Huawei Intelligent Computing Product & Solutions

Date: February 2021

This is a presentation draft to be completed!

20 Years of Strategic Partnership with Intel
Continuously Providing the Latest Technology and Products

2001

In 2001, signed a strategic partnership.

.....

2016

Joint launch of KunLun, 32-socket mission-critical server.

Set up a joint innovation computing center in Hangzhou.

2017

Jointly hold CTO summits to start strategic cooperation in AI, edge computing, and cloud computing.

2019

Jointly released the second-generation Intel Xeon scalable processor-based FusionServer Pro intelligent server.

2020

FusionServer Pro intelligent server based on the third-generation Intel Xeon scalable processor.

Huawei FusionServer Pro Intelligent Server

V5 2017 → V6 2020 → V7 2022

Intel's latest technologies and products

Cascade Lake Refresh Processor

3rd Generation Intel ® Xeon ® scalable processor

Intel® Optane™ Persistent Memory

FUTUREWEI Technologies

Partners Continues Compatibility Certification & Technical Support

- ✓ Partners such as **VMware**, **SUSE**, **CentOS**, and **Ubuntu** will continue to work with Huawei to provide compatibility certification and technical support for **FusionServer Pro** servers.

On May 23, 2020, VMware announced that Huawei FusionServer Pro 2288H V5 and other servers have passed the latest VMware vSphere 7.0 compatibility certification.

***VMware vSphere, vSAN, and vCenter will continue to provide compatibility certification and technical support for Huawei FusionServer Pro servers.**

Huawei Technologies Co., Ltd.	2488 V5	Intel Xeon Platinum 8200 (Cascade-Lake-SP) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	2488H V5	Intel Xeon Gold 6100/5100, Silver 4100, Bronze 3100 (Skylake-SP) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	2488H V5	Intel Xeon Gold 6200/5200 (Cascade-Lake-SP/Refresh) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	2488H V5	Intel Xeon Platinum 8100 (Skylake-SP) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	2488H V5	Intel Xeon Platinum 8200 (Cascade-Lake-SP) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	2488H V6	Intel Xeon Gold 6300/5300 (Cooper-Lake-SP) Series	ESXi	7.0 U1				
Huawei Technologies Co., Ltd.	2488H V6	Intel Xeon Platinum 8300 (Cooper-Lake-SP) Series	ESXi	7.0 U1				
Huawei Technologies Co., Ltd.	5288 V5	Intel Xeon Gold 6100/5100, Silver 4100, Bronze 3100 (Skylake-SP) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	5288 V5	Intel Xeon Gold 6200/5200 (Cascade-Lake-SP/Refresh) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	5288 V5	Intel Xeon Platinum 8100 (Skylake-SP) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	5288 V5	Intel Xeon Platinum 8200 (Cascade-Lake-SP) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	5288 V5	Intel Xeon Silver 4200, Bronze 3200 (Cascade-Lake-SP/Refresh) Series	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	
Huawei Technologies Co., Ltd.	5288X V5	Intel Xeon Gold 6200/5200	ESXi	7.0 U1	7.0	6.7 U3	6.7 U2	

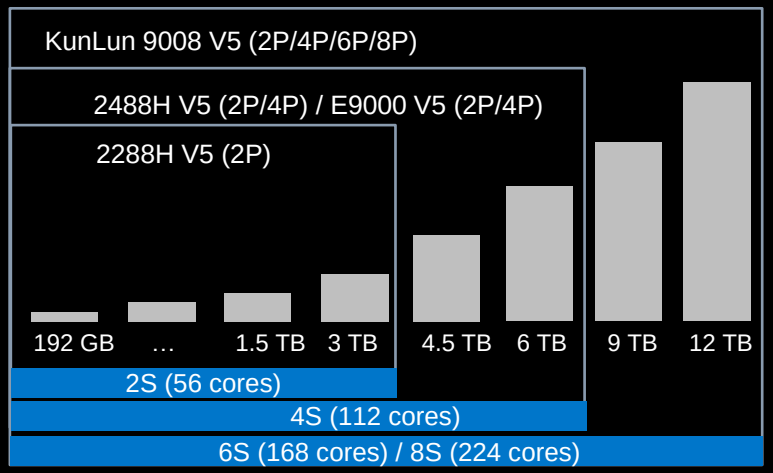
https://www.vmware.com/resources/compatibility/search.php?deviceCategory=server&details=1&partner=242&evcModes=19&page=1&display_interval=50&sortColumn=Partner&sortOrder=Asc

FUTUREWEI Technologies

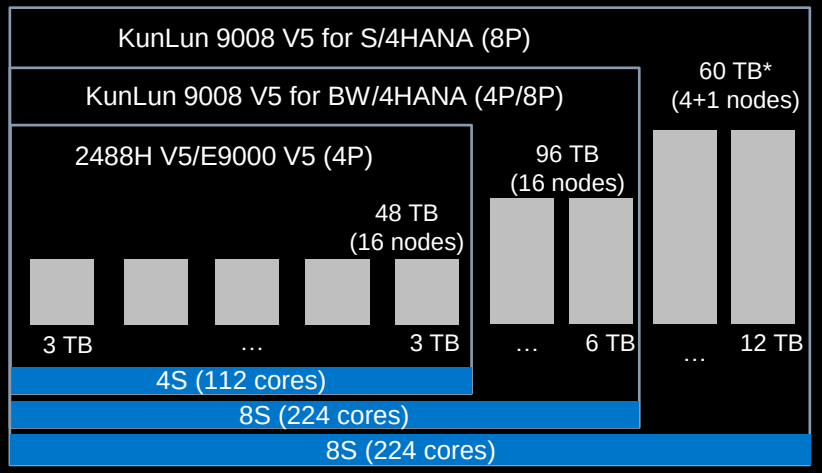
Leading Big Data Solutions with SAP

✓ Our Strategic Partner **SAP** will continue provide compatibility certification and technical support for **FusionServer Pro** servers.

Scale-Up Single-node Solution



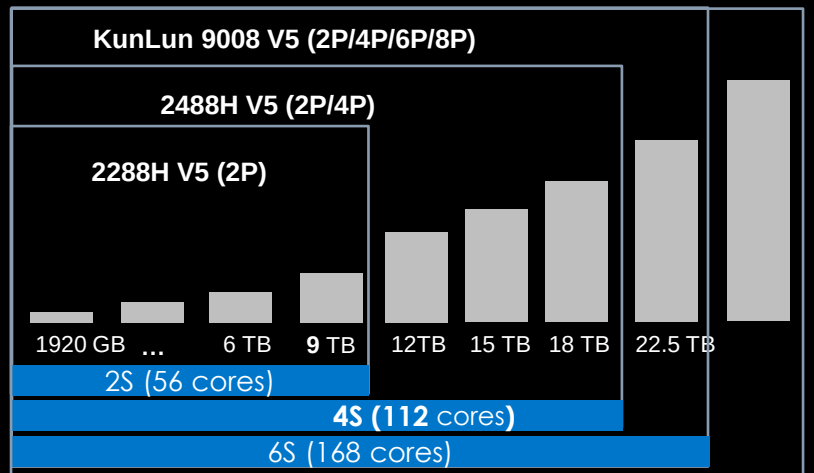
Scale-Out Cluster Solution

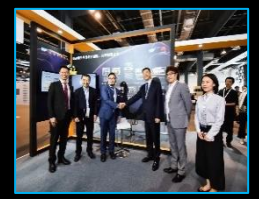


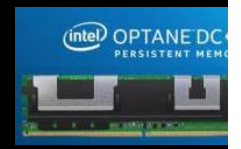
* S/4HANA Scale-out maximum 4+1 nodes, single node minimum size is 6 TB and maximum size is 12 TB

HANA TDI Solution
Support larger memory with Intel PMEM

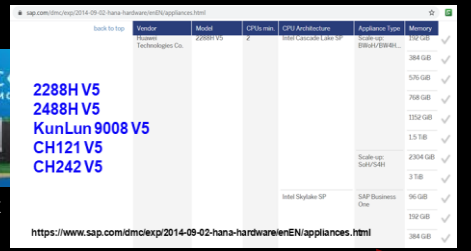
SAP® Certified Hardware for SAP HANA®







512GB/DIMM max
3TB/CPU max



2012 — SAP's first global technology partner in China

2013 — Jointly launched the first server for SAP HANA: Tecal RH5885 V2

2014 — Strategic alliance partnership & launch of FusionCube for SAP HANA solution at CeBIT

2015 — MOU on IoT and joint innovation at CeBIT

2016 — **Top-Class Sponsor at SAP SAPPHIRE in Orlando**

2017 — SAP HANA Innovation Award at SAP SAPPHIRE, Orlando

2018 — SAP as Diamond Sponsor & Key Partner at HUAWEI CONNECT 2018

2019 — Introduction of PMem-based Large Capacity HANA solution

2020 … — Continuous certification and technical support

FUTUREWEI Technologies

Huawei Server Honor List

5+ Million Units

shipped
(2012 - 2020)

400+

SPEC test records

15%

lower failure rate than industry average*

Trusted & Recognized

Best Supplier**
Gartner MQ Challenger

* Source: internal statistics of key accounts in the Internet and finance industries
** Awarded by Tencent and Alibaba

FUTUREWEI Technologies

Winning Extensive Customer Recognition

211 of the Fortune Global 500 companies and 48 of the Top 100 companies have chosen Huawei as their digital transformation partner. Our service products and solutions are widely used in the energy, transportation, finance, and manufacturing industries.

170+
Power companies

211
G500

48
G100

300+
Financial institutes

1,000+
Manufacturing companies

260+
Oil and Gas

220,000+
KM of Highways & Railways

FUTUREWEI Technologies

Rich Portfolio to Meet Your Every Intelligent Computing Need

- ✓ Rich Product Portfolio to meet your needs for different application scenarios:
  - Balanced, Computing or Storage Intensive, Air or Liquid Cooling, Rack or Blade, etc.
- ✓ Full series upgraded to the latest Intel Cascade Lake Refresh / Cooper processor, synchronous to Intel CPU Release Schedule

