

広域網におけるネットワーク応答特性プロ トコルの実装と評価

河合 東 太田 賢 渡辺 尚 水野 忠則

静岡大学大学院理工学研究科

インターネットの利用率が高まるなか、ますますリアルタイムマルチメディアコミュニケーションに対する期待は大きい。そのような通信サービスに対処するため、実時間転送サービスをできる限り保証するネットワーク応答特性プロトコル NPP を提案した。NPP の実装に関して新たにルータの機能を UNIX の BSD/OS2.1 に実装し、プライベートなサブネットを構成してその動作を確認した。また NPP には、ユーザが送信パケットに対し高い優先度のみを与えると機能しないという問題があった。そこで、NPP キューモデルに対しフェア・キューイングモデルを組み合わせることによってこの問題を解決した。

An Implementation of the Network Protocol with Performance and its Evaluation over Wide Area Networks

Azuma Kawai Ken Ohta Takashi Watanabe Tadanori Mizuno

Graduate School of Science and Engineering, Shizuoka University
3-5-1, Johoku, Hamamatsu, 432 Japan

We proposed the Network Protocol with Performance(NPP), which made effort to guarantee real-time transport services for time-critical traffic such as multimedia data and control data of real time systems. We have implemented the routing function and tested it on our test-bed. When a lot of packets are assigned high priority by greedy rogues, NPP doesn't work well. So we've adopted the Fair Queuing Model to avoid such a situation.

1 はじめに

インターネットを利用する人々の増加にともなう、インターネット環境で利用できるさまざまなアプリケーションの開発され市場にでていく。特に最近の通信サービスは、リアルタイム・マルチメディアコミュニケーションが主流になっている。マルチメディアアプリケーションが扱うデータの種類には、テレビ会議システムのような双方向ストリーム型、VODのようなストリーム型、そして、電子カタログ配送サービスなどの非リアルタイム型データの3種類に分けられる。リアルタイム性を持ったデータは低遅延で転送される必要があるが、そのほかにもアプリケーションやシステムの制御メッセージについても低い遅延で転送する必要があるため、タイムクリティカル通信 (Time Critical Communication: TCC) サービスが要求される。しかし現在のパケットネットワークは時間制約を必要としないデータを対象に設計されているためその要求に答えることができない。

この問題に対する現在の一般的なアプローチの一つに、時間制約の保証がなくトラフィックの変動が激しいネットワークに対してアプリケーションがネットワークに適応することである。もう一つのアプローチは、既存のネットワークを時間制約を保証するように改良することである。現在 RSVP[1] や XTP[2] などに代表されるように、ネットワークバンド幅、バッファ領域などの資源予約に関して数多くの研究がなされている。

しかしながらサービスを保証するためには、送受信ホスト間のパス上にある全ての中間ルータ群に対して資源予約を行なう必要がある。また資源予約はマルチメディアストリームに対しては有効であるが、制御メッセージのようなストリームのでないものについてはその処理により即時に送信ができないといった問題が起こる。

既存の TCP/IP パケットネットワークに対して、改良をほどこすことなく、実時間制約を考慮するような高度な転送サービスをインターネット環境で提供するプロトコルであるネットワーク応答特性プロトコル (Network Protocol with Performance: NPP)[3] が提案されている。NPP は IP をベースにしたネットワーク層に位置する

プロトコルであり、パケットスケジューリングなどによって、指定された制限時間内にパケットをできる限り転送することを保証する。なおこの転送サービスは、時間制約を完全に保証するわけではないため、最善努力 (ベストエフォート) 型ソフトタイムクリティカル通信サービスという位置づけである。

本稿では、このネットワーク応答特性プロトコルの NPP キューモデルに関する変更点について述べる。またルータ機能を実装し動作確認を行う。

2 内外の研究

2.1 FIFO+

FIFO+ キューイング [4, 6] は、統計多重の利点をいかしながら、FIFO キューイングでの最悪時の振舞いを補正するアルゴリズムと導入したものであり、そのアプローチの仕方は Jitter-EDD[4] によく似ている。

