

動的資源割り当てを可能とする広域分散型配信システムの提案

神屋 郁子
Yuko KAMIYA

九州産業大学 大学院 情報科学研究科
Graduate School of Information Science, Kyusyu Sangyo University
kamiya@nw.is.kyusan-u.ac.jp

下川 俊彦
Toshihiko SHIMOKAWA

九州産業大学 情報科学部 情報科学科
Faculty of Information Science, Kyusyu Sangyo University
toshi@is.kyusan-u.ac.jp, <http://www.is.kyusan-u.ac.jp/~toshi/>

1. はじめに

インターネットの普及に伴い、家庭でもブロードバンドネットワークの利用が可能となってきた。これにより、動画などの大容量コンテンツの配信が広く利用されている。大容量コンテンツの配信にはネットワーク帯域が必要である。近年は、CDN(Content Delivery Network)[Day 03]を利用して配信用サーバを広域分散配置することでネットワーク帯域を用意している。CDNは配信を行う際、クライアントからのアクセスを事前に想定し、十分なサービスの提供ができるように処理能力やネットワーク帯域を準備する。しかし、事前の想定に基づき資源を準備するため、想定を超えるアクセスが発生した場合、帯域不足や処理能力不足により、全てのクライアントに対しては正常に配信することができない。これはCDNが帯域や処理能力を動的に増やすことができないためである。近年ではクラウドコンピューティング技術が注目を浴びている。この技術を用いることで計算機を動的に増やすことで処理能力を動的に増やすことが可能となる。しかし、現在のクラウドコンピューティング技術は、局所的にしか計算機を増やすことができない。従って、ネットワーク帯域を増やすことができない。広域ライブマイグレーション技術[Robert 07]のように広域環境での計算機移動に関する取り組みもある。これを応用して広域環境での計算機増設も可能である。しかし、これらの取り組みは単純に計算機を増設することを目的としており、どこに増設するべきか、という問題は手付かずである。

本研究では、帯域を動的に増やすことが可能なコンテンツ配信システムを実現する。本システムは、帯域を動的に増やすために配信用のサーバを広域に動的に増設する。本システム構築の基盤として、我々が提案してきた、仮想計算機を利用して配信用サーバを広域に動的に増やす「サーバ増殖」[Kamiya 08, Kamiya 09]を用いる。サーバ増殖では、複数の物理計算機を広域分散配置しておき、必要な時に、物理計算機の中の1台を選択し、その上で仮想計算機を実行する。物理計算機を共有資源としておくことで、コストの削減を図る。また、仮想計算機の広

域複製により、物理計算機より短時間に広域増設可能とする。

我々は、これまでにコンテンツ配信システムにおけるリクエスト誘導システムの研究[Shimokawa 06]において、実運用システムへの適用などの成果を上げてきた。その中で、リクエスト誘導ポリシーの動的な変更に対する様々な要望があることが分かった。これらは静的に設置された配信サーバに対する動的な誘導ポリシーの変更を行うものであった。本研究は、これを進展させ、配信サーバの設置を動的に行うことにより、より柔軟なコンテンツ配信を実現する。すなわち、必要に応じて配信用のサーバを増設し、配信能力を増設する。この際に、増設したサーバの配置位置によってその配信能力は大きく変化する。本研究はこの課題に対する解決策を提案するものである。

本論文では、さまざまな計算機選択ポリシーを利用することで動的資源割り当てを可能とし、構成変更を可能とする、次世代コンテンツ配信システムを提案する。ここで資源割り当てとは、適切な物理計算機を選択し、その上で、仮想計算機を起動することである。この時、物理計算機を選択基準は配信者によって異なる。本論文では物理計算機を選択基準のことを計算機選択ポリシーと呼ぶ。さらに、本システムで利用する様々なポリシーについて提案する。

2. 既存の手法と問題点

ネットワーク帯域を増やす仕組みの一つにCDNがある。2・1節ではCDNについて説明し、2・2節ではCDNの問題点について述べる。

2・1 CDN

CDNは、広域に分散配置した複数のサーバに同一のコンテンツをミラーし、それら複数のサーバから配信を行うシステムである。CDNにおいて、ユーザに対してコンテンツを配信するサーバを配信用サーバと呼ぶ。ユーザは最適なサーバに誘導され、そこからコンテンツを取

得する。配信者は配信能力が足りないと予想される時に配信用サーバを追加しておく。既存の CDN サービスの例として Akamai[aka] や Limelight[lim] がある。

