Tecnologie emergenti per nuovi ambiti applicativi: HPC/A.I./D.L.

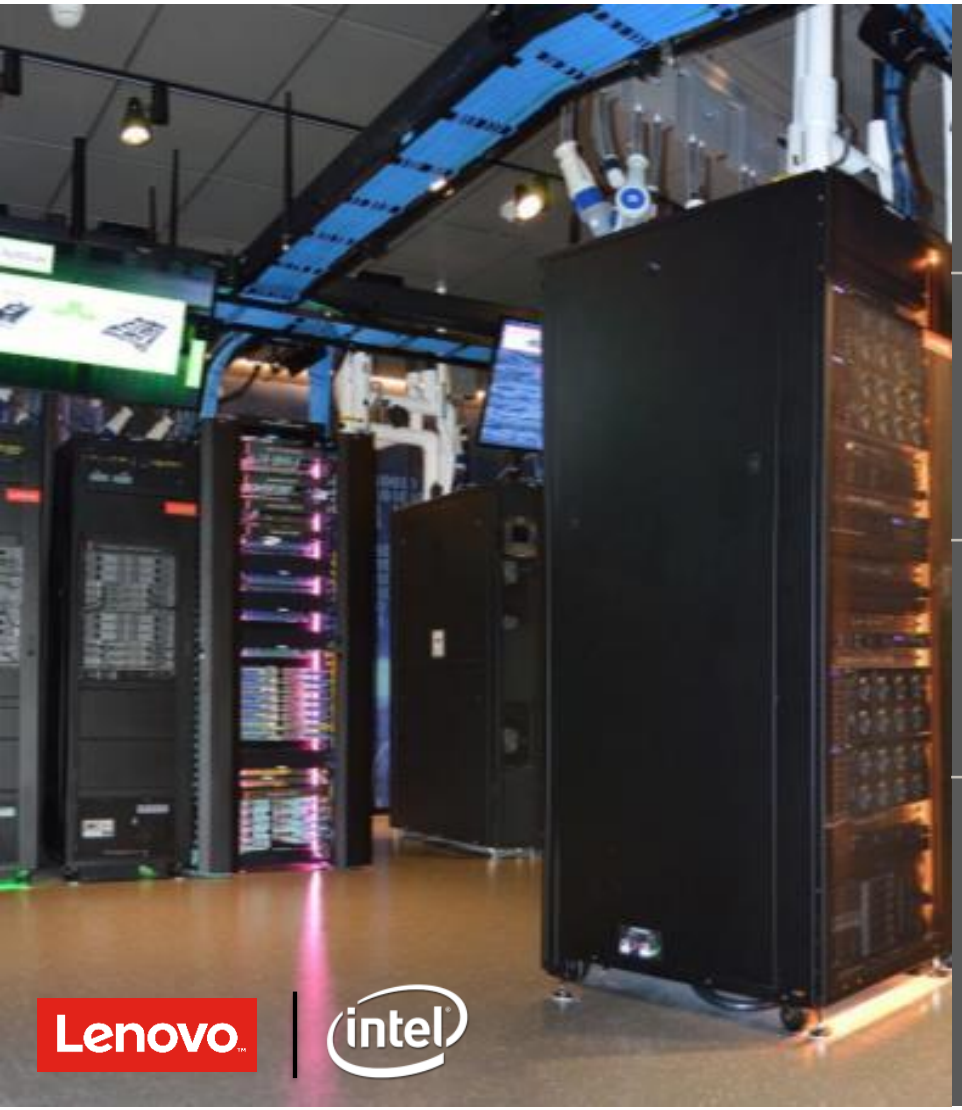
Marco Briscolini, HPC&AI Sales Leader
Data Center Group, Lenovo Global Technology Italy
E-mail: mbriscolini@Lenovo.com

Lenovo™

Best in Class

Best Use of High Performance Data Analytics	Best Use of HPC in Life Sciences	Best HPC Server Product or Technology	Best HPC Storage Product or Technology	Best HPC Collaboration (Academia/ Government/ Industry)
Vestas reducing "Loss Prevention Factor" for turbines	Birmingham real-time Zika genome analysis	Lenovo ThinkSystem SD530	Lenovo DSS-G	Birmingham providing CLIMB cloud computing

DISCOVER AI benefits with the Lenovo AI Innovation Centers



INFRASTRUCTURE

ThinkSystem

PARTNERS

RESEARCH INSTITUTIONS

TECHNOLOGY PARTNERS

ECOSYSTEM PARTNERS

FRAMEWORKS

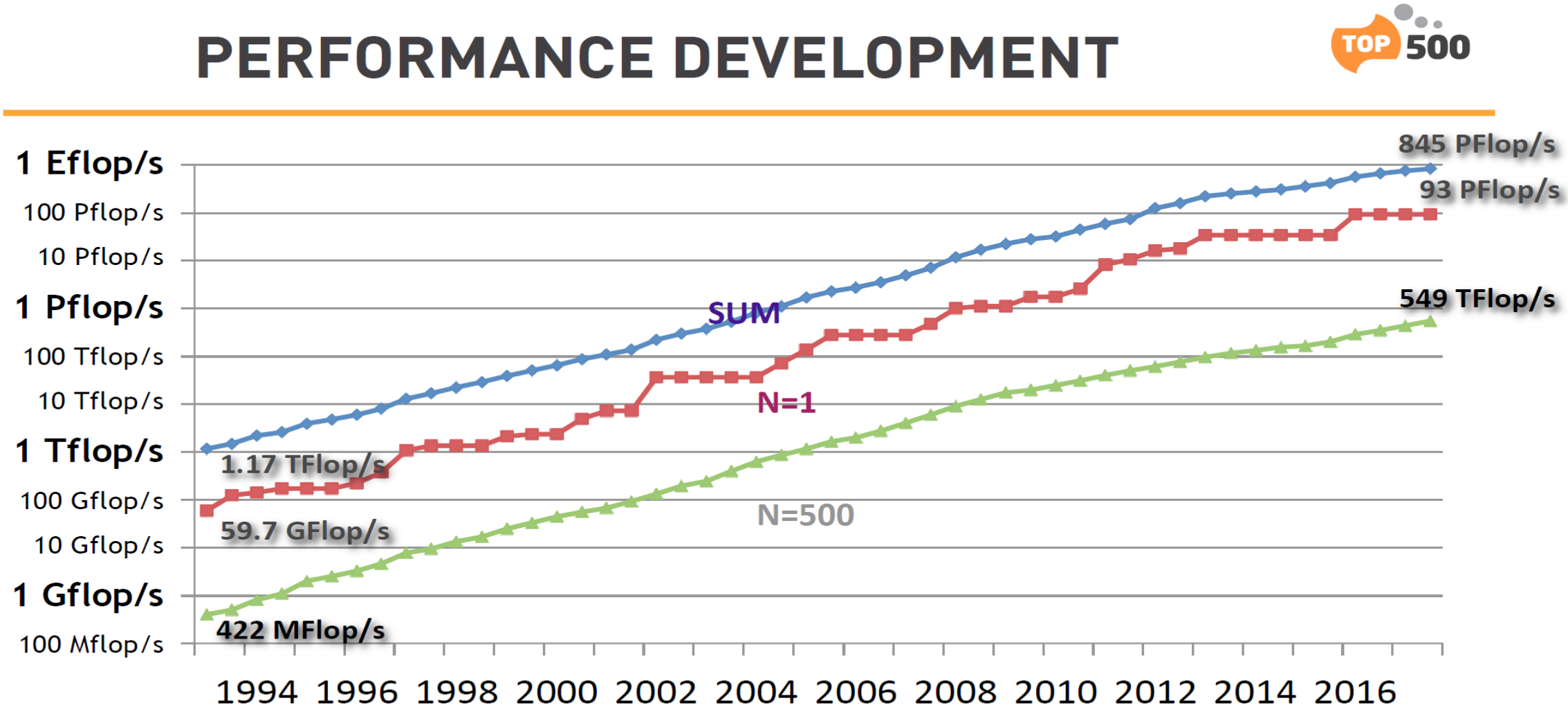
Caffe2 torch theano mxnet

Innovation through Collaboration

Delivering the power to help solve humanity's greatest challenges

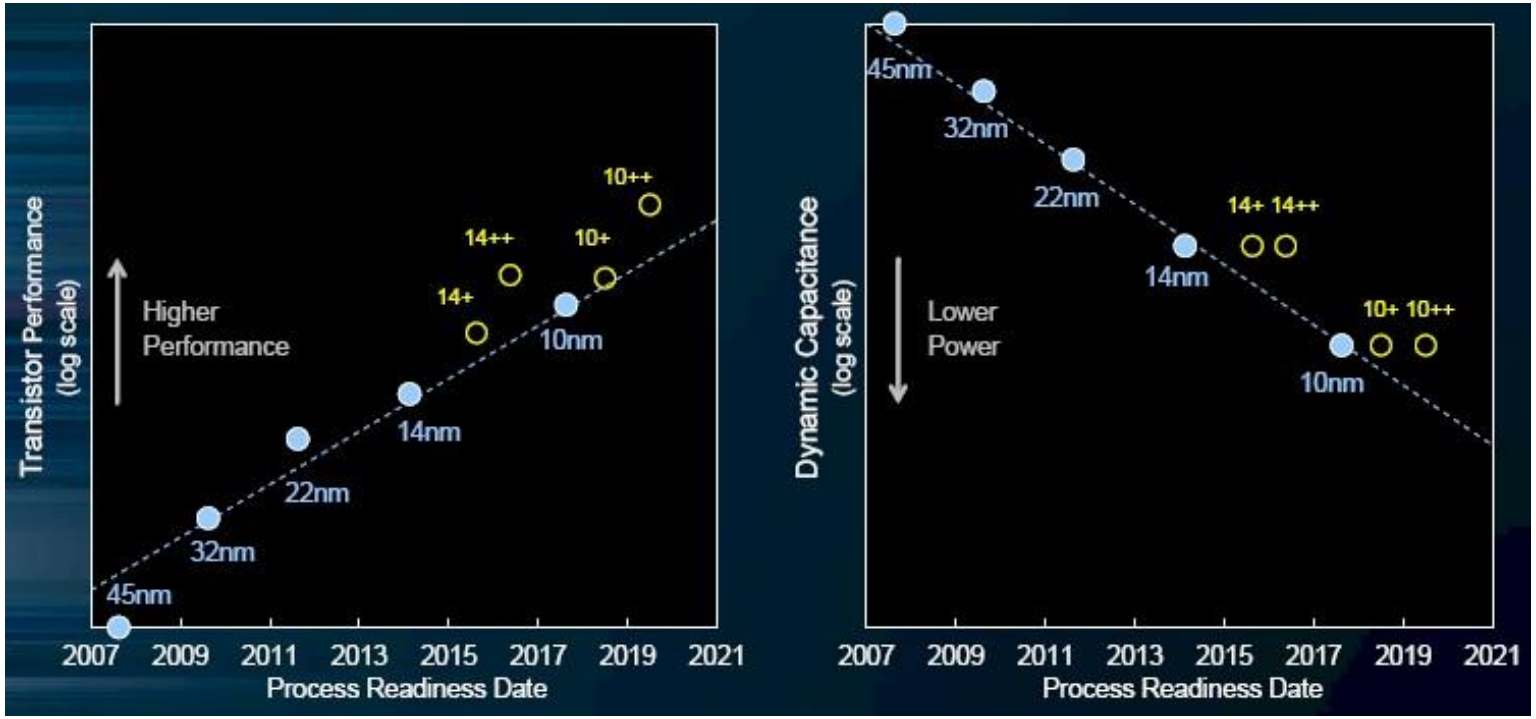
- Collaboration with industry partners and clients
- Barcelona Super Computing Center
  - Energy aware MPI runtime
  - HPC Virtual environment investigation
- CINECA
  - Open System Management
  - GPFS Filesystem optimization on Intel OPA
- LRZ
  - High dense DWC architecture
  - Energy Aware Runtime
- HPC Benchmarking and Innovation Centers
  - Industry leaders to bring the newest technology and skills together for greater outcomes
  - Collaboration with visionary client partners to bring focused knowledge and deep skills in specific areas of science

