEAP

Quantum computing will both be/not be an exciting technology in 2016. Now that Google has confirmed that their system from D-Wave is performing quantum annealing, we expect more headlines to follow, but by the end of the year the industry will still struggle to understand exactly what it is, or what it will be most useful for. This is certainly a new and exciting technology and the excitement it generates is good for the industry, but it still won't be a truly productive technology in 2016, or 2020, or perhaps even 2025 for that matter.

2016 – THE YEAR FOR DATA-TIERING?

We've seen an explosion in the past few years in the usefulness of solid state devices and associated technologies, as well as dramatic price-performance improvements. This shift is causing some concern over how to easily manage data that is moved through multiple and complex layers – from disks to solid state storage to non-volatile memory (NVM) to random-access memory. In spite of these complexities, we will see a lot of cool, new technologies, solutions and systems coming to the market in 2016.

This technology has already created one very positive trend in cost-effectiveness: if you want bandwidth, you buy bandwidth; if you want capacity, you buy capacity. The difficult part is actually managing the data movement between these layers efficiently. So a lot of software development has to be done in order to take advantage of these cost savings.



Solutions like DDN's Infinite Memory Engine, EMC's (oops, I mean Dell's) DSSD-based product and Cray's own DataWarp product are a few examples of attempts to combine software and hardware innovation around data-tiering, and 2016 will see some blazing-fast solutions and advances in this technology.

COHERENCE OF ANALYTICS AND SUPERCOMPUTING

Coherence and convergence. It seems like everyone is talking about it, including the U.S. Government in its National Strategic Computing Initiative, the Human Brain Project in the UK, and every major commercial and academic player in high performance computing. But it's not all just talk. Big data is changing how people use supercomputing, and supercomputing is changing how people handle big data. Whether this is using supercomputers for baseball analytics, or using IoT sensor data to model weather and climate -- we now see that analytics and supercomputing are inextricably and forever intertwined. In 2016 we will continue to see this coherence grow and bear productive fruit.

WE WILL FINALLY ABANDON LINPACK AS THE METRIC FOR THE TOP500... JUST KIDDING

Since 1993, the Top500 has been the metric of choice for press and public to be able to understand how supercomputing is changing and evolving. That being said, it has never truly been a good measure for whether supercomputers are actually important in our lives. It simply doesn't measure whether a supercomputer is productive or not.

The Top500 does allow us to see, retrospectively, the evolution of technologies that are being delivered in the governmental and commercial spaces. Unfortunately, it doesn't tell us much about where technology is going or whether these systems are useful and productive. There is a key benefit of the Top500 that is actually very important and one that should not be minimized -- the Top500 is easy to un-

derstand. As much as I don't believe in the Top500, I must say that I admire its simplicity and its ability to focus discussion on the big computing challenges we face. So my prediction is that the Top500 will still be the genesis for proud, chest-thumping by governments, institutions of higher learning and commercial companies in 2016. All I can say is, Happy New Year Top500! Maybe we will have better metric someday, but for now, you are the best we have.

ABOUT THE AUTHOR

Barry Bolding is Chief Strategy Officer at Cray

BIG DATA AND ANALYTICS



Business intelligence continues to be one of the fastest-moving areas in the enterprise. And not only is the technology moving fast, but the techniques people are using to drive adoption and get value from their data are multiplying. Among these trends are an increasing appetite for more advanced analytics to answer deeper questions, and new approaches emerging for governance of self-service BI. The potential for innovation is far from over. This paper highlights the top trends in business intelligence for 2016.

GOVERNANCE AND SELF-SERVICE ANALYTICS BECOME BEST FRIENDS

People have considered governance and self-service analytics to be natural enemies. Maybe that's why those people were surprised to see governance and self-service having a drink together. The war is over, and the cultural gap between business and technology is closing. Organizations have learned that data governance, when done right, can help nurture a culture of analytics and meet the needs of the business. People are more likely to dig into their data when they have centralized, clean,



and fast data sources, and when they know that someone (IT) is looking out for security and performance.

VISUAL ANALYTICS BECOMES A COMMON LANGUAGE

Data is changing the conversation—in boardrooms, in the media, and in social media. People are visualizing their data to explore questions, uncover insights, and share stories with both data experts and non-experts alike. As data usage grows, even more people will turn to data with both professional and personal questions. And employers will look for candidates who can think critically with data. Visual analytics will serve as the common language, empowering people to reach insights quickly, collaborate meaningfully, and build a community around data.

THE DATA PRODUCT CHAIN BECOMES DEMOCRATIZED

Self-service analytics tools have changed people's expectations for good. In 2016, people will seek empowerment across the data continuum, especially as more millennials enter the workforce. For business users to stay iterative, they must be able to shape certain data on the fly. That's why the demand for self-service data preparation tools and even self-service data warehouses will grow as a natural extension of self-service analytics. This democratization will allow people to respond quickly to shifting priorities.

DATA INTEGRATION GETS EXCITING

These days many companies want agile analytics. They want to get the right data to the right people, and quickly. This is no small challenge, because that data lives in many different places. Working across data sources can be tedious, impossible, or both. In 2016, we'll see a lot of new players in the data integration space. With the rise of sophisticated tools and the addition of new data sources, companies will stop trying to gather every byte of data in the same place. Data explorers will connect to

each data set where it lives and combine, blend, or join with more agile tools and methods.

ADVANCED ANALYTICS IS NO LONGER JUST FOR ANALYSTS

Non-analysts across the organization are becoming more sophisticated. They've come to expect more than a chart on top of their data. They want a deeper, more meaningful analytics experience. Organizations will adopt platforms that let users apply statistics, ask a series of questions, and stay in the flow of their analysis.

CLOUD DATA & CLOUD ANALYTICS TAKE OFF

In 2015, people began embracing the cloud. They realized putting data in the cloud is easy and highly scalable. They also saw that cloud analytics allows them to be agile. In 2016, more people will transition to the cloud thanks, in part, to tools that help them consume web data. Early adopters are already learning from this data, and others are realizing they should. And more companies will use cloud analytics to analyze more data faster. They'll come to rely on it just like any other critical enterprise system.

THE ANALYTICS CENTER OF EXCELLENCE (COE) BECOMES EXCELLENT

An increasing number of organizations will establish a Center of Excellence to foster adoption of self-service analytics. These centers play a critical role in implementing a data-driven culture. Through enablement programs like online forums and one-on-one training, the centers empower even non-experts to incorporate data into their decision-making. Over time, these centers enable data to inform workflow across the entire organization.

MOBILE ANALYTICS STANDS ON ITS OWN

Mobile analytics has grown up and moved out. It's no longer just an interface to legacy business intelligence products. In 2015, products with a fluid, mobile-first experience began to



emerge. Working with data out in the world is going from being a chore to becoming a dynamic part of the analytics process.

PEOPLE BEGIN TO DIG INTO IOT DATA

The Internet of Things is poised to become even more prevalent in 2016. It seems that everything will have a sensor that sends information back to the mothership. Just think of all the data mobile devices are generating around the clock—and that’s just one piece. As the volume of IoT data grows, so does the potential for insights. Companies will look for tools that allow users to explore the data, then share their findings in a secure, governed, and interactive way.

NEW TECHNOLOGIES RISE TO FILL THE GAPS

There are a number of new technologies in the BI ecosystem. As these go to market, we’ll see gaps that need to be filled. There will be new companies to do just that. Hadoop accelerators, NoSQL data integration, IoT data integration, improved social media—each of these provide an opportunity for a new company to start up. In 2016, we’ll see the rise of the gap fillers, leading to a market consolidation. And organizations will continue to shift away from single solutions and embrace an open and flexible stack that includes these new technologies.

BUSINESS INTELLIGENCE



2016 marks the third year that I’ve penned my business intelligence market predictions for Forbes. Upon reviewing what I wrote last December for 2015, I realized that while many of the predictions had already come to fruition, some are proving too soon to tell. So this year, I’ve decided to focus my predictions specifically around a common theme to better be able to measure myself next year. Overall, 2016 will be “The Year of Information on Demand.”

THE BIG BANG YOU HEAR IS THE EXPLOSION OF INNOVATION

The combination of scale out + hyper-converged + elastic cloud architecture + machine learning is creating massive disruption—breaking all existing paradigms for computing—particularly the performance-for-cost ratios. SQL on Spark, for example, allows us achieve 100x the data processing performance at one-tenth the cost of traditional Business Intelligence (BI) architectures, removing the need for relational data warehouses and other old world tools. This new world will create new leaders...

We’ll see the Dinosaurs trying to reinvent themselves by cross-breeding (e.g., SAP + HP, MSFT + SFDC), while some public companies will seek shelter from Wall Street by going private. Teradata, as an example, might be the first to go, and I expect to see others the caliber of EMC + Dell or SFDC +)

Organizations are consolidating traditional fragmented data integration, data quality, ETL tools into a more connected information layer designed for self-service access and enterprise-grade security and scalability. In order to remain relevant, vendors will attempt to “retrofit” their legacy IT-centric offerings but will be thwarted by the head winds of self-service despite their own forays into providing business-consumable front-end capabilities. The winning designs will be business-analyst centric and built from the ground up for today’s velocity, veracity and variety of data, with transparent governance and multiple deployment models.

WE MUST ADAPT OR DIE

Data preparation will be a critical capability of subject matter experts who, traditionally, relied on others to get data ready for them. In order to transform data into information on demand, people doing risk analysis, customer targeting, security monitoring, marketing or sales operations will need the necessary skills and tools to handle self-service data preparation at scale.



Those who don't adapt to the modern paradigm will experience big data blunders, including misclassification of data and embarrassing data quality errors, as the gap widens between all of the data and the people who understand how it should be used. Expect to see more stories like these in 2016.

WHOSE BI WILL YOU BUY?

Several trends will emerge in the Business Intelligence market. Spending has come to a screeching halt in traditional BI platform investments. Gartner sees a decline of more than 20% in net new license for traditional operational BI vendors as organizations are now increasing their spend on self-service BI and data discovery tools, such as those from Tableau Software, which saw incredible growth of 77.7%. This paves the way for more disruption:

The 800-pound data visualization gorillas will subsume basic data blending, relegating that capability as "table stakes" which will leave point solution vendors scrambling to prove differentiation.

Seattle is the city to watch as MSFT Azure and AWS make big moves in the cloud BI market place. How will Mountain View respond? With Diane Greene at the helm, Google has an opportunity to make this a very interesting competition.

AWS QuickSight will also disrupt the playing field for first generation cloud BI vendors who are not highly differentiated and can't compete with AWS margins.

In general, we will see a major movement to adopt Cloud in the financial service industry. In fact, one of the top five financial institutions will announce a Cloud-first or Cloud-only IT philosophy –adopting one or more of the big three players. If the CIA is doing it, why can't the banks?

ENOUGH ABOUT DATA – WHERE'S THE INFORMATION?

The 2016 Presidential elections will be the first time in our history where a new approach to

information on demand will trump traditional polling as the means for predicting voter behavior and intentions. While micro-segmentation and statistical modeling was done in the 2012 Presidential elections, it required that analyses be done in sample sets in order to deal with all of the variables and data sources. We now have the combination of machine learning, artificial intelligence and compute power to instantaneously build Nate Silver-like predictions without sampling. Picture doing predictions on the entire voting population, bringing together social media data, historical preferences, psychographic, sociographic and firmographic data into information that tells a far richer picture. I vote for that!

While there has been a lot of buzz around the Internet of Things (IoT), people will start to recognize that sensor data is abundant, very dirty and mostly useless. Sensor data, by its very nature, is cryptic and often made unusable after transmission. The more sensor data being collected, the more chaos will be created. It will be critical that we have a pragmatic discussion about how data streams can be cleaned in real time, and why it's necessary to merge IoT data with static data to create contextual information that is relevant and valuable. Smart City programs are early examples of how the combination of IoT data from traffic sensors, social networks, transportation logs, census and emergency personnel data can drive rapid insights and community response plans to national security events or natural disasters.

