

毎正秒パケット到着間隔 (PAI) に基づいた時刻同期方式

町澤 朗彦^{†a)} 岩間 司[†] 鳥山 裕史[†]

Packet Arrival Interval (PAI) Based Time Synchronization

Akihiko MACHIZAWA^{†a)}, Tsukasa IWAMA[†], and Hiroshi TORIYAMA[†]

あらまし 近年, ネットワーク計測・分散制御分野では, マイクロ秒以下の同期精度が要求されている. 高精度な時刻同期では, クロックレートが変動するスキューとパケット交換ネットワークにおけるキューイング遅延が問題となっている. そこで, 本論文では, パケットの到着間隔による時刻同期方式 (PAI: Packet Arrival Interval based time synchronization) を提案する. 本方式では, 正確に一定間隔にブロードキャストされる同期パケットの到着間隔の平均がキューイング遅延の影響を受けにくいことを利用し, 高精度にクロックレートを推定する. また, マスタから各クライアントへの片方向遅延は, LAN の構成要素に関する事前知識から高精度に算出可能である. なお, マスタを GPS に同期させれば, 広域同期も可能である. 実験の結果, 回線利用度が 80% 程度のクロストラフィックの存在下でもマイクロ秒以下の同期精度が実現されることを確認した. しかも, 本方式は, 高い計算能力を必要としないため, センサネットワーク等の小さなノードでも利用可能である.

キーワード 高精度時刻同期, クロックレート, スキュー, パケット到着間隔, キューイング遅延

1. ま え が き

近年, ネットワーク計測・分散制御分野では, マイクロ秒精度の時刻同期が求められている [1], [2]. 例えば, 高精度な広域同期が必要な用途として, 時分割多重方式を採用している携帯電話システムでは, 屋外からの電波が届かない地下や大型ショッピングモールなどの屋内と, 屋外との間を通話を継続しながら移動 (ハンドオーバー) するためには, 屋内と屋外の携帯電話の基地局が高精度に同期している必要があり, その要求精度は, CDMA2000 で ± 3 マイクロ秒, WCDMA の時分割多重モードでは ± 1.25 マイクロ秒 [3] である. しかも, 屋内では電波の到達距離が短くなるため多数の基地局を必要とする.

しかし, 現在, インターネットでデファクトスタンダードとなっている NTP [4] ではミリ秒程度の精度しか得られない [5]. これは, (1) ネットワーク遅延時間が未知, (2) 時計の進み具合 (クロックレート) が変

動する (スキュー), (3) キューイング遅延の存在, などが主な原因である. NTP など, ネットワークを介した時刻同期では, 往路と復路の遅延時間が等しいと仮定している方式が多いが, 実際のネットワーク遅延は非対称であり [6], 現在のインターネット技術では, ネットワーク側から遅延時間情報は提供されていないため, 一般のインターネットを介した場合, 絶対的な同期精度は保証されない. 一方, GPS が利用できる環境では, マイクロ秒精度の時刻同期が可能であるが, GPS にはアンテナやコストの問題があるため, すべての機器に GPS を用いることは現実的ではない. また, スレーブが 1 台であったとしても, GPS アンテナに用いられる同軸ケーブルを新たに敷設するよりも, 既設 LAN を用いることができれば, 費用や作業量だけでなく, 構内光ファイバを用いた長距離伝送などメリットが大きい. したがって, GPS に同期した stratum 1 のマスタ時計が接続されている LAN において, 低コストで高精度な時刻同期方式が望まれている.

さて, 前記 NTP の課題 (1)~(3) は, インターネットだけではなく, LAN においても問題となる. まず, ネットワーク遅延時間は LAN でも, 往路と復路でトラフィックが異なるため非対称となるが, LAN では, ネットワークスイッチやケーブルなどの構成要素が既知で

[†] 情報通信研究機構 小金井市

National Institute of Information and Communications Technology, 4-2-1 Nukui-Kitamachi, Koganei-shi, 184-8795 Japan

a) E-mail: machi@nict.go.jp

あるため、遅延時間を事前知識から得ることが可能である。

これまで、スキューに対しては、制御理論に基づいたフィードバック系が多く用いられていたが [7], [8], フィードバック係数が経験的に決定されており、最適値を得ることが困難であった。一方、通信分野で、デジタル PLL よりもフィードバック係数を最適化できるカルマンフィルタが優れていることが示され [9], 最近, Bletsas はクロックレート推定にカルマンフィルタを応用している [10]。しかし、カルマンフィルタでは、計算能力を必要とする。