Mission Critical

KunLun 9008 V5

Full Liquid Cooling Solution

X6000 XH321L V5

E9000 CH121L V5

Converged Blade Servers

E9000

General

CH121 V5

Compute-intensive

CH242 V5

10/25/40/100GE
8/16G FC, FDR/EDR IB

Rack Servers

General

1288H V5

Balanced

2288H V5

Storage-intensive

2298 V5

5288 V5

Compute-intensive

2488/H V5

Mission critical

5885H V5

High-Density Servers

2U 4-node

X6000

Compute-intensive

XH321 V5

XH321L V5

4U 4-node

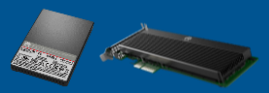
X6800

Storage-intensive

XH628 V5

Intelligent Acceleration Engine

NVMe SSD

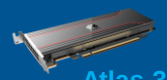


ES3000 V5

Intelligent NIC

IN200/300

AI accelerator card



Atlas 300

Intelligent Management Engine

FusionDirector

5 intelligent management features

Heterogeneous Servers

G5500 full-width server

G5500 half-width server

Nvidia V100 GPU shipping
Ampere GPU coming

FUTUREWEI Technologies

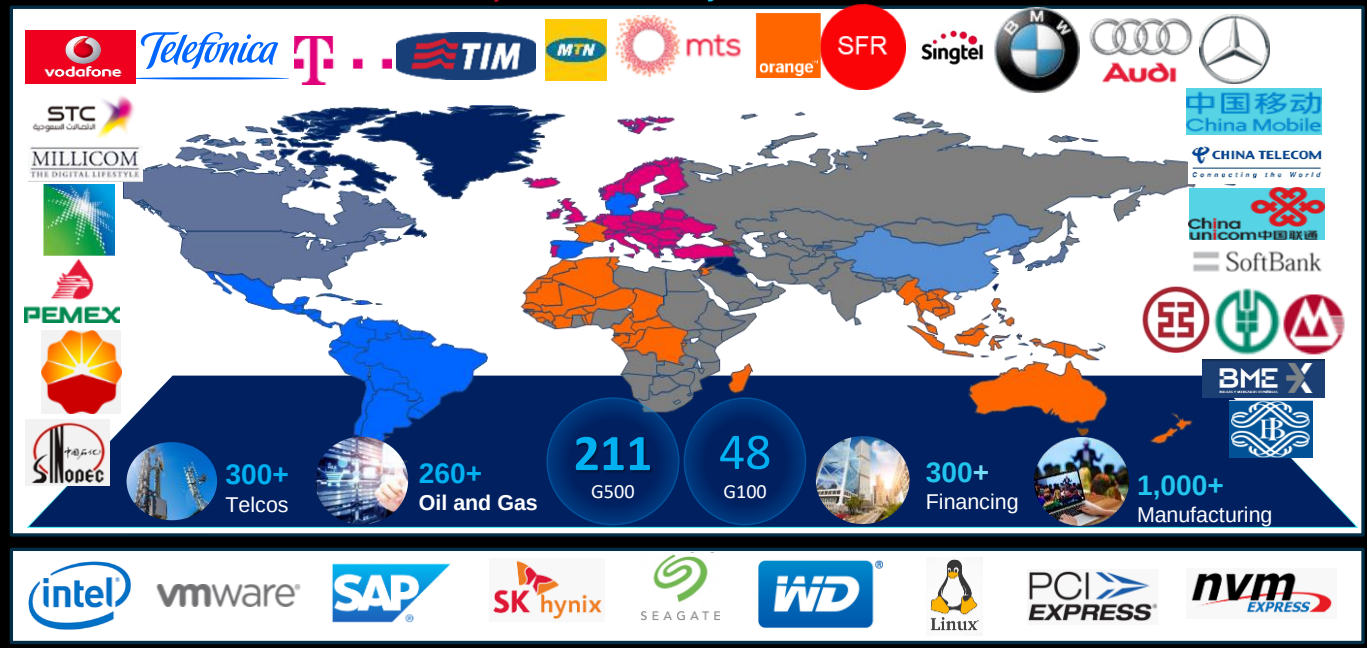
Continuous Evolution Roadmap Synchronous to Intel Release Schedule

HUAWEI

What Customers Considered When Choosing Huawei Computing

- ✓ **Quality Products**
  - DNA of Telecomm Products, E2E Quality Control
- ✓ **Innovative Product Features**
  - Enabled by Self-Developed ASICs for Management, Perf, Deployment & Diagnosis
- ✓ **Knowledgeable & Experienced Team**
  - Both Hardware & Software, Applications & Operations
- ✓ **Long-Term Commitment & Support**
  - Both spare Parts & Software
- ✓ **Local Support**
  - Quick response & Easy to reach
- ✓ **Competitive Pricing for TCO**
  - Not Only CAPEX, but also OPEX
- ✓ **Partnership & Ecosystem**
  - Strategic Partnership with key players, X86 ecosystem
- ✓ **Rich Product Portfolio**
  - No One Fits All, Rich Models to choose
- ✓ **Consolidation of Suppliers**
  - The same supplier for CT & IT, fewer suppliers to Manage

Serves **10,000+** industry customers worldwide



Key Partners & Ecosystem

5+M Units
shipped
(2012 - 2020)

400+
SPEC test records

15%
lower failure rate than industry average*

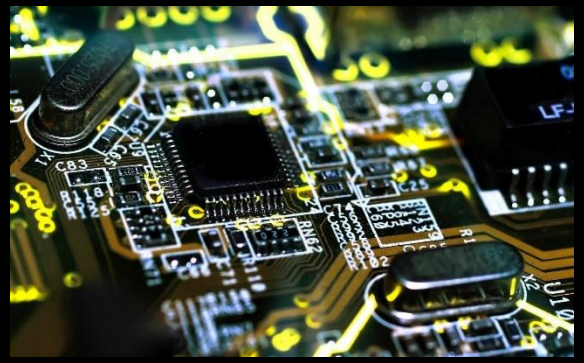
Trusted & Recognized
Best Supplier**
Gartner MQ Challenger

* Source: internal statistics of key accounts in the Internet and finance industries
** Awarded by Tencent and Alibaba

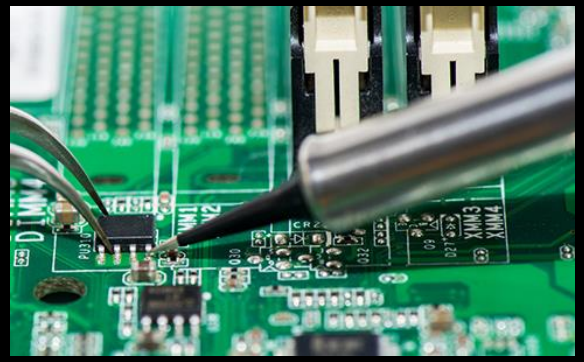
Quality Product is our DNA from Telecom Products

Over 30 years of hardware design, development, and manufacturing capabilities coupled with complete product R&D and test processes to maximize server reliability, reduce downtime and data loss, and enhance device maintainability

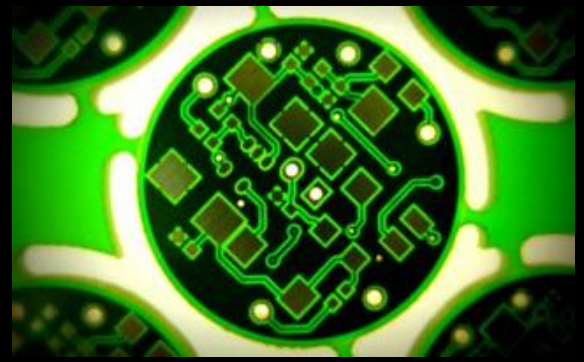
Component Selection and Process Technique



High Reliability Design



Strict Test and Certification



Enhanced O&M Features

Hard Disk Failure rate

40%⬇

Source: Per Internal Testing & O&M Data Collections

Working Temperature

45°C

Source: Internal Lab Tested for most products

Machine Failure rate

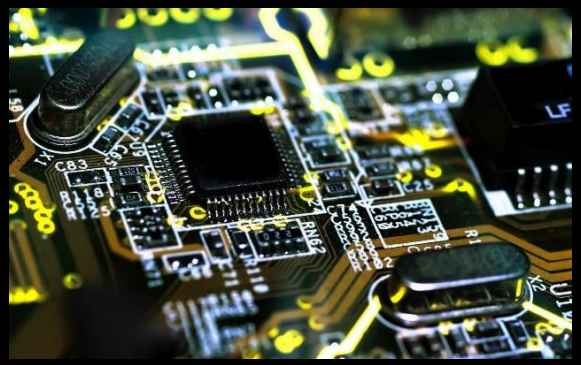
15%⬇

Source: Per a Large Customer Statistics

Quality Product by Design, Process, Sourcing and Testing

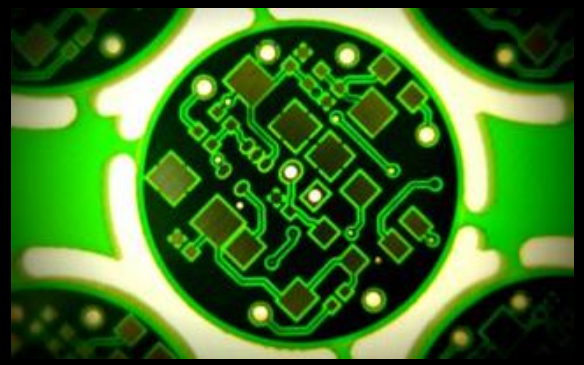
30+ years of hardware design, development, and engineering manufacturing expertise,
combined with leading product R&D and testing processes.

Component Selection and Process Technique



High Reliability Design

Strict Test and Certification



Enhanced O&M Features

Operating temperature **45°C**

Failure rate **30%**

Compared to industry peer's average

FUTUREWEI Technologies

Efficient Heat Dissipation, Stable Operation at 45°C

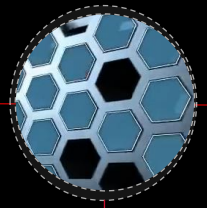
The strict electronic component selection, flow-dividing air ducts and independent air channels, and end-to-end heat dissipation design improve heat dissipation efficiency and prolong the service life of electronic components, allowing servers to run stably even when the air conditioners in equipment room are faulty.

≥30%

Reserved margin for component derating design

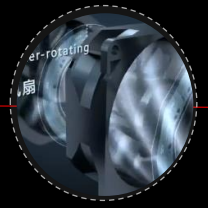
75%

Cellular board design increases porosity rate of the front panel



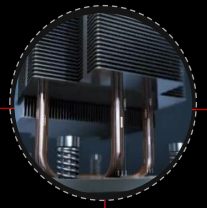
30%

Counter-rotating fans for higher wind speeds



3%

3D VC heat dissipation with higher efficiency than common VC heat dissipation



9%

Flow-dividing air duct and independent air channels for BIO modules improve heat dissipation efficiency

Stable running at high temperature

Related Videos "Efficient Heat Dissipation"

FUTUREWEI Technologies

Triple Anti-Vibration Designs Reduce the Hard Drive Failure Rate by 60%

Triple anti-vibration designs for hard drives: Mechanical hard drives have the highest failure rate among server components. The triple anti-vibration designs for hard drives improve stability and reduce the hard drive failure rate.

0.78%

Industry server drive failure rate

Vibration-damping screws

Vibration source: 8 vibration-damping screws reduce impact of fan vibration on the fan tray

Enhanced rigidity

Vibration path: enhanced chassis rigidity mitigates vibration transfer

Metallic spring washer & high-damping cushion

Receptor protection: vibration-absorbing components added to drive trays

0.31%*

Huawei server drive failure rate

* Based on the fault statistics of 26,000 hard drives from the data center of a major financial customer.

FUTUREWEI Technologies

Accelerating X86 Server Performance

ES3000 series intelligent SSDs

- Optimal performance: PCIe Gen4 1.66M Random read IOPS, 7.0GB/s sequential read bandwidth
- All-scenarios: low latency, high performance, large capacity
- Service-aware: intelligent multi-stream, SR-IOV, atomic write

Performance 100% up

Service life 20% longer

IN200/300 series intelligent NICs

- Converged: 10GE/25GE/40GE/100GE, 8G/16G/32G FC
- Accelerated: RoCE
- Reliable: industry unique to support firmware upgrade without service interruption

