

IP ネットワークにおける通信品質サービス実現の提案

日下 貴義 菅野 政孝

NTTデータ通信株式会社 技術開発本部

〒135 東京都江東区豊洲 3-3-3 豊洲セントビル

Tel: 03(5546)9571 Fax: 03(5546)9572

Email: kusaka@open.rd.nttdata.co.jp

あらまし:

インターネット上で EC やマルチメディアサービスが提供されるに伴い、IP ネットワークには高い品質が要求されてきている。しかし、現状の TCP/IP ベースのネットワークでは、それらのサービスが要求する品質を十分満たすことができないという問題がある。本稿では、上位サービスから要求される通信品質のうち、ネットワークを経由するパケットの遅延時間の保証を行うサービスクラスに注目し、ルータのようなひとつの中継装置を通過するパケットの最大遅延時間の制御方法を提案する。高品質ネットワークを提供する立場から、遅延時間制御に関して既存のパケットスケジューリングの問題点と、ここでの提案方法によって見込まれる改善点を示す。

キーワード: QoS, 通信品質, IP ネットワーク, 遅延時間制御

an Approach of Communication Quality of Service on IP Based Network

Takayoshi Kusaka, Masataka Sugano

NTT DATA CORPORATION

Research and Development Headquarters

Toyosu Center Bldg., 3-3-3, Toyosu, Koto-ku, Tokyo, 135, Japan

Tel: 03(5546)9571 Fax: 03(5546)9572

Email: kusaka@open.rd.nttdata.co.jp

Abstract:

Recently high quality networks are required for the multimedia applications or electronic commerce applications. Meanwhile, the network based on TCP/IP as Internet is not so enough high quality that it can't applied for such applications. In this paper, we focus the service class that guarantees delays through the networks among sorts of network quality of service(QoS) required by applications, and suggest a method to control maximum queuing delays within a packet relay equipment such as routers. As the point of offering network QoS, we describe problems of existent packet scheduling and prospective improvement by our suggested method.

Key words: QoS, Communication Quality, IP networks, Delay Control

1. はじめに

TCP/IP をベースとした今日のインターネットはベストエフォート型のネットワークであり、マルチメディアや電子商取引といった多様化するネットワークサービスの要求する通信品質 (QoS: Quality of Service) を、十分に提供できていない。インターネットの発展のためには、ビジネストラフィックを流すに耐えうる高い通信品質を、ネットワークのサービスとして提供することが重要であると考えられる。

現在の IP ネットワークで、高い通信品質を提供するサービス実現のために必要となる機能は、大別すると、ポリシーコントロール、シグナリング、QoS ルーティング、パケットスケジューリング、以上の4つのフレームワークがあると考えられる[1]。本研究では、これら高品質ネットワークの機能の一部を提供するために、パケットスケジューリングについて改良の提案を行う。

通信品質を表現するファクターに、帯域 (Bandwidth)、遅延 (Delay)、損失率 (Loss Rate) が考えられる。実際にこれらを制御する方法がパケットスケジューリングの仕事である。

IP ネットワークを構成する通信機器では、中継ノードにおいて、通過するパケットの遅延時間を制御しないことが一般的である。制御しても、相対的に小さい遅延時間を提供するものであったり、過剰な帯域を消費して遅延を短く調整するものであったりして、定量的な遅延時間制御は困難である。このようなネットワークでは、エンドーエンド間のパケットの遅延時間が網の構成や混雑具合によって可変で、最大遅延時間は大きくなりがちである。

通信内容が、それほど遅延に厳しくないデータ転送だけであれば、回線の高速化と統計的な多重化によって、十分な帯域を確保すれば、遅延時間がデータ転送にとって重要な問題になることは、まれである。

しかし、対話型リアルタイム通信を行うアプリケーションは単純なデータ転送に比べ遅延時間により敏感であり、大きな遅延時間では使用に耐えない。実際、電話のような音声会話における遅延時間はITU-Tの勧告 (G.114 他) 等で示されており、遅延時間の上限値を保証することは、対話型リアルタイム通信のアプリケーションを快適に使用するために必要である。

