

4

Bandwidth Broker System の開発と運用実験

森島直人* 小川晃通** 穎原桂二郎*** 染川隆司****

技術の発展によってインターネットは着実に安定性を増し、人々の興味は到達性から通信の品質へと移ってきた。現在、この要求を満たすためにDiffservが標準化されている。しかし、標準化は優先配送制御を実現するための機構のみにとどまり、インターネット全体での運用モデルについては触れられていない。そのため、必要なアドミッション制御の機構を決定することができず、現在のところ人間の介在する静的な契約のみを想定したもののみが考えられている。我々は、インターネットにおけるDiffservの階層化された運用モデルを提案している。そのモデルでは、ネットワークの末端部において人間の介在しない動的な契約システムが必要となる。そこで我々は、動的な帯域割り当てシステムの実装および合宿における運用実験を行った。本稿ではその報告および結果と評価について述べる。

通信品質への要求

近年のインターネットの発展は、日々の生活の情報化を加速している。人々は電話やFAXに加え、電子メールとWWWという情報交換のための手段を手に入れ、それに対する依存度は日に日に高まっている。これに従い、展示会や国際会議といった各種イベントにおいてネットワークを必要とすることが多くなった。各種相互実験を行ったり、あるいは単なるライフラインとしてインターネットが必須となりつつあるからである。また、DV over IP¹⁾ やWMT (Windows Media Technology) といったストリーミング技術の発展により、これらのイベントをTVに代わりインターネットにおいて中継・配信することが多くなった。

このようなストリーミングコンテンツを高い品質で配信し続けるためには安定して広帯域を確保し続ける必要がある。しかし、現在のインターネットは最善努力 (Best Effort) の枠組みで動作しており、ストリーミングの帯域を確保するためにストリーミング専用の回線を用意しているというのが現状である。インターネットの黎明期には、ユーザの最大の関心事は到達性であった。当時はさまざまな障害により到達性の失われるこ

とがしばしば発生したためである。しかし、動的経路制御をはじめとするさまざまな技術の発展によりインターネットの安定性が増した今日、人々の関心事は上記のような要求から通信の品質や帯域の確保へと移行してきた。このような現状を受け、RSVPやDiffservをはじめとするさまざまな品質保証のための技術が提案されている。しかし、アドミッション制御やサービスモデルの欠如により、実運用には至っていない。

本稿では品質保証の仕組みとしてDiffservを取り上げる。さらに我々は本連載において取り上げているWIDE合宿で動的帯域予約の運用実験を行った。本稿ではその報告および結果と評価について述べる。

Diffservの概要

Diffserv (Differentiated Service)²⁾ は、インターネットにおける通信品質を保証するための仕組みとして標準化されている。Diffservでは、ISP (Internet Service Provider) が提供する限定されたサービスをDSCP (Diffserv Codepoint)³⁾ と呼ばれる識別子によって分類し、優先制御によって通信品質を統計的に保証する。サービスを受ける顧客はISPとのサービスレベルの契約 (SLA; Service Level Agreement) を結ぶ。これらのサービスはスループットや遅延、遅延の揺らぎなどの統計的な指標によって表される。

