

経路情報を利用した広域分散配信の実証実験

神屋 郁子[†] 下川 俊彦[†] 岡村 耕二^{††} 河合 栄治^{†††}
 寺田 直美^{††††} 岡本 裕子^{†††††} 谷崎 文義^{††††††} 赤藤 倫久^{†††††††}

Case Study on Scalable Contents Delivery System using Routing Information

Yuko KAMIYA[†], Toshihiko SHIMOKAWA[†], Koji OKAMURA^{††}, Eiji KAWAI^{†††}, Naomi TERADA^{††††}, Yuko OKAMOTO^{†††††}, Fuminori TANIZAKI^{††††††}, and Tomohisa AKAFUJI^{†††††††}

あらまし 今後、インターネットにおける大規模コンテンツ配信の重要性はますます高まってくる。我々は、実際の大規模配信システムの運用を通じての技術開発や、次世代技術の開発により、この問題に対処しようと考えている。本研究では、動的な広域分散配信実現の第一歩として、IP Anycast を用いた広域分散配信実験を行った。この結果、サーバの削減時のクライアントの対応が、DNS を用いた場合よりも速やかであり、優れていることが明らかになった。一方で、運用上の問題なども明らかになった。本論文では、全国高校野球選手権大会のインターネット中継の中で実施した広域分散配信実験について報告する。

キーワード BGP, IP Anycast, CDN, 広域負荷分散

1. はじめに

一般家庭へのブロードバンドインターネットの普及に伴い、様々なコンテンツがインターネット上で配信されている。

多数の利用者に高品質なコンテンツを配信するためには大規模な配信システムを構築する必要がある。これは処理能力の高いサーバシステムや広帯域なネットワークから構成される。しかし、サーバシステムが局所的に配置された場合、ネットワークの輻輳が発生しボトルネックとなりうる。この問題を解決するための手法に広域分散配信がある。これはサーバシステムを

広域に分散配置し、それぞれのサーバシステムからコンテンツを配信することで、輻輳の発生を防ぐ。我々は広域分散配信を実現するための基礎技術としてサーバ選択システムや動的サーバ広域分散配置システムを開発してきた。

一方で我々は、全国高校野球選手権大会のインターネット中継などの大規模な配信に取り組み、安定かつ効率的な配信システムを開発してきた。

これらの技術と経験を基盤として、動的に広域分散配置したサーバ群に対して、経路情報を利用して負荷分散を行うことを計画している。すなわち、経路情報を利用することで、各利用者をネットワーク的に近傍のサーバへ誘導することを目的とする。従来、我々はDNSを用いてこの目的を実現してきた。しかしDNSを用いると、キャッシュの影響により意図した誘導が出来ないことが分かってきている。そこで我々は、実現手段としてIP Anycastを用いる。本研究では、静的に広域分散配置したサーバ群に対してIP Anycastによる広域分散配信の実証実験に取り組んだ。

本論文の構成は以下のようになっている。2. で我々がこれまでに開発してきた広域分散配信技術について説明する。3. で本実験の基本的なアイデアである、経路情報を利用した広域分散配信について述べる。4. で

[†]九州産業大学
Kyushu Sangyo University
^{††}九州大学
Kyusyu University
^{†††}情報通信研究機構
National Institute of Information and Communications
Technology
^{††††}奈良先端科学技術大学院大学
Nara Institute of Science and Technology
^{†††††}NTT スマートコネク
NTT Smart Connect
^{††††††}NTT 西日本
NTT West
^{†††††††}朝日放送
Asahi Broadcasting Corporation

本実験の内容および結果を説明し、5. で本論文をまとめる。

2. 広域分散配信

本章では、我々が開発してきた、広域負荷分散配信のための要素技術である、DNS を基盤としたサーバ選択システムや仮想計算機を用いたサーバの動的広域分散配置システムについて概要を説明する。本論文では CDN [1] の用語を用い、分散配置した各サーバのことをサロゲートと呼ぶ。

2.1 DNS を基盤としたサーバ選択システム

広域分散配信システムでは、広域に分散配置した複数のサロゲート間で、利用者からのアクセスを適切に分散させることが重要である。我々は、これを実現するために DNS を基盤としたサーバ選択システム TENBIN を開発してきた [2]。

