

# 次世代ネットワーキングサービスの課題と方向性

串田 高幸

日本アイ・ビー・エム株式会社東京基礎研究所

## 概要

インターネットにおいてネットワークから提供されるネットワークサービスは、エンドホスト間で、パケットのベストエフォートの転送であり、ネットワークのQoSの改善や保証、あるいは付加サービスの提供はない。現在、インターネットからのネットワークサービスを改善するため、この論文は、グリッドコンピューティングの仮想化や動的な資源の提供機能をもとにした新しいアーキテクチャフレームワークを提案している。この論文では、共通基盤となるグリッドネットワーキングインフラストラクチャーに関するアーキテクチャの特徴及び優位性について、具体的な例をもとに説明を行なう。

## Architecture and framework of the next generation networking infrastructure

Takayuki Kushida

IBM Research, Tokyo Research Laboratory

### Abstract

The Internet provides the best effort packet transmission services among end nodes, and this service can be only provided by the Internet. There is no guaranteed QoS improved network services and additional services from the network. This research theme aims to establish the architectural framework for providing the network service based on the Grid technology which can support the virtualized resources and dynamic resource allocation. This paper describes the architecture and framework of the Grid networking infrastructure with the example of its usage.

### 1. はじめに

インターネットでは、一般にIPプロトコルの転送によって、サブネットワークが相互接続されて、エンドホスト間では、TCPあるいはUDPを使って、論理的なデータの仮想ストリーミングのパイプあるいは、パケット転送のサービスをアプリケーションに対して提供している [1].

一方、グリッドコンピューティングは、ヘテロジニアスな分散環境において、資源の共有化を可能にするための技術である。例えば、科学技術において計算量が多い場合、分散されて設置されているノードで計算処理を行なうことによって短時間で結果を得ることができること、あるいは大量データの処理を短時間で終了する必要があるアプリケーションがあった場合に、分散してファイルの処理を行なうことによって、短時間で結果を得ることができるメ

リットがある。

この論文では、従来からのインターネットの環境でのサービスの向上を行なうため、グリッドコンピューティングの資源の仮想化及び処理の分散化の技術をもとにして、「グリッド・ネットワーキング・インフラストラクチャー」のアーキテクチャフレームワークを提案する。この論文では、次の「バックグラウンド」、「提案アーキテクチャ」、「議論」、「おわりに」の構成に分けて述べていく。まず、「バックグラウンド」では、インターネットとグリッドに関して、従来からの経緯及び現状について述べる。また、「提案するアーキテクチャ」では、この論文において提案するアーキテクチャ及びアーキテクチャを使った具体的なシナリオについて述べる。「議論」では、このアーキテクチャの問題点と今後の課題について述べる。最後に「まとめ」について述べる。

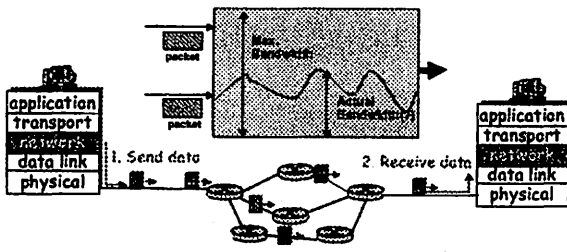


図1: パケット交換ネットワーク

## 2. バックグラウンド

この章では、バックグラウンドの技術として、インターネット (TCP/IP プロトコルに関連した技術) とグリッドコンピューティングの2つについて説明とそれぞれの特徴及び問題点について述べる。

### インターネット

一般に TCP/IP でユーザに対して提供されているサービスは、ベストエフォート型サービスであり、パケットが転送されている途中ルータにおいて輻輳が起これば、そこ時点で転送されていたパケットは棄却されてしまい、その結果、エンドホストにパケットは転送されないことになる。そのため、この損失したパケットをエンドホストの TCP によって再送することによって、最終的には、完全なデータとしてアプリケーションに転送される。

