

省電力化のための動的転送性能制御手法の性能評価

福田 豊¹

1 はじめに

アクセスネットワークのブロードバンド化や、動画像を中心としたリッチコンテンツの提供により、インターネット上を流通するトラフィック量は飛躍的に増加しています。例えば総務省が2010年9月に発表した「我が国のインターネットにおけるトラフィックの集計・試算」では、2010年5月時点の総ダウンロードトラフィックは推定で約1.45 Tb/sであり、1年間で約1.2倍(17.8%増)増加したと報告しています[1]。このような急激なトラフィック量の増加は、当然のことながらトラフィックを転送するネットワーク機器での消費電力の増大を招きます。例えばGuptaらは2000年の米国におけるネットワーク関連機器の消費電力は6 TW-hに達すると報告していますが[2]、Nordmanはこの電力需要が増加し続けており、米国におけるネットワーク分野の消費電力は2006年までに20TW-hまで増加したと報告しています[3]。このような傾向は今後も続くと思われ、ネットワークにおける省電力化技術の重要性が高まっています。

これまでも省電力化については数多くの研究が成されてきましたが、それらは主にネットワーク機器のチップセットなどのデバイスや、ネットワークに接続するパソコンなどの機器を対象としていました。そこでこれらの取り組みに加えて、トラフィックを転送するネットワーク自体が省電力化を実現する機構を導入できれば、全体としてより大きな省電力化を期待することができます。そこで本研究では、ネットワークのトラフィック特性を考慮に入れたネットワークスイッチの省電力化手法の開発を目指しました。

開発を始めるにあたり、まず幾つかのスイッチの消費電力を計測しました。パケット転送性能の差が電力消費に及ぼす影響を調査するために、図1に示すように、2台のスイッチ間をUTPケーブルで接続して稼働リンク数を1つずつ増加させた時の、スイッチ全体の消費電力を電力計で計測しました。計測対象のスイッチはAlaxala AX2430のペアと、Cisco Catalyst 2970のペアです。実験結果より、総消費電力はAlaxa AX2430とCisco Catalyst 2970でそれぞれ異なるものの、1ポート当たりの転送速度が1Gb/sの場合、リンク数に比例してスイッチ全体の電力消費量は2~3Wずつ増加することが分かりました。同様の結果は[4]でも報告されており、トラフィック量に基づく転送性能の動的な制御は、ネットワークの電力消費を減らすために有効な手段であることが分かりました。さらに、既存のスイッチはトラフィックの有無にかかわらず常に最大転送能力を発揮できるよう動作しているため、もしネットワークスイッチのCPUやエンジンの処理速度を、PCにおけるIntelのSpeedStepテクノロジー[5]のよう

¹九州工業大学情報科学センター fukuda@isc.kyutech.ac.jp

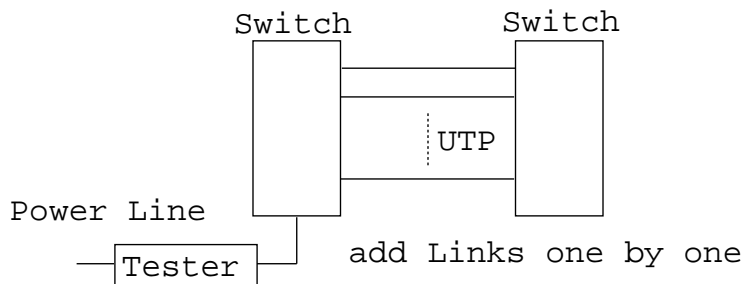


図 1: 電力計測実験 .

にトラヒックの到着レートに応じて制御することができれば，電力消費をさらに削減することができると考えられます .

そこでこの調査結果に基づいて，本研究では消費電力を削減するためにトラヒックの変化に応じてスイッチ間の転送性能を動的に制御する手法を開発しました . 具体的には，リンク集約技術を用いて接続されているスイッチを想定し，そのリンク本数を動的に制御することで，スイッチの省電力化を目指します . ここで，提案手法を現実的な省電力化手法として設計するためには，実際のネットワーク運用における要求事項を考慮に入れる必要があります . ネットワーク管理者は，第一に管理するネットワークが良好な通信品質を提供できるよう維持運用しなければなりません . よって，ネットワークの省電力化は重要な課題ではありますが，その導入による通信品質の低下は避けねばなりません .

