

グリッド・コンピューティングの技術的背景

The technical background to grid computing



日本アイ・ビー・エム株式会社
グリッド・ビジネス事業部
技術理事

関 孝則

Takanori Seki

Distinguished Engineer
Grid Computing Business
IBM Japan, Ltd.

インターネットは、「ネットワークの共有」「メッセージの共有」「情報の共有」「アプリケーションの共有」...と、次々と「共有」の領域を広げてきましたが「コンピューター資源の共有」という概念はありませんでした。それを実現するのがグリッド・コンピューティングです。グリッド・コンピューティングは「プロセッシング」「データ」「弾力性」「オンデマンド」などの幅広い適用領域を持ちますが、技術面からいえば、組織をまたがった異機種混在環境における、インターネット上の分散コンピューティング技術です。オープン・ソースのミドルウェアGlobus Toolkitとその新しいアーキテクチャーOGSA(Open Grid Services Architecture)によって、グリッド化されるIBMのサーバーおよびミドルウェア群は、その中核となっていくでしょう。各地に散在するマルチベンダー、マルチプラットフォームのリソースの差異は、このミドルウェア層で吸収されることになり、通信機能、リソース情報の提供/収集機能のほか、リソース割り当て・予約・監視・制御などのリソース管理、スケジュール管理を行うための機能が提供されます。このほかミドルウェアに欠かせないものとして、ユーザー認証などのセキュリティー管理、オンデマンド時代に重要となる課金管理などがあります。グリッド・コンピューティングの要素技術そのものは、それほど目新しいものではないかもしれませんが、では、何が画期的か。それは、世の中のシステム環境の制約条件に真正面から取り組み、解決する技術体系だからです。技術とは世の中の「何か」を解くためのものだとすれば、グリッド・コンピューティングは間違いなく、「現実解」を導く本物の技術です。

The Internet has been gradually expanding in scope to include more and more fields of "sharing": sharing of networks, sharing of messages, sharing of information, and sharing of applications. The idea of sharing of computer resources has not hitherto existed, but it has at last been realized through grid computing. Grid computing has a wide range of applications including processing, data, flexibility and on-demand. From the technical angle it refers to the technology of decentralized computing on the Internet in a mixed environment incorporating different machines on a horizontal, cross-organizational basis. The core of this technology is occupied by gridded IBM services and groups of middleware involving the Globus Toolkit open-source middleware and its new architecture, OGSA (Open Grid Services Architecture). The difference in multi-vendor and multi-platform resources scattered over many different places is absorbed by this middleware stratum and, in addition to communications functions and functions for presenting and gathering resources information, functions are also present for resources management including allocation, reservation, supervision and control of resources and schedule management. Security management in forms such as user verification is also indispensable for middleware. Management of levying of charges is also important. The factor technology employed in grid computing is in itself not that new. Well then, what's so revolutionary about it? It's because grid computing is a technical system that tackles head-on and provides solutions to the conditions that restrict system environments in the world at large. Assuming that technology is all about solving specific problems in the world at large, then grid computing is authentic technology that will lead to the solution of real problems.

「共有する」技術の進展

インターネットの登場と普及は、コンピューター環境を革新的に変えてきました(図1)。

インターネットの核となる技術は通信プロトコルのTCP/IPで、これにより全世界のネットワーク同士が結ばれ、「ネットワークの共有」が実現しました。やがてインターネット上でのEメールが普及し、コミュニケーション面で「メッセージの共有」が当たり前のようにになりました。

インフォメーション面では、WWW(World Wide Web)が登場。WWWはインターネット上に分散する情報を互いに結び付け、ブラウザを使ってマルチメディアの情報にアクセスする環境をつくりました。すなわち、「情報の共有」です。

