

ルータのバッファサイズがTCPの性能に与える影響についての一考察

富岡 健史[†] 長谷川 剛^{††} 村田 正幸^{†††}

[†] 大阪大学 基礎工学部情報科学科 〒 560-8531 大阪府豊中市待兼山町 1-3

^{††} 大阪大学 サイバーメディアセンター 〒 560-8531 大阪府豊中市待兼山町 1-32

^{†††} 大阪大学 大学院情報科学研究科 〒 565-0871 大阪府吹田市山田丘 1-5

E-mail: [†]tomioka@ics.es.osaka-u.ac.jp, ^{††}hasegawa@cmc.osaka-u.ac.jp, ^{†††}murata@ist.osaka-u.ac.jp

あらまし 現在、インターネットルータのバッファサイズの決定のためには、帯域遅延積を指標とする方法が広く利用されている。これに対し、ルータを通過するフローが多く存在する場合、ネットワークリンクの利用率を維持するには帯域遅延積をフロー数の平方根で除算したサイズで十分であるという主張が提起されている。しかしこの主張には、リンク利用率以外の視点からの評価が行われていない、評価を行う際のパラメータが限られている等の問題が存在する。そこで本報告では、ns-2を用いたシミュレーションにより、ルータのバッファサイズがTCPの性能に与える影響を考察する。その結果、フロー数が十分に大きければ、バッファサイズを小さくしてもリンク利用率を高く維持できるが、帯域遅延積を用いた場合と比較すると、パケット廃棄率に関して性能が劣化していること、50 - 100Kbyte程度の転送データサイズを境としてデータ転送遅延時間が増加し、転送データサイズが小さいフローが多数を占めるネットワークや、エンド端末間の伝播遅延時間が小さいネットワーク以外においては、バッファサイズを小さくすることによる悪影響が現れることが明らかになった。

キーワード ルータ、バッファサイズ、TCP (Transmission Control Protocol)、リンク帯域、伝播遅延時間

Effect of the router's buffer size on TCP performance

Takshi TOMIOKA[†], Go HASEGAWA^{††}, and Masayuki MURATA^{†††}

[†] School of Engineering Science, Osaka University 1-32 Machikaneyama, Toyonaka, Osaka, 560-0043, Japan

^{††} Cybermedia Center, Osaka University 1-32 Machikaneyama, Toyonaka, Osaka, 560-0043, Japan

^{†††} Graduate School of Information Science and Technology, Osaka University 1-32

Yamadaoka, Suita, Osaka, 565-0871, Japan

E-mail: [†]tomioka@ics.es.osaka-u.ac.jp, ^{††}hasegawa@cmc.osaka-u.ac.jp, ^{†††}murata@ist.osaka-u.ac.jp

Abstract The size of router buffers are traditionally determined by the *bandwidth-delay product rule*, which means a product of the link bandwidth and the average round trip time (RTT) of flows passing through the router. On the other hand, the recent research results have revealed that when the number of flows is sufficiently large, the buffer size can be decreased to the bandwidth-delay product divided by the square-root of the number of flows (*sqrt-N rule*), without introducing the under-utilization of the network link. In this paper, we compare the performance of the above two rules for the router buffer size, focusing on the performance of TCP connections traversing the router. Through extensive simulation experiments, we show that it would degrade the TCP performance in terms of packet loss ratio and file transmission delay, and the sqrt-N rule would enhance the network performance only when the transferring file size is between 50-100 Kbytes or when the propagation delay between the sender and the receiver hosts is significantly small.

Key words router, buffer size, TCP, bandwidth, bandwidth delay product

1. はじめに

TCP (Transmission Control Protocol) [1] は現在インター

ネットで広く利用されており、そのトラヒックは現在のインターネットトラヒックの大部分を占める [2]。そのため、ネットワーク上の TCP トラヒックの性能評価を行うことは非常に重

