Converging HPC and Big Data / AI Infrastructures at Scale on TSUBAME3 and ABCI Supercomputers

Satoshi Matsuoka

Professor, GSIC, Tokyo Institute of Technology /
Director, AIST-Tokyo Tech. Big Data Open Innovation Lab /
Fellow, Artificial Intelligence Research Center, AIST, Japan /
Vis. Researcher, Advanced Institute for Computational Science, Riken

NVIDIA Singapore AI 2017
2017/10/24
Singapore

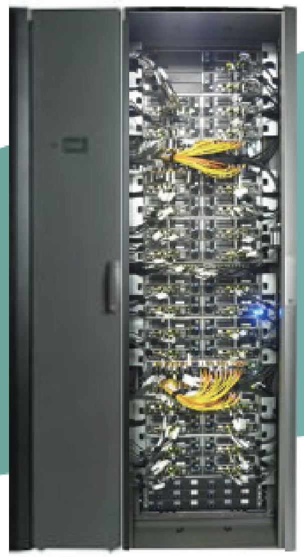
TSUBAME2.0 Nov. 1, 2010
"The Greenest Production Supercomputer in the World"

- GPU-centric (> 4000) high performance & low power
- Small footprint (~200m2 or 2000 sq.ft), low TCO
- High bandwidth memory, optical network, SSD storage…

System
(42 Racks)
1408 GPU Compute Nodes,
34 Nehalem "Fat Memory" Nodes

TSUBAME 2.0
New Development

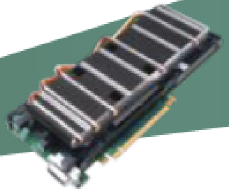
Rack
(8 Node Chassis)



Node Chassis
(4 Compute Nodes)

Compute Node
(2 CPUs,3 GPUs)

Chip
(CPU ,GPU)



CPU(Westmere EP)
76.8 GFLOPS
32nm

GPUs(Tesla M2050)
515 GFLOPS
3 GB
40nm

1.6 TFLOPS
55 GB/103 GB
>400GB/s Mem BW
80Gbps NW BW
~1KW max

6.7 TFLOPS
220 GB/412 GB
>1.6TB/s Mem BW

53.6 TFLOPS
1.7 TB/3.2 TB
>12TB/s Mem BW
35KW Max

2.4 PFLOPS
80 TB
4224 GPUs
>600TB/s Mem BW
220Tbps NW
Bisecion BW
1.4MW Max

Integrated by NEC Corporation

TSUBAME-KFC/DL: TSUBAME3 Prototype [ICPADS2014]

Oil Immersive Cooling＋ Hot Water Cooling + High Density Packaging + Fine-Grained Power Monitoring and Control, *upgrade to /DL Oct. 2015*

Single Rack High Density Oil Immersion
168 NVIDIA K80 GPUs + Xeon
413+TFlops (DFP)
1.5PFlops (SFP)
~60KW/rack

High Temperature Cooling
Oil Loop 35~45℃
⇒ Water Loop 25~35℃
(c.f. TSUBAME2: 7~17℃)

Cooling Tower：
Water 25~35℃
⇒ To Ambient Air

2013年11月/2014年6
Word #1 Green500

Container Facility
20 feet container (16m²)
Fully Unmanned Operation

2017 Q2 TSUBAME3.0 Leading Machine Towards Exa & Big Data

1. "Everybody's Supercomputer" - High Performance (12~24 DP Petaflops, 125~325TB/s Mem, 55~185Tbit/s NW), innovative high cost/performance packaging & design, in mere 180m^2...
2. "Extreme Green" – ~10GFlops/W power-efficient architecture, system-wide power control, advanced cooling, future energy reservoir load leveling & energy recovery
3. "Big Data Convergence" – BYTES-Centric Architecture, Extreme high BW & capacity, deep memory hierarchy, extreme I/O acceleration, Big Data SW Stack for machine learning, graph processing, …
4. "Cloud SC" – dynamic deployment, container-based node co-location & dynamic configuration, resource elasticity, assimilation of public clouds…
5. "Transparency" - full monitoring & user visibility of machine & job state, accountability via reproducibility

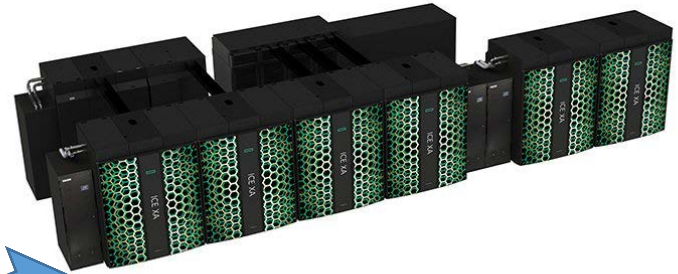
2006 TSUBAME1.0
80 Teraflops, #1 Asia #7 World
"Everybody's Supercomputer"

2010 TSUBAME2.0
2.4 Petaflops #4 World
"Greenest Production SC"

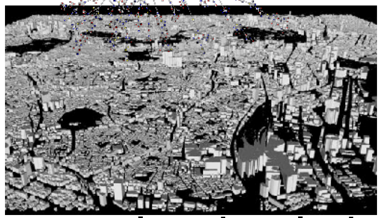
2011 ACM Gordon Bell Prize

2013
TSUBAME2.5
upgrade
5.7PF DFP
/17.1PF SFP
20% power
reduction

2013 TSUBAME-KFC
#1 Green 500



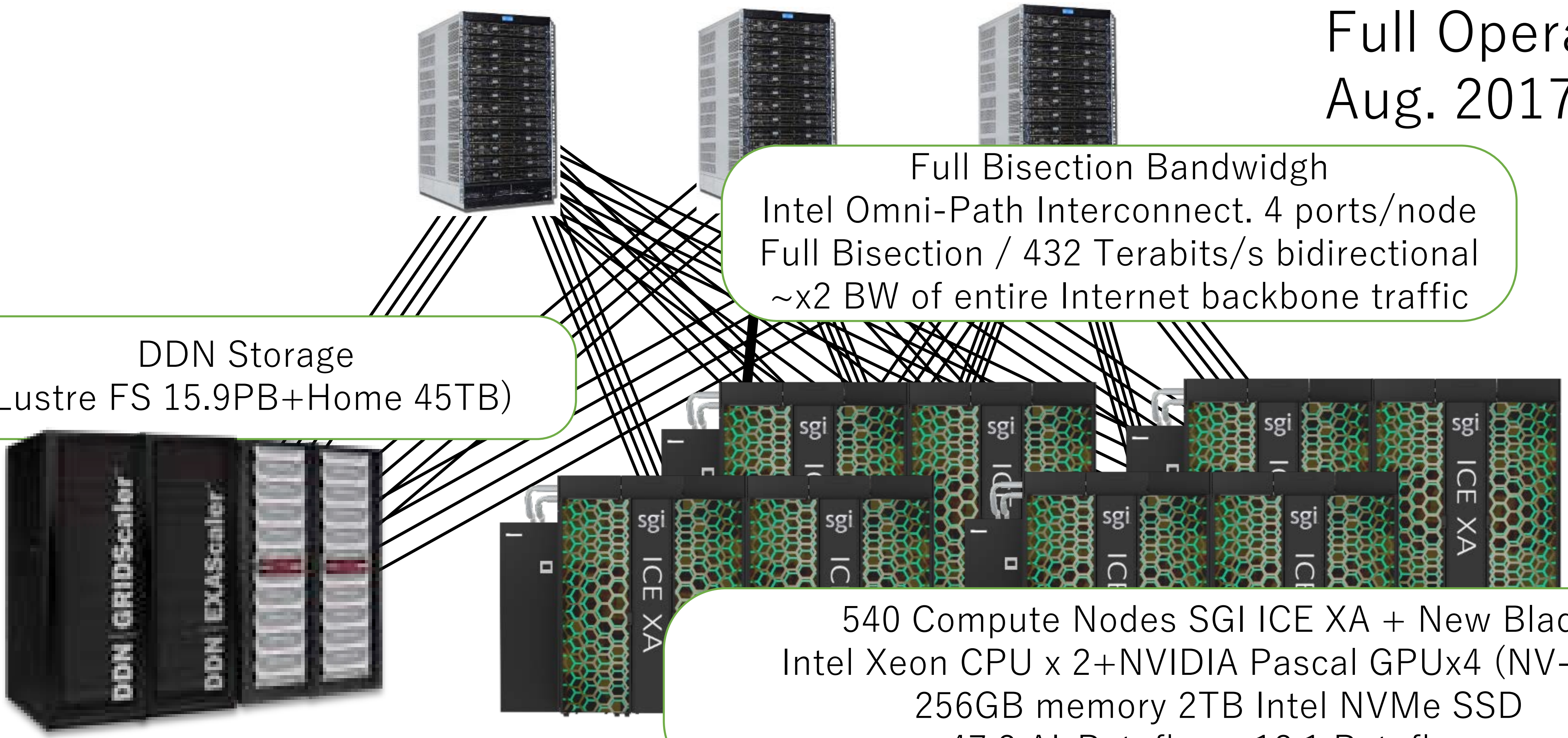
2017 TSUBAME3.0+2.5
~18PF(DFP) 4~5PB/s Mem BW
10GFlops/W power efficiency
Big Data & Cloud Convergence



Large Scale Simulation
Big Data Analytics
Industrial Apps

Overview of TSUBAME3.0
BYTES-centric Architecture, Scalaibility to all 2160 GPUs, all nodes, the entire memory hiearchy

Full Operations
Aug. 2017

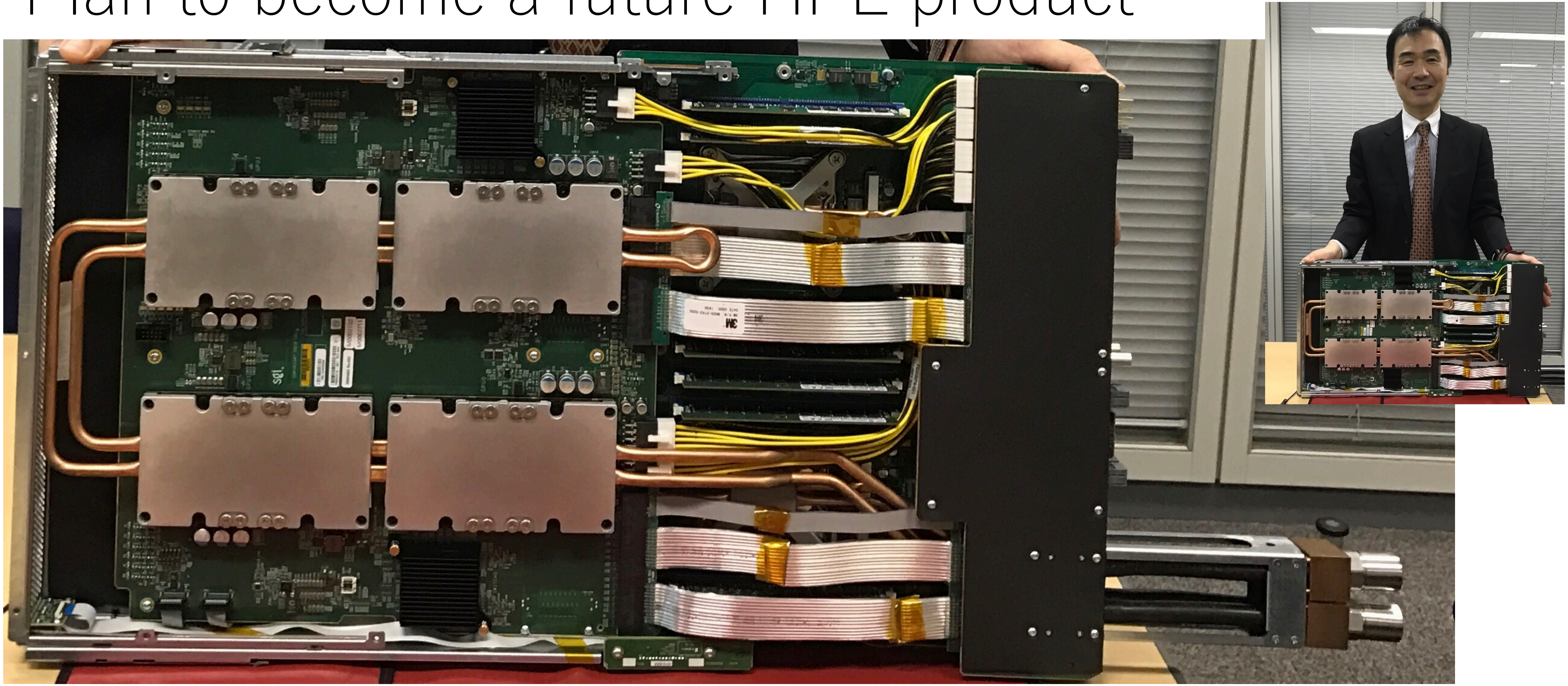


Full Bisection Bandwidgh
Intel Omni-Path Interconnect. 4 ports/node
Full Bisection / 432 Terabits/s bidirectional
~x2 BW of entire Internet backbone traffic

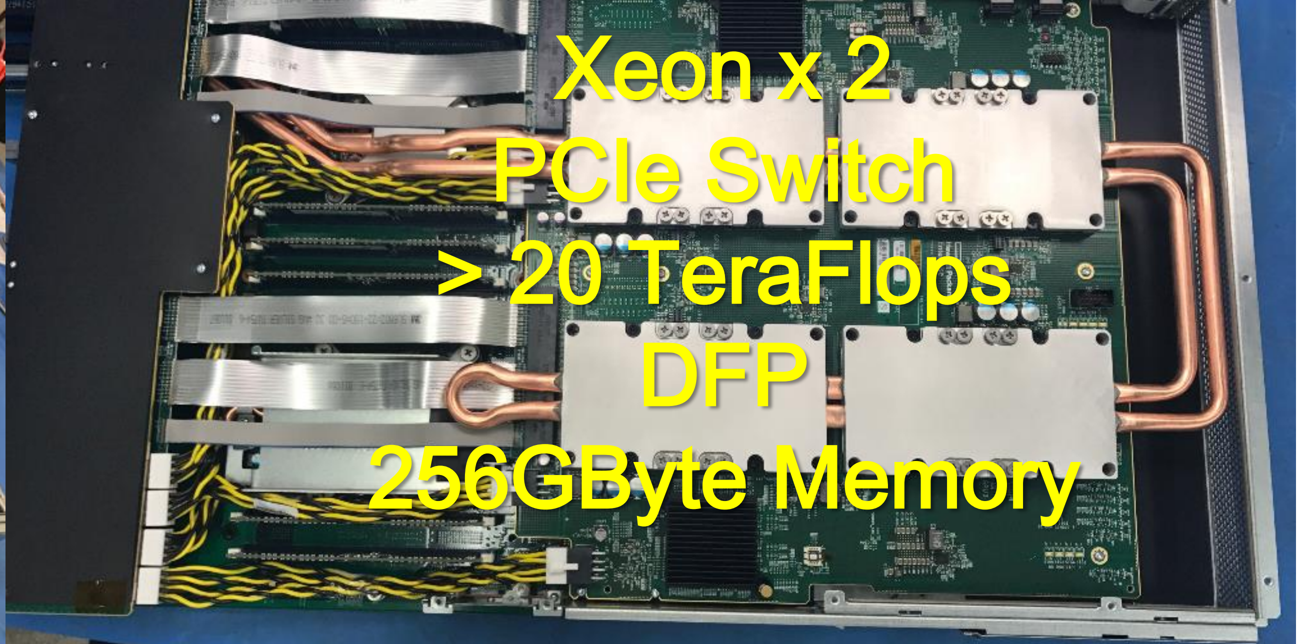
DDN Storage
(Lustre FS 15.9PB+Home 45TB)

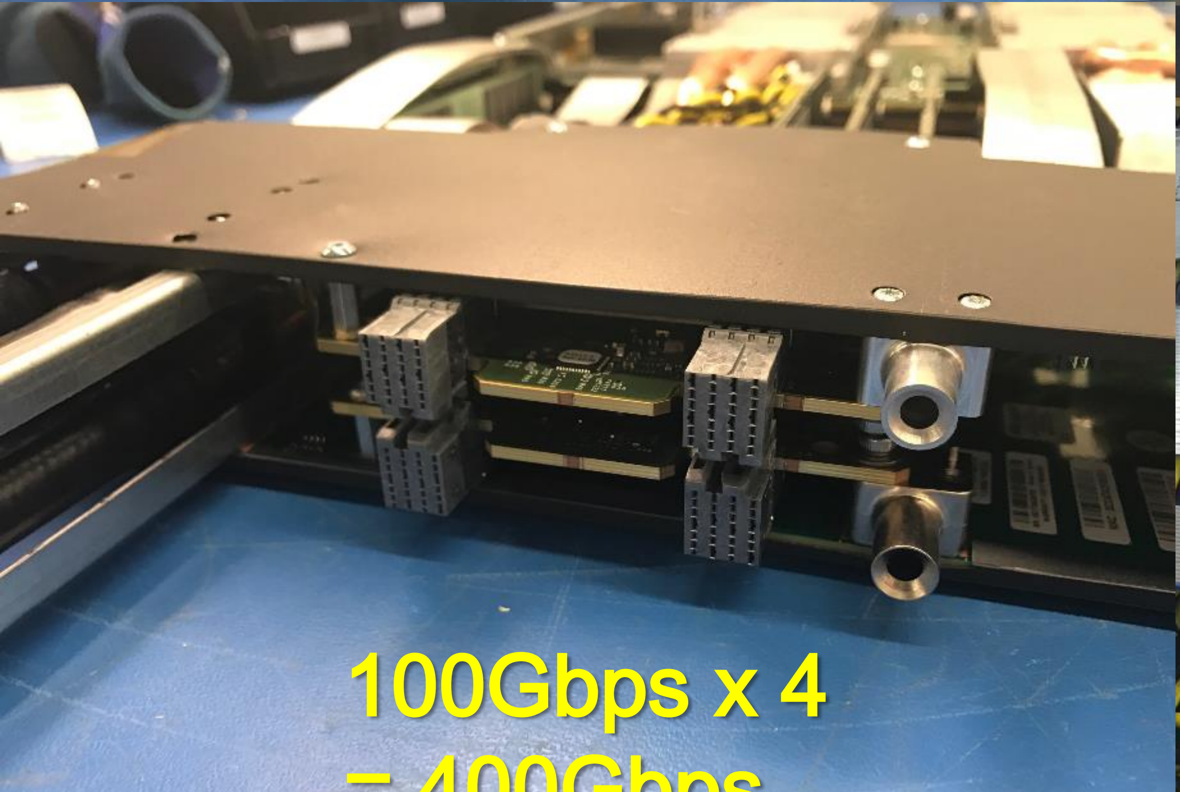
540 Compute Nodes SGI ICE XA + New Blade
Intel Xeon CPU x 2+NVIDIA Pascal GPUx4 (NV-Link)
256GB memory 2TB Intel NVMe SSD
47.2 AI-Petaflops, 12.1 Petaflops

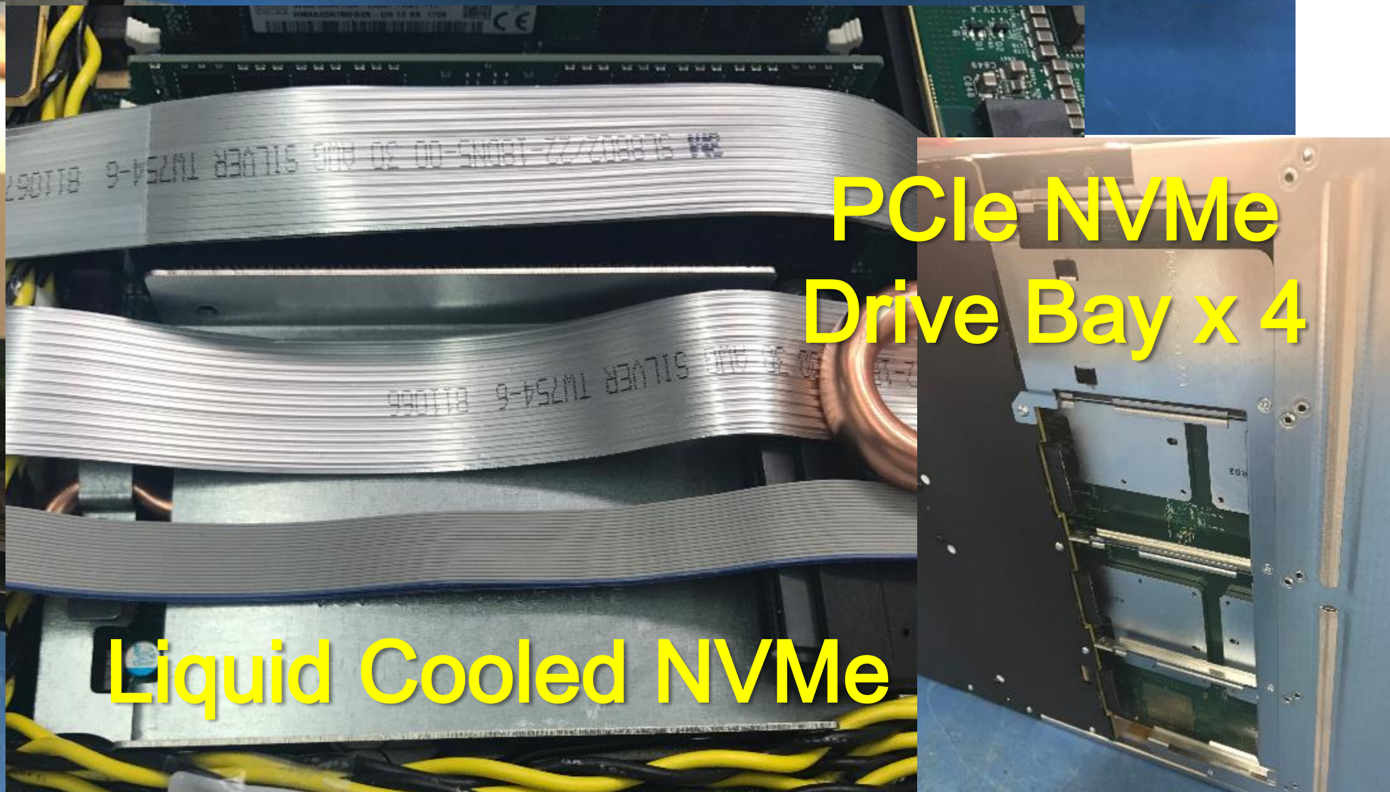
TSUBAME3.0 Co-Designed SGI ICE-XA Blade (new)
- No exterior cable mess (power, NW, water)
- Plan to become a future HPE product

