1 MPLS/L3TE/Diffserv/ネットワーク制御

1.1 実験の背景

インターネットはもともとは構成要素が独立して緩やかな協調分散をおこなう系として設計・運用されてきた。しかしその成功によるユーザ数の増加や多種多様なアプリケーションの登場は、インターネット自体の変化を要求する圧力ともなっている。特に、リアルタイム性や帯域制御といった設計時に考慮されていないネットワーク特性を求めるネットワークアプリケーションや、DDoS といったネットワークの一部だけでの処理では対応しきれない攻撃の登場によって、広域ネットワークを線または面で捉えた『包括的な協調分散制御』の必要性が考察されつつある。

ネットワークを包括的に制御するためのネットワークモデルとして転送層と制御層の分離モデルが提案されている。転送層はパケット転送を実際におこなう部分で Diffserv(Differentiated Services) [1] や MPLS(MultiProtocol Label Switching) [2]、L3TE(Layer 3 Traffic Engineering) などがこれに含まれる。制御層は管理ドメイン内でのポリシに応じて転送層の制御をおこなう。

1.2 実験の目的

このような背景のもとで、今回の合宿では包括的な広域ネットワーク制御の検討を目的とした実証実験として、ユーザからの動的な要求に対するネットワーク資源予約システムを構築した。このような細かな粒度でのネットワーク制御は、既存の第 3 層での (主に経路制御技術によっておこなわれる) 制御では扱いづらく、さまざまな技術的チャレンジが必要である。実験ネットワークの管理ドメインは、合宿ネットワーク内部だけでなく対外接続点である SFC に設置されたルータを含んでおり、広域ネットワークにおける制御に関しても考察している。

本実験は、WIDE 内でアクティビティを持っている『あやめプロジェクト』 [3] と『くまプロジェクト』[4] による共同プロジェクトであるくまあやめプロジェクトによっておこなわれた。あやめプロジェクトは、第 3 層以下技術に着目したネットワーク制御に関する議論を背景に、MPLS の研究と開発を行っている。また、研究開発用 MPLS 環境として、MPLS ルータ (LSR) および MPLS シグナリング機構である LDP(Label Distribution Protocol) [5]/CR-LDP(ConstraintRouting LDP) の設計および実装を。行っている。くまプロジェクトは、多様なポリシを多様な方法で実現するためのアーキテクチャ、コンポーネント間のインタフェイス、ポリシ履行状況確認のための計測、ネットワークプロビジョニングのための計測など、広域ネットワーク制御を実現するためのフレームワークの研究を行っている。また、制御層として COPS(Common Open Policy Service) [6] を、転送層として L3TE を実装し、実証実験を行っている。

1.3 実験の概要

本システムは、さまざまな技術が相互に関連して動作している。以下に利用している技術とその説明を示す。

制御層

- COPS/資源分配/課金/認証

  ユーザからの動的な予約受付および、ネットワーク資源の分配、課金、認証に従ったアドミッション制御を行う。課金には『WIDEUnit』と呼ばれる仮想通貨を導入し、予約の制御および平等性を実現する。さらに、アドミッション制御の結果にしたがってネットワークの制御パラメータを決定する。このパラメータは COPS によって制御の対象となる機器に送信される。

  今回のネットワークは転送層に L3TE と MPLS といった異なる技術で実現される複数の転送層を利用している。このため単一の制御層から特性に合わせて複数の転送層を制御する必要がある。

- NAT-friendly End-to-End communication

  NAT を利用したネットワークでは、アドレス変換の前後で同一のフローに対して観測されるフロースペッ

クが異なり、End-to-End 性が阻害される。このため、Diffserv や MPLS などのフローの識別を本質とするサービスの提供が困難となる場合がある。そこで、NAT の変換表を動的に取得し、アドレス変換の前後のフロースペックを対応づけることによって、単一のアドレスに変換された複数のフローを識別する。

