

インターネットにおけるトラフィック適応型 サーバ選択方式

竹内 大五郎* 小野里 好邦* 山本 潮* 清水 英夫** 木村 博茂**

群馬大学工学部情報工学科*
〒376 群馬県桐生市天神町1-5-1
TEL: 0277-30-1835

東京電力技術開発本部システム研究所通信技術研究室**
〒230 神奈川県横浜市鶴見区江々崎町4番1号
TEL: 045-585-8877

あらまし WWWトラフィックの増加および集中によりインターネット上にホットスポットが発生する問題が生じている。この問題に対処する一つの方法としてミラーサーバの設置による負荷の分散が行なわれているが、その選択方法としてはユーザによる決定またはDNSのラウンドロビン機能による選択方法が取られている。本稿では、インターネットの状態の観測により、適切なミラーサーバへの自動転送方式の提案する。また、シミュレーションモデルを設定し本方式とDNSラウンドロビン方式のアクセス時間の比較を行なった。その結果アクセスを行なうWebページサイズが大きい場合に本方式によるアクセス時間の改善が見られた。

キーワード ミラーサーバ, WWW, トラフィック分散

WWW Server Adaptive Selection Method based on Roundtrip Time

Daigoro TAKEUCHI* Yoshikuni ONOZATO* Ushio YAMAMOTO*
Hideo SHIMIZU** Hiroshige KIMURA**

Dept. of Computer Science, Faculty of Engineering Gunma University*
1-5-1 Tenjin-cho, Kiryu, Gunma, 376 Japan
TEL: 0277-30-1835

Telecommunication Engineering Department, Computer and Communication R&D Center, Tokyo
Electric Power Company**
4-1 Egasaki-cho Tsurumi-ku Yokohama, 230 Japan
TEL: 045-585-8877

Abstract One of the most serious problems on the Internet is occurrence of hot spot by the increase and the concentration of WWW traffic. A way to solve this problem is to establish some mirror servers to be used on behalf of the primary server for load balancing. At present, a mirror server is selected by decision of users or DNS round robin method. In this paper, we propose an automatic selection method of an appropriate mirror server based on the condition of the Internet. We evaluated our method by comparing with the DNS round robin method for access time by using simulation. The result shows our method achieves the improvement of access time in the case Web page size is comparatively large.

Keyword Mirror Server, WWW, Traffic Dispersion

1 はじめに

近年、インターネットにおける WWW はそのグラフィカルなインターフェースと情報への容易さなどにより、急激な成長および発展をとげた。現在日本ではユーザ数 555 万人に達しなおも増加を続けている。また、Web ページ内のコンテンツに画像、音声、動画像等のマルチメディアコンテンツが多用され、Web ページサイズは増加化傾向にある。ユーザ数および Web ページサイズの増加はインターネットにおける WWW トラフィック急増の原因となっている。

しかしながら、インターネットの帯域や Web サーバの回線速度などの資源には制限があるため、トラフィックの増加および集中はネットワーク内にホットスポットを発生させ、WWW ユーザのアクセス環境を著しく悪化させる。このため、トラフィックの分散をおこない、限られた資源を有効的に使用する必要がある。

サーバの設置が行われている。現在、同じ Web ページおよびコンテンツを保持する Web サーバを複数異なる場所に設置することにより、Web サーバの負荷やネットワークトラフィックを分散させることが行なわれて来た。これらの WWW サーバの中心となる Web サーバをプライマリーサーバと呼び、負荷分散のために設置されたサーバをミラーサーバと呼ぶ。ユーザがミラーサーバへのアクセスを行なう際、ミラーサーバの選択はプライマリーサーバに登録されている Web ページにミラーサーバの使用を促すリンクが張られており、そのリンクをユーザがたどることにより行なわれる。しかしながらユーザはもっともネットワーク的に近いミラーサーバを見つけ出すには、リンク先の URL から推定するか実際にミラーサーバにアクセスし Web ページのアクセス時間から体感的に判断するしかない。また、DNS のラウンドロビン機能を利用したミラーサーバの選択方式を採用している Web サーバも存在する [2]。インターネットでは DNS のラウンドロビン機能とはドメイン名 (たとえば `www.nztl.cs.gunma-u.ac.jp`) から IP アドレス (たとえば `133.8.11.4`) を求める際に DNS (Domain name service) を利用するが、この時、ひとつのドメイン名に複数の IP アドレスを DNS に登録しておき、DNS に対して検索要求