同一のポリシーによって運営されるDiffservネットワ

* 奈良先端科学技術大学院大学 情報科学研究科
moley@kyoto.wide.ad.jp

** 慶應義塾大学 政策・メディア研究科 akimichi@sfc.wide.ad.jp

*** 慶應義塾大学 政策・メディア研究科 popo@sfc.wide.ad.jp

**** 奈良先端科学技術大学院大学 情報科学研究科
somegawa@nara.wide.ad.jp

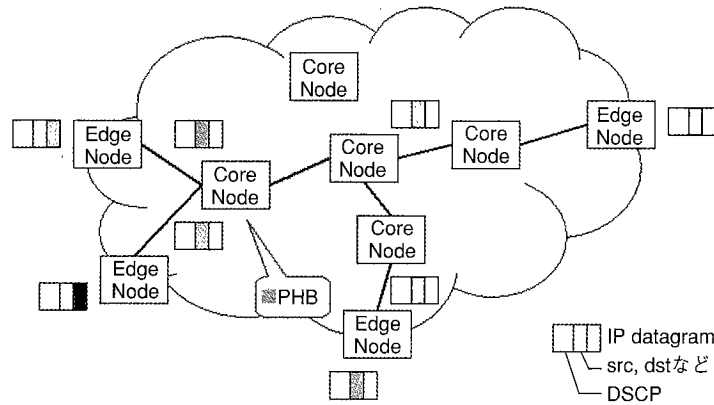


図-1 Diffservの概要

ークをDSドメインと呼ぶ(図-1)。DSドメインの境界には入口/出口境界ノードと呼ばれる機器が設置され、DSドメインに流入するフローはすべて入口境界ノードを通過する。入口境界ノードではSLAに従って分類され、現在のフローの状態を勘案してIPヘッダにDSCPを付加する。DSドメインの内部では、同一のDSCPを持つすべてのフローは集約され、ホップごとにDSCPに関連づけられたキュー制御などの処理をされる。この集約をBA (Behavior Aggregate)、ホップごとの処理をPHB (Per-Hop Behavior)と呼ぶ。

Diffservの枠組みでは、PHBに利用するキュー制御アルゴリズムやアドミッション制御(ポリシーの適用や資源配分機構等)の方式を限定していない。このため、サービス提供者であるISPは、さまざまなキュー制御アルゴリズムとアドミッション制御を組み合わせることにより特色あるサービスクラスを定義し、品質保証サービスを構築・提供することができる。また、複数のフローをDSCPごとに集約することにより、より多くのフローが集中するDSドメイン内部のスケラビリティを実現している。

このように、DiffservではDSドメイン内におけるフローの扱いの枠組みのみを標準化したものであり、運用モデルやアドミッション制御機構と独立している。現在、アドミッション制御機構は静的なものが考えられているが、我々は個々のユーザが動的に資源要求を行うモデルを考えている。しかし、動的な資源要求に関する運用実験はほとんど例をみない。そこで我々は、次章で述べるWIDE合宿の合宿ネットワークで動的帯域予約の運用実験を行った。

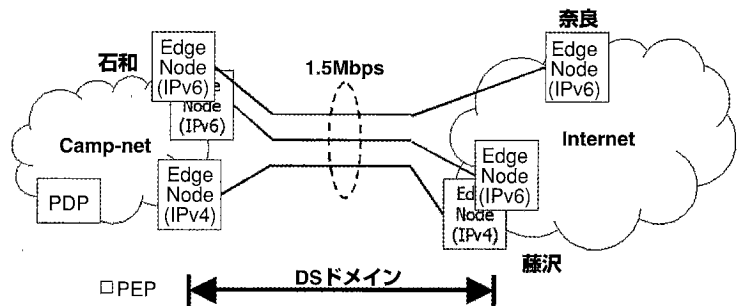


図-2 WIDE Camp-netのトポロジ

WIDE 合宿での運用実験

本連載の第1回目で説明したように、WIDE Project⁴⁾はインターネットとその関連技術に関する研究と開発を行う非営利団体であり、春と秋に4日間の合宿を行っている。合宿では一時的な運用ネットワークを構築し、インターネットとの接続性を確保するとともにさまざまな実験を行う。このネットワークは合宿ネットワークと呼ばれる。2000年春の合宿は山梨県石和町で行われ、参加者は236人であった。この合宿において、我々はユーザが動的に帯域を要求できるDiffservネットワークの運用実験を行った。

■概要

今回の合宿では、インターネットへの接続にATM over T1を利用した。このATM回線上に3本のPVCを設定し、2本をWIDE藤沢NOCに、1本をWIDE奈良NOCに接続した(図-2)。本実験ではこの対外線を1つのDSドメインとして扱い、ユーザからの動的な要求によるトラフィックの優先制御はこの対外線の両端で行った。Diffserv

