

推薦論文

パスベース IP パフォーマンス計測システムの実装と評価

井原 修一[†] 中嶋 卓雄^{††}

本研究では、パスベースな IP パフォーマンス評価手法および指標を提案し、計測システムを実装することにより実験的に評価する。ある End ノードに対する特定のパス上のノードに対して連続的にパケットを送出する双方向のアクティブ計測手法を提案し、一連のパケットの時間的および空間的な流れを表すデータから遅延およびロスに関する指標を提案する。またロスを表すノード特徴ベクトルを導入し、ロスに関する指標の定式化を行う。手法および指標を評価するために、ICMP/NTP TIMESTAMP メッセージを送信するプログラムを実装し、ローカルおよびリモートノードが打刻した時刻から遅延およびパケットロスを計測した。結果として、提案した手法によりパス上のロス、遅延の特徴および安定性に関する特徴を抽出することができた。具体的には、一部のノードの遅延については中間ノードとなるルータの内部処理に依存した定期的な遅延が発生したり、ロスについてはパスに依存したりする傾向があり、一部のパス上でロスが蓄積したり、以降のパスではロスが回復したりする傾向などを抽出することができた。

Implementation and Evaluation of Path Based IP Performance Measurement System

SYUICHI IHARA[†] and TAKUO NAKASHIMA^{††}

In this paper, we propose an implementation and evaluation of an IP performance measuring system based on the path based evaluation method. This method is a bi-directional active measurement method that sends packets sequentially to nodes through the path. We introduce new metrics of packet delay and loss based on special and temporal data. To formalize loss evaluating expressions, we introduce node characteristic vector. In our experiments, we sent ICMP/NTP TIMESTAMP messages to sequential nodes of one selected path and extracted four different time stamps of local and remote nodes. As the results, we could extract periodical delays depending on the inner process of nodes and characteristics of loss that depend on the pass, accumulate losses through the path and recover the occurrence of loss after these accumulated nodes of the path.

1. はじめに

インターネットは大規模・複雑であり、動的にネットワークポロジおよび経路が変わるため、モデル化やデータ解析が困難である。そのため、TCP ではある時点における RTT (Round Trip Time) を計測し、ネットワークの利用可能な資源および性能を評価している。しかし、正確で有効なネットワークを評価するため、IP ネットワークのモデル化により理論的にトラフィックデータを解析する手法が提案されている。

従来、IP ネットワークにおけるパケットの到着をポアソン過程に基づきモデル化する研究¹⁾や、ユーザが起動するセッションの到着分布としてポアソン過程が有効であるとする研究²⁾が提案されてきた。また、大規模なネットワークではなく LAN トラフィックに限定して、ポアソン過程より自己相似 (self-similar) な特徴を示す傾向が得られている³⁾。一方、End-to-End のアプリケーションにとって有益な利用可能帯域やスループットを評価する TReno⁴⁾, patcher⁵⁾, bing⁶⁾などの各種ツールが開発されている。おもにこれらのツールでは、ノードにおいてパケットを抽出するパッ

[†] サン・マイクロシステムズ株式会社
Sun Microsystems

^{††} 九州東海大学
Kyushu Tokai University

本論文の内容は 2001 年 3 月の火の国情報シンポジウム 2001 にて報告され、火の国情報シンポジウム 2001 プログラム委員長により情報処理学会論文誌への掲載が推薦された論文である。

シブ計測 (Passive Measurement) ではなく、ノードに対して UDP, ICMP (Internet Control Message Protocol) パケットを送出して性能を評価するアクティブ計測 (Active Measurement) の手法が用いられている。IETF の IPPM (IP Performance Metrics) WG においては、おもにアクティブ計測を用いた、IP 層における性能・評価の指標および手法が検討されている⁸⁾。一方、IP の流れを解析する IPFIX (IP Flow Information Export) WG の活動では、定点での IP フロー情報の抽出がパッシブ計測によって実現されている。

しかし、インターネット全体を特定のモデルにあてはめることは困難であり、評価目的に対応した詳細な計測が必要と考えられる。またパッシブ計測では特定のノードに注目した解析しか行えないため、評価できる項目にも限界がある。したがって、本研究ではネットワークの特徴を抽出することを目的としたアクティブ計測によるネットワークの性能評価を考える。

アクティブ計測で利用するパケットのタイプとして、前述したツールなどでは、UDP や ICMP echo パケットを利用した計測実験が行われてきた⁹⁾。一方、End-to-End の TCP 性能向上を最終的な目的とした計測においては、TCP パケットを利用した計測実験が行われ、TCP ヘッダを抽出し、その評価が行われてきた¹⁰⁾。また、パケットを一方向の流れとし OTT (one-way transit time) を計測しデータを解析するのか、いったん中間ノードを経由した双方向の流れとして RTT を計測しデータを解析するのか、2つの手法がある。

アクティブ計測でインターネットを評価する場合、まず、ネットワーク経路の動的な変動を検討する必要がある。経路は一定のパスで支配される場合が多いが、実際は行き帰りにおける経路変動も発生し、方向に関する対称性などの考察も行われている¹¹⁾。さらに、各ノードが保持する時間の誤差、ノード間の時間の相対誤差について検討する必要がある。IETF では、特に NTP (Network Time Protocol)¹²⁾ により維持されている絶対時刻の存在を前提とした一方向のパフォーマンス評価指標が詳細に検討されている^{13)~15)}。

本研究では、ICMP TIMESTAMP メッセージによるアクティブ計測を用いることにより、広範囲にわたって計測・評価が可能である双方向性の特徴も生かしながら、一方向のデータも収集する。また、End 間の特徴だけではなく、特定のパスに沿った一連のパケットの流れを解析するパススペースな IP パフォーマンス評価手法を提案する。2章では、従来の評価手法の一部

の問題点を検討し、3章では、提案する評価手法について述べる。4章において、その手法に基づく評価指標を提案する。5章で、具体的な実験により評価した結果を、経路変動および時間誤差について考察しながら検証する。