そこで本研究では，通信品質の低下を避けつつ消費電力の削減を行うために，リンク本数の推定に到着トラヒックの平均スループットではなく，ある区間ごとのピークスループットを利用することにしました . さらに提案手法の有効性を，省電力性能とパケット廃棄などで示される通信品質に及ぼす影響の両面から評価することにしました .

以降，本論文の 2 節では，入力トラヒックのピークスループット値を用いた出力リンク本数の決定方法について説明します . 続く 3 節でシミュレーションモデルを説明し，結果と議論を 4 節に示します . ネットワークにおける省電力化の関連研究を 5 節で述べ，最後に 6 節でまとめます .

2 動的な集約リンク数制御手法

本節では，提案するスイッチの省電力化を実現するための動的なリンク制御手法について説明します . 提案手法では，トラヒックに応じてスイッチ間の転送性能を動的に制御するために，ネットワークの幹線で転送性能向上のために広く利用されているリンク集約技術を用いました . リンク集約は IEEE 802.3ad[6] で標準化されており，そこではスイッチ間で複数の物理リンクを用いて論理的な 1 本のリンクを構成するためのプロトコルとして LACP (Link Aggregation Control Protocol) を規定しています . LACP 機能を有するスイッチは，リンク障害やリンクの追加などの状態変化に従って，LACP DU (Data Unit) パケットを対向スイッチに送信し，自動的に新たな集約リンクを構成します .

しかし，既存の LACP には到着トラヒック量に応じた動的なリンク本数の設定機能は含まれていないため，LACP において図 2 に示す動作の拡張を行いました . 以下，その動作について説明します .

1. スイッチは到着トラヒックを計測 .

2. スイッチはトラヒックの転送に必要となるリンク本数を推定 .
3. スイッチは新たに設定されたリンク本数を LACP DU パケットにより対向スイッチに通知 .
4. 通知に従って , 新しいリンク本数で集約リンクを構成 .

上記の拡張において , 適切なリンク本数を決定し , 適用するために , スループットのピーク値の経験分布を利用しました . 必要な転送性能の推定にピーク値を用いたのは , リンク本数の削減による省電力化を達成しつつ , パケット損失や遅延による通信品質の低下を避けるためです .

図 3 に提案手法における計測および推定方法を示します . まず , 1 ミリ秒単位のようなごく短い区間ごとにスループットを計測し , τ 区間ごとにその内のピークスループットを記録します . 各制御区間内で記録されたピークスループットから経験分布を作成し , 分布の x パーセンタイル値に基づき , 次の制御区間におけるパケット転送性能を決定します . 決定された転送性能はリンク本数に変換され , LACP DU パケットにより対向スイッチに通知されます . その後 , 新たに決まったリンク本数によりスイッチは論理リンクを再構成し , 通信を行います .

3 シミュレーションモデル

本研究では , 2 台のスイッチがリンク集約により接続されているシミュレーションモデルを用いて , トラヒックに応じて動的にスイッチ間のリンク本数を制御する提案手法の有効性を評価しました .

3.1 シミュレーション環境

シミュレーションに使用したパラメータを表 1 に示します . スイッチ間の物理リンクの本数は 8 本 , 各リンクの転送速度は 1 Gb/s 固定としました . よってスイッチ間の最大転送速度は 8 Gb/s となります . 推定手法については , 制御間隔を T とすると , τ は $T/100$ となるように設定し , 各 τ 区間におけるピークスループットは 1 ミリ秒ごとに計測したのから選択しました . また , スイッチ間でのリンク本数の再構成は推定後直ちに反映されるものとしました . スイッチのバッファサイズは , LACP 機能を有する L3 スイッチで実際に製造販売されているものを参考に , 512 MBytes としました .

シミュレーション時間は 40 秒とし , 短い間にトラヒックの到着レートが大きく変化するトラヒックモデルを想定して , 平均到着レートを 0 から 8 Gb/s 間でランダムに変化させ , その継続時間は平均 1 秒の指数分布に従うとしました . またパケットの発生間隔は指数分布に従い , そのパケットサイズは 1,500 Byte としました . 図 4 に本研究で使用したトラヒックモデルの 1 試行を示します .

3.2 評価指標

提案手法の有効性を評価するための指標として , 本研究では (1) リンク切り換え回数 , (2) 平均使用リンク数 , (3) パケット廃棄率 , (4) パケットがスイッチに到着してから対向のスイッチに到着するまでの平均遅延時間 (平均遅延) , の 4 つを用いました . リンク切り替え回数はシミュレーション時間中にリンクの本数が切り替えられた回数で , 安定した通信性能を提供するためには少ない方がよいことになり

表 1: シミュレーションパラメータ.

