

[全国共同利用情報基盤センター研究開発論文集 No.28]より

新並列コンピュータシステムについて

小野敏^{*1} 大泉健治^{*1} 伊藤英一^{*1} 岡部公起^{*2} 小林広明^{*2}

^{*1} 東北大学情報部情報基盤課、^{*2} 東北大学情報シナジーセンター

1. はじめに

東北大学情報シナジーセンター（以下、本センターと略す）では、常に最高・最新鋭の大型計算機を導入し、大規模科学技術計算の要求に応じてきた。本センターの大規模科学計算システムは、主に研究室のワークステーションでは行えないような大規模シミュレーションに用いられている。このうちベクトル演算に適したものは、スーパーコンピュータシステムで、それ以外のもは並列コンピュータシステムで実行されている。これまでの本センターの計算システムの利用状況を見ると、並列コンピュータについては演算性能と主記憶容量の不足は深刻であり、大規模なシミュレーションを行うには強化する必要があった。また、2003年の1月に更新したスーパーコンピュータSX-7上ではベクトル化率の高いプログラムが常に実行中であり、利用者の計算需要はスーパーコンピュータの処理能力の限界を超え、増強をせまられていた。このことから、新並列コンピュータシステムとして、汎用的な利用を目的としたスカラ型並列コンピュータと、既設のスーパーコンピュータシステムを増強するベクトル型並列コンピュータを導入することとした。

本稿では、新並列コンピュータシステムのシステム構成、運用方法、利用者ジョブを用いた演算性能評価について述べる。

2. システム構成

新並列コンピュータシステムは、スカラ型並列コンピュータ TX-7/i9610^[1]（日本電気(株)製）とベクトル型並列コンピュータ SX-7C^[2]（日本電気(株)製）およびファイルサーバから構成される。既設のスーパーコンピュータSX-7を含めたシステム全体の構成を図1に示す。

スカラ型並列コンピュータは、3つのノードから構成される。1つのノードは、Intel最新の64ビットプロセッサ Itanium2(IA-64)を32台(64プロセッサ・コア)と512GBの主記憶容量を搭載している。1CPUあたり6.4GFLOPSの演算性能を有し、ノードあたりの最大性能は409.6GFLOPSである。なおOSはSUSE LINUXを採用している。

ベクトル型並列コンピュータは、5つのノードから構成される。1つのノードは、16GFLOPSのベクトルプロセッサ8台と128GBの主記憶容量を搭載している。各ノードは、ノード間接続装置(IXS)で高速に相互接続されており、1ジョブで5ノードすべて(640GFLOPS、640GB)を占有する大規模な並列処理専用のシステムである。なおOSはUNIX System Vに準拠しているSUPER-UXである。

ファイルサーバは、スカラ型並列コンピュータとベクトル型並列コンピュータの両方にファイバチャネルで接続され、GFS(グローバルファイルシステム)による高速なファイル共有システムを提供している。ファイルサーバは、10 テラバイトの記憶容量を持つ磁気ディスク装置と、15 テラバイトの記憶容量を持つテープバックアップ装置から構成されている。なお、スーパーコンピュータ SX-7 とは、従来通りの NFS によるファイル共有となっている。

バッチ処理としては、スカラ型並列コンピュータ、ベクトル型並列コンピュータ、スーパーコンピュータのすべてに NQSII (Network Queuing System)を導入することで、ジョブの一元管理が可能となり、利便性の高いジョブ投入環境となっている。また、学内ネットワークシステム TAINS/G に接続され学内外からの利用が可能である。