があった場合、登録されている IP アドレスを順に検索結果とする機能である。この機能を使用する事により、ユーザのアクセスを各ミラーサーバに分散することができるが、ネットワーク的な近さや、混雑状況を考慮しないため、ユーザのアクセスしたミラーサーバがより快適なサーバである保証はない。

近年 Web アクセスにおける転送をネットワーク的な近さ (ホップ数等) を考慮しミラーサーバの選択を自動的におこなう方式が実用化された [1]。これは、一つの主要なサーバに対して行なわれたアクセスを自動的にネットワーク的に最も近いミラーサーバへ転送する機能を備えている。この方式のシステムは、自律システム内のルーターに実装されたエージェントおよび、クライアントからのアクセスの受け付け、アクセス転送、エージェントからの情報を処理をおこなう装置からなる。自律システム内のルーターに実装されたエージェントが、各ミラーサーバとクライアント間の距離を BGP および IGP のルーティング情報をもとに調査し、この結果からクライアントのアクセスを転送するミラーサーバを選択する。しかしながら、この方式の利用にはエージェント機能を有するルーターが必要であり、現在のネットワーク機器の変更および、これからのネットワーク構築に制限を加えるものであろう。

本研究では、より簡単に、WWW サーバの変更のみでインターネットの混雑状況から適切なミラーサーバの選択をおこなう方式の提案をおこなう。また、シミュレーションにより、本方式と DNS ラウンドロビンとのアクセス時間の比較をおこなう。

2 提案方式

2.1 概要

本方式 (Server Adaptive Selection) では、ICMP を使用し Web サーバ/クライアント間のネットワーク状況をパケットのラウンドトリップタイムを観測 [3] することにより推測し、各クライアントのアクセスすべきミラーサーバの選択を行なう。本方式ではラウンドトリップタイムの観測をサーバ/クライアント間およびサーバ/キャッシングサーバ間の二つの区間に分けておこなう。これはクライ

アントのWWWサーバへのアクセスは通常キャッシングサーバを介して行なわれることが多いためである。これら二つの区間のラウンドトリップタイムを観測しその観測結果からクライアント/キャッシングサーバ間のネットワーク状況を推測し、クライアントがアクセスすべきサーバの選択を行なう。プライマリサーバはクライアントに対し指定先のURLに転送をするHTTPのステータスコード302やMETAタグの含まれたWebページをクライアントに伝送することによりミラーサーバへクライアントのアクセスを転送する。サーバ/クライアント間のネットワークは帯域が十分であっても、クライアント/キャッシングサーバ間が混雑している場合は、ミラーサーバへの転送を行なうことによる伝送時間の改善は見込めない。このような場合は、プライマリサーバ/キャッシングサーバ間のネットワークの状況が良好である場合は、プライマリサーバがWebページデータをクライアントに伝送する。また、サーバ/キャッシングサーバ間が混雑している場合、プライマリサーバは各ミラーサーバに対してミラーサーバ/クライアント間のラウンドトリップタイムの観測要求を行ない、その結果からミラーサーバ/クライアント間のラウンドトリップタイムが最も小さいミラーサーバに転送する。

2.2 動作

以下に本方式における、WWWサーバおよびクライアントの動作を説明する。

- クライアントがプライマリサーバにたいしてWebページの伝送要求を行う。
- クライアントからの伝送要求を受けたプライマリサーバはプライマリサーバ/クライアント間のパケット往復時間 U 、およびプライマリ/キャッシングサーバ間のパケット往復時間 C が観測をおこなう。(図1)

ここで、閾値 K_u, K_c を導入する。 K_u はクライアント/サーバ間が混雑していると判断されるラウンドトリップタイムの閾値、 K_c はクライアント/キャッシングサーバ間が混雑していると判断されるラウンドトリップタイムの閾値である。

