inspur 浪潮

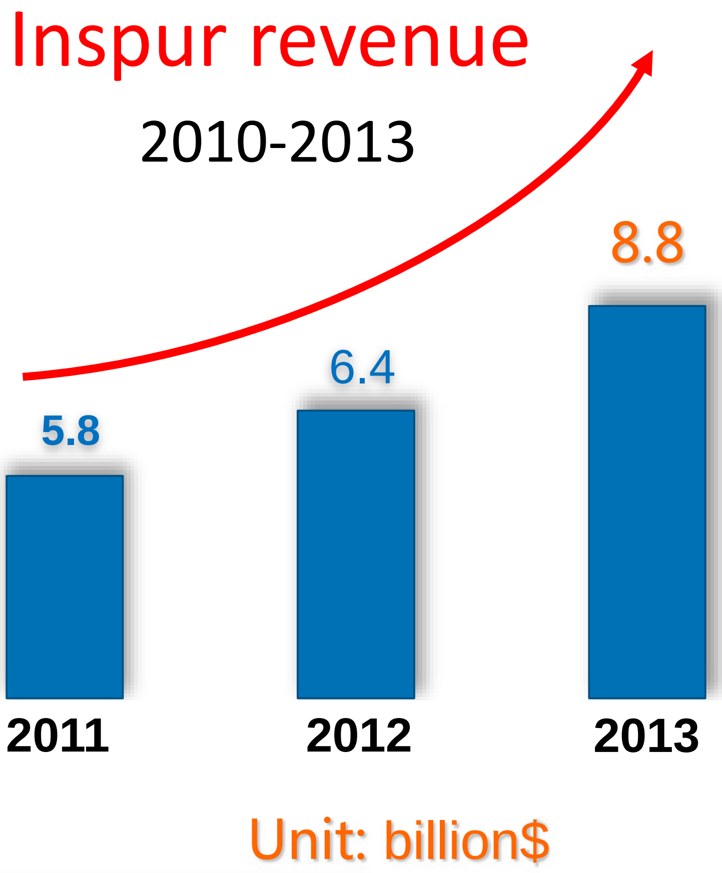
創中造國

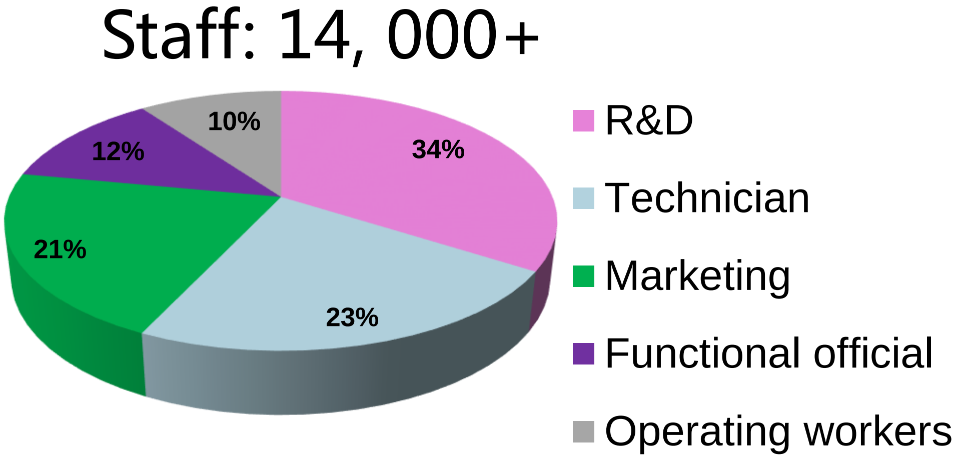
Tianhe-2, the world’s fastest supercomputer

