WWW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[†]
冨沢 考弘[†] サックチャイティプチャクスラット[†]

近年WWWでは，ユーザ数およびコンテンツサイズの増加や新たなサービスの登場により，伝送されるデータ量が急増している．しかし，インターネットの通信回線帯域やWebサーバの能力などの資源には限りがあり，急増するトラフィックを収容できない．特に，トラフィックの集中しやすいインターネットバックボーンへの交換点などでは，ネットワークの処理能力を超えたトラフィックの流入によりホットスポットが発生し，WWWでのアクセス時間を長時間化させてしまう．このため，トラフィックの分散，削減を行いホットスポットの発生頻度を低減させる必要がある．現在WWWではトラフィック分散方法の1つとしてミラーサーバ設置が行われている．しかし，クライアントがミラーサーバにアクセスする際，Webサーバとクライアント間の経路の混雑状況が考慮されることは少ない．本稿では，Webプライマリサーバへアクセスが行われた際，プライマリサーバが最初にサーバとクライアント間の経路の混雑状況を計測し，その結果によりミラーサーバへの転送を行うかを決定し適切なミラーサーバの自動選択を行うTraffic Adaptive Server Selection（TASS）方式について考察する．TASS方式は負荷分散，ホットスポットの回避およびその発生頻度の低減によりユーザのアクセス時間を短縮し，ミラーサーバ選択にともなう煩わしさを解消することを目的とする．シミュレーションにより平均アクセス時間を求め，DNSラウンドロビン方式と比較する．

Traffic Adaptive Server Selection Method for WWW

DAIGORO TAKEUCHI,[†,☆] YOSHIKUNI ONOZATO,[†] USHIO YAMAMOTO,[†]
TAKAHIRO TOMIZAWA[†] and SAKCHAI THIPCHAKSURAT[†]

Since the numbers of users and networks connected to the Internet are rapidly incresing, it may cause the server overload. We call this problem accessibility. To solve this, mirror servers are used to distribute the network loads. In this paper, we present the new server selection method called Traffic Adaptive Server Selection (TASS) to solve the accessibility problem in the Internet. We focus on the round trip time between the server and client, and employ it as the criterion to select the most appropriate server. Primary and mirror servers observe the round trip time between them and the client, and the primary server selects one server based on results of the observations. By using our approach, it is shown that the client will derive the fast service from the server in our simulation results.

1. はじめに

近年，WWW（World Wide Web）はブラウザのGUIによる操作性と全世界からの情報取得の容易性などから急激に発展し，現在もなお成長を続けている．初期にはテキストのみであったWebページも最近では音声，画像，動画像等のマルチメディアコンテンツが組み込まれるようになり，その結果としてWebページサイズが増加傾向にある．ユーザからのアクセス数の増加とページサイズの肥大化がインターネットにおけるWWWトラフィックの増加の原因となり，ネットワークの能力を超えたトラフィックの流入によりホットスポットと呼ばれる局所的な輻輳箇所を発生させる．このホットスポットの発生によりWWWでのアクセス時間が長時間化し，WWWユーザのアクセス環境は著しく悪化する．このような輻輳箇所を回避して，クライアントのサーバに対するアクセス時間の短縮やサーバなどの負荷を軽減させる方法の1つとして，現在ではミラーサーバの設置が行われている．これは，同じWebページおよびコンテンツを保持するWebサーバを異なる場所に複数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を可能にする．ここで，同一コンテンツを持つWWWサーバの中でオリジナルのコンテンツを持つ

† 群馬大学工学部
Faculty of Engineering, Gunma University
☆ 現在，株式会社ナムコ
Presently with Namco Ltd.

サーバをプライマリサーバ，負荷分散のためにコピーを持って設置されたサーバをミラーサーバと呼ぶ．

あるユーザに関してこれらのサーバ群からアクセスするサーバを決定する方法とは，(1) ユーザ自身による手動選択，(2) システムによる自動選択に大別されるが，前者は選択基準の曖昧さやユーザに対して余計な負荷を与えるという問題点がある．後者に関しては，DNS（Domain Name Service）のラウンドロビン機能を利用した方法が提案されている[1]．これは，1つのドメイン名に対して複数のIPアドレスを割り当てておき，そのドメインに対するIPアドレスの照会要求のたびに順次登録されたIPアドレスを返す方式である．しかしながら，この方法は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の状況を考慮しないため，ホットスポットをできるだけ回避してユーザがWebサーバにアクセスすることはできない．森田ら[2]は，ネットニュースやDNSを利用してミラーサーバの情報を配信することにより負荷分散を行う方法について検討しているが，具体的なサーバの決定法については述べられておらず，また配信される内容はネットワークの状況に関する情報ではない．

また，後者の別の方法としてハードウェアレベル，ソフトウェアレベルのそれぞれにおいて商品化されているものもある．しかしながら，これらの方式の共通点として，情報を伝送するのに適切なサーバを決定する権限がサーバ自身ではなく仲介のシステムにあるため，サーバ自身は動いていても仲介システムのトラブルによりサーバが機能しなくなる，またクライアントとサーバの間に仲介システムが存在することによるトラフィックの増加が考えられる．

本稿では，Webサーバ自身がWebサーバとクライアント間の経路の混雑状況を観測し，その結果から各クライアントにとってアクセスに最適なミラーサーバを決定し，そのミラーサーバに自動的に再接続を行うTraffic Adaptive Server Selection（TASS）方式を提案する．TASS方式は負荷分散，ホットスポットの回避およびその発生頻度を低減し，ユーザのアクセス時間を短縮する．また，これらの効果はキャッシングに不向きである動的なコンテンツ，データ量が巨大なコンテンツ，参照頻度の低いコンテンツ等に対しても有効である．

以下，2章では従来のミラーサーバの選択方式に関してその特徴と問題点について述べ，3章では本稿で提案するTASS方式の概念について説明する．4章では2章で述べる従来のミラーサーバ選択方式とTASS方式のアクセス時間についてシミュレーションを行い比較・検討する．最後に，5章で結論を述べる．

2. ミラーサーバの選択方式

図1 に示すように，インターネットを経由して接続されるWebサーバとクライアントのネットワーク構成を仮定する．図中には3つのWebサーバが存在し，そのすべてのサーバに同一のコンテンツが収容されている．また，図1中のWebサーバ1に対して多くのアクセスが発生し，インターネットの一部にホットスポットが発生していると仮定する．このとき，クライアントがWebサーバに対しアクセスを行うことを考える．図1のWebサーバ2のようにサーバ自身へのアクセスは少ないが，その他のトラフィックによりサーバ/クライアント間の経路にホットスポットが発生している場合もアクセスを避ける方が望ましい．結果として，同図中のWebサーバ3のようにホットスポットによって阻害されないサーバにクライアントを接続すればより良いアクセス環境が得られる．このようにインターネットの状態やサーバ回線帯域，サーバの処理能力，サーバ/クライアント間のネットワーク構成などによってWebページデータの伝送速度が速いサーバと，遅いサーバがクライアントごとに存在する．