リンク本数	8
リンク速度	1 Gb/s
バッファサイズ	512 MByte
パケットサイズ	1,500 Byte
シミュレーション時間	40 second

まず、平均使用リンク数はシミュレーションでを使用した物理リンク数の平均で、この本数が少ないほど省電力効果は高くなります。パケット廃棄率と平均遅延は、小さいほどスループット特性の低下など通信品質への影響を抑制できることを示します。

4 シミュレーション結果と考察

本節では最初に提案手法の基本的な性能を評価します。提案手法はトラヒックの計測に基づいて転送性能を推定するため、まずはトラヒックの変化に対する提案手法の挙動について調査しました。

省電力効果を高めるためには、トラヒックの変動に瞬時に反応できるよう制御間隔を短くすることが考えられますが、リンク数の変更が頻繁に生じてしまいます。また、ヒストグラムからリンク本数を決定する際のパーセンタイル値を小さくすることで、転送性能の余剰を抑え、高い省電力化を期待することができますが、予想を上回るトラヒックが到着した際に通信品質が低下してしまう可能性があります。そこで達成可能な省電力化と通信品質の双方の視点から提案手法の性能を評価しました。

4.1 提案手法の基礎性能評価

本節では、スイッチに到着するトラヒックの変化に対する提案手法の挙動について調査しました。転送性能を決定する制御間隔は 0.5 秒と 5 秒とし、集約リンク数は経験分布の 90 パーセンタイル値を基に決定するとしました。

全シミュレーション時間 40 秒の内、5 ~ 25 秒間のトラヒック特性とリンク本数の推定結果について、制御間隔を 0.5 秒に設定した場合を図 5 に、5 秒に設定した場合を図 6 に示します。両図において到着トラヒックは 1 ミリ秒ごとのスループットです。図 5 より、提案手法により推定されたリンク本数はトラヒックの変化に比較的追従していますが、16 秒付近で不必要なリンク本数の変更が生じています。これは到着トラヒックの変化に過敏に反応してしまったためであり、制御期間が短いと提案手法はトラヒック特性の変化に影響を受けやすいことが分かります。スイッチは、新たに決まったリンク本数により論理リンクを構成する際に、パケットの送信処理を扱うスイッチの CPU やエンジンの処理速度を変更するための時間を必要とします。さらに、短い期間でリンク本数が大きく変化すると、トラヒックの伝搬遅延特性に影響を与えてしまう可能性があります。よってスイッチとトラヒック特性の観点からは、非常に短い期間の間にリンク本数が変化することは望ましくありません。

一方で、図 6 の 20 秒付近に示されているように、トラヒックの変化に対して制御間隔が長すぎる場合は、推定リンク本数がトラヒックの変化に追従できないことが分かります。このように制御間隔を長く設定すると、トラヒックの変化に対する反応が遅くなり、パケット廃棄が生じて通信品質が低下してしまう可能性があります。よって、提案手法ではこのような通信品質の低下を避けながら、トラヒック特性に応じた動的なリンク本数制御による省電力化を達成するための適切な制御間隔について検討する必要があります。そこで次節では、制御間隔が推定の確かさと通信品質にどのような影響を与えるかについて調査しました。

4.2 適切な制御間隔，および，パーセンタイル値の検討

本節では制御間隔を 0.1 ~ 10 秒に設定した場合の提案手法の性能を評価します。また、各制御間隔内で作成する、ヒストグラムからリンク本数を決定するためのパーセンタイル値の影響についても調査します。前節で用いた 90 パーセンタイルは、ピークスルーブットから作成されたヒストグラムの 90 パーセンタイルを含む階級を転送性能として設定するため、トラヒックの変化に対して十分な余裕を持っています。そこで本節ではより小さい、50 パーセンタイル値を用いることでより積極的に省電力化を目指した場合についても評価することにしました。

まず最初に提案手法の省電力性能について検討します。図 7 は制御間隔を 0.1 ~ 10 秒とした時のリンク切り替え回数、図 8 は平均使用リンク数を示しています。両方の図中には、変化する到着トラヒックの平均に正確に追従することができた、理想的な場合における結果も図示しています。図 7 と 8 に示すように、トラヒックの変化に理想的に追従した場合は、リンク切り替え回数は 36.57 回、平均使用リンク数は 4.52 本となりました。