図1は、インターネットのもとなっているパケット交換ネットワークである。この図において、送信側のホストからは、パケットがネットワークに送信され、途中の経路では、パケットは経路情報によって受信ホストに転送されている。受信側のホストでは、ネットワークから転送されてきたパケットを受信する。このとき、ネットワークのリソース (それぞれのリンクのバンド幅やルータの能力) によって、パケットが転送される時間、パケットが損失率が異なるために、その結果、最大のバンド幅はあるが、実効バンド幅は、時間とともに常に変動している。このインターネットのパフォーマンスが動的に変化するの、ベストエフォート型ネットワークでの問題となっている。この問題に対処するため、Diffserv のようにネットワーク内部のルータにおいて、少しでも差別化できるサービスを行なうための技術が作られてきている [2]。また、一方では、Intserv として、保証された QoS を統合したサービスとしてを提供することを行なってきた [3]。しかし、これらのサービスを提供するためには、既存のインターネットに設置されたルータに、新しい機能を付

け加える必要があり、End-to-End まですべて新しい方式に統一する必要がある。

また、インターネットにおけるパフォーマンスの問題は、アプリケーションとして Web のコンテンツやストリーミングコンテンツを既存のインターネットにおいてサービスしようとする、顕著な問題である。そのため、既存のインターネットをそのまま使うことを想定すると、コンテンツをキャッシュするためのサーバ (Edge Servers) をネットワークの各所に設置して、実質的な速度を改善する必要がある。この方式は、既存のインターネットのアーキテクチャを改善せずにそのまま利用して、外部に接続するサーバによって付加的な価値をつけて、全体のサービスレベルを高くするというアプローチになる。このキャッシュサーバを利用することによって、ユーザは、本来のサーバと同じコンテンツをユーザからネットワーク上の近い場所にあるキャッシュサーバから転送することができることになり、その結果、改善された転送速度によって、ユーザにコンテンツを転送することが可能になる [4]。

一方、Clark らは、現在のインターネットのアーキテクチャに対して、新しいアーキテクチャ提案として Knowledge Plane と呼ばれる機能を新規に追加することを提案している [5]。この Knowledge Plane が対処する問題の例としては、欠陥の診断とその軽減、自動的な構成、オーバーレーネットワークに対するサポート、知識を使って拡張された侵入の検出がある。この研究では、まだアーキテクチャを作ることがはじまったばかりで、どのようにして Knowledge を表現して使用するか、また、拡張性をどのようにして持たせるか、あるいは Knowledge をどのようにして分散させるか、さらに、どのように経済的にみて Knowledge Plane をどのようにしてネットワークサービスプロバイダーに入れてもらうか、あるいは、悪意があつて信頼できないコンポーネントをどのように扱うか、ということについては、今後、研究して解いていく必要がある。このように、現在のインターネットを拡張して、新しいサービスを提供する研究については、現在も進められている。

### グリッドコンピューティング

「グリッド」という用語は、もともと 1990 年代半ばに、科学技術計算に対する分散コンピューティングを適用するために使われてはじめてのが、最初である。それ以来、グリッドに関するインフラストラクチャーが整備されてきている [6]。グリッドとい

う概念は、調整された資源共有、動的に複数組織の仮想組織における問題解決であり、このグリッド環境では、VO(Virtual Organization)と呼ばれる仮想的な組織が作られる。

グリッドコンピューティングは、一般にネットワーク上に分散しているヘテロジニアスな資源を仮想化して、この資源を有効に活用することを目的としている [7]。分散している資源のうち、特に計算リソース (CPU パワー) を仮想化して、特定のアプリケーションに対して集約して使用することは、グリッドコンピューティングにおいて大きな成功になっている。

例えば、計算によってシミュレーションを行なう場合、単一のコンピュータにおいて非常に高速で計算ができたとしても、複数のシナリオ (パラメータ) を計算するためには計算量が膨大になる。この場合、グリッドの環境を使って複数のコンピュータで同時にシナリオを計算することで、計算可能な総 CPU パワーを増大させることができる。この方法を使うと、実質的な処理能力を飛躍的に増大させることが可能になる。