要である。TCP によるデータ転送の性能は、パケット廃棄率やラウンドトリップ時間 (RTT) によって大きく影響を受ける [3]。現在のインターネットルータの多くは FIFO 規律に基づくバッファを持つが、FIFO 規律に基づくバッファを用いている場合において、ネットワークのパケット廃棄率やデータ転送遅延時間を決定する大きな要素の一つとして、ルータのバッファサイズが挙げられる。リンク帯域を越えるレートのパケット送信がそのリンクに発生した場合、パケットは一時的にそのリンクの出力バッファ内に到着順に蓄積される。バッファが一杯のときに新たなパケットが到着すると、そのパケットは廃棄される。そのため、バッファサイズを大きくすることによりパケット廃棄が発生する確率は小さくなる。しかし、バッファ内に大量のデータが蓄積されると、パケットがバッファに到着してから、処理を受けてリンクへ出力されるまでの時間、すなわちキューイング遅延も増大する。[4] によると、リンク利用率を高い水準で維持しつつ、キューイング遅延を最小にするためのボトルネックリンクのバッファサイズは、リンク帯域 C とそのルータを通過する TCP コネクション (フロー) の平均 RTT (以下 \overline{RTT} と表記する) を用いて $B_n = C \times \overline{RTT}$ であるとされている。このバッファサイズの決定方法は帯域遅延積として知られており、現在多くのルータがこの方法に従ったサイズのバッファを装備している。[5] にも、ルータは一般に帯域遅延積で決定されるサイズのバッファを持つ、と記されている。

しかし [6] によれば、この方法に基づくサイズのバッファを持つルータを構築することは困難になっているという。現在のバックボーンネットワークでは 2.5 Gbps や 10 Gbps といった帯域幅を持つリンクに、10,000 を超えるフローが流れている [7]。10 Gbps の帯域を持つリンクに流れるフローの平均 RTT が 250 ms であれば、必要とされるバッファサイズは $10 \text{ Gbps} \times 250 \text{ ms} = 2.5 \text{ Gbits}$ となる。ルータのバッファに利用されるメモリデバイスは SRAM と DRAM の二種類が存在するが、現在商用に用いられている SRAM チップの最大のサイズは 36 Mbits であり、必要なバッファサイズを確保するためには多数の SRAM チップを並列的に用いて構築する必要がある。たとえば 2.5 Gbits のバッファを構築するためには、およそ 70 個の SRAM チップが必要となる。しかし SRAM は DRAM に比べて非常に高価であるため、SRAM を用いてバッファを構築することは経済的に困難である。また、必要とされる設置面積が増大すること、発熱量が大きくなることなど物理的にも多くの問題がある。一方 DRAM は SRAM よりも安価であり、最大 1 Gbits のサイズを持つが、アクセス時間が 50 ns と大きいということ、使用するにつれて 7%/年の割合でアクセス時間が低下することなどから、やはり巨大なバッファを構成することは難しい。これらのことから、10 Gbps を超えるリンクにおいて、帯域遅延積に相当するサイズのバッファを装備することは難しい、と [6] は述べている。

さらに [6] ではその解決方法として、リンクを通過するフロー数が十分に多い場合 (主張では 500 程度以上) には、バッファサイズは帯域遅延積に対しリンクを通過するフロー数 N の平

方根で除算した値、すなわち $B_s = \frac{C \times \overline{RTT}}{\sqrt{N}}$ で十分であるということを示している。[6] で示されている結果によると、この方法でバッファサイズを設定した場合にもリンク利用率はほとんど低下しない。また、[8] では、アクセスリンクの帯域が非常に小さい場合 (1 Gbps のボトルネックリンク帯域に対して 10 Mbps 程度のアクセスリンク帯域を持つ場合) や、ペーシング TCP (TCP のデータパケットがバースト的に送出されるのを防止した手法) を利用した場合には、さらに小さい数十パケットのバッファサイズで十分であるとされている。

しかし、[6,8] とともに、性能指標としてボトルネックリンクの利用率のみに着目している。リンク利用率を保つことはネットワーク管理者にとっては重要であるが、ネットワークを利用するユーザにとって重要な性能指標である TCP の性能 (フローのパケット廃棄率やデータ転送遅延時間) に関する評価も必要であると考えられる。[6,8] におけるシミュレーションおよび実験の結果はリンク利用率が維持されていることを示すものがほとんどであり、用いられているネットワークに関するパラメータも限定されている。また [6] では、転送データサイズが 14 パケットの場合についてのみ、データ転送遅延時間に関するシミュレーション結果が示されており、バッファサイズを小さくすることによってデータ転送遅延時間が小さくなるとされている。しかし、パケット廃棄率については、TCP の性質から数 RTT に一度しか廃棄が発生しないため影響は少ないのではないか、という見解のみが述べられている。また [8] で示された結果から、TCP がデータを送信する際のバースト性が軽減されるため、パケット廃棄率、データ転送遅延時間が減少すると考えられるが、アクセスネットワークのリンク帯域は急速に増大している上に、現在ペーシング TCP は普及しておらず、今後も急速に普及することは期待できない。また、ペーシング TCP が普及するとしても、普及の途中段階で通常の TCP と混在することが考えられるが、そのような混在環境におけるペーシング TCP の性能や既存 TCP との公平性については検討されていない。

