

COMPUTERWORLD

月刊「コンピュータワールド」

Get **Technology** Right

米国のIT関連職給与調査に見る
ITプロフェッショナルの
「仕事の値打ち」

2004
October
1,280yen
10

Technology Focus

IPセントレックス・サービスの
現状を知る

Emerging Technology

次世代ビデオ会議システム

IT業界の定説「嘘か真実か？」

ヒューマノイドの存在意義



エンタープライズ

機は熟したか?

Linux



HPC

ハイパフォーマンス・コンピューティング

第10回

Grid RPCと
データ・グリッド

遠隔地に点在する計算資源の有効活用

最新事情

今回は、「Grid RPC」と「データ・グリッド」について紹介する。

Grid RPCは、手元のPCから、遠隔地にある計算資源にジョブをリクエストする手順を定めたもので、グリッド・コンピューティングにおいて重要な役割を果たす。

一方、データ・グリッドは、データの複製を広域に複数配置することで、データ・アクセス時の負荷を分散させ、膨大な量のデータを高速に扱えるようにするサービスである。

廣安友之

同志社大学 工学部 知識工学科 助教授

建部修見

産業技術総合研究所 グリッド研究センター

計算資源にジョブを送る
Grid RPC

グリッド・コンピューティングを行ううえで重要な役割を果たすものの1つに「Grid RPC」がある。RPCとはRemote Procedure Callの略で、ネットワークに接続された他の計算機にジョブをリクエストする手順を指す。このRPCは、UNIXの世界では広く普及しており、Windowsにも実装されている。RPCは通常、ジョブを依頼するクライアントとジョブを処理するサーバから構成されるクライアント/サーバ型をとっている。このRPCをグリッド・コンピューティングに拡張したのがGrid RPCである。

Grid RPCは、次のようなシーンで役立つ。あるユーザーが計算能力の低いノートPCを使って作業をしていたとする。ところがある日、そのノートPCでは時間がかかりすぎて到底実行不可能なほど大量の計算処理をしなければならなくなるとしよう。このような場合に、Grid RPCを用いれば、他の計算資源に必要なデータを送信して処理を実行し、その結果を受け取るといったことが可能になる。

また、計算機のようなハードウェアだけでなく、ソフトウェアもGrid RPCの対象とすることができる。例えば、ある高価なアプリケーションが必要となったときに、そのアプリケーションを所有していなくても、そのアプリケーションが稼働するサーバにGrid RPCを利用してデータを送信して処理を実行し、その結果だけを受け取るといった使い方も可能である。

各国で開発されている
Grid RPC

現在、さまざまな団体がGrid RPCの研究開発を行っている。主なものとしては、産業技術総合研究所の「Ninf-G」、米国テネシー大学の「NetSolve」、豪州モナッシュ大学の「Nimrod」、筑波大学の「OmniRPC」などがある。

各団体は、使いやすいAPIの提供、計算資源を効率よく並列稼働させること、セキュリティ、リソース管理、スケジューリング管理など、グリッド・サービスとして提供しなければならない機能の実装を目指して開発を進めている。

ただし、それぞれが独自にGrid RPCの開発を行っているため、APIの違いという問題も生じている。そのため、グリッドの標準化を行っているGGF(注1)内のGrid RPCワーキング・グループによってGrid RPCのAPIに関する提案が出されており、現在ではそれに沿った実装が進められている。

Grid RPCを利用した基本的なプログラムは、図1のようになる。この図では、「mmul」という関数を利用している。ご覧いただければわかるように、まず初期化を行い、利用したいmmul関数を指定する。次に、mmul関数に必要な引数を指定し、Grid RPCによる呼び出しを行って結果を得る。最後に終了処理を行うという流れである。

注1：GGF(Global Grid Forum)とは、グリッド技術の国際標準を定めるための国際会議で、年3回開催されている

産業技術総合研究所の Ninf-G

次に、現在最も注目されているGrid RPCであるNinf-GとNetSolveの特徴を紹介しよう。

Ninf-Gの開発は、1994年に産業技術総合研究所で開始された。Ninf-Gの特徴は、利用者の認証・認可といったセキュリティ機能、計算資源の情報を検索する機能などを実装する際に、グリッドの基盤ソフトウェアであるGlobus Toolkitを利用していることだ。そのほか、通信時のオーバーヘッドを非常に小さく抑えることに成功していることも特筆すべき点である。Ninf-Gはすでに、500台規模のPCを用いた気象予測アプリケーションの稼働実験などで、成果を挙げている。