ネットワークの混雑がどのような状態であろうと，インターネットのように巨大かつオープンなネットワークではクライアントにとって相対的に速いサーバは存在し，そのサーバに接続すればより短時間にアクセスすることが可能である．このようなサーバを選択するための尺度を本稿ではアクセシビリティと呼ぶこ

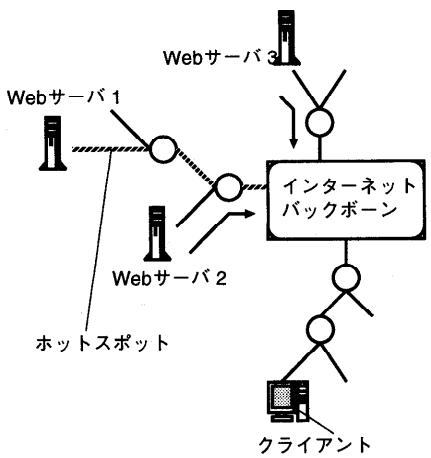


図1 アクセシビリティ
Fig. 1 Accessibility.

とにする．ミラーサーバを設置することで負荷の分散だけでなくアクセシビリティを利用して適切なサーバを選択することによりアクセス時間の短縮を行うことができる．アクセスするサーバを自動的に選択する方式として，DNSのラウンドロビン機能を利用した方式が行われている．また，ハードウェアレベル，ソフトウェアレベルでサーバ選択を行う方式も実用化されている．本章では，これらの方式について簡単に紹介する．特に，ハードウェアレベルの選択方式としてDistributedDirector[3]，ソフトウェアレベルの選択方式としてGlobal Dispatch[4]を紹介する．

2.1 DNSラウンドロビン方式

WWWではコンテンツの場所を示すURL中にWebサーバのホスト名が含まれる．このホスト名からIPアドレスへの変換の問合せにDNSが使用される．DNSラウンドロビン機能を使用し，1つのWebサーバのホスト名にプライマリサーバとすべてのミラーサーバのIPアドレスを割り当てることにより，DNSから返答される最初のIPアドレスはすべてのサーバを順番に指し示す．クライアントはDNSが最初に返答したIPアドレスを持つサーバにアクセスを行うため，アクセスは各Webサーバに分散される．また，DNSラウンドロビン方式では自動的にミラーサーバの選択が行われるためミラーサーバ選択の煩わしさを解消することができる．しかし，DNSラウンドロビン方式による選択ではネットワークの構成や混雑度などの動的状況が考慮されないため，選択されたミラーサーバが適切なサーバである保証はない．

2.2 DistributedDirector

Cisco Systems Inc.により実用化されたDistributedDirector[3]は，インターネットにおけるネットワーク構造のルーティングテーブル情報を使用し，クライアントからホップ数が最も小さいサーバへクライアントを再接続させるシステムである．ここでのホップ数とはクライアントからWebサーバまでの経路にいくつのネットワークが存在するかを表す．

DistributedDirectorはホップ数によるアクセシビリティを考慮することにより，アクセス時間，コストなどを削減する方式である．DistributedDirectorはDRPサーバエージェントを必要とし，DRPサーバエージェントは現在のところCisco社製のルータのみで動作する．このため，DRPの標準化が行われ，各メーカに採用されないかぎりはDistributedDirectorの導入に際しルータの変更等の設備投資が必要となる．

2.3 Global Dispatch

Resonate Inc.により実用化されているGlobal Dispatch[4]は，適切と思われるサーバを選択するために3つの評価基準，クライアント/サーバ間のネットワーク遅延，現在のサーバ状況，サーバの可動性を考慮する．これらの評価基準により選択したサーバに対して，クライアントは直接アクセスを行う．評価基準の1つであるネットワークの遅延は，クライアントから要求を受けるたびにすべてのサーバについて測定する．このソフトウェアが動作するサーバは仲介システムとしてクライアント/サーバ間に配置されるため，クライアント/サーバ間のトラフィックの増加が考えられる．またこの仲介システムに何らかのトラブルが発生した場合は，Webサーバ自身が動作していてもサーバとして機能しなくなる．

3. TASS方式

本論で考察するTASS方式は，仲介システムを存在させずにWebサーバ自身がミラーサーバへの転送の判断と適切なミラーサーバの決定を行う方式である．現在，クライアントからインターネットへのアクセスはWebサーバとクライアント間にあるキャッシングサーバを介して行われることが一般的となっている．キャッシングサーバはWebサーバとクライアント間に介在し，一度クライアントがWebサーバにアクセスし伝送されてきたコンテンツデータを記録（キャッシュ）する．そして再び同一のWebページにクライアントがアクセスを行ったとき，キャッシングサーバはすでにキャッシュされているコンテンツデータをクライアントに伝送する．キャッシングサーバにキャッシュされているコンテンツデータはWebサーバへリクエストされないためWebサーバの負荷は低減し，アクセス時間は短くなる．TASS方式ではプライマリサーバへアクセスが行われた際，プライマリサーバはプライマリサーバ/キャッシングサーバ間，およびプライマリサーバ/クライアント間の2つのパケットの往復時間（**ラウンドトリップタイム**）を計測し，その結果からミラーサーバへの転送が有効か判断し，有効であるならミラーサーバへの転送を行う．転送先のミラーサーバの決定は各ミラーサーバとキャッシングサーバ間のラウンドトリップタイムを考慮して決定される．

3.1 混雑状況の測定方法

インターネット中のトラフィックを正しく測定し解析することができれば，ネットワークの使用率などから混雑状況を推測することができる．しかし，インターネットには非常に多くの組織が相互接続されているために集中したトラフィックの測定および管理を行うことは不可能である．また，インターネットでは現

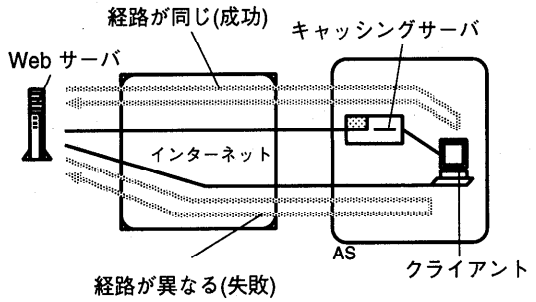


図2 ICMP による測定例
Fig. 2 The examples of ICMP based observation.