転送層

- MPLS/CR-LDP

  合宿ネットワークからインターネットまでLSP(Label Switching Path) を確立し、要求に応じた通信品質を保証する。MPLS のシグナリング機構として代表的なプロトコルは LDP であるが、LDP は IP 層の経路を MPLS 層に写像するためにもちいられるため、トラフィックエンジニアリングには向かない。そのため、特定の制約に従った LSP を確立するために CR-LDP プロトコルをもちいた。

- L3TE

  現在の IP 層の次ホップ検索機構では、フローごとに異なる経路を与えることができない。そこで、送信元アドレスやポート番号等を次ホップの検索メトリックとして利用できるように拡張し、ユーザの要求に応じた経路を選択できるようにする。

- Diffserv

  IP 層での Diffserv および MPLS/Diffserv を利用して、合宿ネットワークから対外線を含んだ帯域的な DS ドメインを構成した。各構成ルータはパケット中の DSCP(Diffserv Code-Point) [7] に応じて、あらかじめ設定されている挙動制御を行う。

図 1 にシステムの概要を示す。

ユーザは予約を発行する際、Web サーバにアクセスする。このとき、User DB に登録されているユーザ情報によって認証する。予約は (src, dst) などのパラメータを Web ページのフォームに投入することによって行う。

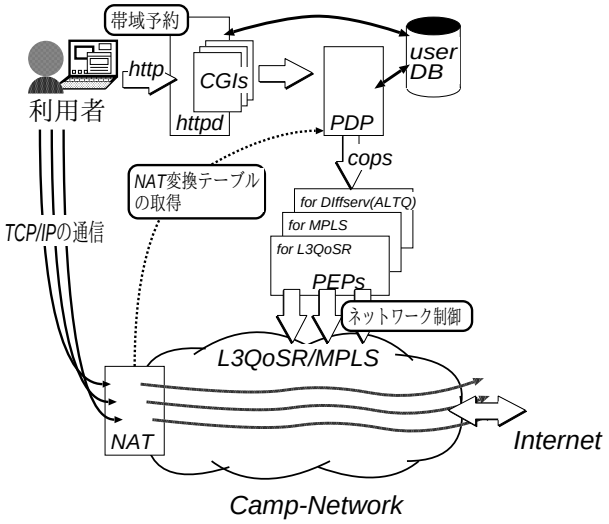

図 1: システムの概要

投入された予約は PDP(Policy Decision Point) により、課金や資源の状況にしたがってアドミッション制御される。予約が受理された場合、各構成ルータの制御パラメータを決定し、PEP(Policy Enforcement Point) に送信する。PEP では、転送層にしたがって制御パラメータを変換し、機器の設定を行う。

MPLS 網では、PEP の指示に従い、CR-LDP をもちいて LSP を確立する。また、L3TE では PEP によって経路を設定する。どちらの場合もユーザからのトラヒックは Diffserv 的に識別されて処理される。

1.4 結果

このように多数の技術を導入して生成されたネットワークであったが、合宿期間中特に問題なく動作し、実験の狙いである『広域ネットワークにおける柔軟な制御』が実現可能であることを実証できた。

1.5 まとめ

本実験を通してネットワーク制御に関するさまざまな知見が得られた。これらは論文としてまとめられる予定である。

[1] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. *An Architecture for Differentiated Service*, December 1998. RFC 2475.

[2] E. Rosen, A. Viswanathan, and R. Callon. *Multiprotocol Label Switching Architecture*, January 2001. RFC 3031.

[3] Ayame project. http://www.ayame.org/.

[4] Moon bear project. http://www.moon-bear.net/.

[5] L. Andersson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas. *LDP Specification*, January 2001. RFC 3036.

[6] J. Boyle, R. Cohen, D. Durham, S. Herzog, R. Rajan, and A. Sastry. *The COPS (Common Open Policy Service) Protocol*, January 2000. RFC 2748.

[7] K. Nichols, S. Blake, F. Baker, and D. Black. *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*, December 1998. RFC 2474.