

Road Space Rationing 方式のインターネットへの適用：ネットワーク仮想化を利用したベストエフォートネットワークにおける負担に応じた帯域スループットの実現に向けて

小林克志†

† 理化学研究所・計算科学研究機構

E-mail: jikob@riken.jp

あらまし ネットワーク中立性の議論では、ヘビーユーザと他の一般ユーザ間のコスト負担の不均衡によってインターネットサービスの持続性が懸念されている。すなわち、現状の指数的なトラフィック需要の伸びは ISP に定常的な帯域投資を要求し続ける一方、ISP にとっては定額制のインターネットアクセスサービスから現在以上の利益を得ることは期待できない。本稿ではアクセス課金の定額制、ベストエフォート型ネットワークの条件下で“負担に応じたサービス品質”を実現するネットワークアーキテクチャを提案する。このアーキテクチャはアテネ、メキシコシティ、サンパウロなどの大都市で採用されている交通需要管理手法の、“road space rationing”をに着想を得たものである。具体的には、ネットワーク全体を多数のネットワークスライスに仮想的に分割しスライス間の帯域公平性を提供するとともに、利用者も同数のグループに分割する。あるグループに属する利用者は対応する一個の仮想スライスのみデータを送信することができる。逆に、受信に関しては全てのスライスから自分あてのデータを受信することができる。ここで、高いスループットを要求する利用者は応分の負担のもとに複数の仮想スライスへのデータ送信を許すことで“負担に応じたサービス品質”が期待できる。本稿では提案方式の実現方法について議論する。

キーワード インターネット、ネットワークアーキテクチャ、ネットワーク中立性、ルータ

Katsushi KOBAYASHI†

† AICS, RIKEN

E-mail: jikob@riken.jp

1. はじめに

電子メール、Web、動画配信、P2P と次々と登場してくるネットワークアプリケーションによって、インターネットトラフィックは著しい伸びを示している。このような増大によって、ネットワークサービス事業者 (ISP) は継続的な設備投資が求められる。しかしながら、先進国の ISP にとってはあらたな投資に見合う収益拡大は困難となりつつある。この理由は、ネットワーク普及が一巡した環境ではネットワーク加入者数の伸びが期待できないうえ、定額制のインターネット利用契約では利用者あたりの収益増大も難しいことによる。国内のネットワーク利用調査によれば、少数の利用者によって帯域の多くが占められていることが報告されている [1]。一日あたり 2.5 GB (または 230 kbps) のデータを転送している上位 4% の利用者がアップロードトラフィックの 75% を、ダウンロードトラフィックの 60% を占めている。一方でその他のユーザのトラフィックは伸びが続いているものの極めて小さく、FTTH アクセスであっ

ても一日あたり 94 MB (または 8.7 Kbps) にとどまっている。すなわち、現状のヘビーユーザとその他のユーザのアクセスコスト負担はトラフィックからみるときわめて不均衡な状況となっている。このような不均衡を放置したまま、ISP の帯域投資を回収する負担増を求めることは困難と言わざるをえない。トラフィック需要に応える帯域投資なしにはネットワーク品質は悪化する一方であり、将来的にはインターネットサービスの持続性も懸念される。このような、トラフィックあたりのアクセスコスト負担の不均衡とインターネットの持続性については“ネットワーク中立性”の議論でも大きな課題とされている。

ネットワーク帯域資源利用の不均衡の是正と、ISP の収益機会の増大には、品質・価格の異なるサービスレベルの実現、すなわち差異化戦略が有効と考えられる。多くの ISP では、居住者・事業者向けでは異なるサービスレベルの品質・価格設定をおこなっている。現在、国内 ISP はネットワーク品質の維持を理由に、居住者向けサービスでアップロードトラフィックに制限を課しているところが多い [2]。そして、それを超える利用者

はより高価な事業者向けのサービスへの移行を求めている。しかしながら、この制限の対象となるユーザ数は極めて小さく、トラフィックあたりのコスト不均衡の是正に対する効果は小さい。この原因は、サービスレベル・価格設定が2レベルに留まり、その品質・価格設定が極端に異なることにある。さらに、居住者・事業者向けのサービスレベルの違いは最寄りISPのネットワーク内に限られ、複数のISPにまたがる環境ではその効果が期待できない。従って、より多くの加入者の要求に細かい粒度で多くのサービスレベルとその効果がインターネット全体で有効な差異化方式が望まれる。その他トラフィック量に応じた従量課金といった手法もあり、これはアクセス回線が投資・帯域ボトルネックとなる、3G/LTE等には有効と考えられる。しかしながら、バックボーン、あるいはアクセスが集中するサービス側がボトルネックとなる場合は従量課金方式では現実の負担と便益を反映することはできない。