さらに、商取引などの信頼性を必要とするビジネスランザクションにおいても、最大遅延時間が不明であったり、短い遅延時間を保証できないことは信頼性の点で問題である。

本稿では、パケットスケジューリングの中でも、既存の帯域制御方式を改良することによって、帯域制御を行う一つ、単一中継ノードにおける最大遅延時間の制御を行う方式を提案する。ひとつのノード内最大遅延時間が制御できれば、エンドーエンド間の最大遅延時間も制御できると考える。第2章で、従来の方式で最大遅延時間を制御する上での問題点を指摘し、第3章以降では改善案の詳細を解説する。

2. 従来のパケットスケジューリングの遅延時間制御に対する問題点

遅延時間に関して優れた特性を持つトラフィック制御の代表的なスケジューリング方式に WFQⁱⁱ (Weighted Fair Queuing) がある。WFQ のようにトラフィックのアイソレーションを行うスケジューリング方式において、資料[2]で参照されるように、エンドーエンド間のあるフローにおける最大遅延時間は一般的に以下の式1で表現できる。

$$\text{Delay_bound} = \frac{(b-M)}{R} \times \frac{(p-R)}{(p-r)} + \frac{(M+Ctot)}{R} + \text{Dtot} \quad \dots \text{式1}$$

r : Token Bucket Rate (Sustainable Rate)
b : Token Bucket Size (Burst Size)
p : Peak Rate
M : Maximum Packet Size
R : Service Rate
Ctot : Total Rate-Dependent Error Term
Dtot : Total Rate-Independent Error Term
 $0 \leq r \leq R \leq p, 0 \leq m \leq M \leq b, 0 \leq Ctot, 0 \leq Dtot$

ただし、電気的な伝播遅延時間や機器内部のプロセッシング性能による処理時間は十分小さいものとして扱う。

式1において、r, b, p はトラフィックの特徴を表しており、M, Ctot は通信経路により決定される値である。Dtot はキューイング遅延時間ⁱⁱⁱを表し、複数のフローが重なっているとき大きくなる。

ここで R を調整することによって遅延時間を調整できる。つまり R を大きくすれば、遅延時間を小さくできる。この操作は遅延時間を小さくすると同時に、実際に必要な帯域 r よりも過剰な帯域が、ひとつのフローのために中継ノードで消費されてしまうという問題点をもつ。

また、QoS 制御を行うフローの数やバースト長が一定で

ⁱ PGPS (Packet by packet Generalized Processor Sharing) とも呼ばれ、トラフィックのアイソレーションに有効な規則をさだめている[5]。

ⁱⁱ パケットが送信用のキューに移動 (挿入) した直後から、実際に送信インタフェースより送信が開始されるまでの時間。

あればDtotも一定となるが、実際運用上、制御対象のフロー数や各々のバースト長はほとんど不定であるため、Dtotはフローの数とそのバースト長に応じて変動する値になる。このことは、扱うフローの数だけ最大遅延時間が大きくなることを示し、見積もる遅延時間も大きくなりがちであるという問題点をもつ。

以上の二つが、遅延時間制御を行う点で従来の方式の問題点と考えられる。

3. 割り込み型パケットスケジューリング方式

3.1. 方式の概要

今回提案する「割り込み型パケットスケジューリング方式」(Interrupt Packet Scheduling, 以下IPS)はキューイング遅延時間の制御により、単一中継ノードにおける最大遅延時間の確保を試みるものである。

IPSは、フロー毎に遅延時間の上限値を制御する。式1における第一項と第二項は帯域R以外定数値であるが、過剰なRを設定することによる遅延時間制御を行わず、 $R=r$ と固定して必要帯域分のみを設定する。このことにより遅延時間制御の対象になるのは、フロー数の増加を見越して大きく見積もらざるえなかった第三項のDtotになり、これを制御することが対象フローの遅延時間制御を行うことになる。

Dtotは複数台中継ノードのキューイング遅延時間の総和であるが、ここでは中継ノード一台における遅延時間Dについてのみ制御を行う。すなわち、Dの制御を行えば、その合計であるDtotの制御も行えることになる。