2.2 CDN の問題点

既存の CDN は、クライアントからのアクセスを予想し、事前に配信用サーバを配置する。そのため、事前に配置した配信用サーバ数が不十分だった場合には、ネットワーク帯域不足や処理能力不足になり、すべてのクライアントに正常に配信ができない。また、CDN では、サーバとして予め設置済みのものを使うので、構成を変更するのは容易ではない。例えば、Web やストリーミングなど既に対応しているアプリケーションや TCP などの対応済みのプロトコルを利用したサービスを導入するのは容易だが、P2P を用いた新たな配信システムや、SCTP のような新しいトランスポートプロトコルを利用した配信システムでの CDN の利用は困難である。

3. 動的資源割り当てとサービスに応じた構成変更が可能な広域分散型配信システム

2.2 節で述べたように、既存の CDN には動的に資源を増やすことができないという問題点と構成変更が容易ではないという問題点がある。そこで本研究では、動的資源割り当てとサービスに応じた構成変更が可能な広域分散型配信システム Soarin を提案する。本システムは、我々が研究を進めている「サーバ増殖」を利用してサーバを動的に広域分散配置する。3.1 節ではサーバ増殖について説明し、3.2 節でサーバ増殖とクラウドコンピューティングを比較し、3.3 節で動的資源割り当てを可能とする広域分散型配信システムについて述べる。

3.1 サーバ増殖

サーバ増殖とは、計算機システムの処理能力や利用可能なネットワーク帯域を動的に増減させることを可能とする技術である。このために、計算機を広域に動的に増減させる。図 1 にサーバ増殖システムの構成を示す。サーバ増殖では、仮想計算機(図中 VM)を用いる。仮想計算機を実行するためのソフトウェアが仮想計算機モニタ(図中 VMM)である。仮想計算機モニタを導入した物理計算機を予めインターネット中に広域分散配置しておく。この物理計算機を実行用サーバと呼ぶ。実行用サーバでは、仮想計算機の実行のみを行う。さらに、仮想計算機の転送専用のサーバである配布用サーバを準備する。これは、仮想計算機の広域動的転送に関する処理を実行用サーバから分離し、仮想計算機の実行に利用可能な処理能力を増やすためである。配布用サーバには仮想計算機のイメージファイル(図中 VM イメージ)を予め配置しておく。サーバの増設を行う際には、配布用サーバから他の実行用サーバへ仮想計算機のイメージファイルを転

送する。

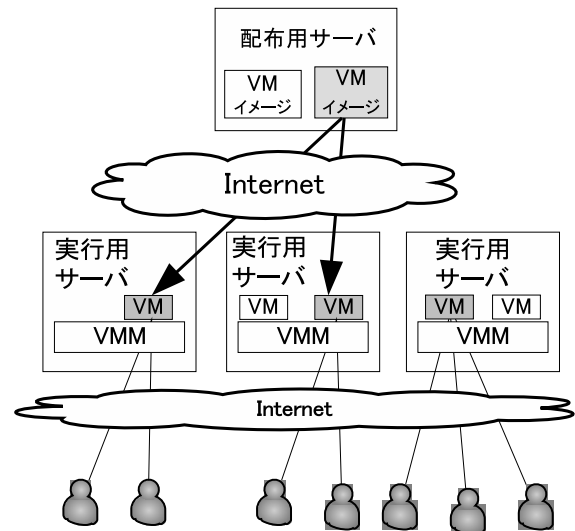


図 1 サーバ増殖の構成

3.2 サーバ増殖とクラウドコンピューティングの比較

計算機を動的に増減し、処理能力を増減する技術としては、クラウドコンピューティング技術がある。サーバ増殖の特徴は、計算機を広域に増減することで、処理能力のみならず、利用可能なネットワーク帯域も増減させることが可能な点である。

計算機を増設するとその計算機の CPU やネットワークが利用可能になる。その結果、処理能力や利用可能なネットワーク帯域が増加する。多くのクラウドコンピューティングシステムでは動的に計算機を増加させ、その処理能力を向上させる。しかし、処理能力に比べて利用可能なネットワーク帯域を増加させることは困難である。一般にクラウドコンピューティングシステムは、ある特定の iDC(インターネットデータセンター)内に設置される。そのシステムがインターネットへのネットワークトラフィックを流すのに利用可能な帯域は、その iDC の対外接続ネットワークの帯域が律速となる。

そこで、サーバ増殖では計算機を広域に増設することにした。これにより、増設前の計算機が利用する対外接続ネットワークとは別の対外接続ネットワークを利用して、インターネットへ接続することが可能となる。この結果、システムが利用可能なネットワーク帯域を増加させることが可能となる。