2. 従来の評価手法の問題点

従来のトラフィックをポアソン分布と対応づける手法¹⁾や、LAN トラフィックの自己相似的な現象を評価する手法³⁾においては、手法自身が持つ平均化操作により特徴的なデータを見逃す場合があると考えられる。また、実際のインターネット上の中継ノードにおいて NTP デーモンは全般的に実装されているが、その相対誤差の調整・管理について統一的に行われているとはいえず、また相対誤差はノードが参照する stratum の階層構造などにも依存しており、一定の誤差の幅を持つと考えられる。以下では、これらの問題について考える。

2.1 平均化に関する問題

ネットワーク全体または End-to-End の1つのパスにおいて平均的なパケットの挙動を評価するとき、平均的な値の算出のためロスや遅延の延長と考え、1つの統合した値として挙動を表したり、少数のパケット送信によりノードまたはパスの特徴を暫定的に表したりすることにより、不正確な挙動を評価する結果となる可能性がある。たとえば、前者の場合、ロスの時間を遅延、たとえば RTT を定数倍することにより、ロスを揺らぎの一部と見なし、遅延の分散を表す場合がある。ここでは、あるノードにおいて一定の時間間隔を持って送出されたパケットがその前後のパケットの RTT にどのように影響するのかを、RTT の時系列に関する自己相関係数により評価する場合を考える。

時刻 t においてノード 1 からノード i にパケットを送信した際の RTT の時間系列を $x_i(t)$ とするとき、ノード i における RTT の自己相関関数 $R^i(t, \tau)$ および自己相関係数 ρ^i は次のように表現される。

$$R^i(\tau) = E[x_i(t)x_i(t+\tau)] \quad (1)$$

$$\rho^i(\tau) = R^i(\tau)/R^i(0) \quad (2)$$

ロスパケットを RTT の最大値の定数倍として評価して、各ノードにおいて定数値を変化させた際の自己相関係数の相違を図 1 に示す。各ノードにおいて自己相関係数が大きく変化しているのが分かる。

これらの結果から、本研究においては、遅延とロスを分離させ数値の大きさとして統合せず、限定的に平均化を用いる。

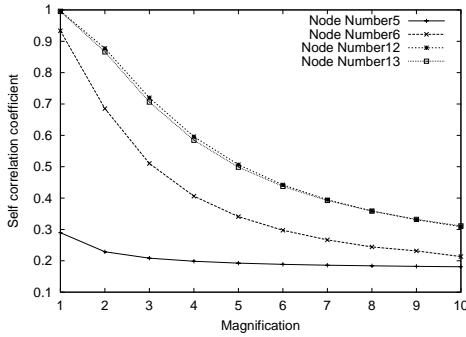


図 1 定数値による自己相関係数の変化

Fig. 1 Variation of coefficient of self-correlation by the magnification.

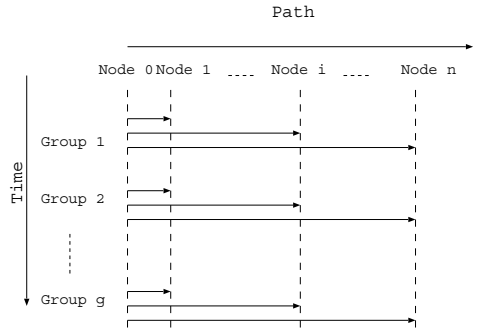


図 2 計測パケット送信パターン

Fig. 2 Measuring packet sending pattern.

2.2 ノードが保持する相対時間誤差に関する問題

IPPMで議論されている一方向性の性能評価においては、NTPデーモンまたはGPS(Global Positioning System)により計測ノード間の時刻が同期していることが前提となる。しかし、時刻補正に関してNTPを用いる場合、RFC2330⁸⁾においても述べられているように、NTPが分や日時程度の長いスケールでの正確さを考えているのに対して、パフォーマンス測定では、秒やミリ秒などの短いスケールでの正確さを求めているため注意が必要である。また、NTPサーバのstratumのレベルなどのNTPの階層的な参照トポロジーに関してノード間の時間の相対的な誤差が発生する可能性がある。

本研究では、広範囲にわたるパスを調査するため双方向の計測手法を用いるとともに、TIMESTAMPメッセージを用いて、ローカルおよびリモートノード上の時刻情報を収集することにより、パス上の中間ノードの相対誤差についても考慮する。

3. 提案する評価手法

本研究では、End-to-Endの性能の大きな要因となるインターネットの性能を評価するため、あるEndノードまでの特定の経路に基づいてパケットを送信する。特定の送信パスに基づいて時間的に連続するパケットの送信によりノード間の特性を解析する。

3.1 計測パケット送信パターン

ノード上の時間的な振舞いのほかに、パス上のノード間の振舞いを解析するため、図2に示す計測パケットの送信パターンを用いる。図2の縦軸方向で経緯する時間を、横軸方向でパス上のノードを示す。ほぼ同時(1[msec]以内)に特定のパス上にあるすべてのノードに対してパケットを送信する。以降このほぼ同時に送信するパケット群をグループと呼ぶ。一定の時

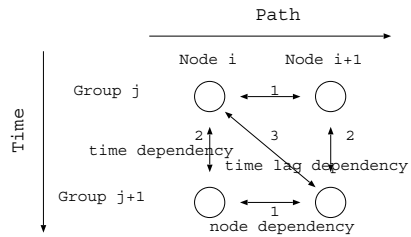


図 3 依存性の分類

Fig. 3 Classification of dependency.

