FUJITSU

shaping tomorrow with you

Key Software Technologies for Future High Performance Computing

Copyright 2013 FUJITSU LIMITED

Key software issues for future HPC

FUJITSU

- Scalability
- Manageability
- Power Efficiency
- Productivity

Copyright 2013 FUJITSU LIMITED

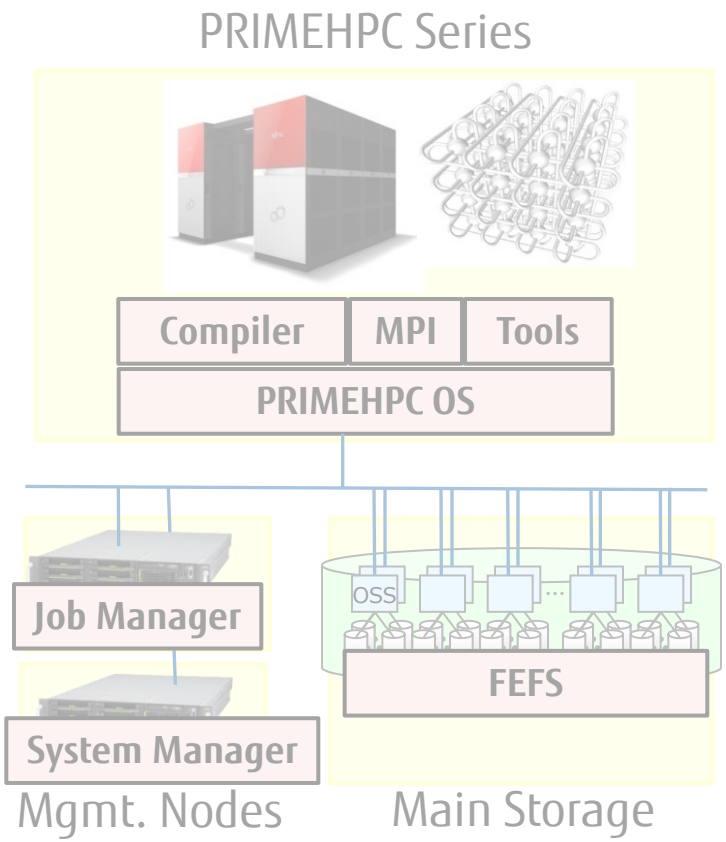
Key software issues for future HPC

FUJITSU

- Scalability
- Manageability
- Power Efficiency
- Productivity

Copyright 2013 FUJITSU LIMITED

Technologies for Scalability



Background:

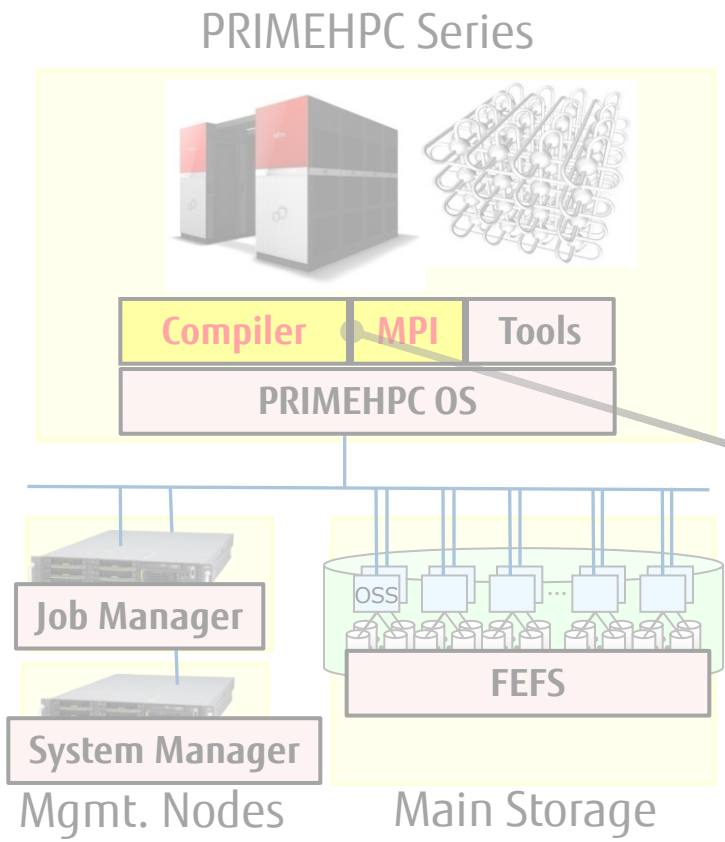
The performance are always top-priority on HPC application. Scalability is the most important for performance on massively parallel computing.