DiffservはIPヘッダのDSCP領域へのマーキングによってサービスクラスの差異化を実現している[3]。Diffservはベストエフォートトラフィックではなく、比較的帯域が小さく、網設計の容易なVoIP、VPNから高い収益を得る設計となっている。今日、ほとんどのサービスがWeb上に構築されておりベストエフォートで差異化が実現できないアプローチではその効果は期待できない。さらに、ISPをまたがるネットワーク品質の確保は今日でも大きな課題であり、Diffservによるサービスレベルの差異化は特定のISP領域に限られることとなる。

我々は、ネットワーク全体にわたって有効なサービスクラス差異化の実現がインターネット持続性の鍵と考えている。本研究の目的は“負担に応じたサービス品質”を実現するネットワークアーキテクチャの確立にある。本稿では、あらたなQoS実現手法として、アテネ、メキシコシティ、サンパウロなどの大都市で採用されている交通需要管理手法の、Road Space Rationing (RSR) インターネットに適用する手法を提案する。本稿では、RSRが既存のインターネット基盤に大きな変更なしに適用できることを示す。併せて、RSRは、我々が提案している並列化ネットワークアーキテクチャ、Flexible Arrays of Inexpensive Network (FAIN) と高い親和性を有することも示す[4]。本稿の構成は次の通りである。第2.章でRSRの紹介とそのインターネットへの適用について述べる。第3.章で我々が既に提案したFAINアーキテクチャを示す。第4.章でFAINアーキテクチャ上におけるRSR構築について議論する。第5.章では運用における課題を示し、第6.に結論をまとめる。

2. 交通需要管理における Road Space Rationing 手法とそのインターネットへの適用

都市における交通需要管理において、Road Space Rationing (RSR) は少ない実装コストで、交通量あるいは空気汚染の削減を達成する有効な手法と知られている。RSRは都心部に流入できる時間帯を、例えば車両ナンバープレートの最後尾の数字によって制限する手法である。たとえば、偶数車両は月、水、金、奇数車両は火、木、土と流入可能な日付を制限した場合、理論的には20%の交通量抑制が期待される。RSRは、アテネ、メ

キシコシティ、サンパウロといった大都市で採用され、その効果も実証されている。しかしながら、RSRは富裕層が優位となる致命的な欠点知られており、他の都市とくに先進国では採用に至っていない。それは、富裕層はセカンドカーを購入することで簡単にRSRの制約を回避することができるためである。我々は、RSRで見られる富裕層の優位性は、従来インターネットで求められてきた“負担に応じたサービス品質”と考える。

本章では、RSRのインターネットへの適用方法について議論することとする。一般的には、RSRは需要および供給を複数のグループに等分し、分割された需要と供給を割り当てる手法である。すなわち、交通流量制御においては需要は車両でありナンバープレートの最後尾の偶奇によって2つのグループに分割される。そして、道路は供給でありそれは時間帯によって、2つのグループに分割されることとなる。一方、インターネットにおけるRSRでは、需要はIPパケット流量であり、ネットワーク帯域が供給となる。IPヘッダ情報を用いることでパケットのグループ分割は容易に達成できる。しかしながら、交通流量制御で採られた利用時間分割をインターネットに適用することは現実的ではない。すなわち、一日おきにインターネットにアクセスが可能という環境は粒度が荒すぎて現実的とは言えない。さらに、高々2つの差異化レベルでは、100倍以上の不均衡が生じているインターネットトラフィックの問題の是正には明らかに不足している。一方で、時分割数を増やして行くことは利用者がネットワークにアクセスできる時間をさらに縮めることとなる。ここでは、利用者の体感品質を悪化させることのない同時アクセスが可能で、多数のグループが同時に提供可能な分割手法を求められる。ネットワーク全体を数千以上のサブネットワークに分割するアーキテクチャとして、DybaBone、Flexible Arrays of Inexpensive Networks(FAIN)などが提案されており、これらはプロトコルの大幅な変更なしに既存ネットワークへの展開が可能である[4],[5]。

