

平成 14 年度

修士学位論文

CDN サービスに適用する映像コンテンツ 巡回配布方式 Relaycast 方式の検討

Relaycast Examination of A Cyclic Video Streaming
Contents Distribution System

1055127 山岡 徹也

指導教員 島村 和典

2003 年 1 月 31 日

高知工科大学大学院 工学研究科 基盤工学専攻
情報システム工学コース

要 旨

CDN サービスに適用する映像コンテンツ巡回配布方式 Relaycast 方式の検討

山岡 徹也

インターネットにおいて、リアルタイム性を有し、一定のビットレートでデータが連続的に発生する映像コンテンツを配信するための技術が盛んに研究されている。最近では、ネットワーク負荷あるいはサーバ負荷を広域に分散させることで品質保証を狙ったコンテンツ配信インフラストラクチャである CDN(Content Distribution Networks) が注目されている。本稿では、CDN におけるコンテンツのサロゲートサーバへの複製配置に着目し、ユニキャストあるいはマルチキャストと比べて必要帯域の規模適応性 (Scalability) に優れた映像コンテンツの巡回配布方式、Relaycast を提案する。提案方式により、コンテンツ 1 個分の帯域で、複数のコンテンツをサーバ間で配布し合うことが可能となる。

キーワード CDN, Relaycast, 巡回配布

Abstract

Relaycast Examination of A Cyclic Video Streaming Contents Distribution System

Tetsuya Yamaoka

Delivering video content generating the code stream in a constant bit rate and having the time constraints, has been widely studied especially for the Internet. Researchers and developers are paying their attentions for the technology called CDN(Content Distribution Networks) in these days. CDN are expecteded as an infrastructure to enhance the quality of content delivery by way of decreasing the network load and the server load. I focus my interests on the content distribution subsystem on CDN. A cyclic distribution of video content among the surrogates is proposed in this paper. Proposed scheme has the advantage compared to unicast or multicast-based distribution in the sense of scalability of bandwidth. It enables to distribute the multiple contents among distributed servers within using the bandwidth that one content consumes .

key words CDN, Relaycast, Cyclic Distribution

目次

第 1 章	序論	1
1.1	研究の目的	1
1.2	研究の背景	2
1.2.1	コンテンツの配布形態の変化	2
1.2.2	ネットワークの高速化・広帯域化	2
	インターネットプロトコル	2
	バックボーンネットワークの高速化・広帯域化	3
	アクセスネットワークの高速伝送技術	4
1.2.3	負荷分散技術	7
	ネットワーク層	8
	トランスポート層	9
	アプリケーション層	10
1.2.4	まとめ	12
第 2 章	CDN におけるコンテンツ配布	13
2.1	CDN	13
2.2	オリジンサーバからサロゲートサーバへの配布方式	14
2.2.1	ユニキャストの場合	14
2.2.2	マルチキャストによる配布	16
2.3	まとめ	17
第 3 章	提案する Relaycast 方式	19
3.1	システム構成	19
3.2	配送手順	19

目次

3.3	優位性	20
3.4	機能要求条件	21
3.4.1	機能要求	22
第 4 章	Relaycast 管理サーバ	24
4.1	Relaycast 動作の為の機能	24
4.1.1	Relaycast 動作のためのサービス機能要求	24
4.1.2	Relaycast サーバの経路決定	24
	発見	25
	経路計算	26
	経路決定についての課題	26
4.1.3	障害時対応	26
	トークンリングや FDDI における 2 重化	27
	提案する障害対応策	27
4.2	ユーザーの為の機能	28
	リクエストルーティング	28
	提案する手法	30
	Content Information Base(CIB)	30
	ユーザの為の機能要求	31
第 5 章	実験	32
5.1	実験系	32
5.1.1	実験内容	33
	予備実験 01	33
	予備実験 02	33
	実験	34
5.1.2	結果	35

目次

5.2	課題・展開・発展	35
	CPU 負荷の問題	35
	帯域保証と IP 通信網上の運用について	36
	リアルタイム多チャンネル送信の実装	36
	アクセス系の検討	36
	コンテンツの内容記述	37
第 6 章	まとめ	38
	謝辞	39
	参考文献	40
付録 A	MPLS トラヒックエンジニアリング	42
A.1	インターネットの発展と MPLS	42
	A.1.1 はじめに	42
	A.1.2 MPLS 技術と動作概要	43
	A.1.3 IP ネットワークにおけるトラヒックエンジニアリング	43
A.2	MPLS トラヒックエンジニアリング	44
	A.2.1 MPLS トラヒックエンジニアリングの 3 つの流れ	45
A.3	まとめ・これから	47
付録 B	IPv6 のフローラベルの使用法	49
B.1	はじめに	49
B.2	IPv6	49
B.3	IPv6 とフローラベル	49
B.4	IPv6 フローラベルの利用法	51
	B.4.1 RSVP	51

目次

B.4.2	Diffserv	51
B.4.3	MPLS	52
B.4.4	MPLS と IPv4 と IPv6 を統合する機能	52
B.4.5	まとめ	52

目次

1.1	ATM	4
1.2	multicast	8
1.3	L4Switch	9
1.4	ユニキャストによる配布	11
1.5	CDN による配布	11
2.1	CDN	14
2.2	ユニキャストによる配布	15
2.3	マルチキャストで 1 台で配布する場合	16
2.4	マルチキャストによる配布	18
3.1	Relaycast 構成	20
3.2	Relaycast 手順	21
4.1	リングネットワークでの 2 重化	27
4.2	障害の場合	28
4.3	CDN のユーザー側動作	29
5.1	実験環境構成図	32
5.2	予備実験 02 結果	33
5.3	測定結果：S2 パケット	34
5.4	測定結果：S3 パケット	34
5.5	測定結果：CPU	35
A.1	MPLS	44
A.2	Explicit ルーティング	46

目次

A.3	Fast ReRoute	46
A.4	Diffserv	47
B.1	IPv6 ヘッダ	50
B.2	MPLS over IPv6	53

表目次

1.1	Fast Ethernet/Gigabit Ethernet/10Gbit Ethernet	6
5.1	予備実験 01・結果	33

第 1 章

序論

本章では、まず研究の目的を述べ、それに至るまでの技術的背景を示す。映像コンテンツのデジタル化とネットワークの高速化、CDN(Content Distribution Network)をはじめとするネットワークの負荷分散技術の進展によって、ネットワークを介した映像利用が盛んに行われるようになるを述べている。

1.1 研究の目的

現在のベストエフォート型のインターネットアーキテクチャによる配信では、リアルタイム性を有する映像や音声の品質を保証できない。従って、リアルタイム性を有する映像や音声の品質を保証できない。よって、インターネット配信の際の品質保証を実現する研究が盛んである。その研究の 1 つに CDN がある。CDN とは、コンテンツのオリジナルを持つオリジンサーバ、その複製を持ち、地理的に分散配置されるサロゲートサーバ、ユーザ端末からの配信要求を受け付け、どのサロゲートサーバからの配信が最も望ましいか決定・誘導を行う Redirector から成る上位層ネットワークである [1]。現在、Web コンテンツ配信における事例が多いが、今後は映像コンテンツを対象とした事例も増えてくるものと思われる。本検討では CDN において、映像コンテンツをオリジンサーバからサロゲートサーバへの配布方式について提案・実証するのが目的である。

1.2 研究の背景

1.2 研究の背景

本研究の背景にある技術としては、大きく分けて3つの要素がある。1つはコンテンツ配布形態の変化、1つは情報ネットワーク(以下ネットワーク)の高速化であり、もう1つは、CDN技術などの負荷分散技術の発展である。ここでこれらの技術の発展の経緯を振り返っておく。

1.2.1 コンテンツの配布形態の変化

現在、新聞、テレビ、音楽、映画など全てのコンテンツが、アナログからデジタルへ、そしてインターネット上のコンテンツへ移り変わろうとしている。1998年と2000年に発表された調査結果 [2][3] によれば、インターネット上に流れている最も多いデータはHTTP[4]であり、以下 Real Video/Audio[5]、DNS(Domain Name System)[6]などが並んでおり、トラフィック総量の伸びとともに、RTP/RTSP等のコンテンツのデータ割合も増加している。このことから、インターネットが主にコンテンツ配信の手段として利用されており、連続メディアによる映像配信がますます顕著になっている。そして利用者の増加と回線の広帯域化は著しく、ADSL、FTTHなどのアクセス網の高速化と普及に目を見張るものがあるが、映像配信の普及に伴っていくつかの問題も起きている。

1.2.2 ネットワークの高速化・広帯域化

インターネットプロトコル

1980年代半ばに生まれたインターネットは、シンプルイズベストを標榜し、接続性を優先したため、品質保証等の付加機能は考えられていなかった。最大限努力するベストエフォート型のプロトコルであるため、特に映像を配信する場合は、パケット損失や遅延が再生品質に大きく影響する。ここでは、TCP/IPの基本となる通信方式であるTCPとUDPについて説明する [7]。

1.2 研究の背景

- TCP

データのやり取りの際に、パケットごとにシーケンス番号を付与しておくことや、ACK パケットのやりとりをすることによって、データの到着や到着順序が保証されている通信方式である。また、チェックサムを用いてデータが正しいかどうかを確認したり、データの転送速度を相手の処理能力やネットワークの状態に応じて調整する機能も持っている。しかし、到着確認などのやりとりをしている分、通信間隔が長くなる。よって、通信時間より信頼性が重要な場合に利用される。画像配信について言えば、次に説明する UDP を用いた通信と比べると、リアルタイム性が低くなる。

- UDP

UDP を用いた通信の場合は、ACK を返さずデータを送信をする。データは、チェックサム認証はするが、エラーが発生すると破棄するだけなので、到着保証されていない。TCP と違い、再送といった処理がない分、高速の通信が可能である。よってリアルタイム性が重視される映像配信に適している。

バックボーンネットワークの高速化・広帯域化

ネットワークの高速化と広帯域化は、光ネットワークがバックボーンの高速化と広帯域化を支える。技術的には ATM (Asynchronous Transfer Mode)[8] や DWDM (Dense Wavelength Division Multiplexing) 等が使われ、数百 G(ギガ)bps 級から T(テラ)bps 級のといった膨大な伝送容量が光ファイバから生み出される。また交換点であるルータにおいても、光ルータと呼ばれる装置により、従来の銅線に落としてのルーティングから直接光のままルーティングを行うものも実用になろうとしており、高速化・広帯域化へと進んでいる。現在のバックボーンにおいては、高速交換技術のひとつである ATM が導入されていることが多いため、ここでは、ATM について述べる。

ATM ATM は、OSI 参照モデルのデータリンク層および物理層に相当するデータ交換技術であり、図 1.1 にその構成を示す。ATM では、53 バイトの“セル”と呼ばれるもので

1.2 研究の背景

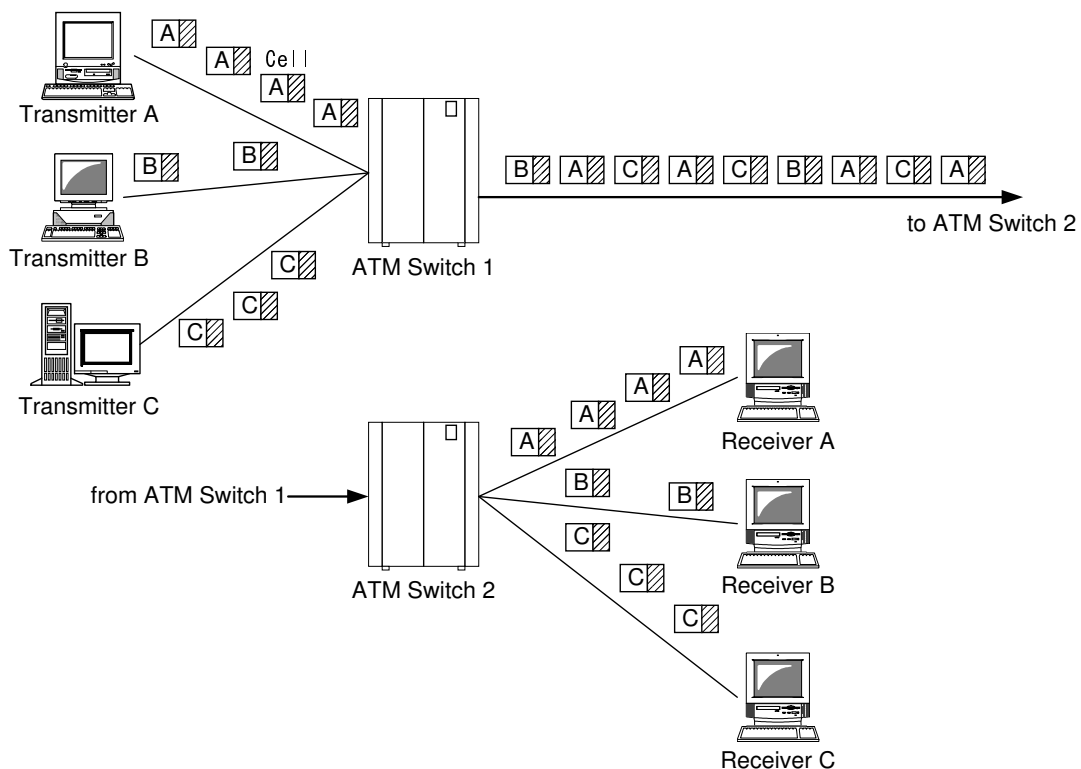


図 1.1 ATM

スイッチングする．セルは固定長のため，効率のよい単純なスイッチングロジックをハードウェアに実装できるため高速化できる．例えば音声や映像といった，あらゆる種類のトラフィックにも適合できる．近年，インターネットの発達により，IP データグラムが ATM 網で転送できる IP over ATM と呼ばれる技術が普及している [9]．

アクセスネットワークの高速伝送技術

アクセス網においても広帯域回線によるインターネット利用の時代が到来しつつある．最近の急速な ADSL(Asynchronous Digital Subscriber Line) や FTTH(Fiber to the Home) の普及により，アクセス網も高速化されている．また Ethernet においても，Gigabit Ethernet 等が普及段階に入り 10Gigabit Ethernet も実用段階に入っている．この結果，ユーザの広帯域を必要とする映像コンテンツ利用への要求は増すばかりである．ここでは，Gigabit Ethernet/10Gigabit Ethernet[10]，xDSL について述べる．