Issues in this area:

- **Improve application parallelization**
- **Reduce OS Jitter affection**
- **Keep I/O performance scalability**

Copyright 2013 FUJITSU LIMITED

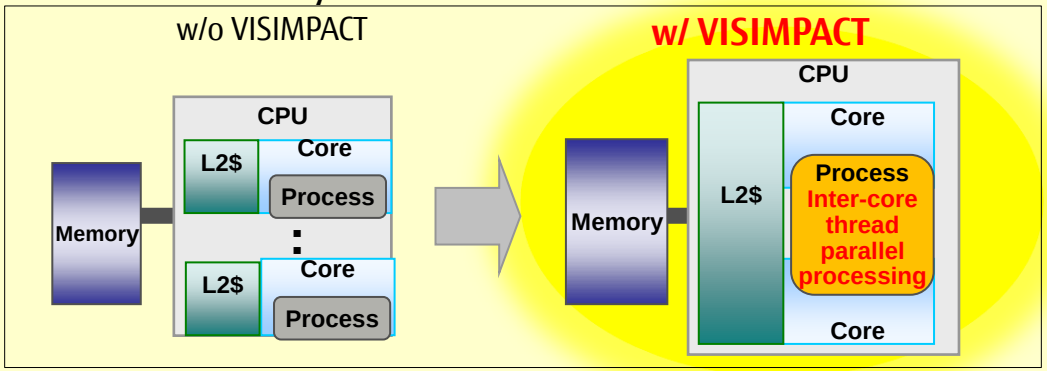
Technologies for Scalability



Improve application parallelization

Technology: VISIMPACT + Tofu Optimized MPI

- Easy and efficient inter-core Parallelization automated by VISIMPACT



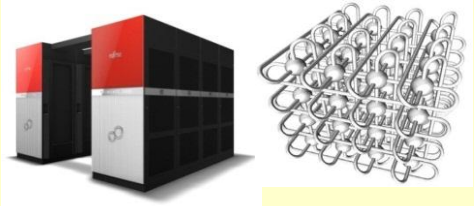
- Threads-Processes hybrid parallelization with Tofu optimized MPI library.

Copyright 2013 FUJITSU LIMITED

Technologies for Scalability

FUJITSU

PRIMEHPC Series



Compiler | MPI | Tools

PRIMEHPC OS

Job Manager

System Manager

Mgmt. Nodes

oss ... FEFS

Main Storage

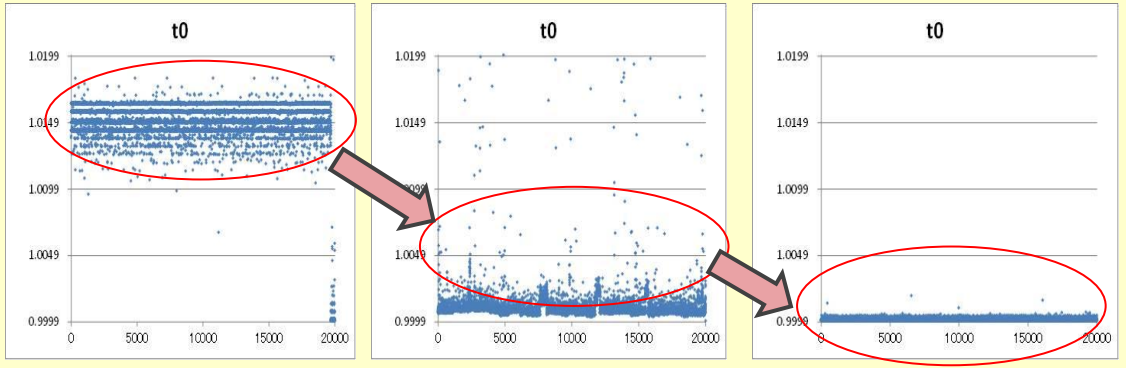
■ Reduce OS Jitter affection

Technology: Tuned Linux OS for PRIMEHPC

minimized negative effect of OS jitter ultimately by:

- Core-binding technology
- Deliberately selected and tuned system service

x86 cluster w/o TCS | x86 cluster w/ TCS | FX10 OS w/ TCS



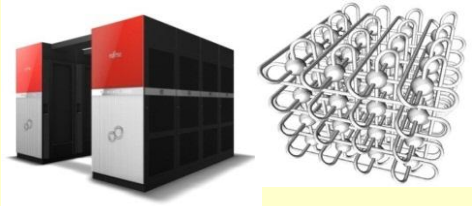
TCS: Technical Computing Suite (Fujitsu's System Software Product)

