

次世代ネットワークとQoS

村田 正幸[†]

[†] 大阪大学サイバーメディアセンター
〒560-0043 豊中市待兼山町1-30

E-mail: †murata@cmc.osaka-u.ac.jp

あらまし 本稿では、次世代ネットワークにおけるQoS (Quality of Service) について述べる。まず、データ系アプリケーションに対するQoS要求をまとめ、その実現可能性について論じる。次に、データ系アプリケーションQoSのためのネットワークプロビジョニングについて議論し、トラフィック測定の重要性について述べる。さらにトラフィック測定に基づいたネットワーク設計手法のシナリオを紹介する。特に、そのためのネットワークインフラストラクチャとしてWDMの波長ルーティングに基づいた方式について述べる。また、ネットワークが提供すべき機能として「公平性」の問題について考え、最近の動向を紹介する。

キーワード QoS、ネットワークディメンジョンング、段階的ネットワーク設計、公平性

Next-Generation Network and QoS

Masayuki MURATA[†]

[†] Cybermedia Center, Osaka University
Toyonaka, Osaka 560-0043, Japan

E-mail: †murata@cmc.osaka-u.ac.jp

Abstract In this article, a network provisioning approach suitable to the next-generation Internet is discussed. Differently from the traditional telephone networks, it is difficult to guarantee QoS for end-users even in a statistical sense due to an essential nature of an end-to-end communication architecture in the Internet. We should therefore adopt another approach, based on the traffic measurement. In the approach, the traffic measurement is performed for monitoring the end-to-end QoS. Then, the network adaptively controls the link capacities to meet the user's QoS demands. For this purpose, the underlying network should support such a capability that the link capacities can be flexibly reused. With the WDM network as an underlying network, an example scenario for network provisioning is illustrated. Finally, we emphasize that the most important functionality supported by the network is "fairness," which cannot be realized by the end-to-end architecture of the Internet.

Key words QoS, network provisioning, incremental network design, fairness

次世代ネットワークとQoS



大阪大学サイバーメディアセンター
先端ネットワーク環境研究部門

村田正幸

e-mail: murata@cmc.osaka-u.ac.jp

http://www-ana.ics.es.osaka-u.ac.jp/

M. Murata

1



*Advanced
Network
Architecture
Research*



*Advanced
Network
Architecture
Research*

目次

- 背景
- インターネットフロー間の不公平
- 公平性向上のための手段
 - プロトコルの改善
 - ルータバッファでの制御
 - エンドホストでの資源管理
- まとめ

M. Murata

2

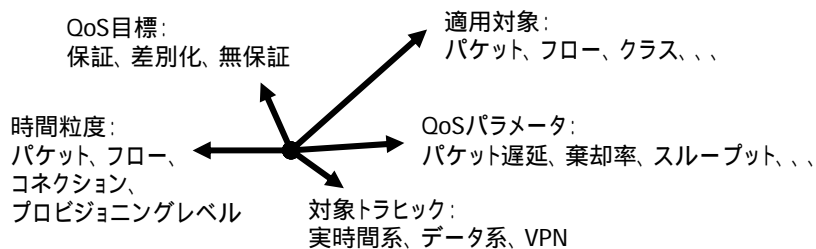


Osaka University

背景

- これまでのインターネット
 - ベストエフォートネットワーク
 - QoS保証は行われない
 - 帯域、遅延等
 - TCPを使えば、「確実に相手に届く」だけは何とかなる
 - 繋がることが重要だった
 - 当然、ユーザー間、フロー間の公平性は期待できない
- バックボーン、アクセス回線速度の飛躍的な向上
 - 「繋がる」以上のサービスへの要求
 - QoS保証
- 「公平性」が重要な要素になりつつある
 - 同じ料金なのに速度が違う
 - 2倍のアクセス速度を2倍の料金で使っているのにスループットは同じ