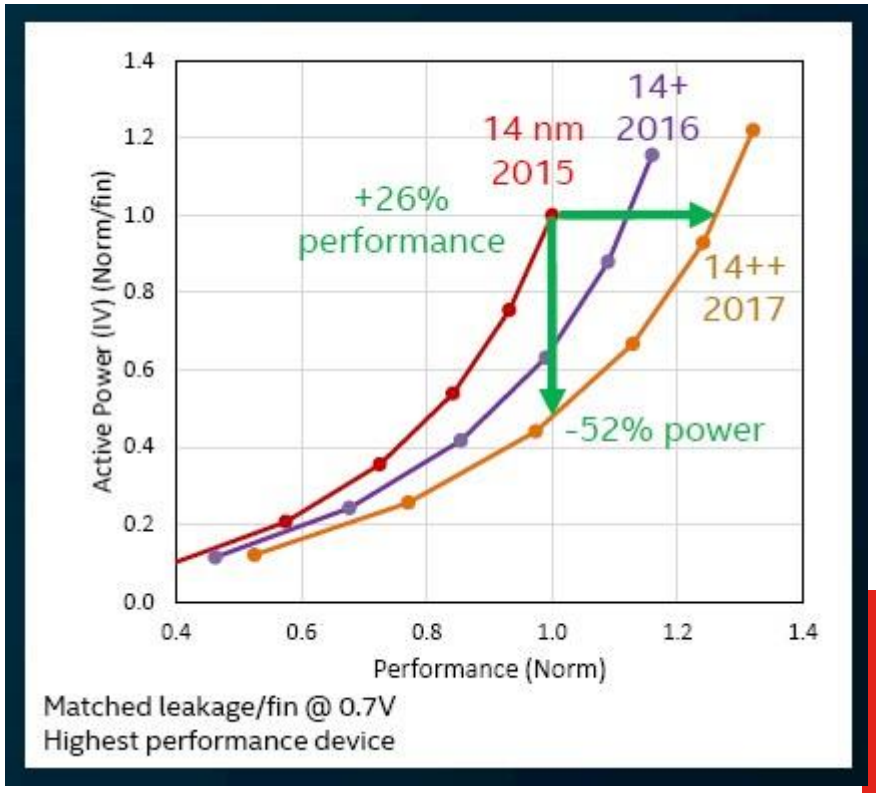
HPC technology trends



HPC technology trends

An example of technology evolution: the Intel x86 micropocessor

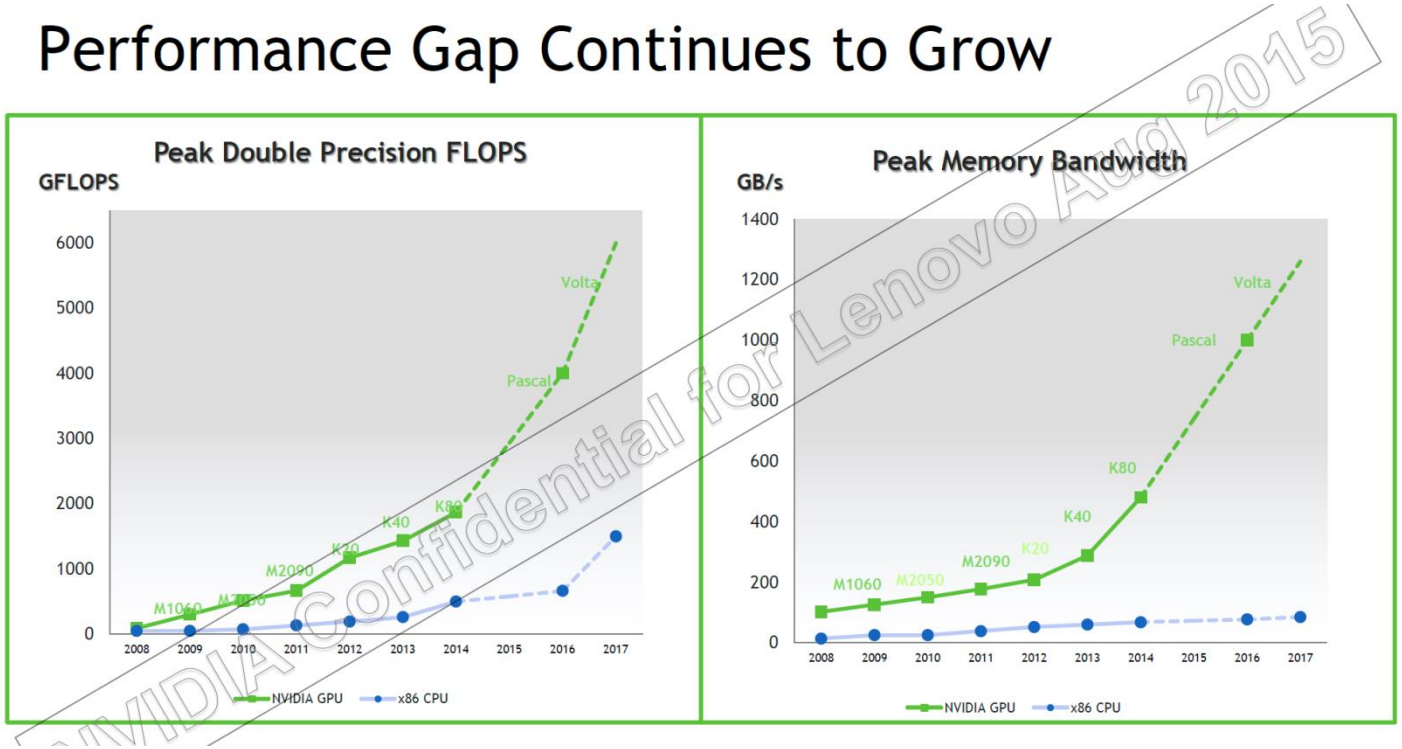


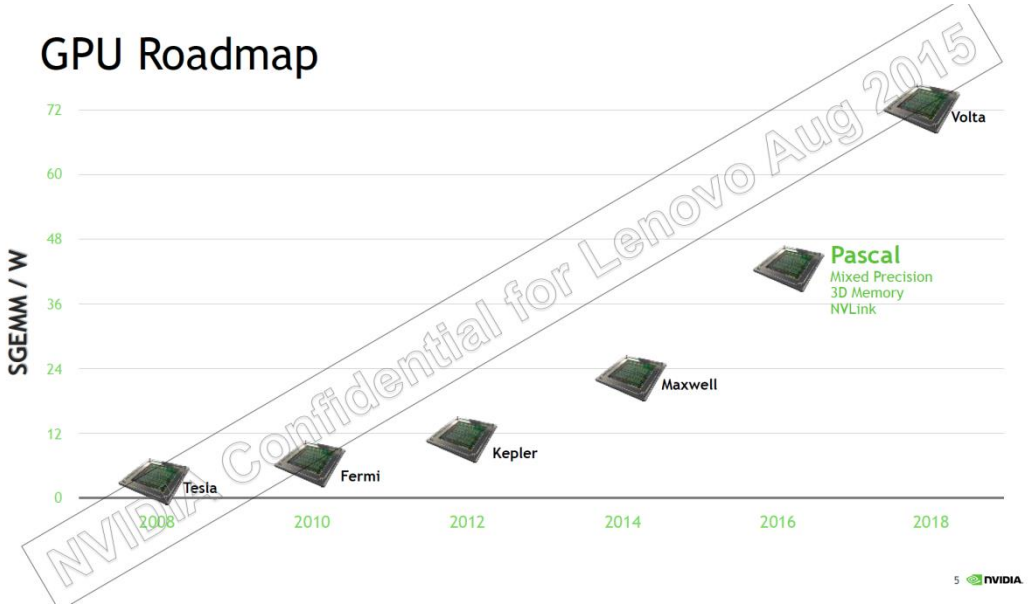


https://www.nextplatform.com/2017/04/27/mapping-intels-tick-tock-clock-onto-xeon-processors/

HPC technology trends

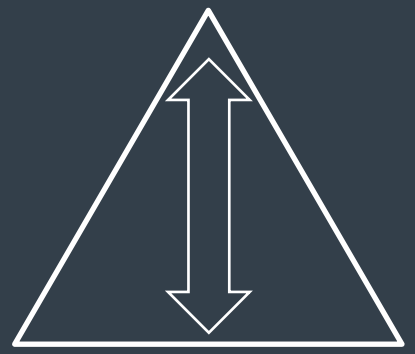
Hybrid computing: CPU+GPU makes the difference as performance and consumption





Lenovo™

HPC in any shape or form



Supercomputing
Divisional
Departmental
Workgroup

Hardware	**Software**
Service	**Solutions**

ACADEMIA

LIFE SCIENCE

OIL & GAS

METEOROLOGIE

ENGINEERING

The Lenovo ThinkSystem Server Portfolio

Datacenter
Business Intelligence
Analytics
Cloud
Virtualization
Retail
Distributed Enterprise
Small/Med Business

