

# グリッドコンピューティングは未来の IT インフラとなり得るか

管理行政学科  
9950037番  
畑山 岳路

## 目次

### 第一章 グリッドコンピューティングとはいかなるものか

- 一、グリッドコンピューティングとは
- 二、グリッドコンピューティングの利点

### 第二章 グリッドコンピューティングの歴史

- 一、インターネットの変遷からみるグリッド
- 二、分散コンピューティングの歩みとグリッド
- 三、グリッドコンピューティング研究の歴史
- 四、まとめ

### 第三章 グリッドコンピューティング実現のための技術

- 一、Globus Toolkit
- 二、Ninf プロジェクト
- 三、Net Solve
- 四、ネットワーク・インフラ
- 五、まとめ

### 第四章 グリッドコンピューティングの課題

- 一、セキュリティ
- 二、スケジューリング
- 三、故障時の対応

### 第五章 グリッド研究の現状

- 一、GGF
- 二、グリッドプロジェクト
- 三、グリッド技術の種類
- 四、民間企業の動向

### 第六章 結論

## 序章

グリッドコンピューティングとは、現在、国際的に活発に研究開発が進められている次世代のインターネット利用技術である。これは、ネットワークに接続された多種多様な情報資源（携帯端末、PC、スーパーコンピュータ、データベース、センサー、観測装置、人、ソフトウェアなど）にアクセスし、動的に変化する環境の中で、大容量データを取り扱うシステムを構築し、高信頼かつ安全に利用するための技術である。

膨大な科学技術計算、様々な種類の情報資源やサービスの享受を受けられるということなどでグリッドコンピューティングは注目を浴びている。新たな物質材料や医薬品などを創出したり、地球規模の異常気象や温暖化現象の予測・解析をしたり、各種産業における新製品設計のサイクルを短縮したりするためには、現在利用可能な超高速計算機のさらに数十倍以上の規模の計算能力が必要とされる。また、これらの研究・開発作業の多くはもはや一個人・一企業の能力を超えつつあり、地球規模の協力が不可欠となりつつある。これらの観点から、国境を越え、遠隔地の研究者たちがチームを組んだり、コラボレーションを図る必要性が高まったことが、特に大きな要因ともなっている。

もともと、グリッドコンピューティングは「スーパーコンピュータを高速なネットワークで繋いで大きな計算をするためのシステム」ということであったが、インターネットの爆発的な普及やコンピュータ性能の向上などのコンピューティングをめぐる環境の変化により利用領域が拡大している。ネットワークに接続された複数の高性能コンピュータを一つの巨大コンピュータのように使用するのとはもとより、多数のパソコンの遊休時間をインターネットを介して収集し、スーパーコンピュータ並の計算をさせたり、大規模なデータや各地に分散配置されたデータに遠隔地からアクセス可能にするなど様々である。

グリッドコンピューティングは、インターネット利用をさらに広げていく技術として世界中の研究者・技術者・コンピュータメーカーなどにより、研究開発、標準化、ビジネス化が勢力的に進められている。本論では、グリッドコンピューティングについて、どういったものか、歴史、課題、現状などについて深く考察していき、グリッドコンピューティングの実体、グリッドコンピューティングの可能性（将来性）、今後の展望、そしてグリッドコンピューティングが未来のITインフラとなり得るのかということの研究していきたいと思う。

## 第一章 グリッドコンピューティングとはいかなるものか

### 一、グリッドコンピューティングとは

グリッドコンピューティングとは、インターネットに接続されている個人情報端末やパソコン、スーパーコンピュータ、大容量データセンター、可視化装置、観測装置等をすべて統合してユーザが扱えるようにするための基盤技術（ソフトウェア、ネットワーク、ハードウェア）とこれを活用するプログラムであり、インフラストラクチャである。「次世代インターネット利用技術」として各分野から高い期待がなされている。グリッドの語源は「電力送電網（Power Grid）」のことで、欲しいサービスがいつでも、どこからでも、誰でもが手軽に利用できるようになることを目指したものである。グリッドコンピューティングの定義は、「ユーザがコンピュータの所在などを意識することなく、いつでも必要なだけ計算資源を利用できる環境」というものである。現代社会では、コンセントにプラグをさすだけで電気を使うことができるインフラストラクチャが提供されている。このとき、自分が使っている電気が、どこの発電所で発電されて、その時の発電の方法、どこを通過してきたかという送電経路などを考える人はまずいないだろう。コンピュータグリッドは、まさに計算資源について同様の環境を提供するインフラストラクチャであり、ネットワークに接続するだけで、欲しいサービスの所在などを意識することなくいつでも利用できる環境を提供しようとするものである。ところがインターネットのサービスは、例えば URL などで「どこのこのファイルを持ってくる」というようにサービスそのものに指定に加えて資源、場所、経路などを明示的に指定するのが通常のやり方である。これを電力のサービスと同じように、欲しいサービスを記述すればそれ以上の情報について末端ユーザは何も気にしなくても必要なサービスが得られるというのがグリッドコンピューティングの目指すものである。

グリッドコンピューティングの概念は、電力供給のユーティリティモデルのように、計算資源や情報資源をネットワークを介してユーザに必要なときに必要なだけ使える環境を提供することであると前述した。これをもう少し厳密に定義すると「グリッドコンピューティングとは、ネットワーク上に分散した多様な計算資源や情報資源（コンピュータ、記憶装置、可視化装置、大規模実験観測装置）を仮想組織のメンバーが一つの仮想コンピュータとして利用する環境」となる。ネットワーク上にある計算資源や情報資源を仮想コンピュータとしてとらえ、それを利用するために、現実の組織を超えてある目的のために作られた仮想的な組織のメンバーとして登録して、利用する。これらの機能は、ミドルウェアが動作し実現する。さらに、例えば自分がプログラミングした所望の計算をしたいと考えたときに計算を実行するだけでなく、他のサイトにある有用なプログラムやデータを利

用して自分の目的にあったデータの処理・加工をする（アプリケーションサービス）ことが可能である。

## 二、グリッドコンピューティングの利点

グリッドコンピューティングを構築することによる利点として三つ考えられる。

第一に、科学者や技術者に分散環境での効率のよい共同作業のツールを与えることである。例えば、EUのData Gridというプロジェクトでは、世界各国の3000人の高エネルギー物理学者がグリッド上でデータやプログラムを共同開発し、共同利用して研究の効率を上げつつ、研究競争が行われている。巨大科学研究では実験装置やデータの巨大化、大規模計算への要求が顕著であり、研究効率向上のためにも研究資源の共有化が求められている。

第二に、分散資源の有効活用、使い勝手の向上である。例えば、ネットワークに接続されているが、使われていない資源などを活用して、スーパーコンピュータ並の計算能力を得たり、資源をネットワーク上で共有化して、一カ所では得られない大規模な資源として利用することを可能とする。ただし、この場合、ネットワークの速度を十分考慮しないと、データ転送に時間がとられ、高速な計算性能が得られないことがある。

第三に、負荷分散と信頼性向上である。資源が分散化されているので、個々に大きな負荷が発生してもネットワーク上の他の装置に負荷を分散できる。個々にはピーク時の最大負荷に耐えられる設備を持つ必要はなく、ネットワーク全体として耐えられる設備を準備をすればよい。ゆえに、高価な高性能計算機が幾つも必要だということでもなくなる。また、個々の装置がダウンしてもシステム全体としては運用が持続しており、信頼性が高くなる。ある部分が集中的にサイバーアタックに狙われた場合も、その部分を切り離し、業務・サービスは他の部分にて対応可能なため、危機対応としても有用である。

ここでいう、資源とは計算資源及び情報資源のことを指す。また、グリッドコンピューティングは、コンピュータの高性能化技術ではなく、コンピュータの利用技術である。「グリッドがあれば大きなスーパーコンピュータはいらない」「大きなスーパーコンピュータがあれば、グリッドコンピューティングなんてものは必要ない」ということではないのである。ネットワークを介すると計算効率が極端に低下する種類の計算も多くあり、スーパーコンピュータの性能向上は必要である。グリッドコンピューティングにおいても個々の高性能コンピュータの性能向上はグリッド全体の計算能力向上になる。