近年ではネットワーク技術の著しい進化とWebサービスの充実により、「アプリケーションの共有」も現実のものとなっています。

さて、このようにインターネットは「共有」の領域を急速に増大させてきたわけですが、いまだに共有されていないものがありました。それがコンピューター資源です。

グリッド・コンピューティングは、高速ネットワークの下で多数の異なる管理下の異機種コンピューター群を接続共有し、その上に高速、巨大、かつ動的で仮想的な、プロセッサ、ストレージ、アプリケーション環境を実現しようとするもの。これにより、「コンピューター資源の共有」の世界が開けます。インターネットという巨大な情報プールが巨大なリソース・プールにもなるわけです。

以上が「共有」という概念から見たインターネットの進展ですが、インターネットをビジネスに活用するe-ビジネスも今後、グリッド・コンピューティングの方向へ進んでいくのは間違いないで

しょう。インターネットに代わって(正確に言えば、インターネットが進化した形で) The Gridという言葉が一般化することも考えられます。

Webサービスとの融合

グリッド・コンピューティングを技術面からいえば、インターネット上の異機種混在環境における分散コンピューティング技術です。

分散処理は、中央のホスト・コンピューターで行っていた処理を、ネットワーク化された複数のコンピューターで分担して行うことです。

分散コンピューティングそのものは、既に世の中で行われています。しかし、分散コンピューティングは、社内や組織内で使われ、その設定や運用スケジュールなどが一元的に集中管理されているケースがほとんどです。その管理をはじめ、業務の変更に伴うアプリケーションの開発には大きな労力が必要でした。

これに対して、それぞれのコンピューターによって管理やセキュリティが異なることを前提に、「非集中化」コンピューティングの考え方を採用し、インターネットを介したアプリケーションの企業間取引を実現するための仕組み(ソフトウェアとサービス)として開発されたのがWebサービスです。

Webサービスは、XMLという文書の論理構造を記述する言語が技術の核になっています。XMLは、タグがデータ内容の項目となっており、そのタグの要素を自分で定義できます。また、その際のデータ内容はアプリケーション・プログラムやコンポーネントであってもいいという特徴があります。

そこで、XML技術をベースに、アプリケーションをやり取りするための通信プロトコルとしてSOAP、情報を記述するための形式としてWSDL、サービスのディレクトリー(目録)の規格としてUDDIが定められ、研究開発が進められてきました。

その普及には、相互運用されるアプリケーションのバージョン管理、ライフ・タイムの管理など、越えなければならない技術的なハードルが幾つか存在します。しかし、Webサービスを企業のビジネス・プロセスに統合することにより、開発期間や運用コストは大幅に短縮・低減させることができ、ニーズは大きなものがあります。

Webサービスの究極の目的は、コンピューティングが電気や水のように、サービスとして安定的に供給されることです。実はこの



図1. グリッド・コンピューティングの背景

ことは、グリッド・コンピューティングの目指すゴールと同じです。

そこでIBMでは、2002年2月、Webサービス技術をグリッド・コンピューティングに融合・進化させていくことを目的の一つとして、オープン・ソース・プロジェクトGlobusとともにグリッド・コンピューティングの標準化団体GGF(Global Grid Forum)にグリッドのオープンなアーキテクチャーの標準化を提唱しました。それがOGSA(Open Grid Services Architecture)です。

今後は、このOGSAをベースとしたグリッド・コンピューティングが、異機種混在の分散コンピューティング環境におけるシステムの統合化と相互運用性のためのインフラストラクチャーになっていくはずで

グリッド技術の適用

グリッド・コンピューティングは、幅広い適用領域とビジョンを持った技術です。

ここでは、その適用を「プロセッシング」「データ」「弾力性」「オンデマンド」に分け、その概要とアプローチ法を紹介します(図2)

(1) プロセッシングのグリッド

プロセッシングのグリッドは、現時点ではグリッド・コンピューティングの代表的なもので、分散環境下での空きCPUを集めたり、サーバーを並列利用したりして、プロセッサに膨大な計算を

させます。コンピューター資源の有効利用とともに、巨大なコンピューター・パワーが安価に、かつ必要時に動的に利用できる価値がもたらされます。各企業では、まず分散ジョブ・スケジューラーや、PCグリッド製品で実行可能なアプリケーションを対象にアプローチするのが有効と思われます。