インターネット上でサービスを利用する場合、クライアントはサーバへ接続する際に、DNS を用いてサーバのホスト名を IP アドレスに変換する。この時、通常の DNS サーバはあらかじめ定められた IP アドレスの候補からラウンドロビンで返答する。しかし、TENBIN では返答するアドレスを選択する際のアルゴリズムを変更することが出来る。このアルゴリズムのことを「選択ポリシー」と呼ぶ。選択ポリシーはホスト名毎に自由に切り替えることが可能である。選択ポリシーの例としては、最も負荷が軽いサーバを選択する、各クライアントとサーバ間の距離を経路情報を用いて比較しもっとも近いサーバを選択する、など様々なものが利用可能である。

我々はこのシステムを、皆既日食ライブ中継システムや大規模 ftp/www サーバに適用し評価を行っている [3], [4]。

2.2 仮想計算機を用いたサーバの動的広域分散配置システム

広域分散配信では複数のサーバを広域に分散配置する。2.1 で述べたサーバ選択システムは、既に配置済みのサロゲートからより良いサロゲートを選択するシステムである。このため、サロゲートの配置状況によっては、適切なサロゲートを選択できない場合がある。

この問題を解決するためには、サロゲートを動的に配置すれば良い。そこで、我々はサーバを広域に動的に増減させることを可能とする技術である「サーバ増殖」を提案している [5]。サーバ増殖はコンテンツ配信以外にも適用可能な技術であるが、本論文ではコンテ

ンツ配信に用いることを前提に説明する。

従来は多くの利用者にとって適切なサロゲートを提供するためには、あらかじめサロゲートを配置しておく必要があった。しかし、それにはコストがかかる。また、負荷の上昇などで設置済みのサロゲートの処理能力が不足した場合やネットワークが輻輳した場合への対応が困難であった。サーバ増殖では、必要な時に動的にサロゲートを増設することで、この問題に対処する。このために仮想計算機技術を利用する。仮想計算機のイメージファイルを事前に用意しておき、サーバの処理能力不足やネットワークが輻輳した場合、そのイメージファイルを適切な物理計算機へと転送し、仮想計算機を起動しサロゲートを増設する。

しかし仮想計算機イメージファイルの転送は物理計算機の処理能力やネットワーク帯域を必要とする。そこで、仮想計算機イメージファイルを転送するために専用の物理計算機を導入した。これにより物理計算機の処理能力やネットワーク帯域に不足が生じた場合でも仮想計算機を転送することができる。

仮想計算機を動的に増設することで、新たに到着するアクセスを新規の仮想計算機に誘導することができ、過負荷を軽減することが出来る。

しかし既存のサロゲートで処理中のリクエストを新規サロゲートに移動させることは容易ではない。これにはいくつかの要因があるが、その中の一つは IP アドレスの問題である。広域分散配信においては、増設したサロゲートは既存のサロゲートとはネットワーク的に離れたところに設置されるのが一般的である。従って、既存の IP アドレスとは異なる IP アドレスを設定することになる。このため接続中の通信を新規サロゲートに移すことは困難である。ただし、HTTP のように生存期間の短い通信の場合、上述の DNS ベースの Anycast との組み合わせなどで、途中から通信を新規サロゲートへ移すことが可能である。

3. 経路情報を利用した広域分散配信

2. で述べたように広域分散配信においては、各利用者を適切なサロゲートに誘導することが重要である。ここで「適切」の判断基準としては様々なものがある。2.1 で述べたように、サロゲートの負荷、クライアントとサロゲートの距離などである。本論文では、判断基準として経路情報を利用する。各クライアントにとって経路が最も短い、すなわち最も近い位置にあるサロゲートを適切なサロゲートと判断する。

以下で、経路情報を利用したサーバ選択の実現手法として、DNS を用いる手法と IP Anycast による手法を説明する。

3.1 DNS による Anycast

2.1 で述べた TENBIN は、DNS を基盤とするサーバ選択システムである。DNS を用いたことにより、利用者は単一のホスト名あるいは URI で複数のサーバの中の一台を利用することが出来る。すなわち利用者は配信システムのサーバが複数のサーバから構成されていることを意識することなく、その中の適切な一台からサービスの提供を受けることが出来る。これは DNS を基盤とした Anycast システムと考えることが出来る。