## 第二章 グリッドコンピューティングの歴史

グリッドコンピューティングは、1980年代の遠隔利用研究や分散コンピューティングの研究から発展してきている。1990年代に入り、理論だけでなく拘束なネットワークを用いて実証することも可能になり、離れた場所にあるコンピュータをつなげ、作業を分担して行くと、求める答えがより早く得られるのではないかという考えのもと、科学者や研究者の間でインターネットを利用した分散コンピューティングの研究は加速していった。ある研究グループは分子の設計やコンピュータ・グラフィックスのレンダリングといった並列処理アプリケーションを何台ものコンピュータを集めて実行し、また他の研究グループでは何台かの大きなスーパーコンピュータを広域ネットワーク上で接続してひとつの仮想的なメタコンピューティングを作り上げ、ユーザへ開放する試みを行った。もちろんこれらの試みはネットワーク、計算能力、ソフトウェア・インフラストラクチャなどの実際の可能性を追求し、いかにそれを発展させるかという意図のもとに行われた。

最近のインターネットのバックボーンは、国内では10 Gbps クラス、国際間で1 Gbps となり、ブロードバンド時代が到来している。また、インターネットの利用拡大とともにネットワークの高速化、信頼性の向上が進み、使い勝手の良いインターネットを介した分散コンピューティング環境を提供できるようになってきた。

ここでは、インターネットの変遷と分散コンピューティングの歩みからグリッドコンピューティングの歴史を見るとともに、グリッドコンピューティング研究の歴史を調べていき、グリッドコンピューティングの今後を考察していきたいと思う。

### 一、インターネットの変遷から見るグリッド

インターネットは広く地理的に離れたコンピュータを通信回線を利用し、相互につなげた ARPANET をはじめとするいくつかの先駆的なコンピュータ・ネットワークが1960年代後半から1970年代前半にかけて開発されてきた。この目的は、データ処理・高速演算・データベース・図形処理機能などの遠隔地からの利用、複数ユーザによる資源の共有・資源の障害時のバックアップなどの機能の提供、といった現在のグリッドコンピューティングの考えと似たものであった。しかし、インターネット利用は、ネットワークを介してメールをやりとりする電子メールの利用や遠隔地の情報を見るなどといった用途に変化していき、Web ブラウザが作られその利便性から爆発的に世界中に広まっていった。この普及は、Web ブラウザが端末のハードやOSによらず情報を閲覧できることにあった。インターネットの普及とそのネットワークの信頼性向上によって、企業の業務でもインターネットが使われるようになってきた。なによりも WWW の登場がインターネットの爆発

的普及の要因となっている。個々の要素技術としては新規性がなかったものの、統一的な取り扱いとしてプロトコルを定めたことによって、誰でもが簡単にネットワークにアクセスしてコンテンツを求めることができるようになった。

これまでインターネットは、どちらかと言えば、電子メールをはじめとして、人と人を結んできたものであったが、これからはコンピュータとコンピュータをも結んでいくものとなっていくであろう。コンピュータ同士が情報交換することによってより広範に自動的に情報検索・交換を可能にする。また、WWW においてインターネット上にあるデータや情報にアクセスするという、どちらかと言えば静的に与えられたデータに対するアクセスであったのに対し、グリッドコンピューティングはネットワーク上に存在するコンピュータ、記憶装置、可視化装置、大規模実験観測装置といったハードウェアの情報資源や人的資源に対する高性能、柔軟かつ安全なアクセス手段を提供することを可能とする。そして、これらの情報資源を必要ときに必要な資源に必要なだけアクセスし、自分の手元の端末を通して自由自在に操ることができるものである。

## 二、分散コンピューティングの歩みとグリッド

分散コンピューティング・アーキテクチャとして、マルチ・プロセッサ、クラスタが挙げられる。そして、これらを発展させていったのがグリッドコンピューティングである。

まず、マルチ・プロセッサとは、一台のコンピュータに複数個のプロセッサを搭載し、各プロセッサに処理を分散させるものである。これには「複数のプロセッサでメモリを共有するタイプ」「メモリを共有しない複数のプロセッサをクロスバー・スイッチなどで接続するタイプ」、そしてこれらを組み合わせたものの3種類がある。

クラスタとは、同一構成の複数台のマシンをネットワーク接続し、各マシンに処理を分散する。ここで、複数のマルチ・プロセッサを使ってクラスタを構成することもできる。クラスタは非常に多量なスループットを必要とする計算や、複数のプロセッサに並列処理とは異なる方式で多くのジョブプロセスを分配し、それぞれの結果を集めてひとつの結果を導くような処理を行う場合に利用されている。

そして、ネットワーク上のシングル・プロセッサ、マルチ・プロセッサ、クラスタの集合体を仮想的な単一コンピュータとして扱うのがグリッドコンピューティングである。グリッドでは、各システムの処理能力にばらつきがあるので、効率よく処理を分散させるのに複雑な並列処理管理機構が必要とされる。その一方で、高度な並列処理で非常に高いスループットを実現することができる。

## 三、グリッドコンピューティング研究の歴史

1992年、米国国立スーパーコンピュータ応用研究所(NCSA: National Center

for Supercomputing Applications)のCharlie Catlett氏とLarry Smarr氏が「メタコンピューティング」という概念を発表し、ネットワーク上に仮想的なコンピュータ環境を構築し、大規模な並列処理を行うシステムの研究を進めた。1995年に、イリノイ大学のTom De Fanti氏とアルゴンヌ国立研究所のRick Stevens氏が共同で行った、I-Way (The Information Wide Area Year)プロジェクトにおいて、グリッドコンピューティング・コンセプトの最初の大規模なメタコンピューティングの実験が行われた。このプロジェクトでは、広域高速ネットワークにて全米の17カ所の計算センターを結合し、多くのアプリケーションを稼働させるデモを行った。このI-Wayプロジェクトをルーツとして、1996年にアルゴンヌ国立研究所のIan Foster氏と南カリフォルニア大学のCarl Kesselman氏のチームがGlobusプロジェクトを開始し、高性能分散コンピューティングのためのミドルウェアが開発された。他にも、I-Wayプロジェクトを契機に、全米科学財団(NSF)のNational Technology Grid、NASAのInformation Power Grid、European Data Grid、Particle Physics Data Grid、Grid Physics Network (GrIPhYn)、Network for Earthquake Engineering Simulation Grid (NEESgrid)といったグリッドプロジェクトが登場してきた。

Globusプロジェクトにて開発されたソフトウェアGlobus Toolkitは、ユーザ認証、グリッド資源配分管理などグリッドミドルウェアの基本機能を提供している。このソフトウェアは、アルゴンヌ国立研究所からオープンソースとして提供され、多くのグリッド構築研究プロジェクトで利用され、事実上の標準となっている。Globus Toolkitについては、この後の第三章において紹介する。

一方、もう一つの研究の流れとして、インターネットを介して遊休CPUを活用する研究がある。1985年に、Wisconsin大学のMiron Livny氏が、分散したワークステーションの遊休CPUを使用して、計算実行を行うことを提案した。1991年にはCondorプロジェクトの中で、400CPU相当を集めることに成功している。1997年には、Scott Kurowski氏がEntropia社を設立し、2年後にはインターネットを介して遊休パソコンを集め、当時のスーパーコンピュータの最高性能に近いTera Flops並の計算を実現した。

メタコンピューティングと遊休CPUの活用は、ネットワーク上に分散されたコンピュータを仮想的に一つのコンピュータとしてとらえる技術であり、グリッドコンピューティングの要素技術となっている。これらについては、第五章にて詳しく説明する。