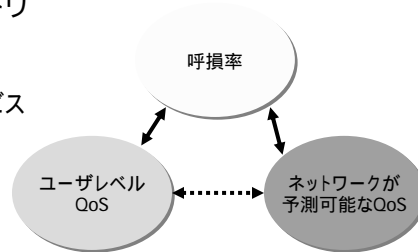
ネットワークQoSのための要素



- 例
 - これまでのインターネット: 全トラフィックを対象、無保証
 - IntServ: フローを単位とした遅延保証(実体はスループット保証)
 - DiffServ: クラスを単位とした遅延・スループットの差別化

電気通信網におけるQoS

1. 過去の統計量に対する蓄積
トラフィック特性
2. (古くは)単一キャリア、単一ネットワーク
3. アーラン呼損式
ローバスト(ポアソン到着、一般サービス
時間分布)
4. QoS測定 = 呼損率
キャリアが測定可能
5. 実時間転送; 音声、動画像
帯域保証のみで十分
エンド間保証が前提

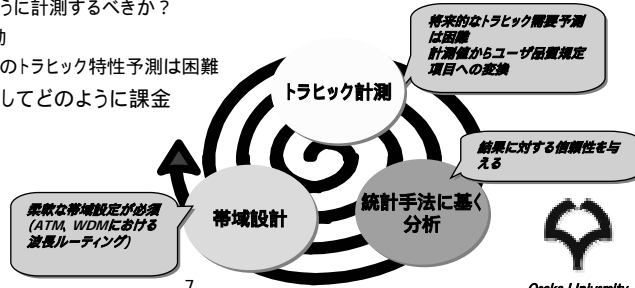


データ系アプリケーションにおけるQoSとは？

- データ系は帯域を食い尽くすアプリケーション
 - アクセス回線、エンドホストの高速化
 - TCP (=エンドホスト) が輻輳制御を行う
 - パックボーンの高速度は解決策にならない
 - これまではアクセス回線がボトルネックになっていた
- データ系に適したQoS機構？
 - IntServによるQoS保証
 - パケット棄却率、パケット遅延を「保証」できるか？
 - トラフィック契約の考え方 (QoSパラメータ、トラフィック特性を事前に申告) とマッチしない
 - RSVPのスケラビリティに対する限界
 - DiffServによるクラスに対するQoS差別化
 - 実現はHOL優先権制御で十分 (AFクラス)
 - 相対的なQoSはユーザのQoS要求とマッチするか？
 - QoS「保証」, 「差別化」なし
 - ただし、ネットワークプロビジョニングレベルでのQoS監視は重要
 - 帯域切り出し (VPN) は意味がある
- データ系アプリケーションQoSの3原則
 1. Data applications try to use the bandwidth as much as possible.
 2. Neither bandwidth nor delay guarantees should be expected. Only network provisioning can satisfy user's QoS requests.
 3. Competed bandwidth should be fairly shared among active users.

データ系アプリケーションに適したQoS制御： スパイラルアプローチ

- 少なくともプロビジョニングレベルにおけるQoS予測が必要
- 電気通信網ではなかった新たな問題
 - データ系QoSとは何か？
 - バケット遅延、棄却率はエンドユーザレベルの性能指標ではない
 - エンド間QoSはユーザしかわからない(エッジルータの可能性はありうる)
 - QoSをどのように計測するべきか？
 - トラフィック変動
 - エンドユーザのトラフィック特性予測は困難
 - 「サービス」に対してどのように課金するべきか？



M. Murata

7

トラフィック計測の2つのアプローチ パッシブ / アクティブ

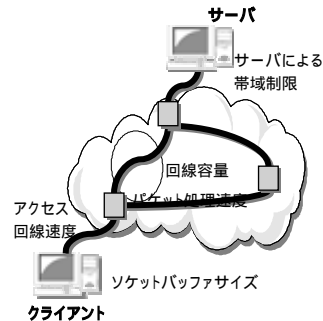
- パッシブな計測
 - OC3MON, OC12MON, ...
 - 点観測
 - 経路制御による経路の不安定性
 - TCPの誤り制御によるセグメント再送
 - 例: 利用率が低いのは輻輳制御のため? エンドユーザのアクセス回線が細いため? エンドホストのパワー不足?
 - ストリーミングメディアのレート制御
 - ユーザQoSは不明
- アクティブな計測
 - Pchar, Netperf, bprobe, ...
 - エンド間ユーザQoSの計測
 - ある特定のユーザのQoSがわかったとしてもネットワーク設計ができるわけではない
 - ネットワークトラフィックの変動への対処