また、分散されているリソースには、ディスクもある。分散環境に置かれたディスクのリソースを有効に使って、実質的にディスク容量を増加させるあるいは、ディスクの実質的な処理能力を増大させることも、グリッド環境を使って可能であり、これをデータグリッドと呼んでいる。また、ディスクへのアクセスを分散させることを、ファイルシステムとして行なうことができる。AFS(Andrew File System)、DFS(Distributed File System)、NFS(Network File System) に代表される分散ファイルシステムが遠隔に分散しているホストのファイルを読み出し、あるいは書き込みを行なうサービスをユーザに提供する。このサービスによって、すべてのデータ転送を伴わずに、仮想的に分散処理における転送を行なうことができる。

一般にグリッドコンピューティングの展開のステップは、図2のように大きくわけて、「イントラグリッド」、「エクストラグリッド」、「インターグリッド」の3つに分かれて進んでいくと考えられる。イントラグリッドは、社内のイントラネットワークにおいて構築されるグリッドであり、ネットワークとしては、主に LAN が使われ、また、セキュリティドメインとしても一箇所で行なうことが基本となるため、比較的簡単に導入して運用することができる。

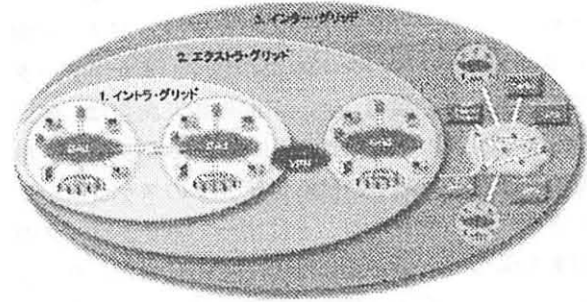


図2: イントラ、エクストラ、インターグリッドの展開

また、ネットワークとしてもバンド幅が十分にあって、また、遅延が少ないのでプログラムやデータの転送を短時間で行なうことができる。一方、エクストラグリッドは、図2のように、VPNによって組織の一部が地理上、別な場所にあるエクストラネットワークによって接続されたグリッドとして定義することができる。このエクストラネットワークは、LAN間接続をVPNで行なっていて、ファイバー接続であれば、最大 100Mbps 程度までの速度を出すことが可能である。また、セキュリティは、エクストラネットであれば、VPNで接続されており、イントラネットとほぼ同じポリシーで運用が可能であり、グリッドのノード間の速度が低速になることを除けば、イントラグリッドと同じ運用が可能になる。また、インターグリッドでは、グリッドノード間がインターネットで相互接続されており、ここに複数の組織間が接続されており、このネットワークのバンド幅は、イントラグリッドやエクストラネットに比べて小さくなると考えられる。また、接続している組織が異なるため、セキュリティも異なるポリシーになり、より高機能なセキュリティ機能が要求される。

#### 要求

ここまで、「インターネット」と「グリッドコンピューティング」について、説明してきたが、グリッドコンピューティングは、インターネットに対して、どのような影響があるかについてここでは、述べる。まず、グリッドコンピューティングが、インターネット上で動作すると、インターネット上でノード間のデータの転送が自動的に行なわれるようになるため、データの転送量が今までに比べて飛躍的に多くなる。しかし、一方でグリッドコンピューティングのノード間の相互のトラフィックの輻輳の調整は、

TCP によってのみによって行なわれている。そのため、インターグリッドのようにグリッドノード間のネットワークのバンド幅が小さく、また、データの転送量が多い場合、ノード間の相互調整は、個別の TCP ではなく、ノード間においてセッション単位での調整や、あるいは転送データのスケジューリングを入れて調整することによってネットワークの均衡を保つことが必要になってくる。また、グリッドコンピューティングにおいて、ジョブの投入やスケジューリングは、サーバとそれぞれのノードにおいて行なわれ、また、データをスケジューリングについては、バンド幅で調整するよりも、むしろ CPU やメモリーが空いているか、どうかによって調整される必要がある。

一方で、今までの研究では、グリッドコンピューティングは、インターネットをデータ転送のリンクとして扱っており、そこでの転送量あるいは転送方式に関する問題点について言及しており、それらの問題点を解決するための提案がなされている。しかし、グリッドコンピューティングの持っている機能を生かして、インターネットのインフラストラクチャーの一部として組み入れて、サービスを提供する共通基盤として考えられていない。