3. Flexible Arrays of Inexpensive Networks

FAINはパケット交換ネットワーク向けのアーキテクチャであり、多数の仮想スライスの配列(Array)と、エッジ側の多重化-分離デバイス(MUX-DEMUX)によって構成されている。ネットワーク内部における同期を排除するため、FAINでは従来ルータ・回線内部に閉じられていたパケット並列転送処理を、広域で行う設計としている。

FAINアーキテクチャでは、ネットワーク全体が複数の仮想スライスに分割される(図1)。端末を除くネットワーク構成要素、すなわちルータ・回線は全て同数の仮想スライスに分割されることとなる。単一のコントロールプレーン、すなわち一つの経路表、が分割されたスライスによって共有されることとなる(図2)。仮想スライスのネットワークトポロジは他のスライスと異なることはなく、仮想化前の物理ネットワークのトポロジと同一である。また、仮想スライスのデータプレーンは仮想化された転送エンジンと仮想回線によって構成されることとなる。そして、異なるスライスの転送エンジンと回線は隔離さ

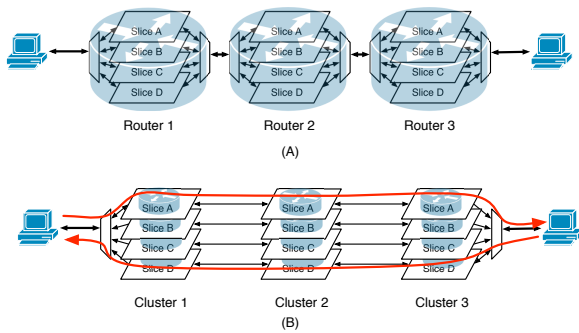


図 1 (A) ルータ・回線内部に閉じた並列パケット処理 (B) パケット転送処理における同期の排除

れており、データパケットは異なるスライスをまたがることはない。

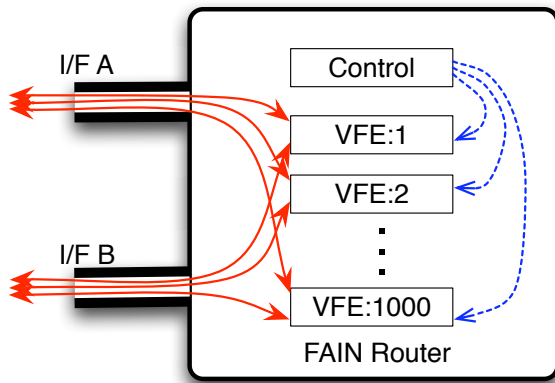


図 2 FAIN ルータにおけるルータクラスタアーキテクチャ。物理インターフェース A および B は数千の仮想リンクから構成されている。FAIN ルータは経路制御部と数千の仮想転送エンジン (VFE) から構成されている。左側の実線はデータプレーンのパケットの流れを示し、右側の破線は経路制御部と VFE の制御関係を示している。

FAIN では仮想転送エンジン・回線を、物理デバイスにそれぞれ割り当てることとなる。ここで、全ての仮想スライスを一つの物理デバイスに割り当てることもできるし、パフォーマンスとコストに応じて複数のグループに分割、複数の物理デバイスに割り当てることもできる。仮想スライ同士は隔離されていることから、異なる仮想スライスが割り当てられる物理デバイス間も隔離、それら相互の同期処理も不要となる。単純に物理デバイスを追加するだけでルータ、回線の帯域性能を向上できることが FAIN の特長である。

利用者を収容する FAIN エッジルータは MUX-DEMUX として動作するため全ての仮想スライスに対してパケットを送受信することができる。ネットワークから利用者向けのパケット送信では、エッジルータは仮想スライスに関係なくパケットを転送する。一方、利用者からネットワーク向けのパケット送信では、エッジルータは ISP の割り当てポリシーに従ってパケットをいずれかの仮想スライスに割り当てる。

FAIN では、end-to-end で仮想スライスを提供することと

なるため、なんらかのスライス識別子がネットワーク層 (の packets ヘッダ) で必要となる。この識別子を利用することで利用者側のポリシーを達成することもできる。たとえば、複数の仮想スライスにパケット送信が許された利用者は同一ストリーム内のパケットを全てのスライスに分散させて送信する、といったことも可能である。

