

Integrated Supercomputer System for Data Analyses and Scientific Simulations: Reedbush

Background and Current Status

The Supercomputing Research Division, Information Technology Center, the University of Tokyo (SCD/ITC/UTokyo, <http://www.cc.u-tokyo.ac.jp>) is operating the following three supercomputer systems:

- Yayoi: Hitachi SR16000, IBM Power7, 54.9 TFLOPS, October 2011~
- Oakleaf-FX: Fujitsu PRIMEHPC FX10, SPARC IXfx, 1.13 PFLOPS, April 2012~
- Oakbridge-FX: Fujitsu PRIMEHPC FX10, SPARC IXfx, 136.2 TFLOPS, April 2014~

Oakleaf-FX and Oakbridge-FX are based on same architecture, and they are commercial version of K computer. Oakbridge-FX is mainly for long-time executions (e.g. > 7 days). Moreover, Post T2K system based on manycore architecture will start its operation in December 2016. Post T2K system is expected to be more than 25 PFLOPS peak performance, and is introduced under collaboration with University of Tsukuba (Fig.1). University of Tsukuba and SCD/ITC/UTokyo established the Joint Center for Advanced High Performance Computing (JCAHPC, <http://jcahpc.jp/>) for the collaboration in designing, installation and operation of the Post T2K system.

The supercomputer systems of SCD/ITC/UTokyo are shared by > 2,000 users, where more than half of them are outside of the University of Tokyo. The systems are utilized for various kinds of research areas, such as engineering, earth/space science, material science, energy science and physics (Fig.2). Recently, number of users related to biology, such as biomechanics and biochemistry, has increased. Our systems are very busy, and combined average usage ratio of Oakleaf/Oakbridge-FX was 80+% in FY.2015.

The Oakleaf-FX system will completed its mission in the end of March 2018. New system (*Post FX10*) is expected to start its operation in Fall 2018. Our current systems including Post T2K are designed for computational science and engineering, and are mostly used for that purpose, as shown in Fig.2. In the Post FX10, we plan to develop new types of users, such as *Big Data*, *Deep Learning* ...

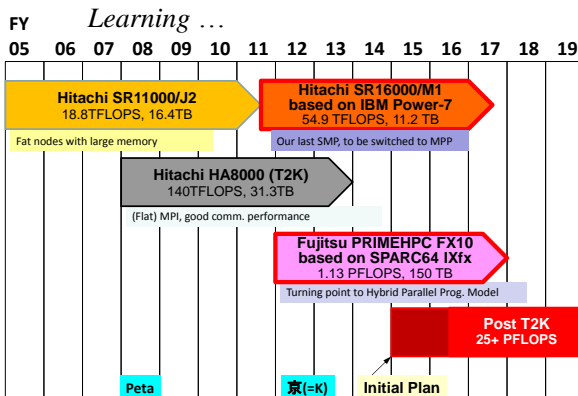


Fig.1 Schedule for Procurement of SCD/ITC/UTokyo

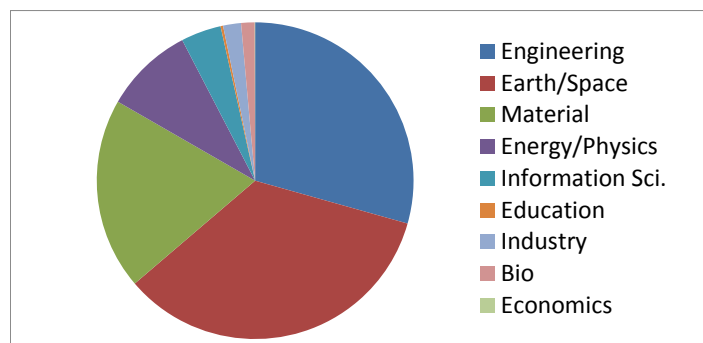


Fig.2 Areas for Usage in FY.2015 based on CPU hours (Oakleaf/Oakbridge/FX)

