

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

下川 俊彦^{†a)} 木場 雄一^{††} 中川 郁夫^{†††} 山本 文治^{††††}
吉田 紀彦^{††††}

Server Selection Mechanism using DNS and Routing Information in Widely Distributed Environment

Toshihiko SHIMOKAWA^{†a)}, Yuichi KOBAYASHI^{††}, Ikuo NAKAGAWA^{†††},
Bunji YAMAMOTO^{††††}, and Norihiko YOSHIDA^{††††}

あらまし 従来の単純なサーバクライアントモデルでは、処理要求がサーバに集中し、十分なサービスを提供することができなくなっている。この問題を解決するための手法の一つに CDN (Content Delivery Networks) がある。CDN を構築するためには、利用者からの処理要求を、分散配置したコンテンツ配布用サーバに適切に誘導する必要がある。これを実現するものがサーバ選択機構である。本研究では DNS を基盤として用い、経路情報を判断材料として利用するサーバ選択機構を開発した。本システムを実システムに適用し、その有効性を確認した。

キーワード CDN, DNS, BGP, サーバ選択, リクエストルーティング

1. ま え が き

FTTH, ADSL, CATV を代表とする、いわゆるブロードバンドネットワークの一般利用者への普及が急速に進んでいる。

従来の単純なサーバクライアントモデルでは、処理要求がサーバへ集中し、サーバシステムの処理性能やネットワーク性能に不足が出てきている。この結果、多数の利用者に対して十分な品質のサービスを提供することが困難になってきている。

一般に、高性能な計算機や、広帯域なネットワークは高価である。そこで、経済的にこの状況に対処する

手法が必要となってきた。この問題に対する解決策の一つが CDN (Content Delivery Networks) である。CDN は、Web やストリーミングのようなコンテンツ配布系のサービスを対象にしている。利用者の近くにコンテンツ配布用サーバ^(注1)を配置することで負荷の分散を図り、高品質なサービスを提供する。

Day らは、CDN を構成する要素は次の四つであると述べている [1]。

- (1) コンテンツ配送基盤
- (2) 要求誘導基盤
- (3) 配布基盤
- (4) 課金基盤

本論文は、このなかの要求誘導基盤を対象としている。本論文では、要求誘導のことをサーバ選択と呼ぶ。

本論文では、DNS を選択基盤として用い、経路情報を選択基準とするサーバ選択機構を提案する。本機構を実装し、インターネット上の実システムに適用し評価を行った。

2. で本論文で提案するサーバ選択機構について説明する。3. で本機構の実現方法を説明し、4. で本機構

[†]九州産業大学情報科学部, 福岡市

Faculty of Information Science, Kyushu Sangyo University,
Fukuoka-shi, 813-8503 Japan

^{††}株式会社東芝 研究開発センター, 川崎市

Corporate Research & Development Center, TOSHIBA,
Kawasaki-shi, 212-8582 Japan

^{†††}インテックネットコア, 東京都

Intec NetCore, Inc., Koto-ku, Tokyo, 136-0075 Japan

^{††††}株式会社アイアイジェイメディアコミュニケーションズ, 東京都

IJ Media Communications Inc, Chiyoda-ku, Tokyo, 101-
0051 Japan

^{†††††}埼玉大学工学部, さいたま市

Faculty of Engineering, Saitama University, Saitama-shi,
338-8570 Japan

a) E-mail: toshi@is.kyusan-u.ac.jp

(注1): CDN では「サロゲート」と呼ぶ。従来、ミラーサーバと呼ばれていたものと同等の役割を果たすと考えてよい。

を, 実システムに適用した結果を述べ評価を行う. 5. に関連研究について述べる. 最後に 6. で本論文のまとめを行う.

2. DNS と経路情報を用いたサーバ選択機構

CDN における要求誘導基盤の目的はクライアントからの処理要求に対して, 最善のサーバを教えることである [2]. 筆者らはこの技術をサーバ選択と呼んでいる. 本論文では, 経路情報を誘導基準として用いることが可能なサーバ選択機構を提案する.

本章では, 本論文で提案するサーバ選択機構について説明する.

2.1 提案するサーバ選択機構の概要

サーバ選択機構の実現においては以下の二つの要素を考慮する必要がある.

- (1) サーバ選択基準
- (2) サーバ選択方式

サーバ選択基準とは, サーバ選択を行う際の選択の基準のことであり, サーバ選択方式とは, どのようにしてサーバ選択を実行し, またその結果を返すかである.

本論文で提案する手法について, 次節以降で説明する.

2.1.1 サーバ選択基準

サーバ選択基準としては, 本研究が CDN の一部であることから経路情報を用いる. 以下に, その理由を述べる.