計算機を動的に増設するにあたって、物理計算機を用いると時間やコストがかかる。この問題を解決するため、サーバ増殖では仮想計算機を利用する。なぜなら、仮想計算機の実体はファイルなので、動的に増減させるのが容易だからである。このファイルを他の物理計算機に転送し、実行することで仮想計算機の動的な増減が可能となる。また、仮想計算機を使うことにより物理計算機資

源の共有を行うことができるので、コストの削減につながる。また、サーバ増殖で増設した仮想計算機以外の仮想計算機も実行できるため、汎用性が高い。さらに、計算機をまるごと複製することで、構成変更を容易に行うことができる。

3.3 動的資源割り当てとサービスに応じた構成変更が可能な広域分散型配信システム Soarin

本研究では、動的な資源割り当てとサービスに応じた構成変更が可能な広域分散型配信システムを提案する。以下、本システムのことを Soarin と呼ぶ。Soarin は、従来型の CDN がもつ、動的に資源を増やすことができないという問題点と構成変更が容易ではないという問題点を解決する。このために、配信用サーバを動的に増設することで資源の動的な増設を実現し、配信用サーバを仮想計算機上に構成することで構成変更を可能とする。

配信用サーバの動的な増設にあたっては、いつどこに配信用サーバを増やすかという新たな問題点が発生する。Soarin では配信用サーバを仮想計算機上に構成するため、この問題は仮想計算機を実行するための物理計算機をどのような基準で選択するか、という問題と等しい。我々はこの選択基準を計算機選択ポリシーと呼ぶ。我々は、これまでに CDN におけるリクエスト誘導技術について研究を進めており [Shimokawa 06]、その中でリクエスト誘導における選択基準には様々なものがあることが分かっている。従って、この計算機選択ポリシーについても様々な要求があることは自明である。すなわち、計算機選択ポリシーは配信者によって異なる。Soarin では、配信者の要求に応じて、さまざまな計算機選択ポリシーを利用可能とした。計算機選択ポリシーの例については 3.4 節で述べる。

配信用サーバを仮想計算機上に構成したため、配信用サーバの動的広域増設自体は、サーバ増殖技術を利用することで実現できる。しかし、計算機選択ポリシーによる増設場所の決定は、現在のサーバ増殖だけでは実現できない。そこで、サーバ増殖における実行用サーバ・配布用サーバに加えて、監視用サーバ・選択用サーバを導入する。監視用サーバでは実行用サーバの負荷やネットワーク帯域の監視、実行用サーバと配布用サーバの距離の計算など、計算機選択ポリシーで利用する様々なメトリックの監視を行う。選択用サーバでは、配信者が決定した計算機選択ポリシーに基づき、監視用サーバが取得したデータを元に、どの実行用サーバに仮想計算機を増設するかを選択する。その後配布用サーバに指示を出し仮想計算機を増設し新たな配信用サーバを起動する。配信用サーバ増設後は、リクエスト誘導システムにおける誘導ポリシーを更新する。リクエスト誘導システムとしては、我々が開発している Tenbin を利用する。

3.4 Soarin における計算機選択ポリシー

本論文では Soarin で利用可能な計算機選択ポリシーとして以下の五つについて説明する。これら五つのポリシーを組み合わせたことや、これら以外の計算機選択ポリシーも利用可能である。

§1 配布用サーバ・実行用サーバ間距離優先

配布用サーバからネットワーク的な距離に近い実行用サーバを選択する。ここでは、配布用サーバと実行用サーバのホップ数を基準とし、ホップ数が少ない実行用サーバを選択する。ネットワーク的に距離に近い実行用サーバを選択することで短時間で実行用サーバ上に仮想計算機を増設することができる。これにより急なアクセス増加に対応できるため、サービス不可能時間を短縮できる。

§2 処理能力優先

処理能力の余力が最も高い実行用サーバを選択する。このために監視用サーバでは、実行用サーバの処理能力の余力を監視しておく。これにより、処理能力不足を解決する。処理能力不足で配信が困難になったとき、処理能力の余力が最も高い実行用サーバを選択することで、配信システムの処理能力を強化し、配信が可能となる。

§3 利用可能帯域優先

利用可能帯域が最も広い実行用サーバを選択する。このために監視用サーバでは、各実行用サーバが接続している対外接続ネットワークの利用可能帯域を監視する。ここで利用可能帯域とは、ネットワークの物理的な帯域から、使用中の帯域を引いた値のこととする。これにより、ネットワーク帯域不足を解決する。動画配信などのネットワーク帯域が必要なサービスにおいて、ネットワーク帯域不足で配信が困難になったとき、利用可能帯域が広い実行用サーバを選択することで配信が可能となる。