Shaohua Wu

Senior HPC application development engineer

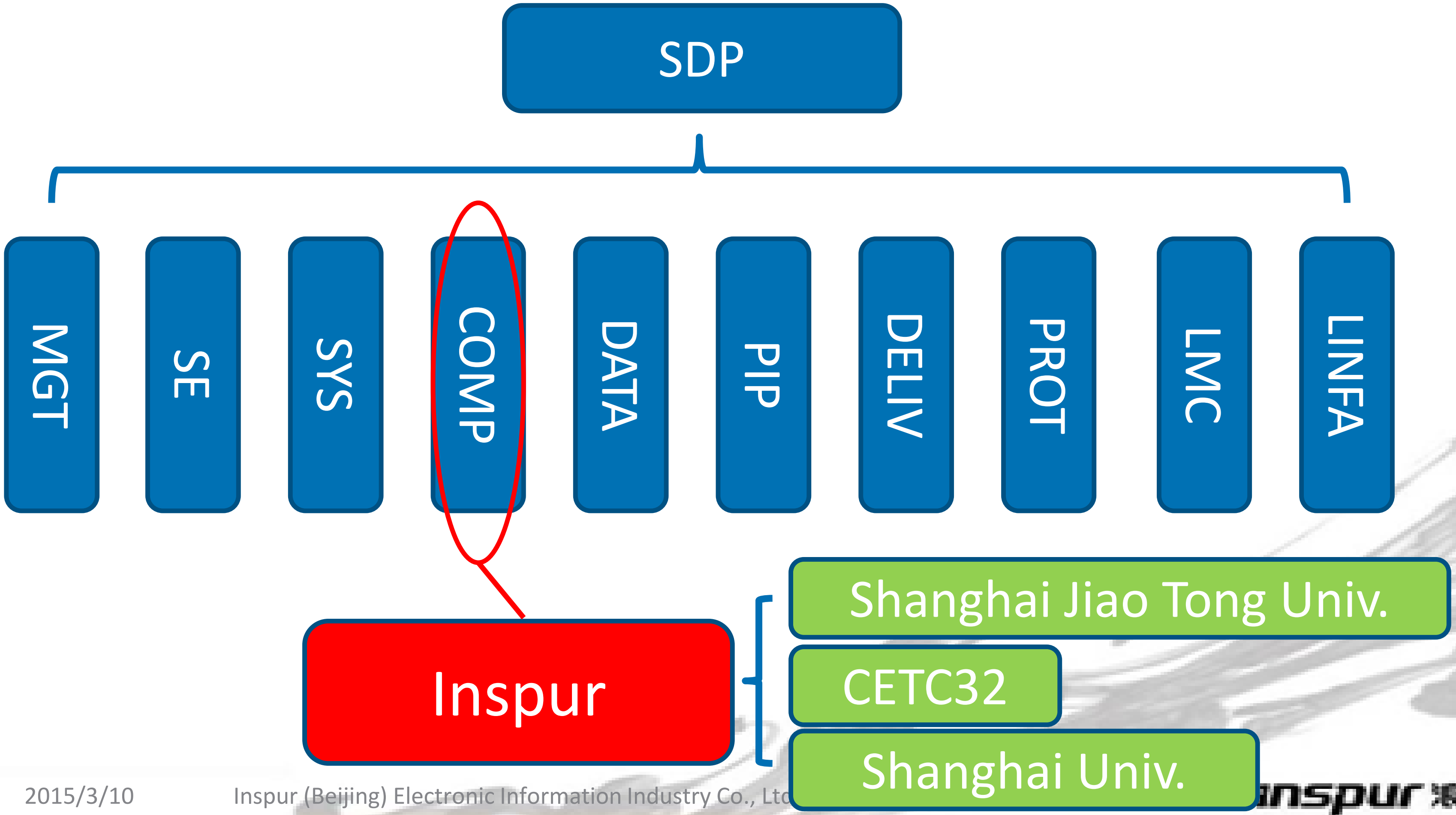
Inspur





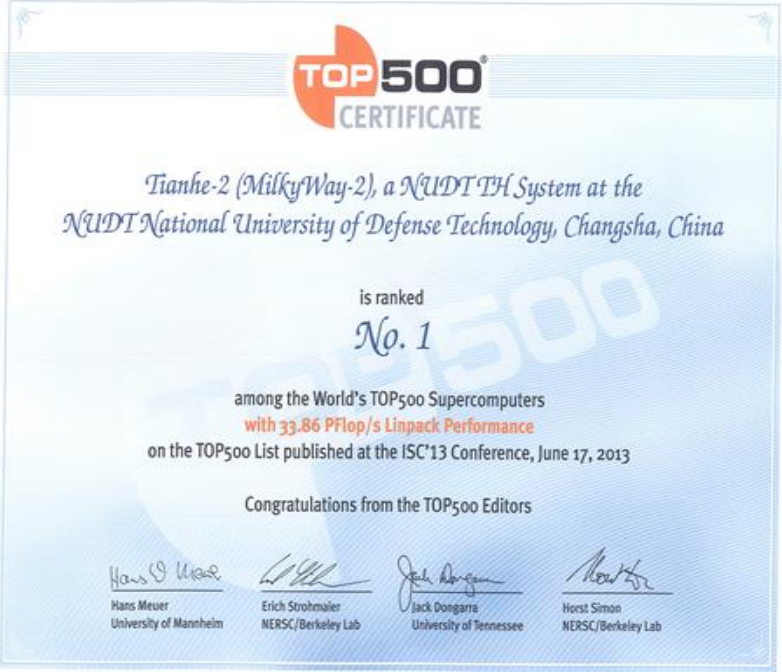
- Leading HPC system vendor in China
- Largest server manufacturer
- #1 native brand server vendor, 17 years
- #1 native brand storage vendor, 9 years

Inspur’s role in Chinese SDP consortium



Tianhe-2 (MilkyWay II)

- No.1 @Top500 since 6, 2013
- Co-Developed by NUDT and Inspur



TOP500 CERTIFICATE

Tianhe-2 (MilkyWay-2), a NUDT TH System at the NUDT National University of Defense Technology, Changsha, China

is ranked

No. 1

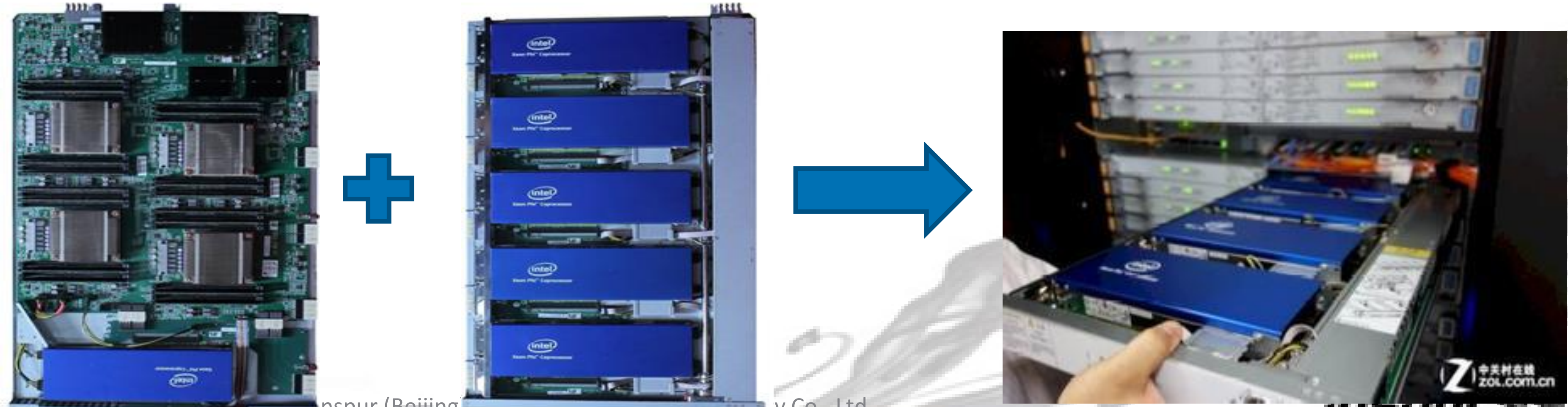
among the World's TOP500 Supercomputers
with 33.86 PFlop/s Linpack Performance
on the TOP500 List published at the ISC'13 Conference, June 17, 2013