Tower

2P

ST550

Racks

2P

2U

1U

SR650

SR630

SR590

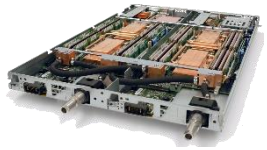
SR570

SR550

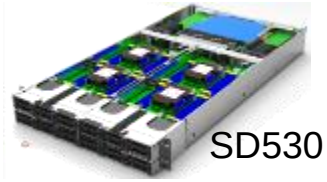
SR530

Dense

2P



SD650



SD530

Blades

4P

SN850

2P

SN550

Mission Critical

8P

SR950

4P

SR860
(EMEA, PRC)

4P

SR850

Lenovo™

ThinkSystem HPC Dense

- Our biggest product launch ever
- Celebrating 25 years in x86 servers
- Top BIN processors and networks
- New GPU capability shipping soon
- Highest performance/lowest cost

ThinkSystem SD530

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) • 4x ApachePass future ready
Storage	• 6x SAS/SATA SF HS HDD/SSD or 2x NVMe, 2x M.2 SATA SSD
NIC	2x 10G SFP+ or 10GBase-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	XCC Stark Mgmt Module, 1 / 2 shared GbE ports 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

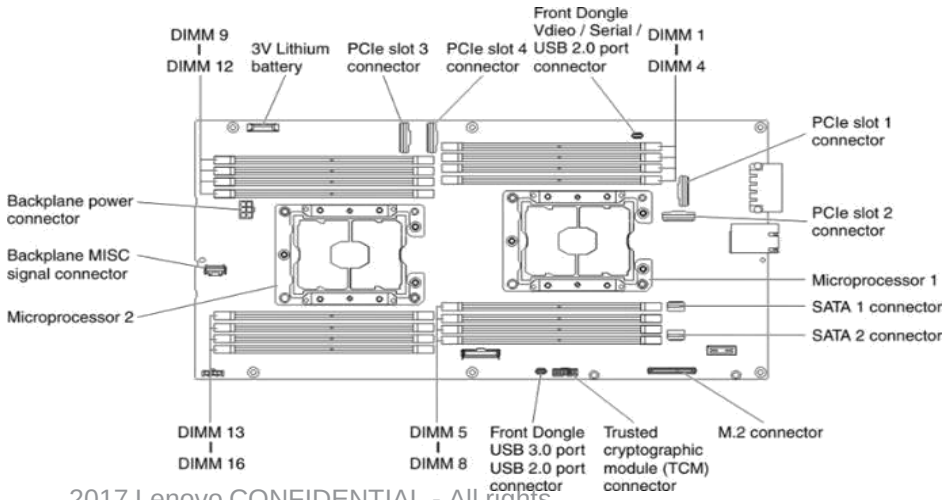
12 DDR4 DIMMs
+ 4 Apache Pass

2x SATA or PCIe M.2 BootSSD

2S Purley CPU max
~ 165W*

Hardware RAID (Optional)

4 or 6x 2.5 front hot swap drives

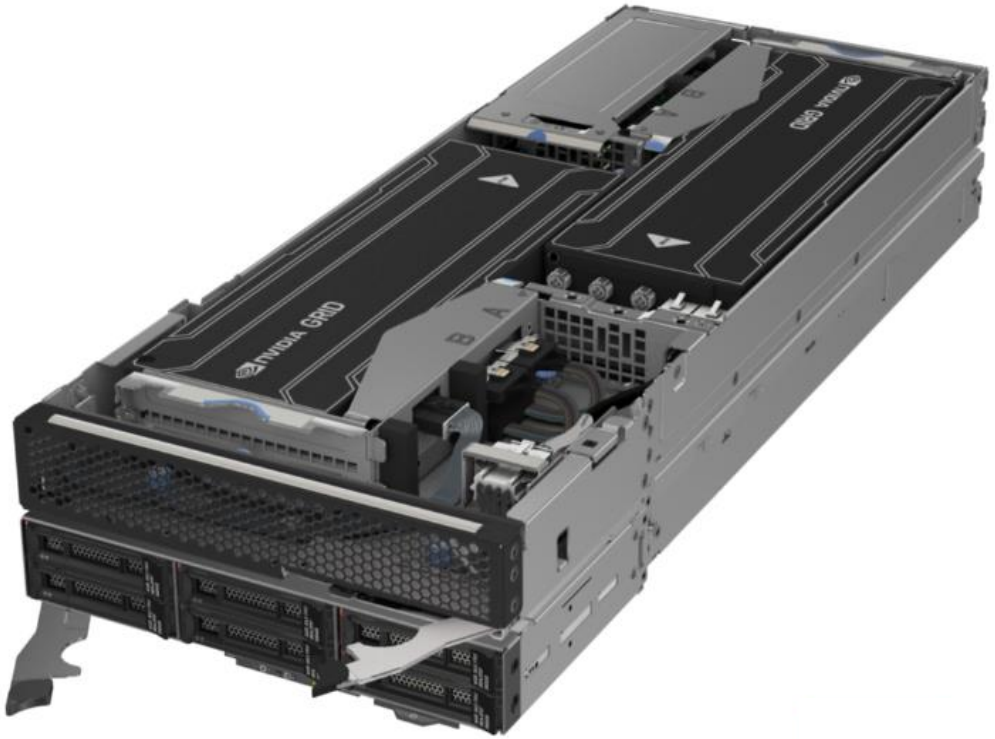


*in 4x drive configuration only

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

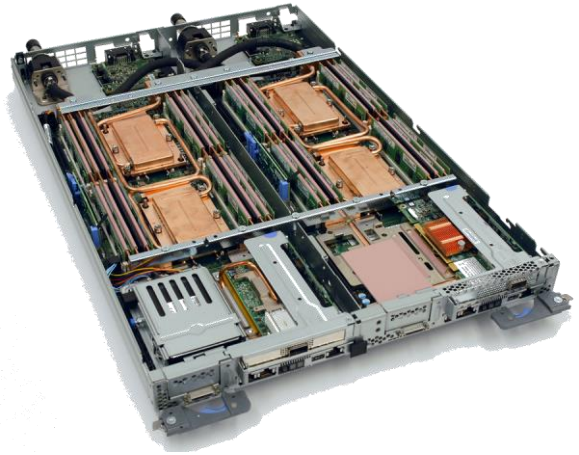


Supported Accelerators
- NVIDIA Tesla M10 32GB (250 W)
- NVIDIA Tesla M60 16GB (300W)
- NVIDIA Tesla P40 24GB (250 W)
- NVIDIA Tesla P100 16GB (250 W)
- NVIDIA Tesla V100 16GB (300 W)

… more planned for the future

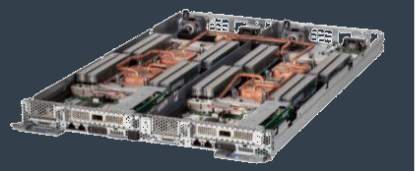
Lenovo™

Direct “Hot” Watercooling

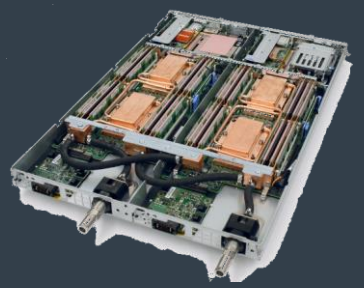


- Up to 50°C Hot Water Cooling
- Up to 90% Heat Removal Efficiency
- World Record Energy Reuse Efficiency
- Many patents on market leading design