DEATH BY DESKTOP

Data silos on desktops have created security risks and scalability challenges for years. Because there were no other options, business teams made a conscious decision to use distributed desktop BI tools leaving their IT team scurrying to add controls and security layers around their users. In 2016, we will see success of solutions where security and scalability get "designed in" from the start. This is critical if we are to achieve governed data discovery, a



THE SHARED ANXIETY AROUND DATA RISK WILL UNIFY BUSINESS AND IT TEAMS AS THE NEW CROP OF SOLUTIONS ENABLE SIMPLER ADMINISTRATION, TRANSPARENT GOVERNANCE AND AGILE COLLABORATION.

term used by Rita Sallam, research vice president at Gartner.

The shared anxiety around data risk will unify business and IT teams as the new crop of solutions enable simpler administration, transparent governance and agile collaboration.

The movement will be further accelerated as more enterprises are seeing elastic cloud architectures as the mainstay of big IT-standardization decisions. In fact, this will become the litmus test for Chief architects and CTOs: if the solution doesn't have a built-in elastic architecture, it won't be deployed enterprise-wide.

WE'LL SEE JUST AS MANY CDO'S HIRED AS FIRED!

The CDO role evolved as organizations wanted to have one throat to choke and one hand to shake for all things data. And while we saw initial excitement in 2014 and 2015 (evidenced by the number of CDO events that cropped up all over the world), the role will have to be better defined with clear responsibility and authority if it is to succeed. This Forbes interview with Jennifer Belissent, principal analyst, and Gene Leganza vice president and principal analyst at Forrester Research, provides a powerful statement: "the democratization of data – both the responsibility for it and the insights from it – across the organization is the ultimate objective." Look for successful CDOs to focus exclusively on quantifiable business objectives.

ABOUT THE AUTHOR

Prakash Nanduri is Cofounder and CEO of Paxata

CLOUD COMPUTING



It doesn't take Nostradamus to foretell that cloud computing will keep booming in 2016. Where the predictions do get interesting, though, is around questions such as which forces will drive cloud growth, how the competitive landscape will change, and how cloud's security reputation will evolve. Here are 11 predictions about the strategic choices and competitive battles that lie ahead.

EIGHTY PERCENT OF PRODUCTION APPLICATIONS ARE HEADED TO THE CLOUD

Oracle's CEO Mark Hurd named this as one of five cloud predictions to come true by 2025. Those predictions are grounded in economic factors, not pure technology ones.

Executives feel an intense pressure to innovate using technology, and they're also under intense cost-cutting pressure, because revenue is barely growing at most big companies. Today, only about 25 percent of production apps are in the cloud, Hurd said, but the adoption pace will accelerate because CEOs and CIOs see no other way to meet those dueling pressures of innovating faster while cutting costs.

The cloud "is the only way to make that leapfrog from where you are to where you have to be and do it at a compelling [total cost of ownership] that's both secure and drives the level of innovation," Hurd said.



SECURITY WILL MOVE FROM A BARRIER TO CLOUD ADOPTION TO ONE OF ITS MAIN BENEFITS

“Today, the #1 reason organizations are not moving to the cloud is security. However, tomorrow, security will be one of the most important drivers to move to the cloud,” said Oracle CIO Mark Sunday.

A survey by Harvard Business Review Analytic Services (sponsored by Oracle) found that 62 percent of respondents thought security issues were by far the biggest barriers to expanding cloud adoption at their companies. Nearly half said data security is harder in the cloud.

But those very same concerns will soon drive organizations to the cloud. Established cloud vendors with a solid security track record have the expertise and resources to deploy layers of defense that many companies can't hope to duplicate in-house.

“This is one factor steering many CIOs toward established vendors for cloud services—they have the resources to invest in state-of-the-art security, both physical and logical,” according to the study. Investments in physical security include biometric scanners, security zones, and video surveillance. Logical security includes privileged user access controls, encryption by default, and regularly scheduled patching and maintenance.

“Cloud vendors like Oracle that have a comprehensive and integrated defense of layered security controls are what can turn security from an inhibitor to an enabler of enterprise cloud deployments,” Sunday said.

DIGITAL DISRUPTION WILL FORCE COMPANIES TO REPLACE THE LAST ERA'S IT SYSTEMS—INCLUDING ENTERPRISE RESOURCE PLANNING

Business and IT leaders are embracing cloud systems, but many still look at ERP and think, “That's the last thing we'll touch.”

That thinking will start changing in 2016, as digital business models demand modern, nimble IT systems. ERP projects of the past

were IT's thankless task—seen as complicated, risky, and lengthy, even if the end results paid off. That's why CIOs hang on to 20-year-old financial, supply chain, and other core applications.

Two forces will drive a change in thinking in the year ahead. One is lower internal risk for taking on a major IT project. Cloud IT implementations aren't as slow and risky as the on-premises software projects of old, so even core IT projects can be done in months instead of years. More importantly, companies now realize they must adapt to changing business models in the digital era. Their processes, performance metrics, channels, and data needs are all evolving fast, and business leaders won't tolerate ERP or other systems that can't keep up with their customers' demands.

Some 81 percent of the participants in Accenture's Technology Vision 2015 study think industry boundaries will dramatically blur as “platforms reshape industries into interconnected ecosystems.” The vast majority (86 percent) of CEOs surveyed for PwC's 2015 Digital IQ report said it's crucial for them to champion the use of digital technologies. Look for companies this year to conclude that they need to modernize their core applications to lead during this time of industry upheaval.

IT'S FOCUS WILL SHIFT FROM SUPPORT TO INNOVATION, THANKS TO THE AUTOMATION BUILT INTO CLOUD APPLICATIONS, PLATFORMS, AND INFRASTRUCTURE.

Using cloud resources, companies don't have to have their own IT pros patching databases, backing them up, or setting up data encryption and disaster recovery. Companies don't have to buy, install, and configure hardware systems to support those efforts, nor manage, cool, and power large data centers.

“With Oracle Database Cloud Service, for example, many of the tasks for provisioning and installation are completely automated,” says Steve Daheb, senior vice president of Oracle Cloud Platform. “We believe that you



should focus your time and resources on doing what matters the most, those things that have the greatest impact on your business. Don't assemble the information; take advantage of it. Who better than to build and manage Oracle Database than Oracle? Let us do that work for you." Automation frees up IT organizations to focus time and resources on developing creative new tools that directly improve how customers buy or employees work—whether that's building mobile apps, harnessing data to improve operations, or working with business units to improve customer engagement.

THANKS TO THE CLEAR ROI, 100 PERCENT OF DEVELOPMENT AND TESTING IS MOVING TO THE CLOUD

One of the top quick-win scenarios CIOs will seize on in 2016 is to move their development and testing to the cloud. Hurd predicts it will be done entirely in the cloud by 2025.

When IT teams use on-premises software and hardware for development and testing, they have to buy, license, and configure everything from servers to databases needed to create a development environment that hopefully matches the environment an app will someday run in production. Hurd estimated that the effort wastes about 80 percent of the resources spent on dev-test, and it slows the work down. IT teams often want to develop and test in a cloud environment and then move the application back on-premises for production, mainly for regulatory reasons. Look for companies to choose cloud development platforms that give them the choice to either scale up in a public cloud or run on-premises, once the app's production-ready.

THE CLOUD WILL RESHAPE THE APPLICATION, PLATFORM, AND INFRASTRUCTURE PROVIDER LANDSCAPE

The cloud is shaking up the ranks of technology providers. Oracle, for example, faces an almost entirely new cadre of rivals in the cloud, where it never sees long-time competitors SAP

and IBM, said Oracle Executive Chairman and CTO Larry Ellison. "This is how much our world has changed," Ellison said. Only Microsoft is competing like Oracle at all three layers of the cloud, in applications, platform, and infrastructure, he said.

How will the cloud provider competition shake out? Ellison didn't put a date on it, but he did forecast where Oracle would land in those three cloud segments. "We're going to win by being #1 in applications and #1 in platform, and I think being a really solid, top-three infrastructure player," Ellison told financial analysts.

SHIFTS IN CUSTOMER EXPECTATIONS AND CORPORATE OPERATING MODELS WILL FUEL A HYPERGROWTH MARKET FOR CLOUD COMPUTING, AS CIOs MOVE CORE BUSINESS FUNCTIONS TO THE CLOUD

Almost every industry is transitioning to a digital business model due to evolving customer expectations and changes in business demands. Consider the video game industry, which has experienced a dramatic shift in distribution over the past three years. Consumers are no longer bound to purchasing games at retail outlets, with nearly three quarters of video games now distributed digitally. This kind of massive shift requires modern technology developed to address these challenges. Unfortunately, the applications most companies run their business on today are more than two decades old; they don't meet current customer expectations or scale to meet the requirements of digital businesses.

"It's clear that most existing business systems were not designed to meet today's business requirements—rapid inorganic growth, efficient post-merger integration, expansion into nontraditional markets, faster innovation cycles, and complete digital transformation of the business. To better support these corporate initiatives CIOs will move core functions to the cloud," said Shawn Price, senior vice president, Oracle Cloud.



BUSINESSES WILL EXPECT TRUE HYBRID CLOUDS—NO WALLS BETWEEN PUBLIC CLOUDS AND ON-PREMISES SYSTEMS—SO CLOUD-BASED INTEGRATION WILL SHOOT TO THE TOP OF CIO PRIORITY LISTS

A data silo in the cloud is still a silo. Companies will attack this problem in 2016, insisting that cloud and on-premises systems share data, and relying on increasingly powerful and simple cloud-based integration tools to do that job.

The beauty of cloud apps, platforms, and infrastructure is companies can implement them in weeks or months; the downside is that too often companies are only building data silos faster. Cloud-based integration promises to remove a lot of the complexity that has gone into custom-coding integrations.

Simpler integration tools mean that data experts—rather than technologists—can do many integrations, letting the people who know what information's most needed build the data links between systems. Such tools free up IT pros to focus on only the most critical or sophisticated integrations, letting them serve as the architects managing the integration process. Clouds and on-premises systems will coexist for a decade or more, so that integration architecture know-how will be in high demand.

CLOUD SYSTEMS WILL PROVIDE AN ALTERNATIVE TO BUILD-IT-YOURSELF BIG DATA

Early adopters had no choice but to build their own big data clusters and environments. But assembling, managing, and maintaining these unique systems built on Hadoop, Spark, and other emerging technologies is costly and time-consuming. And the needed talent is tough to find.

In 2016, we'll see cloud services and appliances mature to the point that they become mainstream for big data initiatives, freeing in-house big data experts to figure out how best to apply that analytical power.

INFRASTRUCTURE VENDORS WITHOUT A CLOUD INFRASTRUCTURE OPTION WILL BE IN TROUBLE

The cloud has fundamentally changed how organizations purchase, consume, and use computing and storage resources. IT now has clear choices for developing, deploying, and running applications and workloads on-premises, in the cloud, or both. But companies will keep some of their existing on-premises environment for many years to come for applications that require sub-microsecond response time. Infrastructure vendors are being disrupted by this change and “need to architect their products to work both in your data center and in the public cloud,” said Dave Donatelli, Oracle executive vice president for converged infrastructure. The pressure will be on hardware and cloud infrastructure providers to deliver complete compatibility, so that workloads can shift between cloud and on-premises with no code changes. “Having a successful public cloud is a requirement for any large-scale hardware company that plans to be successful over the longer term,” Donatelli told Forbes contributor John Furrier. At this point, Oracle, as intellectual property owner from chip to cloud, is the only IT supplier that offers this compatibility, delivering common architectures, common management, and common SLAs, with visibility and support across on-premises and the Oracle Public Cloud.

