

地域ネットワークにおけるサーバ負荷を考慮した サーバ選択システムの構築

水越 一貴^{†1} 羽田 友和^{†2} 林本 雅之^{†3}
八代 一浩^{†4} 安藤 英俊^{†5}

ある特定の地域を対象にした、ネットワーク中継が行われるようになってきている。しかし、新しくネットワーク中継を行うごとにネットワークやサーバを用意するのは、コストや人的資源において負担が大きい。そこで、地域のISP (Internet Service Provider) や組織がすでに所有しているネットワークやサーバ資源を活用したネットワーク中継が求められている。既存の資源を有効に活用するためには、各ISP や組織が所有しているサーバの性能や上位ネットワークとの接続状況を考慮して適切なサーバに利用者を誘導するシステムが必要である。本論文では、サーバ負荷やネットワークトラフィックを測定し、地域のネットワーク内において利用者を最も負荷の少ないサーバに誘導するシステムを提案する。提案システムを実際に構築し、本提案システムの有効性を示す。

Construction of a Server Selection System Which is Designed with Server Load in Regional Network

KAZUTAKA MIZUKOSHI,^{†1} TOMOKAZU HADA,^{†2}
MASAYUKI HAYASHIMOTO,^{†3} KAZUHIRO YATSUSHIRO^{†4}
and HIDETOSHI ANDO^{†5}

Video streaming experiments for a certain specific region have become popular. However, there are problems in the cost and human resources, if the broadcasting system consists of network and servers which are newly built at every experiment. To solve these problems, it is required for video streaming system to use the network and the server resources that ISP (Internet Service Provider) and the organization of the region already own. To use existing resources effectively, the system should be designed with performance of the streaming servers and network status. In this paper, we propose a system which navigates the user to the server with the lowest load in the regional network by measuring server load and network traffic. The effectiveness was confirmed by the experiment on the actual system based on proposed method.

1. はじめに

CATV, FTTH, DSL をはじめとするブロードバンドネットワークが大きく普及してきており、インターネットに接続している世帯でのブロードバンドネット

ワークの普及率は、7割を超えている¹⁾。ブロードバンドが普及するにつれて、配信されるコンテンツもテキストから映像や音楽のような広帯域を使用するものが多くなってきている。また、この環境を利用してイベントを中継するようなことも活発に行われるようになってきた。

このことを利用して地方では、地域情報をコンテンツとして発信し活性化を図っている。効率良く配送を行う仕組みとして地域IX (Internet eXchange) を利用した中継が行われている^{2),3)}。地域IXの近傍にサーバを配置することにより利用者とサーバ間の通信経路を短くすることができるためにRTT (Round Trip Time), 遅延, ジッタなどの通信品質が良くなることが知られている⁴⁾。また、地域外からのアクセスは、地域IXに接続しているISPや組織の上位回線か

†1 山梨大学医学工学総合教育部
Interdisciplinary Graduate School of Medical and Engineering, University of Yamanashi

†2 株式会社 YSK e-com
YSK e-com Corporation

†3 株式会社日本ネットワークサービス
Nihon Network Service, Inc.

†4 山梨県立大学
Yamanashi Prefectural University

†5 山梨大学医学工学総合研究部
Interdisciplinary Graduate School of Medical and Engineering, University of Yamanashi

ら行うことができる。そのため地域 IX から地域外に対して複数の経路を持つことができ、トラフィックを分散させる働きもある。

一方で新たにネットワーク中継をしようとするたびに配信サーバや回線を用意するという配信環境の再構築は人的コストが増加してしまうという問題が指摘されている⁵⁾。そのことから、配信サーバや回線といったリソースを有効活用できる仕組みが必要である。しかし、一般的に地域 IX に参加している ISP や組織が所有しているリソースには偏りがある。具体的に、地域 IX を使ってコンテンツを配信していく際には以下のような問題点がある。

- 各組織の所有する配信サーバの性能が異なる。
- 各組織の上位ネットワークへの回線帯域が異なる。

これらの問題を解決するために、本論文では定期的に配信サーバの負荷を計測しその状況に応じて配信サーバを選択するシステムを提案する。配信サーバの負荷を見ることによって配信サーバの性能に応じた配信を行うことが可能となる。また、配信サーバ選択の際には配信サーバからのトラフィックも測定し、その量が上位ネットワークの回線幅を超える場合には、その配信サーバが選択されないようにする。このことにより ISP や組織により上位ネットワークへ中継可能なトラフィック量を限定することが可能となる。この提案システムを実装し、地域 IX 上の実システムで中継を行った⁶⁾。本論文では、中継を行ったシステムの運用データをもとに評価を行い、システムの有効性を示す。

以下、2章では、既存のサーバ選択の方法について述べる。3章では配信サーバ選択の要件について述べる。4章では提案システムであるサーバの負荷を考慮したサーバ選択システムについて述べ、5章で本システムの実装について説明する。6章では、提案システムを実システムに適用した結果を述べ評価を行う。7章はまとめである。

2. サーバ選択の方法

複数のサーバを使ってコンテンツを配信する際に、地域 IX を使わない場合はユーザへコンテンツを配信するサーバを選択する方法が必要である。複数の配信サーバが同一箇所にあり、同一性能の配信サーバを使用する場合はサーバを均一に選択することで負荷分散が行える。地域 IX を使う場合は、地域外からのアクセスは地域 IX に参加している ISP や組織の上位回線からアクセスされるので配信サーバを分散配置したほうが効率が良い。また、ISP や組織がコンテンツ配信

