

JAXA 新スパコン JSS2 の導入目的と構成概要

藤田 直行
宇宙航空研究開発機構

A purpose of JAXA next generation supercomputer installation and its composition outlines

by
Naoyuki FUJITA (JAXA)

ABSTRACT

JSS(Jaxa Supercomputer System) is the first supercomputer system for Japan Aerospace Exploration Agency(JAXA). On JSS, fruitful numerical simulation have been calculate, for example, launch pad noise analysis, liquid fuel atomization, design of rocket engine, airplane landing gear sound analysis, shock wave on wing, and so on. Five years or more have passed since the system operation start. From that time by the present, aggressive use of numerical simulation on aerospace field is continuing. As a result, supercomputer resource became insufficient. Main requirements are (1) Increase computing power, (2) Correspondence to the analysis of an unsteady phenomenon, (3) Necessity for file system and archive system, (4) Large scale pre/post-processing realization, (5) Realization of the backup function of experimental data and/or observational data. So JAXA decided to replace supercomputer system. This shows a purpose of the installation and JSS generation 2 composition outlines.

1. はじめに

宇宙航空研究開発機構(以下、JAXA)創設後、最初のスーパーコンピュータシステムとして JSS: Jaxa Supercomputer System を導入して5年以上が経過し、数値シミュレーションの航空宇宙分野での活用促進と相まって、計算資源の慢性的な不足に悩まされている。最近では、スパコン利用分野の拡大も起こり、不足傾向に拍車がかかる事態となっている。本稿では、JAXA の新スパコン JSS2 導入の目的や、その構成概要を述べる。

2. 導入の背景・目的

JAXA は、航空宇宙技術研究所、宇宙開発事業団、宇宙科学研究所の宇宙3機関が、2003(平成15)年10月に統合し新たに発足した宇宙開発、航空技術、関連基盤技術に関する我が国唯一の総合研究開発機関である。JAXA では、ロケット・宇宙輸送システム、国際宇宙ステーション・有人宇宙活動、人工衛星・探査機、航空機・航空推進システムに係る開発プロジェクトに取り組むと共に、宇宙科学・地球科学研究や宇宙航空基盤技術強化を進めている。JAXA は、これまで、我が国初のベクトル計算機「FACOM230 -75AP」や世界初・世界最速の分散主記憶型並列ベクトル計算機「数値風洞(NWT)」等を導入し、流体力学や構造力学の数値シミュレーションに取り組んできた。2009(平成21)年4月から全システムが稼働した現有する JAXA のスパコンシステム”JSS”は、JAXA 発足時、調布事業所、角田事業所、及び相模原事業所の3カ所にあったスパコンシステムを統合し整備したものである。スパコンシステムを統合運用することにより、ユーザから見たサービス内容の向上や効率的運用が実現された。これは、JAXA 内部のユーザに留まらず、共同研究や設備貸付のユーザへの支援体制の充実にも繋がるものであった。JAXA 開発のジョブスケジューラ”JaRMAN”の採用をはじめとした運用・利用技術の高度化により、平均稼働率は95%を超えており、スパコンニーズは、今後ますます高まっていくと考えられる。計算機能力の向上と共に、非定常解析や多分野統合解析が行われるようになり、演算性能のみならず、ファイルシステム性能やアーカイバの性能が求められ、また、格子生成や可視化に代表されるプレ・ポスト処理において、大規模データを扱うた

