



東北大学

# News Updates: SX-ACE's Operations and Applications Development for the Future

**Hiroaki Kobayashi**

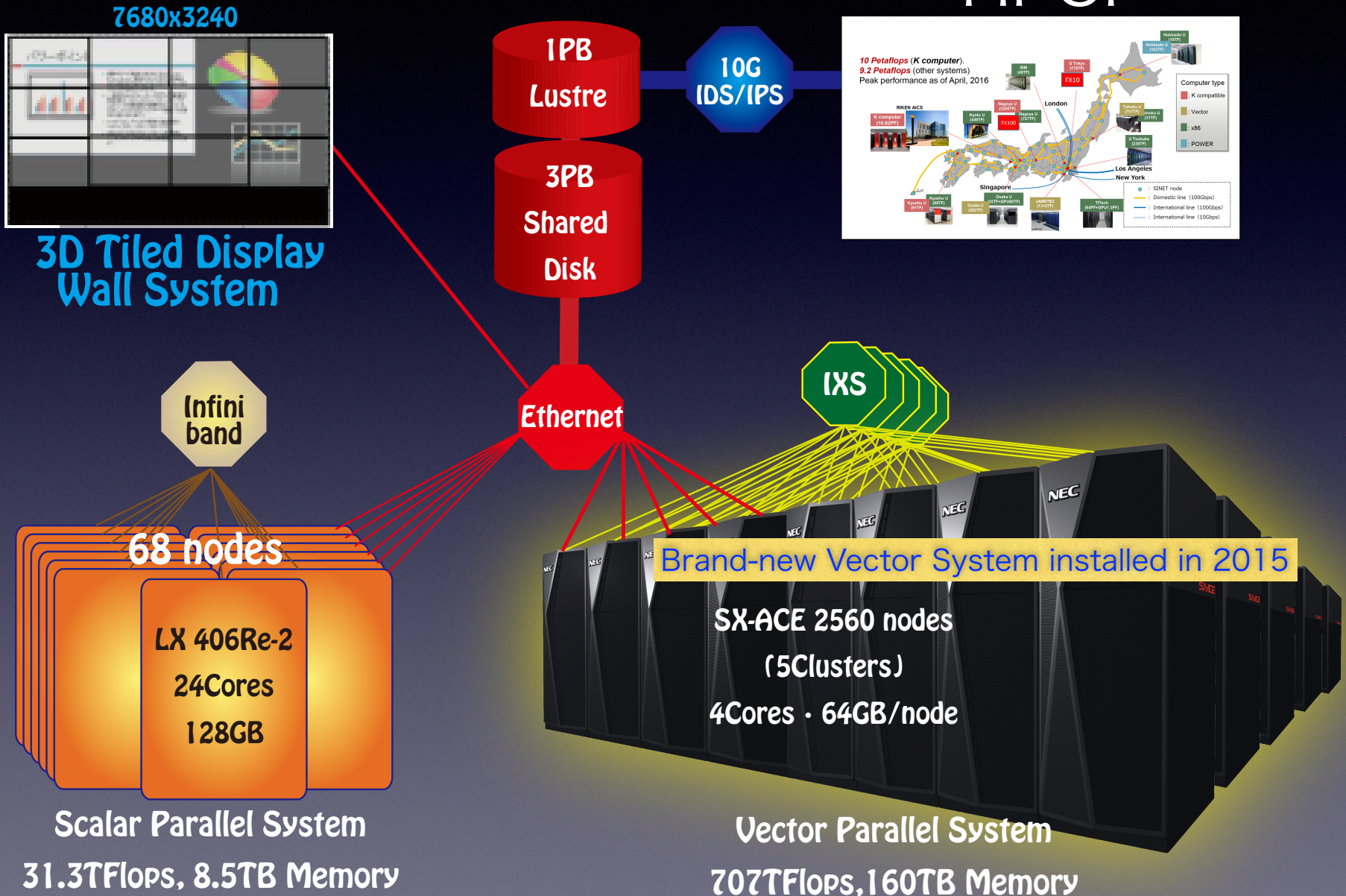
Deputy Director for the HPC Strategy of Cyberscience Center,  
Professor of Graduate School of Information Sciences,  
Tohoku University  
koba@tohoku.ac.jp

24th WSSP  
Dec. 5-6, 2016





# Tohoku Univ.'s New Supercomputer System (2015.2.20~) HPCI

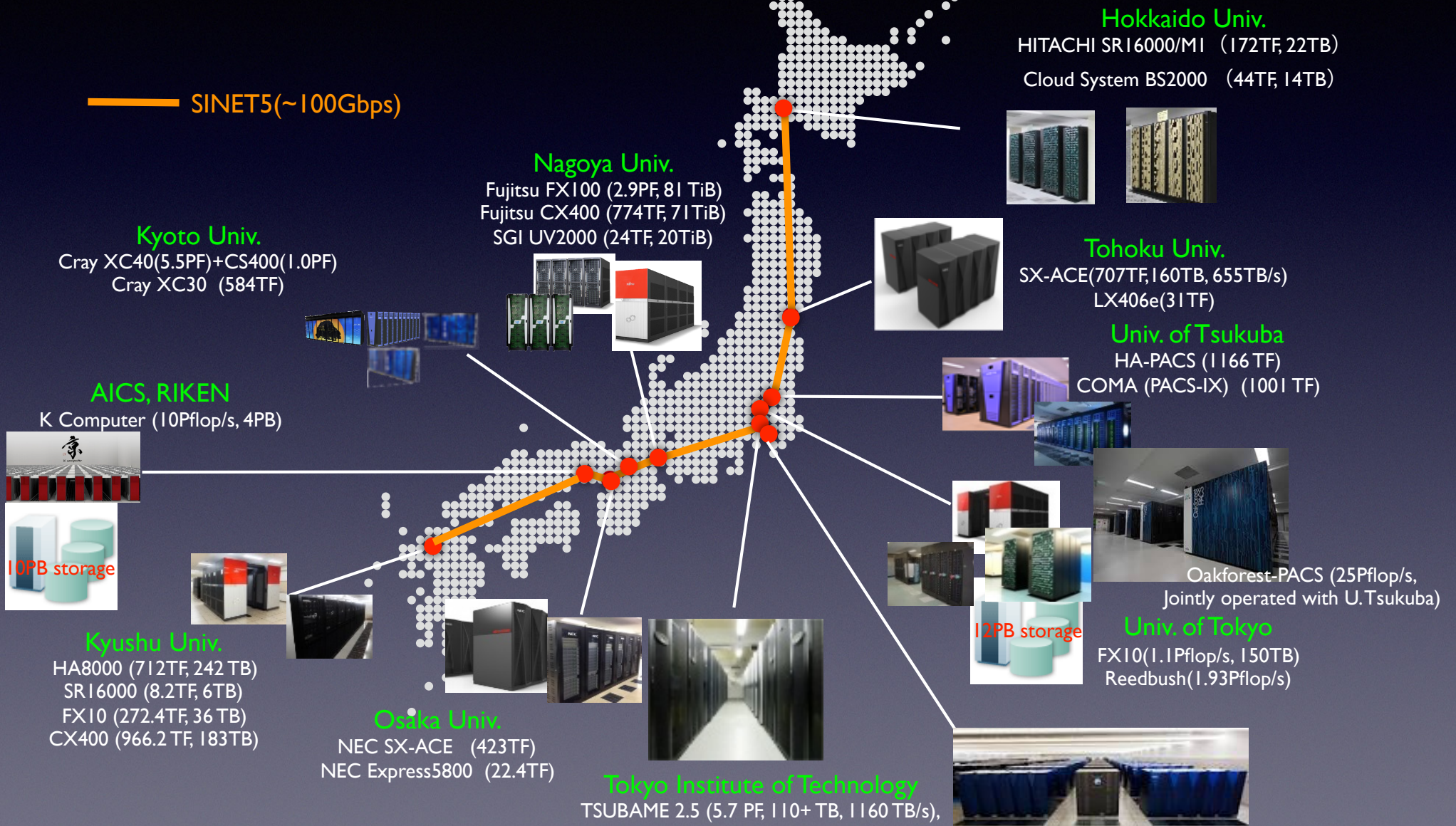




# HPCI: High Performance Computing Infrastructure in Japan

As of Sept., 2016

 SINET5 (~100Gbps)



**Hokkaido Univ.**

HITACHI SR16000/MI (172TF, 22TB)  
Cloud System BS2000 (44TF, 14TB)



**Nagoya Univ.**

Fujitsu FX100 (2.9PF, 81 TiB)  
Fujitsu CX400 (774TF, 71TiB)  
SGI UV2000 (24TF, 20TiB)



**Kyoto Univ.**

Cray XC40(5.5PF)+CS400(1.0PF)  
Cray XC30 (584TF)



**Tohoku Univ.**

SX-ACE(707TF, 160TB, 655TB/s)  
LX406e(31TF)



**Univ. of Tsukuba**

HA-PACS (1166 TF)  
COMA (PACS-IX) (1001 TF)



**AICS, RIKEN**

K Computer (10Pflop/s, 4PB)



Oakforest-PACS (25Pflop/s,  
Jointly operated with U.Tsukuba)



**Univ. of Tokyo**

FX10(1.1Pflop/s, 150TB)  
Reedbush(1.93Pflop/s)



**Osaka Univ.**

NEC SX-ACE (423TF)  
NEC Express5800 (22.4TF)



**Tokyo Institute of Technology**

TSUBAME 2.5 (5.7 PF, 110+ TB, 1160 TB/s),



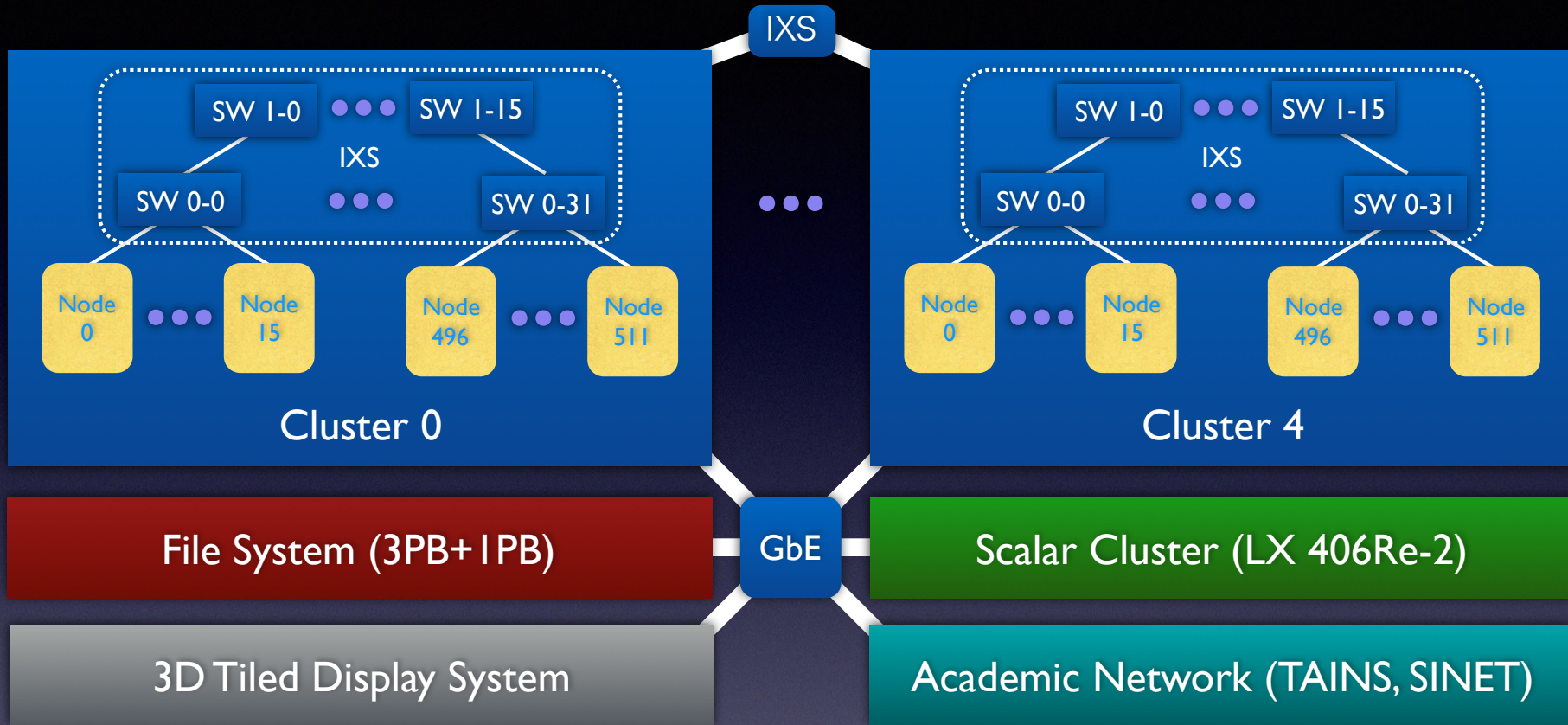
**JAMSTEC**

SX-ACE (1.3 PF, 230 TB),





# Organization of Tohoku Univ. SX-ACE System



	Core	CPU(Socket)	Node	Cluster	Total System
Size	1	4 Cores	1 CPU	512 Nodes	5 Clusters
Performance (VPU+SPU)	69GFlop/s (68GF+1GF)	276GFlop/s (272GF+4GF)		141Tflop/s (139TF+ 2TF)	707Tflop/s (697TF+10TF)
Mem. BW	256GB/s			131TB/s	655TB/s
Memory Cap.	64GB			32TB	160TB
IXS Node BW	-		4GB/s x2		-



# Features of Tohoku Univ. SX-ACE System

## Significant Performance Improvement with Lower Power and Less Space

		SX-9 (2008)	SX-ACE (2014)	Improvement
CPU Performance	Number of Cores	1	4	4x
	Total Flop/s	118.4Gflop/s	276Gflop/s	2.3x
	Memory Bandwidth	256GB/sec	256GB/sec	1
	ADB Capacity	256KB	4MB	16x
Total Performance, Footprint, Power Consumption	Total Flop/s	34.1Tfop/s	706.6Tflop/s	20.7x
	Total Memory Bandwidth	73.7TB/s	655TB/s	8.9x
	Total Memory Capacity	18TB	160TB	8.9x
	Power Consumption (Max)	590kVA	1,080kVA	1.8x
	Footprint	<b>293m<sup>2</sup></b>	<b>430m<sup>2</sup></b>	1.5x

## Powerful CPU/Node Performance and Higher B/F rate

		SX-ACE(2014)	K(2011)	Ratio
CPU (Node) Performance	Clock Frequency	1GHz	2GHz	0.5x
	Flop/s per Core	64Gflop/s	16Gflop/s	4x
	Cores per CPU	4	8	0.5x
	Flop/s per CPU	256Gflop/s	128Gflop/s	2x
	Bandwidth	256GB/s	64GB/s	4x
	Bytes per Flop (B/F)	1	0.5	2x
	Memory Capacity	64GB	16GB	4x