サーバ選択基準は, サーバ選択を実行する主体によって変化する. この主体としては大きく, サービス提供者 (サーバ), サービス利用者 (クライアント), ネットワーク運用者の三つに分けることができる. 本研究では, サービス提供者及びネットワーク運用者によるサーバ選択に着目する. ネットワーク運用者の視点では, 回線の利用が片寄らないことや, いわゆる上流回線を使いつぶさないことが選択の基準となる. サービス提供者にとっては, より高品質なサービスを利用者に届けることが重要である. この場合も, その配送基盤であるネットワークが有効活用されなければ, 実現は困難である.

以下に例を示す. 近年のインターネットは ISP や組織が相互に接続されることで構成されている. ここで三つの ISP と一つの大学が相互接続された, 図 1 に示すようなネットワークを考える. ISP A と ISP B には

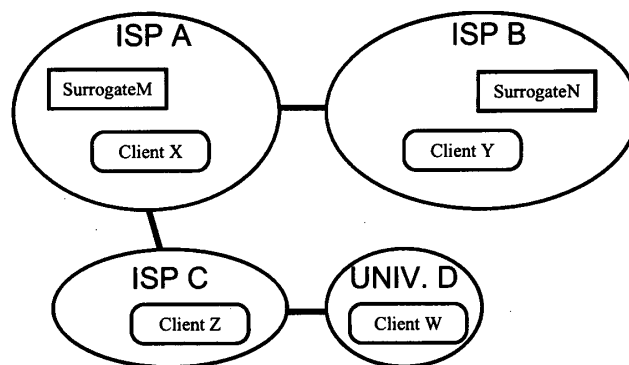


図 1 ISP 相互接続
Fig. 1 ISP interconnection.

配信用サーバ (Surrogate M, Surrogate N) が配置されている. この場合, ISP B に存在する Client Y は, Surrogate N からサービスを受けることで, ISP A, B 間のネットワークを有効活用できる. また, Client W, Client X, Client Z は, Surrogate M からサービスを受けることで, 同様に ISP A, B 間のネットワークを有効活用できる.

そのような制御を行うための判断基準として, 従来 IP アドレスが広く用いられている. すなわち, クライアントの IP アドレスを見て, 適切な配信用サーバを選択する方法である. しかし, この方法には問題がある. 例えば, 図 1 の場合, Client W は Surrogate M が選択されることが望ましいのは上述のとおりである.ところが, UNIV. D が, ISP B との接続を確立 (購入) し, 更に ISP B との回線を主たる回線として利用し始めたとする. すると, Client W が Surrogate M と通信する場合, ISP A と ISP B 間のネットワークを通してしまふ. この接続変更後, Client W は Surrogate N からサービスを受けるほうが望ましいのである.

インターネット上には多数の ISP や組織がつながっている. したがって, それらの接続変更を把握し, 更に, その情報をサーバ選択機構に反映させることは困難である. したがって IP アドレスを用いた方法は, ネットワークの接続状況の変化への対応が困難であることがわかる.

そこで, 本研究では, 経路情報に着目した. 本論文では, 経路情報として広域での経路制御に利用される BGP を利用する. 上述の UNIV. D のように, 接続を変更した場合, インターネットに広告する経路情報を変更するのは本質的な作業の一つである. したがって, 経路情報は必ず更新される. 例えば, UNIV. D の場

合、ISP B に対して経路情報を流し始める。すなわち、経路情報をサーバ選択基準として用いることで、ネットワークの接続状況の変化への対応が容易になる。

2.1.2 サーバ選択方式

本論文では、サーバ選択方式としては文献 [3] での検討をふまえ、DNS を用いる。

ただし文献 [3] では、サービス利用者によるサーバ選択の実現に DNS を用いた。本論文では、前項でも述べたようにサービス提供者・ネットワーク運用者によるサーバ選択に DNS を利用しその評価を行った。

2.2 本機構のモデル

本論文での提案手法のモデルを図 2 に示す。配信用サーバ (図中 ServerA,B,C) が広域に分散配置されている。経路情報サーバ (図中 Routing Server) が、広域分散配置した配信用サーバから経路情報を収集する。経路情報の動的な変更に対応するため、経路情報は継続的に収集する。クライアント (図中 Client1, 2) は、サーバ選択機構の存在を意識することなく、通常の処理どおりに DNS サーバ (図中 DNS Server) に対して名前解決要求を発行する。DNS サーバは、受け付けた名前解決要求に対して、経路情報サーバがもつ各配信用サーバの経路情報をもとに最適なサーバを選択する。選択されたサーバの IP アドレスを受け取ったクライアントは、そのサーバに対して処理要求を発行する。

説明のため、図 2 では、クライアントが本機構を構成する DNS サーバに直接名前解決要求を発行している。この場合、DNS サーバは、クライアントの IP ア

ドレスを知ることができるので、その情報をもとに最適なサーバを選択することができる。

これを基本にして、一般的な状況に展開すると、クライアントは同一あるいは近傍のネットワークに存在する DNS サーバ (ローカル DNS サーバ) に対して、名前解決要求を発行する。このローカル DNS サーバが、DNS の recursive query の仕組みを用いて、本機構の DNS サーバに対して名前解決要求を発行する。この場合、本機構を構成する DNS サーバは、ローカル DNS サーバの IP アドレスしか知ることができない。

