

## 汎用ロードバランサによるスケーラビリティを有する 遠隔教育システムの構築

丹 康雄<sup>†</sup>(ytan@jaist.ac.jp), 廣瀬 正之<sup>‡</sup>,

敷田 幹文<sup>††</sup>, 井口 寧<sup>††</sup>, 藤枝 和宏<sup>††</sup>, 松澤 照男<sup>††</sup>, 熊谷 義隆<sup>††</sup>

<sup>†</sup> 北陸先端科学技術大学院大学 情報科学研究科

<sup>††</sup> 北陸先端科学技術大学院大学 情報科学センター

<sup>‡</sup> 東京エレクトロン株式会社 コンピュータ・ネットワーク部

<sup>†††</sup> 東京エレクトロンFE株式会社 コンピュータ・ネットワークFE部

遠隔教育関連のサービスを提供するシステムを汎用のロードバランサを用いて構築する方法について述べる。学内と学外の両方からアクセスされる遠隔教育システムでは各々に異なるポリシーを適用できるファイウォールが必要であるが、これが性能上のボトルネックとなる。本稿では、このファイウォール装置も含めロードバランサで負荷分散することにより、スケーラビリティが高く運用方法の継続性を有するシステムを提案し、JAIST で実稼働しているシステムを例に、効果と問題点について述べる。

### A scalable distance learning server system by generic load balancers

Y. Tan<sup>†</sup>, M. Hirose<sup>‡</sup>, M. Shikida<sup>††</sup>, Y. Inoguchi<sup>††</sup>, K. Fujieda<sup>††</sup>, T. Matsuzawa<sup>††</sup>, Y. Kumagai<sup>†††</sup>

<sup>†</sup> Japan Advanced Institute of Technology, School of Information Science

<sup>††</sup> Japan Advanced Institute of Technology, Center for Information Science

<sup>‡</sup> Tokyo Electron Ltd., Computer and Network Section

<sup>†††</sup> Tokyo Electron FE Ltd., Computer and Network FE Dept.

This paper presents the design methodology for scalable distance learning server system with generic load balancers. In the distance learning system which serves for clients both inside and outside of the campus network, a firewall system is required to apply different policies for each clients, and the firewall tends to be a bottleneck in the performance. This paper presents a system design that utilizes load balancers for firewalls as well as the servers, and discusses its effect and problems with an example of JAIST distance learning server system.

#### 1. はじめに

教育機関における IT 環境の向上と、社会人再教育やより専門性の高い講義への要求の高まりなどから、Web やストリーミングを用いた教育システムが注目を集めるようになってきた。

これらを実現するための構成要素は既に世の中に出回っているが、実用に供するにはスケーラビリティの確保が重要であり、履修可能な学生の数を現実的なところまで増やすためには様々な工夫が必要とされている。本稿では、Web とビデオストリーミングを対象に、汎用のロードバランサを用いてシステム全体としてのスケーラビリティとアベイラビリティを向上させるシ

ステムについて報告する。

大学における遠隔教育サービスでは、学内からのアクセスと学外からのアクセスが同時に行われるため、それぞれに対して異なったセキュリティポリシーを適用する必要があり、ファイウォール装置を含めた構成でシステムを構築しなければならない。

こうしたファイウォール装置は、多くの場合ネットワークパフォーマンス上のボトルネックとなり、特にストリーミングデータの利用を考える上で、ファイウォール装置の負荷分散が重要な課題となる。また、これらはトランスペアレントな装置でありながら、学内、学外、DMZ の 3 インタフェースを有するため、単純に複数のサーバーに負荷分散をするような構成に比べ、

より複雑な振り分けを必要とする。

本稿では、汎用のファイアウォール装置を複数使い、学内、学外、DMZ の何れにもロードバランサを配置することで、サーバーに対する負荷分散と同時にファイアウォール装置の負荷分散を実現し、ユーザー数に対するスケーラビリティを確保するとともに単一の装置の故障においてもサービスを停止することなく稼働可能なシステムの構成方法について報告する。また、ストリーミングサービスにおけるファイアウォール装置の負荷およびその負荷分散の効果について示す。