米国テネシー大学の NetSolve

続いて、米国テネシー大学で開発されているNetSolveについて紹介しよう。前述のNinf-Gがシステム側からの視点で、ハードウェアの稼働効率を重視して開発されているのに対し、NetSolveはユーザーの利便性を重視していることが特徴である。

NetSolveは図2に示すように、RPCの基本構成要素であるクライアントとサーバ以外に、エージェントを設けている。エージェントは、サーバ情報の管理、簡単なスケジューリング、およびジョブ・マッチングを行う役割を担う。

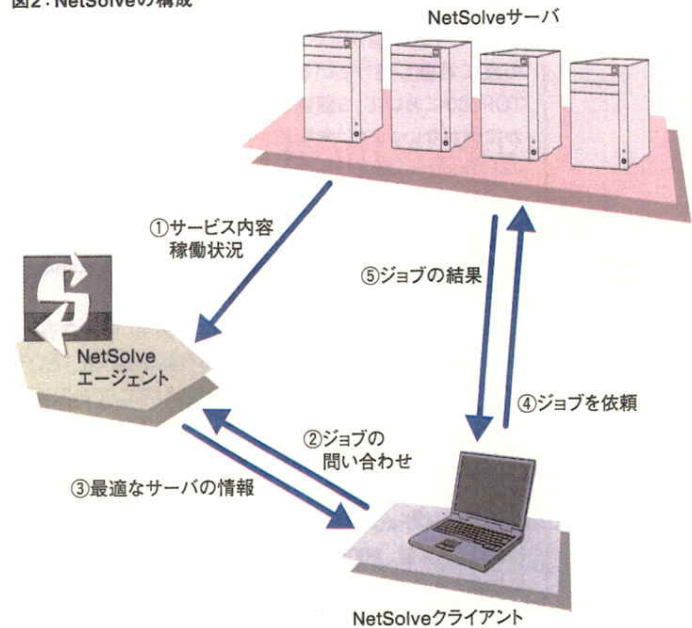
サーバにジョブの実行を依頼するクライアントは、まずエージェントに問い合わせを行う。エージェントは問い合わせのあったジョブに最適なサーバを選び、その所在をクライアントに通知する。クライアントはその情報を基にサーバと通信を行い、サービスを受ける。このような仕組みを実現するために、サーバは対応しているサービスの内容を記述したファイルを用意し、サーバの起動時にそのファイルをエージェントに送信する。

エージェントを設置するメリットは、サービスを提供するサーバの所在をユーザーが知る必要がないという点である。これは些細なことに思われるかもしれないが、実際には非常に重要な仕組みである。な

図1: Grid RPCを使ったプログラムの例



図2: NetSolveの構成



ぜなら、実際にシステムを運用する場合は、サーバを用意する管理者とジョブを依頼するユーザーが異なるからである。

例えば、次のような場合を考えてみよう。あるアプリケーションをあるサーバで稼働させていたとする。そのサーバのメンテナンス時などは、一時的に別のサーバでそのアプリケーションを稼働させなければならない。そのような場合にサーバの管理者は、エージェントに新しいサーバの情報を登録するだけで、各ユーザーに変更を通知する必要はない。というのも、

ユーザーは常に、アプリケーションが稼働しているサーバの所在をエージェントに問い合わせるからだ。

そのほかのNetSolveの特徴的な点としては、PC

上で稼働する「Mathmatica」や「MatLab」といった科学計算用のアプリケーション上から、NetSolveサーバを利用可能にしていることが挙げられる。また、

HPC TOPICS 1

最新TOP500に見る スパコン最新トレンド

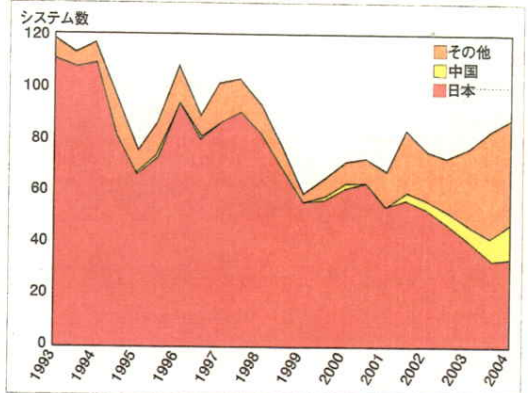
廣安知之

世界中から寄せられたLinpack Benchmarkの結果を基に、年2回のペースでスパコンの上位500を発表している「TOP500 SUPERCOMPUTER SITES」(<http://www.top500.org/>)が今年6月、ランキングの更新を行った(表A)。ここでは、同ランキングから見えてくる、HPCの最新トレンドを紹介したい。