めの能力が再び求められるようになってきている。図1に、1980年代から現在に至る代表的な数値シミュレーション事例と計算機ピーク性能の関係を示す。

JAXA のスパコンシステムには、(1)宇宙航空分野での先導的計算科学研究の推進、(2)数値シミュレーション技術や計算能力による JAXA プロジェクトへの貢献、(3)宇宙航空分野における学術的基礎研究の実施、及び(4)大学共同利用設備、という4つの大きな役割がある。分野別に見ると年度毎に多少の変動はあるものの、航空分野が5割、宇宙分野が3割、基礎研究が2割というのが、ここ5年間の JAXA スパコンシステムの利用状況である。また、近年、乱流・音響・過渡現象といった非定常現象の解析が盛んに行われるようになり、また、数値シミュレーションの対象としては、ロケットエンジンのロバスト設計・コンタミネーション低減や破壊分散予測、機体騒音、エンジン騒音、燃焼室解析、燃焼効率評価、混合比特性評価、燃焼振動評価、噴霧燃焼流解析、飛行安全、音響環境予測と低減技術、超音速混合燃焼、実機空力性能解析、超音速機離着陸時の騒音低減解析、システム動特性評価、宇宙プラズマ等が、今後実施されていくと予想される。図2、図3に、現行スパコン JSS での代表的な数値シミュレーション事例を示す。

更に、今までは、スパコン利用の大半が数値シミュレーションでの利用であったが、今後は、衛星観測データの処理や、実験データ・観測データのバックアップ機能等、今までにない機能がスパコンシステムに求められることが予想される。

以上をまとめると、(1)計算能力の慢性的不足、(2)乱流、音響、過渡現象といった非定常現象の解析の増加、(3)ファイルシステム性能、アーカイバ性能への要求の高まり、(4)大規模なプレ・ポスト処理(格子生成・可視化等)への要求の高まり、(5)数値シミュレーションに加えて、衛星観測データ処理や、実験・観測等のデータのバックアップ機能の要求、が顕在化している。これらを踏まえた上で、宇宙航空分野での先導的計算科学研究を推進し日本を牽引すると共に、スパコンによる JAXA 事業促進を目指して、4年程度の期間を目途に、数値シミュレーションを発展させるスパコンインフラを整備することとなった。

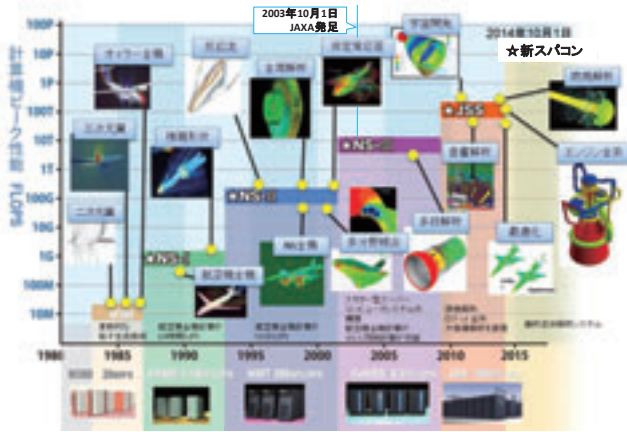
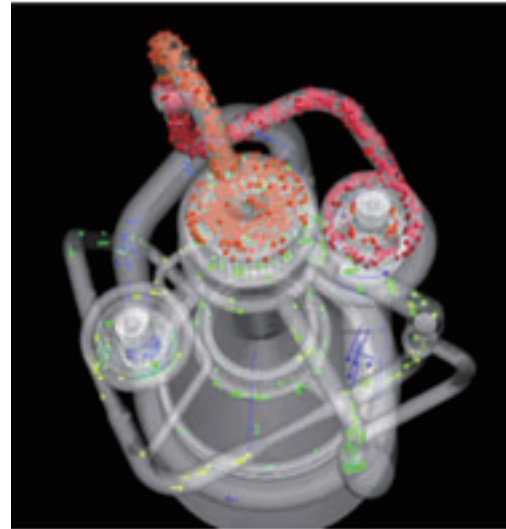
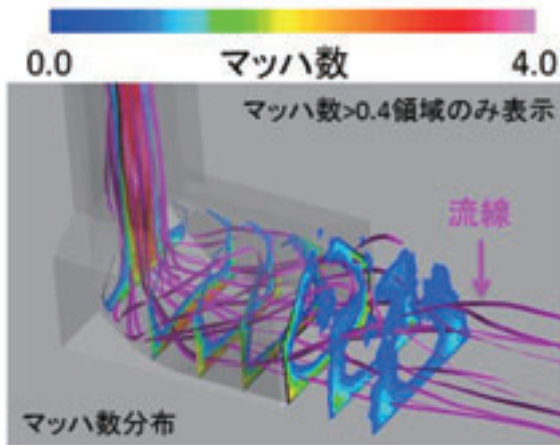