のためのサーバを提供する場合、サーバの性能が異なることがある。そのためコンテンツを地域 IX に参加している ISP や組織の既存リソースを使って配信していくためには、分散配置された配信サーバからサーバの性能を考慮して最適な配信サーバを選択する必要がある。

サーバ選択を実現する方法として、DNS によるラウンドロビンや経路情報を基にした選択、負荷分散装置のような専用機器による方法などがある。本章では、いくつかの方法から主なものについて概説する。

2.1 DNS ラウンドロビン

ユーザが配信サーバにアクセスする際には、通常 IP アドレスではなく、より分かりやすいホスト名を使う。DNS サーバは、このホスト名と IP アドレスの対応を調べるためのサーバである。DNS サーバは通常、1つのホスト名に対して複数の IP アドレスを登録することができる。たとえば、あるホスト名に対して I1, I2, I3 の3つの IP アドレスが登録されているとする。この登録されたホスト名に問合せがきた場合、クライアントに対して DNS サーバは I1 から順番に答えていく。このようにして I1 ~ I3 までの IP アドレスをアクセスされるたびに順に答えていく。この機能を DNS ラウンドロビンと呼ぶ。クライアントからの問合せごとに登録されている IP アドレスを順番に返答するため、各サーバが均等に選択される。この方法では、配信サーバの状態によらず各サーバが均等に選択されるため最適なサーバが選択されるとは限らない。

2.2 RADIX+TENBIN

RADIX+TENBIN^{7),8)} は DNS を基盤としたサーバ選択システムである。このシステムは、サーバ選択の選択基準に基づいてサーバを選択する DNS サーバ TENBIN とサーバ群の経路情報を収集しサーバ選択基準として TENBIN に情報を渡す RADIX からなる。クライアントから要求があると RADIX からの経路情報を基にサーバ群から最適なサーバの IP アドレスを TENBIN が応答する。経路情報は BGP4⁹⁾ を使ってサーバ付近のルータから収集する。経路情報を用いてサーバ選択を行うため、選択されるサーバはクライアントからネットワーク的に近いものとなる。この方式ではネットワーク的にクライアントと近いサーバを選択するが、本研究のように特定の地域におけるネットワーク中継を想定した場合には、ネットワーク距離はいずれもほぼ均しくなるため、本手法を適用しても均一に誘導されてしまうという問題がある。また、配信サーバの負荷を考慮していないので同じネットワークのクライアントから多くのアクセスが集中した場合に

1つの配信サーバに負荷が集中してしまう。

2.3 DNS フィルタ方式

DNS フィルタ方式¹⁰⁾は、上位DNSとクライアントの間にローカルDNSのように設置する。クライアントから本システムに問合せがあったときには、上位DNSに代理で問合せを行う。問合せ結果のサーバに対して調査を行い、不適なサーバを除外してクライアントに答える。本システムにおいて不適なサーバとはICMP ECHO パケットに対する応答速度の遅いサーバ、TCP 接続の接続時間が長いサーバのことを指す。このようにローカルDNSサーバのように振る舞いながら不適なサーバのフィルタリングを行うシステムである。たとえば、あるサーバ群S1, S2, S3があり、DNSサーバには1つのホスト名に対してサーバ群のIPアドレス3つが登録されているとする。クライアントは本システムにIPアドレスの問合せを行う。評価システムはDNSサーバに問合せを行い、そのホスト名にはIPアドレスが3つあることを知る。3つのIPアドレスからサーバ群を調査し、不適なサーバのIPアドレス以外をクライアントに返答する。この方式は、クライアントからみて最適なサーバに導くことを目的としている。そのため、サーバ群にとって必ずしも最適な負荷分散が行われるとは限らない。また、システム自体を上位DNSとクライアントの間にローカルDNSのように設置しなくてはならず、多くの商用ISPではローカルDNSの制御はできないため、配信側からサーバ選択を制御することは難しい。

3. 配信サーバ選択の要件

前章で述べたサーバ選択方法は、広域で有効にサーバを選択する方法であったり、クライアントにとって適切なサーバを選択する方法であったりする。今回の地域IXのように比較的密集したネットワークにおいては、広域でサーバを選択する方法では、必ずしも最適な配信サーバが選択されとは限らない。

そこで、本論文では次のような条件で地域IXでのコンテンツ配信と負荷分散を行うシステムを構築することを目的とする。

- 配信サーバは地域IXに参加する組織においてIX近傍に設置される。
- 設置される配信サーバの性能は不均一。
- 各配信サーバの性能限界を超えない。
- 各組織によって上位ネットワークに中継するトラフィックを限定する。

地域外からのアクセスは地域IXに参加している組織を通して地域IX内に入るのそれぞれ分散してサー

バを配置するとそれぞれの組織の上位ネットワークを使用できるので効率が良い。また、コンテンツを配信するたびに新しいサーバを用意するのではなく各組織が持っているサーバ資源をコスト削減のために有効活用するようにしたい。しかし、各組織の持っているサーバが同じ性能であることはない。そのような状況でも各配信サーバの性能限界を超えてしまえば、コンテンツを配信することができなくなってしまう。そのため、限界を超えないようにする必要がある。

