

パスベースな IP パフォーマンス評価手法の提案

井原 修一, 中嶋 卓雄

熊本大学 工学部

E-mail: {ihara@st.,taku@}cs.kumamoto-u.ac.jp

本研究では、広範囲にわたって測定・評価が可能な双方向によるアクティブ測定を用い、End-to-Endではなく、パスに沿って隣接ノード間の一連のデータパケット流れを解析するパスベースな IP パフォーマンス評価手法について提案する。まず、時間的および空間的な環境の中で RTT を分類しパフォーマンスの特徴を表現するモデルを提案する。実験として、ICMP TIMESTAMP および NTP TIMESTAMP メッセージを用い、ローカルおよびリモートホストが打刻した時刻およびパケットロスを測定した。結果として、(1) ロスを含まない遅延はパスに依存せず、ロスについては依存する傾向がある。(2) ロスを含まない遅延については、パケットを戻す内部処理に依存し、ルータ内の処理に比例して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。

A Proposal of Path-based IP Performance Measurement Method

Syuichi Ihara, Takuo Nakashima

Faculty of Engineering, Kumamoto University

In this paper, we propose the bi-directional active measurement method based on the transmission path. We transmit the ICMP TIMESTAMP packet to every node through one Internet path with packet group, and catch the return packet. We can get the following features. (1) The delayed packet non-including loss packet does not depend on the transmission path, but the loss packet depends on the path. (2) The delayed packet mostly depends on the internal processing on the node, and sometime the internal processing has the periodical load.

1 はじめに

ネットワークの性能評価については、IPPM(IP Performance Metrics) WGによって、トランスポート層における性能・評価の指標および手法が検討されている [1]。特に NTP 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 [2][3]。一方で、Treno, Traceroute などのツールが公開され、部分的な評価も自由に行える環境にある。また、実際のインターネット上の中継ノードにおいては、NTP デモンが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行われているとはいえず、またノードの参照する 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。したがって、ICMP や NTP などを利用したパフォーマンス評価においては、計測値の有効な幅を考慮する必要がある。

従来から利用されている評価手法はリモートノードに対して定期的なパケットを送出することによる、一方向または双方向性の最大利用可能帯域、遅延お

よびパケットロスなどである。また、そのデータの解析は、あるノードに対する遅延の時間的な変化の解析であり、一定の期間の平均を求めたりするなど、変化の特徴を抽出するのではなく、全体の振舞いを均一に考える手法が基本となっており、原因となるノードおよび途中のパス上の問題などについては解析が困難である。また、アプリケーションのパフォーマンスを評価する場合、end-to-end の測定によって得られたデータに基づき評価が行われている。しかし、インターネットの途中の経路による影響が少なからずアプリケーション層におけるパフォーマンスに影響すると考えられるが、ネットワーク層におけるノードの特徴を含めた解析は十分に行われていない。

本研究では、広範囲に渡って測定・評価が可能な双方向によるアクティブ測定を用いるとともに、隣接ノード間のパスに沿って一連のデータパケット流れを解析するパスベースな IP パフォーマンス評価手法を提案する。

具体的には、ICMP TIMESTAMP および NTP TIMESTAMP メッセージを用い、ローカルおよび

リモートホストが打刻した時刻、およびパケットロスを測定する。その後、分類した RTT の基本的な性質に沿ってノードの特徴を抽出する。またパスに依存する要素について考察する。これらの結果から RTT の値の評価に関する総合的な指標を提案する。

2 従来の評価手法の問題点

従来の評価手法に関して、(1) 計測値を平均化することによる問題、(2) ホストが保持する時刻の相対誤差に関係した問題、がある。

2.1 平均化に関する問題

平均化の対象となる計測パケットの母集団の性質とその計測により評価される事項との対応関係が不適切なものがある。例えば、ユーザがインターネットにアクセスする動作を、時間、日、曜日などの単位で大量のパケットにより調査するのは意味があるが、アプリケーションの速度改善のため、秒および分程度の間隔を持つ計測パケットによりネットワーク負荷を評価することは、アプリケーションの動作速度と計測パケットのゆらぎの関係を明確にしない限り問題があると考えられる。