DEMAND FOR SELF-SERVICE APPS WILL RISE, LED BY ANALYTICS TOOLS

The days of relying on IT for tasks like sending out periodic reports are ending. Employees want to do their own analyses, jumping on new opportunities and problems as they pop up. Cloud-based, self-service tools are making such speedy action possible in many realms—collaboration, digital marketing, procurement—but nowhere will the value look more clear than the booming area of visual analytics. With data soaring in volume and complexity, business people are using visual analytics to access it, make sense of it, and take



action. Some 46 percent of respondents to a TDWI survey on visual analytics best practices said they're dependent or extremely dependent on IT for analytics, but only 37 percent thought they would be just as dependent on IT in two years.

Forward-looking IT teams will encourage this cloud-based, self-service independence in 2016. Personalized data visualizations, easier data mashups, and unfettered data access will give business leaders insight at the speed they need, and let IT focus on more innovative work rather than report building.

ABOUT THE AUTHORS

Mark Hurd is CEO of Oracle

Mark Sunday is CIO at Oracle

Steve Daheb is senior vice president of Oracle Cloud Platform

Shawn Price is Oracle Cloud Senior Vice President

CONTEXTUALIZED INFORMATION



The trends recognize the enormous potential that comes with digitization – and the fact that corporate leaders have significant opportunities to deploy next-generation technology across the enterprise to drive change, growth and value for all stakeholders.

“From the boardroom to the C-suite, there’s greater recognition that strategically investing in innovation and transforming IT can improve customer experiences, heighten productivity and efficiency, safeguard infrastructure and brand, and ensure a competitive edge,” said Dan Hushon, CSC’s chief technology officer. “There is a dramatic shift in the way organizations are using digital platforms to do business, and CSC is excited to help our clients succeed on their digital transformation journey.” CSC’s top five IT trends will make 2016 the “Year of Contextualized Information,” with more open, information-driven organi-

zations modernizing their platforms and applications, while simultaneously addressing risk and security requirements.

CONTEXTUAL DATA DRIVES VALUE

Building on the growth of information generated by digital engagement platforms and data supplied by “Internet of Things” (IoT) devices, contextualized analytics will be a prevailing force in 2016. This contextualized information enhances enterprise data with meta-data such as device, location, language and social network. Influencers help enterprises develop better insights, personalize products or services, or even suggest specific actions. Richer context will allow enterprises to create a more integrated and valuable information experience for clients, employees, partners and citizens.

AS CONTEXT INCREASES, CYBER TARGETS INCREASE

As data becomes more contextually rich, it becomes more valuable to the enterprise — and to cybercriminals as well. The growing risk of attack will require next-generation techniques for network defense, identity access management, risk management and now information management. Public clouds will play a role in the integration of contextual data, and these will need to be included in security system architectures. Context-rich environments will force enterprises to give significant thought to acceptable risk levels and controls as they share a growing range of information with key stakeholders.

THE API ECONOMY IS STRONG

Enterprises have discovered the power of Application Programming Interfaces (APIs) and will continue to build on them. APIs enable information access and exchange between systems, often acting as a wrapper around older systems. This enables organizations to combine data from legacy applications and new applications. APIs are more than a development tool; they create new channels for



THIS PARTNERSHIP OR BUSINESS RELATIONSHIP MANAGEMENT (BRM) WILL HELP THE ORGANIZATION GAIN ACCESS TO THE RIGHT INFORMATION AND TECHNOLOGY TO MAKE BETTER DECISIONS AND INTRODUCE COMPETITIVE PRODUCTS, QUICKLY AND AT SCALE.

service integration, information coordination and ecosystems of information sharing. The core of a digital strategy depends on democratizing information access, and APIs play a central role in that process. They are key to developing a business strategy for innovation and providing contextual information.

CIO PARTNER TO BRING INFORMATION INTO CONTEXT

The CIO will increasingly ensure that enterprise controls and governance are in place to manage information. Where in the past CIOs have been more focused on the “T” in “IT,” they will now be more focused on the “I,” partnering closely with the business to drive digital transformation through better information. This partnership or business relationship management (BRM) will help the organization gain access to the right information and technology to make better decisions and introduce competitive products, quickly and at scale.

ENTERPRISE PLATFORM PLAYERS CONTINUE TO CONVERGE

Another trend is the ongoing consolidation of enterprise IT infrastructure players. This consolidation answers the market’s demand for more standard, agile and integrated solutions out of the box, such as converged infrastructure. Look for greater specialization of infrastructure platforms in future products and services.

ABOUT THE AUTHOR

Dan Hushon is CTO, at CSC

CYBERSECURITY



Thought you might like to see Stu Sjouwerman’s predictions for 2016. If you are interested in more information, images or speaking with Stu, please let me know. Thanks and happy new year!

BOARD ROOM

Boards of directors will finally appreciate that information security risk management should be treated as an enterprise risk equivalent to financial, reputational, and legal risk. Too often board members see the light only after a data breach. Having understood how security risk impacts the business, in 2016 they will make sure to change corporate culture into the direction of a good security risk management program. Next year will be a very important year for cyber-insurance; boards are going to ask for this. PwC predicts that the cyber insurance market will triple in the next five years and this will force boards to take a long, hard look at the cost of their continued insecurity.

BUDGET SURVIVAL TIPS

The things that will get approval are projects designed to cut the cost of doing business. Smaller-scale IT security initiatives that have a quick ROI, e.g. new school security awareness training which combines training with simulated phishing attacks, will be popular with management and boards.



CEO FRAUD AKA BUSINESS EMAIL COMPROMISE

Looking at the rapid uptrend of CEO Fraud over the past year, this will be the new scourge in 2016 following ransomware, hitting consumers, small and medium enterprise and large enterprise with competing cyber mafias specializing in verticals like financial institutions, healthcare and manufacturing.

INTEGRITY ATTACKS

Think Stuxnet for the Financial Industry. The data breach attacks we have seen by the hundreds are loud and obvious. They expose data which causes embarrassment, inconvenience, and financial losses. Integrity attacks are stealthy, selective, and can be much more devastating. Instead of doing damage or making off with vast amounts of sensitive data, they instead focus on carefully changing particular elements within transactions, communications, or data to gain a significant benefit.

In 2016, you will see an integrity attack in the financial sector in which millions of dollars will be stolen by cyber thieves who will modify selected data in the transaction stream, resulting in a significant redirection of payment to anonymized accounts. How they'll get in? Spear phishing.

INTERNET INFRASTRUCTURE

2016 is the year that passwords will start to disappear. Biometrics like voice and face recognition go mainstream and 2-factor tech like authentication code generators on mobile phones will spread exponentially. For your own infrastructure, look at containers that host similar 2FA micro-services that you can integrate in your own apps. Nation-states will continue battling for the domination of Internet backbone and infrastructure components.

INTERNET OF THINGS

6.4 billion connected "things" will be in use globally by the end of 2016, but IoT standards related to security are a hodge-podge. There are literally hundreds of standards that poten-

tially touch IoT and precious few that directly accommodate IoT. It's early days and there is no consensus. As long as vendors' #1 concern is "time to market" and not "security by design" you will see a flurry of attacks on IoT devices like Talking Barbie and others. Cutting edge criminal hackers will create the very first BoT - Botnet of Things. Don't leave any kind of Wi-Fi enabled devices or toys in the master bedroom!

MALICIOUS E-COMMERCE GOES SOCIAL

Many traditional social networking sites such as Pinterest, Facebook and Twitter will add "buy" buttons to their platforms in an effort to increase stickiness with their users and help monetize their user base. It's going to be heaven for cyber criminals who will social engineer themselves into millions.

MOBILE MALWARE

Mobile malware, specifically mobile banking Trojans, are on a trajectory to become much more prevalent for banks and financial institutions in 2016. There will be an increase in malware families that are gaining root access rights on users' devices. These attacks will pose a significant problem for many financial institutions, which have thus far mostly ignored the threats mobile devices pose. Commercial malware authors will continue to reinvest at ever greater rates, bringing them towards the "spending power" of nation-state activity. This includes purchasing zero days. The bad guys have lots of cash and they are smart investors.

RANSOMWARE

- A ransomware crime wave will surge across America.
- The use of Cryptowall 4.0 will explode, and Cryptowall V5.0 will add an actually working "feature" that TeslaCrypt only threatened with: extortion by potentially publishing private personal or business files on the Internet.
- Cyber mafias will focus on professional services firms and local government using Cryp-



toward as their tool and extort tens of thousands of dollars from organizations that don't want their business disrupted or their intellectual property compromised.

- Cryptowall will be the first strain of ransomware to hit a billion dollar in damages.
- Ransomware is the new APT: "Annoying Persistent Threat", as it will be increasingly used in double payload attacks combined with other scams.
- Ransomware attacks doubled in 2015 and will double again in 2016. The U.K. is to some extent a bell-weather for the U.S. as they function as a beta test site for Eastern European cyber mafias who can test malicious code in their own time zone. Well, over half (54%) of all malware targeting UK users in 2015 contained some form of ransomware. Buckle up.
- Ransomware-as-a-service hosted on the TOR network and using Bitcoin for ransom payment enables a new generation of cybercrime newbies to make their mark.
- A new sleeper ransomware variant will start to stealthily encrypt data, pull your critical files onto a C&C Server, and wait until a backup been made. At that point they will yank the encryption key and demand a much larger amount of ransom than the current 500 bucks.

SOCIAL ENGINEERING

IDG asked hundreds of high-level InfoSec pros the following question: "What will be the single biggest security threat of 2016?" The number one answer was: "people". With events like the presidential election drumming up a frenzy of social media activity in 2016, you can expect attackers to use the attention given to political campaigns, platforms and candidates as an opportunity to tailor social engineering lures.

ABOUT THE AUTHOR

Stu Sjouerman is the founder and CEO of KnowBe4

INTERCONNECT, CONTAINERS AND ALTERNATIVE PROCESSORS

INTERCONNECT WARS HEAT UP



Intel's Omni-Path interconnect will challenge the dominance of Mellanox InfiniBand in the mid-range HPC space. At SC15 both vendors were strongly asserting their relative superiority, but still based mostly on synthetic benchmarks and projections. With major deployments of both the latest-generation EDR InfiniBand and Omni-Path in the beginning of the year, it will be very interesting to see how the performance stacks up in real life scenarios.

The French supercomputer manufacturer Bull (Atos) is also deploying the first system using their BXI interconnect in the 2nd half of 2016. This is a very promising design especially in the high-end of the market and provides competition for Cray's Aries interconnect that has dominated that segment for a couple of years.

This new competition is very welcome and I predict it will provide more alternatives, drive introduction innovative features while keeping prices down. However having a multi-system environment with a unified fabric for storage, like we have at CSC today, may become more challenging to architect: The different high-speed interconnect fabrics may be incompatible and require gateway devices.

While ubiquitous, Ethernet has long been confined to the low-end and "embarrassingly parallel" clusters with relatively little take-up on the mid- and high-end general purpose clusters. However, the hyperscale industry is deploying ever larger and more scalable Ethernet fabrics with especially the financial community pushing for low latency solutions. Thus I predict that the first very large and fairly capable Ethernet-based HPC cluster for



scientific workloads will be deployed in 2016. I've predicted this for many years and it has yet to emerge so I'm not betting any big money on it :) If such a thing would happen and prove successful, it would set an important precedent and potentially open the floodgates for Ethernet-based solutions in the mid-range.

CONTAINERS MAKE CUSTOM ENVIRONMENTS EASY

Containers can provide a mechanism that's more lightweight than traditional virtualization for supporting custom compute environments in a platform-independent way. For example, users can run their own containers on HPC systems and HPC centers' can provide their software in containers for people to run on their laptops. There are also many other potential use cases to explore both on user application as well as systems management side.