2012



2014

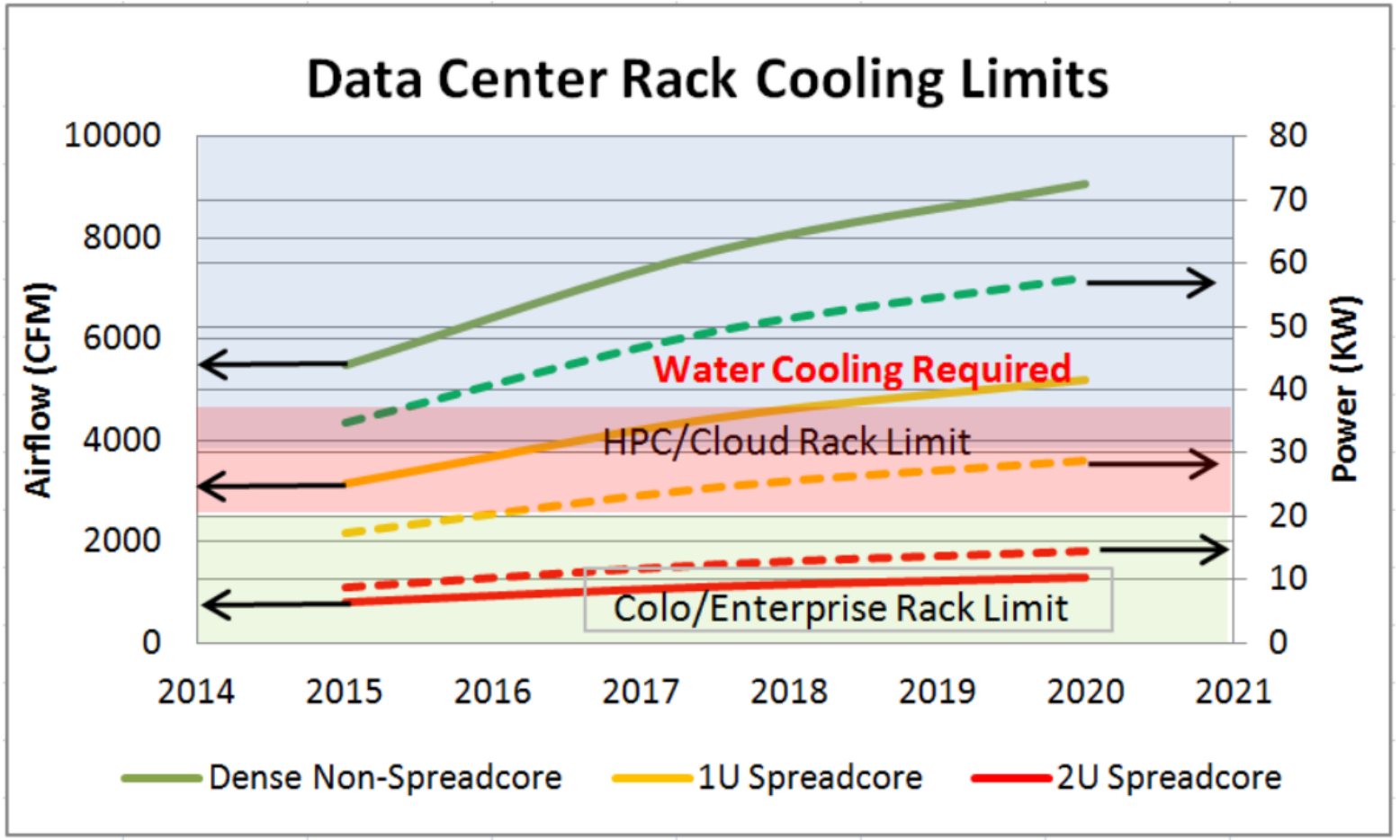


2018

>20.000 nodes globally

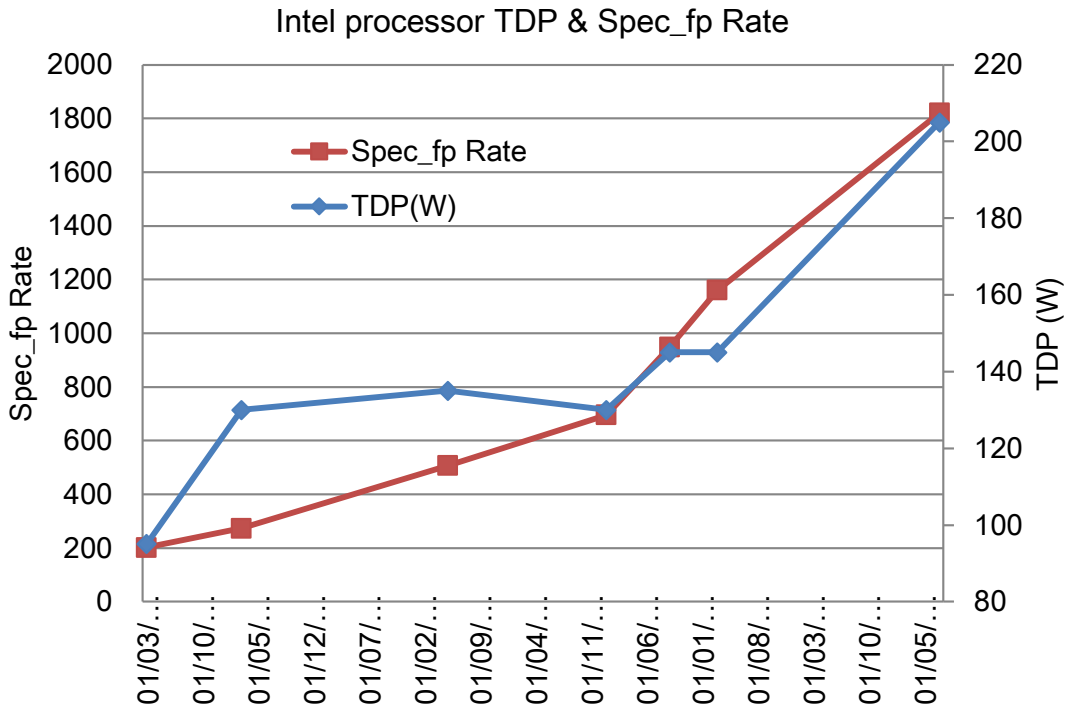
Data Center Level Cooling and Power Limits

- Node power density trends cannot be cooled at data center level
  - Partial rack population or rack level power capping may be required for Dense and 1U



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

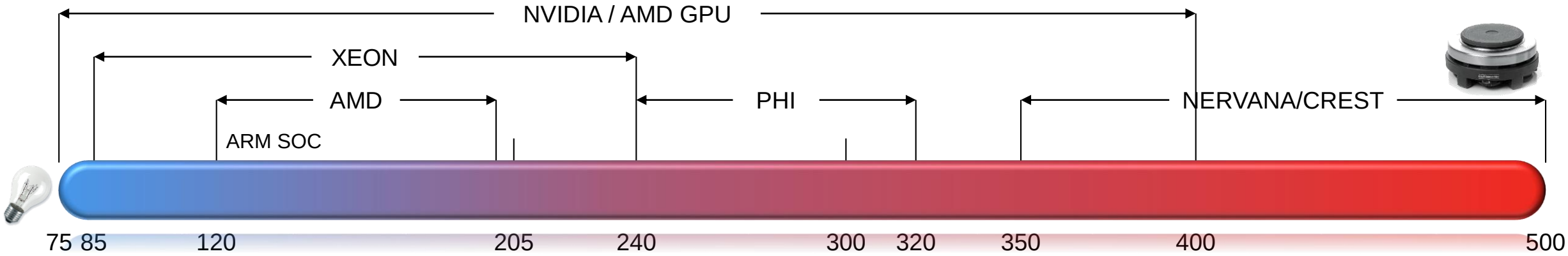


Processor performance trend

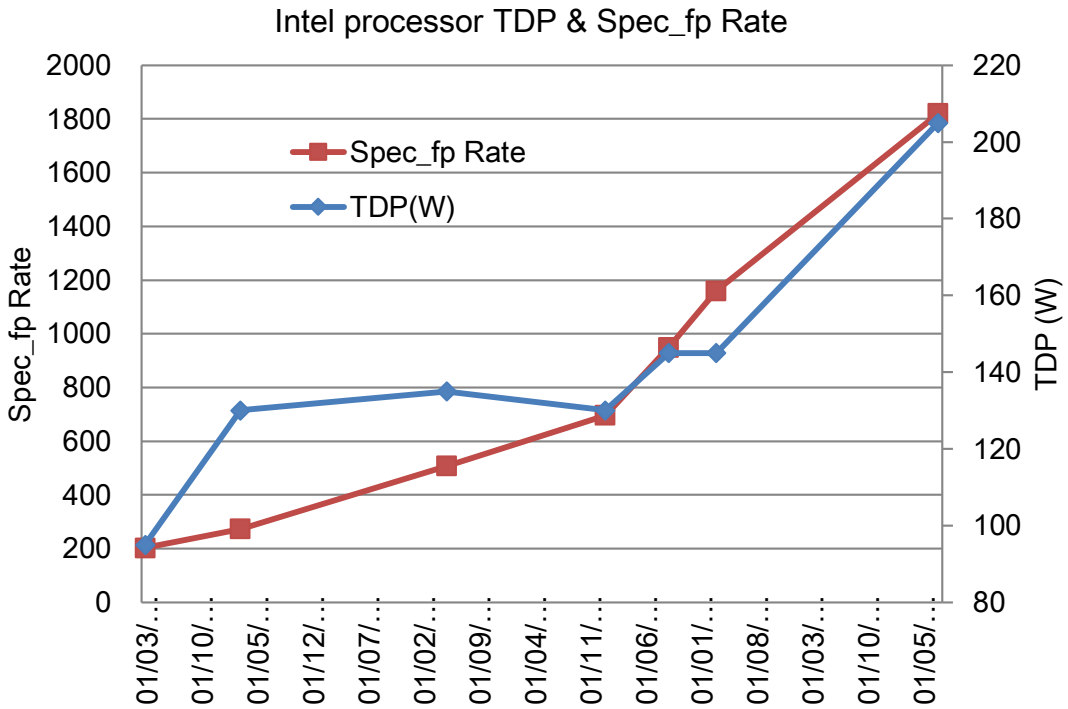
- Spec_fp rate with 2 processors/node has increased 40 times the past 11 years (2006 – 2017).
- The number of cores on the chip increase 14 times.
- After being flat, since 2014 TDP increases linearly with Spec_fp rate.
- Current maximum TDP is 205W. Knighs Mill Xeon phi processor will be 305 W

To sustain increased performance servers will have to be less dense or use new cooling technology

Industry Thermal Challenges



- Maintaining Moore's Law with increased competition is resulting in higher component power
- Increased memory count, NVMe adoption, and I/O requirements are driving packaging and feature tradeoffs (superset of features doesn't fit in 1U)
- Shared cooling fan power savings no longer exist for dense 2S nodes architectures due to non-spreadcore CPU layout high airflow requirements
- Industry moving away from dense nodes (e.g. Open 19 and OCS Olympus)

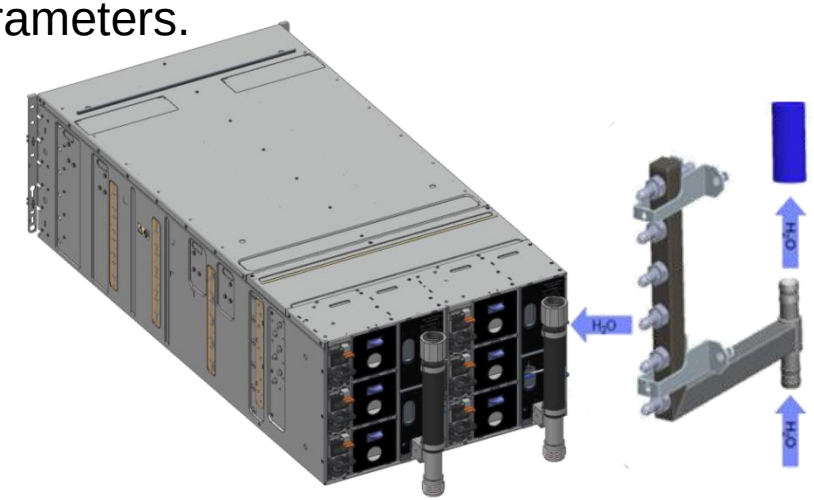


Heat Extraction/Performance Expectations

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

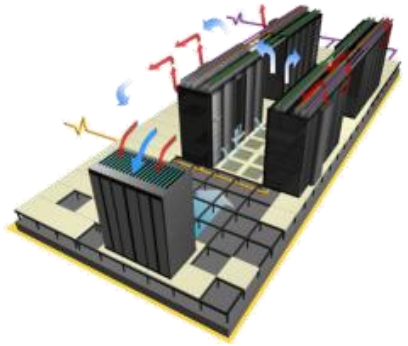
Assumptions: 36x compute trays (dual node), 72 servers + 6x switches / Rack

- 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.