また、ネットワーク遅延計測分野では、スキューを補償するために、キューイング遅延が非負であることを利用し、キューイング遅延の影響を除去する方式が Paxson 及び Moon によって提案されている。Paxson は区分的最小値を利用し [11], Moon は線形計画法によって推定している [12]。しかし、高精度な時刻同期では、キューイング遅延だけではなく、ネットワーク機器による処理遅延の変動の影響が顕在化する。

なお、キューイング遅延そのものを発生させない研究も進められているが、ルータやネットワークスイッチなどに変更が必要であったり [2], [13], ネットワークを時分割で利用する方式 [14], [15] では、既存システムとの併用は不可能である。

そこで、本論文では、LAN における高精度時刻同期を目的として、毎正秒パケット到着間隔による時刻同期方式を提案する。本方式では、一定間隔にブロードキャストされた同期パケットの到着間隔がキューイング遅延の影響を受けにくいことを利用し、高精度な時刻同期を可能とする。なお、NTP のブロードキャストモードや IEEE1588 でも、同期パケットは一定の頻度で送出されるが、送出間隔に精度を要求しない点が提案方式と異なる。本方式では、マスタには同期パケット送出間隔に高い精度を要求するが、スレーブは整数の減算とビットシフト演算及びメジアンフィルタのみで実現されるため、センサネット等の計算能力の小さなスレーブにも実装可能である。実験の結果、回線利用率が 80% 程度のクロストラヒックの存在下でもマイクロ秒以下の同期精度が得られることを確認した。

本論文では、2. でクロックレート推定に関する従来方式の関係を整理し、3. で毎正秒パケット到着間隔時刻同期方式を提案する。4. では事前知識として必要となる、LAN の平均遅延時間について述べ、5. で提案方式の性能について実データを用いて評価する。

2. 関連研究

本章では、クロックレート推定に関連する研究をまとめる。従来提案されてきた方式は、数学的には最小二乗法と線形計画法に大別されることを示す。また、各方式の計算量についても整理する。

2.1 PLL とカルマンフィルタ

NTP で用いられている PLL/FLL は、フィードバック係数が経験的に決定されており、最適値を得ることが困難であったが、Driessen は二次のデジタル PLL と等価なカルマンフィルタを構成し、フィードバック係数を最適化できるカルマンフィルタが優れていることを示し [9], Bletsas はクロックレート推定にカルマンフィルタを応用した [10]。しかし、クロックレートは短期間では一定と考えられるが、推定値が一定となる場合は、カルマンフィルタは最小二乗法の逐次アルゴリズムと一致することが知られている [16]。したがって、Bletsas の方法は、最小二乗法の近似解にすぎない。一方、推定に用いるデータ点数を N とすれば、共役分散行列を作り直すため、計算量は $O(N^2)$ の乗除算を必要とする。

2.2 Paxson のアルゴリズム

Paxson は、キューイング遅延が非負であることを利用して、キューイング遅延の影響を除去した遅延時間の差の単位時間平均を用いている [11]。具体的には、まず、 N 個のデータを、 \sqrt{N} ずつ \sqrt{N} 個の小区間に分割し、各小区間の遅延時間の最小値を選び、任意の二つの最小値の差の単位時間平均を求める。すべての最小値の組合せから得られる平均値のメジアンをクロックレートの推定値とする。ここで、次章で述べるように、二つの遅延時間差の単位時間平均は、その期間のクロックレートの最小二乗法の解となる。ただし、小区間の最小値を選択する際に、スキューを補償していないため、スキューよりも小さいキューイング遅延を排除することはできない。一方、計算量は $O(\sqrt{N}!)$ の除算を必要とする。