我々は TENBIN の選択ポリシーとして、BGP の経路情報を用いるものを実装した。TENBIN は DNS なのでクライアントの IP アドレスを知ることは出来ない。しかし一般的に各クライアントはネットワーク的に近傍にある DNS サーバをリゾルバとして用いる。そこでこの選択ポリシーでは、問い合わせ元の DNS サーバの IP アドレスをクライアントの IP アドレスとして取り扱った。その上でクライアントともっとも近い位置にあるサロゲートを最適と判断する。

あらかじめ各サロゲートが所属する AS の BGP スピーカと、TENBIN 用 BGP スピーカを BGP 的に peering させておく。ただしこれは経路情報の収集だけが目的であるので、TENBIN 用 BGP スピーカは経路情報を広告することはない。従って、他の BGP スピーカは TENBIN 用 BGP スピーカからの経路情報を受け取ることはなく、経路制御には影響を与えない。

この TENBIN 用 BGP スピーカが収集している経路情報を用いて、クライアントと各サロゲート間の AS パス長を得る。これにより、もっとも AS パス長が短いサロゲートを適切なサロゲートと判断することができる。そのサロゲートの IP アドレスを DNS の名前解決要求の答えとして返し、クライアントはそのサロゲートとの通信を開始する。

3.2 BGP による Anycast

IP Anycast [6] は、分散した複数のサロゲートが存在する物理的に異なるセグメント上に同一の IP アドレスを存在させ、その IP アドレスの経路情報をそれぞれ広告することで実現できる。IP Anycast では、同一論理セグメントが物理的に分散することになるので、経路制御プロトコルとして、リンクステート型のものとは相性がよくない。しかし距離ベクトル型のもの

であれば問題ない。一方、分散したサロゲートの経路情報の広告到達範囲の制御を行なうとすると、RIP (RIPv2) では機能不足である。これは BGP であれば、経路情報送信側で E-BGP ピアに対して経路情報の広告をするかしないかによって広告対象の制御が可能であり、さらに、経路情報を受信する側で受信した経路情報の優先度の変更が可能であるので、ある決められたポリシーに従って経路情報の伝播を制御することが可能である。そのため、BGP による IP Anycast では、それぞれのサロゲートからのパケット到達させる範囲の設定が可能である。

4. 全国高校野球選手権大会中継における実験

本研究では、経路情報を利用した広域分散配信の実証実験を行った。

経路情報を利用した広域分散配信を実現するために 3.2 で説明した、BGP による IP Anycast を用いた。BGP を用いることで、AS 単位での利用者の誘導が実現できる。すなわち本実験における利用者情報とは、利用者が使用するネットワークが所属するの AS のことである。

また配信実験のコンテンツとして、全国高校野球選手権大会のインターネット配信を用いた。これにより様々な利用者からの多数のアクセスが期待できる。

本章では、まず、配信実験を行った全国高校野球選手権大会のインターネット配信について説明し、今回の実証実験について述べる。

4.1 全国高校野球選手権大会のインターネット配信

朝日放送では、一般のインターネット環境が整備されつつあった 1996 年より、全国高校野球選手権大会のインターネット配信に取り組んでいる。このインターネット配信では、通常の WEB コンテンツに加えて、試合のハイライトやダイジェストおよびライブ中継など多くの動画コンテンツを有している。この配信での WEB コンテンツの役割は、これら高校野球の多様な動画コンテンツのポータルサイトであるとともに、大会期間中の試合および選手の情報提供が挙げられる。加えて、継続的にテレビや動画を見る環境にない視聴者が多く存在することを考慮すれば、より手軽にアクセスできる WEB コンテンツとしても、即時性や臨場感のある内容が求められている。このような需要に応えるために、テレビ中継用の CG データと同期した正確かつリアルタイム性のあるスコア情報のほかに、中

継映像を常時キャプチャし、その静止画像を数秒おきに更新するコンテンツを用意している。図1が、キャプチャした画像を埋め込んだページである。図中央下部のテレビ画面のような部分が、埋め込まれたキャプチャ画面である。