でのDSCPの付加はDSドメインの境界点である合宿ネットワークの出口ルータで行う必要がある。このため、インターネットに接続するためのルータが複数設置された合宿ネットワークでは、通常の方法ではそれぞれのルータを入口境界ノードとして設定・運用する必要がある。このような設定・運用上の負荷を軽減するため、ATM PVCブリッジングの技術を利用することによって、1つのルータにATMレベルですべてのトラフィックを集中させ、DSCPの付加やトラフィック制御を行う入口境界ノードの数を減少させた(図-3)。また、インターネットから合宿ネットワークへのトラフィックも同時にサービスの対象とした。これに関しても、ATMレベルですべてのPVCをWIDE藤沢NOCに集中させ、PVCブリッジングによる単一の入口境界ノードでDSCPの付加やトラフィック制御を行った後に合宿ネットワークへ接続するという手法を用いた。

入口境界ノード内では、ALTQ⁵⁾を利用してトラフィック制御を行った。合宿ネットワークまたはインターネットからDSドメイン(対外線)へのトラフィックには、石和または藤沢に設置した入口境界ノードの入力インタフェースのキューでマーク付加を行った。また、出カインタフェースのキューでは、入力側で付加されたDSCPに従い、HFSC⁶⁾とRIO⁷⁾を利用したスケジューリングとキュー制御を行った。

ルータへの品質制御パラメータの供給にはCOPS(Common Open Policy Service)⁸⁾、^{☆1}を利用した。品質制御パラメータの管理、アドミッション制御および課金管理(後述)を行うPDP(Policy Decision Point)は合宿ネットワーク内で運用した。また、PDPの決定に従って品質制御パラメータを受け取り、実際に品質制御を行うPEP(Policy Enhancement Point)は対外線の両端に設置した入口境界ノードで運用した。

■ サービスクラス

帯域予約サービスは、合宿ネットワーク内のすべてのユーザに対して提供された。このサービスは、フローを始点と終点の組(src, dst)で分類・集約し、その集約(以下フロー集約と呼ぶ)をAF(Assured Forwarding)⁹⁾方式で優先制御するものである。前節で述べたように、対外線はT1(1.5Mbps)であった。

^{☆1} インターネットにおいて通信品質保証サービスを提供するためには、管理ポリシーに従ってそのネットワークを構成する各機器ごとに通信品質パラメータを決定する必要がある。さらに、決定したパラメータを各機器に配布するとともに、そのパラメータに従って機器の設定を行わなければならない。COPSはIETFによって1999年に標準化が行われた基盤プロトコルである。

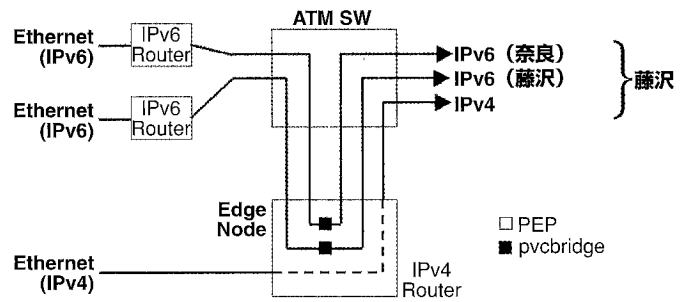


図-3 石和における第2層の設定

まず、予約のためにこの帯域を64kbpsの帯域を持つ19個のブロックに分割した(図-4)。これらのうち、18ブロックがユーザからの予約に利用され、残りの1ブロックは常に優先されるべきネットワークの制御情報(COPSの制御メッセージや経路制御情報など)に利用された。予約した合宿参加者には1ブロックが割り当てられる。したがって、参加者は1つの予約につき64kbpsの帯域を確保することができる。同一のフロー集約に対しては同時に1つの予約しかできない。つまり、参加者は帯域の足し込みを行うことはできない。また、参加者に割り当てた帯域は18ブロックであるため、同時に予約可能な最大値も18であった。