## 2. 遠隔教育システム

IT を活用して、教育における空間的、時間的制約を解消しようとする e-Learning と呼ばれる動きは米、欧では我が国に先行して行われており、英語に対する障壁のない東南アジア諸国では、自国の大学の学位に加えて遠隔教育で得られた学位を積極的に利用するようになるなど、既に国境を越えた教育市場での競争が始まっている。我が国においても法改正により IT を利用した単位取得が可能となり、今後、日本国内でも遠隔教育が日常的に行われる可能性が高い。

大学における遠隔教育には、通常的高等教育と同様、

1. 講師が多数の学生に同一の内容を教授する講義型
2. 講師と限られた人数の学生とが討議しながら教育を行うゼミ型
3. 学生同士が互いに影響を与えながら学習をすすめる協調型
4. 学生が独力で教材を学習する自習型

の各形態が存在する。

究極的にはこれら全てを遠隔教育とすることも可能であるが、現実的にはこれらから効果の高いものを選択して実施していくことになる。表 1 にこれらを実現するシステムの例を挙げる。

これらのうち、その本来の目的から、あまり多人数で実施することはないものと考えられる同期型のゼミ型、協調型を除くと、何れの型においても数十から数百の学生数を対象とする可能性がある。

Table 1. Types of distance learning and examples

	時刻	システム例
講義型	同期	ライブストリーミング
	非同期	ビデオオンデマンド
ゼミ型	同期	テレビ会議システム
	非同期	掲示板、メール会議
協調型	同期	テレビ会議システム
	非同期	掲示板、メール会議
自習型	同期	-
	非同期	WebBasedTraining

実際に遠隔教育として運用を行うためには、上述のような直接的に教育活動を支援するシステムだけではなく、

- A) 進捗状況や成績の管理を行うマネジメントシステム(Learning Management System)
- B) コンテンツ制作を支援するシステム
- C) 各システムの動作を支援するネットワークおよびキャッシングシステム

などが構成要素として必要になる。

A)は空間的にも時間的にも共有するもののない遠隔教育において、ひとつの教育機関としてのまとまりを実現するために必要な要素であり、システム的にはデータベースとそのインタフェース(教師用、事務員用、学生用)となる。B)は講義型や自習型において教師が用意する教材の制作と管理を支援するもので、C)は全システム共通のインフラとなる。これらは遠隔教育を実用規模で実施するためには何れも必須のものであり、また、学習システムと進捗管理など、互に関連性を要求するシステム群を形成する。A)、B)のマネジメント関係のシステムでは標準化の動きもあるが、まだ発展途上で規格に基づいた相互運用性などは望めないのが実情である。

本稿では、表 1 に示された直接的に遠隔教育活動を行うシステムに関して述べていく。

### 3. 遠隔教育サーバー

遠隔教育システムは、以下のような点で通常の e-commerce などのシステムと異なっている。

- (ア) 遠隔教育では講師が時間軸に沿って学生

の学習を誘導するということが重要であり、ストリーミングの要素が必須である

- (イ) 学外からのアクセスと同時に学内からのアクセスがあり、それぞれに対し、異なるセキュリティポリシーを適用する必要がある

後者については、学内も学外同様のサービス対象と考える運用形態をとることも可能であるが、学内からはコンテンツ提供者である教師や利用者の管理を行う事務員のアクセスがあり、これらのアクセスの利便性を考慮すると、ポリシーを分けるのが、多くの場合妥当である。