まず、特筆すべき点は、35.86TFLOPSという記録で2002年6月に1位に躍り出た地球シミュレータが5回連続で1位の座に君臨し続けていることだ。20年以上の歴史を持つTOP500において、5回連続で1位を保持したスパコンはほかに例がない。また、米国ローレンス・リバモア国立研究所の「Thunder」が新たに2位にランクインしたこともニュースである。このシステムは、Itanium 2プロセッサを4,096個搭載し、19.94TFLOPSの性能を記録している。とはいえ、この値は、地球シミュレータが記録したLinpack Benchmark値の約半分にすぎない。

ところで、この地球シミュレータを近い将来追い抜く存在と予想されているのが、IBMが開発している「BlueGene/L」という、PowerPCベースのプロセッサを搭載したスパコンである。同システムは小さな筐体に非常に多くのCPUを搭載し、他のスパコンに比べて動作電力が極めて小さいのが特徴である。このBlueGene/Lのプロトタイプが、今回初めてTOP500に登場し、4位(11.68TFLOPS)と8位(8.65TFLOPS)に入っている。BlueGene/Lは、2005年にはローレンス・リバモア国立研究所に納品される予定だ。同システムが完成すると地球シミュレータをLinpack Benchmark値で抜き去るだけでなく、地球シミュレータの15分の1の消費電力、システム全体のサイズも10分

図A：TOP500にランクインしたアジア各国のシステム数の推移



の1以下というコンパクトなものとなると予想されている。

次に注目したい点は、クラスタへの移行が進んだことである。前回のランキングでは、クラスタは208システムであったが、今回のランキングでは291システムに増加し、初めてTOP500の半数を超えた。日本勢からも、理化学研究所に納品された富士通製の「RIKEN Super Combined Cluster」が7位に、産業技術総合研究所の3台の新しいクラスタもそれぞれTOP500にランクインしている。

最後はアジアの動向である。図Aのグラフからもわかるように、TOP500にランクインしている日本製スパコンの数は、減少傾向にある。それに対して、中国勢はその数を増加させている。

中国製のスパコンとして初めて10位に入った「曙光4000A」(本誌16ページ参照)を筆頭に、TOP500に今回ランクインした中国製のスパコンは15システムに上っている。この結果、中国は韓国の9システムを追い抜き、日本の34システムに次ぐアジア2位の地位を確保した。この中国の勢いは止まらず、TOP500にランクインするシステム数で日本が抜かれてしまうのは時間の問題かもしれない。

表A：「TOP500 SUPERCOMPUTER SITES」の最新トップ10

順位	所有者(国名)	システム名	プロセッサ数	実効性能/ピーク性能(TFLOPS)
1	地球シミュレータセンター(日本)	地球シミュレータ	5,120	35.86/40.96
2	ローレンス・リバモア国立研究所(米国)	Thunder	4,096	19.94/22.93
3	ロス・アラモス国立研究所(米国)	ASCI Q	8,192	13.88/20.48
4	IBM(米国)	BlueGene/L DD1 プロトタイプ	8,192	11.68/16.38
5	米国国立スーパーコンピュータ応用研究所(米国)	Tungsten	2,500	9.81/15.3
6	ヨーロッパ中期気象予報センター(英国)	eServer pSeries 690	2,112	8.95/16.05
7	理化学研究所(日本)	RIKEN Super Combined Cluster	2,048	8.72/12.53
8	IBM(米国)	BlueGene/L DD2 プロトタイプ	4,096	8.65/11.46
9	パシフィック・ノースウエスト国立研究所(米国)	Mpp2	1,936	8.63/11.6
10	上海スーパーコンピュータ・センター(中国)	曙光4000A	2,560	8.06/11.26

*赤字で記しているのが今回がランキング初登場のシステム

Excel上にある行列を直接NetSolveサーバに送信できるようにする仕組みも現在開発中である。

今回取り上げた上記2つのGrid RPC以外にも、すぐれた機能を備えたGrid RPCが存在する。今後、各Grid RPCがGGFで定められたAPIに準拠し、共通のAPIで扱えるようになれば、よりいっそう計算機資源の有効活用が進むだろう。

