

# 宇宙航空研究開発機構様数値シミュレータⅢ を支えるPRIMEPOWER HPC

## Use of PRIMEPOWER HPC for Numerical Simulator III by Japan Aerospace Exploration Agency

### あらまし

宇宙航空研究開発機構様は、1987年以降、スーパーコンピュータを中核とする数値シミュレータによって、CFD（数値流体力学）に代表される数値シミュレーション技術の発展と普及を推進している。さらに、第3世代数値シミュレータとして富士通PRIMEPOWER HPC2500を中核としたシステム（数値シミュレータ）を導入し、2002年10月より運用を開始した。

数値シミュレータは、中央計算システム、中央マスタストレージシステムおよび中央可視化システムの三つの要素から構成され、大規模SMPシステム、卓越した科学技術計算処理性能、大容量・高速ストレージシステムおよび高速リアルタイム可視化といった特徴を有する。

本稿では、数値シミュレータの概要とその特徴について紹介する。

### Abstract

Since 1987, the Japan Aerospace Exploration Agency has studied and operated numerical simulators using core supercomputers, and developed and expanded numerical simulation technology as typified in computational fluid dynamics (CFD). In October 2002, the laboratory began operating the third-generation numerical simulator (Numerical Simulator III), which incorporates Fujitsu's PRIMEPOWER HPC2500 as the core computer. Numerical Simulator III consists of three components: Central Numerical Simulation System (CeNSS), Central Mass Storage System (CeMSS), and Central Visualization System (CeViS), and features large-scale SMP system capabilities, outstanding scientific and technological computation performance, large-capacity, high-speed storage, and high-speed real-time visualization. This paper describes an overview and the features of Numerical Simulator III.



矢澤克巳（やざわ かつみ）  
研究機関ソリューション部 所属  
現在、宇宙航空研究開発機構様の数  
値シミュレータ システム運用業務  
のサポートに従事。



稲荷智英（いなり ともひで）  
研究機関ソリューション部 所属  
現在、宇宙航空研究開発機構様の数  
値シミュレータ システム運用業務  
のサポートに従事。

## ま え が き

宇宙航空研究開発機構様（Japan Aerospace Exploration Agency, 以下JAXA）<sup>(注1)</sup>は、航空宇宙機に関する開発・飛行実証プロジェクトや、風洞技術、CFD（Computational Fluid Dynamics：計算流体力学）技術、先進複合材料評価技術をはじめとした基盤技術開発を行っており、設立以来、日本の航空宇宙機の研究開発を支えてきた。その代表的な活動の一つとして数値シミュレーション技術の研究開発に取り組んでおり、スーパーコンピュータの高速計算処理能力を利用し、CFDに代表される数値シミュレーション技術の発展と普及を推進している。

本稿では、この研究活動を支える計算基盤システムとして、富士通が新たに提供したPRIMEPOWER HPC2500<sup>(1)</sup>を中核とした先進的な第3世代の数値シミュレータ（Numerical Simulator：NS）システムの概要とその特徴について紹介する。

## 数値シミュレータとは

数値シミュレータ<sup>(2)-(5)</sup>とは、JAXAにおいて運営されているスーパーコンピュータ（以下、スパコン）を中核とする共用計算機システムを指す。第1期数値シミュレータNSは、1987年、富士通製スパコンVP400の導入によりスタートした。VP400は、1 GFLOPSの演算処理性能を有し、3次元翼のナビエストークス解析や全機形状の非粘性解析を初めて可能とした。

第2期の数値シミュレータNSは、1993年、「数値風洞（Numerical Wind Tunnel：NWT）」の導入により始動した。NWTは、JAXAと富士通が共同開発したベクトル並列型計算機である。ナビエストークス方程式をベースとするCFD技術の実機開発への適用を目指し、クリーン全機<sup>(注2)</sup>のパラメタ解析を行うのに1 M（= 100万）格子点の計算を10分で行うことを念頭にVP400の100倍以上の性能をターゲットに開発された。ピーク性能280 GFLOPS、主記憶容量44.5 Gバイトの分散メモリ型並列計算機