仮想回線と物理デバイスの対応の柔軟性を高めるためには、多くの仮想スライスが用意されることが望ましい。我々は、IPv6 のフローラベル領域、20 bit、を FAIN スライス識別子として流用する方法を提案した [4]。これによって 100 万の仮想スライス数が提供できることとなり、アクセス可能なスライスによって 100 万段階の差異化レベルを実現している。IPv6 仕様では、フローラベルは端末側で選択することとなっている [6]。しかしながら、後述するような RSR の適用といった用途では、エッジルータ側でフローラベルを書き換え利用者がアクセス可能なスライスを制約する必要がある。フローラベルに関するここでの違いはエッジルータにだけ影響するもので、その他のルータおよび端末は既存のものをそのまま利用することができる。

4. 仮想スライス間の公平性

RSR のインターネットの適用における基本的な考え方は、それぞれの仮想スライスがネットワーク帯域を均等に占め、より多くの数のスライスにパケットを送信できる利用者が高いスループットを提供するという考え方である。そして、ISP はアクセスを許すスライス数に応じた柔軟な価格設定が可能となる。

RSR のインターネット適用には、ルータ・利用者エッジ・端末には FAIN に加えて以下の機能が必要とされる。

ルータ: FAIN スライス間の帯域公平性の提供

エッジ: ISP-利用者間の契約に基づくスライス数の制限、提供

端末: 複数のスライスを活用するトランスポートプロトコルスタックあるいはアプリケーション

以降の章でこれらの機能について議論する。

4.1 ルータ: スライス間の帯域公平性の提供

スライス間の帯域公平性は FAIN 仮想スライスを fair queuing, round robin, deficit round robin といった公平性の提供を目的とするキューイング機構の独立した送信キューに割り当てることで実現できる [7], [8]。複数のフローが同じキューに割り当てられることから、我々の手法は広く研究されてきた stochastic fair queuing をネットワーク全体に拡大したものを見なすこともできる [9]。複数ルータのキューイング機構を連携させるアプローチでは、Eriksson らの Justice では通信経路上のルータに weighted fair queuing の重みパラメータを展開する手法が提案されている [10]。しかしながら、Justice の目的は帯域スループットの公平性の提供であり、我々が目標とする“負担に応じたサービス品質”の達成ではない。

4.2 エッジ: ISP-利用者間の契約に基づくスライス数の制限、提供

これまで述べてきたようにスループットの差異化は、利用者がアクセス可能なスライス数によって実現される。たとえば、価格競争力が必要な居住者向けには単一のスライスのみが提供

される。エッジデバイスは、利用者発の全てのパケットのスライス識別子を上書きするとともに該当する仮想スライスに転送する必要がある。そしてこのような利用者は自身で複数の仮想スライスを管理する必要はない。また、ISP は上書きするスライス識別子を、負荷分散の適正化などの目的で動的に変更することもできる。

一方、複数スライスにアクセス可能な利用者を収容するエッジデバイスでは、契約に基づいたスライス識別子の検証、機能が必要となる。そして、端末側では複数スライスの優位性を活かす機能が新たに求められる。

4.3 端末: 複数のスライスの活用

最も簡単な複数スライスの活用法は、トランスポートコネクションを与えられたスライスに割り当てる方法である。利用者が個々のコネクション毎にスループットを差異化する必要があるれば、スライスに割り当てるコネクション数を制御するという方法が考えられる。すなわち上級顧客むけには複数スライスにコネクションを分散させ、全ての一般顧客向けのコネクションを一スライスだけに集中させるという方法である。また重要通信分野では、FEC で損失耐性を持たせたデータを複数スライスに分散し、特定スライスの輻輳を回避するといった方法が考えられる。さらに、コネクションあたりのデータ転送スループットの向上を目指して、転送データを複数スライスに割り当てる方法も考えられる。ただし、異なるスライスの同時利用ではスライスをまたがるパケットの順序交替が損失とみなされ転送性能がかえって悪化する懸念もある。従って複数コネクションを同時利用する Grid FTP, MPTCP, CMT-SCTP を用いるなど、サブコネクションがスライスをまたがらないような配慮も必要となる。