日本でも遠隔地からコンピュータを利用するグローバルコンピューティングの研究が進められており、1994年からは、電子技術総合研究所(現産業技術総合研究所)の関口智嗣氏、東京工業大学の松岡聡氏等を中心に、Ninf(Network Infrastructure Global Computing)プロジェクトが開始された。クライアントサーバ型のグローバルコンピューティングとして、クライアントがネットワークを介してサーバに計算を実行させるRPCの設計と実証が進められた。この研究成果は、現在グリッドミドルウェアの一部として標



準化が検討されていて、独立して始められた同様の研究プロジェクトであるテネシー大学のNet Solveプロジェクトと協力して、精力的に研究開発が進められている。また、日本原子力研究所、理化学研究所、大学の計算センターなどでは、スーパーコンピュータのネットワーク遠隔利用、共同利用の研究も進められてきている。

#### 四、まとめ

1980年代にネットワークを介した分散処理が注目され始め、90年代に入るとネットワークで接続されたコンピュータを仮想的な並列計算機とみなして、その上で並列計算を行うクラスタシステムに関する研究が盛んになった。また、同時にネットワーク技術もめざましい進歩を遂げ、高い通信性能を持つネットワークが利用できるようになった。こうして、今まで一つの部屋の中に置いておくことが当然であった、クラスタやLANのレベルで利用されていた分散処理などが、広域環境においても十分利用可能となってきた。こうした流れの中で、並列・分散・高性能計算の研究者がグリッドコンピューティングに着目するのはごく自然な流れだと思われる。その結果、グリッドコンピューティングという分野はまだ生まれて間もないのに、世界中で様々なシステムが各研究グループによって開発されるようになった。しかし、グリッドコンピューティングを完全に実現させるためには様々な課題がある。技術的な課題、管理上の問題、社会的な問題などいろいろある。これらに関しては、この後に触れていきたいと思う。

### 第三章 グリッドコンピューティング実現のための技術

グリッドコンピューティングに関する研究やシステムの開発は世界中の研究グループによって行われている。主なものとして、次のものが挙げられる。

アメリカ

- **A k e n t i**  
分散ネットワーク環境でのセキュリティサービスを提供するセキュリティモデルおよびアーキテクチャ。
  
- **A p p L e S**  
アプリケーションレベルでのスケジューリングに着目した研究プロジェクト。
  
- **C o n d o r**  
分散配置されたワークステーション等を使って高スループットな計算を行うための開発環境等の枠組みを提供するシステム。
  
- **G l o b u s**  
アメリカを中心とした、現在最も大きなプロジェクト。グリッドコンピューティング環境のソフトウェア・インフラストラクチャを構築するためのツール群を提供している。G l o b u s T o o l k i tを開発。
  
- **I c e T**  
グリッドコンピューティング環境における仮想的な計算機の併合や分割、コードやデータの可搬性などの機能を備えたシステム。
  
- **I P G**  
コンセントにプラグを差し込むだけで誰もが電力の供給を受けることができると同様に、世界中の情報を簡単に引き出す事ができるような枠組み、仕組みを提供するシステムに関する研究プロジェクト。
  
- **N e t S o l v e**  
ネットワークを介して遠隔地にある科学技術計算用ライブラリを利用するためのクライアント・サーバモデルに基づくシステム。

## ヨーロッパ

- EuroTools

ヨーロッパ全体で高性能計算・ネットワークを利用するためのプロジェクト。

- Albatross

グリッドコンピューティング環境におけるクラスタシステム（広域クラスタシステム）上でのプログラミング方法や、高性能を得るための方法などに関する研究プロジェクト。

- Unicore

ユーザが遠隔地にある計算資源にジョブを送信するためのシステム。

## 日本

- Ninf

沿革手続き呼び出しの手法を利用し、プログラマが容易に遠隔地にあるハードウェア、ソフトウェア、データベースなどを利用することのできるシステム。

この中から、事実上の標準ミドルウェアとなっている、GlobusプロジェクトのGlobus Toolkitと日本のNinfプロジェクト、米国のNetSolveを紹介していく。さらに、グリッドコンピューティングを実現するために必要なネットワーク・インフラの開発を行っている研究も紹介していきたいと思う。

## 一、Globus Toolkit

1996年頃に、Globusプロジェクトが発足した。Globusプロジェクトは並列・分散計算、ネットワーク、セキュリティといった様々な分野の研究者達が参加するプロジェクトであり、現在もグリッドコンピューティングの分野におけるトップレベルの研究者達がこのプロジェクトに参加している。Globus Toolkitは、Globusプロジェクトから生まれた。Globus Toolkitはユーザ認証システム、通信ライブラリ、計算資源管理機構といったグリッドコンピューティングシステムの構築に必要とされる基本的なサービスの集まり(toolkit)であり、グリッドコンピューティングシステムのソフトウェア・インフラストラクチャを構成する要素の事実上の標準になりつつある。以下に、Globus Toolkitが提供するサービスを紹介する。

- 資源管理 (GRAM: Globus Resource Allocation Manager)

ネットワークを介したジョブ管理（遠隔ジョブ管理）の機能を提供する。遠隔ジョブ管理には、プロセスの起動、ポーリング、強制終了が含まれる。

- 通信（Nexus）

Globus Toolkitの通信サービスはNexusという通信ライブラリによって提供されている。Nexusはマルチキャスト機能を含む低レベルな通信機能を提供しており、メッセージ通信や遠隔手続き呼び出しなどの高レベルな機能を構築できる。Nexusは単なる通信プロトコルのみならず、属性を設定することによって、セキュリティや圧縮などの機能もサポートしている。

- 情報サービス（MDS：Metacomputing Directory Service）

アーキテクチャタイプ、OS、メモリ、ネットワークバンド幅、利用可能な通信プロトコル、IPアドレスなどの様々な情報を管理する情報サービスの機能を提供する。

- セキュリティ（GSI：Grid Security Infrastructure）

認証、アカウント、プライバシー保持などのセキュリティに関する構成要素である。GSIは、遠隔プロセスの起動やプロセス間通信の際に使われる。Globus Toolkitでは、ユーザは計算に際し一度だけ認証を行い、その時にプロセスがユーザに代わって資源を獲得できるように証明書が発行される。

- 遠隔データアクセス（GASS：Global Access to Secondary Storage）

遠隔のファイルサーバへのアクセス機能を提供する。

- 実行ファイル管理（GEM：Globus Executable Management）

実行ファイルの構築、転送、配置などの機能を提供する。

## 二、Ninfプロジェクト

スーパーコンピュータや大規模クラスタシステムなどの高性能計算システムは、そう簡単に手に入るものではない。それらのシステムを導入している企業、研究機関、大学などは限られている。それら高性能計算システムの魅力は、高い計算処理能力にある。手元のPCやワークステーションでは何日、何年かかるかわからないような計算も、スーパーコンピュータを使えば数分、数時間で答えを求めることが可能である。処理時間だけでなく、例えばあるスーパーコンピュータに搭載されている計算ライブラリを使えば、非常に精度の高い計算結果を得られることもできる。また、計算のみならず、どこかのサイトが保持

している大規模データベースに手元の端末から簡単にアクセスしたいということもあるかもしれない。このような要求を満たすシステム、手元にある計算資源（PCやワークステーション）上でユーザインタフェースを提供し、遠隔地にある高性能計算機を利用するシステムのことをD a t o r r ( Desktop Access to Remote Resources)という。N i n fはこのD a t o r rの一つであり、電子技術総合研究所（現産業技術総合研究所）で設計、開発された。

N i n fは、Network based Information library for High Performance Computingの略で、ネットワーク数値情報ライブラリという意味である。N i n fの目的は、LANあるいは広域ネットワーク上の数値計算ライブラリや科学技術計算に必要な数値情報データベースを通じて、おもに科学技術計算分野の情報ならびに計算資源を提供・共有する仕組みを提供することである。計算の時間のかかる数値計算ライブラリをスーパーコンピュータなど速い計算機で行うことで、科学技術計算アプリケーションを高速化するというよりも、質の高い、あるいは信頼できるライブラリをネットワーク上の知的資源として、ソースプログラムでなく、使いたいときに、いつでも使えるようにするというに重点を置いている。N i n fの基本システムはクライアント/サーバモデルに基づいて設計されていて、N i n fシステムはクライアント、N i n fサーバ、要素ルーチン、メタサーバの4つの要素から構成される。

