

大規模科学計算システムの構築と運用

情報シナジーセンター スーパーコンピューティング研究部

岡部 公起・小林 広明・後藤 英昭

システム管理掛 伊藤 英一 システム運用掛 加藤 昇

東北大学情報シナジーセンターの大規模科学計算システムについて、その位置づけとシステム構成、スーパーコンピュータSX-4の利用状況、計算機の性能を最大限に引き出すための高速化推進活動について報告する。

1. はじめに

東北大学情報シナジーセンター（本センター）の大規模科学計算システムは、全国の大学等の教員、研究者が学術研究のために利用する全国共同利用のシステムである。本センターは、その前身の一つである大型計算機センターの設立（1969年）以来、研究室のレベルをはるかに超える最高性能の計算機システムを設置して、最先端の学術研究を強力に支援・推進してきた。加えて、利用者にとって使い勝手の良いシステムの構築、他では実行できない大規模プログラムの実行環境の整備、計算機システムの最大限の性能を引き出すための高速化推進活動等を行ってきた。こうした努力の積み重ねの結果、本センターの利用者は2000人を超え、全体の2/3は東北大学の研究者が占めるものの、残り1/3は他大学等の研究者であり、全国共同利用の計算機システムとして果たす役割は極めて大きい。

2. システムの構成

現在、高性能のワークステーションやパーソナルコンピュータが研究室にも導入され、小規模の計算や可視化処理等は研究室で行われている。それに伴い、従来の多種多様なサービスの提供に代わり、大規模科学計算システムには、研究室のレベルをはるかに超える高い演算性能と大容量のメモリが求められるようになってきた。このため、大規模科学計算システムとして、スーパーコンピュータシステムと並列コンピュータシステムの2つの計算機システムを設置している（図1）。

スーパーコンピュータシステムは、最新のベクトル型スーパーコンピュータSX-7（日本電気㈱製、2003年1月導入）で、32個のCPUと256GBのメモリから成るシステム（ノード）が7台、16個のCPUと128GBのメモリのノードが1台の構成である。ベクトル演算機構による高速化に加え、並列処理による高速化の機能を有しているため、ノードあたり282.56GFLOPSという高いベクトル演算性能と256GBのメモリを利用して、研究室等では到底実行し得ないような大規模プログラムの実行が可能である。なお、並列処理の機能としては、共有メモリ型の特徴を活かした自動並列化機能が中心であり、他にOpenMP、MPI、HPFも利用できる。

並列コンピュータシステムは、スカラ型並列サーバTX7/AzusA（日本電気㈱製）で、16個のCPUと32GBのメモリから成るノードが4台、16個のCPUと16GBのメモリのノードが3台の構成である。ベクトル演算に不向きなプログラムを並列処理によって高速に実行するシステムで、計算需要の増大に対応するため2002年1月に、並列演算サーバExemplar/X（HP社製、48CPU、12GB）中心のシステムから更新したものである。ノードあたりの演算性能は以前の11.52GFLOPSから51.2GFLOPSに向

