Extreme Big Data (EBD)

Convergence of Extreme Computing and Big Data Technologies

Global Scientific Information and Computing Center
Tokyo Institute of Technology

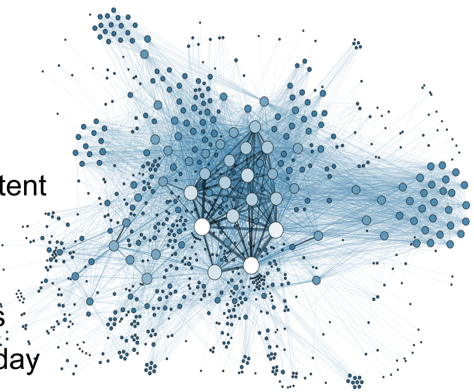
Hitoshi Sato

Big Data Examples

Rates and Volumes are extremely immense

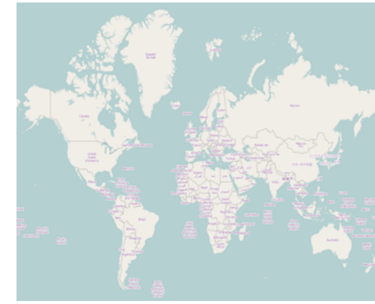
Social NW

- Facebook
  - ~1 billion users
  - Average 130 friends
  - 30 billion pieces of content shared per month
- Twitter
  - 500 million active users
  - 340 million tweets per day
- Internet
  - 300 million new websites per year
  - 48 hours of video to YouTube per minute
  - 30,000 YouTube videos played per second



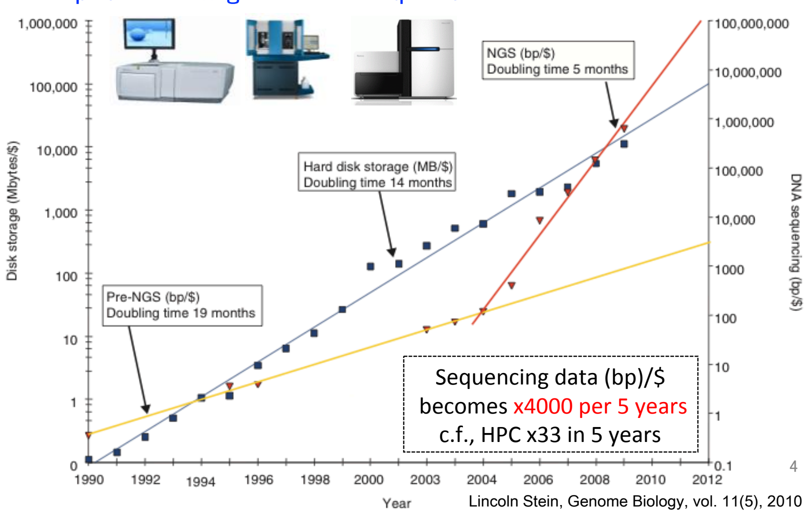
Social Simulation

- Applications
  - Target Area: Planet (Open Street Map)
  - 7 billion people
- Input Data
  - Road Network for Planet: 300GB (XML)
  - Trip data for 7 billion people 10KB (1trip) x 7 billion = 70TB
  - Real-Time Streaming Data (e.g., Social sensor, physical data)
- Simulated Output for 1 Iteration
  - 700TB



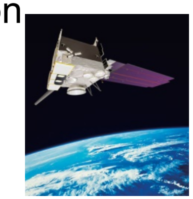
Genomics

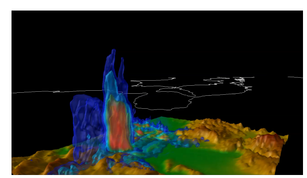
Impact of new generation sequencers

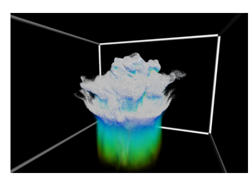


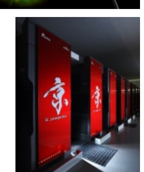
Lincoln Stein, Genome Biology, vol. 11(5), 2010

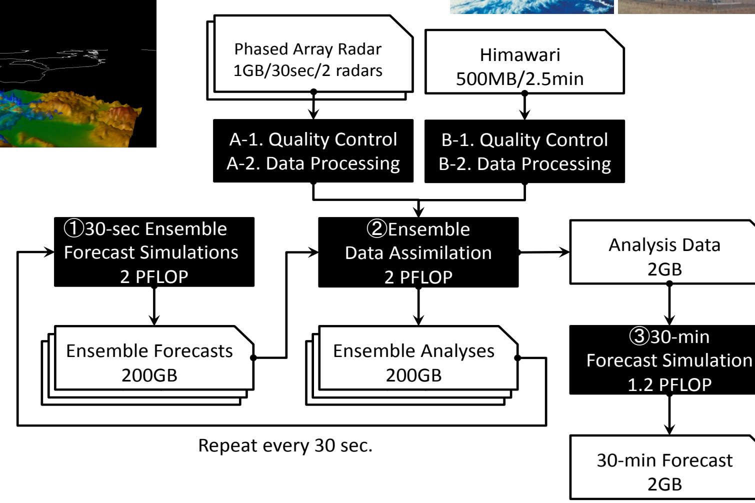
Weather











Big Data Examples

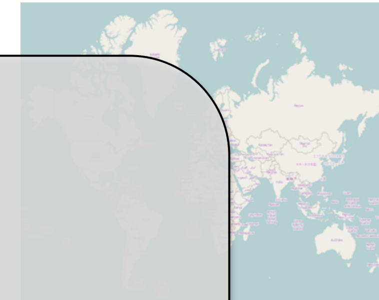
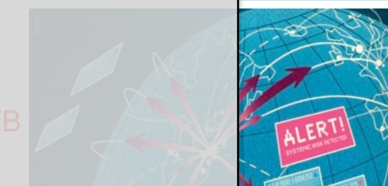
Rates and Volumes are extremely immense