自然科学分野で必要とされたデータ・グリッド

今回のもう1つのテーマである「データ・グリッド」という言葉が初めて公の場で使われたのは、2000年3月に米国サンディエゴで開催された第3回グリッド・フォーラム(注2)でのことであった。スイスのジュネーブ郊外にある欧州合同素粒子原子核研究機構(CERN)で行われていた、LHC実験(注3)に関する発表の中で、データ・グリッドは提案されたのである。

LHC実験では、毎年約10PB(ペタバイト:1PBは10の15乗バイト)の測定データが生成され、数十カ国、数千人規模の素粒子物理学者がデータの解析を行っていた。この規模になると、もはや1カ所にストレージと計算機を集結させることは現実的ではない。地域ごとに計算機センターを階層化した、データ解析を行う世界規模の仕組みが必要であり、それを支えるITとしてグリッドの技術が切に望まれている、という発表内容であった。

当時、グリッドに関する研究は計算処理を行うことがメインであったが、この提案が契機となり、米国や欧州では、データに関するグリッドの応用であるデータ・グリッドの大型プロジェクトが一気に立ち上がった。

データ・グリッドの利用シーンは、高エネルギー物理学、天文学、重力波検出などの自然科学分野が中心となる。というのも、これらの分野で利用されている実験装置、計算機、可視化装置などは広域に分散していることが多いからだ。そのため、大量のデータを長距離間で高速に転送する技術や、広域に分散された複製データを管理する技術の開発が行われた。さらに、データ解析を行うためのテストベッドの構築や、開発されたソフトウェアを簡単にインストールするためのパッケージ化も進められた。

データ・グリッドのより広い応用を目指して

2002年2月に加国トロントで開催された第4回GGFで、OGSA(Open Grid Service Architecture)が発表された。OGSAは、Webサービス技術を基盤とするサービス指向のグリッド・アーキテクチャである。OGSAの登場により、インタフェースの記述が統一され、グリッド技術の商用利用やビジネス・アプリケーションでのグリッド活用の道が開かれた。

また、同フォーラムでは、英国のe-Science Projectを中心に、複数のデータベースをグリッドにより連係させようとする試みについて議論された。

このようにデータ・グリッドは、自然科学分野からの必要に迫られて研究開発が始まったものであったが、ビジネスでの利用に向けた研究開発も始まり、現在に至っている。

データ・グリッドのソフトウェアと国際標準化への取り組み

データ・グリッド関連のソフトウェアとしては、オープンソースのGlobus Toolkitや次節で紹介するグリッド・データファームなどがある。また、オラクルは、組織内のクラスタ環境でグリッド・コンピューティングを実現するRDBMS「Oracle Database 10g」を発表している。

一方、データ・グリッドの国際標準化に関しては、GGF内のGridFTP WG、OGSA Replication Services WG、Database Access and Integration Services WG、Grid File System WG、Grid High-Performance Networking RGといったワーキング・グループやリサーチ・グループが進められている。

計算資源も有効活用するグリッド・データファーム

ここで、データ・グリッドの1つの方向性として、高

注2: グリッド・フォーラムとは、GGFの前身に当たる国際会議で計5回開催された

注3: LHC(Large Hadron Collider:超大型素粒子加速器)実験とは、物質に質量を与えると考えられている未知の素粒子「ヒッグス粒子」の発見を目的とした実験で、陽子を7兆電子ボルトまで加速し、正面衝突させている

図3: 広域仮想ファイルシステム
クライアントB (独国)