Congratulations from the TOP500 Editors

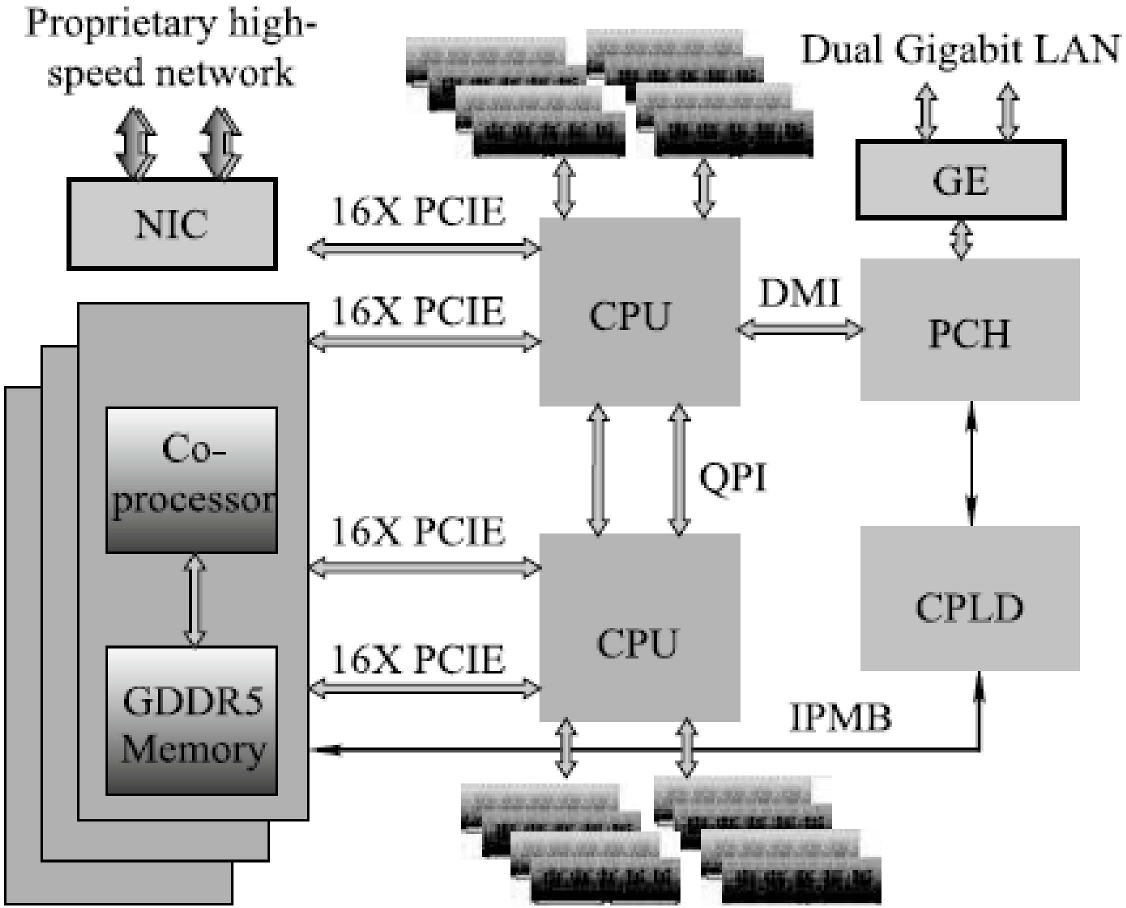
Hans Meuer University of Mannheim	Erich Strohmaier NERSC/Berkeley Lab	Jack Dongarra University of Tennessee	Horst Simon NERSC/Berkeley Lab

Compute blade of Tianhe-2

- 125 Rack
  - Each rack has 8 frame, each frame has 8 blade.
- Compute Blade
  - CPM module + APU module
  - 128GB memory, 2 comm. Ports



Compute node of Tianhe-2



Proprietary interconnection network

- High-radix Network Routing Chips (NRC)
  - Feature size: 90 nm
  - Die size: 17.16mm*17.16mm
  - 2577 pins
  - Throughout of single NRC: 2.56 Tbps

- High-speed Network Interface Chips (NIC)
  - Same feature size and package
  - Die size: 10.75mm*10.76mm
  - 675 pins, PCI-E G2 × 16

