

誰でも使えるPC普及の一方で ハイパフォーマンス

国家間競争になって ハイパフォーマンス



はじめに

地球シミュレータが稼働を開始したのが、平成14年3月1日であるから、2年半が経過した。これまでに、地球シミュレータの見学者は、1万人を超え、大勢の見学者の声をまとめると、「感激した」とか、「日本の技術に自信を持った」、中には「生きてて良かった」などというものであった。

特に「日本の技術に自信を持った」といわれるのは、おおむね60歳台初頭でメーカー系企業を退職された方が多い。この年代の方々は、過去の日本の経済の絶頂期も、このところの低迷期も経験されており、見学の後、地球シミュレータのあるJR新杉田駅の近くで反省会を開き、地球シミュレータを肴に、大いにボルテージが上がり、溜飲を下げたと聞いている。

5年間にわたり、コンピュータ本体系に400億円、建物に70億円をかけた、このビッグプロジェクトに、開発に際しては企業側、そして今は運用側の地球シミュレータセンターに在職しているものとして、地球シミュレータの開発から今日に至るまでを振り返ってみたい。

1.地球シミュレータの概要

1.1 開発経緯

地球シミュレータ開発に到るプロセスの背景には、国際的に地球環境変動予測推進の機運の高まりがあった。有体にいえば、地球温暖化問題である。日本は、国際的な貢献を行うために、1994年に気候変動枠組条約への参加、そして、1995年には、学術審議会で「地球環境科学の推進について」という答申、さらに翌1996年には、航空・電子等技術審議会において「地球変動予測の実現にむけて」答申がなされ、このための体制（モデル研究、観測研究、シミュレータ）の整備を行うこととなった。

なぜ地球シミュレータのような超大型のスーパーコンピュータが必要になったのかは次節に譲るが、世界的にみて、いつの時代も気象・気候関係がスーパーコンピュータのメインプレーヤーあるいはパワーユーザーであったというのは紛れのない事実であった。

他方、1995年頃というのは、日本のスーパーコ

着実に進展している コンピューティングの世界

いる

シミュレーションの最前線

地球シミュレータの筐体を搬送するために臨時に組み立てられたリフトと全景

独立行政法人 海洋研究開発機構

地球シミュレータセンター センター長補佐 平野 哲

ンピュータ状況はどうなっていたかという、当時の航空宇宙技術研究所（現JAXA）のNWT（Numerical Window Tunnel）が1993年から稼働したがTOP500リストで3期連続トップをキープしていた。

しかし、米国のASCI計画が発動されると、その後、瞬間風速的に日本のコンピュータが顔を出すことはあったものの、トップは米国勢に奪われたままとなった。ASCI計画は、核爆発のシミュレーションを第一義的に行うもので、米国にとっては、保存核兵器の信頼性検証や新兵器開発のために地下核実験や臨界前核実験と進めてきた核開発計画にとっては必須のものと位置付けられている。

NWTの開発は、大規模風洞実験を行うのに風洞を構築せずに、コンピュータシミュレーションの中で実現しようというものである。シミュレーションは、Navier-Stokes方程式と呼ばれる流体の方程式を数値的に解くことにより、行われる。

このプロジェクトのリーダーは、後に地球シミュレータ研究開発センターのセンター長となられる三好甫さんであった。三好さんは、NWTの次のシミュレータが日本の科学技術の発展上、欠かせないものであるとの主張を展開しており、科学技術庁の幹部に説いて回っておられたのである。ち

なみに三好さんは、航空宇宙技術研究所にて日本のスーパーコンピュータの黎明期から研究開発を進められてきた方であり、富士通のスーパーコンピュータの生みの親とも言われる方である。

当時の科学技術庁のトップの理解もあり、以上述べた2つのニーズ「地球環境問題に取り組む必要性」「米国に負けないスーパーコンピュータ」から「地球シミュレータ」の発想が生まれ、1997年度から予算化が始まった。図1では、米国の現在開発中の地球シミュレータ対抗機も表示してある。

1.2 アーキテクチャおよび要求仕様

1997年に設定した開発目標（ターゲット）は、当時のCRAYのベクトル型スーパーコンピュータであるC90の1,000倍性能というものであった。当時は、全地球上の大気大循環をシミュレーションしようとしても100Kmのメッシュでしかできなかった。これでは、台風の存在や、より細かい現象をコンピュータ上で再現することはできず、10Kmのメッシュが望まれた最大の理由であった。