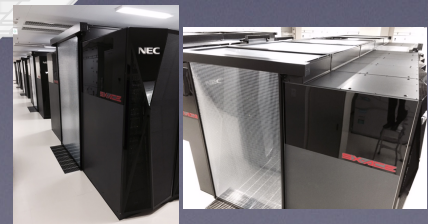
A Balanced System for High Sustained Performance, resulting in High Productivity in the Wide Area of Applications in Academia and Industry





TOHOKU UNIVERSITY

# Cooling Facility of HPC Building



## Evaporation

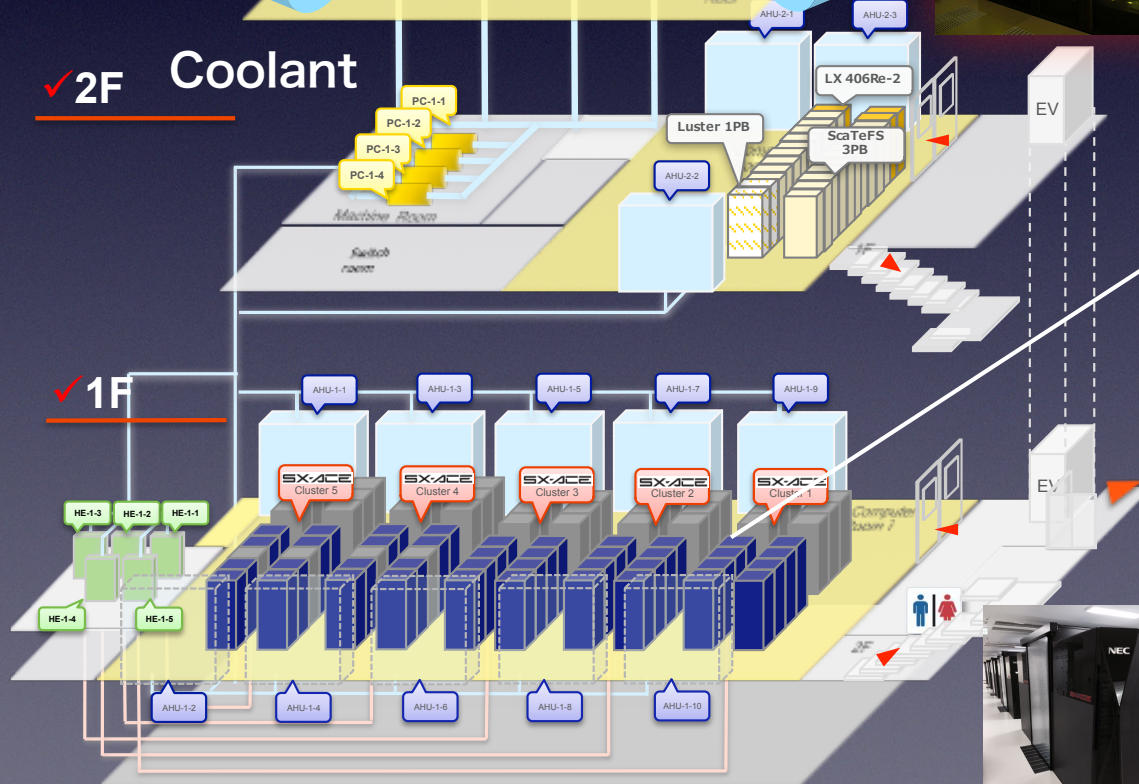
Fresh Air



✓RF

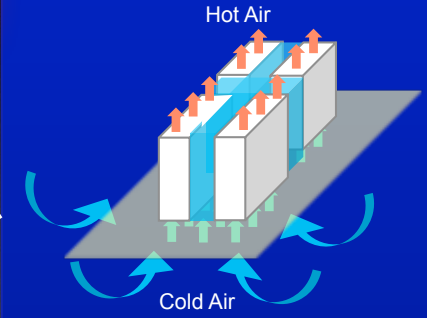
## Coolant

✓2F



✓1F

## Aisle-capping



Capping area of cold-aisle

SX-ACE

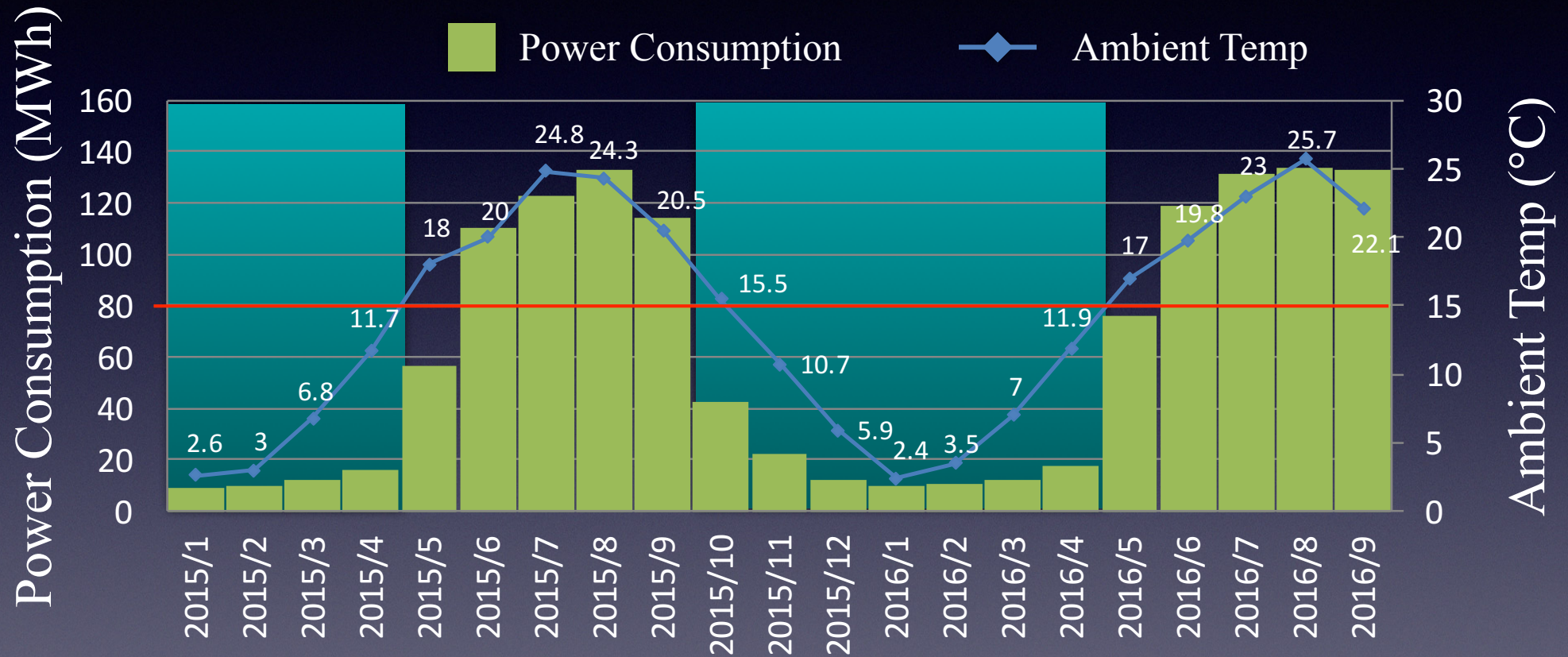
Air-conditioning Equipment

Heat-exchange equipment

coolant pump

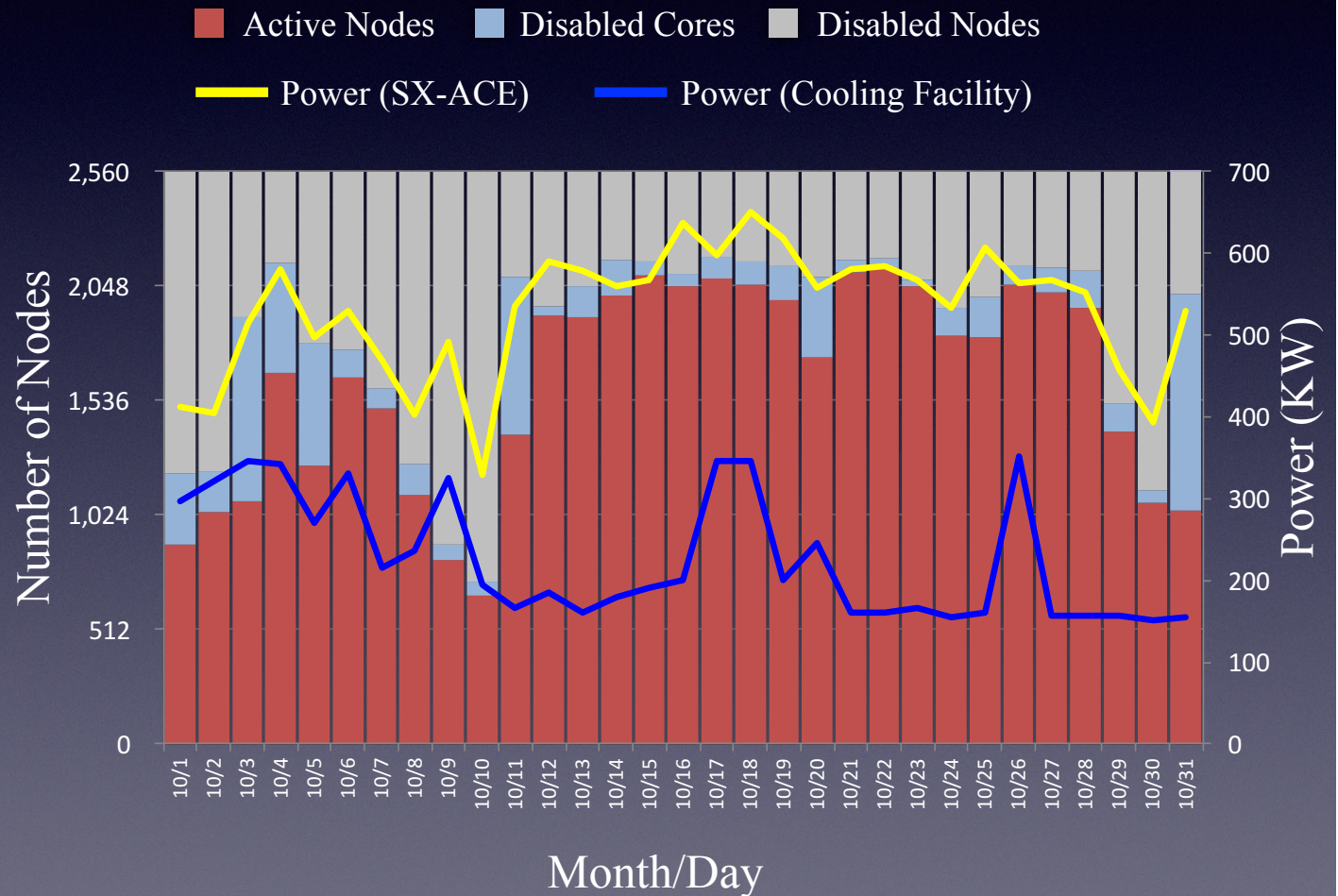
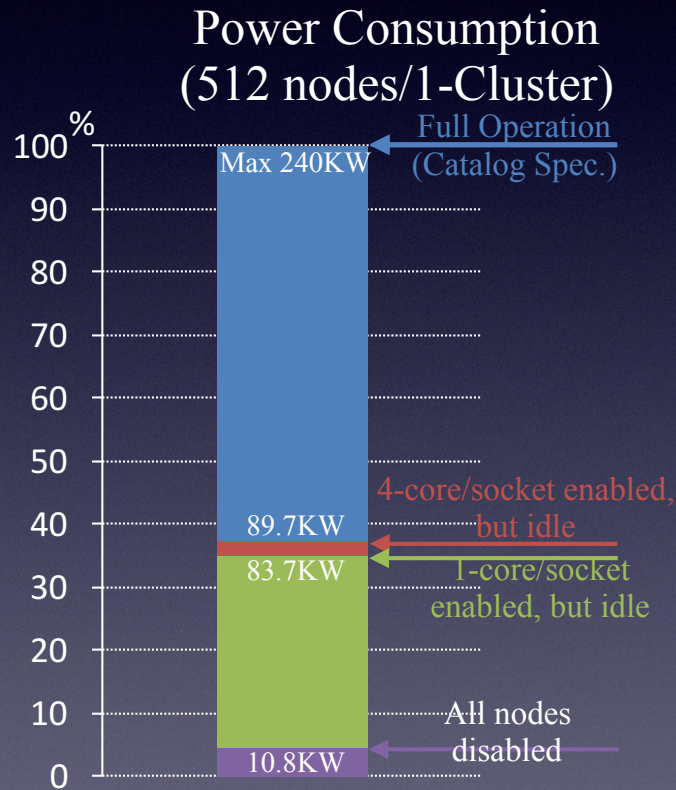


