

第3世代

JAXAスパコンが目指すもの -“TOKI”導入の目的と初期性能評価-



国立研究開発法人 宇宙航空研究開発機構
スーパーコンピュータ活用課長/航空技術部門研究領域主幹

藤田 直行

目次

- “TOKI”導入の目的
- JSS3概要
- 初期性能評価
- おわりに

■三つの柱

■背景

- スパコンと共に発展してきた数値流体力学(CFD)は、振興から実用の時代へ
- スパコン利用分野がCFDのみならず、地球観測データ処理やAI基盤というデータ処理へと広範囲化

- I. 航空宇宙分野の国際競争力を強化する数値シミュレーション実施基盤
- II. 大規模データ解析基盤としてのデータセンター機能
- III. 新たなニーズを受け止める研究開発基盤

目次

- “TOKI”導入の目的
- JSS3概要
- 初期性能評価
- おわりに

JAXAスーパーコンピュータシステム JSS3

稼働開始
2020年12月1日

【コンピュータ基盤】 TOKI: TOKyo and ibaraKI



調布航空宇宙センター



TOKI-SORA: HPCシステム

PRIMEHPC FX1000
 ノード数: 5,760 ノード (15ラック)
 総理論演算性能: 19.4 PFLOPS
 総主記憶容量: 180 TiB (32 GiB/ノード)



TOKI-RURI: 汎用システム

総理論演算性能: 1.24 PFLOPS
 総主記憶容量: 104 TiB

- ST**: PRIMERGY RX2540 M5 x 375 ノード (192 GiB/ノード, Quadro x 1 基)
- GP**: PRIMERGY CX2570 M5 x 32 ノード (384 GiB/ノード, Tesla V100 x 4 基)
- XM**: PRIMERGY RX2540 M5 x 2 ノード (DCPMM 6.0 TiB/ノード, Quadro x 1 基)
- LM**: PRIMERGY RX2540 M5 x 7 ノード (DCPMM 1.5 TiB/ノード, Quadro x 1 基)



TOKI-FS: ファイルシステム

ファイルシステム: FEFS
 オールフラッシュ NVMe ストレージ: 10PB
 ハードディスクドライブ ストレージ: 40PB

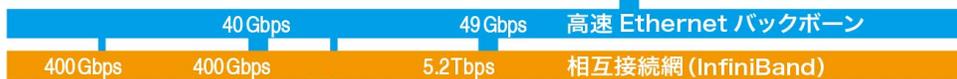
TOKI-LI: ログインシステム

PRIMERGY RX2540 M5 x 最大14 ノード (384 GiB/ノード, Quadro x 1 基)

運用管理システム



筑波宇宙センター



TOKI-TFS: 筑波ファイルシステム

ファイルシステム: FEFS, 総実効容量: 0.4PB

TOKI-TLI: 筑波ログインシステム

PRIMERGY RX2540 M5 x 2 ノード (384 GiB/ノード, Quadro x 1 基)

筑波運用管理制御システム

TOKI-TRURI: 筑波汎用システム

総理論演算性能: 145 TFLOPS
 総主記憶容量: 10.8 TiB

- TST**: PRIMERGY RX2540 M5 x 46 ノード (192 GiB/ノード, Quadro x 1 基)
- TGP**: PRIMERGY CX2570 M5 x 2 ノード (384 GiB/ノード, Tesla V100 x 4 基)
- TLM**: PRIMERGY RX2540 M5 x 1 ノード (DCPMM 1.5 TiB/ノード, Quadro x 1 基)