During the last year the notion of having HPC workloads in containers (especially Docker) has gained increasing traction. Container support has been emerging including IBM's container support with LSF and the User Defined Images service in NERSC, based on their Shifter software. The latter has also led Cray to announce nearly imminent support for containers on their XC systems.

There's still a lot of work to do to and a lot of open questions. Thus I expect that next year will see a boom in pilot projects testing and refining containerization in a HPC context. We at CSC will definitely be testing these features in the upcoming year. I also predict that it will be almost certain that containers will become a standard feature of HPC systems within a couple of years, occupying a middle ground of use cases between full-on virtualization and "traditional" bare-metal HPC.

ALTERNATIVE PROCESSORS INCHING TO THE MAINSTREAM

Let's face it: Intel is completely dominating the processor scene in the HPC server market. However, serious challengers are in the horizon with multiple ARM vendors (Broadcom,

Cavium, etc.) building increasingly capable processors. Also increasingly capable HPC prototype systems are being developed, most prominently within the pan-European Mont Blanc project. Furthermore IBM has opened up their POWER processor ecosystem via the very active OpenPOWER consortium and is closely partnering with NVidia to tightly integrate GPUs. Time seems to be ripe for some of these challengers to really step up.

On the flipside, there is such a large inertia built around Intel and the related ecosystem that I expect that in the near future there will not be a huge surge of systems in the HPC space featuring such alternative processor architectures.

That said, these challengers to Intel's throne will continue build up momentum in the datacenters of web-scale IT companies like Facebook, Rackspace and Google. The massive amounts of hardware that these giants consume will surely drive economies of scale and probably make these chips very attractive from a price/performance standpoint for the HPC market as well.

ABOUT THE AUTHOR

Olli-Pekka Lehto works in various roles involving supercomputers since 2002. These days he manages the Computing Platforms group at CSC.

INTERNET OF THINGS



The "security of things" will take centre stage. In 2015, the market saw tremendous growth in the number of connected devices, and that proliferation gave rise to concerns about the security of IoT. Next year will be the one where IoT security takes centre stage – and the winners will be the solution providers who can help enterprises not only deliver connected services, but secure them, too.

We will stop counting the "things" in favour of more important success metrics. By



2020, we could have 20 billion, 34 billion or maybe even 50 billion devices connected by IoT – depending on the analyst firm you ask. But the internet of things is ultimately not about “things”, it’s about service. In 2016, the analysts watching the IoT industry will shift their focus to put a premium on the quantifiable impact of new services (like revenue generated or new experiences enabled) instead of simply tallying device quantities. In fact, IDC added an amendment to its recent device count forecast to say that “by 2018, there will be 22 billion internet of things devices installed, driving the development of more than 200,000 new IoT apps and services”.

For the first time, more new cars will be connected than not. In 2015, GM announced that every one of its new vehicles in the United States would ship with 4G-LTE embedded – and the automaker already has 1 million connected cars on the road. In 2016, we’ll see this trend expand greatly and, for the first time, the majority of new vehicles produced in the United States will be connected. We won’t hit the same tipping point globally in 2016, but initial strong growth indicates that connected cars will be the norm for all new vehicles produced worldwide within the next couple of years.

Low-power wireless area network (LPWAN) technologies will not go mainstream. LPWAN technologies, such as France’s Sigfox, hold great promise and could one day contribute to the continued proliferation of IoT. But the LPWAN market is still in its infancy and far from ready, so we won’t see wide deployment of low-power technologies this year. Instead, 2016 will be the year of LPWAN partnerships, consolidation and marketing wars to determine which of the many players and standards will be victorious.

The IoT market will align around a more precise lexicon. In 2015, IoT emerged from the hype cycle and began to become mainstream. In an attempt to claim relevance, lots of technology providers attached the IoT acronym to their names, resulting in a proliferation of

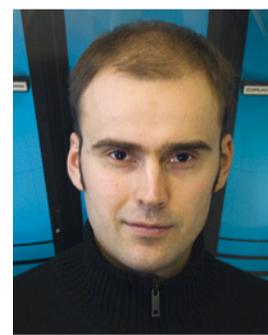
companies talking about their IoT platforms, IoT applications, IoT solutions and more. Business savvy enterprises will demand accuracy and accountability, resulting in clearer definitions from IoT solution providers. Companies will need to articulate exactly what their IoT solution enables – whether that’s IoT application enablement, service lifecycle management, analytics or another function – and clearly demonstrate how they make it easier for enterprises to launch, manage and monetize IoT.

Little doubt remains regarding IoT’s continued growth. As IoT quickly becomes an imperative across industries and adoption continues to accelerate, companies will increasingly deliver services that generate new sustainable revenue sources while enhancing the experiences of their customers. Thousands of businesses worldwide are already reaping IoT’s rewards, and 2016 will be the year that more companies join the IoT revolution.

ABOUT THE AUTHOR

Jahangir Mohammed is Founder and CEO, Jasper Technologies

MANY-CORE PROCESSORS



During the last twenty to thirty years the performance of processors has increased exponentially, first by increasing the clock frequency and the number of instructions per clock (IPC), and in the last ten years through increased parallelism. This trend has led to multicore processors with more than ten cores each able to execute up to w16 double precision floating point operations per cycle, and GPUs with thousands of very lightweight cores able to run tens of thousands of simultaneous threads.

A major new architecture is introduced by Intel in 2016; the latest generation many integrated core (MIC) Xeon Phi processor, Knights Landing (KNL). This processor is not an acce-



erator, but a x86 CPU that is fully compatible with normal x86 processors. It is different in that it is a thoroughbred HPC processor giving in total 3 Tflops of performance per socket. It achieves this high performance through a number of new technologies, which in the future will become commonplace also for normal processors:

A very high core number. KNL contains 72 compute cores connected with a 2D mesh interconnect. This is a true “cluster on a chip”.

- New AVX-512 vector instructions that are able to operate on 8 double precision numbers, enabling each core to execute 32 double precision floating point operations per cycle.
- A new level in the memory hierarchy in the form of a high bandwidth memory sitting on the socket. This memory is 16 GB in size, and will have 5x more bandwidth than normal DDR4 main memory, while latency is comparable to normal main memory.
- Integrated on-socket network interface for Omni-Path.

I predict that KNL in itself will become a successful processor architecture for HPC. I also predict that one should look at KNL as a proxy for what the future will bring. Tuning applications now for KNL by implementing a hybrid MPI + OpenMP parallelization scaling to tens of threads per MPI process, and by enabling the core loops to vectorize well will pay off on any architecture. The new memory level is also something that one should be able to exploit in the algorithms, the new flop-monsters need to be fed and optimizing memory traffic will become ever more important.

ABOUT THE AUTHOR

Dr Sebastian von Alfthan is a senior HPC specialist at CSC with a talent for deep optimization of codes for extremely high-performance as well as exploring novel processor architectures. He has also worked at the Finnish Meteorological Institute developing the space plasma code, Vlasiator.

QUANTUM COMPUTING AND APPLICATION PERFORMANCE

INCREASED APPLICATION PERFORMANCE VIA BURST I/O BUFFERING



In HPC systems of today, there is a horrendous performance gap between DRAM-based main memory and the storage system, typically based on Lustre or similar parallel file system. On Lustre, the I/O bandwidth is very high because of its highly parallel design, but with a downside of pronounced I/O latency. This is due to the access over an interconnect plus the separate metadata server. In practice, this means that small and frequent I/O accesses are poisonous for application performance - and there are several algorithms and codes in scientific computing that require the small but frequent non-contiguous reads and writes.

These bottlenecks have been alleviated in the current generation of HPC systems by introducing some kind of local storage on a compute node; earlier by spinning disks and more recently by SSD (solid state disk) based technologies. Typically these options have been available only on mid-range cluster systems, since designs aimed towards the Top-100 systems cannot host these due to packaging density considerations. Furthermore, the local disk solutions are separate mount points i.e. separate directories. It has been up to the user or the application code to use the proper storage location.

Next year will bring some improvements to the situation by introducing more levels to the storage hierarchy, between the main memory and the Lustre-type file system. For instance, Cray is introducing to the current XC family of supercomputers (for which addition of compute node disks of any kind has not been possible) their DataWarp technology, which is a



flash-based burst I/O buffering and a cache level for Lustre. The final version of the technology will be fully transparent to the application, DataWarp absorbing all the bursty I/O and taking it to the Lustre in the background, and being also capable of prefetching data. DDN will hit the market with their Infinite Memory Engine solution, aiming at providing similar outcome but with a different philosophy. Early versions of these technologies have shown very promising performance benefits for real-world applications. CSC's Cray XC "Sisu" will get two DataWarp blades in mid-January in order to experiment with the burst buffer technology. The prediction? These technologies will become more or less an auto-include in systems aimed for hosting real-world, heterogeneous workloads deployed in 2016!

FIRST STEPS TOWARDS PRACTICAL QUANTUM COMPUTING

Rather than storing information as 0's and 1's as conventional computers (from tablets to supercomputers) do, a quantum computer uses qubits, which can be 0 and 1 at the same time. This quantum superposition, along with the quantum mechanical phenomena of entanglement and tunneling, enable quantum computers to consider and manipulate all combinations of bits simultaneously.

The first commercial realization of a quantum computer by Canadian startup company D-Wave Systems has gained interest during this year, announcing installations by e.g. Lockheed Martin, a joint procurement by Google and NASA and a joint procurement of Los Alamos and Sandia National Laboratories. Their solution is a quantum annealing computer, consisting of 1024 (in fact a bit more) qubits based on superconducting loops. It is capable of finding an energy minimum of more or less arbitrarily complex optimization problem in a constant time (exactly one operation). Early benchmarks by Google show 108-fold speedup in quantum system simulation as compared to the Quantum Monte Carlo method calculation on a conventional computer.

My related prediction for the year 2016 is that the D-Wave Systems will have their order book full, announcing more and more installations by national laboratories and leading computing centres across the globe. In addition, we are going to see first computational challenges (intractable by conventional computers) solved by quantum computers published in scientific journals. Even if it is unlikely that CSC will be installing a quantum computer in the near future, some introductory activity on this topic may be coming up. Stay tuned!

ABOUT THE AUTHOR

Dr Pekka Manninen is a senior HPC specialist at CSC. He works with support and optimization of end-user applications on CSC supercomputers.

SOFTWARE DEVELOPMENT



Hardware and the Top 500 ranking always grabs effort and headlines in HPC, but for a sometimes Pyrrhic victory of raw power over utility. It has often been at the expense of training, tools and code development that would unleash far more scientific output from the same "tin". 2016 is set to change all that.

Just as representative benchmarks like HPCG are set to replace Linpack, so a focus on software is taking over. From industry analysts to users at SC15 we heard that software is the number one challenge and the number one opportunity to have world-class impact. The key drivers of this attention?

CODE MAY RUN SLOWER IN NEWER HARDWARE

Funding bodies and sponsors will not accept a significant HPC investment delivering less scientific output. The Knights Landing Platform will provide many-core nodes – but without software preparation, anticipated lower



IN THE TUG-OF-WAR BETWEEN NATIONWIDE CONSOLIDATION, LOCAL SYSTEMS AND ON-DEMAND CLOUD DEPLOYMENT, THE OUTCOME ISN'T CLEAR.

INITIATIVES SUCH AS OPEN HPC WILL ENABLE EASIER CONSTRUCTION OF ON-DEMAND AND PHYSICAL CLUSTERS. STAYING RELEVANT FOR ALL OUTCOMES MEANS RECOGNIZING WHERE THE VALUE LIES.

clock-speeds will hamper performance. KNL has great potential but code modernization is vital. Using threads, SIMD vector units and HBM memory should be transformational.

