Lenovo
HPC
PORTFOLIO

Lenovo™
ThinkSystem

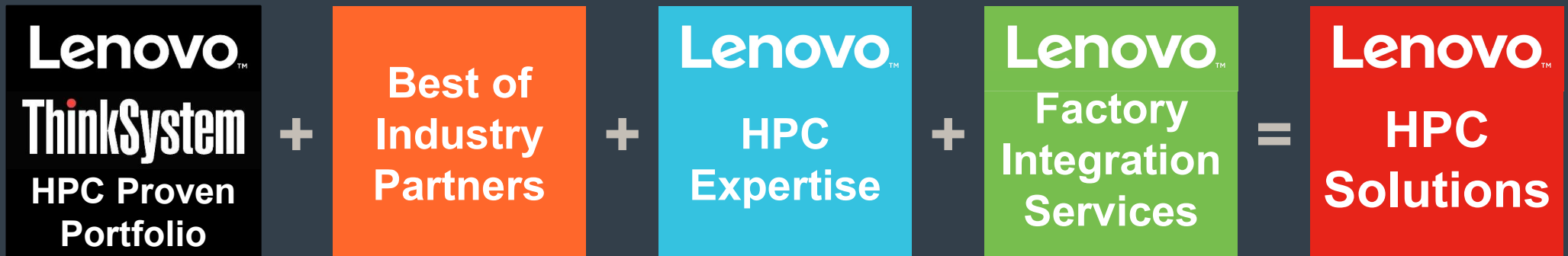SERVER

NETWORKING

STORAGE

Lenovo™

Scalable Infrastructure

Lenovo™ ThinkSystem HPC Proven Portfolio + **Best of Industry Partners** + **Lenovo™ HPC Expertise** + **Lenovo™ Factory Integration Services** = **Lenovo™ HPC Solutions**

Lenovo's HPC Server

The Lenovo ThinkSystem Server Portfolio

Datacenter

Business Intelligence

Analytics

Cloud

Virtualization

Retail

Distributed Enterprise

Small/Med Business

Tower

2P

ST550

Racks

for HPC

2P

2U

1U

SR650

SR630

SR590

SR570

SR550

SR530

Dense

for HPC

2P

SD650

SD530

Blades

4P

SN850

2P

SN550

Mission Critical

8P

SR950

4P

SR860
(EMEA, PRC)

4P

SR850

for HPC
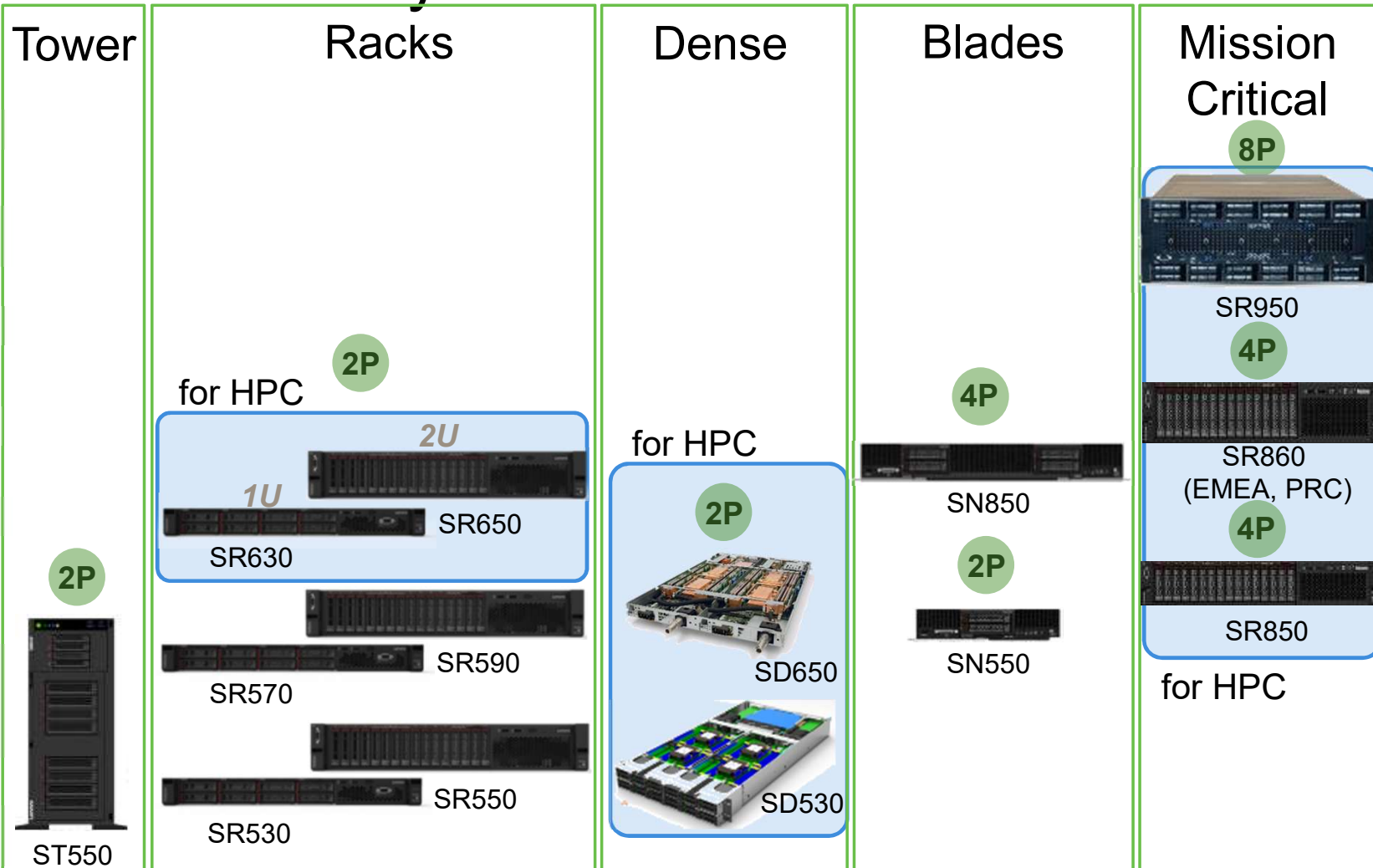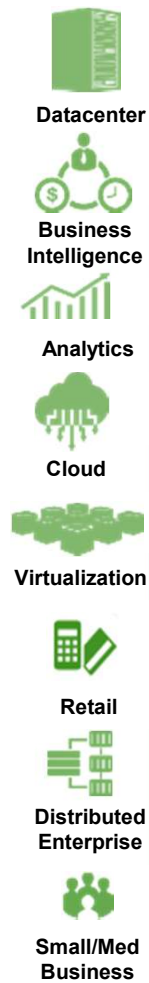
Intel Xeon Server processor history