また、以前の研究 [5] において示したように、ロスパケットを RTT が揺らいでいるパケットと同様に扱うため、定数倍することにより、一括して扱うことが可能となるが、揺らぎの特徴を表す自己相関などもその定数値により大幅に変化する現象がみられる。これも平均化の対象としてロスパケットも同様に扱ったことが問題と考えられる。

2.2 相対誤差に関する問題

ICMP ECHO REQUEST パケットのみによる遅延の計測では、行きのパケットの経路と帰りのパケットの経路が異なる問題については解決能力をもたないが、一方で、厳密に一方方向のみによる計測においても、NTP デーモンの相対誤差をその NTP サーバのトポロジー構造なども考慮して議論しないと厳密な評価にはつながらない。しかし、パケットが流れるパス上の各ホストは異なる NTP トポロジーを持つ場合が多いと考えら、複雑な解析が必要だと考えるので、この問題に対しては本研究では一定の誤差があるという前提で議論することにする。

3 提案する評価指標と評価手法

ICMP を利用した RTT の測定においては、従来リモートノード上の時間的な振舞を RTT の値により類推していたが、送出パケットの途中の経路による影響も無視できない。本研究では、ノード上の時間的な振舞の他にパス上のノード間の振舞を解析することにより、詳細に特徴を分析する。

3.1 計測手法

3.1.1 パケット送信パターン

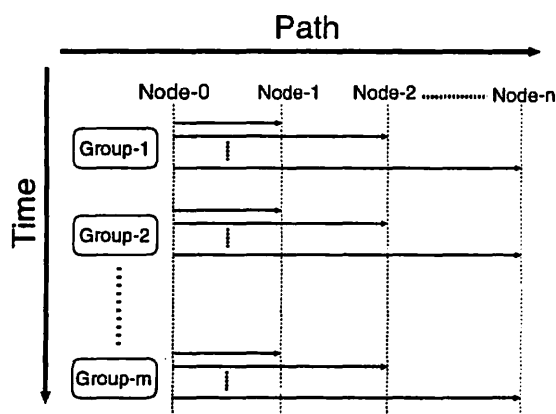


図1. パケット送信パターン

図1にパケットを送信する時のパターンを示す。縦軸方向に経緯する時間を、横軸方向にパスとパス上にあるノードを示す。ほぼ同時(1msec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 RTT に影響する基本的性質の分類

ノード上における RTT の揺らぎ、およびパス上のノード間における RTT の相互依存性について考える。

3.2.1 揺らぎ

ノード上の時間軸に関して揺らぎを、(1) パケットロス、(2) 不安定領域(周期的揺らぎ、非周期的揺らぎ)、(3) 安定領域(最小 RTT に近い領域の揺らぎ、の3つに分類する。

以下のパラメータを導入することにより、揺らぎを説明する。

$\alpha =$ 安定化係数 ($\alpha > 1$)
 $J_{period}(i) =$ ノード i における周期的揺らぎの周期
 $J_{span}(i) =$ ノード i における周期的揺らぎの間隔

ここで、(1) 安定化係数はノードに依存しない、(2) $[\alpha * \text{最小遅延}]$ に収まる揺らぎを安定領域とし、それ以外でロスしない領域を不安定領域とする、(3) 周期的揺らぎはノードに依存する、と仮定する。

3.2.2 パス依存性

パスにおける依存性を、(1) 前方のノードの揺らぎが後方へ影響を及ぼす後方依存性、(2) 後方のノードにおける揺らぎが時間をずらして前方に影響を及ぼす前方依存性、の2種類について検討する。揺らぎの中でロスとならず RTT が変化する場合は、前後1つのノード間での揺らぎの関係を考える。しかし、ロスが発生する場合は、例えばあるノード n へ向けて送信したパケットがロスした場合には、ロスが発生したノードが $1 \sim n$ のどれか決定できない。本研究では、1つのグループ中のパケット群でパスを構成しているので、そのパケット群中の異なるノード宛に送信したパケットがロスしたか否かによって全体的なパス上のロスを考察する。