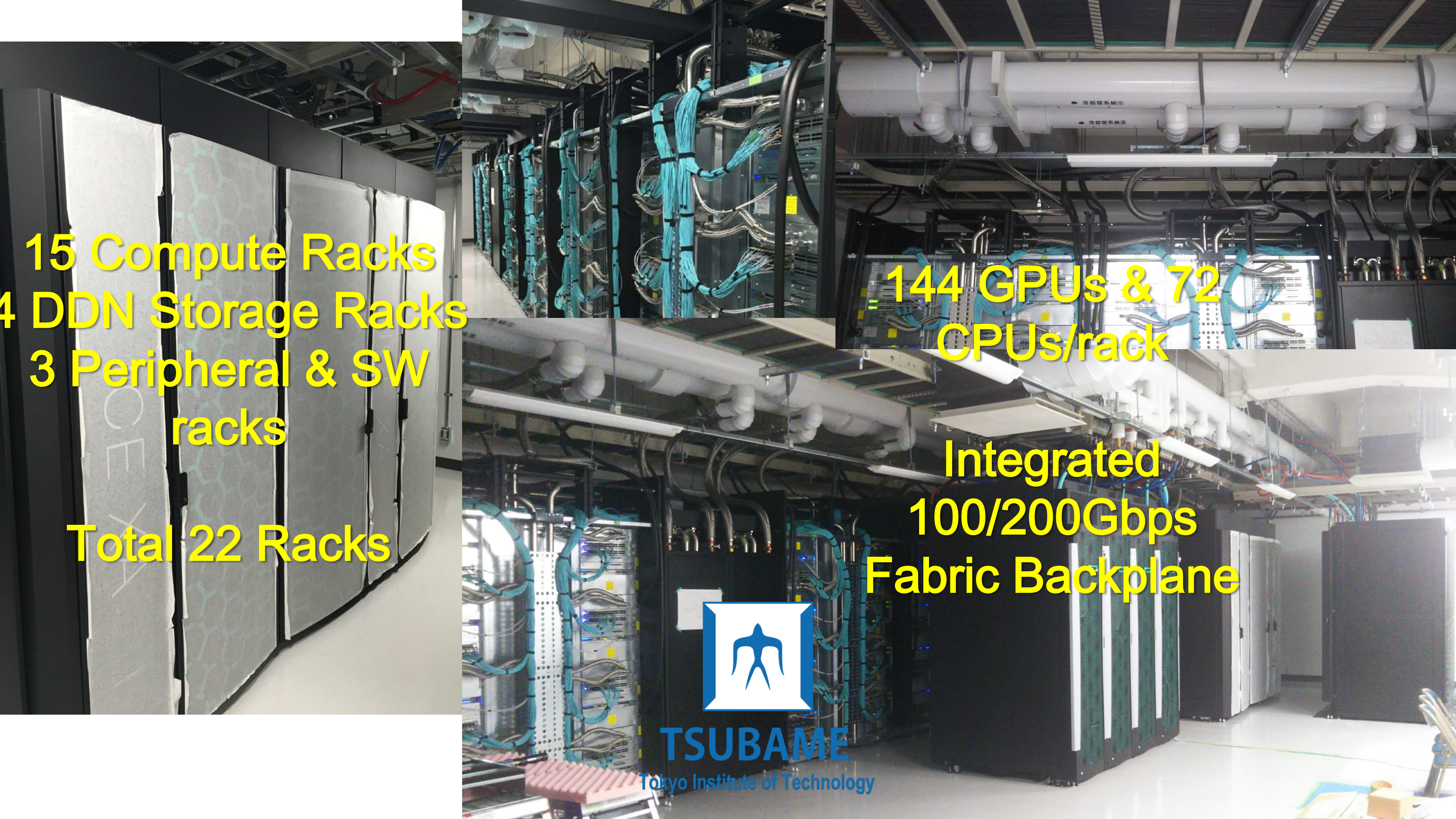



Liquid Cooled
"Hot Pluggable" ICE-XA Blade
Smaller than 1U server, no cables or pipes


Xeon x 2
PCIe Switch
> 20 TeraFlops
DFP
256GByte Memory


100Gbps x 4
= 400Gbps


PCIe NVMe
Drive Bay x 4
Liquid Cooled NVMe


15 Compute Racks
4 DDN Storage Racks
3 Peripheral & SW racks
Total 22 Racks
144 GPUs & 72 CPUs/rack
Integrated 100/200Gbps Fabric Backplane
TSUBAME
Tokyo Institute of Technology

Tsubame3 Highly Efficient Datacenter Machine PUE ~= 1.03 (~1.1 w/storage)

Over 100t total floor load

Reinforced “Slab-Like” flat floor surface.
Over 1t/m² floor load

Ultra High Density

Space Efficient

Power, water, and cabling are all above with ceiling support, for space efficiency and freedom of layout
IDC space ~= 130 m2

Piping and cabling hang from the ceiling

Max 32 degrees Celsius water

Year-round free “warm-water” cooling with cooling tower, PUE = 1.03, machine power ~= facility power

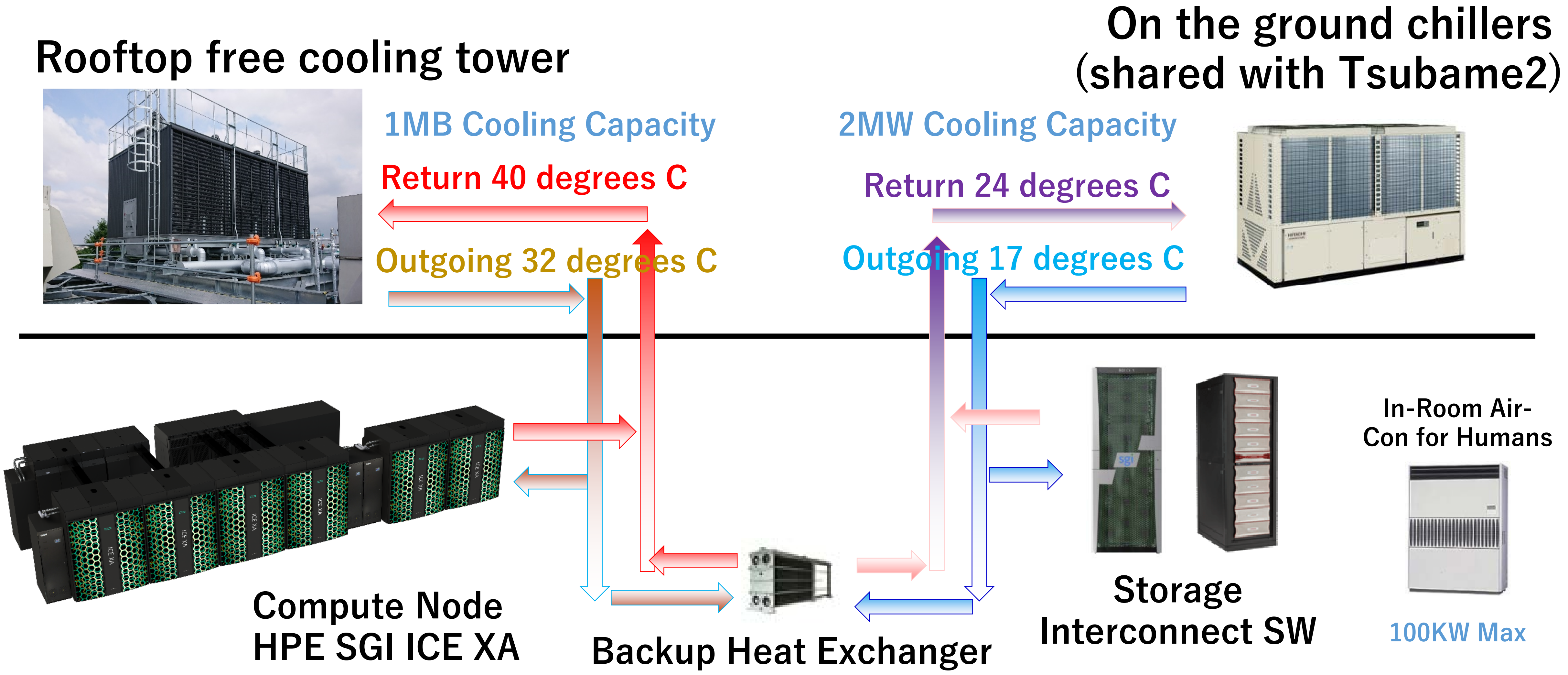
Efficient Warm-Water Cooling

Low Electrical Distribution Loss

420V High Voltage to minimize electrical distribution loss and cheaper cabling

420V Tri-Phase AC Power

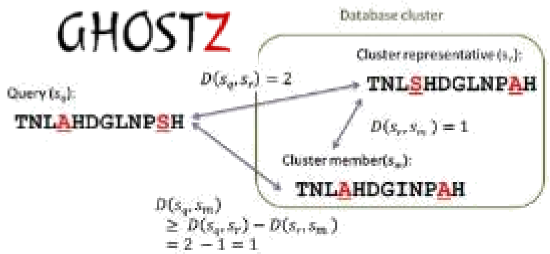
Warm Water Cooling Distribution in T3



HPC and BD/AI Convergence Example [Yutaka Akiyama, Tokyo Tech]

Genomics

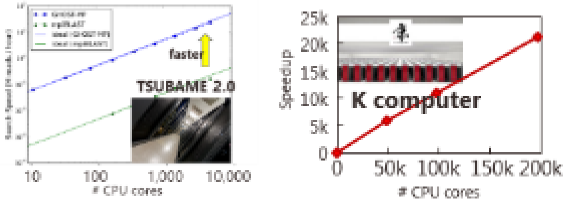
Ultra-fast Seq. Analysis



$D(s_q, s_r) = 2$

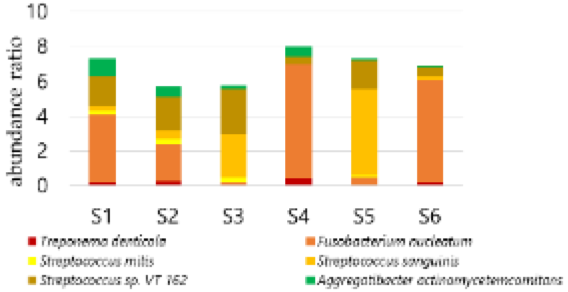
$D(s_r, s_m) = 1$

$D(s_q, s_m) \geq D(s_q, s_r) - D(s_r, s_m) = 2 - 1 = 1$



- Suzuki *et al., Bioinformatics* (2015)
- Suzuki *et al., PLOS ONE* (2016)

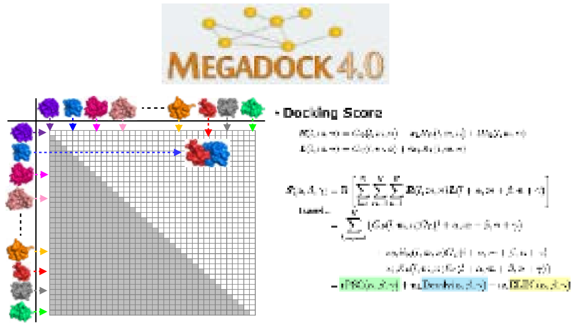
Oral/Gut Metagenomics



- Yamasawa *et al., IIBMP* (2016)

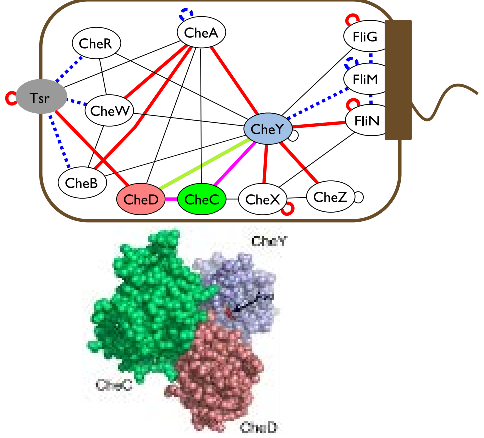
Protein-Protein Interactions

Exhaustive PPI Prediction System



- Ohue *et al., Bioinformatics* (2014)

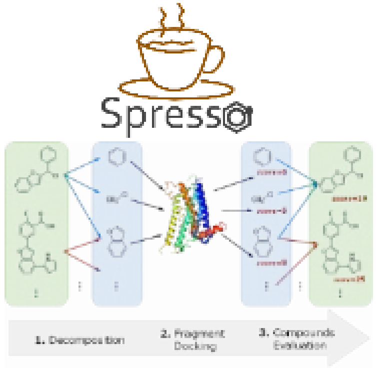
Pathway Predictions



- Matsuzaki *et al., Protein Pept Lett* (2014)

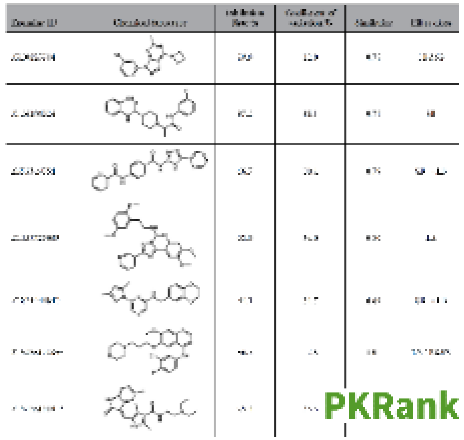
Drug Discovery

Fragment-based Virtual Screening



- Yanagisawa *et al., GIW* (2016)

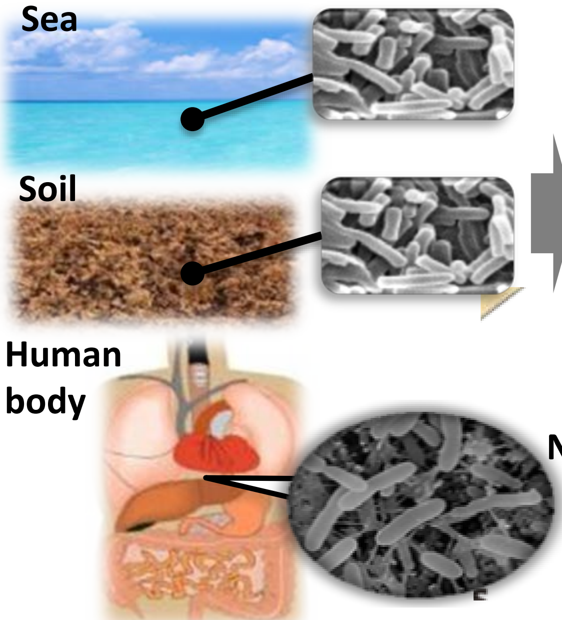
Learning-to-Rank VS



- Suzuki *et al., AROB2017* (2017)

EBD vs. EBD : Large Scale Homology Search for Metagenomics

- Revealing uncultured microbiomes and finding novel genes in various environments
- Applied for human health in recent years



Various environments

...CCTTATCTTCG...

...CCACATAAACT...

...ATGGTCGATGTT...

Next generation sequencer

increasing

EBD

O(n)

Meas. data

increasing

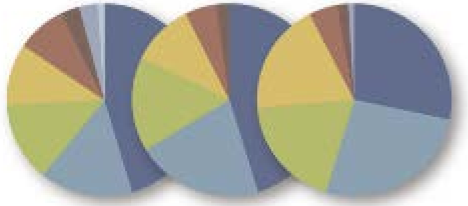
O(m) *Reference Database*

EBD

O(m n) calculation

Correlation, Similarity search

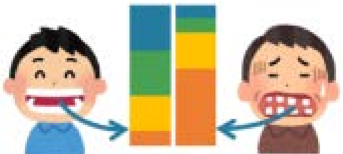
Metabolic Pathway



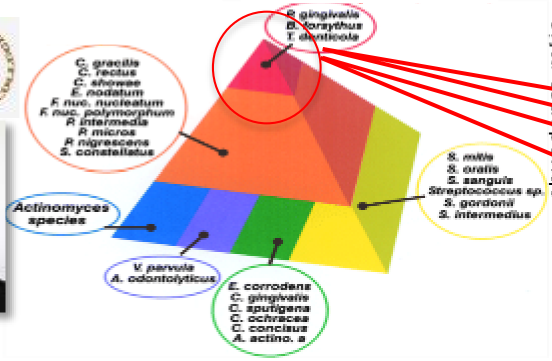
Taxonomic composition

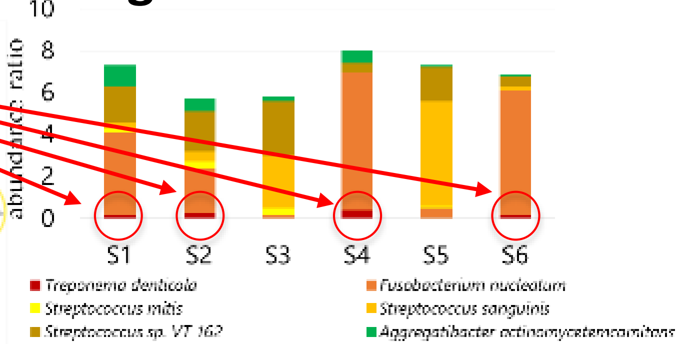
Metagenomic analysis of periodontitis patients

- with Tokyo Dental College, Prof. Kazuyuki Ishihara
- Comparative metagenomic analysis bewtween healthy persons and patients



High risk microorganisms are detected.





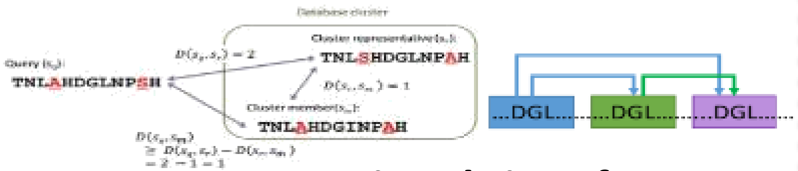
Development of Ultra-fast Homology Search Tools

x100,000 ~ x1,000,000 c.f. high-end BLAST WS (both FLOPS and BYTES)

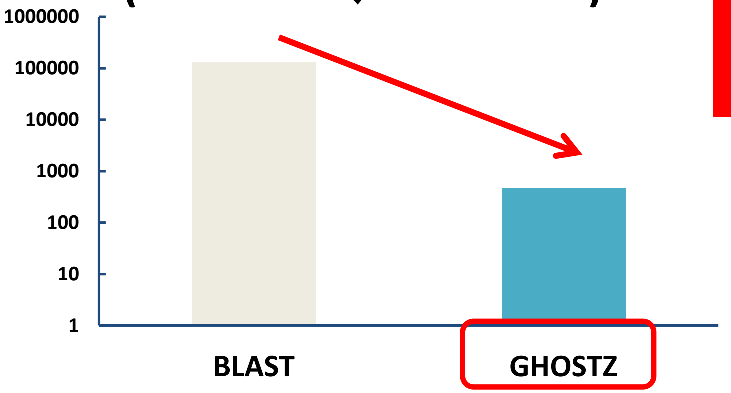
GHOSTZ

Suzuki, *et al*. *Bioinformatics*, 2015.

Subsequence sequence clustering



computational time for 10,000 sequences (sec.) (3.9 GB DB、1CPU core)

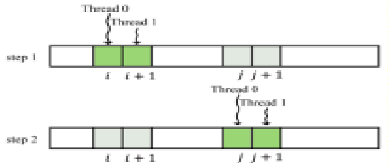


×240 faster than conventional algorithm

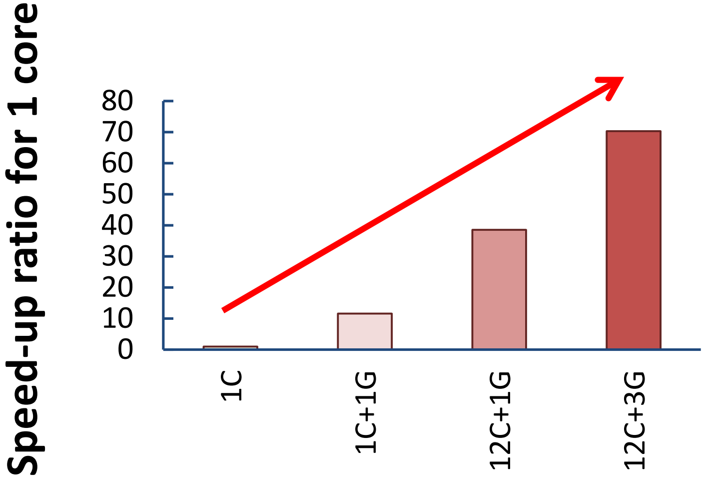
GHOSTZ-GPU

Suzuki, *et al*. *PLOS ONE*, 2016.

Multithread on GPU



TSUBAME 2.5 Thin node GPU



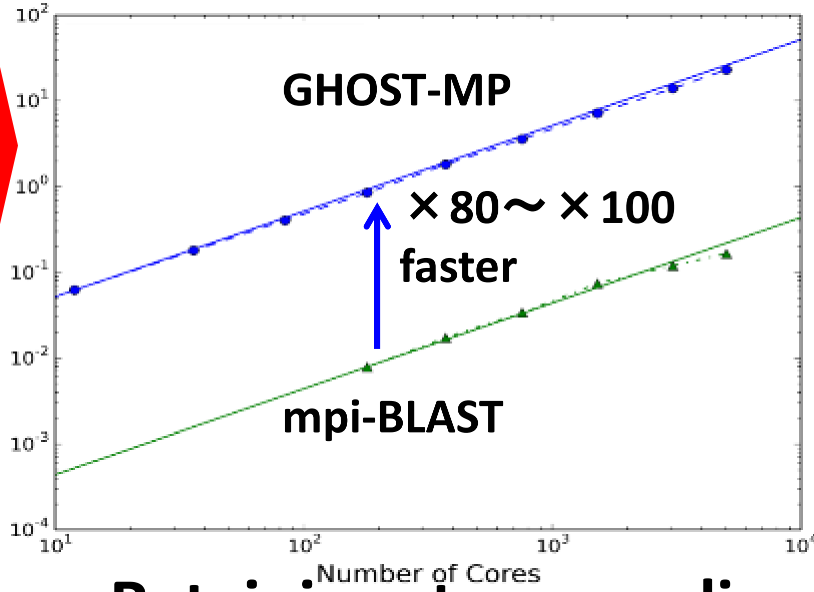
×70 faster than 1 core using 12 cores + 3 GPUs

GHOST-MP

Kakuta, *et al*. (submitted)

MPI + OpenMP hybrid pallelization

TSUBAME 2.5



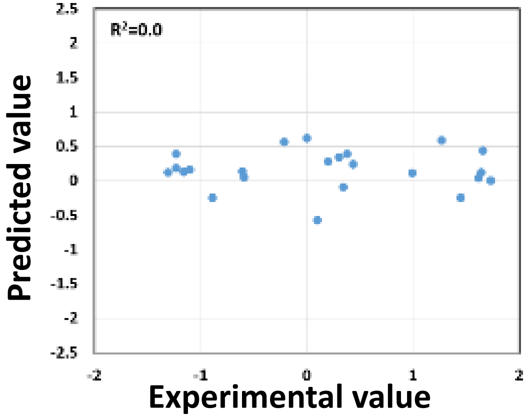
Retaining strong scaling up to 100,000 cores

Plasma Protein Binding (PPB) Prediction by Machine Learning
Application for peptide drug discovery

Problems

	Small molecule drug	Peptide drug	Antibody drug
Molecular weight	~1,000	600~2,500	150,000~
Number of targets	◎	○	△
Target specificity	△	○	◎
PPI inhibition	×	○	○
Bio-stability	○	△	◎

- Candidate peptides are tend to be degraded and excreted faster than small molecule drugs
- Strong needs to design bio-stable peptides for drug candidates

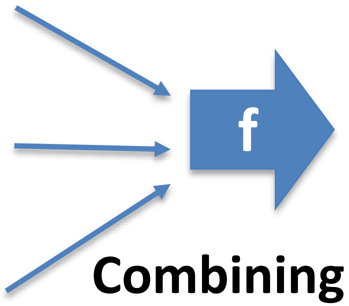


Previous PPB prediction software for small molecule can not predict peptide PPB

Solutions

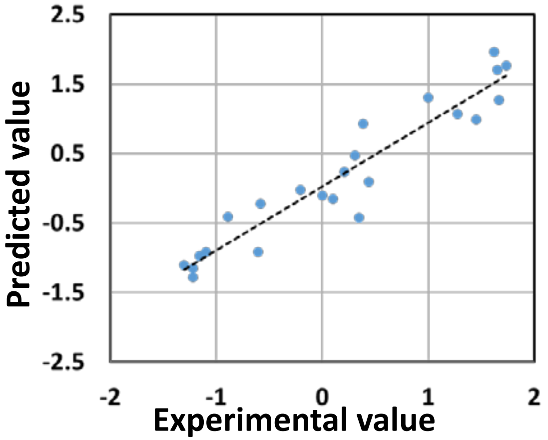
Compute Feature Values
(more than 500 features)

LogS
LogP
:
MolWeight
:
SASA
polarity



PPB value

Combining feature values for building a predictive model



R^2 = 0.905

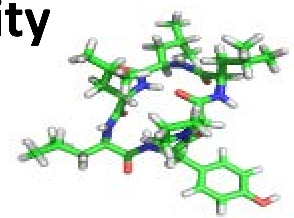
A constructed model can explain peptide PPB well

Molecular Dynamics Simulation for Membrane Permeability
Application for peptide drug discovery

Problems

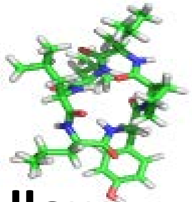
1) Single residue mutation can drastically change membrane permeability

Sequence : D-Pro, D-Leu, D-Leu, L-Leu, D-Leu,
Membrane permeability : 7.9 × 10^{-6} cm/s

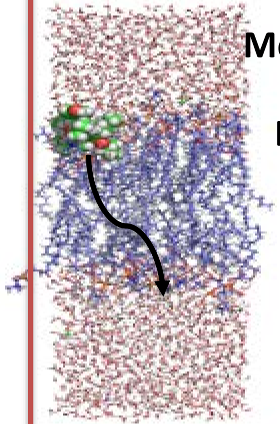


× 0.006

Sequence : D-Pro, D-Leu, D-Leu, D-Leu, D-Leu, L-Tyr
Membrane permeability : 0.045 × 10^{-6} cm/s



2) Standard MD simulation can not follow membrane permeation.