Developers and computational scientists are essential – and to apply their skills requires tools such as Allinea Forge. Forge finds vectorization, threading and memory access optimizations – as well I/O and MPI bottlenecks – and can debug the changes!

CODE MAY NOT RUN AT ALL ON NEW HARDWARE

We're set for a wider range of viable architectures than at any time in recent memory: NVIDIA GPU on ARM or OpenPOWER, or an Intel Xeon or Intel Xeon Phi.

Performance and correctness on one platform does not necessarily transfer to another. It's complex work and requires deep software insight that only tools can provide.

Allinea Forge is HPC's cross-platform development tool suite. Users can switch platforms easily and compare systems like-for-like, and can ensure changes for one architecture do not degrade performance on others.

AWARENESS OF THE PERFORMANCE PROBLEM IS RISING

As more codes are modernized, more hero stories emerge – which leads to competition amongst codes and highlights the potential for improvement elsewhere! Again the developer is the key.

Faster, better codes attract larger user-bases with greater numbers of publications, and, for commercial codes, significant customer wins or business gains. Our Performance Reports tool helps users, admins and developers to understand performance problems quickly and how to tune them – and helps to target the development efforts quickly when necessary.

ACCESS TO HPC IS CHANGING AND WIDENING

In the tug-of-war between nationwide consolidation, local systems and on-demand cloud deployment, the outcome isn't clear. Consolidation may be a trend, but initiatives such as Open HPC will enable easier construction of on-demand and physical clusters. Staying relevant for all outcomes means recognizing where the value lies.

At SC15 many centers spoke of a drive to achieve recognition through software excellence on their new machine. The success of a center is the outstanding science that its members and users create. That means empowering those who develop and use the software to succeed – through training and providing the tools they need to do the job well! While many thought-leading sites have pushed software excellence alongside excellent systems, this thinking is finally going mainstream. For that reason, 2016 will be a great year for HPC.

ABOUT THE AUTHOR

David Lecomber is CEO of Allinea



SUPERCOMPUTING



More powerful supercomputers can accomplish some incredible things. These devices hold the key to improving health care, making transportation safer and giving a more accurate weather forecast. But how does it all happen?

David Yates, PhD, an associate professor in the computer information systems degree program at Bentley University, says a supercomputer's power is based on how many instructions they can carry out per second. Some of the most powerful computers today can complete 20 quadrillion instructions within that time span. That's a number with 16 zeros. To put that into perspective, if you devoted all of your time to counting right now, you wouldn't reach 2 billion by the end of your lifetime. These computers can complete over 10 million times that amount of commands in a heartbeat.

Yates would know. He's made a living in the computer science field and earned some impressive awards with Bentley University's information technology degree students in the process. Just last year, they came in first place in a cluster computer competition in the Supercomputing Conference in Colorado. The victory earned them a spot in the 2014 International Supercomputing Conference in Leipzig, Germany.

But what's the point of building a computer that can do quadrillions of things in a smaller and smaller amount of time? How exactly can it make our lives better? The applications are near-infinite, but here are four specifics with huge potential to impact your life:

A MORE ACCURATE WEATHER FORECAST

Meteorologists don't always get it right. Sometimes that means a wedding planned on a rainy day. Other times, it can put lives in jeopardy by misjudging the time or place a storm is predicted to hit. But supercomputers are closing

the gap between forecasts and actual weather events. Today's meteorologists depend on incredibly powerful supercomputers to predict where a storm is headed. The more powerful the computer, the more equations it can solve simultaneously and the more accurate the prediction will be. The National Oceanic and Atmospheric Administration (NOAA) has recently announced a \$44.5 million investment in supercomputing to increase forecast accuracy. "We continue to make significant, critical investments in our supercomputers and observational platforms," Louis Uccellini, PhD, director of NOAA's National Weather Service, says. "By increasing our overall capacity, we'll be able to process quadrillions of calculations per second that all feed into our forecasts and predictions. This boost in processing power is essential."

PREVENTING THE NEXT GREAT DEPRESSION

In 2030, we may suffer a disastrous population decline and economic collapse. At least, that's what research by Australian physicist Graham Turner predicts in his study based off of a 1972 academic report that used supercomputers to predict outcomes according to changing populations and resource consumption.

In 2012, Turner said that the economy is right on track for disaster as predicted more than 40 years ago. However, there's still time to commit to more sustainable practices and delay catastrophe. If the economy really is on track for collapse, the supercomputer predictions could serve as a valuable warning sign.

MAKING TRANSPORTATION AS SAFE AS POSSIBLE

According to the Insurance Institute of Highway Safety, more than 30,000 people die in car crashes every year in America. To a large extent, collisions occur when unsuspecting drivers come face-to-face with unexpected obstacles. But what if a small-scale supercomputer in a car could sense these things and immediately maneuver away?



That's exactly what autonomous car hopefuls aim to create, and they're already making great strides. An autonomous vehicle recently drove from Silicon Valley to Las Vegas — a trip more than 500 miles long — using an autopilot feature.

The system relied on two laser scanners, a front-mounted 3-D camera, four side-mounted cameras and a main supercomputer. The supercomputer would constantly predict driving outcomes on the road and react accordingly, dodging any potential dangers at the same time. It's just one more step toward the self-driving cars of the future that will rely on supercomputers, which are immune to distractions on the road.

IMPROVING HEALTH CARE

Since the outbreak of avian flu 10 years ago, scientists have used supercomputer simulations to understand the efficacy of different vaccinations against the most threatening and most common flu viruses.

“These simulations can take days or even weeks to complete. Therefore, having more processing power is essential to making routine flu vaccinations more effective and protecting the public against the most deadly strains of the flu,” Yates said.

Some supercomputers are becoming successful diagnosticians, too.

If you missed the “Jeopardy!” showdown between renowned champion Ken Jennings and IBM supercomputer “Watson,” here's the gist: man was no match against the machine. Watson crushed the competition, collecting \$77,147 compared to Jennings's \$24,000. IBM also understands the massive impact that computers can have on our collective health. After its 2011 win, Watson left showbiz and took up a career in medicine. Two forms of Watson can be rented by hospitals and insurers: one which siphons through thousands of health records to help diagnose and treat lung cancer, and another that helps insurance companies make decisions.

Supercomputers, which give actual diagnos-

tic and treatment advice to doctors on top of displaying numbers and facts, will be an important part of improving health care in the near future.

WEATHER FORECASTING

3 WAYS BIG DATA, SUPERCOMPUTING CHANGE WEATHER FORECASTING



Early, granular, and accurate forecasts could benefit everyone from governments to retailers and healthcare providers.

“You can't control the weather.” It's a phrase we often utter as we plan weddings, family picnics, and beach vacations. While it's true that we can't control the weather, scientists and technologists are working hard to improve our ability to predict it.

What that means for enterprise IT isn't just the ability to foresee disasters that could wipe out a data center. Granular and early forecasts can create business opportunities -- for example, alerting a brewery to supply distributors and points-of-sale greater inventory in anticipation of an unusually warm spring weekend in the Northeast. Or suppose a blizzard is due to hit Atlanta on Black Friday -- with enough notice, retailers could adjust their plans.

Beyond the burgeoning data services industry, weather has massive economic and safety implications. Weather Analytics, a company that provides climate data, estimates that weather impacts more than 33% of worldwide GDP, affecting the agriculture, tourism, fishing, recreation, and airline industries, to name just a few.

Uncertain weather conditions also impact small business owners, such as the local painter who can't complete his job on time when foul weather persists. In addition, public safety is of vital concern when officials aim to understand the impact of extreme weather events such as hurricanes, tsunamis, or wildfires.



WHILE WE'RE NOT ABLE TO CONTROL THE WEATHER, **BETTER FORECASTING WILL ALLOW US TO MAKE MORE INFORMED PLANS THAT CAN LIMIT FINANCIAL LOSSES**, PROVIDE NEW BUSINESS OPPORTUNITIES, REDUCE GOVERNMENT SPENDING, AND EVEN SAVE LIVES.

Costs associated with extreme weather across the world totaled more \$125 billion in 2013, and the frequency of these events is on the rise. Beyond the private sector, without specific data about a forthcoming event, governments may waste money casting a net that's too wide. For example, every additional mile of evacuation associated with an impending storm can result in millions of dollars in lost revenue, wages, and relocation expenses. And hospitals and emergency facilities that anticipate a severe storm that could shut down power could stock up on extra fuel for generators. So while we're not able to control the weather, better forecasting will allow us to make more informed plans that can limit financial losses, provide new business opportunities, reduce government spending, and even save lives.

Unfortunately, improving our ability to predict the weather is challenging, both scientifically and computationally. Supercomputing has played a major role in enabling predictive models since the 1950s and remains at the cornerstone of today's weather and climate modeling. Constantly improving computational capabilities have allowed scientists and forecasters to produce results faster than ever while also investigating increasingly complex phenomena and producing specialized forecast products. From model performance to system and data management, weather prediction presents unique high-performance computing challenges.

Supercomputing, along with big data, can meet the future demands of weather forecasting in three key areas:

1. MANAGING AND UTILIZING ENORMOUS DATA SETS:

The volume and diversity of environmental data is increasing exponentially, placing great demand on the infrastructure to transport, manage, and store this data, and requiring ever-greater computational power for simulations that use it. This creates new opportunities for specialized services, developed with researchers in public and private institutions. One example is leveraging new sources of observation, such as sensors placed on automobiles. Imagine thousands of sensors in an urban area providing real-time meteorological information. Models are also evolving to analyze this tsunami of data and augment traditional physics-based simulations.

2. INCREASING MODEL RESOLUTION: Higher-resolution models are a critical element to better estimate the long-term state of climate systems and to improve weather forecasting, particularly for severe weather events. Recent simulations of Hurricane Sandy by researchers at the National Center for Atmospheric Research and the University of Illinois using Blue Waters supercomputers have zeroed in to a 500-meter resolution -- the equivalent of a few city blocks.

3. ADDRESSING TECHNOLOGY HURDLES:

As weather modeling and analytics become more data-intensive and computationally demanding, researchers must watch for performance bottlenecks such as memory, I/O, and interconnect latencies and bandwidths. Weather simulation requires thousands of microprocessors to run in parallel, pushing hardware and software to its scalability limits. In



addition, scalable operating systems, compilers, and application libraries play an essential role in achieving sustained performance. Ultimately, the underlying technology infrastructure must be tightly integrated to support simulation and analytics workflows.

Infrastructures offering simulation and data-driven analytics capabilities to support routine execution of high-resolution forecasts will combine with advanced research to promote a whole new array of specialized meteorological services for public and private sectors. The future of weather forecasting requires capabilities we couldn't even conceive of when we began predicting the weather 64 years ago. Supercomputing innovation has so far kept pace with the demands of the community, and it is poised to offer new solutions in the years to come.

ABOUT THE AUTHOR

Per Nyberg is Senior Director of Business Development at Cray, responsible for worldwide strategic planning, market development, and solutions development across Cray's vertical segments.

STORAGE



In 2015, DDN achieved another successful year of storage innovation in the Enterprise, Big Data, Cloud and HPC markets, and we continue to see high demand from companies that want to accelerate their applications and performance. As we look ahead, the emergence of hyper-converged storage solutions and the explosion of flash storage, are paving the way for more advanced approaches to solving analytics, workflow and data access needs at large scale. Here are some of the storage trends we expect to see in 2016.

HYPER-CONVERGED STORAGE SOLUTIONS ABOUND

The storage industry will take the lead on converging the newest components that drive speed and density into a “hyper-converged”

infrastructure platform. This platform will take full advantage of advanced technologies to achieve the highest performance and lowest latency in a form factor that is extremely efficient, easy to maintain and will scale without bound. This hyper-converged infrastructure will be more about the performance gain you get providing faster time-to-results and a competitive edge, than just about simplicity and cost saving.