Social NW

- Facebook
  - Average 130 friends
  - 30 billion pieces of content shared per month
- Twitter
  - 500 million active users
  - 340 million tweets per day
- Internet
  - 300 million new websites per year
  - 48 hours of video to YouTube per minute
  - 30,000 YouTube videos played per second

Social Simulation

- Applications
  - Target Area: Planet (Open Street Map)
  - 7 billion people
- Input Data
  - Road Network for Planet: 300GB (XML)
  - Trip data for 7 billion people 10KB (1trip) x 7 billion = 70TB
  - Real-Time Streaming Data (e.g., Social sensor, physical data)
- Simulated Output for 1 Iteration

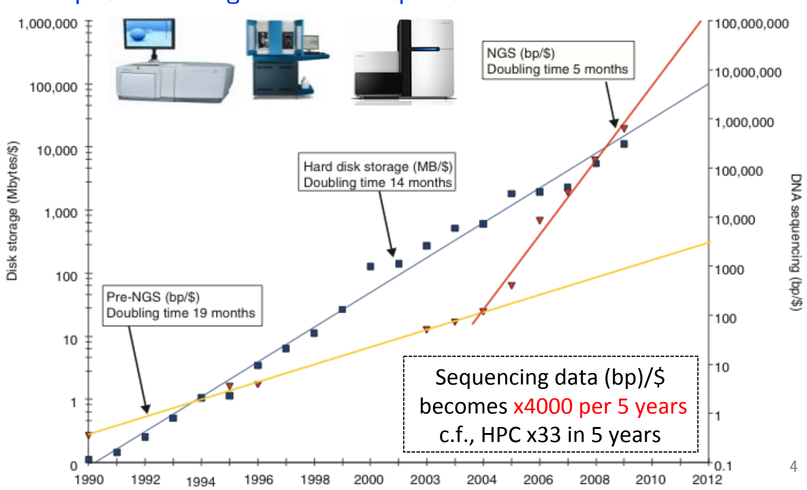



Future "**Extreme Big Data**"

- NOT mining Tbytes Silo Data
- Peta~Zetabytes of Data
- Ultra High-BW Data Stream
- Highly Unstructured, Irregular
- Complex correlations between data from multiple sources
- Extreme Capacity, Bandwidth, Compute All Required

Genomics

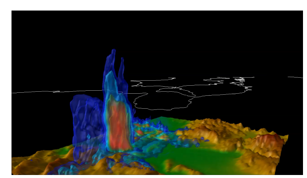
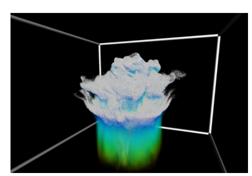
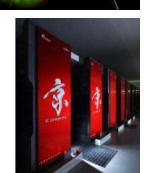
Impact of new generation sequencers

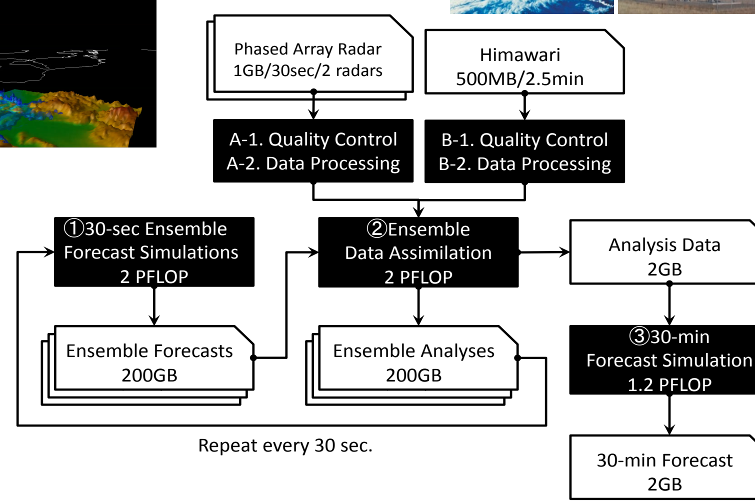


Lincoln Stein, Genome Biology, vol. 11(5), 2010

Weather

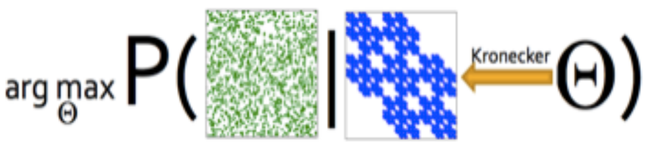




Graph500 "Big Data" Benchmark

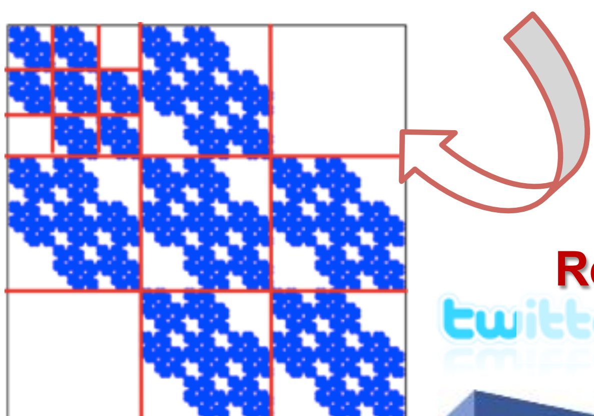
Kronecker graph BSP Problem



A: 0.57, B: 0.19
C: 0.19, D: 0.05

$$\begin{array}{|c|c|c|}\hline 1 & 1 & 0 \\ \hline 1 & 1 & 1 \\ \hline 0 & 1 & 1 \\ \hline\end{array}$$