Membrane permeation is millisecond order phenomenon.

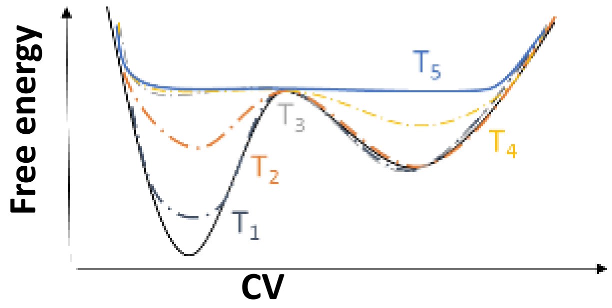
Ex) Membrane thickness : 40 Å
Peptide membrane permeability : 7.9 × 10^{-6} cm/s

Typical peptide membrane permeation takes
40 Å / 7.9 × 10^{-6} cm/s = 0.5 millisecond

Solutions

1) Apply enhanced sampling

Metadynamics (MTD)



Supervised MD (SuMD)

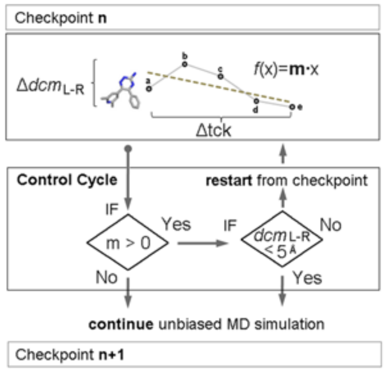
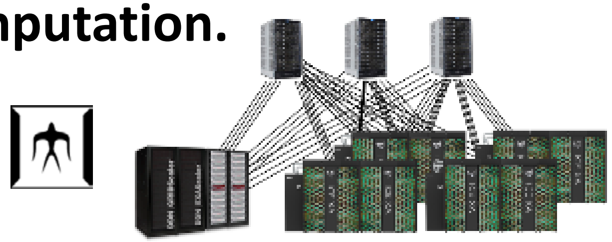


Figure 1. Scheme of the ligand−receptor distance vector ($dcm_{L\text{-}R}$) supervision algorithm implemented in the supervised molecular dynamics (SuMD) technique.

2) GPU acceleration and massively parallel computation.



MD engine
GROMACS
DESMOND
} on GPU

- Millisecond order phenomenon can be simulated.
- Hundreds of peptides can be calculated simultaneously on TSUBAME.

RWBC-OIL 2-3: Tokyo Tech IT-Drug Discovery Factory Simulation & Big Data & AI at Top HPC Scale

(Tonomachi, Kawasaki-city: planned 2017, PI Yutaka Akiyama)

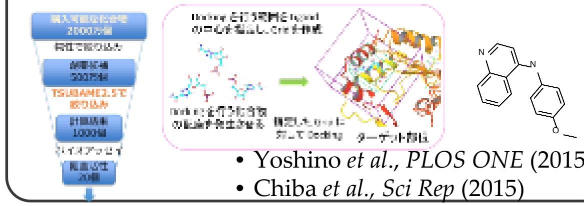
Tokyo Tech's research seeds

①Drug Target selection system

Minister of Health, Labour and Welfare Award of the 11th annual Merit Awards for Industry-Academia-Government Collaboration

②Glide-based Virtual Screening

TSUBAME's GPU-environment allows
World's top-tier Virtual Screening



- Yoshino *et al.*, *PLOS ONE* (2015)
- Chiba *et al.*, *Sci Rep* (2015)

③Novel Algorithms for fast virtual screening against huge databases

Fragment-based efficient algorithm designed for **100-millions cmpds data**

- Yanagisawa *et al.*, *GIW* (2016)

Drug Discovery platform powered by Supercomputing and Machine Learning

Application projects

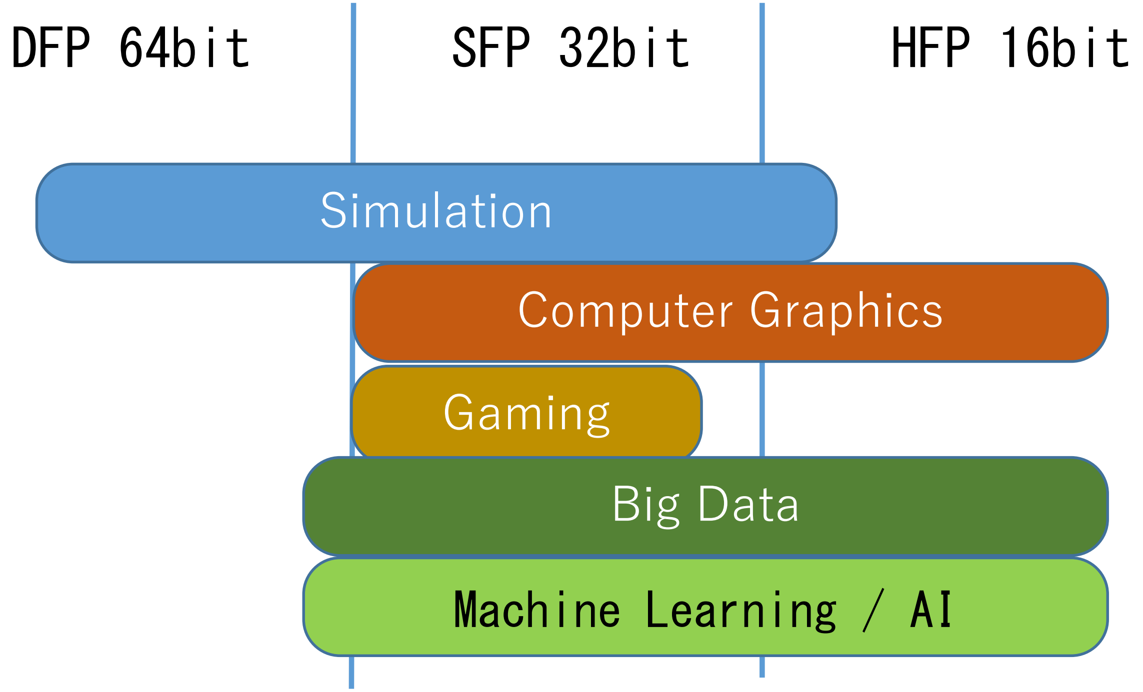
New Drug Discovery platform especially for specialty peptide and nucl. acids.

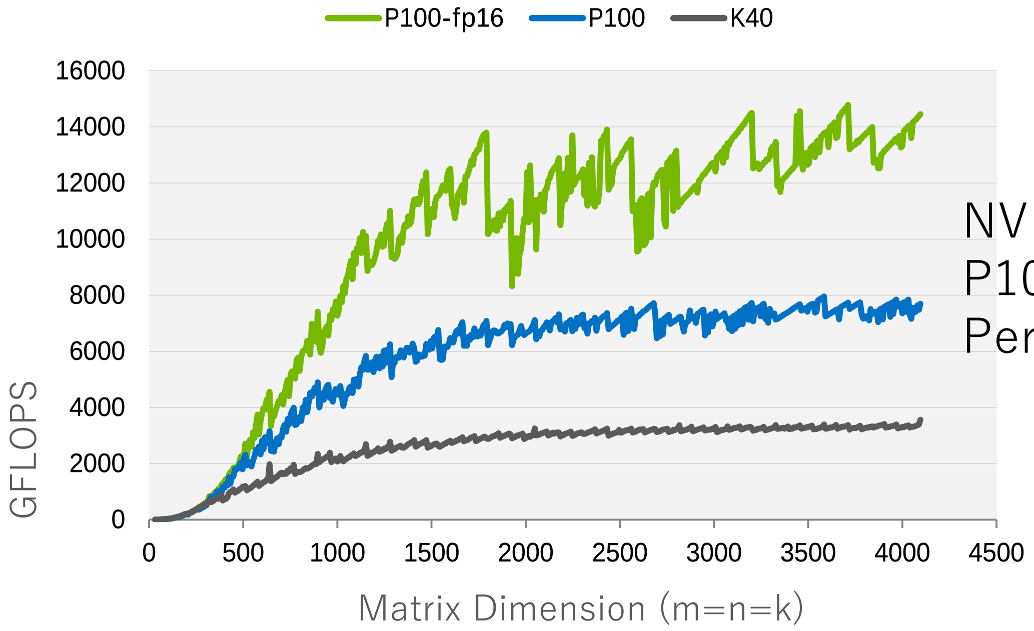
Multi-Petaflops Compute
Peta~Exabytes Data
Processing Continuously

Cutting Edge, Large-Scale HPC & BD/AI Infrastructure Absolutely Necessary

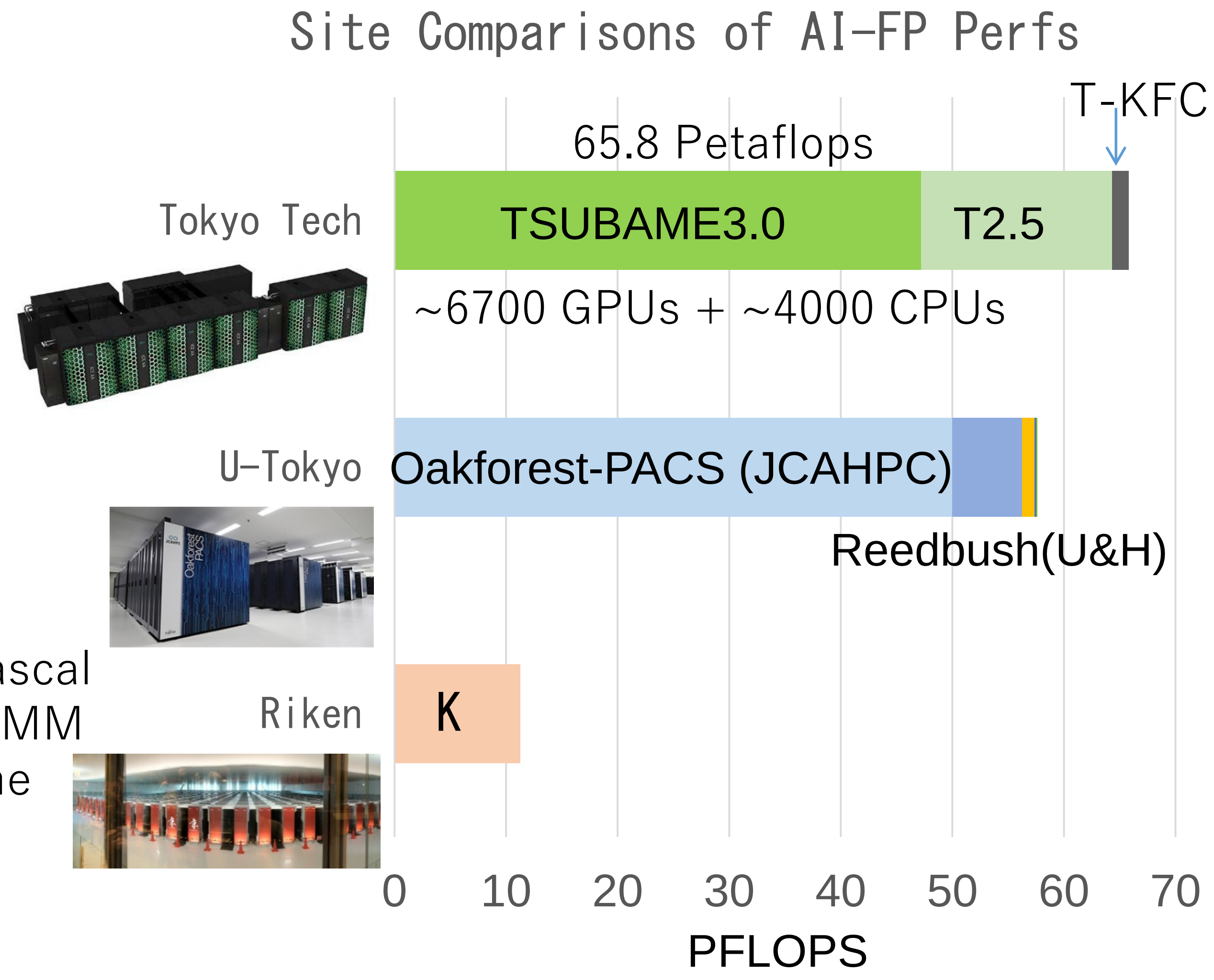
Investments from JP Govt., Tokyo Tech. (TSUBAME SC) Munincipłe Govt (Kawasaki), JP & US Pharma

Tokyo Tech GSIC leads Japan in aggregated AI-capable FLOPS TSUBAME3+2.5+KFC, in all Supercomuters and CloudsNV





NVIDIA Pascal P100 DGEMM Performane



Tremendous Recent Rise in Interest by the Japanese Government on Big Data, DL, AI, and IoT

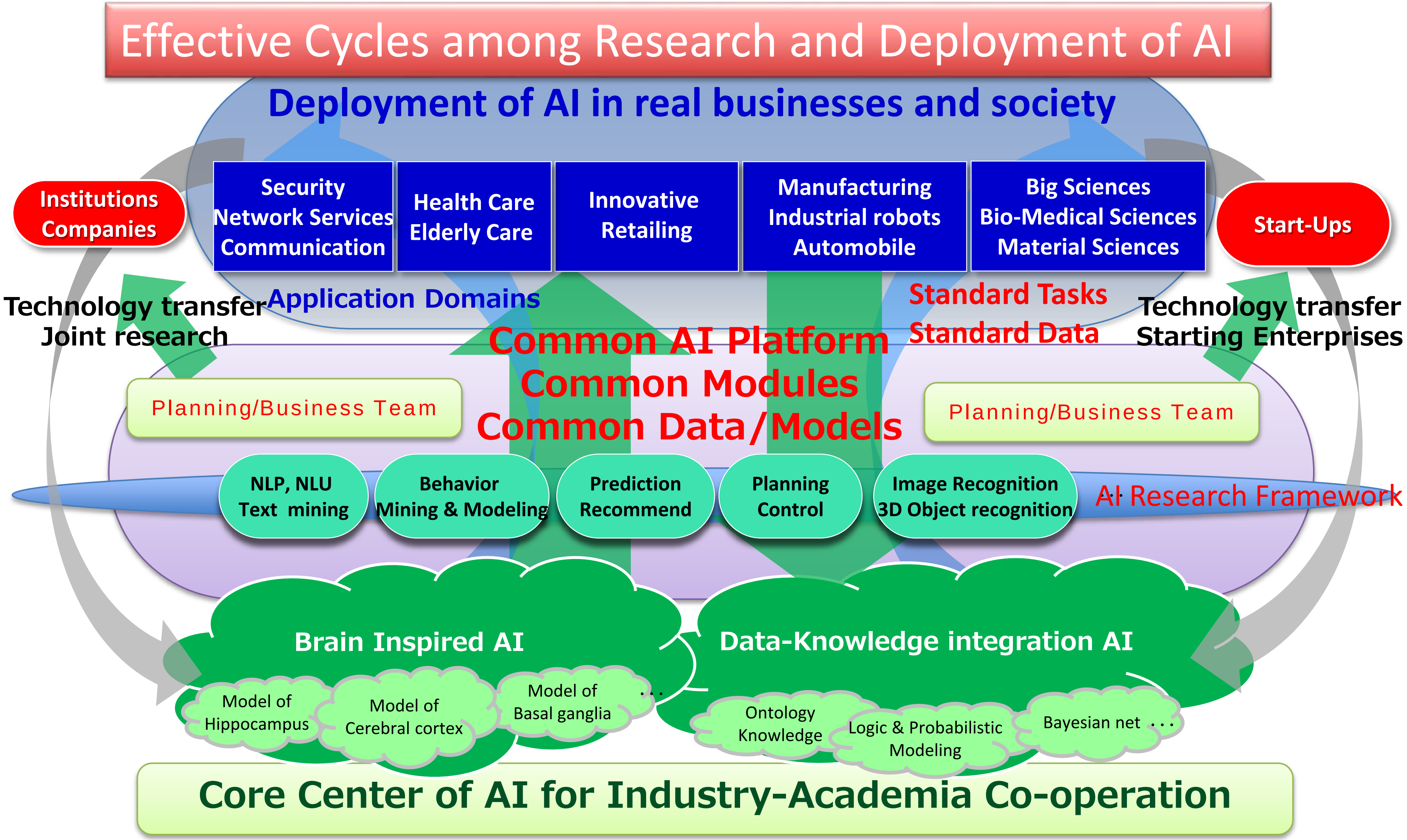
- Three national centers on Big Data and AI launched by three competing Ministries for FY 2016 (Apr 2015-)
  - METI – AIRC (Artificial Intelligence Research Center): AIST (AIST internal budget + > $200 million FY 2017), April 2015
    - Broad AI/BD/IoT, industry focus
  - MEXT – AIP (Artificial Intelligence Platform): Riken and other institutions ($~50 mil), April 2016
    - A separate Post-K related AI funding as well.
    - Narrowly focused on DNN
  - MOST – Universal Communication Lab: NICT ($50~55 mil)
    - Brain –related AI
  - $1 billion commitment on inter-ministry AI research over 10 years => Supplanting HPC activities?

Vice Minsiter Tsuchiya@MEXT Annoucing AIP estabishment

2015- AI Research Center (AIRC), AIST

Now > 400+ FTEs

Director:
Jun-ichi Tsujii



Matsuoka : Joint appointment as "Designated" Fellow since July 2017

Joint Lab established Feb. 2017 to pursue BD/AI joint research using large-scale HPC BD/AI infrastructure

National Institute of Advanced Industrial Science and Technology AIST
独立行政法人
産業技術総合研究所

National Institute for Advanced Industrial Science and Technology (AIST)

経済産業省
Ministry of Economy, Trade and Industry

Ministry of Economics Trade and Industry (METI)

AIST Artificial Intelligence Research Center (AIRC)

AIRC

Application Area
Natural Langauge Processing
Robotics
Security

ABCI
AI Bridging Cloud Infrastructure

Joint Research on AI / Big Data and applications

REAL WORLD BIGDATA
RWBC-OIL

AIST-Tokyo Tech
Real World Big-Data Computation
Open Innovation Laboratory
(RWBC-OIL)

Director: Satoshi Matsuoka

Industrial Collaboration in data, applications

Industry

YAHOO! JAPAN ITLAB DENSO ●●

DENSO IT LABORATORY, INC.

Tokyo Institute of Technology / GSIC

TSUBAME
Tokyo Institute of Technology

GSIC
Global Scientific Information and Computing Center

Resources and Acceleration of AI / Big Data, systems research

Tsubame 3.0/2.5
Big Data /AI resources

ITCS Departments

Basic Research in Big Data / AI algorithms and methodologies

Other Big Data / AI research organizations and proposals
JST BigData CREST
JST AI CREST
Etc.

The current status of AI & Big Data in Japan

We need the triage of advanced algorithms/infrastructure/data but we lack the cutting edge infrastructure dedicated to AI & Big Data (c.f. HPC)

Tokyo Tech

AIRC

Joint RWBC Open Innov. Lab (OIL) (Director: Matsuoka)

AIST
NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY (AIST)
AIST-AIRC

東京工業大学
Tokyo Institute of Technology

Riken -AIP

NICT-UCRI

AI Venture Startups

UEI

Preferred Networks

AI/BD Centers & Labs in National Labs & Universities

R&D ML Algorithms & SW

Big Companies AI/BD R&D (also Science)

IT LAB
DENSO IT LABORATORY, INC.

Panasonic

ABEJA

MIZUHO
みずほ情報総研

DeNA

Seeking Innovative Application of AI & Data

ZMP

Use of Massive Scale Data now Wasted

DENSO

Petabytes of Drive Recording Video

FA&ロボット&ロボマシン
FANUC

FA&Robots

TOYOTA

JARI
一般財団法人 日本自動車研究所

Y! YAHOO! JAPAN

SoftBank

NTT

Web access and merchandice

IoT Communication, location & other data

"Big" Data

Over $1B Govt. AI investment over 10 years

Massive Rise in Computing Requirements (1 AI-PF/person?)