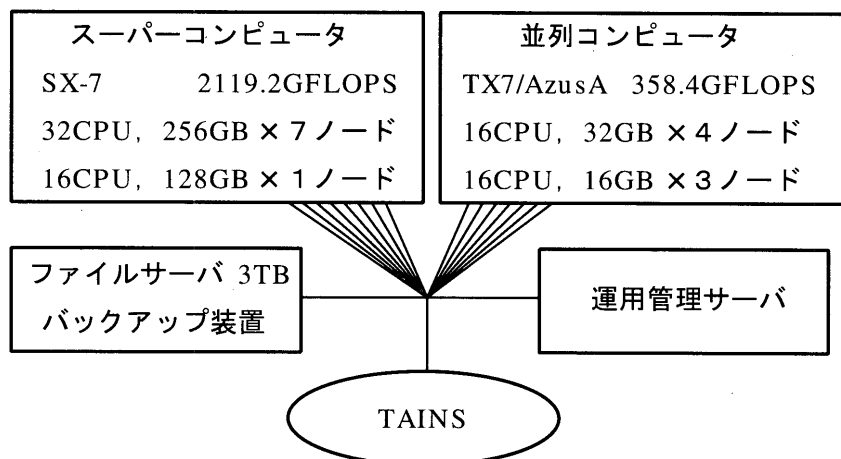


図1 大規模科学計算システム

上しており、さらなる利用の増大が予想されている。並列処理機能としては、自動並列化機能が中心であり、他にOpenMP、MPIが利用できる。

ファイルサーバは、スーパーコンピュータと並列コンピュータに共通のファイル装置であり、ベクトル化率の程度に応じて、両システムを容易に使い分けられることのできるシステム構成としている。また、運用管理サーバは、システム全体にわたる、利用者情報の登録・管理、課金情報の管理を一元的に行う。

3. 利用状況

スーパーコンピュータの利用状況として、旧機種であるベクトル型スーパーコンピュータSX-4（日本電気㈱製、(32CPU,8GB) × 4ノード、1998年1月～2002年12月）の5年間について、ジョブクラスごとのCPU時間の分布を表1に示す。ここで、ジョブクラスp8、p16、p32は、それぞれ8CPU、16CPU、32CPUによる並列処理のクラスであり、sは逐次処理（非並列）のクラスである。各年ともp32のCPU時間が最も多く、1998年と2002年は60%を超えている。また、p8、p16、p32の合計は、毎年90%以上であり、2002年は95.9%にも達している。SX-4は現SX-7同様共有メモリ型のシステム（1ノード32CPU）であり、自動並列化機能を有していたことがこの結果に現れているとともに、今後も自動並列化機能が不可欠であることを示している。

表1 ジョブクラス別のCPU時間分布 (%)
(スーパーコンピュータSX-4)

表2 は会話型処理以外のジョブについて、使用したCPU時間の分布を年ごとに示したものである。100時間未満のジョブについては、1998年の21.3%をピークに減少の一途をたどり、2002年には14.5%にまで低下した。また、100時間以上～1000時間未満のジョブについては、2000年までは40%を超え利用の中心であったが、2001年以降は

年	1998	1999	2000	2001	2002
会話型	0.7	0.8	0.6	0.3	0.2
s (非並列)	7.7	6.2	8.9	3.8	3.9
p8 (8CPU)	12.3	13.6	12.6	11.4	11.3
p16 (16CPU)	18.6	23.8	23.5	28.8	24.4
p32 (32CPU)	60.7	55.6	54.4	55.7	60.2

2000時間以上のジョブが急激に増えて利用の中心を成すようになり、ジョブの大規模化が進んでいることを示している。短時間のジョブは研究室のワークステーションで実行し、研究室で実行できないような大規模のジョブを本センターで実行するという利用の形態が日常化してきているものと思われ、今後の整備・拡充の方向を指し示していると考えられる。

表2 ジョブのCPU時間の分布 (%)

年	1998	1999	2000	2001	2002
1 時間未満	0.4	0.3	0.2	0.3	0.2
1 時間以上	3.3	2.4	2.9	2.3	1.7
10 時間以上	17.6	14.7	16.8	12.7	12.6
100 時間以上	44.5	40.8	40.5	26.0	36.3
1000 時間以上	15.8	13.3	13.0	13.2	11.8
2000 時間以上	18.4	28.5	26.6	45.5	37.4