これまで述べた複数の機構の組み合わせによって RSR のインターネット適用が実現する。ここで重要な点は個々の機構は独立に動作すればよく、機構間でなんら連携は不要なことにある。例えば、個々の利用者がアクセス可能なスライスは ISP - 利用者間の契約によって決定される。その適用にあたっては利用者を収容するエッジデバイスの設定だけで完了するため実装、運用にあたって大きなコスト上昇は考えられない。

5. 運用の問題について

今日のルータが持つ送信キューは数個から数千個程度であり、100 万個 (あるいは 20 bit) の仮想スライスには明らかに不足している [11], [12]。このようなケースでは複数の仮想スライスをまとめて、例えば上位ビットのみの情報で、限られた送信キューに割り当てることで問題を回避できる。さらに、隣接する FAIN ルータ同士でキューの数が異なる場合でもキューへの割り当ては独立に行われるため問題は生じない。少なくとも利用者は経路上のルータに実装されたキュー数に相当する差異化レベルを享受することができる。そして、高いスループットが必要な利用者は実効的に期待される差異化レベルに相当する仮想スライスへのアクセスを契約すれば良い。

6. まとめ

本稿では、交通需要管理で使われている RSR に着想を得たアプローチでインターネット上に“負担に応じたサービス品質”の差異化サービス提供についての概念を示した。そして、RSR を我々が提案した仮想スライスを利用する並列ネットワークアーキテクチャ、FAIN に適用する手法について議論した。我々は、スライス数による差異化によって、ネットワーク中立性の議論で提起されたヘビーユーザとその他のユーザとのコスト負担の不均衡を改善できると考えている。本稿では我々の主張を裏付ける具体的な証拠を示していないが、今後シミュレーション、計測データを用いて我々の主張を裏付けていく。

謝辞 本稿の内容には、情報通信機構が実施した新世代ネットワーク仮想化ソフトウェア研究開発委託の成果が含まれる。

文 献

- [1] K. Cho, K. Fukuda, H. Esaki, and A. Kato, “Observing slow crustal movement in residential user traffic,” Proceedings of the 2008 ACM CoNEXT conference ACM New York, NY, USA, pp.12:1–12:12 2008.
- [2] “NTT Communications: New Bandwidth Controls to Help Ensure OCN Network Quality,” http://www.ntt.com/release_e/news08/0006/0625.html 2008. http://www.ntt.com/release_e/news08/0006/0625.html
- [3] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, “RFC2475: An Architecture for Differentiated Service,” 1998.
- [4] K. Kobayashi, “Flexible arrays of inexpensive network (FAIN): toward global parallelism in the internet to satisfy future traffic growth,” Proceedings of the 2008 ACM CoNEXT Conference ACM, pp.1–6 2008.
- [5] J.D. Touch, G.G. Finn, Y.S. Wang, and L. Eggert, “DynaBone: dynamic defense using multi-layer Internet overlays,” DARPA Information Survivability Conference and Exposition, 2003. Proceedings, vol.2 IEEE, pp.271–276 2003.
- [6] S. Deering and R. Hinden, “RFC2460: Internet Protocol, Version 6 (IPv6) Specification,” 1998.
- [7] J. Nagle, “On packet switches with infinite storage,” Communications, IEEE Transactions on, vol.35, no.4, pp.435–438, 1987.
- [8] M. Shreedhar and G. Varghese, “Efficient fair queueing using deficit round-robin,” IEEE/ACM Transactions on Networking (TON), vol.4, no.3, p.385, 1996.
- [9] N. Shacham and P. McKenney, “Packet recovery in high-speed networks using coding and buffer management,” INFOCOM’90. Ninth Annual Joint Conference of the IEEE Computer and Communication Societies. The Multiple Facets of Integration. Proceedings., IEEE/IEEE, pp.124–131 1990.
- [10] J. Eriksson, M. Faloutsos, and S. Krishnamurthy, “Justice: Flexible and enforceable per-source bandwidth allocation,” NETWORKING 2005, pp.1206–1218, 2005.
- [11] “Cisco CRS-3 Modular Services Card (Line Card),” Cisco Systems, 2010.
- [12] “Junos® OS, Class of Service Configuration Guide, Release 10.3,” Juniper Networks, 2010.