- MPI Performance
  - Broadcast ~6.36 GB/s. Latency: ~ 2 us

I/O system: H2FS

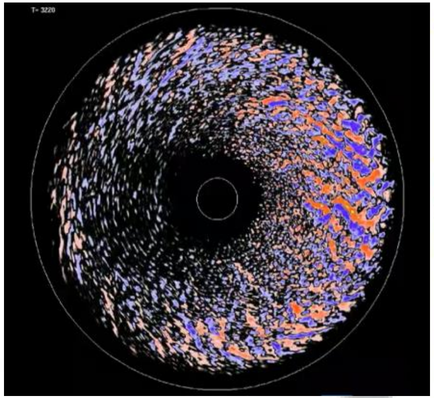
- Hybrid hierarchy file system
  - Co-operates node-local storage and shared storage
- Storage subsystem
  - 256 I/O nodes
  - 64 storage servers
  - 24 storage racks
  - 12.4 PB

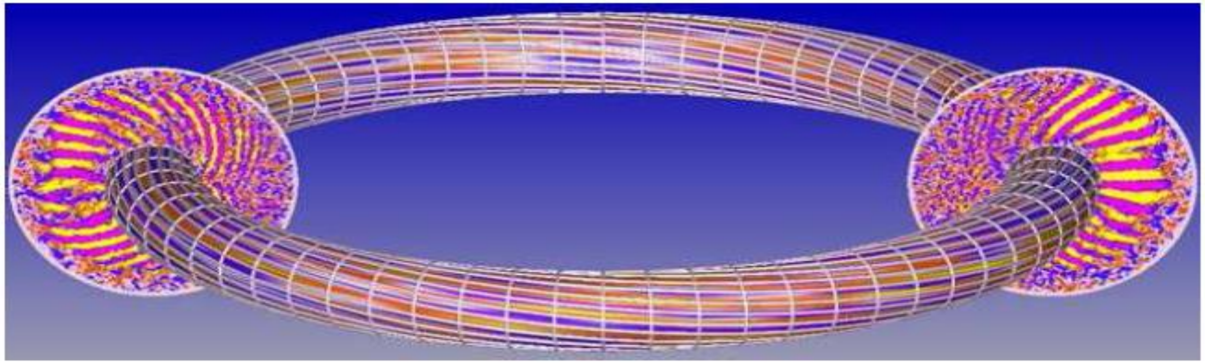
Cooling system

- Cooling type
  - Close-coupled chilled water cooling
- Customized Liquid Cooling Unit
  - High Cooling Capacity: 80kW
- NSCC-GZ uses city cooling system to supply cool water to LCUs
- Power consumption
  - 17.6MW
  - 24 MW including the cooling system