4. 高速化推進活動

スーパーコンピュータSX-4を運用した5年間における、並列ジョブのCPU時間の分布をベクトル演算率と並列化率について千分率で示したのが表3である。ここで、ベクトル演算率と並列化率は下に示す式で求めたものであり、特にベクトル演算率はベクトル化率の代用となるものである。また、表を見やすくするために数値0は空白に置き換えている。

$$\text{ベクトル演算率} = \frac{\text{ベクトル演算要素数}}{\text{ベクトル演算要素数} + \text{スカラ演算要素数}} \times 100 (\%)$$

$$\text{並列化率} = \frac{\text{総 CPU 時間}}{\text{占有 CPU 時間}} \times 100 (\%)$$

表より、ベクトル演算率90%以上でかつ並列化率90%以上のジョブで占めるCPU時間は千分の552である。また、ベクトル演算率90%以上のジョブのCPU時間は千分の905、並列化率80%以上のジョブのCPU時間は千分の749である。ベクトル演算率と並列化率が共に80%未満のジョブのCPU時間は千分の13に過ぎず、全体としては効率的な利用が行われていることが窺える。

本センターでは、1997年3月の並列演算サーバの導入および1997年8月のスーパーコンピュータSX-4導入決定を契機に、計算機の性能を最大限に引き出すための高速化推進活動が今後の最重要課題と考え、利用者およびメーカーの協力を得て高速化推進活動プロジェクトを1997年9月に立ち上げた。その目的は、プログラムの高速化による実行時間の短縮と大規模化の促進、計算機の効率的な運用とジョブのターンアラウンドタイムの短縮、コンパイラを中心とした自動ベクトル化技術と自動並列化技術の強化をはかること、および、高速化技術の教育・指導法の確立にある。

1997年から2002年までのスーパーコンピュータSX-4での活動成果を表4に示す。単体性能向上比は1 CPU使用時のオリジナルに対する性能向上倍率、並列性能向上比は8 CPU使用時のオリジナルに対する性能向上倍率であり、ともに高速化を試みたジョブの平均である。2000年以降の性能向上比がそれまでにくらべ低くなっているが、これは利用者自身の手による高速化がすでに行われていたり、コンパイラの強化等によりさらなる高速化の余地が少なかったためであり、本活動が有効に機能していることの顕れであると考えている。先の表3は本活動の成果を含むものであり、本活動を行わなかつ

た場合は異なった結果になっていたであろうことは容易に推測できる。なお、各年とも高速化に取り組んだ件数が10件以下であるのは、改善すべき箇所が局所的でなく、プログラム全体にわたることが多く人手と時間を要するためである。

表3 並列ジョブのCPU時間分布（千分率）

	90 ≤	7	8	10	11	17	35	44	84	137	552	
ベ ク ト ル 演 算 率 %	80 ≤		2		1	1	3	2	6	4	21	
	70 ≤			2	1		1	4	2	1	9	
	60 ≤							1		3	8	
	50 ≤										7	
	40 ≤										2	
	30 ≤							1			1	
	20 ≤										1	
	10 ≤										1	
	0 ≤	1										2
	%	0 ≤	10 ≤	20 ≤	30 ≤	40 ≤	50 ≤	60 ≤	70 ≤	80 ≤	90 ≤	

並列化率%

5. おわりに

計算機上で模擬実験を行うシミュレーションは、理論、実験とともに科学を支える3本柱の一角を成しており、本センターの大規模科学計算システムは先端的な研究分野において、不可欠の存在となっている。このことは、利用状況にも如実に現れており、他では実行できないような規模の演算性能と

メモリを備えたシステムの重要性は一層増している。また、それとともに計算機の性能を最大限に引き出すための高速化推進活動の必要性もより高まっている。今後とも、学術研究を支える重要かつ強力な基盤である大規模科学計算システムの整備・運用に努めていきたいと考えている。

表4 活動成果（スーパーコンピュータ SX-4分）

年	1997	1998	1999	2000	2001	2002
件数	2	8	8	9	10	7
単体性能向上比	1.9	46.7	4.5	2.5	1.6	2.2
並列性能向上比	11.1	18.4	31.7	8.6	4.9	2.8

性能向上比はオリジナルを1としたときの性能倍率