図 7 より、50 パーセンタイルと 90 パーセンタイルの差は殆ど無く、制御間隔が 200 ミリ秒以上である場合は、リンク切り替え回数は理想的にトラヒックの変化に追従した場合よりも低いことが分かります。一方で、図 8 ではパーセンタイル値による差が見られますが、傾向として、制御間隔を短くするほど平均使用リンク数は小さくなっています。特に制御間隔が 1 秒である場合、50 パーセンタイルでは平均使用リンク数は 4.85、90 パーセンタイルでは 5.83 です。本シミュレーションの最大使用リンク数は 8 ですので、削減率はそれぞれで 40 % と 27 % になります。ここで削減率が異なるのは、90 パーセンタイルの方が到着トラヒックに対する余裕をやや多く見積もっているためです。これらの結果より、達成できる省電力効果は制御間隔とパーセンタイル値により異なるものの、提案手法により十分な省電力化が可能であることが分かりました。

提案手法の導入、つまり省電力化を達成するために通信品質が低下することは避けなければなりません。そこで、次に、通信品質の観点から提案手法を評価しました。図 9 にパケット廃棄率、図 10 に平均遅延を示します。図 9 より、パケット廃棄を避けるためには、制御間隔は 50 パーセンタイルでは 500 ミリ秒以内、また 90 パーセンタイルでは 800 ミリ秒以内に設定する必要があることが分かります。

さらに、平均遅延は制御間隔が長くなるにつれ増加していくため、良好な通信品質を提供するためには、制御間隔は短い方が望ましいこととなります。一方で、図 10 では制御間隔がある閾値に達すると、平均遅延は小さくなっていくことが分かります。リンク数はスイッチ間の転送速度を示しており、図 8 に示すように制御間隔が長くなるほど増加します。これは制御間隔が長くなるほど、区間内で高いピー

クスルーットが記録されやすくなるためです。結果として、制御間隔が長くなるほど高速な通信速度を推定するようになり、図 10 の制御間隔が 3 秒以上である場合のように、制御区間が長くなるほど遅延は短くなります。

上述したように、提案手法はパケット廃棄による通信品質の低下を避けながら省電力化を達成しなければなりません。パケット廃棄を避けるためには、50 パーセントイルを利用する場合は制御間隔を 500 ミリ秒よりも小さく設定しなければなりません。高い省電力化を達成できることが分かりました。90 パーセントイルを利用する場合は、省電力効果は 50 パーセントイルよりも幾分低くなるものの、制御間隔は 800 ミリ秒まで長くすることが可能でした。スイッチ製造ベンダーの技術者との議論によれば、スイッチの転送性能を動的に変更するためには少なくとも 100 ミリ秒必要であるとのことでしたので、800 ミリ秒の制御間隔は既存のスイッチでも十分可能だと言えます。これらの結果より、提案手法は十分な省電力性能と実現性を有することが分かりました。

5 関連研究

省電力化を実現する研究としては、ICT (Information Communication Technology) を活用した省電力化 (Greening by ICT) だけでなく、ICT そのものを対象とした省電力化 (Greening of ICT) についても取り組みが成されています。本節では、特にネットワークにおける省電力化の関連研究について述べます。

Gupta らは省電力化のために選択的にスイッチのポートをスリープ状態に遷移させる手法を提案しており [2, 7]、シミュレーションにより省電力効果を検証しています。Gupta らの手法ではパケットの到着が無い場合に直ちにスリープ状態に入りますが、この手法ではスリープ状態とオン状態を頻繁に遷移することになり、スイッチの電気的な負荷を高める可能性があります。そこで著者を含む研究グループではスリープ状態への遷移に待機時間を設けることで、過度の状態遷移を防ぎながら省電力化を達成できることを解析的に明らかにしました [8]。

ネットワークにおける省電力化の精力的な取り組みとしては、南フロリダ大の Christensen らが NSF からの資金を得て進めている Energy Efficient Internet Project があります。彼らの提案する省電力化手法は主に PC などの接続機器の省電力化を達成するためにネットワークがどのような機能を提供すればよいかに着目しており、具体的には proxy, split TCP connections, また PC とスイッチ間のリンク速度の適応制御を提案しています [4]。これらの提案では、スイッチや NIC に省電力化機能を追加することでスリープ時間を増加させ、PC の電力消費の抑制を実現しています。また ALR (Adaptive Link Rate)[9] はスイッチと PC 間のリンク速度をリンクの利用率を元に変更する手法で、この提案手法については、IEEE 802.3 の Energy Efficient Ethernet タスクフォース (IEEE 802.3az[10]) で標準化が進められています。