そこで本稿では、TCP-Reno が利用される環境の下で、ルータのバッファサイズがネットワークのリンク利用率やパケット廃棄率、および個々のフローのパケット廃棄率やデータ転送時間に与える影響を、計算機上のシミュレーションによって評価する。評価に際してはボトルネックリンクの伝播遅延時間やリンク帯域、データ転送を行うホスト数などをパラメータとしてシミュレーションを行い、[6] で提案された手法でバッファサイズを決定した場合と、帯域遅延積を用いて決定した場合を比較することで、バッファサイズが TCP の性能にどのような影響を与えるかを明らかにする。

本稿の構成は以下の通りである。2 章で両バッファサイズの決定方式について詳細を示す。次に 3 章でシミュレーションに用いたネットワークの構成と各種パラメータについて説明を行う。4 章ではシミュレーション結果を示し、考察を行う。最後に 5 章でまとめと今後の課題を示す。

2. ルータのバッファサイズの決定方法

2.1 帯域遅延積によるバッファサイズの決定

帯域遅延積によるバッファサイズの決定は [4] で提唱され、現在広く用いられている。本稿では以降この方法を normal 方式と称する。図 1 は一本の TCP フローが輻輳回避フェーズでデータ転送を行っている場合の、TCP フローの輻輳ウィンドウサイズとバッファ内パケット数の変化を示している。ここで B_{max} はバッファサイズをあらわす。時刻 t_1 から TCP のパケット送信レートがリンク帯域を超え、リンク帯域を越えて到着したパケットがバッファ内に蓄積され始める。TCP は ACK パケットを受け取るたびに輻輳ウィンドウサイズを増加させるため、最終的にはバッファあふれが発生しパケットが廃棄される（時刻 t_2 ）。パケットが廃棄されてから約 $1RTT$ 後には、送信側 TCP が重複 ACK を受信することによってパケット廃棄を検知し、Fast Retransmit 機構によりそのパケットを再送すると同時に、輻輳ウィンドウサイズを廃棄が発生したときの値 W_{max} から $W_{max}/2$ に減少させる（時刻 t_3 ）。この結果、送信側 TCP から送出されるパケット数が減少し、ルータへのパケット到着レートがリンク帯域を下回るため、バッファ内に蓄積されていたパケット数は減少する（時刻 $t_3 - t_5$ ）。その後、送信側 TCP は輻輳ウィンドウサイズを再び増加させるため、時刻 t_5 からバッファ内のパケット数は再び増加に転じる。

このとき、バッファ内のパケット数が時刻 t_5 に達する前に 0 になると、ルータへのパケットの流入量がリンク帯域を下回りかつバッファ内にパケットが存在しない状態、すなわち送信すべきデータが存在するにもかかわらずリンク帯域が 100% 利用されていない状態が生じる。

リンク帯域 C のボトルネックリンクにおいてルータのバッファが空にならない場合、送信ノードへの ACK パケット到着レートは C で表すことができる。一方、輻輳ウィンドウサイズ W_{max} から $W_{max}/2$ に減少した場合、送信側 TCP は $W_{max}/2$ 個の ACK パケットが帰ってくるまで、新しいパケットを送信することができない。送信ノードに $W_{max}/2$ 個の ACK パケットが到着するまでにかかる時間は $(W_{max}/2)/C$ で表すことができる。また、バッファサイズ B_{max} のバッファが空になるまでには B_{max}/C の時間がかかるために、 $B_{max}/C < (W_{max}/2)/C$ であればバッファが空になる。逆に $B_{max}/C \geq (W_{max}/2)/C$ であればバッファが空になる前に送信側からのパケット送信が再開されるため、リンク利用率を維持することができる。すなわち、バッファが空になる前に新しいパケットが送信側 TCP から到着するためには、次式で表されるバッファサイズが必要となる。

$$B_{max} \geq W_{max}/2 \quad (1)$$

また、リンク上に存在するデータの総量は \overline{RTT} と C を用いて $C \times \overline{RTT}$ として表すことができるが、輻輳ウィンドウサイズの値が、半減した後も $C \times \overline{RTT}$ 以上であれば、リンクに常にデータが流れるためにリンク利用率を維持することができる。すなわち、リンク利用率を維持するためには、