### 3. 提案アーキテクチャ

この研究では、グリッドコンピューティングの資源の仮想化を使って、分散環境においてサービスを仮想化して、ネットワークのサービスをユーザに提供するアーキテクチャとフレームワークを提案する。

図3は、コンピュータでの計算分野とネットワークの今までの経緯について示している。図3の左側は、もともとホストコンピュータに複数の端末が直接接続されていて使用していることを表わしている。次の世代にスーパーコンピュータが出てきて、ネットワークを使った遠隔ワークステーションが接続されている。また、左側の3つ目の図は、グリッドのノードがネットワークに相互接続されて分散コンピューティング環境を作っている構成である。この構成では、ノードが高速のネットワークに相互接続されて、資源の仮想化が行なわれ、大きな一つの資源として利用されている。

一方、図3の右側は、ネットワークの今までの経緯を表している。最初、それぞれの組織が個別にネットワークアーキテクチャを使っていた。そのとき、TCP/IP もまだ一般的ではなかった。その後、OSI が TCP/IP のアーキテクチャをさらに一般化した形

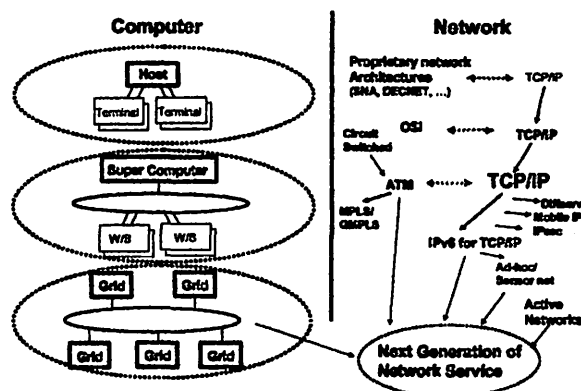


図3: グリッドコンピューティングとネットワークのバックグラウンド

で、ISOにより標準化され、アーキテクチャをもとにしたプロトコルが実証実験的に利用されたが、本格的な実用には至らなかった。また、TCP/IPは、これらとは異なるアーキテクチャとして、ARPANET, NSFNETをはじめとして研究教育機関で使われきて、相互接続するためのネットワークとして利用された。一方、データネットワークとして、音声やデータを統合してデジタル化して転送する方式として、小さいデータの単位であるセルをスイッチによって転送するATMがITU-Tで標準化され、使われてきた。この技術によって、統合された新しいサービスを取り入れることができたが、しかし、高速転送で多くのセルをサービスのグレードを分けて転送するため、ハードウェアスイッチに対するコストがかかり、また、エンドツーエンドでATMを使わないと、高機能なサービスを享受できないために大きな展開が起こらなかった。一方で、TCP/IPは、多くのISPが、商用接続サービスを開始して、インフラストラクチャーになってきた。しかし、TCP/IPは、ベストエフォートサービスしか提供しないため、ネットワーク内部により高機能なサービスをいれる必要性が出てきた。そのため、Diffserv, Mobile IP, IPsecのような個別の技術の標準化がIETFでなされている。また、ワイヤレスネットワークや小型の機器の発達によってAd-hocネットワーク, Sensorネットワークのような新しいネットワークに関連する技術も研究されている。さらに、IPv4のアドレスの枯渇問題や限定されているヘッダー機能のため、新しい拡張バージョンとしてIPv6がIETFで標準化され、現在、ISPにおいても一般のサービスがはじまっている。