6 まとめ

本研究では、スイッチにおける動的な転送性能制御による省電力化について調査を行いました。その可能性を調査するために、まず初めに幾つかのスイッチの消費電力を計測し、1 ポート当たりの転送速度

が 1Gb/s の場合、リンク数に比例してスイッチ全体の電力消費量は 2~3W ずつ増加することが分かりました。もしネットワークスイッチの CPU やエンジンの処理速度を、PC における Intel の SpeedStep テクノロジー [5] のようにトラヒックの到着レートに応じて制御することができれば、ネットワークの利用率に応じて電力消費をさらに削減できると考えられます。この調査に基づいて、本研究では消費電力を削減するためにトラヒックの変化に応じてスイッチ間の転送性能を動的に制御する手法を提案しました。具体的には、リンク集約技術で接続されているスイッチを想定し、そのリンク本数を動的に制御することで、スイッチの省電力化を目指しました。提案手法の基礎性能評価では、スイッチに到着するトラヒックの変化に対する提案手法の挙動について調査し、制御間隔が短い場合、トラヒックの変化に対する提案推定手法の追随性は比較的良好ですが、不必要なリンク本数の変更も生じることが分かりました。また、制御間隔が短いと提案手法はトラヒック特性の変化に影響を受けやすいことも分かりました。一方で、トラヒックの変化に対して制御間隔が長すぎる場合は、推定リンク本数がトラヒックの変化に追随できないことが分かりました。このように制御間隔を長く設定すると、トラヒックの変化に対する反応が遅くなり、パケット廃棄が生じて通信品質が低下してしまう可能性があります。そこで制御間隔が提案手法の推定の確かさと通信品質にどのような影響を与えるかの調査を行いました。その結果、制御間隔を 800 ミリ秒以下に設定できるならば、ヒストグラムからリンク本数を決定するための閾値が 50 パーセントでも 90 パーセントでも十分な省電力性能を期待できることを明らかにしました。これらの結果より、提案手法は十分な省電力性能と実現性を有することが分かりました。

今後の課題としては、提案手法でスイッチのバッファの長さも考慮して転送性能を決定することで、制御間隔をより長く設定できるようにすることがあげられます。また提案手法を他のトラヒックモデルや、ネットワークの幹線の実計測データにより評価することを検討しています。

参考文献

- [1] 総務省, “我が国のインターネットにおけるトラヒックの集計・試算,”
http://www.soumu.go.jp/menu_news/s-news/01kiban04_1000001.html, 2010 年 9 月 9 日
- [2] M. Gupta, S. Singh, “Greening of the Internet,” *Proc. of ACM SIGCOMM '03*, pp. 19–26, August 2003.
- [3] B. Nordman, “Networks, Energy, and Energy Efficiency,” presentation at Cisco Green Research Symposium, March 2008.
- [4] C. Gunaratne, K. Christensen, B. Nordman, “Managing energy consumption costs in desktop PCs and LAN switches with proxying, split TCP connections, and scaling of link speed,” *International Journal of Network Management Volume 15*, Issue 5, pp. 297–310, September 2005.
- [5] S. Gochman, R. Ronen, I. Anati, A. Berkovits, T. Kurts, A. Naveh, A. Saeed, Z. Sperber, and R. C. Valentine, “The Intel Pentium M Processor: Microarchitecture and Performance,” *Intel Technology Journal*, V.7, Issue 2, 2003.

- [6] IEEE 802.3ad Link Aggregation Task Force,
<http://www.ieee802.org/3/ad/index.html>
- [7] M. Gupta, S. Grover, S. Singh, “A Feasibility Study for Power Management in LAN Switches,” *Proc. of the 12th IEEE International Conference on Network Protocols*, pp. 361–371, October 2004.
- [8] H. Tamura, Y. Yahiro, Y. Fukuda, K. Kawahara, and Y. Oie, “Performance Analysis of Energy Saving scheme with Extra Active Period for LAN Switches,” *Proc. of IEEE Globecom 2007 General Symposium*, November 2007
- [9] C. Gunaratne, K. Christensen, S. Suen, and B. Nordman, “Reducing the Energy Consumption of Ethernet with an Adaptive Link Rate (ALR),” *IEEE Transactions on Computers*, Vol. 57, No. 4, pp. 448-461, April 2008
- [10] IEEE 802.3 Energy Efficient Ethernet Study Group,
http://grouper.ieee.org/groups/802/3/eee_study/index.html, 2006.