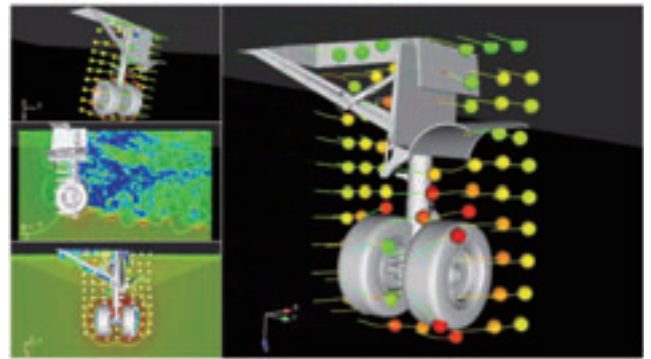
図1 代表的な数値シミュレーション事例と計算機性能



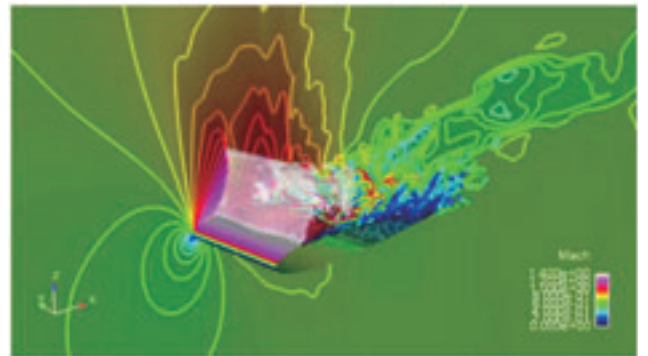
(c)ロケットエンジン設計



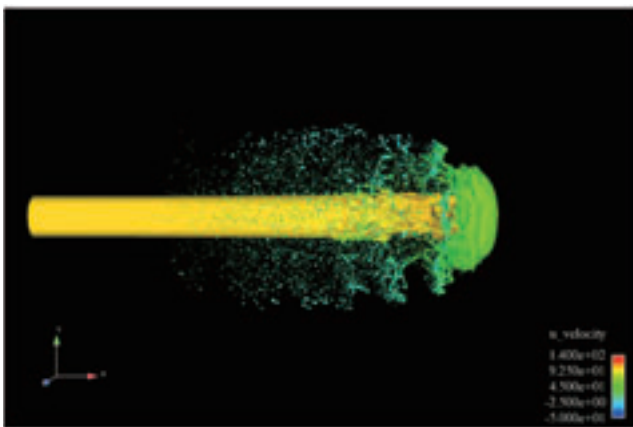
(a)イプシロンロケット射場設計



(d)航空機ランディングギア騒音



(e)翼上面衝撃波の可視化



(b)液体燃料微粒化過程

図2 現行スパコン JSS での代表的な数値シミュレーション例 (その1)

図3 現行スパコン JSS での代表的な数値シミュレーション例 (その2)

3. 構成概要

今回導入する新スパコンシステムを JSS2: Jaxa Supercomputer System generation 2 と呼ぶが、これは、計算エンジン部、ファイルシステム部、アーカイバ部、遠隔部、管理・制御部、スーパーコンピュータネットワーク及び、IO 結合スイッチ、並びに関連するハードウェアとソフトウェアからなる総合システムであり、計算エンジン部は更に、計算システム(呼称、SORA-MA)、プレポストシステム(SORA-PP)、大メモリ計算システム(SORA-LM)、ログインシステム(SORA-LI)から構成される。既出の呼称も含め、表1に呼称の由来を示す。また、図4に呼称も含めた全体の構成概要図を示す。