G_1



G_4 adjacency matrix

November 15, 2010

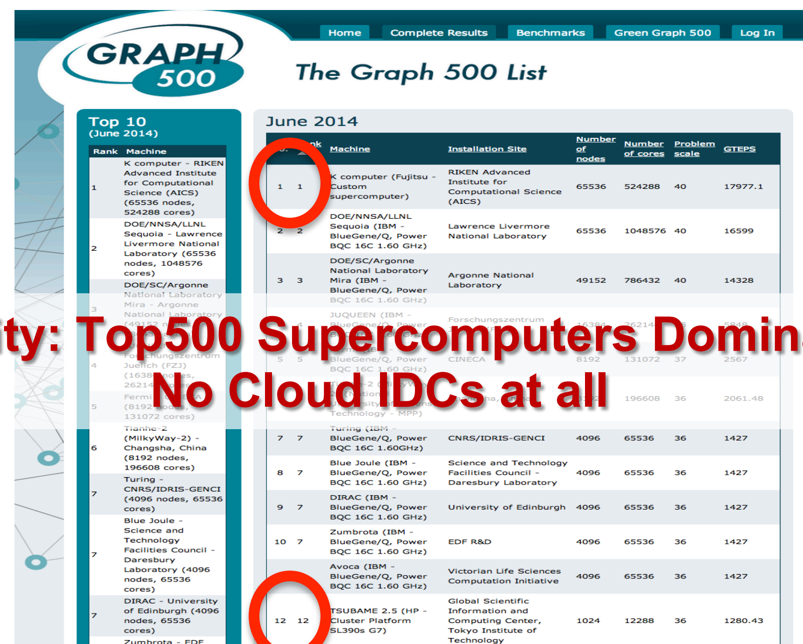
Graph 500 Takes Aim at a New Kind of HPC

Richard Murphy (Sandia NL => Micron)

" I expect that this ranking may at times look very different from the TOP500 list. Cloud architectures will almost certainly dominate a major chunk of part of the list."

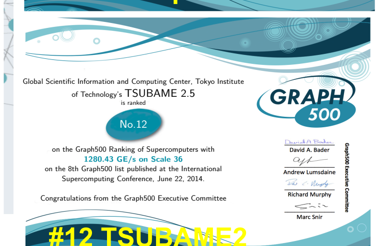
The 8th Graph500 List (June2014): K Computer #1, TSUBAME2 #12

Koji Ueno, Tokyo Institute of Technology/RIKEN AICS



No.	Rank	Machine	Installation Site	Number of nodes	Number of cores	Problem scale	GTEPS
1	1	K computer (Fujitsu - Custom supercomputer)	RIKEN Advanced Institute for Computational Science (AICS)	65536	524288	40	17977.1
2	2	DOE/NNSA/LLNL Sequoia (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Lawrence Livermore National Laboratory	65536	1048576	40	16599
3	3	DOE/SC/Argonne National Laboratory Mira (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Argonne National Laboratory	49152	786432	40	14328
[illegible]	4	JUQUEEN (IBM - BlueGene/Q, Power [illegible])	Forschungszentrum [illegible]	16384	[illegible]	[illegible]	5848
5	5	[illegible] BlueGene/Q, Power BQC 16C 1.60 GHz)	CINECA	8192	131072	37	2567
[illegible]	[illegible]	[illegible] Technology - MPP)	[illegible]	[illegible]	196608	36	2061.48
7	7	Turing (IBM - BlueGene/Q, Power BQC 16C 1.60GHz)	CNRS/IDRIS-GENCI	4096	65536	36	1427
8	7	Blue Joule (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Science and Technology Facilities Council - Daresbury Laboratory	4096	65536	36	1427
9	7	DIRAC (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	University of Edinburgh	4096	65536	36	1427
10	7	Zumbrota (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	EDF R&D	4096	65536	36	1427
		Avoca (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Victorian Life Sciences Computation Initiative	4096	65536	36	1427
12	12	TSUBAME 2.5 (HP - Cluster Platform SL390s G7)	Global Scientific Information and Computing Center, Tokyo Institute of Technology	1024	12288	36	1280.43

Reality: Top500 Supercomputers Dominate
No Cloud IDCs at all



GPU-based Heterogeneous Supercomputer: TSUABME2 (TokyoTech)

17PF(SFP)/5.7PF(DFP) #13 on Top500 (Jun 2014)