SAKURA Internet

Google Cloud Platform

Windows Azure

amazon web services

In HPC, Cloud continues to be insufficient for cutting edge research => dedicated SCs dominate & racing to Exascale

AI&Data Infrastructures

Massive "Big" Data in Training

JST-REST "Development and Integration of Artificial Intelligence Technologies for Innovation Acceleration"

Fast and cost-effective deep learning algorithm platform for video processing in social infrastructure

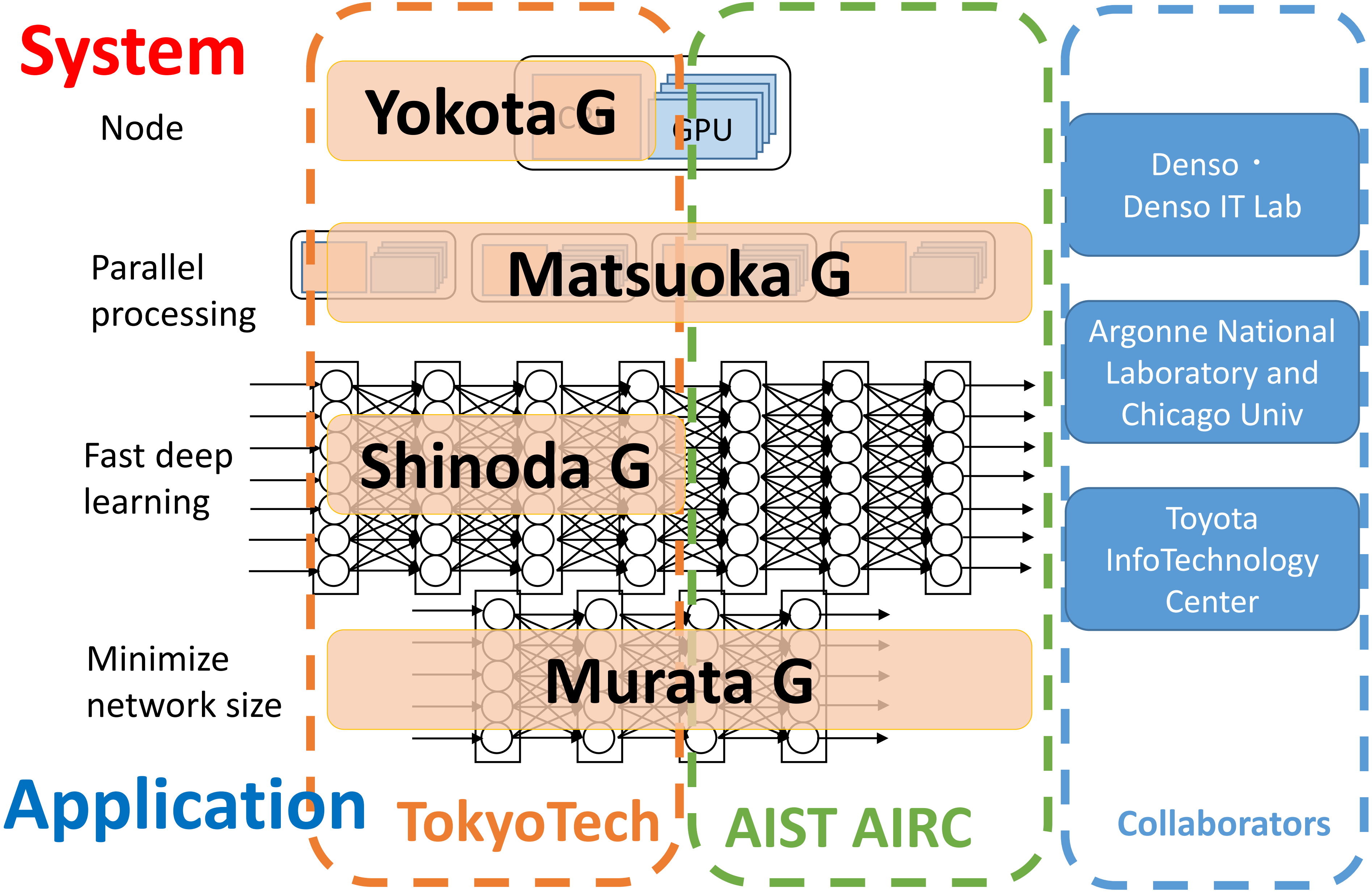
Principal Investigator: Koichi Shinoda

Collaborators: Satoshi Matsuoka
Tsuyoshi Murata
Rio Yokota

Tokyo Institute of Technology
(Members RWBC-OIL 1-1 and 2-1)

Research team

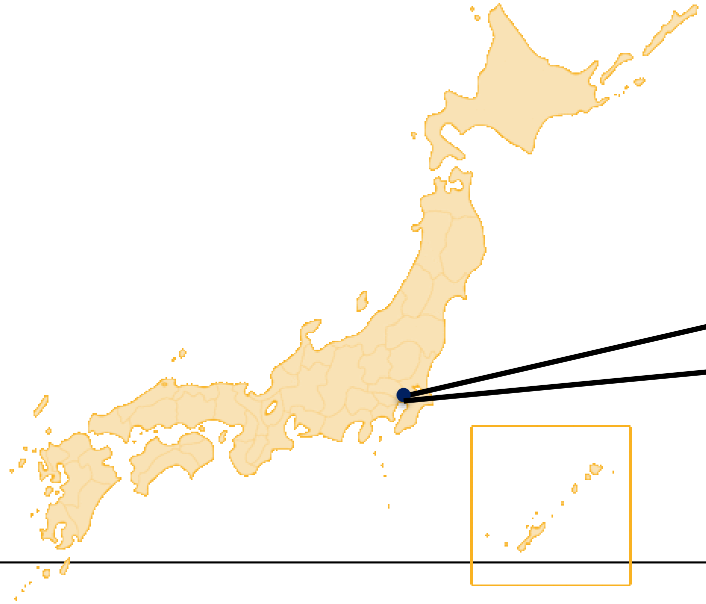
Reference



METI AIST-AIRC ABCI

as the *worlds first large-scale OPEN AI Infrastructure*

- **ABCI:** AI Bridging Cloud Infrastructure
  - Top-Level SC compute & data capability for DNN (>130 AI-Petaflops)
  - Open Public & Dedicated infrastructure for AI & Big Data Algorithms, Software and Applications
  - Platform to accelerate joint academic-industry R&D for AI in Japan



- >130~ AI-Petaflops
- < 3MW Power
- < 1.1 Avg. PUE
- Operational 2017Q4 ~2018Q1

東京大学 THE UNIVERSITY OF TOKYO

Univ. Tokyo Kashiwa Campus

The “Real” ABCI – 2018Q1

- **Extreme computing power**
  - w/ >**130 AI-PFlops** (likely several 100s AI-Pflops) for AI/ML especially DNN
  - **several million speedup** over high-end PC: 1 Day training for 10,000-Year DNN training job
  - TSUBAME-KFC (1.4 AI-Pflops) x 90 users (T2 avg) min
- **Big Data and HPC converged modern design**
  - For advanced data analytics (Big Data) and scientific simulation (HPC), etc.
  - Leverage Tokyo Tech’s “TSUBAME3” design, **but differences/enhancements being AI/BD centric**
- **Ultra high BW & Low latency memory, network, and storage**
  - For accelerating various AI/BD workloads
  - Data-centric architecture, optimizes data movement
- **Big Data/AI and HPC SW Stack Convergence**
  - **Incl. results from JST-CREST EBD**
  - **Wide contributions from the PC Cluster community desirable.**
- **Ultra-Green (PUE<1.1), High Thermal (60KW) Rack**
  - **Custom, warehouse-like IDC building and internal pods**
  - **Final “commoditization” of HPC technologies into Clouds**

ABCI Cloud Infrastructure

ABCI AI-IDC CG Image

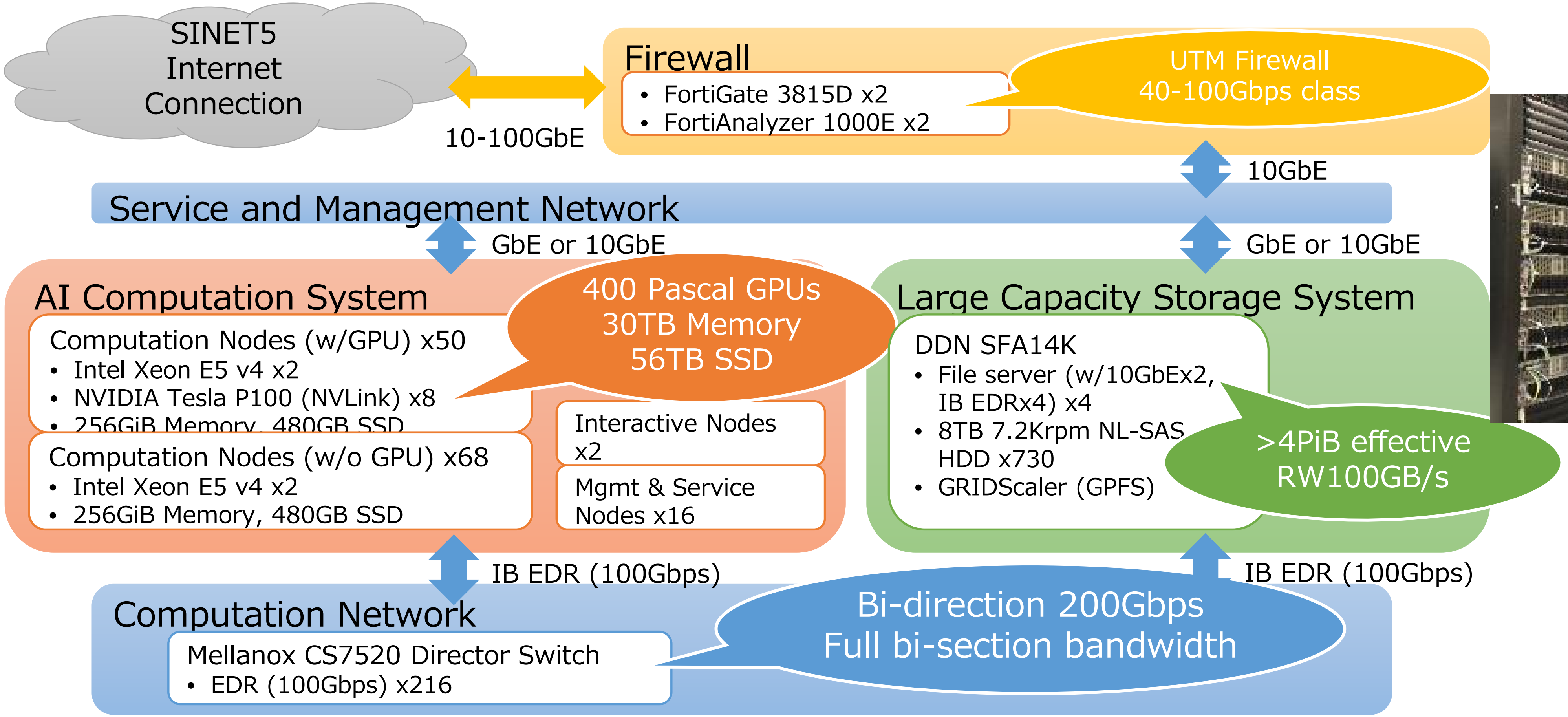
- **Ultra-dense IDC design from ground-up**
  - Custom inexpensive lightweight “warehouse” building w/ substantial earthquake tolerance
  - **x20 thermal density of standard IDC**
- **Extreme green**
  - Ambient warm liquid cooling, large Li-ion battery storage, and high-efficiency power supplies, etc.
  - **Commoditizing supercomputer cooling technologies to Clouds (60KW/rack)**
- **Cloud ecosystem**
  - Wide-ranging Big Data and HPC standard software stacks
- **Advanced cloud-based operation**
  - Incl. dynamic deployment, container-based virtualized provisioning, multitenant partitioning, and automatic failure recovery, etc.
  - Joining HPC and Cloud Software stack for real
- **Final piece in the commoditization of HPC (into IDC)**
- **Open Sourcing of Next-Gen IDC Architecture for AI**

Reference Image

引用元: NEC導入事例

ABCI Prototype: AIST AI Cloud (AAIC) March 2017 (#3 June 2017 Green 500)

- **400x NVIDIA Tesla P100s and Infiniband EDR** accelerate various AI workloads including ML (Machine Learning) and DL (Deep Learning).
- Advanced data analytics leveraged by **4PiB shared Big Data Storage and Apache Spark** w/ its ecosystem.



ABCI Update Oct. 1st

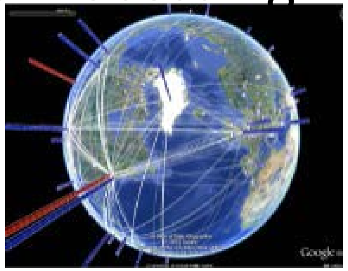
- Fujitsu has won the contract, subject to signing and delivery
- Cannot disclose the details until formal announcement (hopefully by SC17)

- The compute nodes are Fujitsu-designed next gen GPU nodes with 4 NVIDIA Voltas
- Interconnected by two Mellanox EDR links, tapered FatTree
- Large capacity DRAM + Intel NVMe per node
- Large capacity HDD+SSD DDN storage, GPFS+S3+Swift+… and BeeOND for temporary store
- Various HPC+Cloud+AI software in the software stack
- 100Gbps external connectivity + firewall
- Warm water liquid cooling, very low PUE

Characteristics of Big Data and AI Computing

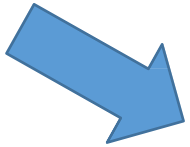
As BD / AI

Graph Analytics e.g. Social Networks
Sort, Hash, e.g. DB, log analysis
Symbolic Processing: Traditional AI



As HPC Task

Integer Ops & Sparse Matrices
Data Movement, Large Memory
Sparse and Random Data, Low Locality



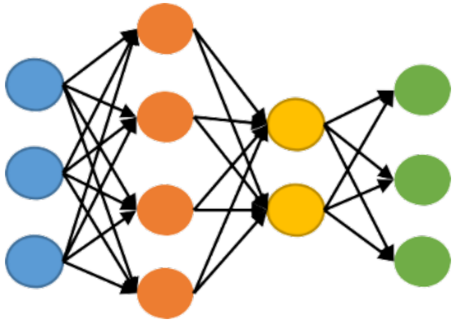
Acceleration, Scaling

Opposite ends of HPC computing spectrum, but HPC simulation apps can also be categorized likewise

Acceleration via Supercomputers adapted to AI/BD

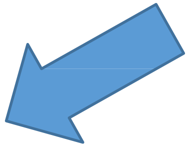
As BD / AI

Dense LA: DNN
Inference, Training, Generation



As HPC Task

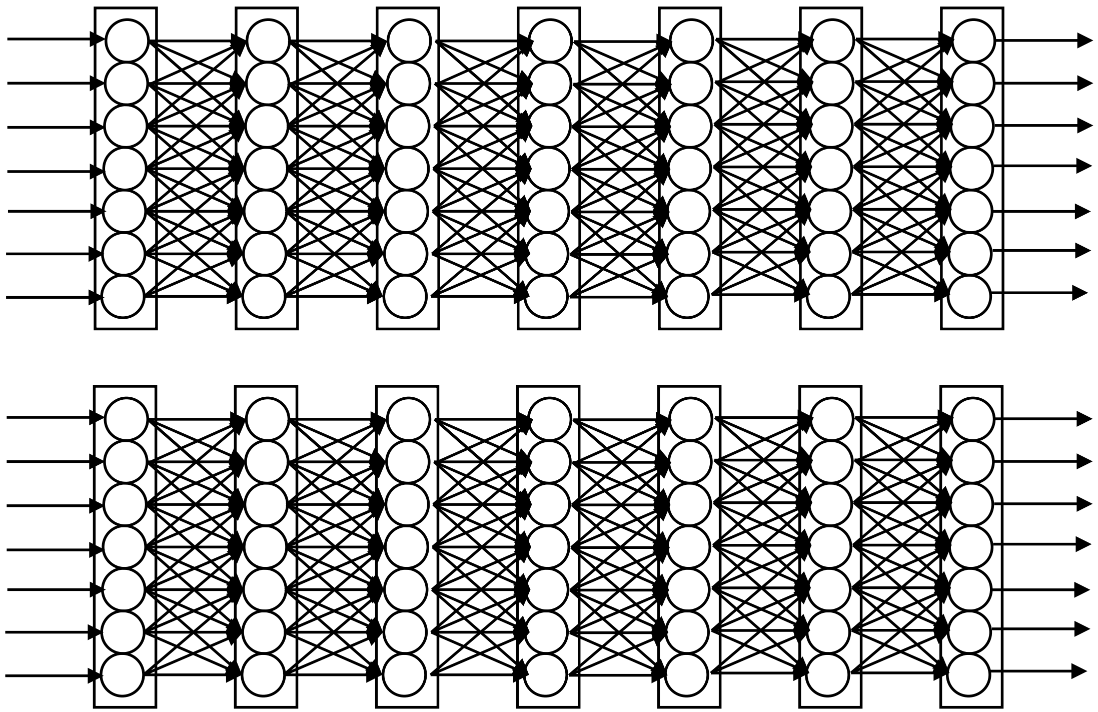
Dense Matrices, Reduced Precision
Dense and well organized neworks and Data



Acceleration, Scaling

4 Layers of Parallelism in DNN Training

- Hyper Parameter Search
  - Searching optimal network configurations and parameters
  - Often use evolutionary algorithms
- Data Parallelism
  - Split and parallelize the batch data
  - Synchronous, asynchronous, hybrid, …
- Model Parallelism
  - Split and parallelize the layer calculations in forward/backward propagation
- ILP and other low level Parallelism
  - Parallelize the convolution operations etc. (in reality tensor op / matrix multiply)



What about the other layers?

How do we co-Design?

Deep Learning is “All about Scale”

- **Andrew Ng:**
  - “Deep Learning is scalable”
  - “Performance just gets better if you feed in more data”
- **Data-parallel training with (Asynchronous) Stochastic Gradient Descent**

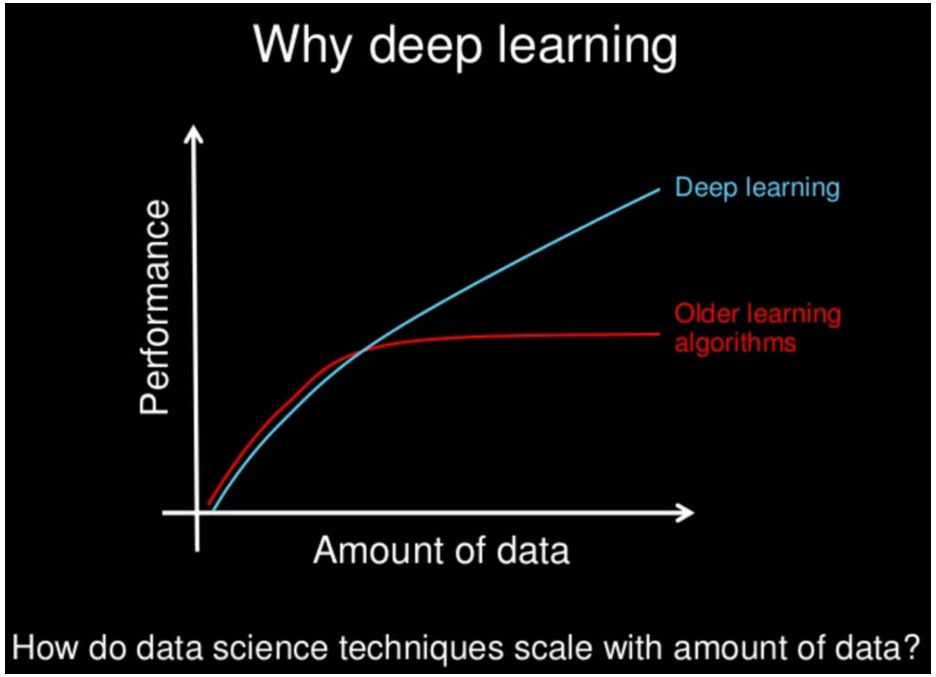


Fig. 2: Andrew Ng (Baidu) “What Data Scientists Should Know about Deep Learning”

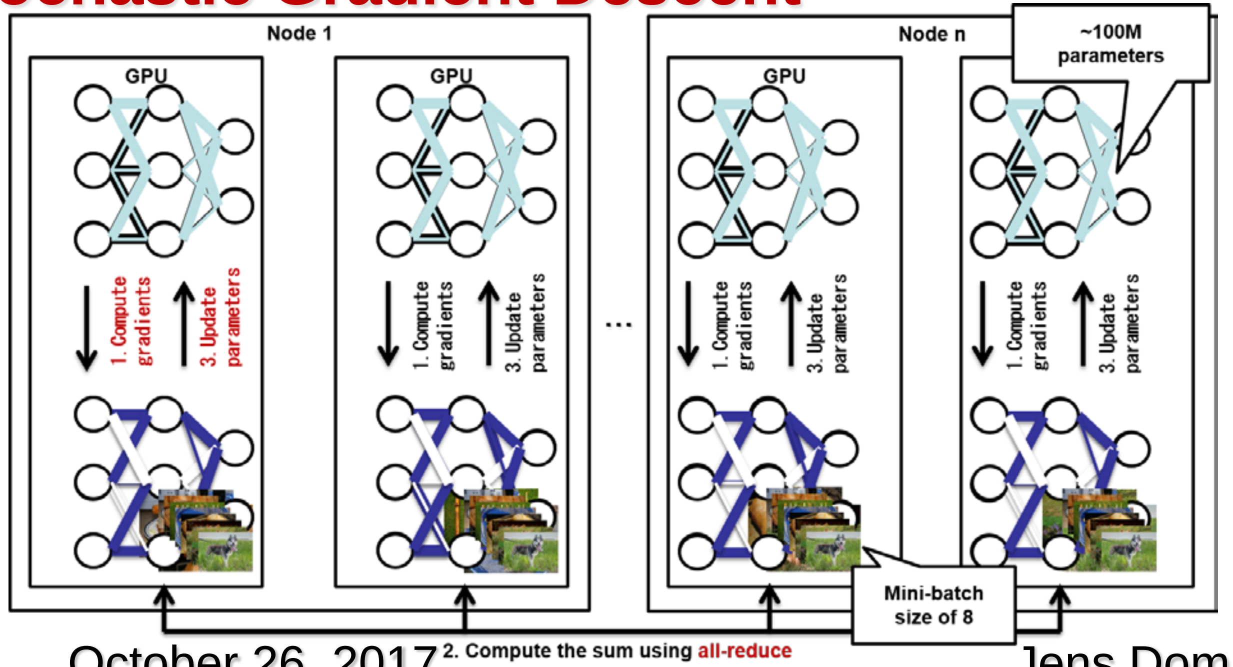


Fig. 3: Simplified DL workflow with ASGD per iteration:
1. Compute gradient
2. Exchange gradients via all-reduce; and
3. Update network parameters

Deep Learning is “All about Scale”

- **In Strong scale training (w/ fixed mini-batch size) inter-GPU and inter-node communication is bottleneck**
  - T_{Comp} ($\propto$ 1/#GPUs) $\ll$ T_{Comm} ($\propto$ log(#GPUs))
- **In Weak scale training (w/ fixed batch size per GPU) a large mini-batch may harm the DL network’s accuracy**
  - T_{Comp} (= const.) $\gg$ T_{Comm} ($\propto$ log(#GPUs))
  - Related work on distributed DL usually focus on weak scaling

➔ **Shortening communication time is essential to accelerate DL without any accuracy loss**

