

## 研究課題別研究評価

1. 研究課題名: 超広域高性能計算環境の基礎的研究

2. 研究者名: 松岡 聡

3. 研究のねらい:

科学技術一般において、計算機を用い、従来には不可能だと思われていたレベルの計算パワーを湯水のように使ういわゆる E-Science の要求が高まっている。これらは、TeraFlops 級の大規模計算およびデータ解析用の数百 Tera ~ Peta バイト級のストレージのニーズがあり、分野としては高エネルギー物理、原子力、量子コンピュータ設計、量子化学、分子レベル設計、オペレーションリサーチ、など、多岐にわたる。例えば、計算化学により抗がん剤の創薬には数十 TeraFlops の計算能力が必要とされており、高エネルギー物理学において新世代の加速器のデータ解析にはペタバイト級のオンライン(テープではない)ストレージが必要とされる。

一方、従来の大学の計算機センターなどを中心としたスパコン群は、新世代の E-Science のアプリケーションをサポートするには計算およびストレージ両面にわたり能力的に不十分で、かつ圧倒的にコストが高い。特に、ストレージに対しては最上級の計算機センターでもようやくテラバイト級のオンラインストレージが提供される程度であり、これらの E-Science の要求とは数千倍の開きがある。従って、従来型のスパコンを多少増強しても全く意味が無く、新たな先進的技術による学内の E-Science の手法とそのインフラが必要である。

世界的には、莫大な数の計算機を超高速ネットワークで結合することによって、超高性能計算(HPC)を分散的かつ仮想的に実現する超広域高性能計算手法、いわゆる「グリッド(Grid)」が上記の問題解決手法として注目されている。特に米国ではこの 2-3 年間で急速に研究開発が進み、NGI および Internet2 などの超広帯域次世代インターネットと呼応して、急激にハード、ソフト、ストレージ、および人的なインフラの整備が進行していて、各国立研の連携研究の N S F Alliance/NPACI/ TeraGrid/ iVDGL/GridCenter, DOE の DOE2000/DISCOM, NASA の IPG などの Grid の国家プロジェクトに莫大な研究資金がつかぎ込まれ、TeraFlops 級のインフラが構築されている。欧州では、CERN を中心に EU DataGrid プロジェクトにおいて 10Peta バイト級の加速器のデータ処理を行うための Grid の研究開発をはじめ、E-Science 実現のために Grid 計算を最重要研究開発課題の一つと位置付けている。わが国のメーカ・大学・研究所はこの流れからは取り残されてきたが、近年ようやくその重要性が諸方面で認識されつつある。

本研究では、グリッドに関する諸基礎技術、特に PC 技術を主体とするコモディティテクノロジ中心の基礎的なテーマを扱い、それらを次世代の高性能計算としての「コモディティグリッド」によるテラスケール計算基盤の礎とすることを目的とする。具体的なテーマは、グリッド上の有効なプログラミングモデルやプログラム言語の構築のみならず、定量的な実行モデルやシミュレーションモデル、さらには実際の性能予測のための並列シミュレータの開発を狙う。また、コモディティグリッドの資源としてのコモディティクラスタの構築技術を探求し、それを基盤にさらにシミュレーションの並列実行、さらにはグリッドアプリケーション用のテストベッド開発を目指す。このために、計算資源たるクラスタの計算資源・負荷・ネットワーク接続速度・計算ライブラリの利用可能性などの情報収集および予測、計算の適切な負荷分散アルゴリズムおよびシステム、計算環境の仮想化技術、さらには広域のグリッドのテストベッド構築および管理技術などが必要となる。これらの基礎的技術の構築により、21 世紀の新たな