JSS2の導入は3段階で行い、2014年10月1日から、その第1段階である第1期JSS2が稼働開始した。第2期は2015年4月1日から、SORA-MAシステムの1PFLOPS分が追加され、第3期は2016年4月1日からSORA-MAシステム3PFLOPSが揃い、JSS2システム全系が稼働を開始する。

3.1 計算エンジン部

計算エンジン部とは、スパコンシステムにおける演算を担う部分で、演算の特性に合わせた複数のシステムから構成される。

(1) 計算システム(SORA-MA)

計算システムとは、スパコンシステムの中で最大の演算性能を持つ部分である。高度にチューニングされたプログラム資産をユーザに負担をかけることなく、例えばリコンパイル程度の移行作業で高速に実行することが必要であり、高並列で大規模にバッチ型実行するためのものである。

(2) プレポストシステム(SORA-PP)

プレポストシステムとは、格子生成、可視化アプリケーション及び、ソルバー等の市販アプリケーションをバッチ型及び会話型で実行するためのものである。

また、大規模化する可視化作業等を効率的に実施するために、計算エンジン部で行った数値シミュレーションやデータ処理の結果ファイルを他の場所にコピーすることなく直接操作する遠隔可視化機能を備える。

(3) 大メモリ計算システム(SORA-LM)

大メモリ計算システムとは、分散メモリ並列化が困難で大きなメモリを必要とするアプリケーションをバッチ型ないしは会話型で実行するためのものである。また、プレポストシステムと同様の目的で、遠隔可視化機能を備える。

3.2 ファイルシステム部(SORA-FS)

ファイルシステム部は、大規模な数値シミュレーションやデータ処理のファイル操作及び、ソースコードの保存等を行うための高速・高可用性なクラスター型のファイルシステムである。ディレクトリ単位、ユーザ単位、グループ単位でのQuotaの設定が可能である。

3.3 アーカイバ部(J-SPACE)

アーカイバ部は、階層管理型ストレージシステムとして提供され、計算エンジン部やファイルシステム部等とは独立して稼働するクラスター型のアーカイブ装置で、大規模ファイルや長期保存が必要なファイルの保管庫であり、高速なファイル入出力能力と複数のユーザインターフェースを持つ。

3.4 遠隔部(SORA-TPP,-TLI,-TFS,-SFS,-KFS)

遠隔部とは、角田事業所、つくば事業所、相模原事業所から、計算エンジン部、ファイルシステム部やアーカイバ部を効果的に利用するためのものである。

3.5 管理・制御部

管理・制御部とは、ジョブ管理、課金・統計管理、空調設備や冷却設備等の外部設備を制御する自動運転機能を提供するものである。

3.6 スーパーコンピュータネットワーク

スーパーコンピュータネットワークとは、機構既設ネットワークとスパコンシステムを相互接続するものである。SINET4のL2-VPN機能を用いて遠隔地との相互接続も実現する。

表1 呼称一覧

呼称	呼称の由来
JSS2	Jaxa Supercomputer System generation 2
SORA	Supercomputer for earth Observation, Rockets, and Aeronautics
-FS	- File System
-TFS	- Tsukuba File System
-KFS	- Kakuda File System
-SFS	- Sagamihara File System
-MA	- Main system
-PP	- Pre Post system
-TPP	- Tsukuba Pre Post system
-LM	- Large Memory
-LI	- LogIn system
-TLI	- Tsukuba LogIn system
J-SPACE	Jaxa's - Storage Platform for Archiving, Computing, and Exploring