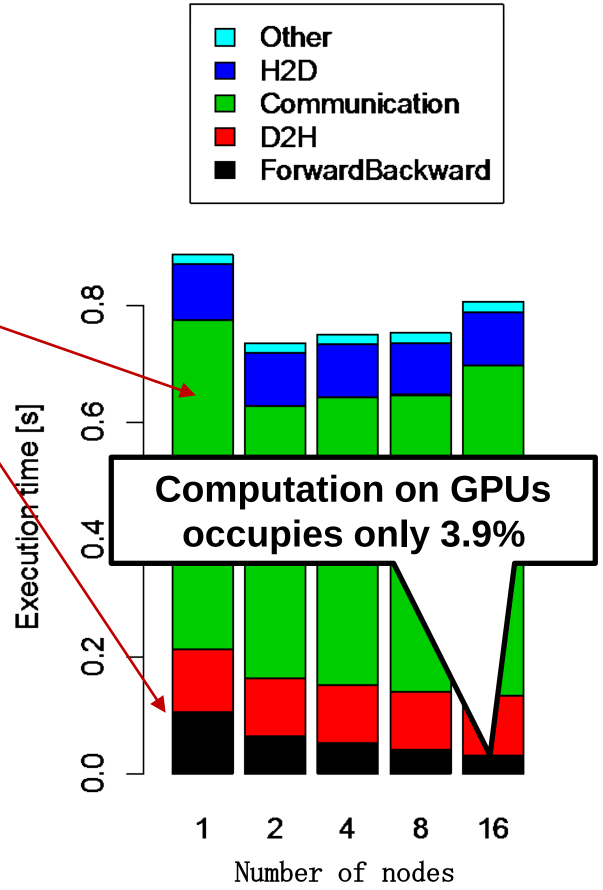
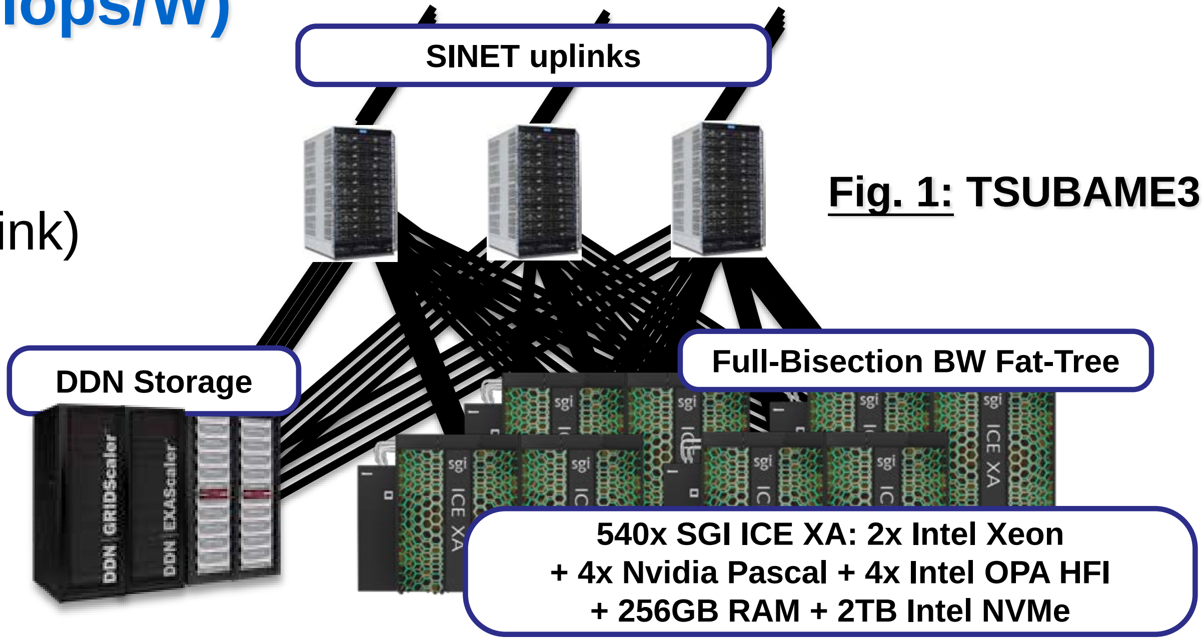


Fig. 3: Breakdown of one iteration of CaffeNet training on TSUBAME-KFC/DL (8 GPUs/node, Mini-batch size of 256)

Huge Injection BW (400G) of T3 enough?

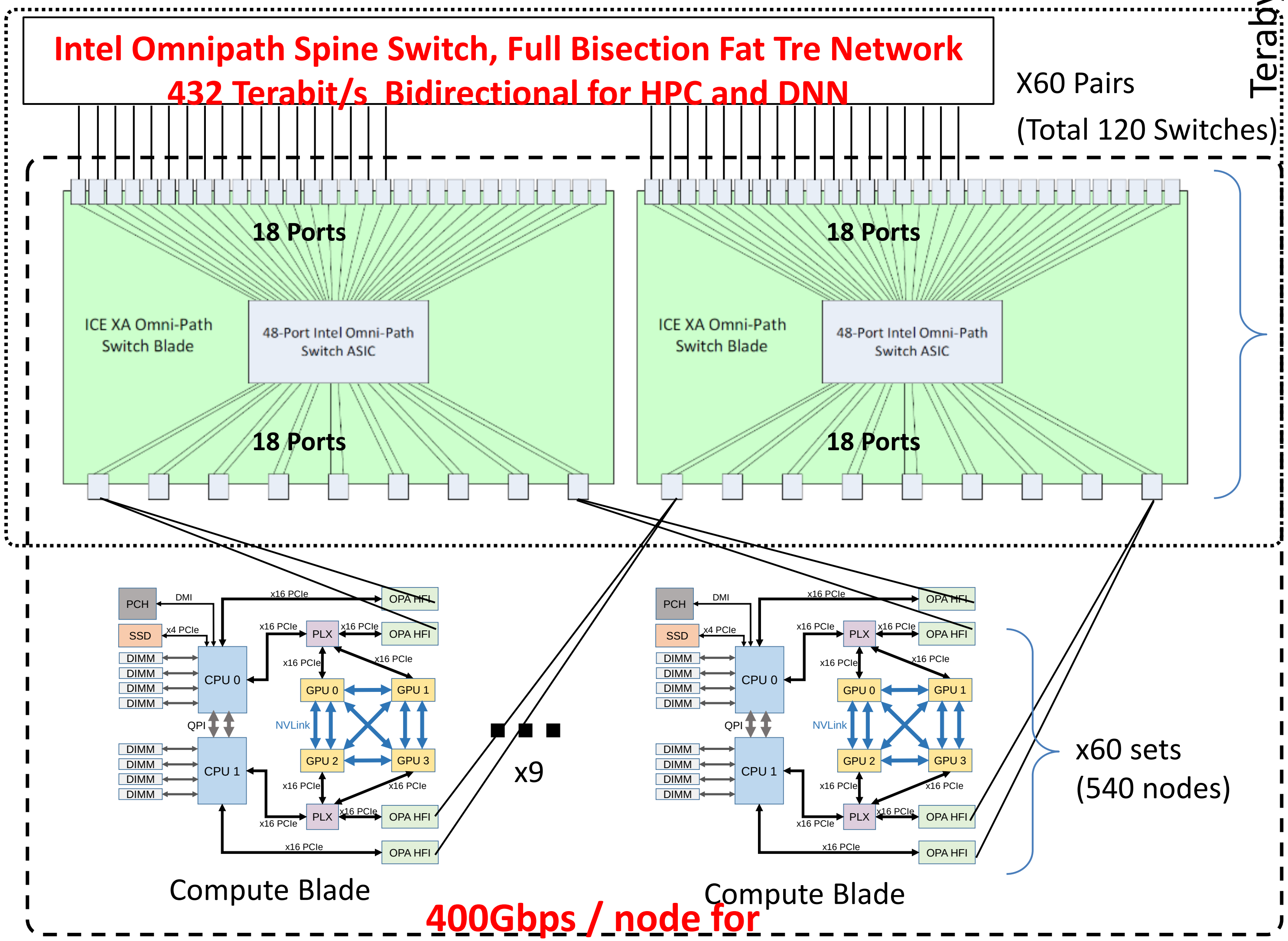
- **TSUBAME2.5 (T2)**
  - 1408 nodes with 3x Nvidia K20X
  - Dual-port QDR IB attached to 2 full-bisection bandwidth fat-trees
  - ➔ 3x 16 GB/s PCIe x16 vs. 2x 4 GB/s IB (6 : 1 bandwidth ratio)
- **TSUBAME3 (Green500 #1: 14.110 GFlops/W)**
  - **12.1 Pflop/s (DP) or 47.2 AI-Pflop/s**
    - 15 racks (36 CNs; 4x HFI; 16x OPA SW)
  - 540 nodes w/ 4x Nvidia P100 (+ all2all NVLink)
  - 4x 100 Gbps Intel OPA (gen.1) injection ports into single full-bisec. BW fat-tree
  - ➔ 6x 80 GB/s NVLink vs. 4x 16 GB/s PCIe vs. 4x 12.5 GB/s OPA (≈ 10 : 1.3 : 1 ratio)
  - ➔ Getting worse; need to get data out!!!



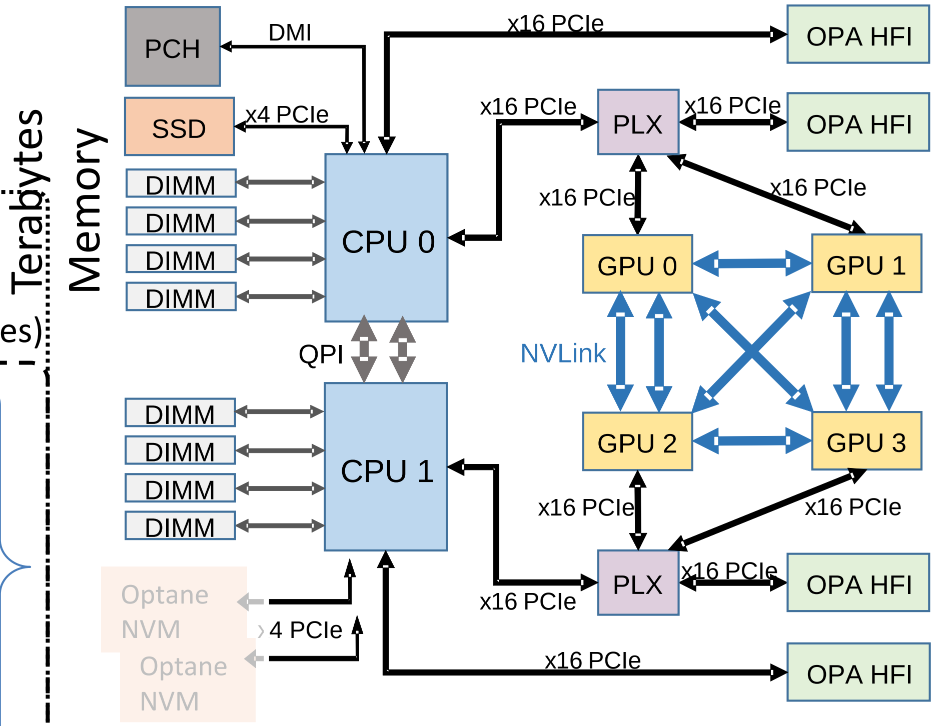
<u>Fig. 1:</u> **TSUBAME3**

TSUBAME3.0 Compute Node SGI ICE-XA, **a New GPU Compute Blade Co-Designed by SGI and Tokyo Tech GSIC**

SGI ICE XA Infrastructure



400Gbps / node for HPC and DNN



Ultra high performance & bandwidth "Fat Node"

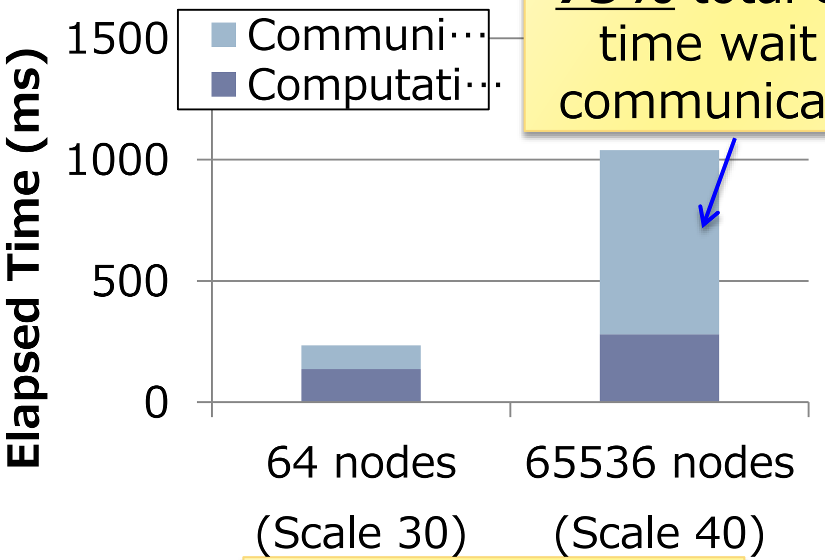
- High Performance: 4 SXM2(NVLink) NVIDIA Pascal P100 GPU + 2 Intel Xeon 84 AI-TFLops
- High Network Bandwidth – Intel Omnipath 100GBps x 4 = 400Gbps (100Gbps per GPU)
- High I/O Bandwidth - Intel 2 TeraByte NVMe
  - > 1PB & 1.5~2TB/s system total
  - Future Octane 3D-Xpoint memory Petabyte or more directly accessible
- Ultra High Density, Hot Water Cooled Blades
  - 36 blades / rack = 144 GPU + 72 CPU, 50-60KW, x10 thermals c.f. IDC

HPCG Top 10 ranking June 2017

#	T	Site	Manufacturer	Computer	Country	HPCG [Pflop/s]	Rmax [Pflop/s]	HPCG/ Peak	HPCG/ HPL
1	8	RIKEN Advanced Institute for Computational Science	Fujitsu	K Computer SPARC64 VIIIfx 2.0GHz, Tofu Interconnect	Japan	0.6027	10.5	5.3%	5.7%
2	2	National University of Defense Technology	NUDT	Tianhe-2 NUDT TH-IVB-FEP, Xeon 12C 2.2GHz, IntelXeon Phi	China	0.5801	33.9	1.1%	1.7%
3	3	Swiss National Supercomputing Centre (CSCS)	Cray	Piz Daint Cray XC50, Xeon E5 12C 2.6GHz, Aries, NVIDIA Tesla P100	Switzerland	0.4700	19.6	1.9%	2.4%
4	7	JCAHPC Joint Center for Advanced HPC	Fujitsu	Oakforest-PACS PRIMERGY CX1640 M1, Intel Xeons Phi 7250 68C 1.4 GHz, OmniPath	Japan	0.3855	13.6	1.5%	2.8%
5	1	National Supercomputing Center in Wuxi	NRCPC	Sunway TaihuLight NRCPC Sunway SW26010, 260C 1.45GHz	China	0.3712	93.0	0.3%	0.4%
6	6	Lawrence Berkeley National Laboratory	Cray	Cori Cray XC40, Intel Xeons Phi 7250 68C 1.4 GHz, Aries	USA	0.3554	14.0	1.3%	2.5%
7	5	Lawrence Livermore National Laboratory	IBM	Sequoia BlueGene/Q, Power BQC 16C 1.6GHz, Custom	USA	0.3304	17.2	1.6%	1.9%
8	4	Oak Ridge National Laboratory	Cray	Titan Cray XK7, Opteron 16C 2.2GHz, Gemini, NVIDIA K20x	USA	0.3223	17.6	1.2%	1.8%
9	10	Los Alamos NL / Sandia NL	Cray	Trinity Cray XC40, Xeon E5 16C 2.3GHz, Aries	USA	0.1826	8.10	1.6%	2.3%
10	15	NASA/ Ames Research Center/NAS	HPE	Pleiades SGI ICE X, Xeon E5 10C 2.4-2.8GHz, Infiniband FDR	USA	0.1750	5.95	2.5%	2.9%

Sparse BYTES: The Graph500 – 2015~2016 – world #1 x 4

K Computer #1 Tokyo Tech[Matsuoka EBD CREST] Univ. Kyushu [Fujisawa Graph CREST], Riken AICS, Fujitsu



73% total exec time wait in communication

88,000 nodes,
660,000 CPU Cores
1.3 Petabyte mem
20GB/s Tofu NW

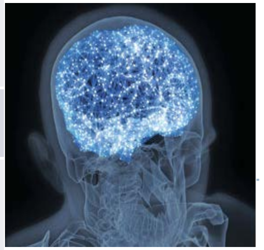
K computer

Effective x13 performance c.f. Linpack

≫

List	Rank	GTEPS	Implementat
November 2013	4	5524.12	Top-down o
June 2014	1	17977.05	Efficient hybrid
November 2014	2	19585.2	Efficient hybrid
June, Nov 2015 June Nov 2016	1	38621.4	Hybrid + Node Compression

BYTES Rich Machine + Superior BYTES algoithm



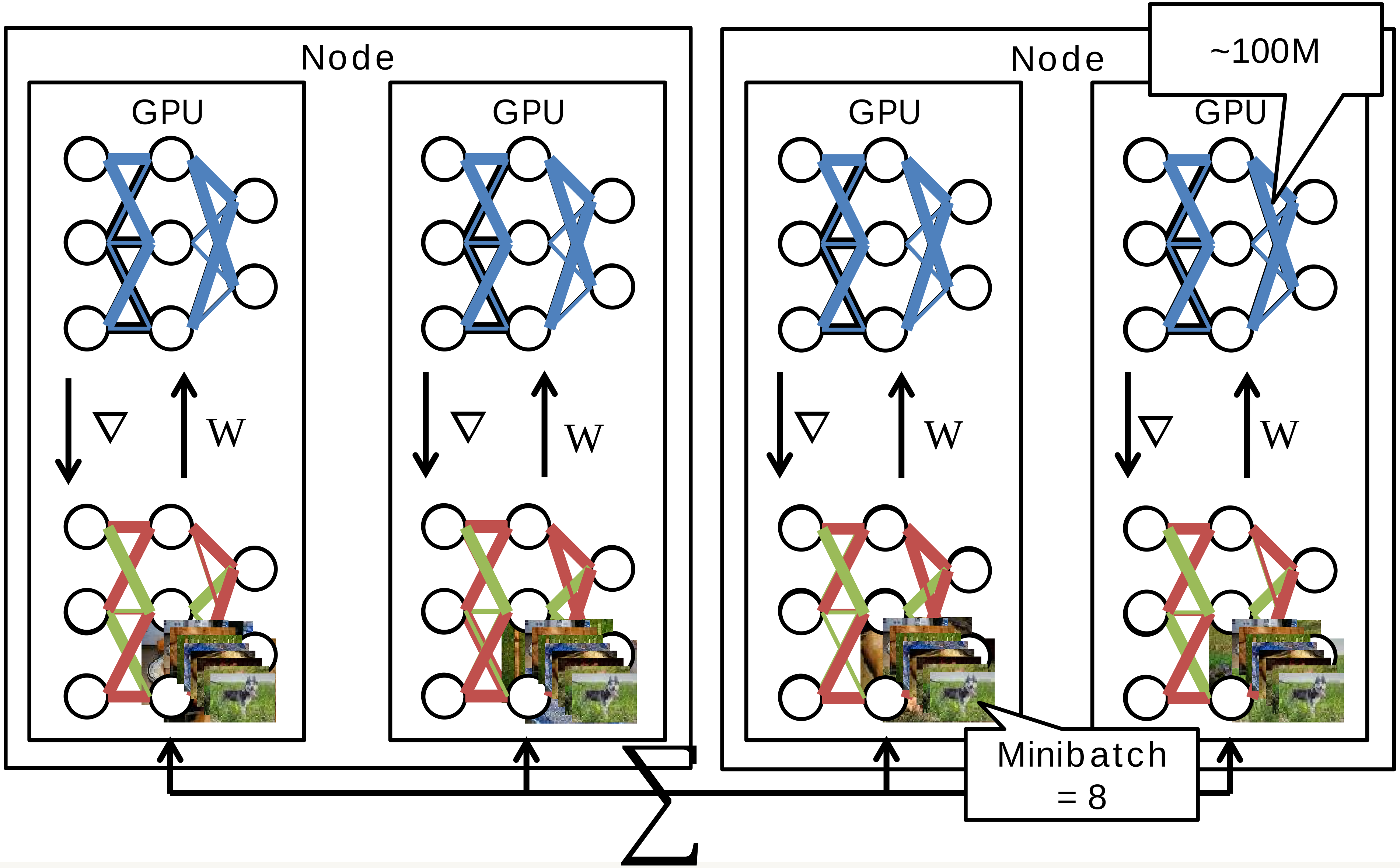
LLNL-IBM Sequoia
1.6 million CPUs
1.6 Petabyte mem

TaihuLight
10 million CPUs
1.3 Petabyte mem

BYTES, not FLOPS!