こうした要求を満たすためには、遠隔教育システムのサーバーはファイヤウォールの DMZ に設置することになる。

一般に、ファイヤウォール装置はセッションなどの複雑なルールを実現するため、ソフトウェアにより実装されていることが多く、パフォーマンス上のボトルネックとなりやすい。これは、サーバーサイドでの処理が行われる場合、同一のクライアントからのリクエストは継続して同じサーバーに届かねばならないため、状態を保持した振り分けが必要になることによる。

遠隔教育においても Web Based Training を中心に、単なる read only のコンテンツに留まらないものが増えているため、同様の問題が発生する。

つまり、遠隔教育用サーバーにおいては、ストリーミングのようにファイヤウォールに対して複雑な処理は要求しないが高いスループットを求めるアプリケーションと、e-commerce 同様のセッションを保持したサーバー負荷分散を要求するアプリケーションとが混在することになる。

これに対し、

- I. ハードウェアを援用した高速なファイヤウォール装置を用いる
- II. アプリケーションの種類ごとにセグメントを分割する
- III. ファイヤウォール装置も負荷分散の対象とする

といった対応方法が考えられる。I は構造が単純で運用も容易であると考えられるが、性能や信頼性はファイヤウォール装置に依存することとなる。また、記述できるルールの複雑さと実

効性能、およびコストの点が問題となってくる。II は複数のグループに分けてファイヤウォールや負荷分散をそれぞれについて行うことになるため、全体としては規模が大きくなるものの、各グループを見た場合には比較的単純なシステムとして運用できるものと思われる。一方、こうした運用を行うには、最初から各サーバーの役割が明確で、かつ、各々がどのような動作をするのかを把握できる管理者が構築・運営せねばならない。III の方法はファイヤウォールもロードバランサによって負荷分散されるサービスのひとつとして捉えてしまう方法である。全体としての性能と信頼性はロードバランサが決めることになる。これは I や II とは違った意味で単純さを有しているが、DMZ つきのファイヤウォールを負荷分散するためには全てのインタフェースにロードバランサを必要とし、通常、耐故障性のために 2 台一組のロードバランサを用いることを考えるとロードバランサだけで 6 式を必要とする。

これらの得失を考慮すると、対象となる遠隔教育システムの内容、管理体制など、各サイトごとに最適な方法は異なってくるものと考えられるが、III の方式は、一度構築してしまった後は各要素の能力が不足した際に同様のものを増設するという形で必要に応じて対応できることと、サーバーの数や内容が変化しても運用方法があまり変わらないといった特徴を持つ。

これは、純粋に技術的な問題だけではなく、教育機関においては多くの場合、システムの予算面から計画的、継続的な投資が難しく、管理運用にかかる費用を十分に計上することができないといった点も実運用上重要な問題であり、不定期につく予算を有効に活用して拡張を行えることや、故障が生じた際、補修のための予算が獲得できるまで運用を止めないでも済むことを考えれば、III の方法は有効であるものと考えられる。

#### 4. JAIST 遠隔教育システム

JAIST(北陸先端科学技術大学院大学)においては、1997 年からビデオオンデマンド、ビデオストリーミングをはじめとした遠隔教育システムを開発、整備してきた。教師と学生が同期してライブ通信する講義型やゼミ型については、家庭用デジタルビデオなどで用いられている DV データを ATM 上で伝送する装置を開発し、2000

年からは総務省総合通信研究所と接続し、講義を JGN 経由の遠隔の形態で行ない、単位を認定してきている。蓄積型の非同期講義型システムについては、2000 年からはビデオと Web を用いた本格的なコンテンツ開発が行われ、現時点で 4 教科については一学期分全ての内容が収録されている。また、2001 年には学内組織として遠隔教育センターも設置され、体制的にも本格的な立ち上げを目前に控えている。