ユーザはクライアントプログラムをCやFortran、Javaなどの既存のプログラミング言語を用いて作成する。N i n fサーバは分散システムを構成する各々の計算機上で待機し、クライアントからのリクエストに応じてあらかじめ用意されている要素ルーチンを実行し、その結果をクライアントに返す。このようにN i n fシステムは、「計算サービスを提供する側（サーバ）」に対し、手元の端末（クライアント）から計算要求を発信することにより、遠隔地にある計算資源を利用することを可能とするシステムである。

### 三、NetSolve

科学技術計算のための数値ライブラリは通常ソースファイル、あるいはオブジェクトファイルを手元のコンピュータにインストールし、それらを利用するユーザプログラムとリンクし実行します。あるいは、計算機センターなどのスーパーコンピュータにログインし、すでにインストールしてあるそれらのライブラリを利用します。手元のコンピュータにライブラリをインストールする場合は、ライブラリのバージョンが上がるたび、あるいはソフトウェアのバグが見つかるたびに、ライブラリを新しくする必要があります。また、計算センターの場合は、利用するためにわざわざ計算センターのスーパーコンピュータにログインし、実行する必要があります。

これらの手間を軽減するためのソフトウェアがNetSolveである。NetSolveは、手元のコンピュータでほぼ今まで通り数値ライブラリを呼び出すと、ネットワー

クで接続されたスーパーコンピュータでその数値ライブラリを実行するためのソフトウェアで、この時、手元のコンピュータに数値ライブラリをいちいちインストールするまでもなく、常に最新のライブラリで計算することができる。また、計算センターのスーパーコンピュータにいちいちログインする必要もなくなるわけである。

NetSolveは、NetSolveクライアント、NetSolveエージェント、NetSolveリソースで構成されている。NetSolveクライアント（以下、クライアント）は、手元のコンピュータで動き、NetSolveエージェント（以下、エージェント）を通じて実際の計算はNetSolveリソース（サーバ）（以下、計算リソース）で行われる。計算リソースはネットワークで接続されたコンピュータである。実行の流れとしては、クライアントが実行の要求をエージェントに対し行う。エージェントは要求を受け取り、適切な計算リソースを選ぶ。クライアントは選ばれた計算リソースにデータを送る。計算リソースは、クライアントからデータを受け取り、数値ライブラリを実行し、結果をクライアントに返す。というようなものである。

NetSolveは、ネットワークに接続されたコンピュータ（計算リソース）にライブラリ実行の要求をするわけだから、NetSolveサーバが起動していれば世界中のコンピュータで実行できる可能性がある。現在、NetSolveのインターネットを用いた環境におけるテストベッドとして、デンマークコンピューティングセンター（Danish Computing Center for Research and Education）、NCSA（National Center for Supercomputing Applications）、カリフォルニア大学サンディエゴ校（University of CALIFORNIA, San Diego）、テネシー大学ノックスビル校（University of Tennessee, Knoxville）の協力により、それぞれのサイトでNetSolveサーバが起動し、いつでも誰でもどこからでも実験が行える環境が整っている。

計算リソースは、ワークステーション、スーパーコンピュータ、計算機クラスタなど、さまざまな計算機であり、エージェントはその時々ネットワーク状況、計算機の負荷を見ながら最適な計算リソースを選びます。計算リソースの種類に関係なく同一のインターフェース利用することができるということで、たとえ計算リソースが分散メモリ型の並列計算機であっても並列プログラミングをすることなく並列計算ライブラリを利用することができるということになる。

NetSolveとNinfは非常に類似している。開発時期もほぼ同時期に行われている。これらのプロジェクトはお互いに協力し合い、例えばNinfサーバはNetSolveの計算リソースを利用する機能を持っていて、NinfクライアントでNetSolveの計算リソースを利用するなどが可能になっている。今後は、両者のインターフェースも極めて類似していることから、ネットワークの転送プロトコルの共通化を進めることにより、プロトコルの変換なしでそれぞれのクライアントからそれぞれの計算サーバを呼び出せるようにしようとしている。

## 四、ネットワーク・インフラ

グリッドコンピューティングを実現させるためには、今よりもバンド幅の大きいネットワークが必要となってくる。そこで、次世代の高速インターネットの構築、次世代のネットワークアプリケーションの開発を行っている I 2 - D S I ( Internet2 Distributed Storage Infrastructure ) というプロジェクトがある。世界中のどこからでも大容量の記憶装置に高速にアクセスし、世界中にある記憶装置の所在を意識することなく、高速にアクセスできるようにするというものである。単一の記憶装置を用いる場合、ネットワーク的に近い所であれば問題はないが、遠く離れた所からのインターネット経由では高速にアクセスすることができない。単一の記憶装置でも多大な費用をかけて高速で品質が保証されている専用のネットワークを導入すればできるのであろうが、そういうネットワークを世界中に張り巡らせるのは現時点では無理であろう。I 2 - D S I では、複数の記憶装置にコピーを持たせ、ネットワーク的に近い所にアクセスできるようにすることで、どこからでも高速なアクセスを実現しようとしている。そのためには、いかに高速に、効率良く複数の記憶装置間で同期をとる ( ファイルの内容を一致させる ) かということと、ファイルのアクセスをいかに近くの記憶装置に行わせるかということである。

他にも、Internet Backbone Protocol ( I B P ) の研究を行っている研究グループもいる。I B P はファイル転送のためのプロトコルで、現在ネットワークがどういうふうにつながっているかわからなくてもつかえているのと同じように、記憶装置においても、どこの記憶装置に存在するかを意識することなく、かつ効率よくアクセスできることを目指している。I B P を用いたアプリケーションとして、I B P - M a i l という巨大な添付ファイルを通常の電子メールで送るためのアプリケーションがある。電子メールでは、メールサーバの容量を超える巨大なサイズのメールを送ることはできないが、I B P - M a i l では送信時に巨大な添付ファイルを I B P を用いてある記憶装置に書き込み、実際に受信するまでの間に受信者の近くの記憶装置にデータを移しておく。受信時には添付ファイルは近くの記憶装置から I B P を用い読み込ませる。したがって、メールの送信者、受信者はそれらの行いを意識することなく巨大な添付ファイルを電子メールでやりとりすることができるのである。

## 五、まとめ

これまでに紹介してきたもの以外にもグリッドコンピューティングを実現させるための研究はたくさんある。しかし、紹介してきたものも含めてまだまだ研究段階といったところで、まだ完全に実行可能なものとしては成り立っていない。これらを実行可能なものとしていくためには、さらなる研究が必要なのと同時に、これらの研究者が共同に研究して

コラボレーションを図っていく必要があるだろう。それは、グリッドコンピューティングが世界的に稼働させるためにはとても重要なことなのである。今後のさらなる研究に期待したいと思う。

実現のための技術を研究するとともに、グリッドコンピューティングの問題点に対しても前向きに検討していかなければならないだろう。次の章では、解決すべきグリッドコンピューティングの課題について述べていく。



## 第四章 グリッドコンピューティングの課題

現在、Globus Toolkitをはじめとする様々なミドルウェアが存在し、多くのプロジェクトでグリッドコンピューティングの構築や構築されたグリッドコンピューティング上の応用開発が行われている。しかし、簡単にシステムを作れるほどまだソフトウェアは整備されてはいない。また、現在のグリッド技術にはいろいろな問題がある。

その中でも一番重要なのが、セキュリティに関する問題。グリッドコンピューティングは常にネットワークに接続されているわけであり、常時接続されているからには通信内容が盗聴される可能性もあるわけである。そういった問題をどのように解決するかがとても重要なことである。