Choice of Cooling

Air Cooled



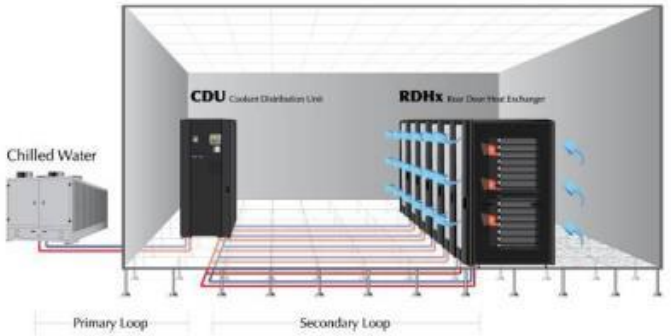
- Standard air flow with internal fans
- Fits in any datacenter
- Maximum flexibility
- Broadest choice of configurable options supported
- Supports Native Expansion nodes (Storage NeX, PCI NeX)

PUE ~1.5

ERE ~ 1.5

Choose for broadest choice of customizable options

Air Cooled with Rear Door Heat Exchangers



- Air cool, supplemented with RDHX door on rack
- Uses chilled water with economizer (18C water)
- Enables extremely tight rack placement

PUE ~1.2

ERE ~ 1.2

Choose for balance between configuration flexibility and energy efficiency

Direct Water Cooled

- Direct water cooling with no internal fans
- Higher performance per watt
- Free cooling (45C water)
- **Energy re-use**
- Densest footprint
- Ideal for geos with high electricity costs and new data centers
- Supports highest wattage processors

PUE <= 1.1

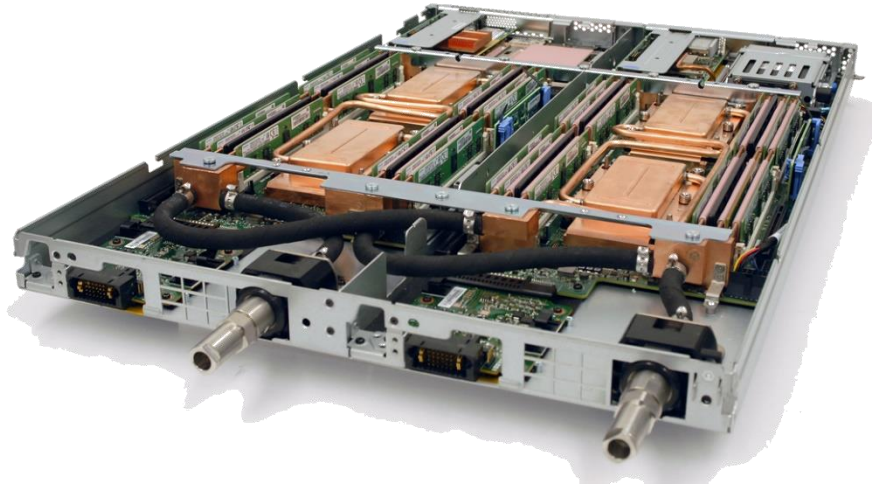
ERE ~ 0.3 **with hot water**

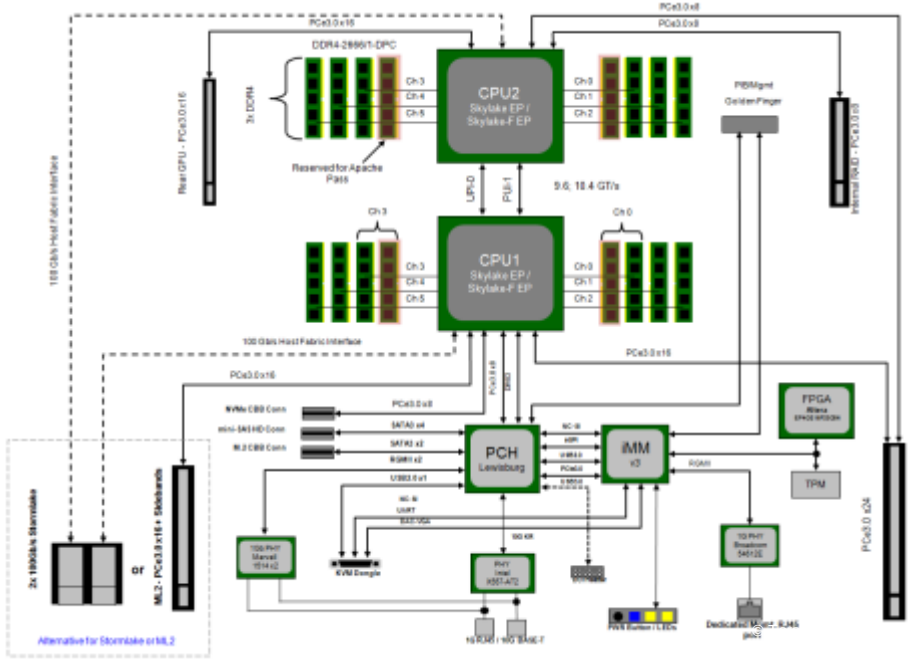
Choose for highest performance and energy efficiency

ThinkSystem SD650 (OceanCat)

GA Q1/2018

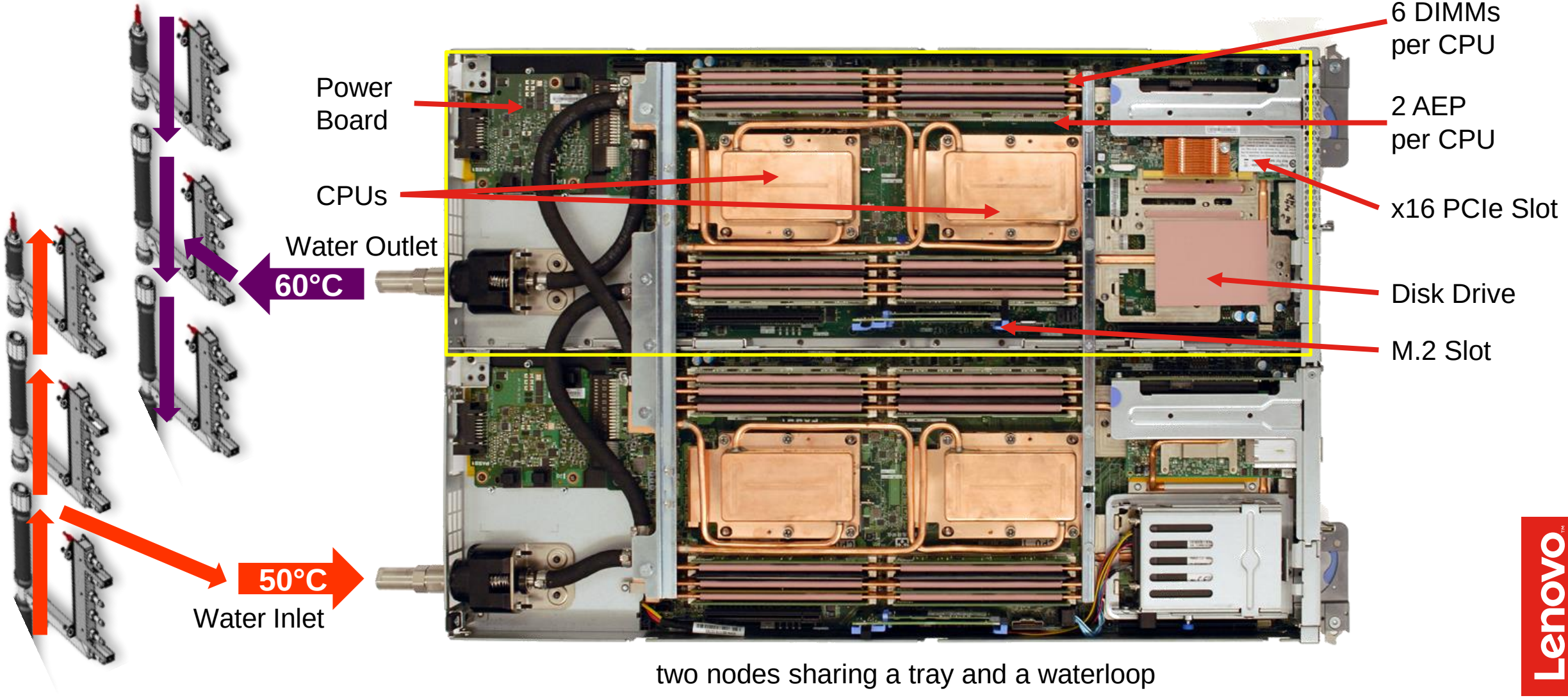
Feature	SD650
Processors	2 Intel “Purley” Generation processors per node • Socket-F for Intel Omnipath supported • >120W all Skylake Shelves supported
Form factor	1U Full wide tray double-node / 6U12N Chassis
Memory Slots Max Memory	• 12x DDR4 (R/LR) 2667MHz DIMM • 4x Intel Apache Pass DIMM
Storage	• 2x SATA slim SSD, 1x NVMe, 2x M.2 SATA SSD
NIC	1x 1 GBaseT, 1x 1 GbE XCC dedicated
PCIe	1x x16 PCIe EDR Infiniband, HDR with internal x16 1x x16 ML2 for 10Gbit Ethernet Internal: 1x x8 for RAID, 1x x16 for pot. expansions
Power	1300W/1500W/2000W Platinum and 1300W Titanium
USB ports	Up to 1x front via dongle cable + 1x internal (2.0)
Cooling	• No fans on chassis, PSU fans only • Up to 50°C warm water circulated through cooling tubes for component level cooling
System MGMT / TPM	XCC, dedicated port or shared TPM, Pluggable TCM
Dimensions	915mm depth, front access w/ front I/O





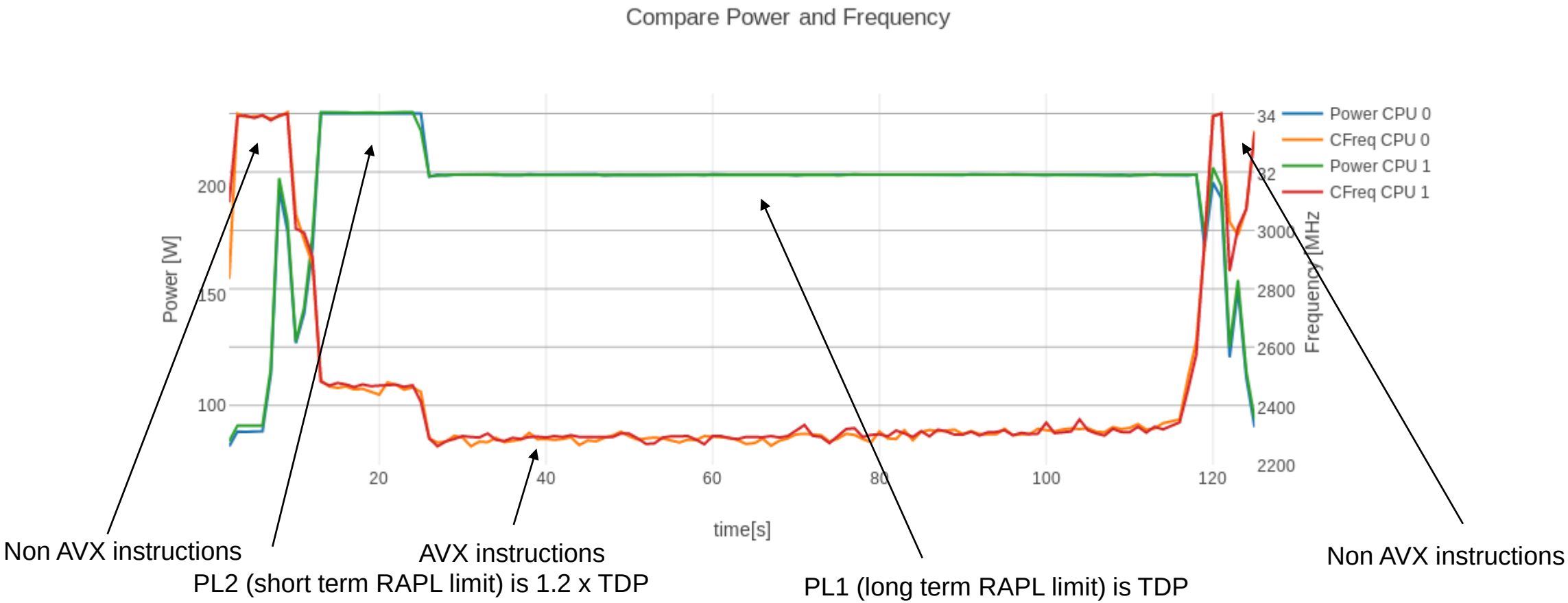
ThinkSystem SD650 (OceanCat)