在膨大な量のパケットが転送されており，トラフィック状況は刻々と変化している．したがって，ネットワークのトラフィックおよび混雑の詳細な測定は非常に難しく，また解析にも時間がかかる[5]．TASS 方式では「短時間」かつ「そこそこな精度」でインターネットの混雑状況を計測する必要がある．

ネットワークは回線速度や交換機の能力など，ネットワーク資源の能力から伝送可能な帯域には上限が存在する．一時的な超過トラフィックはルータなどに流入した際バッファに蓄えられ，資源が利用可能になるに従い伝送される．バッファに蓄えられた場合，資源が利用可能になるまでの時間は伝送遅延時間となる．このことから伝送遅延時間が大きい場合はネットワークが混雑していると考えられる．そこで，TASS 方式ではネットワークの 2 点間のラウンドトリップタイムから，混雑状況を把握する．

具体的な測定方法を述べる．まず，Web サーバとキャッシュサーバ間，および Web サーバ/クライアント間のラウンドトリップタイムを計測する．計測方法は UNIX 等で実装されている ping コマンドと同様に ICMP を使用し計測する．しかし，これでは，**図 2** のように，計測経路と Web ページのアクセス経路が異なる場合が発生する．そこで，IP の LRS オプション[6]を使用する．LRS オプションの使用により IP パケットを任意の中継点を経由し伝送を行うことができる．そして，中継点としてキャッシングサーバを指定することで，Web サーバからクライアントまでのラウンドトリップタイムをより正確に測定できる．

3.2 TASS 方式で想定するネットワーク

今日インターネットは多くの独自に運営されるネットワークである**自律システム（AS）**[6]が相互に接続され，インターネット全体を構成している．自律システムは LAN，WAN およびそれらを結ぶルータの集合であり，1 つの管理機関の管理下にある．また，交換ポイント（Exchange Point）と呼ばれる自律システム間を

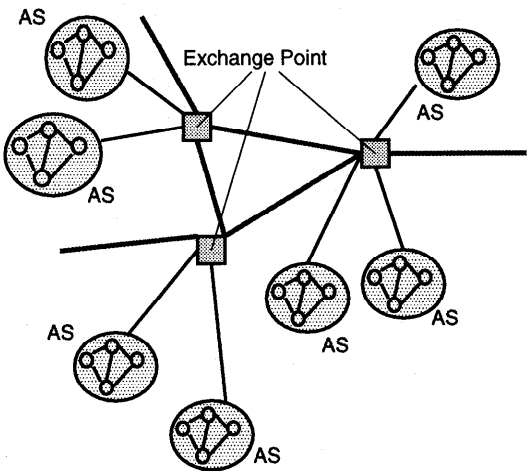


図3 交換ポイントを使用したインターネット
Fig. 3 Internet with exchange points.

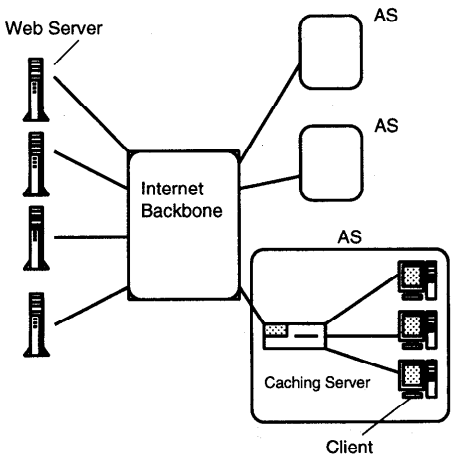


図4 TASS 方式で想定するネットワーク
Fig. 4 Network model using TASS.

相互に接続するポイントの設置と，交換ポイント間を相互接続するインターネットバックボーン（Internet Backbone）という形態が進められた．

図 3 のように自律システムから交換ポイントへの接続，交換ポイント間の相互接続という接続形態をとる．

図 4 に示すように，TASS 方式で想定するネットワークはインターネットバックボーンにいくつかの Web サーバおよび自律システムが接続されている．自律システム内にクライアントは設置されており，クライアントは自律システム内のある 1 つのキャッシングサーバを介して WWW にアクセスする．ここで，Web サーバが N 台接続されていた場合，それぞれの Web サーバに 0 から $N-1$ まで番号をつける．ただし，0 番目の Web サーバはプライマリサーバを表す．

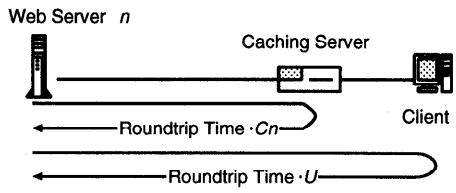


図5 ラウンドトリップタイム U と C_n
Fig. 5 Roundtrip time U and C_n.

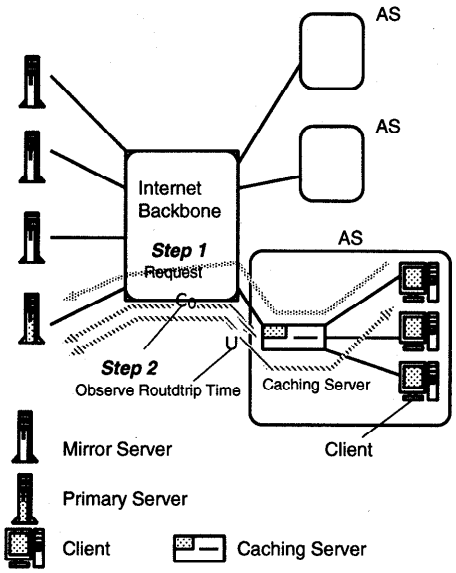


図6 TASSの動作 (Step1～2)
Fig. 6 TASS-step execution (Step1～2).

図5においてWebサーバ n $(n=0,1,\cdots,N-1)$ からキャッシングサーバ間のラウンドトリップタイムを Cn と定義し，また，プライマリサーバからクライアントまでのラウンドトリップタイムを U と定義する．

3.3 TASS方式の動作

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．

Step1：クライアントがプライマリサーバに対してWebページの伝送要求を行う（**図6** 中のStep1）．

Step2：プライマリサーバはラウンドトリップタイム C_0 および U を観測を行う（図6中のStep2）．

ここで，閾値 Ku，Kc を導入する．Ku はWebページの提供者が決定する値であり，Webサーバ/クライアント間が混雑していないと判断される最大のラウンドトリップタイム U である．また，Kc はWebサーバ/キャッシングサーバ間が混雑していないと判断される最大の C_0 である．これら2つの閾値と観測された U，C_0 から，次の4つの状態が存在する．

この Ku の導入によりミラーサーバへの問合せに要する時間，ミラーサーバのネットワーク観測時間を特定の場合において省略することができる．ミラーサーバが高負荷である場合やTCPのコネクションセットアップ時間が非常に長い場合など，ミラーサーバの応答までの時間が非常に大きくなるときに有効である．

状態1 $U \le Ku$ かつ $C_0 \le Kc$

状態1ではサーバ/クライアント間およびキャッシングサーバ/クライアント間の回線はともに混雑していない．クライアントにとりプライマリサーバより状態の良いミラーサーバが存在する可能性はあるが，プライマリサーバによるデータの伝送は良好であり，目的の時間内でのWebページ伝送が期待できるため，プライマリサーバがクライアントに対してデータの伝送を行う．

状態2 $U \le Ku$ かつ $C_0 > Kc$

状態2では，サーバ/キャッシングサーバ間は比較的混雑しているが，サーバ/クライアント間は混雑していない．状態2において，プライマリサーバより状態の良いミラーサーバが存在する可能性はあるが，プライマリサーバによるデータ伝送は目的の時間内での伝送完了を期待できるため，プライマリサーバがクライアントに対してデータを伝送する．