Release date	Code	Processor	core/chip	TDP(W)	Spec FP	Spec_fp Rate
2006/6/26	Woodcrest	Intel Xeon 5160	2	80	17.7	45.5
2007/11/12	Harpertown	Intel Xeon x5460	4	120	25.4	79.6
2009/3/30	Nehalem	Intel Xeon x5570	4	95	43.8	202
2010/3/16	Westmere-EP	Intel Xeon x5690	6	130	63.7	273
2012/5/1	SandyBridge	Intel Xeon E5-2690	8	135	94.8	507
2014/1/9	IvyBridge	Intel Xeon E5-2697v2	12	130	104	696
2014/9/9	Haswell	Intel Xeon E5-2699v3	18	145	116	949
2015/3/9	Bradwell	Intel Xeon E5-2699v4	22	145	128	1160
2017/7/11	Skylake	Intel Xeon Platinum 8180	28	205	155	1820

Intel processor TDP & Spec_fp Rate
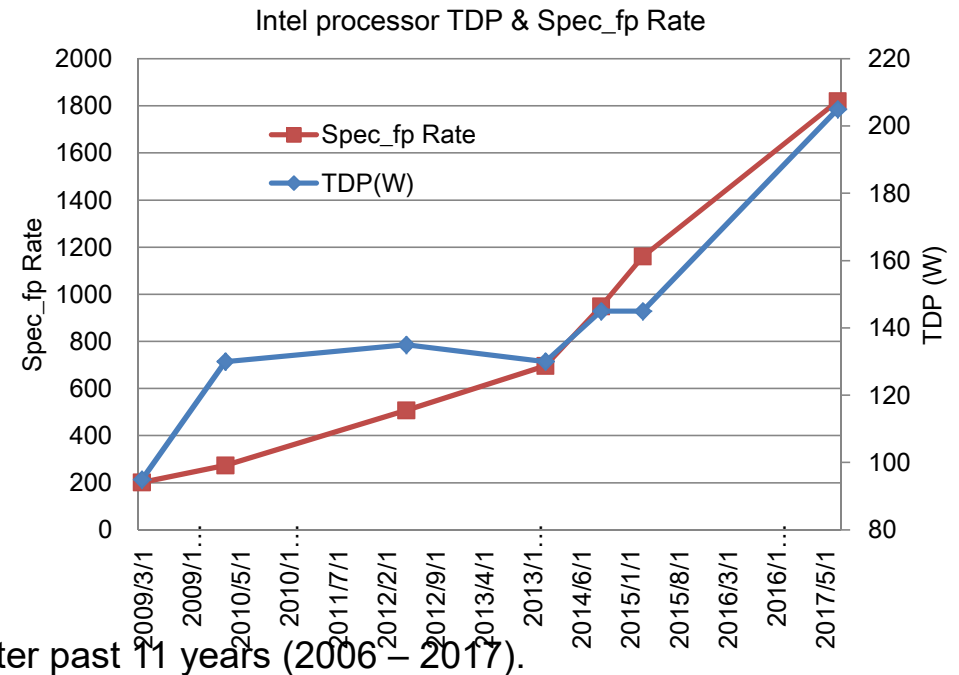
Processor performance dependency

- Spec_fp Rate with 2 processor chpis/node is grouing up 40 times faster past 11 years (2006 – 2017).
- The number of cores on the chpi increase 14 times.
- TDP watts increads 2.56 times bigger.
- Current maximum TDP is 205W. Knighs Mill Xeon phi processer will be 305 W

Big issue:
For increasing performance will be **processor / server cooling technology.**.

Moving forward into a decade of dense HPC

Lenovo continues development on both Air and Water cooling system

Air Cooling System

Direct Water Cooling System

iDataPlex – Air / DWC
dx360 - Harpertown
dx360M2 - Nehalem
dx360M3 - Westmere
dx360M4 - Sandy Bridge

2011~

NeXtScale – Air / DWC
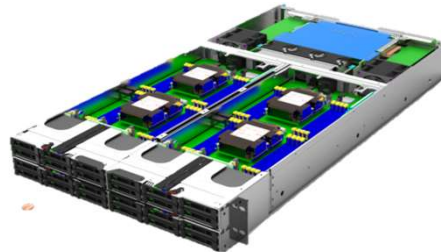nx360M4 - Sandy Bridge/Ivy Bridge
nx360M5 - Haswell/Broadwell

2013~

ThinkSystem – Air
SD530 - Skylake/Icelake

2017/18~

ThinkSystem – DWC
SD650 - Skylake/Icelake
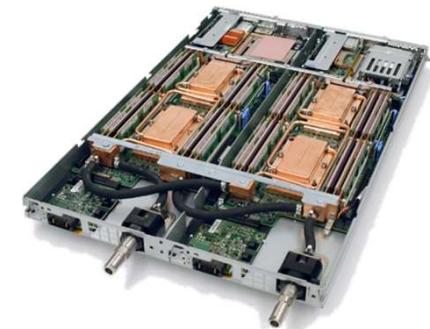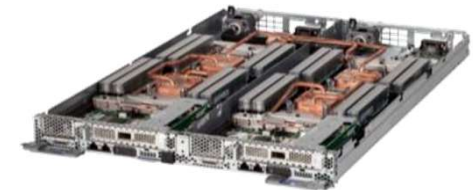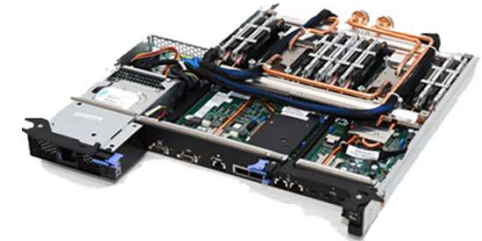
Lenovo cooling technology keeps 0.5U/server