Copyright 2013 FUJITSU LIMITED

Technologies for Scalability

FUJITSU

PRIMEHPC Series



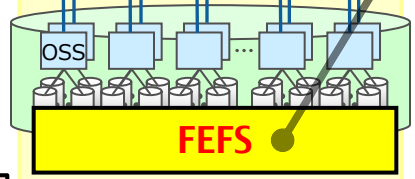
Compiler | MPI | Tools

PRIMEHPC OS

Job Manager

System Manager

Mgmt. Nodes



Main Storage

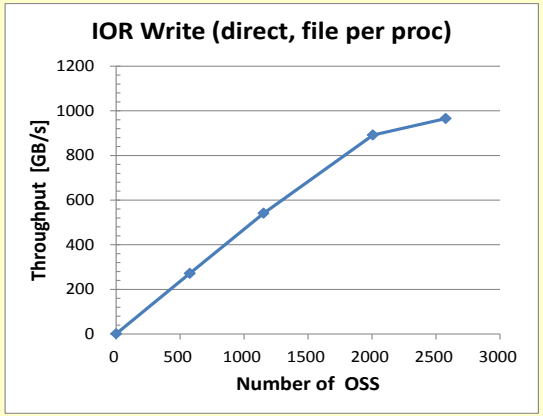
■ Keep I/O performance scalability

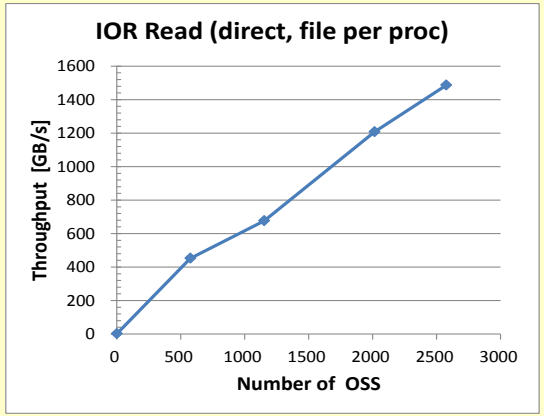
Technology: FEFS (Fujitsu Exabyte File System)

Lustre based scalable file system

- Supports up to 8 Exa bytes capacity
- Achieves superb performance on K computer

  Write : **965 GB/s** Read : **1,486 GB/s**





Collaborative work with RIKEN

Copyright 2013 FUJITSU LIMITED

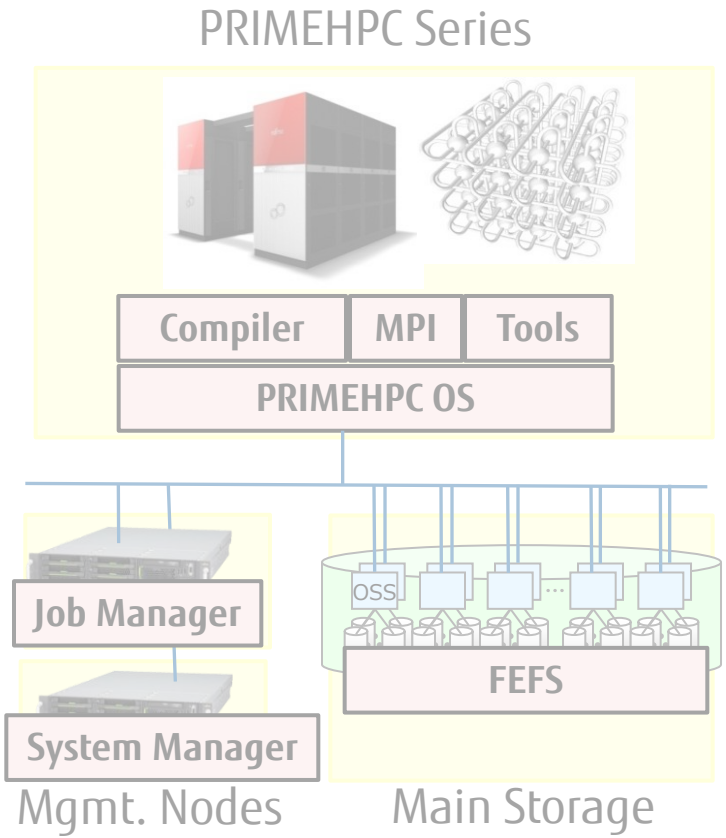
Key software issues for future HPC

FUJITSU

- Scalability
- **Manageability**
- Power Efficiency
- Productivity

Copyright 2013 FUJITSU LIMITED

Technologies for Manageability



Background:

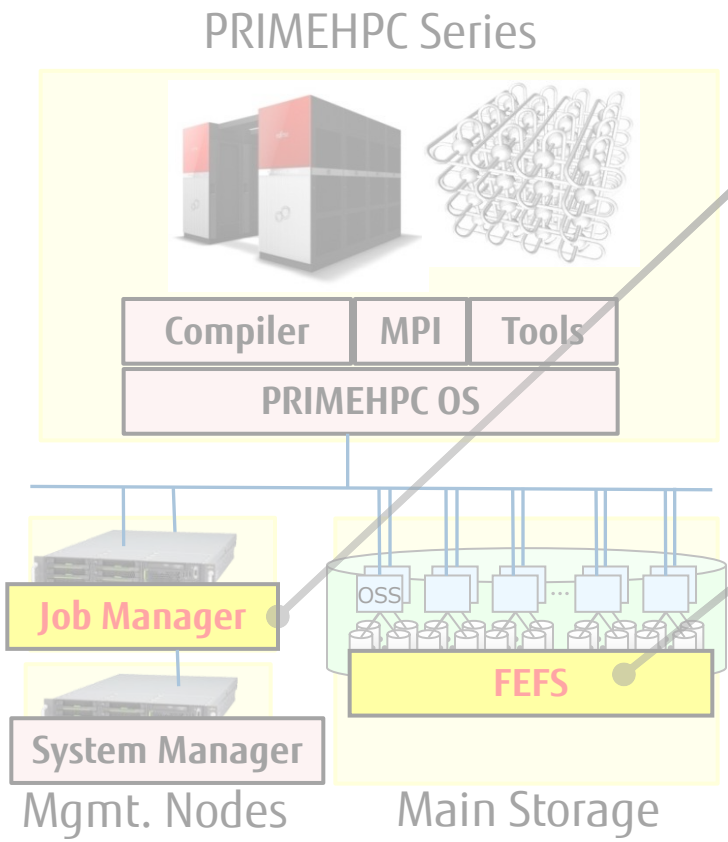
For system administrator, managing system is a key issue. We have already achieved nearly 100,000 nodes system on K computer.

Issues in this area:

- **Availability**
- **Operability**

Copyright 2013 FUJITSU LIMITED

Technologies for Manageability



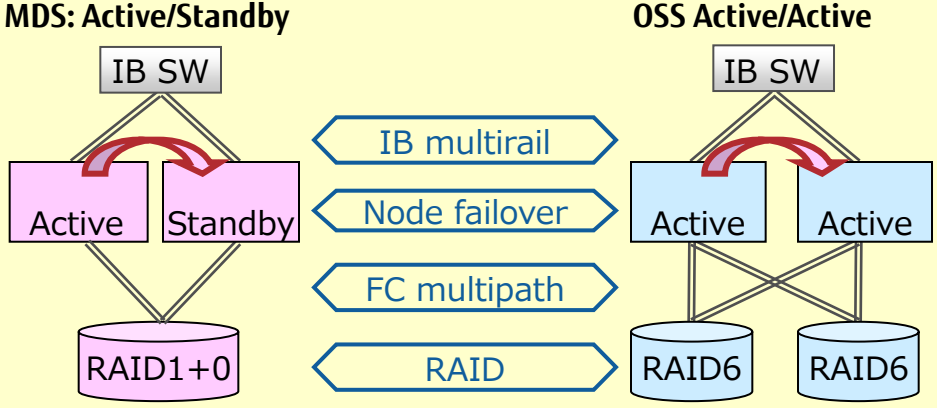
■ Availability

Technology: Automatic Management Node failover

- Immediate failover by Hot Stand-by

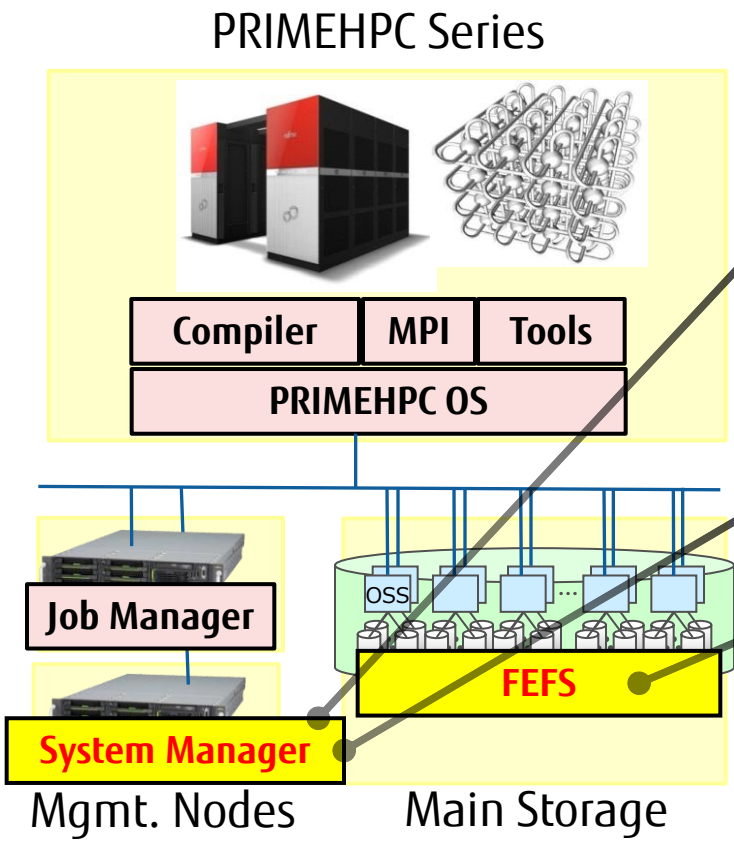
Technology: Automatic MDS/OSS failover

- Monitored by TCS system manager



Copyright 2013 FUJITSU LIMITED

Technologies for Manageability



■ Operability

Technology: Centric Management provides single system image for:

- ➢System Installation / Update
- ➢Node status Monitoring (Hardware / Software)
- ➢Power Control / Monitoring
- ➢Support PRIMEHPC / x86 Hybrid Cluster system

Technology: Flexible cluster management provides various physical/logical partitioning.