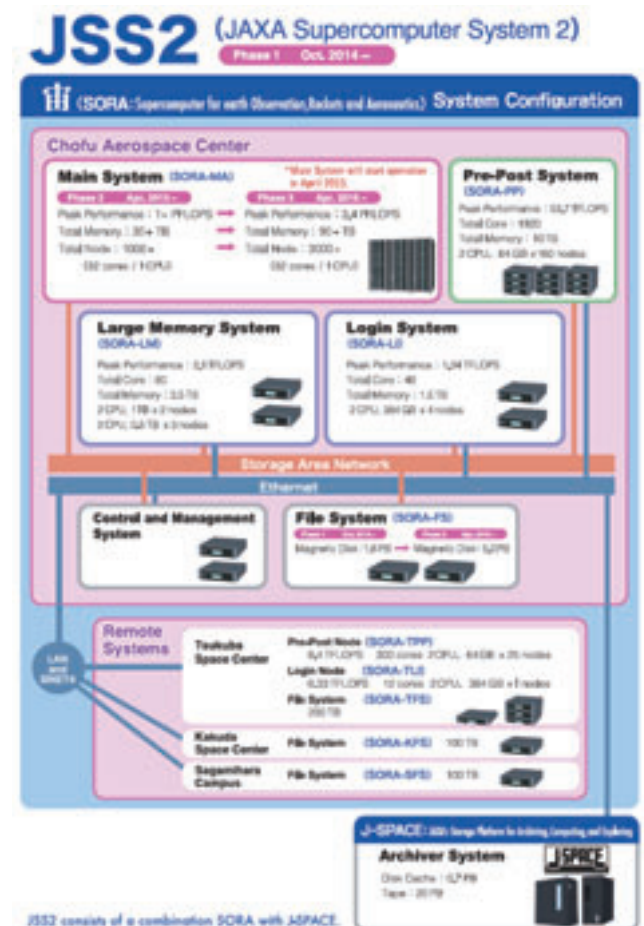


図4 JSS2全体概要図

3.7 IO結合スイッチ

IO結合スイッチとは、計算エンジン部及び管理・制御部とファイルシステム部を、高速かつ可用性を持って相互結合するためのものである。

3. 8 演算ノードの構成

SORA の演算用資源は、-MA,-PP,-TPP,-LM である。-LI も含めて、各資源のノード構成を表2に示す。

MA は、約 3000 ノードからなり、インターコネクトは、富士通株式会社固有の TOFU2 というトーラス構造のネットワークを使用し、ノード当たりのバンド幅は 100GB/s である。各ノードは 32 コアの SPARC アーキテクチャ CPU が 1 台搭載されており、1 ノード当たり 1TFLOPS 以上の演算性能を有する。

PP, TPP は、intel アーキテクチャの CPU がノード当たり 2 台搭載されており、ノード当たりのメモリ量は 64GB のものが 160 ノード(PP)、ないしは 25 ノード(TPP)あり、フルバイセクションの InfiniBand FDR ファットツリー構造のインターコネクトで結合されている。

LM は、PP,TPP と同じく intel アーキテクチャの CPU を搭載しているが、ノード当たりのメモリ量が 512GB ないしは 1TB と大きいのが特徴である。

LI は、ユーザがログインするためのノードで 4 台で構成されており、MA 用クロスコンパイル環境及び、PP, TPP 用コンパイル環境を提供している。また、シェル環境によるファイルシステムへのアクセスもこのノードから行う。

表2 各資源のノード構成

項目	SORA				
	-MA	-PP	-TPP	-LM	-LI
ノード数 [ノード]	約3000	160	25	2 1024GB×メモリ 3: 512GB×メモリ	4
インターコネクト	"Tofu2"	FDR-IB(*) 2段FatTree	FDR-IB 1段	(*)	(*)
CPU数 [個/ノード]	1	2	2	2	2
コア数 [コア/ノード]	32	6 ×2cpu	6 ×2cpu	8 ×2cpu	6 ×2cpu
メモリ量 [GB/ノード]		64	64	1024(2ノード) 512(3ノード)	384
演算性能 [TFLOPS/ノード]	1以上	0.168 ×2cpu	0.168 ×2cpu	0.211 ×2cpu	0.168 ×2cpu
メモリバンド幅 [GB/s/ノード]		59.7 ×2cpu	59.7 ×2cpu	59.7 ×2cpu	59.7 ×2cpu
インターコネクト 幅 [GB/s/ノード]	100	6.8	6.8	6.8 ×2本	6.8 ×2本
備考		GPU	GPU	GPU	