広域計算インフラが構築可能となる。

#### 4. 研究結果及び自己評価:

##### 研究結果

本研究では、グリッド(グローバルコンピューティング)の定量性のある基礎的モデルを確立し、その有効性を検討することを目指し、GridRPC と呼ぶグリッド上のタスクパラレルのプログラミングモデルの確立、および GridRPC の定量的な性能モデルの提案並びにシミュレーションを行った。さらに、グリッドの計算資源としての大規模 PC クラスタの構築を行い、数理最適化やゲノム関係のプログラムのグリッド上での稼働を行った。特に近年ではグリッドの基本的な計算資源として PC クラスタ計算機の研究も推進しているが、本研究では Presto Cluster I, II, III と三台の大規模クラスタを構築し、600 プロセッサ以上、合算ピーク性能 800GigaFlops を達成した。これは、東工大の計算機センター(GSIC)の二倍にもあたる計算パワーであり、また、Presto IIIは実性能でも当初77.4GigaFlopsを達成して、世界最速の計算機の公式リストである Top500 の 2001 年 6 月版に 439 位にランキングされ、さらに 2001 年 10 月においては改良により 331.7Gigaflops の性能に達した。これは最新の 2001 年 11 月版の Top500 では現状で 86 位で、PC クラスタとしては世界で 4 位に入るものである。クラスタ上でグリッドのソフトウェアの基本構築技術の研究も行い、特に Presto Cluster を用いて定量的モデルに基づいたグリッドの大規模なシミュレーションを行った。また、実際の性能計測が可能となるテストベッドおよびツール群 RealGrid の研究開発を行い、実環境の定量的性能評価、およびシミュレーションとの比較を行った。また、東工大内部およびUCSDと協力して Grid 上での並列アプリケーションを Presto Cluster 上で実際に稼働させ、実問題で有効な結果を得た。さらに、高エネルギー物理学研究所と電子総合技術研究所とともに、2005 年から CERN の LHC 加速器から出力されるペタバイト級のデータ処理を Grid 技術を用いて行う Grid DataFarm プロジェクトをスタートさせた。

その他、様々なグリッドの研究成果をあげたことが評価され、国際的な Grid 構築の活動のため、Global Grid Forum の Steering Committee (実行委員)委員として選出され、Global Grid Forum の Advanced Programming WG の担当として、幾つかのプログラミングモデルに関する White Paper を GGF で出版し、性能モデルや測定に関する講演を行っている。

##### 自己評価

成果は上記に述べたとおりであり、かつ詳細は報告書にある。

自己評価としては、総括的・客観的に見ても、本さきがけ研究は当初の目的を遥かに上回る成果を達成し、大変満足しており、かつ研究分野的に我が国のグリッド研究全体の立ち上げ、並びに国際的な活動に大きな貢献をしたと述べたい。

成果は上記に述べているが、これらによる分野的ならびに社会的なインパクトは大きく、研究開始時にはほとんど知られていなかったグリッドというキーワードおよび研究分野が一般的な市民権を得たと感じられる。

上記に述べたが、当初は Presto Cluster は性能評価およびシミュレーションのみに用いる予定であったが、下記に述べる実際のクラスタおよびグリッドアプリケーションを共同研究でターゲットすることにより、世界 4 位の PC クラスタ計算機、並びに大学の情報系の研究室としては世界唯一であろう莫大な計算インフラを構築し、その技術を獲得できた。すでに研究室では東工大の計算機センターのスパコンを完全に凌駕する計算パワーを有しており、後継プロジェクトでは世界一の PC クラスタ、並びに我が国では 5-6 位(世界では 20-30 位程度)のクラスタが構築できる予定である。並びに、クラスタの

ミドルウェアの研究を種々行うことが可能になり、研究の幅が大幅に広がった。

当初の予定より、遥かに多くの外部のアプリケーションならびにグリッドシステムの研究者と共同研究を行うことができた。特に、OR では世界記録を幾つか達成し、また高エネルギーの分野では情報分野と高エネルギー物理学者が「史上はじめて」本格的な共同研究を行う体制の一端を担うことができた。また、Global Grid Forum に対しても、我が国から技術的に貢献し、それが認められることにより体制づくりの重要なパートナーとして迎え入れられた。これらにより、他分野並びに海外からの本プロジェクトならびにさきがけに対する評価を最終的に大変高いものにできたと信ずる。