Dtotを要求される遅延時間になるよう実現するための経路選択や中継ノードに設定するDは、QoSルーティングの仕事である。DはQoSルーティングのために、遅延に関するパラメータを提供することになる。

Dは、ある中継ノード内で、送信を開始しようとしたフローが、他のフローの送信が先に行われるため、そのフローのバーストが終わるまで送信できずに待つ時間の合計である。そのため、最悪で中継ノードが抱えている残り全てのフローの数だけ、送信を待たねばならない事態が発生する。特にバースト性の強い通信が混在していると、遅延時間の最大値は大きくなりがちである[3]。従来の方式では、トークンパケットなどでシェーピング(ポリシング)されていても、出力キューで一時的に物理帯域を越えることによって、Dについて大きな遅延が発生していたため、Rの増加によって遅延が小さくなるよう補っている。そこで、

IPSでは、より小さい遅延時間を要求するフローは出力キューのより前方(出力口により近い)に、大きい遅延時間を要求するフローは出力キューの後方にキューイングさせる。つまり、遅延時間について上限があるフローについては、決められた規則にしたがって作成された優先順位の高いキューにキューイングすることによって最大遅延時間を制御する。IPSはDの上限を制御するものである。

ただし、Dtotを調整しても実現できないような遅延時間が要求されたとき、すなわち、

$$\text{要求遅延時間 } D_{\text{request}} < b/r + Ctot/r$$

のときは $R=r$ では実現できないので、 R は $r < R \leq p$ の値を取る必要がある。

3.2. 動作

3.2.1. 準備

ここで実現する遅延時間の制御は、受信側の制御と送信キューの制御の大きく二つから構成される。受信側の制御はポリシング(帯域制御)を行っている。遅延時間要求があった特定のフローに対応するキューの、送信キューでの配置方法が割り込み型パケットスケジューリング方式になる。

パケットの発生源(送信端末など)は式1で表されるパラメータのうち、 r , b , m , M といったトークンパケットのパラメータは最低でも指定する必要がある。この指定をせずに送信を開始する場合、ポリシングによりIPSの対象にならないパケットが発生する可能性がある。

Dtotに求められる遅延時間は以下のようにして求める。

$$D_{\text{tot}} = D_{\text{request}} - b/r - Ctot/r$$

(D_{request} : 要求遅延時間)

この後、Dtotは通信経路上の中継ノードで分配されるが、この分配方法はQoSルーティングの一部として実施されるものであり、ここでは言及しない。分配された値はDとなり、IPSの制御パラメータとして利用される。

受信されたパケットは、中継ノード内部で出力インタフェースの選択を終えたあと、遅延時間制御の対象になるパケットかどうか選別される(QoSフィルタ)。遅延時間制御の対象になるパケットの場合は、IPSの処理が行われる。なんの制御の対象にもならないパケット(ベストエフォート)は、最終的に送信キューを構成するキューのひとつであるデフォルトキューへ移動することになる。(図1各機能の構成図参照)

送信キューの全体バッファサイズ(B_{transmit})は、その送信インタフェースで発生する最大キューイング遅延時間を与えるものになる。各々の制御対象フローが持つパー

ストサイズ b の総和は、この B_{transmit} より小さくなる。

中継ノード一台における最大パケットサイズ M の許容値は、その中継ノードで実現できる最小キューイング遅延時間と C_{tot} とを与える。したがって、最小遅延時間を与える M の許容値は必要以上に大きくしないことが望ましい。

(一般的には物理リンク層の MTU : Maximum Transmit Unit と同じサイズになる)

3.2.2. 遅延時間制御の処理手順

図1 (各機能の構成図) に全体的な流れを示す。

(1). 受信側制御

中継ノードの受信側で、QoS フィルタによって選別され、QoS の必要なパケットは帯域制御 (ポリシング) を行う。この帯域制御手法は、Parekh[4]等によるアプローチと同様であり、「リキーパケットつきトークンパケット」で正規化するものである。QoS の不要なパケットは、ベストエフォートとしてデフォルトキューに移動する。また、ポリシングでの違反パケットについても、デフォルトキューへ移動する。