一方、輻輳が発生していないときには、使われていないブロックは他のトラフィックで共有される。合宿参加者がブロックを予約すると、TRTCM(Two Rate Three Color Marker)¹⁰⁾によってDSCPが付加される。参加者のフロー集約が64kbps以内の場合は青色にマークされ、優先的に処理されるようになる。また、64kbps以上96kbps未満のものについては黄色にマークされる。さらにそれ以上については赤色にマークされ、予約を行っていないその他のトラフィックと同様に扱われる。

■ 課金

特定のユーザによる帯域の占有を防止するため、長期にわたる帯域予約を抑制する仕組みが必要である。本実験では、合宿参加者全員にアカウントを発行し、一定量の仮想通貨を配布した。この仮想通貨は帯域の予約時間に応じて減少し、仮想通貨がなくなると予約ができなくなる。仮想通貨の単位はWU(WIDE Unit)とし、合宿の参加者には合宿開始時に一律2000WUを配布した。また、1ブロックの予約を1分間行うためには10WU必要であるとした。さらに、現在の仮想通貨の残量や自分の予約状況を確認したいというユーザ要求が予想される。本運用実験では、Webを通して現在の帯域予約状

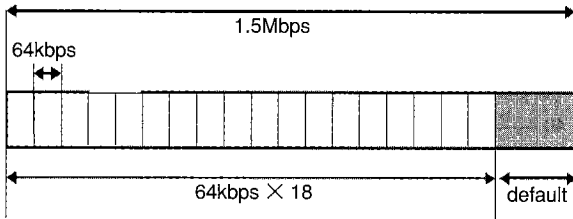


図-4 対外線の分割

開始時刻	終了時刻
3月15日 18:14:50	3月15日 18:19:05
3月16日 12:33:37	3月16日 13:39:28
3月16日 13:42:54	3月16日 13:49:03
3月16日 16:19:47	3月16日 19:39:47
3月16日 20:55:27	3月16日 23:49:24

表-1 人為的輻輳の発生時間

況および仮想通貨の残量を知ることができるようにした。

■予約クライアント

現在のところ、ユーザが網に直接要求するようなプロトコルでCOPSと強調して動作することが想定されているものはない。そこで、この実験のためにユーザが網に予約要求を行うための一時的なプロトコルを設計した。このプロトコルにより、ユーザはDiffservネットワークに対し、最寄りのPEPを通して帯域を要求することができる。また、ユーザの認証情報や予約フローに関する詳細情報は予約メッセージ内に含まれる。さらに、ユーザにはこのプロトコルを実装した帯域予約クライアント「PEPe」を提供した。このクライアントはIPv4およびIPv6の両方に対応した。Diffservネットワークでの予約プロセスが終了すると、PEPeはPEPから予約要求に対する結果を受け取る。これには予約の成功・失敗および失敗した場合にはその原因が含まれている。

結果と考察

前述したように、WIDE合宿の参加者数は236人であった。その一方、DSドメインはT1(1.5Mbps)の対外線であり、参加ユーザ数と比較して広帯域回線であった。このため、各ユーザは十分な帯域を利用することが可能であり、帯域を予約したユーザが優先制御の利益を得ることができない。そこで、本運用実験では合宿参加者