CPU offload 15%

Latency 30% less

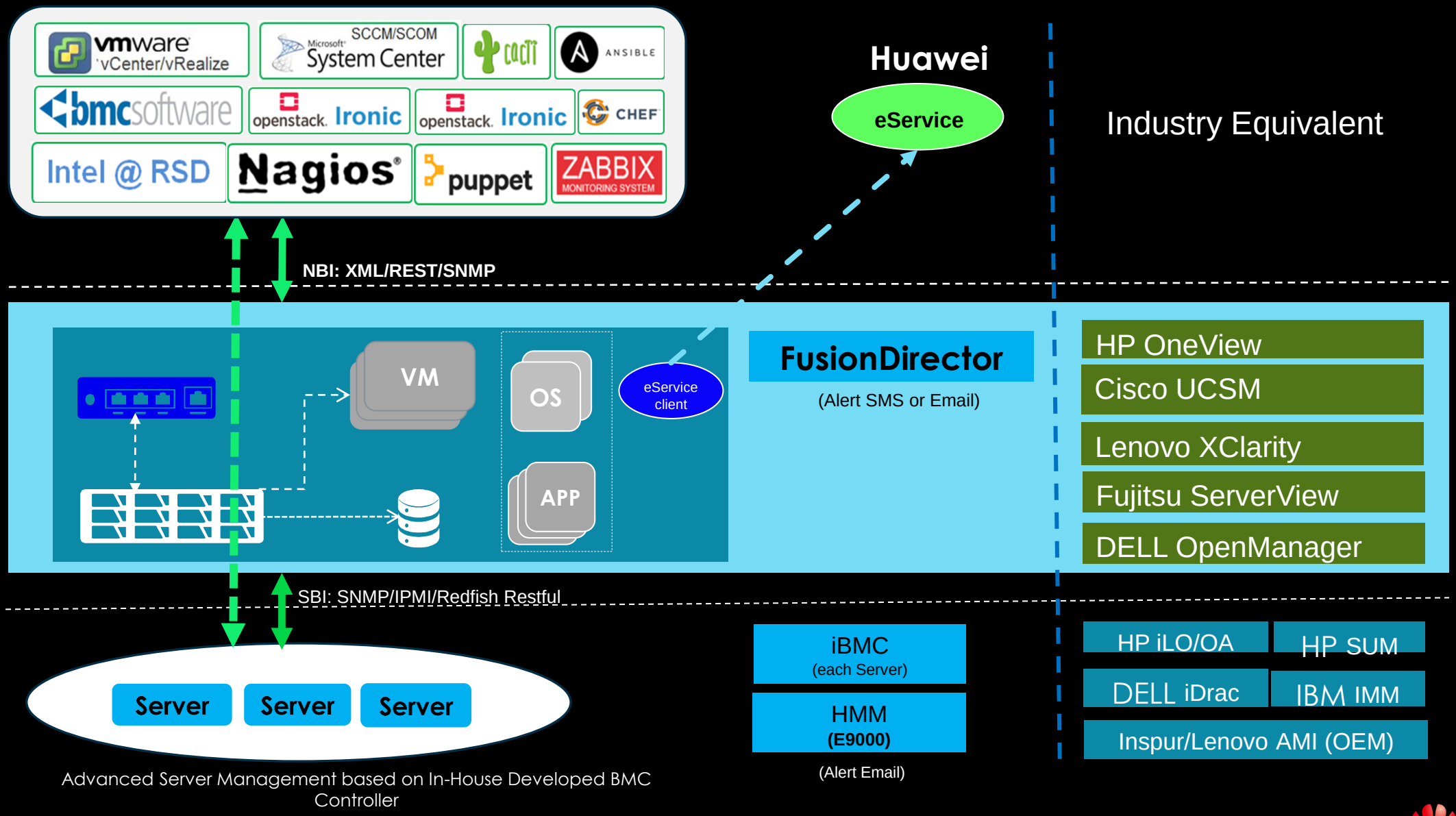
Atlas 300 AI accelerator card

- Powerful computing: 64 TOPS of INT8
- High-density video parsing: 64-channel HD video real-time analytics

CPU offload 15%

Performance 64-Ch HD

Open & Standard-Based Server Management



Advanced Server Management based on In-House Developed BMC Controller

HUAWEI

FusionDirector: 5 Intelligent Management Features Reduce OPEX

Intelligent Asset Management

Server model, configuration, inventorying, and retirement
Inventorying in seconds, improving space utilization by 10%+

Intelligent Deployment Management

Automatic deployment in data center cabinets
10x device rollout efficiency

Intelligent Version Management

Automatic version matching, one-click update
Preparation done in seconds, reducing update steps from 20 to 3

Intelligent Fault Management

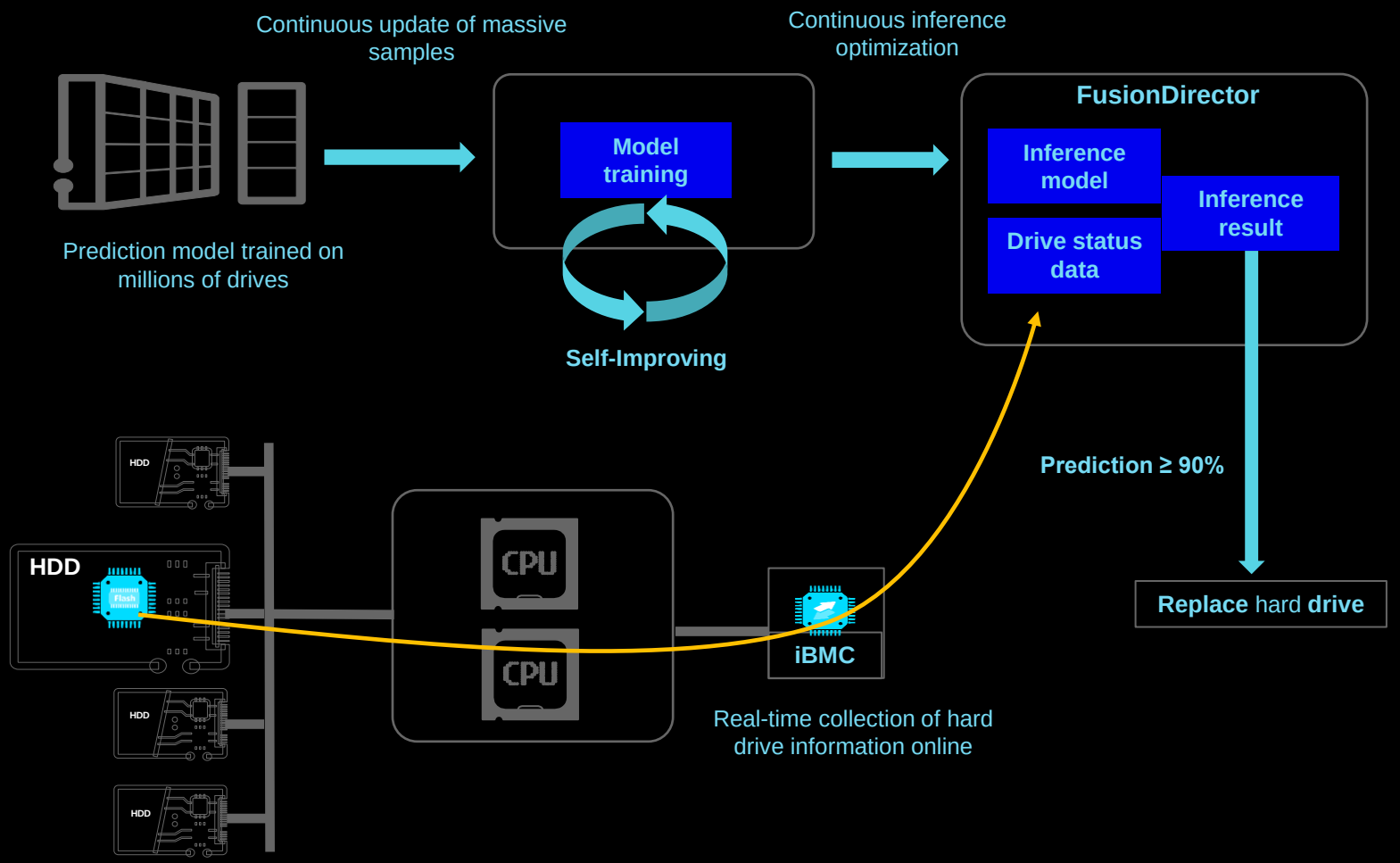
Fault prewarning, diagnosis, and locating
7–30 days in advance for fault prediction,
93% fault locating accuracy

Intelligent Energy Efficiency Management

Server DEMT, cabinet power capping, and data center linked control
10% server energy saving, 20% higher deployment density

HUAWEI

Intelligent Fault Prediction: Predicts Risky Hard Drives 7–30 Days in Advance



Real-time — Real-time collection of drive SMART information for model analysis on the cloud

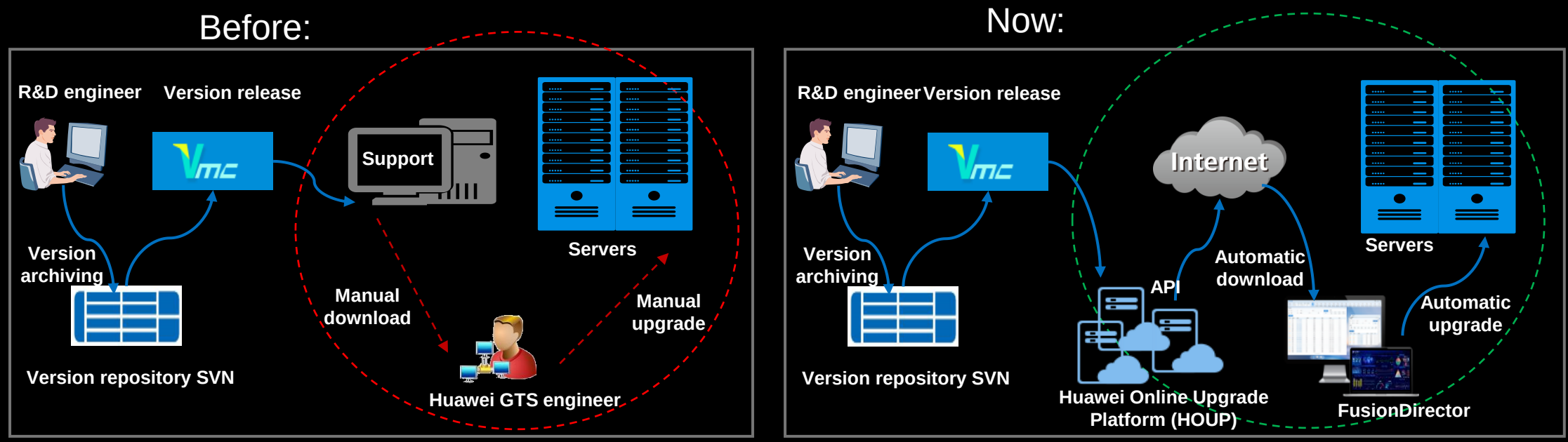
Accurate — Prediction Accuracy ≥ 90%

Reliable — Lower multi-drive simultaneous failure rate

Efficient — 90% less manpower for fault locating

HUAWEI

Intelligent Version Management: Automatic Upgrade Of Device Software



- Automatic download: Automatically detects the firmware and drivers of servers of different models and manages firmware versions comprehensively
- O&M engineers do not need to manually download, upload, or upgrade firmware
- Independent out-of-band upgrade, without occupying service bandwidth

HUAWEI

1288H V5: High-Density Deployment

- Virtualization
- Cloud computing
- High-performance computing (HPC)
- Web application
- Software-defined storage (SDS)
- AI inference

High reliability

- Carrier-class component specifications and 100% derating design
- Redundancy design: PSU modules in 1+1 mode, fan modules in N+1 mode, and memory module sparing/mirroring
- 99.99% availability
- 15% fewer failures than the industry average

Ultimate performance

- 2 x 2nd Gen Intel® Xeon® Scalable processors, up to 28 cores and TDP 205 W per CPU
- 24 DDR4 DIMMs, up to 2933 MT/s; a maximum of 12 Optane™ Persistent Memory (PMem), up to 2666 MT/s; total memory capacity up to 7.5 TB on a single server
- No. 1 in SPECint®_base2006 and SPECfp®_base2006 performance tests
- No. 1 in SPECpower ssj2008 Benchmark energy efficiency test

Unique value

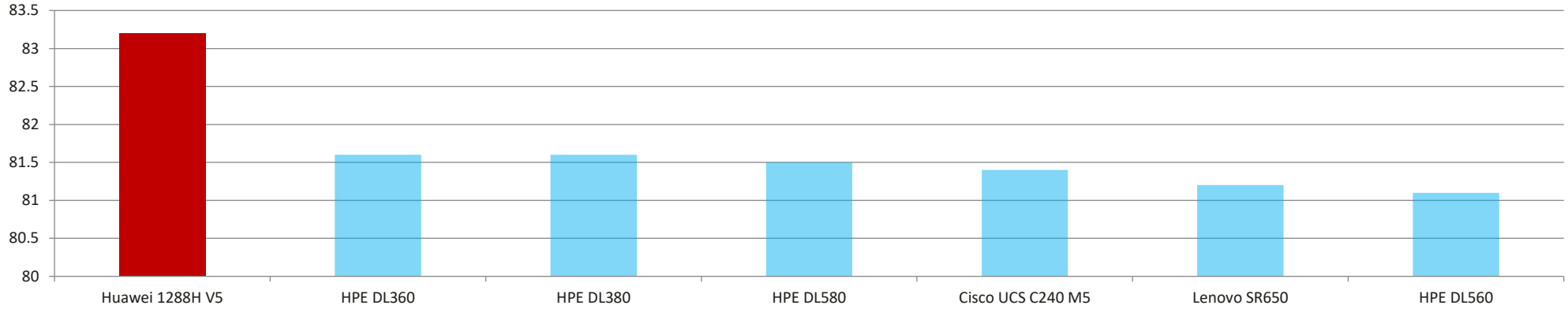
- **FDM:** The patented Fault Diagnosis & Management (FDM) technology provides the fault locating accuracy up to 93% and predicts a fault 7 to 30 days before it occurs, reducing OPEX
- **DEMT:** The patented Dynamic Energy Management Technology (DEMT) reduces energy consumption by 15% without compromising services
- **BSST:** Boot Speedup Storage Technology (BSST), 2 M.2 SSDs as high-speed OS boot drives, industry-unique hot plug and hardware RAID