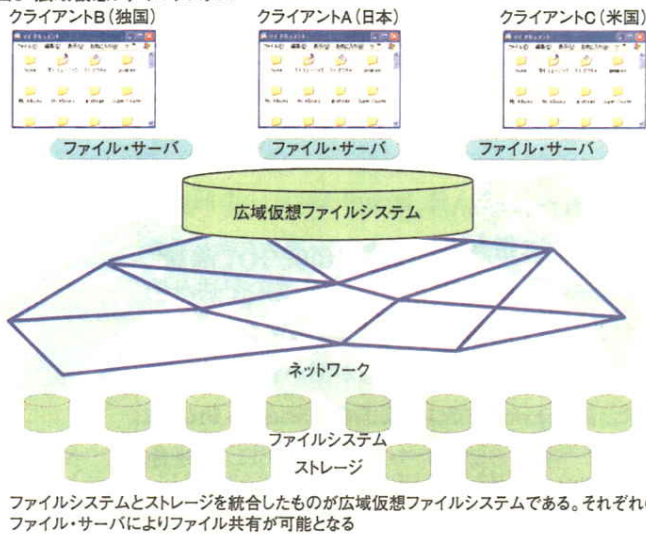
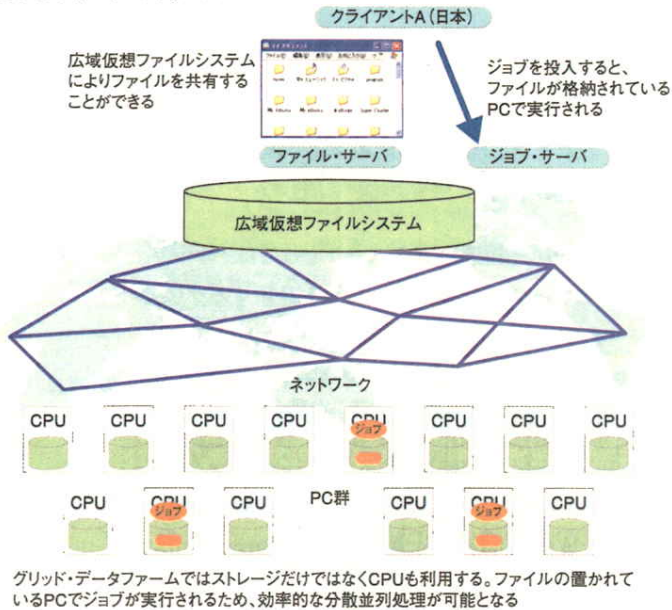


図4: グリッド・データファーム



エネルギー加速機研究機構、東京工業大学、東京大学が共同で研究を行っているグリッド・データファームについて紹介しよう。

グリッド・データファームは、複数の組織から成るグリッド環境において、高い信頼性を持つファイル共有機能だけではなく、同時に高速データ処理も実現することを目的としたデータ・グリッドである。広域仮想ファイルシステムを実現するだけではなく、同時に計算資源を提供する計算グリッドの機能も備えているところが、グリッド・データファームの特徴なのだ。

広域仮想ファイルシステムは、グリッド・ファイルシステムとも呼ばれ、広域に分散するストレージ・システムの統合利用を可能にするファイルシステムである。図3に、同ファイルシステム の概念図を示した。ストレージ・システムに格納されているファイルに対し、その物理的な格納位置とは無関係に、仮想的な階層ディレクトリを用いてアクセス可能としている。同ファイルシステムは、プログラムの変更をすることなくアプリケーションから利用可能である。また、FTPやscp (Secure Copy:SSHを用いたセキュアなファイル転送機能)によるファイル共有、およびsambaによるWindowsからのファイル共有もサポートしている。

広域仮想ファイルシステムとグリッド・データファームの違い

広域仮想ファイルシステムは、必要に応じてファイルを複製する仕組みを備えている。ファイルの複製は利用者には意識されず、ファイル・アクセス時には、目的のファイルの数ある複製の中から、利用可能な状態にあり、最も応答が速く、かつ、最もアクセスされていないファイルが自動的に選択される。これにより、ストレージ・システムとしての耐故障性を高めると同時に、なるべく近いファイル複製を選択することで、ファイル・アクセスの集中を回避している。

同ファイルシステムを利用するクライアントは、ネットワークを通じてファイルを格納するストレージからデータを受け取ってから処理を行うことになる。ただし、この仕組みの場合、大量のデータを処理しようとすると、いくら高速なネットワークを用意しても、データの転送にかかる時間がボトルネックとなり、パフォーマンスの向上に限界があるという問題がある。

一方、グリッド・データファームでは、ストレージを提供するPCが持つCPUも計算処理に利用することで、この問題の解決を図ろうとしている(図4)。つまり遠隔地のストレージに格納されているデータをクライアントに転送するのではなく、データが格納されているPC上でデータを処理しようというわけだ。いわば、ファイルシステムのためのPCクラスタと、計算のためのPCクラスタを効率的に融合したものがグリッド・データファームなのである。