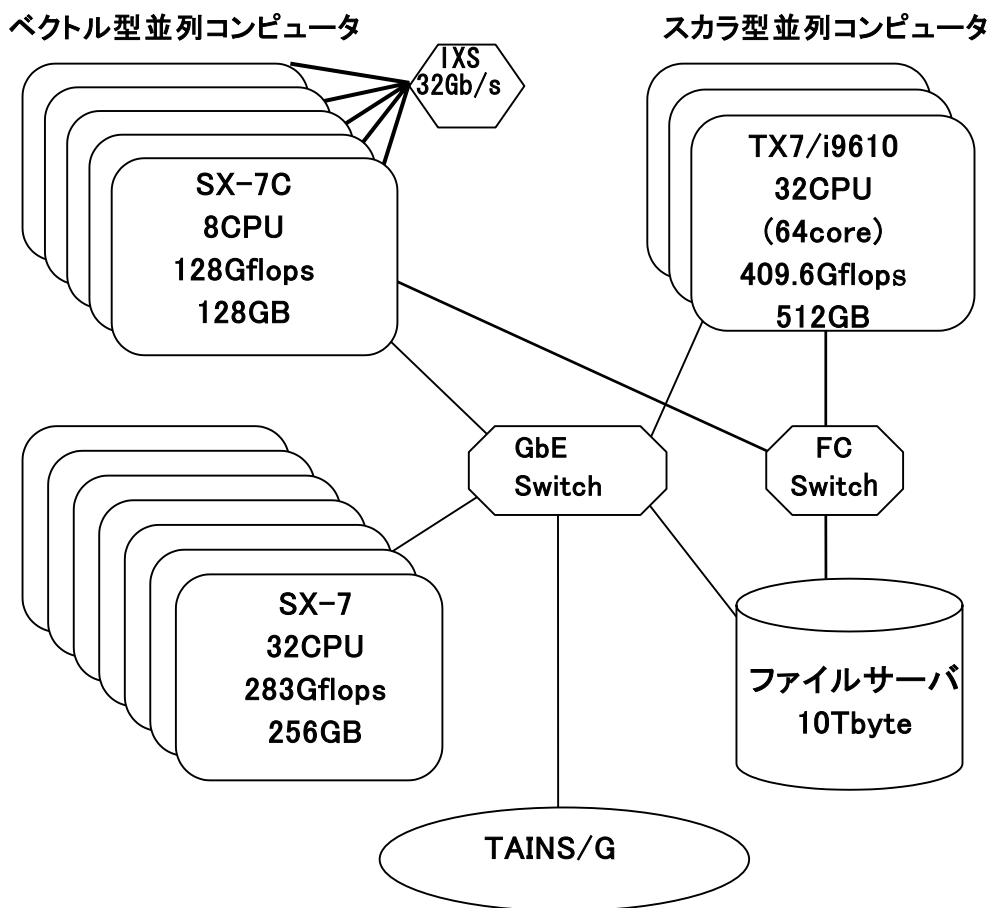


図 1 新並列コンピュータシステムの構成

3. 運用方法

利用者がログインするマシンはスカラ型並列コンピュータ中の1ノードであり、スカラ型並列コンピュータの他のノードとベクトル型並列コンピュータは、すべてNQSⅡを用いその資源を利用するようにしている。会話型処理用として8CPU/64GBを割り当てており、コンパイル、デバッグ、バッチジョブの投入、アプリケーションおよびフリーソフトウェアの実行が可能である^{[3][4]}。スカラ型並列コンピュータは、ジョブクラスとして非並列のasクラス、8並列用のa8クラス、16並列用のa16クラス、32並列処理用のa32クラス、64並列処理用のa64クラスを用意している。ベクトル型並列コンピュータは、既設のスーパーコンピュータと一体的に運用し、MPIにより5ノードすべてを占有する、40並列処理専用のジョブクラスpxとして位置づけている。なお、ジョブの大規模化、長時間化に対応するために、ssクラス以外のすべてのジョブクラスは従来同様CPU時間を制限していない。

並列処理機能としては、スカラ型並列コンピュータとスーパーコンピュータはFortran、C、C++の自動並列化機能を有しているため、コンパイル時に自動並列化のオプションを指定するだけで良い。また、OpenMP、MPIによる並列プログラミングも可能である。なお、ベクトル型並列コンピュータはMPIの使用を前提とするが、自動並列処理と組み合わせたハイブリッド並列処理も可能である。

表1 スカラ型並列コンピュータのジョブクラス

ジョブクラス名	並列数	CPU 時間制限	メモリ制限
as	1	無制限	8GB
a8	8	無制限	64GB
a16	16	無制限	128GB
a32	32	無制限	256GB
a64	64	無制限	512GB

表2 スーパーコンピュータのジョブクラス

ジョブクラス名	並列数	CPU 時間制限	メモリ制限
ss	1	1 時間	8GB
s	1	無制限	8GB
p8	8	無制限	64GB
p16	16	無制限	128GB
p32	32	無制限	256GB
px(MPI 専用)	40	無制限	ノードあたり 128GB 合計 640GB

4. 演算性能の評価

センター利用者が開発した 5 本のプログラムと Gaussian について、旧並列コンピュータシステム TX7/AzusA の実行性能を 1 とした時の、スカラ型並列コンピュータ TX7/i9610 の実行性能比を図 2 に示し、スーパーコンピュータ SX-7 の実行性能を 1 とした時の、ベクトル型並列コンピュータ SX-7C の実行性能比を図 3 に示す。なお、利用者プログラムについては 1CPU の時の実行時間の比較、Gaussian については 8CPU の時の実行時間の比較である。各プログラムの内容は、不均質媒質中における電磁波伝播・散乱の解析、超帯域アンテナの解析、二次元混合火炎の不安定性の解析、非定常剥離流れの熱伝導解析、プレート境界面上における滑りの解析である。評価に用いた計算機の理論性能比は、TX7/AzusA を 1 とした場合 TX7/i9610 では 2 倍であり、SX-7 を 1 とした場合 SX-7C では 1.8 倍である。