しかし、この場合でも本モデルの有効性は崩れない。上述のように、一般的に、クライアントとローカル DNS サーバはネットワーク的に近い位置にある。したがって、ローカル DNS サーバの IP アドレスをもとにサーバ選択を行っても、クライアントの IP アドレスをもとにサーバ選択を行う場合とほぼ同一の結果を得ることが可能となる。

このモデルでは、図中の DNS サーバと経路情報サーバに負荷が集中する可能性がある。これらについて説明する。

2.2.1 DNS サーバの負荷分散

DNS サーバについては、DNS 自体がもつ機能で回避可能である。DNS では、特定のドメインの DNS サーバとして、複数の計算機を登録しておくことができる。クライアントはこれらのうちのどこに問い合わせても同一の回答を得ることができる。例えば、A, B, C と三つの DNS サーバが登録されている場合、DNS では、特定の DNS サーバにのみ集中させるのではなく、A, B, C に順に問合せを出す。これで、DNS サーバの負荷を分散させることができる。この機能のことを NS ラウンドロビンと呼ぶ。

2.2.2 経路情報サーバの負荷分散

経路情報サーバについては、分散配置が可能である。これは、利用者から直接アクセスされるものではないので、配置場所・配置数についての自由度が高い。例えば、分散配置した DNS サーバのそれぞれの近傍に経路情報サーバを置くことで、経路情報サーバへの負荷を分散させることが可能である。

この場合の問題点は、経路情報のため負荷の増加である。各配信用サーバから、すべての経路情報サーバへ経路情報を提供するとなると、運用上の手間が無視できない。この問題に対しては、経路情報の再配布 (redistribute) で解決が可能である。すなわち、各配信用サーバからの経路情報を、いったん中央の経路情

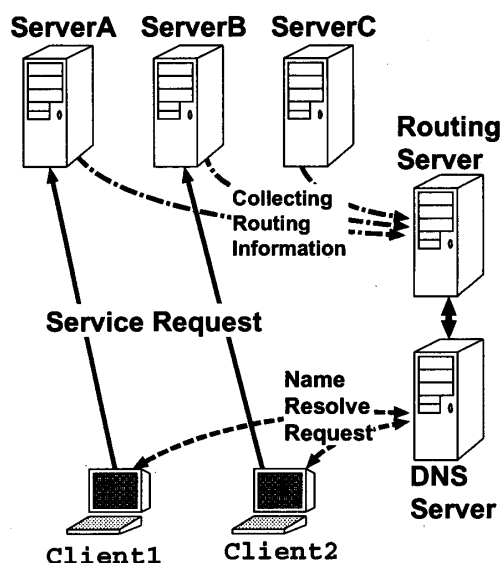


図 2 提案方式のモデル
Fig. 2 Model of our system.

報サーバに集信し、そこから各地の経路情報サーバに再配布する。中央の集信用経路情報サーバの多重化も可能であるため、ここがボトルネックとなることも回避可能である。

3. DNS と経路情報を用いたサーバ選択機構の実現

本章では、本論文で提案するサーバ選択機構の実現方法について述べる。

3.1 DNS サーバ TENBIN

DNS サーバとしては、筆者らが開発を進めている TENBIN [3] を用いる。TENBIN は柔軟なサーバ選択ポリシーを適用可能なサーバ選択機構として開発を進めている。ここで選択ポリシーとは、選択基準、及びそれをどう評価するかを定めるルールのことである。

TENBIN では、サーバ選択の選択基準となる情報を外部プログラムを通して収集できる。本論文においては、後述する経路情報サーバ RADIX から経路情報を取得するモジュールを開発することで、経路情報を選択基準として用いるサーバ選択を実現した。

TENBIN では、サーバ選択のルールをプログラムとして記述することができ、そのプログラムのことをポリシーモジュールと呼んでいる。本論文では、上述のように RADIX から経路情報を得て、その情報を用いて判断する。4. で述べるように、経路情報の判断方法には様々なものがあり、その結果、選択ポリシーも多様なものとなる。そこで、TENBIN が複数の選択ポリシーを柔軟に取り扱えることを生かし、多様なポリシーモジュールを開発し適用した。

3.2 経路情報サーバ RADIX

2.1.1 で述べたように、収集する経路情報としては BGP を用いる。これは、配布用サーバが広域に分散配置されていることを想定しているからである。しかし、同一 AS 内に複数の配布用サーバが配置されることもありうるため、将来的には BGP のコミュニティ属性の活用や IGP の利用も視野に入れている。