状態3 $U > Ku$ かつ $C_0 \le Kc$

状態3は，サーバ/クライアント間の回線が混雑しているため，目的の時間内にデータの伝送することは期待できない．しかし，サーバ/キャッシングサーバ間の回線は良好である．このことは，混雑がキャッシングサーバ/クライアント間で発生していることを示している．したがって，ミラーサーバへの転送を行ったとしても伝送時間の改善は期待できない．したがって，プライマリサーバがクライアントに対してデータを伝送する．

状態4 $U > Ku$ かつ $C_0 > Kc$

状態4ではサーバ/クライアントおよびサーバ/キャッシングサーバ間がともに混雑している．この状態では，プライマリサーバによる伝送では目的の時間内にデータを伝送することは期待できない．しかし，ミラーサーバに接続することにより，サーバ/キャッシングサーバ間の混雑が解消される可能性がある．この場合のみプライマリサーバはミラーサーバへアクセスを転送することを試みる．

状態1，2，3の場合，プライマリサーバがデータの伝送を行い，TASS方式の動作は終了する．この場合，ラウンドトリップタイム U と C の観測は同時に行われ，また通常 $U > C$ あるため，観測のためのオーバ

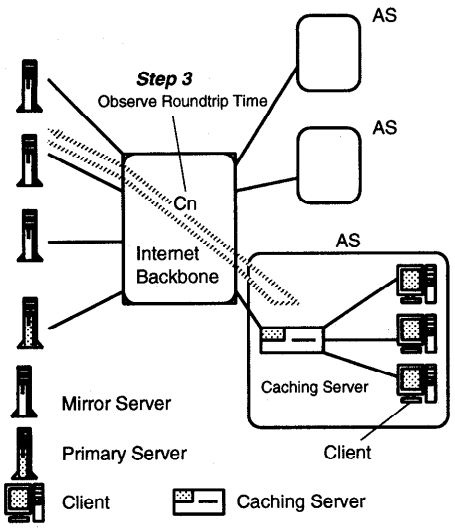


図7 TASSの動作（Step3）
Fig. 7 TASS-step execution (Step3).

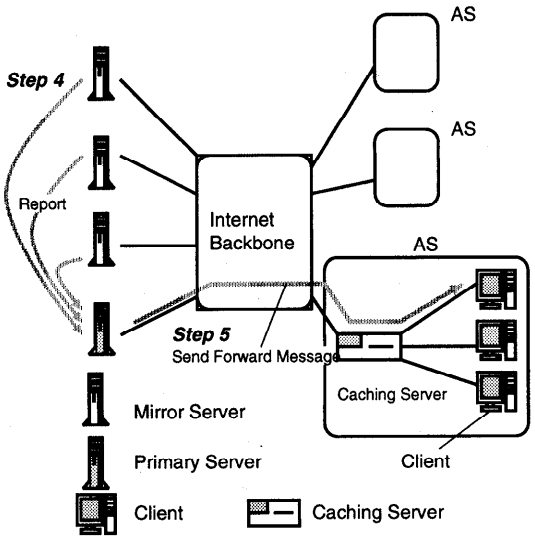


図8 TASSの動作（Step4～5）
Fig. 8 TASS-step execution (Step4～5).

ヘッド時間は U であることが分かる．

状態4の場合はStep3以降の動作を行う．

Step3：プライマリサーバは全ミラーサーバに対して Cn の観測要求を行う．これを受けて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ラウンドトリップタイム Cn の観測を行う（**図7**）．

Step4：ミラーサーバは観測を行ったラウンドトリップタイムをプライマリサーバに報告する（**図8**中のStep4）．

Step5：プライマリサーバは最小の Cn を報告したミラーサーバ min を選択し，そのミラーサーバへの転送要求をクライアントに対して伝送する（図8中のStep5）．この転送要求はHTTPの302コード，あるいは転送命令含むMETAタグを使用したWebページを伝送することにより実現される．

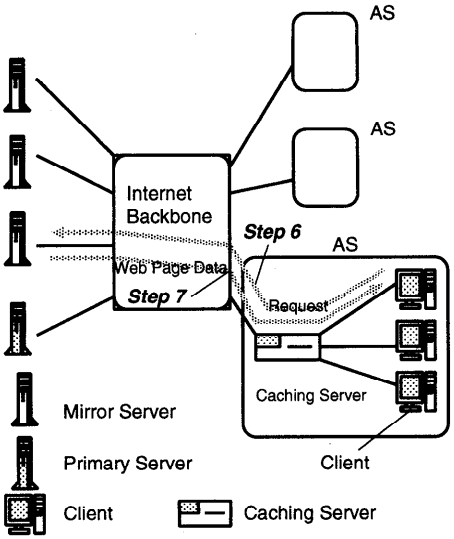


図9 TASSの動作（Step6～7）
Fig. 9 TASS-step execution (Step6～7).

Step6：クライアントは選択されたミラーサーバ min に対してリクエストを行う（**図9**中のStep6）．

Step7：ミラーサーバ min はクライアントに対してデータの伝送を行う（図9中のStep7）．

以上で状態4の場合，TASS方式の動作は終了する．

3.4 メディアスケーリングへの応用

TASS方式は適切なサーバに転送を行うことでアクセス時間の短縮を行うが，インターネットの混雑度やサーバ/クライアントの場所によっては目的の時間以内にWebページの伝送を行えないことも考えられる．この場合，メディアスケーリング[7]を行うことにより，Webページを目的の時間内に伝送できる可能性がある．メディアスケーリングとはコンテンツの品質と伝送量のトレードオフを行うことができる仕組みである．画像におけるメディアスケーリングは画質（解像度，色の再現性）と伝送量におけるトレードオフであり，画像サイズを縮小することなどで実現できる．また，動画像圧縮技術の1つであるMPEG2[8]はこのメディアスケーリングを効率良く行うために，動画像の1つのストリーム内に基本的情報を含む基本レイヤとその基本レイヤを元に，より高品質化を行う高位レイヤを含ませている．MPEG2は基本レイヤだけでも再生は可能であるので，高位レイヤの伝送をするかしないか

で伝送量と品質のトレードオフが達成できる[8]．また，1つのWebページに含まれるコンテンツの数を増減させることによって，Webページ全体のデータ量とWebページ含まれる情報量のトレードオフを行う方式も提案されている[9]．しかしながらユーザはメディアスケーリングにより低品質化されたコンテンツより高品質なコンテンツを好み，またアクセス時間はより短いことを望む．このため，ユーザが許容できるアクセス時間内に伝送可能である最も品質の高いコンテンツを伝送することが望まれる．