TX7/i9610 の性能は、図 2 より電磁波の伝播・散乱の解析において TX7/AzusA の 2.1 倍、超帯域アンテナの解析において 3 倍、混合火炎の不安定性の解析において 3.7 倍、非定常剥離解析において 3.4 倍、プレート境界面の滑り解析において 4.4 倍、Gaussian の解析で 3.1 倍となっており、理論性能比(2 倍)以上の性能を示している。これは、表 3 に示すように、演算性能が 2 倍になったことに加え、キャッシュやメモリバンド幅が大きくなった効果であると考えられる。

SX-7C の性能は、図 3 より電磁波の伝播・散乱の解析において SX-7 の 1.8 倍、超帯域アンテナの解析において 1.7 倍、混合火炎の不安定性の解析において 1.8 倍、非定常剥離解析において 2 倍、プレート境界面の滑り解析において 1.8 倍となっており、理論性能程度の性能を示している。

電磁波の伝播・散乱解析のプログラムにおいて、TX7/AzusA の性能を 1 とした時の TX7/i9610 の性能は 2.1 倍程度であるのに対し、SX-7 の性能は 71.6 倍、SX-7C の性能は 129.2 倍と大きくなっている。このプログラムは、大きな配列データに対して不連続なアクセスを頻繁におこなっているために、スカラ型の並列コンピュータではキャッシュミスが多発しているのに対し、ベクトル型コンピュータでは強力なメモリバンド幅の効果が顕著に現れていると考えられる。

超帯域アンテナ解析のプログラムにおいて、各コンピュータの 1CPU の実行性能を 1 とした時の並列性能を図 4 に示す。この図から TX7/i9610 の並列性能は 8CPU、16CPU とともに TX7/AzusA を上回り、4 機種の中で最も並列性能が良いことがわかる。これは、このプログラムの並列化率が 99.5% であり、オンキャッシュでほとんどの処理が行われている(キャッシュヒット率 99.9%) ためであると考えられる。

表 3 TX7/AzusA と TX7/i9610 のスペック比較

	TX7/AzusA	TX7/i9610
CPU 数/ノード	16CPU	64CPU
最大主記憶容量	32GB	512GB
単一 CPU 性能	3.2GFLOPS	6.4GFLOPS
キャッシュサイズ(L3)	4MB	12MB
メモリバンド幅	2.1GB/s	8.5GB/s

