

[全国共同利用情報基盤センター研究開発論文集 No.30] より

スーパーコンピュータシステム SX-9の紹介と性能評価

大泉健治[†] 江川隆輔[‡] 伊藤英一[†] 岡部公起[‡] 小林広明[‡]

[†]東北大学 情報部情報基盤課システム管理係

[‡]東北大学 サイバーサイエンスセンター スーパーコンピューティング研究部

1. はじめに

東北大学サイバーサイエンスセンターでは、2008年3月にスーパーコンピュータシステムSX-7を、最新鋭のベクトル型スーパーコンピュータSX-9(日本電気(株)製)にリプレースした。今回導入したSX-9システムは全16ノードで構成され、CPU単体で102.4Gflop/s、システム全体では26.2Tflop/sの理論性能を有し、並列処理はもちろん単一CPUでも高い性能を発揮するベクトルプロセッサを特徴としている。メインメモリは、1ノードあたり1Tバイト、システム全体で16Tバイトと圧倒的な大容量共有メモリを搭載し、大規模モデルの計算にも対応可能である。

本稿は、SX-9システムのハードウェア、プログラミングを支えるソフトウェアの紹介をはじめ、本センターの運用構成、そして実アプリケーションを用いた性能評価について説明する。

2. SX-9システムの紹介

2.1. ハードウェアの特徴

スーパーコンピュータSX-9(図1)は、前システムSX-7から引き続きベクトル型並列スーパーコンピュータを継承している。世界最速の1チップベクトルプロセッサは1CPUあたり102.4Gflop/sのベクトル演算性能を持ち、前システムに比べ大幅に性能向上している。これはCPUが3.2GHzの高い周波数で動作することや、ベクトルユニットが持っている浮動小数点演算器のうち加算器と乗算器が、それぞれ2器に増強されたことによる(図2)。この高い性能を持つCPUを16個搭載し1ノードを構成する。

メインメモリはノードあたり1Tバイトの巨大なサイズを登載しており、16個のCPUで共有するSMP(Symmetrical Multi Processing)構成をとる。また、CPUとメモリ間のデータ転送速度(メモリバンド幅)は1CPUあたり256Gバイト/秒の圧倒的に高い転送速度を持つため、大規模かつ高性能な演算を可能にしている。前システムのSX-7とSX-9の性能を表1に示す。

本センターのスーパーコンピュータシステムは16ノードのSX-9で構成されており、前システムのSX-7に比べて約12.5倍の性能向上を実現している。



図1 本センターの SX-9 システム

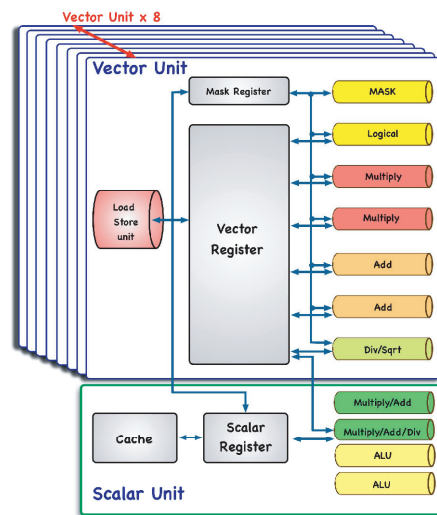


図2 SX-9 ベクトルユニット

表1 SX-7とSX-9の性能比較

		SX-7	SX-9	向上比
CPU	動作周波数	0.551 GHz	3.2 GHz	5.6倍
	ベクトル演算性能	8.83 Gflop/s	102.4 Gflop/s	11.6倍
	メモリバンド幅	64 GB/s	256 GB/s	4倍
SMP ノード	ベクトル演算性能	282 Gflop/s	1.6 Tflop/s	5.8倍
	メモリ容量	256 GB	1 TB	4倍
	メモリバンド幅	1.13 TB/s	4TB/s	3.5倍
	バンクサイズ	16 K	32 K	2.0倍
	ノード間通信速度	32 GB/s	256 GB/s	8.0倍
システム	ベクトル演算性能	2.1 Tflop/s	26.2 Tflop/s	12.5倍
	メモリ容量	2 TB	16 TB	8.0倍