であり、NWTは富士通VPPシリーズのプロトタイプとなった。NWTの処理性能は当時としては破格であり、1990年代中盤から後半にかけて日本における航空宇宙CFDの研究開発活動をリードし、3次元ナビエストークス解析を定着させた。しかし、NWTをもってしても航空宇宙における計算需要を満たすには十分とは言えず、内外から新システム更新への期待が高まっていた。

このような背景のもと、第3世代の数値シミュレータとしてNSでは計算処理能力として、「30 M点（1 M点 = 100万点）の多分野統合解析を1日のうちに10ケース処理する」を目標とし、NWTの30倍程度の処理性能がJAXAから要求された。さらに、NSでは、計算能力だけでなく、大規模ジョブの柔軟な運用や更なる利用操作性を兼ね備えたトータルなシステムとしての高い能力も要求された。富士通は、これらの要求に対し、SE・開発部門が一体となったインテグレーションによって第3世代の数値シミュレータとしてPRIMEPOWER HPC2500を中核としたNSシステムを構築し、2002年10月運用を開始した。

## 数値シミュレータ（NS）の概要と特徴

NSは、超高速計算エンジンとなる中央計算システム（Central Numerical Simulation System：CeNSS）、大容量・高速ストレージである中央マスタストレージシステム（Central Mass Storage System：CeMSS）、大規模ビジュアライゼーションサーバである中央可視化システム（Central Visualization System：CeViS）の三つの要素から構成される<sup>(2)</sup>

CeNSSは、9.3 TFLOPSの演算性能と3.5 Tバイトのメモリを持ち、日本最大、世界においても最大級のスカラ型計算エンジンである。また、OSとしてSolaris8を採用し、ISV（Independent Software Vendor）、OSS（Open Source Software）などによるオープン化・標準化を追求している。CeMSSはCeNSSにおける大規模計算能力に見合う十分な容量と高速性を提供し、CeViSでは大規模なプレポスト処理とリアルタイム可視化機能を提供している。NSの構成概念を図-1に示す。

NSでは、以下の大きな特徴を有している。

(1) 高利用性・高信頼性を持つ大規模対称型マル

(注1) 平成15年10月1日に航空宇宙技術研究所、宇宙科学研究所、宇宙開発事業団の3機関が、宇宙航空研究開発機構に統合された。

(注2) 胴体と翼だけの機体。

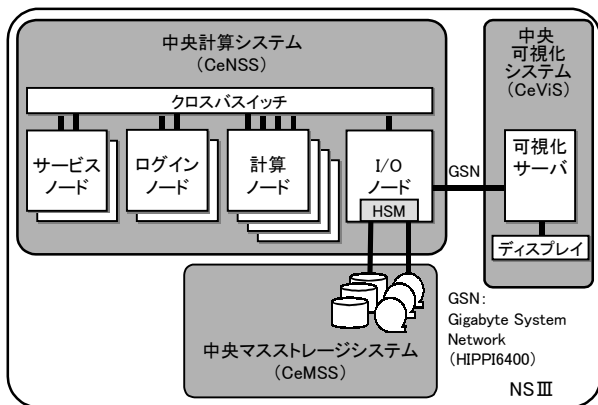


図-1 NS 構成概念  
Fig.1-NS composition concept.

チ プロセッサ 計算機 ( Symmetric Multi Processor : SMP ) システム

- (2) 卓越した科学技術計算処理性能
- (3) 大容量・高速ストレージシステム ( CeMSS )
- (4) 高速リアルタイム可視化

それぞれの内容については、次章以降で説明する。

**高利用率・高信頼性を持つ大規模SMPシステム**

NS の超高速計算エンジンであるCeNSSでは、14筐体のPRIMEPOWER HPC2500を計算ノードとして割り当てている。各筐体は、それぞれ128個のCPU、256 Gバイトのメモリを共有するSMPである。計算ノード以外の4筐体については各種サービスを行う専用ノードとして割り当て、ノードごとに効率的な用途分割を行っている。PRIMEPOWER HPC2500では、各筐体をノードというOSが動作する単位に分割可能であり、CeNSSでは運用要件や利用状況を考慮し、63のノードに分割して運用を行っている。各ノード構成一覧を表-1に示す<sup>(2),(3)</sup>