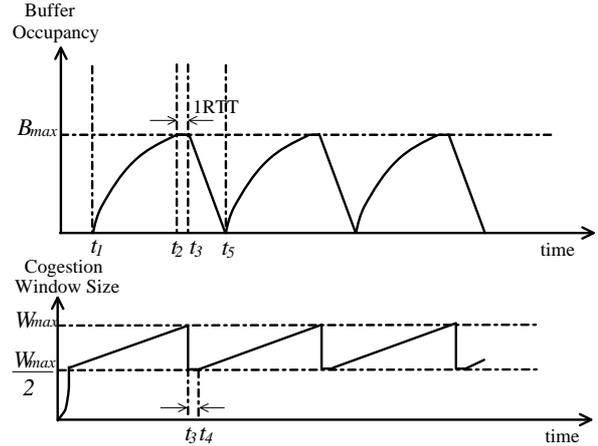


図 1 一本の TCP フローが流れている場合の、輻輳ウィンドウサイズ及びバッファ内パケット数の変化

$$W_{max}/2 = C \times \overline{RTT} \quad (2)$$

であればよい。(1)(2) 式から

$$B_{max} \geq W_{max}/2 = C \times \overline{RTT} \quad (3)$$

となる。すなわち、1本の TCP フローが存在する場合にリンク利用率を維持するためには、 $C \times \overline{RTT}$ のサイズのバッファが必要であるといえる。

また、TCP フローが複数本存在する場合にも、それらが同期して動作する、すなわち、バッファあふれが発生したときにすべてのフローのパケットが等しく廃棄される場合には、必要となるバッファサイズは変化しない。これは、TCP フローが N 本存在する場合には、個々のフローは $1/N$ の帯域を持つリンクに 1本のフローを流したときと同様の挙動を示すため、 $C/N \times \overline{RTT}$ のバッファが N 本分必要と見なすことができ、全体として必要となるバッファサイズは、 $(C/N \times \overline{RTT}) \times N = C \times \overline{RTT}$ となるためである。

2.2 [6] で提案された手法によるバッファサイズの決定

この方式は、帯域遅延積をリンクを通過するフロー数の平方根で除算したものをバッファサイズとするものである。本稿ではこの方式を sqrtN 方式と称する。以下では、[6] で示された sqrtN 方式の根拠の概要を示す。詳細については [6] を参照されたい。

TCP の輻輳ウィンドウサイズは、ボトルネックリンクを通過するフロー数が 100 以下の場合には、同期して変動することが知られている。すなわち、あるフローの輻輳ウィンドウサイズが半減するときには、他のフローの輻輳ウィンドウサイズも半減する。これは、ルータのバッファあふれによってパケット廃棄が発生する際に、通過するフローのパケットが等しく廃棄されることによる。しかし、フロー数が 500 を超えるような場合は、輻輳ウィンドウサイズが同期しない。これは、バッファあふれによるパケット廃棄が全フローに同期して発生しないことを意味する。このとき、ルータを通過する TCP コネクションの輻輳ウィンドウサイズの分布はほぼ正規分布にしたがう。

N 本のフローが通過するルータの時刻 t におけるキュー長 (バッファ内のパケット数) $Q(t)$ は、その時の各フローの輻輳ウィンドウサイズ $W_i(t)$ を用いて

$$Q(t) = \max(0, \sum_{i=1}^N W_i(t) - (\overline{RTT} \times C))$$

と表すことができる。よって、 $\sum_{i=1}^N \overline{W_i(t)} = \overline{W}$ を用いて、ルータにおける平均のキュー長 \overline{Q} は

$$\overline{Q} = \max(0, \overline{W} - (\overline{RTT} \times C)) \quad (4)$$

となる。すなわち $\overline{Q} > 0$ のとき、各フローの輻輳ウィンドウサイズの平均値 $\overline{W_i}$ は (4) 式を用いて

$$\overline{W_i} = \overline{W}/N = \frac{\overline{RTT} \times C + \overline{Q}}{N} \leq \frac{\overline{RTT} \times C + B_{max}}{N} \quad (5)$$

と求めることができる [6]。また TCP の輻輳ウィンドウサイズの平均値 $\overline{W_i}$ に対する標準偏差 σ_{w_i} は、TCP フローのウィンドウサイズの変動をモデル化することによって、

$$\sigma_{w_i} = \frac{1}{3\sqrt{3}} \overline{W_i} \quad (6)$$

となり、正規分布の性質より、フロー数が十分に多い場合は、輻輳ウィンドウサイズの合計値の標準偏差 σ_w が

$$\sigma_w \leq \sqrt{n} \sigma_{w_i} \quad (7)$$

となることから、ルータにおけるキュー長 Q の標準偏差 σ_Q は、(5)-(7) 式より

$$\sigma_Q = \sigma_w \leq \frac{1}{3\sqrt{3}} \frac{\overline{RTT} \times C + \overline{Q}}{\sqrt{n}} \leq \frac{\overline{RTT} \times C + B_{max}}{\sqrt{n}} \quad (8)$$