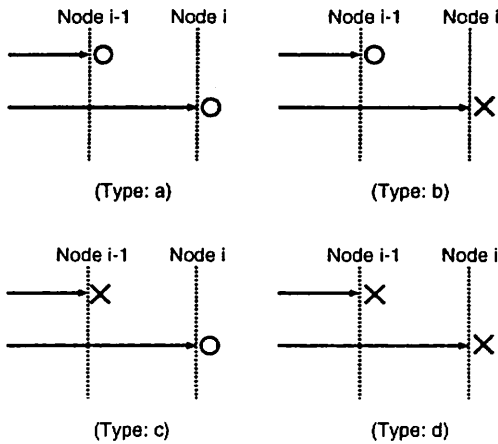


図2. パス上のパケット到着パターン

まず、同一のグループ内の前後1つのノード間におけるロスの発生するまたは発生しない場合の組合せパターンを図2に示す。

従来は、あるノードへ向けたパケットの解析であったため、どのノードにロスの原因があったのかを判別することが困難であったが、パスベースでは、どのグループのどのノードでロスをしたのか明確にな

るので、パス全体に渡った拡張した分類を前記のパターンに基づき次のように定義する。(1) 通常の場合 (normal case): 実際にノードへパケットが到着したか否かによりカウントする場合。この場合は従来の手法と同様にパス上のロスをノード上のロスとして考えた場合である。(2) 厳密な場合 (strict case): 同じグループ送信中、そのノード以前のノードでパケットがロスした場合。すなわち、図2における (type :d) の場合に相当し、そのノードにはパケットが到着しなかった、すなわちそもそもパケットの送信を行わず、ロスしたとも見なさない場合。この場合を考察することによりロスの原因となるノードが明確になる。

3.3 基本的性質に基づく評価指標

前述した基本的な性質は相互に影響を及ぼす可能性がある。以下では、複合した指標を定義する。

3.3.1 ロス率

ロス率 (loss rate) については、通常の場合および厳密な場合について考えることができる。通常の場合のロス率は、[type1:] ノードにおいてパケットが到着したか否かに依存したロス率、[type2:] ロス率を定義するノード以前のパス上のノードで同じグループのパケットがロスしていた場合、それ以降のパスにおいてもすべてロスしたと考えた場合のロス率、の2種類について考察する。

3.3.2 ロスに関連した他の指標

Draft[4] で定義されているロス距離 (loss distance) およびロス期間 (loss period) に基づいて、次のように本研究におけるパスベースな定義に変更する。なお以下では通常の場合にのみ意味を持つ。ノード i におけるロス距離 $L_d(i)$ とは、ロスが発生したグループ番号と次にロスが発生したグループ番号の差分をロス距離と呼ぶ。ロスの長さではなくロスの発生がない安定した長さを表す指標となる。ロス距離に基づき、ロスのない安定性を示す事前安定性 (pre-stability) $P_s(i)$ を次のように定義する。

$$P_s(i) = C_{loss}(i, L_d(i), \delta) / C_{loss}(i) \quad (1)$$

ここで,

$$\begin{aligned} \delta &= \text{ロス距離のサンプル値} \\ C_{loss}(i, L_d(i), \delta) &= \delta \text{より大きな値を持つ} \\ &\quad L_d \text{の総数} \\ C_{loss}(i) &= \text{通常の場合のロスをしたパケ} \\ &\quad \text{ットの総数} \end{aligned}$$

4 実装と評価

4.1 環境

実験環境は FreeBSD-2.2.8 上に, NTP TIMES-TAMP および ICMP TIMESTAMP を利用するアプリケーションを実装し評価した. パケットのサイズを 40[byte]とし, 前述したように, まずパ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]とした.