地域IXに参加している組織では、それぞれインターネットへ接続するための上位回線を独自に持っている。地域IX以外からの配信サーバへのアクセスが集中するとコンテンツ配信のため上位回線の帯域がひっばくする。これにより、他のサービスに影響が出ることは許されない。そこで、各組織で上位ネットワークに中継できるトラフィックを限定する必要がある。

上述した条件を実現するためには次のような要件を満たす必要がある。

配信サーバの性能が不均一な状態で各配信サーバの性能限界を超えないようにするためには配信サーバの負荷を知る必要がある。負荷を測定し、配信サーバとして動作できる負荷を超えた場合にはその配信サーバにはそれ以上新しい接続をしないようにする。

上位ネットワークへのトラフィックを限定するために利用可能帯域を設定できるようにする。トラフィックは配信サーバが設置されているサイトから地域IX外のネットワークへ送信される。そこで配信サーバごとに利用可能帯域を設定することによって、地域IX外のネットワークへ送信されるトラフィックを限定することができる。また、配信するコンテンツによっては同時接続数を限定したい場合もある。そのための接続数上限も設定できるようにし、接続数によっても制御を行えるようにする。

配信サーバの選択では、配信サーバにコンテンツを配信することが可能かどうかを負荷状況から判断する。その後、その中で利用可能帯域と接続数上限を超えない配信サーバを判断し、負荷の最も低い配信サーバを選択する。

4. サーバ負荷を考慮したサーバ選択システム

地域IXからコンテンツを配信するのに最適な、配信サーバの負荷を考慮したサーバ選択システムを提案する。本章では、提案するサーバ選択システムについて説明する。

4.1 提案システムの構成

提案するサーバ選択システムの構成は図1の点線

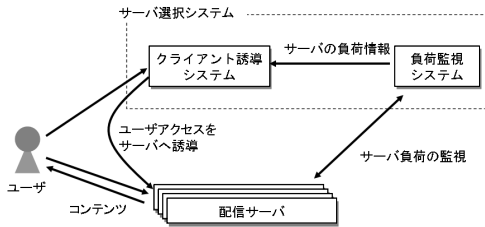


図1 提案システムの構成
Fig. 1 Configuration of proposed system.

枠内の2つの処理システムからなる。1つはクライアント誘導システムであり、もう1つは負荷監視システムである。

負荷監視システムは常時配信サーバの負荷を測定し、その状況をクライアント誘導システムに報告する。ユーザがクライアント誘導システムにアクセスしてきた場合、負荷監視システムからの報告を基にユーザのアクセスを最適な配信サーバに誘導する。誘導後、ユーザは誘導された配信サーバからコンテンツを受信する。

この提案システムは、コンテンツ配信を行う地域IX内に設置することを前提としている。地域IX内ならば高速回線で接続されているので提案システムをどこに設置しても大きな違いはない。また、提案システム自体を負荷分散する場合は提案システム全体を同じサイトに複数設置しても各サイトに分散して設置してもかまわない。

4.2 負荷監視システム

負荷監視システムは配信サーバの負荷を測定する。測定した負荷情報から配信サーバの負荷を総合的に見て、順位付けを行う。付けた順位とクライアント誘導に必要な情報をクライアント誘導システムに送る。

負荷監視システムが測定する負荷情報は次の4つを使用する。

- トラフィック
- コネクション数
- CPU使用率
- メモリ使用率

トラフィックは、配信サーバから送信されるデータ量である。コンテンツを配信する場合、配信サーバが送信するのはコンテンツであり、受信するのはコンテンツ要求メッセージである。一般に要求メッセージのデータはコンテンツに比べて小さい。よって配信サーバが受信するトラフィックが負荷に与える影響は無視できると考えられる。コネクション数は、配信サーバに現在接続しているクライアントの数である。コネクション数が増加するほどコンテンツを配信するクライ

アントが多くなるので配信サーバへの負荷は大きいと考えられる。また、クライアント誘導システムが接続数上限を判断する際にも使用する。CPU使用率とメモリ使用率によって配信サーバのハードウェア負荷が分かる。これら4つの値を使うことによって配信サーバのハードウェア負荷とネットワーク負荷を評価する。

配信サーバの順位付けは測定した4つの負荷情報を使って行う。4つの負荷情報から最も負荷が低い配信サーバから順番にリストにする。クライアント誘導システムでは、利用可能帯域と接続数上限を判断する。利用可能帯域を超えないようにするためには現在のトラフィックを使うことは難しい。なぜならば利用可能帯域の上限に近づいた状態で新たな接続を誘導すると利用可能帯域を超えてしまうからである。そこで現在のトラフィックから新しい接続が行われた場合にどの程度のトラフィックになるかを予測する。よってクライアント誘導システムに送信する負荷情報は配信サーバリストと予測トラフィック、コネクション数となる。

4.3 クライアント誘導システム

クライアント誘導システムは随時、負荷監視システムから配信サーバの負荷情報を受け取る。負荷情報は、負荷の低い配信サーバから順にリストになっている。ユーザからのアクセスがあった場合、負荷監視システムから受け取ったリストの最上位にある配信サーバの予測トラフィックとコネクション数情報から判断を行う。誘導の判断に用いる情報は次の3つである。