OPTIMIZED STORAGE HARDWARE MAKES A COME BACK

Flash is being used to increase performance for many enterprise workloads. A lot of end users are embracing flash-based solutions, while ignoring the cost-efficiency side of the equation. Maximizing efficiency and performance of flash requires the optimization of hardware as well as software. Organizations will increasingly recognize the importance and necessity of purpose-built architectures to encompass greater efficiency and exploit the full capabilities of their expensive flash investments.

HPC MOVES INTO THE ENTERPRISE

To meet the evolving requirements of performance and scale, storage, data management and application acceleration technologies from the high performance computing (HPC) industry are being tapped at an accelerated rate and will replace traditional IT infrastructures at an even a faster pace in 2016. Large-scale commercial enterprises are learning from the successes of early pioneers in the adoption of HPC technology into enterprise to solve a myriad of challenges: rapidly growing data sets, “Big Data” democratization, analytics initiatives, and data sharing across geographies. These enterprises will accelerate their adoption of storage solutions that historically have almost exclusively been deployed in supercomputing labs.

END-TO-END, CONSOLIDATED STORAGE SYSTEMS TAKE ROOT

As storage buyers look to achieve performance, capacity and cost objectives, they increasingly



do so by consolidating storage silos into much larger, shared data depots or data lakes. These will be automated, tiered systems that will handle data across a wide and increasingly varied set of applications, including the ability to ingest, process and archive data from multiple locations in a single, unified namespace. Environments where this preference will be particularly pronounced include Finance, Life Science, Manufacturing, Video Surveillance and Oil & Gas.

FLASH BLURS THE LINE BETWEEN CACHING AND STORAGE

Flash will increasingly be placed outside of the storage array and closer to compute to further reduce latency and increase application performance by orders of magnitude. This move will continue to blur the line between caching and storage. The flash layer will naturally become more persistent as performance-intensive applications rely upon this ultra-fast tier to deliver results at speeds that outpace their competition.

PROLIFERATION OF FLASH MEDIA TYPES

Organizations are deploying flash to accelerate application and I/O performance. Storage vendors will have to look beyond applications and accelerate entire workflows—which means selecting very specific flash media, whether that be NVMe, PCIe and SSD, among others.

DDN is poised to be at the forefront of these trends. We remain focused on delivering the next generation of storage innovations that will help our customers solve their most complex data and analytics challenges faster and more efficiently than ever thought possible. Here's to a great 2016!

ABOUT THE AUTHOR

Molly Rector is CEO, EVP Product Management and Worldwide Marketing at DDN Storage

STORAGE AND HYPERCONVERGENCE



PARALLEL PROCESSING IS A "GAME-CHANGER"

The greatest innovations of the future are often built on those of the past. Flash back to the 1970's during the birth of the modern microprocessor when along with the help of Moore's Law, microprocessors set the stage for two major paths of technology advances. The first resulted in faster, more efficient uniprocessors which led to the PC revolution and to the use of microprocessors today in everything from smartphones to intelligent devices. The second was parallel computing, which set out to harness the power of "many microprocessors" working together.

However, parallel processing advances were stifled by a lack of affordable commoditized parallel computing hardware, the rapid pace of advances in uniprocessor clock speeds resulting from Moore's Law, and by the lack of software to run parallel workloads without the need to rewrite applications or require specialized / exotic gear. So although parallel processing power became available, the parallel computing revolution as a generic IT capability was put on hold because there was no software available to make it happen.

Fast forward to today, and new breakthroughs in software are on pace to make 2016 a turning point year for parallel computing. Application performance, enterprise workloads and the ability to achieve greater consolidation densities on virtual platforms has been held back for years by the growing gap between compute and I/O. Multicore processing has evolved and has made parallel processing platforms a commodity that remains untapped without software to do so. That too has changed.

Parallel I/O software can schedule I/O from virtualization and application workloads effectively across today's readily available mul-



ticore server platforms. It can and will overcome the I/O bottleneck currently holding back the industry by harnessing the power of multicores to dramatically increase productivity and reset what can be accomplished by consolidating more workloads and reducing inefficient server sprawl. This will allow much greater cost savings and productivity by taking consolidation to the next level and allowing systems to truly “do far more with less.” As a result, we’re now on the cusp of that promised parallel computing revolution.

In 2016, there will also be dramatic performance and productivity gains that will transform hyper-converged and Software-Defined Storage as the need grows for solutions that bring ease-of-use benefits with technologies that can also be easily integrated within company infrastructures for both existing investments and future technologies. Although there has been much hype around hyper-converged lately, the technology creates new silos to manage and the current offerings present several limitations, especially when it comes to scaling and performance to handle enterprise-class workloads. Users should be able to advance to the next stage of hyper-converged technology deployments where performance and interoperability with the rest of their investments aren’t sacrificed.

SERVERS ARE THE NEW STORAGE

Only a Software-Defined Storage layer combined with parallel I/O software can effectively manage the power of multicore servers, migrate and manage data across the entire storage infrastructure, incorporate flash and hyper-converged systems without adding extra silos, and effectively utilize data stored anywhere in the enterprise or in the cloud. By untapping the power within standard multicore servers, data infrastructures will realize tremendous consolidation and productivity benefits from parallel I/O technologies.

As this convergence of software and servers is realized, servers will continue to become the dominant model for doing storage – “ser-

vers are the new storage.” Why use specialized gear when smart software and today’s servers can drive greater functionality, more automation and provide comprehensive services to productively manage and store data across the entire data infrastructure? Because traditional storage vendors with specialized systems can no longer keep up with Moore’s Law and the pace of innovations and cost savings that generic server platforms deliver, these new server-based storage systems have and will continue to transform our industry and significantly make IT more productive.

PARALLEL I/O AND SOFTWARE-DEFINED STORAGE ARE REVOLUTIONIZING IT INFRASTRUCTURES AND PRODUCTIVITY

Best of all, this revolution is happening right now. DataCore recently set the new world record on price-performance and did it on a hyper-converged platform (on the Storage Performance Council’s peer reviewed SPC-1 benchmark). DataCore also reported the best performance per footprint and the fastest response times ever. Bottom-line, today’s multicore servers and software can “do far more with less” and dramatically change the economics and productivity one can achieve. See for yourself the latest benchmark results: [Top 10 SPC1 Price-Performance Results](#) and feel free to learn more by checking out the new DataCore ebook: [How to Unlock the Full Power of Your Servers with DataCore’s Adaptive Parallel I/O Software](#).

Parallel I/O software in combination with Software-Defined Storage will lead to a new revolution in productivity – and together they will revolutionize IT in 2016.

ABOUT THE AUTHOR

George Teixeira is President and CEO, DataCore Software



HPCToday

The global media in high-performance IT

Subscribe now > for FREE

Subscribe now and receive, every month, an expert, actionable coverage of HPC and Big Data news, technologies, uses and research...



+ get yourself instant access to exclusive contents and services

www.hpctoday.com



CONTROLLING THE TIDE OF WEATHER AND CLIMATE DATA WITH TIERED STORAGE

There's an old cliché that everyone talks about the weather, but no one does anything about it. While Cray can't (yet) prevent droughts or cool off hot spells, we can help make the lives of weather professionals easier.

AN ABUNDANCE OF DATA, BUT WHERE TO STORE IT?

Weather and climate modeling centers strive to improve the accuracy of their models by gathering and assimilating more diverse input data and by increasing model resolution and complexity. As a result, these models are in-

gesting and producing ever-increasing volumes of data. These weather and climate organizations often find themselves challenged by the sheer volume of data, trying to manage various ways it may be used and simply trying to find the resources, financial and otherwise, to store and access it over the long term. Many

IF A DATASET ISN'T USED ANY MORE, THIS IS DETECTED, AND IT'S AUTOMATICALLY MIGRATED TO A TAPE LIBRARY FOR LONG-TERM SAFEKEEPING UNTIL IT'S NEEDED AGAIN.

numerical weather prediction centers have tens to hundreds of petabytes of data, and are facing annual data growth rates in excess of 40%. This level of data growth makes a data archive solution a critical component for maintaining order and producing the reliable weather and climate forecasts the world needs. Confronted with massive amounts of data, researchers need to easily archive it and then just as easily retrieve it for quicker model refinement and analysis. A cost-effective and easily manageable archive solution is demanded, one that keeps data readily accessible. Typical archiving solutions often fall short: proprietary hardware compromises data movement, scalability is inadequate for increasing data volumes, and integration into workflows requires expensive support services. Any one of these factors can obstruct a numerical weather prediction or climate research workflow.

TRANSPARENTLY BLENDING FAST DISK AND ECONOMICAL TAPE

Historically, weather centers have used a variety of storage technologies for online and archival storage. It often involved the manual construction of elaborate processes to copy data from one to the other and back, locking organizations into barely functional process and products for lack of better alternatives. By thinking of storage, online and archival, as a unified whole, intelligence can be built in to keep the complexity under the surface and out of the way of users. This intelligence is provided by a hierarchical storage system (HSM), which blends fast disk and economical tape into a single transparent filesystem. Users can see and access all files in their huge libraries, but only the files in use are kept on fast, expensive disk. If a dataset isn't used any more, this is detected, and it's automatically migrated to a tape library for long-term safekeeping until it's needed again.

Existing and often proprietary archive systems tend to lock data into internal formats, making it very difficult to retrieve in a future beyond the system's lifetime. By comparison, open archival systems like Cray® Tiered Adaptive Storage (TAS) store their archives in well-documented formats that should be easily retrievable in a future era.

Even with all the current hype about disk being dead, tape libraries still offer the lowest storage cost per terabyte and consume 10x less power than disk-based solutions. As data and the need to process it expands, the need for larger facilities with more power and cooling capacity is often overlooked. By keeping archives in compact, green tape libraries, at least some of this pressure can be mitigated.

Data residing on tape isn't a solution by itself. In order to be effectively used, it must be accessible to the compute infrastructure. Scripting is often used to orchestrate large-scale copying from archive to online storage, but this is no longer sufficient. By using a modern open archiving solution like TAS, all files — archive and active — are presented to users as a single file system, leaving the magic behind the scenes. TAS also has the virtue of integration with high-performance Lustre® storage, allowing transparent movement from the fastest to the slowest storage without manual intervention by a systems administrator.

By bringing all weather and climate data, including vast archives, under the management of an open archival system like Cray TAS, weather and climate centers can accomplish more research, refine more models and simplify operation of their facilities.

JEFFREY KATCHER

IS BUSINESS DEVELOPMENT MANAGER IN CHARGE OF SOLUTIONS AND STRATEGIES AT CRAY, INC.



HOW TECHNOLOGY WILL DRIVE THE NEXT WAVE OF CIRCULAR AND SERVICE-BASED ECONOMICS

The main topic of discussion at the World Economic Forum in Davos in January this year revolved around the impact of digital technology and the Internet of Things, which is set to revolutionize the way businesses serve their customers. In this new age, there are great commercial opportunities for those that utilize this new technology effectively.

B

usiness leaders today face unprecedented challenges, in part due to the astonishing rate of technological advancement. At the moment, much of this potential is being lost because the data is not being fully utilized. However, if businesses do draw on it effectively, the new industrial wave could generate up to \$11.1 trillion a year by 2025,

according to McKinsey^[1].

TECHNOLOGY UNLEASHES NEW POTENTIAL

Harvard professor Michael Porter, a globally recognized authority on competitiveness, believes the Internet of Things is reshaping business and society. The rise of smart, digitally connected products will, he writes in Harvard Business Review, “generate real-time readings that are unprecedented in their variety and volume.