2.3 線形計画法

Moon は線形計画法によって推定している [12]。すべての遅延時間データが、クロックレート及び時刻オフセットをパラメータとする直線より上方に存在する制約条件下において、データと直線との距離の総和を目的関数とする。しかし、高精度に遅延時間を計測する場合、キューイング遅延だけではなく、処理遅延の変動も影響するため、同制約条件で得られる最適解

は、実際のクロックレートとは異なる。一方、計算量は $O(N)$ の乗除算を必要とする。

3. PAI 時刻同期方式の原理

本章では、パケット到着間隔 (PAI: Packet Arrival Interval) に基づいた時刻同期方式の原理について述べる。3.1 でネットワークを介した時刻比較をモデル化し、3.2 で比較時に問題となるネットワーク遅延をモデル化し、3.3 でキューイング遅延を除去してクロックレートを推定する手法を提案し、3.4 でキューイング遅延及び処理遅延の変動を除去して時刻オフセットを推定する手法を提案する。

3.1 時刻比較モデル

マスタの時刻とスレーブの時刻をネットワークを介して比較した例を図 1 に示す。クロックレートが合っていないため、時間の経過に伴って、時刻差が広がっている (時刻オフセットのドリフト)。また、比較開始後 70 秒付近より、正方向にキューイング遅延が見られる。時刻同期ではスキュー及びキューイング遅延によるノイズの除去が主要課題である。

まず、マスタの時刻 t_i におけるスレーブ PC の示す時刻 $s(t_i)$ の間には、周波数源の種類に応じた短期間について、次式のような線形関係が存在する [17]。

$$s(t_i) = Rt_i + O \quad (1)$$

ここで、定数 R 及び O はそれぞれクロックレートと時刻オフセットである。さて、理想的に正確な時計では、時計の進む速さはマスタに等しく、 $R = 1$ であるが、現実の時計のクロックレートは周波数偏差 ρ を含む。

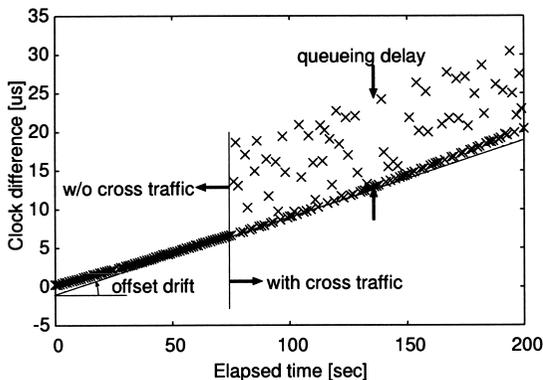


図 1 ネットワークを介した二つの時計の比較
Fig. 1 Clock difference via network.

$$R = 1 + \rho \quad (2)$$

なお、温度補償されていない水晶では、周波数偏差は 10 ~ 100 ppm 程度であるため、 $|\rho| \ll 1$ とする [11]。ただし、 $| \cdot |$ は絶対値を表す。通常温度補償されていない水晶によって駆動されている PC では、線形関係が保たれるのは数十秒以下である。

さて、マスタが毎正秒に時刻同期パケットをブロードキャストする場合を考える。時刻 t_i にマスタから送出されたパケットは、遅延時間 $d_i (\ll 1)$ 後にスレーブに到着する (図 2)。このとき、スレーブの時計の値 $s(t_i + d_i)$ は、

$$\begin{aligned} s(t_i + d_i) &= R(t_i + d_i) + O \\ &= s(t_i) + Rd_i \\ &= s(t_i) + d_i \quad (\because \rho d_i \cong 0) \end{aligned} \quad (3)$$

となる。ここで、簡便のため、以下、スレーブの受信時刻 $s(t_i + d_i)$ を y_i と記述する。

3.2 ネットワーク遅延モデル

本節では、式 (3) に示されるように、受信時刻に大きな影響を与える遅延時間をモデル化する。まず、遅延時間 d_i を固定成分 ξ 、キューイング遅延成分 η_i そしてその他の変動成分 ψ_i に分解する。