現時点における遠隔教育サーバーセグメントを図 1 に示す。このシステムにはストリーミングサーバー、Web サーバー、Java アプリケーションサーバーなどがあり、サーバーセグメント内でこれらのサーバーから利用されるファイルサーバーやデータベースサーバーが存在している。前節で述べた方法の III を用い、ロードバランサによりファイアウォール装置と各サーバーの負荷分散を行う設計である。用いているロードバランサは f5 社製 BigIP、ファイアウォール装置は WatchGuard 社製 FireBox である。

ファイアウォールの内側は学内に接続されているが、現時点では学外に対する公開を行っていないため、外側は実験運用を共同で行っているサイトに専用線や JGN を用いて接続されている状態である。

ノンブロッキングな Layer2 および Layer3 スイッチを用いて構成することにより、中間経路でのボトルネックを廃している。経路制御は L3 スイッチ、ファイアウォール装置、ロードバランサにて static 記述することにより行っている。

受講者は Web によるインタフェースでサインインした後、WBT、VOD などのサーバーに誘導される。VOD においては、MPEG2(6Mbps)や MPEG1(1.5Mbps)による高品位映像で講義の様子を収録したピュア VOD 型と、SMIL により RealServer の 250Kbps 程度の映像に講義中で使われているスライドと文字起こしをした講義内容とが同期して表示される Web 併用型が用いられており、現在のところ、全ての講義について両方のコンテンツが開発されている。

現在の構成では VOD サーバーが増強された結果、ストリームポンピング能力は合計すると 400Mbps 程度となり、ファイアウォールはもとより、ロードバランサのネットワークインタフェースの速度自体も超えるため、MPEG2 デコーダを有する専用端末は入室できるメンバーに限

られているラボに設置してサーバーと同一のサブネットに接続するなどの措置で緩和をはかっている。

## 5. ロードバランサ方式の効果と課題

JAIST 遠隔教育システムでは、当初の目的通りファイアウォール越しの VOD 利用をロードバランサによる複数のファイアウォール装置への分散という形で実現することができた。単一のファイアウォール装置では MPEG1 を利用するクライアントが 10 を越えたあたりからファイアウォール装置の負荷によりほとんどのサービスに障害をきたすようになったが、現在のところ、4 台までのファイアウォール装置を増設して、ほぼ比例した利用可能クライアント数を得られている。但し、これは単純に vod クライアントのみが存在したときの数であり、ほとんどがファイアウォール装置のパケットフォワーディング能力に依存しているものと考えられるため、トランザクション処理が平行して行われた場合にはまた異なった挙動を示すものと考えられる。

耐故障性に関しても、今までの実験運用期間内でもファイアウォール装置や、ロードバランサが数回障害を起こし、また、VOD サーバー、Web サーバーのバージョンアップが行われたが、何れも正常に稼働しているコンポーネントを用いた継続運用が可能であった。

しかしながら、いくつかの問題も存在する。

まず、一部の VOD サーバー(Ver. 4 までの MediaBase など)では負荷分散を行うにはやや特殊な設定を必要とした。これは、ロードバランサで負荷分散を行う際にはロードバランサが持つ仮想サーバー IP アドレスに届いたリクエストをロードバランサに登録されている該当サーバー群のいずれかのアドレスに再度転送することになるが、問題のサーバーではアプリケーション層の http ヘッダ内に書かれている hostname 部が仮想サーバー IP アドレスのものであって、VOD サーバー自身のものではないことを理由にリクエストを拒否するためである。このようなチェックを行うサーバーはすなわち NAT にも対応していないことを意味しており、逆に NAT に対応できないサーバーは負荷分散装置との運用で同様の問題を起こす可能性があると考えられる。対処としては、該当する全てのサ