実際には、水平方向の解像度が各10倍、垂直方向は5倍程度、水平方向の解像度を上げたことにより、計算を安定させ、時間刻みを短くし2倍程

われる演算，アルゴリズムが中心であり，多くの場合，LINPACKの性能も実アプリケーションの性能とかけ離れているからである。

また，プロセッサの台数を増やすとアムダールの法則により，効率はダウンしていくので実効効率は12.5%とし，このためのピーク性能としては32Tflopsを必要とし，ベクトル型スーパーコンピュータで実現するとした。メモリ容量も同様な計算を実行するには数千倍のものが必要だとされ，4TB（テラバイト）以上とした。

1.3 開発スケジュール

開発スケジュールは，地球シミュレータが完成した後でも鮮度を保てるよう，慎重かつ開発リスクを考慮しつつ，年々変わるSIAのロードマップ，メーカーからのヒヤリングを加味して決定された。地球シミュレータ向けのチップの製作には，最新鋭の微細加工のできる製造ラインを必要としたためである。

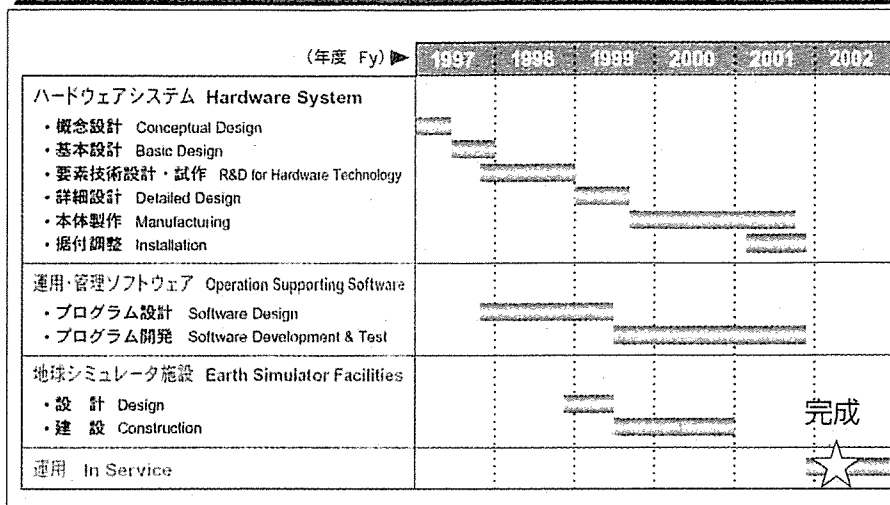
目標として，地球シミュレータの稼働後，ピーク性能で2年，LINPACK性能で3年，アプリケーションの実行で5年はもたせたいとされた。

平成9年4月，地球シミュレータ研究開発センターが発足した。場所は浜松町である。中心メンバーは，三好さんがセンター長となり，当時の動力炉核燃料事業団，宇宙開発事業団からの出向者で構成された。その後，動力炉核燃料事業団は，日本原子力研究所に引き継がれ，さらに建物の立地に伴い，海洋科学技術センターからの出向者も加わった。

早速，概念設計の提案が公示され，富士通と日本電気がこれに応じ，前述のターゲットをベース

地球シミュレータの開発工程

開発スケジュール Development Schedule



地球シミュレータ研究開発センター (ESRDC : NASDA, JAERI, JAMSTEC) が研究開発担当

に設計が開始された。そしてこの概念設計の締め括りとして，メーカーへ性能と費用，最終納期の提出を求め，最終的にコストパフォーマンスに優れた日本電気の提案が採用されたのである。

その後，日本電気との共同で基本設計が進み，開発目標として，性能は40Tflops，これを構成するノード数は640ノード，1ノードあたり8個のベクトルプロセッサと16GBのメモリ，1プロセッサの性能は8 Gflops，ノード間の最大データ転送性能も12.3GB/秒×2と決まった。

この基本設計以降は，要素技術，試作，詳細設計，本体製造など費用も増大していくので，フェーズを進める前，平成10年8月に，中間評価委員会を開催し，有識者の判断を仰いで了解を得た。

一方，ソフトウェアに関しては，超巨大なシステムでもあり，NECの標準インターフェイスの上にセンターが運用ソフトウェアを構築することとなり，センター側の開発体制も整えた。