$$d_i = \xi + \eta_i + \psi_i \quad (4)$$

なお、図 2 では、 ψ_i は省略している。次に、クロストラフィックが存在せず、 $\eta_i = 0$ の場合の遅延時間平均を ξ_0 とする。

$$\xi_0 = \frac{1}{N} \sum_{i=1}^N d_i \quad (5)$$

ただし、 N は標本数である。また、同様に、 $\eta_i = 0$ の場合の変動成分 $\psi_{0,i}$ の平均と分散は以下のとおりで

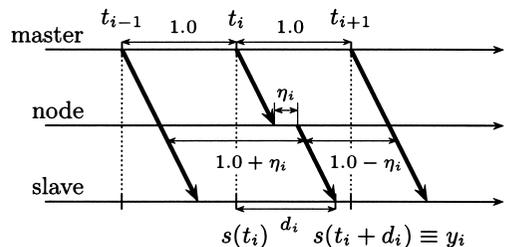


図 2 クロックモデル
Fig. 2 Clock model.

ある .

$$\bar{\psi}_0 = \frac{1}{N} \sum_{i=1}^N \psi_{0,i} = \frac{1}{N} \sum_{i=1}^N (d_i - \xi_0) = 0 \quad (6)$$

$$\sigma_{\psi_0}^2 = \frac{1}{N} \sum_{i=1}^N \psi_{0,i}^2 = \frac{1}{N} \sum_{i=1}^N (d_i - \xi_0)^2 \quad (7)$$

なお、本論文では LAN を対象としており、ネットワークを構成するスイッチの通過遅延の分布は平均値付近に集中しており、 $\pm 3\sigma_{\psi_0}$ の範囲にほとんど入っている (例えば、図 7)。また、これらのスイッチはハードウェアによって処理するため、通過遅延の平均も分散も負荷の影響は小さく (図 8 参照)、特に、負荷 80% 以下では、ほとんど一定であるため、遅延時間の固定成分とキューイング遅延以外の変動成分の統計的性質は一定と考える。

$$\xi = \xi_0 \quad (8)$$

$$\sigma_{\psi} = \sigma_{\psi_0} \quad (9)$$

一方、キューイング遅延は非負であり、

$$\eta_i \geq 0 \quad (10)$$

その大きさは σ_{ψ} よりも大きい場合が多く、平均は σ_{ψ} よりも十分大きい (付録参照)。

$$\bar{\eta} \gg \sigma_{\psi} \quad (11)$$

式 (1)、(4) より、次式を得る。

$$y_i = Rt_i + O + \xi + \eta_i + \psi_i \quad (12)$$

したがって、時刻同期問題は、ノイズ成分 η_i 及び ψ_i の存在下で、毎正秒にブロードキャストされる同期パケットの到着時を観測値 y_i として、スレーブのクロックレート R 及び時刻オフセット O を推定する問題に帰着する。また、本方式は、LAN を対象としており、固定遅延成分 ξ は事前知識として 4. で示す推定法により与えられる。ただし、クロックレートを一定と考えられるのは、温度補償されていない水晶では数十秒程度であり、ネットワークトラヒックへの影響を考慮すると同期パケットは毎秒一つ程度となるため、統計処理に用いることのできるデータ数が少ない。なお、ここでは、同期パケットが 1 秒間隔で送出される場合を示しているが、2 のべきで等間隔であれば、1 秒である必要はない。

3.3 クロックレート推定

ここで、パケット到着間隔に注目する。提案方式では、毎正秒にパケットは送出されるため、送出間隔は常に正確に 1 秒となり、

$$t_i - t_{i-1} = 1 \quad (13)$$

一方、図 2 では、時刻 t_i に送出された時刻同期パケットは、ネットワーク上の途中ノード (node) で η_i のキューイング遅延が生じている。このとき、先行時刻同期パケットとの間隔は η_i だけ伸びるが、逆に後続時刻同期パケットとの間隔は η_i だけ縮むことになる。もし、前後の時刻同期パケットにはキューイング遅延が含まれていなければ (処理遅延は小さいので、ここでは無視する)、時刻同期パケットの到着間隔を平均化すれば、キューイング遅延を除去することが可能となる。