フローは各パス上のルータの平均キューイング時間を要求し、各パケットに対しルータは、そこでのキューの待ち時間とパケットクラスの平均待ち時間の差をとり、その差をパケットのオフセットフィールドから足す(あるいは引く)。こうしてパケットのオフセットフィールドには、平均よりどれだけ進んでいるのか、あるいはどれだけ遅れているかが記録される。これとパケットの実際の到着時間を加えた時間を指標にし、パケットスケジューリングを行う。平均より進んでいるパケットは待ち行列に押し戻され、反対に遅れているパケットは押し出される。

FIFO+ の利点は、後に述べる WFQ よりも最悪時の遅延時間の限界が小さいことである。問題点としては FIFO+ は統計多重化法であるので、厳密にフローごとに分けてはいないため遅延の保証を満足できないことがあるという点と、実装が容易ではないという点である。

2.2 WFQ

重み付けフェアキューイング (Weighted Fair Queuing: WFQ)[4, 6] はフェア・キューイング (Fair Queuing) を改良したものである。フェア

キューイングは、全てのホスト N に帯域の割り当てを公平 ($1/N$) にし、ラウンドロビン方式を用いてホストを選択しそのキューからパケットを転送する。ラウンドロビン方式はパケットごとのラウンドロビン方式とビットごとのラウンドロビン方式があるが、ビットごとのほうがより公平である。

WFQ はビットごとのラウンドロビン方式を使用し、ラウンドロビン順のキューを処理するのに、送信パケットが入っている個々のキューから 1 ビットずつ送信する。その際帯域を m ビット (m はキューのサイズ (n) より大きい) のサイクルに分割し、帯域をより多く必要とするホストに多めのビットを割り当てる。

WFQ は遅延の最大値を決定することが可能で、シェーピングアルゴリズムと組み合わせると、ばらつきの多いトラフィック発生源に対しても遅延の限界を保証することができる。しかし、WFQ は実装にかなりのコストがかかり、また遅延の上限が大きすぎるのが問題点である。

2.3 RSVP

RSVP(Resource ReSerVation Protocol) は、WAN において 1 対 n 通信において安定した遅延とスループットを提供するプロトコルで、データの送受信の際のスループットや遅延の安定性などを要求する実時間アプリケーションに対し、効率的なデータ転送サービスを提供するものである。RSVP は IPv6 と共に資源予約を行う。また、マルチキャストにも対応しており、マルチキャスト上のすべての経路上のルータに対して帯域予約ができる。しかも、それぞれの経路で回線速度が異なっても、異なる帯域幅を予約して通信することが可能である。資源予約に関する大きな特徴として、予約資源は受信者主導で行われることがあげられる。

3 ネットワーク応答特性プロトコル

ネットワーク応答特性プロトコル NPP は、インターネット環境において実時間制約を考慮するような高度な転送サービスを提供するプロトコルで、ベストエフォート型ソフトタイムクリティカル通信サービスという位置付けである。NPP

の特徴は、資源予約は行わず、NPP パケットスケジューリングと呼ばれるスケジューリング方式で、これは各パケットに与えられた優先度とタイムウィンドウの値を使用してスケジューリングを行い、ソフトな実時間パケット転送サービスを提供する。