他にもスケジューリングの問題や故障時の対応などがあるが、それらを検証していきたいと思う。

### 一、セキュリティ

セキュリティというものは大きく分けて、認証、資源の制御、情報の秘匿からなる。

認証とは、資源を利用する主体を識別することである。ログイン時でのパスワード入力などがこれに相当する。

資源の制御とは、識別された主体に応じて、資源の利用に制約を課すことである。ここでいう資源とは、計算資源、ストレージ資源、ネットワーク資源などすべての資源をいう。

情報の秘匿とは、資源の利用に関する情報を秘密にすることである。ネットワークを介した通信を行うと、その通信内容が盗聴される可能性がある。情報を暗号化するなどをして、盗聴を防ぐことが情報の秘匿にあたる。

グリッドコンピューティングには、1．計算は常にネットワークを介して行われる。2．計算には非常に多くの資源が関与し、資源の構成は動的に変化する。3．計算に用いられる資源は多数の組織から提供される。4．計算は複数の資源にまたがって行われる場合もある。という特徴がある。こういった中で、認証、資源の制御、情報の秘匿の3つを実現することは通常的环境に比べてはるかに困難になる。

これらの特徴を持つグリッドコンピューティング環境におけるセキュリティシステムには以下の性質が必要となる。

#### 1．盗聴やサーバソフトウェアへのアタックといった攻撃への対処

通常的环境でも問題になることだが、グリッドコンピューティングにおいては、常にネットワークに接続されているということもあり、通常の場合よりも被害を受ける機会が多い。

## 2. 各組織の管理ポリシーを変更しないこと

グリッドコンピューティングでは、大規模計算をしたいという時に、大学や研究所、もしくは企業の計算センターにスーパーコンピュータを提供してもらうことがある。このような組織にはそれぞれの固有の管理ポリシー、セキュリティシステムを採用している。グリッドコンピューティングシステムをインストールするためだけに、組織の管理ポリシーを変更することは困難である。

## 3. プロセスへのユーザ権限の付与

計算は複数の資源にまたがって行われる場合もある。ユーザがある計算機上で起動したプロセスAが、さらに別の計算機上にプロセスBを起動する、といったことが可能でなければならない。この場合、プロセスAがユーザに代わってプロセス起動を行うことになる。このためには、プロセスにはユーザのけんげんを付与することができなければならない。

## 4. シングル・サインオン

プログラムが個々の資源にアクセスするたびに、ユーザにパスワード入力を求めるようなシステムでは、グリッドコンピューティングシステムは普及していかないだろう。一度パスワードを入力してグリッドシステムへログインしたら、計算をしている間は、再入力が必要ないようなシステムでなければならない。

セキュリティの問題にはこういったことがある。また、外部からの不正アクセスから守るファイアウォールとのセキュリティ制御の問題もある。グリッドコンピューティングは外部からアクセスしてくるわけであるから、ファイアウォールがあるとその内側の計算資源にアクセスできないわけである。そのため、ファイアウォールを導入している組織は、グリッドコンピューティングシステムを導入するために、その管理ポリシーを変更しなければならないという問題が生じてしまうのである。

## 二、スケジューリング

グリッドコンピューティングシステムのように複数の計算機上でプログラムを実行する環境において、プログラムの一部または全体をどのように計算機に割り当てるのか、というスケジューリングがとても重要なものになってくる。スケジューリングがうまくいかないと、最悪の場合、手元のパソコンで全ての計算をした方が早かったなんて状況にもなりかねないのである。上手なスケジューリングを行うためには、利用可能な計算機環境や解きたいプログラムに関する様々な情報を収集し、効率の良いプログラムへの割り当て方法を考える必要がある。

スケジューリングには、情報収集と予測、情報蓄積、プログラム割り当ての3つの処理が必要である。ここでグリッドコンピューティングシステムをネットワークに接続された、クライアント（プログラム処理を依頼する人、コンピュータ）、サーバ（依頼されたプログ

ラム処理を実行する計算機)、スケジューリングシステムの3つから構成されるとする。スケジューリングシステムは、クライアントのプログラムを実行するサーバを決定するために、グリッドコンピューティングシステム上のサーバやネットワークの状態に関する情報を収集することが必要である。例えば、大量の計算パワーを必要とするプログラムをできるだけ短時間に実行するためには、計算スピードの高いサーバへプログラムを割り当てなければならないし、大量のデータ転送を必要とするプログラムを短時間に実行するためには、クライアントとデータ転送能力の高いネットワークで接続されたサーバにプログラムを割り当てなければならない。また、大量の計算と大量のデータ転送の両方を必要とするプログラムにはサーバの計算スピードとネットワークのデータ転送能力の両方を考えなければならない。さらに、計算スピードやデータ転送能力は、サーバの性能やネットワークのバンド幅だけでは決まらない。どんなに高性能なサーバでも負荷が高い場合には仕様書通りの性能は出せない。ネットワークの場合もプログラムをサーバに割り当てる場合はバンド幅だけでなく混雑度も考慮する必要がある。

スケジューリングシステムは、各サーバのハードウェア情報(CPU性能、メモリ容量、等)や負荷情報(CPU上で実行されているプロセス数、メモリ空き容量等)、ネットワークのバンド幅や混雑度、といった情報を収集する必要がある。サーバの負荷やネットワークの状況は時々刻々と変化するので、定期的に収集しなければならない。そして、収集した情報から最も負荷の低いサーバ、混雑度の低いネットワークに割り当てるために将来のサーバの負荷やネットワークの混雑度を予測する。さらに、収集したサーバやネットワークの情報、予測値を蓄積し、今後の予測のための情報として使う。そして、最適なサーバ、ネットワークを予測したら、それを割り当てる。これが、スケジューリングシステムの役割である。

動的に変化する状況の管理をどのように開発していこうかというのが今後の課題である。

### 三、故障時の対応

グリッドコンピューティングシステムで、たくさんの遊休PCを利用して大規模な計算をしようという時に、PCがいきなりダウンした場合やあるPCが接続されているネットワークの故障でデータが送れなくなったなどの状況に陥った時に、処理途中のデータをどうするか、復帰してきた時にどう対応するかなどの問題がある。

これらの課題を解決しなければグリッドコンピューティング環境実現はありえないのである。各方面でこれらの課題を解決しようと研究が行われているが、解決するにはまだまだ時間がかかりそうである。

## 第五章 グリッド研究の現状

ここでは、グリッドコンピューティングの研究の現状を知るために、グリッドコンピューティングの標準化団体であるGGF (Global Grid Forum) と各国のグリッドプロジェクトを紹介していき、グリッドコンピューティングの類型 (応用開発)、民間企業のグリッドコンピューティングに対する動向をみていきたいと思う。

### 一、GGF (Global Grid Forum)

グリッドコンピューティングは、世界各国と接続できて初めて機能するものであるため、インターフェースの標準化が重要な課題となっている。

グリッド標準化活動は、1998年に、米国でグリッド研究プロジェクト関係者が集まりグリッドフォーラム (GF) が結成されたことを契機に、開始された。その後、E-Grid (European Grid Forum)、アジア太平洋ではAP-Grid (Asia-Pacific Grid Forum) がそれぞれの地区のグリッドコンピューティング開発を推進していた。2001年には、グリッドフォーラム (GF) が、E-GridとAP-Gridの標準化活動を吸収合併し、全世界的なグリッドコンピューティングの標準化団体グローバルグリッドフォーラム (GGF: Global Grid Forum) が結成された。

GGFの標準化の進め方は、個別技術ごとに研究グループまたはワーキンググループが作られ、目的・目標を明確にして議論を進めていく。現在は、7つの領域に21のワーキンググループと14の研究グループがある。年3回の全体会議の他に、個別会議、それぞれ個別に意見交換を行うことにより議論を進めていて、標準化の議論の進展は早い。このGGFの標準化は強制力があるものとはならないが、事実上の業界標準になっている。ここで、GGFにおける活動グループの一覧を以下に示しておく。