パケット到着間隔 (PAI) を δ_i とすれば、式 (12)、(13) より、次のようにクロックレートにノイズの差分が重畳している。

$$\begin{aligned} \delta_i &= y_i - y_{i-1} \\ &= R(t_i - t_{i-1}) + \eta_i - \eta_{i-1} + \psi_i - \psi_{i-1} \\ &= R + \eta_i - \eta_{i-1} + \psi_i - \psi_{i-1} \end{aligned} \quad (14)$$

ここで、 δ_i の例を図 3 に示す。ただし、ダイナミックレンジが狭いため、縦軸は 1 秒からの差分 (PAI-1) でプロットしている。なお、図 3 は、図 1 と同じデータを用いている。図 3 より、 δ_i が一定値の上下にほぼ均等に散らばっていることが分かる。この一定値が求めるクロックレートである。そこで、最小二乗法に基づいてクロックレートを推定する。まず、 δ_i の推定値

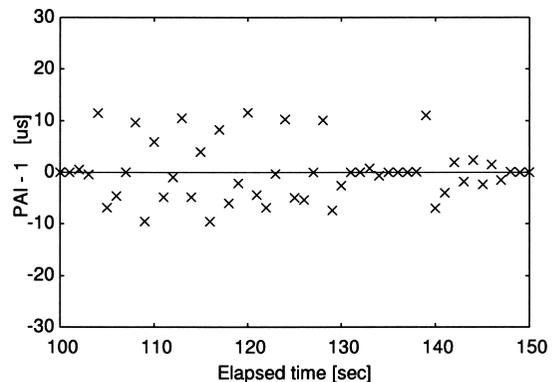


図 3 パケット到着間隔とクロックレート
Fig. 3 Packet arrival interval and clock rate.

を $\hat{\delta}_i$ とし、回帰直線を次式のように傾きが 0 の一定値とおく．

$$\hat{\delta}_i = R \quad (15)$$

次に、測定値 δ_i と推定値 $\hat{\delta}_i$ の予測残差の平方和を最小化する R を求める．クロックレートが一定と考えられるのは短時間であるため、近傍 N 点のデータに対する予測残差の平方和 J_i を最小化させる．

$$J_i = \sum_{k=0}^{N-1} (\delta_{i-k} - \hat{\delta}_{i-k})^2 \quad (16)$$

上式を R で偏微分し、正規方程式を得る．

$$\frac{\partial J_i}{\partial R} = -2 \sum_{k=0}^{N-1} (\delta_{i-k} - R) = 0 \quad (17)$$

したがって、クロックレートの推定値 \hat{R} は、

$$\begin{aligned} \hat{R} &= \frac{1}{N} \sum_{k=0}^{N-1} \delta_{i-k} = \frac{1}{N} \sum_{k=0}^{N-1} (y_{i-k} - y_{i-k-1}) \\ &= \frac{1}{N} (y_i - y_{i-N}) \end{aligned} \quad (18)$$

で与えられる．これは、区間 $[i - N : i]$ の両端での時計の差を、区間長 N で平均している．したがって、 N を 2 のべきとすれば、二乗誤差最小化規範による推定を、整数の減算とビットシフトのみによって実現可能である．

ところで、式 (18) では、区間の両端の値 y_i 及び y_{i-N} のみから推定しているため、ノイズの影響を受けやすい．そこで、次のように、同期パケット受信ごとに、近傍 $2N$ 点のデータを用いて、 N 個の推定値候補 $\hat{R}_i[k]$ ($k = 0, \dots, N-1$) を算出し、メジアンフィルタを施す．

$$\hat{R}_i[k] = \frac{1}{N} (y_{i-k} - y_{i-N-k}) \quad (19)$$

$$\hat{R}_i = \text{MEDIAN}(\hat{R}_i[k]) \quad (20)$$

なお、メジアンフィルタは算術平均よりもロバストであることが知られている．

3.4 時刻オフセット推定

まず、前節で得られたクロックレート推定値を用いて、式 (12) から導出される次式によって時刻オフセットのドリフトを補償した後、時刻オフセットを推定する．

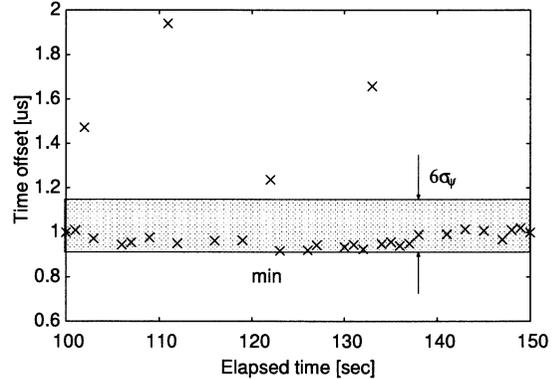


図 4 時刻オフセット推定
Fig. 4 Time offset estimation.