(2) データのグリッド

大量のデータが分散している環境においても、必要なデータを必要なときにユーザーに透過的に供給するのがデータのグリッドです。組織を超えた大量のデータの共有と、それによる協調作業(コラボレーション)の実現が価値となります。今日のデータベース技術は必ずしもグリッド・コンピューティングに対応しているとはいえ、データの共時性などが強く求められるビジネスへの適用においては幾つかのブレークスルーが必要ですが、急速に進展するストレージの管理技術がその課題を解いてくれるはずで

(3) 弾力性

不測の事態に対しても、動的に資源を割り当てたり、回復したりして、サービス品質を維持するのが弾力性のグリッドです。高いサービス品質をe-ビジネスで実現します。その機能は主として後述するグリッド・コンピューティングのモジュールウェアが担います。

(4) オンデマンド

企業アプリケーションにおいて必要となる計算リソースがオンデマンドで入手され、実行されるのがオンデマンドのグリッドです。計算リソースなどのコンピューター・パワーをユーティリティーのように利用できるようになります。それが社会的なインフラストラクチャーとして定着するのは数年先のことと思われますが、企業にあっては将来のユーティリティー化に備え、ASP(Application Service Provider)やアウトソーシングの活用によって、計算リソースの分離体制をいち早く整えておくことが大切になります。

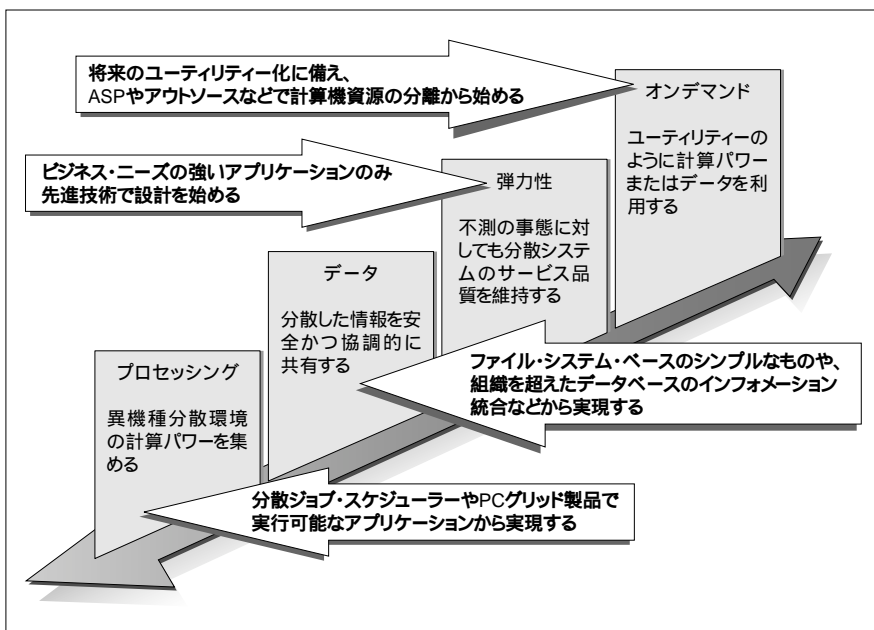


図2. グリッド・技術適用のステップ 現在のアプローチ

IBMにおけるグリッドの製品化

IBMは、グリッド・コンピューティングの技術はオープン・スタンダードであるべきだと認識し、スタンダード確立のための研究開発を世界中の大学・研究機関・他ベンダーなどと共同で進めています。まずグリッド・コンピューティングの市場を開拓し、その中でグリッド対応の製品化を進めながら、利用方法のコンサルティングなどに競争力を発揮したいと考えています。

図3は、IBM製品のグリッド化の概要を示したものです。

ここで中核をなすのは、OGSAに準拠したミドルウェアで、IBM製品では、WebSphere®、Tivoli®、DB2®などが該当します。

ミドルウェアでは、通信機能、リソース情報の提供・収集機能のほか、リソース割り当て・予約・監視・制御などのリソース管理、スケジュール管理を行うための機能が提供されます。各地に存在するマルチベンダー、マルチプラットフォームのリソースの差異は、このミドルウェア層で吸収されることになります。

