

Reedbush : データ解析・シミュレーション融合 スーパーコンピュータシステム

背景

東京大学情報基盤センター (SCD/ITC/UTokyo, <http://www.cc.u-tokyo.ac.jp>) では、現在以下の3つのスーパーコンピュータを運用しています。

- ・ Yayoi: Hitachi SR16000, IBM Power7, 54.9 TFLOPS, 2011 年 10 月~
- ・ Oakleaf-FX: Fujitsu PRIMEHPC FX10, SPARC IXfx, 1.13 PFLOPS, 2012 年 4 月~
- ・ Oakbridge-FX: Fujitsu PRIMEHPC FX10, SPARC IXfx, 136.2 TFLOPS, 2014 年 4 月~

Oakleaf-FX と Oakbridge-FX は、現在日本最速を誇るスーパーコンピュータ「京」の商用版である FX10、それぞれ 50 ラック、6 ラックから構成されています。Oakleaf-FX は一般ジョブ実行用に、Oakbridge-FX は 1 週間以上掛かるような長時間ジョブ実行用として運用しております。また、本センターは筑波大学と共同で JCAHPC (Joint Center for Advanced High Performance Computing, <http://jcahpc.jp/>) を立ち上げ、Post T2K システムを導入することに致しました。Post T2K は理論演算性能 20~25PFLOPS (京は約 10PF) に及び、メニーコアアーキテクチャを搭載したシステムを予定しており、本年 12 月より運用開始予定です。

本センターのスーパーコンピュータシステムは、Fig.2 に示す通り、工学、地球・宇宙科学、材料科学といった様々な分野の、2000 を超える利用者に活用されております。利用者の半数以上が東京大学外部からの利用者であり、また最近では生物学、生体力学、生化学などの分野の利用者が増加しています。このような計算機利用需要の増加から、Oakleaf/Oakbridge-FX は大変混雑しており、2015 年のシステム利用率は 80%以上に及びました。

Oakleaf-FX は 2018 年 3 月に運用を停止する予定であり、2018 年の秋頃より新しいシステム(*Post FX10*)の運用開始を目指しています。Fig.2 からわかる通り、現在のシステムは主に計算科学や工学向けに利用されており、また FX10 システムもそのような目的で使っていただくことを想定し設計されております。Post FX10 システムにおいては、さらに多くの方々に利用していただくために、ビッグデータ解析や人工知能といった、近年盛り上がりを見せる新たな分野の要求を満たすシステムの開発を目指しています。

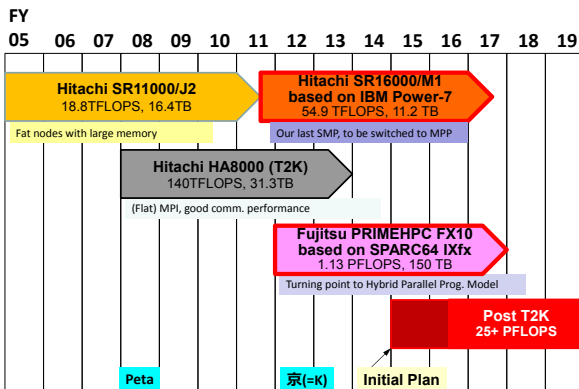


Fig.1 調達スケジュール

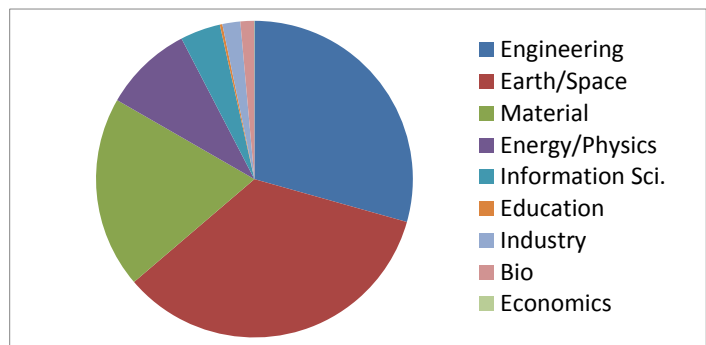


Fig.2 Oakleaf/Oakbridge-FX の 2015 年の分野ごと使用内訳(CPU 時間)