Top-Down View



two nodes sharing a tray and a waterloop

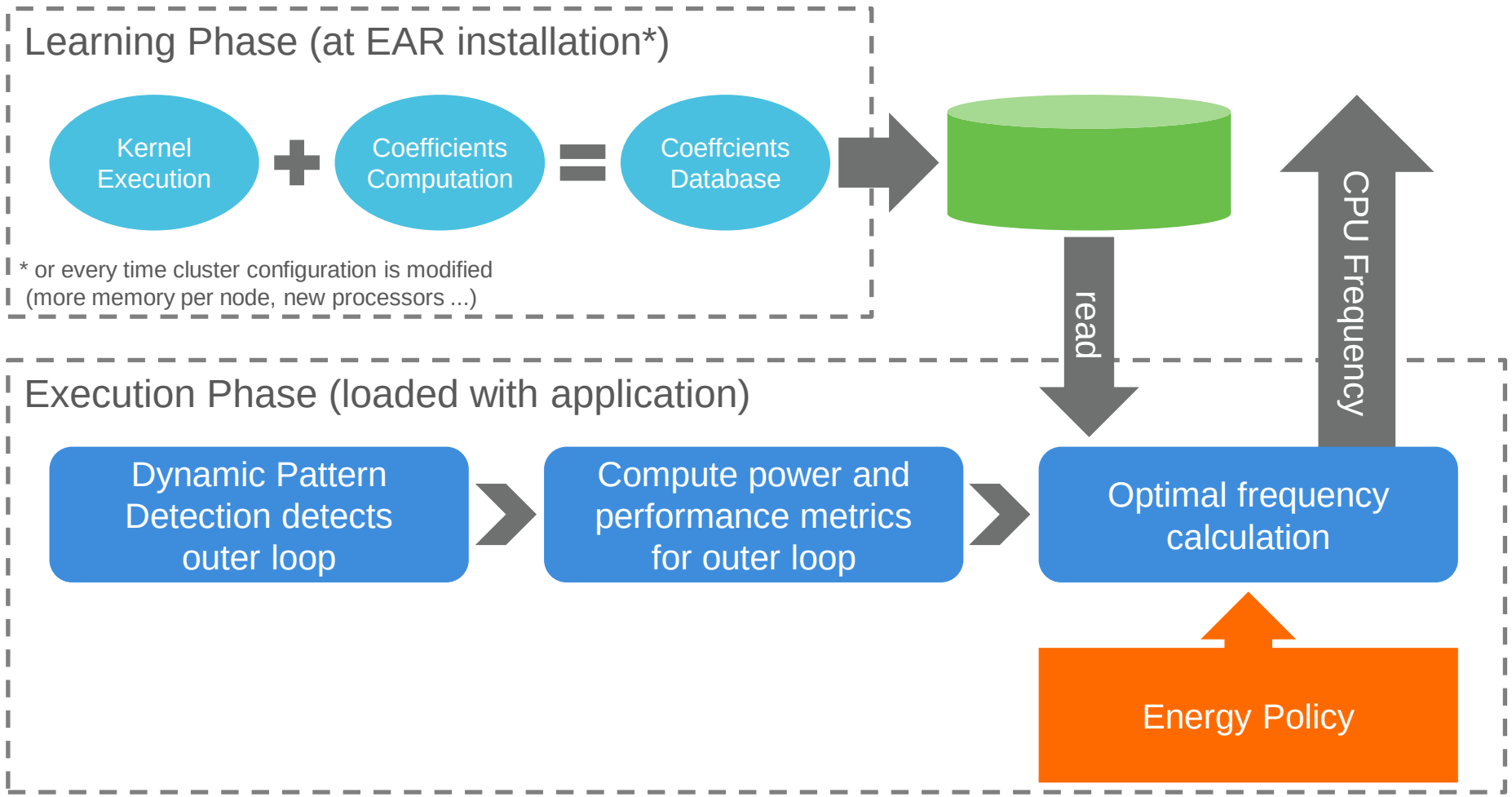
HPL DC Power & frequency on SD650 with 8168



SD650 with 2 sockets 8168 and 6 x 16GB DIMMs; room temp = 21°C, inlet water = 40°C, 1.5 lpm/tray

Lenovo Energy Aware Runtime (LEAR)

Functional Overview



Lenovo™

ThinkSystem HPC & AI

Common AI/ML Terminology

Inference	• Using data training instead of programming to generate output
Intelligence Agent(s)	• A device that interacts with its environment to achieve a goal
Ontology	• A set of objects, concepts and data, about which a device is aware
Machine Learning	• Utilizing algorithms to train data that improves results over time via experience
Perception	• Using input from external devices (cameras, sensors) to make decisions
Deep Learning	• Utilizing neural networks to train data in a way similar to the way human learning occurs
Cognitive computing	• Systems that mimic human brain (self-learning systems)