$$y_i - \hat{R}_i t_i = O + \xi + \eta_i + \psi_i \quad (21)$$

なお、遅延の固定成分 ξ は、次章で後述する事前知識により与えられる．図 1 からスキューを補償した結果を図 4 に示す．最小値付近に集中しているデータとそれらの上方に離散しているデータに分離されている．これは、キューイング遅延が非負であることによる．さて、処理遅延は平均値付近に集中しており、 $\pm 3\sigma_\psi$ の範囲にほとんど存在するが (図 7 参照)、式 (11) より、同範囲に含まれるキューイング遅延の確率は非常に小さい．なお、 σ_ψ も、次章で後述するように、事前知識により与えられる．したがって、最小値付近 $[\min(y_i - \hat{R}_i t_i), \min(y_i - \hat{R}_i t_i) + 6\sigma_\psi]$ のデータに対してメジアンフィルタを施すことにより、キューイング遅延と接続機器の処理遅延ノイズを取り除き、時刻オフセットを推定する．ここでも前節と同様に算術平均よりもロバストなメジアンフィルタを用いている．

4. 遅延の固定成分

本章では遅延時間から、キューイング遅延や処理遅延の変動成分などを除いた固定成分について検討する．経路に沿った固定成分の総和は、前章で導入したネットワーク無負荷時の平均遅延時間 ξ に相当する．まず、4.1 では、固定成分を構成要素に分解し、4.2 では、スイッチングハブの通過遅延について述べ、4.3 では、ケーブル遅延について述べる．最後に、4.4 で、 ξ 及び σ_ψ を求める計算式について述べる．

4.1 固定成分の構成要素

マスタからクライアントまでの遅延時間を分解すると、図 5 のようになる．マスタ内部で生成またはマスタ外部より与えられる正秒信号 (1 pps/s) を起点とし

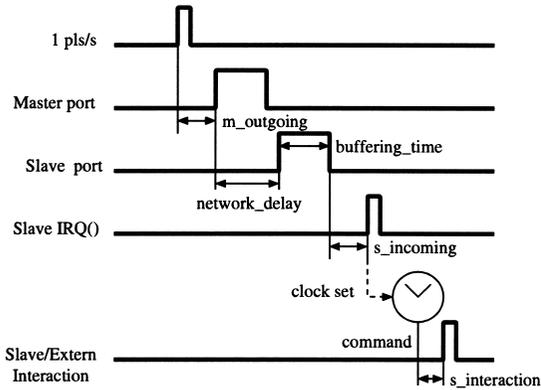


図 5 遅延の構成要素
Fig. 5 Delay components.

て、以下のような成分により構成されている。

- m_outgoing* : マスタでの正秒とネットワークポートからフレーム先頭が出力されるまでの時間
- network_delay* : ケーブルの伝搬遅延と接続装置の通過遅延
- buffering_time* : スレーブでのネットワークポートにフレーム先頭が到着してから末尾までの時間
- s_incoming* : スレーブでのネットワークポートにフレーム末尾が到着してから割り込みハンドラが呼出されるまでの時間
- s_interaction* : スレーブでの入出力命令発呼から実際の外界との入出力までの時間

これらの固定成分が事前知識として与えられていれば、式 (12) によりオフセットを決定することができる。本論文では、固定成分を実測するが、将来的には、構成機器製造ベンダより提供されるべきである。

ただし、スレーブの *buffering_time* はリンク速度とフレームサイズによって計算により求めることができ、 $buffering_time = \text{フレームサイズ [bit]} / \text{リンク速度 [bit/s]}$ である。一方、スレーブ PC は外界との入出力（インタラクション）によってのみ、内部の状態（時計）を知ることができる。したがって、入出力遅延の和 $s_incoming + s_interaction$ として扱う。

以下の節で、*network_delay* にかかわるスイッチングハブ及びケーブルについて詳細に検討する。

4.2 スwitchingハブ

スイッチングハブの通過遅延を推定するために必要

表 1 テストしたハブの遅延特性
Table 1 Latency of HUBs.