System

1408 nodes, 44 racks

CPU: Intel Xeon X5670 (6 cores) × 3

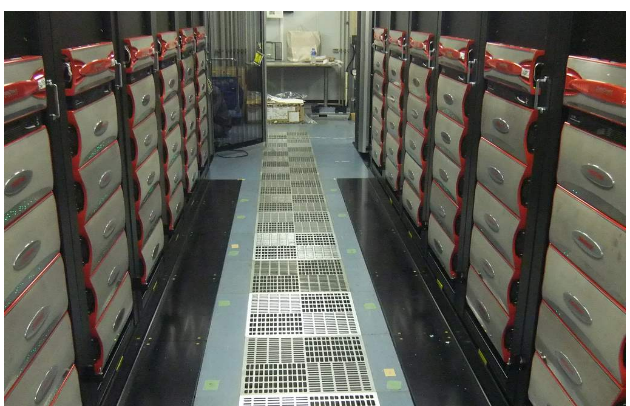
GPU: NVIDIA Tesla K20X × 3

MEM: 58 GB DDR3 1333MHz

SSD: 60 GB × 2

Network

4× QDR Full bisection Optical Infiniband



Storage

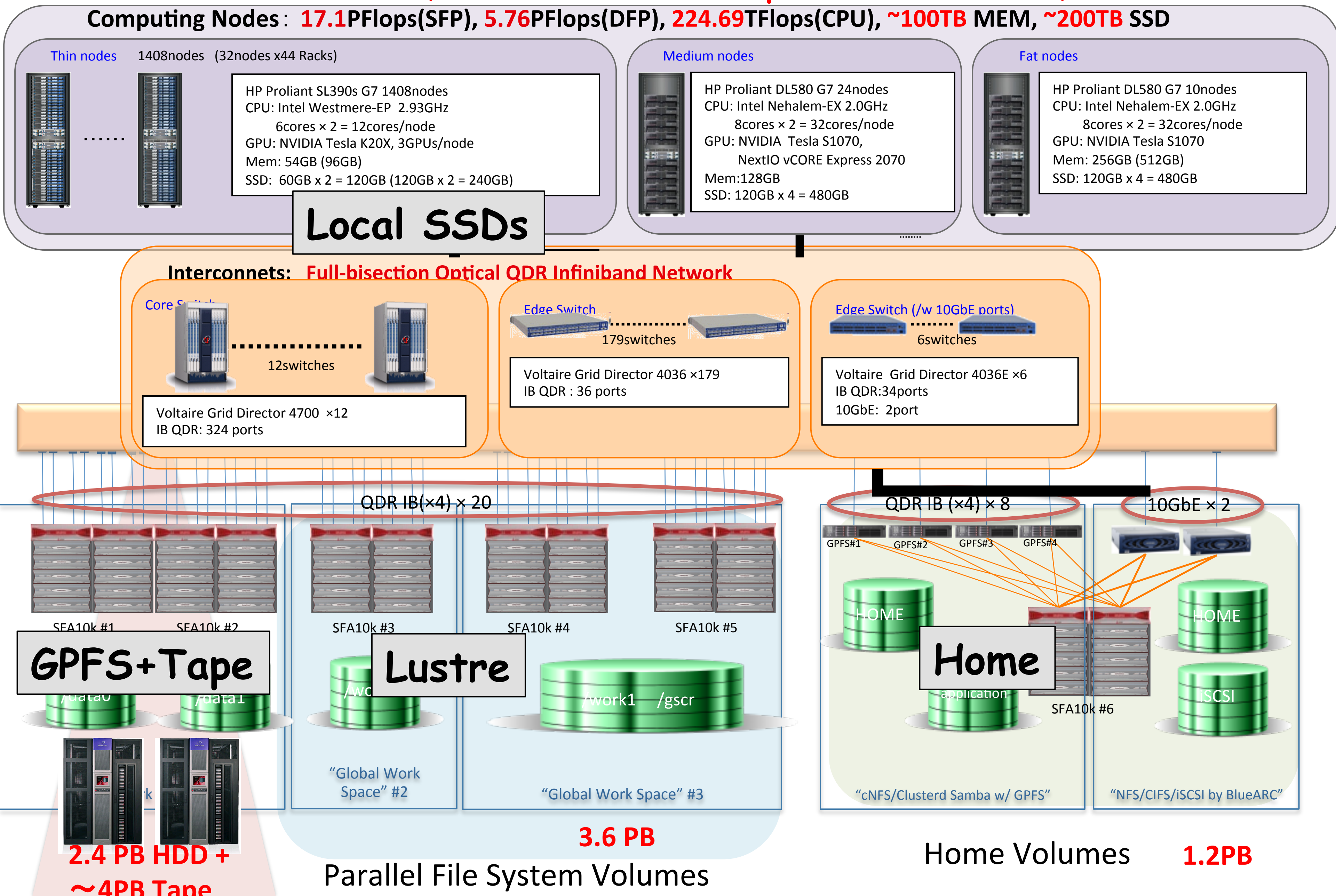
SSD: 200TB

HDD: 7PB (Lustre, GPFS)

Tape: 4PB (SorageTek SL8500× 2)