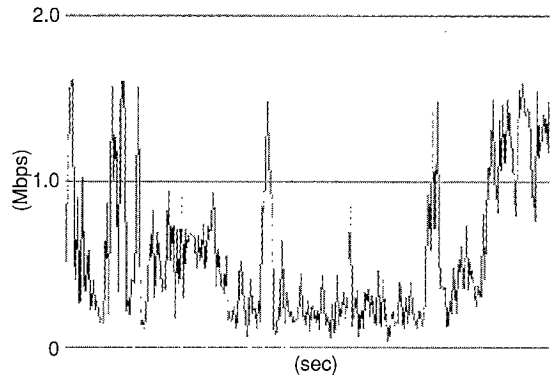


図-5 通常時のトラフィック

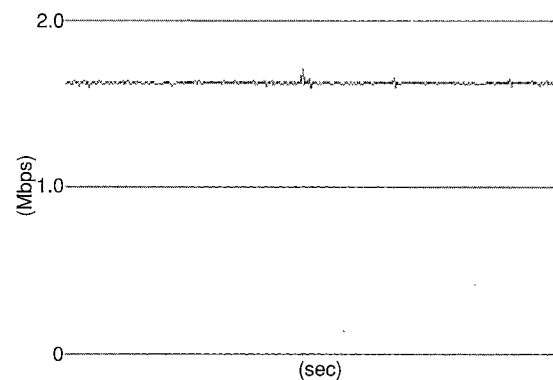


図-6 輻輳時のトラフィック

の帯域予約サービスの利用を促進するために、対外線上に突発的で人為的な輻輳を発生させた。輻輳は高帯域を消費するUDPトラフィックをインターネット側から合宿ネットワークに向けて送信することによって発生させた。合宿参加者の利用するサービスはWWW・POP・FTPといったインターネットから合宿ネットワークへのトラフィックが発生するものが多い。そのため、前述のようなUDPトラフィックを発生させることによって輻輳を参加者により一層強いものとして感じさせることができる。UDPトラフィックの送信開始時刻と終了時刻を表-1に示す。

次に、インターネットから合宿ネットワークへのトラフィックを図-5および図-6に示す。これらの図はそれぞれ輻輳が発生していないときおよびUDPトラフィックによって輻輳を発生させたときのトラフィックを示している。輻輳中は対外線上でパケットの喪失が発生した。

次に、合宿参加者からの予約要求の数を図-7に示す。UDPトラフィックによって輻輳が発生しているときに予約要求が集中している様子が分かる。

また、予約エラーの数を図-8に示す。これは何らかの

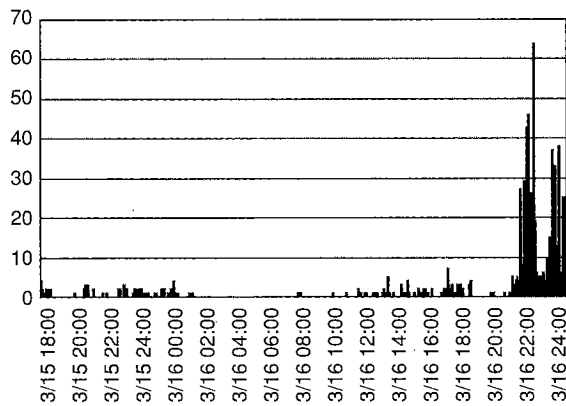


図-7 予約要求数

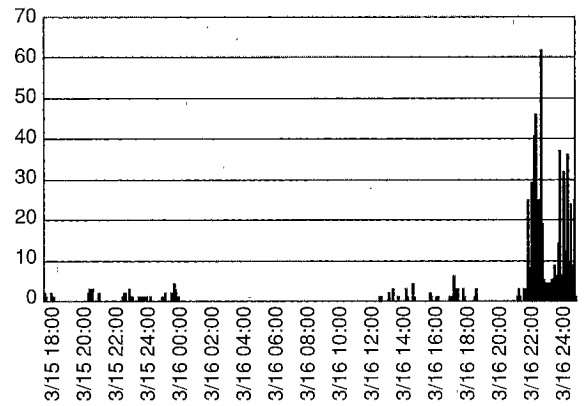


図-8 予約エラーの数

理由によりPDPで拒否された予約要求の数を表している。表-1や図-7と比較すると、予約エラーの多くは対外線の輻輳時に発生している。前述のように、本運用実験ではDSドメインの帯域を分割し、そのうち合宿参加者の帯域予約用に18ブロックを割り当てた。図の予約エラーのほとんどは、上記の帯域予約用ブロックの枯渇により発生したものである。