System

Heat sink

2006 HS21 Brade server Air cooling
Woodcrest 2cores TDP=80W
Blade server E chassis 7U 14servers
(1server = 0.5U)

Air Cooling 0.5U /server

2017 Thinksystem SR650 2U 1server Air cooling
SkyLake Platinum 28cores TDP=205W
(1 server = 2U)

Air Cooling 2U /server

2018 Thinksystem SD650 1U 2servers DWC
SkyLake Platinum 28cores TDP=205W
(1server = 0.5U)

Direct Water Cooling 0.5U /server

⊕ WCT remove heat by water

NextScale Chassis w/ 12 Nodes
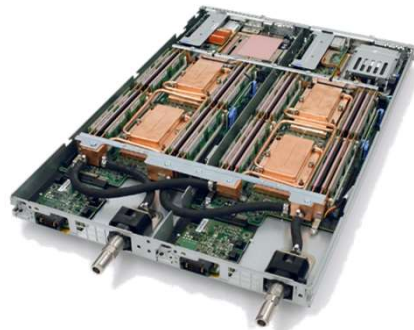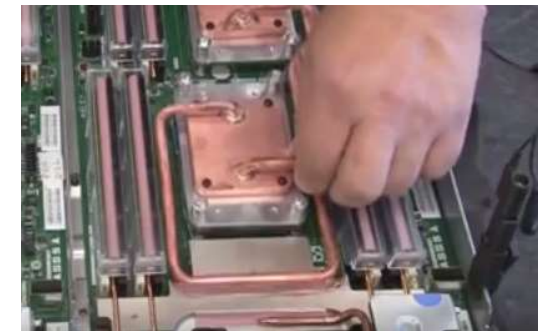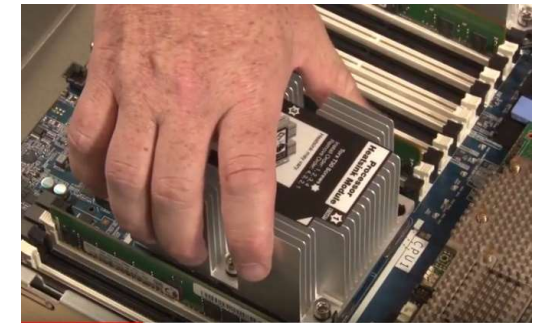
Chassis Power 6 kW-7kW capable

Maximum node power w/ standard memory

- ~ 450W w/165W CPUs
- ~540W w/ 205W CPUs

With 3D-Cross Point memory

- ~510W w/ 165W CPUs
- ~600W w/ 205W CPUs

Heat Recovery target with air cooled PSU is 90%+

Scalable water manifold to enable 1-6 chassis in a rack

Chassis water flow 6 – 9 liters/min,

Rack 36 – 54 liters/min for 6 Chassis (72 Nodes)

Example Node Power Map with 165W CPUs

Node Power		Estimates				
	Component Specifics (QSPEC, Model, rev, P/N, etc.)	Specification Power (W)	12V Power	Qty	Max. Power (W)	Weighted Power (W)
CPU	Intel Skylake CPU (165W)	165	175.5	2	351.1	351.1
Memory	64GB RDIMM (1DPC)	7	7.8	12	93.3	65.3
	256/512GB Apache Pass (2nd DIMM on channel on 12 channels)	18	20.0	0	0.0	0.0
Chipset	Lewisburg PCH (LWB-2) 2x1Gbe+ 2x10GbE capable	13	14.4	1	14.4	10.1
Planar	Emulex Pilot 4	4	4.4	1	4.4	3.1
	1GbE Mangement Port	2.2	2.4	1	2.4	1.2
	System component	Item Power (W)	12V Power	Qty	Max. Power (W)	Weighted Power (W)
I/O	PCIE slot(15W) - General Purpose Dual port infiniband EDR	15	15.8	1	15.8	15.8
Storage	M..2 Module HW RAID	2	2.0	0	2.2	1.6
TPM	Pluggable module	1	1.0	1	1.1	0.6
Total					493.4	455.0

**Direct to Water**

Conduction to Water loop

- Water cooled PSU evaluation in plan for POC effort with Lenovo STIC team to increase heat recovery to 95%+.
- Rack and manifold thermal insulation for external walls could be evaluated.

Lenovo™

✚ 7.Heat Extraction/Performance Expectations

Compute Tray Configuration (dual compute nodes)	Single Rack Steady State Power (w/Linpack, Turbo ON/OFF)	90% Heat Removal		85% Heat Removal	
		Heat-to-water	Heat-to-air*	Heat-to-water	Heat-to-air*
2 server / tray with TDP=165W CPU, 16x16GB DDR4 Memory, 2x NIC	33.5kW	29.3kW	4.2kW	27.6kW	5.9kW
2 server / tray with TDP=145W CPU, 16x 16GB DDR4 Memory, 2x NIC	30.7kW	26.7kW	4.0kW	25.3kW	5.4kW
2 server / tray with TDP=135W CPU, 16x 16GB DDR4 Memory, 2x NIC	29.4kW	25.6kW	3.8kW	24.1kW	5.3kW
2 server / tray with TDP=120W CPU, 16x 16GB DDR4 Memory, 2x NIC	27.3kW	23.7kW	3.6kW	22.4kW	4.9kW

Assumptions: 36x compute trays (dual node), 72 servers + 6x switches / Rack

- Performance is dependent on many facility variables such as ambient temp, water temp, flow rate, etc… Based on testing, it is reasonable to assume 85-90% heat-to-water given the availability of typical environmental input parameters.

Lenovo
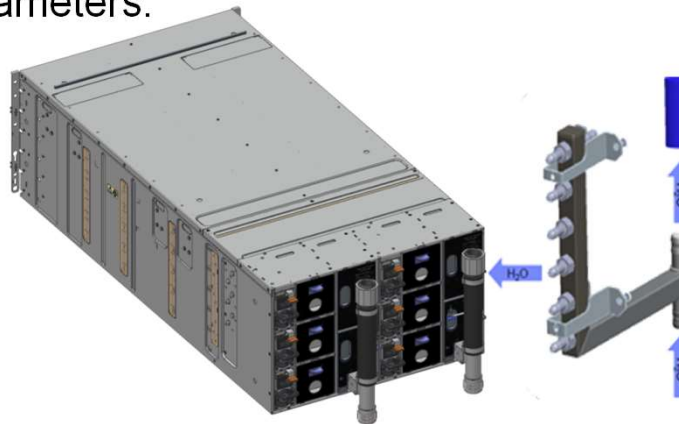
ThinkSystem SD530 (Air Cooling)

Feature	SD530
Processors	2 Intel "Purley" Generation processors per node • Up to 165W and 28 Cores
Form factor	1U Half wide tray node / 2U4N Chassis
Memory Slots Max Memory	• 12x DDR4 (R/LR) 2667MHz DIMM (up to 64GB)
Storage	• 6x SAS/SATA SF HS HDD/SSD or 2x NVMe, 2x M.2 SATA SSD
NIC	2x 10G SFP+ or GBase-T
PCIe	2x x8 or 1x x16 PCIe (HS) in Rear Shuttle 2x x16 internal for expansions
Power	HS/1+1 redundant 1100W/1600W/2000W Platinum
USB ports	1x USB 2.0
Cooling	• 5x hot swap redundant fans • Support ASHRAE A2/3/4 (config limitation A3/4)
System MGMT / TPM	Stark Mgmt Module, 1 shared GbE port TPM, Pluggable TCM, Opt. KVM break-out
Expansions	GPU Expansions for 2 GPUs per node
Dimensions	846mm depth, front access w/ rear I/O