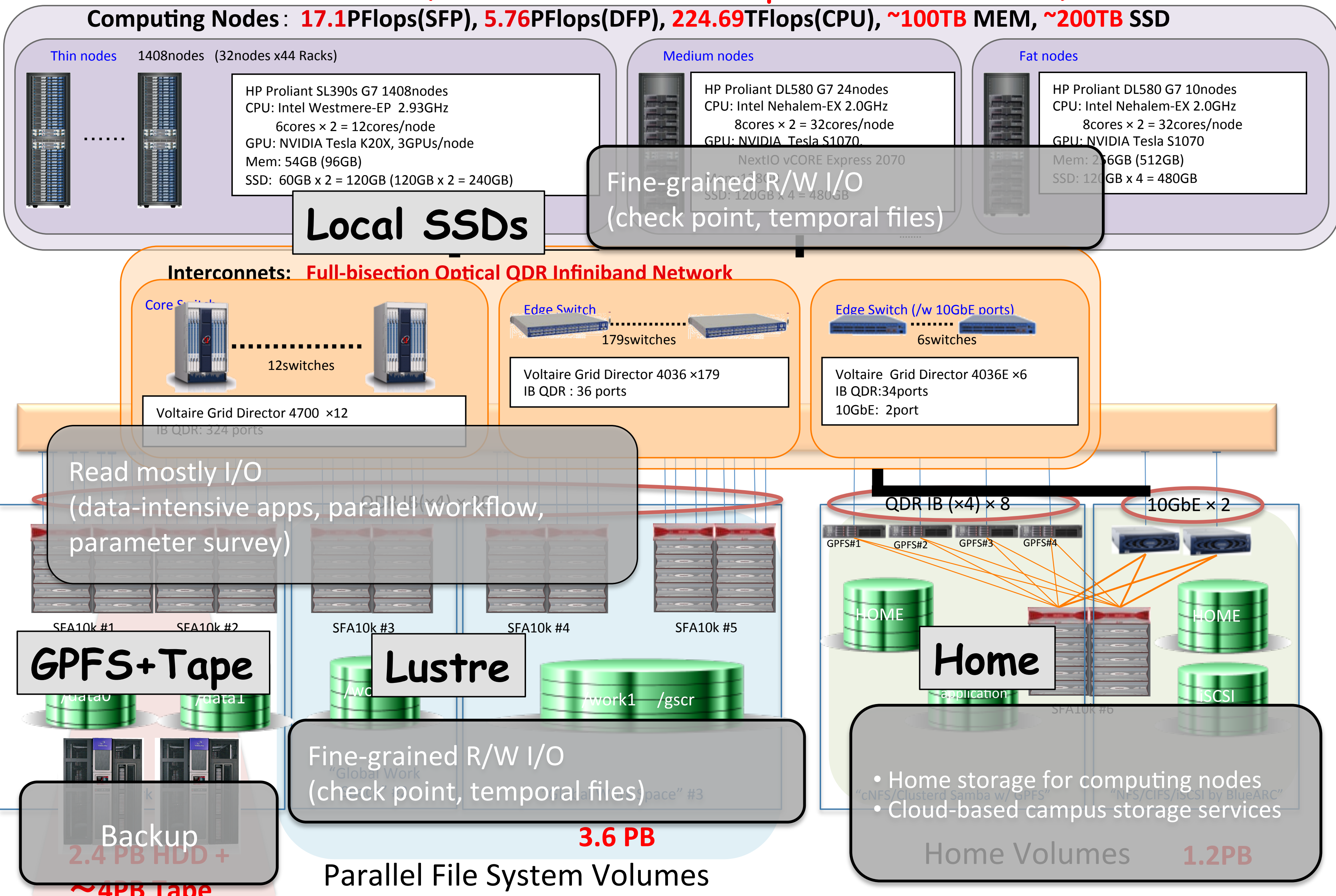
TSUBAME2 System Overview

11PB (7PB HDD, 4PB Tape, 200TB SSD)