各ノードは、往復それぞれ4 Gバイト/秒の転送性能を有する高速光インタコネク装置 ( クロスバスイッチ ) によって密結合されており、ノード間での高速データ転送を実現している。<sup>(1)</sup> さらに、Parallelnavi<sup>(注3)</sup> によって、大規模かつ柔軟な並列ジョブ実行環境も実現している。プログラム開発言語としては、Fortran, C, C++を提供し、並列環

(注3) 高速光インタコネク装置で接続されたPRIMEPOWER HPC2500システムのハードウェア性能を最大限に引出すためのプログラム開発環境と高速実行環境を提供する富士通製ソフトウェア。

表-1 ノード構成一覧

ノード名	数	各ノードのCPU数	用途
計算	56	32	ジョブ実行
ログイン	2	64	ログインほか
サービス	4	64	ISVほか
IO	1	128	IO処理

境としては、XPFortran, MPI ( Message Passing Interface ), OpenMP ( SMP向けの並列言語仕様 ), 自動並列を提供している。XPFortranは、NWT-Fortran ( VPP-Fortran ) と互換性を持ち、これにより、NWTで開発したプログラムはコンパイルするだけでCeNSSで利用可能となり、既存プログラム資産の継承をスムーズに行うことを可能としている<sup>(6)</sup> また、Parallelnaviは、CPUなどの資源管理とジョブスケジューリングを柔軟にカスタマイズ可能なインタフェースを持ち、JAXA独自のジョブプライオリティ制御、ノード割当方式などのきめ細かなジョブ運用に対する要望を実現している。

さらに、信頼性・管理面においても機能の向上を図り、PRIMEPOWER自身が持つ高信頼性とネットワークなどの各コンポーネントの2重化、CentricMGRによる一元管理を行い、システムの安定化、運用の効率化を実現している<sup>(7)</sup>

**卓越した科学技術計算処理性能**

CeNSSでは、ラージページ機能による大規模データの高速アクセス、ハードウェアバリア、DTU ( Data Transfer Unit ) による直接ノード間転送機能といったPRIMEPOWER HPC2500の持つハードウェア・ソフトウェア機能によって、高性能を実現している。また、CPUは、SPARC64 Vチップを採用し、クロック1.3 GHz、プリフェッチ、アウトオブオーダー実行、浮動小数点演算の4命令同時発行などのスカラ高速化技術を取り入れている<sup>(1)</sup>

ここでは、CeNSSの実際の演算性能の指標となる標準ベンチマークテスト、およびJAXA所有のコードによる並列アプリケーション実効性能の測定結果について紹介する<sup>(3),(4),(8)</sup>

**標準ベンチマークテスト**

R&D分野の世界標準ベンチマークテストであるLinpackベンチマークを測定し、実効性能5.406 TFLOPSを実現した。この結果は、Linpack

表-2 SPEC OMPM2001登録値

CPU数	スレッド	測定値	スレッド別順位
128	32	22,280	第1位
128	64	28,533	第1位
128	96	30,561	第1位
128	128	31,812	第1位

表-3 JAXA CFDコード概要

名称	処理内容	並列化方式
コード1	化学反応を含むDNS (Direct Numerical Simulation)	MPI + OpenMP
コード2	圧縮性流のLES (Large Eddy Simulation)	MPI + OpenMP

ベンチマークの結果をランク付けする“21st TOP500 List”<sup>(9)</sup>に登録されており、CeNSSは世界のスパコン第7位(2003年6月登録時)に位置付けられている。

また、CeNSSの構成要素であるPRIMEPOWER HPC2500は、R&D分野の標準的な並列処理ベンチマークテストである、SPEC OMPM2001<sup>(10)</sup>において、32、64、96、124スレッドの世界最高値(2003年6月公開時)を記録した。登録された測定値を表-2に示す。本結果より、CeNSSのスカラSMPとしての高い処理性能とスケーラビリティが実証されたと言える。