受信側制御を終えたパケットは、IPS の処理が行われる。

(2). 割り込み型パケットスケジューリング (IPS) の処理

図1, 図2 (処理概念図) 参照。

ポリシングを行ったパケットのうち、遅延時間制御が必要なパケットは、IPS で作成された送信キューに移動する。

IPS は基本的に、新しい遅延時間要求 (QoS シグナリングなど) により、優先順位を持った送信キュー (プライオリティキュー) 中で、適切な優先順位の位置を計算して、新しい優先順位を持つキューを挿入するものである。

遅延時間制御キューの作成方法と、挿入手順を以下に示す。(IPS の処理手順)

①. 遅延時間制御キューの作成/挿入位置計算方法

遅延時間制御キューを挿入すべき、送信キューの先頭からの位置を、パラメータ D によって、以下の計算で求める。

$$\text{Insert_position_Head} = \text{Tx} \times D - b$$

$$\text{Insert_position_Tail} = \text{Tx} \times D$$

Insert_position_Head (IH) : 遅延時間制御キュー先頭の位置

Insert_position_Tail (IT) : 遅延時間制御キュー末尾の位置

Tx : 送信インタフェースの物理的な最大送信速度

②. 遅延時間制御キューの挿入方法

ここでは実際に遅延時間制御キューが送信キュー群に挿入可能かどうかを判断し、可能な場合は、新しく優先順位をつけたキューが作成され送信キューでソートされる。

(図2 参照)

具体的には、すでに送信キュー群に存在するキューを持つ IH や IT と、これから挿入する遅延制御キューの IH (IH_{new}) や IT (IT_{new}) とを比較することによって、挿入の可否と優先順位の位置を求める。

区間 [IH_{new}, IT_{new}] に既存の IH や IT が存在しない場合、IH_{new} に最も近い IT を持つキューと、IT_{new} に最も近い IH を持つキューの間の優先順位を付けて、新しい遅延制御キューの挿入が完了となる。

また、既存のキューの IH, IT の値を変更ⁱⁱⁱ、あるいは挿入キューの IH_{new}, IT_{new} の値を変更することにより挿入ができる場合は、挿入キュー前後にあるキューの中間の優先順位を付けて、新しい遅延制御キューの挿入を行う。

上記以外の場合は、既存の送信キューの状態では新しい遅延時間制御キューを挿入できないことになる^{iv}。

③. 遅延時間制御キューの削除方法

削除は、基本的に挿入の逆手順となる。挿入のために押しのけたキュー群は、もとの位置に戻される。これにより、詰められていた位置を元のように空けるため、別のキューの再挿入をしやすくする。

(3). 送信キューの処理

IPS により、IH が小さいフローのキューが高い優先順位を持つことになる。すなわち、小さな遅延時間を要求しているキューが高い優先順位を持つことができる。送信キューはプライオリティキューイングにより送信を行うので、要求遅延時間が小さいフローの順番に優先して送信される。

なお、QoS の制御を行わないデフォルトキューは最も優先順位の低いキューとし、帯域制御のみのキューは二番目に優先順位の低いキューとして優先順位を付ける。

4. 遅延に関する特性と他方式との比較

IPS はフロー毎のパケットの最大キューイング遅延時間に注目した方式であり、フローが持つバースト (サイズ b) を許しながら^v、要求遅延時間とバースト長によりプライオリティキューイングの優先順位を決定し、単一中継ノード中のフロー毎にキューイング遅延時間の上限を確保するものである。すなわち、これらの上限値を合計すると、式1の D_{tot} の上限値を与えることになる。送信キュー中にお

ⁱⁱⁱ 主に送信キューの先頭に近い位置へ変更、ただしバーストサイズは変更しない。

^{iv} QoS 制御ができない (この場合要求遅延時間を満たす処理ができない) がわかった後の処理は、QoS に関するポリシーコントロールによって決定される。