The Node

x16 PCIe Out

2x10 GbE Out

12 DDR4 DIMMs
+ 4 Apache Pass

Hardware RAID (Optional)

4 or 6x 2.5 front hot swap drives

2x SATA or PCIe M.2 BootSSD

2S Purley CPU max
~ 165W*

DIMM 9
|
DIMM 12

3V Lithium
battery

PCIe slot 3
connector

PCIe slot 4
connector

Front Dongle
Vdieo / Serial /
USB 2.0 port
connector

DIMM 1
|
DIMM 4

PCIe slot 1
connector

Backplane power
connector

PCIe slot 2
connector

Backplane MISC
signal connector

Microprocessor 1

Microprocessor 2

SATA 1 connector

SATA 2 connector

DIMM 13
|
DIMM 16

DIMM 5
|
DIMM 8

Front Dongle
USB 3.0 port
USB 2.0 port
connector

Trusted
cryptographic
module (TCM)
connector

M.2 connector

*in 4x drive configuration only

Front View

Node front view: Six 2.5-inch hot-swap HDD model

Activity LED (green)   Status LED (yellow)

NVMe Bays

HDD 0   HDD 2   HDD 4

HDD 1   HDD 3   HDD 5
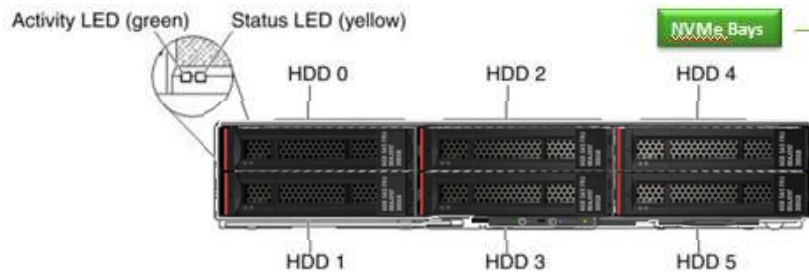
Node front view: Four 2.5-inch hot-swap HDD model

Node Storage bay options

- Six 2.5-inch hot-swap SAS/SATA HDD/SSD bays
- Four 2.5-inch hot-swap SAS/SATA HDD/SSD bays
- Four 2.5-inch hot-swap SAS/SATA HDD bays + two 2.5-inch hot-swap SAS/NVMe drive bays

HDD 0   HDD 2   KVM connector   USB 3.0 connector

KVM breakout module

HDD 1   HDD 3

Node 3   Node 4

NVME Bays
(2 per node)

NVME Bays
(2 per node)

NVME Bays
(2 per node)

NVME Bays
(2 per node)

Node 1   Node 2

The Shuttle – I/O

Hot-swap cooling fans

Edge connectors used to connect 10GbE on IO Riser to NIC Module

Two PCIe x8 slots per node
Or 1 PCIe x16 slot per node

8x10GbE Ethernet ports - either 10GBASE-T or SFP+

System Management Module connection

Eight low-profile PCIe 3.0 x8 cards

10GbE cage (SFP+)

10GBASE-T cage (RJ-45)

PCIe slot 4-B   Ethernet filler   PCIe slot 3-B

PCIe slot 4-A   PCIe slot 3-A

PCIe slot 2-A   PCIe slot 1-A

PCIe slot 2-B   PCIe slot 1-B

System Management Module   Power supply 1   Power supply 2
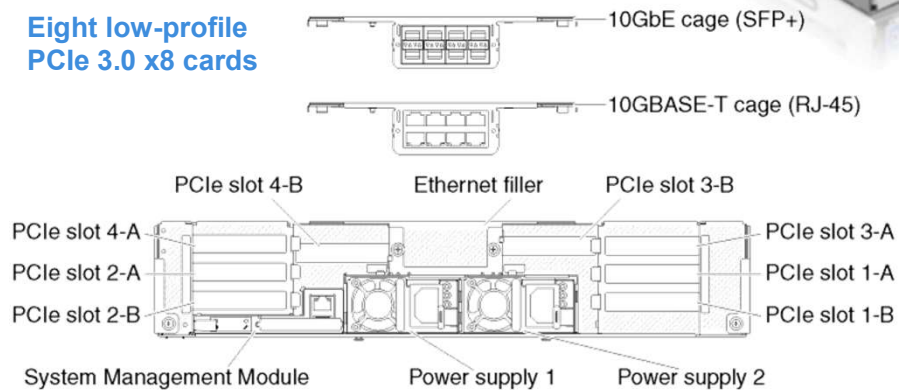
Four low-profile PCIe 3.0 x16 cards

10GbE cage (SFP+)

10GBASE-T cage (RJ-45)

Ethernet filler

PCIe slot 4   PCIe slot 3

PCIe slot 2   PCIe slot 1

System Management Module   Power supply 1   Power supply 2

14
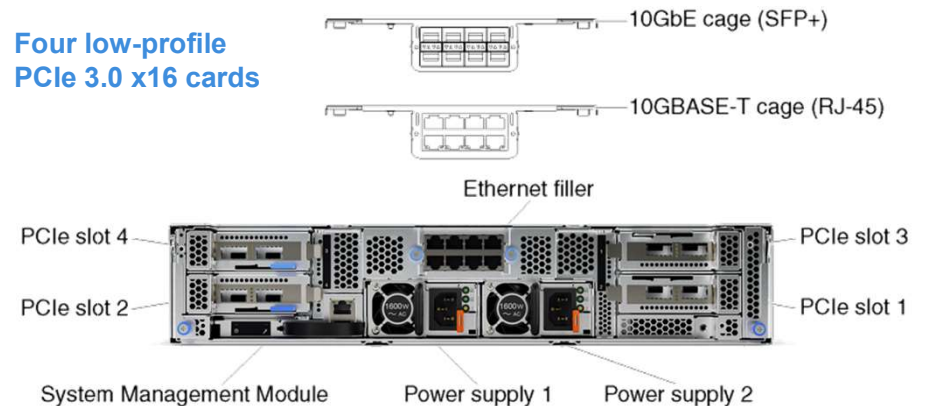
GPU Tray

Node limitation
- Only support 2x2 Backplane (4 HDD config)
- Must be identical GPU cards in the same GPU tray

Chassis limitation
- Must have 2 x GPU tray in a chassis for air flow balance
- Only support 2000W PSU with GPU tray