図1 キャプチャした静止画が埋め込まれたページ

ここで、コンテンツのデータサイズが大きく、更新頻度の高いWEBコンテンツの配信では、サーバ負荷やネットワーク帯域がボトルネックとなることが多い。特に、インターネット接続環境の普及と通信性能の改善とともに、視聴者の絶対数は年々増加し、配信負荷も増大していることから、当初より配信サーバの性能面でのチューニングに加えて、安定かつ効率的な配信への様々な取り組みを行っている。コンテンツ制作系の取り組みでは、中継映像の静止画像を数秒おきに配信する際に、ファイルの形で配信サーバにデータファイルをアップロードすると、そのデータ読み出しの際のディスクI/Oが性能のボトルネックになることが判明した。そのため、配信サーバへのデータ提供では、配信サーバが確保しているメモリ領域内にネットワーク経由で直接書き込んで使用する、共有メモリによる負荷分散(I/O分散)を行った。一方、コンテンツ配信系では、DNSラウンドロビンによる負荷分散(機能分散およびサーバ分散)から、L4スイッチ導入による頑健性のあるサーバクラスタの運用に移行し、その上でサーバ資源をアクセス傾向の変化に応じて、動的に随時再配分し、最適化を図るシステムの開発[7]等を行ってきた。

4.2 実験の概要

本実験では、全国高校野球選手権大会インターネッ

ト配信の一部のコンテンツを、広域分散配信した。

具体的には、コンテンツとしては、中継映像をキャプチャした静止画を対象とした。全国高校野球選手権大会インターネット配信ではweb系のサーバを2系統用いている。一つは4.1で述べた、中継映像をキャプチャした静止画像配信専用のサーバである。もう一系統のサーバでこれ以外のwebコンテンツを配信している。今回はこのうちの前者のwebサーバを、本来の配信元であるCKP(Cyber Kansai Project)に加えて、QGPOP(Kyushu GigaPOP Project)に設置した。

このコンテンツを採用した理由は、コンテンツの同期が容易だからである。一般にwebサーバが配信するコンテンツはファイルシステム上に格納されており、そのデータを同期しなければいけない。しかし、4.1で述べたように、この静止画のデータはファイルシステムを用いない。キャプチャサーバで生成した静止画像を、配信用webサーバにネットワークで伝送し、配信用サーバでは共有メモリに格納する。従来、CKPに設置したキャプチャサーバで生成した静止画像データを、CKPに設置した配信用webサーバに伝送していたのに加えて、QGPOPに設置した配信用webサーバにも伝送するようにした。

実験で広域負荷分散を行うに当たって、できるだけ通常の配信に影響を与えないという要求があった。そこで、負荷分散の対象とする範囲を限定することにした。本実験ではBGPを用いた制御を行うため、AS単位での要求元の割合を調査した。調査対象としては2007年度の全国高校野球選手権大会での中継映像キャプチャ静止画の配信記録を用いた。この結果SINETからの要求が全体の3.1%であることが分かった。これは相当数のアクセス数に上るが、配信全体から見ると大きな影響ではないと判断した。さらにSINETはQGPOPとピアリングしている。そこで負荷分散を行う対象をSINET ASおよびその接続先のASと決定した。

実験は8/11、8/13、8/14の三日間実施した。8/11,13は4試合、8/14は3試合が行われた。ただしQGPOPに置いたサロゲートからの配信は、8/11,13は第2試合~第4試合の3試合、8/14は全3試合の時間帯に実施した。

このうち、初日である8/11は第2試合(正確には11:16)からQGPOPからの経路広告を始めた。その後SINETからの経路広告のフィルタの調整を行っ

た。これは当初想定していなかったネットワークに対して経路が広告されていることが判明したためである。14:30 頃に、最終的なフィルタの調整が終了した。このため 8/11 については、QGPOP からの配信開始から 3 時間程度、予定していた範囲より広い範囲に対して配信した。

4.3 実験システムの構成

本実験システムは、BGP による IP Anycast を用いて、SINET 加入組織が本実験の当該の IP address と通信する際に、CKP のサロゲートではなく、QGPOP にあるサロゲートと通信が行なえるようにするように設計、設定を行なった。今回構築した実験ネットワークの概要を図 2 に示す。