TASS方式では観測されるサーバ/クライアント間のパケットのラウンドトリップタイムUがKuを超えると目的の時間内にそのWebページを伝送することは期待できない．そこで，メディアスケーリングによるWebページの情報量の削減やコンテンツ品質の劣化なしでクライアントへWebページを伝送可能な他のサーバへの転送を試みる．しかし，ユーザ回線の帯域がボトルネックになっている場合などは他のサーバに転送を行っても伝送時間が改善される見込みは少ない．このとき，メディアスケーリングを行い，Webページのデータ量の削減を行うことが考えられる．

メディアスケーリング自体の技術がすでにある場合には，TASS方式の動作のStep2，Step6を以下のStep2'，Step6'に変更することで実現される．

Step2'：

- $U > Ku$ かつ $C_0 \leq Kc$
  コンテンツあるいはWebページにメディアスケーリングを施しデータ量の削減を行ったのち，プライマリサーバがクライアントにWebページデータの伝送を行う．

Step6'： $C_{min} + (U - C_0) > Ku$ であればコンテンツあるいはWebページにメディアスケーリングを施しデータ量の削減を行ったのち，ミラーサーバ min がクライアントにWebページデータの伝送を行う．

4. シミュレーション

本章では，シミュレーションモデルを設定し，TASS方式およびDNSラウンドロビン方式を使用した場合のそれぞれの平均アクセス時間を求める．ここで平均アクセス時間は，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．さらに，Kuの値による影響についても考察を行う．

4.1 シミュレーションモデルの仮定

シミュレーションモデルの仮定を以下に示す．

クライアントの動作

- クライアントは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Webページへのアクセスは行わない．
- WWWサーバからWebページデータが完全に伝送されたのち，指数分布に従う時間をおいて新たなアクセスを行う．
- TASS方式において，クライアントはWebページリクエストを初めにプライマリサーバに対して行う．
- 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再度伝送要求を行う．

Webページサイズ

- Webページサイズは指数分布に従う．

DNSサーバ

- クライアントは伝送要求を行う際，DNSに対して最初にリクエストを行うべきWWWサーバの問合せを行い，DNSにより指定されたWWWサーバに伝送要求を行う．
- DNSはクライアントからの問合せを受けるたびに，登録されているWWWサーバを順に返す．

その他，Webサーバの動作はTASSの動作に従う．

4.2 ネットワークモデル

ネットワークモデルでは各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，サーバの接続されているプロバイダー回線中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2つに分割されている．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1つ存在し，自律システム内のクライアントはすべ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．回線はすべてFIFOの待ち行列とする．Webページデータの伝送は2048 byteのパケットに分割され伝送されるものとし，1つのパケットが待ち行列の1つのジョブとする．このため，各回線に対応する待ち行列のサービス時間S[sec]は回線の帯域をW[bps]とすると，$S = \frac{2048 \times 8}{W}$となる．1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1つずつクライアントに送られ，1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われないものとする．これはTCP/IPにおけるウインドウサイズがつねに1である場合を表している．ま

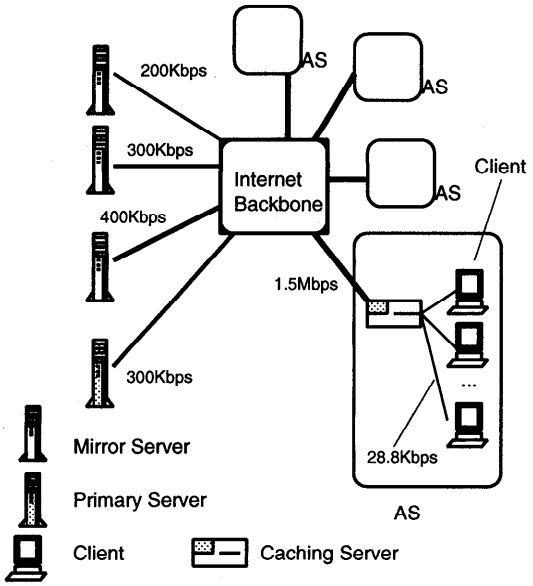


図 10 ネットワークモデル 1
Fig. 10 Network model 1.

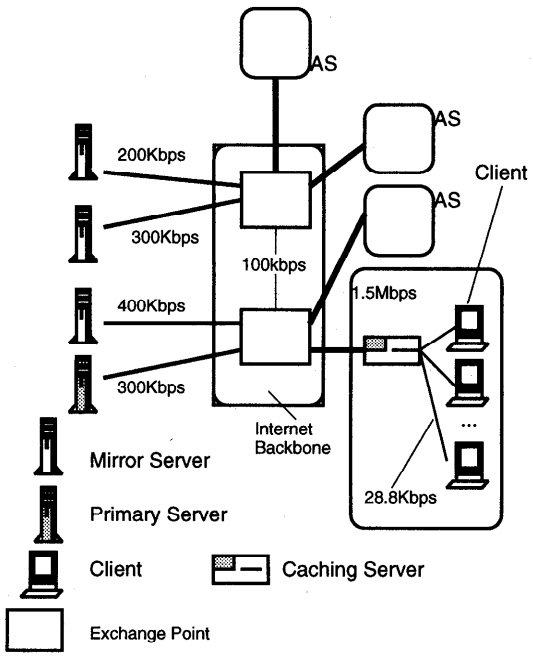


図 11 ネットワークモデル 2
Fig. 11 Network model 2.

た，TCP/IP のコネクション設定時間は考慮しない．

4.2.1 ネットワークモデル 1

ネットワークモデル 1（図 10）は 4 つの WWW サーバとクライアントを有する 4 つの自律システムにより構成される．WWW サーバはインターネットバックボーンへ接続される帯域が 200 kbps，400 kbps が 1 つ，300 kbps が 2 つあり，プライマリサーバはそのうちの 300 kbps である．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は 1.5 Mbps の帯域で接続されており，各クライアントは 28.8 kbps でキャッシングサーバに接続されている．このモデルは各サーバのインターネットバックボーンへの帯域に偏りを作ることにより各サーバのアクセシビリティを発生させている．

4.2.2 ネットワークモデル 2

ネットワークモデル 2（図 11）はネットワークの相互接続ポイント間のボトルネックを示したモデルである．Web サーバはそれぞれネットワークモデル 1 と同じ回線帯域で接続されている．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は 1.5 Mbps の帯域で接続されており，各クライアントは 28.8 kbps でキャッシングサーバに接続されている．

また相互接続ポイントは 2 つあり，Web サーバと自律システムはそれぞれ 2 つずつ相互接続ポイントに接続されている．相互接続ポイント間の回線は混雑状態を想定し 100 kbps に設定する．

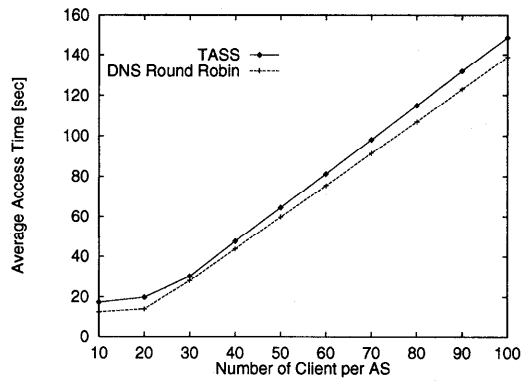


図 12 ネットワークモデル 1 における平均アクセス時間（平均 Web ページサイズ 40 kbyte，$Ku=2$ sec，$Kc=100$ msec，クライアント平均待ち時間 20 sec）
Fig. 12 Average access time in network model 1 (average Web page size 40 kbyte, $Ku = 2$ sec, $Kc = 100$ msec, client average waiting time 20 sec).