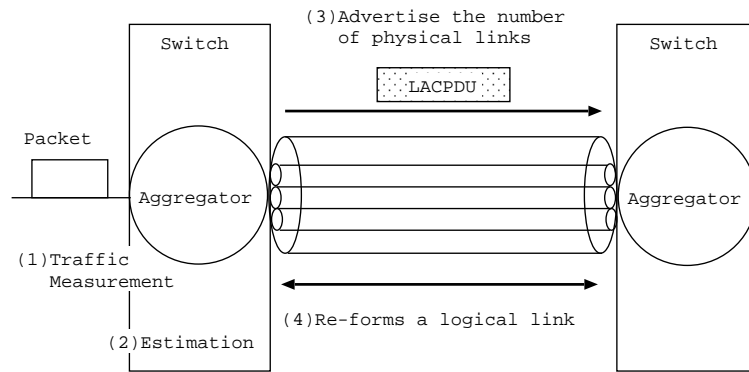


図 2: LACP への拡張機能 .

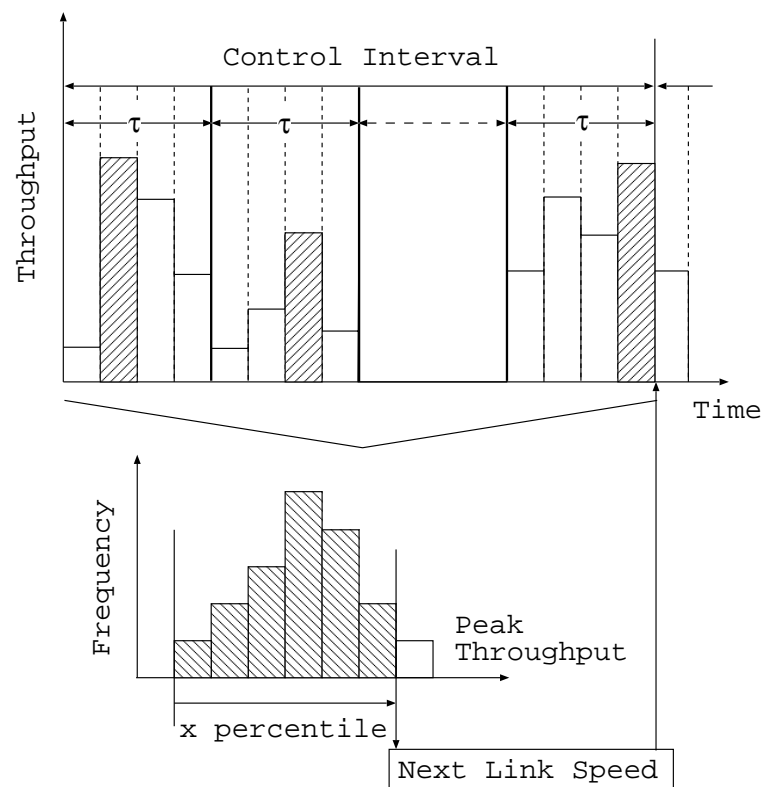


図 3: 推定手法 .

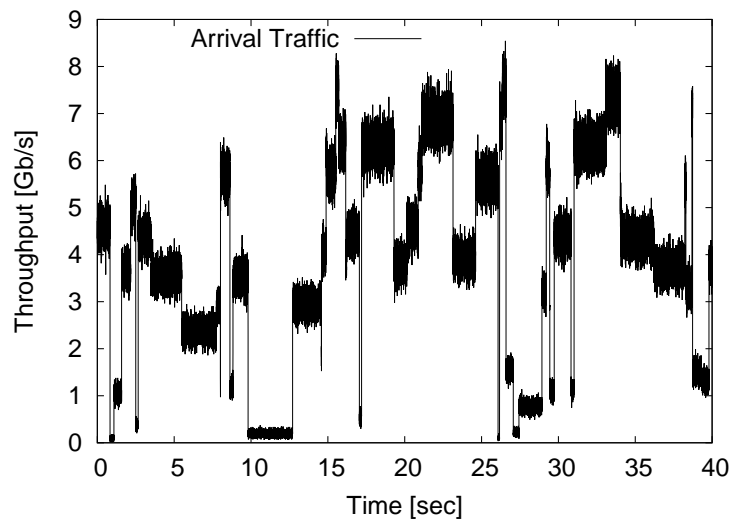


図 4: トラフィックモデルの一試行 .