- 配信サーバの負荷順位
- 予測トラフィック
- コネクション数

まず、リストの上位の配信サーバから予測トラフィックが利用可能帯域を超えてないか比較する。また、コネクション数が接続数上限を超えていないかも比較する。それぞれ比較を行った結果が両方とも真のときはその配信サーバにクライアントアクセスを誘導する。それ以外の場合、その配信サーバにはクライアントアクセスを誘導せずにリストの次にあるサーバに対して同様の比較を行っていく。

5. サーバ負荷を考慮したサーバ選択システムの実装

前章までの設計から、システムの実装を行った。前章4.2節の負荷監視システムをLAFS (Load Average Forecast System) と命名し、前章4.3節のクライアント誘導システムをHADA (High Available Dynamic Access-control) と命名した。以下、全体のシステムについて説明する。

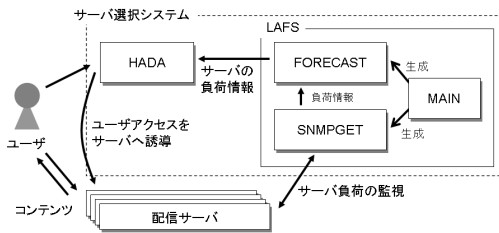


図 2 LAFS+HADA によるサーバ選択システム

Fig.2 Server selection system: LAFS+HADA.

LAFS+HADA の構成を図 2 に示す。まず、LAFS は、定期的に配信サーバの負荷情報を取得し、分析する。その結果を基に優先的に割り当てる配信サーバを導き出し、リストにして HADA に送信する。HADA は LAFS から送られてきたリストを使って配信サーバを選択する。選択順位は基本的にリストの上位にある配信サーバを選択するが、各組織の上位回線の利用可能帯域や接続上限数によって選択するサーバを決定する。

以下の各節では、LAFS+HADA の詳細について説明する。

5.1 LAFS

LAFS は配信サーバの負荷情報を取得し分析する。実装は実行速度を考慮して C 言語で記述した。

配信サーバの負荷情報の取得には、ネットワーク管理に広く利用されている SNMP¹¹⁾ を使った。SNMP のエージェントは多くの機器に対応しており、配信サーバに導入することで様々な情報を取得できるようになる。LAFS による負荷情報の取得は、モジュール化してあるので他の方法で情報を取得するプログラムと容易に差し替えることが可能となっている。

LAFS は MAIN, FORECAST, SNMPGET の 3 つのモジュールから構成される。

5.1.1 MAIN

MAIN では、SNMPGET と FORECAST の管理を行う。設定ファイルから配信サーバの情報を取得し、それを基にして SNMPGET と FORECAST のプロセスを作成する。プロセスの生成は、設定した時間ごとに行う。この時間によって配信サーバの負荷を測定する時間と HADA に送るサーバ選択リストの精度を調整することができる。配信サーバの負荷変動が起こりやすいコンテンツの場合は時間を短くし、変動の少ない場合は長くする。

5.1.2 SNMPGET

SNMPGET は、MAIN によってプロセスが作成される。SNMP マネージャとして動作し、配信サーバから負荷を測定するのに必要な情報を SNMP GET リ

クエストを使って取得する。4.2 節で検討した項目を知るのに必要なデータを MIB¹²⁾ (Management Information Base) から取得する。使用する MIB 値は次のとおりである。

- トラフィック

出力バイト数 (ifOutOctets)

1.3.6.1.2.1.2.2.1.16

(iso.org.dod.internet.mgmt.mib-2.interfaces.ifTable.ifEntry.ifOutOctets)

- コネクション数

TCP コネクション数 (tcpCurrEstab)

1.3.6.1.2.1.6.9

(iso.org.dod.internet.mgmt.mib-2.tcp.tcpCurrEstab)

- CPU 使用率

CPU 使用率 (hrProcessorLoad)

1.3.6.1.2.1.25.3.3.1.2

(iso.org.dod.internet.mgmt.mib-2.host.hrDevice.hrProcessorTable.hrProcessorEntry.hrProcessorLoad)

- メモリ使用率

メモリ全体量 (hrStorageSize)

1.3.6.1.2.1.25.2.3.1.5

(iso.org.dod.internet.mgmt.mib-2.host.hrStorage.hrStorageTable.hrStorageEntry.hrStorageSize)

メモリ使用量 (hrStorageUsed)

1.3.6.1.2.1.25.2.3.1.6

(iso.org.dod.internet.mgmt.mib-2.host.hrStorage.hrStorageTable.hrStorageEntry.hrStorageUsed)

トラフィック量については、配信サーバのネットワークインタフェースを通して出力されたバイト数を使用する。取得した値は累積トラフィックなので前回取得した値を保持しておき、次に取得した値から引くことによって時間あたりのトラフィックを求める。コネクション数は、TCP のコネクション数を使用する。tcpCurrEstab 項目では、配信サーバが ESTABLISHED または CLOSE-WAIT しているコネクション数が測定できる。配信に UDP を使用した場合は、コネクション数を直接取得することはできない。そのため新たに MIB を定義するか、他の方法で接続数を測定する必要がある。CPU 使用率は、前 1 分間平均の値である。メモリ使用率は、MIB の項目にはないのでメモリ使用量とメモリ全体量によって求める。