4.3 数 値 例

4.3.1 DNS ラウンドロビンとの比較

まずネットワークモデル 1 について数値例を示す．図 12 は TASS 方式および DNS ラウンドロビン方式を使用した場合の平均アクセス時間と AS あたりのクライアント数のグラフである．平均 Web ページサイズは 40 kbyte，閾値 Ku，Kc はそれぞれ 2 sec，100 msec，ユーザの平均待ち時間は 20 sec である．この場合では，TASS 方式では DNS ラウンドロビン方

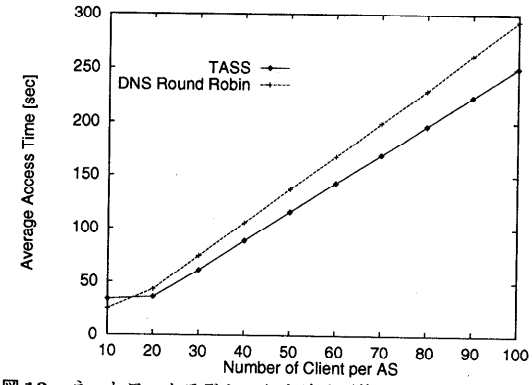


図13 ネットワークモデル1における平均アクセス時間（平均Webページサイズ80 kbyte, $Ku = 1000$ msec, $Kc = 100$ msec, クライアント平均待ち時間20 sec）

Fig. 13 Average access time in network model 1 (average Web page size 80 kbyte, $Ku = 1000$ msec, $Kc = 100$ msec, client average waiting time 20 sec).

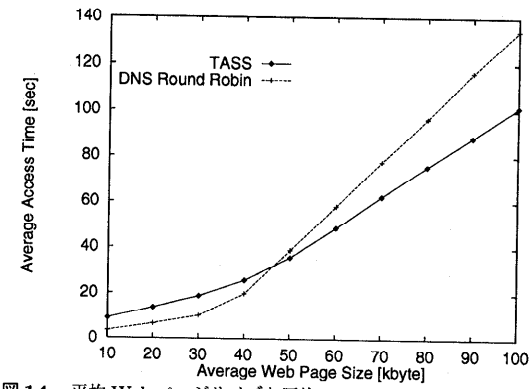


図14 平均Webページサイズと平均アクセス時間（ASあたりのクライアント数50, $Ku = 1000$ msec, $Kc = 100$ msec, 平均クライアント待ち時間60 sec）

Fig. 14 Average Web page size and average access time (number of client per AS 50, $Ku = 1000$ msec, $Kc = 100$ msec, client average waiting time 60 sec).

式を使用した場合と比較し，平均アクセス時間が長くなっている．

TASS方式ではインターネットの混雑状況を観測するためのオーバヘッドが発生する．TASS方式はWebページの伝送時間自体はDNSラウンドロビン方式より小さくなっているが，伝送されるWebページサイズが小さいためその差は小さい．このためネットワーク観測のために発生するオーバヘッドがTASS方式により短縮されたWebページの伝送時間より大きくなってしまうため，平均アクセス時間はDNSラウンドロビン方式よりTASS方式の方が長くなる．ASあたりのクライアント数30以上の場合において，平均アクセス時間とASあたりのクライアント数とに線形的な関係が存在することが分かる．負荷の増大によりアクセス時間が長くなると，クライアントのほぼすべてがいずれかのサーバに対してアクセス中になる．この場合，回線の使用率はほぼ100%となる．サーバの回線帯域は一定であり，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，平均アクセス時間とASあたりのクライアント数とは線形的な関係となる．

図13 に平均Webページサイズが80 kbyteのときの，TASS方式およびDNSラウンドロビン方式を使用した場合の平均アクセス時間を示す．この場合においてはDNSラウンドロビン方式を使用した場合と比較し，TASS方式を用いることでより短い時間でアクセスを行えることが分かる．これはDNSラウンドロビン方式では各クライアントからのアクセスが帯域の狭い低速サーバに対しても均等に分散されるため，低速サーバにおける伝送時間が増加し，システム全体のアクセス時間に大きく影響するからである．それに対しTASS方式では高速サーバには多く，低速サーバには少なくアクセスが分散される．このことにより，全体的なWebページの伝送時間が短縮される．Webページサイズが大きいほど，Webページ伝送時間の短縮効果がオーバヘッドに対して大きくなるため，DNSラウンドロビン方式と比較した際のアクセス時間の短縮につながる．

図14 は平均Webページサイズの違いによる平均アクセス時間への影響を示したグラフである．図14より，このモデルにおいてはTASS方式は平均Webページサイズが約45 kbyte以上で有効であることが分かる．

次にネットワークモデル2について数値例を示す．**図15** はTASS方式およびDNSラウンドロビン方式を使用した場合の平均アクセス時間のグラフである．平均Webページサイズは40 kbyteである．ネットワークモデル2ではインターネットバックボーン間に狭帯域回線が存在する．Webページデータがこの狭帯域回線を経由しない経路で伝送がされるようなWebサーバにクライアントがリクエストを行えば，より快適にアクセスができる．しかしながら，DNSラウンドロビン方式ではアクセシビリティを考慮しないためインターネットバックボーン中の狭帯域回線を含む経路でWebページ伝送が確率0.5で行われる．TASS方式ではこの狭帯域回線を含むWebサーバ/キャッシングサーバ間のパケットのラウンドトリップタイムは大きくなるため，クライアントはWebページデータが狭帯域回線を経由せずに伝送される最寄りのWebサーバに転送される．このため，TASS方式はDNSラウ

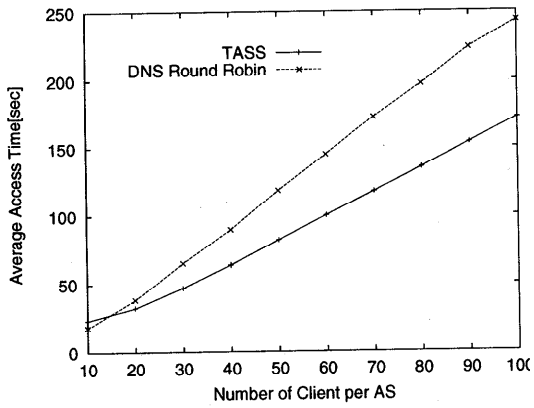


図15 ネットワークモデル2における平均アクセス時間（平均Webページサイズ40 kbyte, Ku=2 sec, Kc=100 msec, クライアント平均待ち時間20 sec）

Fig. 15 Average access time in network model 2 (average Web page size 40 kbyte, Ku = 1000 msec, Kc=100 msec, client average waiting time 20 sec).