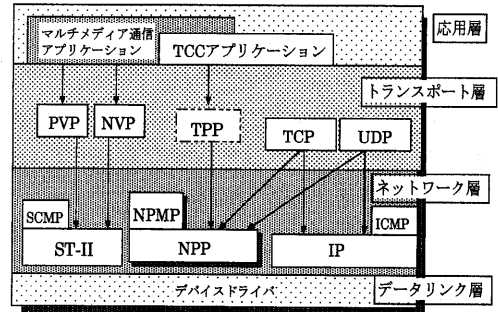


図 1: NPP プロトコル階層

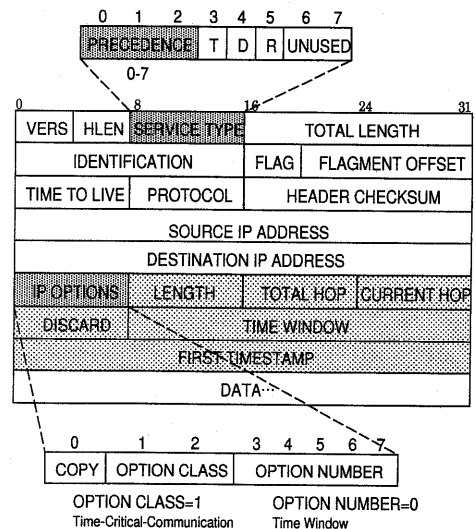


図 2: NPP ヘッドフォーマット

プロトコル階層は NPP は IP と同じネットワーク層のプロトコルであり、各種ネットワークを相互接続するコネクションレスのプロトコルである。しかしながら、タイムクリティカルなデータグラムを扱う点が IP と異なる点である (図 1)。ネットワーク応答特性獲得機構は、ネットワーク

応答特性モニタリングプロトコル (NPMP: Network Performance Monitoring Protocol) により実現される。また TPP (TransPort Protocol for NPP) は、NPP の上位層のプロトコルであり、TCC アプリケーションに対するインターフェースと、タイムクリティカルメッセージが損失した場合の処理 (再送など) を行うものである。

以下にアルゴリズムの概要を示す。

NPP のパケットスケジューリングアルゴリズムは、

『優先度が最も高いパケットを選択する。同じ優先度のパケットが複数あれば、時間制約 (タイムウィンドウ) をもつパケットを優先する。時間制約を持つパケットが複数ある場合は、始点ホストではその中で最初にバッファに入ったパケットを、終点ホスト、ルータでは時間制約の厳しいパケットを選択する。』

で定義され、CURRENT HOP フィールド、TOTAL HOP フィールドと FIRST TIMESTAMP フィールドを用いて、時間制約の厳しいパケットを選択する。選択の基準は D/T を時間制約の厳しさと定義し、この値が小さいならばそれは時間制約が厳しいことを意味する。以下に D と T を求める計算式を示す。

$$T = \frac{NOW - FIRST\ TIME\ STAMP}{TIME\ WINDOW}$$

$$D = \frac{CURRENT\ HOP + 1}{TOTAL\ HOP + 1}$$

さらにこのパケットスケジューリングアルゴリズムに対し、スタベーション回避、重要パケットのロス回避、不要パケットの廃棄という機能が付け加えられている。

3.1 NPP キューモデルの変更

NPP キューはレベル 0 からレベル 7 までの優先度ごとにキューを設けている。NPP のパケットスケジューリングアルゴリズムでは、優先度の高いキューから順に送信するパケットを選択する。しかしながら、ユーザが送信するパケットを全てレベル 7 に設定する恐れがある。実際優先度を設定するのはユーザ自身で行うため、ネットワーク資源の有効利用はユーザの良識に強く依

存する。そのため全てのユーザがパケットの優先度をレベル 7 に設定し送信すると、NPP のパケットスケジューリングアルゴリズムはあまり機能しない。このことを未然に防ぐため、フェア・キューイング (Fair Queuing) [5] (図 3) 方式を組み合わせることを考える。

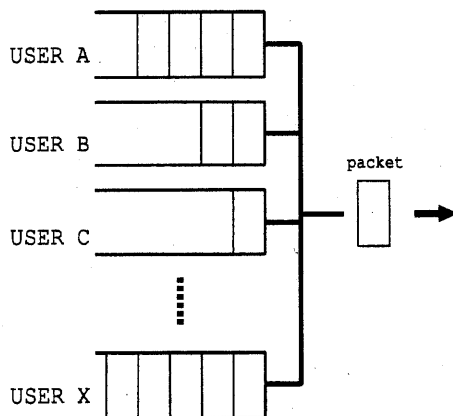


図 3: フェア・キューイング

フェア・キューイングの基本的なアイデアは、キュー資源を複数のフローに対し公平に割り当て、パケットをフローに割り当てられたキューに分類することである。具体的に説明すると、ルータのキュー資源を発信ホスト別に分類して管理し、そのキューの最初のパケットをラウンドロビン方式で送信するというものである。