Innovative design

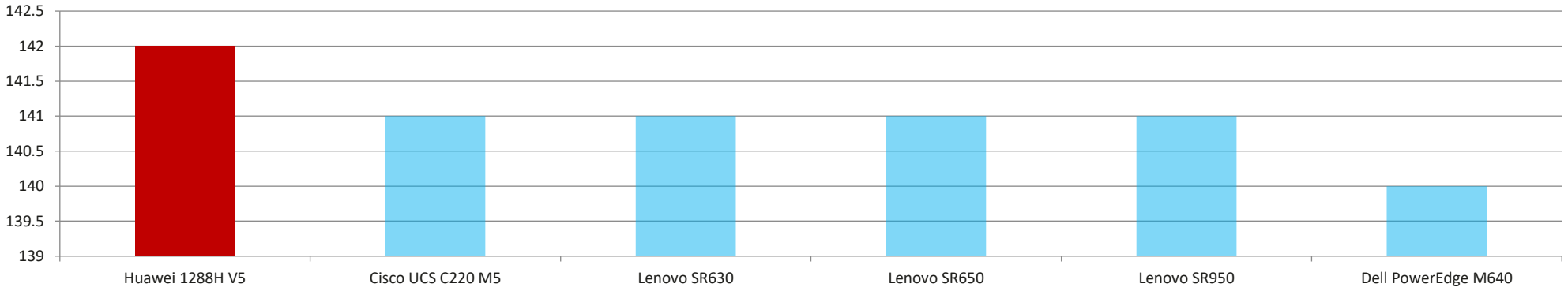
- 2 x 10GE and 2 x GE LOM ports, reducing CAPEX
- Flexible NICs for higher networking capabilities
- 10 x 2.5" SSAS/SATA HDDs or SSDs, providing ultra-large storage capacity
- 8 NVMe SSDs for optimal service experience
- 2 NVIDIA Tesla T4 GPUs for AI inference

1288H V5: Multiple New SPEC Performance Records

SPECint_base2006 Test Result



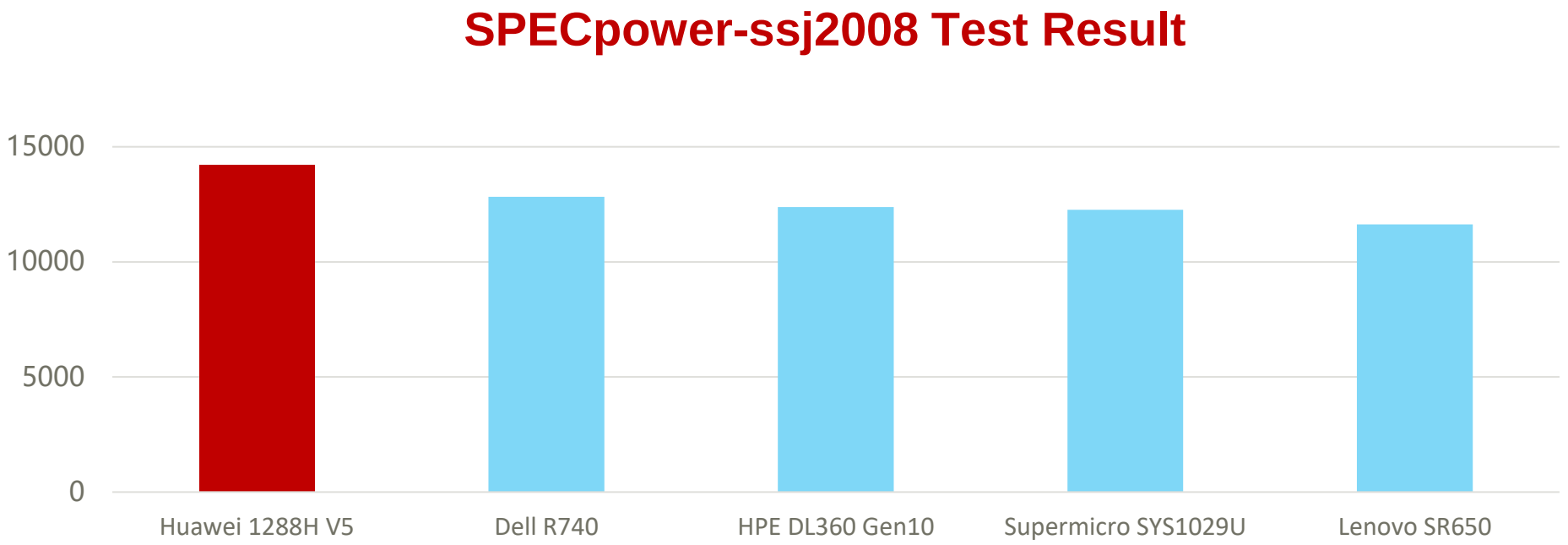
SPECfp®_base2006 Test Result



Data released on http://www.spec.org/ as of Apr 2018

FUTUREWEI Technologies

1288H V5: New SPEC Energy Efficiency Records



Data released on http://www.spec.org/ as of Apr 2019

FUTUREWEI Technologies

2288H V5:Flexible Configurations Adaptive to Various Workloads

- Virtualization and cloud computing
- Virtual desktop interface (VDI)
- SDS
- HPC
- Big data and hyper-convergence
- Entry-level in-memory database (SAP HANA)
- AI inference or training

High reliability

- Carrier-class component specifications and 100% derating design
- Redundancy design: PSU modules in 1+1 mode, fan modules in N+1 mode, and memory module sparing/mirroring
- 99.99% availability
- 15% fewer failures than the industry average

Ultimate performance

- 2 x 2nd Gen Intel® Xeon® Scalable processors, up to 28 cores and TDP 205 W per CPU
- 24 DDR4 DIMMs, up to 2933 MT/s; a maximum of 12 Optane™ Persistent Memory (PMem), up to 2666 MT/s; total memory capacity up to 7.5 TB on a single server
- No. 1 in SPECint®_base2006 performance test
- No. 1 in SPECpower ssj2008 Benchmark energy efficiency test

Unique value

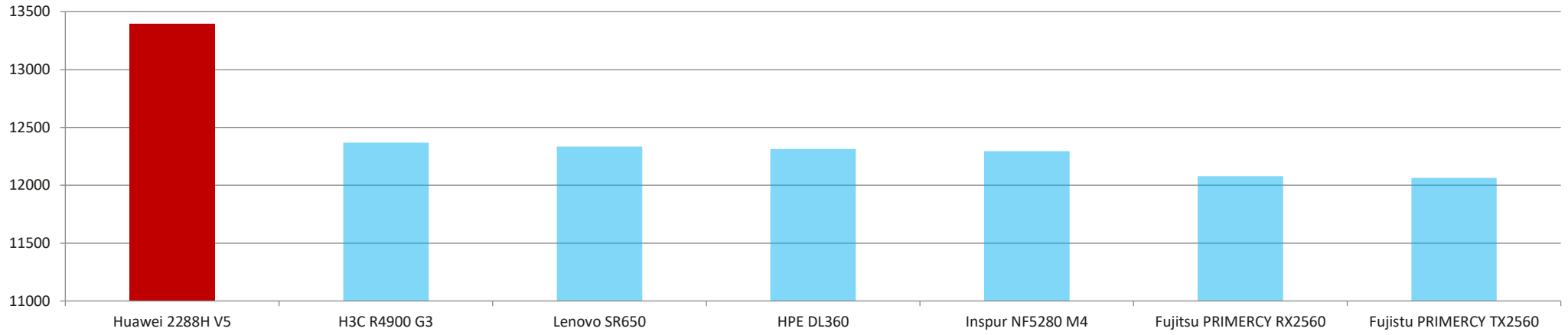
- **FDM:** The patented Fault Diagnosis & Management (FDM) technology provides the fault locating accuracy up to 93% and predicts a fault 7 to 30 days before it occurs, reducing OPEX
- **DEMT:** The patented Dynamic Energy Management Technology (DEMT) reduces energy consumption by 15% without compromising services
- **BSST:** Boot Speedup Storage Technology (BSST), 2 M.2 SSDs as high-speed OS boot drives, industry-unique hot plug and hardware RAID

Innovative design

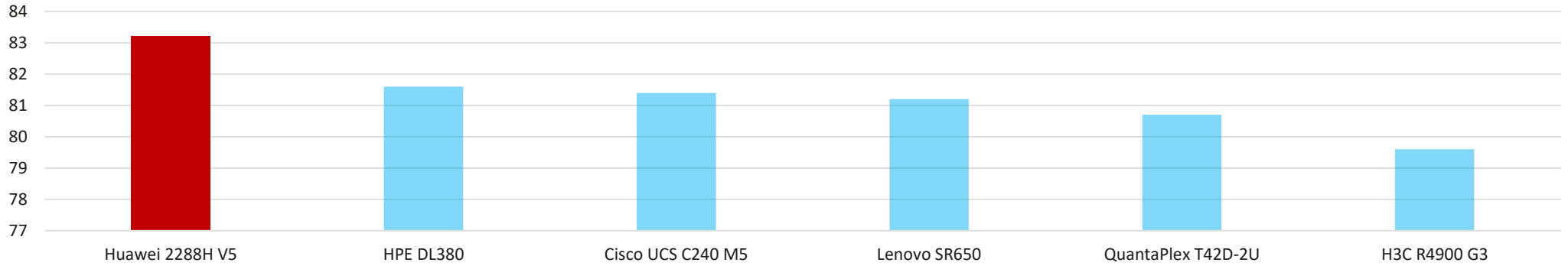
- 2 x 10GE and 2 x GE LOM ports, reducing CAPEX
- Flexible NICs for higher networking capabilities
- 20 x 3.5" or 31 x 2.5" drives for industry-leading storage capability
- 28 NVMe SSDs for optimal service experience
- 7 Atlas 300 inference cards/NVIDIA Tesla T4 GPUs for AI inference
- 2 NVIDIA Tesla V100 cards for AI training

2288H V5: New SPEC Energy Efficiency & Performance Records

SPECpower-ssj2008 Test Result



SPECint®_base2006 Test Result



Data released on http://www.spec.org/ as of Apr 2018

FUTUREWEI Technologies

5288 V5: Ultra-Large Storage Capacity

- Tiered storage of hot, warm, and cold data
- Historical data archiving
- Video surveillance, CDN, and big data
- E-mail OA application

High reliability

- Carrier-class component specifications and 100% derating design
- Redundancy design: PSU modules in 1+1 mode, fan modules in N+1 mode, and memory module sparing/mirroring
- 99.99% availability
- 15% fewer failures than the industry average

Ultra-large capacity design

- Hybrid storage architecture for tiered data storage
- 44 x 3.5'' SAS/SATA drives + 4 x 2.5'' SAS/SATA/NVMe SSDs in a 4U space, providing ultra-large local storage capacity
- Industry-leading storage capacity, up to 765 TB on a single server

Unique value

- **FDM:** The patented Fault Diagnosis & Management (FDM) technology provides the fault locating accuracy up to 93% and predicts a fault 7 to 30 days before it occurs, reducing OPEX
- **DEMT:** The patented Dynamic Energy Management Technology (DEMT) reduces energy consumption by 15% without compromising services
- **BSST:** Boot Speedup Storage Technology (BSST), 2 M.2 SSDs as high-speed OS boot drives, industry-unique hot plug and hardware RAID

Innovative design

- 2 x 10GE LOM ports, reducing CAPEX
- Flexible NICs for higher networking capabilities
- 6 Atlas 300 inference cards/NVIDIA Tesla T4 GPUs for AI inference, an industry-leading feature
- Dual RAID controller cards, doubling IOPS