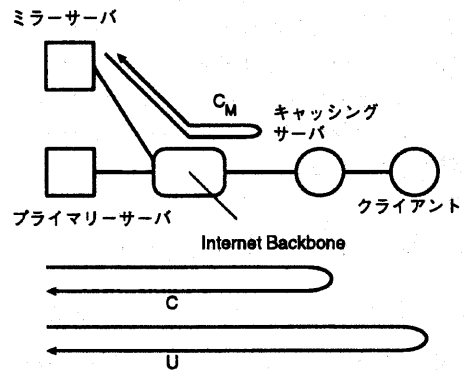


図1: ネットワークの観測

- $U \leq K_u$ または、 $U > K_u, C < K_c$
プライマリからクライアントへWebページデータを伝送
- $U > K_u, C \geq K_c$
ミラーサーバに対して観測要求を行ない適切なミラーサーバに転送を行なう。この動作を以下に示す。
 - Step1: プライマリサーバは各ミラーサーバへミラーサーバ/キャッシングサーバ間の観測要求を行う。
 - Step2: 観測要求を受けた各ミラーサーバはミラーサーバ/キャッシングサーバ間のパケット往復時間の観測 C_M の観測
 - Step3: ミラーサーバからプライマリサーバへの返答
 - Step4: プライマリサーバは最小の C_M を返したミラーサーバへの転送をクライアントに指示するためのWebページの伝送をおこなう。
 - Step5: クライアントがミラーサーバへアクセス (R_M)
 - Step6: ミラーサーバからWebページを伝送 (T_M)

通常 $U > C$ であることから、転送がない場合における、オーバーヘッドは U のみとなることがわかる。転送が発生した場合最低でも、プライマリサーバ

バでのネットワーク観測時間 U 、ミラーサーバでのネットワーク観測時間 C_M 、プライマリサーバ/クライアント間の情報交換 M_S, M_R 、クライアントへ転送を指示する Web ページの伝送 T_t 、およびクライアントがミラーサーバへの要求を行う時間 R_M がかかる。

3 シミュレーションモデル

3.1 シミュレーションモデルの仮定

本方式におけるシミュレーションモデルの仮定を以下に示す。

- プライマリサーバからミラーサーバへの転送要求が送信されて来た場合、クライアントは指定されたミラーサーバに対して伝送要求を行なう。
- サーバから Web ページデータが完全に伝送されるまで、新たな伝送要求は行なわない
- WWW サーバから Web ページデータが完全に伝送されたのち、平均 20 秒の指数分布に従う時間において新たな伝送要求をプライマリサーバに対しておこなう
- Web ページサイズは指数分布に従う

DNS ラウンドロビン方式におけるシミュレーションモデルの仮定を以下に示す。

- クライアントは伝送要求をおこなう際、DNS に対して伝送要求を行なうべき WWW サーバの問い合わせを行ない、DNS により指定された WWW サーバに伝送要求をおこなう
- DNS はクライアントからの問い合わせを受けるたびに、登録されている WWW サーバを順に返す
- WWW サーバから Web ページデータが完全に伝送されるまで、新たな伝送要求は行なわない。
- サーバから Web ページデータが完全に伝送されたのち、平均 20 秒の指数分布に従う時間において新たな伝送要求をプライマリサーバに対しておこなう。

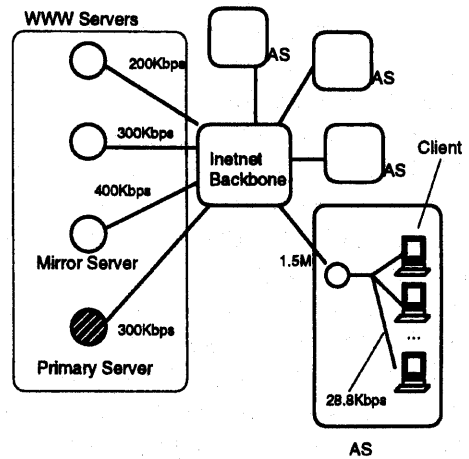


図 2: ネットワークモデル 1

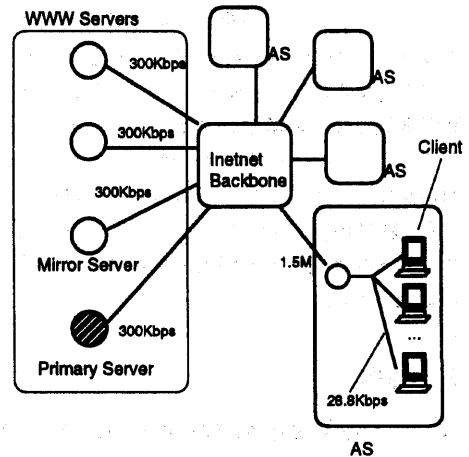


図 3: ネットワークモデル 2

- Web ページサイズは指数分布に従う

3.2 ネットワークモデル

ネットワークモデルを図 2, 図 3 にしめす。各 WWW サーバはインターネットバックボーンにそれぞれ固有の帯域で接続されている。この帯域が小さい場合、サーバ回線能力の不足や、自律システム内のホットスポットの存在を表している。クライアント側の回線はインターネットバックボーン/キャッシングサーバとキャッシングサーバ/クライアント群、クライアント回線の三分割されている。キャッシングサーバはクライアント側の