Supported Accelerators

NVIDIA Tesla P40 24GB (250 W)	NVIDIA Tesla P100 16GB (250 W)
NVIDIA Tesla M10 32GB (250 W)	Intel Xeon Phi 7240P 16GB, 1.3 GHz, 68 Core (275 W)
NVIDIA Tesla V100 16GB (300 W)	NVIDIA Tesla M60
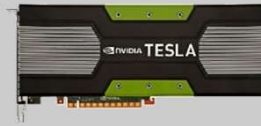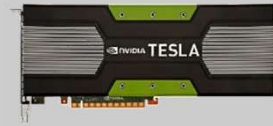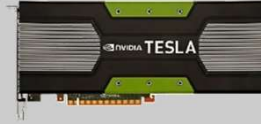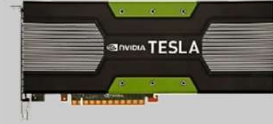
SD530 Embedded Management Engine

XClarity Controller delivers modernized deployment and inter-operability
for data centers who put a premium on efficiency and precision

All-new management engine
- Common in all ThinkSystem servers

Elegant User Experience
- Fresh, uncluttered GUI

Modernized inter-operability
- Built on Redfish-compliant APIs

Vastly Improved Performance
- 2x faster boot time to OS screen
- 6x faster firmware updates

Lenovo's HPC
Storage & Solution

Storage Requirements Devouring Resources

Every day, 15 petabytes of new information are being generated: 8x more than the information in all U.S. libraries

Internet of Things

Internet, Web 2.0

Distributed Computing

Data Center

Storage budgets up 1-5%

Zettabytes
1000 X
Exabytes
1000 X
Petabytes
1000 X
Terabytes
1000 X
Gigabytes

1980   1990   2000   2010   2020

18
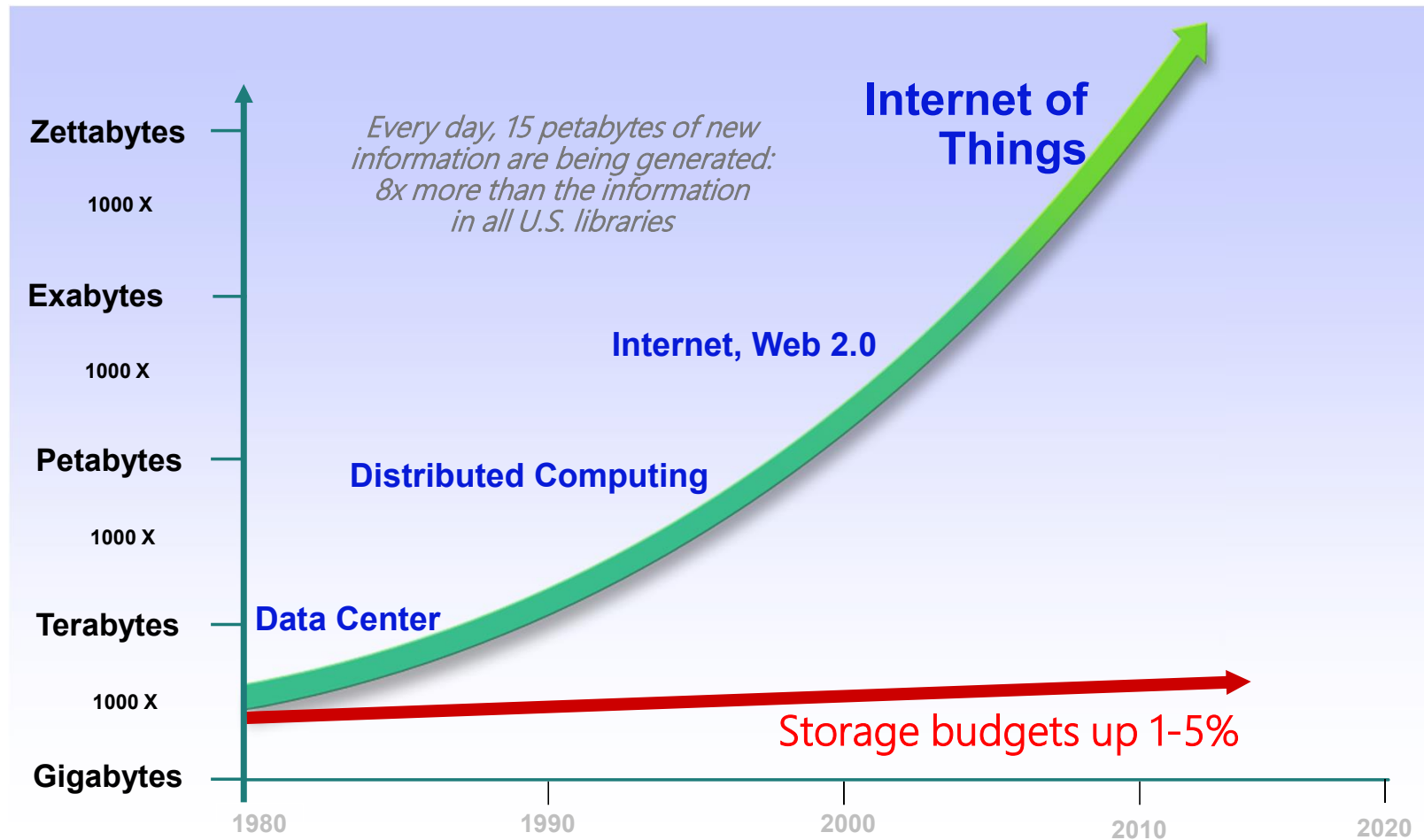
Storage is Essential and We Have the Tools

Taking our off-the-shelf server and storage portfolio

marrying it with leading HPC Storage Software
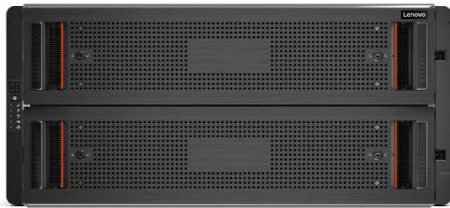
Distributed Storage Solution for IBM Spectrum Scale	Distributed Storage Solution for SUSE Enterprise Storage	Distributed Storage Architecture for SUSE Enterprise Storage / Red Hat Ceph Storage	Distributed Storage Architecture for Intel Lustre EE
DSS - G	DSS - C	DSA - C	DSA - L
Defined **Solution** especially for large capacity, high performance workloads in HPC environments	Defined **Solution** especially for interaction with Lenovo scale-out HANA solutions.*	Tested **architecture** as entry point and mid range CEPH offering in HPC environments.	Tested **architecture** as entry point and mid range Lustre offering in HPC environments.