2488 V5H: High Performance Computing

- Virtualization
- Cloud computing
- HPC
- Database
- In-memory database (SAP HANA)
- AI inference

High reliability

- Carrier-class component specifications and 100% derating design
- Redundancy design: PSU modules in 1+1 mode, fan modules in N+1 mode, and memory module sparing/mirroring
- 65 RAS features
- 99.99% availability
- 15% fewer failures than the industry average

Ultimate performance

- 4 x 2nd Gen Intel® Xeon® Scalable processors, up to 28 cores and TDP 205 W per CPU
- 48 DDR4 DIMMs, up to 2933 MT/s; a maximum of 24 Optane™ Persistent Memory (PMem), up to 2666 MT/s; total memory capacity up to 15 TB on a single server
- No.1 in SPECrate2017_int_base/SPECvirt_sc 2013 benchmark/SAP B4H benchmark tests
- Full mesh CPU architecture design, improving performance by 14.9%

Unique value

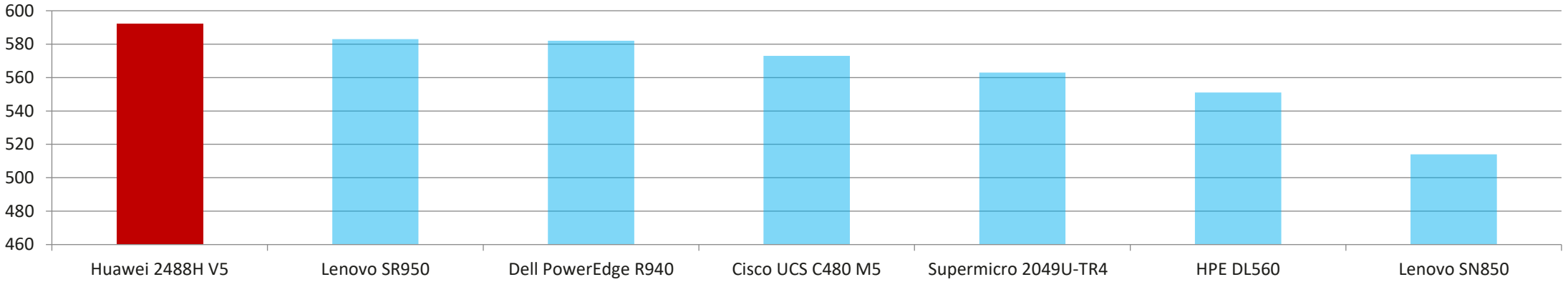
- **FDM:** The patented Fault Diagnosis & Management (FDM) technology provides the fault locating accuracy up to 93% and predicts a fault 7 to 30 days before it occurs, reducing OPEX
- **DEMT:** The patented Dynamic Energy Management Technology (DEMT) reduces energy consumption by 15% without compromising services
- **BSST:** Boot Speedup Storage Technology (BSST), 2 M.2 SSDs as high-speed OS boot drives, industry-unique hot plug and hardware RAID

Innovative design

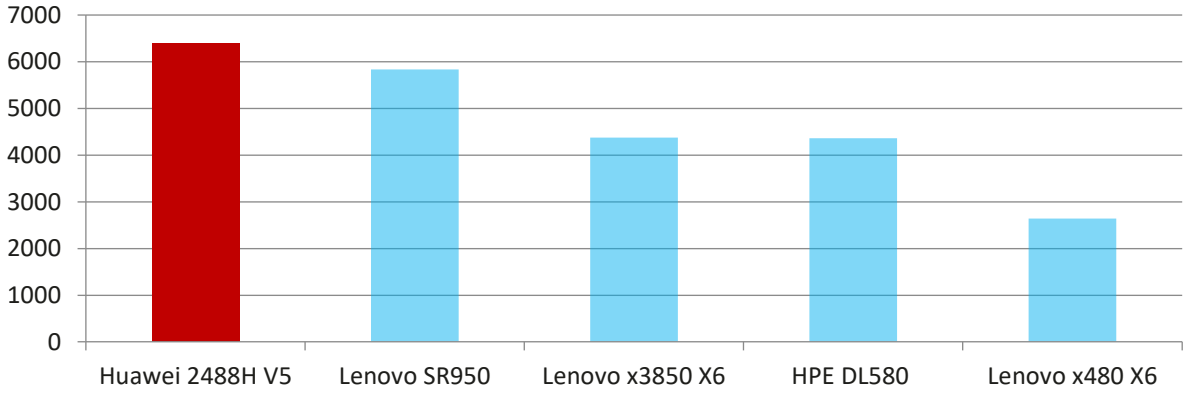
- 2 x 10GE and 2 x GE LOM ports, reducing CAPEX
- Flexible NICs for higher networking capabilities
- 25 x 2.5'' drives, providing the industry's largest storage capacity
- 3 NVIDIA Tesla T4 GPUs for image and video inference and analytics
- Up to 11 PCIe 3.0 slots, the most in the industry for diverse applications

2488H V5: New SPEC/SAP HAHA Performance Records

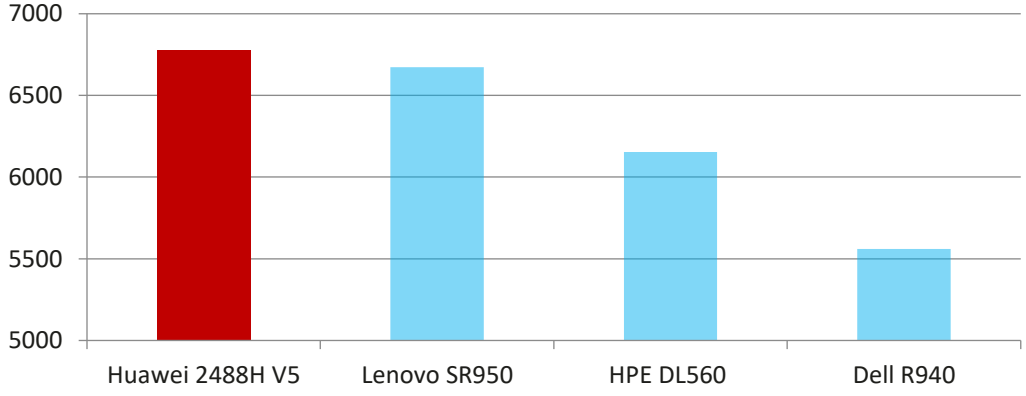
SPECrate2017_int_base Test Result



SPECvirt_sc2013 Test Result



SAP B4H Benchmark Test Result



Data released on http://www.spec.org/ as of Apr 2018

FUTUREWEI Technologies

2488H V6: Stable and Reliable Mission Critical Server

- Virtualization
- Cloud computing
- High-performance computing (HPC)
- Database
- In-memory database (SAP HANA)
- AI inference and training

High reliability

- Carrier-class component specifications and 100% derating design
- Redundancy design: PSUs in 1+1 mode, fan modules in N+1 mode, and memory module sparing/mirroring
- 69 RAS features
- 99.999% availability
- 15% fewer failures than the industry average

Ultimate performance

- 4 x 3rd Gen Intel® Xeon® Scalable processors, up to 28 cores and TDP 250 W per CPU
- 48 DDR4 DIMMs, up to 3,200 MT/s; 24 Intel® Optane™ persistent memory (Optane™ PMem) modules (200 series), up to 2,666 MT/s; total memory capacity of a single server up to 18 TB
- 2 x 300 W high-performance dual-width GPU accelerator cards, providing powerful computing for AI inference and training

Unique value

- **FDM:** The patented Fault Diagnosis & Management (FDM) technology provides the fault locating accuracy up to 93% and predicts a fault 7 to 30 days before it occurs, reducing the OPEX
- **DEMT:** The patented Dynamic Energy Management Technology (DEMT) reduces energy consumption by 15% without compromising services
- **BSST:** The Boot Speedup Storage Technology (BSST) uses 2 M.2 SSDs as high-speed OS boot drives, supporting the hardware RAID and the industry-unique hot plug for M.2 SSDs

Innovative design

- 2 FHFL dual-width GPU AI accelerator cards, such as Tesla V100
- 4 HHHL single-width GPU AI accelerator cards, such as Tesla T4
- 11 PCIe 3.0 slots, including one dedicated PCIe slot for the OCP 3.0 NIC (hot-swappable)
- 25 x 2.5'' SAS/SATA drives, up to 24 NVMe SSDs[1]
- Access to iBMC using Type-C, enabling agile O&M

[1] The configuration of 24 NVMe SSDs is estimated to be released in Q4 2020.

5885H V5: High Performance and Scalability

- Virtualization
- Cloud computing
- HPC
- Database
- AI inference or training

High reliability

- Carrier-class component specifications and 100% derating design
- Redundancy design: PSU modules in 1+1 mode, fan modules in N+1 mode, and memory module sparing/mirroring
- Maintenance of front fan modules without opening the chassis cover
- 65 RAS features
- 99.99% availability
- 15% fewer failures than the industry average

Ultimate performance

- 4 x 2nd Gen Intel® Xeon® Scalable processors, up to 28 cores and TDP 205 W per CPU
- 48 DDR4 DIMMs, up to 2933 MT/s; a maximum of 24 Optane™ Persistent Memory (PMem), up to 2666 MT/s
- Full mesh CPU architecture design, improving performance by 14.9%

Unique value

- **FDM:** The patented Fault Diagnosis & Management (FDM) technology provides the fault locating accuracy up to 93% and predicts a fault 7 to 30 days before it occurs, reducing OPEX
- **DEMT:** The patented Dynamic Energy Management Technology (DEMT) reduces energy consumption by 15% without compromising services
- **BSST:** Boot Speedup Storage Technology (BSST), 2 M.2 SSDs as high-speed OS boot drives, industry-unique hot plug and hardware RAID

Innovative design

- 2 x 10GE and 2 x GE LOM ports, reducing CAPEX
- Flexible NICs for higher networking capabilities
- 4 NVIDIA Tesla T4 GPUs for image and video inference and analytics
- 2 NVIDIA Tesla V100 cards for AI training
- Up to 15 PCIe 3.0 slots, the most in the industry for diverse applications

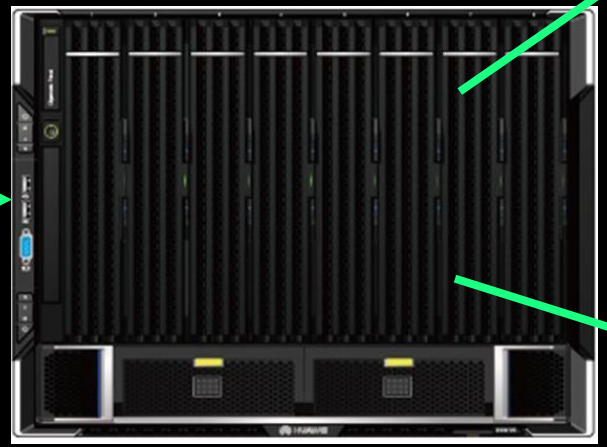
9008 V5 for SAP HANA: Reliable & Scalable

- Xeon Scale-up w/ Direct UPI Interconnect
- Up to 8X Xeon Platinum CPUs
- Up to 96 DDR4 DIMMs, Max 12TB DDR4 for V5
- Rich I/O, Storage & Expansion
- Modular Design w/ One CPU + Memory / Blade
- Scalable Power & Cooling for Xeon in 2022+
- Physical & Logical Partitions Support
- Synchronous to Intel Xeon Release