2.2. プログラミング言語と並列処理

プログラミング言語は、従来どおりFortranとC/C++を用意している。それぞれ自動ベクトル化機能、自動並列化機能を有しているため、既存の逐次処理プログラムを修正することなくベクトル化や並列化をコンパイラ任せで実行することができる。このため、ノード内の並列処理であれば比較的簡単に16並列までの並列処理が実現可能である。その他OpenMPによるノード内並列化にも対応している。また複数ノードを使用した並列処理も可能であり、この場合はMPIによる並列化プログラムを用い並列処理を行う。

表2 SX-9で扱えるプログラミング言語とMPI

Fortran90/SX (日本電気)	ISO/IEC 1539-1:1997 自動ベクトル化、自動並列化、OpenMP 対応
C++/SX (日本電気)	ISO/IEC 14882:1998 自動ベクトル化、自動並列化、OpenMP 対応
MPI/SX (日本電気)	MPI-2仕様準拠

2.3. システム構成

本センターのシステムは、今回導入したスーパーコンピュータシステムと既設の並列コンピュータシステムの2つのシステムから構成されている。

SX-9システムを利用するには、並列コンピュータシステム上でプログラミングを行う。ソースファイルの作成、コンパイル、ジョブの実行等、並列コンピュータで全ての作業が可能のようにSX-9向けプログラミング環境を構築してある。これはコンパイル等のプログラミングと演算実行の負荷を分散し、SX-9のパフォーマンスをジョブの実行に専念させるためである。

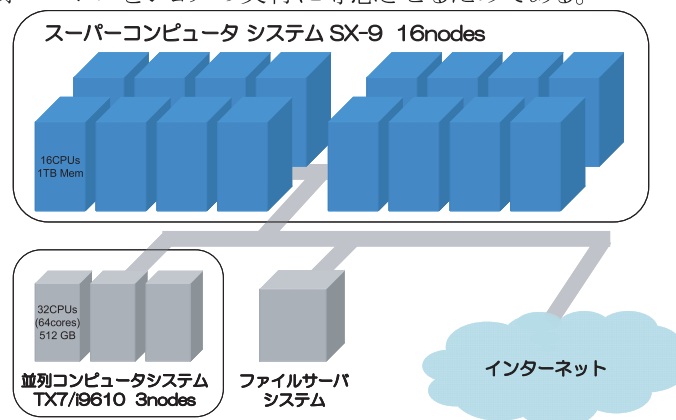


図2 大規模科学計算システムの構成

2. 4. ジョブの運用

SX-9システムのジョブは、ジョブ管理システムNQS II (Network Queuing System II)で管理される。設定したジョブクラスは表3に示す6クラスである。基本的には、逐次処理と並列処理に大きく分けられ、並列処理の場合はさらに並列数でクラスを分けた。実行時間は従来どおり無制限としているので、制限にとらわれることなく長時間のジョブを実行できる。

sクラスとssクラスは、単一CPUで実行する逐次処理プログラムに用いる。ssクラスだけは唯一時間制限が設けてあり、経過CPU時間が1時間以内の短いプログラムの実行を目的としている。このため少ない待ち時間で実行状態に入ることができる。またデバッグ作業も想定しており、4並列までの並列処理も可能でメモリサイズも256GBと大きめに設定している。並列化プログラムは並列数に応じてp8,p16,p32,p64のいずれかのクラスに投入する。p8,p16クラスはノード内の共有メモリを利用できるため自動並列化、OpenMPそしてMPI並列化プログラムが実行できる。p32,p64クラスは複数ノードで実行するため、MPIによる並列化プログラムでノード間を並列処理する。

表3 SX-9システムのジョブクラス

クラス名	利用可能なCPU数	CPU時間制限	メモリサイズ制限
ss	4	1時間	256 GB
s	4	なし	32 GB
p8	8	なし	512 GB
p16	16	なし	1 TB
p32	32	なし	1 TB × 2
p64	64	なし	1 TB × 4

3. 性能評価

3. 1. シングルCPU性能

前章ではSX-9システムの理論性能値について説明したが、この章では実効性能を評価することでその有用性を検証する。ここではSX-9のシングルCPUの実効性能について評価する。評価方法は6つの実アプリケーションをベンチマークプログラムとして用い、本センターのSX-9でコンパイル、実行することで行う。それぞれのアプリケーションは本センターで実際利用されているプログラムなので、他のプログラムを実行するときもひとつの目安として評価できるであろう。6つのアプリケーションは、Earthquake, Turbulent Flow, Antenna, Land Mine, Turbine, Plasmaを用いる。