2017 April Announcement

Future Aspirations

*formerly published as „ThinkStorage for SAP HANA TDI"

19

Lenovo Scalable Infrastructure (LeSI)

Lenovo Scalable Infrastructure (LeSI) is a framework for development, configuration, build, delivery and support of integrated data center solutions

- Complete HPC data center portfolio with the best-of-breed partner technology
- Collaborate on OpenSource HPC software in true commitment to Openess
- End-to-end expert-designed, tested, integrated and supported HPC solutions

Product Portfolio
Server
Storage
Networking
Options
Software

Scalable Infrastructure	
Scalable Infrastructure **Components**	Scalable Infrastructure **Solutions**
HPC Network systems/options	LeSI 1410 Solutions
Infrastructure options	LeSI 7X74 Solutions
OpenSource HPC Software (LiCO, xCat/Confluent, Antilles, ...)	Distributed Storage Solution for IBM Spectrum Scale

20

LeSI Solutions
From Complete Portfolio to Results Driven Designs

Effective 11/21/2017 for DSS-G

Best of the Lenovo Portfolio	+	Scalable Infrastructure OEM Portfolio	+	Scalable Infrastructure HPC Support	+	Factory Integration Services	=	LeSI Solutions

Best of the Lenovo Portfolio

Selection from off-the-shelf HPC proven Lenovo Portfolio

- Server
- Storage
- Networking
- Options
- Software

Scalable Infrastructure OEM Portfolio

Bringing in best-of-breed HPC partner components, especially HPC switches and adapters

- Mellanox FDR/ EDR Infiniband
- Intel Omnipath 100 Series

Scalable Infrastructure HPC Support

Testing hardware and software interoperability resulting in best recipe code level:

- SLES, RHEL, MOFED, IFS xCat/Confluent, Spectrum Scale

Factory Integration Services

1410 physical rack ship racked/cabled testing power & reduncancy, p2p cabling, best recipe, factory settings, run Linpack and load client xCat image or solution tested as above and shipped ready to install by Lenovo in customer choice of rack

LeSI Solutions

A single cluster, end-2-end designed & configured, built and shipped ready to accept the client images through xCat or alternatives.

Value Add Choice:

- Installation Services
- Managed Services
- Full HPC SW Suite

Customizing solutions for each client's unique needs

✚ Lenovo Distributed Storage Solutions : **DSS-G** for IBM Spectrum Scale

Lenovo DSS-G
Our first organic HPC Storage Offering

Lenovo Distributed Storage Solution for **IBM Spectrum Scale**

- Designed for extreme use cases

- High storage density and I/O performance with superior availability, reliability and resiliency

- HPC, BigData and Cloud

DSS-G integrates many components, but has **No** storage controller!

Servers:
x3650 M5 (2U)
Broadwell

Network:
DSS-G will work with any LAN supported by LeSI (e.g., EDR, OPA, Ethernet). While adapters are included with the solution, switches must be purchased separately; this allows customers to use existing fabrics
EDR
OPA

Storage Enclosures/Disks:
D1224 (2J24)
24 x 2.5" SAS/SSD

D3284 (5U84)
84 x NL-SAS

Disk Options: Many!
- GNR MD – (2) 12Gb 2.5" SAS, 3DWD: 400GB
- 15K, 12Gb 2.5" SAS: 300GD, 600GB
- 10K, 12Gb 2.5" SAS: 600GB, 900GB, 1.2 TB, 1.8TB
- 7.2K, 12Gb 2.5" SAS 1TB, 2TB
- 3DWD, 12Gb 2.5" SSD 400GB, 800GB, 1.6TB 800 GB SED (FIPS)
- 1DWD, 12Gb 2.5" SED 3.84TB

Disk Options:
- GNR MD – (2) 12Gb 2.5" SAS, 3DWD: 400GB
- Data – 12Gbit NL-SAS: 4TB, 6TB, 8TB, 10TB

Lenovo Distributed Storage Solution for **IBM Spectrum Scale**

- Acceptable short names
  - Lenovo DSS for Spectrum Scale
  - DSS-G (G= GPFS/Spectrum Scale)

- DSS-Gxyz
  - X = the number of DSS-G servers (always 2)
  - Y = the number of large storage enclosures (D3284)
    - Ex.: DSS-G220 = 2 servers & 2 D3284 enclosures
  - Z = the number of small storage enclosures (D1224)
    - Ex.: DSS-G204 = 2 servers & 4 D1224 enclosures

What is DSS-G?

DSS-G refers to the *Lenovo Distributed Storage Solution for Spectrum Scale^{TM}*(aka GPFS). It is a storage *solution* based on Spectrum Scale RAID(aka, GPFS Native RAID or GNR). DSS-G integrates many components as part of a *complete storage solution*.

DSS G202
SSD or SAS

DSS G206
SAS

DSS G240
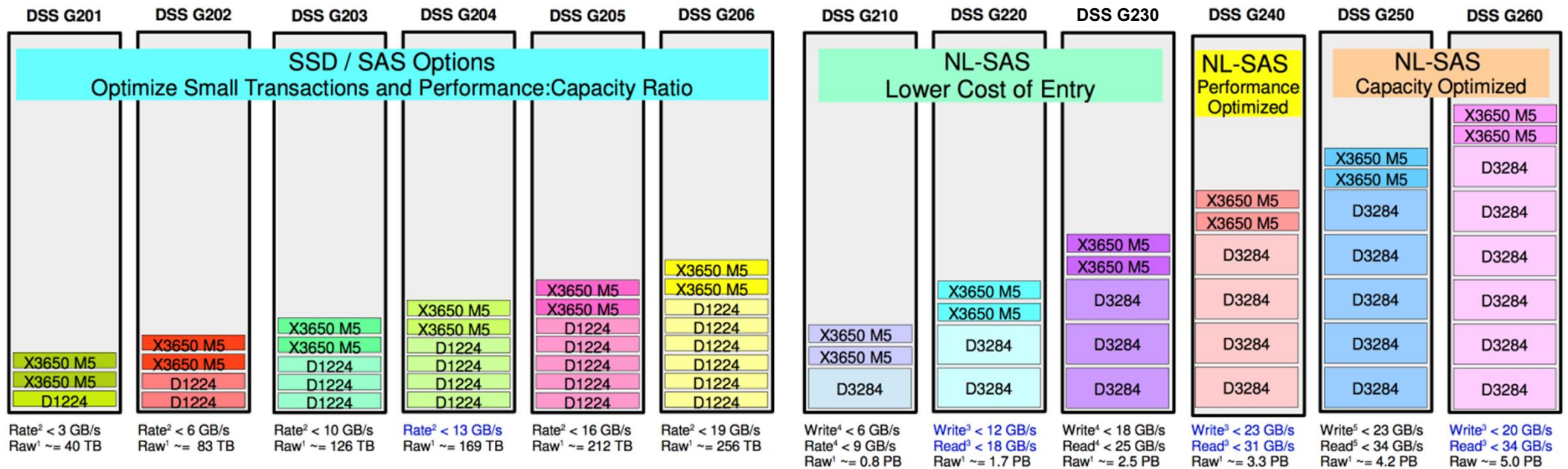NL-SAS

DSS-G Head node

2 x D1224

6 x D1224

Lenovo Storage D1224 enclosures/disk
2U rack mount enclosure SAS direct-attached storage
Holds 24x 2.5-inch hot-swap drive bays
Selectable May capacity SAS HDD / SSD