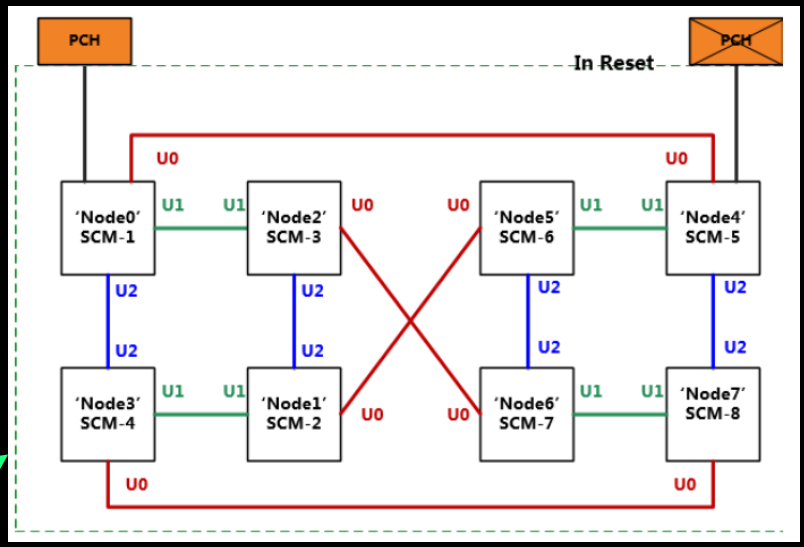
Central Management Enclosure (CME) (2U)

System Computing Enclosure (SCE) (8U)

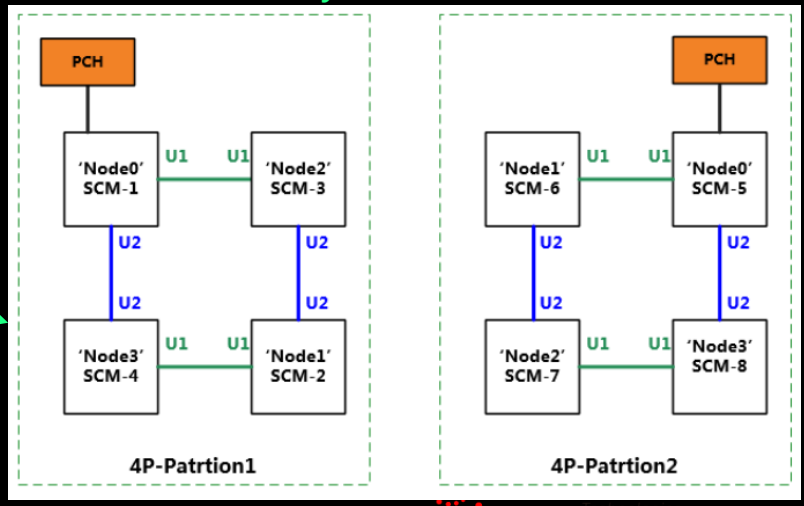
Resource Expansion Enclosure (REE) (4U)



One-System Mode



Dual-System Mode



Huawei Servers Enter STC's Procurement Shortlist (2018)

Saudi Telecom is a telecommunications company based in Saudi Arabia that provides landline, mobile and Internet services.

The Saudi Telecom Group provides integrated mobile, landline and broadband communications services to more than 1.6 billion customers worldwide. The company is headquartered in Riyadh, Saudi Arabia. Saudi Telecommunications Corporation (STC) is the largest telecommunications company in the Arab region in terms of market capitalization, total revenue and number of employees. In the past few years, STC has transcended its borders into international markets, forming commercial investment networks in many Gulf countries, Asia and Africa.

Challenges

- Improve the layered decoupling ratio of network and IT services, and need to rapidly increase the scale of collection and increase the supply capacity.
- Enterprises face digital transformation, need to find end-to-end strong partners, break the monopoly of existing manufacturers。
- The overall collection of the server is 40M, and this time we need to bid for the 10M blade server framework.

Solution

- E9000 2-way and 4-way blade servers for network and IT resource pool deployment。
- E9000 supports full-width and half-width blade hybrid insertion, supporting 40GE/100Gb/s IB EDR。
- E9000 chassis supports multi-generation evolution, and equipment upgrades do not need to change frames to protect customer investment

Customer Benefits

- E9000 is deployed in network and IT computing resource pools for integrated deployment to improve efficiency.
- Powerful computing performance and network capabilities provide customers with critical business carrying capabilities and lower customer TCO。
- Industry-unique chassis evolution technology to maximize customer protection

Huawei Servers Bolster Digital Transformation for the Iraq Branch of Gazprom

Gazprom is the world's largest natural gas company. Gazprom provides almost all of the natural gas needed in Central and Eastern Europe and the former CIS countries. Gazprom ranked No. 49 on Fortune Global 500 2018.

Challenges

- Oil and gas field mining is increasingly dependent on digitalization to improve service processing capabilities. Service rollout demands fast delivery and rapid deployment.
- Servers, as a major type of device to facilitate digital transformation, are critical to business continuity.
- The continuous growth of services poses higher requirements on server performance and energy saving.

Solution

- The 2-socket 2288H V5 rack server provides dozens of RAS features and hot swap for key components, ensuring stable and reliable services.
- 2288H V5 supports long-term stable running at 45°C with processors of full power consumption. The servers have undergone strict reliability tests, and use high-reliability components.
- Based on load pressure and resource utilization, 2288H V5 dynamically adjusts the working status of components, such as processors, DIMMs, and PSUs, to achieve the highest system power efficiency.

Customer Benefits

- Thanks to various RAS features and strict manufacturing process, the failure rate of 2288H V5 is 15% lower than the industry average. 2288H V5 were deployed in batches without alarms.
- Leveraging multiple technologies, such as Titanium PSUs with 96% energy efficiency, intra-board high-voltage DC PSUs, and DEMT, 2288H V5 delivers superior heat dissipation efficiency. These technologies save up to 15% energy without compromising server performance and help reduce power costs.
- With massive delivery and fast service rollout, Gazprom can deploy new services quickly, getting a head start in new business domains.

3 Key Oil & Gas Activities – Exploration, Production & Prediction

Exploration - Seismic Data Processing

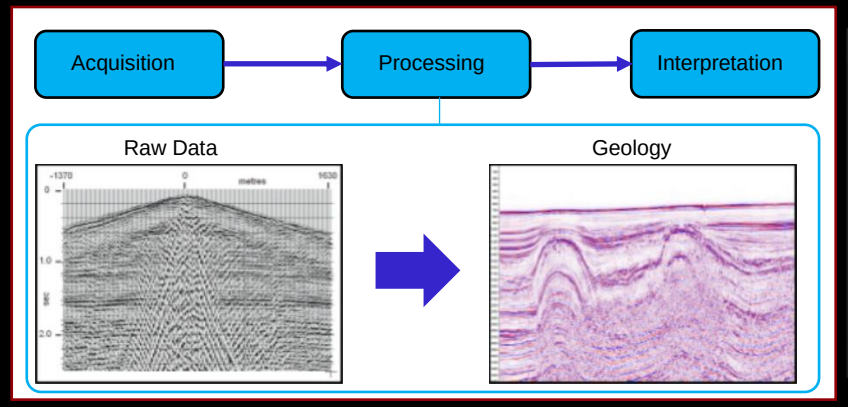
- Where and How Much is the Oil & Gas?
  - ✓ Build accurate HD earth Subsurface Models
  - ✓ Interpret the Models Automatically

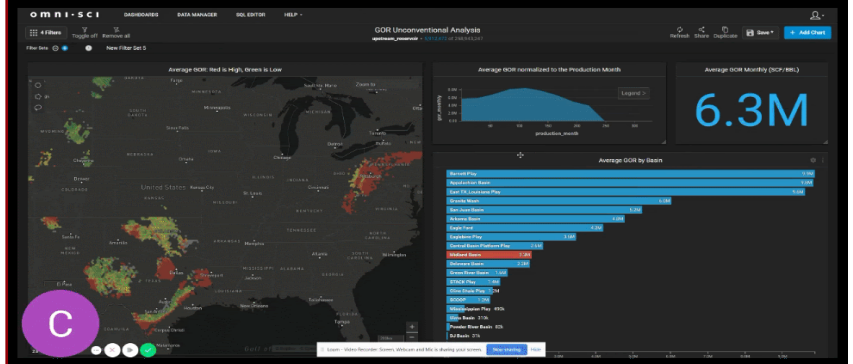
Production - Drill Operations

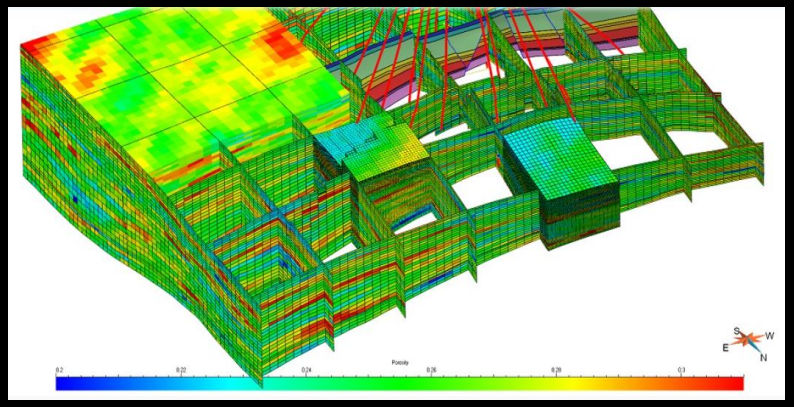
- How are the wells going in Real-Time?
  - ✓ Well Operation & equipment status Monitoring
  - ✓ Predictive Maintenance to avoid operation disruptions

Prediction - Reservoir Simulation

- How Much Oil & Gas Left & How would the reservoir be changing?







- 32-bit FP intensive operations
- FWI algorithm becoming popular but requires 10x compute w/ every frequency doubling
- DNN may for automated image Analysis

- Real-time Big-Data monitoring & Analysis
- In-Memory Computing w/ SAP HANA or HBM-based GPU will help

- Aramco TeraPOWERS is Powerful (X86 CPU based)
- DNN may for automated image analysis & Open-source DEEPSEISMIC is active

FUTUREWEI Technologies

CPU or Heterogeneous Computing for HPC, HPDA & AI?

HPC Fabric Evolution Trends: Ethernet-Based Growing

✓ Data Center Ethernet port speed in pace with or surpasses InfiniBand technologies

❖ IB w/ more In-Network Computing features (such as SHARP)

InfiniBand →

- NVIDIA
  - 25.6T & 400GE now w/ 50G PAM4 LR
  - CPO optical coming in 2022
- intel
  - Discontinued 200G OmniPath (IB-based) Development in 2019

Ethernet & Ethernet-Based ↗

- BROADCOM
  - Tomahawk-4: 25.6T & 400GE now w/ 50G PAM4 LR
  - CPO optical coming in 2022
- NVIDIA
  - Spectrum-3: 12.8T & 400GE now w/ 50G PAM4 LR
  - CPO optical may come in 2022
- Hewlett Packard Enterprise — Slingshot (Ethernet-Based)
  - 12.8T & 200GE now w/ 56G PAM4 LR
  - Used in 3 near-term Exascale systems
- HUAWEI
  - CloudEngine Family lossless 100/200/400GE Data Center Switches
- intel — Teralynx 8
  - 12.8T & 400GE now w/ 56G PAM4 LR
  - CPO optics demoed in 2020
- Innovium
  - 25.6T & 800GE sampled in 2H2020
  - 112G PAM4 LR SerDes
- Xsight — X1
  - 25.6T & 800GE sampled in Dec. 2020
  - 100G PAM4 LR SerDes