1.2 研究の背景

Ethernet LAN では、1976年に発表された Ethernet が、IEEE802 委員会で標準化されて以降、現在に至るまで最も広く利用されているデータリンク層技術である。当初、Ethernet 規格の伝送速度は 10Mbps であったが、1995年、IEEE802.3u において、伝送速度を 100Mbps にした Fast Ethernet が標準化された。現在は、Fast Ethernet のネットワークアダプタやリピータも安価となり、Fast Ethernet は LAN 環境に欠かせないものとなった。元々 Ethernet は、伝送媒体をスイッチ制御によらず共有していたが、Fast Ethernet では、全二重通信やフロー制御もサポートするようになり、スイッチ型のネットワークとして利用できるようになった。

さらに 1998 年には、IEEE802.3z において、1G bps の伝送速度を目指した Gigabit Ethernet が標準化された。

そして 2002 年、IEEE802.3ae において 10Gbps の伝送速度を目指した 10Gigabit Ethernet が標準化された。

Fast Ethernet/ Gigabit Ethernet/ 10Gigabit Ethernet の特徴 [表 1.1] についてだが、Fast Ethernet は、物理媒体に非シールド撚り対線 (UTP:Unshielded Twist Pair Cable) が多く使われている。そして Gigabit Ethernet の多くに、光ファイバが使われており、最大伝送距離は 1000BASE-LX の 5km である。さらに、10Gigabit Ethernet では、SONET/SDH や ATM で実現されていた信頼性や通信品質制御機能が盛り込まれ、伝播遅延の一因となっていた CSMA/CD 方式は廃止された。そして最大伝送距離は 10GBASE-ER/EW で 40km と小規模な WAN での利用もできる仕様になっており、アクセス網技術としてのみならず、安価で取り扱いが容易であることから、ATM に代わる広域 LAN 環境のバックボーンとして期待されている。

DSL デジタル加入者線 企業や学校では比較的高速なインターネット接続が提供されている場合が多いが、各家庭からインターネット接続する場合、ごく最近までアナログモデムを利用した接続や、ISDN 網による接続が一般的であった。

専用のモデム (MODEM:変復調器) によって、より高い標本化周波数でデジタル信号

1.2 研究の背景

表 1.1 Fast Ethernet/Gigabit Ethernet/10Gbit Ethernet

	Fast Ethernet	Gigabit Ethernet	10Gbit Ethernet
伝送速度	100Mbps	1Gbps	10Gbps
最大伝送距離	100m	5km	40km
フロー制御	対称フロー制御	非対称フロー制御	非対称フロー制御
主な伝送媒体	UTP	光ファイバ/UTP	光ファイバのみ
主な用途	LAN	LAN	LAN/WAN
その他	半/全 2 重	半/全 2 重	全 2 重のみ

を変復調することで、これまでのアナログ電話用の 1 対の銅線のままであっても、数 Mbps 程度の伝送速度を得ることが可能となる技術である ADSL であるが、その後、用途や最大転送レートに応じていくつかの派生技術が生まれた。これを総じて xDSL 技術と言う。

DSL の技術の種類であるが、上りが遅く下りが速い非対称な ADSL(Asymmetric Digital Subscriber Line)、ADSL に転送レートを動的に制御して常に最大転送レートが出ることを目指す RADSL(Rate Adaptive Asymmetric Digital Subscriber Line)、複数対の加入者線を使い 1.5M を目指した HDSL(High bit-rate Digital Subscriber Line)、ADSL を高速化させ、下り 50Mbps 相当の速度を持つ VDSL(Very High bit-rate Digital Subscriber Line) などがある。

ADSL の技術は、2002 年 3 月現在、日本で 238 万人に使用され、アクセス系における現在の映像コンテンツ配送の中心となっている [11]。

その他の技術 現在日本の ADSL のアクセスネットワークは、NTT の局舎までは、ATM による光ファイバでの伝送である、それを局舎まででなく、光ファイバを加入者宅まで敷設し、全ての伝送路に光ファイバを用いる FTTH(Fiber To The Home) など、アクセス網の伝送設備を光ファイバ化する技術が都市部で普及しつつある。

1.2 研究の背景

1.2.3 負荷分散技術

トラヒック負荷分散の重要性

電話網に代表される回線交換方式では、通信の両端を結ぶ回線の中継システムが管理することで、信頼性のある通信を保証している。しかし、1回線あたり流れるデータ量は様々で常に1回線全てを使っているわけではないので、総利用帯域は効率的とは言えない。また、同時に利用できる通信の端点の組の数は中継システムの能力によって限定されるため、中継システムの配置は、資源が競合しないように考慮する必要がある。一方、インターネットに代表されるパケット交換方式では、回線をパケット単位で共有するため、回線交換網と比べ多数の利用者が同時にネットワークを利用できる。つまり、多くの人と同時に利用するような広域ネットワークに適している。しかし、パケット交換方式による中継システムはパケットの確実な配送を保証しない。このため、中継システムの処理負荷が大きい場合や、中継システムに接続されたネットワーク資源が過度の競合状態にある状況では、パケットが破棄されることがある。このように、パケット交換方式は、回線の占有が行われぬ反面、通信を行う終端ノードからネットワークに流入するデータ量の総量が、ネットワークの許容量を超えることがあるという欠点がある。例えば、オリンピックのようなアクセス頻度の高いコンテンツを持つサイトには世界中からのアクセスが集中し、中継回線には過度の輻輳が起きる。今日のインターネット環境において、このような状況は特別なことではない。従って、アクセスの集中を回避し安定的にデータ転送を行うために、インターネットを介して転送されるトラヒックの負荷分散機構が重要になる。また、インターネットでは1対多の通信が可能のため、終端サーバにクライアントからのアクセスが集中した場合、終端サーバの処理能力を超えることがある。サービスを安定的に提供するためには、終端サーバへのアクセスの負荷分散もまた重要である。

様々なレイヤーでのトラヒック負荷分散は以下のように、OSI 参照モデルにおける各階層で負荷分散の手法が提案されている。

1.2 研究の背景

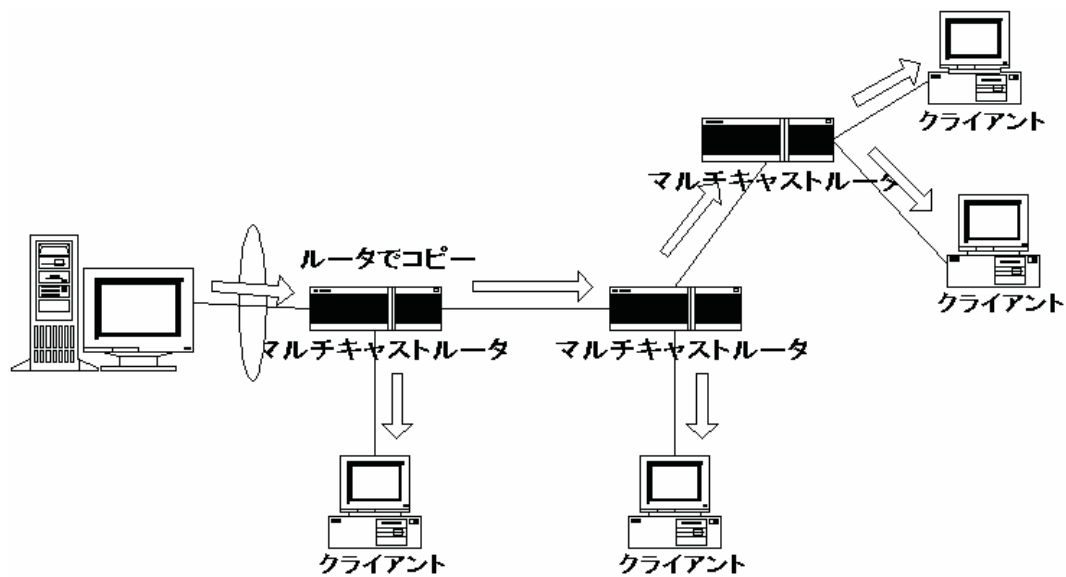


図 1.2 multicast

ネットワーク層

IP マルチキャスト IP マルチキャスト (以降マルチキャスト) を図 1.2 で示す。マルチキャストでは、マルチキャストルーターがデータをコピーする。従って 1 本分の帯域で複数の相手に同時にデータを送信することができる場合に効果がある。送信には UDP を用いており、データの到着と到着順序は保証されていない。マルチキャストを用いるには、マルチキャストアドレスをマルチキャストグループと対応付けさせる。各クライアントは、マルチキャストグループに参加することによって、そのグループに対応づけられたデータを受信することができる。

マルチキャストの問題点は、TCP でなく UDP を使用しているので信頼性のあるデータの送受信が難しいという問題、マルチキャストは、マルチキャストに対応したルーターが必要になるが、その普及が進んでいないという問題やマルチキャストグループへの参加・不参加時に配送ツリーを作り直すという問題などがある。

1.2 研究の背景

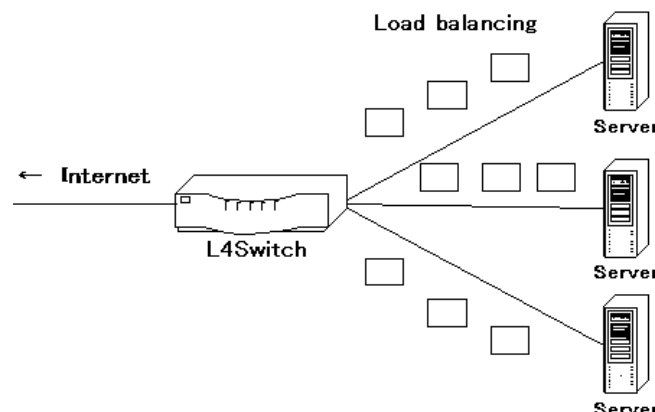


図 1.3 L4Switch

トランスポート層

TCP の輻輳制御 TCP は、転送時に中継回線で輻輳が起きた場合、終端ノード間で輻輳を回避するメカニズムを持つ。そして、終端ノード間の通信でパケットの喪失が起きなければ転送速度を徐々に上げていき、パケット喪失が起こるまで転送速度を上げる。その仕組みにより通信経路の限界をパケット喪失で調べることができ、その値をもとに、適切な転送速度を決定する。

L4 スイッチ 複数のサーバを一つのアドレスを持ったサーバとして仮想的に運用し、ある拠点内での分散を実現する。振り分けにより終端ノードからのアクセスの集中を分散することができる。L4 スイッチの主な機能として、TCP/UDP ポート番号の値をもとにハッシュを行い、コンテンツの分散を図るクラスタリング機能、信頼性を高めるフェイルオーバー機能、スイッチ自体が強制的にリクエストをキャッシュに送り込む透過型キャッシュ機能などがある [図 1.3]。

1.2 研究の背景

アプリケーション層

Content Distribution Network(CDN) CDN とは、ユーザアクセスの高速化とネットワークボトルネックの解消を目的とし、マルチキャストのような配送技術だけでなく、様々な技術を持つアプリケーションの集合体である。負荷分散で言えば、OSI 参照モデルにおけるトランスポート層で行うよりも、CDN 技術に代表される地理的に分散されたキャッシュサーバを利用したアプリケーション層で行われる方が効率が良い。なぜなら、インターネットトラフィックの増加傾向は 1 年で 2 倍になっている [1,2] のに対し、ムーアの法則からホストの性能は 18 ヶ月で 2 倍となるので、ルータやサーバの性能を向上させるよりも、トラフィック自体を分散させた方が効果的である。CDN は 2000 年初頭から急速に立ち上がり、CDN サービスの市場規模は、年々倍のペースで増加している。現在 Web コンテンツ配信による事例が多いが、今後は映像コンテンツ配信が普及してくると思われる。

CDN によるコンテンツ配布 現在、コンテンツ配信において最も利用されているのは、ユニキャスト通信である。ユニキャストの場合、同じコンテンツのリクエストがあると図 1.4 のように経路中のどこかでトラフィックの重複が起きる。クライアントが N 台ある場合、回線には同じトラフィックが N 回流れてしまう。また中継回線でも同様の重複が起こりうる [図 1.4]。

このような重複を最小限にとどめるためにネットワーク上に地理的に分散させたサロゲートサーバをおく。図 1.5 にあるように CDN を利用することによってオリジンサーバはサロゲートサーバだけに配信すればよく、中継回線のトラフィックの重複を軽減する。また、トラフィックの重複だけでなく、オリジンサーバとの接続数も軽減することができるので、ユーザの応答時間も短くすることができる。CDN を利用することによってトラフィックの負荷分散を行い、アクセス頻度の高いコンテンツのトラフィック集中を回避することができる。