【アーカイバ基盤】 J-SPACE

ディスクキャッシュ容量: 3PB
 テープ容量: 70PB

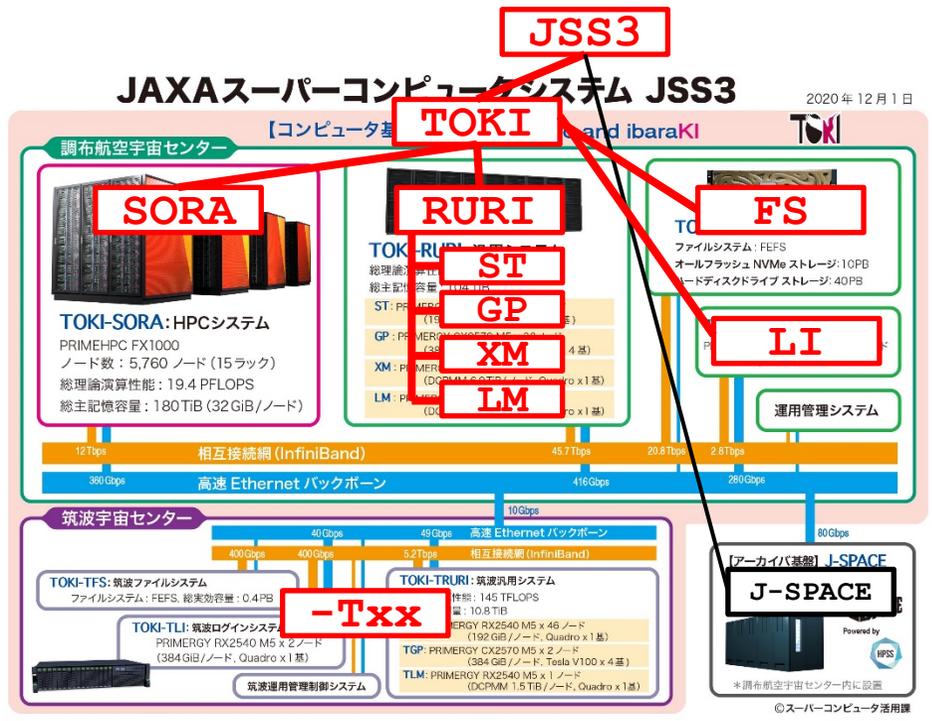


* 調布航空宇宙センター内に設置

©スーパーコンピュータ活用課

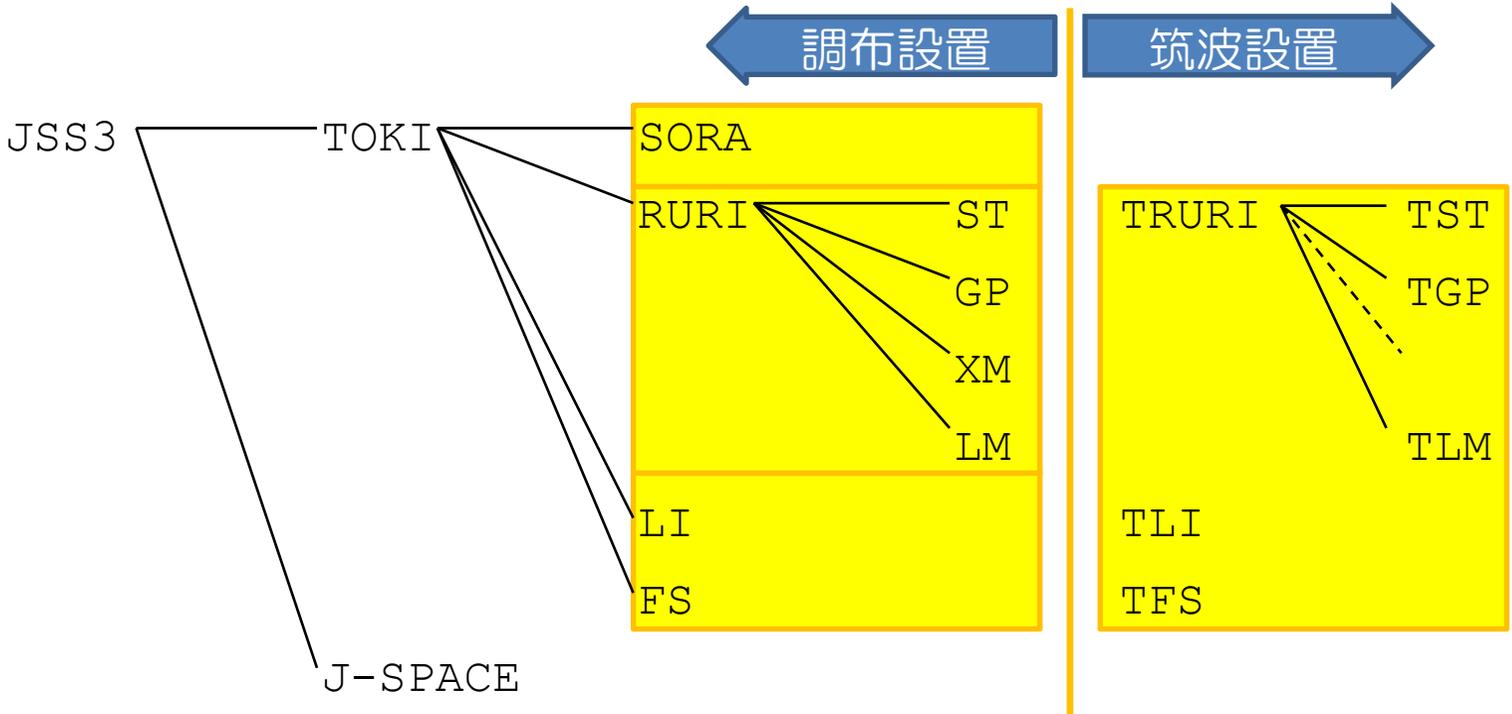
■ 名前に込めた思い

- JSS3: **J**axa **S**upercomputer **S**ystem generation **3**
- TOKI: **T**Okyo and **i**bara**KI**、時 (SORAと合わせて時空)、解き、朱鷺
- SORA: **S**upercomputer for earth **O**bservation, **R**ockets, and **A**eronautics
- RURI: all-**R**o**U**nd **R**ole **I**nfrasturcture
 - ST : **S**Tandard
 - GP : **G**Pgpu
 - XM : e**X**tra large **M**emory
 - LM : **L**arge **M**emory
- FS : **F**ile **S**ystem
- LI : **L**og**I**n
- -Txx : -**T**sukuba xx



TOKI-SORA

Supercomputer for
earth Observation, Rockets, and Aeronautics



TOKI-SORA

Supercomputer for
earth Observation, Rockets, and Aeronautics

PRIMEHPC FX1000



1ノードあたりの構成

CPU	プロセッサ	A64FX(Armv8.2-A + SVE)
	プロセッサ数	1プロセッサ
	理論演算性能	3.3792TFLOPS
主記憶	種別	HBM2
	容量	32GiB
	メモリバンド幅	1,024GB/s
インタコネク	種別	Tofuインターコネク
	理論性能	最大40.8GB/s × 双方向