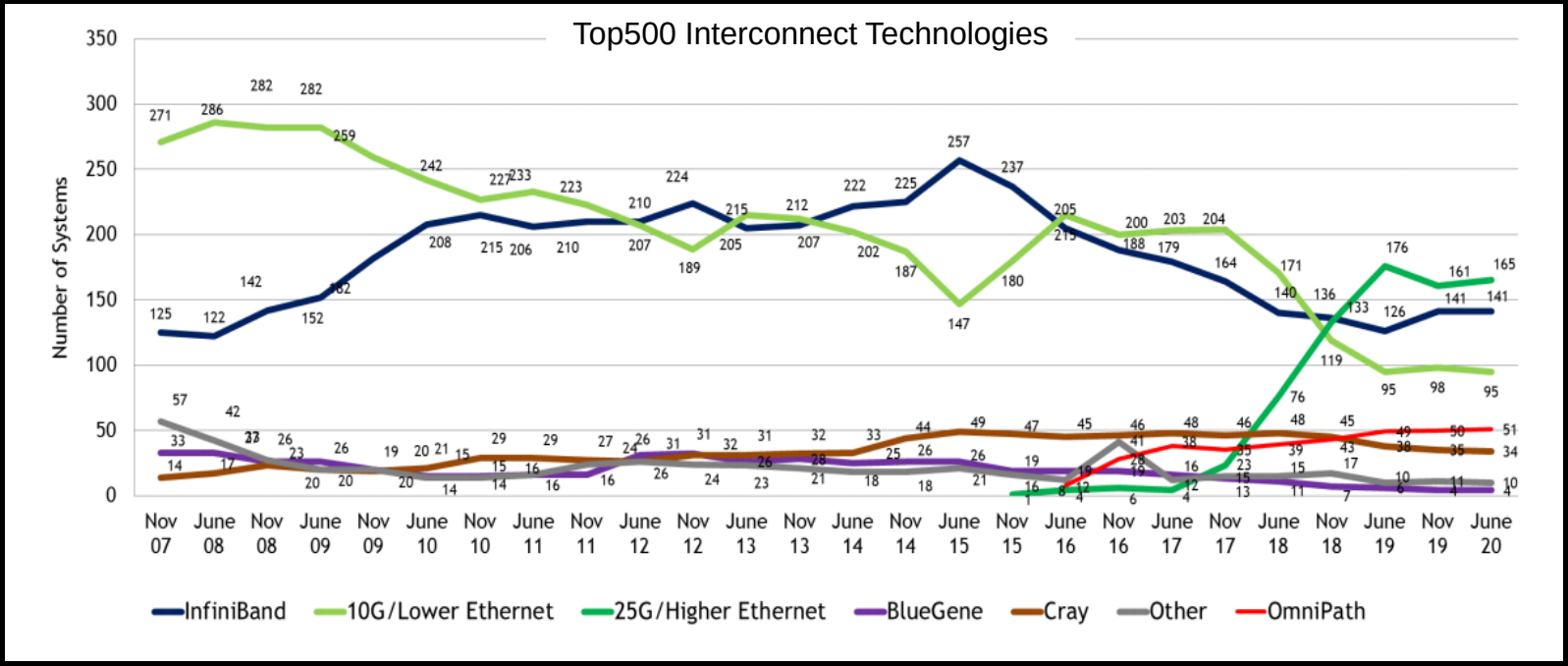
Other proprietary Interconnect →

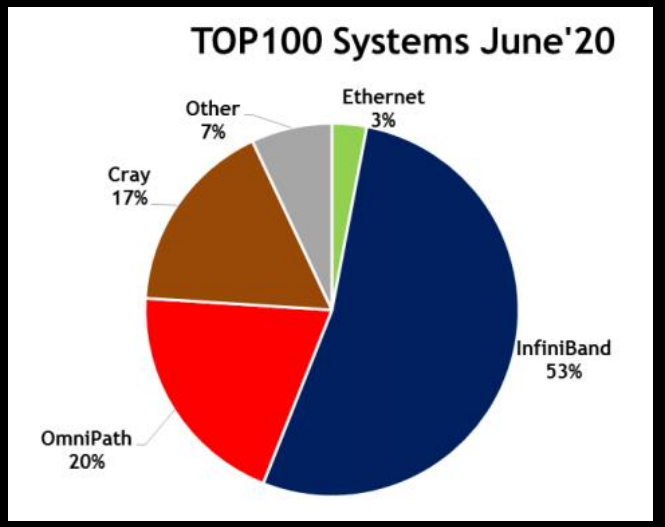
✓ Fujitsu Fugaku Tofu-D 6D Mesh

HPC Interconnect Technologies

Top500 HPC Interconnect Trends: Ethernet is No.1 & Growing

- ✓ 25GE & Higher Growing Rapidly, 10GE & Higher Surpass IB a lot (260 vs. 141 in June 2020)
- ✓ In August 2019, Intel discontinued 200G OmniPath development, but acquired Barefoot Networks to enter Ethernet Switch market
- ✓ IB is still dominating Top10 and Top100 Interconnect





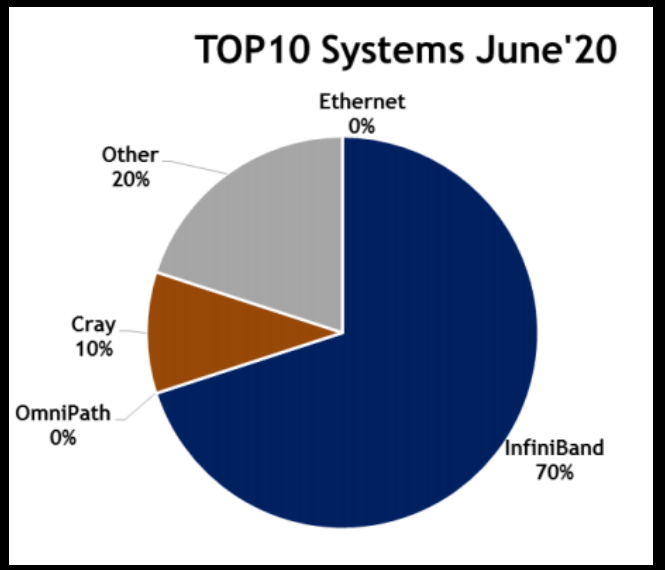


Figure Top500 Interconnect Technologies & Trends (image source & credit: Mellanox)

FUTUREWEI Technologies

Near-Term Exascale Systems Interconnect

✓ All the 3 Near-Term USA Exascale Systems adopt Ethernet-based HPE/Cray Slingshot Interconnect, not IB

Frontier at ORNL
- On in 2021-2022
- AMD EPYC CPU + AMD Radeon GPU (future-gen)
- ≥ 1.5 Exaflops DP
- System Interconnect: HPE/Cray Slingshot, an enhanced Data Center Ethernet, 25.6Tbps per Switch chip, 200G/Port

Aurora at ANL
- On in 2022-2023
- Intel Sapphire Rapids CPU + Intel Xe GPU
- ≥ 1 Exaflops DP
- System Interconnect: HPE/Cray Slingshot, an enhanced Data Center Ethernet, 25.6Tbps per Switch chip, 200G/Port

EI Capitan at LLNL
- On in 2022-2023
- AMD EPYC CPU + AMD Radeon GPU
- ≥ 1.5 Exaflops DP
- System Interconnect: HPE/Cray Slingshot, an enhanced Data Center Ethernet, 25.6Tbps per Switch chip, 200G/Port

Fugaku in Japan
- On in 200-2021
- Fujitsu A64FX ARMv8 CPU only
- > 400PF DP, >800PF SP, 1.6EP HP
- System Interconnect: Fujitsu Tofu-D 6D Mesh, 28Gx2x10 Per Node

Figure Near-term Exascale System Interconnects (image source & credit: www.hpcwire.com)

FUTUREWEI Technologies

Ethernet Switch ASIC Evolving Rapidly, CPO coming

- ✓ **Broadcom announced industry first Co-Packaged Optics Ethernet Switch ASIC to be available in 2022**
- ✓ **400GE Ethernet Switch is shipping, and 800GE is coming soon, in parallel with IB Speed Evolution**

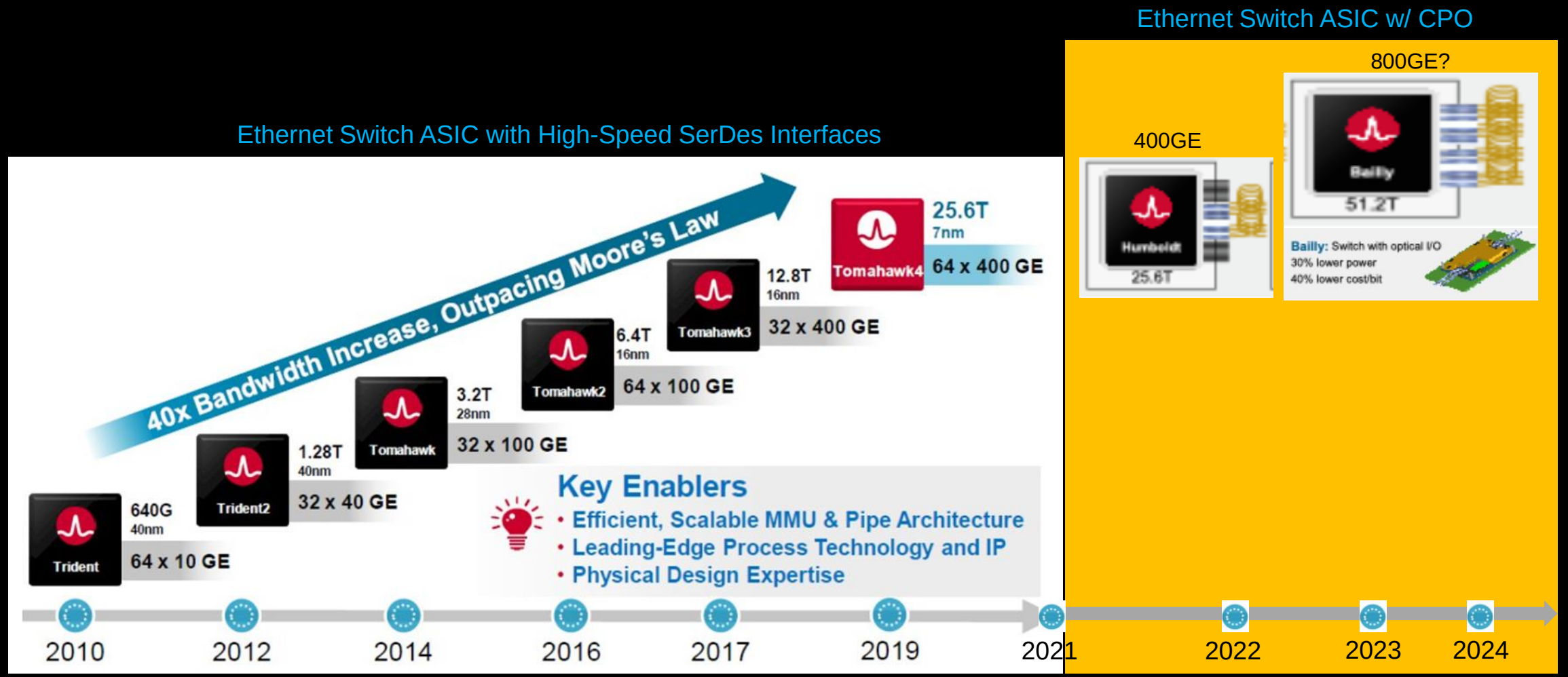


Figure Broadcom Data Center Ethernet Switch Evolution (Basic image source & credit: Broadcom)

IB: Speed Evolution Parallel w/ Ethernet

- ✓ Speed evolution parallel with Ethernet, used to be ahead of Ethernet, but IB w/ more In-Network Computing features (SHARP)
- ✓ Intel discontinued 200G OmniPath (IB) development in 2019, leaving Nvidia/Mellanox the ONLY IB vendor
- ✓ For the same data rates, IB is still delivering better performance for some applications