と表すことができる。これらのことから、リンク利用率 $Util$ について次式を導くことができる。

$$Util \geq \text{erf} \left(\frac{3\sqrt{3}}{2\sqrt{2}} \frac{B_{max}}{\overline{RTT} \times C + B_{max}} \right) \quad (9)$$

$B_{max} = \frac{\overline{RTT} \times C}{\sqrt{n}}$ であれば $Util \geq \text{erf} \left(\frac{3\sqrt{3}}{2\sqrt{2}} \right) \simeq 0.9899$ となるため、フロー数が十分に多い場合は $B_s = \frac{\overline{RTT} \times C}{\sqrt{n}}$ で十分である。

また [6] では、スロースタートフェーズ中にデータ転送が終わるような転送データサイズの小さいフロー (short-lived フロー) が存在する場合においても、必要となるバッファサイズは short-lived フローにはほとんど影響を受けず、 $B_s = \frac{\overline{RTT} \times C}{\sqrt{n}}$ で導かれるバッファサイズよい、とされている。

3. シミュレーション環境

3.1 ネットワークとトラフィックモデル

シミュレーションでは図 2 のネットワークモデルを用いる。モデルは送信側ホスト $S_1 \sim S_n$ 、受信側ホスト $R_1 \sim R_n$ と両者をつなぐアクセスリンク、ボトルネックリンク、ルータで構成される。ボトルネックリンクの伝播遅延時間および帯域をそれぞれ D [ms] および C [Mbps] とし、アクセスリンクの帯域は

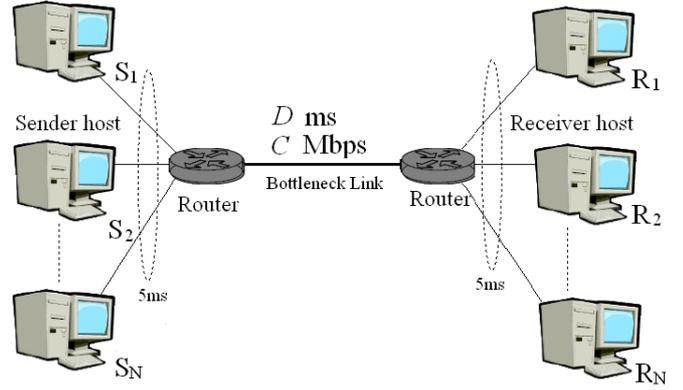


図 2 シミュレーションで用いるネットワークモデル

要素	生成関数	パラメータ
サイズ - Body	$p(x) = \frac{e^{-(\ln x - \mu)^2 / 2\sigma^2}}{x\sigma\sqrt{2\pi}}$	$\mu = 9.357; \sigma = 1.318$
サイズ - Tail	$p(x) = \alpha k^\alpha x^{-\alpha+1}$	$k = 133K; \alpha = 1.1$
生成間隔	$p(x) = \alpha k^\alpha x^{-\alpha+1}$	$k = 1; \alpha = 1.5$

表 1 SURGE モデル [9]

ボトルネックリンクと等しいものとする。シミュレーションにおいては、ホスト数 N を変化させ、ボトルネックリンクを通過するフロー数を変動させる。また、 D 、 C についても変化させ、各条件の下でバッファサイズが TCP の性能に与える影響を評価する。各ホストから送られるトラフィックには、各ホストが無限大の転送データサイズをもつ P2P トラフィック、および [9] で述べられている SURGE モデル (表 1) にしたがって、各ホストの転送データサイズや転送間隔などが決定される Web トラフィックを用いる。以降、明記していない場合は Web トラフィックを使用する。シミュレーションには ns-2 [10] を用いる。