M. Murata

8

トラフィック計測によるボトルネック特定

- トラフィック計測によるボトルネック特定
 - ネットワーク提供者による容量設計の限界
 - エンドユーザ(エッジルータ)による性能向上策
 - ボトルネックの特定
 - 回線増強へのフィードバック
 - TCP Calculus
 - RTT、パケット棄却率を用いたTCP性能特性の定式化
- エンドユーザによるQoS測定
 - アクティブ計測ツール(帯域推定ツール) pchar, pathchar
 - 計測ホストからProbeパケットを送出
 - RTTを計測
 - 最小RTTから各リンクの帯域(回線容量)を推定



TCP Calculus: TCPの理論値と測定値の比較

- TCP コネクションのウィンドウサイズの期待値 $E[W]$

$$E[W] = \frac{2+b}{3b} + \sqrt{\frac{8(1-p)}{3bp} + \left(\frac{2+b}{3b}\right)^2}$$

$E[W] \geq W_{\max}$ なら受信側のバッファサイズが不足

- RTT、パケットロス率、ウィンドウサイズの最大値、タイムアウト時の再送間隔を用いて予測

$$E[W] \geq W_{\max}$$

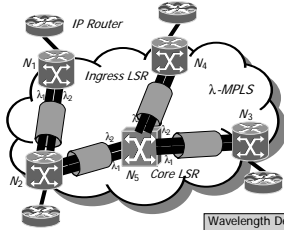
$$\frac{\frac{1-p}{p} + W_{\max} + \hat{Q}(W_{\max})}{1-p}}{RTT \left(\frac{b}{8W_{\max}} + \frac{1-p}{pW_{\max}} + 2 \right) + \hat{Q}(W_{\max})T_0 \frac{f(p)}{1-p}}$$

- 実測値と予測値からボトルネックを特定
 - 回線容量増大へのフィードバック

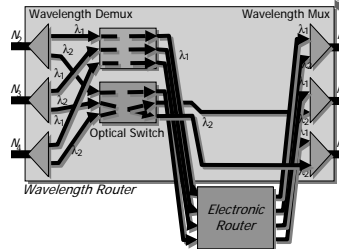
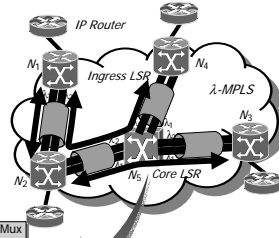