1.2 研究の背景

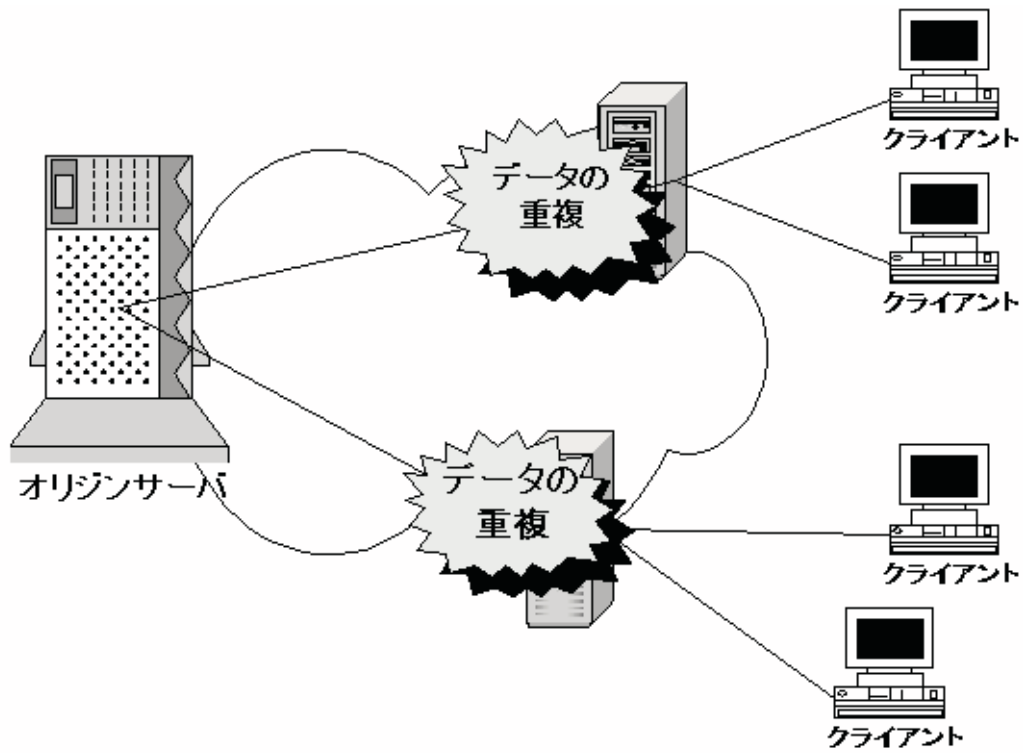


図 1.4 ユニキャストによる配布

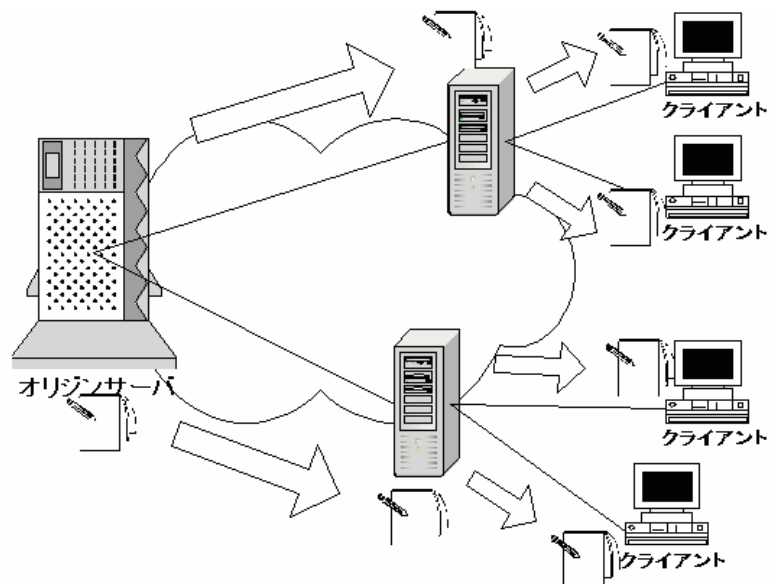


図 1.5 CDN による配布

1.2 研究の背景

1.2.4 まとめ

本章では，研究の目的を述べ，コンテンツの流通形態，ネットワークの高速化，負荷分散の進展について述べた．

インターネットが普及し，放送のデジタル化やその他メディアも IP ベースへと移行しようという動きが進んでいる．インターネットの爆発的普及による中経路の輻輳などに対応するため負荷分散技術が進み，今後益々インターネット上での映像コンテンツ流通が促進されるだろう．

第 2 章

CDN におけるコンテンツ配布

この章では、一般的な CDN の構成を述べ、そして特に、オリジンサーバからサロゲートサーバまでの配布方式の 2 つ、最も単純な CDN コンテンツ配布形態であるユニキャストによる配信と現在、研究提案されている CDN Multicast[12] について比較検討を行う。

2.1 CDN

CDN のシステムモデルを図 2.1 に示す。CDN の構成は、コンテンツのオリジナルを持つオリジンサーバ (Origin Server)、その複製を持ち、広域に分散配置されるサロゲートサーバ、そしてユーザ端末からの配信要求を受け付け、どのサロゲートサーバからの配信が最も適当なものかを決定し、そこへ要求を誘導するリクエストルーティング機能を持つ Redirector から成る。

その動作を図 2.1 で示す。クライアントは、Redirector に要求を出す。そして、オリジンサーバから配布されたデータを、どのサロゲートサーバからダウンロードすれば良いか最適なものを Redirector が選択する。その Redirector に選択されたサロゲートサーバからクライアントまで配送する。このときクライアントからの要求に対し、迅速にダウンロードすることが求められる。このとき、Redirector は、各サーバの負荷状態、クライアントまでの距離 (ホップ数)、RTT(Round-Trip-Time) といった評価値に基づき、どのサーバから転送を行うか決定する。

2.2 オリジンサーバからサロゲートサーバへの配布方式

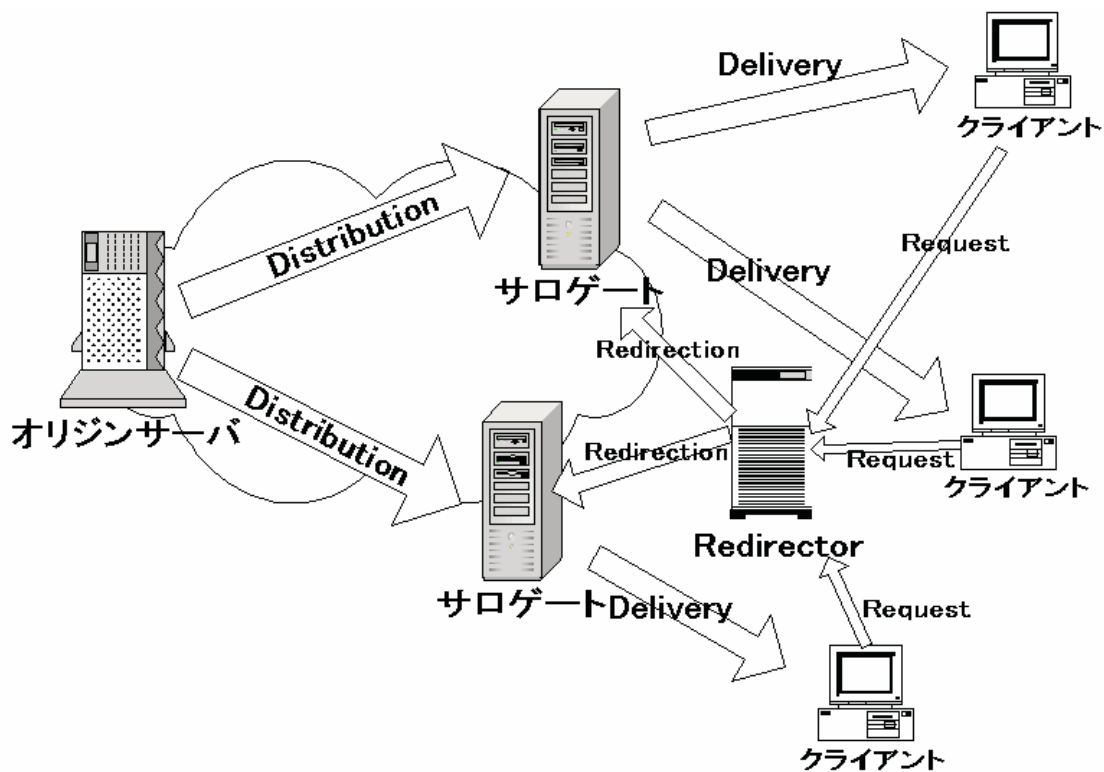


図 2.1 CDN

2.2 オリジンサーバからサロゲートサーバへの配布方式

ここでは、CDN のオリジンサーバからサロゲートサーバまでの配布方式に注目し、ユニキャストによる配布と現在、研究提案されている CDN Multicast による配布について比較検討を行う。

2.2.1 ユニキャストの場合

CDN のオリジンサーバからサロゲートサーバまでの配布とは、オリジンサーバにあるコンテンツをサロゲートサーバに配布・蓄積することを指す。最も簡単な配布方式は、オリジンサーバから各サロゲートサーバに対し、ユニキャストにより送信することである。ユニキャストによる通信は、無駄が多い、そのコストを定量化し示すことにする。

トラフィックに流れる負荷を計算する場合、コネクションパス本数、サーバからサロゲート

2.2 オリジンサーバからサロゲートサーバへの配布方式

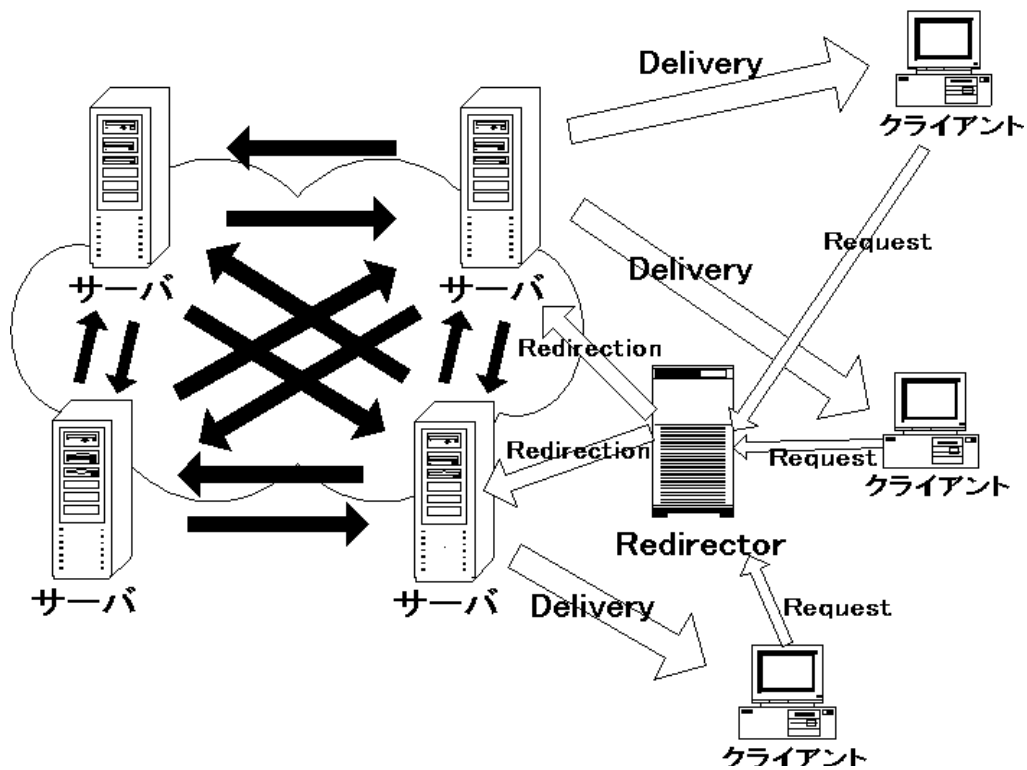


図 2.2 ユニキャストによる配布

までの距離，コンテンツ容量の 3 つが重要になってくる．そこでその 3 つを基準とするトラフィック総量指数を設定し，それに加えトラフィック配信に要する総時間を比較検討する．

N 台のオリジンサーバがあり，それぞれ持っているデータを互いに配布したい．すなわち，各サーバはオリジンサーバかつ他のサーバのサロゲートサーバとなる場合，配布したいコンテンツの容量を V_i ，サーバ間距離を A_i という形に表すと，コネクションパスの本数は，図 2.2 にあるように

$$N(N - 1)$$

になり， N は N^2 に比例する．トラフィック総量指数 bt は，

$$bt = \sum_{i=1}^N V_i(N - 1)A_i$$

となる．

各サーバが， $(N - 1)$ 台あるサーバより同時に送受信可能なら，総転送時間 T は，

$$T = \text{Max} \left\{ \frac{V_i}{b_c} \right\} [s]$$

2.2 オリジンサーバからサロゲートサーバへの配布方式

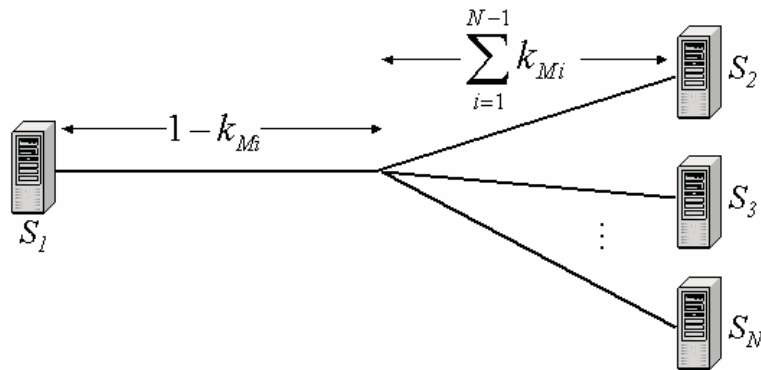


図 2.3 マルチキャストで 1 台で配布する場合

となる．しかし，同時に送受信するための回線帯域は，コンテンツ 1 つあたりの通信帯域を b_c とすると， $2b_c(N - 1)$ の回線帯域が必要になり，規模適用性に問題がある．

また，送受信 1 つずつのみ転送可能とすると総転送時間 T は，

$$T = \frac{N}{b_c} \sum_{i=1}^N V_i[s]$$

となる．

すなわち，サーバ 1 台あたりでも $N - 1$ 台に配布することから規模適用性に問題があり，ソースを持つサーバ数が多くなると，さらに帯域の面で規模適用性に問題が生じることがわかる．特に，高品質の映像になると b_c が大きくなるため，ユニキャストによる配布は難しくなる．

2.2.2 マルチキャストによる配布

経路の分岐とコピーの技術を使った IP マルチキャストを利用をした場合を考える．IP マルチキャストを利用すれば， N 台のサーバがそれぞれ持っているデータを互いに配布し合う場合，マルチキャストによる経路長削減効果が得られる．

その指数を k_{Mi} とすると，パス本数は N に比例する，そしてトラヒック総量指数 bt は，

2.3 まとめ

図 2.3 のようになるので

$$bt = \sum_{i=1}^N V_i (1 - k_{Mi} + \sum_{i=1}^{N-1} k_{Mi}) A_i$$

となる．そして，その総転送時間 T は，各サーバが $(N - 1)$ 台あるサーバより同時に送受信可能なら，ユニキャストと同様

$$T = \text{Max} \left\{ \frac{V_i}{b_c} \right\} [s]$$

以上のような時間となるが，各サーバは $b_c \times N$ の回線帯域が必要になり，規模適応性に問題がある．また送受信 1 つのみ転送可能だとすると総転送時間 T は

$$T = \frac{1}{b_c} \sum_{i=1}^N V_i [s]$$

になる [図 2.4] ．

従って，マルチキャストはユニキャストに比べ経路削減効果のある分有利である．しかし，配送ツリー管理など，マルチキャスト自体の課題も多い．

2.3 まとめ

この章では，CDN の配布方式を説明し，その中でもオリジンサーバからサロゲートサーバまでの配布について，ユニキャスト，マルチキャストでの配布の比較検討を行った．結果，マルチキャストは，ユニキャストより，経路削減効果のある分利点が多い．しかし，マルチキャスト自体の複雑な問題も残っている．