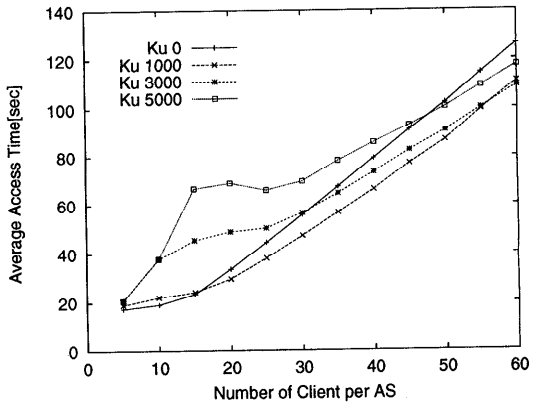


図16 ネットワークモデル2における平均アクセス時間（平均Webページサイズ50 kbyte, Kc=0 msec, クライアント平均待ち時間20 sec）

Fig. 16 Average access time in network model 2 (average Web page size 50 kbyte, Kc=0 msec, client average waiting time 20 sec).

ンドロビン方式より平均アクセス時間を短縮できる．

最後にモデル1とモデル2の結果について比較する．モデル1はクライアントと各サーバ間のネットワークの状況にあまり違いがないので，サーバ群がLAN内に配置されるような場合としてみることができる．それに対してモデル2は，サーバが広範囲に分散されている場合としてみることができる．これにより，本方式においては特にサーバを距離的に離れた位置に配置した場合に対して効果が得られる．図12および図15の結果から，モデル2の方が効果的に動作していることが分かる．

4.3.2 Ku の影響

ネットワークモデル2においてTASS方式を使用した場合の平均アクセス時間のグラフとオーバヘッドのグラフを**図16**，**図17**に示す．ここで示すオーバヘッドとは観測要求，結果報告，ラウンドトリップタイム観測等，サーバを選択する際に必要な時間である．$Ku = 0$ のときは，必ずミラーサーバに転送するため，プライマリサーバはWebページデータの伝送を行わない．ゆえに，プライマリサーバはアクセスが行われるたびに毎回ミラーサーバへ観測要求を行うので，オーバヘッドによるトラフィックが最も増加する．ネットワークモデル2ではこのトラフィックにより相互接続ポイント間の回線が混雑しやすい．相互接続ポイント間の回線が混雑すると，今回のように制御パケットにタイムアウトを導入しない場合は，観測要求と結果報告に時間がかかりアクセス時間に影響を及ぼす．またタイムアウトを導入すると，時間内に結果報告を返

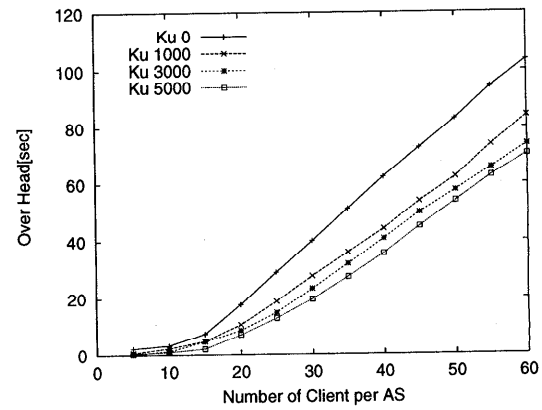


図17 ネットワークモデル2におけるオーバヘッド（平均Webページサイズ50 kbyte, Kc=0 msec, クライアント平均待ち時間20 sec）

Fig. 17 Over head in network model 2 (average Web page size 50 kbyte, Kc=0 msec, client average waiting time 20 sec).

したサーバの中に適切なサーバが含まれない状況が生じる．

Ku を大きくするとミラーサーバへの観測要求が減少するためオーバヘッドが減少する．これにより相互接続ポイント間の回線の混雑を多少おさえることができるが，必要以上に大きくしてしまうと，プライマリサーバが伝送することが多くなるため，今度はプライマリサーバ回線が混雑してしまい，アクセス時間が増加してしまう．

閾値の設定法については本論では明確に定義を行わ

ないが，ある初期値を閾値として設定し，統計的結果から閾値を変化させていく方法などが考えられる．ただし，閾値の設定法についてはさらに検討していく必要がある．

4.4 考　　察

シミュレーション結果からTASS方式はアクセシビリティを考慮することによりアクセス時間を短縮することが可能であることが確認された．TASS方式はデータ量が肥大化傾向にある近年のWebページ伝送に大きく役立つと考えられる．しかしTASS方式の問題点として観測にかかるオーバヘッドがあげられる．DNSラウンドロビン方式はアクセシビリティを考慮しないためオーバヘッドがなく，またDistributedDirectorはDNSキャッシングモードでは，クライアントが一度アクセスを行えば，そのクライアントにとって最もネットワークの構成上近いとされる（ホップ数が少ない）サーバのIPアドレスがローカルDNSにキャッシュされる．このため最も近いWebサーバへのIPアドレスのキャッシュが存在する間，リクエスト開始からWebページデータ伝送までのオーバヘッドをなくすことができる．

TASS方式は伝送されるデータ量が小さい場合はオーバヘッドのためDNSラウンドロビン方式よりアクセス時間が長くなることがシミュレーションにより確認された．しかし，オーバヘッドが発生するのはプライマリサーバに対してクライアントがアクセスを行ったときのみである．このことは，一度ミラーサーバに転送された後，ミラーサーバ内のリンクをたどるのであれば，オーバヘッドはかからず，アクセシビリティの考慮によるアクセス時間短縮の効果も継続して得られることを示している．通常Webページの制作者は複数のページをリンクし網型，階層型，線形型の構造を作成する．このため，ユーザは同一のサーバに対し複数回のリクエストを行うことが多い．このようなユーザの動作はTASS方式にとって非常に有利である．今回シミュレーションを行ったネットワークモデルのようにWebページへのアクセスごとに必ずプライマリにリクエストするユーザ動作はTASS方式にとって最悪のケースである．アクセスログからユーザ動作をシミュレートしたトレースシミュレーションを行えばさらなるアクセス時間の短縮が得られるものと思われる．これは今後の課題である．

また，実際にWebサーバプログラムにTASS方式を実装する際，観測パケット自体の損失による問題も予想される．観測パケットがネットワーク中で損失するとクライアントは永久にWebデータを得られない．このため，観測パケットのタイムアウトを導入する必要がある．観測パケットがタイムアウトした場合その回線にホットスポットが発生したと見なし，動作を継続することで，Webデータ伝送のデッドロックは回避できる．

プライマリサーバ/ミラーサーバ回線の混雑はより深刻な問題を引き起こす．プライマリサーバからミラーサーバへの問合せやミラーサーバからプライマリサーバへのリポートが阻害された場合，プライマリサーバはそのミラーサーバへの転送は行えなくなる．