SNMP は、配信サーバにエージェントを導入し

LAFS からの負荷情報要求に回答するように動作する．測定したトラフィックには配信サーバが負荷情報要求に回答する際のトラフィックも含まれている．このトラフィックは，出力バイト数，TCP コネクション数，CPU 使用率，メモリ使用量，メモリ使用量の 5 つの項目を回答するパケットでも 1 KByte 以下である．毎秒負荷を測定したとしても 8 Kbps に満たない．よって測定したトラフィックに含まれる SNMP のトラフィックは無視できる程度であるといえる．

配信サーバの負荷は配信するコンテンツによって異なる．たとえば，動画などのようにアクセスが集中したり，継続したりするようなコンテンツを配信する場合は負荷が変動しやすいので負荷の測定を短い間隔で行っていく必要がある．Web ページのようにアクセスが分散するようなコンテンツを配信する場合は配信サーバの負荷変動が比較的ゆるやかなので測定間隔を長くしても有効な負荷分散が可能である．このような負荷変動の違いに対応できるように，配信サーバの負荷を測定する間隔は変更できるようにした．

5.1.3 FORECAST

FORECAST は，MAIN によってプロセスが作成される．SNMPGET で取得した配信サーバの負荷情報を使って，配信サーバの負荷を順位付けする．

最初にトラフィックとコネクション数から次のコネクションが増えたときのトラフィックを予測する．トラフィックの予測は次の式を用いた．

$$FT = T + (T/C) \quad (1)$$

ただし，FT は予測トラフィック，T は現在のトラフィック，C は現在のコネクション数である．たとえば，現在のトラフィックが 500 Kbps で，接続されているコネクション数が 10 の場合には予測トラフィックは 550 Kbps となる．この式により配信サーバに新たな接続を許可した場合，トラフィックがどの程度になるかを予測することができる．

その後，各配信サーバの負荷量である weight を算出する．weight の算出には次の式を用いた．

$$weight = FT * (\alpha PL + \beta MU) \quad (2)$$

ただし，FT は式 (1) で求めた予測トラフィック，PL は CPU 使用率，MU はメモリ使用率である． α と β は配信するコンテンツによって配信アプリケーションが CPU やメモリの使用率に及ぼす影響が違うのでそれを調整するための変数である．weight によって配信サーバの順位付けを行う．weight の大きい配信サーバほど負荷が高い．HADA に送られるデータは，配信サーバの優先度リスト，予測トラフィック，コネクション数である．これらのデータを weight の小さい

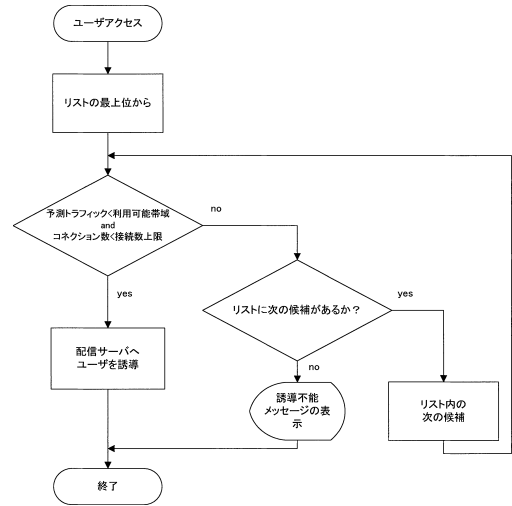


図 3 HADA のフローチャート

Fig. 3 Flowchart: HADA.

配信サーバから並べてリストを作成し，HADA に送信する．また，weight が 0 になるのは FT または PL と MU の両方が 0 の場合である．そのような状況になるのは，ネットワークまたは配信サーバ自身になんらかの問題が発生していると考えられるのでリストには含めないようにする．よって，配信サーバすべてのリストが HADA に送られるわけではない．配信できるサーバがない場合は空のリストが HADA に送られる．そのときの対処は HADA によって行われる．

5.2 HADA

HADA はユーザからのアクセスを HTTP で受け，動的リンクを作成することによって配信サーバへ誘導する CGI プログラムとして実装した．実装は，CGI プログラムで広く使われている Perl を使って行った．HADA は LAFS から定期的に送られてくる配信サーバの優先度リストと予測トラフィック，コネクション数を基にユーザを配信サーバに誘導する．各配信サーバの使用可能帯域と接続数上限の情報を使い，ユーザからのアクセスを制御し配信サーバへ誘導する．

HADA のフローチャートを図 3 に示す．HADA は，LAFS からサーバの優先度リストと各配信サーバの予測トラフィック，TCP コネクション数を受け取る．ユーザアクセスがあった時点で最も優先度が高いサーバの予測トラフィックと利用可能帯域，TCP コネクション数と接続数上限をそれぞれ比較する．そのときに，閾値を上回っていなければその配信サーバを選択する．その後，選択された配信サーバへ誘導できるように Web ページを生成する．もし，上回っていた場合は次に優先度の高いサーバについて同様の比較を行う．す