ーバーが仮想サーバーIP アドレスを自分の名前として認識できるようにする必要があり、やや特殊な運用を強いられることになる。

次に、サーバーの管理運用において、学内からの ssh, telnet, ftp などのアクセスを許すことは可能であるが、これらの TCP セッションがロードバランサのタイムアウトにあつて切られてしまうことが挙げられる。これは、使わないでしばらく放置したセッションが切られるのみで多くの場合に実害はないが、端末で行っている内容に依存せずネットワークの途中で切られてしまうという点で、ファイルのオープンを含めた操作を行っていた場合などに問題となることがある。特に、コンテンツ管理などの面でこうした使い方をする教師ユーザーは必ずしも計算機システムの専門家ではないことを考えると問題として取り上げざるを得ないものである。我々が用いているロードバランサではプロトコルに依存したタイムアウト時間の設定ができないため一律に値を決めねばならず、多数の http リクエストがある可能性を考えると極端に長い時間に設定することは望ましくない。ロードバランサ内のテーブルエントリの消費量を見ながら適切などころまで延ばすようなチューニングを必要とする。

その他、トランスペアレントなデバイスの生存確認をするために適切な宛先と経路を選択して設定する必要があることや、機器の数とサブネットの数が多く、ワイヤリングを単純にするためには一つの Layer3 スイッチ内の VLAN で構成すると楽になるが、信頼性の点でここがボトルネックになってしまうこと、トランスペアレントなデバイスは生存確認はできていても負荷状況は容易にはわからないこと、といった一般的な実装上の問題も残されている。

## 6. 今後の課題

本システムは概ね想定通りの効果を上げているが、いくつか課題が残されている。

まず、前述したように、VOD サーバーのようなストリーミングを多用するシステムではロードバランサがパフォーマンス上のボトルネックになる。これは、ロードバランサ自体の能力を上げることや、2 台のロードバランサを Active-Standby 運用ではなく Active-Active 運用にしてロードバランサ自体の負荷を分散することが考

えられるが、問題の一つの原因は学内と学外に対して同様の構成となっているところにある。

学外への接続は、多くの場合帯域が限られており、これを対象とした場合にはコンテンツそのものの帯域を見直すことになるため、現在のファイヤウォール装置やロードバランサでも対応が可能となる。そこで、現在のような内、外、DMZ の 3 インタフェース型のファイヤウォール装置ではなく、学内と学外の両側にファイヤウォール装置を置き、中間のセグメントを DMZ とする構成をとれば、学内と学外の必要帯域に応じて機器を変えることが可能となる。もっとも、学外側もダークファイバによる接続や、物理的には近接した LAN 技術で到達できるオープンスペースなどであった場合には、結局学内と同様の設備が必要になってくる。

また、各機器に必要とされる機能の絞り込みもコストを考慮する上では必要な検討事項である。上記のような配置の場合にはロードバランサが 8 台(両方のファイヤウォール装置の両側に 2 台ずつ)となるが、最近のルーターには簡単なロードバランサ機能が含まれているものもあり、また、ロードバランサにも高度なサーバー選択アルゴリズムは持たない安価なものもでてきている。これらをうまく組み合わせることで、全体としての信頼性、性能、コストのバランスを向上させることができるものと思われる。

当然のことながら、運用上のデータ収集も重要な課題である。JAIST 遠隔教育サーバーも次第に実験段階から実運用に入ろうとしており、定常利用時のデータ収集を行う予定である。

## 7. まとめ

本稿では遠隔教育関連のサービスを行うシステムに必要な要件をまとめ、汎用のロードバランサによりサーバーのみならずファイヤウォール装置も含めた負荷分散を行うことでスケラビリティと運用方法の継続性のあるシステムを構築する方法を提案した。

### 参考資料

1. <http://www.aicc.org/>
2. <http://www.aic.gr.jp/>
3. <http://www.nacse.com/>
4. <http://www.nacse.co.jp/>
5. <http://www.webct.com/>
6. <http://www.f5.com/>
7. <http://www.extreme.com/>
8. <http://www.tel.co.jp/cnd/>
9. <http://www.jaist.ac.jp/~ytan/>