間(以降、タイムスライスと呼ぶ)間隔を空けた後、グループ単位にパケットを繰り返し送信する。

3.2 依存性

提案する送信パターンから計測可能な依存性を図3に示すように3つのタイプに分類する。各タイプは以下に示す特徴を持つと考える。

1. ノード依存性(Node dependency): 同じパス上のノード間における、後方ノードまたは前方ノードへの依存関係を表す。
2. 時間依存性(Time dependency): 同じノード上におけるタイムスライス間隔を空けた場合の依存関係を表す。
3. 時間差依存性(Time lag dependency): 同じパス上の後方に隣接するノードへのタイムスライス間隔を空けた依存関係を表す。

3.3 ノード特徴ベクトル

N をパス P のノード集合とし、 G をパス P のグループ集合とする(図4)。まず、グループ $g \in G$ 、ノード $n \in N$ において、ロスが発生したか否かを、 $l_n(g)$ によって表す。 $l_n(g)$ は値として1または0をとり、1の値をとるとき、パケットロスが発生し、0の値を持つとき、ロスは発生していないとする。

ノード n における、ノード特徴ベクトル L_n を次のように定義する。

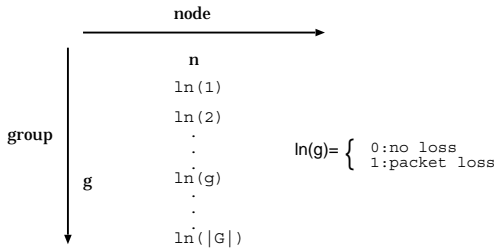


図 4 ノード特徴ベクトル
Fig. 4 Node characteristic vector.

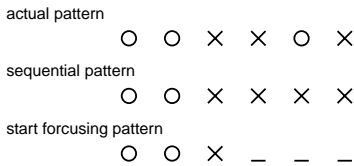


図 5 ロス計測パターンの分類
Fig. 5 Classification of loss measurement pattern.

$$L_n = (l_n(1), l_n(2), \dots, l_n(|G|)) \quad (3)$$

ここで、 $|G|$ をグループのサイズとする。

3.4 パス方向のロス計測パターンの分類

End-to-End の計測では、途中でロスが発生した場合において、どの中間ノードでロスが発生したのかは決定できない。本研究では、1つのグループの packets がほぼ同時に送信されており、同一の packet と見なすことによりロスが発生する中間ノードに注目することが可能である。そこでパス方向のロス計測パターンを次のように 3 つに分類し、具体例を図 5 に示す。
Actual pattern: 実際にノードへ packet が到着したか否かによりカウントする場合。

Sequential pattern: 同一パス上の以前のノードでロスが発生した場合は、そのノードに packet が到着した場合においても、ロスと見なす場合。packet グループを同一の packet と見なしたロスに関する最悪の場合。

Start focusing pattern: ロスが発生した最初のノードに注目し、パス上で以前ロスが発生した場合には、それ以降のノードではそもそも packet が存在しなかったと見なし、カウントしない場合。

4. 提案する評価指標

本研究では、基本的にパス方向の指標を定義する。時間軸方向の指標はノード上の平均化のための指標として利用する。

4.1 遅延に関する指標

従来、遅延に関する指標として、End-to-End にお

ける TCP の性能向上を目的として、最大 RTT と最小 RTT との関係、利用可能帯域と RTT の積による BDP (bandwidth-delay product) によって帯域を有効に利用できるウィンドウサイズの推定および OTT とコネクションに関係したパラメータとの関係などが評価されてきた¹⁶⁾。しかし、End-to-End の計測データからではリンク全体の有効帯域は解析することができるが、パス上のボトルネックとなる可能性の高いインターネット上の中間ノードの検出や隣接ノード間の影響などについて解析が困難である。

本研究では、まずパス上の異なるノードにおいて、同じ指標により比較できる遅延に関する特性を、行きおよび帰りの経路変動を考えて、両方向の最小 OTT に基づいて評価することを考えた。同一パス上の始点からあるノードまでの両方向の OTT を同一の比較対象とするため、方向ごとに最小 OTT を基準にして、次に示す安定エリアと不安定エリアを定義する。

$$\text{安定エリア (Stable area)} : d < \alpha * d_{min} \quad (4)$$

$$\text{不安定エリア (Unstable area)} : d \geq \alpha * d_{min} \quad (5)$$

ここで、

d : OTT

d_{min} : 最小 OTT

(minimum one-way transit time)

α : 安定化係数

(stabilization coefficient) ($\alpha > 1$)

とする。

ノード i において安定エリアに含まれる packet の割合を次式に示す正規化を行った安定化度 (Stable degree) S_i^α によって表す。

$$S_i^\alpha = \frac{\text{安定エリアに含まれる packet 数}}{\text{総 packet 数}} \quad (6)$$

また、不安定エリアにおける遅延は周期的な遅延と非周期的な遅延に分類できる。非周期的な遅延はバースト的な変化など複雑であるが、周期的な遅延は以下のような指標が定義できる。

$$P_{time}(n) : \text{ノード } n \text{ における周期的遅延のピーク点間の時刻差} \quad (7)$$

$$P_{span}(n) : \text{ノード } n \text{ における周期的遅延のピークの幅となる時刻差} \quad (8)$$

4.2 ロスに関する指標

従来、packet ロスは伝統的なネットワークモデルであるポアソン過程としてモデル化されており^{2),17)}、そのため一定時間間隔におけるロス率の計測などの平均的な評価指標しか導入されてこなかった。このよう

な指標では End-to-End の平均的なロスの振舞いは把握できるが、インターネットの隣接するノード間およびそれらに影響するエリアなどに対する考察は困難である。

本研究では、パス上のロスの振舞いを関連づけ、またロスの影響の範囲を表現する指標を導入する。

4.2.1 ノード関連度

同じパケットグループ内の隣接ノード間のロスの振舞いを表す指標として、ノード関連度 (Node Relation Degree) を定義する。ノード i とノード j とのノード関連度 $R_{i,j}$ を次のように定義する。

$$R_{i,j} = (L_i * L_j) / |L_i| \quad (9)$$

ここで、 $*$ は内積を表し、 $||$ はノルムを表す。

ノード関連度 $R_{i,j}$ は、ノード i とノード j の同一パケットグループにおけるロスの依存性を示す。ここで、 $i < j$ のとき、 $R_{i,j}$ を事前依存性 (Pre-dependency) と呼び、 $R_{j,i}$ を事後依存性 (Post-dependency) と呼ぶことにする。