2.3 まとめ

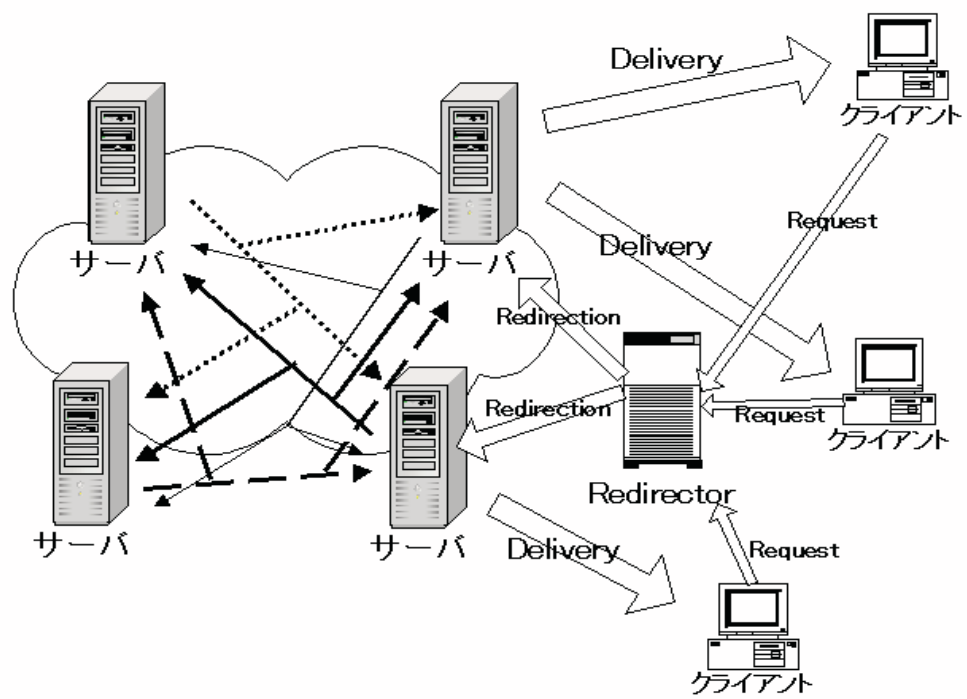


図 2.4 マルチキャストによる配布

第 3 章

提案する Relaycast 方式

これまで挙げたように，複数サーバに入力されている各映像ソースを配布し合うには，従来の方式では，帯域の制約，及びネットワーク負荷の問題に直面する．この問題に対してあらかじめ，ストリーム配布順序を決め，サーバ間を巡回させて配布する Relaycast 方式を提案する [13]．本章では，提案する Relaycast 方式の構成と手順を示し，優位性を示した上で機能要求などを挙げる．

3.1 システム構成

まず，システム構成を図 3.1 に示す．リングを作り出すために自動的に配布順を各サーバに知らせる機能を持つ 1 台の管理サーバ，転送機能，その他諸々の機能を持ち， N 台を地理的に分散している Relaycast サーバ，そして Ethernet 網上でデータのやり取りをする．

3.2 配送手順

N 台のサーバがあり，各映像ソースを配布し合う場合，従来の方式ではネットワーク負荷・総転送時間等に問題がある．これに対してストリームデータをサーバ間で巡回させ，1 ストリーム分の帯域で複数映像コンテンツを各サーバに配布する方式，Relaycast を提案する．Relaycast サーバが S_N までであった場合の送信手順を図 3.2 に示す．配布順は管理サーバが各サーバに知らせる．このとき知らせるのは順序のみで管理も容易である

- (1) S_1 が S_2 へそのデータを送信する．
- (2) 以後， S_2, S_3, \dots, S_N へと送信する．

3.3 優位性

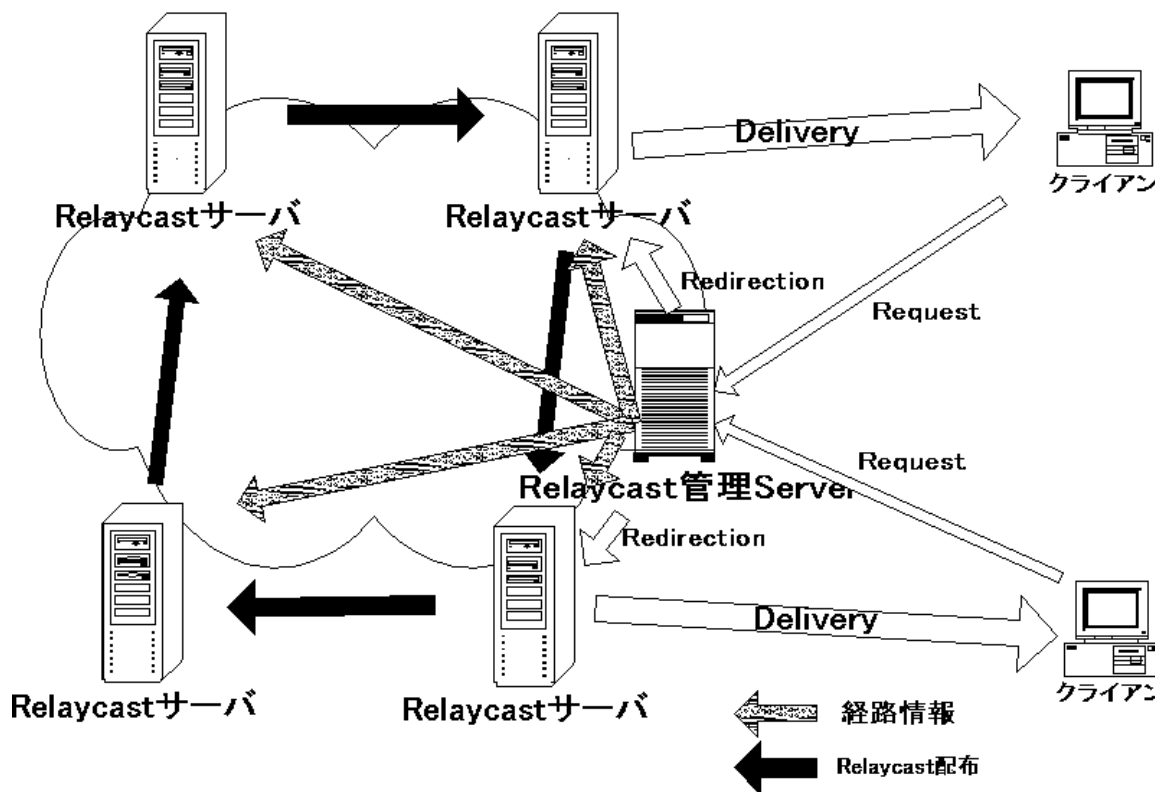


図 3.1 Relaycast 構成

- (3) S_N に到着し終わったら S_N は S_1 にその旨を知らせ、さらに S_1 は S_2 にストリームの送信を指示する。
- (4) S_2 からの巡回配布を実行する。このときの配布順は、 $S_2 \ S_3 \ \dots \ S_N \ S_1$ となる。以後、(2)-(5) と同様の処理を S_N の配布まで繰り返す。

3.3 優位性

Relaycast は経路は隣接サーバ 2 つのみであるから経路削減指数を示すことができる。その Relaycast における経路削減指数を k_{Ri} とすると、前章でも用いたネットワークトラヒックの総量指数 bt は、

$$bt = \sum_{i=1}^N V_i(N-1)k_{Ri}A_i$$

3.4 機能要求条件

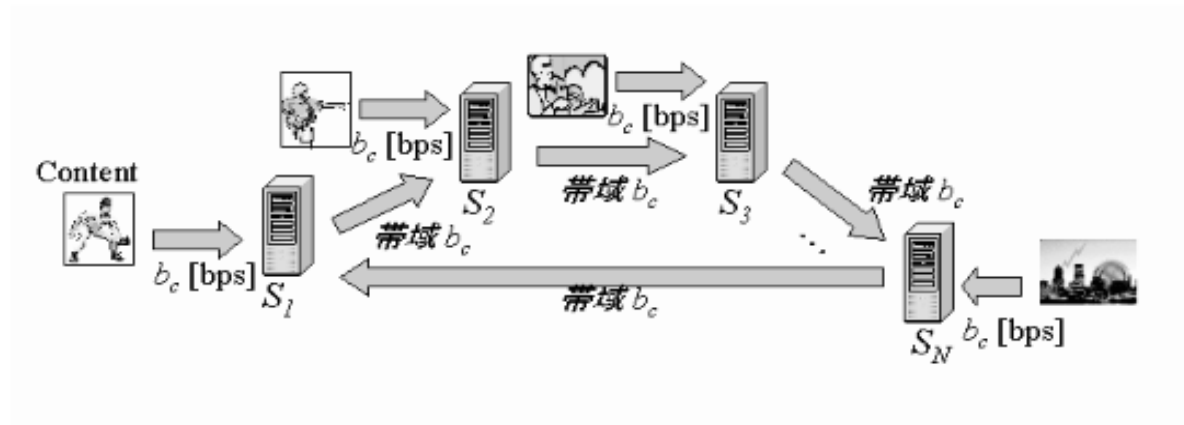


図 3.2 Relaycast 手順

であり経路削減指数 k_{Ri} は，隣接サーバ間にパスを張る状態の指数である．総転送時間 T は，各サーバが 1 つ以上送受信可能であれ不可能であれ，

$$T = \frac{1}{b_c} \sum_{i=1}^N V_i [s]$$

となる．各サーバのパスの通信帯域において，通常のユニキャストでは N^2 に比例し，マルチキャストでも N に比例するが，この巡回配布方式では N がいくら増加しようとも各サーバ隣接間の送受信通信帯域は $2b_c$ で済む．よって，トラフィック総量指数，総転送時間，規模的適用性を比較した結果，一度に送信できる場合のみ時間的に Relaycast 方式が遅くなるが，規模適用性等，他の部分ではどんな場合でもユニキャスト，マルチキャストより優れ，問題を解決できる．

3.4 機能要求条件

Relaycast 方式を満たすための次のような要求を満たす必要がある．

3.4 機能要求条件

3.4.1 機能要求

初期状態決定部

- 管理サーバの機能については 4 章にて述べる。

各 Relaycast サーバは、隣接の 2 つのサーバである Next サーバと Before サーバのみを初期状態から知ってなければならない。管理サーバから送信先、受信元を Relaycast サーバに転送する。

送信部

- 次の Relaycast サーバのアドレス、送信元アドレスしか持たない。
- Next メッセージが来たら次のファイルを送信する機能
- ファイル送信の重複を避ける機能

送信パケットに、できるだけ多くビデオデータを格納し、他は最低限の情報しか持たせないようにする。リングのネットワークを構成する中の隣接サーバの情報以外の情報は、各 Relaycast サーバは持たない。そして送信パケットには、宛先アドレスに、次の Relaycast サーバのアドレス、送信元アドレスに、送信元を入れる

リレー部

- リレーの形でコピー
- コピー時にその Relaycast サーバが持っている隣の Relaycast サーバのアドレスに変換
- サーバの持つ Next Address と宛先アドレスが一致すると送信をやめる。データが来たら宛先アドレスを次の Relaycast サーバのアドレスに変換する。

メッセージ部

- サーバが持つ Next Address と送信元アドレスが一致し転送をやめたら、次のサーバに到着メッセージを流す機能

3.4 機能要求条件

- 到着メッセージが来た場合 Next メッセージを送信する機能
- Next メッセージが来た場合，サロゲートサーバに送信データがあるかどうか判断し，なければ，その次のサーバに Next メッセージを送信する機能

保存部

- ファイル名を付けて保存する機能

第 4 章

Relaycast 管理サーバ

この章では，Relaycast 管理サーバについての提案を行う．管理サーバとは，Relaycast サーバ側では Relaycast の経路制御・障害復旧を行う．また，ユーザに対してはコンテンツ内容を加えた Redirector としての働きを持つ．

4.1 Relaycast 動作の為の機能

4.1.1 Relaycast 動作のためのサービス機能要求

- ・ Relaycast の経路決定し，各 Relaycast サーバへ配信する機能
- ・ 障害時の対応

Relaycast 動作の為の管理サーバ機能要求として以上 2 点を検討する．

4.1.2 Relaycast サーバの経路決定

本提案方式での経路は，パケットの方には次のサーバの送信先アドレス，送信元アドレスのみを格納し，経路情報は各サーバが持つ．このことにより，パケットは最低限の宛先アドレスのみ持つことになり，映像データを効率的に送信できる．また，各 Relaycast サーバはパケットが来たときにパケツリレー方式で送信先アドレスを次のサーバのアドレスに書き換える．

また，小さなシステムの場合は宛先を明示的に指定できるが，大規模システムになった場合，サーバひとつひとつ手動で隣接サーバの設定を行うことは考えにくい．よって，管理

4.1 Relaycast 動作の為の機能

サーバを置き、自動でリング状に連なる各サーバの経路情報を知らせることが望ましいと考える。

Relaycast サーバ同士のリンクは、互いにサーバを見つけ、やりとりするリング状の経路を見つけることができなければならない。各 Relaycast サーバはピア・ツー・ピア (P2P) モデルにおいての「発見 (discovery)」と呼ばれるタスクを利用し、サーバを見つけ MIB による情報や各 Relaycast サーバ間のリンク状況を取得し、得られた情報を管理サーバに送信し、管理サーバ上で判断し経路を決定する。

発見

Relaycast サーバ発見にあたって効果的な 2 種類の方法が考えられる。Napster モデルによる明示的な発見方法と Gnutella モデルによる動的な発見方法である。

明示的な発見

明示的な発見は、Relaycast サーバに他の Relaycast サーバの存在についての情報を持たせ管理サーバ上で管理する方法が最も簡単で非常に効果的である方法である。まず、Relaycast サーバは、Relaycast のリングネットワークに参加したい場合には、管理サーバに参加要求を出す。その後、管理サーバは、Relaycast の参加サーバのアドレスを与え、参加要求を出したサーバに返し、サーバ自身の状況や他の参加 Relaycast サーバとのリンク状況を返し、管理サーバに伝え、経路計算させ、適切な経路を与えるというものである。