以下に NPP キューイングモデルとフェア・キューイング方式を組み合わせたモデルを示す (図 4)。

パケットはルータに到着すると、各出力インターフェースごとのキューに入れられるが、このとき発信ホストごとのキューに分類する。発信ホストの識別には発信ホストの IP アドレスを利用する。この時のアルゴリズムは以下の通りである。

ラウンドロビン方式により対象となるユーザを選択しパケットの有無を調べる。選ばれたユーザのキュー内にパケットがあれば、NPP パケットスケジューリングを行い、転送するパケットを決定する。もし、ユーザのキューが空であれば次のユーザが選択される。またルータは、到着した

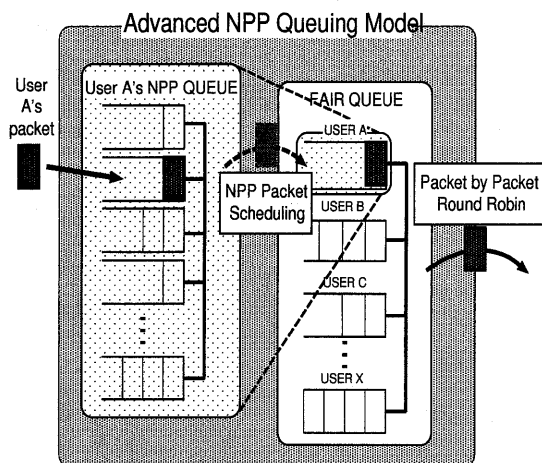


図 4: Advanced NPP Queuing Model

パケットの IP アドレスを調べ、一致したユーザのキューに入れる。もし一致しなかった場合は、新たにそのユーザのためのキューを作成する。パケットを転送したあとキュー内が空であればそのユーザのキューを削除する。

このフェア・キューイング方式の利点は、各発信ホストのトラフィックが自分以外のトラフィックに影響を及ぼさないことである。これにより変更後のキューは、発信ホストごとにキューが設けられているため、ユーザが送信するパケットのレベルをすべてレベル 7 に設定しても他のユーザに迷惑をかけることはない。従ってユーザは送信するパケットを適切なレベルに設定せざるを得なくなる。

しかしながら、フェアキューイング方式には以下のような問題点がある。

1. パケットの長さに関して考慮していない。
2. 全てのユーザに対し公平な割り当てを行うことについての疑問。

NPP パケットスケジューリングの目的は、できるだけ速くパケットを送信することである。フェア・キューイング方式を組み合わせることにより、ルーティング時の処理が複雑化し NPP の性能は低下するが、優先度の設定は必然的に適切に

設定せざるを得ないため、優先度の乱用は防ぐことができる。また、性能低下を最小限に押えるように処理が複雑にならないようなアルゴリズムを選択する必要がある。

フェア・キューイングを用いることは、すべてのホストに対して公平な割り当てを行うことを意味する。しかしながら、例えば、単に WWW でネットサーフィンを楽しんでいるユーザと、インターネットを利用してロボットを遠隔操作を行うユーザがいる場合、かれらに対して同等のサービスを提供するというのは公平ではない。

この問題を回避するために、クラス別にフェア・キューイングを設けるようにする。ユーザの要求に応じてパケットをそれぞれ決められたクラスに分類する。例えば、低遅延を望むユーザであれば、パケットを他のクラスよりも優先的に送信するクラスに入れる。しかし、クラス分けをどのように行うかが問題点として残る。従って今回はフェア・キューイング方式を組み合わせるのにとどめ、残りの部分は今後の課題とする。

4 広域網における NPP の実装

これまでいくつかの未実装の部分があったが、新たにタイムウィンドウの指定とルータの機能を BSD/OS2.1 上に実装した。

4.1 ルータ機能の実現

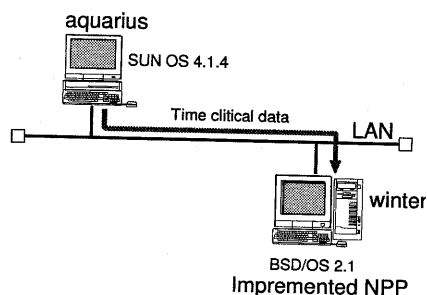


図 5: 動作テスト環境

まず NPP にタイムウィンドウを持ったパケットも扱えるようにパケットスケジューリングを変更し、その動作も確認した。タイムウィンドウの指定はアプリケーションから行い、IP オプショ