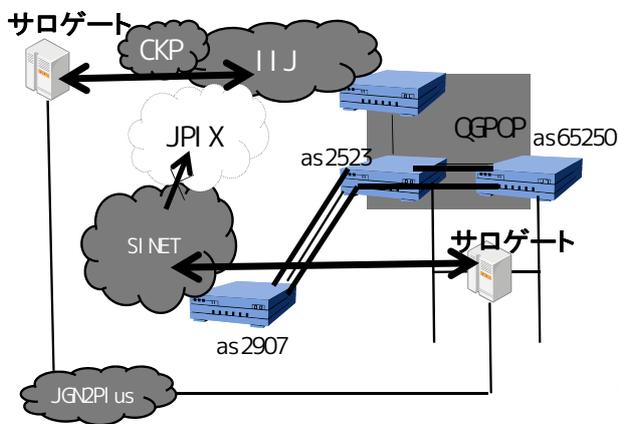


図 2 実験ネットワークの概要

当初は、SINET (AS 番号 2907) と直接 BGP ピアリングを行なっている既設の AS 番号 2523 (九州ギガポッププロジェクト) から実験用セグメントの prefix を広告し、SINET の他の E-BGP ピアで、その prefix の広告を制御する方法を検討した。しかし、運用上の問題により、既設の AS 番号 2523 による BGP ピアリングとは別に新規にプライベートに BGP ピアを SINET との間に設け、その BGP ピアから実験用セグメントの広告を SINET に対して行ない、SINET 側で他の E-BGP への広告のフィルタ制御を行なうようにした。

この非常に単純かつ明快なネットワーク構成で、本実験は遂行できると期待されたが、SINET の加入機関は、我々の予想とは異なり SINET の持つ AS 番号 2907 に所属するものだけでなく、SINET に対して BGP ピア接続をしている組織が多く、SINET 側での

フィルタ制御は予想よりも困難であった。

本実験の思惑としては、SINET 加入機関のアクセスを QGPOP のサーバへ誘導することであったが、SINET から見て、BGP ピアリングしている加入機関のネットワーク上のコンピュータと、SINET がピアを持つ ISP をトランジットして通信ができる商用ネットワーク上のコンピュータの区別は事実上自動的に行なうことはできなかった。

そのため、実験開始後、トラフィックをフローデータで計測し、アクセスに含まれる AS 番号を調査し、商用ネットワークからのアクセスを発見すると、そのネットワークへの SINET からの広告をフィルタすることで、最終的に我々の思惑を反映した経路広告の状態にすることができた。

本実験ネットワークにおいて、新規でプライベートピア用の BGP ルータはネットワークとしては、既設の AS 番号 2523 の AS 内部に存在し、SINET とは AS 番号 2523 を経由してマルチホップで BGP 接続していた。そのため、QGPOP にある本実験ネットワークからの通信は、AS 番号 2523 のポリシーに従っていた。このような状況で、SINET の加入機関で、商用ネットワークとも BGP ピアを持っている組織は、本実験の prefix は SINET から広告され、QGPOP のサーバをアクセスするものの、QGPOP からそのその組織へは AS 番号 2523 の持つ商用ネットワーク経由で通信が行なわれる、という状況になっていた。これは、本実験ネットワークがこの実験ではプライベート AS を用いたマルチホップ接続であったため、固有のポリシーを持つことができなかったことに起因する。

4.4 ネットワークトラフィック

実験時における QGPOP 側サロゲートのトラフィックを図 3,4,5 に、CKP 側のサロゲートのトラフィックを図 6,7,8 に示す。MRTG の設定の関係で、QGPOP 側のトラフィックと CKP 側のトラフィックで in と out が逆になっている。

QGPOP 側の三つのグラフ (図 3,4, 5) を比較すると 8/11 のトラフィック (図 3) が多いことが分かる。これは、4.2 で述べたように、SINET における経路広告フィルタ調整の影響である。調整終了後となる 8/11 の 15 時以降は、最大 4.4Mbps 程度に落ち着いた。