はじめに、並列コンピュータシステムに搭載されているItanium IIとSX-7, SX-8, SX-9、それぞれの1CPUあたりの性能を評価する。図3の結果が示すとおり、全てのアプリケーションにおいてベクトル型プロセッサがItanium IIを上回ることが分かる。また、SX-9は、理論性能値どおり実効性能においても圧倒的な演算速度を有していることを確認できる。前システムのSX-7とは、最大で7倍、平均で約5倍の実効性能差がある。理論性能値の差ほど高速ではないが、容易に5倍から7倍高速に演算ができることが分かる。

つぎに、Itanium II 1CPUとSX-9 1CPU、Itanium II 16CPUとSX-9 1CPUの比較を行う。理論性能値においては、表4が示すとおりItanium IIに対しSX-9は16倍の性能差があることになるが、図4では最大で158倍、平均でも約67倍の差が見られる。一方、Itanium II 16CPUと比較した場合でも、理論性能値では102.4Gflop/sで全く等しいはずだが全てのアプリケーションでSX-9が上回っている。最大で28.5倍、平均で約9.5倍、SX-9が高速であることが分かる。

Land Mineで顕著に性能差が見られる(図4,図5)が、これはキャッシュヒット率が低く約90%以上をメモリアクセス時間が占めるためである。このようなメモリ転送に大きく依存しているアプリケーションは、SX-9のように高いメモリバンド幅を持つベクトル型計算機が有効であると言える。