各自律システムに一つ存在し、自律システム内のクライアントは全てそのキャッシングサーバに接続されているものとする。Web ページデータの伝送は 2048Kbyte のパケットに分割され伝送されるものとする。一つのクライアントへの Web ページデータの伝送において、Web ページデータはパケットとして一つづつクライアントに送られ、一つのパケットがクライアントに到着するまで新たなパケットの伝送は行なわれないものとする。このパケット伝送の動作はウィンドウサイズが常に 1 である場合を表している。

3.2.1 ネットワークモデル 1

ネットワークモデル 1(図 2) は 4 つの WWW サーバと 4 つのクライアントを有する自律システムにより構成される。WWW サーバは帯域 200Kbps, 400Kbps がそれぞれ一つ、300Kbps が二つあり、プライマリーサーバはそのうちの 300Kbps でインターネットバックボーンに接続されている。クライアント側の自律システム内ではインターネットバックボーンからキャッシングサーバ、およびキャッシングサーバから、クライアント群までは 1.5Mbps の帯域で接続されており、各クライアントは 28.8Kbps でアクセスを行なっている。

3.2.2 ネットワークモデル 2

ネットワークモデル 2(図 3) はモデル 1 と同様に 4 つの WWW サーバと 4 つのクライアントを有する自律システムにより構成される。WWW サーバとインターネットバックボーン回線の帯域がすべて 300Kbps であることがネットワークモデル 1 と異なる。

4 数値例

4.1 ネットワークモデル 1

図 4 に Web ページサイズが 40Kbyte の時の、本方式 (Server Adaptive Selection) および DNS ラウンドロビンを使用した場合のアクセス時間 (要求の発生から Web ページデータの伝送完了までの時間) を示す。閾値 K_u , K_c はそれぞれ 2sec, 100msec である。この場合において本方式では DNS ラウ

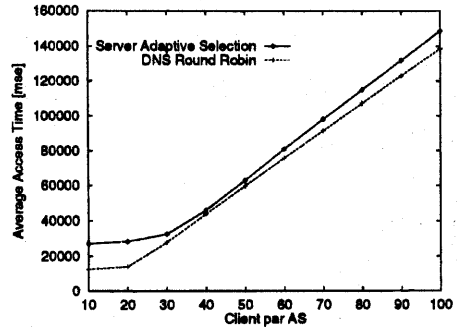


図 4: ネットワークモデル 1 におけるアクセス時間 (Web ページサイズ 40KByte、 $K_u=2\text{sec}$, $K_c=100\text{msec}$)

ンドロビンを使用した場合と比較し、アクセス時間が長くなっている。これは、Web ページデータ小さいため本方式のネットワークの観測のために発生するオーバーヘッド時間が本方式の負荷分散による利得を上回ってしまったためである。また、各自律システムのクライアント数 30 以上の場合において、アクセス時間と各自律システムのクライアント数と線形的な関係が存在することが分かる。これは、負荷の増大によりサーバのクライアントのほぼ全てがいずれかのサーバに対してアクセス中になってしまった状態である。本モデルではクライアント数が有限であると仮定したため、各クライアントは Web ページデータの伝送が完了するまで次のリクエストを行なわない。サーバの回線速度は一定であり、コンテンツ伝送待ちのクライアント数はほぼクライアントの総計となるためアクセス時間と各自律システムのクライアント数と線形的な関係となる。

図 5 にコンテンツデータが 80Kbyte の時の、本方式および DNS ラウンドロビンを使用した場合のアクセス時間を示す。この場合においては DNS ラウンドロビンを使用した場合と比較し、本方式の使用により短い時間でアクセスを行なえることが分かる。これは DNS ラウンドロビンでは各クライアントからのアクセスが処理に時間のかかるインターネットへの回線速度が低速なサーバでの伝送待ち時間が大きな影響を与えてしまうためである。逆に高速サーバへの負荷は非常に小さく、高