8/13 のグラフ (図 4, 7) を比較すると、QGPOP 側のトラフィックのピークが 3 回なのに対して、CKP 側のトラフィックのピークが 4 回であることが分かる。これは、この日は 4 試合実施されたにも関わらず、

QGPOP 側のサロゲートが第 2 試合から第 4 試合の 3 試合分しか配信しなかったため、4.2 で述べたように、計画的なものである。

8/14 のグラフ (図 5, 8) を比較すると、どちらも似たようなカーブを描いていることが分かる。

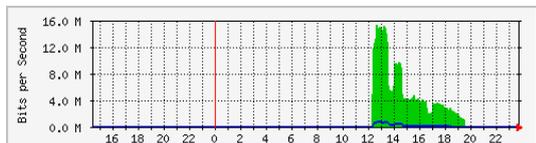


図 3 QGPOP 側の 8/11 のトラフィック

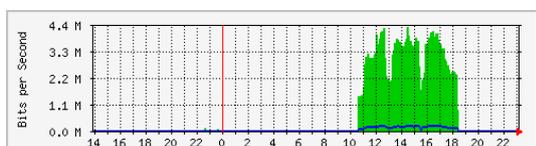


図 4 QGPOP 側の 8/13 のトラフィック

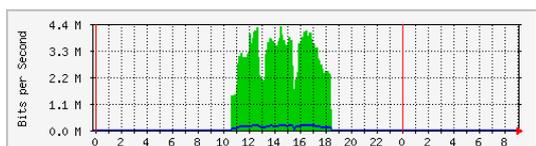


図 5 QGPOP 側の 8/14 のトラフィック

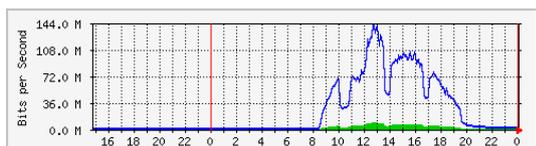


図 6 CKP 側の 8/11 のトラフィック

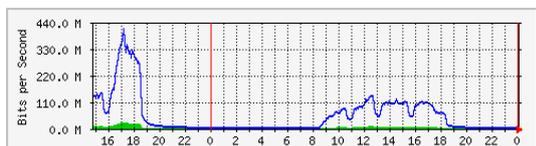


図 7 CKP 側の 8/13 のトラフィック

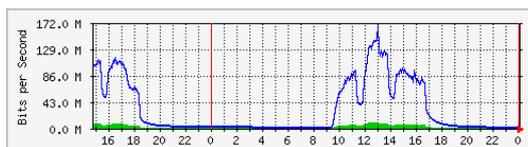


図 8 CKP 側の 8/14 のトラフィック

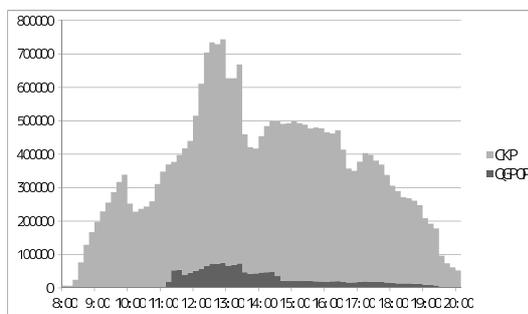


図 9 8/11(火) のアクセス数

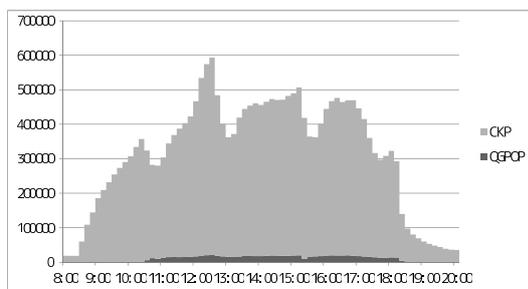


図 10 8/13(木) のアクセス数

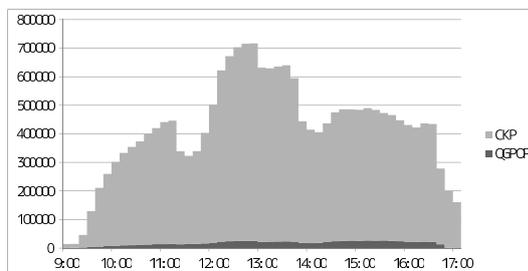


図 11 8/14(金) のアクセス数

4.5 アクセス数

実験時における各サロゲートへのアクセス数を図 9, 10, 11 に示す。これは各サロゲートへのアクセス数を毎日に 10 分単位で集計したものである。

アクセス数の変化は、基本的にトラフィックと同様の変化を示している。また QGPOP と CKP のアクセス数を合計した総アクセス数の変化は、試合時間な

どと一致しており、従来のアクセス傾向と比較して大きな変化はない。

4.6 考察

4.6.1 DNS anycast との比較

2.1 で述べたように、従来我々は DNS を用いた Anycast による実験を行ってきたので、これとの比較を述べる。図 12 は、以前我々が DNS による Anycast を行った際の結果である。この実験では、11 番から 21 番までの 11 台の http サーバを用意し DNS を用いて局所的な負荷分散を行った。15:00 から 16:00 までは 11 番の http サーバ 1 台のみがサービスを提供した。16:00 から 17:40 までは 12 番と 13 番のサーバ 2 台でサービスを提供した。17:40 以降はまず 14 番から 17 番までのサーバ 4 台でサービスを提供し、18:00 に 18 番から 21 番までの 4 台を追加しサーバ 8 台でサービスを行った。