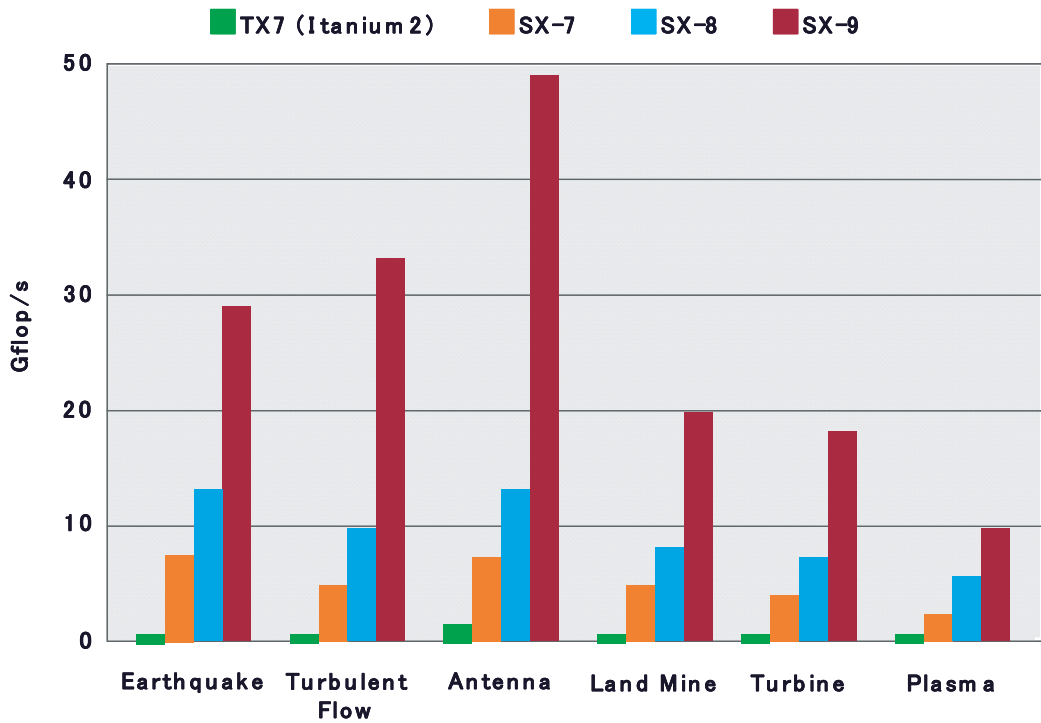


図3 SX-9とSXシリーズ、Itanium II の1CPUあたりの性能

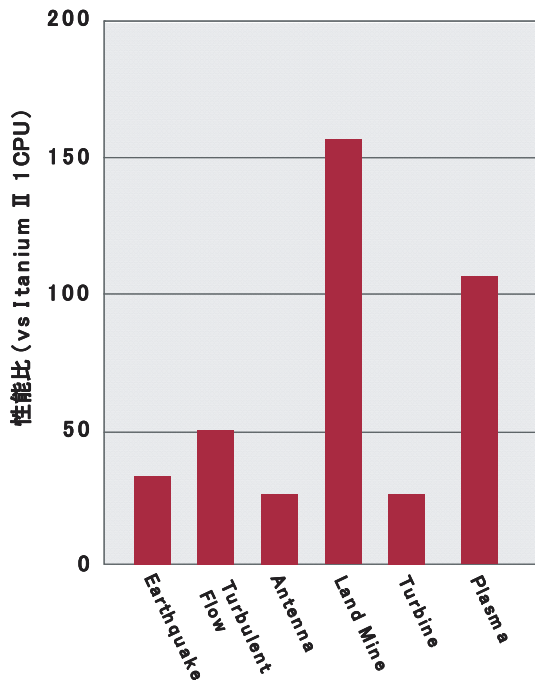


図4 SX-9とItanium II 1CPU

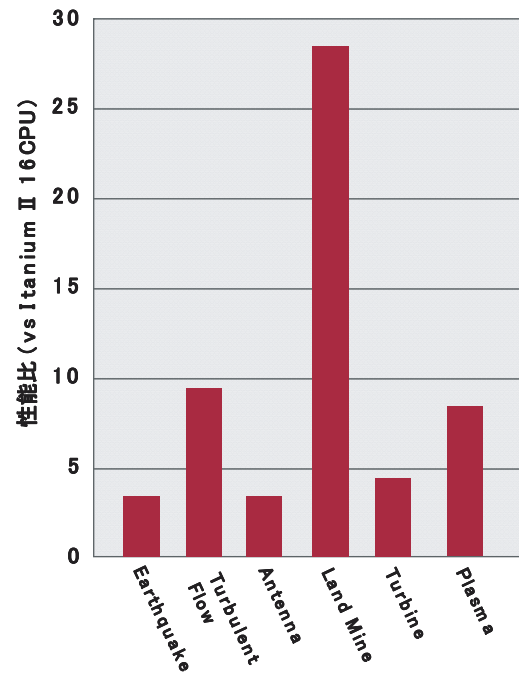


図5 SX-9とItanium II 16CPU

表4 各プロセッサの理論性能

プロセッサ名	理論性能 (1CPUあたり)
SX-9 Vector	102.4 Gflop/s
SX-8 Vector	16 Gflop/s
SX-7 Vector	8.83 Gflop/s
Itanium II (Montecito)	6.4 Gflop/s

3. 2. ノード性能

つぎに、SX-9のノード単位での実効性能について評価する。評価アプリケーションは、前節のAntennaを用いる。1ノードを構成しているCPU数はSX-9が16CPU、SX-7が32CPUのため、評価はSX-9の16CPU、8CPU、1CPUと、SX-7の32CPU、16CPU、8CPU、1CPUで行う。図6は、SX-7の1CPUでの測定値を1としたときの比率で表している。

ノードあたりの性能で見ると、SX-9のCPU数が半分にもかかわらず約5倍、同じCPU数では約10倍の性能を持つことが分かる。また、並列数を増やしたときの性能アップの割合もSX-9の方が大きいことが確認できる。このことからノード内並列処理においてもSX-7に比べ高速な処理が可能で、並列処理のスケーラビリティも高いことが分かる。SX-7用に最適化されたアプリケーションであれば、SX-9でも容易に高い実効性能を達成できると言える。