高速なデータ・アクセスが必要なデータ処理は、

バッチ・ジョブの形で投入され、その計算処理は、入力データが格納されているPCに割り当てられる。これと同様のデータ処理を、あるまとまったデータの集合に対して行くと、データが格納されている各PCにより、分散並列処理が行われることになる。

このとき、頻繁に参照されるデータを格納しているPCでは負荷が高まるおそれがあるが、それを回避するために、そのデータの複製を適宜負荷の低い別のPC上に作成し、負荷の分散を図っている。ストレージを提供するPCとしては、一般的なデスクトップPCやPCクラスタのノードを想定しており、ほとんど利用されることのないローカル・ディスクを有効に活用することもできる。

グリッド・データファームの現状と今後の展望

現在、グリッド・データファームを実現するためのグ

リッド基本ソフトウェアとして、「Gfarm version 1.0.3.1」(<http://datafarm.apgrid.org/>)が公開されている。

Gfarm version 1は、大規模データ・コンピューティングを想定して、POSIX標準(言語のインタフェースを規定した仕様)のファイル入出力APIのサブセットを提供している。これにより、高速なデータ処理という要求は満たしているが、その一方でファイル内容の書き換えができないなどの制限を持っており、利用できるアプリケーションが限られている。

そのため、POSIXに完全準拠したGfarm version 2の実装が現在進められている。同ソフトウェアでは、ファイルの更新やロック、ユーザー/グループごとのアクセス権限の管理、不当なファイル操作の禁止といった機能が追加される。これにより、汎用的なアプリケーションからグリッド・データファームを利用できるようになる。よって、ビジネスでのグリッド・データファームの活用も進むだろう。

CW

HPC TOPICS 2

HPC ASIA 2004で浮き彫りとなった日本と中国の“勢いの差”

廣安知之

今年7月20日から22日までの3日間、埼玉県の大宮ソニックシティで「HPC ASIA 2004」(<http://www.hpcc.jp/HPCAsia2004/>)が開催された。HPC ASIAは、アジア各国が持ち回りで主催国となり、約18か月に1度のペースで開催されている、スパコンに関する国際会議で、日本での開催は今回が初めてである。

同会議の主な内容は、基調講演、パネル・ディスカッション、学術講演、ワークショップおよび展示会で、今回は国内外から約300名のHPC関連の研究者、開発者、ベンダーなどが参加した。特に、展示会場では、日本初公開となったIBMの「BlueGene/L」の計算ノードと新しいブレードサーバ、ミロムの新しいスイッチ製品、筑波大学が研究している省電力サーバなど興味深い製品が数多く展示され、大いににぎわった。

著者は、インドで開催された前回(第6回)のHPC ASIAには参加していないが、第4回および第5回には参加した。過去2回のHPC ASIAと今回とを比べて強く感じたことは、HPC分野での中国の躍進と日本の停滞である。

アジアでは、各国の協力体制がいまだ確立されておらず、HPCへの取り組みに関する各国の温度差も感じられる。特に、大型計算機センターの存在意義が問われ始めている

日本と、今後さらにHPCに力を入れていこうとしている中国とでは、スパコンへの取り組みに大きな差があるように思えてならないのだ。

確かに、NECの渡辺貞氏がスピーカーを務めたHPC ASIA 2004の基調講演やその他の学術講演では、依然として日本はすぐれたHPC技術を有しており、他のアジア各国を牽引する立場にあることがアピールされた。しかしながら、今後も日本がHPC分野でリードを保てるかという点、樂觀視していただける雰囲気ではないように思える。技術は有しているものの、それを活用する道筋が見えず、中国にも激しく追い上げられている——。HPC分野に限って言えば、これが今、日本が置かれている状況ではなかろうか。

話は変わるが、米国「フォーチュン」誌(<http://www.fortune.com/fortune/>)が先ごろ発表した、年間売上高の上位500社のリスト(フォーチュン500)を見ると、日本企業で10位以内にランクインしているのはトヨタ自動車(8位)のみである。しかも、ランクインしている日本企業の数も年を追うごとに減ってきている。それに対して、中国企業は年々増え続け、今回は2003年の11社から15社に増加した。

フォーチュン500やTOP500のリスト、HPC ASIA 2004の状況から、日本の抱える問題が見えてきたように思う。日本は高い技術力を有しているが、それを将来に結び付けることがなかなかできず、中国の台頭におびえているという構図である。HPC分野に限らず、産業界全体として、こういった問題を解決していかなければ、日本の未来は明るくないのかもしれない。