地球シミュレータの開発予算の獲得は，事務当局の獅子奮迅の動きと，折からの度重なる補正予算と科学技術基本法の施行が追い風となり，順調に推移した。まさに，地球シミュレータにとってこの時期にスタートできたのは大変ラッキーなことであったと思われる。

本体製造においては，予算は，原子力研究所，宇宙開発事業団，海洋科学技術センターで案分さ

れた。したがって資産番号もそれぞれの法人のものが筐体ごとに貼り付けられた。なお、平成16年4月からの海洋研究開発機構の発足に伴い、すべての資産は海洋研究開発機構に一本化されている。

これに先立つこと平成13年2月、三好センター長を交えた6名で、NEC山形の鶴岡工場視察が実施された。この視察のタイミングは、地球シミュレータ心臓部のベクトルプロセッサの出来具合のヒヤリングが目的であった。

チップは、一辺21mm角の正方形に6,000万個のトランジスタを詰め込み、しかも0.15ミクロン幅の技術を使った最新のものであり、チップの修正には2ヶ月の時間を必要とするものである。この年の6月には、3月に完成するセンターの建物にプロセッサノードの搬入を控えており、チップの作り直しは、あと1回、頑張って2回というところまで来ており、まさにギリギリのタイミングであった。

これがうまくいかないと、すべてが後にずれかねない。予定された納入期限は平成14年2月末であり、これが遅れることはメーカーにとっては莫大なペナルティが発生するし、センター長の責任問題にまで発展しかねない恐れもあった。

この時のNEC側の説明は、「バグはすべて判明しており、次回のチップでは修正できるので、安

心してください」というものであった。三好センター長は、「来年の5月の連休には、五能線の旅行をぜひ楽しみたいものだ」と視察の後、話したが、帰京すると持病が悪化し、入退院を繰り返すこととなり、以後、病院から指揮を執られた。

それでも体調が悪いにもかかわらず、10月29日には、センターにおいてになった。しかし、センター長室に入られるまでが体力の限界であり、結局、このときすでに納入されていた数十台のプロセッサノードをご覧になることもなく、進捗についてのグループリーダーからの報告を聞くに止まった。三好センター長は、その年の11月17日に亡くなった。

地球シミュレータに命を賭けたと言わざるを得ない壮絶な死であったと思う。センター長からは生前、「死者は生者を走らすな」という言葉を何度も聞いたが、これもなかなか実行できることではなく、唯唯、頭の下がる思いである。

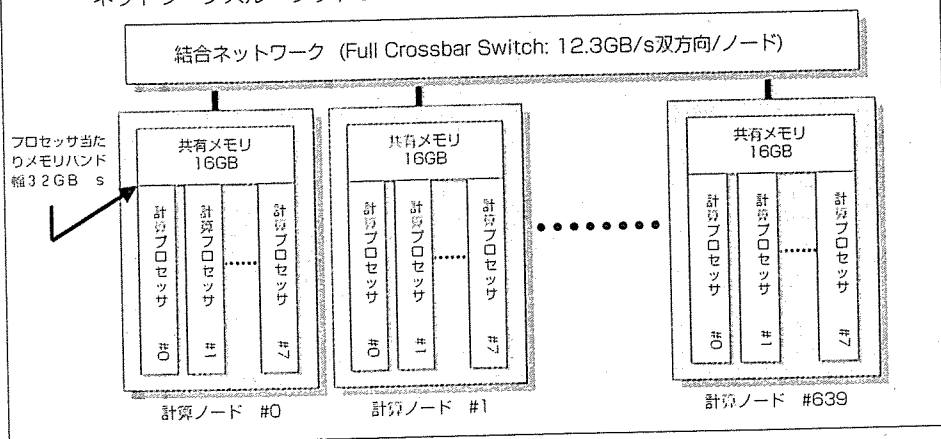
地球シミュレータを本番稼働／運用して2年半、チップそのものには、大きなトラブルもなく安定稼働していることを報告しておきたい。また、故三好センター長の業績を称えて、海洋研究開発機構横浜研究所には、三好記念講堂が設けられていることを付記しておきたい。

マシン搬入が粛々と行われている中で、三好セ

ンター長の後任として、日本原子力研究所理事の浅井さんが地球シミュレータ研究開発センター長として就任された。そして平成14年4月から店開きする海洋科学技術センター地球シミュレータセンターのセンター長には、核融合研究所で理論・シミュレーションセンター長を努めておられた佐藤哲也先生の就任が決まっ