経路情報を収集するための経路情報サーバとしては、我々が開発を進めている RADIX [4] を用いる。経路情報の収集元としては、広域分散配置した配布用サーバと同一 AS 内の BGP ルータを用いる。通常、配布用サーバが直接 BGP の経路情報を処理することはない。一般的には、その配布用サーバが対外接続するためのルータが BGP を処理する。つまり、各配布用サーバと同一 AS 内に存在する BGP ルータから経路

情報を収集することで、必要な情報を得ることが可能となる。

本機構において RADIX は、通常のルータ同様経路情報収集元のルータから経路情報の提供を受ける。ただし、その経路情報を RADIX が動作しているホストの経路表に反映させることはしない。また、RADIX 側から経路情報を流すこともしない。RADIX が経路情報の提供を受けるルータ側からは、RADIX は他と同様のルータの一つに見えるので、情報提供に際して経路情報交換相手が増えるだけである。管理運用の時間も小さく、本機構の導入も容易である。また、他のルータと同様に継続的に経路情報の提供を受けるので、経路の状態に変化があった場合、即座にその情報が伝搬してくるという利点もある。

RADIX はソケット経由で他のプロセスとの通信が可能であり、TENBIN に対して経路情報を受け渡すことが可能である。また、RADIX はアプリケーションから利用しやすい形で経路情報を提供するという目的もっており、本サーバ選択機構はその一応用例ともいえる。本機構用に、ある IP アドレスまでの AS パスのみを表示する機能を追加した。また、本機構以外の他のアプリケーションとの連携例としては、CGI から経路情報を利用するシステムも開発している。

3.2.1 RADIX による経路検索

RADIX を用いた経路表検索の例を図 3 に示す。RADIX が、サイト A、サイト B、サイト C、サイト D の四つのサイトから経路情報を集めていたとする。このとき、133.5.22.47 という IP アドレスまでの経路を調べた例である。この結果より、サイト A,C,D から 133.5.22.47 までの AS パスがわかる。また、各 AS パスの最終要素はクライアントが属する AS (以下、オリジン AS と呼ぶ) を示す。サイト B からの AS パスが表示されていない。これは以下のいずれかである。

- (1) サイト B から 133.5.22.47 には到達できない。
- (2) サイト B から 133.5.22.47 に到達するにはデフォルト経路を通る。

```
radix>find ip as-path string 133.5.22.47
siteA 1111 4444 2500 2508
siteC 2222 2907 2508
siteD 3333 5555 2500 2508
```

図 3 経路検索例

Fig. 3 Route lookup example.

(3) サイト B は 133.5.22.47 への経路情報を RADIX に提供していない。

上述の(3)に示すように、各ルータは、全経路情報を RADIX に提供しない場合もある。この場合のサーバ選択結果は選択ポリシーに依存する。

3.3 TENBIN+RADIX

TENBIN と RADIX によるサーバ選択機構を図 4 に示す。本機構では、TENBIN と RADIX 間での通信プロトコルを定義することにより、サーバ選択機構からの経路情報の利用を可能とした。図に従って本機構の動作を説明する。

(1) RADIX は各配布用サーバの近傍ルータから経路情報を収集する。

(2) クライアントは配布用サーバの IP アドレスを得るため、ローカル DNS に問合せを送る。

(3) ローカル DNS は、DNS の仕様に従って名前解決を行い、TENBIN へ名前解決要求を送る。

(4) TENBIN は、各配布用サーバからローカル DNS までの経路を RADIX に問い合わせる。

(5) RADIX は(1)で収集している情報をもとに、経路情報を返す。

(6) TENBIN は、得られた経路情報を選択ポリシーに従って評価し、最適な配布用サーバを決定し、その IP アドレスを返す。

(7) ローカル DNS は、得られた IP アドレスをクライアントに返す。

(8) クライアントは、得られた IP アドレスあてに処理要求を送る。

以上により、利用者はサーバ選択機構の存在を意識

することなく、適切な配布用サーバに処理要求を送信することが可能となる。

4. 実運用システムへの適用

筆者らが開発を進めている TENBIN+RADIX を、これまでに以下の実運用システムに対して適用し、その有効性を確認した。

- 2001 年 1 月月食中継におけるストリームサーバ選択
- 2001 年 6 月皆既日食中継における Web サーバ選択・ストリームサーバ選択
- 2001 年 11 月しし座流星群中継における Web サーバ選択・ストリームサーバ選択
- 2002 年 6 月金環日食中継における Web サーバ選択
- RingServer Project [5] における Web サーバ選択・ftp サーバ選択

本論文では、2001 年 6 月日食中継における Web サーバ選択の結果をまとめ考察を行う。

4.1 日食中継

2001 年 6 月 21 日には、今世紀初の皆既日食が起こった。ライブ!エクリプス実行委員会[6]では、この皆既日食をインターネットでライブ中継を実施した。本中継では Web サーバを六つのサイトに広域分散配置した。また、ストリームサーバは 11 の ASP に広域分散配置した。この日食中継において、Web サーバとストリームサーバのサーバ選択に TENBIN+RADIX を適用した。