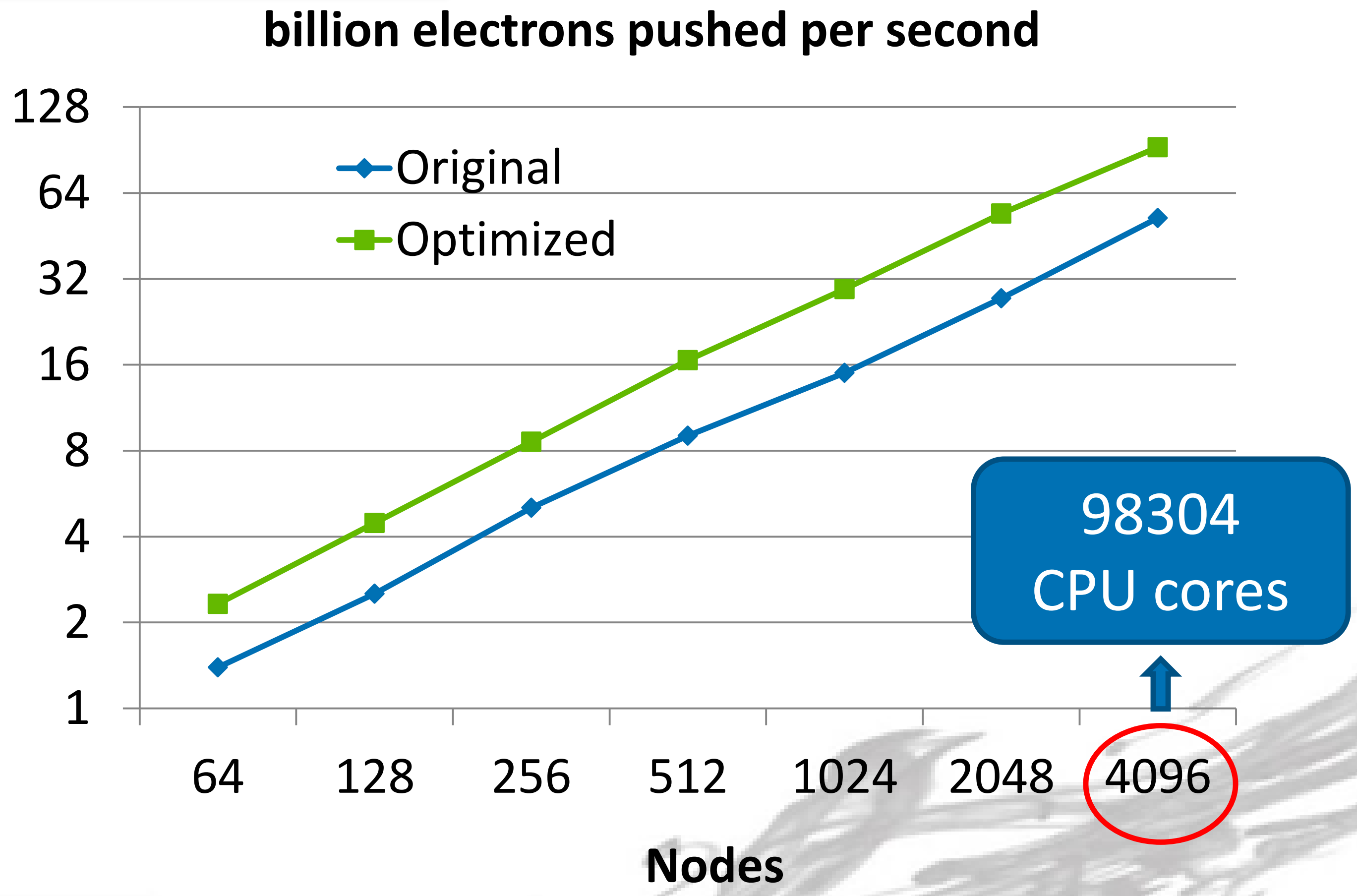
Optimization of GTC on Tianhe-2

- Gyrokinetic toroidal code (GTC) is a massively parallel code for turbulence simulation in support of the burning plasma experiment in international fusion collaboration (ITER), the crucial next step in the quest for fusion energy.
- we **innovatively redesign** the MPI communication in GTC that simplifies the original multiple MPI communication into once, which greatly improve the MPI communication efficiency.