4.2.2 計測パターンに基づくロス率 (loss rate)

すでに定義した計測パターンに基づいてノード i におけるロス率を定義する。

まず、Actual pattern の場合のロス率 $Rate_{i,act}$ について考える。ノード i において $|G|$ により総パケット数が表される。またロスが発生したとき、 $l_i(g)$ は 1 となるので、 $\sum_{g=1}^{|G|} l_i(g)$ によってロスの総数を表すことができる。したがって、ロス率 $Rate_{i,act}$ は

$$Rate_{i,act} = \left(\sum_{g=1}^{|G|} l_i(g) \right) / |G| \quad (10)$$

と表すことができる。

次に、ノード i における sequential pattern の場合のロス率 $Rate_{i,seq}$ について考える。この場合においても送信したグループすべての場合についてカウントするので総数は $|G|$ で表すことができる。ここで、1, 0 の値を逆にする否定記号 “ \sim ” を用い、

$$\widetilde{l}_n(g) = \begin{cases} 1 & l_n(g) = 0 \\ 0 & l_n(g) = 1 \end{cases}$$

を導入する。

このとき、あるグループ g において最初のノード 1 から評価したいノード i までの積 $\prod_{s=1}^i \widetilde{l}_s(g)$ により、ノード i まで少なくとも 1 度以上のロスが発生したか否かを表すことができる。すなわち、 $\prod_{s=1}^i \widetilde{l}_s(g) = 1$ となるのはノード 1 から i まで $l_s(g)$ が 0、すなわちロスが発生しなかった場合であり、 $\prod_{s=1}^i \widetilde{l}_s(g) = 0$ となるのは、ノード i を含み 1 度以上のロスが発生した

ことを表す。したがって、1 との排他的論理和 \oplus をとることによって、ノード 1 から i まで少なくとも 1 度以上のロスが発生した場合の数を 1 として取り出すことが可能であり、総和を求めることによって、ロスの総数を求めることができる。したがって、sequential pattern の場合のロス率 $Rate_{i,seq}$ は次のように表すことができる。

$$Rate_{i,seq} = \left(\sum_{g=1}^{|G|} \left(\prod_{s=1}^i \widetilde{l}_s(g) \right) \oplus 1 \right) / |G| \quad (11)$$

最後に start focusing pattern の場合におけるロス率を定式化する。式 (11) における説明から、 $\sum_{g=1}^{|G|} \left(\prod_{s=1}^{i-1} \widetilde{l}_s(g) \right) \oplus 1$ によって、ノード $i-1$ までの sequential pattern の場合のロスの総数が計算できる。Start focusing pattern の場合は定義から以前のノードまでにおいてロスが発生していたらカウントしないので、全パケット総数 $|G|$ からの減算によりパケットの総数が算出できる。また同様に sequential pattern における $i-1$ までのロスの総数を i までのロスの総数から引くことにより、 i における start focusing pattern な場合のロスの総数が算出でき、次のようにロス率 $Rate_{i,sta}$ が定義できる。

$$Rate_{i,sta} = \frac{\left(\sum_{g=1}^{|G|} \left(\prod_{s=1}^i \widetilde{l}_s(g) \right) \oplus 1 - \left(\prod_{s=1}^{i-1} \widetilde{l}_s(g) \right) \oplus 1 \right)}{\left(|G| - \sum_{g=1}^{|G|} \left(\prod_{s=1}^{i-1} \widetilde{l}_s(g) \right) \oplus 1 \right)} \quad (12)$$

4.2.3 事前安定性

Draft¹⁵⁾ に準じてロス距離 (loss distance) を定義すると、たとえば図 6 に示すような場合にはノード i においてロス距離が 2 となる。

ここで、ノード i においてロスが発生し、 δ 以上のロス距離となる事前安定性 (Pre-stability) $Pre_{i,\delta}$ を導入する。

ノード i から δ 以上のロス距離がある場合は、ノード $i-\delta$ から $i-1$ までの間はロスが発生していないため、 $\prod_{s=i-\delta}^{i-1} l_s(g) = 1$ となる。そのとき、 $l_i(g)$ と積をとることによって、ノード i でロスが発生しロス距離が δ 以上の場合の数を計算することができる。したがって、事前安定性 $Pre_{i,\delta}$ は次のように定義できる。

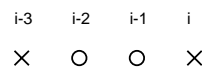


図 6 ノード i においてロス距離が 2 となる場合
Fig. 6 Case of loss distance = 2 on node i .

$$Pre_{i,\delta} = \frac{\sum_{g=1}^{|G|} (\prod_{s=i-\delta}^{i-1} l_s(g)) \cdot l_i(g)}{\sum_{g=1}^{|G|} l_i(g)} \quad (13)$$

ここで、 δ をロス距離の閾値とする。

5. 実装と評価

5.1 実装プログラム

プラットフォームとして FreeBSD-2.2.8 上に、ICMP TIMESTAMP および NTP TIMESTAMP を送信するアプリケーションを実装した。図 7 に ICMP TIMESTAMP⁷⁾ のパケットシーケンスを示す。まず、計測ノードから ICMP TIMESTAMP フィールドの送信元タイムスタンプ (Originate Timestamp) フィールドに現時刻を打刻し、計測されるリモートノードへメッセージを送信する。リモートノードでは、TIMESTAMP メッセージを受信した段階でその時刻を受信タイムスタンプ (Receive Timestamp) フィールドに打刻し、ICMP TIMESTAMP REPLY メッセージとして構成しなおし、送信する直前に伝送タイムスタンプ (Transmit Timestamp) を刻印し、計測ノードへメッセージを返信する。計測ノードでは ICMP TIMESTAMP REPLY メッセージを受信した段階でローカル時刻を求めれば 4 カ所の時刻情報を得ることができる。