Area	Working Groups	Research Groups
<a href="#">Applications and Programming Environments</a>	?? Grid Checkpoint/Recovery	?? <a href="#">Advanced Collaborative Environments (ACE-RG)</a> ?? <a href="#">Advanced Programming Models (APM-RG)</a>

		<p>?? <a href="#">Applications and Test Beds (APPS-RG)</a></p> <p>?? <a href="#">Grid Computing Environments (GCE-RG)</a></p> <p>?? <a href="#">Grid User Services (GUS-RG)</a></p> <p>?? <a href="#">Life Sciences Grid (LSG-RG)</a></p> <p>?? <a href="#">Production Grid Management RG</a></p>
<a href="#">Architecture</a>	<p>?? <a href="#">Open Grid Services Infrastructure (OGSI-WG)</a></p> <p>?? <a href="#">Open Source Software (OSS-WG)</a></p> <p>?? <a href="#">New Productivity Initiative (NPI-WG)</a></p> <p>?? <a href="#">Open Grid Services Architecture (OGSA-WG)</a></p>	<p>?? <a href="#">Accounting Models (ACCT-RG)</a></p> <p>?? <a href="#">Grid Protocol Architecture (GPA-RG)</a></p> <p>?? <a href="#">Service Management Frameworks (JINI-RG)</a></p> <p>?? <a href="#">Semantic Grid RG</a></p>
<a href="#">Data</a>	<p>?? <a href="#">GridFTP-WG</a></p> <p>?? <a href="#">Data Access and Integration Services (DAIS-WG)</a></p>	<p>?? <a href="#">Data Replication (REP-RG)</a></p> <p>?? <a href="#">Persistent Archives (PA-RG)</a></p> <p>?? <a href="#">Grid High-Performance Networking (RG)</a></p> <p>?? <a href="#">Data Transport</a></p>

<a href="#">Information Systems and Performance</a>	<p>?? <a href="#">Discovery and Monitoring Event Description (DAMED-WG)</a></p> <p>?? <a href="#">Network Measurement (NM-WG)</a></p> <p>?? Grid Information Retrieval (GIR-WG)</p> <p>?? <a href="#">CIM based Grid Schema (CGS-WG)</a></p>	<p>?? <a href="#">Relational Grid Information Services (RGIS-RG)</a></p> <p>?? <a href="#">Grid Benchmarking (GB-RG)</a></p>
<a href="#">Peer-to-Peer</a>	<p>?? <a href="#">Previous activities of the Peer to Peer Working Group</a></p>	<p>?? <a href="#">Appliance Aggregation</a></p> <p>?? <a href="#">OGSA-P2P-Security</a></p>
<a href="#">Scheduling and Resource Management</a>	<p>?? <a href="#">Scheduling Attributes (SA-WG)</a></p> <p>?? <a href="#">Scheduling Dictionary (SD-WG)</a></p> <p>?? <a href="#">Distributed Resource Management Application API (DRMAA-WG)</a></p> <p>?? <a href="#">Grid Resource Allocation Agreement Protocol (GRAAP-WG)</a></p> <p>?? <a href="#">OGSA Resource Usage Service (RUS-WG)</a></p> <p>?? <a href="#">Grid Economic Services Architecture (GESAWG)</a></p> <p>?? <a href="#">Usage Record (UR-WG)</a></p>	
<a href="#">Security</a>	<p>?? <a href="#">Grid Security Infrastructure (GSI-WG)</a></p> <p>?? <a href="#">Grid Certificate Policy (GCP-WG)</a></p> <p>?? <a href="#">Open Grid Service Architecture Security</a></p>	<p>?? <a href="#">Site Authentication, Authorization, and Accounting Requirements RG</a></p>

	<a href="#">Architecture Security (OGSA-SEC)</a> ?? <a href="#">CA Ops</a> ?? <a href="#">Authorization Frameworks and Mechanisms WG</a>	
--	--	--

(出典：GGFホームページより)

## 二、グリッドプロジェクト

グリッドコンピューティングは、巨大科学を進める上で重要な基盤となるだけでなく、ITとバイオ、ナノとの融合領域において重要な基盤となるであろうという考え方から欧米では積極的に国家プロジェクトが実施されている。日本、アジアにおいてもいくつかのプロジェクトが実施されてきてはいるが、その勢いは欧米に比べると多少弱く感じる。ここで、米国、欧州、日本、アジアにおいての取り組みを紹介していく。

### 1. 米国のプロジェクト

スーパーコンピュータをネットワークを介して利用して計算しようというグリッドコンピューティングの根幹となる研究は米国から始まっており、米国にはグリッドコンピューティング技術開発と実証システム開発の過去の実績がある。さらに、グリッドコンピューティング技術開発、大規模な実証システム、グリッド応用開発が積極的に進められている。注目すべきものとしては、全米科学財団(NSF)がサイバーインフラの一環として整備する計画のTera Gridプロジェクトがある。テラバイト級の大容量データ処理の研究が行われており、全米の4つの計算センターが参加し、40Gbpsの高速ネットワークで接続される計画である。この内、SDSC(San Diego Supercomputer Center)とNCSA(National Computational Science Alliance)という全米1位、2位のスーパーコンピューティングセンターに接続される。もう一つ注目すべきものとしては、BIRN(Biomedical Informatics Research Network)プロジェクトがある。これは、カリフォルニア大学を中心として全米10カ所の医療研究機関によって、アルツハイマー病などの研究のため、脳の様子を可視化してデータを共有し、研究を活発化させようという目的でグリッドコンピューティングを開発している。これには、医療の現場を含め多くの研究者が参加している。他には、核実験の代わりに超高速スーパーコンピュータを構築して数値シミュレーションで解決するという、エネルギー省(DOE)のASCI(Accelerated Strategic Computing Initiative)プロジェクトやNASAのInformation Power Gridプロジェクトなどがある。

## 2．欧州のプロジェクト

欧州の中でも、英国が特にグリッドコンピューティングの研究をしている。英国では、2000年度からe-Scienceプロジェクトを開始している。このプロジェクトは科学研究をITにより推進加速させることにあり、グリッドコンピューティング技術をベースにして、高性能コンピュータや大規模実験装置の共同利用、世界的に分散した大規模データの共同利用を実現するプロジェクトを精力的に進めている。グリッド基盤として、英国9カ所に国立e-Scienceセンターを設立し、高性能コンピュータを高速ネットワークにより接続したグリッドコンピューティングを構築している。このグリッドコンピューティング上に作られる応用は、高エネルギー物理、ゲノム、バイオ、たんぱく質構造解析、医療・保健、環境、気候、宇宙、科学・材料などたくさんある。

EUとしての代表的なプロジェクトとしては、大規模データを取り扱うEU Data Gridがある。欧州原子核共同研究機関(CERN)が中心となって、米国、日本とも連携を取って研究が進められている。

欧州の英国以外の各国のプロジェクトとしては、ドイツのUnicore、イタリアのINFN Grid (Italian National Institute for Research in Nuclear Physics)、オランダのGrid-based Virtual Laboratory Amsterdam (VLAM)などのプロジェクトがある。

## 3．日本のプロジェクト

国内では、2000年から開始された、大学や国立研究所の計算センターを高速ネットワークによって接続するスーパーSINETプロジェクトや2001年から開始された旧科学技術庁の研究所を高速ネットワークで接続するITBL (IT-Based Laboratory) プロジェクトがある。ITBLプロジェクトでは、日本原子力研究所と理化学研究所を中心に仮想研究所の構築を目指しており、スーパーコンピュータをネットワークで接続し共同研究の環境を作っている。これらは、グリッドコンピューティング構築のための重要な基盤となっている。

他の主なプロジェクトとしては、2001年度補正予算にて、産業総合研究所グリッド研究センター設立、東京工業大学キャンパスグリッド構築。2002年度、大阪大学を中心としたバイオグリッドプロジェクトが開始され、また、東大、北陸先端大、徳島大、理化学研究所などによるバイオインフォマティクス用グリッド構築が開始されている。2003年度には世界最高水準のグリッド構築を目指したナショナル・リサーチグリッド・イニシアティブやグリッドコンピューティングによる柔軟なサービスの提供をするソフトウェアを開発するビジネスグリッドプロジェクトなどが行われる。