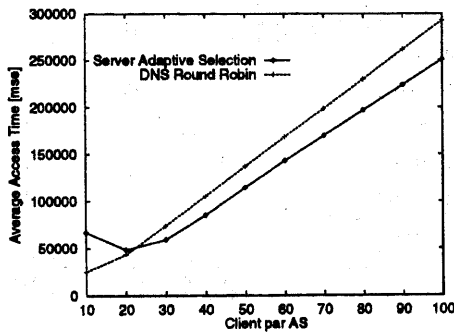


図 5: ネットワークモデル 1 におけるアクセスタイム (Web ページサイズ 80KByte, $K_u=2\text{sec}, K_c=100\text{msec}$)

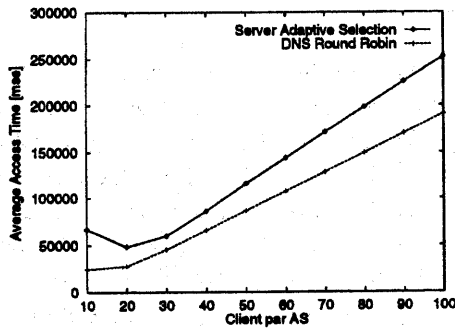


図 6: ネットワークモデル 2 におけるアクセスタイム (Web ページサイズ 80Kbyte, $K_u=2\text{sec}, K_c=100\text{msec}$)

速サーバにアクセスを行なったクライアントのアクセス時間は短い。それに対し、本方式ではインターネットへの回線が高速なサーバ、低速なサーバにともにアクセスが分散される。このことにより、全体的なコンテンツの伝送時間が短縮される。コンテンツデータが大きい程、コンテンツの伝送時間の短縮がオーバーヘッドに対して大きくなるため、DNS ラウンドロビンと比較した際のアクセス時間の短縮につながる。しかし、本方式ではプライマリサーバに最も多くの負荷がかかるため、プライマリサーバの処理能力不足は全体的な能力の低下につながる。

4.2 ネットワークモデル 2

図 6 に各サーバのインターネットへの回線速度が全て 300Kbps である場合の DNS ラウンドロビン方式と本方式によるアクセス時間を示す。平均

Web ページサイズは 80Kbyte である。DNS ラウンドロビンではサーバの回線速度が平均している場合はサーバの回線速度が偏っている場合と異なり、クライアントからのアクセスを各サーバに対し均等に分散が行なわれる。そのため本方式による、トラフィックの分散によっては DNS ラウンドロビンよりアクセス時間を短縮することはできない。ネットワークモデル 2 において本方式では、ネットワーク観測のためのオーバーヘッドがあるため、アクセス時間が DNS ラウンドロビンと比較し増加する。

5 まとめ

本稿ではラウンドトリップタイムの観測により、クライアントを最適なミラーサーバに自動的に転送を行なう方式の提案を行なった。またシミュレーションモデルを設定しシミュレーションを行ない DNS ラウンドロビンとの比較を行なった。その結果、本方式は WWW サーバからインターネットバックボーン間の帯域にホットスポットなどの影響により偏りが存在するモデルにおいては、Web ページサイズが比較的大きい場合に DNS ラウンドロビンと比較しアクセス時間の改善がみられた。しかしながら、Web ページサイズの小さい場合においては、ネットワーク観測のためのオーバーヘッドおよび、観測用パケットのトラフィックのため、DNS ラウンドロビン方式と比較しアクセス時間を増加させる結果となった。今後、WWW サーバプログラムに本方式を実装し、インターネットを使用した実験をおこなうことが課題である。

参考文献

- [1] Cisco System, Inc. *Cisco Distributed Director*. <http://www.cisco.com/warp/public/751/distdir/dd.wp.html>.
- [2] 藤浦, 内藤. トラフィック適応型ネットワーク資源最適配置法の提案. 情報処理学会第 55 回全国大会, 5T-6, 1997.
- [3] 竹内, 小野里, 山本, 清水, 木村. WWW における伝送時間推定に関する一考察. 情報処理学会第 55 回全国大会, 5U-2, 1997.