5.2 計測実験

計測実験は次のような手順により行った。まず、送信元は熊本大学内のノードに限定し、End ノードに関しては国内および国際回線、および異なる回線を利用するノードを複数抽出した。送信元ノードについては NTP を stratum 1 として動作させ、絶対時刻との誤差を極力取り除いた。次に、End ノードに対して数度にわたって traceroutet によって経路を調べ、ほとんど変化しないと思われる経路を計測パスとして抽出した。その後、その計測パスを構成する各ノードに対して、前述した計測パケットと送信パターンによりパケットを送信しデータを収集した。

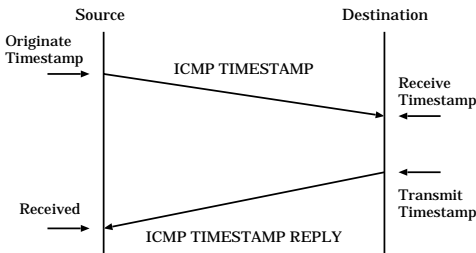


図 7 ICMP TIMESTAMP パケットシーケンス
Fig. 7 ICMP TIMESTAMP packet sequence.

計測パケットのサイズはネットワークへの負荷を最小にするため、サイズを 40 [byte] とし、送信パターンは、まずバス上のすべてのノードへ 1 [msec] 以内にパケットを送信し、グループ間の先頭パケットの間隔を 200 [msec] とし、2,000 回のグループ送信を行った。その計測実験を異なる時間、異なる曜日について、数度にわたって計測した。対象エリアを変えて 2 度実験を行い、実験 I では、全 13 パス (そのうち、海外を含むパスは 2 パス)、および全ノード数は 170 ノード (内訳は国内が 150 ノード、海外が 20 ノード)、実験 II では、全体で 21 パス、242 ノードについて実験を行った。

5.3 方向性と経路変動に関する仮定

経路の動的な変動に関して、TTL (Time To Live) を変化させながら、ある特定の経路について変動を検出し安定性を評価する試みも行われているが¹⁶⁾、TTL を決め対象となるノードへアクティブパケットを送出した場合においても、同じ TTL で同じノードであっても、途中の経路は異なる可能性があることを考えると、全体として経路の動的な変動は誤差として考察することにした。

遅延に関して、送信パケットである ICMP TIMESTAMP により時間的な長さの点から計測パス上の行き帰りの両方向の経路変動については考察が可能であるが、一方向の経路変動については実験データに含まれる誤差として個別に検討を行う。

ロスに関しては、たとえば、単一の ICMP パケットであれば、ネットワーク上のどのノードで、さらには行きまたは帰りのどちらの方向で、ロスが発生したのかを判定することはできない。しかし、本研究では、ほぼ同時に送信させたパケットのグループを構成し、それらを 1 つのパケットと見なすことにより、行き方向の経路でロスが発生したと仮定する。行き方向に限定した視点をとるのは、帰り方向の経路では、パケットの間隔も分散化し、健全なネットワークであればロスの発生は少なく、相対的にロスが発生しにくい状況になっていると考えたからである。

経路の動的な変動により、実験段階において当初選択した計測パスと異なる経路を通ったり、ある End ノードへの経路としては、計測パス上の中間ノードを通らない可能性や、計測パスの隣接関係が実際には成立していない可能性もあるが、選択した計測パスは End ノードまでの妥当な経路と仮定し、提案した計測手法は経路パス上の一連のノード間の特性などを評価できると考える。経路の変動およびノードが保持する相対時間の誤差についての影響などは、個別実験デー

sec	node	stratum
1	133.95.102.254	-
2	133.95.14.217	-
3	kumamoto-1-A8-0-9.sinet.ad.jp	3
4	kyoto-10-A6-0-3.sinet.ad.jp	3
5	kyoto-gate-FE1-0.sinet.ad.jp	3
6	if-12-0-0-1.bb2.PaloAlto.Teleglobe.net	2
7	if-2-3.core1.PaloAlto.Teleglobe.net	2
8	p3-2.paix-bi1.bbnpplanet.net	3
9	p7-0.paloalto-nbr1.bbnpplanet.net	3
10	p4-0-0.paloalto-br2.bbnpplanet.net	3
11	f1-0-0.paloalto-cr11.bbnpplanet.net	3
12	h2-0.oracle2.bbnpplanet.net	3
13	bigip-www.us.oracle.com	-

図 8 計測パスの一例

Fig. 8 Sample of the measuring path.

タの説明の箇所です述べることにする。また、本研究では、ノードを分類するための指標および特徴などを抽出することを目的としているため、代表的な計測データを例として説明している。

5.4 計測パスの一例

図 8 に示す計測パスは、熊本大学から通過ノード数が 13 である SINET と国際回線を含む海外までのパスである。他のパスでも同様の結果が得られているので、実験データの例示については、この例を代表例として図で示す。また、ノードが起動している NTP デーモンの stratum 番号を右に併記する。ただし、“-” は NTP デーモンが起動していないノードを示す。実験 I では、170 ノード中でほぼ 80% のノードが NTP デーモンを起動していた。

5.5 最小 Trip Time

RTT の値についてはノードの相対時間誤差については考慮する必要はないが、経路変動による影響はある。また、バラツキが大きく、どの RTT の値を基準と考えるのかは目的によって異なる。

本研究では、ノードにおける特徴をパス上で比較することを目的としているので、始点ノードと注目するノード間において、経路変動の発生の有無に関係なく、最良の値として最もネットワーク的な距離が近くなった値と考えられる最小 Trip Time を基準にする。

計測パス上の各ノードにおける最小 RTT の変化、および TIMESTAMP メッセージから得られた、2 つの Trip Time のうち、行きに要した fptt (forwarding path trip time) の最小値を図 9 に示す。ここでは、計測パケットの送信時間帯を (1) 9 時、(2) 14 時、(3) 17 時の 3 種類変え比較する。図から、行きまたは帰りの経路変動が発生した場合においても、計測パス上の一連のノードは、最良の場合において漸近的にネットワーク的な距離を増加させていることが分かる。ま