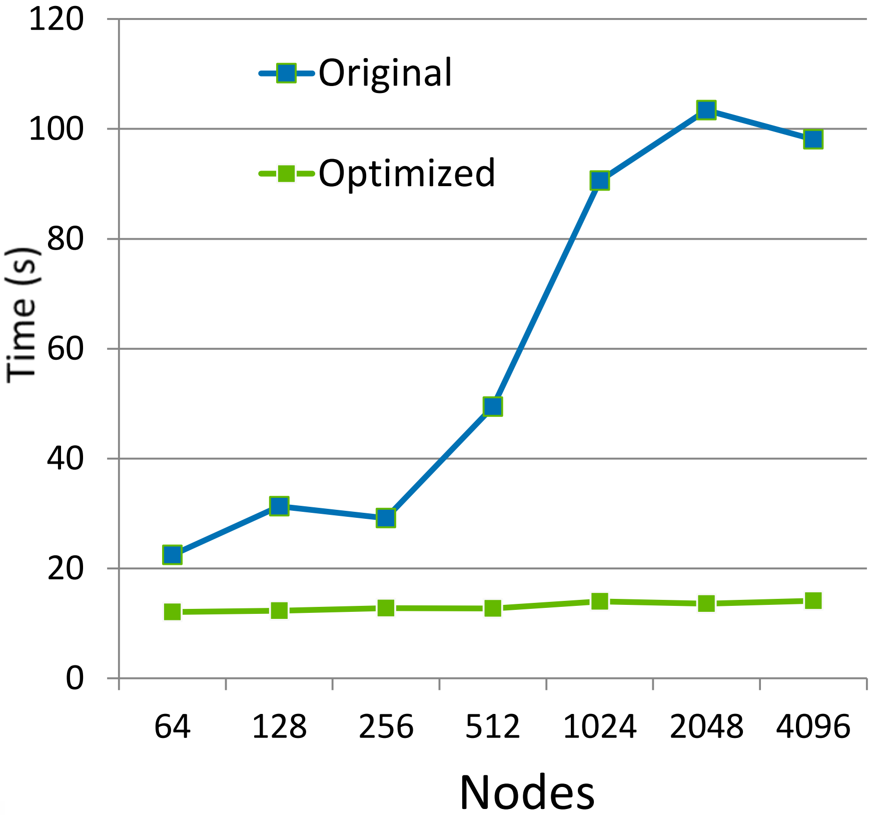


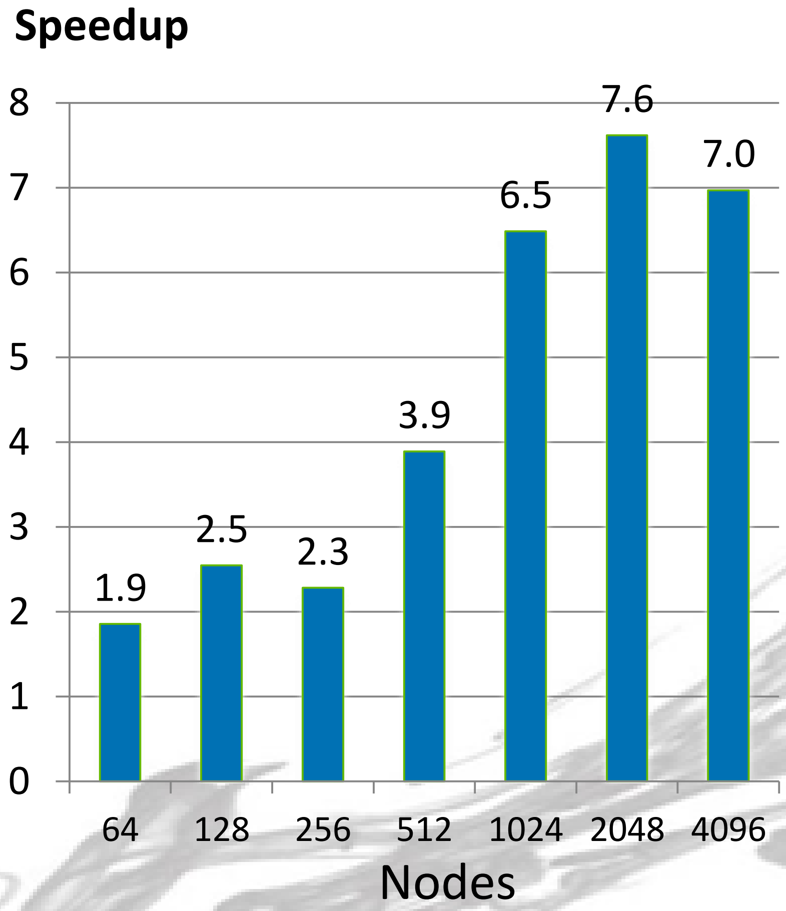


Testing results of Tianhe-2: Scalability

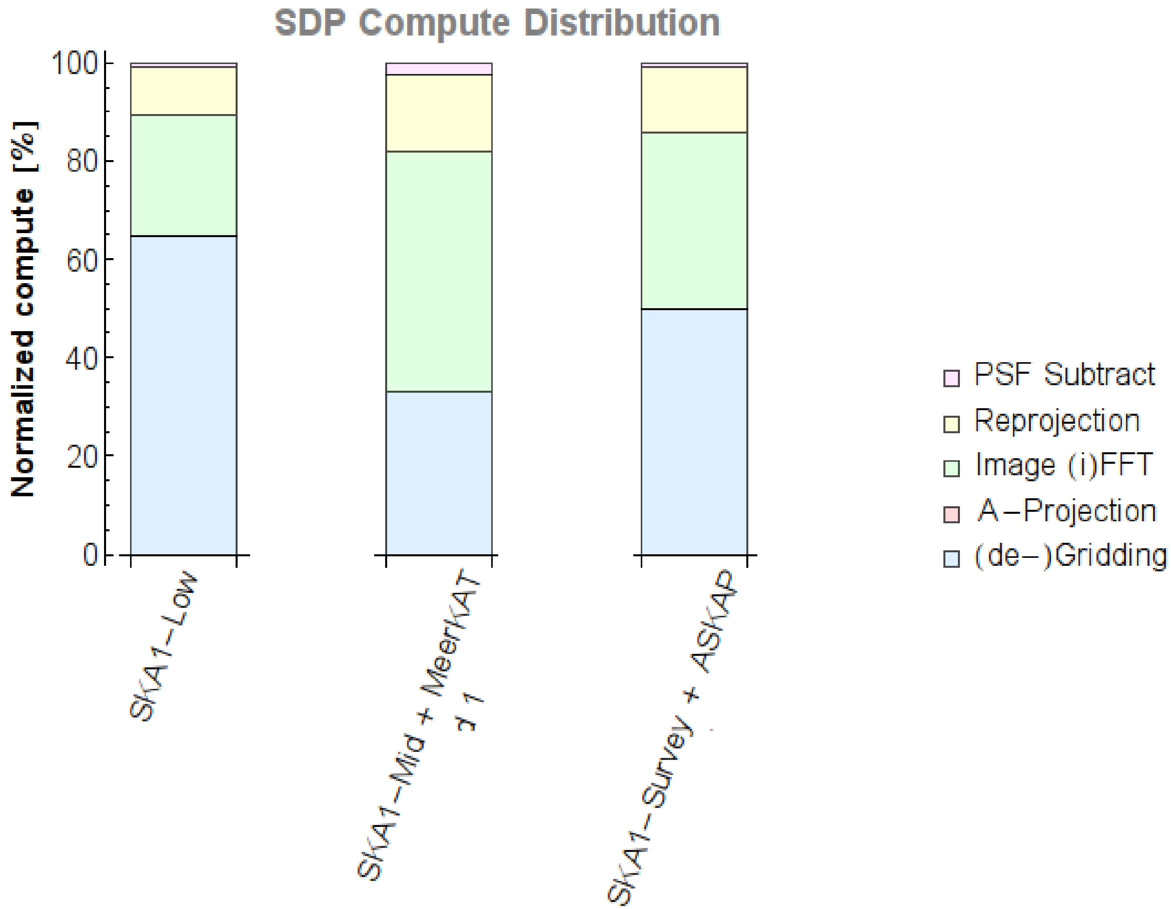


Testing results of Tianhe-2: Speedup





Optimization of Gridding algorithm on CPU+MIC



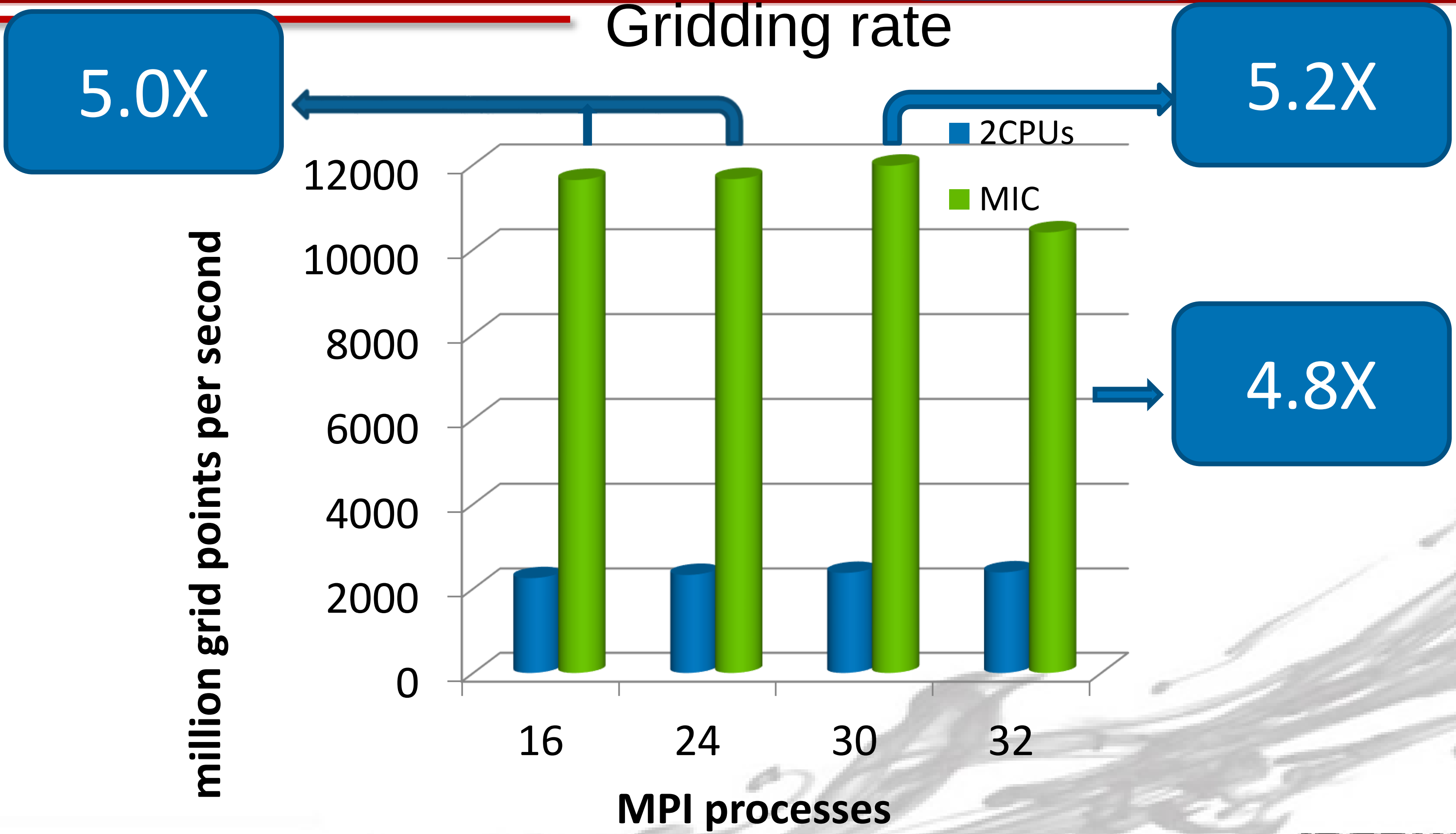
Testing platform

CPU	Intel$^{(R)}$ Xeon$^{(R)}$ E5-2650 v2, 8cores, 2.6GHz
Memory	128GB memory， 1333MHz
MIC	Intel Xeon Phi 7120P， 61cores， Frequency: 1.25GHz， GDDR Speed: 5.5GT/s， 16GB memory
Network	FDR InfiniBand 56Gb/s

OS		Red Hat 6.4, 2.6.32-358.el6.x86_64
MIC driver		MPSS： 3.2.1-1， Flash Version: 2.1.02.0390
Compiler	icpc	Intel(R) 64, Version 14.0.2.144 Build 20140120
	mpi	Intel(R) MPI Library for Linux* OS, Version 4.1 Update 3 Build 20140226