また，DNSラウンドロビン方式は負荷分散やそれにともなうアクセス時間の短縮だけでなく，可動性を向上させる手段として用いられている．DNSラウンドロビン方式の場合，1台のWebサーバが故障により使用不可能であったとしても，その他のサーバは通常に使用できる．ユーザが故障したサーバにリクエストを行うこともあるが，数回リクエストを繰り返すことにより，ユーザは正常なサーバに接続することができる．一方，TASS方式ではプライマリサーバの故障により，システム全体がダウンしてしまう．このため，TASS方式における可動性はシングルサーバの場合と同等になる．

TASS方式は，すべてのサーバが均等に伝送するのではなく，プライマリサーバが優先的に伝送し，伝送効率が悪化してきた時点で初めて他のミラーサーバが伝送する方式である．そのため，TASS方式は頻繁に情報を更新しているサーバに対して有効であると考えられる．情報更新がプライマリサーバで行われていると仮定した場合，最新の情報はつねにプライマリサーバが保持しているため，クライアントに対してより良いサービスを提供するにはプライマリサーバからの情報伝送を優先すべきであると考えられるからである．

5. 結　　論

本稿で提案されたTASS方式は，プライマリサーバが最初にサーバとクライアント間の経路の混雑状況を計測し，その結果によりミラーサーバへの転送を行うかを決定し適切なミラーサーバの自動選択を行うことによりインターネットにおけるアクセス時間を短縮でき，かつアクセスにおける制御用通信のオーバヘッドをおさえることができる．TASS方式はWebページのサイズが大きく，またWebサーバ間のアクセシビリティの差が大きい場合にアクセス時間の短縮に特に有効であり，データの伝送サイズが非常に小さい場合およびサーバ間にアクセシビリティの差がない場合には効果的なアクセス時間短縮が行えないことが分かっ

た．TASS 方式の問題点として観測パケットの損失，DNS ラウンドロビン方式と比較した場合の可動性の低さ，Ku などの閾値の設定法などがあげられる．また，現在の方式においては伝送するサーバが一度決定されるとその後はそのサーバのみからの伝送になるが，伝送速度の変化にともないアクセス時間が他のサーバより時間がかかる可能性が生ずる．これを回避する方法の 1 つとしては，伝送されてきた Web ページにもう一度 TASS 方式を行うトリガを組み込んでおく方法が考えられる．これらに関しては今後の課題である．

参考文献

1) Albitz, P. and Liu, C.: *DNS and BIND*, O'Reilly & Associates (1992).
2) 森田昌宏，速水治夫：WWW における複製情報資源の一致制御とリクエストリダイレクションの提案，第 25 回 jus UNIX シンポジウム論文集 (1995).
3) Cisco System: Cisco DistributedDirector, http://www.cisco.com/warp/public/751/distdir/dd_wp.html.
4) Resonate: Resonate Global Dispatch, http://www.resonateinc.com.
5) 串田高幸：インターネットにおけるトラフィックの測定方法，情報処理学会マルチメディア通信と分散処理研究会，86-43 (1998).
6) Washburn, K. and Evans, J.: *TCP/IP: Running a Successful Network*, Addison-Wesley (1993).
7) Balachandran, A.: Wireless Media Scaling, http://www.ctr.columbia.edu/~anand/research/m_scaling.html.
8) 藤原　洋：最新 MPEG 教科書，ASCII (1994).
9) 玉井詩子：インターネットの混雑度を考慮したコンテンツ表示形式，第 55 回情報処理学会全国大会論文集，5U-1 (1997).

付　録

A.1 ラウンドトリップタイムと Web サーバからの情報伝送時間の相関関係

ラウンドトリップタイムと Web サーバからの情報伝送時間の関係を調べるために，静岡大学に実験用 Web サーバを設置し，群馬大学と静岡大学の間で ping によるラウンドトリップタイムと群馬大学から HTTP によるアクセスを行ったときの情報伝送時間を計測した．ラウンドトリップタイムを横軸に，Web サーバからの伝送時間を縦軸にしたグラフを**図 18** に示す．

この図より，若干のばらつきがあるものの，各点が右上がりの曲線上に並ぶことを確認することができる．

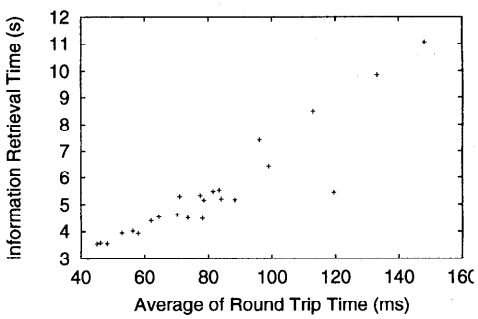
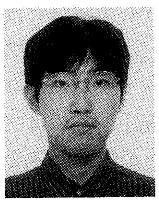


図 18 RTT と Web サーバからの情報伝送時間の関係
Fig. 18 Relation between round trip time and information retrieval time from Web server.

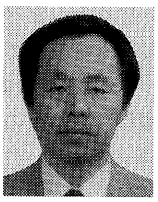
つまり，ラウンドトリップタイムが小さいほど情報の伝送時間が小さくなることが分かるこの結果，アクセス時間を短くするためのサーバ選択の基準としてラウンドトリップタイムを利用することは妥当であると考えられる．

(平成 10 年 5 月 8 日受付)
(平成 10 年 11 月 9 日採録)



竹内大五郎

1996 年群馬大学工学部情報工学科卒業．1998 年同大学院工学研究科情報工学専攻修士課程修了．同年（株）ナムコ入社．



小野里好邦（正会員）

1974 年東北大学工学部通信工学科卒業．1981 年同大学院博士課程修了．工学博士．同年電気通信大学電気通信学部助手．現在，群馬大学工学部情報工学科教授．この間，コンピュータネットワーク，衛星通信システム等の研究に従事．日本 OR 学会，ACM，IEEE 各会員．

山本　潮（正会員）

1969 年生．1992 年東北大学工学部情報工学科卒業．1997 年同大学院博士後期課程修了．同年より群馬大学工学部情報工学科助手．情報科学博士．ソフトウェア開発支援，マルチエージェントシステム，WWW サーバアクセスに関する研究に従事．人工知能学会会員．

冨沢 考弘

1998年群馬大学工学部情報工学科卒業．現在，同大学院修士課程1年在学．ネットワークの性能評価，サーバ選択方式等の研究に従事．

Sakchai Thipchaksurat

He was born in Suphanburi, Thailand. He received the B.Sc. degree in statistics from Srinakaritwirot Prasarnmitr University in 1988 and M.Eng. degree in electrical engineering from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 1996. From 1993 to 1997 he was with the Computer Research and Service Center, King Mongkut's Institute of Technology Ladkrabang. Since 1997, he has been with the Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Thailand. He is currently working toward the Ph.D degree at the department of computer science, Faculty of Engineering, Gunma University, Japan. His current interests are in the area of performance evaluation of computer network, wireless and mobile computing.