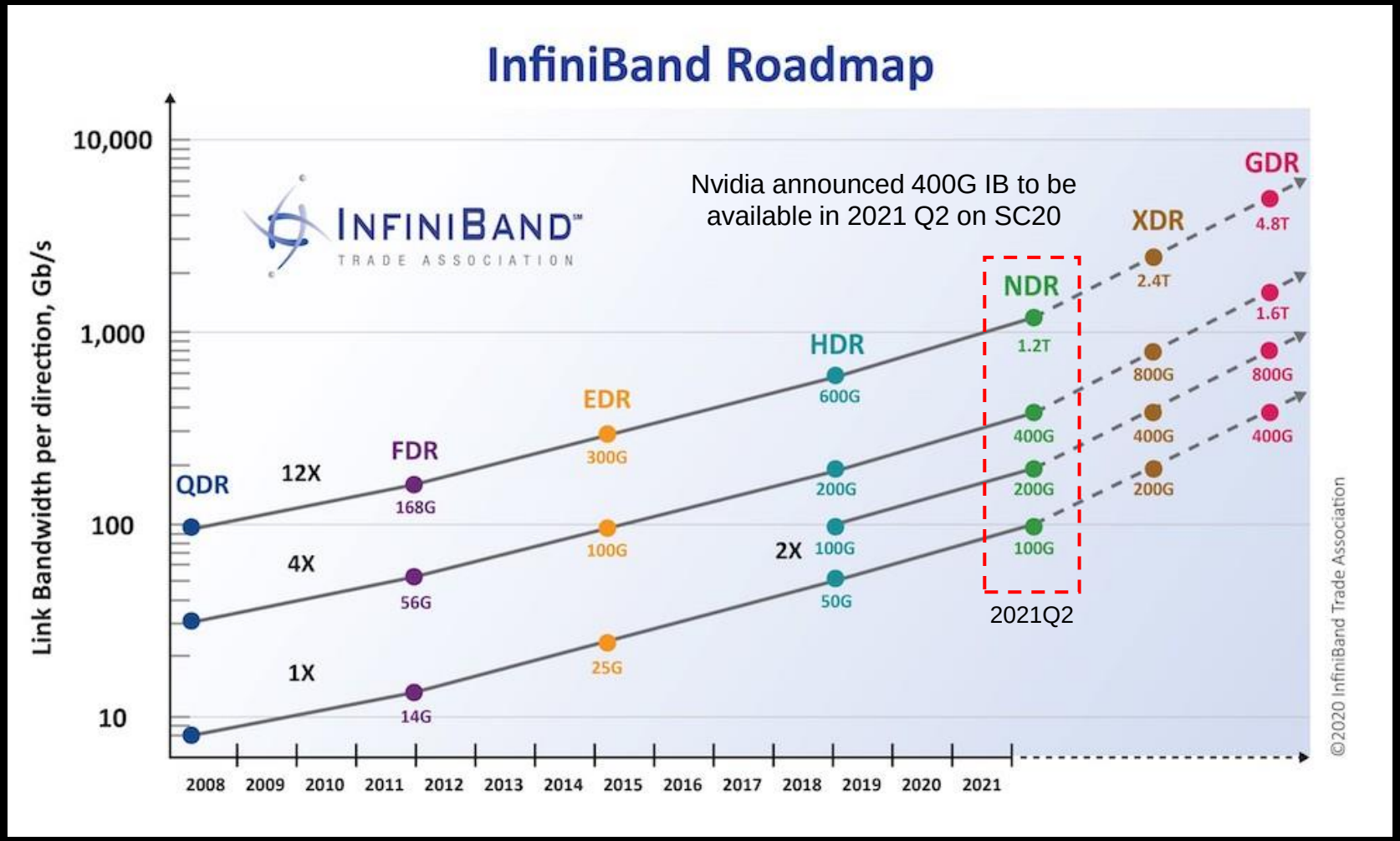


Figure InfiniBand Evolution (Basic image source & credit: IBTA)

FUTUREWEI Technologies

Why Ethernet-Based HPC Interconnect Becoming Popular?

- ✓ Latency caused by bad Congestion Control is much higher than Switch's P2P Latency where IB is a little better; Verified by Industry and our internal HPC & AI application benchmarkings
- ✓ Why Ethernet for HPC Popular: Comparable Performance, Unified Fabric for Lower TCO, No Need for IB talents & Costly High-Performance IB-Ethernet GW for storage, Native Network for NVMe-over-Fabric

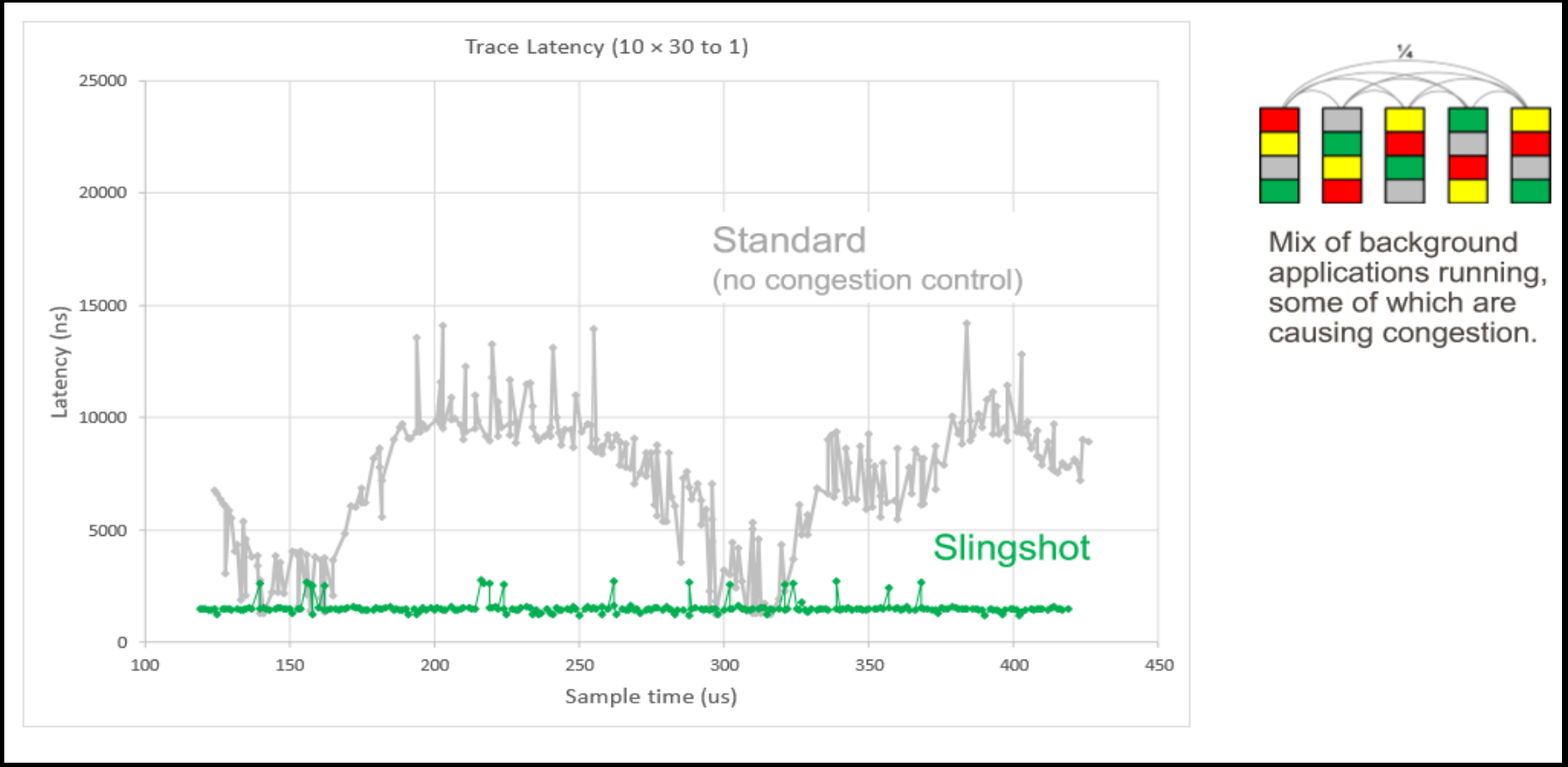


Figure Congestion Control is Key to End-to-End Latency (image source & credit: Cray via www.Nextplatform.com)

FUTUREWEI Technologies

Energy HPC POC w/ Huawei Lossless Data Center Switch

- ✓ We are doing a HPC POC with our lossless & enhanced Congestion-Control 100GE Switch with one of our energy customers in Middle East; and the initial result showing minor performance difference with the same speed IB
- ✓ Similar Result for Our performance benchmarking with GROMACS, LAMMPS, VASP, Quantum Espresso

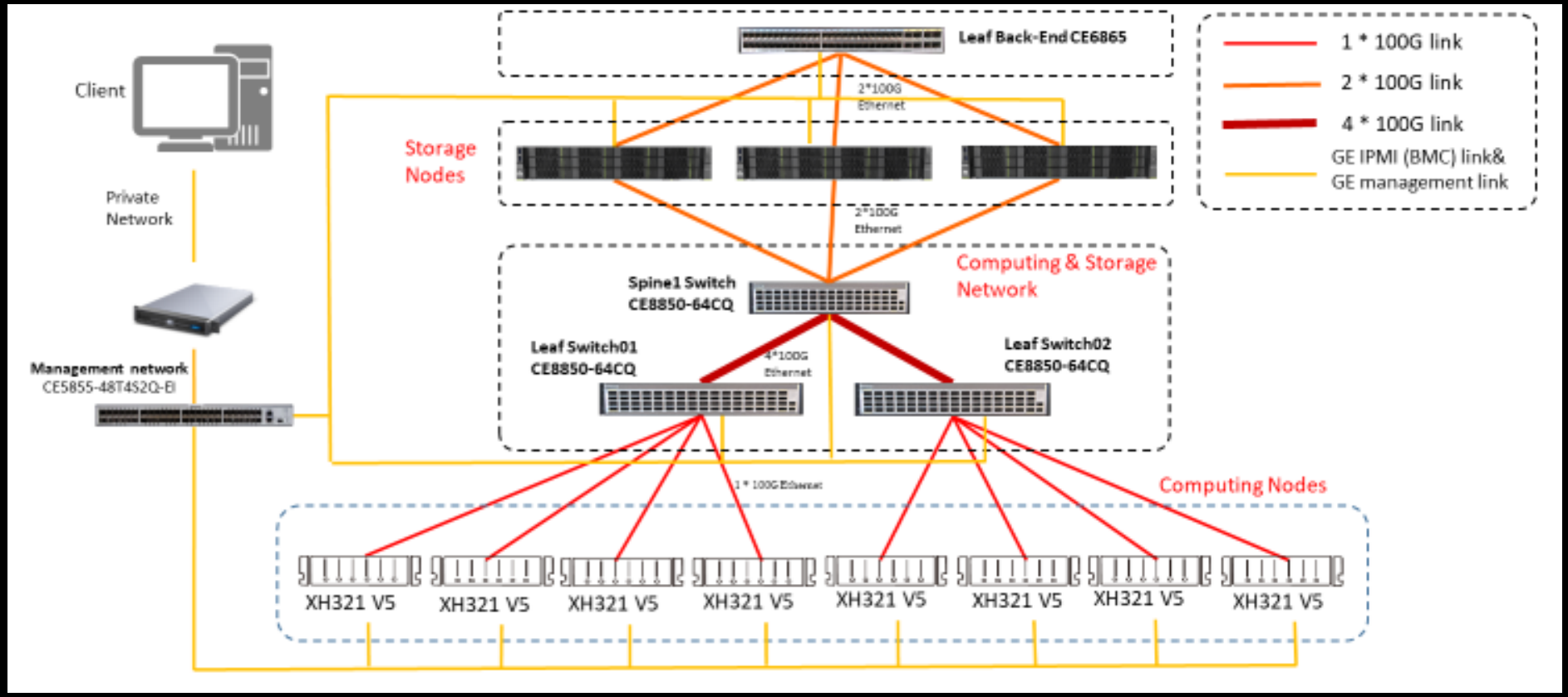


Figure Energy HPC Application POC with Huawei Lossless Data Center Switches (image source & credit: Huawei)

FUTUREWEI Technologies

Take-Aways

- ✓ CPU, especially Intel X86, will continue play a key role in HPC, Big Data Analytics and AI Applications
- ✓ CPU, especially Intel X86, are improving HPC, HPDA & AI Performance with Denser Vector (AVX512) & Matrix units
- ✓ New CPUs with HBM as System Memory and enhanced Vector and Matrix acceleration units will come soon, providing competitive performance to GPU for HPC, HPDA & AI with simpler programming model
- ✓ Data Center Ethernet is evolving in speed parallel to or surpassing InfiniBand, and is taking more market share in HPC, HPDA and AI applications
- ✓ In working with partners like Intel, VMware, SAP, Huawei has been offering a rich and innovative X86 Product Portfolio, delivering industry leading performance, reliability and manageability with elegant engineering, complete end-to-end process, strict sourcing and rigorous testing; as well as lossless & AI-based ECN flow-control CloudEngine Data Center 100/200GE Switch products

Thank you