# JAIST Distance Learning Server Segments

2001. fall

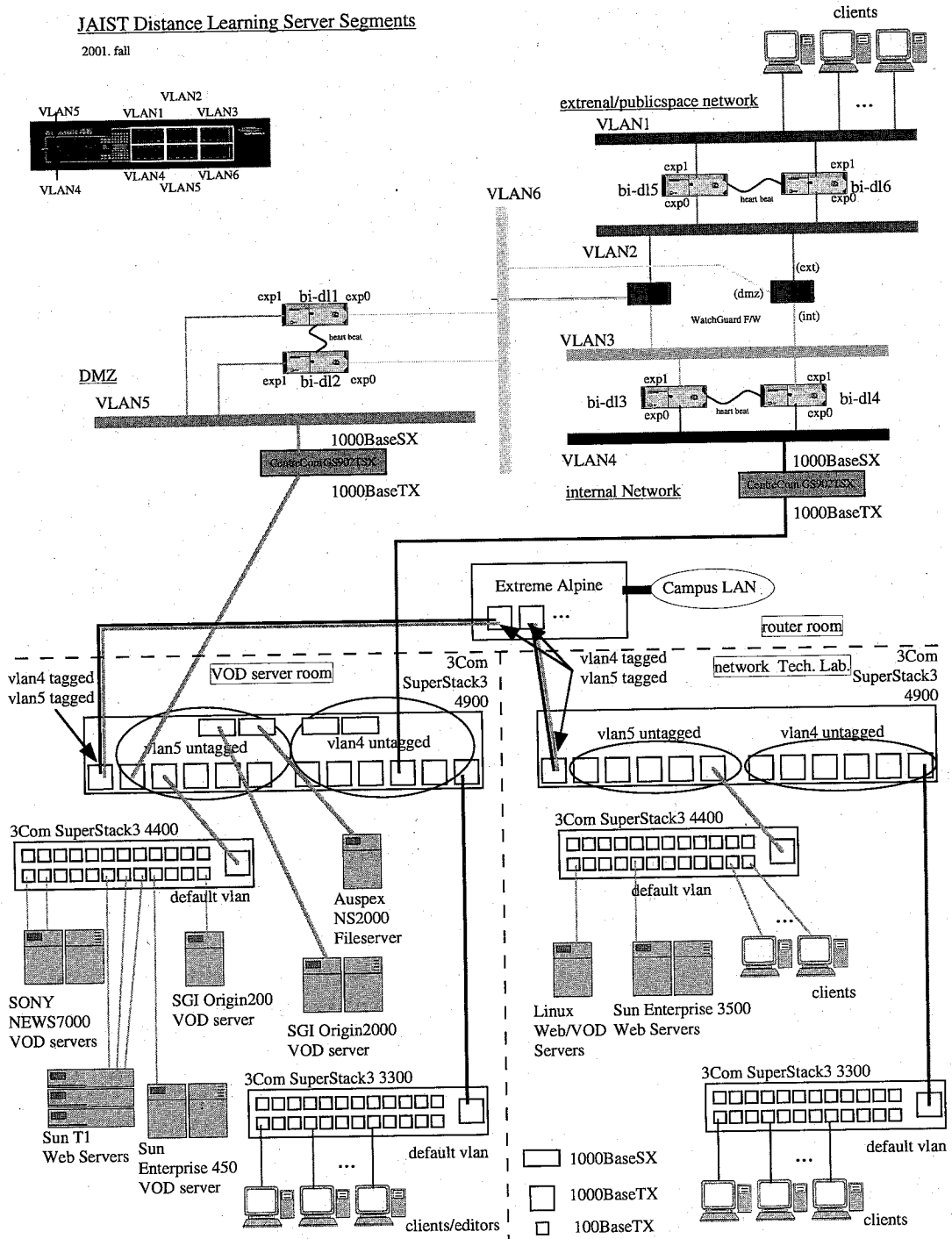


Fig. 1 JAIST distance learning server segments