すべての配信サーバが不適だった場合は、その旨を伝える Web ページを作成する。

6. 実運用による評価

前章で実装した LAFS+HADA を山梨県市川大門町で行われた第 17 回神明の花火大会のインターネットライブ中継におけるストリームサーバ選択に使用して、その有効性を確認した。この中継は、NAVEL プロジェクト¹³⁾ によって行われたものである。2005 年 8 月 7 日 19 時 30 分から 21 時の間中継が行われた。

6.1 構成

コンテンツ配信には Microsoft 社の Windows-Media¹⁴⁾ を使用した。全体のネットワーク構成図を図 4 に示す。配信サーバは山梨県の商用地域 IX である BeX-J¹⁵⁾ の参加組織内のネットワークに分散して 5 台設置した。設置した配信サーバの仕様は、表 1 のようにそれぞれ異なっている。また、配信サーバの負荷を SNMP で測定するために各配信サーバに Windows 標準の SNMP エージェントを設定した。そのほかに 5 台の配信サーバに再送信を行う配信サーバ (WMS-D) を 1 台設置した。撮影した映像はエンコードを行ってから配信元となる WMS-D に送り、そこから分散配置された 5 台の配信サーバへ再配信が行われる。配信レートは 512 kbps とした。

LAFS+HADA は配信サーバの 1 台を設置したネッ

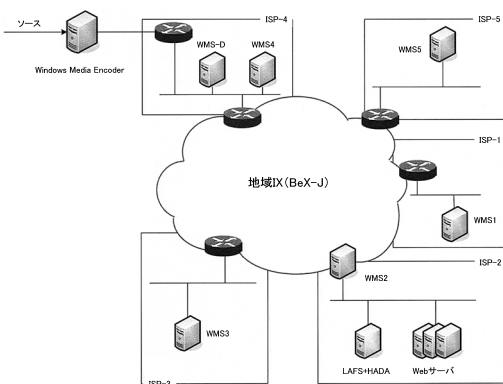


図 4 システム構成

Fig. 4 System configuration.

表 1 配信サーバの仕様

Table 1 Specifications of servers.

| 配信サーバ | CPU | Memory | HDD |
|-------|-----------------|---------|----------|
| WMS1 | Xeon 3.2 GHz | 1.0 GB | 40 GB |
| WMS2 | Pentium3 1 GHz | 1.25 GB | 17 GB |
| WMS3 | Pentium3 1 GHz | 512 MB | 10 GB |
| WMS4 | Xeon 2.4 GHz x2 | 1.5 GB | 80 GB x2 |
| WMS5 | Celeron 2.6 GHz | 1.0 GB | 20 GB |

トワークと同じところに設置した。利用可能帯域と接続数上限については、表 2 のように設定した。そのほかに Weight を求める際の α と β については、 α を 2、 β を 1 にした。これは、事前に配信サーバを用意したテストを行った結果から決定した。テストに用いた配信サーバは CPU が Celeron の 2.6 GHz、メモリが 1 Gbyte の機器である。接続クライアント数を 0、10、30、50 で CPU 負荷率とメモリ使用率を測定した。その結果を表 3 に示す。CPU 負荷率、メモリ使用率はともに線形に増加している。しかし、CPU 負荷率の増加率の方が高い。これはストリーム配信における負荷はメモリよりも CPU の影響を大きく受けることを表している。しかしながら、実機においては CPU の性能は様々であり、すべての場合について事前にテストを行うことは困難である。そのことから CPU 負荷率を重視し CPU 負荷率とメモリ使用率の比を 2 対 1 になるようにした。この比は、CPU 負荷率を重視するために設定した比であり、どのような比が最も動画配信に適しているかは今後検討していく必要がある。また、映像配信のサーバ誘導においては短い時間間隔で負荷を測定するため¹⁶⁾、SNMP による負荷測定間隔を 1 秒とした。

6.2 中継におけるサーバ選択の結果と考察

花火中継の間のユーザアクセス数と誘導数を図 5 に示す。総アクセス数は 5116 で、19 時 42 分には最大同時アクセス数が 204 あった。ユーザの誘導が 20 時 12 分までうまくできていなかった。それは HADA の設定にミスがあり LAFS からのリストで優先度が 1 番目のものがつねに使用される状態だったためである。設定を修正したところうまくユーザの誘導が行えるよ

表 2 配信サーバの利用可能帯域と接続数上限

Table 2 Available band of transmission and maximum number of connection.

| 配信サーバ | 利用可能帯域 | 接続数上限 |
|-------|---------|-------|
| WMS1 | 15 Mbps | 27 |
| WMS2 | 80 Mbps | 142 |
| WMS3 | 25 Mbps | 42 |
| WMS4 | 40 Mbps | 72 |
| WMS5 | 80 Mbps | 142 |

表 3 配信テストにおける配信サーバの平均 CPU 負荷率と平均メモリ使用率

Table 3 Server's average CPU load rate and average memory use rate on broadcasting test.

| クライアント数 | 平均 CPU 負荷率 | 平均メモリ使用率 |
|---------|------------|----------|
| 0 | 2.0% | 32.8% |
| 10 | 2.6% | 32.8% |
| 30 | 4.0% | 33.1% |
| 50 | 5.3% | 33.3% |

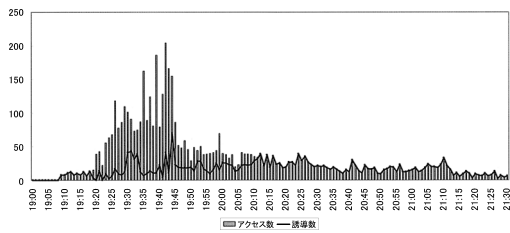


図 5 ユーザアクセス数と誘導数

Fig. 5 The number of user access and the number of user navigation.