4.2 一般性

アクティブ計測による計測データは, (1) パケットを送信する時間帯, (2) パケットの送信総時間, および本手法で採用しているパスベース計測手法では, (3) パスの特異性, によって大きな影響を受ける. (3)については, 全 13 パス(その内, 海外を含むパスは 2 パス), および全ノード数は 170 ノード(その内, 国内が 150 ノード, 海外が 20 ノード)の実験により, ある程度特異性は取り除かれていると思われる. なお, (1),(2)については, その特異性が影響をおよぼしたパラメータについて, 以下のパラメータの説明の箇所で述べることにする. しかし, 本研究では, ノードを分類するための指標, 評価式, 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用いて評価を行っている.

4.3 測定パス

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.bbnplanet.net	3
9	p7-0.paloalto-nbr1.bbnplanet.net	3
10	p4-0-0.paloalto-br2.bbnplanet.net	3
11	f1-0-0.paloalto-cr11.bbnplanet.net	3
12	h2-0.oracle2.bbnplanet.net	3
13	bigip-www.us.oracle.com	-

測定パスは, 熊本大学から通過ノード数 13 で SINET と国際回線を含む海外のホストとした. 他

のホストでも同様の結果が得られているので, 以下ではこの例に基づきデータを解析する. 以下にそのパス上のホストを示す. また, そのホストが持つ NTP デーモンの stratum 番号を横に併記する. ただし, "-" は NTP デーモンが起動していないホストを示す. また, 今回実験をした 170 ノード中での独立なノードについてはほぼ 80%が NTP デーモンを起動している.

4.4 最小遅延

まず, 最大利用可能帯域が計算できる最小遅延について測定した. 最小遅延は帯域の大きさおよび送信ノードからの距離が相互に関係した指標である.

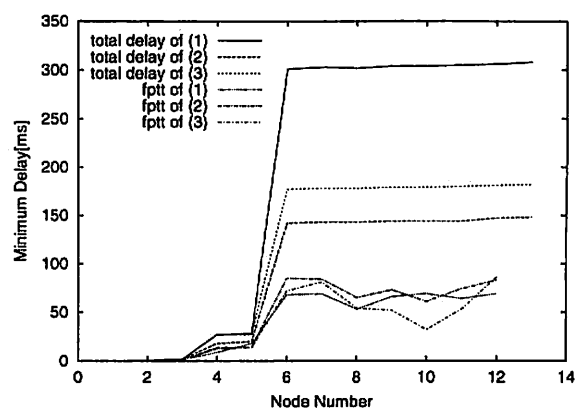


図 3. ノードによる最小遅延の変化

パス上のノードに関する最小遅延の変化を次の図に示す. 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. また, TIMESTAMP メッセージから得られた, 行きに要した最小遅延 (forwarding path trip time(fptt)) も図示する.

RTT に関する最小遅延の結果から, 国内から国外に出るノード 5 から 6 までの空間的距離が大きいところでの変化が大きいことがわかる. 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. 最小遅延は行だけの経路を計測した場合においても減少することはないことを前提とすると, NTP が動作しているノードにおいて, ノードがどの程度の誤差を保持している可能性があるかを fptt の計測により抽出することができる. 結果から, NTP が動作している場合においても, 30msec 程度の相対誤差が存在することがわかる. ここでは示していないが, 相対誤差は時間の推移によって大きくなる部

分もあったので、以下の実験では、最小遅延は減少しないという前提で、相対誤差の少ない部分のデータをサンプルデータとして利用した。

4.5 遅延の周期性

TIMESTAMPにより遅延を行き、および帰りに分離させると、行きの遅延については、次のよう周期的な変化となった。ここではノード 10 におけるパケットに付与した順序番号に対する遅延の変化を示す。ノード 8,9,10,11 についても同じように約 17sec の周期的な遅延が発生し、今回調査した国内 150 ノード中 82 ノード (54%)、海外 20 ノード中 9 ノード (45%) のノードが周期的な遅延が発生していた。