3. 9 利用形態

表3に SORA の利用形態を示す。ジョブキューイングシステムを使用する利用形態としてバッチジョブと会話型ジョブが、使用しない利用形態としてリモートデスクトップがある。

バッチジョブは JSS でも運用していた利用形態で、ジョブをキューイングシステムに登録することにより資源を利用する。

会話型ジョブもキューイングシステムを介して資源を利用する点はバッチジョブと同様であるが、ジョブ投入時のシェルを通じてバッチジョブとの間で会話型形式でジョブを使用することができる利用形態である。複数のノードを使用した可視化や計算処理を会話型で制御したい時の利用形態である。

リモートデスクトップも、会話型でジョブを使用するための形態であるが、1 ノードのみを使用する場合に使用する。ノードの資源管理は、Web インターフェースを持つリモートデスクトップ用ノードの資源管理システムを用いるため、キューイングシステムでのノード管理の対象ではない。160 ノードの何ノードをリモートデスクトップ利用形態として運用するかは、利用状況を観察しながら、適切に調整していく予定である。

表3 SORA 利用形態

利用形態		ジョブ実行に必要なリソース	説明	利用可能なシステム
キューイングジョブ	バッチジョブ	・ノード (・ISVライセン)	“普通の”バッチ実行	MA PP LM
	会話型ジョブ	・ノード (・ISVライセン) ・オペレータ	開始時刻を予約できるバッチ実行	PP LM
インタラクティブジョブ	リモートデスクトップ	・ノード (・ISVライセン) ・オペレータ	キューイングシステムを使用しない実行	PP LM

3. 10 事業別集計

JSS2 では、その利用状況をきめ細かく分析し、効率的運用、利便性やユーザサポートの向上に資するため、事業別集計という機能を実現している。JSS2 で実行される各ジョブと予め登録された事業コード(JAXA の事業を一定の細かさで整理したもの)を関連付け、各事業で JSS2 をどれくらい使用しているかを集計できるようにするものである。また、事業コードとは別の軸として、利用分類という軸での整理も行う。利用分類とは、JSS での特別利用制度、大学共同利用や設備貸付等、利用に際して優先度や利用の仕組みの異なるジョブを分類するためのものである。表4に事業別集計の整理の例を示す。

表4 事業別集計の例

事業コード	利用分類					
	一般	優先	大学共同利用	設備貸付	運用	合計
○×技術の研究開発	100	1,000	-	-	-	1,200
△□の非定常CFD解析技術に関する研究	1,000	0	-	-	-	1,000
▽☆の遷音速動特性予測技術の研究	2,000	1,000	-	-	-	3,100
:	:	:	:	:	:	:
大学共同利用01	-	-	500	-	-	500
設備貸付01	-	-	-	1,000	-	1,000
運用管理利用	-	-	-	-	300	300
合計	3,100	2,000	500	1,000	300	7,100

4. まとめ

JAXA の新スパコン JSS2 の導入目的と構成概要について述べた。

(1) 計算能力の慢性的不足、(2) 乱流、音響、過渡現象といった非定常現象の解析の増加、(3) ファイルシステム性能、アーカイバ性能への要求の高まり、(4) 大規模なプレ・ポスト処理(格子生成・可視化等)への要求の高まり、(5) 数値シミュレーションに加えて、衛星観測データ処理や、実験・観測等のデータのバックアップ機能の要求等に対応し、4 年程度の期間を目途に、数値シミュレーションを進展させ、スパコンによる JAXA 事業促進を目指すスパコンインフラを整備しており、一部稼働を開始した。今後、スパコンの利用実態の詳細な把握を行い、より効果的にスパコンを活用するために、稼働情報を蓄積・処理する仕組み、大規模可視化、遠隔可視化機能の実装を予定している。これらの整備については機会を改めて、初期の利用成果と共に述べたいと思う。

なお、JSS2 は施設設備供用、大学共同利用等の制度を利用し JAXA 外の利用者も使用可能である。JSS2 公開ページ <https://www.jss.jaxa.jp/how2use/>等からお問い合わせいただきたい。