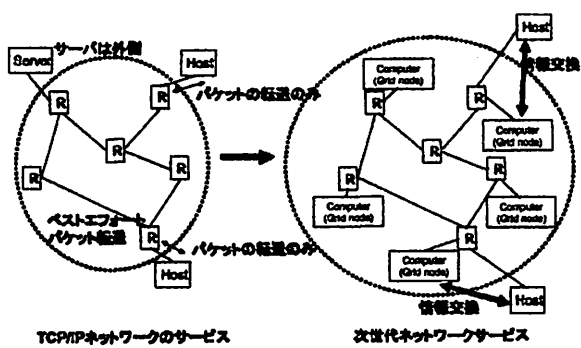


図4: 現在TCP/IPの構成とグリッドネットワークの構成

このような背景から、この論文では、次世代のネットワークサービスを行なうために、グリッドコンピューティングの技術を使った既存のネットワークサービスや新規サービスをユーザに提供するため、次世代ネットワークアーキテクチャを提案する。図4は、従来のTCP/IPネットワークとこの論文で提案しているグリッドを利用したネットワークサービスの2つを表している。図4の左側は、TCP/IPネットワークサービスであるが、基本的にはネットワークは、ベストエフォートでパケット転送をしており、エンドのホストは、ネットワーク内部ではデータパケット転送だけを行なっているだけである。また、ネットワークで行なっている主要なサービス(DNSとメールのような基本サービス)は、ネットワークの外側にあるサーバによって行なわれており、通常、そのサーバは、クライアントからのリクエストを通常のデータ転送と同じ方法で受け付けている。

一方、図4の右側は、グリッドノードがネットワークの内部に位置して、それぞれのグリッドノードが協調しながら、ネットワークのサービスをエンドホストに対して提供する。つまり、ネットワークのサービスの供給は、ネットワーク内部であり、そのサービスを楽しむのがエンドのホストである。また、それぞれのグリッドノードは、仮想化された資源(CPU、あるいはディスク)を使ってサービスを行なうため、ネットワーク上のどの場所からでも利用できるようにしている。

図5は、共通のグリッド分散環境によって、ネットワークサービスがエンドのホストに対して提供されることを表している。この環境において、ネットワーク上にある複数のグリッドノードによって、資源やサービスの仮想化が行なわれており、共通の分散環境として動作している。このサービスを利用す

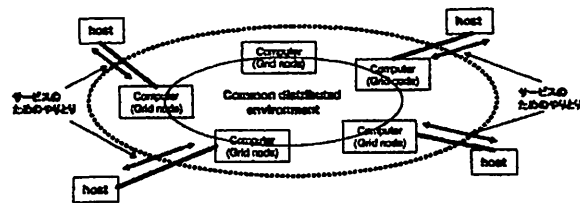


図5: 共通のグリッド分散環境とサービスプロバイダー

るため、エンドのホストは、サービスの提供を受けるために、最も近いグリッドノードと情報のやりとりを行なう。このやりとりは、通常のデータ転送をはじめる前に、グリッドノードとの間で事前に行なわれ、また、データ転送の途中であっても必要に応じて随時行なわれる。

#### 具体例

ここでは、グリッドコンピューティングを使って、どのようにサービスを提供するかについて、例として、ネットワーク側から提供するサービスとして、Webのキャッシュサーバを使って述べる。

現在のインターネットでは、Webのキャッシュサーバは、プロキシサーバとして、ブラウザにおいて定義されるか、あるいは、ブラウザがネットワークの構成に従うように作られていれば、DHCPで定義されたアドレスに対して、プロキシサーバがセットされることになる。そして、通常の構成では、図6のように、1. クライアント、2. プロキシサーバ、3. Webサーバという順番で要求が転送されていき、2. にコンテンツのキャッシュがない場合、3., 2., 1. と逆方向に返答が送られる。

ネットワークのバンド幅あるいは、サーバの混雑によって、2. のプロキシサーバのアドレスを動的に変更したい場合がある。このような状態を想定する。クライアントは、まず、ネットワークサービスの提供者に対して、プロキシサーバに関する情報を問い合わせる。また、ネットワークサービスの提供者は、クライアントに最も近く、空いているプロキシサーバの情報を提供する。クライアントは、プロキシサーバを使用してアプリケーションを利用する。動的なプロキシサーバの割り当てとなるために、同じネットワークに接続されていたとしても、別なクライアントであれば、別なプロキシサーバが割り当てをされることがある。また、一定時間経過すれば、再度、プロキシサーバの情報をネットワークサービスの提供者に問合せを行なう。