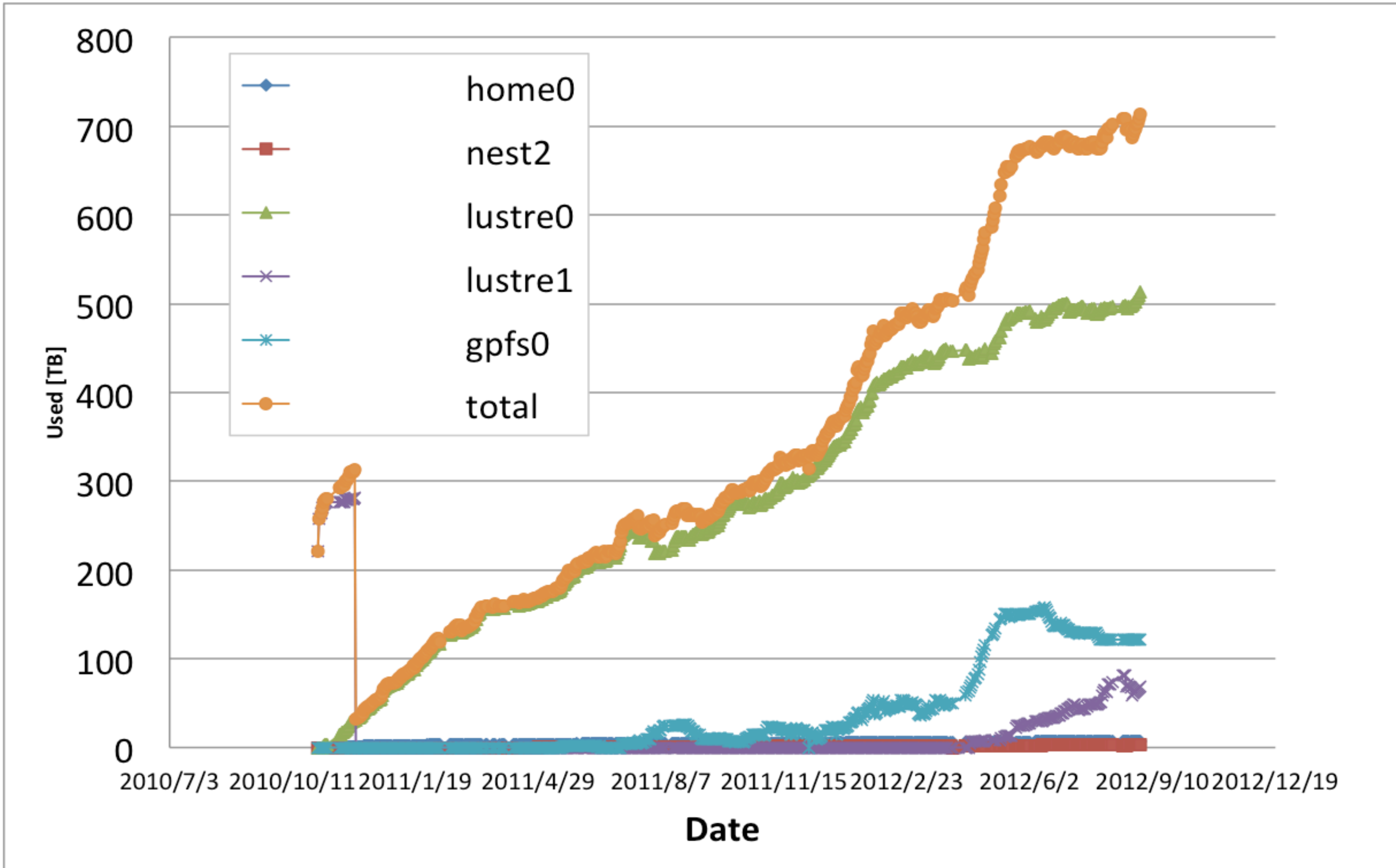


TSUBAME2 System Overview

11PB (7PB HDD, 4PB Tape, 200TB SSD)



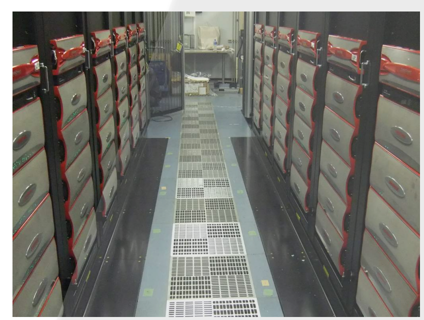
TSUBAME2 Storage Usage Since Nov. 2010



TSUBAME2.5 as a BigData Infrastructure

- Huge Memory Volumes per Node
  – 58GB per node × 1408 nodes

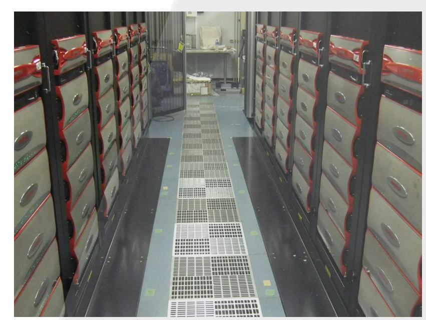
- Fat Tree Dual-Rail QDR InfiniBand
  – 200Tbps of full bisection bandwidth



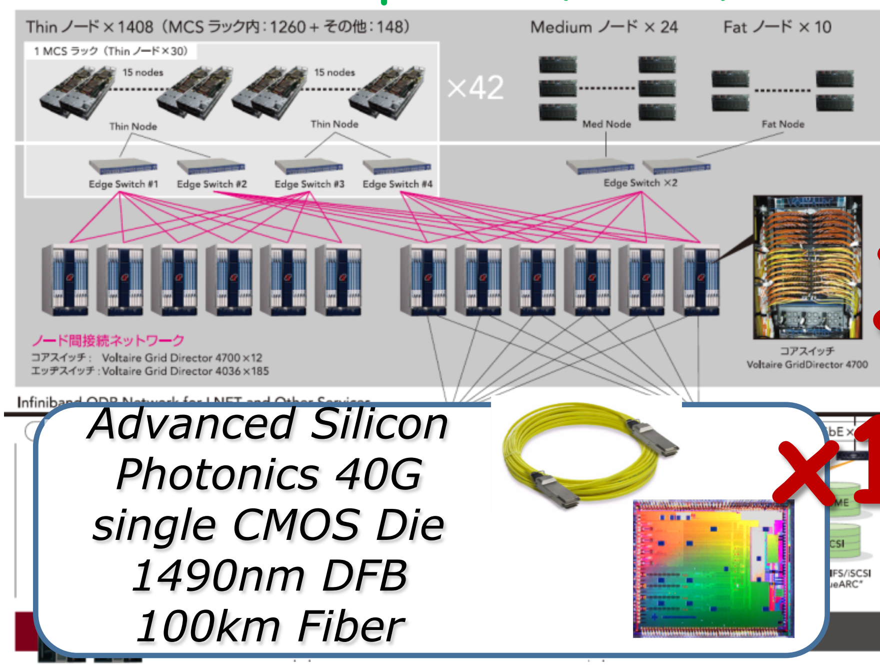
- Large-scale Storage Systems
  – Lustre, GPFS
  – 7PB of HDDs, 4PB of Tapes

TSUBAME2.5 as a BigData Infrastructure

- GPU-based Many-core Accelerator
  - NVIDIA Tesla K20X 4224 cards
- Huge Memory Volumes per Node
  - 58GB per node × 1408 nodes
- Fat Tree Dual-Rail QDR InfiniBand
  - 200Tbps of full bisection bandwidth
- Local SSD devices per node
  - 200TB in total
- Large-scale Storage Systems
  - Lustre, GPFS
  - 7PB of HDDs, 4PB of Tapes