ラックあたりの構成

計算ノード数	384ノード
--------	--------

HPCシステム全体構成

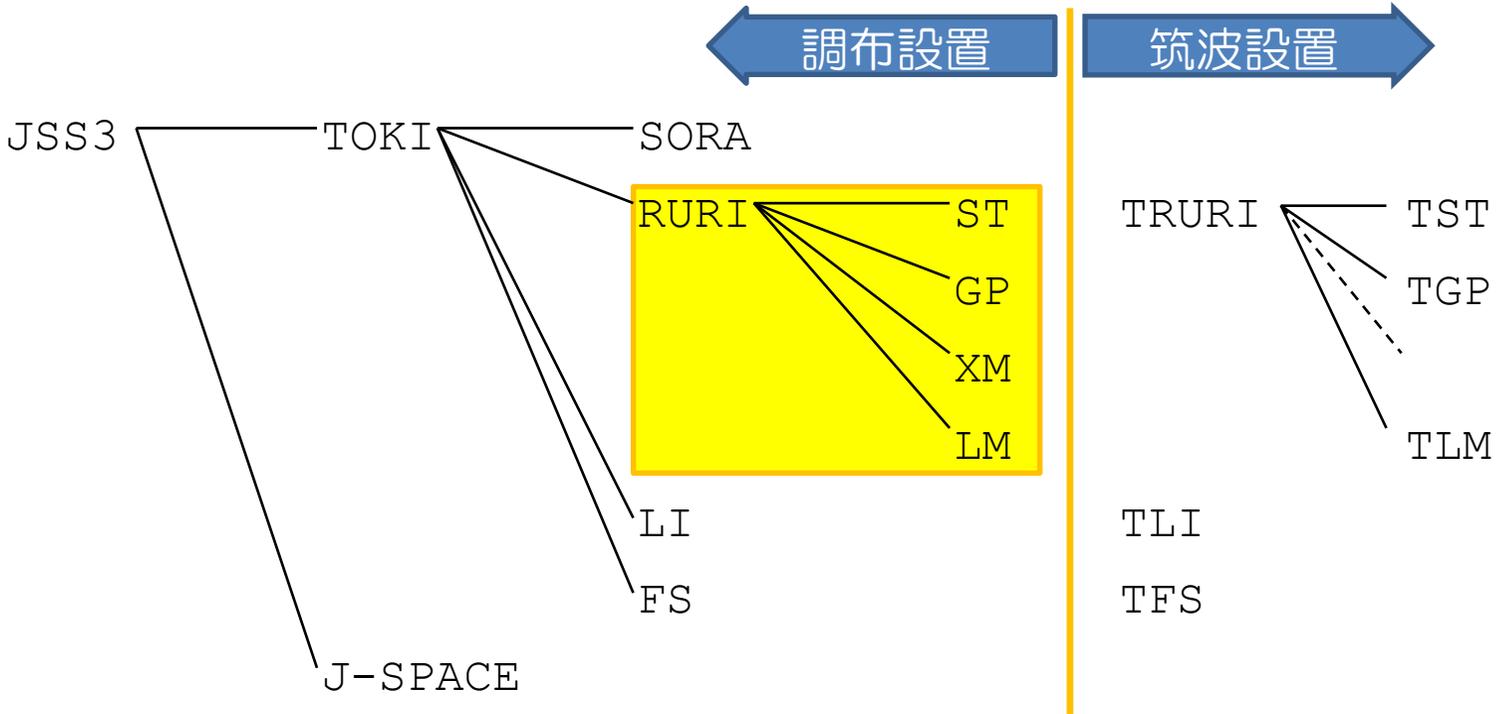
- 総理論演算性能 19.4 PFLOPS

$3.3792\text{TFLOPS}/\text{ノード} \times 384\text{ノード}/\text{ラック} \times 15\text{ラック}$
 $\doteq 19.464\text{PFLOPs}$

- 総主記憶容量 180TiB

$32\text{GiB}/\text{ノード} \times 5,760\text{ノード}$
 $= 180\text{TiB}$

TOKI-RURI all-RoUnd Role Infrastructure



TOKI-ST STandard

PRIMERGY RX2540 M5



x375

1ノードあたりの構成

CPU	プロセッサ	Intel Xeon Gold 6240 (2.6GHz/18コア)
	プロセッサ数	2プロセッサ
	理論演算性能	2.9952TFLOPS (2.6GHz × 18コア × 32命令 × 2CPU)
GPU	種別	NVIDIA Quadro P4000(1,792 CUDAコア, 8GiBメモリ)
	GPU数	1基
主記憶	種別	16GiB DDR4 2933MHz RDIMM (ECC)
	容量	192GiB (16GiB × 12枚)
	メモリバンド幅	281.5GB/s (2933MHz × 8byte × 12)
内蔵ディスク	種別	M.2 Flash モジュール
	実効容量	240GB
インターコネク	種別	InfiniBand HDR100 × 1ポート
	理論性能	100Gbps (双方向)
ネットワークインターフェース		1000Base-T×2ポート
筐体内監視		iRMC (CPU,メモリ,HDD,カード,ファン,電源,温度,電圧等)
高さ		2U
電源		800W×2 (80PLUS PLATINUM) ※冗長化、活性交換可能

TOKI-GP(1/2) GPgpu

PRIMERGY CX2570 M5 (水冷)



x32

PRIMERGY CX400 M4 シャーシ(水冷)



x16

PRIMERGY CX2570 M5 1台あたりの構成

CPU	プロセッサ	Intel Xeon Gold 6240 (2.6GHz/18コア)
	プロセッサ数	2プロセッサ
	理論演算性能	2.9952TFLOPS (2.6GHz x 18コア x 32命令 x 2CPU)
GPU	種別	SXM2版 NVIDIA Tesla V100 (5,120 CUDAコア, 32GiB HBM2)
	GPU数	4基
	理論演算性能	倍精度 31.2TFLOPS (7.8TFLOPS x4基) 単精度 62.8TFLOPS (15.7TFLOPS x4基) Tensor 500TFLOPS (125TFLOPS x4基, 半精度・単精度混合)
主記憶	種別	32GiB DDR4 2933MHz RDIMM (ECC)
	容量	384GiB (32GiB x 12)
	メモリバンド幅	281.5 GB/s (2933MHz x 8byte x 12チャネル)
内蔵ディスク	種別	M.2 Flash モジュール
	実効容量	256GB
インターコネクト	種別	InfiniBand EDR x 2ポート
	理論性能	200Gbps (双方向)
ネットワークインターフェース		1000BASE-T x 1ポート
筐体内監視		iRMC (CPU,メモリ,HDD,カード,ファン,電源,温度,電圧等)