並列アプリケーション実効性能

CeNSSで実行される並列アプリケーションの実例として、JAXA所有のCFDコード2本の性能を測定した。各コードの概要を表-3に示す。各コードごとに、プロセス数の増加に比例して格子データサイズをスケールアップさせた場合の演算性能を測定した。OpenMPによるスレッド並列数を30に固定し、MPIによるプロセス並列数を変化させて実行した結果を図-2と図-3に示す。いずれも高いスケーラビリティを実測し、最大規模でそれぞれ1,225 GFLOPS、1,014 GFLOPSといった高性能を実現した。これにより、目標としている多分野統合解析の実現にめどがたった。

大容量・高速ストレージシステム (CeMSS)

NSの大容量・高速ストレージであるCeMSS<sup>(11),(12)</sup>は、57TバイトのRAID5ディスク(富士通製PW-D500B1、RAIDディスク80台)と総

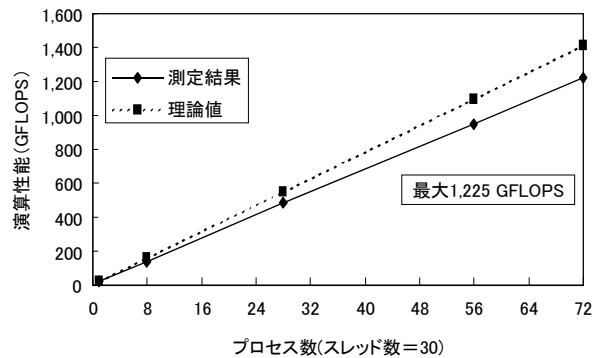


図-2 コード1の性能  
Fig.2-Performance of Code1.

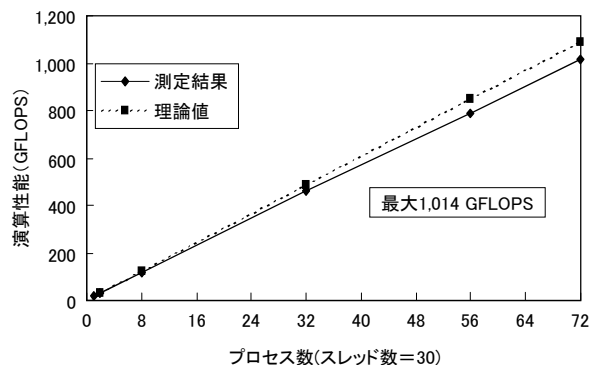


図-3 コード2の性能  
Fig.3-Performance of Code2.

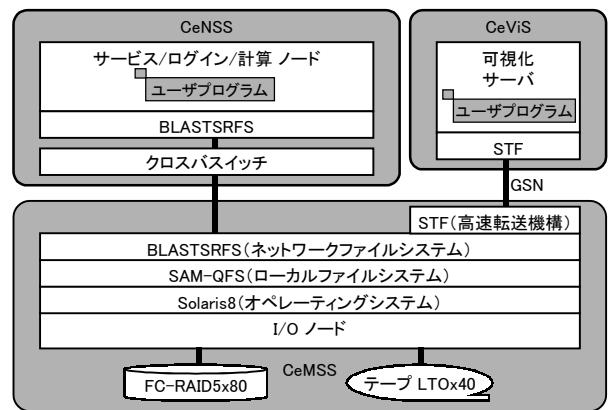


図-4 NS ストレージ概念  
Fig.4-NS Storage concept.