# Power Consumption of the Cooling System Effect of Fresh-Air Cooling



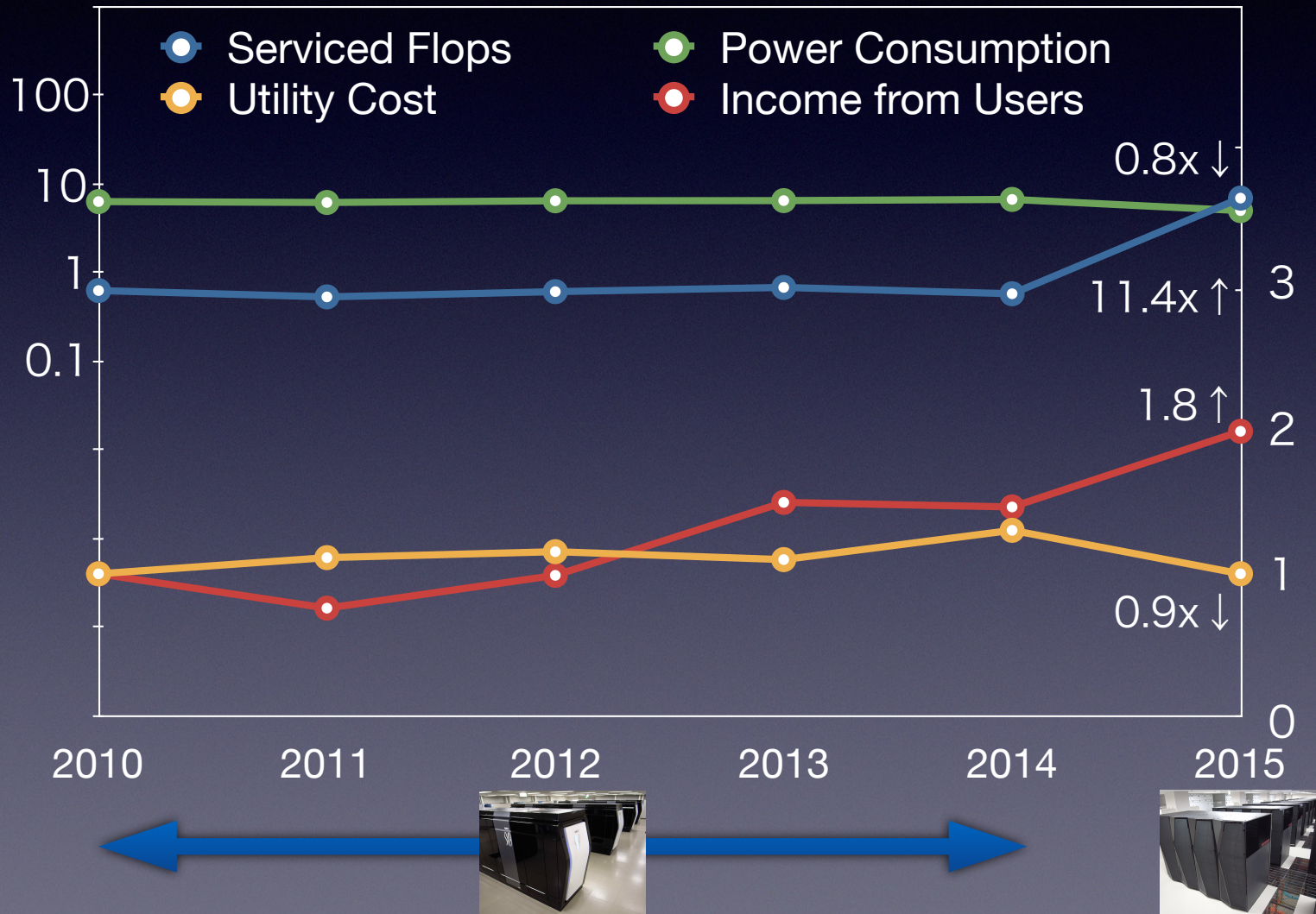


# Node-Core Activity: Effect of Automatic Core-Node Activation/Deactivation Control





# Operation Statistics of SX-ACE (Normalized by SX-9 Data)





# Performance Evaluation of SX-ACE by Using the HPCG Benchmark

- ★ HPCG (High Performance Conjugate Gradients) is designed to exercise computational and data access patterns that more closely match a broad set of important applications,
  - ✓ HPL for top500 is increasingly unreliable as a true measure of system performance for a growing collection of important science and engineering applications.
- ★ HPCG is a complete, stand-alone code that measures the performance of basic operations in a unified code:
  - ✓ Sparse matrix-vector multiplication.
  - ✓ Sparse triangular solve.
  - ✓ Vector updates.
  - ✓ Global dot products.
  - ✓ Local symmetric Gauss-Seidel smoother.
  - ✓ Driven by multigrid preconditioned conjugate gradient algorithm that exercises the key kernels on a nested set of coarse grids.
  - ✓ Reference implementation is written in C++ with MPI and OpenMP support.



# Features of the SX-ACE Vector Processor

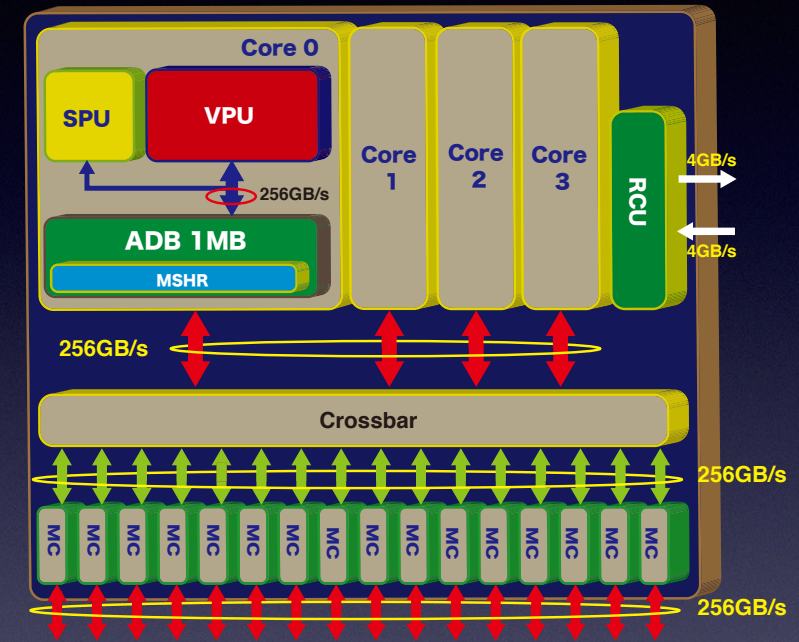
## 4 high-performance core Configuration, each with High-Performance Vector-Processing Unit and Scalar Processing Unit

- 272Gflop/s of VPU + 4Gflop/s of SPU per socket
  - 68Gflop/s + 1Gflop/s per core
- 1MB private ADB per core (4MB per socket)
  - Software-controlled on-chip memory for vector load/store
  - 4x compared with SX-9
  - 4-way set-associative
  - 4-way set-associative
  - MSHR with 512 entries (address+data)
  - 256GB/s to/from Vec. Reg.
    - 4B/F for Multiply-Add operations
- 256 GB/s memory bandwidth, Shared with 4 cores
  - 1B/F in 4-core Multiply-Add operations
  - ~ 4B/F in 1-core Multiply-Add operations
- 128 memory banks per socket

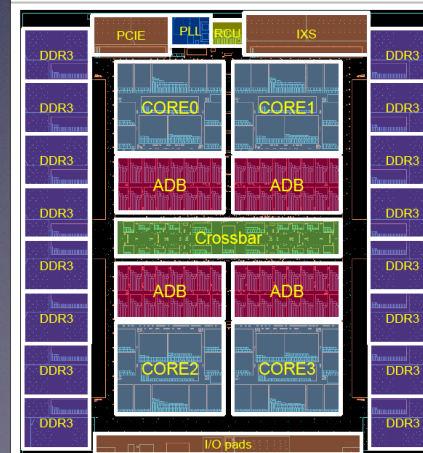
## Other improvement and new mechanisms to enhance vector processing capability, especially for efficient handling of short vectors operations and indirect memory accesses

- Out of Order execution for vector load/store operations
- Advanced data forwarding in vector pipes chaining
- Shorter memory latency than SX-9

## Source: NEC SX-ACE Processor Architecture



### Floor Plan of the CPU



#### “Memory access” focused layout

#### Specifications

- Process rule: 28nm
- Clock speed: 1GHz
- Die size: 23.05 x 24.75mm
- # of transistors: 2BTr.

#### I/F

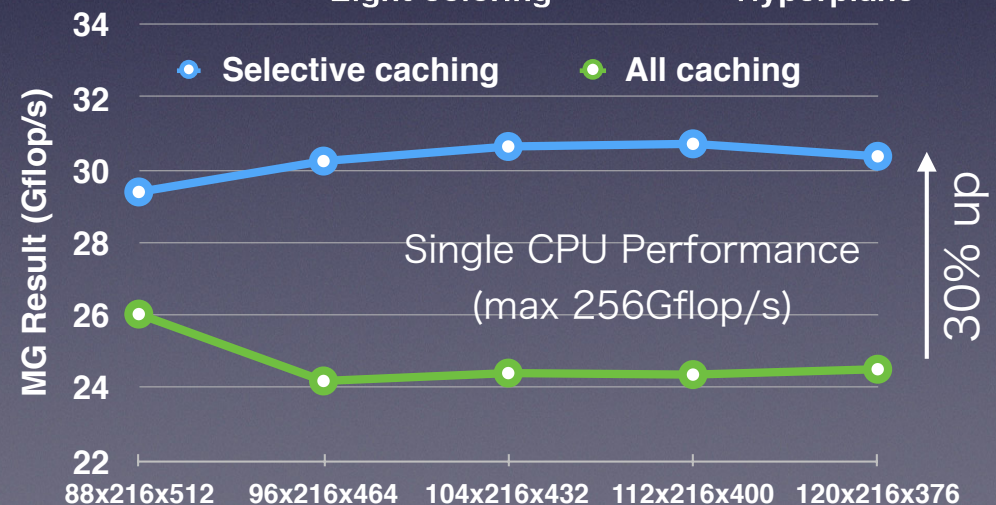
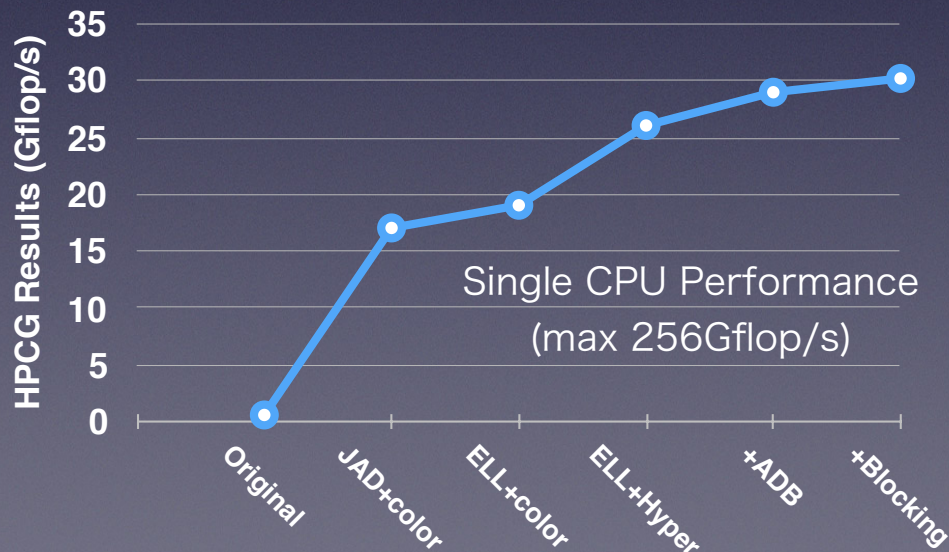
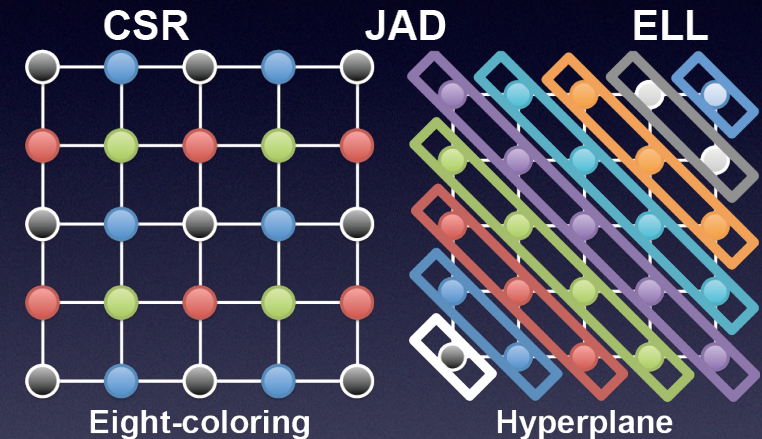
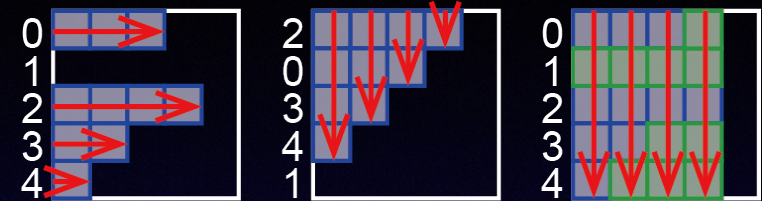
- 16ch DDR3 I/F
- IXS 8GB/s x 2
- 2ch PCIe8 I/F



# Optimizations of the HPCG Benchmark for SX-ACE

- Data packing for vector-friendly matrix memory allocation of sparse matrices
- Parallelization of 27-point stencil computation by using coloring and hyperplane methods
- Selective reusable-data caching and blocking for effective use of ADB

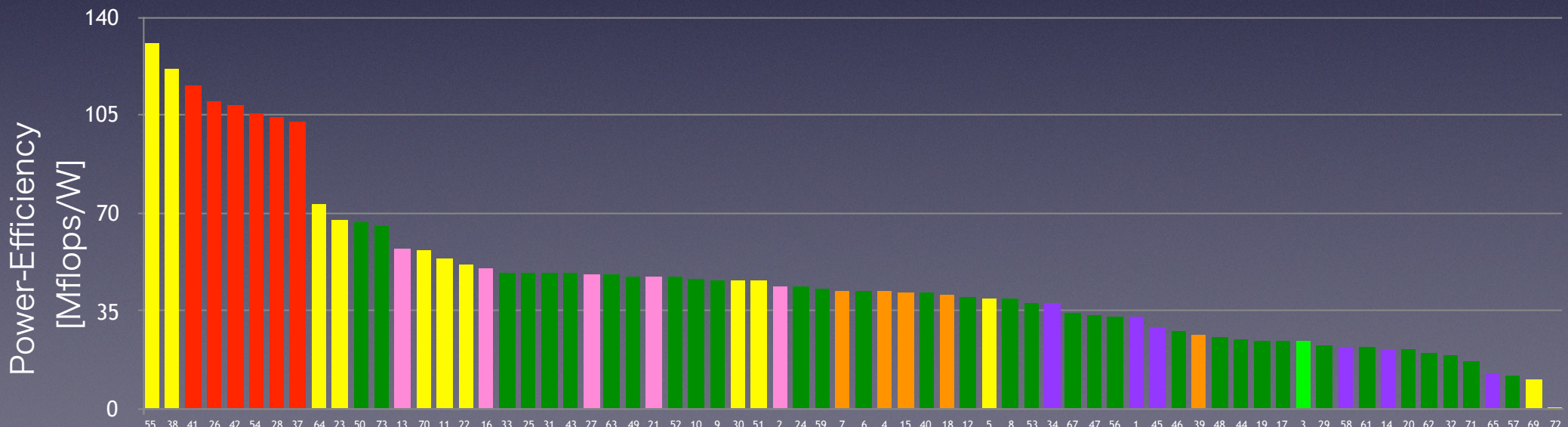
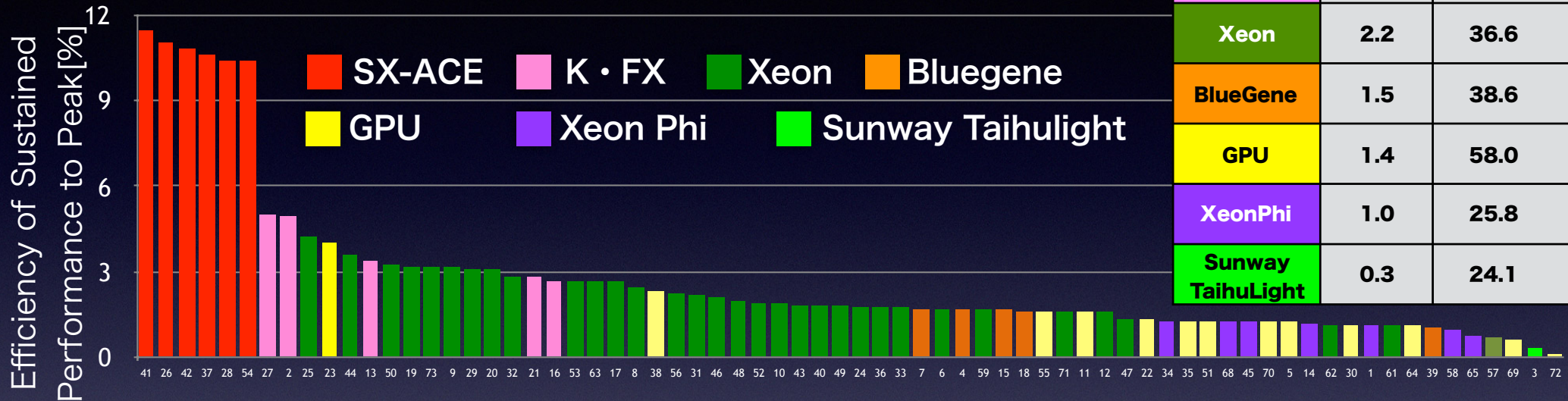
\*Komatsu et al.@SC15





