

# HOKUSAI System

October 11, 2017

Information Systems Division, RIKEN

## 1.1 System Overview

The HOKUSAI system consists of the following key components:

- Massively Parallel Computer(GWMPC,BWMPC)
- Application Computing Server including ACS with Large Memory and ACS with GPU
- Front end servers that provide the users with the application interface for the system
- Two types of storages with different purposes, one of which is the Online Storage and the other of which is the Hierarchical Storage.

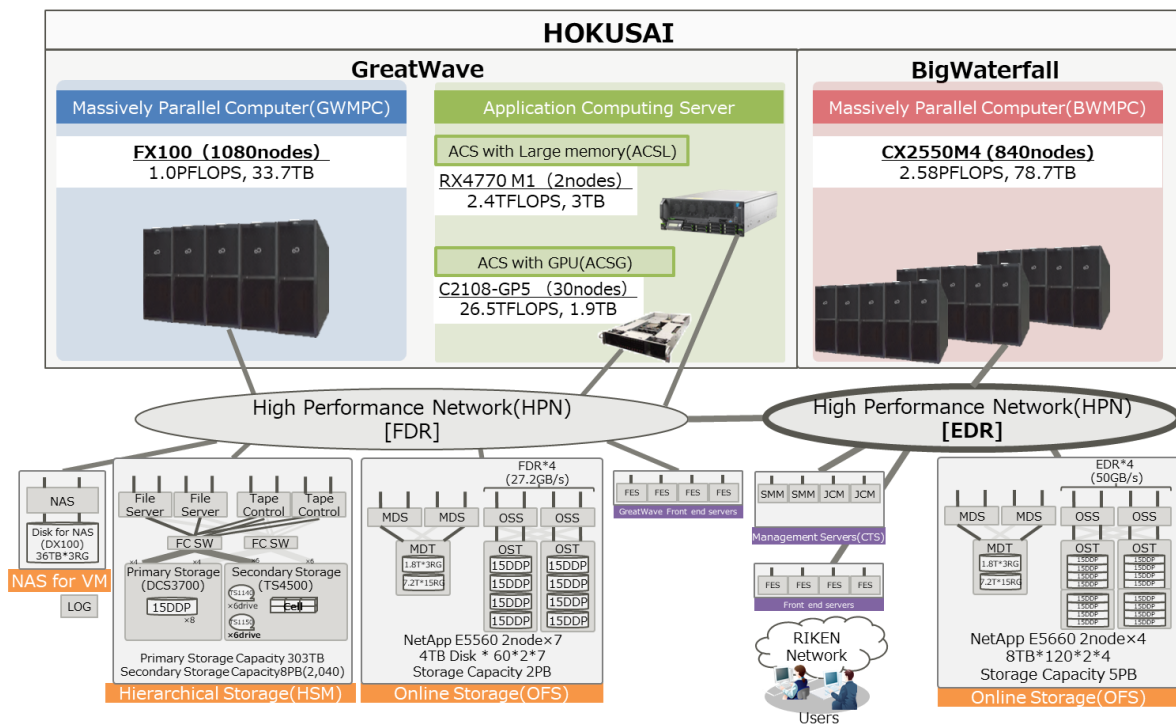


Figure 0-1 System diagram

The Massively Parallel Computer (GWMPC) comprises FUJITSU Supercomputer PRIMEHPC FX100. FX100, with high performance processors SPARC64 Xlfx and high performance memory systems, provides a theoretical peak performance of 1 TFLOPS (double precision) and the high memory bandwidth of 480 GB/s per one node with 32 cores/CPU's. The Massively Parallel Computer of 1,080 nodes provides

the a theoretical peak performance of 1 PFLOPS and a total memory capacity of 33.7 TB and uses the 6D mesh/torus network (Torus Fusion Interconnect 2\*<sup>1</sup>) to tightly connect each node with 12.5 GB/s high-speed link.