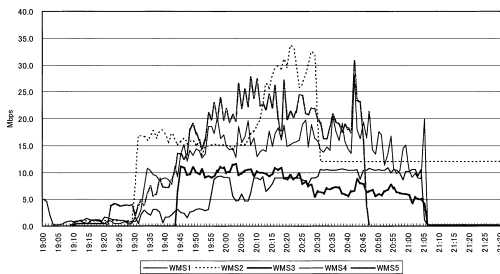


図 8 配信サーバのトラフィック

Fig. 8 Traffic of broadcasting servers.

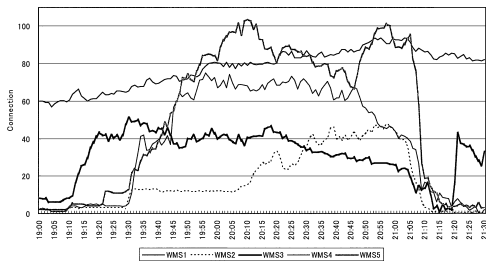


図 6 配信サーバ接続数

Fig. 6 The number of connection server provide.

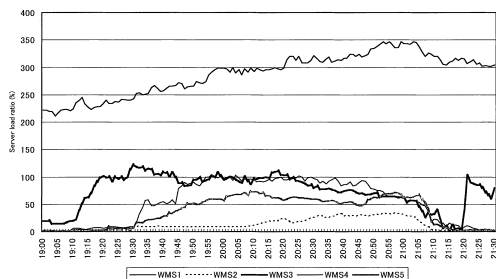


図 7 サーバの接続数上限に対する接続数の割合

Fig. 7 Rate for connections against the maximum connection on servers.

うになった。

本中継において、各配信サーバへの接続数と接続数上限に対する接続数の割合を図 6 と図 7 に示す。図 6 において 19 時 30 分の中継開始から各配信サーバへのアクセスが行われている。WMS1 で常時 60 以上の接続が発生しているのは、他のアプリケーションが違う目的で接続を行ってしまっていたためである。また、WMS3 が中継開始前から接続数が増加しているのは HADA の設定ミスによりしばらくの間 WMS3 への誘導が続いたためである。他の 3 つの配信サーバは、同じ割合で接続数が増加している。これは誘導が偏らず行われたためだといえる。また、WMS4 の 19 時 55 分から 20 時 30 分の接続数を見ると接続数上限 (100%) に達したときにはその配信サーバには新たな誘導が行われていないことも分かる。花火大会終了後、

21 時 05 分に配信サーバを停止させた。WMS3 の接続数が 21 時 20 分すぎに急増しているのは、配信サーバ停止後にログ収集や L3 スイッチの設定変更などの作業を行っていたためである。

図 7 から接続数においてほとんど接続数上限を超えずに各配信サーバに誘導が行われているといえる。ここでも WMS1 が 200%以上の大きな接続率を示しているが、前述した接続によって接続数上限を超えてしまっているからである。WMS2 の接続割合が低いのは、前述した HADA の設定ミスによるものである。20 時すぎからは WMS2 にも誘導が行われるようになったが、そのときには中継への新しいユーザアクセスが少なくなっていた。また、同時刻から他の配信サーバへの接続が減少していることからこの時間からのユーザアクセスはほとんど WMS2 に誘導されており、最も負荷の少ないサーバが選択されていたことになる。

各配信サーバのトラフィックを図 8 に示す。中継が始まったときには WMS2 の優先度が高かったが、その後 WMS5、WMS4 の順で誘導されている。20 時 30 分ごろから WMS2 のトラフィックが一定になっているのは LAFS のプログラム内でトラフィック値を格納している変数の桁あふれによるものである。WMS2 のトラフィックにおいて他の WMS よりも接続数が少ないにもかかわらずトラフィックが多いのは、図 4 に示すように WMS2 が設置した組織と地域 IX 間のトラフィック中継も行っており、中継 Web ページへのアクセスや中継以外のトラフィックが流入してためである。配信サーバを停止した 21 時 05 分以降はトラフィックがなくなっている。図 6 において接続数がトラフィックのなくなったあとも緩やかに減少しているのは、接続数を測定している `tepCurrEstab` が `ESTABLISHED` の状態以外に `CLOSE-WAIT` を含んでいるため完全に `CLOSE` するまでは接続数としてカウントされてしまうためである。どの配信サーバのトラフィックも利用可能帯域を超えることなく動作している。

7. おわりに

本論文では地域 IX を使ってコンテンツを配信する際に有効なサーバ選択システムを提案した。提案システムは、地域 IX 内に分散配置された配信サーバの負荷を調査して最適なサーバに誘導するものである。また、提案システムを実際のコンテンツ配信に用い、配信サーバの性能や上位回線の違う環境における有効性を確認した。