4.2 日食中継における選択ポリシー

本中継においては、ストリームサーバの選択ポリシーは Web サーバの選択ポリシーのサブセットといえるものであった。したがって、本論文では Web サーバのサーバ選択についてのみ述べる。

基本的な選択ポリシーは単純なものである。各利用者と全 Web サーバ間の距離を比較し、最も近いものを選択する。

しかし、本中継においては、Web サーバ設置サイトによっては、特定の条件に合致する場合にのみコンテンツの提供を行いたいという要望があった。

各サイトからの要求を以下に示す。

- 自サイト (ISP) の顧客、及び、いわゆる下流サイトのクライアントのみにサービスを提供したい。
- 国内のみ、あるいは国外のみのクライアントに対してサービスを提供したい。

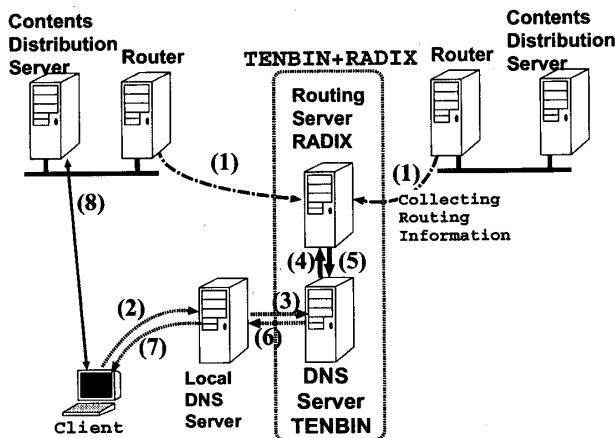


図 4 TENBIN+RADIX によるサーバ選択機構
Fig. 4 Server selection system: TENBIN+RADIX.

(iii) アクセス状況全般を調査するためのサンプリングを行いたい。

これらのうち(iii)について解説を加える。これは、最適なサーバへの誘導を試みるものではない。今回のような大規模イベントにおける、クライアントからのアクセス傾向を調査したいという要望に基づいたものである。また、このルールを加えることで、イベント終了直後に、配信数の速報を可能にすることが期待できる。すなわち、通常は配信数を調査するためには、広域分散配置した配信用サーバすべての配信ログを調査する必要がある。ところが、サンプリングサーバを用意することで、そのサーバのログのみから全配信数を推測することが可能となる。このようなイベントにおいては、配信数などの情報を速やかに公開することが望まれているので、これは非常に重要である。

4.3 日食中継における選択ポリシーの実現

各々のサイトの要望に答えるためのルールを作成した。その際、経路情報を用いることで、これらの要望を実現することができた。最終的に、これら複数のルールの組合せで選択ポリシーを構成した。各ルールの詳細な説明は省略するが、概略を示す。以下のような順でサーバ選択を行う。

前提として、各サイトは RADIX に対して経路情報の提供を行った。ただし、提供する情報は、インターネット全体に対する経路（いわゆる full route）の場合と、そのサイトが配信を希望するネットワークへの経路だけの場合がある。

(1) 固定確率選択

このルールは、「最適」なサーバの選択を行うものではない。前節で述べた要求の(iii)を実現するものである。固定の確率でこのサイトを選択するようにし、サンプリングを行う。本中継においては、この確率は 1/10 とした。

(2) 経路情報問合せ

要求パケットより、クライアントの IP アドレスを知ることができる。この情報をもとに、RADIX に問合せを行い、各サイトからそのクライアントまでの経路情報を取得する。この経路情報には、図 3 に示したように、各サイトからそのクライアントまでの AS パス情報が含まれる。

(3) 特定サイト優先選択

前節の要求のなかの(i)を実現するものである。各サイトは、配信を希望する範囲のネットワークに対する経路情報を RADIX に対して提供する。範囲を制限

しないサイトはフルルートを提供する。すなわち、本節の(2)で取得した経路情報に、あるサイトからあるクライアントへの経路情報が含まれていなければ、そのサイトは、そのクライアントへの配信を希望しないことがわかる。

あらかじめ、(i)の要求をもつサイトについてはリストを作成した。そのリストをもとに、そのサイトからの経路が含まれている場合、他の条件を判断せずに、このサイトを選択した。

(4) オリジン AS に基づく選択

これも、前節の要求のなかの(i)を実現するものである。本節の(2)で取得した情報から、クライアントのオリジン AS が判断できる。この情報をもとに判断する。

経路情報の有無にかかわらず、特定の AS に対する配信を希望したサイトがあり、その要求に答えるためのルールである。

上述の(3)は、サイト側が経路情報の提供の有無によって、選択されるかどうかの制御を動的に行えるという利点をもつが、経路情報の提供が必須になるという欠点をもつ。一方、(4)は経路情報を提供しなくてもよいという利点をもつが、動的制御が困難であるという欠点をもつ。各サイトの事情により、この二つを使い分けた。