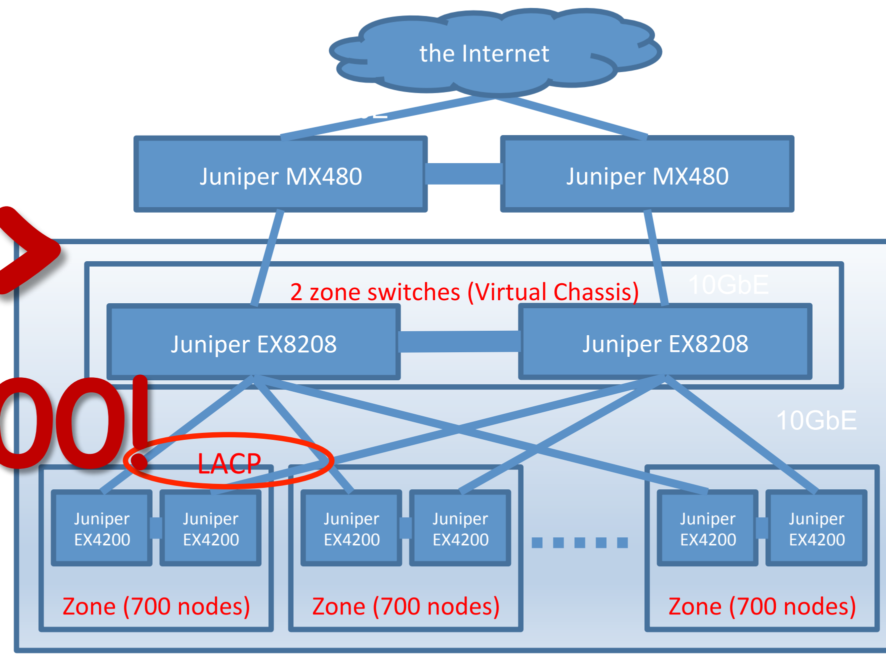


Supercomputer Tokyo Tech.
Tsubame 2.0
#4 Top500 (2010)



x1000!

A Major Northern Japanese Cloud Datacenter (2013)

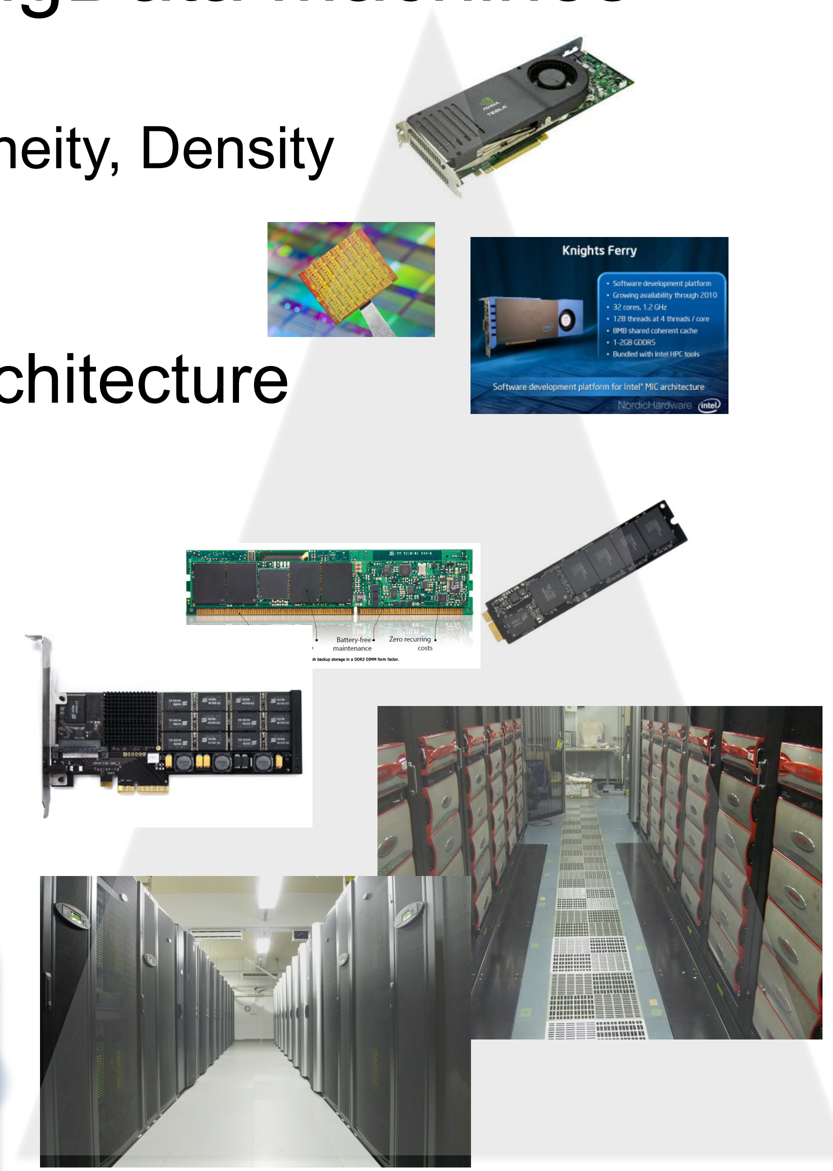


~1500 nodes compute & storage
Full Bisection Multi-Rail
Optical Network
Injection 80GBps/Node
Bisection 220Terabps