地球シミュレータのシステム構成

- 総計算ノード数: 640
 - 総プロセッサ数: 5120
 - ピーク性能: 40TFLOPS
 - 計算プロセッサのピーク性能: 8GFLOPS
 - 主記憶容量: 10TB
 - 計算ノードのピーク性能: 64GFLOPS
 - 計算ノードの主記憶容量: 16GB
- ネットワークスループット 8 TB/s



たのもこの頃である。

残りの約3ヶ月で地球シミュレータの運用組織、利用規定など決めなければならないことが山積みであり、この間の事務に関わった方々のご苦労は大変なものであったと思う。

1.4 達成性能

地球シミュレータの搬入は、平成13年6月から始まった。この搬入のためにセンターの建物横に臨時にリフトを建てて、このリフトにトラックを横付けして、工場から届くPN筐体、IN筐体をトラックの床面を合わせてリフトにスライドし、マシンルーム入り口（研究棟の4階の高さ）まで吊り上げる方法をとった。

マシンルーム入り口の床面は、マシンルーム内フリーアクセスフロアの面と一致しており、後は、所定の場所まで押していくのである。平成13年12月末までにすべての筐体の搬入が完了した。最初は、16ノードを束ねてSクラスと称して、INの調整が済むまでgEther接続により、OS、ジョブスケジューラ、アプリケーションのデバッグが行われた。

アプリケーションとしては、目標となった大気大循環モデルで5Tflopsを実現するためのプログラムである「TF5」、米国GFDLの海洋大循環モデルMOM3を地球シミュレータ向けに最適化した「MOM」の2本のプログラムである。

前者は、日本のCCSR（東京大学気候研究センター）/NIES（国立環境研究所）により開発されたプログラムであるが、これを地球シミュレータ向けに最適化したものである。この時期は、メーカー技術者も常に100名近くは来ていて、連日の搬入と進捗管理で大わらわとなっていた。

12月中旬にMOMが7ノード使って、初めて全球10Kmメッシュの海洋大循環の処理結果を見せてくれた。これを、海洋科学に関わる研究者に見せたところ、啞然として見入っていたことを思い出す。従来の全球モデルでは現れなかった細かい中規模渦と呼ばれる渦が随所にあられ、しかも海流までが明確に姿を現したからである。

この処理結果の表示には、GRADSというグラフィックパッケージが用いられたが、地球シミュ

レータの出力するデータ量が余りに大きく、扱えるデータのサイズである2GBを超えるため、ソースにかなり手を入れざるを得ず、担当者は想定しない苦勞で泣かされた。

この2GB超え問題は、その後、いろいろな商用パッケージが抱えている問題として入力データ処理、出力データ処理に影を落としていくことになる。

さらに平成14年1月に入ると、INを使った並列処理環境も整い、センターが開発してきたジョブスケジューラもなんとか動き始め、ワークディスクとテープライブラリとの間でマイグレーション/リコールも可能になってきた。まだ、PNやINは初期トラブルが多く、ジョブの実行も、とてもスムーズに流れるとは言い難い状況ではあったが、1月下旬には、160ノードを使って、TF5で7.6Tflopsの性能を確認した。

つまり、全体の約1/4の構成でも、ターゲットをクリアしてしまったのである。この段階で、まずは一安心であったが、640ノード全部が動くには、ほど遠い状況には変わりなく、以後、システムの安定化が最大テーマとされて、毎日の進捗確認が行われた。

連日の会議に出されるチェックリストは、瞬く間に厚さを増し、こんなに出るトラブルを収束できるのかと、大いに気を揉んだものである。圧倒的にプロセッサノードのトラブルが多く、次いでINのインターフェイスのトラブルが多かった。それでも、2月28日の検収までには、640全ノードの同時稼働が確認できた。

この時には、TF5の性能は320ノードで16Tflopsにまで達していた。これは、ターゲットの実に3倍の性能が出たことを意味している。そして地球シミュレータは、2月28日に検収に合格し、地球シミュレータ研究開発センターに引き渡され、翌日から本番稼働/運用運転に入った。

平成14年3月からは、システム安定性の評価を兼ねて、日本原子力研究所、宇宙開発事業団、海洋科学技術センターの研究者による利用が始まった。また、3月15日には、地球シミュレータが稼働した旨のプレスリリースを行った。