今後の課題として、weight を計算する際の各配信アプリケーションに適した CPU 負荷率とメモリ使用率の比を検証していく必要がある。また、サーバ選択基準として負荷以外の項目を組み合わせた場合のサーバ選択手法の検討がある。

謝辞 本論文の執筆にあたり、本システムの運用、評価に協力いただいた NAVEL プロジェクトの皆さんに感謝する。

なお、本論文の一部は笹川科学研究助成によるものである。

参 考 文 献

- 1) 財団法人インターネット協会：インターネット白書 2006，インプレス R&D (2006).
- 2) 八代一浩，笹本正樹，平川寛之，山本芳彦，林英輔：地域 IX を基盤とするネットワーク中継，情報処理学会研究報告，Vol.1998-DSM-010，pp.49-54 (1998).
- 3) 太田麻衣子：地域向けコンテンツ配信方式に関する考察．<http://e-inet.wag.ad.jp/meeting/file/2003112103.pdf>
- 4) 八代一浩，笹本正樹，平川寛之，山本芳彦，林英輔：地域 IX を用いた通信環境改善手法の実現と評価，情報処理学会論文誌，Vol.41，No.12，pp.3245-3254 (2000).
- 5) 豊島修平，菊池 豊：地域間相互接続実験プロジェクトにおける定常的な動画配信ネットワーク，情報処理学会研究報告，Vol.2003-DSM-30，pp.41-46 (2003).
- 6) 水越一貴，羽田友和，林本雅之，八代一浩，安藤英俊：動画配信における負荷分散システムの構築，分散システム/インターネット運用技術研究会シンポジウム 2005，情報処理学会，pp.123-128 (2005).
- 7) 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 B，Vol.J86-B，No.8，pp.1454-1462 (2003).
- 8) 下川俊彦，木場雄一，中川郁夫，山本文治，吉田紀彦：広域分散ストリーム配信における経路情報の利用，電子情報通信学会技術研究報告，

Vol.IA2001-8，pp.1-8 (2001).

- 9) Rekhter, Y. and Li, T.: A Border Gateway Protocol 4 (BGP-4), RFC 4271 (2006).
- 10) 横田裕思，木村成伴，海老原義彦：DNS フィルタ方式によるミラーサーバ選択法の提案と実装，情報処理学会論文誌，Vol.44，No.3，pp.682-691 (2003).
- 11) Case, J., Fedor, M., Schoffstall, M. and Davin, J.: A Simple Network Management Protocol (SNMP), RFC 1157 (1990).
- 12) Case, J. and Rose, M.: Management Information Base for Network Management of TCP/IP-based internets: MIB-II, RFC 1213 (1991).
- 13) NAVEL プロジェクト：山梨地域情報ネットワーク相互接続機構．<http://www.navel-y.jp/>
- 14) Microsoft Co. Windows Media．<http://www.microsoft.com/windows/windowsmedia/default.msp>
- 15) BeX-J: Broadband or Business EXchange Japan. <http://www.bex-j.net/>
- 16) 水越一貴，牧野 晋，林 英輔：通信トラフィック監視システムの試作とパーストラフィックの検出，情報処理学会研究報告，Vol.2004-DSM-34，pp.31-36 (2004).

(平成 18 年 7 月 7 日受付)

(平成 19 年 1 月 9 日採録)



水越 一貴 (学生会員)

2003 年麗澤大学国際経済学部国際産業情報学科卒業。2005 年同大学大学院国際経済研究科政策管理専攻修士課程修了。現在、山梨大学大学院医学工学総合教育部博士課程在学中。ネットワーク運用管理技術、サーバ負荷分散技術に関する研究に従事。



羽田 友和

2003 年山梨大学工学部コンピュータメディア工学科卒業。同年、山梨ソフトウェア株式会社 (現、株式会社 YSK e-com) 入社。同年 10 月より株式会社インテック・ネットコアへ出向後、MPLS ネットワークに特化した監視システムの開発に従事。2006 年より株式会社デジタルアライアンスへ出向。山梨県情報ハイウェイの運営に従事。次世代 IX 研究会コアメンバ。



林本 雅之

2002年山梨大学工学部コンピュータメディア工学科卒業。2004年山梨大学大学院コンピュータメディア工学専攻博士前期課程修了。同年より株式会社日本ネットワークサービス

勤務。ネットワーク運用管理技術に興味を持つ。



八代 一浩（正会員）

1987年山梨大学大学院工学研究科電気工学専攻学科修了。1997年より山梨県立女子短期大学助教授。2005年より山梨県立大学助教授。インターネットシステムの運用技術に

関する研究に従事。電子情報通信学会，日本教育工学会各会員。



安藤 英俊（正会員）

1986年東京大学工学部精密機械工学科卒業。1988年東京大学大学院工学系研究科情報工学専攻修士課程修了。1990年より日本学術振興会特別研究員。1991年東京大学大学院

工学系研究科情報工学専攻博士課程修了。1997年山梨大学総合情報処理センター助教授。2003年より山梨大学工学部コンピュータ・メディア工学科助教授。現在に至る。工学博士。形状モデリング，コンピュータ・グラフィクス，数値シミュレーション，ネットワークシステムの研究に従事。日本バーチャルリアリティ学会会員。1989年情報処理学会研究賞受賞。2006年情報処理学会優秀研究発表賞受賞。