Reedbush

このような状況を鑑みて、我々は新たに **Reedbush** と呼ばれるシステムを導入することに決定いたしました。Reedbushには、

1. Oakleaf/Oakbridge-FX の混雑緩和
2. Post FX10 システムに向けてのテストベッドシステム

としての2つの役割がございます。また Reedbush は CPU のみのノードからなる **Reedbush-U** と、演算アクセラレータとして GPU を搭載したノードからなる **Reedbush-H** の2つのサブシステムから構成され、それぞれは独立のシステムとして運用可能です。以下に Reedbush の概要を示します。

(1) Reedbush-U

- ・ ノード単体: Intel Xeon E5-2695 v4 (Broadwell-EP 2.1 GHz 18core) x2 socket, 1,209.6 GF, 256 GB Mem, 153.6 GB/sec
- ・ サブシステム全体: 508.03 TF, 107 TB, full bisection バンド幅 Fat-tree (100 Gbps/ノード)

(2) Reedbush-H

- ・ CPU: Reedbush-U と同様
- ・ GPU: NVIDIA Tesla P100 x2, 単体性能 4.8~5.3 TF, 16 GB, 720 GB/sec
- ・ CPU-GPU 間通信: PCIe Gen3 x16 (15.7 GB/s)
- ・ GPU-GPU 間通信: NVLink (20GB/s) x2 brick
- ・ サブシステム全体: 1.14~1.27 PF (GPU のみ), 1.29~1.42 PF (CPU+GPU), full bisection バンド幅 Fat-tree (56 Gbps x2/ノード)

(3) ストレージ

- ・ 共有ストレージ: Lustre Filesystem 5.04 PB, 145.2 GB/sec.
- ・ 高速ファイルキャッシュシステム: DDN IME 209TB, 436.2 GB/sec.

(4) 消費電力、冷却システム

- ・ 377.56 kVA (冷却電力除く)
- ・ 空冷

(5) スケジュール

- ・ **Reedbush-U 運用開始: 2016年7月**
- ・ **Reedbush-H 運用開始: 2017年3月**

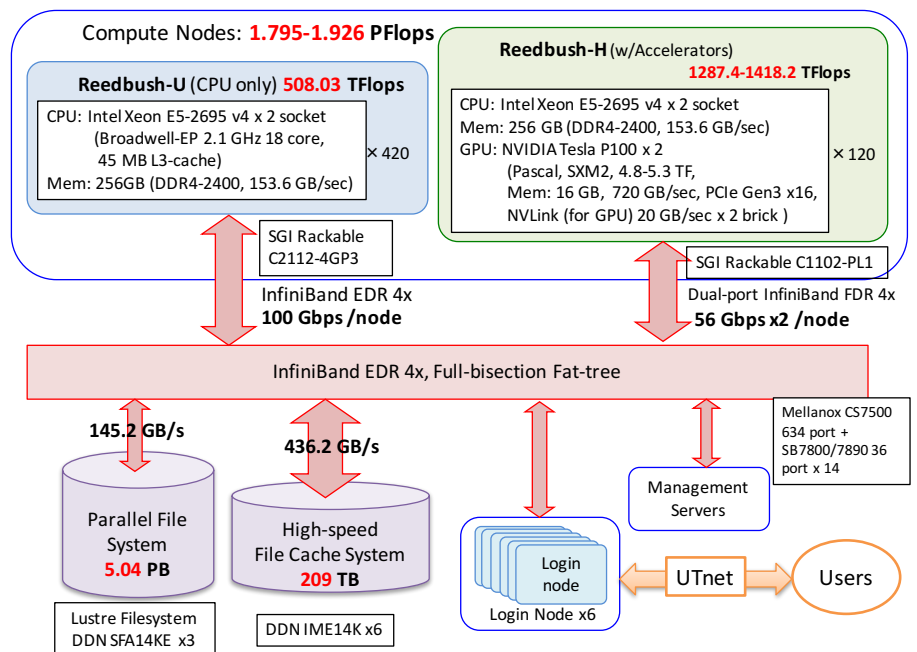


Fig.3 Overview of MPT2K System

質問ございましたら中島研吾または埴敏博 (nakajima/hanawa(at)cc.u-tokyo.ac.jp)までお問い合わせください。