Technology: QoS/Directory Quota on FEFS facilitates sharing global storage across multi cluster system.

Copyright 2013 FUJITSU LIMITED

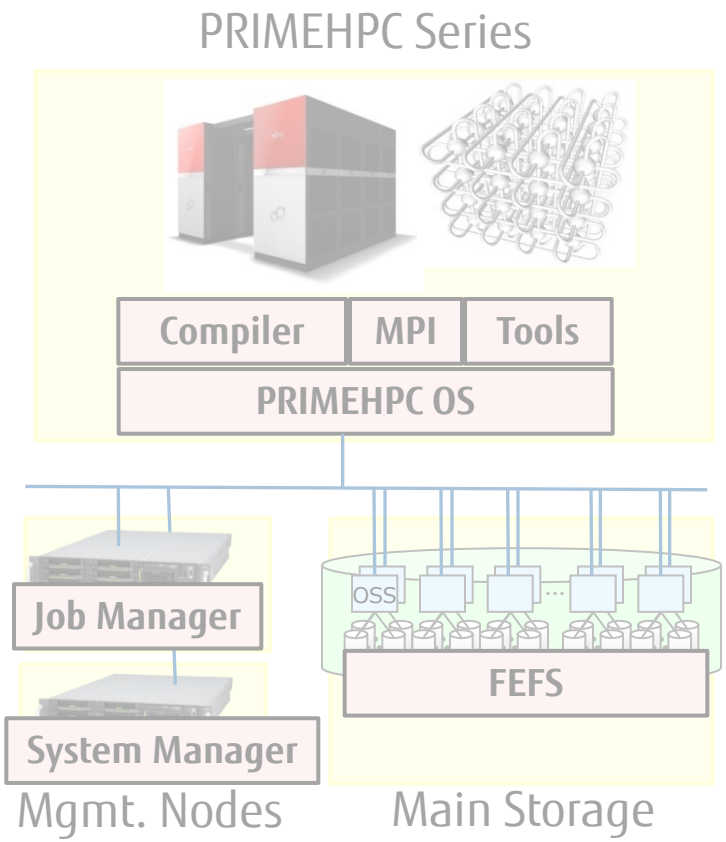
Key software issues for future HPC

FUJITSU

- Scalability
- Manageability
- **Power Efficiency**
- Productivity

Copyright 2013 FUJITSU LIMITED

Technologies for Power Efficiency



Background:

$$\text{Power Efficiency} = \frac{\text{“Actual” Throughput}}{\text{Total system power}}$$

Customer Requirement:

✓Increasing “actual” system throughput

✓Keeping total power at target value

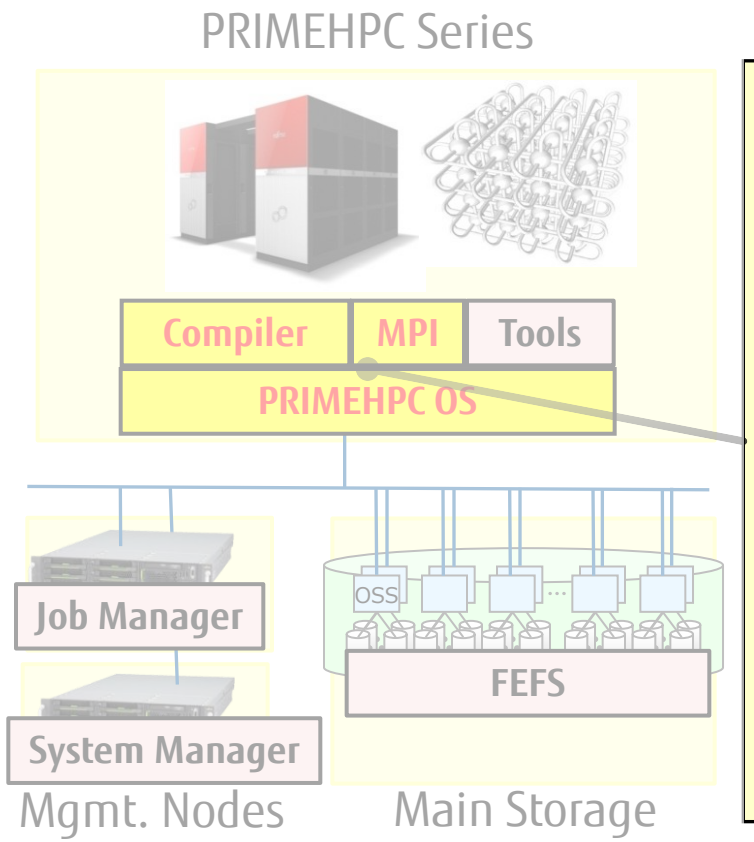
Issues in this area:

- Maximize application efficiency
- Maximize Resource utilization
- System-wide Power Management

Copyright 2013 FUJITSU LIMITED

Technologies for Power Efficiency

FUJITSU



■ Maximize application efficiency

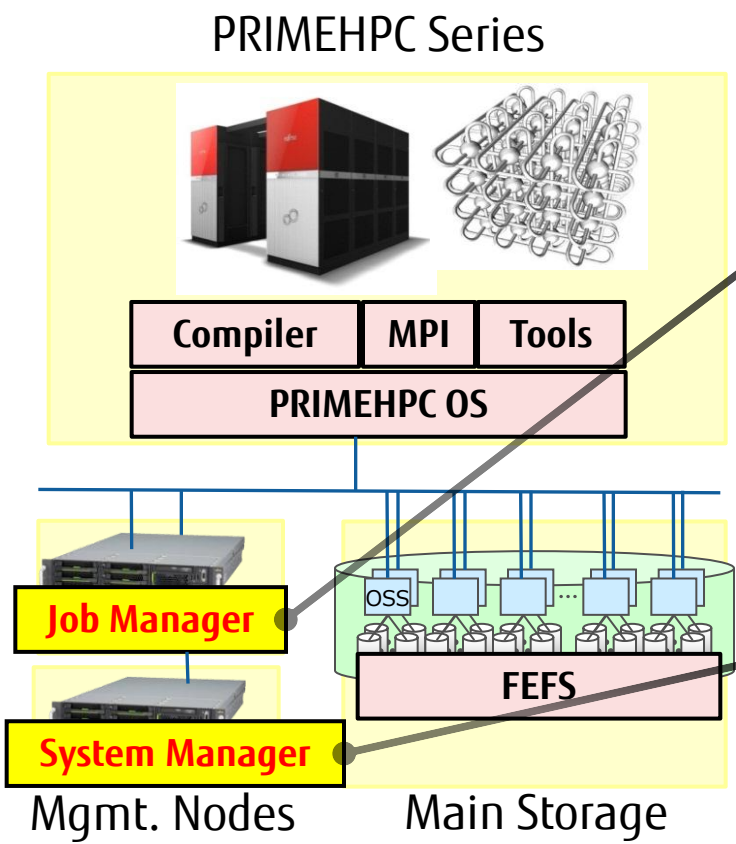
Technology: **Optimized OS and languages** achieves good efficiency for many applications.

Application	Nodes	Efficiency
LINPACK	88,128	93%
NICAM	81,920	8%
Seism3D	82,944	18%
PHASE	82,944	20%
RSDFT	82,944	52%
FrontFlow/blue	80,000	3%
Lattice QCD	82,944	16%
ZZ-EFSI	82,944	46%

Table is provided by Dr. Minami of RIKEN

Copyright 2013 FUJITSU LIMITED

Technologies for Power Efficiency



Maximize Resource Utilization

Technology: **Various job allocation** method to increase node/core utilization even on Torus system

- Torus mode/Mesh mode allocation
- Node simplex/share allocation
- Heterogeneous hybrid parallel job allocation

System-wide Power Management

Technology: **Centric Power Control** helps to integrate center-wide power capping with:

- Interface to control power of nodes or storages
- Power consumption monitoring
- Control power saving mode w/ job manager