(5) 国内・海外選択

前節の要求の(ii)を実現するものである。国内のみ、若しくは海外のみにサービス提供したいというサイトに対応するために適用した。判断材料としては、オリジン AS が JPNIC が発行した AS かどうかで判断した。この選択を行った後で、国内と海外に分けて、以下の(6)と(7)を適用した。

(6) AS パス長選択

今回の選択ポリシーの基本である。クライアントとの距離で判断を行う。「距離」としては、AS パス長を用いた。AS パス長は、本節の(2)で取得した経路情報から求めることができる。最短 AS パス長が 3 以下の場合にのみ、本ルールを適用した。それは、荻野 [7] らによって、AS パス長が長い場合には選択基準として用いることが不適切なことが明らかにされているからである。

(7) ラウンドロビン選択

他のすべてのルールでの選択に失敗した場合には、ラウンドロビンでサーバを選択した。

4.4 日食中継における選択結果と考察

2001年6月19日21:11~2001年6月25日16:00の間のTENBINのログを集計した。この間、日食のライブストリームの中継を行ったのは6月20日20:20~6月20日22:35である。この前後も、ライブストリーム以外の日食情報等のコンテンツは提供された。

4.2で述べた各ルールが適用された数を表1に示す。「経路情報なし」はRADIXが要求元への経路情報をもたなかった場合である。「通信エラー」は、TENBINとRADIX間の通信に失敗した場合である。(1)で述べたように、固定確率選択は最適なサーバの選択を行うものではない。したがって、以下の評価においては固定確率選択を除いて検討を行う。

表2に各ルールの割合を示す。表からわかるように、全処理要求の80.9%を近傍のサーバへ誘導することができた。TENBIN+RADIXを用いることで、ASパス長を基準として、より近傍のサーバに誘導できたことを示している。

残りの19.1%の内訳は以下のようにになっている。

- 最適なサーバがなかったため、海外向けラウンドロビンサーバが選択された：11.8%。
- 最適なサーバがなかったため、国内向けラウンドロビンサーバが選択された：1.7%。

表1 適用された選択ルール
Table 1 Applied selection rule.

選択ルール	回数
ASパス長(海外)	113,149
ASパス長(国内)	70,259
特定サイト優先	22,428
オリジンAS	5,035
ラウンドロビン(海外)	30,779
ラウンドロビン(国内)	4,527
経路情報なし	10,393
通信エラー	4,005
固定確率	28,943
合計	289,518

表2 経路情報によるヒット率
Table 2 Hit ratio using routing information.

選択ルール	回数	割合 [%]
ASパス長(海外)	113,149	43.4
ASパス長(国内)	70,259	27.0
特定サイト優先	22,428	8.6
要求元AS	5,035	1.9
その他	49,704	19.1
合計	260,575	100.0

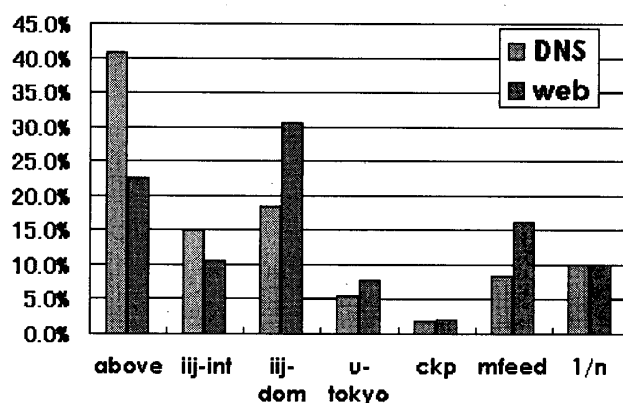


図5 DNSとWebへのアクセス数比較
Fig.5 Access for DNS and Web.

● 経路情報サーバに経路情報がなかったため、海外向けラウンドロビンサーバが選択された：4.0%。

● DNSサーバが経路情報サーバとの通信に失敗したため、海外向けラウンドロビンサーバが選択された：1.5%。

今回の中継においては、ほとんどのサイトが国内に設置された。このため、海外からのアクセスについて最適なサーバを選択できなかった場合が11.8%にもなってしまったと考える。

これら、19.1%の要求に対しては、最適なサーバを選択することができていない。しかし、ラウンドロビン選択によっていずれかのサーバは選択しているため、サービスの提供はできている。