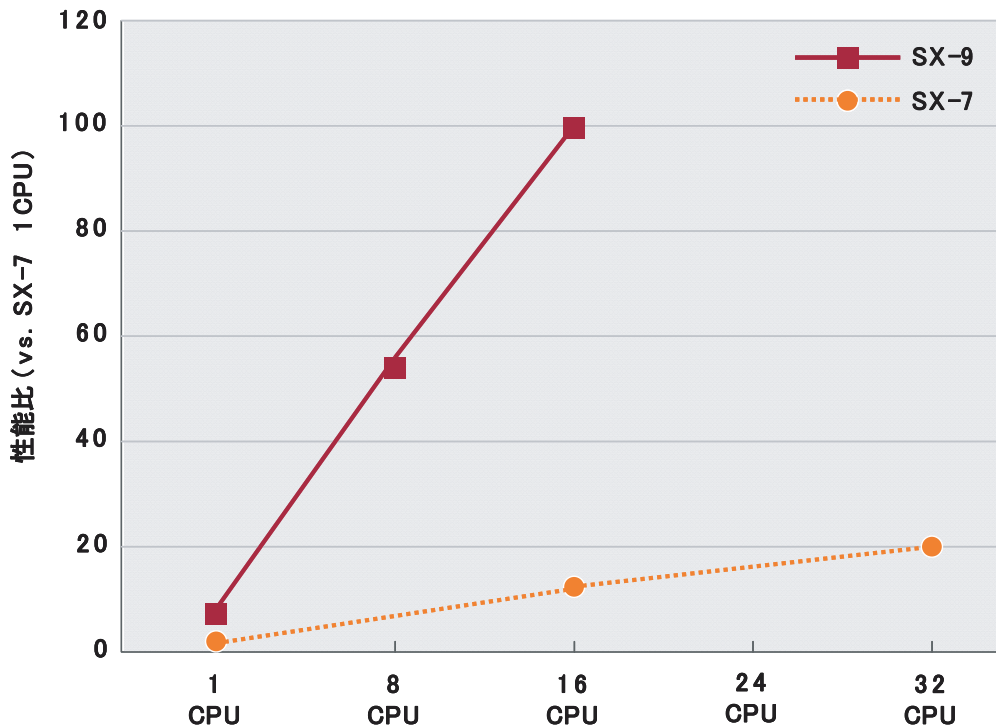


図6 SX-9とSX-7のノード性能

3. 3. マルチノード性能

ここでは、SX-9を複数ノード利用した場合の性能評価を行う。より大規模なシミュレーションを行うためには、複数ノードを用いた並列処理が不可欠となる。評価には、東北大学工学研究科中橋研究室と共同研究により開発された非圧縮性流体シミュレーションを用いる。このアプリケーションは、全てのタスクをMPIで並列処理する方法(Flat MPI)と、ノード間はMPIを用いノード内はOpenMPで並列化するハイブリッド並列処理(Hybrid)の2通りの並列処理が可能である。Flat MPI方式とHybrid方式、それぞれの並列処理についてモデルサイズの小さい版(Coarse)と大きい版(Fine)の評価を行う。図7は、マルチノードの並列処理64CPU,32CPUで実行した場合と、シングルノードの16CPUで実行した場合の性能を、1CPUでの性能を1としたときの性能比として表している。

Coarse版、Fine版いずれのモデルにおいてもFlat MPI方式がHybrid方式より高い実効性能に達しておりスケーラビリティが高いことが分かる。64CPUにおいては、約48倍の性能を持つことが確認できる。Hybrid方式の場合、MPIとOpenMPそれぞれの通信処理が特定のタスクに集中してオーバーヘッド時間が大きくなるが、Flat MPI方式の場合は各タスクが並列に通信処理を行えるため効率の良い並列処理ができると言える。