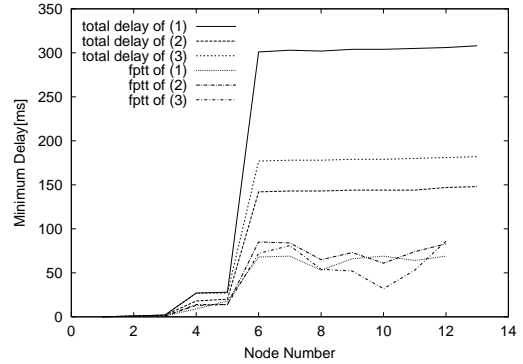


図 9 ノードによる最小 Trip Time の変化

Fig. 9 Variation of minimum RTT on the path.

た、国内の最後のノードである 5 から国外のノードである 6 までの空間的距離が大きいところにおいて変化が大きいことが分かる。時間帯による影響は最小遅延が急激に大きくなるノード間の影響に依存して、基本的な遅延が決定していると考えられる。

最小 RTT の計測結果から、経路の動的な変動があったとしても計測パス上の一連のノードはネットワーク的な距離が増加していると考えることができ、したがって fptt についても、増加する傾向となるはずであるが、一部のノードでは減少するなどバラツキがある。したがって、NTP が動作している場合においても、各ノードと送信ノードの間には、30 [msec] 程度の相対誤差があり、送信ノードが stratum 1 であることから、計測パス上のノードの時間の進みが存在する可能性がある。この相対誤差による影響から、bptt (backwarding path trip time) は fptt と比べて相対的に長くなっている。

5.6 遅延に関するノードの安定性

最良の値である最小 Trip Time に基づき、遅延に関するノードの安定性の指標を式 (6) の安定化度として定義した。NTP デーモンが起動しているノードに対して、行き方向と帰り方向の両方向において最小 OTT に基づいた式 (6) の安定化度を示したのが、図 10 である。ここで、安定化係数 $\alpha = 1.1, 1.3$ とした。

図 10 から、行きについては国際回線に出るノード 5 において、最小 OTT の 1.1~1.3 倍の範囲に揺らぎが存在し、他のノードにおいては、1.1 倍の範囲で揺らぎがおさえられている、帰りについては、すべてのノードにおいて 1.1~1.3 倍の範囲に揺らぎが存在していることが分かる。全体として、最小 RTT が大きく変化するノードにおいては行き方向の揺らぎが大きく、帰り方向はすべてのノードについて揺らぎが大きくなることが分かった。したがって、最小 RTT が大き

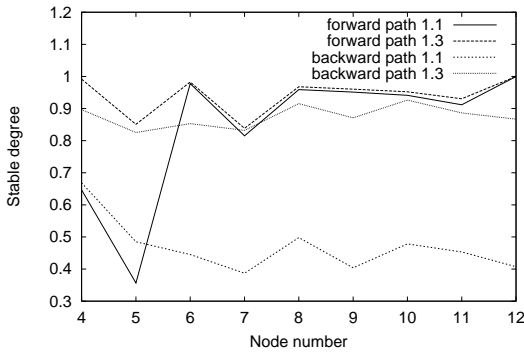


図 10 行きおよび帰りにおける安定な領域の割合

Fig. 10 Stable degree of the forwarding and backwarding path.

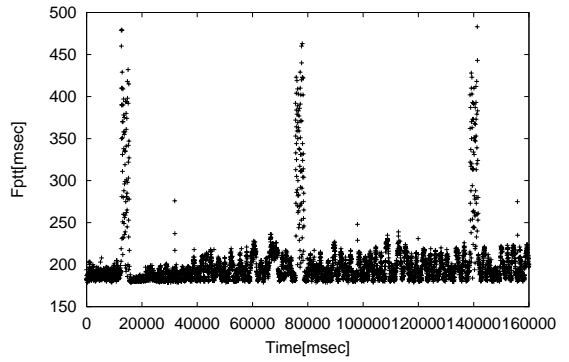


図 11 ノード 10 における行きに要した遅延の変化

Fig. 11 Variation of fpitt on the node 10.

く変化するノードはトラフィックが集中し揺らぎが大きくなり、また、経路の方向によってノードの特徴が異なることから、帰りの方向は経路変動またはパケットの時間的なバラツキの拡大による影響と考えられる。結果から、一定の安定化度を保つ、すなわち全体的な揺らぎを一定の範囲におさめるときの α の値も実験的に得ることが可能である。

5.7 遅延の周期性

実験 I において、TIMESTAMP により RTT を行き、および帰りの遅延に分離させたとき、行きの遅延についてのみ、図 11 に示す周期的な変化が得られた。ここではノード 10 における遅延の変化を示す。具体的には、式 (7) における $P_{time}(10) = 60 [sec]$ となり式 (8) における $P_{span}(10) = 2 [sec]$ となった。

ノード 8, 9, 10, 11 についても同じように約 60 [sec] の周期的な遅延が発生し、今回調査した実験 I の国内 150 ノード中 82 ノード (54%)、海外 20 ノード中 9 ノード (45%) のノードにおいて周期的な遅延が発生していた。一方、帰りに要した遅延においては、このような周期性はまったく見られなかった。

システムの実装として、TIMESTAMP は行きの遅延とノード上での処理時間を加算し、戻る前に押されていることから、ノードにおける他の優先的な処理が周期的であり、fpitt の値に影響していると考えられる。また、ここではデータとして示していないが、個々のノードが異なるタイミングで周期的な動きをしていることから、他のノードへの依存関係が見られず、パスに関しては独立でありノードを通るパケットについては、ノード上の優先的な処理が影響していないことが明らかになった。ノード上の他の処理は ICMP の処理より優先度が高く周期的なことを考えると、隣接 IS 間で送信される BGP-4¹⁸⁾ の KeepAlive メッセージである可能性もある。

表 1 平均ノード関連度 $R_{i,j}$ の例

Table 1 An example of average node relation degree $R_{i,j}$.