容量620 Tバイトのテープライブラリ (IBM製 IBM3584、LTOドライブ40台) から構成される。PRIMEPOWER HPC2500が持つIOモジュールの柔軟な拡張性によって、これらすべてのディスクとテープライブラリは、120本のFC (Fibre Channel) で一つのIOノードにダイレクトに接続

されている。NS ストレージ概念を図-4に示す。

CeMSSのローカルファイルシステムには、ディスクの高速性とテープの階層管理（Hierarchical Storage Management：HSM）機能を持つSAM-QFS<sup>(注4)</sup>を採用している。SAM-QFSは、ソフトウェアによるディスクストライピング機能とテープストライピング機能を持ち、高速なアクセス環境を実現する。ユーザはこれらの機能により、自分のファイルがディスクとテープどちらに存在するかを意識することなく、透過的かつ高速なファイルアクセスが可能となっている。

CeNSSでのノード間共用ファイルシステムには、高速光インタコネクタ装置の直接転送機構を利用した高速アクセスを実現するBLASTSRFS<sup>(注5)</sup>を採用している。BLASTSRFSでは、UNIX標準API（NFS V3互換）に加え、複数ノードからの更新に対するデータの一貫性・整合性の保証を実現している。

CeViSからCeMSSへのアクセスにおいては、GSN/STP<sup>(注6)</sup>による高速転送を実現するソフトウェア機構としてSTF<sup>(注7)</sup>を採用している。この機構により、CeViSからローカルディスクと同等な操作性と高速なアクセスを実現している。

CeMSSの高速性能を実証するための基礎性能として、SAM-QFS（ディスクキャッシュ性能）、BLASTSRFS、STF、Fortran、それぞれについてシーケンシャルRead、Write性能を測定した。IO性能実測値を図-5に、またシーケンシャルIO最大性能を表-4に示す<sup>(12)</sup>

この結果、ローカルファイルシステムでは最大6.6 Gバイト/秒、ノード間では最大3.2 Gバイト/秒といった世界トップレベルの超高速ストレージ性能を実現している。

SAM-QFSは、HSM動作においてテープストラ

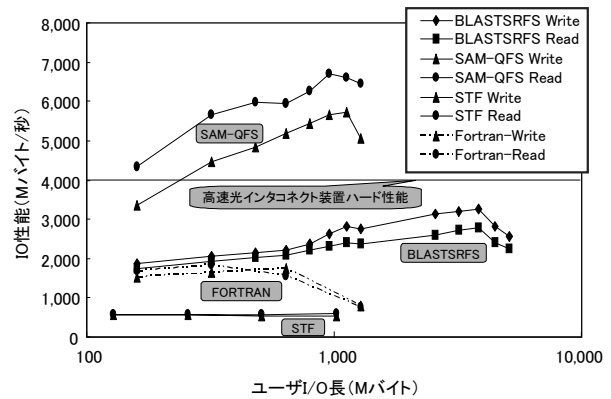


図-5 IO性能実測値

Fig.5-IO performance measurement value.

表-4 シーケンシャルIO最大性能

名称	Mバイト/秒	備考
SAM-QFS (ディスク)	Read	6,623.3 FC×80バス
	Write	5,725.8 FC×80バス
BLASTSRFS	Read	2,774.6
	Write	3,250.4
STF	Read	576.0
	Write	582.3
Fortran	Read	1,822.1 24CPU
	Write	1,648.0 24CPU

表-5 テープストライピング性能（32ドライブ、非圧縮）

名称	Mバイト/秒	備考
SAM-QFS (ディスク)	Archive	353.1 FC×32バス（ドライブ）
	Stage	386.8 FC×32バス（ドライブ）

イピング機能を有しており、Archive（ディスクキャッシュからテープへファイルを書き出す）、Stage（テープからディスクキャッシュへファイルを読み出す）の高速化を実現している。テープストライピング性能を表-5に示す。

CeMSSでは、ストライピング技術と、IOノードにストレージを集中配置することにより、ユーザが特別な意識をすることなく、容易に大容量かつ高速にアクセス可能なストレージシステムを実現している。

### 高速リアルタイム可視化

CeViSは、富士通とSGI社が共同開発した可視化システムであり、SGI社製のOnyx3400（6パイプ<sup>(注8)</sup>、

(注4) 高速・大容量かつ、階層管理（HSM）が可能なファイルシステム。Sun Microsystems社製ソフトウェア。

(注5) 高速光インタコネクタ装置で接続されたPRIMEPOWER HPC2500システム上で動作する分散ファイルシステム。富士通製ソフトウェア。

(注6) GSNはHIPPI6400の技術をベースに、800 Mバイト/秒の高速転送を実現するインタコネクタであり、ANSI規格（HIPPI6400, ANSI NCITS 323-1998）準拠。STP（Scheduled Transfer Protocol）は、GSN上で高速転送を実現するプロトコル。