8 zones, Total 5600 nodes,
Injection 1GBps/Node
Bisection 160Gigabps

Towards Extreme-scale Supercomputers and BigData Machines

- Computation
  - Increase in Parallelism, Heterogeneity, Density
    - Multi-core, Many-core processors
    - Heterogeneous processors
- Hierarchial Memory/Storage Architecture
  - NVM (Non-Volatile Memory), SCM (Storage Class Memory)
    - FLASH, PCM, STT-MRAM, ReRAM, HMC, etc.
  - Next-gen HDDs (SMR), Tapes (LTFS)



Problems

Algorithm
Locality
Scalability
Power
Network
Storage Hierarchy
Heterogeneity
FT
Productivity
I/O

Extreme Big Data (EBD)

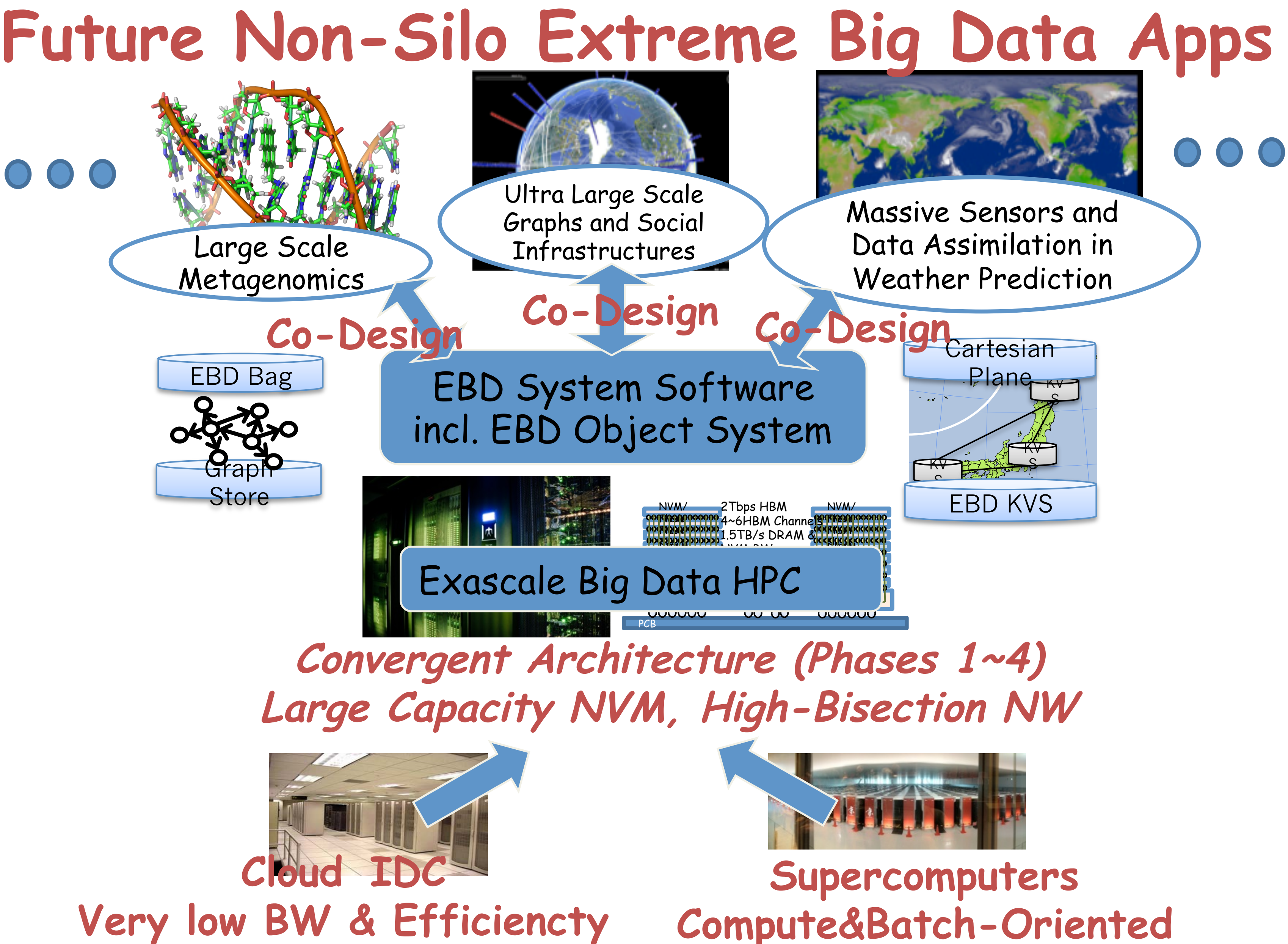
Next Generation Big Data Infrastructure Technologies Towards Yottabyte/Year

Principal Investigator
Satoshi Matsuoka

Global Scientific Information and Computing Center
Tokyo Institute of Technology

2014/11/05 JST CREST Big Data Symposium

EBE Research Scheme



Extreme Big Data (EBD) Team

Co-Design EHPC and EBD Apps

- **Satoshi Matsuoka (PI), Toshio Endo, Hitoshi Sato (Tokyo Tech.) (Tasks 1, 3, 4, 6)**



- **Osamu Tatebe (Univ. Tsukuba) (Tasks 2, 3)**

- **Michihiro Koibuchi (NII) (Tasks 1, 2)**

- **Yutaka Akiyama, Ken Kurokawa (Tokyo Tech, 5-1)**


- **Toyotaro Suzumura (IBM Lab, 5-2)**

- **Takemasa Miyoshi (Riken AICS, 5-3)**