3月末で、まる5年間続いた地球シミュレータ研究開発センターは、役目を完遂し解散。4月か

らは海洋科学技術センター(現海洋研究開発機構)内に発足した地球シミュレータセンターが、佐藤哲也センター長のもとでスタートした。

まず、世界最高速を証明するためのLINPACKベンチマークに挑戦することとなった。

4月14日に3回のテストを行い2回は失敗し、1回は638ノード35.61Tflopsの計測に成功した。このベンチマークをパスするためには、すべてのノードが9時間近く安定して動かねばならないので、この時点では難しいと思われたが、6月のISC2002(ハイデルベルグ)で認定してもらうには、この時点しかなかったので敢えて挑戦したのであった。

とりあえず、この性能値をテネシー大学のドンガラ先生に送付した。この性能値は、第2位のASCI White Pacificの実に5倍の性能であり、ダントツの1位で6年ぶりに日本がスーパーコンピュータの性能で世界に返り咲いた瞬間であった。

さらに5月3日には、640フルノードでの計測に成功し、性能値35.86Tflopsを送付した。この結果、ISC2002では、この値が世界最高速機として認定されたのである。スーパーコンピュータの専門家は、その性能の高さはもとより、それまでピーク性能の60~70%台であった実効性能比が、地球シミュレータの場合には88%に達したことに驚き、ベクトル機は過去のモノというイメージが払拭されたのであった。

ドンガラ先生の談話は、4月20日のニューヨークタイムズのトップ記事となり、当時のTOP500を占める米国の上位20台までのコンピュータの処理能力に匹敵するとし、スプートニックショックならぬコンピュータニックショックだという新語を生み出した。

地球シミュレータの稼働直後、3月の第1週には、米国大使館の科学アタッシュが数名で現れたのを皮切りにLINPACKの性能が公表されるや、続々、見学が相次いだ。

特に、この性能に驚いた米国の対応はまことに素早いものがあった。5月中旬には、全米のスーパーコンピュータの著名な研究者、政策担当者36名を集めて、地球シミュレータ対抗策が話し合われているのである。そして、圧倒的な地球シミュレータの性能差が今後、米国の科学技術にどのよ

うな影響を与えるのか分析し、その早期対抗策を政府に勧告しているのである。

これが引き金となって15年6月の米国下院における公聴会、11月中旬のエネルギー省エブラハム長官の打倒・地球シミュレータ宣言につながっていくのである。

2006年までに対抗機開発のために投じられる費用は600億円以上と見られる。

ベクトル機で世界を席捲したにも拘らず、ベクトル機からスカラーパラレル機への転身を模索し、低迷していたCRAY社は、地球シミュレータの成功をみて急遽ベクトル機X1に力を入れて、矢継ぎ早に米国内のスーパーコンピュータの受注に成功し、経営状況も改善した。

CRAY社のプロモーションビデオに地球シミュレータの画像が、戦車の次に登場するのには大変驚いたものである。

TF5は、その後改良を重ねプログラム名もAFES(AGCM for Earth Simulator)と変えて、640フルノードでの計算に挑戦し、26.58Tflops、ピーク性能比65%を達成した。これは、ターゲットの5倍以上で2002年11月のSC2002において、他の2件のプログラムとともに、HPCの最高の栄誉であるゴードンベル賞の最高性能賞に輝いたのである。