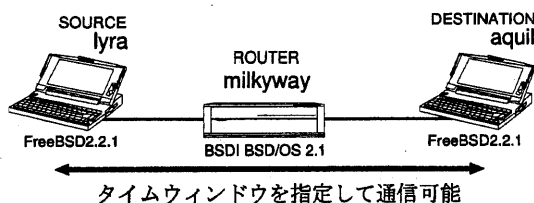


図 6: 実験環境

ンの設定を行う `setsockopt()` を利用した。動作テストは、lyra と aquil の 2 点間でタイムクリティカル通信を行い機能の動作テストを行った。時計合わせは NTP を使い、パケットスケジューリング機能が動作しているのを確認した (図 5)。

つぎに NPP に対してルータの機能を実装した。パケットはルータに到着すると、IP ソフトウェアに渡される。ルータ宛のパケットならば、上位層に渡され、そうでなければパケットを経路づけし転送を行う。その際パケットがタイムウィンドウを指定したパケットであれば IP オプションフィールドの CURRENT HOP フィールドの値を更新する。ルータ機能の動作確認のために図 6 のようなサブネットを構築した。

通信は lyra と aquil でルータ越しで行い、タイムクリティカルデータを送信しルータ機能の動作を確認した。

5 おわりに

本稿では、タイムウィンドウ要求を持つデータ転送に対して、インターネット環境において実時間転送サービスを提供するプロトコルであるネットワーク応答特性プロトコル (NPP) のキューモデルの変更、及びルータ機能を BSD/OS2.1 に実装しその動作を確認したことについて述べた。

NPP ではユーザがパケットを優先度レベル 7 設定して送信を行うことが問題であったが、NPP キューモデルにフェア・キューイングモデルを組み合わせるによりこれを解決した。また NPP パケットスケジューリングでタイムクリティカルデータを扱うよう実装を変更し、ルータの機能を合わせて実装を行い、サブネットを構築してその環境内で動作確認を行った。

今後の課題としては、本稿ではルータとしての機能を実装しテストを行い動作を確認したことしかふれていないが、現在複数のルータも視野にいられた環境で実験を行なっている。また、NPP キューモデルの変更にともない NPP パケットスケジューリングのアルゴリズムが複雑になると予想されるため、これを実装して実験を行い、変更前の NPP との比較を行う必要がある。またフェア・キューイングの追加にともない、そのアルゴリズムについても高速なアルゴリズムを検討することが重要である。

参考文献

- [1] 河井 保博: マルチメディア化進む TCP/IP 帯域管理用の新プロトコル登場, 日経コミュニケーション, 日経 BP 社, pp.71-76(1995 11.6)
- [2] W.T.Strayer, B.J.Dempsey, A.C.Weaver 著, (株) 日本ヒューレットパカード訳: マルチメディア時代のプロトコル XTP, (株) トップラン (1995.9)
- [3] 太田, 坂下, 水野: ネットワーク応答特性プロトコルの提案とマルチメディア通信システムへの適用, 情処研報, Vol.95, No.61, pp.43-48(1995.7).
- [4] Craig Partridge, "Gigabit Networking", Addison-Wesley, 1994.
- [5] クリスチャン・ウイテマ 著, 村井 順 監修, WIDE プロジェクト IPv6 分科会 監訳, 松島英樹 訳: IPv6 次世代インターネット・プロトコル, プレンスティスホール (株)(1997.2)
- [6] David D.Clark, Scott Shenker, and Lixia Zhang, "Supporting Real-Time Application in an Integrated Service Packet Network: Architecture and Mechanism", Proc.ACM SIGCOMM '92, pp.14-26, 1992.
- [7] S.J.Leffler, M.K.McKusick, M.J.Karels, J.S.Quartermann 著, 中村, 相田, 計, 小池 訳: UNIX 4.3BSD の設計と実装, 丸善 (株)(1995.4)