# Efficiency Evaluation of HPCG Performance (AS of Ranking at ISC16)

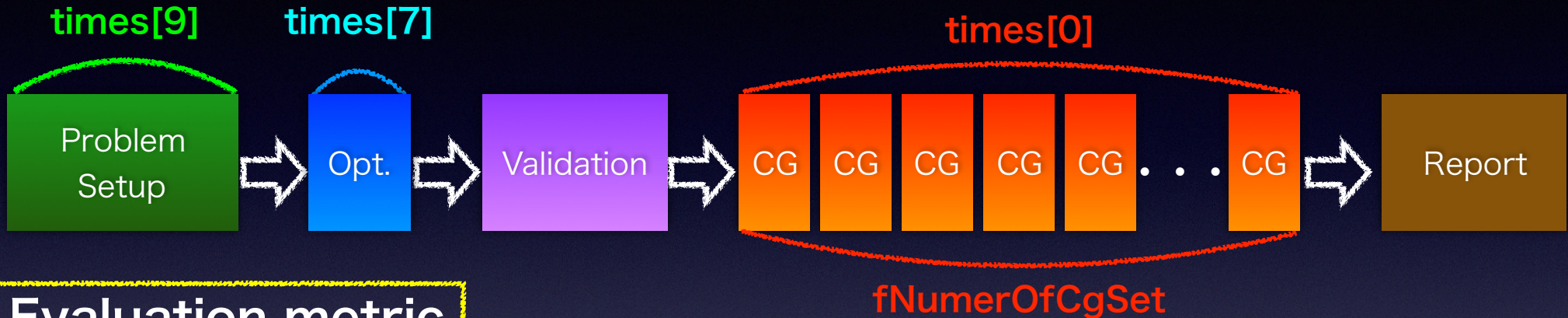
System	Eff.(%)	HPCG/W (MF/W)
<b>SX-ACE</b>	<b>10.8</b>	<b>107.5</b>
<b>K/FX</b>	<b>3.7</b>	<b>49.2</b>
<b>Xeon</b>	<b>2.2</b>	<b>36.6</b>
<b>BlueGene</b>	<b>1.5</b>	<b>38.6</b>
<b>GPU</b>	<b>1.4</b>	<b>58.0</b>
<b>XeonPhi</b>	<b>1.0</b>	<b>25.8</b>
<b>Sunway TaihuLight</b>	<b>0.3</b>	<b>24.1</b>





# HPCG Updates: Evaluation of HPCG Ver3.0 on SX-ACE

## Benchmarking Flow



## Evaluation metric

\* frefnops : total number of floating point operations for CG (# of iterations = 50)

ver. 2.4

$$\text{GFlop/s} = \text{frefnops} / (\text{times}[0] + \text{fNumberOfCgSets} * \text{times}[7] / 10.0) / 1.0\text{E}9$$

ver. 3.0

$$\text{GFlop/s} = \text{frefnops} / (\text{times}[0] + \text{fNumberOfCgSets} * (\text{times}[7] / 10.0 + \text{times}[9] / 10.0)) / 1.0\text{E}9$$

ver. 2.4 → ver. 3.0:

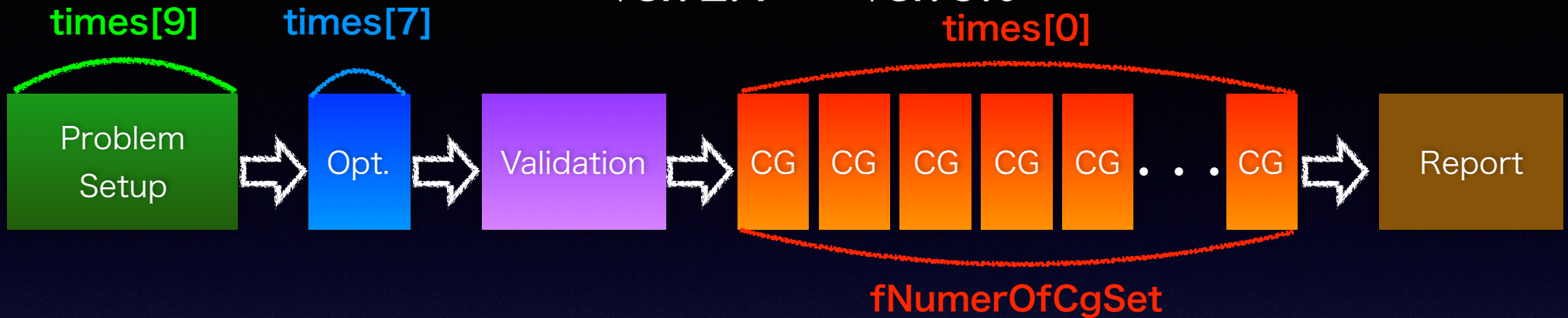
24 Setup overhead considered for individual CG iterations! .5-6, 2016





# Cost Breakdown of HPCB Benchmark

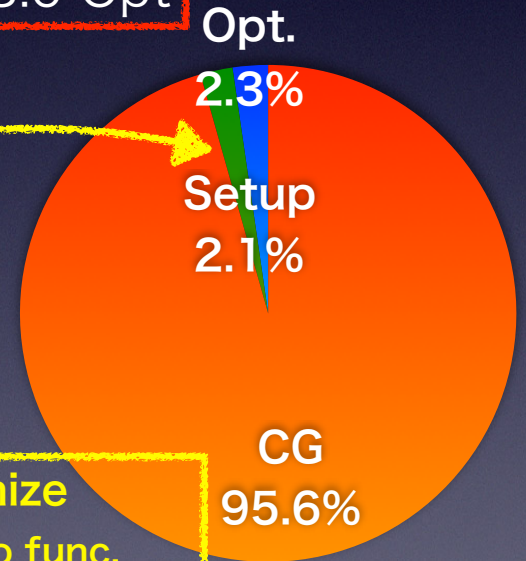
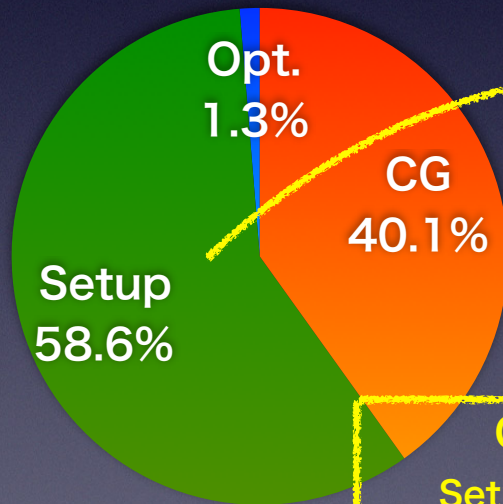
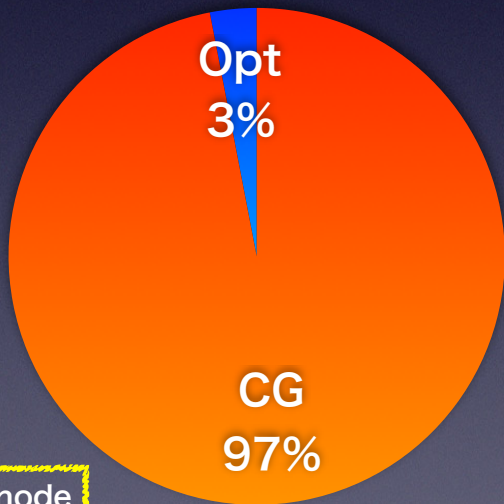
## ver. 2.4 → ver. 3.0



ver. 2.4

ver. 3.0

ver. 3.0 Opt



Performance 1 node

CG
Setup
Opt

62 sec

-

1.94 sec

62.00 sec

43.19 sec

0.925 sec

60.88 sec

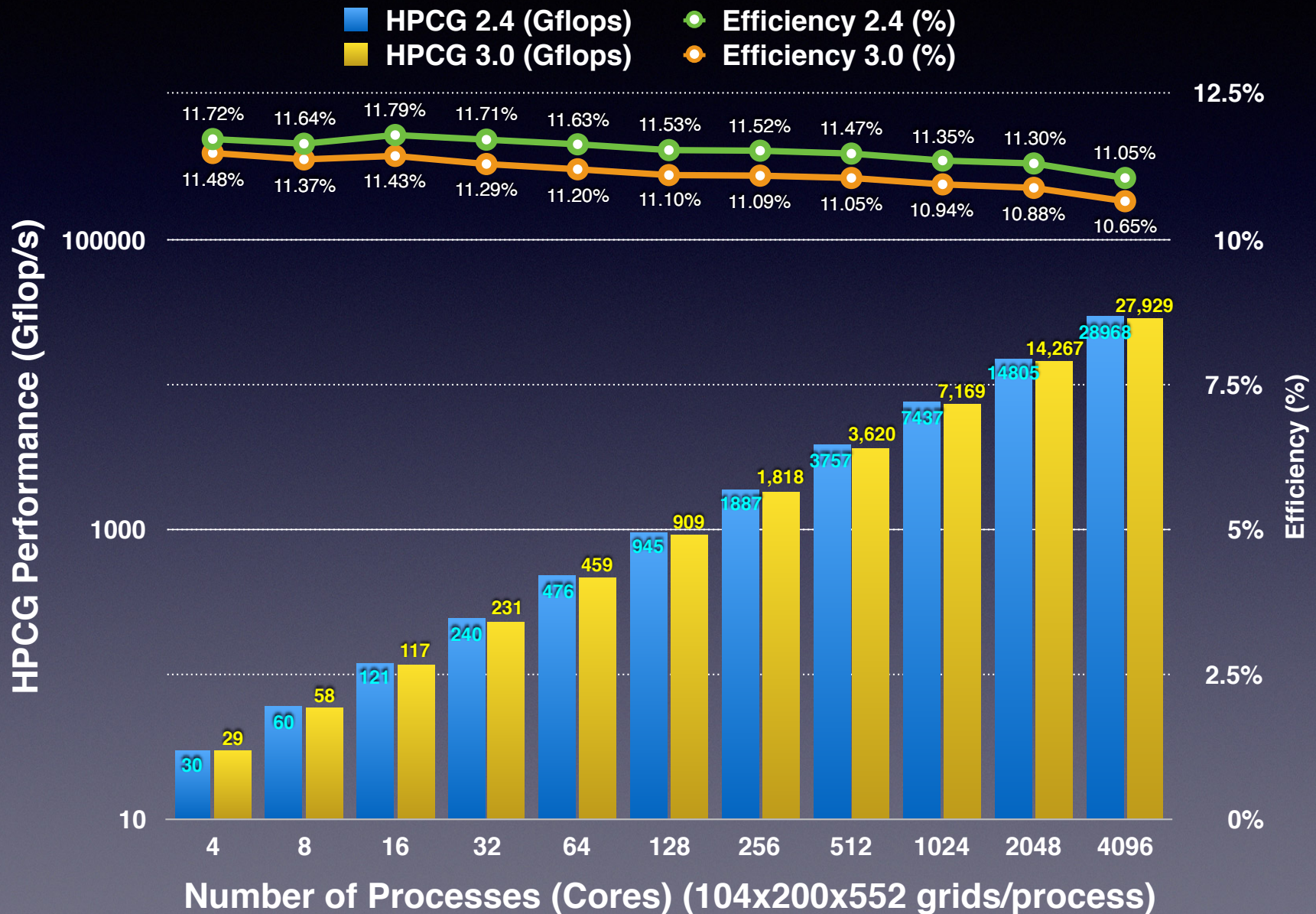
0.637 sec

0.701 sec

Optimize  
SetupHalo func.  
GenerateProblem func.




# Scalability of the HPCG Benchmark





# Performance Evaluation by Using HPGMG (High Performance Geometric Multi-Grid)

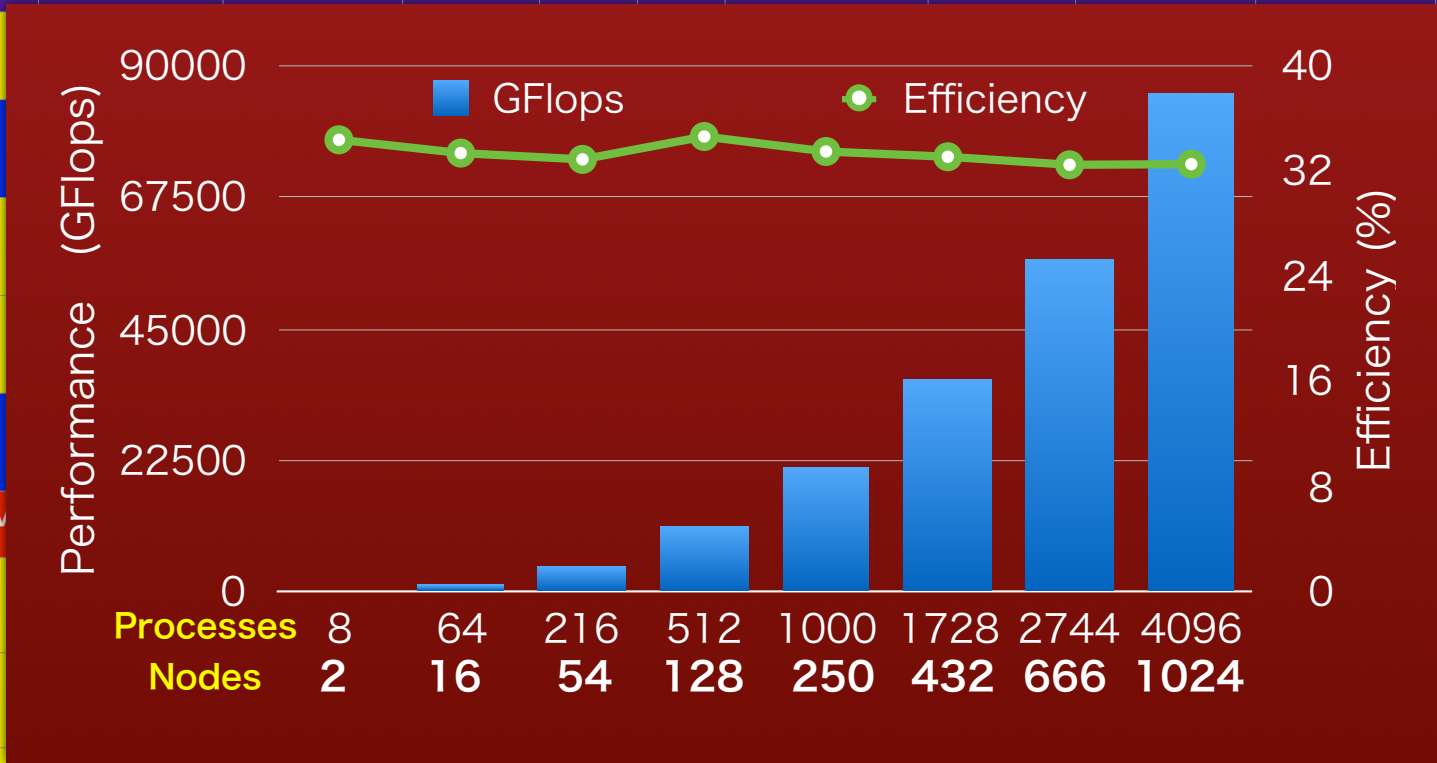
- HPGMG-FV solves variable-coefficient elliptic problems on isotropic cartesian grids
  - Using the **finite volume method (FV)** and **Full Multigrid (FMG)**.
- **Filling the gap between HPL and HPCG**