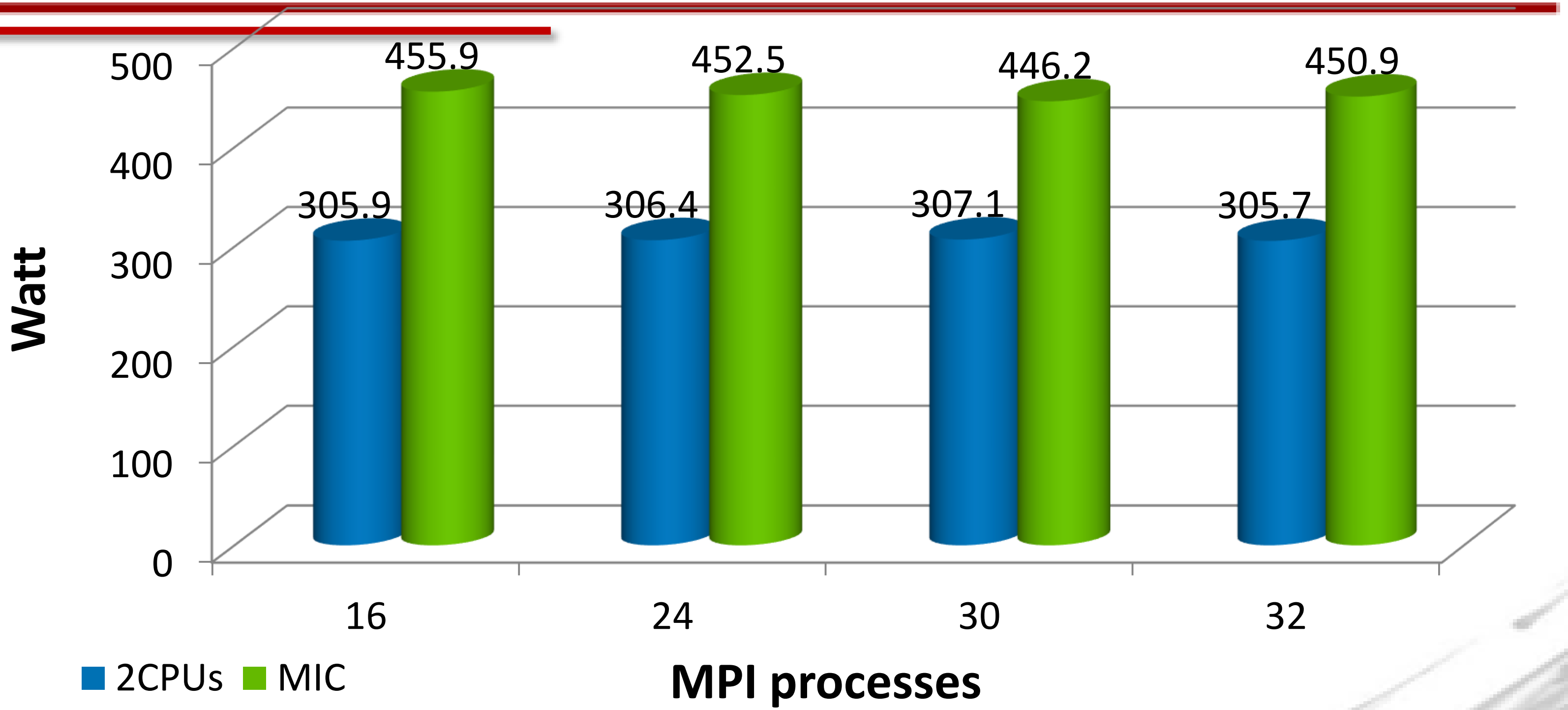
1000 Pflops

200 Pflops

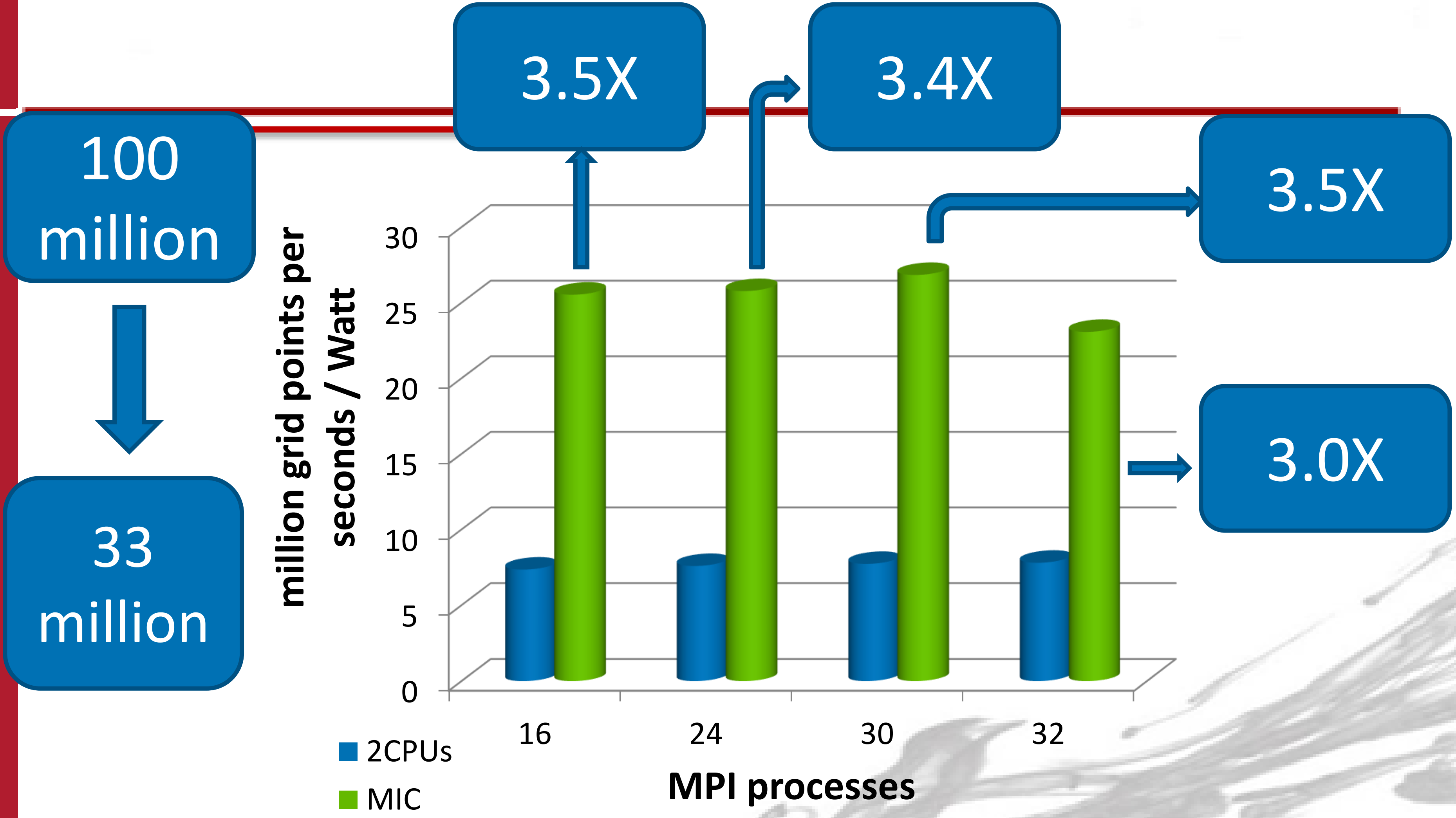


Gridding rate

Energy consumption (Watt)



- The idle state of CPU+MIC is: 159.7 Watt
- The idle state of CPU (removing the MIC) is： 136.2 Watt
- The idle state of MIC： 23.5 Watt



100 million

33 million

3.5X

3.4X

3.5X

3.0X

中國創造

inspur 浪潮

Shaohua Wu
wushh@inspur.com