HUB	Variance of residuals [ns ²]	b [ns]	a [ns/bit]
(C)CG-SW05GTPV2	21	736	1.000
(B)LSW-GT-5NS	29	759	1.000
(N)GS105JP	30	737	1.000
(G)GH-EHG5MRX	5000	1042	1.000
(L)LB-GS05TP	7375	1013	1.001

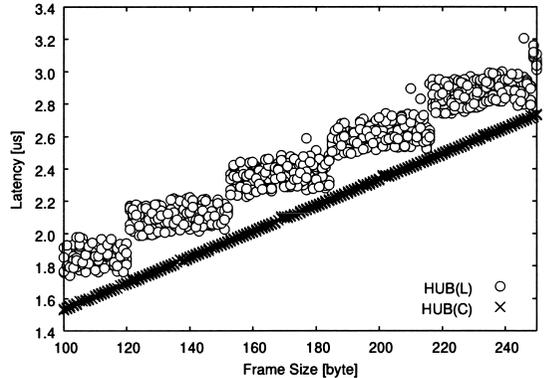


図 6 スwitchingハブによる遅延時間特性の相違
Fig. 6 Latency of switching hubs.

な項目は次のとおりである。

- パケット長依存性
- ポート依存性
- 方向依存性
- ブロードキャスト/ユニキャスト依存性
- 負荷依存性

表 1 に、現在入手可能な普及価格帯の主な 5 ポートのギガビット対応ノンインテリジェントスイッチングハブについて、通過遅延時間を計測した結果を示す。なお、伝搬時間の計測には PUTS [18] を用いた。一般に、通過遅延時間は、パケットサイズに比例するため、表 1 では、各ハブについて、パケットサイズと通過遅延時間に対する回帰直線の係数 (a : 傾き, b : y 切片) と分散を示している。また、パケットサイズをパラメータとした通過遅延時間の例を図 6 に示す。ハブ (C) は通過遅延がパケットサイズに比例しているが、ハブ (L) はブロック状の分布を示している。なお、通過遅延の測定では、各ハブの 1 番ポートから入力し、2 番ポートから出力させている。ここで、ハブ (C) 及びハブ (L) の通過遅延の分布を図 7 に示す。横軸は通過遅延時間、縦軸は頻度を表している。測定には、パケット長 1500 バイトの IPv4/UDP パケットを用い、

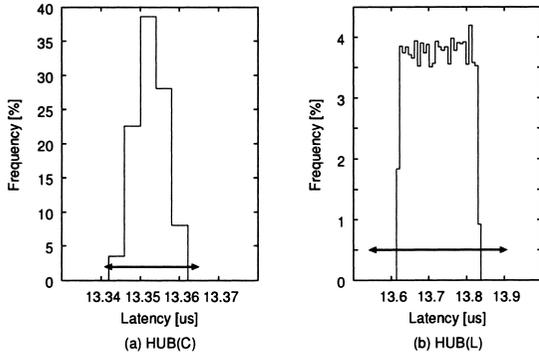


図 7 スイッチングハブ通過遅延分布

Fig. 7 Latency distribution of switching hubs.

表 2 テストしたハブ (C) のポート遅延特性
Table 2 Latency of the port of HUB (C).