$i \setminus j$	1	2	3	4	5	6	7	8	9
1	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	-	0.37	0.36	0.35	0.34	0.33	0.33
4	0.00	0.00	0.54	-	0.53	0.53	0.50	0.49	0.48
5	0.00	0.00	0.54	0.55	-	0.73	0.68	0.66	0.59
6	0.00	0.00	0.45	0.46	0.63	-	0.59	0.58	0.48
7	0.00	0.00	0.49	0.49	0.70	0.73	-	0.77	0.65
8	0.00	0.00	0.30	0.30	0.50	0.51	0.52	-	0.63
9	0.00	0.00	0.33	0.33	0.43	0.40	0.45	0.65	-

5.8 ロスに関するノード依存性

図 8 に示した計測パスにおけるノード 9 までの平均ノード関連度を表 1 に示す。ノード関連度は最小 Trip Time が大きくなるノード 3 と 4 の間から発生し、国際回線に出るノード 5 と 6 の間において高くなり計測パス上を伝わっている。ただし、国内、たとえば 202.232.2.xxx (IIJ のノード) との平均ノード関連度は、ノード 6 が SINET の最後のノードであるが、 $R_{3,4} = 0.46, R_{4,5} = 0.65, R_{5,6} = 0.67, R_{6,7} = 0.64$ となるなど、プロバイダ間の変化点で大きく変化するのはなく、最小 Trip Time がある程度大きいノードに関係した中間的な経路において値が高くなっていた。これはロスが発生したパスに注目し、その平均値をとっているためであり、平均ノード関連度は計測パス上のノード間のロスの傾向を表すと考えられる。

次に、ロスが発生した以降の影響が計測パス上にどの程度継続するかを調べるため、実験 II のすべてのパスについて、最も事前安定性のノード関連度が高いノード間距離とノード関連度との関係を求め図 12 に示した。

この結果から、大きな関連度を持つノードは隣接した位置にあり、6, 7 割の関連度を持つ場合が多い。

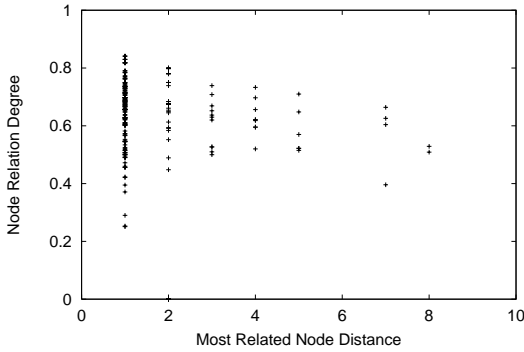


図 12 ノード間距離とノード関連度

Fig. 12 Relation between node distance and node relation degree.

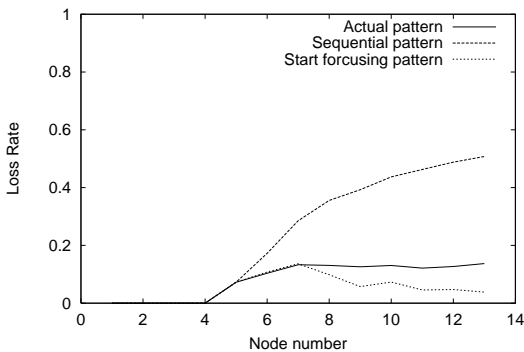


図 13 バケットロス率の例

Fig. 13 An example of packet loss rate.

方、隣接しない場合において最大となるときは、その中間ノードにおいてもノード関連度は大きな値をとることも別のデータから判明している。これらの結果から、経路変動を考慮しても計測パスの隣接ノードは隣接関係を保持している可能性が高く、計測パス上でロスの影響が伝わり、大きな値のまま継続する場合があることが明らかになった。

5.9 ロス率

ノード関連度はロスの発生に注目したデータであるが、一般的なネットワークにおいてロスの発生は少ない。ここでは計測パターンによるロス率 $Rate_{i,act}$ 、 $Rate_{i,seq}$ および $Rate_{i,sta}$ の相違を特徴づけるため、ロスが頻繁に発生した場合の例を図 13 に示す。

Actual pattern が実際の場合であり、sequential pattern が最悪な場合と考えられるため、その 2 つの pattern の間の領域はインターネットが高速に回復する能力を示していると考えられることもできる。理由として、経路変動がなければ、ほぼ同時に送信したパケットは各ノードで同様に扱われるであろうと考えられるが、実際には経路変動やパス上のノードを経由す

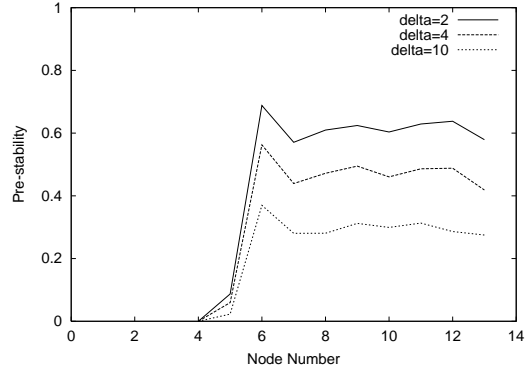


図 14 事前安定性の例

Fig. 14 An example of pre-stability.