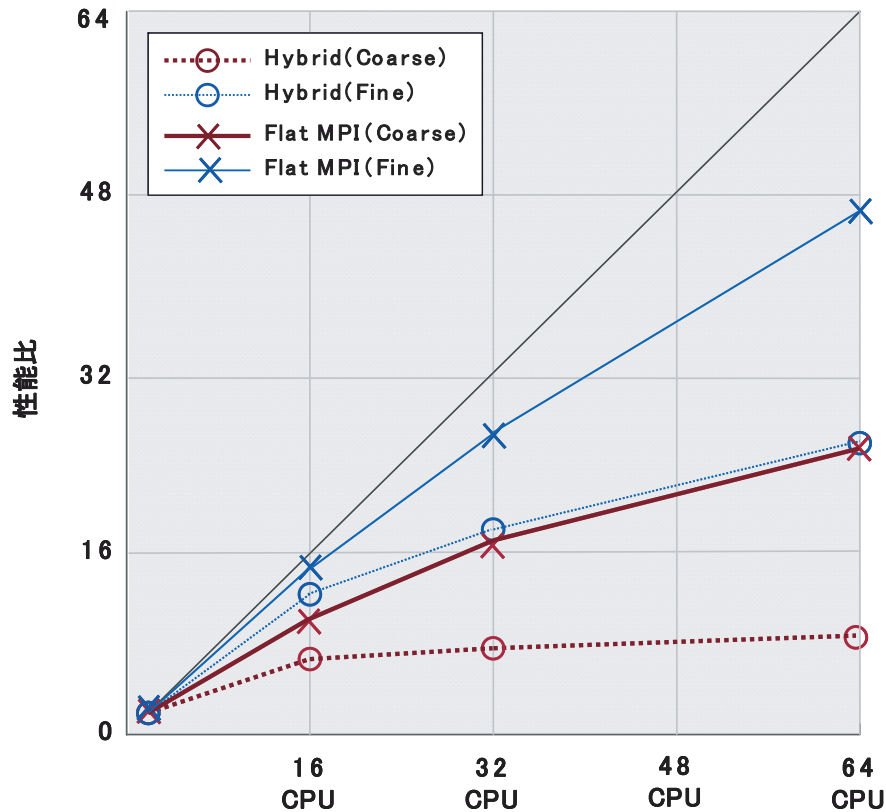


図7 SX-9のマルチノード性能

4. 終わりに

本稿では、今年3月に導入されたスーパーコンピュータSX-9について紹介し、当センターでの運用構成と性能評価について説明した。性能評価には、本センターで利用されている実アプリケーションをベンチマークに用い、前システムのSX-7に対してシングルCPU性能で5倍から7倍、ノード性能で約5倍高速に演算ができることを示した。実効性能を知ること、利用者がプログラミングする際にも有用な指標となると思われる。

新機種が導入されたこの機会に、ぜひベクトル型スーパーコンピュータSX-9をお試しいただき、さらには研究の強力なツールとしてご活用いただけたら幸いである。

謝辞

本稿を執筆するにあたり多くの方々にご協力ご支援を賜った。東北大学の長谷川昭教授、中橋和博教授、澤谷邦男教授、佐藤源之教授、山本悟教授、笹尾泰洋教授、土屋史紀教授、東京理科大学の塚原隆裕研究員、海洋研究開発機構海洋工学センターの有吉慶介研究員には評価プログラムを提供していただいた。また、日本電気㈱の撫佐昭裕様、NECシステムテクノロジーの曾我隆志様、塩田和永様、NECソフトウェア東北の下村陽一様にはシミュレーションにおいて多大なる協力をいただいた。皆様に心より感謝申し上げます。

参考文献

- [1] 江川隆輔, “新大規模科学計算システム スーパーコンピュータSX-9の性能評価”, 東北大学サイバーサイエンスセンター大規模科学計算システム広報SENAC, Vol.41, No.2, pp.11-18, 2008.
- [2] Ariyoshi, K., T. Matsuzawa, and A. Hasegawa, “The key frictional parameters controlling spatial variation in the speed of postseismic slip propagation on a sub-duction plate boundary,” *Earth and Planetary Science Letters*, vol. 256, pp.136--146, (2007)
- [3] 塚原 隆裕, 岩本 薫, 河村 洋, “平行平板間乱流における流体線の直接数値シミュレーション”, 東北大学情報シナジーセンター大規模科学計算システム広報 SENAC, Vol. 39, No. 1, pp. 47--58,(2006).
- [4] Takagi, Y., H. Sato, Y. Wagatsuma, K. Mizuno, and K. Sawaya, “Study of high gain and broadband antipodal fermi anetnna with corrugation,” *In proceedings of 2004 International Symposium on Antennas and Propagation, Volume 1*, pp. 69--27, (2004).
- [5] Kobayashi, T., X. Feng, and M. Sato, “FDTD simulation on array antenna sar-gpr for land mine detection,” *In Proceedings of SSR2003: 1st International Symposium on Systems and Human Science*, pp. 279--283, (2003).
- [6] 笹尾泰洋, 堀田翼, 山本悟, “タービン多段静動翼列流れの大規模並列計算を実現する数値タービンの研究,” 東北大学情報シナジーセンター大規模科学計算システム広報 SENAC, Vol. 40 No.3, pp.15--24,(2007).
- [7] 加藤雄人, 小野高幸, 飯島雅英, “不均質媒質中でのプラズマ波動の伝搬についての計算機実験,” 東北大学情報シナジーセンター大規模科学計算システム広報 SENAC, Vol. 37, No.1, pp. 13 -- 19, (2004).
- [8] Takahashi, Nakahashi, et al., “Large-Scaled computation of Incompressive Flows on Cartesian Mesh Using a Vector-Parallel Supercomputer,” *presented at Parallel CFD 2008*, Lyon France, 5/18-22, 2008