PRIMERGY CX400 M4 1シャーシあたりの構成

搭載サーバード	PRIMERGY CX2570 M5 x 2ノード
シャーシの高さ	2U
電源	電源ユニット(2400W) x 2

TOKI-XM/LM(1/2) eXtra large Memory / Large Memory

PRIMERGY RX2540 M5



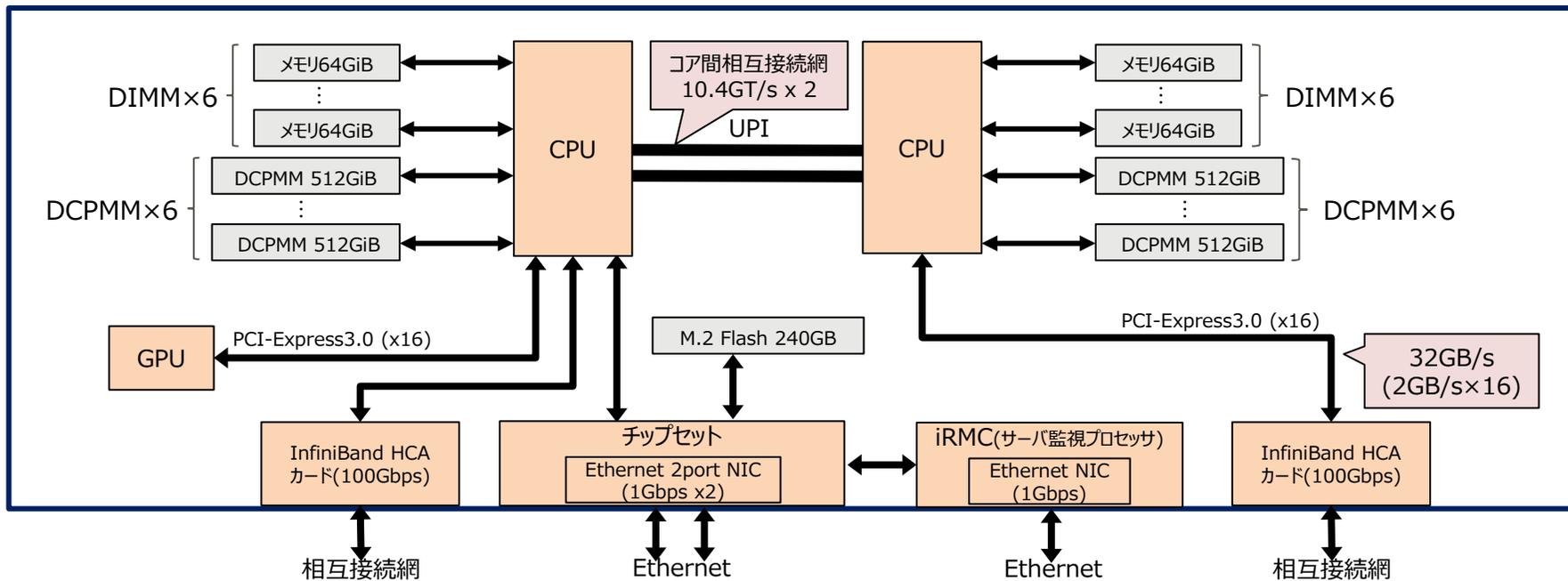
PRIMERGY RX2540 M5



1ノードあたりの構成

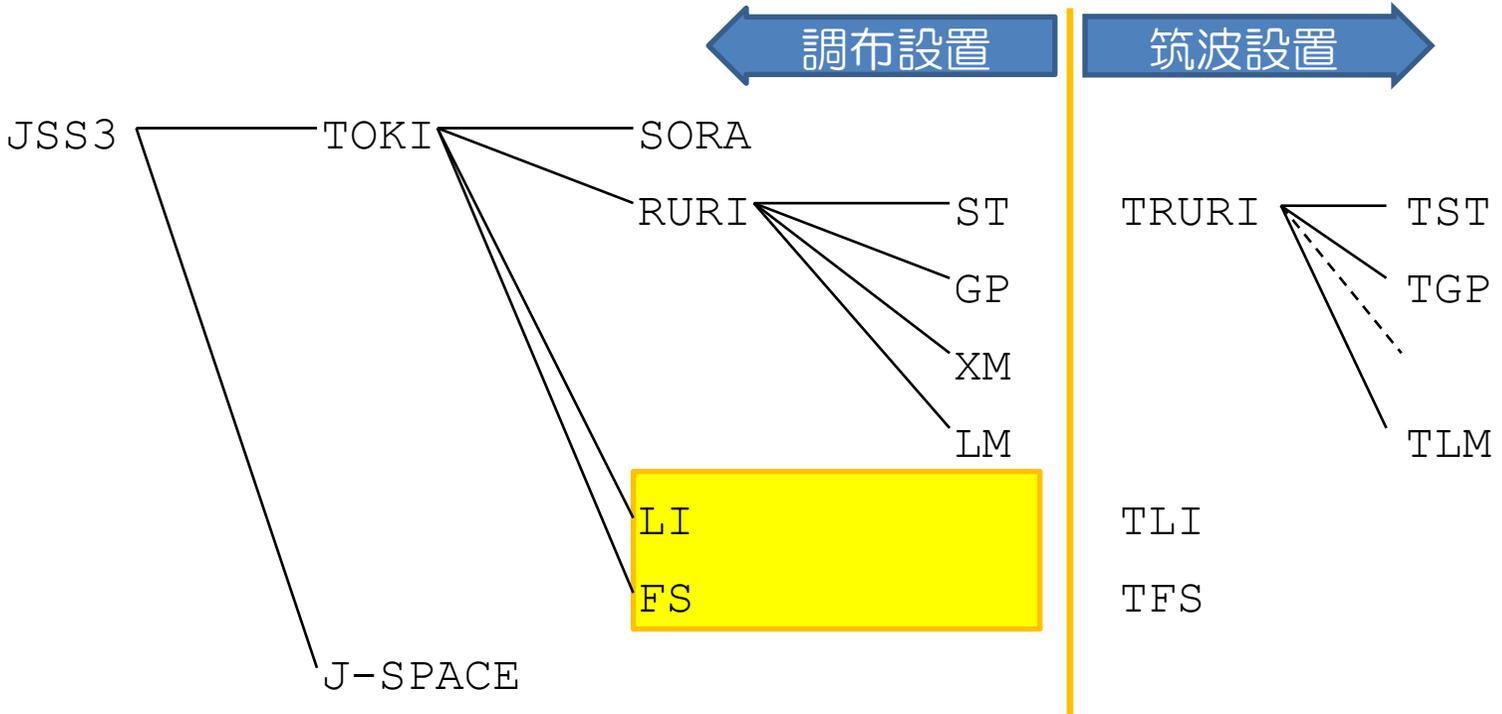
CPU	プロセッサ	Intel Xeon Gold 6240 (2.6GHz/18コア)
	プロセッサ数	2プロセッサ
	理論演算性能	2.9952TFLOPS (2.6GHz x 18コア x 32命令 x 2CPU)
GPU	種別	NVIDIA Quadro P4000 (1,792 CUDAコア, 8 GiBメモリ)
	GPU数	1基
主記憶	種別	16GiB 64GiB DDR4 2933MHz RDIMM (ECC)