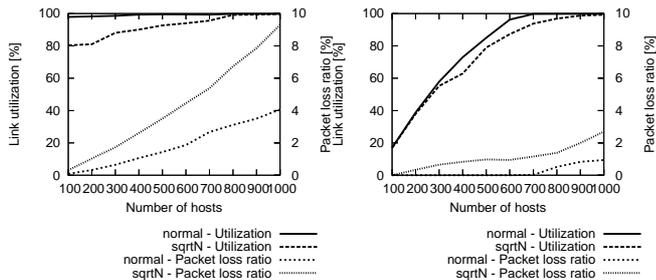
3.2 シミュレーション結果の評価方法

ボトルネックルータのリンクにおいて、通過するパケットの挙動を観測し、単位時間当たり通過したパケット数からリンク利用率を求める。また、ルータに到着したパケット数とバッファあふれによって廃棄されたパケット数から、パケット廃棄率を導出する。Web トラフィックを用いたシミュレーションの場合には、以下に示す各フローの転送に関する性能指標を用いる。データ転送遅延時間は、各フローが転送するデータの最初のパケットの送信が開始されてから、最後のパケットに対する ACK を受信するまでの時間とし、各ホストのパケット送受信のログより求める。また、転送したデータサイズを記録し、データ転送が終了するまでに送信したパケット数との比較から、各転送において廃棄されたパケット数を求め、各データサイズにおけるパケット廃棄率を導出する。このようにして得られたリンク利用率、パケット廃棄率、各サイズにおけるパケット廃棄率、およびデータ転送遅延時間を性能評価指数として用いる。

4. シミュレーション結果と考察

4.1 リンク利用率とパケット廃棄率

図 3 は $C=100$ Mbps、 $D=90$ ms のネットワークに P2P ト



(a) P2P トラフィックの場合 (b) Web トラフィックの場合

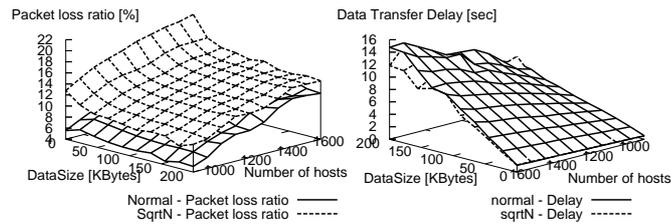
図 3 ネットワークのリンク利用率とパケット廃棄率

ラフィックおよび Web トラフィックをそれぞれ与えた環境において、ホスト数を変化させた場合のリンク利用率とパケット廃棄率の変化を示したものである。図 3(a) から、normal 方式でバッファサイズを決定した場合は、ホスト数が少ない場合においても高いリンク利用率を示しているが、SqrtN 方式を用いた場合は、ホスト数が小さい場合にリンク利用率が低下していることがわかる。また、図 3(b) から、normal 方式ではホスト数が 700 程度より小さい場合にはパケット廃棄が発生していないのに対し、SqrtN 方式ではホスト数に関わらず廃棄が発生していることが確認できる。これは、sqrtN 方式を用いるとバッファサイズが小さくなるため、パケットがバースト的に流入した時にバッファで吸収しきれずにバッファあふれが発生するためである。しかし、トラフィックの種類に関わらず、ホスト数が多い場合にはリンク利用率はほぼ 100% であることから、[6] における主張が正しいことが裏づけられた。

4.2 転送データサイズごとのパケット廃棄率と転送遅延時間

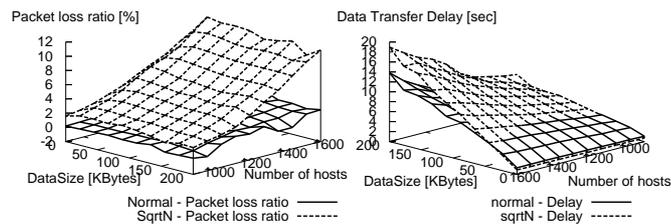
4.2.1 ホスト数と与える影響

図 4 および図 5 は、それぞれ $C=100$ Mbps、 $D=20$ ms および $C=100$ Mbps、 $D=90$ ms の場合においてホスト数を変化させたときの、伝送データサイズに対するパケット廃棄率およびデータ転送遅延時間の変化を示したものである。図 4(a) および図 5(a) から、sqrtN 方式を用いた場合は、転送データサイズに関係なくパケット廃棄率が normal 方式にくらべて高いことがわかる。特に図 5(a) において、ホスト数が少ない場合には小さかった両方式間のパケット廃棄率の差が、ホスト数の差が広がると共に大きく開いていっている。図 4(a) においては、ホスト数にかかわらず小さい転送データサイズの場合にパケット廃棄が多く発生している。このことから、sqrtN 方式を利用した場合、バースト的に発生したパケットがバッファからあふれて廃棄されているということがわかる。データ転送遅延時間（図 4(b)、図 5(b)）においては、ホスト数の増加とデータ転送遅延時間の増加はほぼ比例している。これは normal 方式、sqrtN 方式のいずれにおいても同じであり、両方式間のホスト数の変化による差異は認めがたい。図 5(b) ではホスト数が増加するに従い、データ転送遅延時間の差が大きくなっているが、これはパケット廃棄率の差が増大することに起因すると考えられる。対して図 4(b) ではホスト数の変化によるパケット廃棄率の差が変化していないため、データ転送遅延時間の差にも大きな変