また、日食中継の際に用いたサーバ選択ポリシーの複雑さから明らかなように、TENBINにおける選択ポリシーの柔軟性を示すことができた。

次に、今回TENBINへの名前解決を問合せすべてに関して、全WebサーバとのASパス長について調査した。最小ASパス長はその定義からも明らかであるが1であった。最大ASパス長は21であった。平均ASパス長は3.1であった。一方、TENBIN+RADIXによって誘導に成功した場合の平均ASパス長を求めたところ2.3であった。

4.5 TENBIN+RADIXによる選択結果とWebアクセス結果の比較

本中継における、TENBIN+RADIXでのサーバ選択の結果と、Webサーバへのアクセス数の比較を図5に示す。図中のx軸は今回のWebをサーバを配置したサイトの名称である。このうち、1/nとは、4.3の(1)で述べた固定確率選択を提供したサイトである。

グラフより、1/nサイトに対してはTENBIN

+RADIX で選択したのとほぼ同割合のアクセスが Web サーバに来ていることがわかる。すなわち、固定確率選択を用いた全アクセスのサンプリングはうまく機能したことがわかる。

しかし、その他のサーバについては必ずしも一致していない。傾向としては、海外向けの Web サービスを行ったサイト (above, iij-int) については、TENBIN+RADIX で選択したよりも少数の Web アクセスしか来ていない。一方、国内向けの Web サービスを行ったサイト (iij-dom, u-tokyo, ckp, mfeed) については、TENBIN+RADIX で選択したよりも多数の Web アクセスが来ている。これは、サービス利用者の国内と海外での分散度合の差が原因ではないかと考えている。TENBIN+RADIX は DNS として選択結果を返すので、DNS サーバにより結果がキャッシュされる。国内においては、いくつかの大手 ISP からの利用者が多いことがわかっており、これら ISP の DNS サーバにキャッシュされた結果をもとに Web アクセスを行うため、Web アクセス数が増加する。一方、海外においては、利用者が広く分散されているため、DNS のキャッシュがあまり機能せず、結果として TENBIN+RADIX で選択したよりも Web アクセスが少なくなるのではないかと考えることができる。

このように、固定確率選択とそれ以外で異なる傾向が明らかになった。この違いについては、今後より詳細な調査が必要である。

5. 関連研究

CDN では、配信用サーバを広域に分散配置する。したがって、本論文が対象とする環境も広域分散環境である。現在 CDN は、AKAMAI [8] に代表とされる専門業者や、既存の ISP によって構築が進められつつある。これらの CDN は独自の技術により構築されている。例えば、AKAMAI を利用するためには、コンテンツに対して AKAMIZE と呼ばれる変換を施す必要がある。したがって、異なる CDN 間を接続することが困難である。

本研究は、インターネットで標準的に利用されている技術を基盤として用いるため、それら単独の CDN におけるサーバ選択に適用できるとともに、複数の CDN を相互接続した環境においても適用可能である。このような CDN の相互接続は、今後の高品質コンテンツの提供に必須なものとして、IETF (Internet Engineering Task Force) でも Content Internetworking [1] とし

て積極的に研究開発が進められているものである。

DNS を用いたサーバ選択機構としては、Cluster DNS [9]、Distributed Director [10] や 3-DNS [11] がある。これらは、選択基準として、システム組込みのものしか利用することができない。特に複数の CDN を接続した場合には、運用上の問題より、複雑なポリシーが要求される場合がある。このような場合に対処することが困難となる。一方、本論文での提案手法では、TENBIN に柔軟性の高い選択ポリシーを組み込むことが可能であり、従来のものより広い適用範囲をもつ。

経路情報を用いたサーバ選択機構としては SSLB [12] がある。しかし、SSLB では、経路表からサーバ選択のための表を、システム起動前に静的に作成する。したがって、経路情報の変化に対応することが不可能である。本論文での提案手法は、RADIX が継続的に経路情報を収集し、サーバ選択時にその時点での経路情報を利用する。したがって、経路情報の変化への対応が可能である。

経路情報サーバとしては、zebra [13]、MRT [14] などがある。これらは、他の経路サーバに対して経路情報を再配布するときに、最適経路しか配布できない。RADIX は、すべての経路情報を配布することが可能である。この機能がない場合、サーバ選択に利用可能な情報量が減少してしまう。

6. む す び

本論文では、DNS を基盤として用いた経路情報を選択基準とするサーバ選択機構を開発した。本機構は、サーバを広域に分散配置した CDN の要求誘導基盤として用いることができる。本機構を、実運用システムに適用し、その有効性を確認することができた。本論文で述べた例以外に、本機構を RingServer Project が運営する FTP、WWW サーバにも適用し、評価実験を行っている。このサーバは 30 台を超えるミラーサーバから構成されている、国内でも最大級のソフトウェアライブラリシステムである。本評価実験についての詳細は、別に改めて報告したい。

今後の課題としては、同一 AS 内に複数のサーバがある際のサーバ選択手法の検討がある。

謝辞 本実験を行う機会を与えて下さった、ライブ! エクリプス実行委員会の皆さんに感謝する。本論文を進めるにあたり有意義な議論をして頂いた九州ガガポップ (QGPOP) プロジェクト、日本学術振興会