Data now stands on par with people, technology, and capital as a core asset of the corporation and in many businesses is perhaps becoming the decisive asset.”^[2]

While technology might unleash a new industrial revolution, we are also living in an age where our economic activities are threatening the planet on which we all depend. The great challenge, therefore, is for business leaders also to develop models that draw less and less on finite, non-renewable resources such as metals and minerals and reduces carbon emissions, offsetting two of the biggest environmental challenges we face: resource scarcity and global warming.

SUSTAINABLE ECONOMIC GROWTH IS CIRCULAR

There is a way forward which addresses both these issues and could help businesses to both engage their customers and develop a more environmentally friendly approach to commerce known as the circular economy, in which goods are designed to be recycled, reused, refurbished

THE CIRCULAR ECONOMY CONCEPT, INFORMED BY SUCH IDEAS AS CRADLE TO CRADLE, IS ALSO ABOUT DRIVING TOWARDS A LOW CARBON ECONOMY AND UTILIZING HEALTHY MATERIALS TO MAKE PRODUCTS, ALL OF WHICH LENDS ITSELF TO CREATING POSITIVE CLOSED LOOP MANUFACTURING SYSTEMS.

or remanufactured, extending their product life cycles. The circular economy concept, informed by such ideas as cradle to cradle, is also about driving towards a low carbon economy and utilizing healthy materials to make products, all of which lends itself to creating positive closed loop manufacturing systems.

This is the ideal. But how do we get there? The answer could be through developing service-based models, facilitated by servitization, in which customers demand smarter, added value services. Instead of selling an asset in a one-off transaction, the supplier provides a service such as selling the service of light rather than supplying lighting products or providing on-going medical scans rather than a scanner. This model is already familiar in some sectors such as office technology.

CUSTOMER SERVICE IS AT THE HEART OF THE NEW ECONOMY

But the revolutionary shift is that more and more suppliers can go down this road. Porter describes how the service-based model, underpinned by the Internet of Things, will boost customer engagement: “The ability to remain connected to the product and track how it’s being used shifts the focus of a company’s customer relationship from selling—often a predominantly onetime transaction—to maximizing the customer’s value from the product over time. This opens up important new requirements and opportunities for marketing and sales.” They will maintain ownership of the product and will often be responsible, as part of the deal, for servicing it during use. The In-

ternet of Things means that most suppliers will be able to monitor and keep in constant touch with their equipment digitally, facilitating an enhanced service: they can optimize the equipment, predict problems before they occur and fix them without breakdowns and radically improve minute by minute, their understanding of customer behaviour and needs.

The second thing this service model does is enable the producer to maintain control over the assets in service and plan for their return and next life cycle. They will be incentivized to make things that are durable and perform well and that can be remanufactured, refurbished or recycled at the end of their first lifecycle. The process is, therefore, both profitable, competitive and environmentally responsible. Peter Lacy and Jakob Rutqvist, authors of the new book, *Waste to Wealth*, see such service-based models as one of the clear routes to the circular economy and describe how manufacturers should, “consider the entire product lifecycle when setting strategy”. In this scenario, products “must be designed for optimal use, maintenance, reuse, remanufacture, and recycling to avoid issues such as fast quality degradation, short lifespan, low utilization rate, and low recycling/return, which can directly impact the company’s bottom line.”^[3]

MAKING THE SERVICE ECONOMY A CIRCULAR ONE TOO

Putting together these deals is, however, more complex than a straightforward sale of an asset. First, you need to decide on the model to

PUTTING TOGETHER THESE DEALS IS, HOWEVER, MORE COMPLEX THAN A STRAIGHTFORWARD SALE OF AN ASSET. FIRST, YOU NEED TO DECIDE ON THE MODEL TO BE USED SUCH AS PAY FOR USE, LEASING AND RENTAL, AND THEN FIND WAYS TO AGREE ON TERMS THAT ARE VIABLE FOR BOTH PARTIES, AND ENABLE THE SUPPLIER TO TAKE BACK THE ASSET AT THE END OF USE AND RECAPTURE ITS VALUE IN A SECOND OR THIRD LIFE CYCLE. THIS IS AN AREA THAT DLL HAS BEEN HEAVILY INVOLVED IN AND FROM OUR EXPERIENCE, WE CAN SEE THAT THE RIGHT FINANCIAL SOLUTIONS WILL DRIVE FORWARD THESE NEW CUSTOMER-CENTRIC CIRCULAR ECONOMY MODELS.

be used such as pay for use, leasing and rental, and then find ways to agree on terms that are viable for both parties, and enable the supplier to take back the asset at the end of use and recapture its value in a second or third life cycle. This is an area that DLL has been heavily involved in and from our experience, we can see that the right financial solutions will drive forward these new customer-centric circular economy models.

Indeed, Lacy and Rutqvist explain that companies thinking of adopting a service-based model may “need to think creatively about financing, potentially collaborating with financial institutions such as banks and insurers to make it practically viable.” At DLL we are seeing that financial institutions can do more, working as collaborators and partners in the journey to circularity. Lacy and Rutqvist add that more than 80 percent of the service based-models they have studied are “blending it with one or more circular business models”, often paired with Product Lifecycle Extension models to repair or upgrade products. Indeed, we have built a Lifecycle Asset Management division, which facilitates revenue generating second and third life refurbishment business models. It is here that we have been expanding our capabilities to provide asset data, giving our customers more insights into the

usage of their assets through the whole product lifecycle. This service-based model helps customers reduce the costs of ownership, get more efficient usage of their assets, and is beneficial for business and the environment. In this brave new world, financial solutions like our own lifecycle asset management division can provide the wiring for this new infrastructure: the Internet of Things and service-based models linked to the circular economy. And this is driving a revolutionary shift in the way we do business.

REFERENCES

[1] James Manyika, Michael Chui, Peter Bisson, Jonathan Woetzel, Richard Dobbs, Jacques Bughin, and Dan Aharon, *Unlocking the potential of the Internet of Things*, McKinsey Insights, June 2015
http://www.mckinsey.com/insights/business_technology/the_internet_of_things_the_value_of_digitizing_the_physical_world

[2] Michael Porter and James Heppleman, *How smart, connected products are transforming companies Part II*, HBR, October 2015.

<https://hbr.org/2015/10/how-smart-connected-products-are-transforming-companies>

[3] *Waste to Wealth*

For more information on DLL's Lifecycle Asset Management please visit: www.dllgroup.com/LCAM

IDENTIFYING THE KEY FEATURES OF ENTERPRISE-CLASS SSDS

Introducing the Saber 1000 HMS Platform

A solid-state drive (SSD) has many background tasks that occur at the drive level. Host Managed SSD (HMS) technology adds user controls to those drive level background tasks (via APIs) enabling system-level management of SSD operations. By coordinating the background tasks of each and every SSD in the pool, the entire population can share the burden of a given workload, delivering higher performance, greater functionality and better value.

The HMS platform developed by OCZ Storage Solutions will first support its Saber 1000 SSD Series and usable capacities of 480GB and 960GB in a 2.5" x 7mm enclosure. The technology includes a complementary Programmers Guide that enables developers and programmers to control each Saber 1000 SSD in the pool using HMS APIs (Application Programming Interfaces) and includes source code so that any developer or programmer can freely modify the APIs into their software stack. A complementary Reference Design is also avail-

able that provides a foundation from which any custom development could be built from.

Using new primitives, the management layer can control such flash functions as garbage collection, log dumps and other drive-level capabilities. On their own, each of these background tasks represents a considerable tax to drive performance. Therefore, the goal of HMS is to coordinate those tasks across the whole pool to remove those drive performance taxes.

DESIGNED FOR HYPERSCALE DATACENTERS

As hyperscale and large datacenters continue to grow, the cost of administration will increase particularly as more SSDs are attached at the server level. HMS technology delivers central administration across the increasing pools of server attached storage and represents a new area of SSD design that enables IT managers to better control storage administration costs while maximizing SSD pool performance and overall efficiency.

THE NEED FOR HOST MANAGED SSDS

As SSDs gain increasing traction in the enterprise market, storage challenges are raised creating opportunities for technological in-

novation. Case in point, the Fresh Out-of-Box (FOB) performance of an enterprise SSD is not representative of performance when the drive is under constant load. The constant load is referred to as 'Steady State' and as the SSD achieves this status, user I/O is now competing with background operations for drive performance. As solid-state storage has matured as a technology, SSD vendors have produced even more efficient ways of managing background operations through improved designs and optimized firmware architectures. Part of that maturity is an emergence of software managed storage as the market looks to increase performance consistency while reducing cost. HMS is designed to align with that trend by giving the market more tools to manage storage exactly as they choose. In quick review, some of the background operations that SSD firmware manages include the following:

GARBAGE COLLECTION

Garbage collection is a background SSD process by which the SSD controller procedurally determines which flash cells have flagged-old data and either consolidates or erases those blocks to reclaim usable capacities. Only pages with good data in that block are read and rewritten into a previously erased page and those free pages that remain are now available for new data. As the operation releases blocks for incoming writes, it includes a copy of data pages from the blocks that have many dirty pages and erasures of these blocks. As a result, incoming read and write commands are pending to this garbage collection process that could increase latency responses up to a few microseconds. Therefore, the solution is to erase unneeded data in flash, which in turn, optimizes performance and reduces flash wear. As a drive's usable capacity becomes full, garbage collection becomes an urgent process as there are fewer places to store incoming writes.

LATENCY

Latency is a measurement that determines the

time it takes for a packet of data to get from one designated point to another and assumes that data should be transmitted instantly with no delays. Within an enterprise network, a data packet may be subject to storage and I/O access delays from intermediate devices such as switches and bridges. Specific contributors to latency include mismatches in data speed between a server processor and input/output devices such as SSDs, HDDs and others.

When running write I/Os on an SSD over a sustained period of time, the write history evolves and a change in performance becomes evident. The latency changes on each read or write I/O never return to a level of consistency as delivered in the FOB state. Therefore, the solution is to deliver consistent I/O responses and more predictable, efficient I/O access across the system to reduce potential bottlenecks and help improve end-user productivity. From a TCO perspective, I/O consistency and predictability helps reduce overall power consumption and associated cooling costs, as well as maintenance and support costs.

LOG DUMPS

All metadata changes within the Flash Translation Layer (FTL) are logged in device memory, SRAM or DRAM. Once this buffer is saturated, these changes are flushed into persistent media. The time between log dumps depends on the workloads, background garbage collection, trim operations, etc. In order to assure that a log dump will not delay incoming commands, methods are required to stop, resume and force log dumps for the host.

To achieve the best performance, the SSD controller logs all ongoing changes of write commands and for each measured time, dumps the changes to NAND flash memory. During this dump process, the controller is occupied and may cause delays to other commands resulting in performance spikes.

WRITE CLIFFS

A write cliff occurs when NAND flash memory erasures become slower to perform than

BY COORDINATING THE BACKGROUND TASKS OF EVERY SSD IN THE POOL, THE ENTIRE POPULATION SHARES THE BURDEN OF A GIVEN WORKLOAD AND WORKS COLLABORATIVELY WITH THE HOST TO IMPROVE SYSTEM PERFORMANCE AND EFFICIENCY

actual write (program) operations resulting in performance drops and degradation of write operations. This means that if enough write operations are performed, eventually any finite overprovisioned buffer space will be exhausted and at that point, write operations **MUST BE** slowed down to the speed of the erasures.

The write cliff effect is most commonly seen in unmanaged SSDs once they've been operating for a period of time and occurs when all of the flash cells within an individual SSD have been written to at least once. As part of the Read-Erase- Write process, incoming data is buffered until the flash cell it is destined to be written to has been erased and the new data, along with any retained data (if the page is partially filled), is rewritten back to the flash cell (*Author's Note: only entire pages can be deleted, not portions, which is why new data that includes partially filled retained data are rewritten back to the flash cell*). The write cliff causes degradation in write performance and often in overall performance (depending on the queue depth). The Figure 1 graph depicts the write cliff when a 480GB capacity Saber 1000 HMS SSD reaches full capacity of writes.