明示的な発見方法は、柔軟性は欠如するものの、管理サーバは一定の Relaycast サーバにしかリンクしない上に、認証も容易なので、一定のセキュリティがもたらされる。Relaycast サーバの情報リストによってネットワーク内の各サーバをリング状に構成することによって、各 Relaycast サーバは互いを認識して送受信を行う。そしてインプットに関しては参加サーバ以外から受け付けない仕組みにすることにより、ネットワークは外側からの攻撃に対しても強固になる。

4.1 Relaycast 動作の為の機能

動的な発見

動的な発見方法は、管理サーバが、四方八方に探索パケットを送信し、参加者を探するという方法である。まず Relaycast サーバは四方八方に参加 Relaycast サーバを認識する為のパケットを送信し、その受け取った Relaycast サーバは、サーバ自身の状況や他の参加 Relaycast サーバとのリンク状況を返し、その情報を管理サーバに伝え、経路計算させ、適切な経路を与えるというものである。

この方法は、Relaycast のリングネットワーク参加は、一定の範囲内であるならどこへでも移動可能であり柔軟なものになる。そして、サロゲート同士は隣やその近辺しか知らないということも実現できることから、Relaycast サーバ間同士秘匿性が必要な場合に用いると良いと思われる。

経路計算

経路計算についてはまだ検討中である。各 Relaycast サーバから送信された MIB による情報やリンク状況を利用し、リングネットワークになるよう計算させ、その経路情報を各 Relaycast サーバに配布する。

経路決定についての課題

コンテンツとリンクしてのコンテンツごとの経路計算も考えられる。リングネットワークにて、サーバによっては、既に持っていたり、不要なコンテンツもあるかもしれない、その場合、後述する Content Information Base を利用して、必要なサーバのみでリングを構成し、コンテンツごとにネットワークを構成する方法も考えられる。

4.1.3 障害時対応

障害時対応について、リンクが切れたり極端に細くなった瞬間、一時リングネットワークの次の次のサーバにまず繋ぎ、その後、もう一度各 Relaycast サーバで計算させ、その値を

4.1 Relaycast 動作の為の機能

もとに経路を再計算させ、それを適用させる。

提案方式は、リングネットワークであるので、全てのリンクが正常な状態で送信できていなければ性能が出ない。もし輻輳や障害が起こった場合について提案・検討する。

トークンリングや FDDI における 2 重化

トークンリングや FDDI でのリングネットワークにおけるの 2 重化 [図 4.1] のように、リンクが切れたら検知し、切れた所から逆側へ流すようにする (ループバック)。しかし回線コストは 2 倍になる上、通常時には一方の回線は使用されないため効率が悪い。

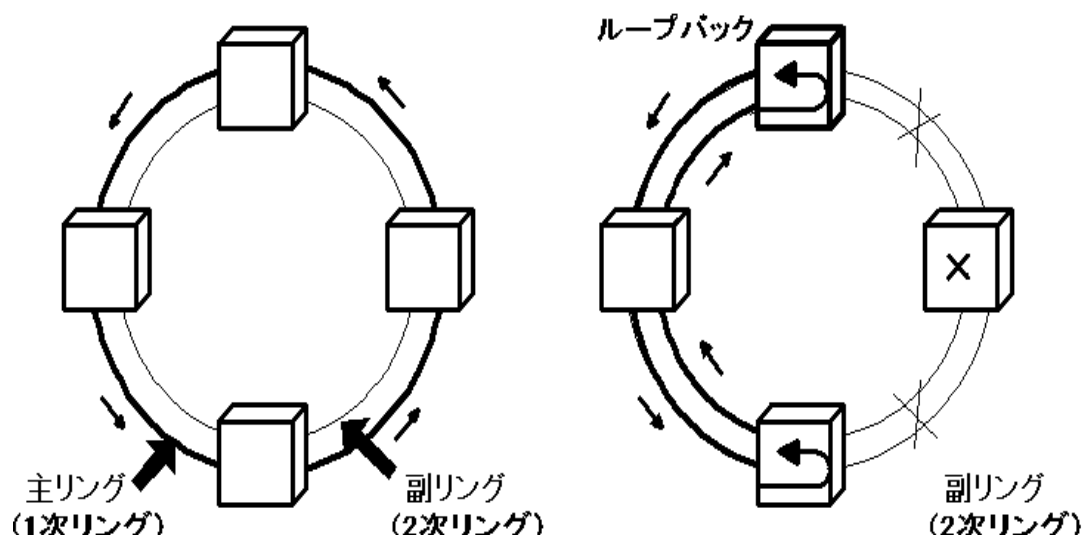


図 4.1 リングネットワークでの 2 重化

提案する障害対応策

そこで輻輳や切断が起きれば各サーバが検知した時点で、経路情報を配布する管理サーバへ知らせ、再配置するという方式が望ましい。ただ、この処理には時間が掛かる可能性がある。よって、リンクが切れたり輻輳した時点で、今リンクしている 2 つ先のサーバへ繋いだ上で、計算した後再配置をするのが望ましい。そのための経路情報は、管理サーバから各サーバへ、1 つ前と 2 つ先のサーバの IP アドレスを渡せばよい。障害の場合の動きを図 4.2 に示す。

4.2 ユーザーの為の機能

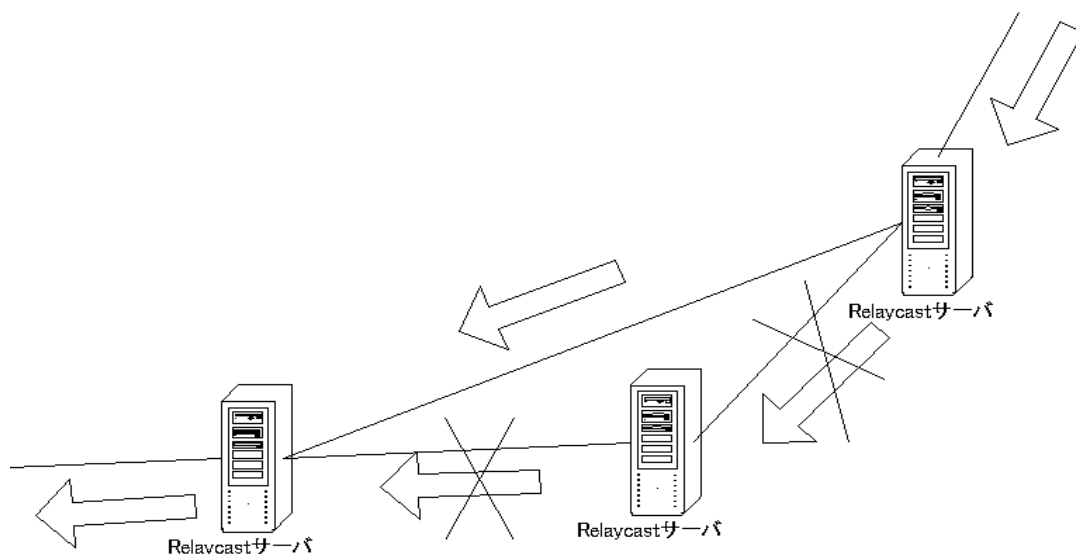


図 4.2 障害の場合

そのサーバは次のサーバとのリンク切れたのを検知した時点でそのサーバは、数 ms 秒後、ひとつ先へリンクするようにする。

障害を管理サーバへ知らせ、ルーティング再計算させた値を適用させる。

障害時には、あらかじめ設定しておいた次の次のサーバへまず経路を変更し、その後、経路計算をする機能を持つ管理サーバにて再計算した値に変更する。これによりスムーズな経路再構築が行われる。

また、障害からの復旧と途中からの参加については、経路管理の機能を持つ管理サーバに、参加問い合わせをすることにより、経路再計算を行い、その値を各サーバに渡す。

4.2 ユーザーの為の機能

リクエストルーティング

通常 CDN におけるリクエストルーティング機能とは、Redirector がユーザ端末からの配信要求を受け付け、ユーザとの近接度やサーバ負荷、サーバー処理能力などの重み付けを行い、どのサロゲートサーバからの配信が最も適当なものを決定し、そこへ要求を誘導するものである。これは、以下のような動作を行う [図 4.3]。

クライアントからコンテンツ要求をあげる。

4.2 ユーザーの為の機能

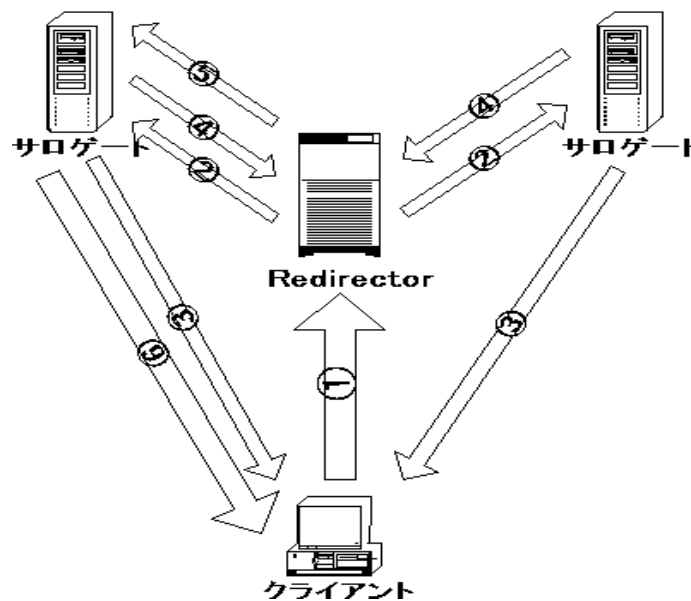


図 4.3 CDN のユーザー側動作

クライアントからの要求に対してコンテンツを持つサロゲートサーバに対して要求をあげる

サロゲートはクライアントに対してリンク容量を測る

ネットワーク回線スピードや混雑度、CPU 負荷、RTT による応答時間などのデータを Redirector に送信する

Redirector は 4 のデータをもとに、クライアントまでどのサロゲートサーバが配信するのが最適かを判断し、送信要求を出す。

サロゲートサーバはクライアントに配信する。

しかし、現在のリクエストルーティングの機能では、OSI 参照モデルにおけるトランスポート層以下のレベルでの判断だけであり、5 層以上を意識した選択はなされていない。そこでコンテンツ内容を考慮したリクエストルーティング手法を提案する。これにより、今までの MIB などによる情報に加えコンテンツ内容や容量等を管理サーバ上でデータベース化することによって、ユーザは映像コンテンツの途中からの視聴をダイレクトに行うことや、帯域の足りない場合、2 つの Relaycast サーバからコンテンツを流す、といったことが可能となる。

4.2 ユーザーの為の機能

提案する手法

通常 CDN における Redirector は、ユーザ端末からの配信要求を受け付け、最も適切なサーバへ要求を誘導するものである。

今回、提案する方式は、コンテンツを持つ各 Relaycast サーバにより収集される MIB による各サーバの帯域状況や Port の使用状況などを収集するだけでなく、コンテンツ情報もやりとりする方式である。各 Relaycast サーバは、コンテンツ情報をやりとりする記述 (Content Information Base(以下 CIB)) による情報によって、管理サーバのコンテンツ情報の要求に従って従来の MIB による情報だけでなく、コンテンツ情報も知らせ、より正確な情報をもとにユーザからの要求に対する適切なサーバ選択を行う。

Relaycast サーバからユーザ端末までの間で輻輳や遅延が起こった場合、複数 Relaycast サーバからのダウンロードが可能となる。またリンク切断時、ユーザはもう一度コンテンツ要求を管理サーバに出せば、ファイルの途中からダウンロード可能である。

Content Information Base(CIB)

映像ファイルの内容や種類について、管理サーバ上の Web サーバ機能で XML にて記述されたコンテンツ情報がユーザのために提供される。これにより、どんなコンテンツが現在あるかのカタログ機能を持つ。CIB には、管理サーバの要求に応じてファイル名とファイルサイズ、映像の Codec 名・ビットレート (VBR の場合)、音声の Codec 名・ビットレート (VBR の場合)・その Relaycast サーバにあるファイルをカタログしたデータベースファイル等を Relaycast サーバから返す機能等がある。

そしてマシンやネットワーク情報を知らせる MIB、コンテンツ内容の記述をする MPEG-7、そしてコンテンツファイルの情報を表す CIB を合わせて使うことにより最大限の効果を発揮する。

4.2 ユーザーの為の機能

ユーザの為の機能要求

対ユーザ側の機能要求

1. Redirector 機能としての機能
2. クライアントが Relaycast サーバの持つコンテンツとその種類を認識する .
3. クライアントはコンテンツの一覧を見ることができる .
4. 各サロゲートの持つコンテンツ・帯域情報・端末情報を知る .
5. クライアントからのファイルのレジューム対応の機能・分割ダウンロード機能
6. 帯域が足りない場合の 2 つの Relaycast サーバから取得できる機能

第 5 章

実験

5.1 実験系

図 5.1 にあるようなネットワークでの最小構成 3 台で、最初のコンテンツ配送の動作確認のための実証実験を行った。S₁ サーバは、入力された DV データを送信する機能を持ち、S₂ は、S₃ へコピーし、S₃ は受信し保存する。