Copyright 2013 FUJITSU LIMITED

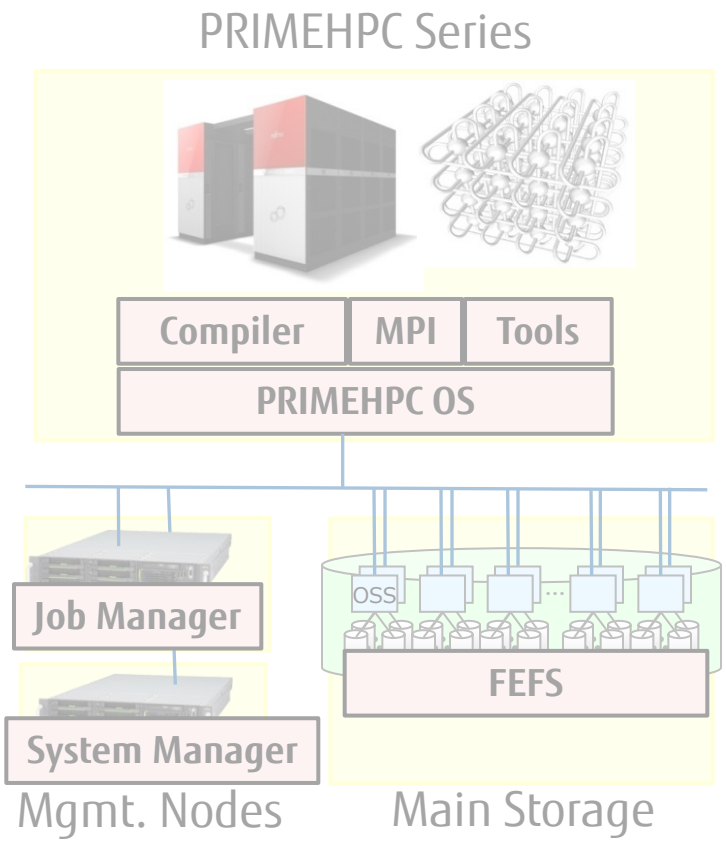
Key software issues for future HPC

FUJITSU

- Scalability
- Manageability
- Power Efficiency
- **Productivity**

Copyright 2013 FUJITSU LIMITED

Technologies for Productivity



Background:

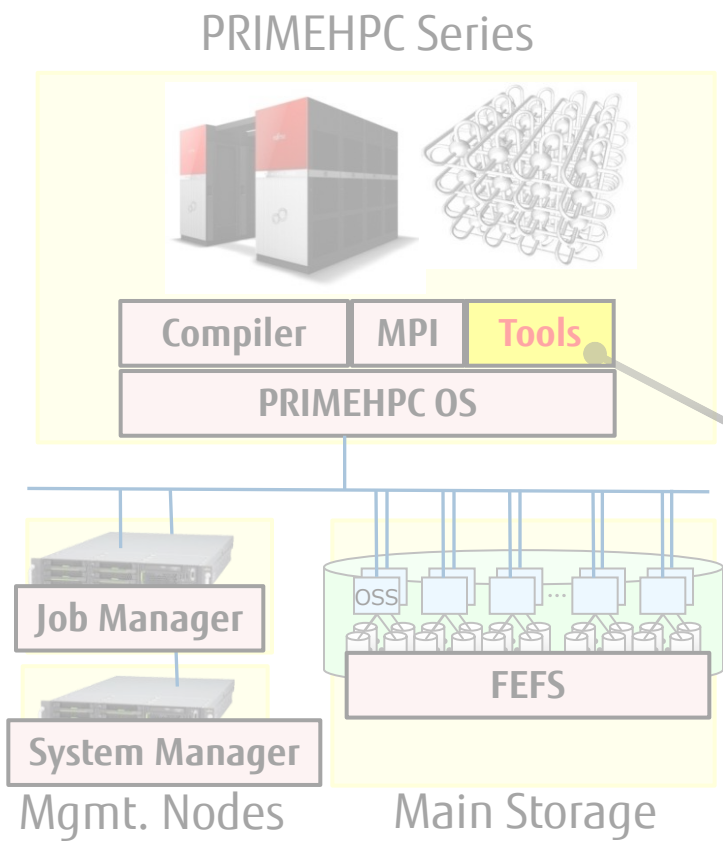
Because of increasing complexity of node and network architecture, developing applications become more difficult.

Issues in this area:

- **Tuning and Debugging**
- **Portability**

Copyright 2013 FUJITSU LIMITED

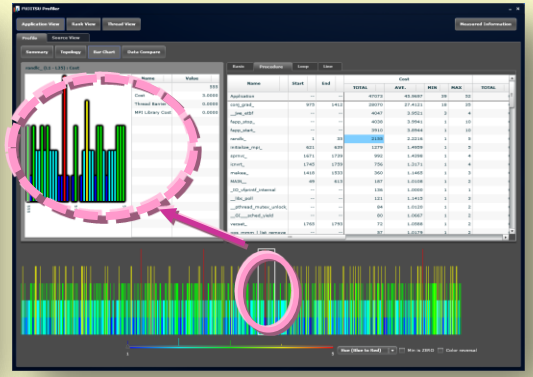
Technologies for Productivity



■ Tuning and Debugging

Technology:

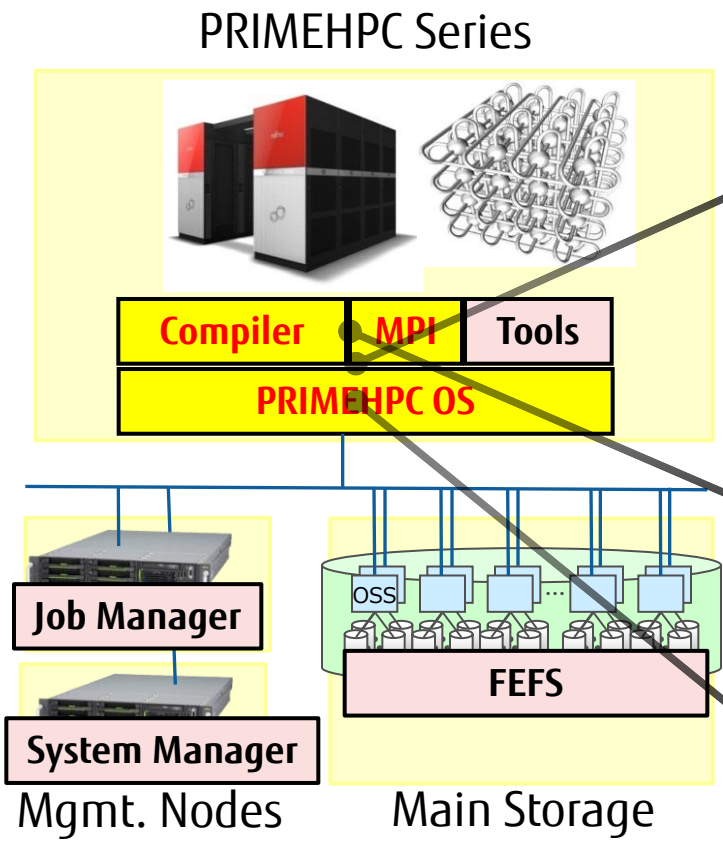
- Supports world's standard debugger (DDT)
- Profiler
  - GUI Detailed PA information
  - Optimize communication on Tofu
- Rank Mapping Optimization (RMATT)



Copyright 2013 FUJITSU LIMITED

Technologies for Productivity

FUJITSU



■ Portability

Technology: **Compatibility with K computer**
- Binary compatibility
- Same Architecture applied
  1cpu/node, VISIMPACT, Tofu interconnect

Technology: Continues to supports:
- The latest international standards
  Fortran 2008, C 11, C++ 11
- De facto standards
  GNU C/C++ extensions, OpenMP 4.0, MPI 3.0

Technology: Supports **Generic Linux OS**
- POSIX compliant Linux and generic libraries

Copyright 2013 FUJITSU LIMITED

FUJITSU

Now, we are ready for 100 PFlops!

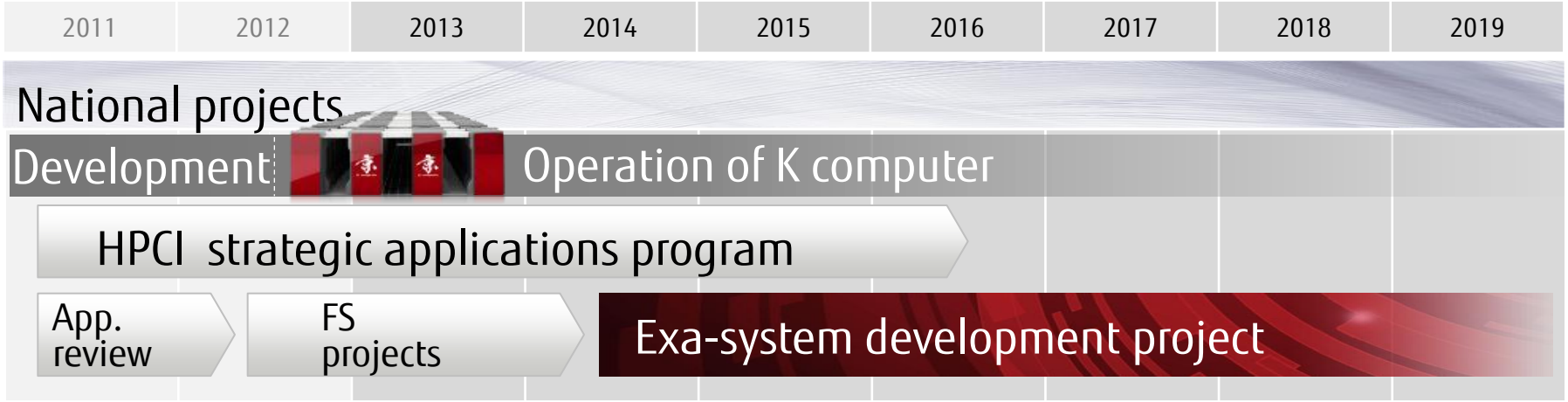
What's next?

Copyright 2013 FUJITSU LIMITED

Activities for Exascale Computing

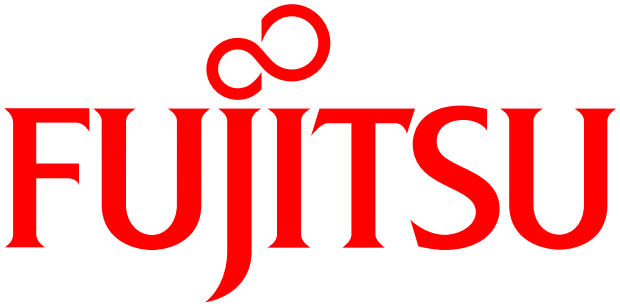
FUJITSU

- Roadmap of Japan's National Project



- Participate co-design for Exascale System software
  - Light-weight Micro kernel next to Linux
  - File I/O performance improvement

Copyright 2013 FUJITSU LIMITED



shaping tomorrow with you