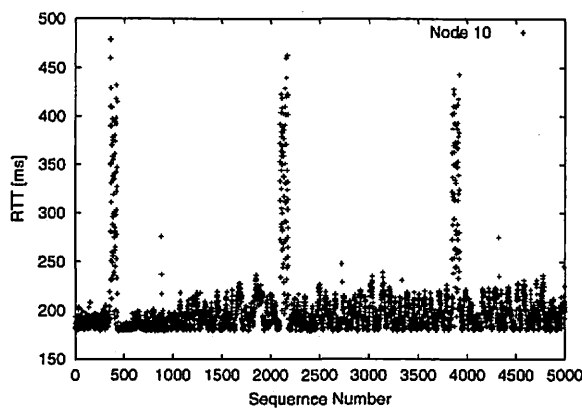


図 4. ノード 10 における行きに要した遅延の変化

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ことから、(1)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。(2) ノードにおける他の処理が周期的であるため、RTT の値に直接遅延として影響している。(3) 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る。ことが明確になった。したがって、ノードをスルーするパケットについては、(2) の処理が影響していない。ここで、ノード上の他の処理とは、BGP-4[6] における隣接 IS 間で送信される KeepAlive メッセージと考えられる。

4.6 ロスを含まない揺らぎ

すべてのパケットについて周期的な部分も含めて安定なエリアに存在するパケットが全体の中での程度の割合になるかを示したのが図 5 である。縦軸

を、安定領域に含まれるパケット数/全パケット数とした。横軸を NTPd が起動しているノードとした。ここで、行きおよび帰りについては $\alpha = 1.1, 1.3$ とした。

この結果から、行きについては、国際回線に出るノード 5 において、最小遅延の 1.1~1.3 倍の範囲において揺らぎが存在し、帰りは全ノードにおいて 1.1~1.3 倍の範囲において揺らぎが存在していることがわかった。

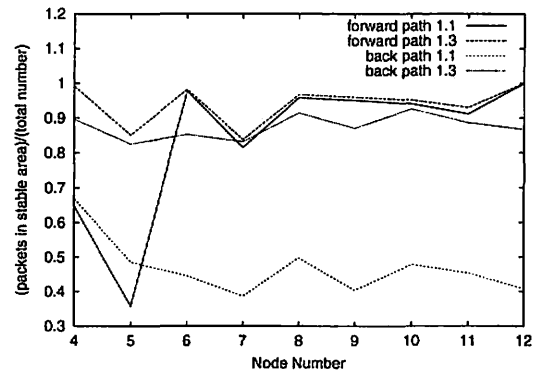


図 5. 行きおよび帰りにおける安定な領域の割合

4.7 ロスに関係した測定値

すべてのノードに対してロス率は約 1.1 という長さで一定となった。また、連続的なパケットロスも発生している。

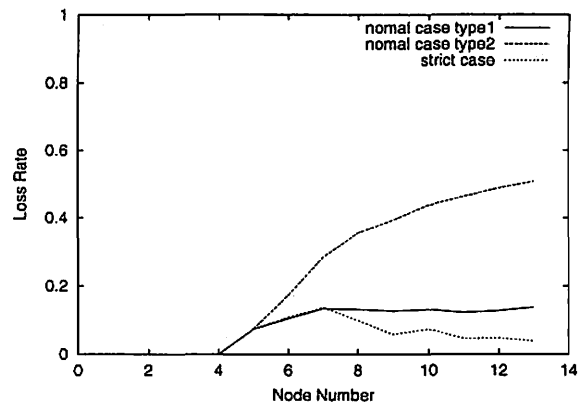


図 6. パケットロス率

通常および厳密な場合のロス率を図 6 に示す。type2 の場合、パスが進むにつれてロス率は上昇する。type1 の場合、LAN から外に出てノード 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。これより、パス上のノード 7 以降においてロス集中していることが分かる。厳密な場合、