(a) パケット廃棄率 (b) データ転送遅延時間

図 4 $C = 100$ Mbps、 $D = 20$ ms



(a) パケット廃棄率 (b) データ転送遅延時間

図 5 $C = 100$ Mbps、 $D = 90$ ms

化は見られない。 D が小さい場合や転送データサイズが小さい場合には、sqrtN 方式を利用したほうがデータ転送遅延時間が短くなっているが、これについては次節で述べる。

4.2.2 ボトルネックリンクの伝播遅延時間の与える影響

図 6 は $C = 100$ Mbps、 $N = 1500$ の場合においてボトルネックリンク伝播遅延時間を変化させたときの、伝送データサイズに対するパケット廃棄率およびデータ転送遅延時間の変化を示したものである。図 6(a) から sqrtN 方式が常にパケット廃棄率が高いこと、 D が大きくなるにつれ両方式間のパケット廃棄率の差が大きくなることがわかる。図 6(b) においても、 D が大きくなるにしたがって sqrtN 方式のデータ転送遅延時間が大きくなっている。しかし、前節で述べたように、sqrtN 方式はパケット廃棄率が高いにもかかわらず、転送データサイズが小さい、あるいは D が小さい場合には、normal 方式に比べてデータ転送遅延時間が小さい。これは、sqrtN 方式を利用することでバッファサイズが小さくなった結果、キューイング遅延が減少（normal 方式に比べて、ホスト数が 900 の場合は 30 分の 1、1600 の場合は 40 分の 1）することにより、パケット廃棄率は高いものの、廃棄されたパケットの再送にかかる時間が短縮されるためであると考えられる。図 4(a) が示す通り、 D が小さい場合は両方式間でのパケット廃棄率の差が小さくなるため、廃棄されたパケットの再送にかかる時間や、輻輳ウィンドウサイズが小さくなることによる転送遅延時間の増加に差が見られなくなる。また、転送データサイズが小さい場合には輻輳ウィンドウサイズが大きくなるため、その減少がデータ転送遅延時間に与える影響は小さい。すなわち、リンク伝播遅延時間や、転送データサイズが小さい場合には、パケット廃棄率よりもキューイング遅延がデータ転送遅延時間に与える影響が