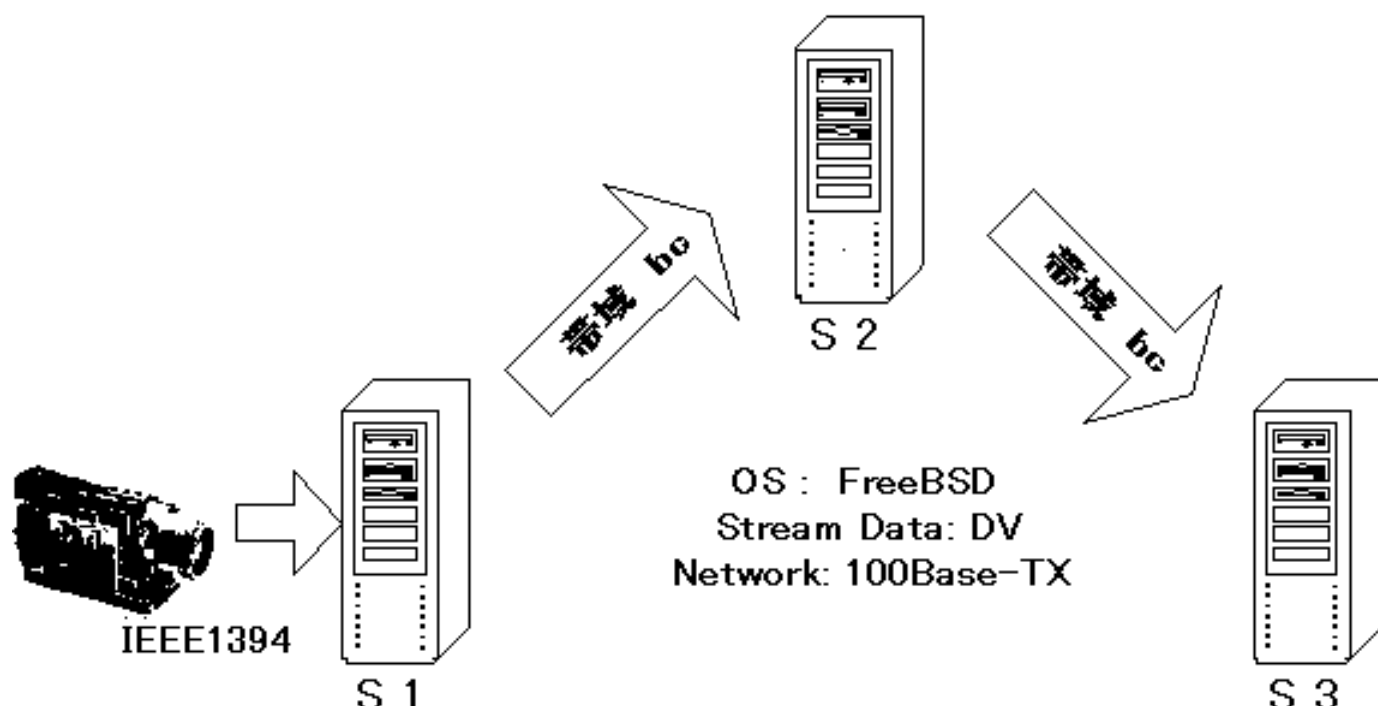


図 5.1 実験環境構成図

5.1 実験系

表 5.1 予備実験 01・結果

	最大帯域
$S_1 - S_2$	89.6M[bps]
$S_2 - S_3$	89.1M[bps]

5.1.1 実験内容

予備実験 01

このシステムにおいては、DV を利用した実験環境の場合、利用帯域が 60Mbps 必要である。十分な利用帯域があるか計るため最大帯域測定をした。

予備実験 02

他の端末から 60Mbps ネットワークカードに負荷をかけた場合の CPU 負荷を調べた。

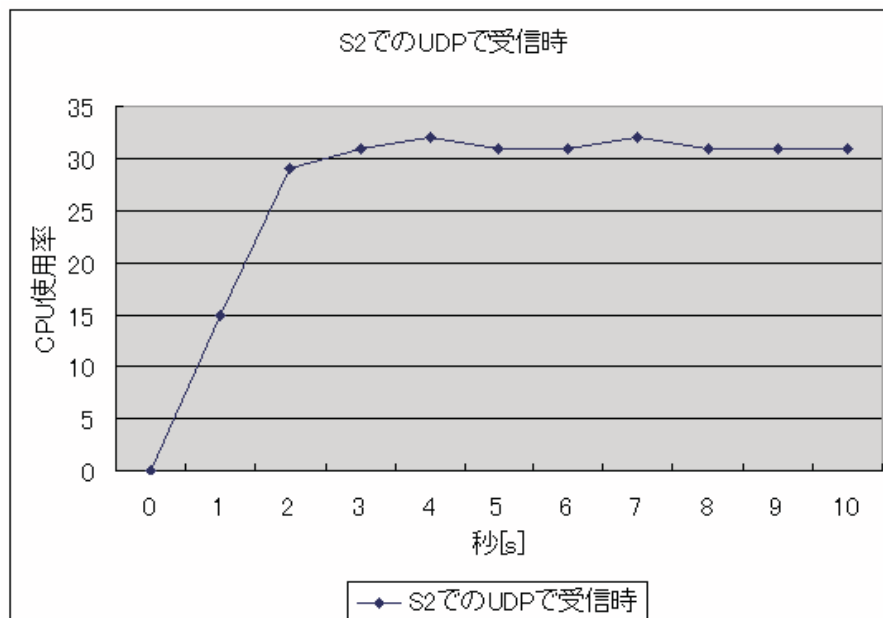


図 5.2 予備実験 02 結果

5.1 実験系

実験

DV から流れるストリームデータを $S_1 - S_2 - S_3$ と流し, S_2, S_3 それぞれ到達パケット数, パケットロス率を計り, S_2 における CPU 負荷を測った.

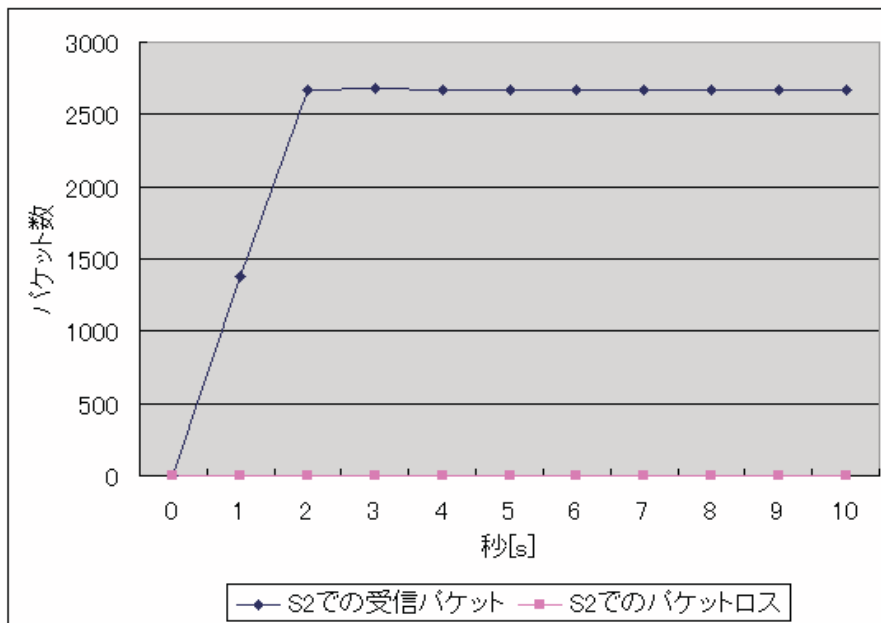


図 5.3 測定結果 : S2 パケット

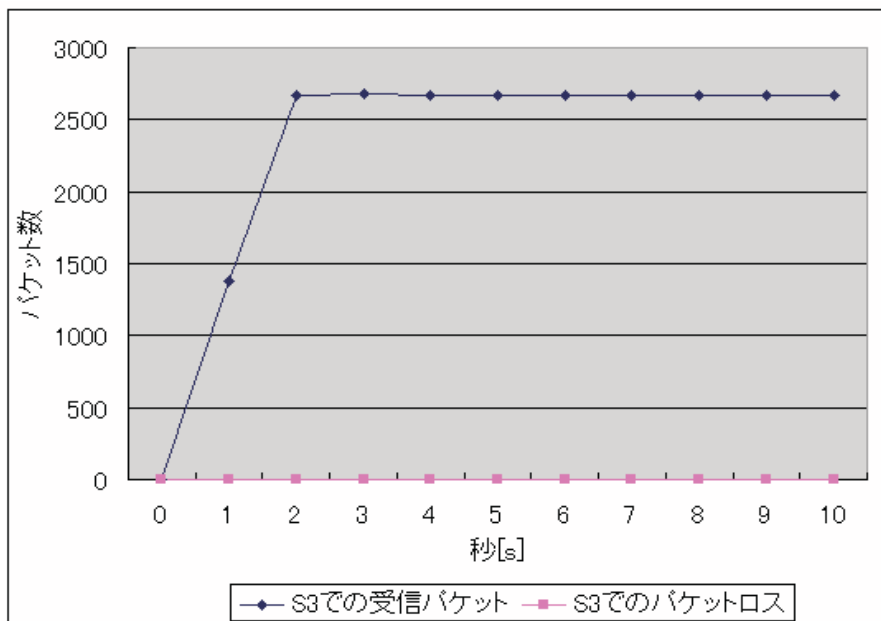


図 5.4 測定結果 : S3 パケット

5.2 課題・展開・発展

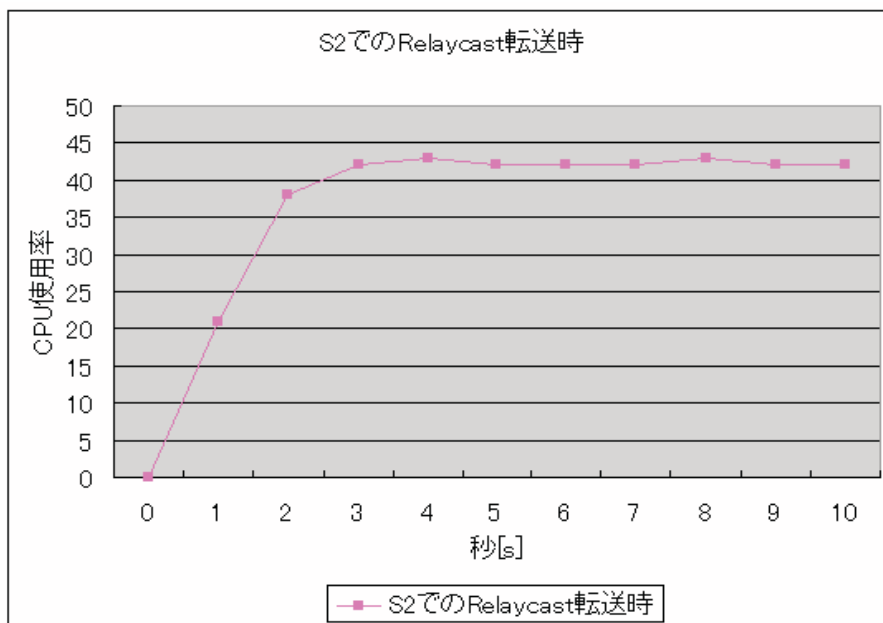


図 5.5 測定結果：CPU

5.1.2 結果

事前実験 01 の結果 , $S_1 - S_2$, $S_2 - S_3$ 間の UDP 送信の場合の帯域はそれぞれ 89.6Mbps , 89.1Mbps であり , このシステム動作には十分な帯域であった . 事前実験 02 の結果 , 60Mbps の負荷をネットワークカードに負荷をかけた S_2 の場合 31% の CPU 使用率であった . S_2 , S_3 における受信帯域は , はじめの 1 秒以外平均 2670 パケットであった . S_1, \dots, S_3 間は , S_2 , S_3 での測定でパケットロスなしで送信することができた .

5.2 課題・展開・発展

CPU 負荷の問題

S_2 における Relaycast 送信時の各サーバの CPU 負荷であるが , 処理負荷を小さくすることが複数ソースでリアルタイム送信する上で重要な点なので仕組みを改善していきたい .

5.2 課題・展開・発展

帯域保証と IP 通信網上の運用について

提案方式の実証では、Fast Ethernet で検討を行った。これは、現在広く普及している通信網は IP 通信が主であり、しかもその発展系である 10Gigabit Ethernet などは WAN まで見据えた設計になっているなどの理由からである。

そしてこのシステム上では各リンクは IP 網状で QoS 保証技術などを使うこともできる。それを実現する手法として RSVP や Diffserv などがあるが、広域に渡って一定の品質を保つには全て同じ機能・性能を持たなければならないなど、様々な問題がある。しかし、あるドメイン内、ISP 内に限って行う場合は、各サーバへのフロー制御などを行えば Diffserv が有効であると考えられる。また RSVP などの End - End 間での帯域保証を行う方式では、双方のサーバの所有者との合意が必要になってくるので、異なるドメイン間の接続部分に利用するのが適当かと思われる。このように RSVP と Diffserv の双方の利点を活用することで ATM に近い QoS 保証型通信が実現できる。

リアルタイム多チャンネル送信の実装

また、Gigabit/10Gigabit Ethernet や ATM 等を使い十分な帯域のある場合、周期 $\{ (N + 1 \text{ 周回}) \text{ データの遅延} + \text{サーバの処理時間} + \text{伝播遅延時間} \}$ に応じたデータを巡回させることによってリアルタイム多チャンネル配布ができる。

アクセス系の検討

今回、サロゲートからユーザまでの環境についてのアクセス系については、Redirector 機能以外はほとんど言及しなかったが、アクセス系には様々な FTTH の広帯域からから携帯等狭帯域の回線までがあり、デジタル映像コンテンツのビットレートの要求には格差がある。現在では、その格差のために異なる必要帯域別にコンテンツを作成し、CDN を利用して各サロゲートサーバに送信していることが多い。

5.2 課題・展開・発展

Relaycast では、バックボーンにおいて広帯域を利用する。そして様々なクライアント要求のために、デジタル映像のビットレートを変化させて配信することが有効である。デジタル映像のビットレートを変更させたり、例えば MPEG2 で圧縮されたものから MPEG4 へ圧縮したり、圧縮方式を変換することをトランスコーディングという。現在のリアルタイムトランスコーディングには、専用ハードウェアを使わなければならないが、将来、十分な CPU 速度が得ることができれば、Relaycast サーバ側でトランスコーディングすることも可能になる。

また、その中間点として、Relaycast サーバ間では、より高画質な映像配信をし、Relaycast サーバにファイルを受信した時点で必要帯域別でトランスコーディングを行い、グループ化して管理する。Relaycast サーバは必要帯域を検知し、ユーザは配信経路の帯域に応じたデジタル圧縮された映像コンテンツを配信することもできる。

コンテンツの内容記述

コンテンツ内容記述に MPEG-7 を利用することを考慮する [14]。これにより一般に時間検索が遅いと言われる avi ファイル等にも MPEG-7 を利用することによって、ユーザの要求に応じて、どのファイルのどの場所からでも即座に途中再生してストリームを受け取ることができるようになる。MPEG-7 は、長時間コンテンツにおいて、部位ごとに何が記録されているか記述することができるので、スポーツ番組のフラインプレーやニュースなど、情報収集の手段として特定のショートクリップで観賞することが多い物について非常に有用である。