Key Terms for A.I. in your case

Predictive Analytics

Image or voice recognition

Natural language processing

Autonomous machine interaction or Robotics

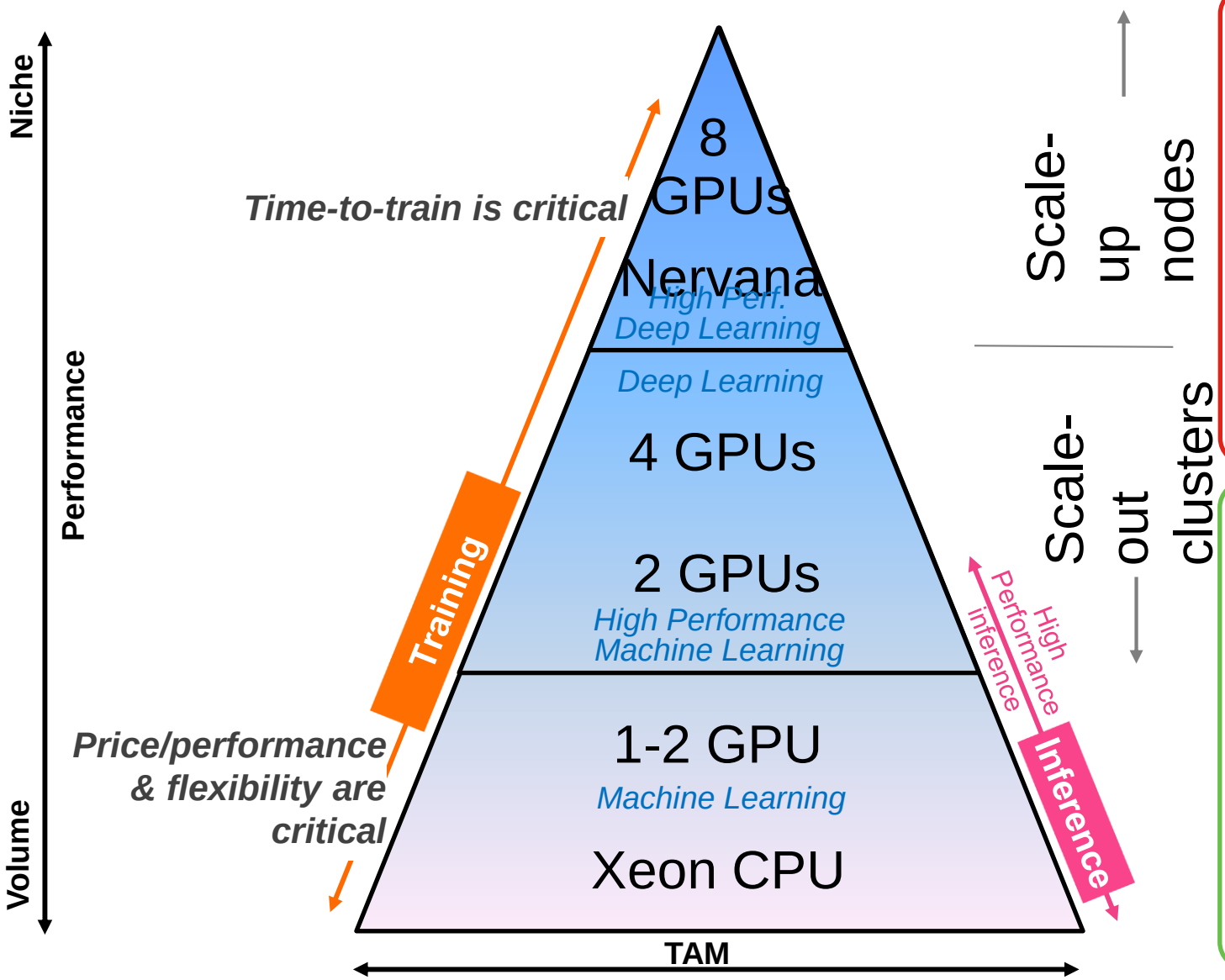
Unstructured data processing

IoT

Diagnostics

Training systems or scoring or inference systems

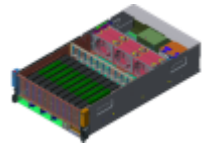
DCG AI and Deep Learning HW portfolio



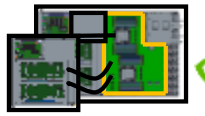
Scale-up nodes

Scale-out clusters

Working Plan

DGX-1/2 (VLH)  NVIDIA

Future Nervana ASIC intel

Greenhill: 4 GPUs  NVIDIA

Products In Plan

SD530 + 2 GPUs  NVIDIA

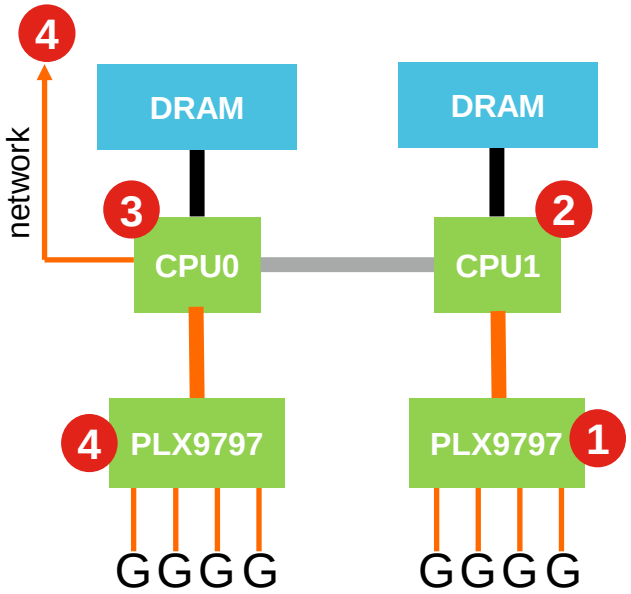
SR650 + 2 GPUs  NVIDIA

SD530/SR650  intel

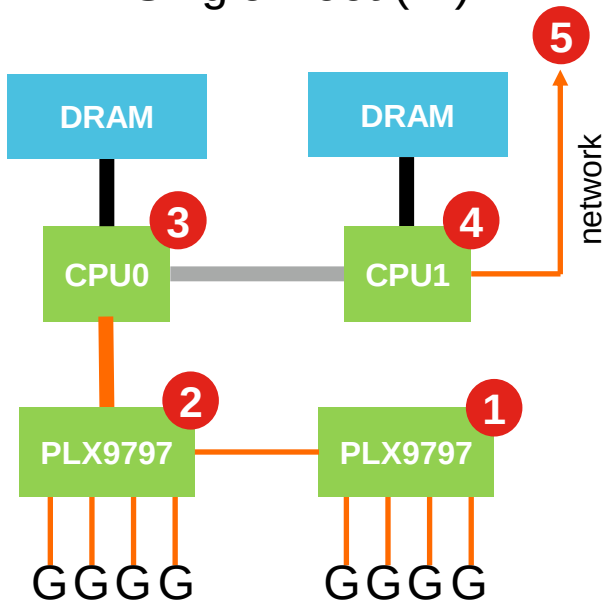
One Size Doesn't Fit All

4:1 GPU Server

Balanced Mode (HPC)



Single Root (AI)

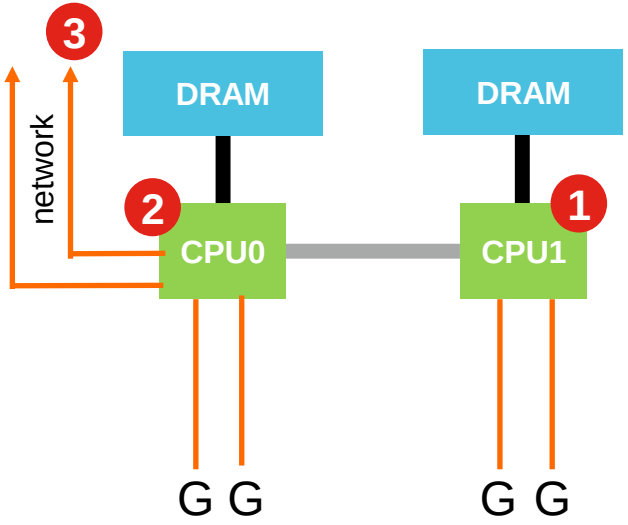


Designing a single 8 GPU server optimized for AI and HPC is not possible. 8 GPU server only for DL/ML using very large datasets

Requires PCIe Switches

2:1 GPU Server

Balanced and "Single Root"



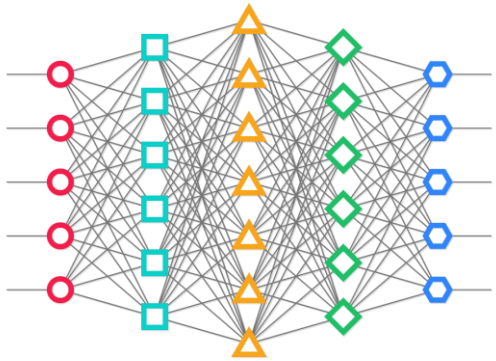
4 GPU systems are ideal for both HPC and AI. Minimal hops to network and from GPU-GPU

No PCIe Switch required

=Latency Hops from farthest GPU

Lenovo™

Artificial Intelligence



1

Discover
Experience AI

Worldwide AI Innovation Centers

2

Develop
Tools for fast start

HW&SW Platforms for app development

3

Deploy
Simplified Deployment

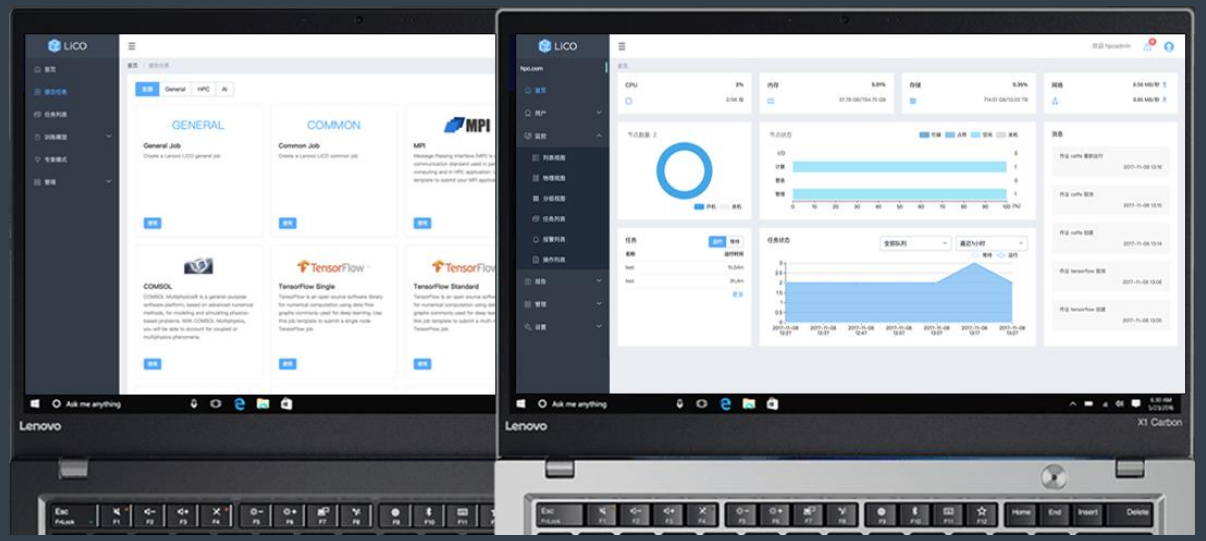
Solution guidance & Professional Services offerings

Lenovo™

Intelligent Computing Ochestration

Integrated SW suite
for HPC & AI

- Strong OpenHPC base
- Highly customizable
- Easy to use new design



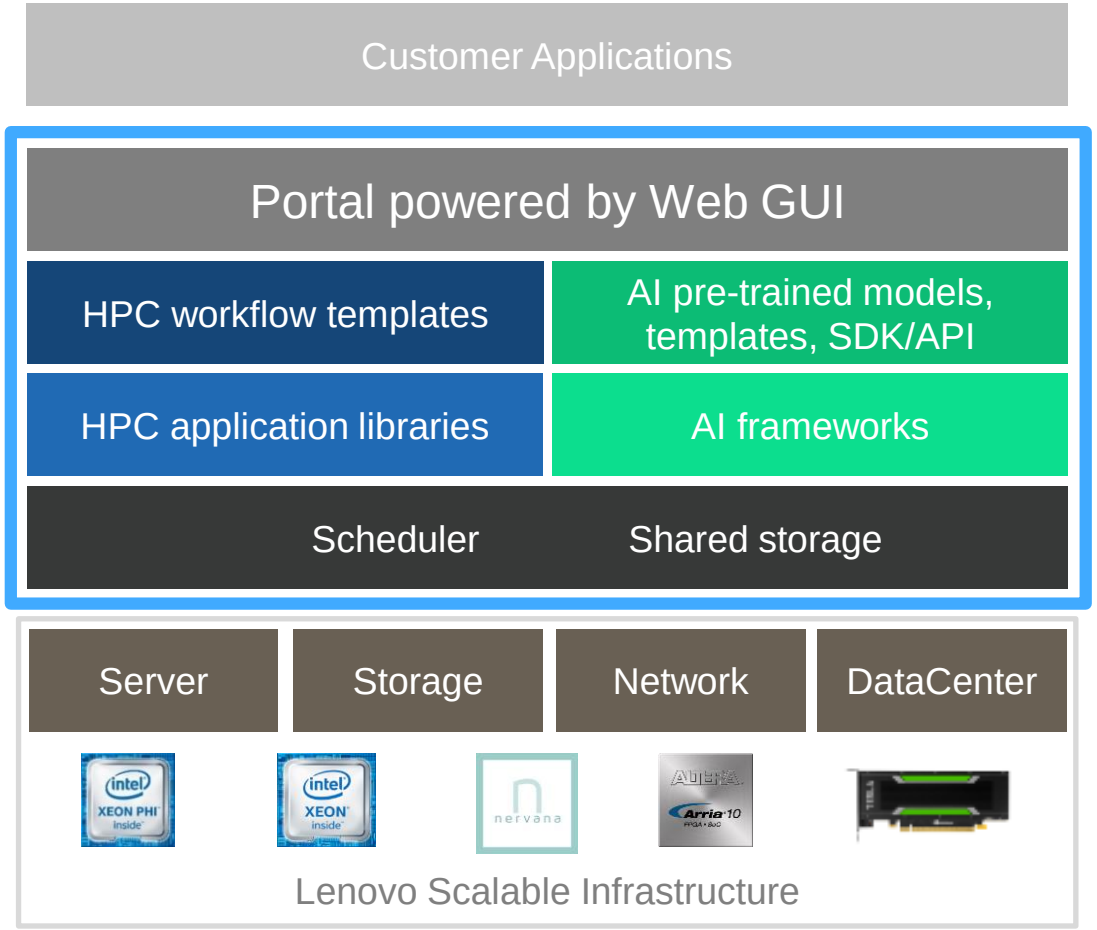
LiCO - Lenovo Intelligent Computing Orchestration

A single software stack to efficiently manage both HPC & AI workloads

For HPC

Simplify monitoring and job submission for HPC

- Easy-to-use interface for users to submit and manage jobs
- Full access to native tools in the stack for more technical users
- Built on an OpenHPC software base, with Lenovo value-add capabilities and optimizations