これらから、合宿参加者からの予約要求は輻輳時に集中しており、逆に輻輳が発生していないときには予約がほとんどなかったことが分かる。また、擬似的な輻輳が回を重ねるごとに予約要求の数が増えていることが観測される。この現象は、合宿参加者が輻輳時の対外線の通信品質に価値を見出していること、時間の経過とともに動的な帯域予約システムの有用性を理解するようになることを示していると考えられる。インターネット上での突発的な輻輳を予測することは困難である。特にイベントなどでは、ユーザの挙動によってストリーミングコンテンツの品質が劣化するという状況が起こりやすい。また、イベントでは少数のユーザが狭帯域を共有することが多い。このため、ユーザ自身も他のユーザの挙動に大きく影響され、通信品質の劣化を感じる可能性も高いと考えられる。そのため、今回の運用実験に見られるようなユーザの挙動を考えれば、個々のユーザに動的な帯域予約システムを提供することは有用であると考えられる。

おわりに

今回は、WIDE合宿におけるDiffservネットワークの運用実験について説明した。VoIPなどの新しいインターネットアプリケーションの登場とともに、通信品質へ

の要求が高まっている。このことは、輻輳時にユーザが集中して帯域予約要求を行うという、本実験の結果からも見て取ることができる。特に、狭帯域を少数のユーザで共有するイベントの場合、通信品質への要求はさらに顕著なものである。Diffservは統計的に通信品質を保証するための枠組みであるが、その運用モデルの曖昧さから実用段階まではほど遠い。特に、動的な資源予約に関してはその有用性やスケーラビリティに疑問の声も多い。しかし、今回の運用実験はDiffservの実運用に向けた大きな一歩であるといえよう。

今後は、動的資源予約についてさらに実験を重ねるとともにDiffserv運用モデルに関する議論を深め、DSドメイン内でのトラフィックエンジニアリングやDSドメイン間での帯域取引の可能性などについて検討および実験を行う予定である。

次回は、甲子園でお馴染みのWeb Castingや次世代ストリーミング技術であるDV over IPといった、イベントなどの中継システムについて説明する。

参考文献

- 1) Ogawa, A., Kobayashi, K., Sugiura, K., Nakamura, O. and Murai, J.: Design and Implementation of DV Based Video over RTP, Packet Video Workshop 2000 (2000).
- 2) Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and Weiss, W.: An Architecture for Differentiated Service, RFC 2475 (Dec. 1998).
- 3) Nichols, K., Blake, S., Baker, F. and Black, D.: Definition of the Differentiated Service Field (DS Field) in the IPv4 and IPv6 Headers, RFC 2474 (Dec. 1998).
- 4) WIDE Project: <http://www.wide.ad.jp/>
- 5) Cho, K.: A Framework for Alternate Queueing: Towards Traffic Management by PC-UNIX Based Routers, USENIX (1999).
- 6) Stoica, I., Zhang, H. and Eugene Ng, T.S.: A Hierarchical Fair Service Curve Algorithm for Link-Sharing, SIGCOMM '97, Real-Time and Priority Service (1997).
- 7) Clark, D.D. and Fang, W.: Explicit Allocation of Best Effort Packet Delivery Service, the IEEE ACM Transactions on Networking, Vol.6, No.4, pp.362-373 (Aug. 1998).
- 8) Boyle, J., Cohen, R., Durham, D., Herzog, S., Rajan, R. and Sastry, A.: The COPS (Common Open Policy Service) Protocol, RFC 2748 (Jan. 2000).
- 9) Heinanen, J., Baker, F., Weiss, W. and Wroclawski, J.: Assured Forwarding PHB Group, RFC 2597 (June 1999).
- 10) Heinanen, J. and Guerin, R.: A Two Rate Three Color Marker, RFC 2698 (Sep. 1999).

(平成12年8月28日受付)