  - Tracking real application's behavior
    - memory bound, but cache friendly
      - 120 points stencil of Gauss- Seidel Red-Black (GSRB)
    - MPI, OpenMP, OpenACC implementations are available
      - Enabling fair comparison with GPUs, Accelerators

Benchmark	Kernel	Required B/F
HPL	DGEMM	< 0.1
<b>HPGMG</b>	<b>GSRB</b>	<b>&gt; 1</b>
HPCG	SpMV, SYMGS	> 4



# HPGMG Results (As of Nov. 2016 at SC16)

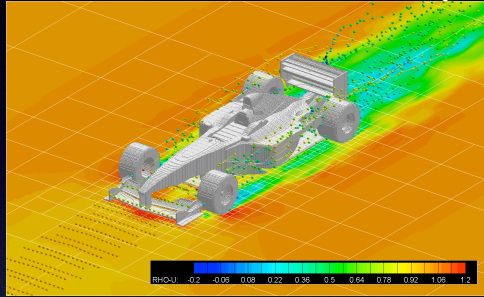
Rank	Site	System	10 <sup>9</sup> DOF/s	MPI	OMP	Acc	DOF/Process	Top500 Rank	HPCG Efficiency	Efficiency	System Arch.
1	ALNL	Mira	500	49152	64	0	36M	9	1.66	5.96	BlueGene
2	HLRS										CrayXC40 (Xeon)
3	OLNL										CRAY (Xeon+K20)
4	KAUST										CrayXC40 (Xeon)
5	NERSC										CrayXC30 (Xeon)
6	CSCS										CrayXC30 (Xeon+K20)
7	Tohoku Univ										NEC SX
8	LRZ										Idataplex (Xeon)
9	NREL										Apollo 8000 (Xeon)
10	NREL	Peregrine	5.29	512	12	0	16M	-	N/A	5.38	Apple 8000 (Xeon)
11	HLRS	KABUKI	3.24	256	1	0	32M	-	11.45	23.71	NEC SX
12	NERSEC	Babagge	0.726	256	45	0	8M	-	N/A	N/A	Intel Xeon Phi



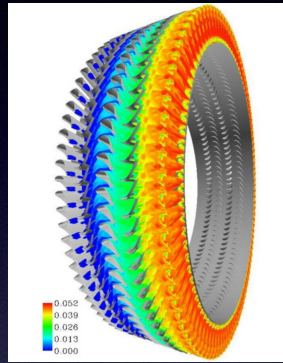


# Leading Science and Engineering Fields supported by the Supercomputer of Tohoku University

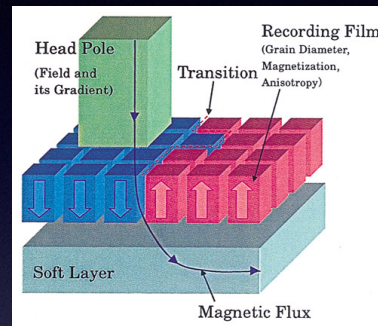
Next-Generation CFD Analysis



Turbine Design



Perpendicular Magnetic Recording Medium Design



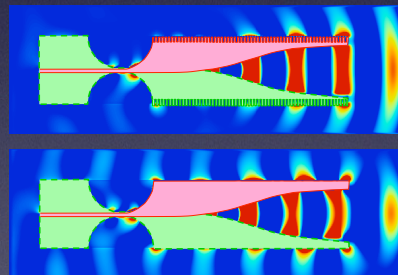
Nano Material Design



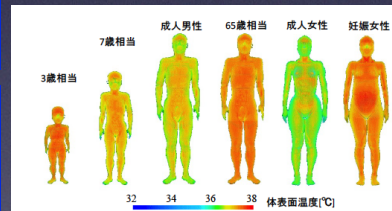
Industrial Use



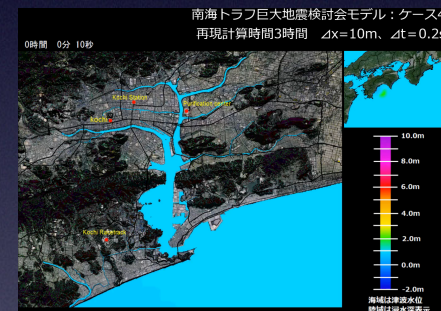
Antenna Analysis



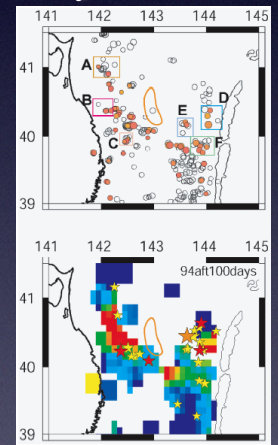
Heat Shock Analysis



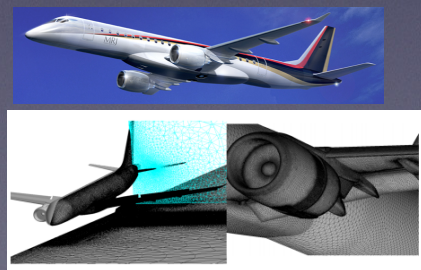
Tsunami Inundation Analysis



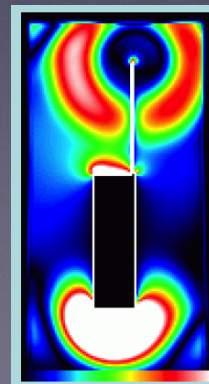
Earthquake Analysis



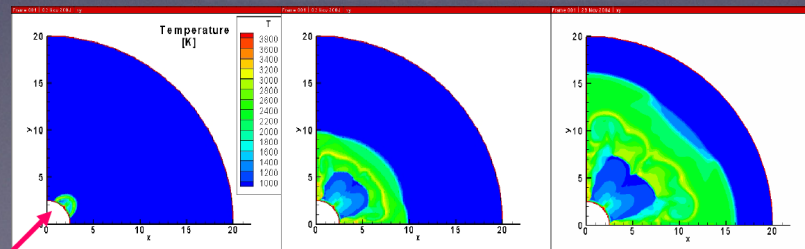
MRJ



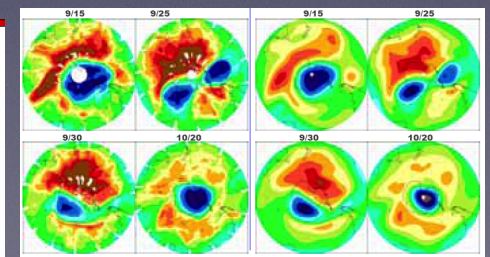
24th WSSP



Combustion Flow Simulation



Ozone-hole Analysis



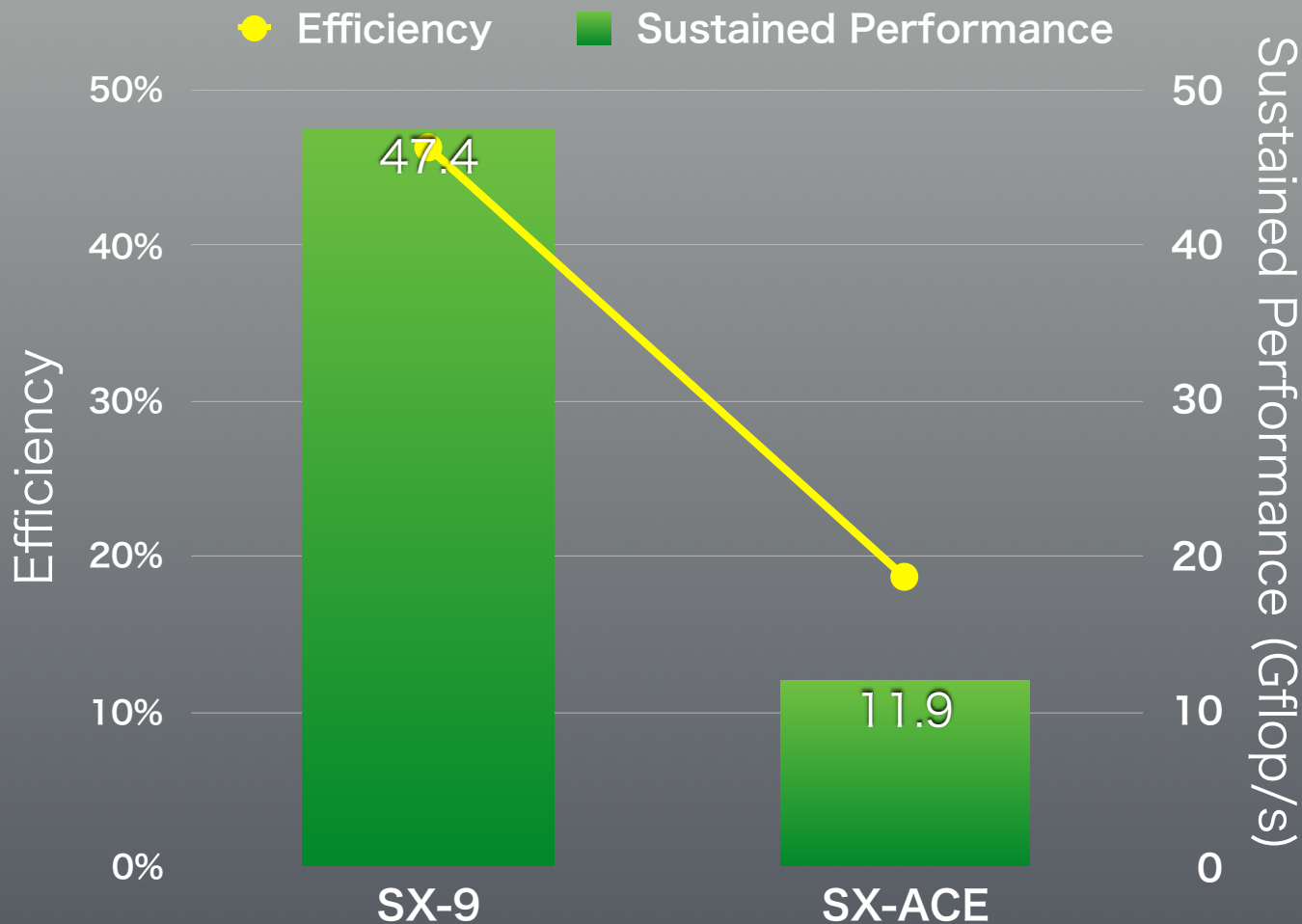


# Case I: Trade off between Vector Length and Stride

## Performance of QSFDM GLOBE codes on SX-ACE

In SX-9, Increasing vector length is effective even in the case of stride memory accesses

single core execution



Loop length :  
2187

Stride length :  
3165



# Case I: Trade off between Vector Length and Stride

## Performance of QSFDM GLOBE codes on SX-ACE

### • Interchange loops to reduce the stride length

```

+---> do jz=max(npol01+n2o+1,mz1b),min(idix11-1,mz1e)
| !cdir select(vector)
|V--> do jx=3,nx1-n2o
||
|:
|| !cdir expand=nl
||*-> do jl=1,nl
||| work1=work1+rxx1(jl,jx,jz)
|:
||| rxx1(jl,jx,jz) =
||| & ((2.0e0*tau1(jl,jx,jz)-dt)/(2.0e0*tau1(jl,jx,jz)+dt))
||| & *rxx1(jl,jx,jz)
||| & -(2.0e0/(2.0e0*tau1(jl,jx,jz)+dt))
||| & *( + coeff1*( dpai1(jl,jx,jz)-2.0e0*damu1(jl,jx,jz))
||| & + coeff2*(-dpai1(jl,jx,jz)+4.0e0*damu1(jl,jx,jz))
||| & + coeff3*( dpai1(jl,jx,jz)-2.0e0*damu1(jl,jx,jz))
||| & + coeff5*( dpai1(jl,jx,jz)-2.0e0*damu1(jl,jx,jz))
||| & - coeff6*( dpai1(jl,jx,jz)-2.0e0*damu1(jl,jx,jz))
|:

```

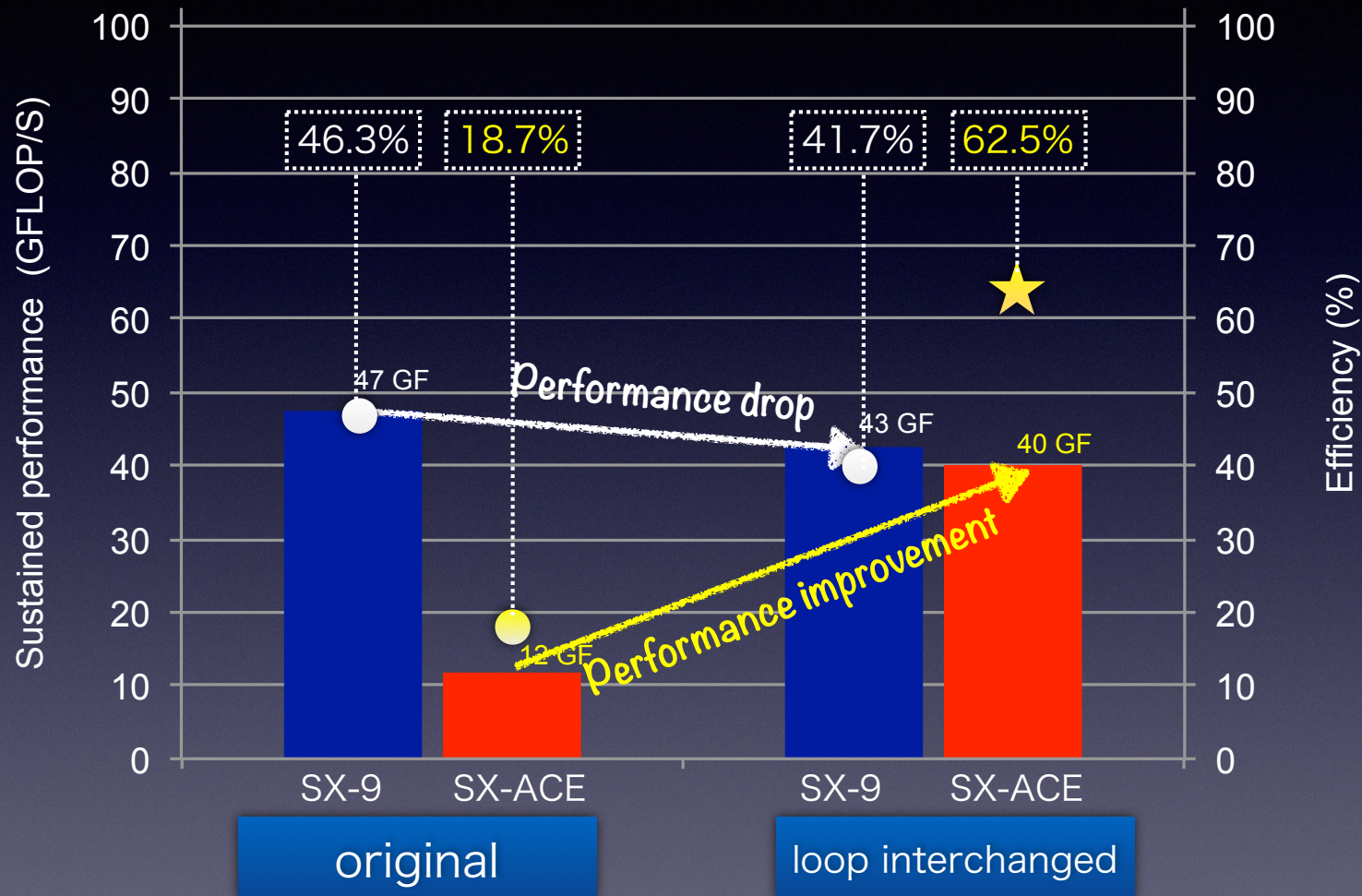
Loop length :  
629

Stride length :  
5



# Case I: Trade off between Vector Length and Stride

## Performance of QSFDM GLOBE codes on SX-ACE

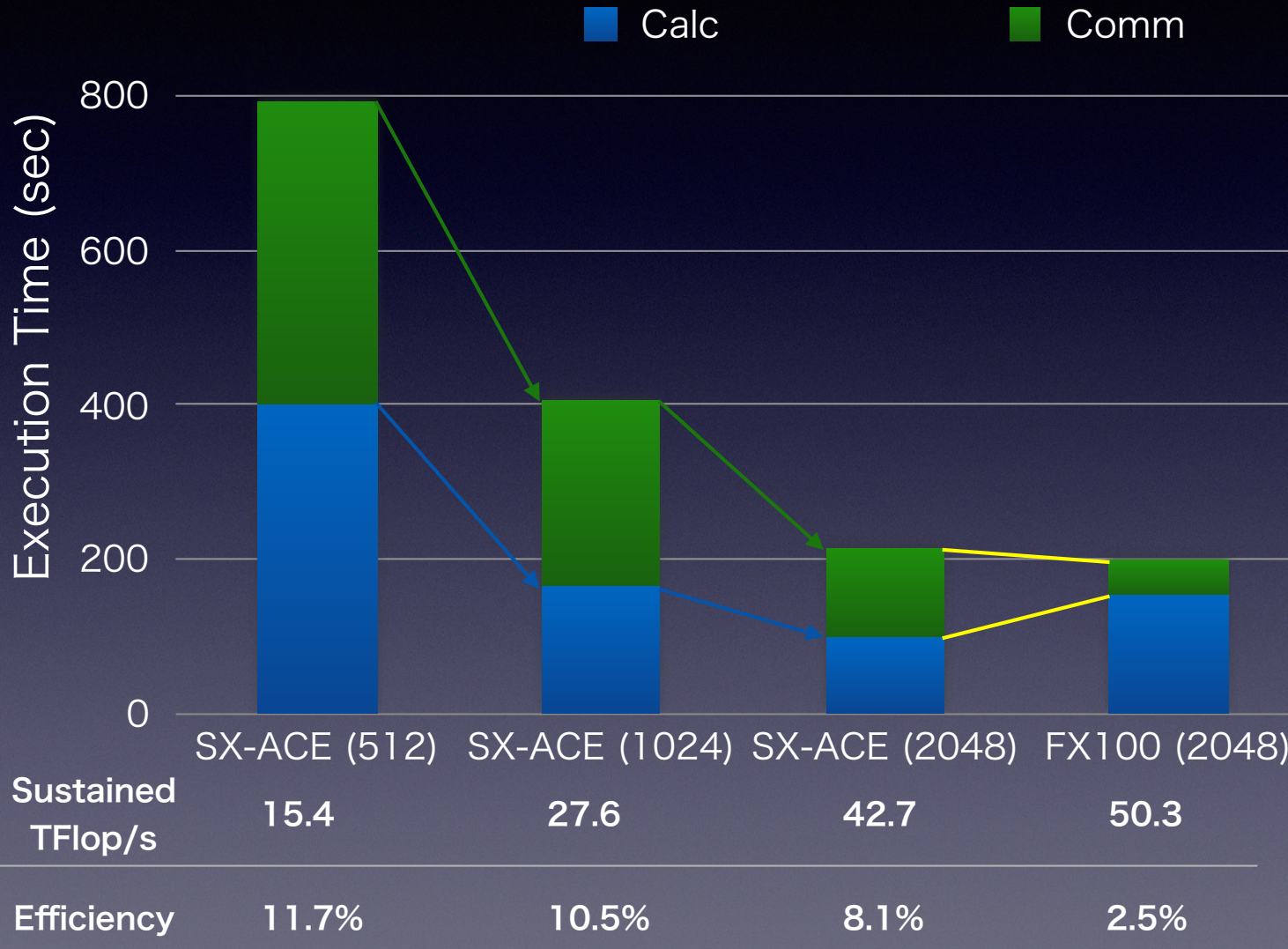
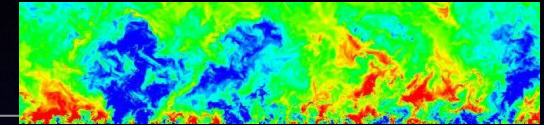
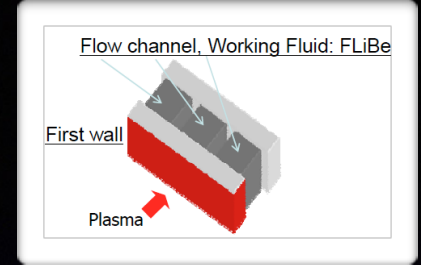


Even in the same system series, HPC codes should be re-optimized according to the evolutions of system architecture!



# Case 2: Balance between Computing Perf. and Network Perf.

## Performance of MHD codes on SX-ACE



	FX100	SX-ACE
Arch	SPARCv9_HPC-ACE	SX
CPU/node	1	1
Cores/CPU	32	4
Node Perf.	1.011 Tflop/s	256 Gflop/s
Mem cap./node	32GB	64GB
mem BW	480 GB/s	256 GB/s
NW BW	12.5 GB/s	4GB/s

Ref : Y. Yamamoto, R. Egawa, Y. Isobe , and Y. Tsuji, "Performance evaluation of DNS code based on high-order accuracy finite difference methods," Japan-Russia Workshop @ Nagoya, Dec 10, 2015.



# Case 2: Optimize Decomposition Size:

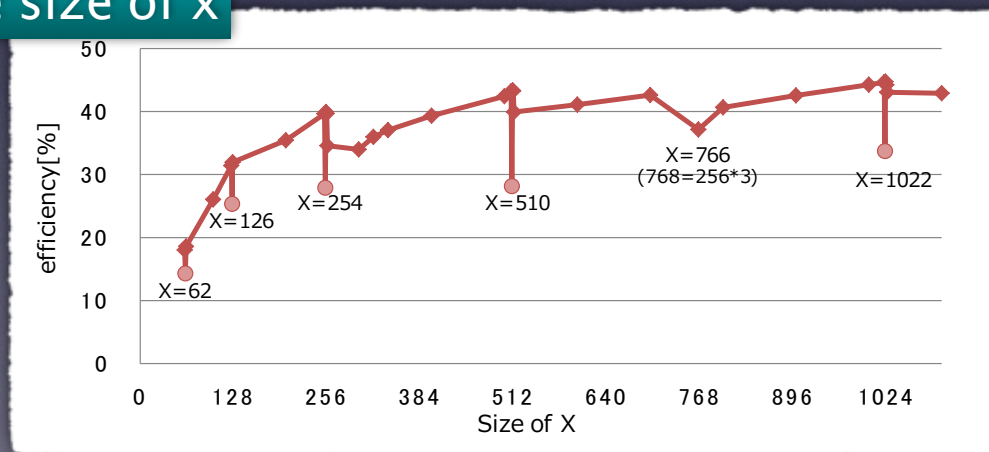
## Increasing Vectorization and Reducing Communications

changing the size of y, z

x	y	z	npex	npey	npez	Communication Elements				Communication		GFLOPS/proc (効率:%)
						x*y	y*z	z*x	Total	[sec]	[%]	
600	25	400	2	8	8	15,000	10,000	240,000	265,000	5.54	25.8	24.01(37.5%)
600	25	400	2	4	16					5.42	25.3	24.09(37.6%)
600	50	200	2	8	8	30,000	10,000	120,000	160,000	4.93	23.6	24.86(38.8%)
600	50	200	2	4	16					4.88	23.5	24.89(38.9%)
600	100	100	2	8	8	60,000	10,000	60,000	130,000	4.05	19.2	24.71(38.6%)
600	100	100	2	4	16					4.14	19.7	24.78(38.7%)
600	200	50	2	8	8	120,000	10,000	30,000	160,000	5.12	23.7	23.95(37.4%)
600	200	50	2	4	16					4.95	23.0	24.14(37.7%)
600	400	25	2	8	8	240,000	10,000	15,000	265,000	7.26	30.4	21.46(33.5%)
600	400	25	2	4	16					7.17	30.0	21.61(33.8%)