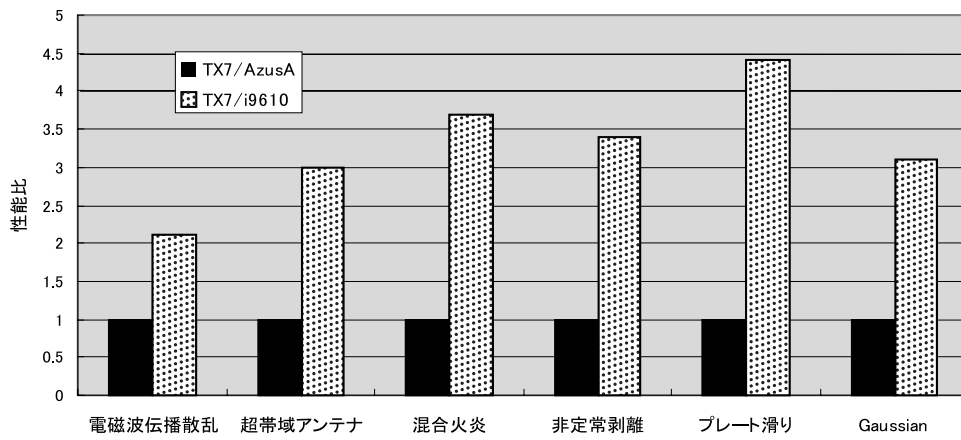


図 2 TX7/AzusA を 1 とした時の TX7/i9610 の性能比

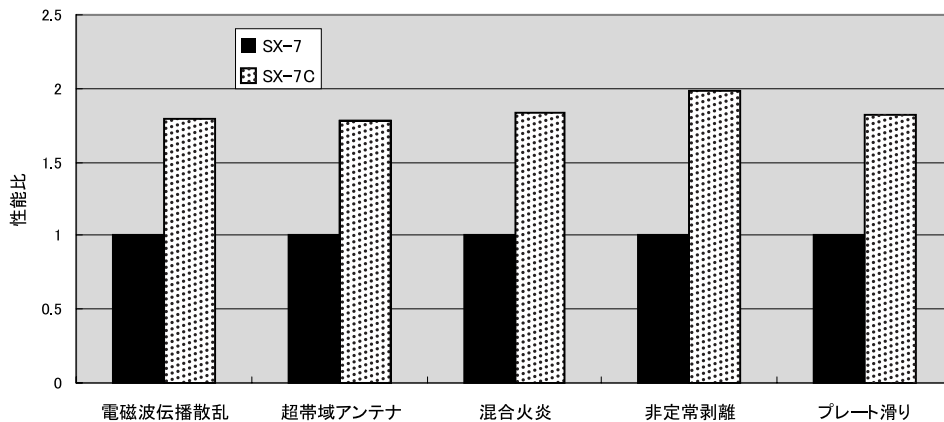


図 3 SX-7 を 1 とした時の SX-7C の性能比

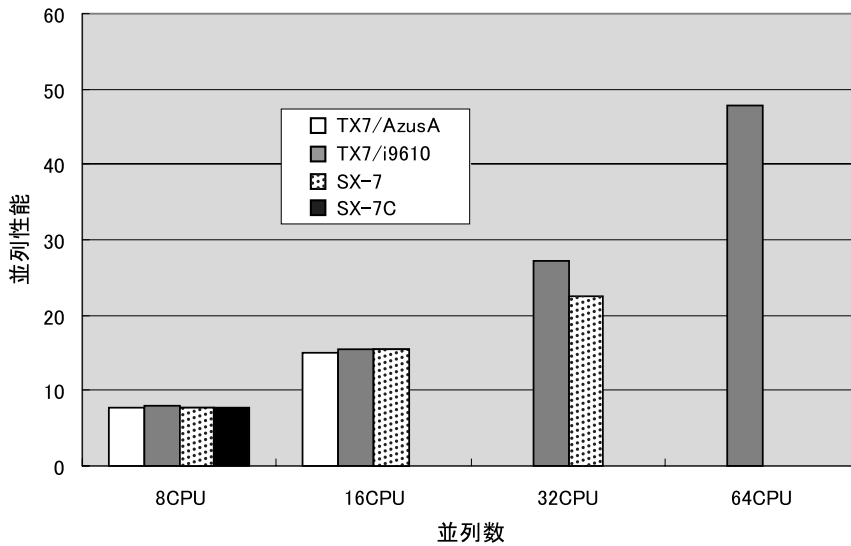


図4 並列性能

5. おわりに

本稿では、新並列コンピュータシステム TX7/i9610、SX-7C についてシステム構成、運用方法と利用者ジョブを用いた演算性能評価について紹介した。スカラ型並列コンピュータ TX7/i9610、ベクトル型並列コンピュータ SX-7C ともに期待通りの性能向上があった。また、本センターにおいては、システム更新によるハードウェアの性能向上を超える勢いで、利用者のジョブは大規模・長時間化してきている。このため、1997 年から高速化推進活動を行い利用者プログラムのベクトル化率、並列化率の向上を図ってきたが、この高速化推進活動をより一層強化する必要があると考える。

謝辞

本プログラム評価において快くプログラム提供を行っていただきました、東北大学東北アジア研究センター佐藤研究室、東北大学大学院理学研究科地震・噴火予知研究観測センター長谷川研究室、東北大学工学研究科航空宇宙工学専攻升谷研究室、東北大学工学研究科電気通信工学専攻澤谷研究室、東北大学工学研究科機械システムデザイン工学専攻太田研究室に感謝いたします。

参考文献

[1] 日本電気株式会社 渋谷俊輝、木村真行、近藤秀俊「並列コンピュータ TX7/i9610 のハードウェア」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 Vol. 39 No. 2 2006-4

[2] 日本電気株式会社 百瀬真太郎、幅田伸一、萩原孝「スーパーコンピュータ SX-7C ハードウェア」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 Vol. 39 No. 2 2006-4

[3] システム管理係、システム運用係、スーパーコンピューティング研究部「スーパーコンピュータ SX-7、SX-7C の利用法」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 Vol. 39 No. 2 2006-4

[4] システム管理係、システム運用係、スーパーコンピューティング研究部「並列コンピュータ TX7/i9610 の利用法」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 Vol. 39 No. 2 2006-4