Parallelizing Deep Neural Network Training

Data Parallel SGD(Stochastic Gradient Descent)



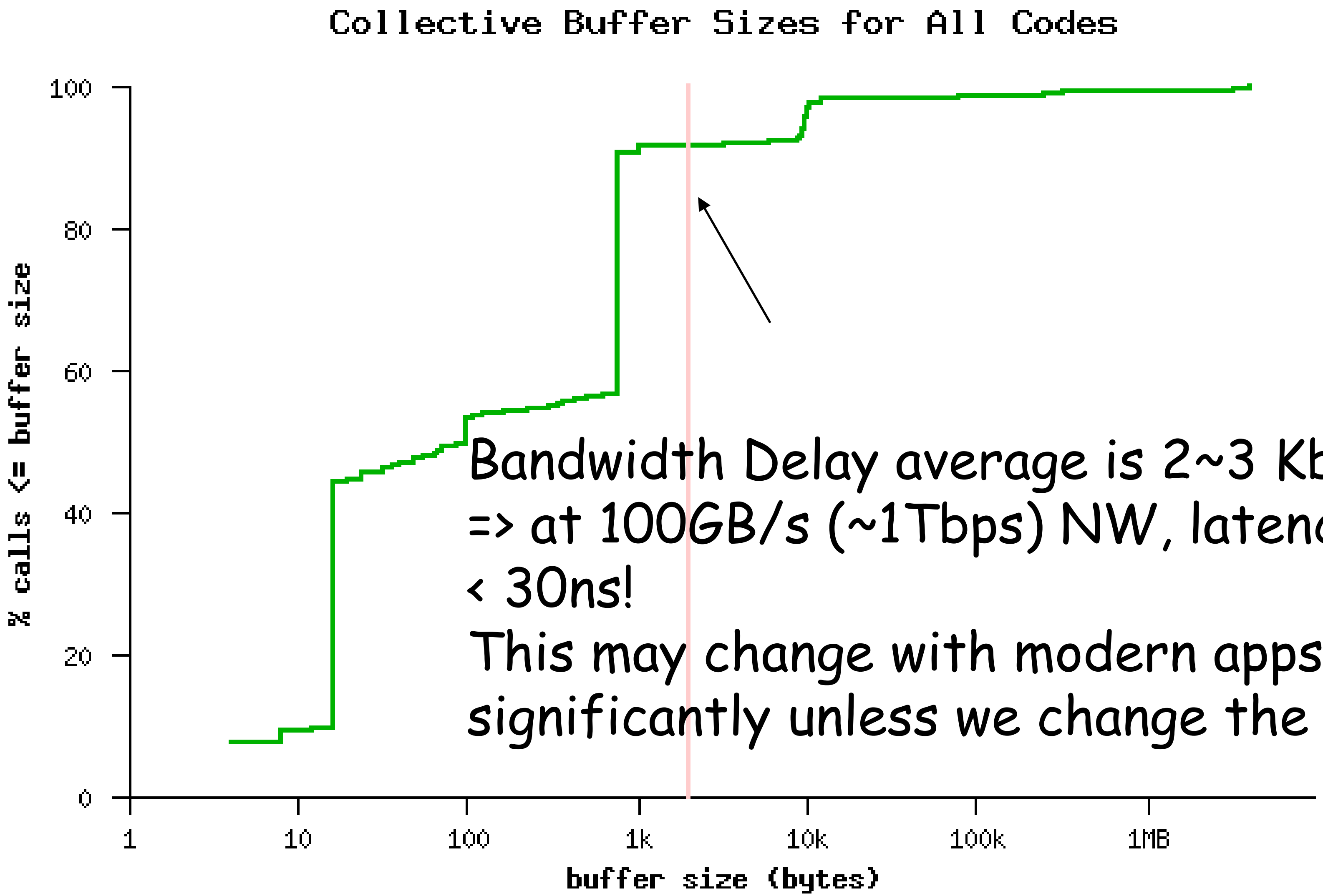
DoE SC Applications Communication Analysis

(以下 courtesy John Shalf @ LNBL)

Kamil et. al. "Communication Requirements and Interconnect Optimizations for High-End Scientific Applications", IEEE Trans. Parallel and Distributed Systems, 2010

NAME	Discipline	Problem/Method	Structure
MADCAP	Cosmology	CMB Analysis	Dense Matrix
FVCAM	Climate Modeling	AGCM	3D Grid
CACTUS	Astrophysics	General Relativity	3D Grid
LBMHD	Plasma Physics	MHD	2D/3D Lattice
GTC	Magnetic Fusion	Vlasov-Poisson	Particle in Cell
PARATEC	Material Science	DFT	Fourier/Grid
SuperLU	Multi-Discipline	LU Factorization	Sparse Matrix
PMEMD	Life Sciences	Molecular Dynamics	Particle

Collective Buffer Sizes Average is Small!



Bandwidth Delay average is 2~3 Kbytes
=> at 100GB/s (~1Tbps) NW, latency needs to be < 30ns!
This may change with modern apps but not significantly unless we change the algorithms

(Original slide courtesy John Shalf @ LBL)

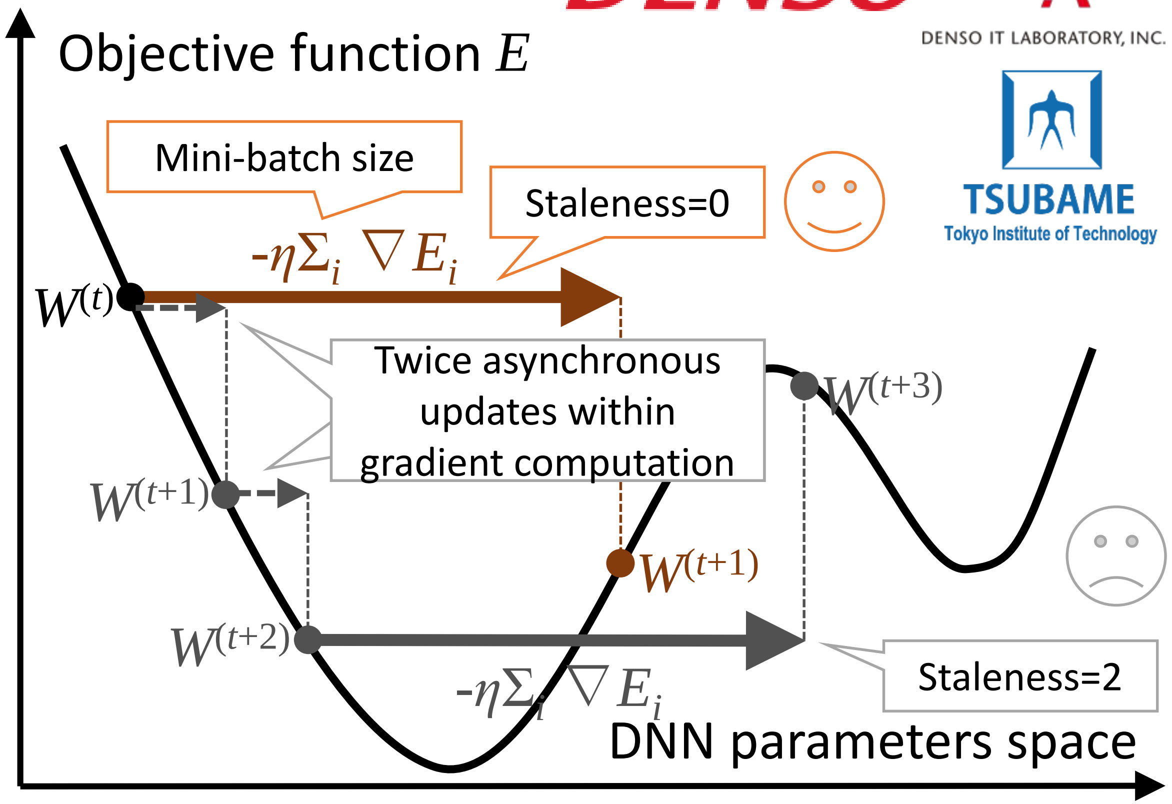
Example AI Research: Predicting Statistics of Asynchronous SGD Parameters for a Large-Scale Distributed Deep Learning System on GPU Supercomputers

Background

- In large-scale Asynchronous Stochastic Gradient Descent (ASGD), mini-batch size and gradient staleness tend to be large and unpredictable, which increase the error of trained DNN

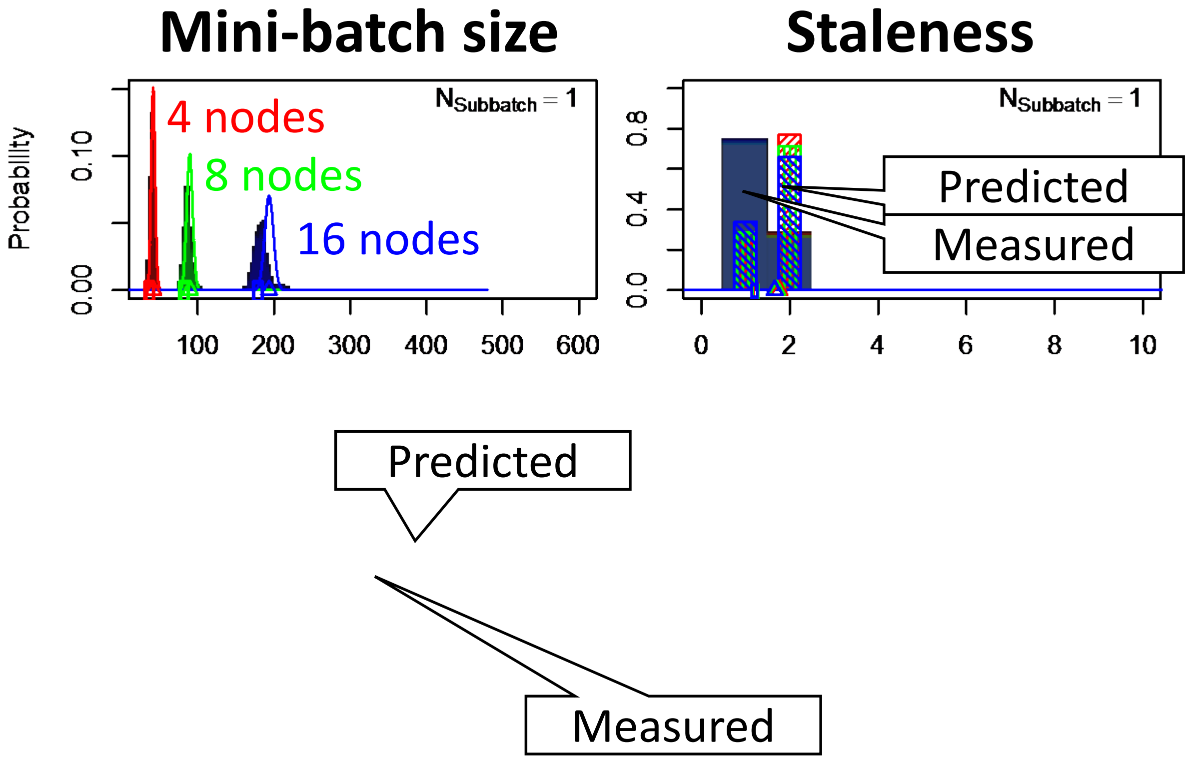
DENSO IT LABORATORY, INC.

TSUBAME
Tokyo Institute of Technology



Proposal

- We propose a empirical performance model for an ASGD deep learning system SPRINT which considers probability distribution of mini-batch size and staleness



($N_{Subbatch}$: # of samples per one GPU iteration)

- Yosuke Oyama, Akihiro Nomura, Ikuro Sato, Hiroki Nishimura, Yukimasa Tamatsu, and Satoshi Matsuoka, "**Predicting Statistics of Asynchronous SGD Parameters for a Large-Scale Distributed Deep Learning System on GPU Supercomputers**", in proceedings of 2016 IEEE International Conference on Big Data (IEEE BigData 2016), Washington D.C., Dec. 5-8, 2016

Interconnect Performance as important as GPU Performance to accelerate DL

- **ASGD DL system SPRINT (by DENSO IT Lab) and DL speedup prediction with performance model**

$$T_{Epoch} = \frac{N_{File} \times T_{GPU}}{N_{Node} \times N_{GPU} \times N_{Subbatch}}$$

  - Data measured on T2 and KFC (both FDR) fitted to formulas
  - Allreduce time (∈ T_{GPU}) dep. on #nodes and #DL_parameters

$$T_{Barrier} + (\alpha \log_2(N_{Node}) + \beta) \times N_{Param}$$

The Optimal Predicted Configurations of CNN-A on TSUBAME-KFC/DL

	N_{Node}	$N_{Subbatch}$	Average mini-batch size	Epoch time[s]	Speedup
Baseline	8	8	165.1	1779	-
FP16	7	22	170.1	1462	**1.22**
EDR IB	12	11	166.6	1245	**1.43**
FP16 + EDR IB	8	15	171.5	1128	**1.58**

Fig. 4: Oyama et al. "Predicting Statistics of Asynchronous SGD Parameters for a Large-Scale Distributed Deep Learning System on GPU Supercomputers

- **Other approaches == similar improvements:**
  - Cuda-Aware CNTK optimizes communication pipeline ➔ 15%—23% speedup
    (Banerjee et al. "Re-designing CNTK Deep Learning Framework on Modern GPU Enabled Clusters")
  - Reduced precision (FP[16|8|1]) to minimize msg. size w/ no or minor accuracy loss

Allreduce of Huge Arrays of Gradients

- **Msg. sizes ≫ 100 MB common even for small networks**
  ➔ potentially [G|T]Bytes of data to exchange per epoch
  and comm. time may dominate the time required per epoch (up to 78%)
  (Gewande et al. "Scaling Deep Learning Workloads: NVIDIA DGX-1/Pascal and Intel Knights Landing")
- **Approaches for efficient all-reduce ops**
  - Reduction trees *(deprecated)* **(Iandola et al. "FireCaffe: near-linear acceleration of deep neural network training on compute clusters")**
  - (Ring-based) streaming reduction →
  - Linear pipeline **(Wang et al. "Efficient Communication in Training Large Scale Neural Networks")**

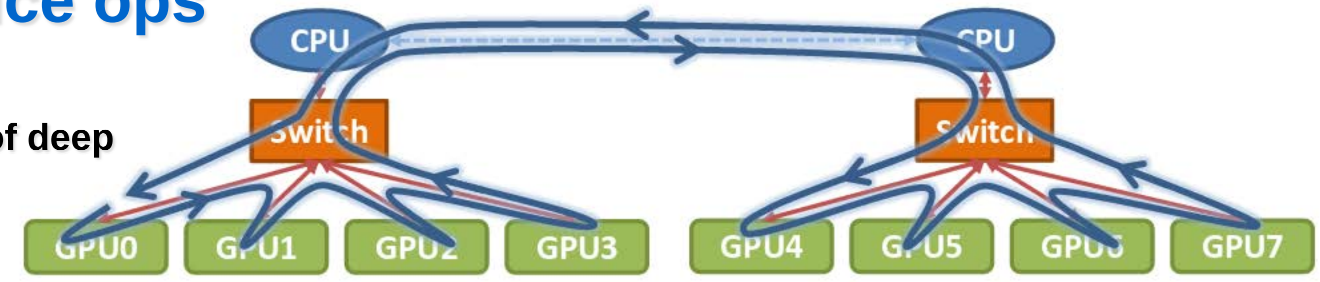


Fig. 5: Luehr et al. "NCCL: Accelerated Collective Communications for GPUs" and Gibiansky "Effectively Scaling Deep Learning Frameworks"

- **Large data (streaming) reductions only bound by bandwidth**
  - Network injection usually maxed out at ≈1MB (see http://mvapich.cse.ohio-state.edu/performance/pt_to_pt/) per port/NIC ➔ theoretically: more injection ports == faster reduction !!!

Higher Injection ➔ Faster Reduction

- **AIST AI Cloud (nodes similar to DGX-1)**
  - Each node: 2x Intel Xeon; 8x Nvidia P100; 1x IB EDR; 256GB RAM; 480GB SSD
- **Benchmarking "Inj. BW"-hypothesis w/ Microsoft Cognitive Toolkit (CNTK) and a 50-layer residual network**
  - Test 1 w/ 32 GPUs on 4 nodes **(blue graph)**
  - Test 2 w/ 32 GPUs on 8 nodes (red graph)
    ➔ relative inj. BW available per GPU doubles
  - Same input; def. MPI-based all-reduce (no streaming)

➔ **Doubling injection BW shortens time to reach top-1 accuracy by >20%**

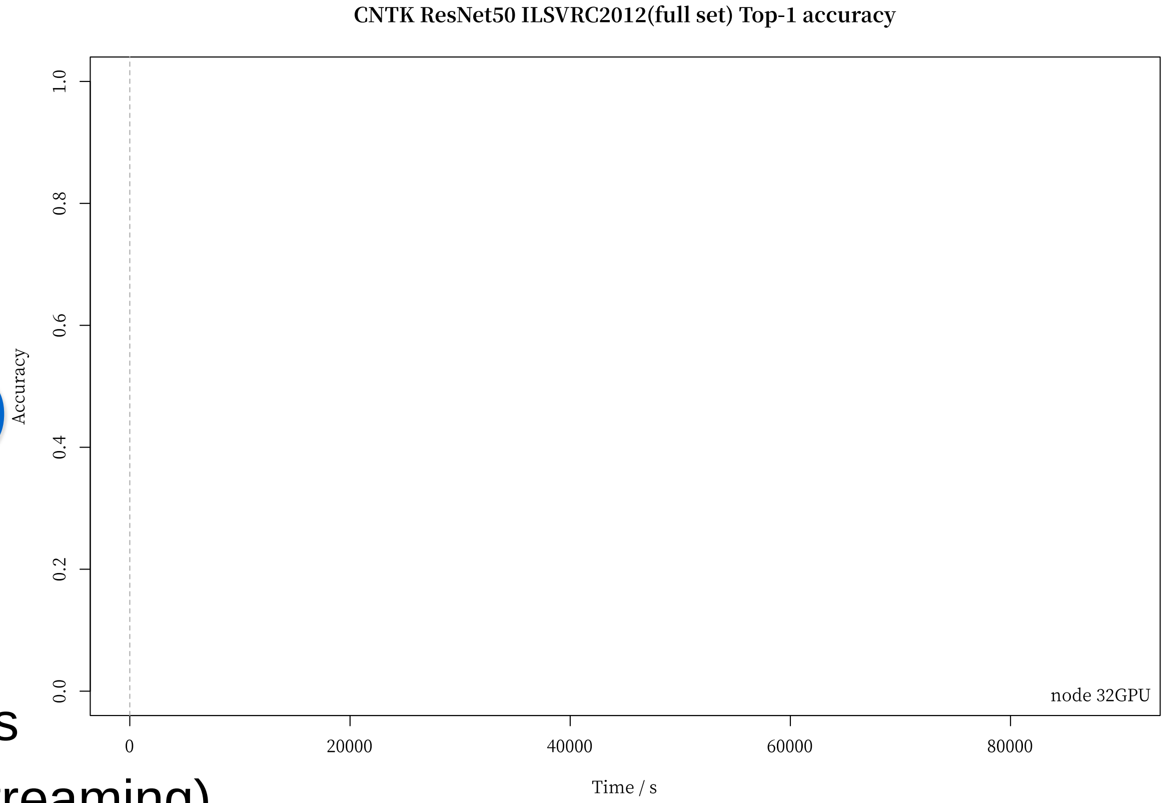
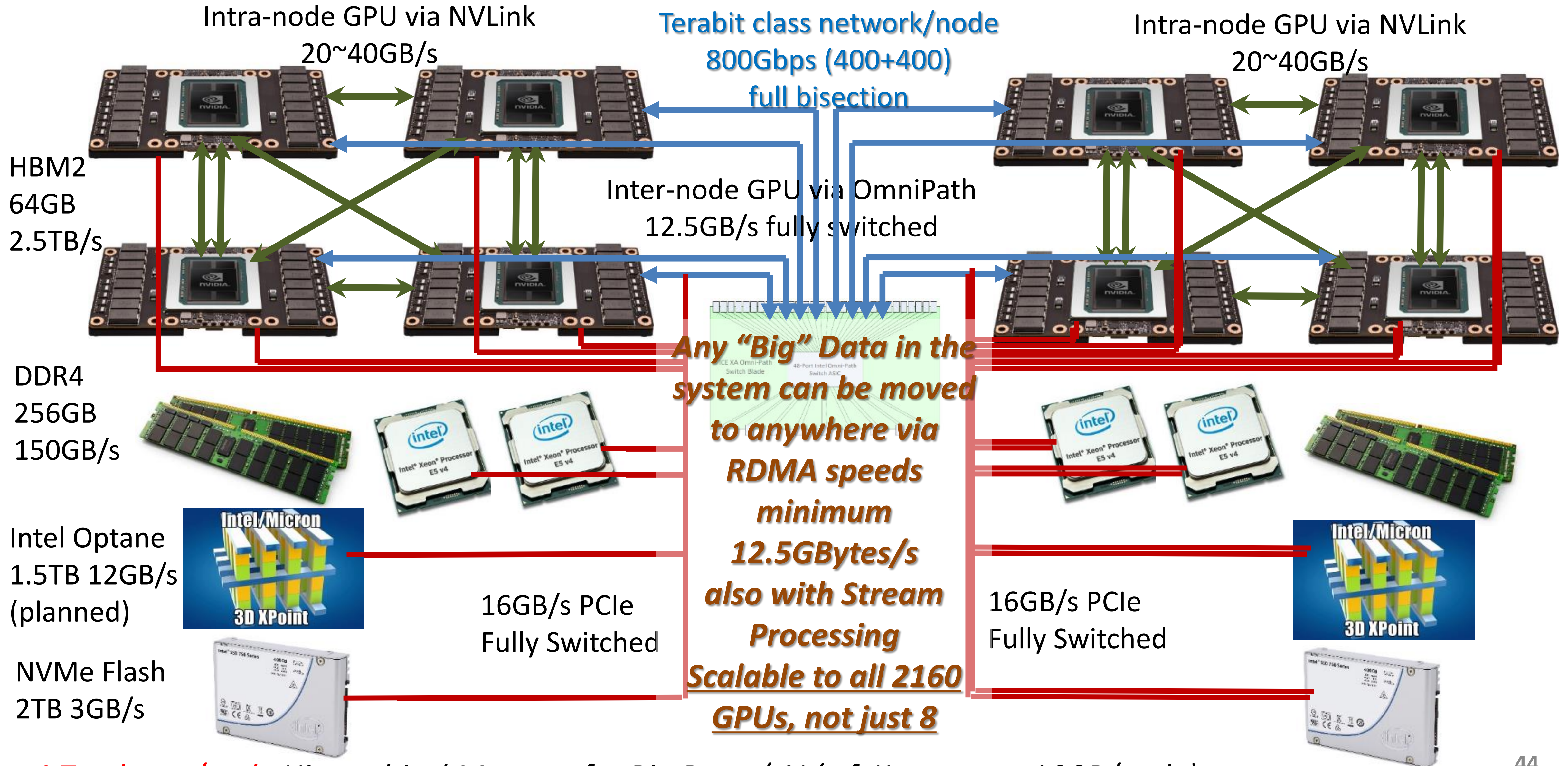


Fig. 6: Measuring time to reach top-1 accuracy for ResNet-50 on 4 or 8 nodes, respectively; but w/ same #{GPUs} ➔ hence, no (strong) scaling