柔軟な帯域割当を可能とするインフラ： WDM波長ルーティングの利用

物理トポロジー



論理トポロジー



M. Murata



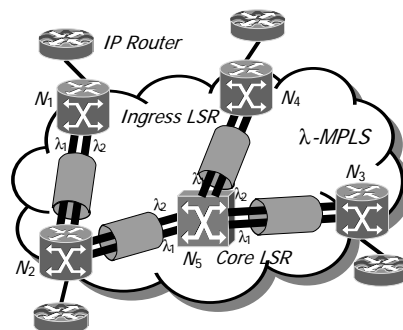
Osaka University



これまでの研究

経路 / 波長割り当て問題 (Routing and Wavelength Assignment: RWA)の例

- 与条件：
トラフィック量既知
- 目的関数：
利用可能波長を使い切って
各波長ごとのトラフィック量を
最小化



M. Murata

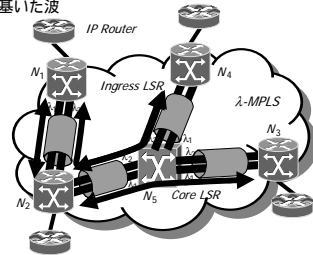
12



Osaka University

スパイラルアプローチの実現

- 段階的ネットワーク設計
 - 初期ステップ
 - 与えられたトラフィック量に基いたトポロジー設計; 従来の設計手法が適用可能; ただし、トラフィック予測が間違っていたとしても、追加ステップで修正可能
 - 追加ステップ
 - トラフィック測定 (パッシブ)、エンドユーザ品質測定 (アクティブ) に基いた波長設定
 - 波長の追加、削減のみ
 - バックアップパスの有効利用
 - 調整ステップ
 - 全体の波長有効利用を考慮したトポロジー再設計
 - サービスの継続性を考慮した1波長ルートごとの変更
- WDM技術による高信頼化: IP & WDM Integration
 - 共有プロテクション方式
 - 複数の障害には対応不可、必要波長数小
 - 追加ステップにおけるバックアップパスの有効利用
 - QoP (Quality of Protection)
- 今後の課題
 - 複数の設定要求をいかに調整するか?
 - 波長不足により設定失敗に終わった時の処理
 - 処理量の把握



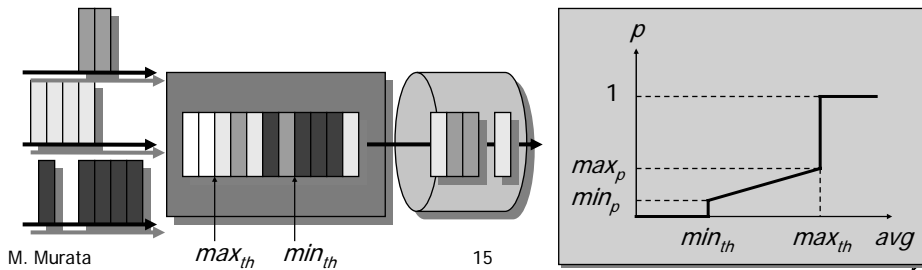
ネットワーク機能の再配分

- エンドホストに頼りすぎ
 - TCPが輻輳制御に責任を持っている
 - 輻輳制御は本来ネットワーク機能
 - 公平なサービスの実現を困難にしている
 - ソフトウェアのバグやコードの書き換えによって、輻輳制御を行わないホストも存在する
 - サービスの有料化を阻害する
- どの機能をネットワークに再配分すべきか?
 - フロー制御、誤り制御、輻輳制御、経路制御
 - RED, DRR, ECN, diff-serv, int-serv (RSVP)
 - ネットワークに過度に頼るのはインターネットのメリットを失う



REDのメカニズム

- 平均待ちパケット数に基づいてパケット廃棄率を制御する
 - $avg < min_{th}$ の時、すべてのパケットを受け付ける
 - $min_{th} < avg < max_{th}$ の時、確率 $p(x)$ でパケット廃棄
 - $max_{th} < avg$ の時、すべてのパケット廃棄
- バースト的に到着する同一フローのパケットのすべてを失うことを予防する: Fast retransmitへの対応