る過程で遅延が発生しパラツキが大きくなりロスが回避されている可能性がある。

Actual pattern の場合、ノード 4 からノード 7 までは単調増加し、その後は 1 割程度のロス率となる。この漸的に増加する部分は、経路変動による影響はなくパス上の隣接関係によるロスの影響が発生していると考えられる。Start focusing pattern の場合がロスのスタートポイントに注目していることから、ノード 7 までロスが蓄積しパス全体のロスの原因を作り、それ以降は回復しながら部分的にロスが始まっていることを示している。このように計測パターンを変えることによりロスの特徴を抽出することができる。

5.10 事前安定性

ロス距離の閾値 δ がそれぞれ 2, 4, 10 をとる場合における式 (13) のパケットの事前安定性を図 14 に示す。ノード上のロスの発生に注目したノード関連度においてはパス上のノード間の間隔は把握できないが、事前安定性により注目するノードより前のノードに対する連続的な距離の関係を表すことができる。図から $\delta = 2$ となる場合がロスの事前状態の中で 6 割程度存在することや、 $\delta = 4$ まで δ を大きくしても、その状態が 5 割ほど存在することから、過半数はバースト的なロスや比較的短い間隔でロスが生じていることが分かる。また、各ノードにおいて δ 値に対して事前安定性は指数関数的に減少している。このように計測パスにおけるロスの継続性はあまりなく、経路変動の影響やパケット転送による遅延の増加によってロスの影響が分散することなどが考えられる。

6. おわりに

本研究では、1 [msec] 以下という短い間隔のパケットをグループ化しパス上に送信するアクティブ計測によって、パス上の振舞いに限らずノード上の時間的な

振舞いも評価することができた。遅延については、一部の間中ノードとなるルータの内部処理に依存した定期的な遅延が発生するなどの傾向を抽出することができた。ロスについては、計測パターンの分類によりパス上のロスの振舞いと特徴を抽出することができた。また、ノード特徴ベクトルの導入によりロス率などの定式化を行うことができ、パス方向にノード間の依存性があることが抽出できた。

今後は NTP による時刻の相対誤差の幅を検討するとともに、ノードの特徴を構成する要因を実験的に抽出し、トラフィックの予測を可能とするモデルに発展させたい。

参 考 文 献

- 1) Frost, V. and Melamed, B.: Traffic Modeling for Telecommunications Networks, *IEEE Communications Magazine*, Vol.32, No.3, pp.70-80 (1994).
- 2) Paxson, V. and Floyd, S.: Wide-Area Traffic: The Failure of Poisson Modeling, *IEEE/ACM Trans. Networking*, Vol.3, No.3, pp.226-244 (1995).
- 3) Leland, W., Taqqu, M., Willinger, W. and Wilson, D.: On the Self-Similar Nature of Ethernet Traffic, Extended Version, *IEEE/ACM Trans. Networking*, Vol.2, No.1, pp.1-15 (1994).
- 4) Mathis, M. and Mahdavi, J. (2002). http://www.psc.edu/networking/treno_info.html
- 5) Jacobson, V. (1997). <http://www.caida.org/tools/utilities/others/pathchar>
- 6) Beyssac, P. (1995). <http://www.cnam.fr/reseau/bing.html>
- 7) Postel, J.: Internet Control Message Protocol, RFC792 (Sep. 1981).
- 8) Paxson, V., Almes, G., Mahdavi, J. and Mathis, M.: Framework for IP Performance Metrics, RFC2330 (May 1998).
- 9) Bolot, J.C.: End-to-end packet delay and loss behavior in the Internet, *Proc. SIGCOM '93*, pp.289-298 (1993).
- 10) Paxson, V.: End-to-End Internet Packet Dynamics: *IEEE/ACM Trans. Networking*, Vol.7, No.3, pp.277-292 (1999).
- 11) Paxson, V.: End-to-End Routing Behavior in the Internet: *IEEE/ACM Trans. Networking*, Vol.5, No.5, pp.601-615 (1997).
- 12) Mills, D.L.: Network Time Protocol, Version 3, RFC1305 (Mar. 1992).
- 13) Almes, G., Kalidindi, S. and Zekauskas, M.: A One-way Delay Metric for IPPM, RFC2679 (Sep. 1999).
- 14) Almes, G., Kalidindi, S. and Zekauskas, M.: A One-way Packet Loss Metric for IPPM, RFC2680 (Sep. 1999).
- 15) Koodli, R. and Ravikanth, R.: One-way Loss Pattern Sample Metrics, <draft-ietf-ippm-loss-pattern-04.txt> (Nov. 2000).
- 16) Paxson, V.: Measurements and Analysis of End-to-End Internet Dynamics, Ph.D. dissertation, U.C. Berkeley (1997). <ftp://ftp.ee.lbl.gov/papers/vs-thesis/dis.ps.gz>
- 17) Altman, E., Avrachenkov, K. and Barakat, C.: A Stochastic Model of TCP/IP with Stationary Random Losses, *ACM SIGCOMM Computer Communication Review, Proc. Conference on Applications, Technologies, Architectures and Protocols for Computer Communication*, pp.231-242 (Aug. 2000).
- 18) Rekhter, Y.: A Border Gateway Protocol 4 (BGP-4), RFC1771 (Mar. 1995).

(平成 13 年 9 月 14 日受付)

(平成 14 年 12 月 3 日採録)

推 薦 文

パスを考慮した性能計測手法の提案および実際のネットワークでの検証を行っており、有用性も高いので、推薦に値すると認めた。

(火の国情報シンポジウム2001プログラム委員長
岡田 直之)



井原 修一

1976 年生。1997 年熊本電波工業高等専門学校卒業。1999 年熊本大学工学部電気情報工学科卒業。2001 年熊本大学大学院自然科学研究科修士課程修了。2001 年サン・マイクロシステムズ株式会社入社。現在に至る。一方向性における IP パフォーマンスの評価に興味を持っている。



中嶋 卓雄 (正会員)

1956 年生。1986 年熊本大学大学院工学研究科情報工学専攻修士課程修了。富士通(株)を経て、1991 年熊本大学大学院自然科学研究科博士課程単位修得後退学。熊本大学工学部助手を経て、2001 年より、九州東海大学応用情報学部情報システム学科専任講師。ネットワーク・システムのパフォーマンスの評価、テキストマイニング等に興味を持っている。ACM, IEEE 各会員。