この切り替えにおいては、サーバのホスト名の DNS の A レコードを順次変更していくことで実現した。15:00 から 16:00 までは DNS には 11 番の IP アドレスが登録されている。16:00 から 17:40 までは DNS には 12 番と 13 番の IP アドレスだけが登録されている。17:40 から 18:00 までは DNS には 14 番から 17 番までの IP アドレスだけが登録されている。18:00 以降は DNS には 14 番から 21 番までの IP アドレスだけが登録されている。この際、DNS のキャッシュの影響を受けないように TTL は 0 を指定した。

グラフを見ると、概ね DNS の A レコードに登録したサーバへアクセスが来ていることが分かる。16:00 以前は 11 番にのみアクセスが来ている。16:00 から 17:40 までは 12 番と 13 番へのアクセスが大半である。17:40 から 18:00 までは 14 番から 17 番へのアクセスが大半である。しかし、切り替えた後も継続して旧サーバにアクセスが来ていることがわかる。例えば、17:40 以降、DNS には 12 番のアドレスは登録されていないにも関わらず、その後 20 時過ぎまで 12 番へのアクセスがある。

これより DNS Anycast では、サーバを減少させることが容易ではないと言える。DNS への登録を変更すると、ほとんどのクライアントは変更後の設定に従う。しかし、一部のクライアントが変更前の設定を使ってアクセスしてくる。もしも、サーバを停止させてしまうと、そのサーバへアクセスしてくるクライアントに対してサービスを提供できなくなる。このように、TTL を無視してくるクライアントについて、OS

やブラウザなどなんらかの特徴を見いだすことが出来ないか調査を行ったが、現時点では共通点を見いだすことが出来ていない。

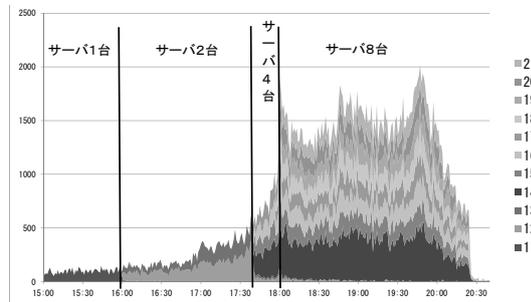


図 12 DNS Anycast におけるサーバの切り替わり状況

一方で、今回用いた IP Anycast では、図 3,4,5 を見れば分かるように、経路の広報を終了すると直後からリクエストが来なくなる。これは IP の経路制御の仕組み上当然の結果である。すなわち、IP Anycast ではサーバを減少させたい場合、経路情報の広報を止めるだけでよいことが分かる。

なお、サーバを増設した場合の新サーバへのリクエストの到達開始については、DNS Anycast の場合と IP Anycast の場合では有意な差は見いだせなかった。