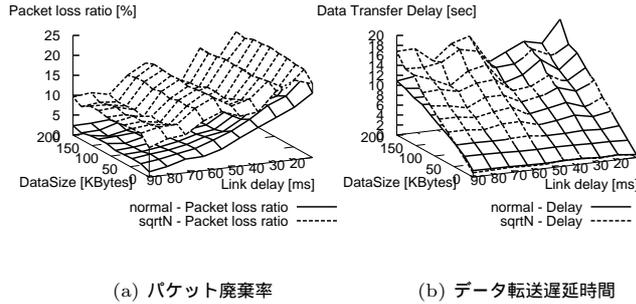


図 6 $C = 100$ Mbps、 $N = 1500$

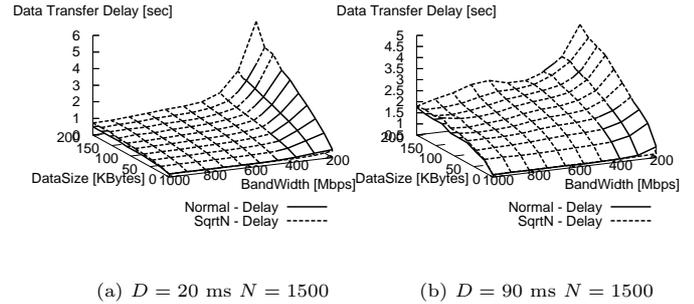


図 8 データ転送遅延時間

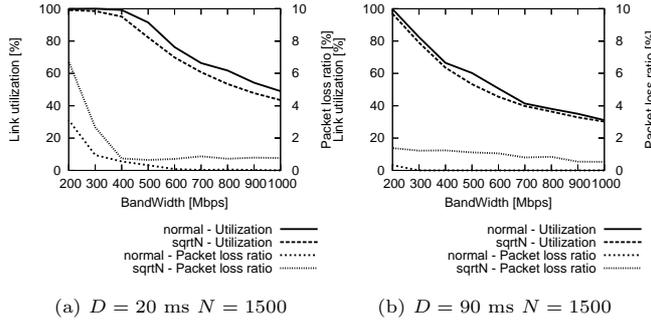


図 7 ネットワークのリンク利用率とパケット廃棄率

大きいため、sqrtN 方式が有利であるといえる。

4.2.3 ボトルネックリンク帯域の与える影響

図 7 は、それぞれ $D=20$ ms、 $N=1500$ および $D=90$ ms、 $N=1500$ の場合において、ボトルネックリンクの帯域を変化させた場合のリンク利用率とパケット廃棄率の変化を示している。また、図 8 は、伝送データサイズに対するデータ転送遅延時間の変化を示したものである。図 7 から、リンク利用率、パケット廃棄率のいずれにおいても sqrtN 方式の性能が normal 方式を下回っていることがわかる。特に、ボトルネックリンク帯域が大きい場合においても、パケット廃棄率がある値より小さくならないことが図 7(a) よりわかる。図 7(b) においても、normal 方式ではボトルネックリンク帯域が約 300 Mbps の場合にはパケット廃棄率が 0% になっているが、sqrtN 方式ではリンク帯域が 1 Gbps の場合においてもパケット廃棄が発生している。これは、sqrtN 方式はバッファサイズが小さいため、パケットがバースト的に到着した場合にそれらのパケットをバッファで吸収し切れず、パケット廃棄が発生するためである。また、パケット廃棄率が大きい場合、リンクの伝播遅延時間や転送データサイズが小さい場合においても sqrtN 方式を用いるほうが転送遅延時間が大きくなっている (図 8)。以上より、リンク帯域が大きく、平均のリンク利用率が低い場合には、sqrtN 方式はリンク利用率、パケット廃棄率、およびデータ転送時間の全てにおいて normal 方式に劣ることが明らかとなった。

5. おわりに

本稿では ns-2 によるシミュレーションを用いて、ルータのバッファサイズが TCP の性能に与える影響の評価を行なった。その結果、sqrtN 方式を用いてバッファサイズを決定した場

合は、フロー数が多く平均のリンク利用率が 100% に近い場合は、normal 方式とほぼ同等のリンク利用率が得られる半面、パケット廃棄率が高くなるために、リンクの伝播遅延時間が小さい場合や、転送データサイズが小さい場合を除いて、normal 方式に比べて TCP の性能が劣化することがわかった。さらに、ネットワーク負荷が低い場合においては、リンク帯域が大きい場合においてもパケット廃棄率が低下しないため、normal 方式に比べてリンク利用率やデータ転送性能が劣ることが明らかとなった。今後は、ペーシング TCP と TCP-Reno が混在する環境における、バッファサイズが各 TCP の性能に及ぼす影響を考察したい。また、必要なバッファサイズを決定するための重要な要素となる、TCP コネクションが同期して動作するための条件を明らかにしたい。

文 献

- [1] J. Postel, "Transmission Control Protocol," *RFC 793*, Sept. 1981.
- [2] M. Fomenkov, K. Keys, D. Moore, and k. claffy, "Longitudinal study of Internet traffic from 1998-2003," *Winter International Symposium on Information and Communication Technologies (WISICT 2004)*, Jan. 2004.
- [3] N. Cardwell, S. Savage, and T. Anderson, "Modeling TCP latency," in *Proceedings of IEEE INFOCOM*, pp. 1742-1751, Mar. 2000.
- [4] C. Villamizar and C. Song, "High performance TCP in ANSNET," *SIGCOMM Computer Communications Review*, vol. 24, pp. 45-60, Oct. 1994.
- [5] R. Bush and D. Meyer, "Some Internet architectural guidelines and philosophy," *RFC 3439*, Dec. 2003.
- [6] G. Appenzeller, I. Keslassy, and N. McKeown, "Sizing router buffers," in *Proceedings of the 2004 Conference on Applications, Technologies, Architectures, and Protocols For Computer Communications*, Sept. 2004.
- [7] C. J. Fraleigh, *Provisioning Internet backbone Networks to support latency sensitive applications*. PhD thesis, Stanford University, Department of Electrical Engineering, June 2002. Adviser-Fouad A. Tobagi.
- [8] M. Enachescu, Y. Ganjali, A. Goel, N. McKeown, and T. Roughgarden, "Part III: routers with very small buffers," *SIGCOMM Comput. Commun. Rev.*, vol. 35, pp. 83-90, July 2005.
- [9] P. Barford and M. Crovella, "Generating representative Web workloads for Network and server performance evaluation," *Measurement and Modeling of Computer Systems*, pp. 151-160, July 1998.
- [10] T. V. Project, "UCB/LBNL/VINT network simulator - ns (version 2)." available at <http://www.isi.edu/nsnam/ns/>.