§4 アクセスの多いクライアントの地域優先

クライアントの IP アドレスを元に、クライアントが存在する地域を判定する。その上で、もっとも多くのクライアントが存在する地域の実行用サーバを選択する。ここで近い地域とは、ネットワーク的な近さと、地理的な近さの二つを判定材料として判断する。

ネットワーク的な近さとしては、クライアントの IP アドレスからアクセス元の AS (Autonomous System) を識別する。AS はインターネット内での経路制御に用いられる単位であり、多くの場合プロバイダ (ISP) や大きな企業などがその単位となる。

地理的な近さとしては、国あるいは国が存在する地域を判断の単位として用いる。IP アドレスの国の識別は、RIR (Regional Internet Registry) が公開しているアドレス割り当て情報を用いることで実現可能である。RIR とは、ARIN、RIPE NCC、APNIC、LACNIC、AfrinIC の五つであり、ARIN は北米、RIPE NCC はヨーロッパ・中東・中央アジア、APNIC はアジア・太平洋地域、LACNIC はラテンアメリカ・カリブ海地域、AfrinIC はアフリカをそれぞれ管轄し、それぞれの管轄地域におい

て IPv4 アドレスなどの登録を管理している。例えば、133.0.0.0/8 は APNIC により日本に割り当てられていることが公開されている。これにより 133.0.0.0/8 の範囲にある IP アドレスは、日本国内であること、およびそれがアジア・太平洋地域にあることが判断できる。国が存在する地域とは、これら各 RIR の管轄地域を単位として用いる。例えば、ある二つのアドレスが、ともに RIPE NCC により、割り当てがなされている場合、これらはともにヨーロッパ・中東・中央アジア地域にあることが分かり、比較的近い地域に存在することが期待できる。

一般に、近い地域内の方がネットワーク帯域が太い。これは一般に、物理的に近い距離の方がネットワークを構築しやすいからである。例えば、ヨーロッパ内からのアクセスに対して、北米やアフリカではなく、ヨーロッパ内にある実行用サーバを選択することは合理的である。

§ 5 配信者の選択した実行用サーバ優先

実行用サーバの中で、配信者の使いたい実行用サーバを選択する。例えば、現時点では、課金制度などは考慮していないが、実行用サーバに課金制度が導入された場合、パフォーマンスだけでなく、コストも問題となってくる可能性がある。そこで、配信者はできるだけ低コストで配信のできる実行用サーバを選択したい可能性がある。また、接続されているネットワークを元にして実行用サーバの選択を行う場合もある。これにより、トラフィックエンジニアリングが可能となる。このように、パフォーマンス以外で配信者の望む実行用サーバの選択基準に合わせて実行用サーバを選択することも可能である。

4. ま と め

近年、広帯域コンテンツの配信には CDN が広く利用されている。しかし、既存の CDN は、配信能力の増減について柔軟性に欠けるといった問題点がある。そこで本論文では、その問題を解決するために、動的資源割当てとサービスに応じた構成変更が可能な広域分散型配信システム Soarin を提案した。物理計算機の計算機資源を動的に割り当てて配信用サーバを増設することで、配信能力の増設が可能である。また、割り当てる物理計算機を選択する際に利用する計算機選択ポリシーについて考察した。

今後は、Soarin の実装および本論文で提案した計算機選択ポリシーの評価を行う。

◇ 参 考 文 献 ◇

- [aka] Akamai Technology Inc; The Business Internet. <http://www.akamai.com/>.
- [Day 03] Day, M., Cain, B., Tomlinson, G., and Rzewski, P.: A Model for Content Internetworking(CDI), *RFC3466* (2003).
- [Kamiya 08] Kamiya, Y. and Shimokawa, T.: Scalable Server Construction Method Based On Virtual Machine Transfer and Duplication, in *Proceedings of International Multi-Conference on Engineering and Technological Innovation IMETI 2008*, pp. 273–278 (2008).
- [Kamiya 09] Kamiya, Y., Shimokawa, T., and Yoshida, N.: Scalable Server System Based on Virtual Machine Duplication in Wide Area, in *Proceedings of The 3rd International Conference on Ubiquitous Information Management and Communication*, pp. 432–436 (2009).
- [lim] Limelight: <http://www.limelightnetworks.com/>.
- [Robert 07] Bradford, R., Kotsovinos, E., Feldmann, A., Harald: Live wide-area migration of virtual machines including local persistent state, *Proceedings of the 3rd international conference on Virtual execution environments*, pp. 169–179 (2007).
- [Shimokawa 06] Shimokawa, T., Yoshida, N., and Ushijima, K.: Server Selection Mechanism with Pluggable Selection Policies, *Electronics and Communications in Japan, Part III(Fundamental Electronic Science)*, Vol. 89, pp. 53–61 (2006).