The ACS with Large memory (ACSL) comprises two nodes of PRIMERGY RX4770 M1. Each node provides a theoretical peak performance of 1.2 TFLOPS and a memory capacity of 1.5 TB. The ACS with GPU (ACSG) consists of thirty nodes of SGI C2108-GP5. Each node provides a theoretical peak performance of 883.2 GFLOPS and a memory capacity of 64 GB. Four NVIDIA Tesla K20X accelerators will be installed on each node of ACS with GPU. The InfiniBand FDR of 6.8 GB/s is used to connect each node to enable high performance communication and file sharing.

The Massively Parallel Computer (BWMPC) comprises 840 nodes of CX2550 M4. Each node provides a theoretical peak performance of 3.07 TFLOPS and a memory capacity of 96 GB. The InfiniBand EDR of 12.6 GB/s is used to connect each node to enable high performance communication and file sharing.

The storage environment consists of the Online Storage (OFS) and the Hierarchical Storage (HSM).

The Online Storage (OFS) is a high bandwidth online file system used for the users' home directories, the shared directories for projects and so on, and can be accessed from the Massively Parallel Computer, the Application Computing Server, and the front end servers. The total capacity is 7.2 PB.

The Hierarchical Storage (HSM) consists of the primary storage (cache disks) of 300 TB and the secondary storage (tape library devices) of 7.9 PB (uncompressed) and is the file system used to store large volumes of data files that should be retained for a long term. The users can read or write data to the tapes without manipulating the tape library devices.

You can access the HOKUSAI system using ssh/scp for login/file transfer, or using HTTPS for the User Portal and FUJITSU Software Development Tools (Development Tools). On the front end servers, you can mainly do the following:

- create and edit programs
- compile and link programs
- manage batch jobs and launch interactive jobs
- tune and debug programs

---

\*<sup>1</sup> Tofu Fusion Interconnect 2 is Fujitsu's proprietary high speed interconnect.



## 1.2 Hardware Overview

### 1.2.1 Massively Parallel Computer (GWMPC)

- Computing performance  
CPU: SPARC64™ XIfx (1.975 GHz) 1,080 units (1,080 CPUs, 34,560 cores)  
Theoretical peak performance: 1.092PFLOPS (1.975 GHz x 16 floating-point operations x 32 cores x 1,080 CPUs)
- Memory  
Memory capacity: 33.7 TB (32 GB x 1,080 units)  
Memory bandwidth: 480 GB/s/CPU  
Memory bandwidth/FLOP: 0.47 Byte/FLOP
- Interconnect (Tofu Interconnect 2)  
6D mesh/torus  
Theoretical link throughput: 12.5 GB/s x 2 (bidirectional)

### 1.2.2 Application Computing Server (ACS)

The Application Computing Server (ACS) consists of the ACS with Large memory (ACSL) and the ACS with GPU (ACSG).

#### 1.2.2.1 ACS with Large Memory (ACSL)

- Computing performance  
CPU: Intel Xeon E7-4880v2 (2.50 GHz) 2units (8 CPUs, 120 cores)  
Theoretical peak performance: 2.4 TFLOPS (2.5 GHz x 8 floating-point operations x 15 cores x 8 CPUs)
- Memory  
Memory capacity: 2 TB (1TB x 2 units)  
Memory bandwidth: 85.3 GB/s/CPU  
Memory bandwidth/FLOP: 0.28 Byte/FLOP
- Local disk  
Disk capacity: 3.6 TB ((300 GB x 2 + 1.2 TB) x 2 units)
- Interconnect  
FDR InfiniBand  
Theoretical link throughput: 6.8 GB/s x 2 paths x 2 (bidirectional)

#### 1.2.2.2 ACS with GPU (ACSG)

- Computing performance

CPU: Intel Xeon E5-2670 v3 (2.30GHz) 30 units (60 CPUs, 720 cores)