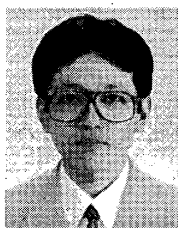
インターネット技術第163委員会(ITRC), WIDEプロジェクトの皆さんに感謝する。

なお, 本論文の一部は平成13年度科学研究費補助金奨励研究(A) 13780249によるものである。

文 献

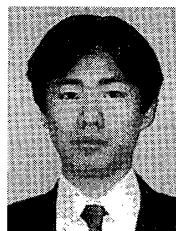
- [1] M. Day, B. Cain, G. Tomlinson, and P. Rzewski, "A Model for Content Internetworking (CDI)," RFC3466, Feb. 2003.
- [2] A. Barbir, B. Cain, F. Douglis, M. Green, M. Hofmann, R. Nair, D. Potter, and O. Spatscheck, "Known CN Request-Routing Mechanisms," draft-ietf-cdi-known-request-routing-02.txt, May 2003.
- [3] 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," 信学論(D-I), J84-D-I, no.9, pp.1396-1403, Sept. 2001.
- [4] 中川郁夫, "経路の抽出を用いた地域内経路制御の実現," ソフトウェア科学会 第一回インターネットテクノロジーワークショップ, Aug. 1998.
- [5] RingServerプロジェクト, <http://www.ring.gr.jp/>
- [6] ライブ!エクリプス実行委員会, <http://www.live-eclipse.org/>
- [7] 荻野 司, 松田和宏, 須藤一顕, 針山欣之, 向阪正彦, 殖栗俊明, "広域分散配置 Web サーバにおける最適サーバ探索システムの検討," 分散システム/インターネット運用技術シンポジウム 2001 論文集, pp.41-50, Feb. 2001.
- [8] Akamai Technologies, Inc., <http://www.akamai.com/>
- [9] V. Cardellini, M. Colajanni, and P.S. Yu, "DNS Dispatching Algorithms with State Estimators for Scalable Web-server Clusters," World Wide Web J., Baltzer Sci., vol.2, no.3, pp.101-113, July 1999.
- [10] CISCO Systems Inc., DistributedDirector, <http://www.cisco.com/>
- [11] F5 Inc., 3-DNS, <http://www.f5.com/>
- [12] SSLB(Streaming Media Server Load Balancer), ネットコム, <http://www.nextcom.co.jp/>
- [13] GNU Zebra, Kunihiro Ishiguro, <http://www.zebra.org/>
- [14] Multi-Threaded Routing Toolkit MRT, Merit Network, <http://www.mrtd.net/>

(平成14年12月9日受付, 15年3月13日再受付)



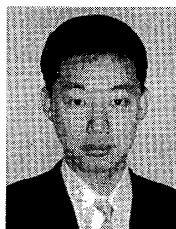
下川 俊彦

平2九大・工・情報卒。平4同大大学院工学研究科情報工学専攻修士課程了。同年(株)東芝入社。平9九大大学院システム情報科学研究科情報工学専攻助手, 平12同システム情報科学研究院情報工学部門助手。平14より九州産業大情報科学部助教授。博士(情報科学)。広域分散協調処理, インターネットなどの研究に従事。情報処理学会会員。



木場 雄一

平11九大工・情報卒。平13同大大学院システム情報科学研究科情報工学専攻修士課程了。同年(株)東芝入社。同社研究所で広域分散システム技術, インターネットに関する研究に従事。



中川 郁夫

平3東工大・理・数学卒。平5同大大学院総合理工学科システム科学先攻修士課程了。同年(株)インテック入社。同社研究所にてネットワーク管理, 大規模経路制御技術, 次世代インターネットに関する研究に従事。平14インテックネットコア, 現在に至る。次世代IXの研究開発に従事。理修。情報処理学会会員。



山本 文治

平7早大・第二文・西洋文化専修卒。同年(株)IIJメディアコミュニケーションズ入社。以来, 同社技術部にて主にストリーミング業務に従事。



吉田 紀彦 (正員)

昭56東大大学院工学系研究科計数工学専攻修士課程了。(株)三菱総合研究所, 九大助手, 同助教授, 長崎大教授などを経て, 平14より埼玉大・工・情報システム・教授。平5~6スタンフォード大客員研究員。博士(工学)。プログラミング方法論, 並列分散協調処理などの研究に従事。IEEE Computer Society Asia-Pacific Activity 委員, 各種国際会議プログラム委員・運営委員などを歴任。International Journal of Cooperative Information Systems 常任編集委員。本会情報システムソサエティ/情報処理学会第1回情報科学技術フォーラム論文賞受賞。情報処理学会, 日本ソフトウェア科学会, ACM, IEEE 各会員。