^v 式1で表されるトラフィックの特徴は維持される。

図1 遅延時間制御 中継ノード内部各機能の構成図

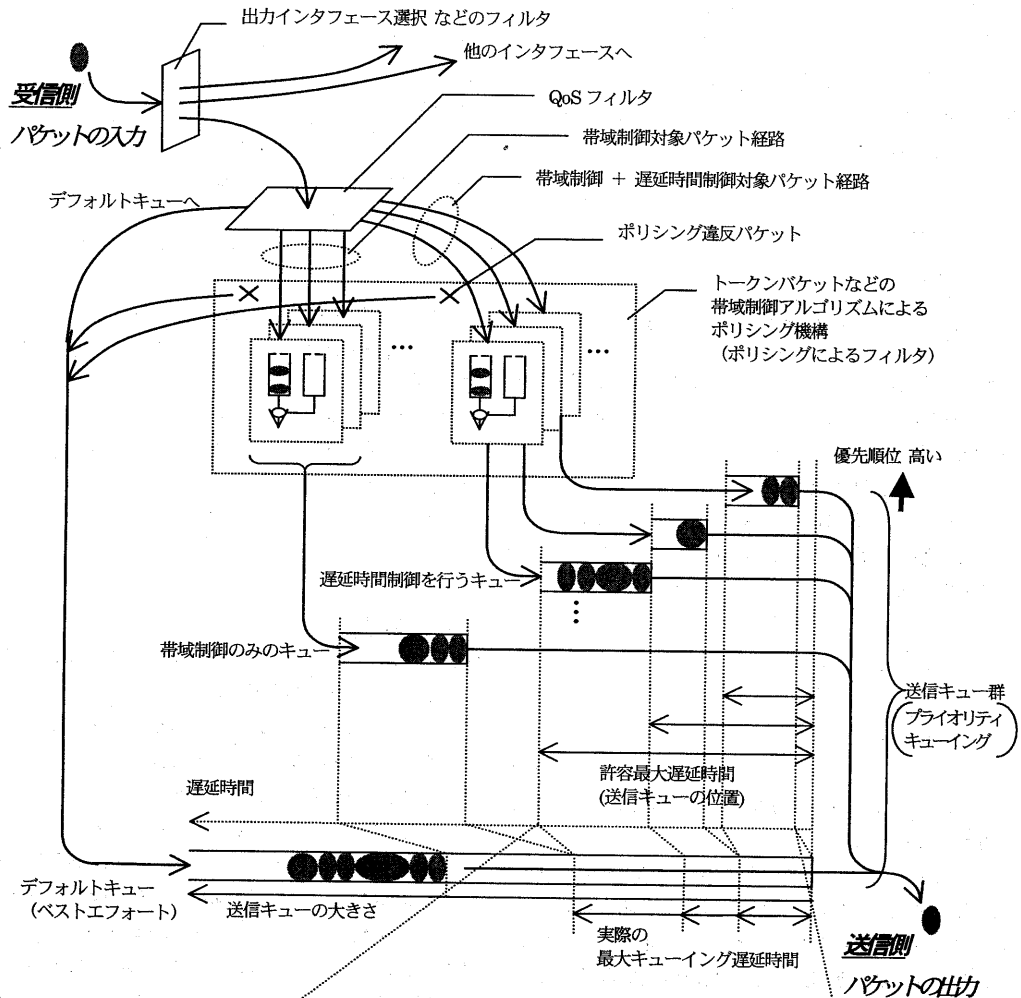
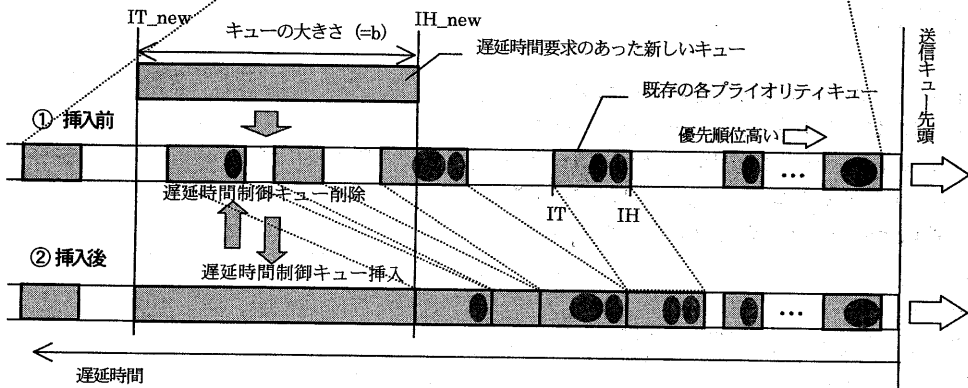