	容量	192GiB 768GiB (16GiB x 12枚)
	メモリバンド幅	255.9 GB/s (2666MHz x 8byte x 12)
不揮発性メモリ	種別	Intel Optane DC パーシステント・メモリ
	容量	1.5TiB 6TiB (128GiB x 12枚)
内蔵ディスク	種別	M.2 Flash モジュール
	容量	240GB
インターコネク	種別	InfiniBand HDR(100Gbps) x 2ポート
	理論性能	200Gbps (双方向)
ネットワークインタフェース		1000Base-T x 2ポート

TOKI-XM/LM(2/2) eXtra large Memory / Large Memory



上図はXMの場合

TOKI-LI、 TOKI-FS



TOKI-LI LogIn

PRIMERGY RX2540 M5



x14

ログインシステム全体構成

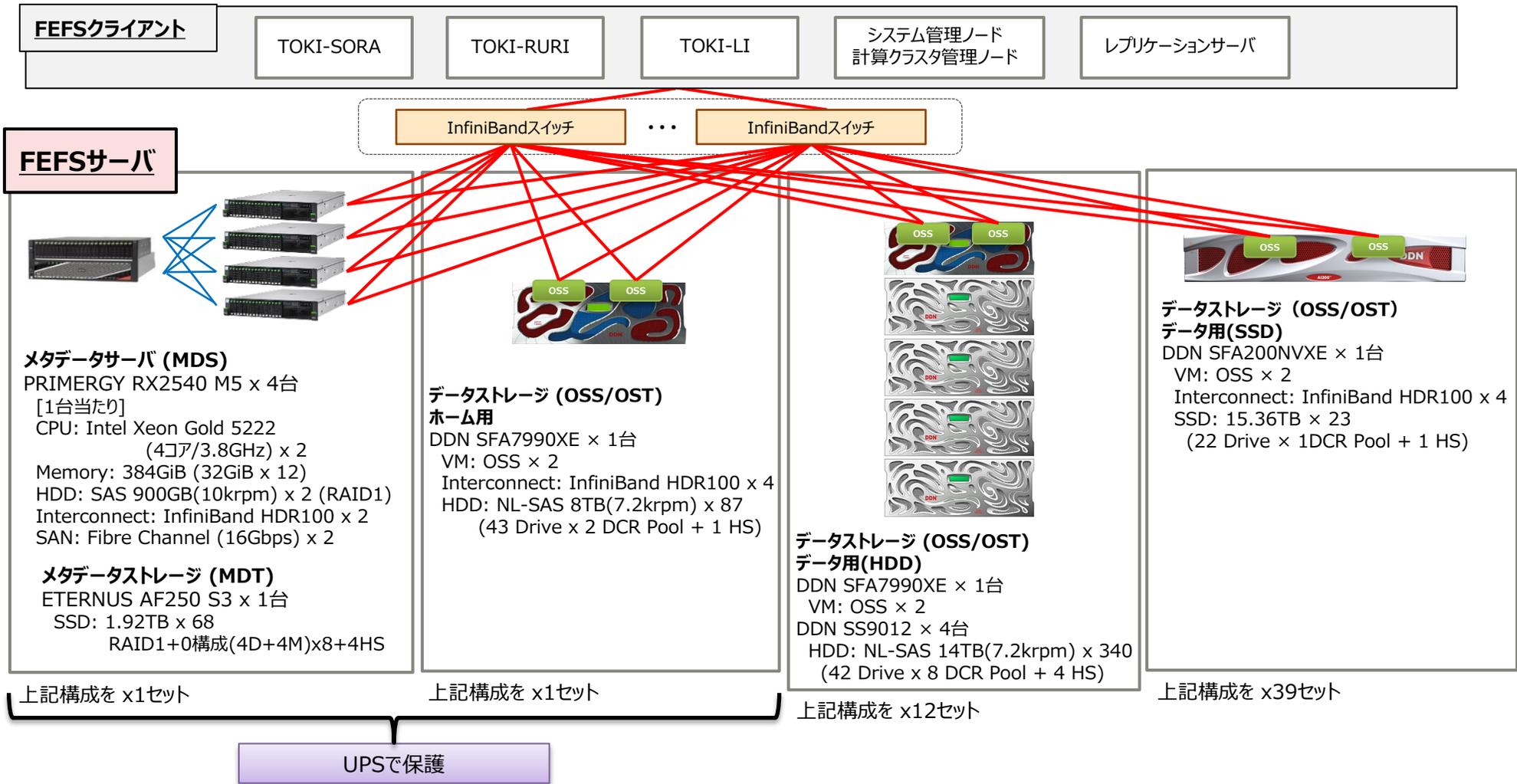
- 総理論演算性能 **47.9 TFLOPS**
 2.9952TFLOPS/ノード x 16ノード
 ≒ 47.9 TFLOPS