グリッドおよびさきがけの研究に関して、内外の様々なシンポジウムやセミナーなどで、基調講演や招待講演を行った。種々のマスコミの取材を受け、テレビ・ラジオなどに出演し、グリッドおよびさきがけ研究のPRを(一部)行うことができた。

- テレビ: NHK 「クローズアップ現代」
- テレビ: NHK 教育「サイエンスアイスペシャル」
- テレビ: CS コンピュータチャンネル (PRESTO Cluster を中心としたさきがけ研究のプロジェクトの紹介)
- ラジオ: BBC World ラジオ (グリッドに関するラジオ)
- 日経コンピュータ、日経新聞など、各種新聞・雑誌
- AMD 社の Presto III に関する全世界プレスリリース、および ZDNet, Yahoo などのインターネット系海外マスコミによる記事

などである。

また、米国 Computerworld 誌が主催する、米国の一般情報ビジネス系の大きな賞である Computerworld Honors Program を Presto Cluster の功績により 2002 年 4 月に受賞することになった。

これらの活動により、松岡は平成 13 年より、東京工業大学・学術国際情報センターにおいて Grid および HPC 技術の研究を推進する教授職に昇進した。引き続きセンターにおいて、グリッド化の推進プロジェクト提案などを行い、平成 13 年度補正予算で我が国ではじめてセンターにグリッドのインフラの予算として「東工大キャンパスグリッド」(約 2 億円、800 プロセッサ/1.2TeraFlops のキャンパス内グリッドインフラ)が認められた。また、米国の二大グリッドセンターの一つである NSF San Diego Supercomputer Center (SDSC) とグリッド研究における包括的な MOU を結ぶことを我が国の計算機センターとしてはじめて行ったなど、数々のセンターとしての成果につながっている。

#### 5. 領域総括の見解:

高性能計算を超広域で行う計算機システム構築に必要な多くの基礎技術を確立し、また、計算機システムを開発し、世界的レベルの性能評価値を得たことは高く評価できる。超広域高性能計算環境の研究は米国が進んでおり、わが国においても大変重要な研究テーマとなっているが、松岡聡はすでにそのリーダーの一人となっており、今後の研究発展が期待される。

#### 6. 主な論文等:

- T. Suzumura, T. Nakagawa, S. Matsuoka, H. Nakada, and S. Sekiguchi. Are Global Computing Systems Useful? Comparison of Client-server Global Computing Systems Ninf, NetSolve Versus CORBA. In Proc. 14th IEEE Intl. Parallel & Distributed Processing Symp., IEEE Computer Society Press, pp. 547-556. 2000.

- Satoshi Matsuoka, Hidemoto Nakada, Mitsuhsa Sato, Satoshi Sekiguchi. Design issues of Network Enabled Server Systems for the Grid, Proc. GRID'2000: International Workshop on Grid Computing, Bangalore, India, Springer LNCS 1971, pp.4-17, Dec. 2000 (Invited paper).
- \*Henri Casanova, Satoshi Matsuoka, and Jack Dongarra. Network-Enabled Server Systems and the Computational Grid. Proc. High Performance Computing Symposium (HPC'01), Advanced Simulation Technologies Conference, April 22-26 in Seattle, Washington (USA), 2001.
- \*Atsuko Takefusa, Satoshi Matsuoka, Henri Casanova, and Fancine Berman. A Study of Deadline Scheduling for Client-Server Systems on the Computational Grid. The 10th IEEE International Symposium on High Performance Distributed Computing (HPDC10), IEEE Computer Society Press ,pp. 406--415, Aug. 2001.
- Toyotaro Suzumura, Satoshi Matsuoka, and Hidemoto Nakada. A Jini-based Computing Portal System, Proceedings of IEEE/ACM Supercomputing '2001, IEEE Computer Society, Denver, Colorado, Nov. 2001 (to appear, CD-ROM Proceedings).