(注7) PRIMEPOWERとOnyx間をGSN/STPにより高速転送を実現するライブラリを含めた富士通製ソフトウェア。

(注8) 高速に可視化するためのハード機構である、SGI社製 Infinite Reality3グラフィックスパイプライン。

32CPU, 64 Gバイトメモリ), 大規模3次元表示装置(画面サイズ4.6 m × 1.5 m)から構成される。NS では, VisLink<sup>(注9)</sup>の転送部分をGSN/STプロトコルに対応させることにより, CeNSSとCeViS間において500 Mバイト/秒以上の高速データ転送を実現している。これにより, CeNSSで計算した結果の高速かつリアルタイムな3次元可視化が可能となっている。

## む す び

本稿ではPRIMEPOWER HPC2500を中核とした, 宇宙航空研究開発機構様の数値シミュレータの概要とその特徴について紹介した。

宇宙航空研究開発機構様では, NS を用いた大規模数値シミュレーションによって, 世界初のジェットエンジン内の複雑な流れの全段流動解析<sup>(注10)</sup>に成功するなど, 数多くの成果を既に生み出している。また, 2003年10月宇宙航空研究開発機構への統合を終え, さらなる研究活動の推進が期待されている。今後, 著者らはPRIMEPOWERを中核としたNS システムの運用・利用操作性の向上に努め, 日本の航空宇宙分野の研究活動の推進と発展に寄与していきたいと考える。

本稿執筆に当たり, 数値シミュレータの歴史, NS での計算結果など, 数多くの情報を宇宙航空研究開発機構 松尾 裕一様にご提供いただきました。この場をお借りして深く感謝致します。

## 参 考 文 献

- (1) 草野義博ほか：ハイパフォーマンスコンピュータ：PRIMEPOWER HPC . *FUJITSU* , Vol.53 , No.6 , p.444-449 ( 2002 ) .
- (2) 松尾裕一：航技研次期数値シミュレータシステム ( NSIII ) の概要 . 航空宇宙技術研究所特別資料SP-57 , 2003 , p.15-21 .
- (3) Yuichi Matsuo : Numerical Simulator III - A Terascale SMP-cluster System for Aerospace Science and Engineering: Its Design and Performance Issue . International Symposium on High Performance Computing , October 2003 .

- (4) 松尾裕一：PRIMEPOWER HPC2500によるポスト数値風洞システムの構築とその性能概要 . サイエントフィック・システム研究会 計算科学技術分科会第1回会合資料 , August 2003 .
- (5) 数値シミュレータ ホームページ .  
<https://censs.nal.go.jp/>
- (6) 高木亮治：NSIIIにおけるソフトウェア開発環境とユーザ利用環境 . 航空宇宙技術研究所特別資料SP-57 , 2003 , p.28-32 .
- (7) 大川博文：NSIIIにおけるネットワークの設計と実装 . 航空宇宙技術研究所特別資料SP-57 , 2003 , p.33-36 .
- (8) 松尾裕一：航技研数値シミュレータIIIの性能と特性 . 航空宇宙数値シミュレーション技術シンポジウム 2003 <概要集> .
- (9) TOP500 SuperComputer ホームページ .  
<http://www.top500.org/>
- (10) SPEC OMP Benchmark ホームページ .  
<http://www.spec.org/omp/>
- (11) 藤田直行：NSIIIにおける大規模ストレージシステムの設計と性能 . 航空宇宙技術研究所研特別資料SP-57 , 2003 , p.22-27 .
- (12) N. FUJITA et al . : Storage Devices, Local File System and Crossbar Network File System Characteristics, and 1 Terabyte File IO Benchmark on the "Numerical Simulator III" . MSST2003 .

(注9) AVSと連携したリアルタイム可視化を実現する富士通製ソフトウェア。

(注10) 入口から出口までの流動解析。