このほかミドルウェアに欠かせないものに、ユーザー認証などのセキュリティー管理があります。オンデマンド時代には、課金管理も重要になります。

OGSAを実現するオープン・ソースのミドルウェアがGlobus Toolkitです。グリッドにおけるアプリケーション開発や運用に必要な基本要素技術を実装したソフトウェア・ツール群です。WebSphereを中心に、IBMのミドルウェア群はこのGlobus Toolkitに2003年以降、実装対応していきます。

ハードウェアの領域でもOGSA対応のほか、自律管理機能を持つサーバーやストレージの開発を進めています。すべての@serverには今後Globus Toolkitを実装し、ストレージの領域では、グリッド構築に欠かせないIPベースの製品の開発、Storage Tank製品での自律機能の強化を進めています。サーバーやストレージが自らを自律的に管理するオートミック・コンピューティングは、グリッド・コンピューティングの普及を補完する重要な技術です。

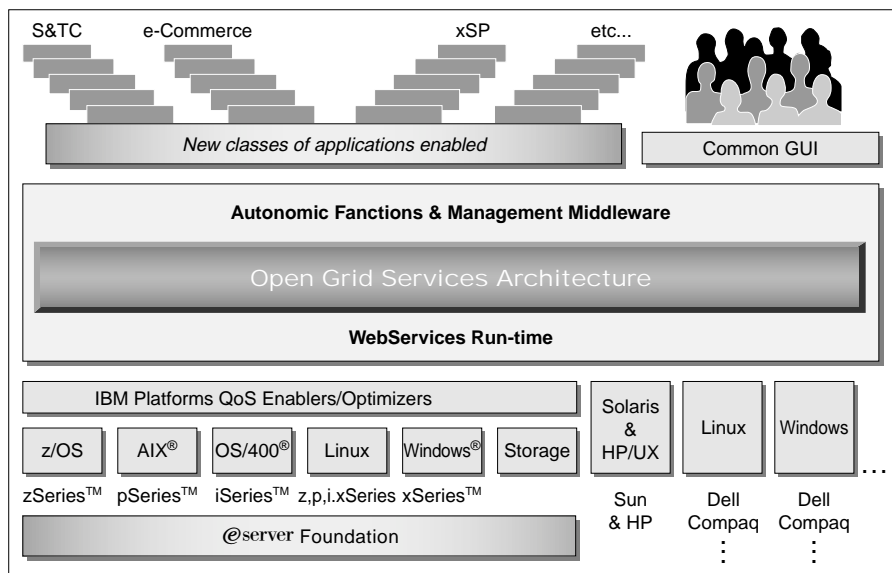


図3 .OGSAベースのグリッドのアーキテクチャー

「現実解」の技術

グリッド・コンピューティングの要素技術そのものは、それほど目新しいものではないかもしれませんが、では、何が画期的か。それは、世の中のシステム環境の制約条件に真正面から取り組み、解決する技術体系だからです。

現実のシステムには、「異機種混在が当たり前になっている」「リソースの分散は加速するばかりだ」「企業ごとにシステムの管理体系が違う」といった問題点があり、センターで集中管理する方法では、これらがシステム構築の大きな制約要件になっています。その上で、「データやアプリケーションをもっとダイナミックに扱いたい」「CPUの使っていない領域や時間が多く、もったいない。これをなんとかできないか」といった実際的なニーズもあります。

これらの制約要件をクリアし、ニーズにこたえるシステムを、シンプルな形で実現するのがグリッド・コンピューティングです。

技術とは、世の中の「何か」を解くためのものです。例えば経営者が、次々と登場してくる技術が本物かどうかを見極めるには、その技術が経営課題にどうかかわってくるかを考えれば分かります。技術が現実フィットしないのであれば、それは技術ではありません。単なる理論です。

その意味でグリッド・コンピューティングは、「現実解」を導く本物の技術です。