第 6 章

まとめ

本検討では、CDN のソースサーバからの配信系に着目し、映像コンテンツをユニキャスト、マルチキャスト、そして提案する方式 Relaycast の、ネットワークの総トラフィック総量指数・総転送時間等について比較検討を行った。Relaycast とは、ストリームを巡回させることでリアルタイムに複数ストリームの配布を実行できる方式である。最初のコンテンツ配布を行った場合において実証した。

謝辞

本研究を行うに際し，多大なるご指導，ご鞭撻を頂いた，本情報システム工学科の島村 和典教授に深く感謝致します，また島村研究室 mAV サブグループのリーダー 中平 拓司氏をはじめとしまして，本講座の皆様，そしてご指導頂いた副査の坂本明雄 教授，菊地 豊助教授に心より感謝致します．

参考文献

- [1] 田代秀一，西角直樹，西木健哉，小島富彦，神原顕文，“インターネットコンテンツ配信技術の最新動向”，情報処理，Vol.42, No.11, pp.1082-1091, Nov. 2001.
- [2] kc claffy, G.Miller, K. Thompson. ”The nature of the beast: recent traffic measurements from an Internet backbone”, INET98.
- [3] Sean McCreary and kc claffy. ”Trends in Wide Area IP Traffic Patterns”. ITC Specialist Seminar on IP Traffic Modeling, Measurement and Management, September 2000.
- [4] T.Berners-Lee, R.Fielding, and H.Frystyk, ”Hypertext transfer protocol - HTTP/1.0” Internet Engineering Task Force, RFC1945 , May., 1996.
- [5] 2003 年 1 月 31 日 時 点 ”RealOne Player”, Real Inc., www.real.com/realone/index.html
- [6] P.V.Mockaperis, ”Domain names: Concepts and facilities. Internet Engineering Task Force, RFC882, Nov., 1983.
- [7] 竹下隆史, 村山公保, 荒井透, 苅田幸雄 共著,”マスタリング TCP/IP 入門編第 2 版”, オーム社, 1998.
- [8] マルチメディア通信研究会編, “標準 ATM 教科書”, アスキー, Mar., 1995.
- [9] Juha Heinanen, ”Multiprotocol Encapsulation over ATM Adaptation Layer 5”, IETF Network Working Group, RFC1483, Jul., 1993.
- [10] スティーブン サンドース 編集, Stephen Saunders 著, 井早 優子, 林田 朋之, 米沢 寿員 翻訳, “ギガビットインターネット徹底解説”, 日経 BP 社, 1999.
- [11] 総務省編, “情報通信に関する現状報告 (情報通信白書)”, 総務省, Jul., 2002.
- [12] 2003 年 1 月 31 日 時 点, 門林雄基, “CDN を用いた次世代インターネット技術の開発・普及の試み”, www.ieice.or.jp/cs/ia/jpn/conference/200111/Slides/kadobayashi.pdf

参考文献

- [13] 中平拓司, 山岡徹也, 島村和典, “CDN における映像コンテンツ巡回配布方式“Relaycast”の提案”, 信学技法, vol.102, No.156, pp13-16, Dec., 2002.
- [14] 菊地義浩, 堀修, “ストリームメディア通信サービス”, 情報処理, Vol.42, No.12, pp.1216-1220, Dec. 2001.

付録 A

MPLS トラヒックエンジニアリ ング

A.1 インターネットの発展と MPLS

A.1.1 はじめに

近年のネットワーク伝送技術の発展に伴い、利用者は通信品質の向上に目を向けるようになってきている。しかし現在のインターネットアーキテクチャは、接続性を重視し、最適かつ最小限の構成となっており、接続を重視した既存のアーキテクチャでは通信品質の向上といったことは想定されていない。その想定していない機能の導入のためにインターネットアーキテクチャ自体の見直しを計るといった流れがある。そのひとつとして、現在のインターネットワークのレイヤリングの見直しを始める動きがあり、今まで第3層によって提供されていたパケット転送機能と経路制御に代表されるネットワーク制御機能を分離して扱うことで、より柔軟なネットワーク制御を行う試みがある。この概念を実現する技術として、IP 経路制御によって決定される既存のフレームワークの制約を受けないマルチプロトコルレベルスイッチング (MPLS) 技術が、注目されている。

MPLS は、開発当初従来の overATM のルータの高速化を目指す技術であった。現在では、L3 スイッチと呼ばれる製品を見れば分かるように、ハードウェアによる IP パケットの高速なフォワーディング技術が実現され、MPLS がこの役割を果たすことは重要でなくなった。

A.1 インターネットの発展と MPLS

現在では、第3層プロトコルである IP と第2層の ATM などの諸技術との緊密な連携により従来の IP ネットワークでは不可能であった高度な通信品質制御 (QoS) やネットワーク構成に依存しないプライベートネットワーク (IP-VPN) の構築などの利用が実現されている。

本検討では、MPLS の主要な利用法のひとつであるトラフィックエンジニアリングについて述べる。

A.1.2 MPLS 技術と動作概要

MPLS では (図 A.1)、従来のルータの動作とは異なり、あて先 IP アドレスを見る代わりに、“ラベル” と呼ばれるパケットに付けられたフィールドのみを見てフォワーディングを行う。それと同時に、どの IP アドレスがどのラベルを使うかの情報を配布する。各ルータはルーティング・テーブルと共にラベル・テーブルを学習する。つまり MPLS では、従来のルータの動作に加え、新たにこのラベル情報を配布し学習する動作が加わる。ルータに入ってきたパケットには、まず最初にあて先 IP アドレスとラベル・テーブルを参照してラベルが付けられ、後はラベルのみを見てフォワーディングが繰り返され、まで到着したらラベルが外される。結果的に、MPLS によりフォワーディングされるラベル・パケットの道筋を、1本のパス(道)のように扱うことができる。

A.1.3 IP ネットワークにおけるトラフィックエンジニアリング

IP ネットワークにおいて、輻輳が起こった場合、回線を増やして負荷分散させるか、RIP や OSPF の経路のメトリックを変更させてトラフィックを別回線に迂回させるなどが考えられる。しかし、RIP だと Hop 数によるメトリックのため経路の輻輳には向かない。また OSPF でも、リンクのコスト値を変更してもどうしても偏りができてしまう。また、ATM やフレームリレーなら実現できるが、別に ATM やフレームリレーのネットワークを用意しなければならなかったり、オーバーヘッドが大きい等、問題も多い。

A.2 MPLS トラフィックエンジニアリング

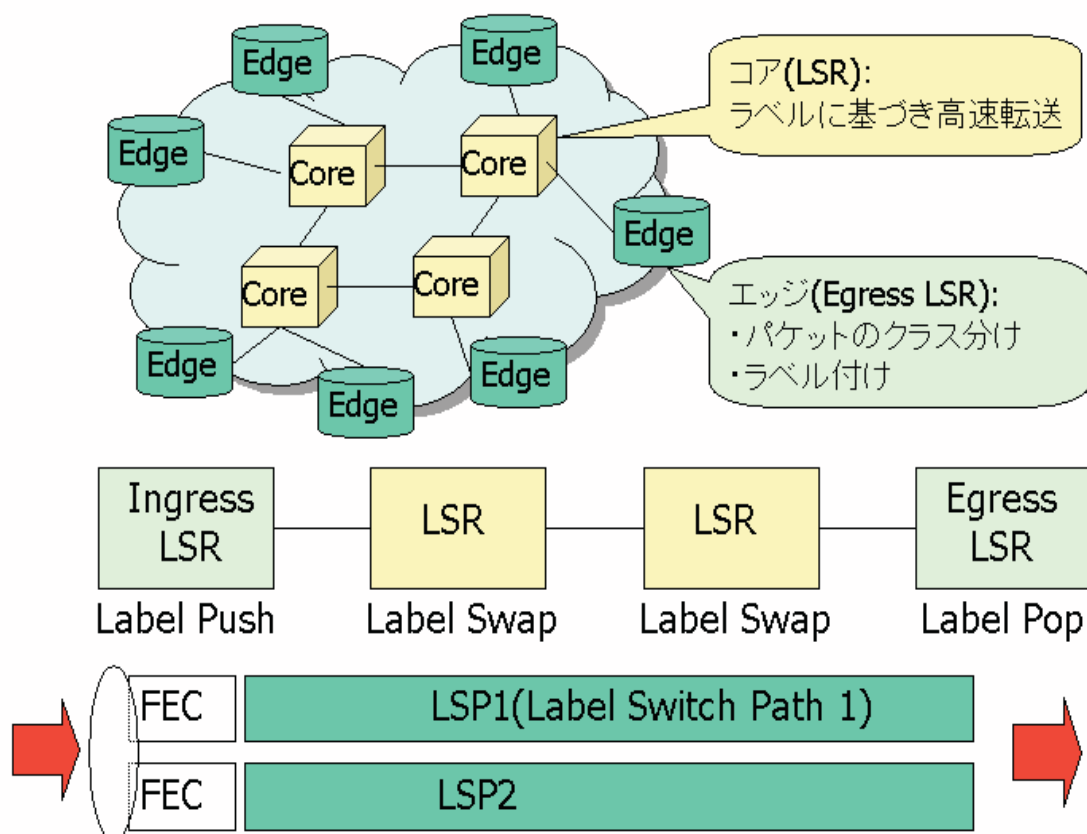


図 A.1 MPLS

従って、IP ネットワークでのトラフィックエンジニアリングは、ネットワーク資源を有効に活用できないことや、利用率の高い回線が障害になったときのリスクが大きいなど、コストパフォーマンス、信頼性、可用性に欠けたネットワークだと言える。

A.2 MPLS トラフィックエンジニアリング

近年、複雑なネットワーク構成を持つ IP ネットワークにおいて、柔軟な経路選択を行う手法のひとつとして、MPLS を使ったトラフィックエンジニアリングが注目されている。一部の機関や ISP では運用が始まっており、その技術の標準化と発展が期待されている。MPLS でトラフィックエンジニアリングを行う目的としては下記のようなものがある。

A.2 MPLS トラフィックエンジニアリング

- (1) 空き回線の有効利用のための明示的な経路指定によるルーティングを行う。
(Explicit ルーティング)
- (2) バックアップ経路をあらかじめ指定しておき、障害発生時に迅速に経路切り替えを行う。
(Fast ReRoute)
- (3) 帯域を確保した経路設定を行う。

A.2.1 MPLS トラフィックエンジニアリングの 3 つの流れ

前項で紹介した MPLS を利用する 3 つの目的についてそれぞれ説明をする。

(1) Explicit ルーティング

通常、RIP や OSPF などの IGP を用いた場合、あて先 IP アドレスへの経路はルータのホップ数、各リンクのコスト（帯域幅に基づいて計算）を足して、一番少ない経路が最短経路として選ばれる。ところが、選択された経路があるリンクに片寄ることで帯域幅の使用効率が悪くなり、最短経路を選んでしまえば、ホップ数やコスト値の最も少ない経路が利用されるため、経路の偏りによる帯域不足が発生することがある。このような場合、片寄りを回避する目的で明示的な経路を指定して迂回する機能が求められる。MPLS により、従来のルータでは難しかったパケットの道筋の操作を簡単に実現し、さらにリンクの使用効率も高めることができる（図 A.2）

(2) Fast ReRoute

Fast ReRoute とは、通常ダイナミックルーティングを使用し、ネットワークを冗長化している場合に、一方の回線に障害が発生すると、もう一つの回線に迂回する。このとき障害を検知し、迂回経路が決まるまでに IGP にて経路情報の再計算が行われる。大規模なネットワークだと計算が終わるまで数分間の時間がかかることがあり、その間はパケットは正常にルーティングされず、パケットロスに繋がることになる。そこで、この計算中に、あらかじめ設定しておいたバックアップ回線を使用することによって、正常にルーティングされない時間を極力短くしようとする技術のことである。この技術を使うと回線切断などの障害を

A.2 MPLS トラフィックエンジニアリング

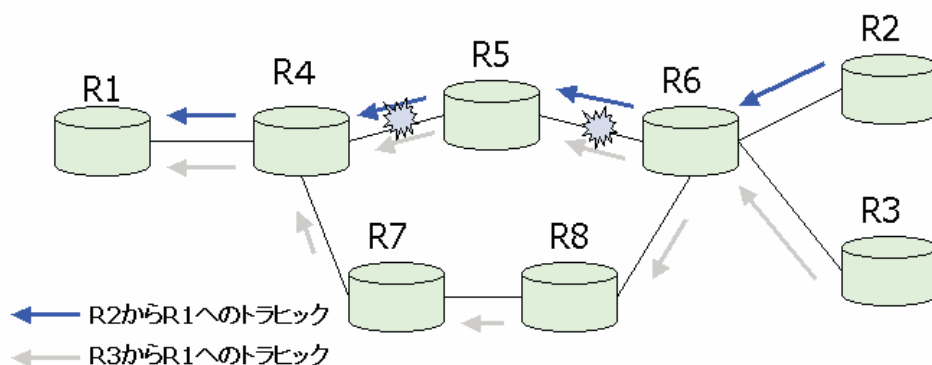


図 A.2 Explicit ルーティング

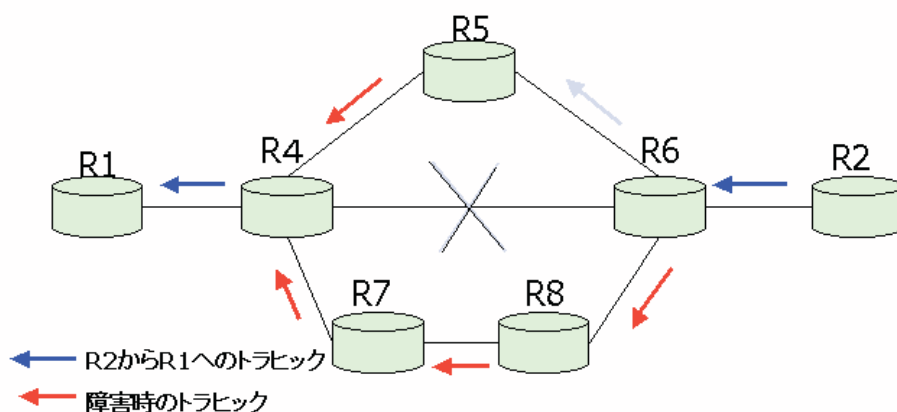


図 A.3 Fast ReRoute

検知すると、通常の IP ルーティングなら数秒～数分かかっていた時間がわずかに数 ms 以内に正常にルーティングされることになる (図 A.3)。

(3) Quality of Service (QoS)