図2 割り込み型パケットスケジューリング処理概念図



ける制御対象になるキューの適切な位置を前章のようにして求め、最低でもその位置から後方に制御キューが動かないように優先順位を付けると、後はプライオリティキューイングによって、挿入位置で確保された遅延時間より短い遅延が保証できる。優先順位 p (p は小さいほど優先順位が高い) を持つキューの実際の最大遅延時間 D_{max} は以下のように表される。

$$D_{max} = \frac{\sum_i^p b_i}{T_x}$$

b_i : 優先順位 i のキューのバーストサイズ
 p : 遅延時間制御キューの優先順位
 T_x : 送信インタフェースの物理的的最大送信速度

最大遅延時間を与えることができる WFQ スケジューリング方式の場合、複数のフローが同時にバケットをバースト^{vi}したとき、最大で全フロー数のバースト量だけキューイングによる遅延が発生する。これは、各フローに対応するキューが送信をほぼ同時に開始しようとする場合、フローによる送信順番に規則がないために発生するキューイング遅延である。

$$D_{max} = \frac{\sum_i^n b_i}{T_x}$$

b_i : フロー i のバーストサイズ
 n : 全フローの数

この最大遅延はフローの数によって増加し、ときに大きくなりがちである。

それに対し、IPS は要求された遅延時間によってフローに対し優先順位をつけ、その優先順位より前になることが許されるキューの大きさの合計は常に一定値以下 (IH 以下) であるので、最大遅延時間は要求最大遅延時間 (IT) を超えることがない。

5. まとめ

IPS は、単一ノードにおける最大遅延時間を制御し、要求遅延時間を確保する手法である。トークンバケットのようなシェーピングを使った帯域管理を前提としているため、スケラビリティの点ではプライオリティコントロールを加えた WFQ と同様になる。WFQ の実装は大量のフローを

同時に扱う場合、多くの処理を必要とするため、スケラビリティの点で良い方法とはいえない。したがって、IPS も同様なことが言える。スケラビリティは、QoS コントロールを行うときの宿命としてトレードオフになりがちである。それに対し、IPS の場合、ひとつひとつのフロー毎に制御を行うのではなく、同じ遅延時間を要求するフローどうしはまとめることによって、ひとつの優先順位で複数のバースを管理する方法が考えられる。

IPS は、高品質ネットワーク構築のフレームワークのうち、バケットスケジューリングについて、最大遅延時間制御を行うという通信品質サービスの提供を提案している。最大遅延時間が制御できれば、QoS ルーティングにおいて必要となる遅延に関するパラメータをひとつ提供できると考えられる。

今後は、中継ノード (ルータ) において、IPS を実装することにより、実装上の最大遅延時間や平均遅延時間の評価を行い、更に RSVP といったシグナリングとの連携を行う予定である。

【参考文献】

- [1] 日下,木幡,松田,菅野,"IP ネットワークにおける通信品質の検証",情処研報 Vol97 No35.
- [2] S.Shenker,"RFC2212 Specification of Guaranteed Quality of Service",IETF,Sep.1997.
- [3] C.Partridge,"Gigabit Network",Aug.1993.
- [4] A.K.J.Parekh, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks", LIDS-TH-2089, MIT Laboratory for Information and Decision Systems, 1992.
- [5] S.Floyd,"Link-sharing and Resource Management Models for Packet Networks",IEEE/ACM Transactions on Networking,Aug.1995.

^{vi} これはトークンバケットによるシェーピングにおいても、ピークレートが大きい場合 (たいていは物理的的最大送信速度)、サイズ b に近いバーストが発生する。