Theoretical peak performance: 26.4 TFLOPS (2.3 GHz x 16 floating-point operations x 12 cores x 60 CPUs)

- Memory

Memory capacity: 1.8 TB (64 GB x 30 units)

Memory bandwidth: 68.2 GB/s/CPU

Memory bandwidth/FLOP: 0.15 Byte/FLOP

- Local disk

Disk capacity: 18 TB ((300 GB x 2) x 30 units)

- Interconnect

FDR InfiniBand

Theoretical link throughput: 6.8 GB/s x 2 (bidirectional)

- Accelerator

NVIDIA Tesla K20X x 4 devices/node

#### 1.2.2.3 Massively Parallel Computer (BWMPC)

- Computing performance

CPU: Intel Xeon Gold 6148 (2.4GHz) 840 units (1680 CPUs, 33,600 cores)

Theoretical peak performance: 2.58 PFLOPS (2.4 GHz x 32 floating-point operations x 20 cores x 1680 CPUs)

- Memory

Memory capacity: 78.7 TB (96 GB x 840 units)

Memory bandwidth: 255GB/s/NODE

Memory bandwidth/FLOP: 0.08Byte/FLOP

- Local disk

Disk capacity: 100.8TB (120GB x 30 units)

- Interconnect

InfiniBand EDR

Theoretical link throughput: 12.6 GB/s x 2 (bidirectional)

### 1.3 Software Overview

The softwares available on the HOKUSAI system are listed as follows:

Table 0-1 Software overview

Category	Massively Parallel Computer (GWMP)	Massively Parallel Computer (BWMP)	Application Computing Server (ACS)	Front End Servers
OS	XTCOS (OS for FX100) (Linux kernel version 2.6)	Red Hat Enterprise Linux 7(x 56nodes) CentOS7(x 784nodes) (Linux kernel version 3.10)	Red Hat Enterprise Linux 7 (Linux kernel version 3.10)	Red Hat Enterprise Linux 7 (Linux kernel version 3.10)
Compiler	Technical Computing Language (Fujitsu)	Intel Parallel Studio XE Cluster Edition Intel C/C++ and Fortran compiler Intel TBB Intel Distribution for Python		Technical Computing Language (Fujitsu) Intel Parallel Studio XE Cluster Edition Intel C/C++ and Fortran compiler Intel TBB Intel Distribution for Python
Library	Technical Computing Language (Fujitsu) - BLAS, LAPACK, ScaLAPACK, MPI, SSLII, C-SSLII, SSLII/MPI, Fast Basic Operations Library for Quadruple Precision	IntelParallel Studio XE Cluster Edition Intel MKL Intel MPI Intel MPI Intel IPP Intel DAAL		Technical Computing Language (Fujitsu) IntelParallel Studio XE Cluster Edition Intel MKL Intel MPI Intel MPI Intel IPP Intel DAAL
Tool	FUJITSU Software Development Tools	Intel Parallel Studio XE Cluster Edition Intel VTunr Amplifier XE Intel Advisor Intel Inspector Intel Trace Analyzer & Collector		FUJITSU Software Development Tools Intel Parallel Studio XE Cluster Edition Intel VTunr Amplifier XE Intel Advisor Intel Inspector Intel Trace Analyzer & Collector
Job management system	Technical Computing Suite			

Application	Gaussian, GROMACS	Gaussian(Only supported Red Hat Enterprise Linux 7(x 56nodes), ADF, AMBER, GAMESS, Q-Chem, GROMACS, NAMD, ROOT	Gaussian, ADF, AMBER, ANSYS(only gwacsl), GAMESS, MATLAB(only gwacsg) Q-Chem, GROMACS, NAMD, ROOT	GaussView, ANSYS(preppost) VMD, ROOT
-------------	----------------------	--	---	--------------------------------------

You can develop programs on the front end servers for the GWMPC(SPARC) and BWMPC(Intel), the Application Computing Server (Intel) although these systems have different architectures.