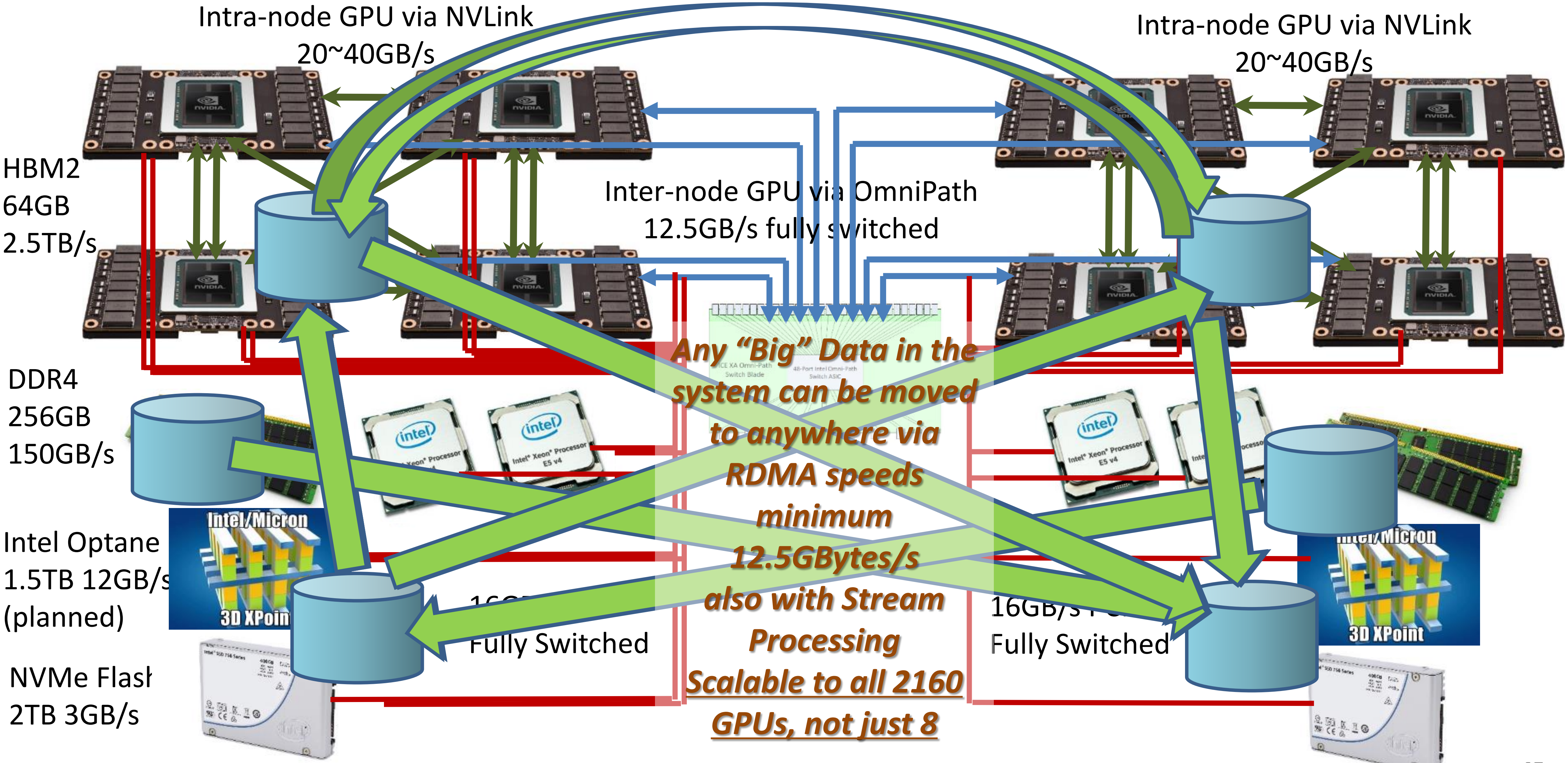
TSUBAME3: *A Massively BYTES Centric Architecture* for Converged BD/AI and HPC



~4 Terabytes/node Hierarchical Memory for Big Data / AI (c.f. K-compuer 16GB/node)

➔ *Over 2 Petabytes in TSUBAME3, Can be moved at 54 Terabyte/s or 1.7 Zetabytes / year*

TSUBAME3: *A Massively BYTES Centric Architecture* for Converged BD/AI and HPC



~4 Terabytes/node Hierarchical Memory for Big Data / AI (c.f. K-compuer 16GB/node)

➔ *Over 2 Petabytes in TSUBAME3, Can be moved at 54 Terabyte/s or 1.7 Zetabytes / year*

Some ABCI Metrics

- 1088 post Fujitsu Multi-GPU server
  - 4 NVIDIA Volta SXM2 GPU + 2 Intel Xeon Gold (20 cores) + 384GB DRAM + 1.6PB Intel NVMe SSD + 2 Mellanox EDR
  - Warm Water (30+C) cooling, PUE < 1.1 (estimate)
- 550 Peta AI-Flops (FP16) Peak performance
  - C.f. Pascal DGX-1: ~x3000 inference, x1500 training
- 22 PetaByte DDN GridScaler
  - Multi-protocol storage: GPFS, S3, SWIFT, …
  - BeOND to federate on-node NVMes (1.6 Petabyte)
- 80 slots PCI-e future AI experiments
  - New AI chips, e.g. Fujitsu DLU, FPGAs, etc.
- 100Gbs external connectivity
  - 80 Gbps Fortigate firewall throughput

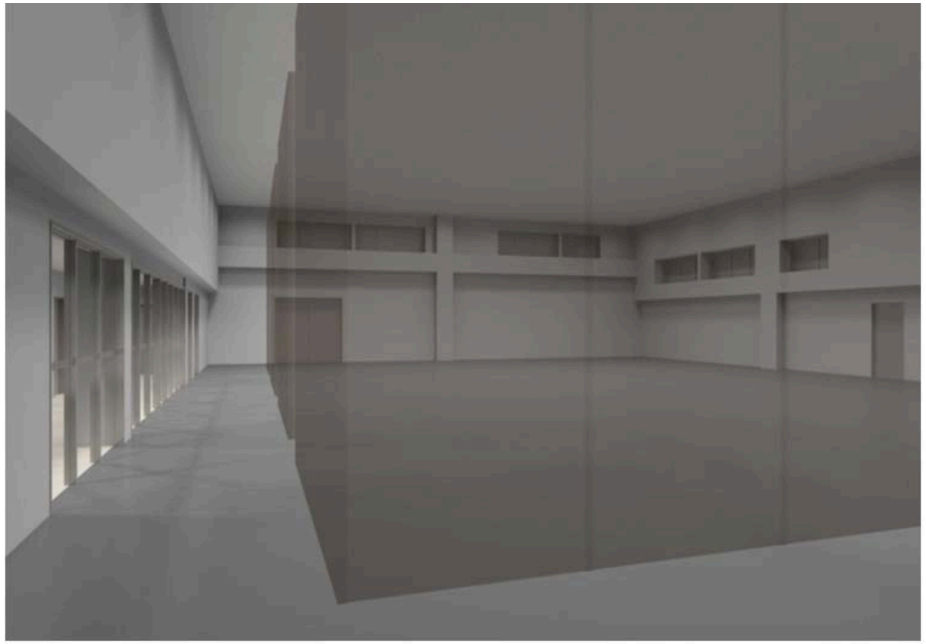
ABCI Procurement Benchmarks

- **Big Data Benchmarks**
  - (SPEC CPU Rate)
  - Graph 500
  - MinuteSort
  - Node Local Storage I/O
  - Parallel FS I/O

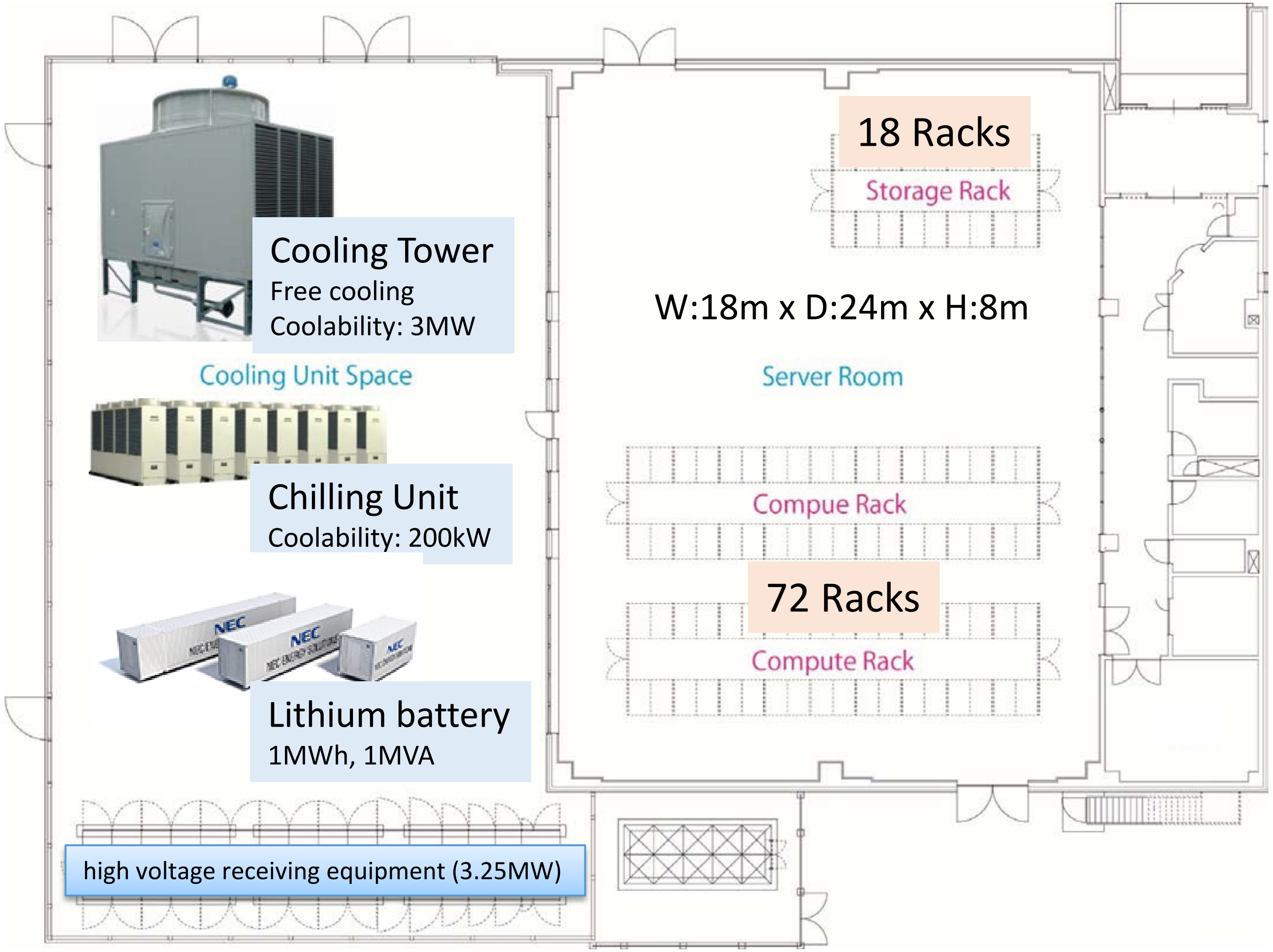
No traditional HPC Simulation Benchmarks except SPEC CPU. Plan on “open-sourcing”

- **AI/ML Benchmarks**
  - Low precision GEMM
    - CNN Kernel, defines “AI-Flops”
  - Single Node CNN
    - AlexNet and GoogLeNet
    - ILSVRC2012 Dataset
  - Multi-Node Scalable CNN
    - Caffe+MPI
  - Large Memory CNN
    - Convnet on Chainer
  - RNN / LSTM
    - Neural Machine Translation on Torch

ABCI Datacenter ~$10 million (Just broke ground, to be completed late 2017)

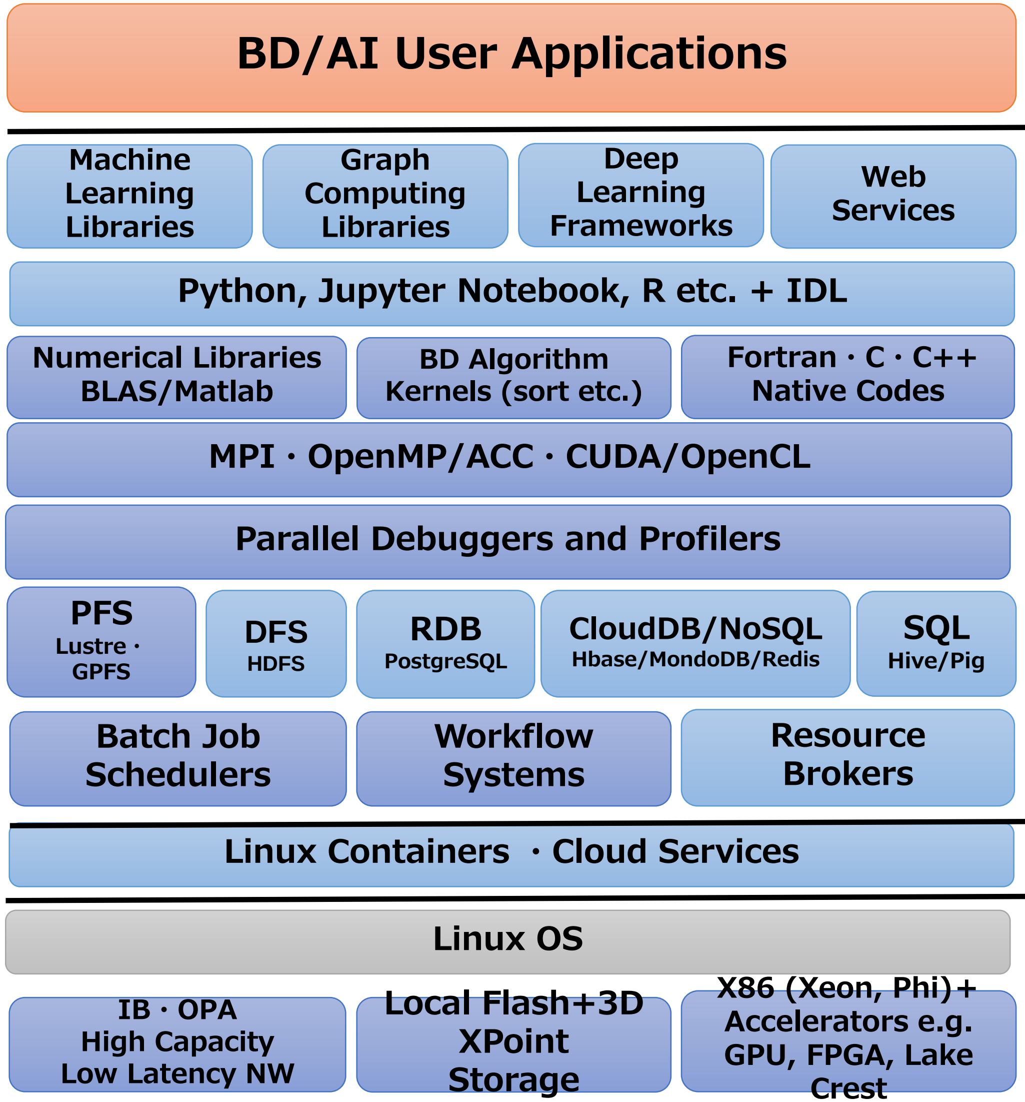


ABCI: Data Center



- Single Floor, inexpensive build
- Hard concrete floor 2 tonnes/m2 weight tolerance for racks and cooling pods
- Number of Racks
  - Initial: 90
  - Max: 144
- Power Capacity
  - 3.25 MW (MAX)
- Cooling Capacity
  - 3.2 MW (min in Summer)

Basic Requirements for AI Cloud System



Application

- ✓ Easy use of various ML/DL/Graph frameworks from Python, Jupyter Notebook, R, etc.
- ✓ Web-based applications and services provision

System Software

- ✓ HPC-oriented techniques for numerical libraries, BD Algorithm kernels, etc.
- ✓ Supporting long running jobs / workflow for DL
- ✓ Accelerated I/O and secure data access to large data sets
- ✓ User-customized environment based on Linux containers for easy deployment and reproducibility

OS

Hardware

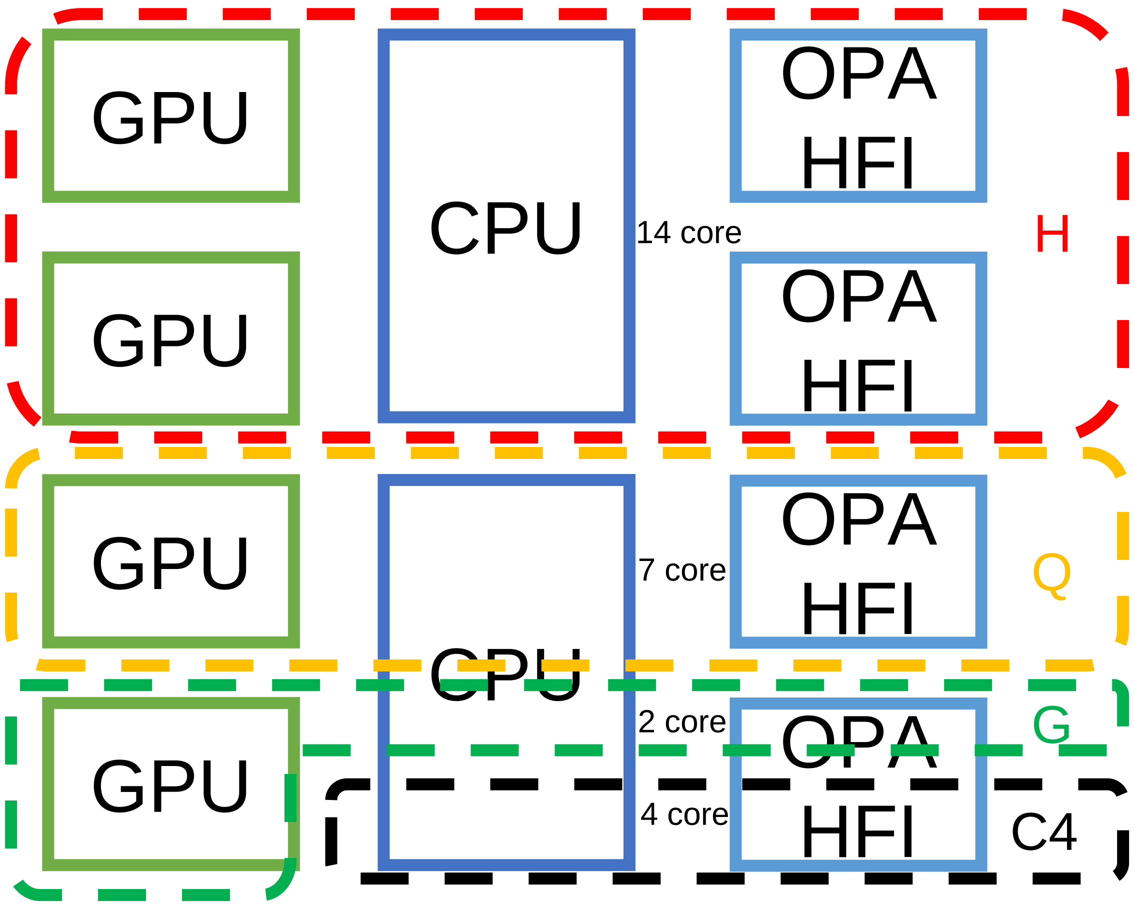
- ✓ Modern supercomputing facilities based on commodity components

Preliminary ABCI Software Stack

- Cent OS 7.3Red Hat Enterprise
- Univa Grid Engine
- Docker and other container engines
- Zabbix
- LifeKeeper
- Intel Parallel Studio XE Cluster Edition
- PGI Professional Edition
- NVIDIA CUDA SDK including NCCL
- Python, Ruby, R, Java, Scala, Lua, Perl, …
- Caffe, Caffe2, TensorFlow, Theano, Torch/PyTorch, CNTK, Mxnet, Chainer, Keras, …
- DDN GRIDScaler / GPFS, S3, Swift
- BeeGFS、BeeOND

- Others are in consideration
- How do you federate all the software consistently?

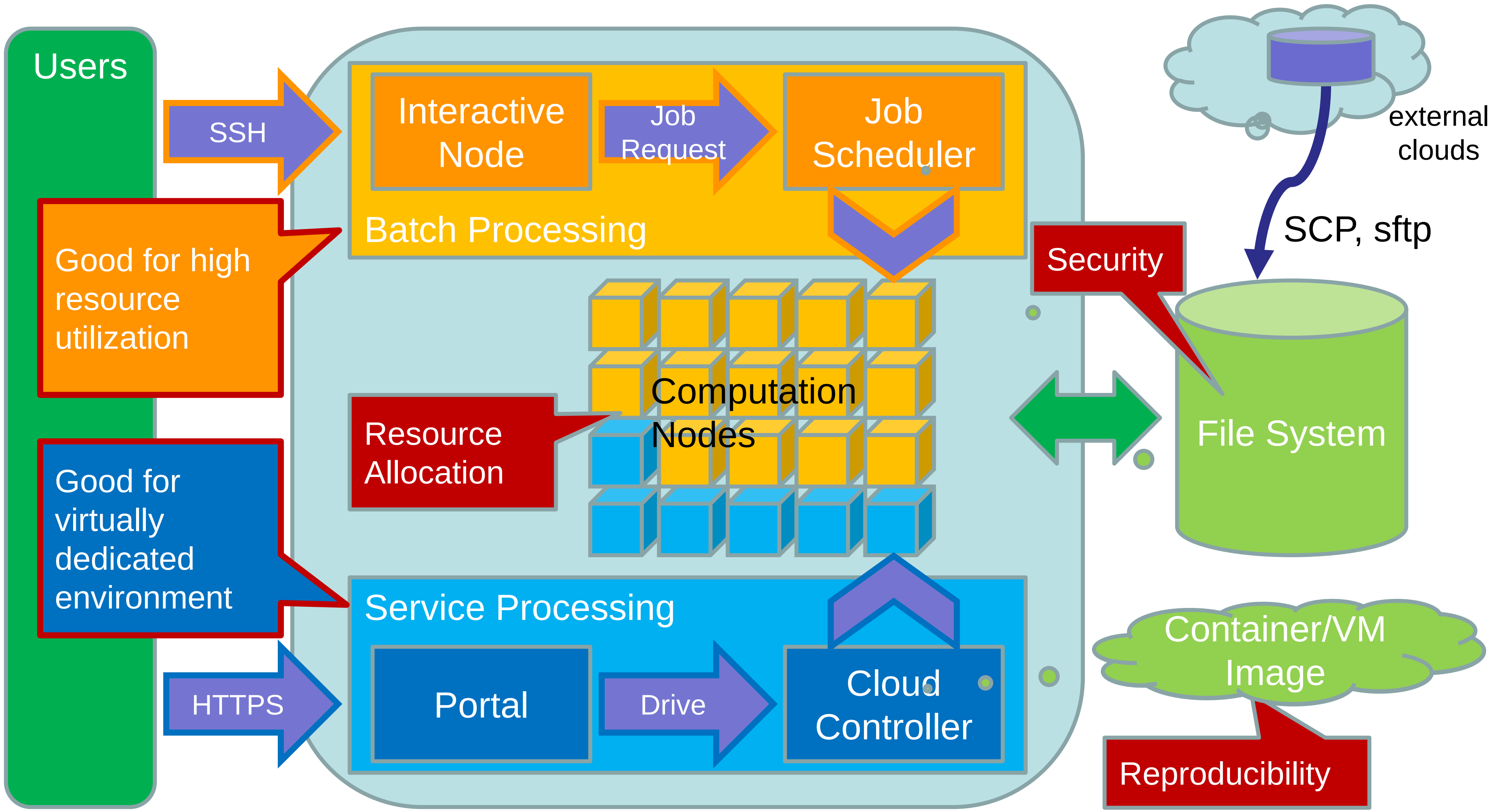
Resource Partitioning with Container in TSUBAME 3.0