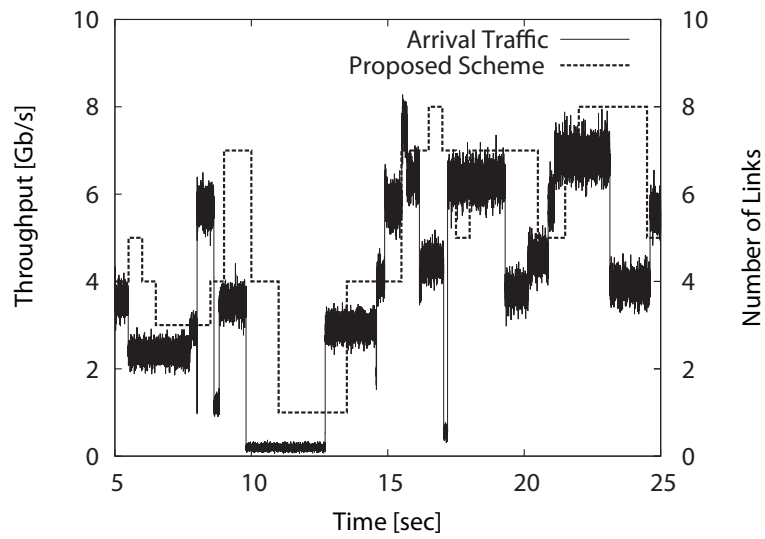


図 5: 集約リンク本数の時間変化 (制御間隔 = 0.5 秒の場合) .

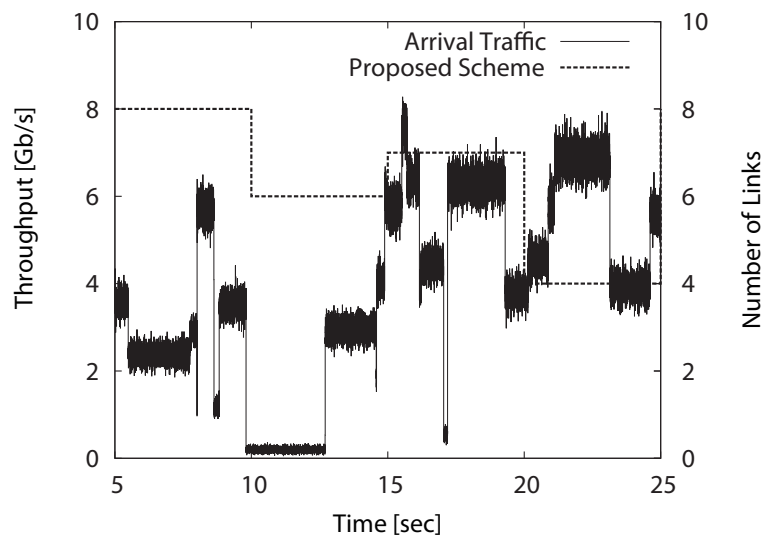


図 6: 集約リンク本数の時間変化 (制御間隔 = 5 秒の場合) .

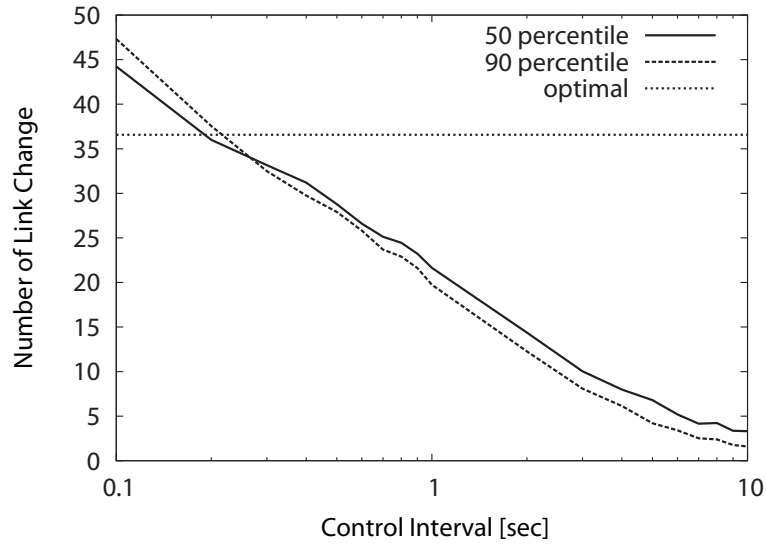


図 7: リンク切り替え回数 .