For AI

Easy access to train and optimize AI models

- Execute jobs, monitor training progress through a single GUI
- Out of the box scaling for both Intel and NVIDIA environments
- Easily try different frameworks, system types to determine best fit

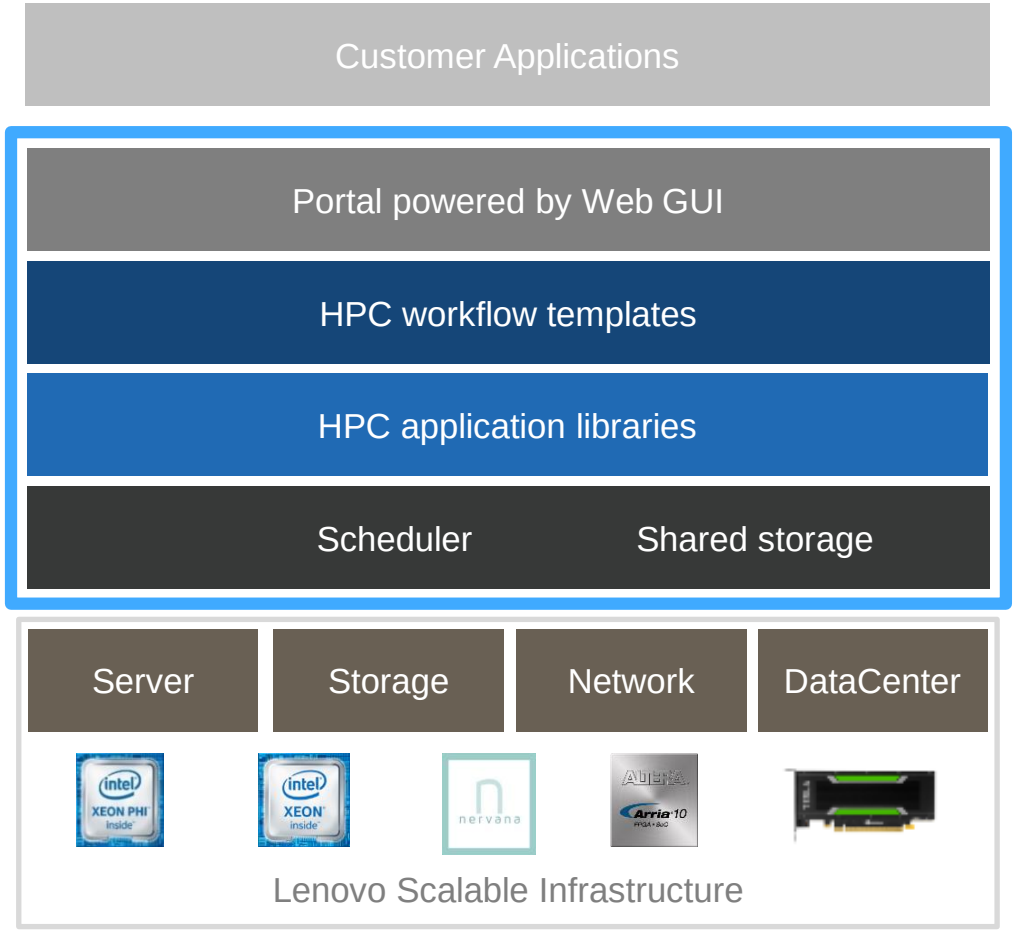
LiCO for High Performance Computing

Validated stack of open tools to simplify cluster environments

- Single GUI
  - consolidates functionality for both administrators and users
- Built on an OpenHPC foundation
  - Ganglia, Slurm/Munge, Torque/Maui, Singularity, MPI, xCAT, Confluent

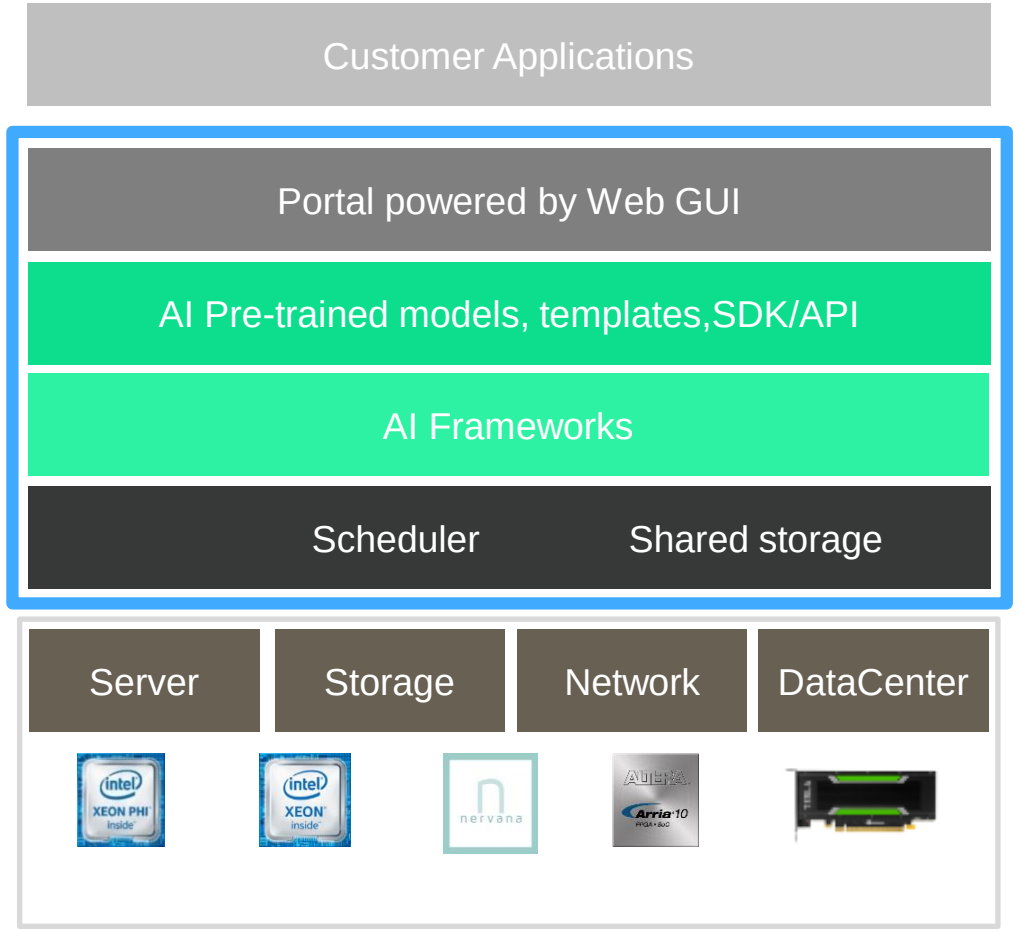
+

- Lenovo value-added capabilities
  - developed through client collaboration
  - Open web portal (Oxford & South Hampton)
  - Energy-Aware Runtime (BSC)



LiCO for Artificial Intelligence

Enables easy experimentation to optimize training models



- Intuitive web portal
  - manage workflow and application development for Data Scientists
- Pre-validated stack of open AI tools and libraries, ready to deploy
  - Caffe, TensorFlow, CUDA, NVIDIA cuDNN, Intel MKL-DNN, etc.
- Optimized for scalability, regardless of underlying hardware processors

AI Use Cases that Lenovo is exploring today

A more efficient and effective

Radiology Practice

CodaLab

Lenovo

For better quality control on

Manufacturing Lines

Mark III Systems

Lenovo

For proactive actions improving

Water & Energy Conservation

NC STATE UNIVERSITY

Lenovo

To advance particle physics at

CERN's Large Hadron Collider

UCL

Lenovo

AI: Better quality control for Manufacturing

Breakthrough innovation in Manufacturing with Lenovo AI

Current State of Manufacturing:

- Better Quality Control directly related to:
  - More yield at higher speed
  - Lower production costs by faster adjustments to process

Where AI can help:

- Leverage cameras and sensors through-out the production lifecycle
- Better manage quality through product age of customizations

AI Inference and Training:

- Image / Pattern recognition and analysis
- ML training of products at different stages of production

See demo in action at SC17, Booth 1353

AI: To advance particle physics in the Large Hadron Collider

Lenovo AI progressing advanced Particle Physics

Current State and Challenges :

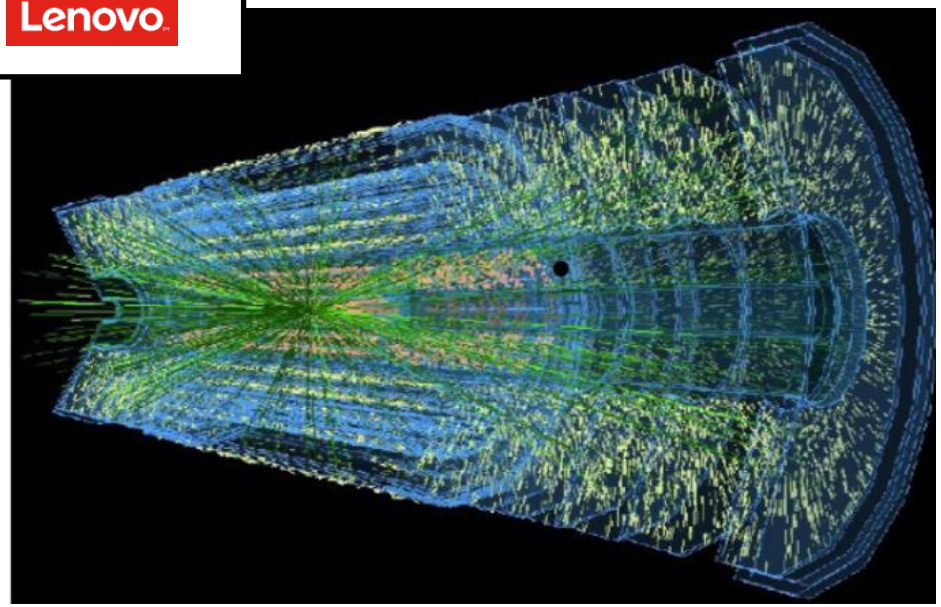
- Reconstructing particle trajectories from thousands of sensor measurements in the detector is an important data analytics task
- Traditional computational methods demand high amounts of resources and don't scale well as the LHC pushes to higher collision frequencies

Where AI can help:

- Pattern recognition: Reconstructing particle trajectories using imaging data from the collider much more efficiently than traditional methods

AI Inference and Training:

- Machine learning methods using binary image data from experiments combined with integral transforms for pattern recognition



Different is better

Lenovo™