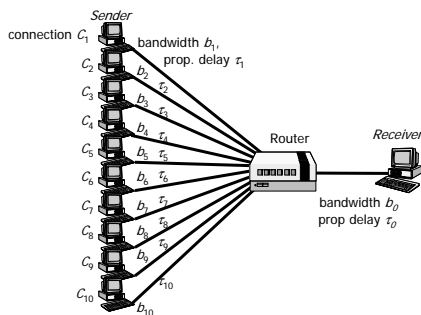


M. Murata

15

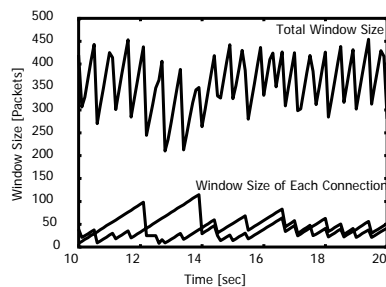
REDの効果

- fast retransmitの起動によってパケット損失の影響を最小限にとどめる
- 一時的に発生する不公平性を緩和

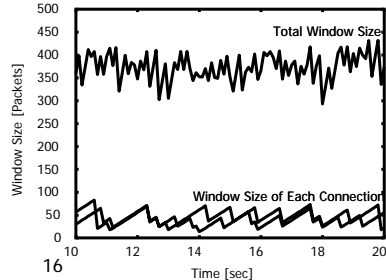


M. Murata

Drop-Tail Router



RED Router



Osaka University

コネクション間の公平性

- TCPは常に帯域に対して貪欲
- 帯域の公平な分配はTCPでは困難
 - ウィンドウサイズの変化による短期的な不公平性
 - 帯域、距離の異なるコネクション間の長期的な不公平性

TCP Renoの場合

$$S = \frac{1.3 \times \text{MTU}}{\text{RTT} \times \sqrt{p}}$$

TCP Reno + REDの場合: RTTにのみ反比例

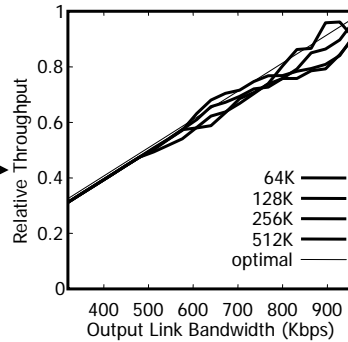
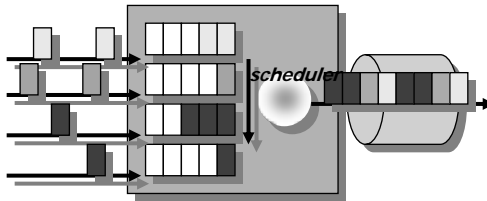
- ルータバッファでのパケット処理方式
 - プロトコルを意識しないで実現できるか?
 - TCP, UDP間
 - 異なるTCPバージョン間
 - ルータにおけるバッファ管理、スケジューリング; RED, DRR
- 既存プロトコルとの親和性を考慮したプロトコル設計
 - ルータに頼ることなく実現できるか?
 - 実時間アプリケーション (UDP) とデータ系アプリケーション (TCP)の公平性?
 - TCP-Friendlyな輻輳制御

公平性を実現するための 第4層スケジューリング

- ステート情報をフルに持つ
 - DRR
- ステートは保持せずにフロー識別を実現(フロー単位のバッファリングではない)
 - FRED
 - アクティブフローごとに到着 / 処理パケット数をカウントし、バッファ内パケット数を計算。それをREDのパケット棄却率に反映させる
 - Stabilized RED
 - Core Stateless Fair Queueing
 - エッジルータではフローごとのレートを計算、パケットヘッダに書き込む。コアルータでは、アクティブなフロー数を計測、パケットヘッダのレートに基づいて、公平な帯域割り当てを計算。それに基づいてパケット棄却・処理を決定する

DRRのメカニズム

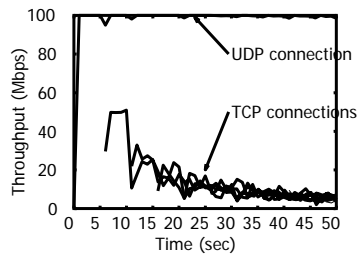
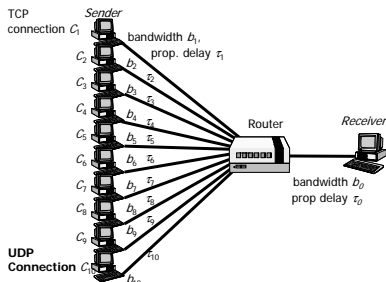
- ウェイトを考慮したビットレベルでのプロセスシェアリングを模倣