DSS-G
JBOD Storage

Lenovo Storage D3284 Drive enclosures
5U rack mount enclosure with 12 Gbps SAS direct-attached storage
Holds 84x 3.5-inch hot-swap drive bays in two drawers.

23

DSS-G line up

The complete set of configurations supported in later releases this year.

DSS G201	DSS G202	DSS G203	DSS G204	DSS G205	DSS G206	DSS G210	DSS G220	DSS G230	DSS G240	DSS G250	DSS G260
SSD / SAS Options Optimize Small Transactions and Performance:Capacity Ratio						**NL-SAS** Lower Cost of Entry			**NL-SAS** Performance Optimized	**NL-SAS** Capacity Optimized	
											X3650 M5
											X3650 M5
										X3650 M5	D3284
									X3650 M5	X3650 M5	D3284
								X3650 M5	X3650 M5	D3284	D3284
					X3650 M5			X3650 M5	D3284	D3284	D3284
			X3650 M5	X3650 M5	X3650 M5		X3650 M5	D3284	D3284	D3284	D3284
		X3650 M5	X3650 M5	D1224	D1224	X3650 M5	X3650 M5	D3284	D3284	D3284	D3284
	X3650 M5	X3650 M5	D1224	D1224	D1224	X3650 M5	D3284	D3284	D3284	D3284	D3284
X3650 M5	X3650 M5	D1224	D1224	D1224	D1224	D3284	D3284	D3284	D3284	D3284	D3284
X3650 M5	D1224	D1224	D1224	D1224	D1224						
D1224	D1224	D1224	D1224	D1224	D1224						

| DSS G201	DSS G202	DSS G203	DSS G204	DSS G205	DSS G206	DSS G210	DSS G220	DSS G230	DSS G240	DSS G250	DSS G260
Rate[2] < 3 GB/s	Rate[2] < 6 GB/s	Rate[2] < 10 GB/s	Rate[2] < 13 GB/s	Rate[2] < 16 GB/s	Rate[2] < 19 GB/s	Write[4] < 6 GB/s Rate[4] < 9 GB/s	Write[3] < 12 GB/s Read[3] < 18 GB/s	Write[4] < 18 GB/s Read[4] < 25 GB/s	Write[3] < 23 GB/s Read[3] < 31 GB/s	Write[5] < 23 GB/s Read[5] < 34 GB/s	Write[3] < 20 GB/s Read[3] < 34 GB/s
Raw[1] ~= 40 TB	Raw[1] ~= 83 TB	Raw[1] ~= 126 TB	Raw[1] ~= 169 TB	Raw[1] ~= 212 TB	Raw[1] ~= 256 TB	Raw[1] ~= 0.8 PB	Raw[1] ~= 1.7 PB	Raw[1] ~= 2.5 PB	Raw[1] ~= 3.3 PB	Raw[1] ~= 4.2 PB	Raw ~= 5.0 PB

Comments:
1. Capacity for D1224 is based on 1.8TB, 2.5", 10KRPM SAS.
   Capacity for D3284 is based on 10TB NL-SAS
2. Performance is an *estimate* based on raw disk tests(n.b., no filesystem) on G204(1.2TB, 10KRPM, 2.5" SAS)
3. Performance based on *actual* G220 and G240 benchmarks
4. Performance is an *estimate* based G220 and G240 benchmarks
5. Performance is an *estimate* based G220 and G260 benchmarks