Keeping  
y = z

changing the size of x



larger x is better





東北大学



Cyberscience  
Center

# Emergency Computing on SX-ACE for Disaster Analysis and Mitigation of Tsunami



IRIDeS  
International Research Institute  
of Disaster Science  
災害科学国際研究所



AOB **NEC**



**KOKUSAI KOGYO CO., LTD.**

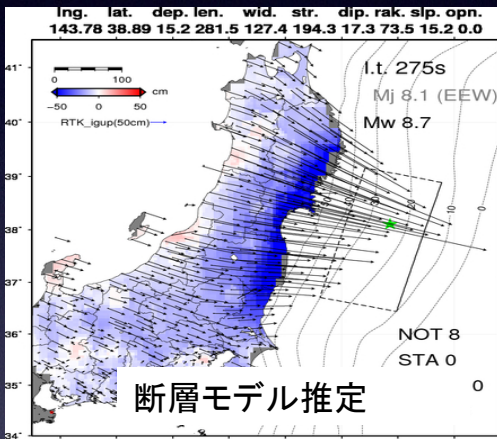


# Design and Development of A Real-Time Tsunami Inundation Forecasting System

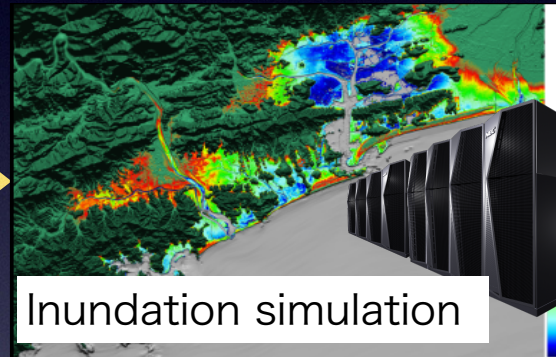
GPS-Observation

Simulation on SX-ACE

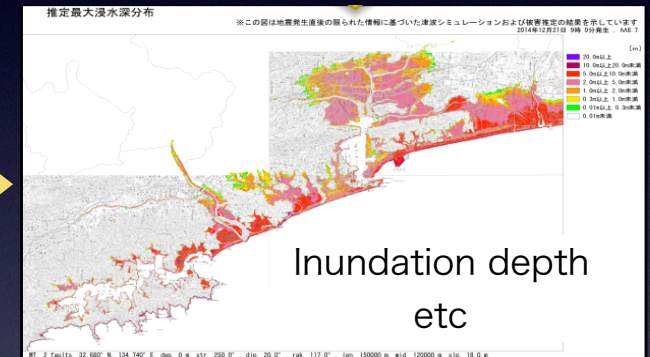
Information Delivery



Fault estimation based on GPS data



10-m mesh models of coastal cities



Just-In-Time access of Visualized information by local governments

< 8 min

< 8 min

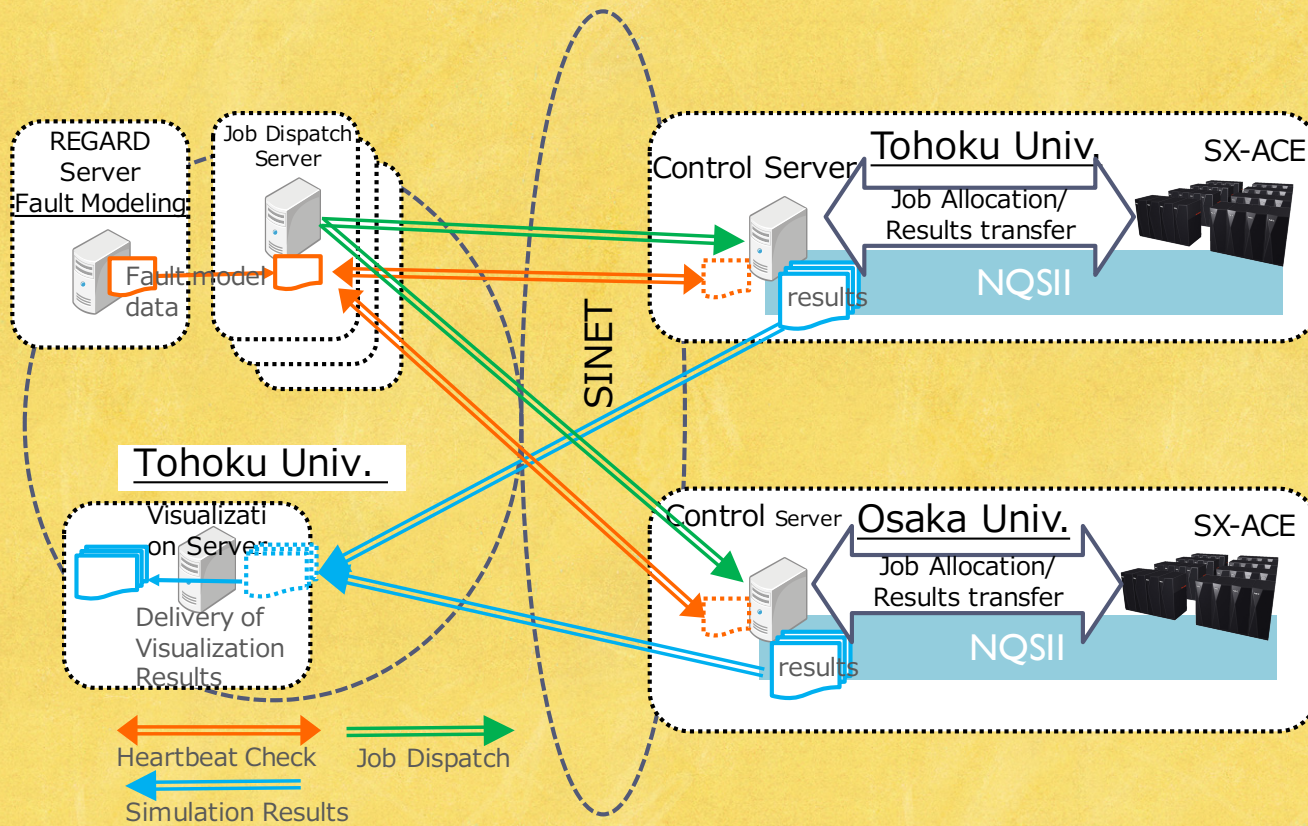
< 4 min

< 20 min

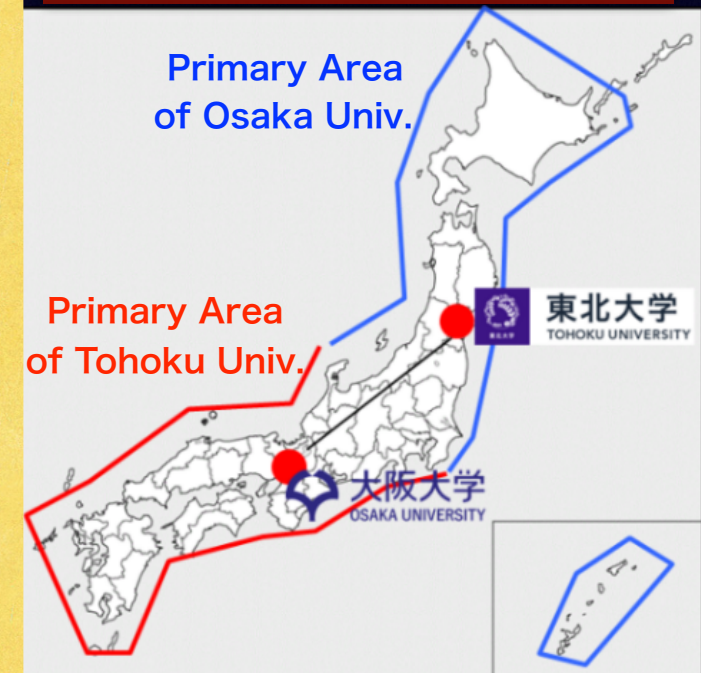


# System Extension for Coverage of the Entire Japan and Improvement of Dependency

## Dual-System with Collaborative Operations of Distant Systems



## Full coverage of Japan





# What Happened on Nov. 22

## Magnitude-7.4 quake likely an aftershock from five years ago

### Events in Simulation Timeline

6:00:34 Estimation Started  
(Fault Estimation) **10minutes**

6:07:26 Job submission  
(Simulation at Tohoku)

6:09:52 Simulations Completed

6:07:27 Job submission  
(Simulation at Osaka)

6:10:00 Simulations Completed

7:31 Visualization Started

Visualization was performed manually, because Sendai and vicinity areas are not automatic visualization areas in the current implementation

**Information delivery**

After the visualization was completed, the inundation information available on the Web is sent to Cabinet Office, Government of Japan and Higashi-Matsushima-Shi and Ishinomaki-Shi.

### Events in Actual Timeline

5:59:47 Earthquake

6:00:06 1<sup>ST</sup> EEW(M6.0)

6:00:34 8<sup>th</sup> EEW(M7.1)

6:01:14 11<sup>th</sup> EEW(M7.3)

6:02:23 Tsunami Information

Tsunami Warning(Fukushima)



6:29 1<sup>st</sup> Tsunami Arrival(Iwaki Onahama)

⋮

7:09 Tsunami Arrival(Sendai Port)

7:26 Tsunami Warning (Miyagi)

8:03 Max Water Level(1.4m)  
at Sendai Port

(arbitrary) 12:50 Alert Withdrawal



# What Happened on Nov. 22

Real-Time Tsunami Inundation Forecasting TOP Monitoring DEMO 日本語 Sign-out

server status: ONLINE connection status: Connected Connected Amount: 1

Area

Display Full Range (Miyagi:RED Shizuoka:BLUE Kochi:GREEN)  
 Miyagi(Ishinomaki and Higashi-Matsushima)  Shizuoka  Kochi

Log CLEAR

2016/11/22 08:00:09	[20161122055958]	Ignored for Low Reliability (443)	[miyagi/ishinomakihigashimat
2016/11/22 08:00:12	[20161122055958]	JMA EEW SEQ 4 [M6.3]	
2016/11/22 08:00:12	[20161122055958]	Ignored for Low Reliability (444)	[miyagi/ishinomakihigashimat
2016/11/22 08:00:13	[20161122055958]	JMA EEW SEQ 5 [M6.2]	
2016/11/22 08:00:13	[20161122055958]	Ignored for Magnitude Under Threshold (6.2)	[miyagi/ishinomaki
2016/11/22 08:00:14	[20161122055958]	JMA EEW SEQ 6 [M6.5]	
2016/11/22 08:00:15	[20161122055958]	Ignored for Magnitude Under Threshold (6.5)	[miyagi/ishinomaki
2016/11/22 08:00:28	[20161122055958]	JMA EEW SEQ 7 [M6.8]	
2016/11/22 08:00:28	[20161122055958]	Ignored for Magnitude Under Threshold (6.8)	[miyagi/ishinomaki
2016/11/22 08:00:34	[20161122055958]	JMA EEW SEQ 8 [M7.1]	
2016/11/22 08:00:34	[20161122055958]	Waiting Job Trigger	(miyagi/ishinomakihigashimatsushima)

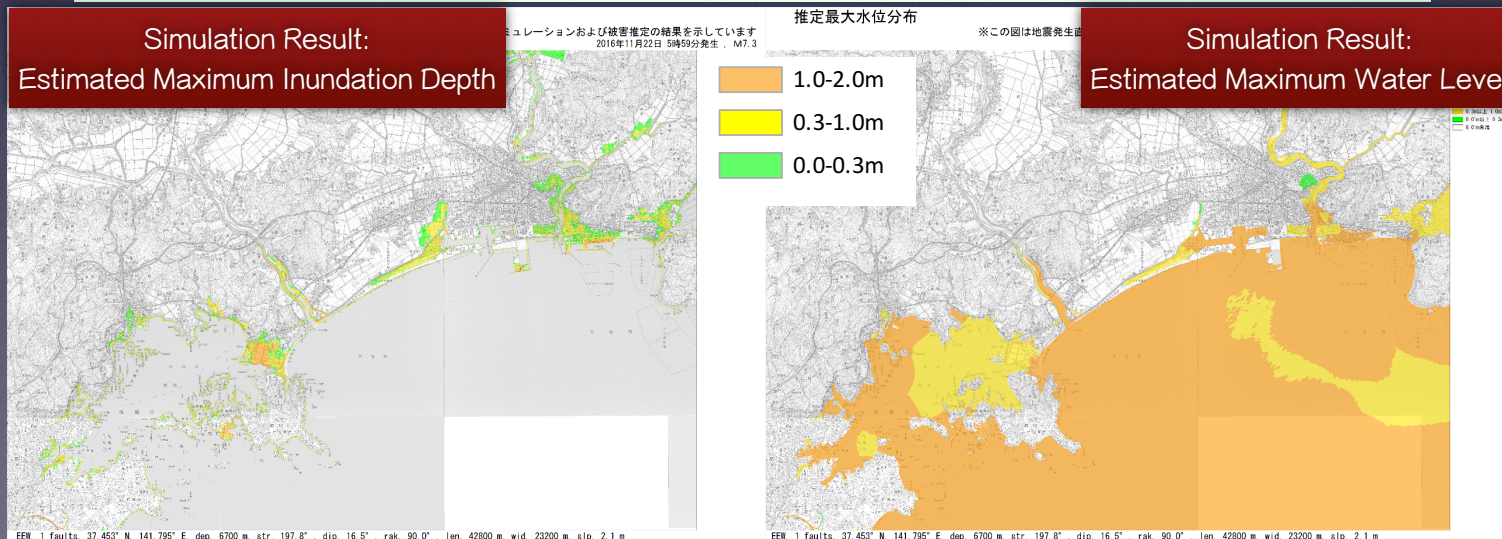
GREEN CIRCLE denotes P-wave arrival area in the map.  
 RED CIRCLE denotes S-wave arrival area in the map.  
 P-wave & S-wave circle disappears after 120 seconds.  
 Earthquake detail appears by clicking marker on the map.

Hypocenter  
 Latitude 37.3  
 Longitude 141.6  
 Depth 10.0km  
 Magnitude 7.1

Lapse time from the earthquake: 52 Seconds  
 Remaining time before calculation of flood estimate start: 367 Seconds

Empowered by Innovation  
 NEC  
 IRIDeS  
 AZ Corporation  
 KOKUSAI KOGYO CO., LTD.  
 Hitz  
 Hitachi Zosen

Copyright© 2016 A2 Corporation.







東北大学

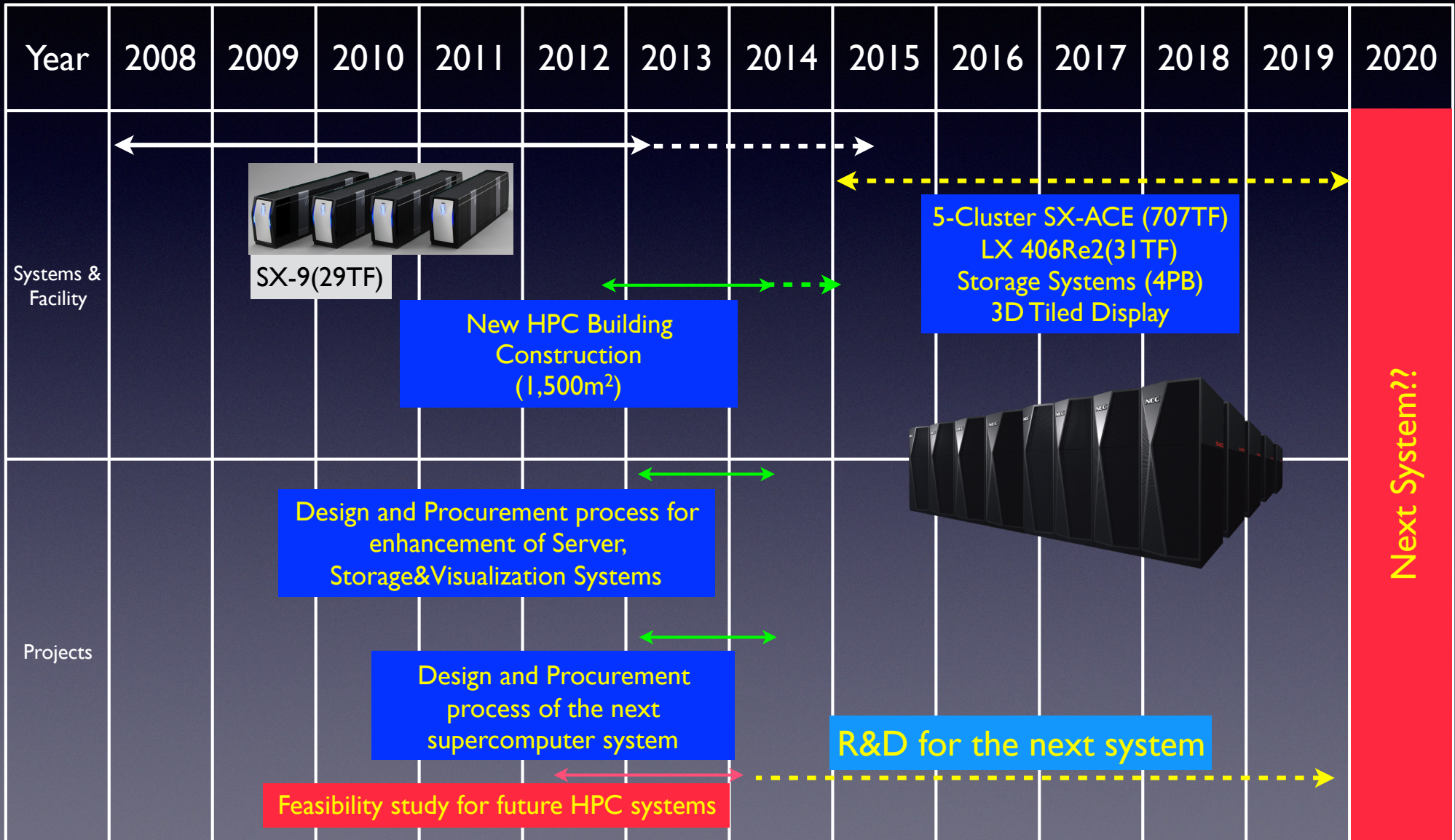
# Future Vector Systems R&D\*

**\*This work is partially conducted with  
NEC, but the contents do not reflect  
any future products of NEC**





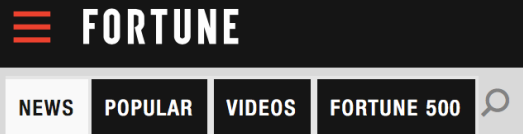
# Timeline of the Cyberscience Center HPC System Development and R&D For the Future





# Future Plan of HPCI Deployment in Japan

Fiscal Year	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
Hokkaido			HITACHI SR16K/M1 (54.9 TF)									
Tohoku			NEC SX-9 (60TF)									
Tsukuba			HA-PACS (1166 TF)									
Tokyo			COMA (PACS-IX)									
Tokyo Tech.												
Nagoya			Fujitsu FX10 (Oakleaf/O) (1.27PFlops, 168TB, 46 TB/s)									
Kyoto			Hitachi SR16K/M1 (54.9 TF)									
Osaka			TSUBAME 2.5 (5.7 PFlops, 1160 TB/s, 1.4MW)									
Kyushu			FX10(90TF)									
JAMSTEC			CX400(470TF)									
			SJI UV2000 (24TF)									
			Cray: XE6 + GB8 (983TF)									
			Cray XC30 (584TF)									
			NEC SX-AC (423TF)									
			HA8000 (712TF, 242 TB/s)									
			SR16000 (8.2TF, 6TB/s)									
			FX10 (272.4TF, 36 TB/s)									
			CX400 (966.2 TF, 18 TB/s)									
			NEC SX-9 (131TF) 5MW									



FORTUNE NEWS POPULAR VIDEOS FORTUNE 500

**This Country Wants to Build the World's Fastest-Known Supercomputer** NOVEMBER 25, 2016

**Here's What China Is Planning Amid Worries About Capital Outflow** 4:15 AM EST

**Trump Aides Say Cuba Must Change or the Door Obama Opened Will Close** 4:04 AM EST

**You Can Now Officially Invest in Alibaba's Singles' Day** 3:59 AM EST

130PF in HP?, 33PF in DP?  
4.4PB/s, 480TB  
3MW  
available 1Q2018,  
\$173 M including the building and facility

Japan plans to build the world's fastest-known supercomputer in a bid to arm the country's manufacture with a platform for research that could help them develop and improve driverless cars, robotics, and medical diagnostics.