Divide a compute node into some partitions in a hierarchical manner

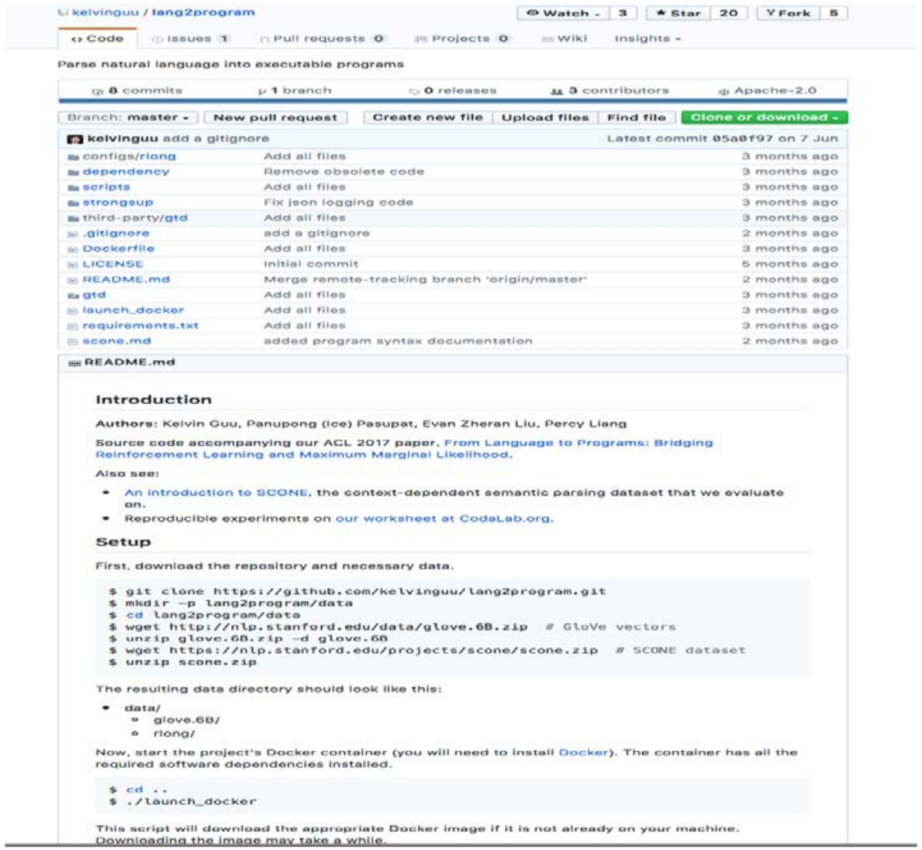
- F: Full node
- H: 1/2 node
- Q: 1/4 node
- G: 1 GPU + 2 CPU Core
- C4: 4 CPU Core
- C1: 1 CPU Core

ABCI Usage Model and Operations Challenges



Modern AI apps are Complex Beasts - need preconfigured containers -

- lang2program (referred in ACL2017)
  - https://github.com/kelvinguu/lang2program
  - Provided as a Dockerfile
  - Bunch of software needed to be run
    - Tensorflow, PostgresQL, Python Pip Packages, etc.
- Traditional large-scale HPC systems usually don't allow
  - OS updates => chaos esp. ISV SW
  - Docker => security bleach
  - Arbitrary/Voluntary installation of software => chaos w/ userland libraries
- We're developing an easy-to-manage and flexible-to-use platform for deploying AI apps as "modules"



kelvinguu / lang2program

Watch 3 | Star 20 | Fork 5

Code | Issues 1 | Pull requests 0 | Projects 0 | Wiki | Insights

Parse natural language into executable programs

8 commits | 1 branch | 0 releases | 3 contributors | Apache-2.0

Branch: master | New pull request | Create new file | Upload files | Find file | Clone or download

kelvinguu add a gitignore		Latest commit 05a0f97 on 7 Jun
configs/rlong	Add all files	3 months ago
dependency	Remove obsolete code	3 months ago
scripts	Add all files	3 months ago
strongsup	Fix json logging code	3 months ago
third-party/gtd	Add all files	3 months ago
.gitignore	add a gitignore	2 months ago
Dockerfile	Add all files	3 months ago
LICENSE	Initial commit	5 months ago
README.md	Merge remote-tracking branch 'origin/master'	2 months ago
gtd	Add all files	3 months ago
launch_docker	Add all files	3 months ago
requirements.txt	Add all files	3 months ago
scone.md	added program syntax documentation	2 months ago

README.md

Introduction

Authors: Kelvin Guu, Panupong (Ice) Pasupat, Evan Zheran Liu, Percy Liang

Source code accompanying our ACL 2017 paper, From Language to Programs: Bridging Reinforcement Learning and Maximum Marginal Likelihood.

Also see:

- An introduction to SCONE, the context-dependent semantic parsing dataset that we evaluate on.
- Reproducible experiments on our worksheet at CodaLab.org.

Setup

First, download the repository and necessary data.

```
$ git clone https://github.com/kelvinguu/lang2program.git
$ mkdir -p lang2program/data
$ cd lang2program/data
$ wget http://nlp.stanford.edu/data/glove.6B.zip  # GloVe vectors
$ unzip glove.6B.zip -d glove.6B
$ wget https://nlp.stanford.edu/projects/scone/scone.zip  # SCONE dataset
$ unzip scone.zip
```

The resulting data directory should look like this:

- data/
  - glove.6B/
  - rlong/

Now, start the project's Docker container (you will need to install Docker). The container has all the required software dependencies installed.

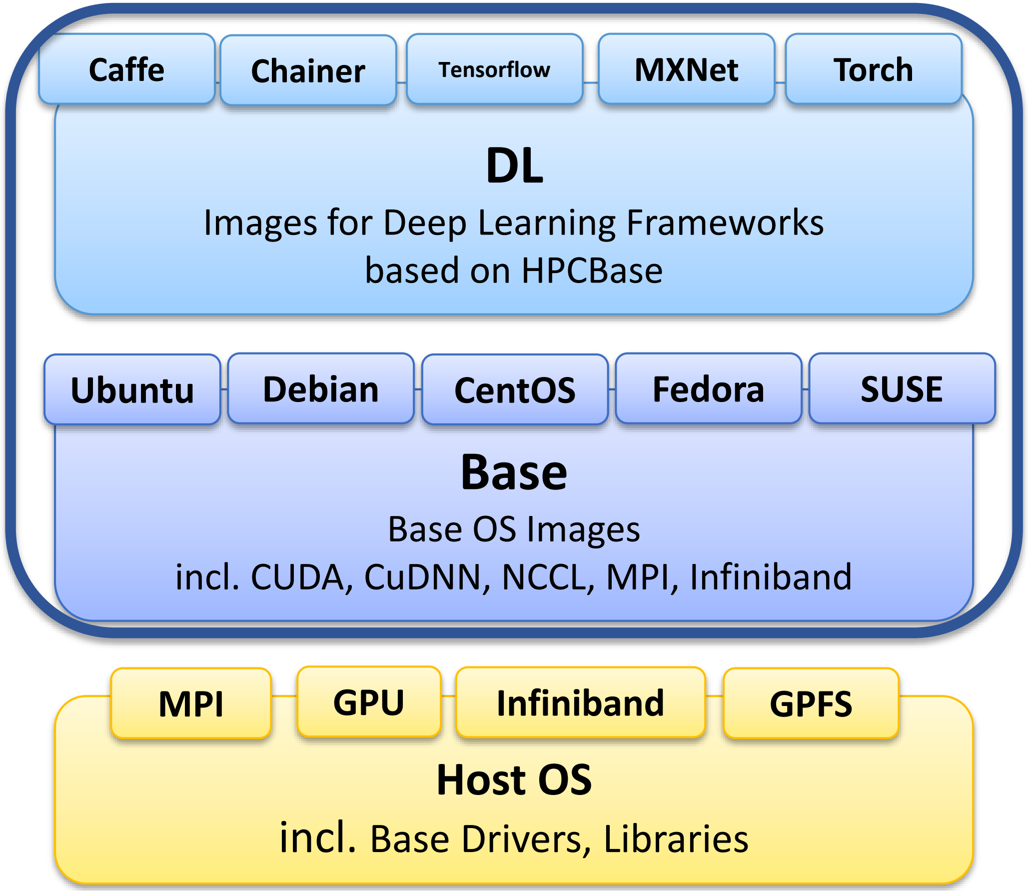
```
$ cd ..
$ ./launch_docker
```

This script will download the appropriate Docker image if it is not already on your machine. Downloading the image may take a while.

AI Frameworks & modules

- We employ LBL Singularity or Shifter as the basis of our AI platform
- Only introduces minimum set of software into Host OS
  - Base Drivers, Libraries
- "Base" repository provides customized OS images as "modules" including:
  - CUDA, cuDNN, NCCL, MPI, etc.
- "DL" repository provides DL frameworks and apps which extend "Base" images
- Our public repositories

  https://github.com/aistairc/aimodules
  https://hub.docker.com/r/aistairc/{base,dl}



Multiple Container Support Mechanisms in ABCI

Traditional

User-level
(user privilege, e.g., Singularity, Shifter)

- **Application** single binary executable, MPI
- **DL module**
- **Userland Base Module**
- Batch-Q System + LW container runtime
- **Host OS**

System-level
(root privilege, e.g., Docker)

- **Application** Everything incl. GUI, Linux services
- **DL module**
- **Userland Base Module**
- Docker runtime
- **Host OS**

ABCI

Bare-metal container
(direct deployment AIST technology)

- **Application** Everything incl. GUI, Linux services
- **DL module**
- **Userland Base Module**
- Various runtimes
- **Host OS**

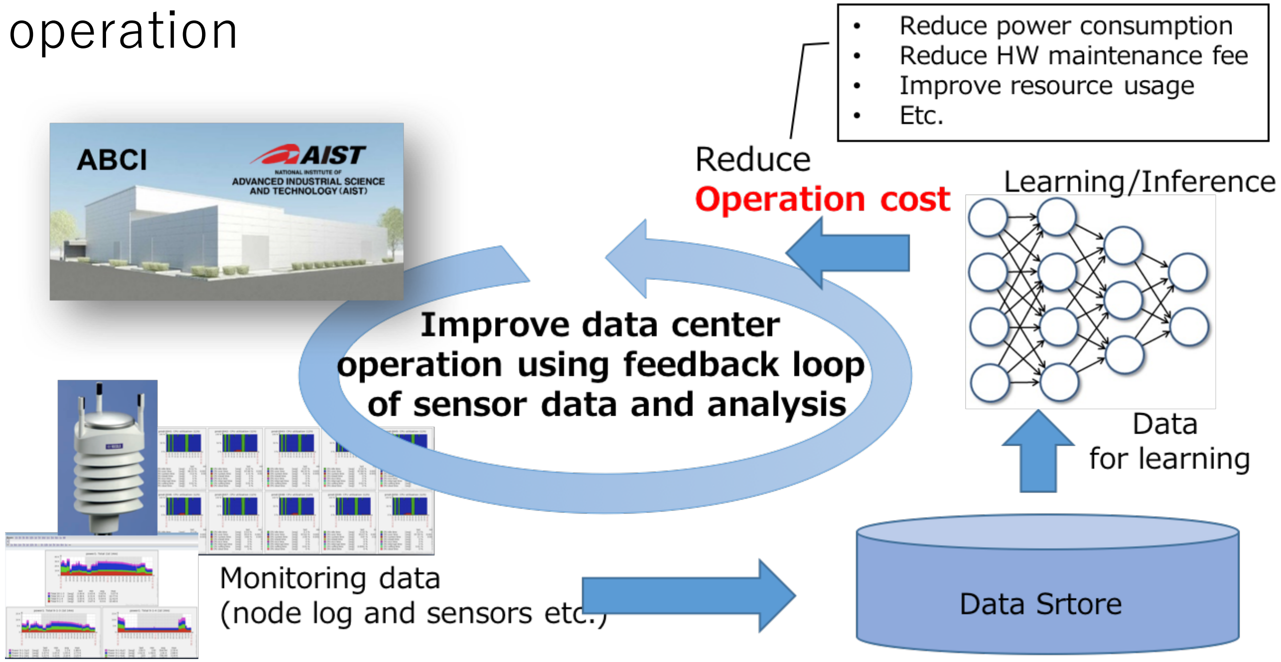
Double Deckered (Under Plan)

- **Application** single binary executable, MPI
- **DL module**
- Batch-Q System + LW container runtime
- **Userland Base Module**
- Docker runtime
- **Host OS**

Smart Data Center Operation for ABCI
(NEDO Project 2017-)

- Started to develop a system that achieves a self-sustainable operation and reduce operation cost of data center, especially for ABCI system
  - Data storage for storing sensor data from node, cooling system, etc.
  - ML/DL algorithms to analyze the data and model data center behavior
    - Reduce power consumption, detect errors, etc.
  - Apply the algorithms to improve operation

- Current status
  - Started from Aug. 2017
  - Designing/developing sensor data collector and its storage



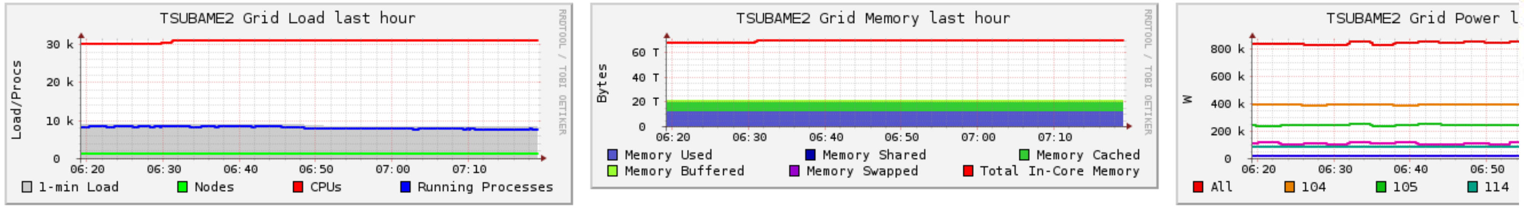
TSUBAME2 Grid Report for Tue, 13 Jun 2017 07:19:06 +0900

Last [hour] BasedOn ◉ *Now* ○ [ ] [0:00] [apply] Sorted [descending]

Grid > [--Choose a Source] > [--Input a Node] [go]

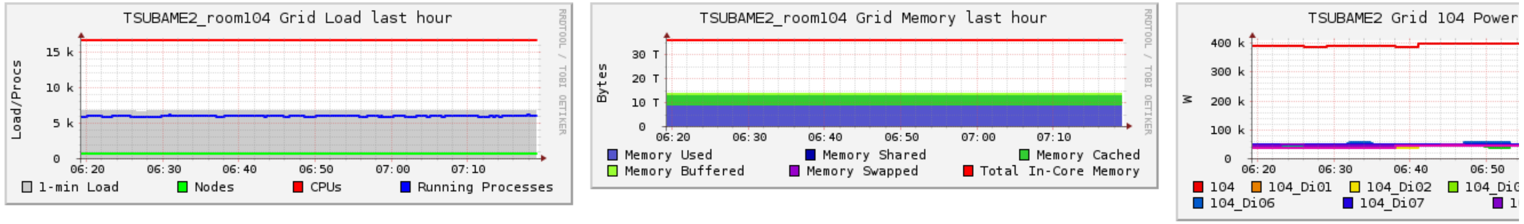
rid (4 sources) (tree view)

30956
1275
13



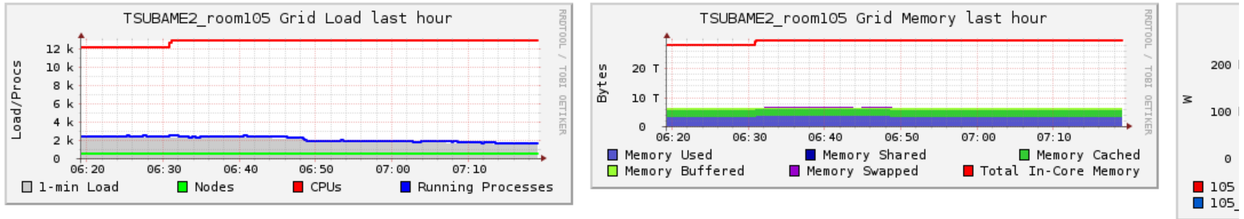
oom104 Grid (tree view)

16584
591
9



oom105 Grid (tree view)

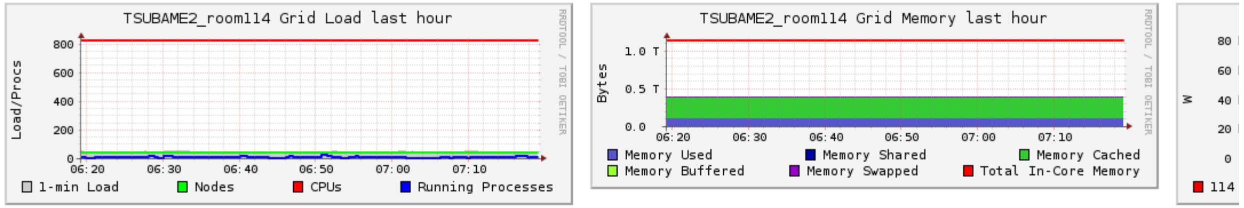
12864
536
0



oom114 Grid (tree view)

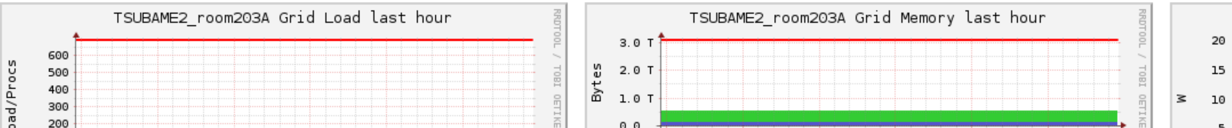
い、2016年6月7日以降は203B室の稼働状況をroom114として表示しています。

820
36
2



oom203A Grid (tree view)

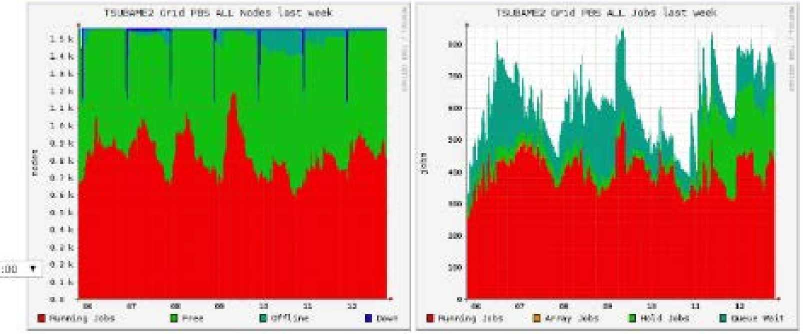
688
12
2



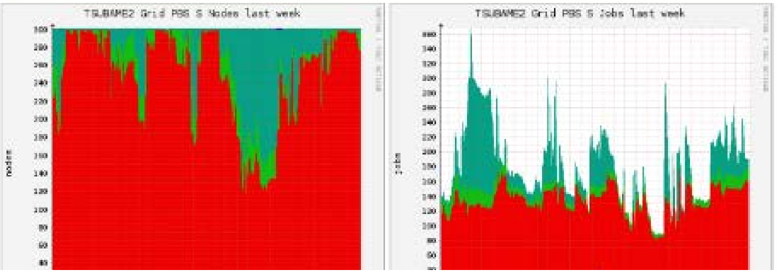
TSUBAME 2.5 ALL
Queue S
Queue S96
Queue G
Queue V
Queue U
Queue L128
Queue L128F
Queue L256
Queue L512
Queue H
Queue X

Last [week]
basedon
◉ now
○ [ ] [0:00]
[apply]

TSUBAME 2.5 All



Queue S



Service List

service	assigned nodes			max nodes	running jobs		user
S	93%	279 / 297 nodes		300 nodes	84%	162 / 192 jobs	
S96	46%	18 / 39 nodes		39 nodes	100%	11 / 11 jobs	
G	83%	356 / 427 nodes		435 nodes	23%	84 / 359 jobs	
V	20%	75 / 371 nodes		385 nodes	99%	145 / 146 jobs	
U	11%	2 / 18 nodes		195 nodes	100%	2 / 2 jobs	
L256	0%	0 / 8 nodes		8 nodes	0%	0 / 0 jobs	
L512	0%	0 / 2 nodes		2 nodes	0%	0 / 0 jobs	
X	11%	48 / 419 nodes		420 nodes	100%	21 / 21 jobs	
ALL	49%	780 / 1581 nodes			58%	425 / 731 jobs	

Node List of G Service

* *These parameters are collected from the job scheduler. Click host links to see the actual load status.*

service	host	assigned CPU	assigned GPU	assigned MEM	running jobs	users	used scr	status
G	t2a001121	1 / 8 CPUs	3 / 3 GPUs	10.0 / 25.0 GB	1	1	5%	enable
G	t2a001122	1 / 8 CPUs	3 / 3 GPUs	5.0 / 25.0 GB	1	1	6%	enable

Cutting Edge Research AI Infrastructures in Japan
Accelerating BD/AI with HPC
(and my effort to design & build them)

In Production

x5.8

Oct. 2015
TSUBAME-KFC/DL
(Tokyo Tech./NEC)
1.4 AI-PF(Petaflops)

In Production

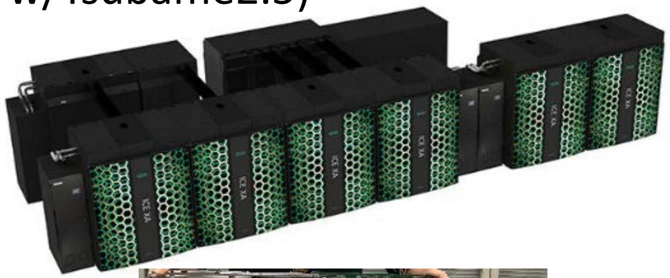
Mar. 2017
AIST AI Cloud
(AIST-AIRC/NEC)
8.2 AI-PF

x5.8

Mar. 2017
AI Supercomputer
Riken AIP/Fujitsu
4.1 AI-PF

In Production

Aug. 2017
TSUBAME3.0 (Tokyo Tech./HPE)
47.2 AI-PF (65.8 AI-PF w/Tsubame2.5)



x11.7

In Construction

1H 2018
ABCI (AIST-AIRC)
550 AI-PF

Draft RFC out
IDC under construction

X4~6?

~x400 in 3 years

Built/funded

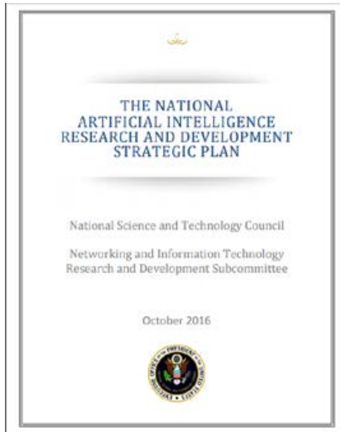
1H 2019?
“**ExaAI**”
~2~3 AI-ExaFlop

Undergoing Engineering Study

Also Post-K Multi AI-Exaflops

still under plans

R&D Investments into world leading AI/BD HW & SW & Algorithms and their co-design for cutting edge Infrastructure absolutely necessary (just as is with Japan Post-K and US ECP in HPC)


THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN

National Science and Technology Council

Networking and Information Technology Research and Development Subcommittee

October 2016

Fujitsu Deep Learning Processor (DLU™)

FUJITSU

Supercomputer K technologies

FY2018～

DLU ™

(Deep Learning Unit)

DLU™ features

- **Architecture designed for Deep Learning**
- **High performance HBM2 memory**
- **Low power design**

→ **Goal: 10x Performance/Watt compared to others**

- **Massively parallel : Apply supercomputer interconnect technology**

→ **Ability to handle large scale neural networks**

→ **TOFU Network derivative for massive scaling**

"Exascale" AI possible in 1H2019

Designed for Scalable Learning, technically superior to Google TPU2