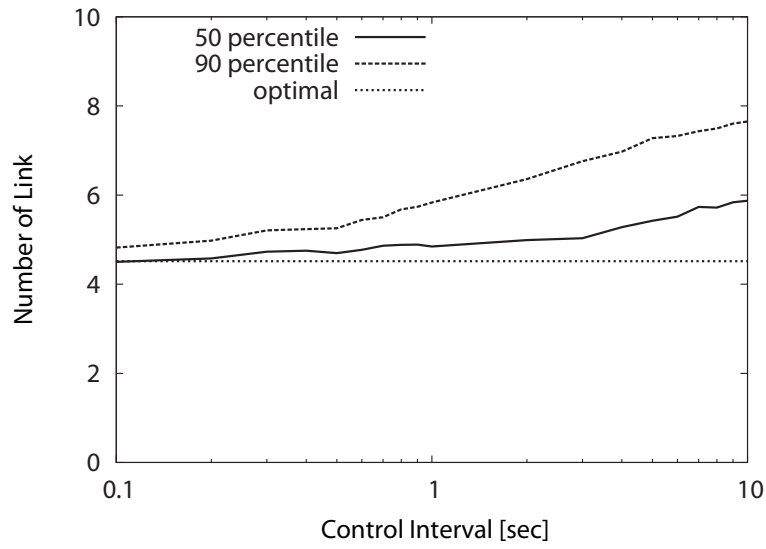


図 8: 平均使用リンク数 .

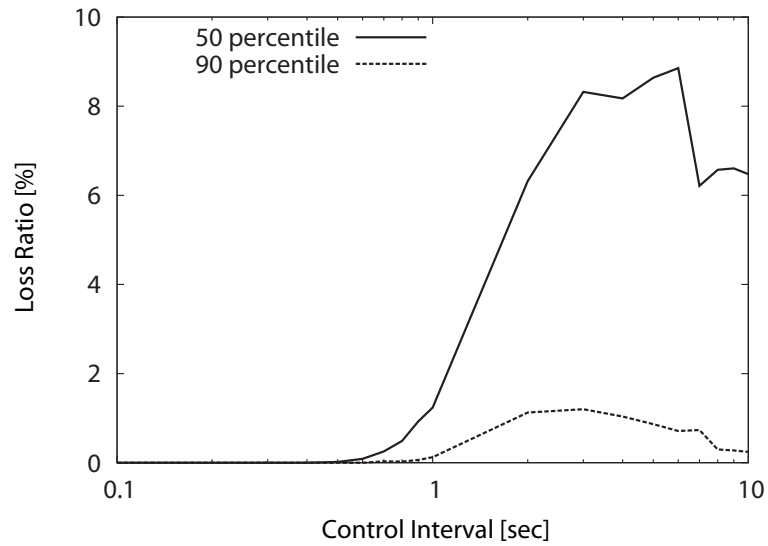


図 9: パケット廃棄率 .

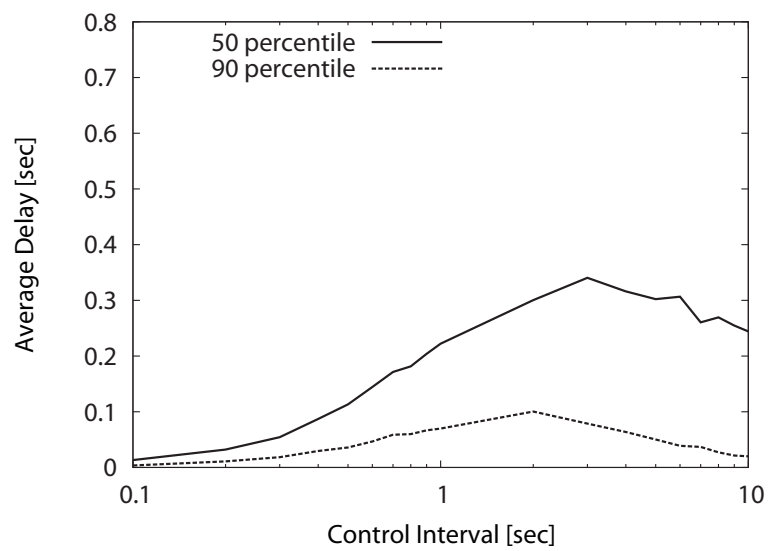


図 10: 平均遅延 .