The Ministry of Economy, Trade and Industry will spend 19.5 billion yen (\$173 million) on the previously unrevealed project, a budget breakdown shows, as part of a government policy to get back Japan's mojo in the world technology. The country has lost its edge in many electrical fields amid intensifying competition from South Korea and China, home to the world's current best-performing

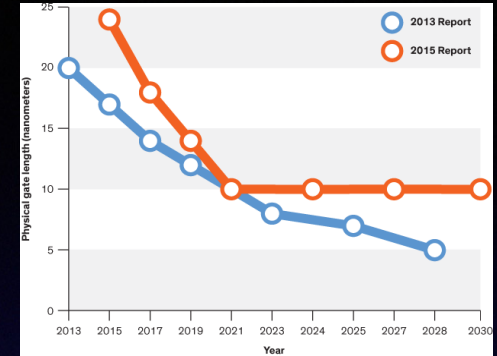
...that is expected to vault Japan to the top of the computing heap, its engineers will be tasked with building a machine that can make 130 quadrillion calculations per second – or 130 petaflops in scientific parlance – as early as next year, sources involved in the project told Reuters.



# But the Road to Exascale is Not So Easy...

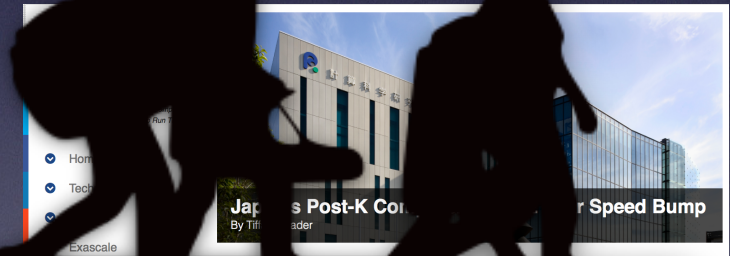
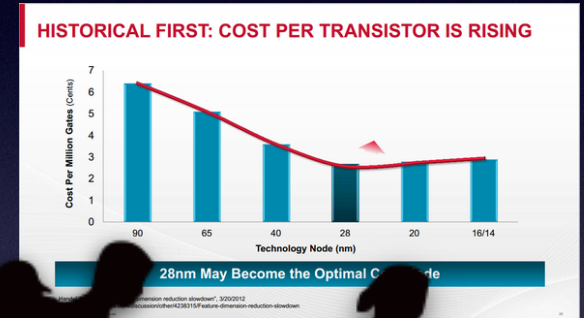
## ● End of Moore's Law???

- ✓ Cost-reduction is no longer available!?
- ✓ Post-K delayed max 2 years due to semiconductor/device technology problems.



## ● Seeking flop/s-oriented, accelerator-based exotic architectures?

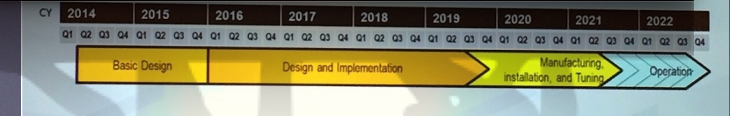
- with heterogeneity in computing and memory models, in particular, large-gap between local and remote, and between layers in the deep memory hierarchy
- co-design is intensively considered, but excessive co-design systems and their applications much more special
- ✓ efficiency of KNL in PL
- ✓ 0.3% efficiency of Sunway Taihulight and 1.5% of Exascale



## Will The Current Approach to Exascale Really Become a Help for HPC Community Entirely?!

## ● Still suffer from high operational cost mainly due to electricity expense, and who pay?

- ✓ The operation cost for 40MW of Post-K is affordable!?





# Scaling may be End, but Silicon is not End! And Use it Smart!

- ✓ We are facing the end of Moore's law due to the physical limitations, and the transistor cost is hard to reduce, however
- ✓ Silicon is still fundamental constructing material for computing platforms such as plastic, steel and concrete for automobiles, buildings and home appliances.

★ So, we have to become much more smart for design of Future HEC systems.

- Exploit sleeping flop/s efficiently by redesign/reinvent of memory subsystems to protect HEC systems from "Brain Infarction"



Use precious silicon budget (+ advanced device technologies) to effectively design mechanisms that can supply data to computing units smoothly.

From Brute Force to Smart Force!

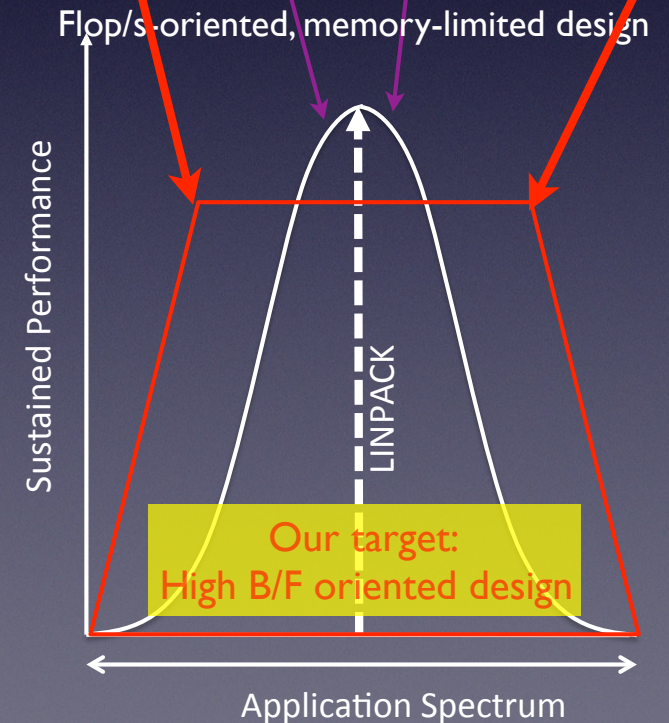
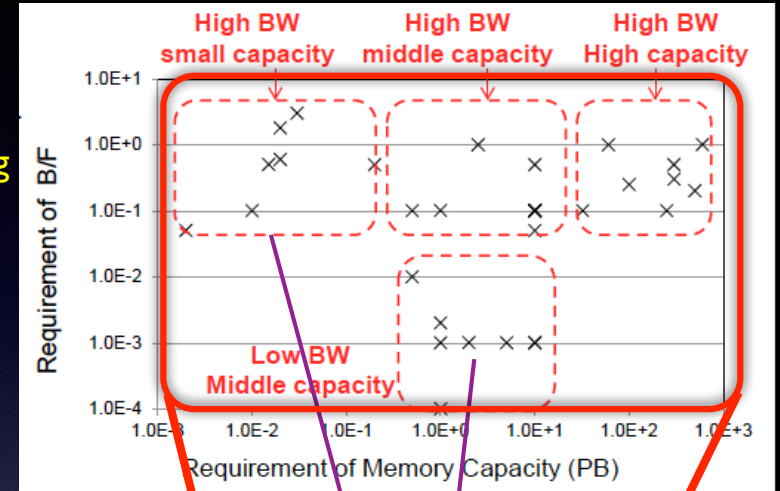
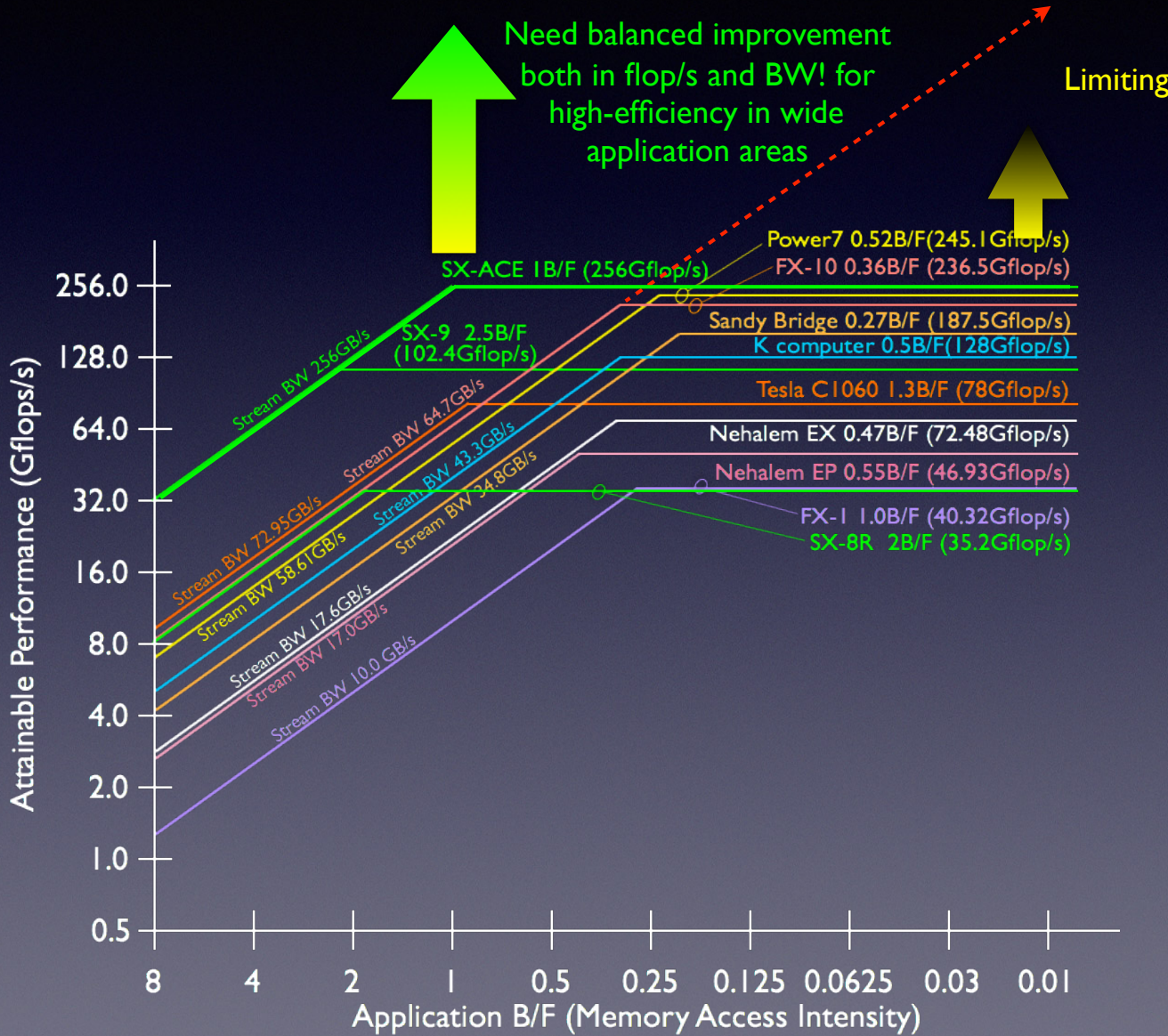
New Moore's law would be  
Productivity doubles every two year?!



Source: Toyota



# Not Peak Performance, Turn Memory-BW into Sustained Performance!





# Tohoku Univ-NEC Joint Research Division of HPC Technologies and Applications

★ **Founded in June, 2014, 4-year period**

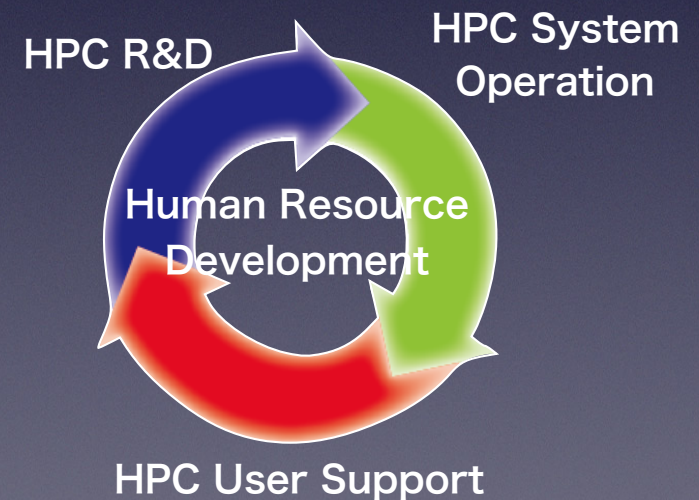
★ **Objectives**

- R&D on HPC technologies to exploit high-sustained performance of science and engineering applications on current HPC Systems and to realize Future HPC Systems targeting at 2020.
- ✓ Evaluation and Improvement of the current HPC environments through migration of SX-9 applications to SX-ACE
- ✓ Detailed Evaluation and Analysis of Modern HPC Systems, not only Vector Systems but also Scalar-Parallel and Accelerator-Based Systems
- ✓ Feasibility study of a future highly balanced HPC system for high sustained performance of practical applications in the post-peta scale era

★ **Faculty Members**

- Hiroaki Kobayashi, Professor and division director
- Hiroyuki Takizawa, Associate Professor
- Ryusuke Egawa, Associate Professor
- Akihiko Musa (NEC), Visiting Professor
- Mitsuo Yokokawa (Kobe Univ), Visiting Professor
- Shintaro Momose (NEC), Visiting Associate Professor
- Masayuki Sato, Assistant Professor

- ✓ In collaboration with visiting researchers from NEC and the technical staff of Cyberscience Center





## Summary

- ★ **SX-ACE shows high sustained performance compared with SX-9 and other modern HEC systems**
  - ✓ achieved the same single core performance in practical applications even with 60% of peak performance of SX-9
  - ✓ No l. computing-efficiency and power-efficiency in the HPCG Benchmark ranking
  - ✓ Pave the way to a new social infrastructure for homeland safety in Japan
- ★ **Well balanced HEC systems regarding memory performance is the key to success for realizing high productivity in science and engineering simulations**
  - ✓ Think different with Smart Force from Brute Force in HPC design
  - ✓ Quality, not Quantity for productive HPC!
  - ✓ Demands for Supercomputers for the rest of us, especially for 2020 and beyond!