- 総主記憶容量 **6.0 TiB**
 384GiB/ノード x 16ノード
 = 6.0 TiB

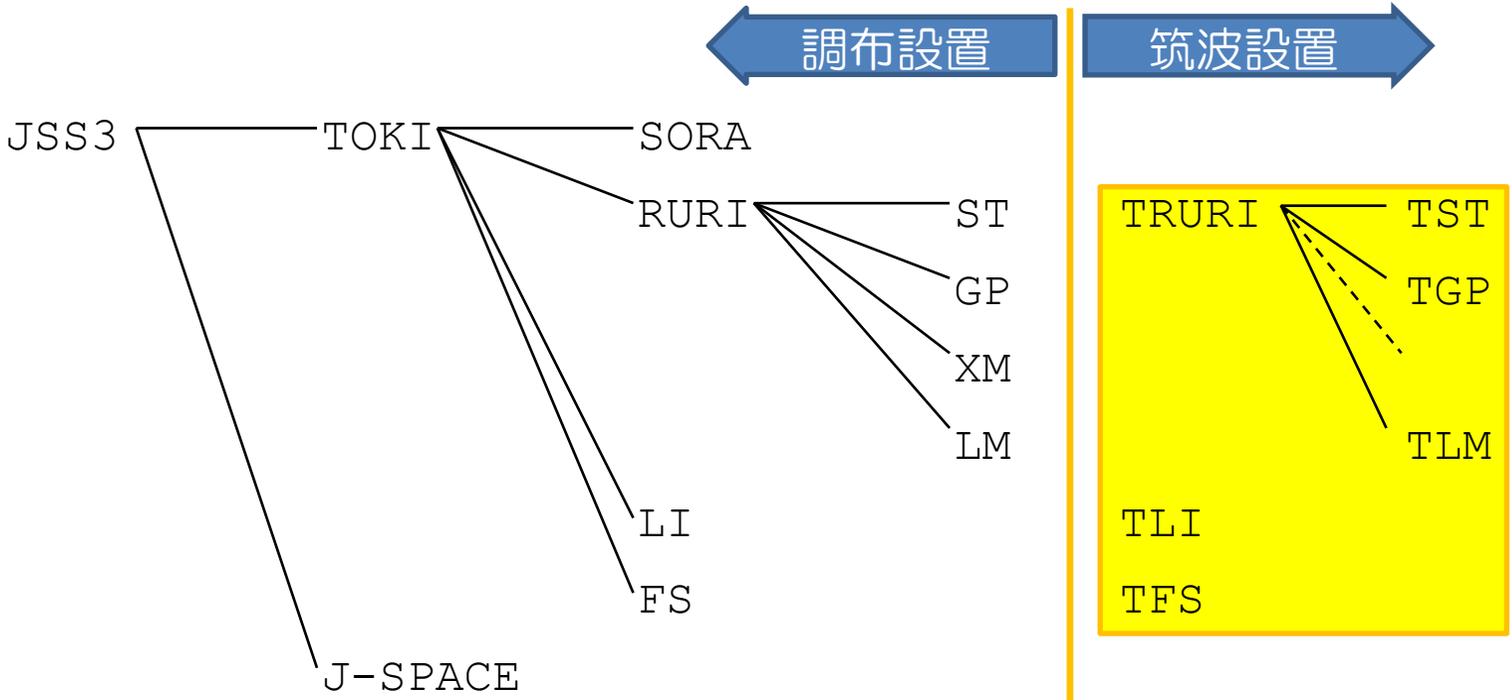
1台あたりの構成

CPU	プロセッサ	Intel Xeon Gold 6240 (2.6GHz/18コア)
	プロセッサ数	2プロセッサ
	理論演算性能	2.9952TFLOPS (2.6GHz x 18コア x 32命令 x 2CPU)
GPU	種別	NVIDIA Quadro P4000 (1792 CUDAコア, 8GiB, GDDR5)
	GPU数	1基
主記憶	種別	32GiB DDR4 2933MHz x 12 (ECC)
	容量	384GiB (32GiB x 12枚)
	メモリバンド幅	281.5 GB/s (2933MHz x 8byte x 12 / 1000)
内蔵ディスク	種別	600GB SAS HDD (15krpm) x 2 (RAID1)
	実効容量	600GB
インターコネク	種別	InfiniBand HDR(100Gbps) x 2枚 (2ポート)
	理論性能	200Gbps (双方向)
ネットワークインターフェース		10GBASE-T x 2ポート、1000BASE-T x 2ポート
電源		電源ユニット(800W) x 2