4.6.2 負荷分散の結果

本実験の主眼である IP Anycast による広域負荷分散については、期待通りの成果が出た。SINET による経路フィルタの調整が終了した 8/11 14:00 以降での、QGPOP 側のサロゲートからの配信を行っている期間での、QGPOP 側サロゲートからの配信数は、総配信数の 4.2%であった。これは、4.2 で述べた 2007 年度配信の際の SINET AS を経由した配信における 3.1%と比較すると、期待通りの結果といえる。

また、利用者には負荷分散を行ったことを一切意識させる必要はなかった。分散配信を行っている時間帯とそうでない時間帯において、利用者からのアクセス方法には変更がなかった。配信元である朝日放送への問い合わせもなかった。このことから、本実験によりアクセス不可能などの障害も発生せず、正常に配信が行えたと判断している。

4.6.3 今後の課題

ただし、IP Anycast の適用性については、まだ検討が必要である。今回は QGPOP 側サロゲートに誘導するために /28 の経路情報を広報した。このように

プレフィックスが短い経路情報はフィルタされてしまう可能性がある。また、多くのサービスで IP Anycast が利用されると、ルータが保持しなければならない経路数が増加する。従って、その影響を考察しなければいけない。

また、今回の実験で当初経路情報のフィルターに問題があったように、従来の BGP オペレーションと比較すると複雑なオペレーションを必要とする可能性がある。運用性についても今後の検討が必要である。

一方で、IP Anycast による配信サーバの切り替えが非常にスムーズに行えたことから、IP Anycast は広域負荷分散以外にも利用可能ではないかと考えている。例えば AS 毎に配信するコンテンツを切り替えることを考えている。

5. おわりに

今後、インターネットにおける大規模コンテンツ配信の重要性はますます高まってくる。我々は、実際の大規模配信システムの運用を通じての技術開発や、次世代技術の開発により、この問題に対処しようと考えている。

本研究では、動的な広域分散配信実現の第一歩として、IP Anycast を用いた広域分散配信実験を行った。この結果、サーバの削減時のクライアントの対応が、DNS を用いた場合よりも速やかであり、優れていることが明らかになった。一方で、運用上の問題なども明らかになった。

今後は、我々が開発している動的なサーバ分散配置技術との組み合わせによる広域分散配信実験を行う予定である。

謝辞

本実験を行う機会を与えてくださった、サイバー関西プロジェクトおよび九州ギガポッププロジェクトの皆さんに感謝する。

本研究の一部は科学研究費補助金，特定領域研究 21013008 及び基盤研究 (B) 20300024 の支援を受けている。

文 献

- [1] M. Day, B. Cain, G. Tomlinson, and P. Rzewski, "A Model for Content Internetworking (CDI)," RFC3466, February, 2003
- [2] 下川俊彦, 吉田紀彦, 牛島和夫, "多様な選択ポリシーを利用可能なサーバ選択機構," 電子情報通信学会論文誌, J84-D-I:9, 1396-1403, September, 2001
- [3] 下川俊彦, 木場雄一, 中川郁夫, 山本文治, 吉田紀彦, "広

- 域分散環境における DNS と経路情報を利用したサーバ選択機構," 電子情報通信学会論文誌, J86-B:8, 1454-1462, August, 2003
- [4] 岩佐宗幸, 吉田充, 下川俊彦, "コンテンツ同期状態を利用したサーバ選択機構," 情報処理学会/電子情報通信学会 情報科学技術フォーラム 2006 論文集, Vol. 4, 7-8, September, 2006
- [5] Yuko Kamiya, Toshihiko Shimokawa, and Norihiko Yoshida, "Scalable Server System Based on Virtual Machine Duplication in Wide Area," Proceedings of The 3rd International Conference on Ubiquitous Information Management and Communication, pp.432-436, January, 2009
- [6] C. Partridge, T. Mendez, and W. Milliken, "Host Anycasting Service," RFC1546, November 1993
- [7] 藤部 修平, 知念 賢一, 河合 栄治, 赤藤 倫久, 香取 啓志, 山口 英, 山本 平一, "Web サーバファームにおけるスタンバイ状態を導入したサーバ切り替え方式の開発", 情報処理学会 分散システム/インターネット運用技術 シンポジウム 2003, January 2003.