QoS とは、ネットワーク内で提供されるサービスに対して品質を保証するための技術である。ある特定のトラフィックに対して必要な帯域確保を行ったり、優先度付けなどを行う。MPLS の QoS による品質保証の形態として Intserv, Diffserv という 2 つのモデルがある。前者は、データフローごとに特定の帯域を確保するモデルであり、後者は、フローをいくつかのクラスに分け、クラス毎に帯域の制御を行うモデルである。また、後者は、前者のような帯域制御だけでなく、トラフィックのシェイピングやポリシングや優先制御ができる。

A.3 まとめ・これから

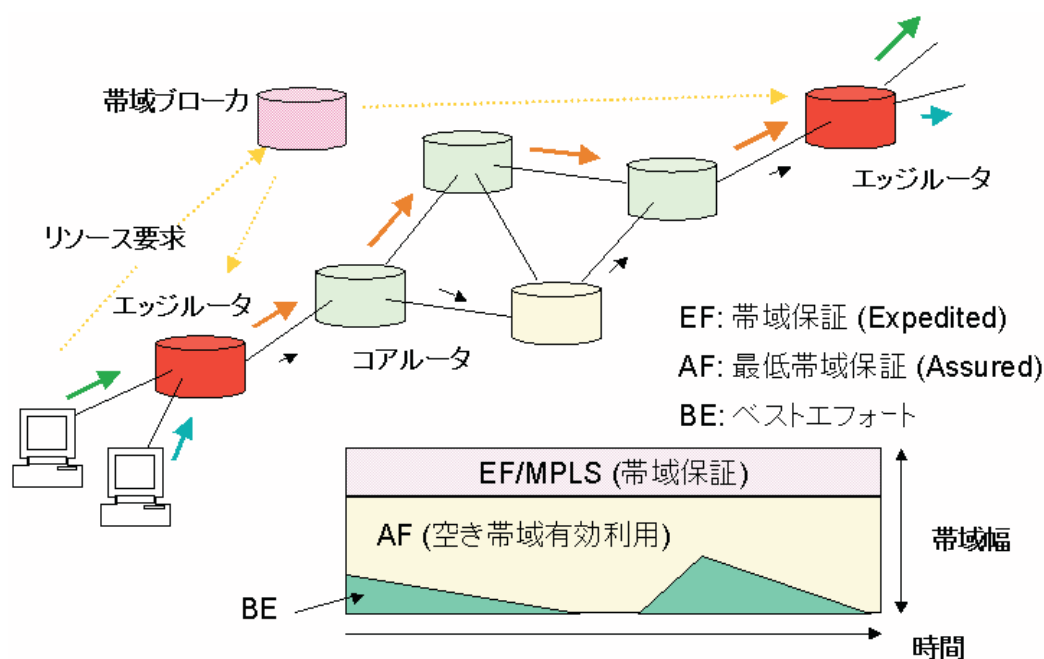


図 A.4 Diffserv

MPLS では、RSVP-Extention の QoS パラメータ交換が可能であり、これによりあるラベルパス用にある一定の帯域を確保しておくといったことが可能になる。MPLS-Diffserv によって、トラフィッククラス毎に帯域を予約したり、トラフィックフローが集約された特定フローに対してのみ、帯域を予約する (図 A.4)。MPLS Diffserv とは、Diffserv over MPLS のことで、Diffserv の MPLS ネットワーク上へのマッピング方法のことである。

A.3 まとめ・これから

MPLS テクニカルエンジニアリングとは、MPLS の技術を使い、トラフィックの流れを見、ネットワーク構成を最適化することである。そして、MPLS を利用したテクニカルエンジニアリングの主な流れとしてネットワークの資源を有効利用できる Explicit ルーティング、障害早期復旧のための技術の Fast ReRoute、ネットワークの品質を保証する技術の QoS(Diffserv) がある。そして、MPLS 自体、IPv4 環境がある限りにおいては発展していくと思われる。IPv6 になると不透明ではあるが、発展しつづけるだろう。ただ、IPv6 フロー

A.3 まとめ・これから

ラベルの使い方によっては、アイデアは残るが MPLS 自体は消えていくのかもしれない。

付録 B

IPv6 のフローラベルの使用法

B.1 はじめに

IPv6 フローラベルは、IPv6 の開発当初、ストリーム系トラヒック等の取り扱いを異にしたいパケットを識別するために設けられたフィールドである。現在、トラヒッククラスや DiffServ, RSVP, MPLS の技術の応用などが期待されている。

B.2 IPv6

近年の急速なインターネットの普及に伴い、IPv4 のアドレスは、2006 年にも枯渇すると言われている、IP アドレスの枯渇の問題の解決のため、現在の 32bit 43 億個から 128bit 3.4×10^{38} 個のアドレスを持つよう開発されたのが IPv6 である。アドレスの拡張、プラグアンドプレイによる簡単な設定 (アドレス自動生成) や効率的な階層状のアドレス体系、追加機能はオプションで行う構造、IPsec の標準搭載、QoS サポートなど多くの機能が盛り込まれた。今回、その中でも IPv6 ヘッダの 13~32bit の計 20bit に割り当てられているフローラベルはまだ、使用法がはっきり決まっておらず、様々な提案がされている。今回、IPv6 フローラベルについて調査・検討する。

B.3 IPv6 とフローラベル

フローラベルは IPv6 パケットのヘッダ情報の一部として実装され、トラヒックのフローを途中のノードが認識するためのものである IPv4 ではフローを直接認識する手段がなかつ

B.3 IPv6 とフローラベル

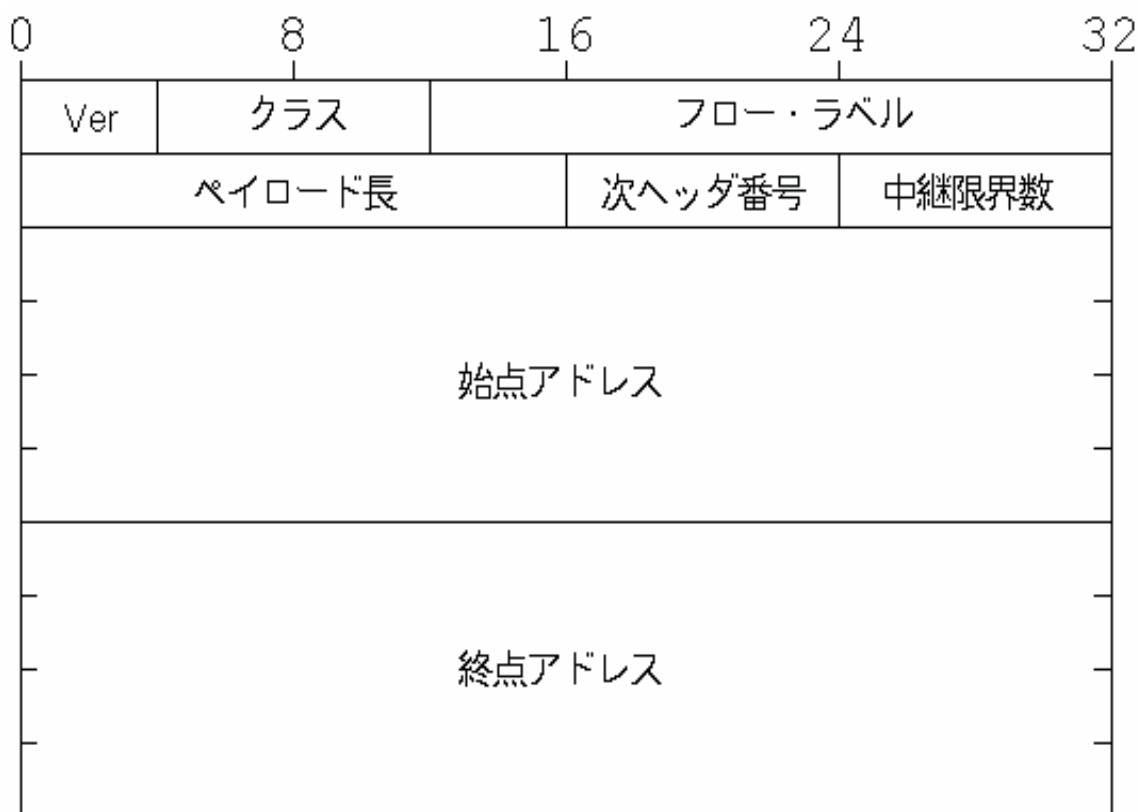


図 B.1 IPv6 ヘッダ

たため、送受信側それぞれの IP アドレスとポート番号の組み合わせや、Diffserv で用いる DS フィールドを代替的に利用していた。フローを直接に扱うことにより、より明確な QoS 処理が可能になる。

ただし IPv6 あるいはフローラベル自体は QoS を実現するものではない。実際の QoS の現場では、優先的に扱うべきパケットを先に送り出すようにキューイング処理を施す必要がある。フローラベルは、このトラヒックの優先度合いをフローごとに区別するための識別子として使われる。

IPv6 では送信元が中継ルータ（あるいはネットワーク）に対し、自分の送信する特定のトラヒックフローに対する特別な扱いをリクエストできるようになっている。IPv4 では、トラヒックフローを認識するために、送信元と送信先の IP アドレスに加え、ポート番号も確かめなければならない。この問題を解決し、フロー単位で RSVP や Diffserv に代表される

B.4 IPv6 フローラベルの利用法

ような資源予約型の QoS 技術などを活用できるようにするために用意されたのがフローラベルである。

B.4 IPv6 フローラベルの利用法

フローラベルの利用法で主に実装されているのは RSVP と Diffserv である。ここではその説明と今後の展開について述べる

B.4.1 RSVP

インターネットの資源予約プロトコルとして RSVP がある。RSVP での資源予約は、映像や音声などを行うマルチキャストアプリケーションに利用されることを前提としている。RSVP は受信者主導のプロトコルとして設計され、どれだけの帯域幅を予約できるか決めるのは全て受信者である。RSVP メッセージには、受信者がパケットを受けたい始点ホスト等が記述され、いくつかの品質保証パラメータを選択することができる。IPv4 の場合、RSVP メッセージに選択したデータを振るいにかけるためには、TCP や UDP のポート番号を指定する必要がある。IPv6 ではこの場合フローラベルを利用すると受信者は、ポート番号を指定する代わりに、優先したいフローのフローラベルをフィルタに記述する。フローラベルは、パケットの先頭にある IPv6 ヘッダ上に存在するので、簡単にいつでも参照できる。従って、フィルタによるパケットふるい分け処理を効率化でき、かつ実装を簡略化できる。しかし、RSVP 自体の複雑さなどから実際利用する例は少ない。

B.4.2 Diffserv

ネットワーク機器のパケット転送処理において、優先度や保証帯域などの差別化を行うことにより、QoS を保証する技術であり、RSVP の複雑さなどもあり、Diffserv も IPv4 における TOS フィールドと同じ使い方をする。

B.4 IPv6 フローラベルの利用法

B.4.3 MPLS

MPLS (Multiprotocol Label Switching) について述べる。従来の IP パケットの中継は受信先アドレスをもとに経路表を調べて中継先を決定するが、この処理に時間がかかる。MPLS はラベルからすぐに中継先を決定できるようにする技術である。ラベルは IP ヘッダにあるものを使用してもいいし、データリンク層のもの (たとえば ATM の VPI/VCI) を使用してもよい。MPLS のヘッダを IP ヘッダのラベルとして、フローラベルを利用しようという提案がある。

B.4.4 MPLS と IPv4 と IPv6 を統合する機能

この機能をもつルータは IPv4/IPv6 デュアルスタック対応のバックボーンへ移行させる最も効率的な手段となる。ユーザは、MPLS ネットワークのエッジ部分の Provider Edge Router (PE ルータ) において、IPv6 対応 PE ルータへの対応および IPv6/IPv4 のデュアルスタックサポート対応ルータへ移行を行うことにより、IPv6 over IPv4 トンネルなどの設定を行うことなく、既存の MPLS バックボーンネットワーク上で IPv6 パケットと IPv4 パケットを混在させたフォワーディングの同時運用が達成される。

IPv6 パケットの到達性は、v6 対応 PE ルータ間で情報交換を行うことにより実現し、IPv6 パケット自身は MPLS ラベルを使用してカプセル化し転送される。そのため、MPLS コアネットワーク内に位置する Provider Core Router では IPv6 パケットを認識する必要はない。

B.4.5 まとめ

IPv6 フローラベルは、フローすなわちストリーミングのために設けられた IPv6 ヘッダである。使用法はまだ検討中であるが様々な提案がされており、QoS のための RSVP や Diffserv、MPLS 等に用いられる予定である。MPLS のような特別なハードウェアを使わず IPv6 フローラベルを利用できるようになれば、IPv6 の大きなアドバンテージとなり、普及

B.4 IPv6 フローラベルの利用法

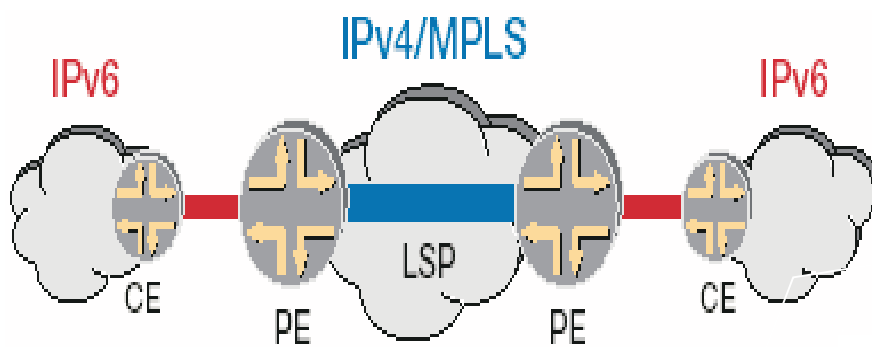


図 B.2 MPLS over IPv6

するのではないかとと思われる。