Lenovo

Lenovo's Event information
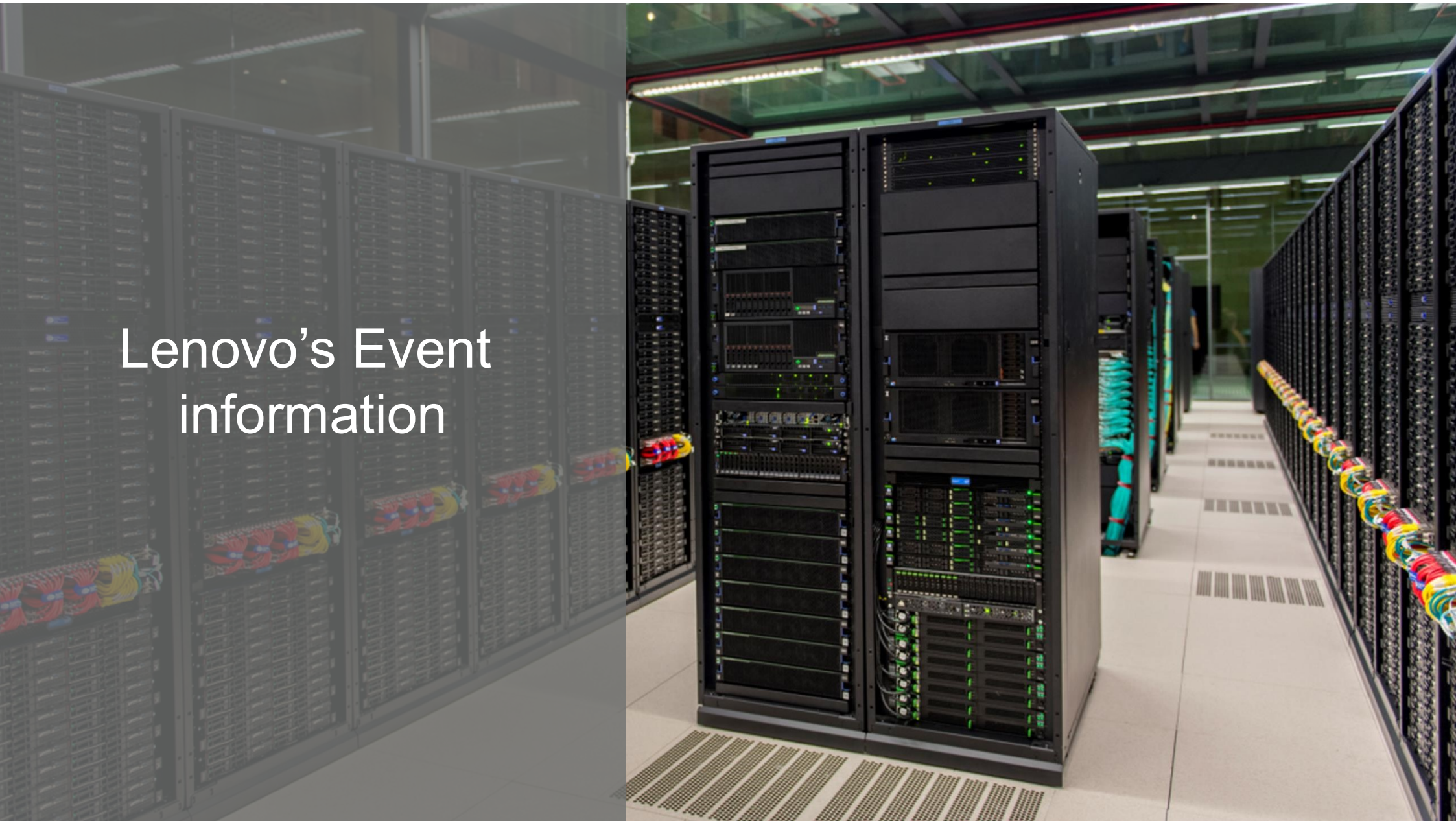
Please join **KEK Facility Tours**
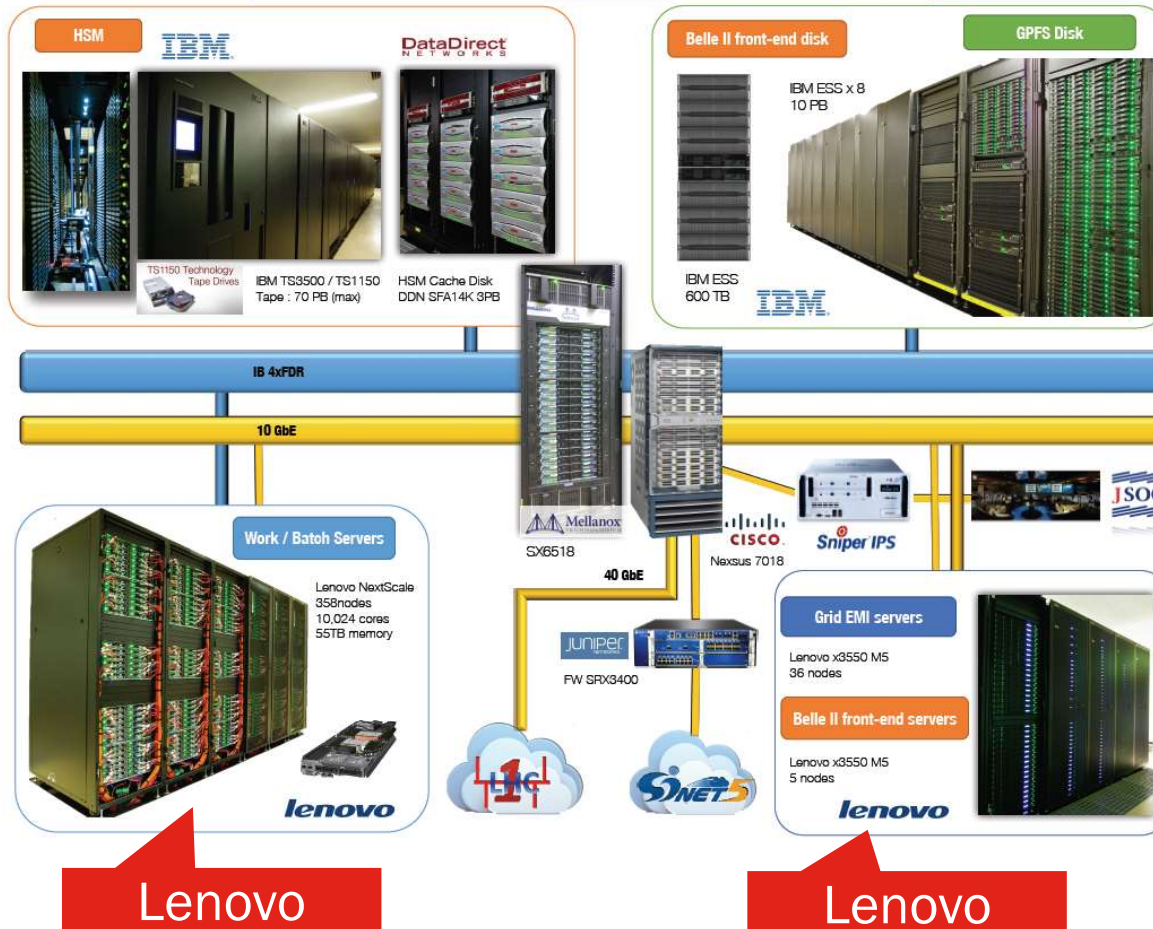
Lenovo provided "KEK Central Computer System : Work/Computing Server"
Lenovo NeXtScale nx360 M5 10,000 cores

The Data Analysis System provides high-performance computing servers and a large-capacity storage system for accumulation and analysis of large amounts of experimental data.

The computing servers, which comprise approximately 10,000 CPU cores, can be used for both interactive and batch analyses.
Ref. https://www.kek.jp/en/PublicRelations/DigitalLibrary/2017/01/17/Eng.pdf

Please join KEK Facility Tours.
Friday after lunch,
October 20th 13:00 - 14:30

Please Visit Lenovo at **Super Computing 2017**

Solve humanity's greatest challenges with a future-defined data center.

Visit Lenovo Data Center Group at SC17

November 12-17, 2017
Denver, CO

Booth #1353

ThinkAgile
ThinkSystem

Lenovo helps our customers solve the greatest challenges of today's world with our HPC solutions.
As the *fastest-growing supercomputing and HPC company and research*, Lenovo is uniquely positioned to help you accomplish your research and development goals.

At SC17, visit Lenovo at **Booth #1353** to see exclusive demos, meet with Lenovo Data Center experts, and envision your future-defined data center.

REMIND ME TO VISIT LENOVO ▸

http://eblasts.cleandesign.com/lenovo/2017-10-06-super-computing/Visit-Lenovo-Booth-1353.ics

BOOK A BRIEFING ▸

https://esg.formstack.com/forms/sc17

thanks.

Different is better

Lenovo™