## 4．アジアのプロジェクト

アジア地区では、韓国、中国を初め、台湾、シンガポールなどでグリッドコンピューテ



ィング研究が進められている。米国、オーストラリアを含めたアジア太平洋地区のグリッド活動を推進する母胎として、アジア太平洋グリッド A P - G r i d とアジア太平洋 G R I D 応用会議 ( P R A G M A : Pacific - Rim Application & Grid - Middleware Assembly ) があり、日本が先頭に立ってグリッド活動を推進してきている。また、A P A N ( Asia - Pacific Advanced Network ) の活動によって、グリッドコンピューティングの重要な基盤なる高速ネットワークが構築されてきている。主なプロジェクトとしては、韓国の Korea Grid Initiative や中国の China Computational Grid などがある。

### 三、グリッド技術の種類

#### メタコンピューティング

ネットワーク上に分散・配置されたスーパーコンピュータなどの高性能コンピュータを複数台同時に使用して、一台では得られない大規模な計算を行う。仮想的な巨大コンピュータを実現させようとする。

#### データグリッド

一つの場所には格納しきれないような大規模なデータや各地に分散配置されたデータをネットワークを介して遠隔地からアクセス可能とするものであり、高エネルギー実験装置、実大規模 3 次元振動台、高感度望遠鏡等の大型実験装置から得られるデータを処理したりする。

例えば、高エネルギー物理学の実験で陽子同士を衝突させてその時に発生する微粒子を解析することで宇宙の始まりを解明しようという実験が行われている。スイスのジュネーブにある欧州原子核共同研究機関 ( C E R N ) の巨大加速器の実験では、500 研究機関の研究者 7000 名が年間に数億回の実験を行っている。このような、高エネルギー物理の実験データを解析し研究者間で共有するためのグリッドを欧州の E U D a t a G r i d やアメリカの Grid Physics Network ( G r i P h y N ) プロジェクトで開発中である。この他、複数の宇宙観測所にて得られるデータから全宇宙を記述しようとする研究でもグリッドが使われており、データグリッドにより巨大科学を効率良く推進することができる。

#### 仮想計算センター

ある組織内の複数の計算サーバをネットワークに接続し、計算サーバのタイプを意識せずに必要な時に必要なだけ計算パワーを利用できるような環境を提供する。いわば、仮想コンピューティングセンターを作ろうという試みである。例としては、東工大キャンパスグリッドがある。キャンパス内に分散設置された P C クラスタと 25 テラバイトのストレージをギガビットクラスの高速度キャンパスネットワークを介して利用

する。この他、企業内において、計算センターを企業内イントラネットで結合してグリッドを構築している例もある。

計算能力を利用するだけでなく、スーパーコンピュータや高性能PCクラスタなどに整備された高精度・高機能ライブラリ、アプリケーションソフトなどがある短時間だけ利用するというようなGrid ASP (Application Service Provider) を構築するという考えもある。ある特定のパッケージ化されたアプリケーションをWeb画面から指定して、入力データを与えるだけで自動的に最適な計算機資源を選択し、結果を戻してくれる。

ここで、コストの概念を入れることで、スーパーコンピュータのように時間単価が高価だが高速なものとPCクラスタのように安価に計算を実行する選択が可能となる。それも、そういう要望を出せばそれに見合った計算機資源を自動的に割り振ってくれる。

Grid ASPは高速性が得られるだけでなく、高品質なソフトウェアの維持コストが低減すること、稀にしか用いられない関数ライブラリを使用できるといったメリットがある。

#### デスクトップコンピューティング

個人のPCには空き時間が結構ある。キーが打たれる、ネットワークからパケットが到着するなどのイベントが発生すると一瞬働きますが、ほとんど時間はイベントを待っているだけです。計算能力の9割が活かされていないとされている。一方で、計算能力はいくらあっても足りない、という非常に大きな問題を抱えている人達もいる。PCの余剰計算能力をかき集めて、大きな問題を抱えている人に提供しようというものの。このことから、ボランティアコンピューティングとも言われる。大きな問題を分割して、計算能力を余らせているPCに処理させます。この例としては、宇宙観測データ解析 (SETI@home: Search Extraterrestrial Intelligence at home) や癌、エイズ、白血病の新薬開発 (Parabon's Compute - against - Cancer、Entropia's FightAidsAtHome、United Devices' Cancer Research Project、Intel's Philanthropic Peer - to - Peer Program)、暗号方式の鍵探索のdistributed.netなどがある。SETI@homeでは、400万人のボランティアがPCの空き時間を提供しており、40テラフロップス性能のスーパーコンピュータと同等の計算能力を得ている。

また、別に能力が余っているわけではないけれど、大きな問題を扱うために組織内の多量のPCを束ねようという種類の取り組みもある。これを最近では、企業が取り込んで、企業内のPCのパワーを利用して受発注業務を実行させている例もある。

#### Access Grid

Access Gridとは、遠隔地の共同研究者と同じ画面、同じ計算結果を共有して、共同研究を円滑にスピーディに進める環境を提供するものである。基本部分は大規模ビデオ会議システムということになり、高速なネットワークを介したマルチキャスト通信を利用して、従来のテレビ会議システムに比べ、画質もよくファイル共有など共同作業を効率よく進められるのである。

Access Gridの基本部分はビデオ会議システムなので、ここで一般のビデオ会議システムと比較してみたいと思う。

ビデオ会議システムはTCP/IP上ではH.323というITU(International Telecommunication Union)勧告に準拠している。ビデオ会議の手段としてだけ考えた場合、これらH.323端末に対するAccess Gridの利点は、IPマルチキャストを利用するので拠点数が増えても映像、音声といったメディアの質が損なわれないということである。H.323は基本的に一対一のビデオ会議のための規格であり、端末間に呼び出し側と応答側という関係を仮定している。そのため、複数の拠点で会議を行うためにはMCU(多地点接続装置)が必要となる。多拠点会議では、全H.323端末がMCUを相手にユニキャストで通信し、MCUは、全端末からの映像と音声をミックスしてその結果を各端末に送る。この際に各端末がMCUから受け取る映像は、あくまで一画面分に限られてしまい、画面を分割したり、各拠点からの映像を適当なタイミングで切り替えることで一画面に収める。

一方、Access Gridは当初より多拠点を指向しており、IPマルチキャストの利用を前提としている。どの拠点も対等であり、どの拠点も全拠点からの完全な映像と音声を受け取ることが可能であり、各拠点からの映像はPCのウィンドウに表示されるので、どういった大きさ、配置で表示するかは各拠点の自由である。

しかし、Access Gridを構築するには、難問がある。それは、IPマルチキャストを利用できる環境の構築である。Access Gridは比較的広い帯域幅を必要とする。それなりの帯域幅のあるネットワーク接続を確保し、接続相手を見つけ、ルータの設定を行わねばならない。さらに、拠点数が多ければ多いほど必要な帯域は増えてしまうのである。

#### 四、民間企業の動向

近年、先に述べたデスクトップコンピューティング、大規模分散処理をビジネスに応用しようという動きが出始めている。ビジネスモデルとしては、集めた計算能力を大きな問題(新薬の開発、遺伝子配列情報解析、経済のシミュレーション等)を抱えている企業に売る、顧客企業が所有している大量のPCを束ねて大規模計算をさせるためのサポートやソフトウェアを商品にしている。