Reedbush

Considering these situations, we decided to introduce another small system, *Reedbush*. The Reedbush system has following two missions:

- Supplementary computational resource for reducing congestion of Oakleaf/Oakbridge-FX systems
- Pilot system towards the *Post FX10* system

We have two types of groups of compute nodes, (1) **Reedbush-U** (with CPU only), and (2) **Reedbush-H** (with accelerators). (1) and (2) could be operated separately. Summary of the specification is as follows:

(1) Reedbush-U

- Each Node: Intel Xeon E5-2695v4 (Broadwell-EP 2.1 GHz 18core) x2 socket, 1,209.6 GF, 256 GB Mem, 153.6 GB/sec
- Total Subsystem: 508.03 TF, 107 TB, Fat-tree with full bisection bandwidth (100 Gbps/node)

(2) Reedbush-H

- Host CPU: Same as Reedbush-U.
- Accelerators: NVIDIA Tesla P100 x2, For each: 4.8-5.3 TF, 16 GB, 720 GB/sec
- Host CPU~Accelerator: PCIe Gen3 x16 (15.7 GB/sec)
- Accelerator~Accelerator: NVLink (20GB/s) x2 brick
- Total Subsystem: 1.14-1.27 PF by Accelerators (excluded host CPU's), 1.29-1.42 PF (CPU+GPU), Fat-tree with full bisection bandwidth (56 Gbps x 2 /node)

(3) Storage

- Shared Storage: Luster Filesystem 5.04 PB, 145.2 GB/sec.
- System for Fast File Cache: DDN IME 209TB, 436.2 GB/sec.

(4) Power, Cooling

- Air cooling only, 377.56 kVA (without A/C)

(5) Schedule

- Phase-I operation for Reedbush-H: July 1, 2016
- Phase-II operation for full system: March 1, 2017

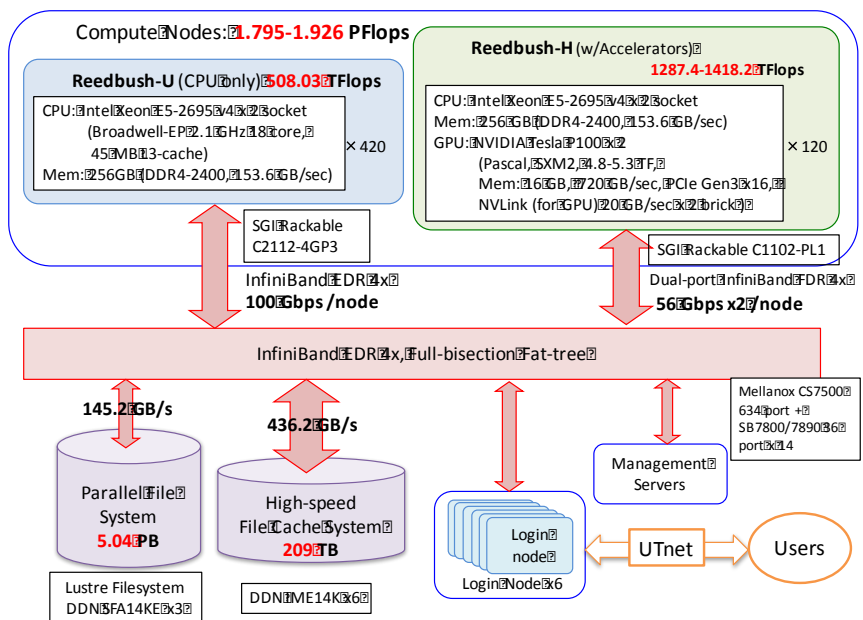


Fig.3 Overview of MPT2K System

If you have any questions, please feel free to contact Kengo Nakajima or Toshihiro Hanawa (nakajima/hanawa(at)cc.u-tokyo.ac.jp).