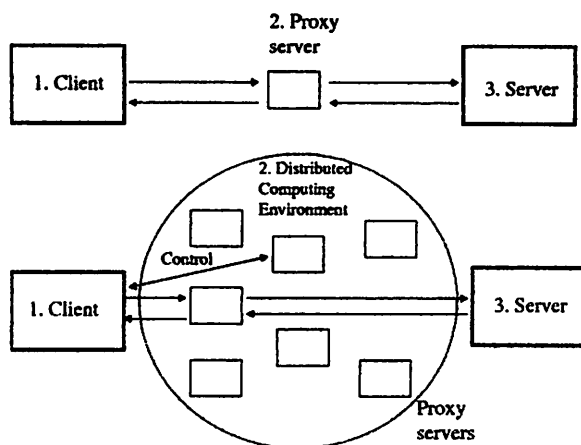


図 6: プロキシサーバの構成

一方、ネットワークサービスの提供者は、例えば、図6のように、クライアントに最も近くバンド幅があるプロキシサーバ、サーバに近いプロキシサーバ、最も空いているプロキシサーバ、キャッシュされたコンテンツを持っているプロキシサーバの中から、ポリシーによって情報を選択して、クライアントに提供することになる。

さらに、プロキシサーバが一台動作して、すでに多くのクライアントへのサービスを行っており、十分に使用されていれば、ネットワークサービス提供者は、グリッドコンピューティングの資源管理機構を使って、別なホストにおいて、追加で次のプロキシサーバのプログラムを新規動作させて稼動状態にする。そして、クライアントからプロキシサーバとしてサービスを利用できるようにする。セキュアな環境において、このような方式を使った動的なジョブの投入及び実行は、グリッドコンピューティングの技術を使うことによって可能になる。

この例は、プロキシサーバの動的な利用方法であるが、また、この方式は、ネットワーク上の別なサービスに対して利用可能である。このようにして資源の仮想化を行なうことによって、動的なサービスの提供が可能になり、この結果、ユーザに対して柔軟でパフォーマンスのよいサービスを提供することが可能になる。

#### 4. 議論

議論として、まず、第一に、このアーキテクチャを実現するための必要な条件は、エンドホストとグリッドノードによって構築される共通の分散環境の間のサービスの交換についてのプロトコルとその内容の定義である。このサービスの交換によって、エ

ンドホストに対して柔軟で多様なサービスを提供できるようにしなければならない。

次にネットワーク内に作られた共通の分散環境のノード間において交換されるべき情報及び共通機能を定義する必要がある。ノード間において共有される情報には、システムの現在のリソースの状況、ノードの処理能力、ノードごとに提供可能なサービスの種類、エンドホストへのネットワークの観点での距離（近いか遠いか）がある。これらの情報を共通の分散環境において情報として共有しておくことによって、エンドホストからのリクエストに対して適切な返答によるサービスの提供を行なうことができる。

#### 5. おわりに

この論文では、ネットワークのサービスを提供するために、グリッドコンピューティングの資源の仮想化とジョブを柔軟に走らせる機能を利用したアーキテクチャフレームワークを提案した。

今後、このアーキテクチャをもとに、ネットワークサービスを提供するための定義及びシステムの設計を行ない、提案しているアーキテクチャの有効性の証明及び改善を行なってゆく。

#### 参考文献

- [1] D. D. Clark, "The Design Philosophy of the DARPA Internet Protocols," In Proc. of ACM SIGCOMM'88, pp. 102-111, 1988.
- [2] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An Architecture for Differentiated Services," IETF Request for Comments 2475, Dec. 1998.
- [3] S. Shenker, C. Partridge, R. Guerin, "Specification of Guaranteed Quality of Service," IETF Request for Comments 2212, Sep. 1997.
- [4] B. Krishnamurthy, C. Wills, Y. Zhang, "On the Use and Performance of Content Distribution Network," In Proc. of ACM SIGCOMM IMW 2001, pp.169-182, Nov. 2001.
- [5] D. D. Clark, C. Partridge, J. C. Ramming and J. T. Wroclawski, "A Knowledge Plane for the Internet," In Proc. of ACM SIGCOMM2003, pp. 3-10, 2003.
- [6] I. Foster, C. Kesselman, S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations", Lecture Notes in Computer Science, vol. 2150, 2001.
- [7] F. Berman, G. C. Fox, A. J. G. Hey ed., "Grid Computing - Making the Global Infrastructure a Reality," John Wiley and Sons, Ltd., 2003.