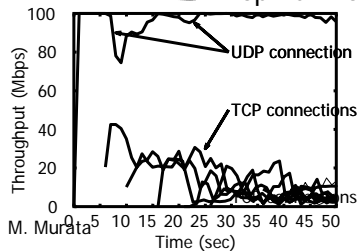
- 物理的(または論理的)にフローごとのパケットバッファリング
- IPアドレスとアクセス回線容量のマッピングを知っておく必要がある

FREDの効果

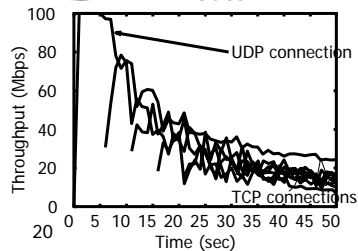
- RED Router



- Drop-Tail Router

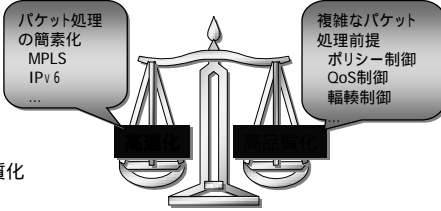


- FRED Router



今後？

- QoSに関して
 - バックボーンの高速度化: フォトニックインターネット
 - GMPLSに基づくルーティング + MPLSに基づく波長ルーティング
 - GMPLSに基づくルーティング + フォトニックパケットスイッチに基づく波長ルーティング
 - フォトニックIPルータ
 - 高機能フォトニックIPルータ
 - エッジルータ、ゲートウェイにおける高品質化
 - プログラマブルルータの活用
 - エンドホストの高速度化
 - ムーアの法則: CPUのコストパフォーマンスは18ヶ月で2倍に向上 (10年で100倍)
 - ビルジョイの法則(?): 回線容量は9ヶ月~1年で2倍に向上 (10年で1,000倍)
- インターネットが目の前にあったからこそ、それに適したWebというアプリケーションが生まれた
 - 背景: 画像圧縮技術、GUI、画像表示能力
 - にわとりと卵(?)
 - napster, gnutella
- 波長の有効利用
 - 波長をどれだけエンドユーザの近いところに持ってこれるか?
 - 多重波長数に依存



今後？

- エンド間QoSを保証、差別化することに意味があるか?
 - IntServ、DiffServの前提; 回線固定、ノード固定、サーバ固定
 - P2P; サーバが突発的に現れる
 - モバイル環境; 情報源が突発的に現れる
- ネットワーク資源の変動を前提とした、アダプティブなエンドホストによるQoS制御
 - 例: ストリーミングサービス vs. リアルタイム動画配信サービス
 - エンドシステムにとって利用可能な資源の実時間推定
 - ネットワークの資源管理は補助的な役割



参考文献

- Masayuki Murata, "Challenges for the Next Generation Internet and the Role of IP over Photonic Networks," *IEICE Transactions on Communications*, Vol. E83-B, No. 10, October 2000.
- Go Hasegawa and Masayuki Murata, "Survey on Fairness Issues in TCP Congestion Control Mechanisms," to appear in *IEICE Transactions on Communications*, January 2001.
- Masayuki Murata, "On A Network Dimensioning Approach for the Internet," to appear in *IEICE Transactions on Communications*, 2001.
- Masayuki Murata and Ken-ichi Kitayama, "A perspective on photonic multi-protocol label switching," to appear in *IEEE Network Magazine*, August 2001.
- Shin'ichi Arakawa and Masayuki Murata, "Lightpath Management of Logical Topology with Incremental Traffic Changes for Reliable IP over WDM Networks," to be presented at *OptiComm*, 2001.
- Go Hasegawa, Tatsuhiko Terai and Masayuki Murata, "Scalable socket buffer tuning for high-performance Web servers," to be presented at *IEEE ICNP 2001*.
- Kazumine Matoba, Shingo Ata, Masayuki Murata, "Capacity Dimensioning Based on Traffic Measurement in the Internet," to be presented at *IEEE Globecom*, 2001.