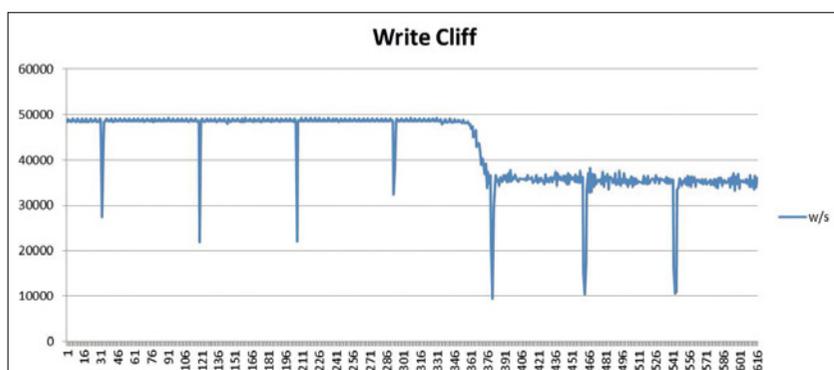


Figure 1 depicts the write cliff when a 480GB capacity Saber 1000 HMS SSD reaches full capacity of writes

THE HOST MANAGED SSD SOLUTION

Adding user controls to SSDs by means of APIs to enable system-level management is the basis of HMS technology. By coordinating the background tasks of every SSD in the pool, the entire population shares the burden of a given workload and works collaboratively with the host to improve overall system performance and efficiency. The system-level knowledge obtained from the internal state of pooled SSDs is opposite to a single SSD HMS concept where no system level knowledge is available. Therefore, the primary use case for OCZ HMS technology is to obtain consistent and predictable latency at low over-provisioning (OP). For many users, consistency is the important parameter. Feedback from large hyperscale customers mentions that 1ms average latency is sufficient but a spike to 6ms for example, would be unacceptable. In enterprise SSDs, this is achieved by adjusting the OP to a high 28%, adversely increasing the average cost of the drive as well. Therefore, the main objective of OCZ HMS technology is to bring entry-level enterprise SSDs to the required HMS working-point;

SABER 1000 HMS OVERVIEW

The main concept of the Saber 1000 HMS solution is to enable storage system vendors to optimize pools of SSDs with an API algorithm of their choosing. These API primitives include Enabling/Disabling HMS Functionality, Starting/Stopping Background Operations, Getting Background Operation Status, Starting/Stopping/Forcing Metadata Log Dumps, Getting Geometry/Endurance/Free Block/ Device Status.

BACKGROUND OPERATIONS THAT SSD FIRMWARE MANAGES INCLUDE GARBAGE COLLECTION, LATENCY, LOG DUMPS AND WRITE CLIFFS

By controlling such flash functions as scheduled garbage collection, log dumps and other system-level capabilities via the management layer, better aggregated efficiency is achieved that translates into improved system performance, increased endurance and improved energy usage. In the future, additional points of control could be added for other background operations such as: Dynamic OP (Over-Provisioning), Dynamic power control or multi-streaming.

In the event of a sudden power loss or failure, Saber 1000 supports Power Failure Management (PFM) so that data written to it will still reside in the drive when power is restored. The data is kept intact (and not lost) for IT management to continue using, and for many hyperscale, web-hosting and distributed computing environments, PFM is all that is required.

SABER 1000 HMS TECHNICAL REQUIREMENTS

This section addresses the HMS technical requirements that affect Saber 1000 SSD performance and latency, and includes techniques on how to enable HMS technology to positively influence background garbage collection, inconsistent I/O responses, log dumps, write cliffs, etc. A quick review of these technical requirements now follows:

OVERCOMING BACKGROUND PROCESSING IMPACT ON LATENCY

One of the main features of HMS technology is designed to avoid background operations affecting performance consistency. The HMS Reference Design assures consistent performance in the aggregated SSD pool using the following APIs:

Primitive	Description
Get Free Pages	Retrieves the number of free pages currently available for programming from the SSD(data that can be written to the SSD without erasing any block)
Disable GC	Stop garbage collection
Enable GC	Resume normal background garbage collection operation

The management layer aggregates the storage devices into a pool and provides volumes over it. The mapping between the volumes and the Saber SSDs is dynamic such that each volume can span over the entire pool of SSDs and change mapping requirements. This is a fundamental aspect of the LUN management structure whereby each arriving write command can be forwarded to every device, balancing the workload across the device pool according to any policy. Thus, for each arriving write command the management layer can apply this rule-base:

- 1** Get the free pages available for each Saber 1000 SSD
- 2** Select the Saber 1000 SSD(s) with maximum free pages (or enough free pages) and write the data

The above rule-base eliminates the chance of writing data that cannot be served due to free blocks shortage. Taking a more active approach, garbage collection can be stopped on some drives, enabling writes and reads from these drives and minimal garbage collection latency impact. Therefore, the HMS rule-base can be modified to overcome the background processing impact on latency as follows:

- 1** Get the free pages available for each Saber 1000 SSD
- 2** Select the Saber 1000 SSD(s) with maximum

THE MAIN CONCEPT OF THE SABER 1000 HMS SOLUTION IS TO ENABLE STORAGE SYSTEM VENDORS TO OPTIMIZE POOLS OF SSDS WITH AN API ALGORITHM OF THEIR CHOOSING

- free pages (or enough free pages) and stop each respective garbage collection operation
- 3** Force garbage collection on the other Saber 1000 SSDs
- 4** Every incoming write command is directed to stopped-GC SSD(s)
- 5** Every read command is fetched from stopped-GC SSD

OVERCOMING ENDURANCE IMPACT ON LATENCY

Flash media is limited by the number of erase cycles that the memory supports before degradation and is determined by the density of the flash itself. As deployed SSDs are subject to malfunctions, errors and failures, this translates to non-successful commands and applications failures. Therefore, each SSD controller applies a mechanism of wear levelling whereby blocks are erased evenly to maintain as little deviation as possible. By levelling erase count across all blocks, the drive write potential is maximized. The following API supports this implementation:

Primitive	Description
Get Erase Count	Get the erase count of maximal/median block
Read Device Geometry	Retrieve information about the device geometry
Read Device Endurance	Retrieve device endurance information
Read debug statistics	Retrieve debug statistics from the device

Therefore, the HMS rule-base can be modified to overcome the endurance impact on latency as follows:

- 1** Get the free pages available for each Saber 1000 SSD
- 2** Select the Saber 1000 SSD(s) with minimal erase count and maximum free pages (or enough free pages) and direct write commands to it.

OVERCOMING LATENCY VARIATIONS DUE TO METADATA UPDATES

On some SSDs, metadata is written frequently as this operation impacts latency. In the Saber 1000 Series for example, this operation occurs about every 14 seconds and its effect is clearly evident. During this metadata write operation, SSD Read/Write performance is impacted and any operation that coincides with this log dump will experience a latency spike. The following API can be used to control the log dump operation:

Primitive	Description
Force log dump	Trigger a log dump operation

System software can use this API to schedule a log dump while preventing any access to the Saber 1000 SSD until the operation has completed.

SUMMARY

Hyperscale, Cloud and Big Data datacenter requirements have diverged from traditional server or SAN/NAS storage to a new usage mo-

MODEL NUMBERS, SKUS A WARRANTY

The Saber 1000 HMS solutions conform to the standard 5-year enterprise SSD warranty and support. The endurance, or 'Life Left,' is tracked using the average P/E cycle count across each Saber 1000 HMS SSD. Each Saber 1000 HMS SSD will remain in warranty until such time that 5-years have expired from the date of purchase, or the remaining 'Life Left' SMART attribute reaches zero, whichever comes

first. Because HMS controls allow the system administrator unfettered control over drive upkeep activities, drive endurance may be exhausted faster than the 5 years of normal rated DWPD. As a result, OCZ has implemented a drive warranty calculation that

guarantees the rated P/E cycle count of the NAND, allowing the administrator freedom to manage storage as they see fit. The Saber 1000 with HMS capabilities has a different name than the standard saber 1000 SSD Series. The model number and SKU are as follows:

Model Number	User Capacity	SKU
Saber 1000 HMS 480GB	480GB	SBHCSK31MT570-0480 Saber 1000
HMS, 960GB	960GB	SBHCSK31MT5D0-0960

THE PRIMARY USE CASE FOR THIS INITIAL HMS OFFERING IS FOR OBTAINING CONSISTENT AND PREDICTABLE LATENCY AT LOW OVER-PROVISIONING

del that adds controls to SSDs. By understanding the system-level information on the internal state of SSDs provides IT managers with the ability to direct I/O accordingly to improve overall performance, endurance or efficiency. In the classic sense of the whole being greater than the sum of its parts, the population of deployed SSDs can share the burden of critical requirements, and by working collaboratively with the host, deliver high value and greater functionality.

The primary use case for this initial HMS offering is for obtaining consistent and predictable latency at low over-provisioning. Reaching the highest IOPS performance is not as important as achieving consistent latency of I/O responses. HMS will become a requirement for SSD vendors enabling IT managers to control/manage these costs and to maximize SSD pool performance. The Saber 1000 HMS solution will enable Toshiba/OCZ to gain early mindshare with hyperscale and large data-center customers and work with them on embedding the HMS enabled API SSDs into their

datacenter management stacks (or HMS reference designs), or for those enterprise customers who want to control their own software management layer.

There are three core development tools that are available with Saber 1000 HMS solutions and include:

- 1** HMS Reference Design (foundation from which a custom solution could be built)
- 2** HMS Programmer's Guide (instruction set to integrate HMS APIs into software stacks)
- 3** HMS Software API Library (adds source code-level freedom to achieve custom solutions)



SCOTT HARLIN
IS MARKETING
COMMUNICATIONS
DIRECTOR AT OCZ

GRANT VAN PATTEN
IS PRODUCT
MANAGER AT OCZ



HPC ACADEMY

Developing High-Performance Computing Expertise in Europe



NONPROFIT ORGANIZATION

Our missions:

- To promote HPC (High Performance Technologies) across Europe
- To develop HPC training courses for industrial and scientific audiences.
- To develop collaborative projects and codes with committed partners
- To initiate the academic and student communities to parallel programming in order to favor career evolutions
- To promote professional insertion among our members
- To encourage those who dare create and innovate.

Our first partner members:



Become a member
www.hpc.academy

NEXT ISSUE



#12
FEBRUARY
29th, 2016

VIEWPOINT
**BUILDING
A MODERN DATA
CENTER**
PRINCIPLES
AND STRATEGIES
OF DESIGN

COVER STORY

**TO INFINITY
AND BEYOND !**

HOW HPC
HELPS INTERSTELLAR
RESEARCH

HOW TO

10 THINGS
TO KNOW
BEFORE
**DEPLOYING
10 GIGABIT
ETHERNET**

 **HPC Media**
The HPC and Big Data Global Media

HPC Media
11, rue du Mont Valérien
F-92210
St-Cloud, France

CONTACTS
editorial@hpcreview.com
subscriptions@hpcreview.com
sales@hpcreview.com

PUBLISHER
Frédéric Milliot

EDITOR IN CHIEF
Joscelyn Flores

EXECUTIVE EDITOR
Ramon Lafleur

CONTRIBUTORS
TO THIS ISSUE
Amaury de Cizancourt
Steve Conway

Roberto Donigni
Ramon Lafleur
Stacie Loving
Solange Belkhayat-Fuchs
Marc Paris

ART DIRECTOR
Bertrand Grousset

SUBMISSIONS
We welcome submissions. Articles must
be original and are subject to editing for
style and clarity