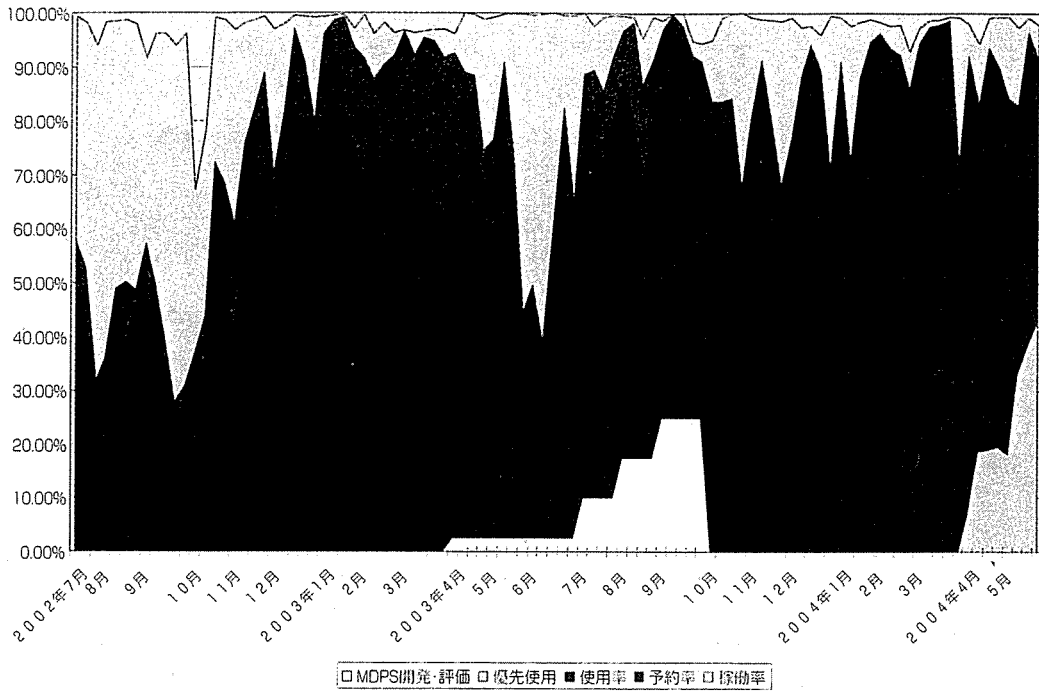
1.5 運用状況

平成13年12月、地球シミュレータの運営に際し産業界、学界、研究機関などの代表者からなる地球シミュレータ運営委員会が開催され、運営基本計画が承認された。

画期的なことは、地球シミュレータが大気・海洋、固体地球の分野だけではなく、計算科学や先進創造分野にも利用の道を開いたことであろう。そして平成14年4月にはいと、利用計画委員会、課題選定委員会が設置され、地球シミュレータを利用する研究分野と割り当てられるリソースの比率、公募方法などが決定され、短期間のうちに課題の公募が行われ、課題選定後、7月15日から正式利用が開始された。

現在の分野別のリソース割り当ては、大気・海洋分野35%、固体地球分野20%、計算機科学分野10%、先進分野15%残りは海外との共同研究や民

2002年7月~2004年5月
使用状況



グラフ

間等との共同研究の試験的利用に当てられている。

グラフに今日までの稼働率、予約率、利用率を示す。ノード(624ノード)×時間(1週間分)を100%として記録されたものである。16ノードは小規模バッチ、会話型に使っているのでこの統計には含まれていない。当初は、利用者のプログラム開発、最適化がなかなか進まず、利用率も低かったが、早くから経験してきたセンター研究員や技術員がプログラム相談を行うことで次第に利用率が向上してきた。

そして12月には、利用率が80%を超えてしまうところまで高まったのである。

2002年10月の稼働率の落ち込みは、INのトラブルで集中的に部品の交換を行ったためである。したがって利用率も低くなっている。また、2003年6月の利用率の落ち込みは、2002年度後半は、相当混みあってきたので、分野別のリソース制限をかけていたものを2003年はプロジェクト別に厳しくカウントすると通知したところ、大事なリソースは後にとっておこうということで「使い控え」が起こってしまったためである。

地球シミュレータは、6000KW/hの電力を消費

する。マシンは使わなくても電源はいれておかねばならないので、こんなことは本意でないので、上期は、予定の倍使っても良いとボーナスを付けたところ、速やかに利用率は高くなった。利用者は敏感に反応するものである。

2003年4月から9月末まで、それまでのワークディスク(容量460TB)とテープライブラリ(容量1.5PB)間のマイグレーション/リコールに代わり、ユーザーディスク(240TB)とテープライブラリを階層型ファイルシステムとして持つMDPS(マスタデータプロセッシングシステム)の導入を図るため、段階的に移行のためのノード占有が行われた。

これによりシステム全体のTATの向上とスループットの向上が図られた。MDPSについては、次号で説明予定。また、2004年4月から9月までは、IPCC(Inter-governmental Panel for Climate Change)向け報告のために、最大、全体の半分近くのノードが3グループで占有利用している。

この報告が国際的に高く評価されれば、地球シミュレータ開発の際の大きな目的が2つとも達成されたことになる。(つづく)(Hiroshi Hirano)