TOKI-FS File System

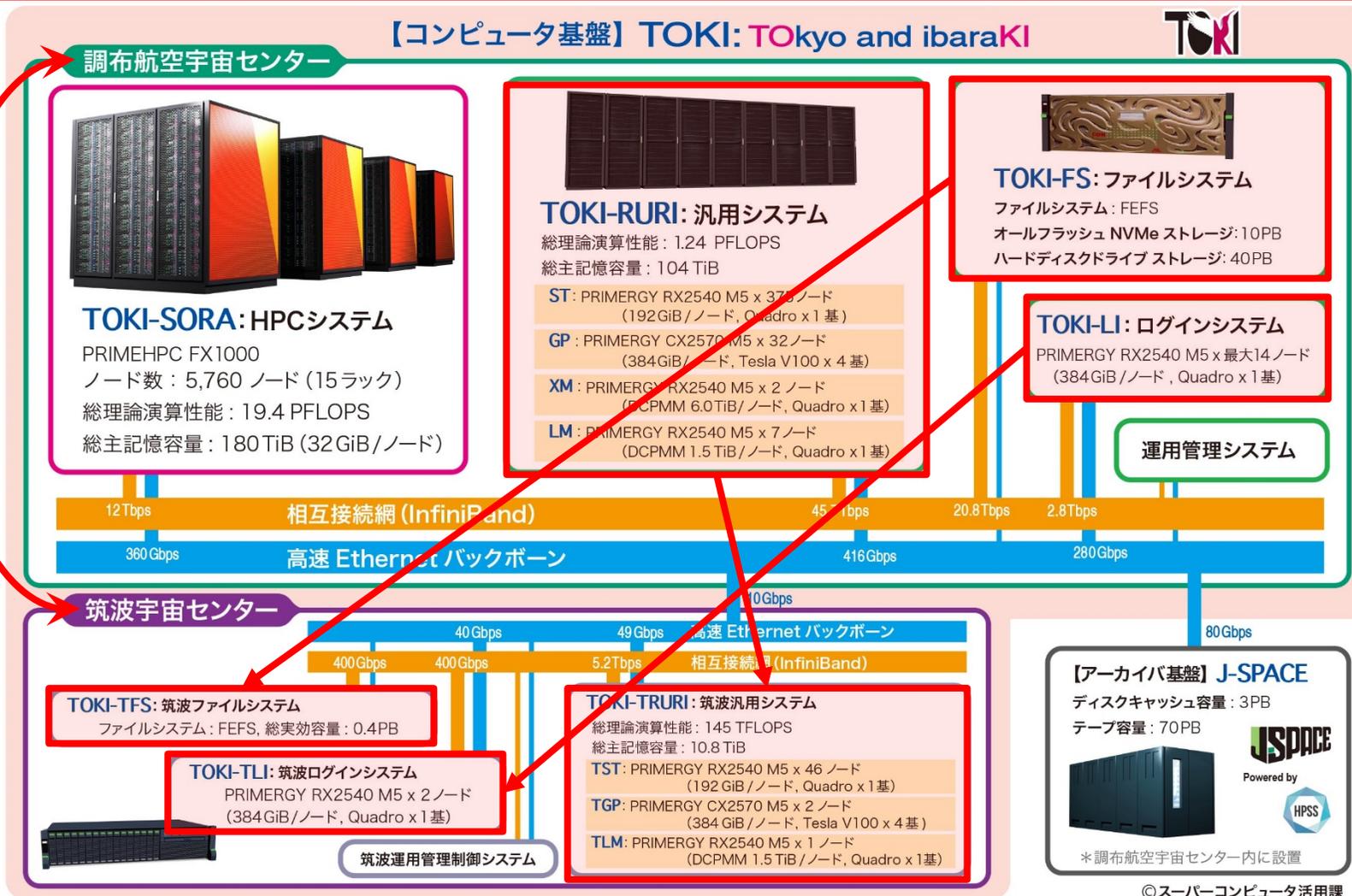


TOKI-Txx - Tsukuba xx



TOKI-Txx - Tsukuba xx

止まらないスパコン

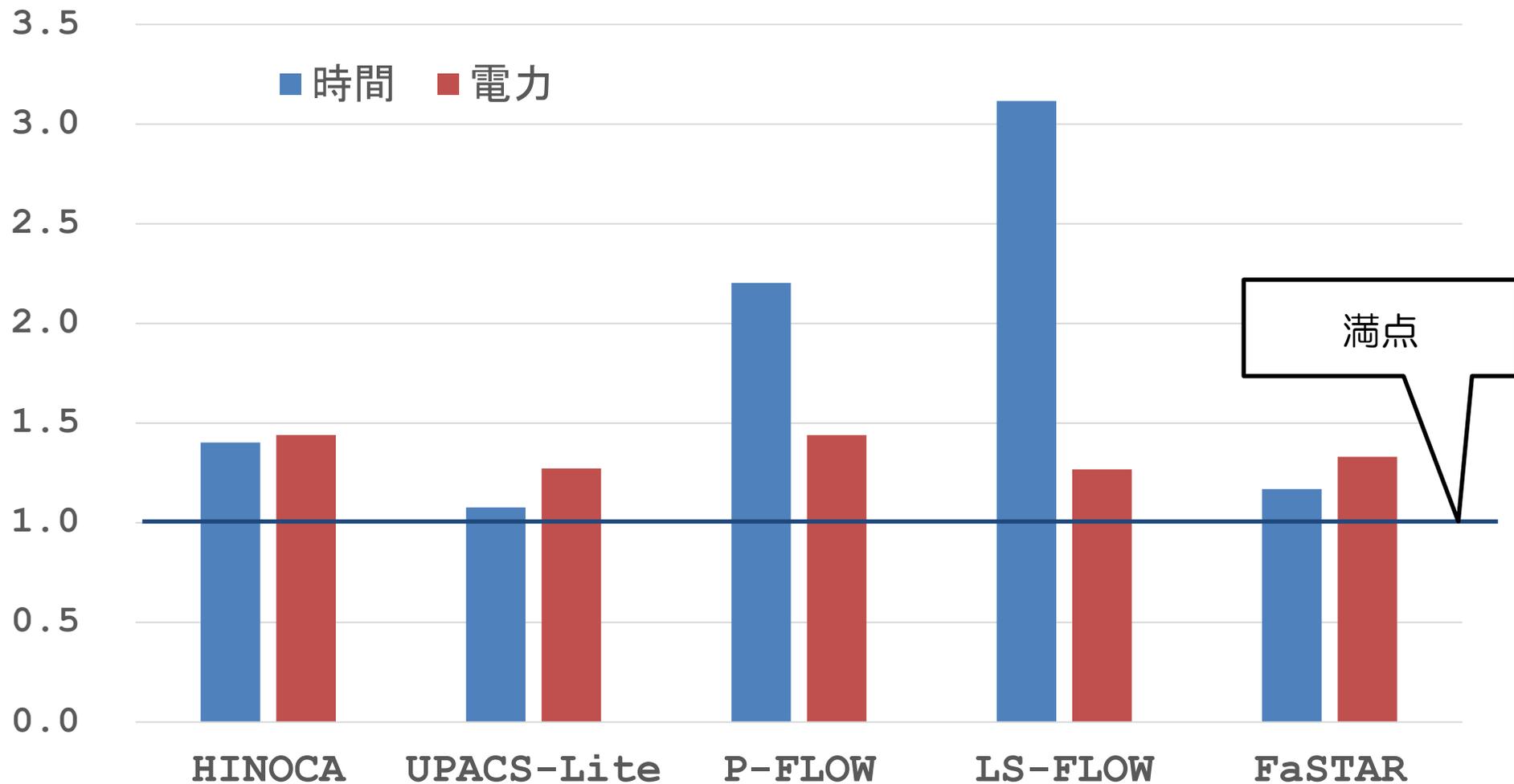


©スーパーコンピュータ活用課

目次

- “TOKI”導入の目的
- JSS3概要
- 初期性能評価
- おわりに

■ 導入時JAXAアプリベンチマーク



■ TOP500

■ 測定条件

- 実行ノード数: 5,760ノード (TOKI-SORA全系)
 - 総コア数: 276,480 (5,760ノード×(4プロセス×12スレッド/ノード))
 - 理論演算性能: 19.4642 PFLOPS
 - ノード形状: 20x18x16:torus
- 実行時パラメータ
 - N(問題規模)=4,208,640、P=96、Q=240

■ 測定結果

- 性能: 16.592 PFLOPS
- 実行効率: 85.24%
- 順位: 19位 (国内 3位)

■ HPCG

■ 測定条件

- 実行ノード数: 5,760ノード (TOKI-SORA全系)
 - 総コア数: 276,480 (5,760ノード×(4プロセス×12スレッド/ノード))
 - 理論演算性能: 19.4642 PFLOPS
 - ノード形状: 20x18x16:torus
- 実行時パラメータ
 - Local Domain Dimensions: nx=160, ny=160, nz=160