ノード7でピークになっていることから、パケットロスの原因がノード7にあることがわかる。

4.8 ロスのパス依存性

揺らぎに関しては、ノード間の依存性がほとんど見られなかったが、ロスが多いノード7以降について同じグループ内におけるロスの連続性を調査した。図7は、同じグループに2個以上の連続するパケットの個数をノード別に集計した図である。開始パケットとは、そのノードから連続が始まっていることを示している。

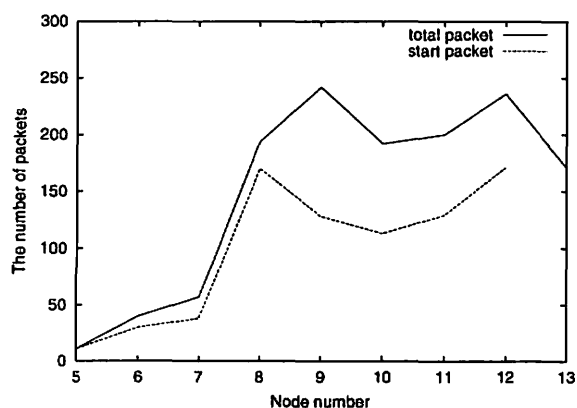


図7. 連続的なパケットロス

ロス率が低い状況でロスがパスに依存して発生していることから、ロスに関しては、後方依存性が存在すると考えられる。

4.8.1 事前安定性

$\delta = 2, 4, 10$ それぞれの場合におけるパケットの事前安定性を図8に示す。図から $\delta = 2$ 、すなわちパケットロスの間隔が3以上となる場合がロスパケットの事前状態の中で7割から8割存在することを意味する。また $\delta = 4$ 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パケットロスや比較的短い間隔でパケット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。したがって、観測する時間的な長さが指数関数的に増加すれば、それに比例して安定しない状態が発生することが分かる。

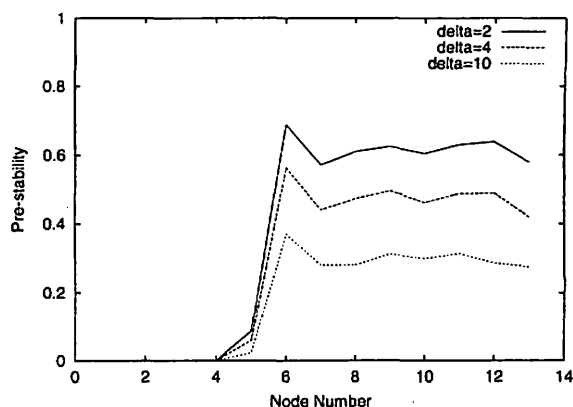


図8. 通常な場合のパケット事前安定性

5 おわりに

本研究では、1msec 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投げ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。特徴としては、(1) ロスを含まない遅延はパスに依存せず、ロスに関しては依存する傾向がある。(2) ロスを含まない遅延については、パケットを戻す処理に依存し、ルータ内の処理に比例して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。今後はNTPによる時間誤差による評価の幅を検討するとともに、数多くのノードの特徴を構成する要因を分析し、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] V. Paxson, G. Almes, J. Mahdavi and M. Mathis: "Framework for IP Performance Metrics", RFC2330, May, 1998.
- [2] G. Almes, S. Kalidindi, M. Zekauskas: "A One-way Delay Metric for IPPM", RFC2679, September, 1999.
- [3] G. Almes, S. Kalidindi, M. Zekauskas: "A One-way Packet Loss Metric for IPPM", RFC2680, September, 1999.
- [4] R. Koodli, R. Ravikanth: "One-way Loss Pattern Sample Metrics", <draft-ietf-ippm-loss-pattern-02.txt>, October, 1999.
- [5] 井原修一, 中嶋卓雄: "バースト的なパフォーマンス評価手法について", マルチメディア, 分散, 協調とモバイル (DICOMO 2000) シンポジウム論文集, pp.487-492, 2000.
- [6] Y. Rekhter: "A Border Gateway Protocol 4 (BGP-4)", RFC1771, March, 1995.