民間企業がグリッドを導入した時の利点としては、1.個人の生産性を高め、計算リソ

ースへのアクセスを向上する。2．企業の生産性を高め、同じ期間内により多くの製品を産出できる。3．反復的な開発プロセスを支援し、製品の質を向上する。4．複数のマシンで同時に開発を進め、市場投入をスピードアップさせる。5．ネットワークと計算リソースの最適稼働、可用性、スケーラビリティ、信頼性の向上を通じて投資利益率を改善する。というものがある。ここで、投資利益率の改善とあるが、これは、企業がスーパーコンピュータなどの高性能コンピュータを導入する際、ハードウェアの性能を業務においての最大負荷を想定してどのような高性能コンピュータを導入するか決めるが、そのような負荷は頻繁には発生しないので、どうしてもオーバースペックなものを導入してしまうのである。また、コンピュータの発達は早いものなので、多額の費用をかけて導入したコンピュータもすぐに時代遅れなものとなってしまう。しかし、あまりにも多額なのですぐに買い換えるということもできないので、どうしても割に合わないものになってしまうのである。

グリッドコンピューティングに特に積極的な企業は、IBM、サン・マイクロシステムズ、プラットフォームコンピューティング、HP、AVAKI、United Devices、Entropia社などがある。これらの企業は、GGF（Global Grid Forum）においても活発に活動しており、企業内の情報資源を束ねる技術とそのソフトウェアを商品としてビジネスを行っている。

この中の、Entropia社はFight AIDS@Homeという新薬開発のプロジェクトを運営しており、AIDSの新薬開発に協力してくださいと呼びかけ、参加者を募って、多くのコンピュータが参加するにいたっている。同様に、United Devices社は、癌の新薬開発への協力を呼びかけている。これらはどちらも、ボランティアとして参加者を募っている。

特にグリッドコンピューティングに力を入れているIBMが高性能コンピューティング・サービスを開始することを発表した。このサービスでは、顧客は巨大なコンピュータ群に接続し、そのときの仕事に必要な分だけのプロセッシング・パワーを受け取って、利用した分についてだけ料金を支払うというものである。IBMが顧客として想定しているのは、大きなプロセッシング・パワーが時々必要になるが、専用のスーパーコンピュータを所有することで発生する固定費や管理の手間を省きたいと考えている企業である。このサービスには、もう早くも顧客がついている。

このIBMの事業が成功すれば、他の企業も乗ってくると思われる。まだ始まったばかりなので、どうなるかはわからないが、今後に期待したいと思う。

## 第六章 結論

グリッドコンピューティングは、ネットワーク上の計算資源を一つのコンピュータと見立てて、容易にかつ効率的に利用するための技術であり、次世代インターネット利用技術として注目を浴びている。また、コンピュータ、大規模データベース、大規模実験装置を遠隔地から利用することもできるようになるので、これらの設備を共同利用することに研究効率向上が見込まれるだろう。

しかし、グリッドコンピューティング研究の歴史はまだ浅く、それゆえに、実現のための基盤、ミドルウェア、ソフトウェアが完全に整っているわけではない。さらに解決すべき問題点がまだまだたくさん存在する。IT分野の進歩がいくら早いとはいえ、グリッドコンピューティング実現はもう少し先になるであろう。

歴史は浅いが、グリッドコンピューティングの研究をしている人は多い。GGFというグリッドコンピューティングの標準化と普及を推進するフォーラムには、数多くの研究者が参加しているし、その数は回をおうごとに増している。各国でも国レベル、研究所レベル、大学レベルでグリッドプロジェクトが立ち上がっている。民間レベルでも民間企業のいくつかがグリッドコンピューティング開発に関わったり、現時点でのグリッド技術をビジネスとして売り出しているところもある。

グリッドコンピューティングは、まだ問題点を抱えながらも着実に実現のために動いているが、グリッドコンピューティングが実現することにより共同研究が活発化するであろう。ITとバイオ、ナノとの融合研究領域においては、データ、解析プログラムの共同開発、共同利用が可能となる。複数の異なる専門分野にまたがる境界領域にて共同研究が活性化すれば、新しい研究分野が誕生する可能性がある。国と民間企業の共同研究もさらに活発化するかもしれない。

従来の並列処理環境とグリッドコンピューティングとは大きく異なる。これまでのものは、主に同じプロセッサを用いて、外乱の少ないネットワークで構成されていた。しかし、グリッドコンピューティングでの並列処理環境は、計算資源として用いられるプロセッサが異なるヘテロなシステムであり、高速のネットワークが利用できるが、そのバンド幅は変動する。基本的にヘテロで動的な環境を対象としなくてはならないが、自分の手元に大量の計算資源を得ることができる。本来、コンピュータシステムに期待されている役割はユーザの問題を素早く解決してくれることである。プログラムを書くか書かないかに関わらず、また、どこで計算されるかに関わらず、なるべく早く、わかりやすく答えを出してくれることであり、その際、ネットワークについて意識することない方がいいだろう。グリッドコンピューティングの究極の姿は、快適で有用な計算環境を提供することであると

考える。

しかし、グリッドコンピューティングが実現したからと言って、個人レベルではグリッドコンピューティングのサービスを受けられるかは不明だし、受けられなかったとしても特に問題はないのである。いくらインターネットが普及したと言ってもまだそれを使わずに生活をしている人たちは大勢いるように、実現しても我々の生活には特に変化はないであろう。だが、社会基盤レベルでのグリッドコンピューティングという位置づけをした時に、直接はサービスを受けないにしろ、間接的にサービスを受けるであろうと考える。例えば、交通機関が発達した時に、それを使ってない人はそのサービスを受けてないかというと、そうではなく、交通機関の発達に伴い、物流が便利になり、街が活性化するというようなものと考えるとよいと思う。

グリッドコンピューティングの分散の必然性やなんのためにこうした環境が必要なのか、だれもコンピュータを簡単に貸してくれないのではないかと、という批判的意見もある。これは、大規模な科学技術計算や実際の産業活動においての新製品の開発において、だんだんと高性能な計算機が必要となってきたのである。しかし、スーパーコンピュータや高性能計算機は一朝一夕には導入できないうえに、第五章の四で述べたように価格性能比がよくないのである。効率化・合理化が叫ばれる今の時代においては、重要なことなのである。また、これからますます大規模計算能力が必要となってくるので、前述したIBMのように計算能力を提供するサービスもこれからは増えていき需要が高まれば、より高品位なサービスを提供しようと努力するだろう。

実際に、グリッドコンピューティングが実現するかといたら、それはむずかしいだろう。現時点では、広域ネット上でのコンピューティングは、技術的にも管理上の問題（セキュリティ、スケジューリング等）的にも実現にはまだまだという段階である。「ユーザがコンピュータの所在などを意識することなく、いつでも必要なだけ計算資源を利用できる環境を提供」というグリッドコンピューティングの目標を現実のものにするのは、時間がかかるだろう。むしろ、応用分野であるデータグリッド、仮想計算センター、デスクトップコンピューティング、Access Gridの方がすでに実現の一手手前まできている。よって、本来のグリッドコンピューティングがITインフラとなるまで身近なものになるとは思えない。ただ、計算機科学の集大成としてのインフラストラクチャの実現という意味で、計算機科学に携わる人々の期待が高いのも事実であるし、Webが新しいITの世界を開いたように、グリッドコンピューティングでも着実にシステム開発、応用の開拓をすすめることによって、これまでとは違う新しい計算分野、応用、そして研究開発の世界が開けるのではないだろうか。

参考文献

Globus <http://www.globus.org>

GGF <http://www.gridforum.org>

Ninf <http://ninf.apgrid.org>

『コンピュータソフトウェア』 Vol.17 No.4 日本ソフトウェア科学会

「次世代 HPC インフラとしての GRID」日本数値流体力学会誌 第 9 巻第 4 号 2001 年

『日経バイオビジネス』 2002 年 5 月号 日経 BP 社

関口智嗣 『Computer Today』 サイエンス社 2001 年 1 月号 No.95

川戸正裕、蒲池恒彦、妹尾義樹、田中良夫、関口智嗣「Globus Toolkit の NEC sx-4/5 への移植および評価」情報処理学会研究報告 Vol.2001-HPC-88 No.5

須藤一幸、田中良夫、小松弘幸、松岡聡、南里豪志、岡村耕二、関口智嗣「Access Grid の構築と Grid 上での国際会議」情報処理学会研究報告 Vol.2002-HPC-89 No.6