 - Processor Dimensions: npx=40, npy=36, npz=16

■ 測定結果

- 性能: 0.61422 PFLOPS
- 実行効率: 3.15%
- 順位: 8位 (国内2位)

■ Graph500

■ 測定条件

- 実行ノード数: 5,760ノード (TOKI-SORA全系)
 - 総コア数: 276,480 (5,760ノード×(4プロセス×12スレッド/ノード))
 - 理論演算性能: 19.4642 PFLOPS
 - ノード形状: 60x96
- 実行時パラメータ
 - SCALE(測定規模): 36

■ 測定結果

- 性能: 10,813 GTEPS
- 順位: 3位 (国内2位)

■ 10500

今後計測予定

目次

- “TOKI”導入の目的
- JSS3概要
- 初期性能評価
- おわりに

■ おわりに(1/2)

- “TOKI”導入の目的
- JSS3概要
- 初期性能評価

■ おわりに(2/2)

- “TOKI”導入の目的
- JSS3概要
- 初期性能評価
- これからのJSS3～新たな利用～
 - TOKI-FSのランダムIO性能
 - IOPSを必要とする処理の本格実行
 - TOKI-SORA/RURIの仮想化技術
 - 外部コンピュータ環境との連携加速
 - TOKI-GPのGPGPU利用
 - JAXA初の大型GPGPU基盤
 - 演算性能の向上→大規模チャレンジ
 - 大規模3D可視化
 - 液滴蒸発のDNS
 - ロケットエンジン燃焼器のLES

ありがとうございました。

JAXAスパコン公開Webサイト

<https://www.jss.jaxa.jp/>