		Output Port				
		1	2	3	4	5
Input Port	1		736	743	739	755
	2	737		740	742	752
	3	741	743		738	761
	4	742	734	739		764
	5	761	753	749	751	

1 万パケットの通過遅延からヒストグラムを作成した。ジッタの小さなハブ (C) ではヒストグラムのピン幅を 4 ns とし (図 7(a)), ハブ (L) ではヒストグラムのピン幅を 8 ns としている (図 7(b)). また, 図中, 水平方向の矢印は平均から $\pm 3\sigma_{\psi}$ の範囲を示している。通過遅延の分布は平均値付近に集中しており, $\pm 3\sigma_{\psi}$ の範囲にほとんど入っていることが分かる。

次に, ポート依存性について実験した。ここでは, 通過遅延が最小であるスイッチングハブ (C) の結果についてのみ報告する。ハブ (C) は五つのポートを有しており, 各ポート間での通過遅延について, 回帰分析した結果を表 2 に示す。1 番~4 番ポートでは特性に相違は見られず, 5 番ポートを利用した場合のみ, 10~20 ns ほどの遅延時間の増大が見られた。しかし, 分散及びパケットサイズに対する係数 a に差異は認められなかったため, 表には示していない。また, 同じ二つのポートの組合せでは, 方向による差異は見られなかった。

また, ユニキャスト及びブロードキャストフレームでの遅延時間も計測したが, 両者に差は見られなかった。しかも, リンクアップしているポートの個数のブロードキャストパケットフレームの遅延時間への影響も見られなかった。

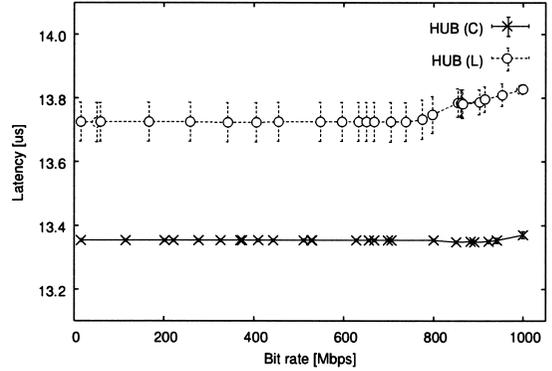


図 8 スイッチングハブ通過遅延負荷特性

Fig. 8 Load characteristics of switching hub latency.

表 3 ケーブル遅延特性
Table 3 Latency of cable.

Length [m]	Outward [ns]	homeward [ns]
20	100	104
40	208	211
60	316	305
80	416	419
100	520	519

更に, ハブ (C) 及びハブ (L) の, 負荷依存性を図 8 に示す。横軸は通過するデータ量, 縦軸は通過遅延時間, エラーバーは通過遅延時間の標準偏差を示している。ハブを介して 2 台の PC を接続し, 片方の PC からパケットサイズ 1500 バイトの IPv4/UDP パケットを一定間隔で送出している。パケット送出間隔を短くすると, 単位時間当りのデータ量が増加する。図 8 より, 通過遅延の平均も分散も負荷の影響は小さく, 特に, 負荷 80%以下では, ほとんど一定である。

4.3 ケーブル

CAT6 ケーブルの伝搬遅延時間は, 100 m の長さに対し 538 ns 以下と規定されている。例として, DTL-C6X (沖電線製) について実測した。表 3 に, ケーブル長を変化させて, 伝搬遅延を実測したデータを示す。更に, この測定データから, 回帰分析によって 1 m 当りの伝搬遅延を推定したと, 5.2 [ns/m] となった。

次に, 理論的に伝搬遅延を求める。まず, UTP ケーブルにおける電気信号の伝搬速度の理論値は, 被覆絶縁材料の比誘電率を ϵ とすれば, 次式で与えられる。

$$v = c/\sqrt{\epsilon} \tag{22}$$

ここで, c は光速。UTP ケーブルに用いられている FEP の比誘電率は, 約 2.1 である。また, 芯線の長さはよりの強さに影響されるが, 信号の干渉を防ぐ

ために芯線ペアごとによりの強さが異なるため、1本のケーブルでも、芯線の長さは異なっている。1mのDTI-C6Xの芯線を抜き取って、長さを測ってみたところ、次のようになっていた。

- 橙 107\,cm 橙/白 105\,cm
- 緑 105\,cm 緑/白 104\,cm
- 青 111\,cm 青/白 111\,cm
- 茶 110\,cm 茶/白 109\,cm

したがって、1mのUTPの信号伝達遅延の理論値は、最短ペアで5.02 ns/m、最長ペアで5.36 ns/mである。結局、5.2 ns/mが妥当と考えられる。

4.4 固定成分の推定値

本章のこれまでの議論により、遅延の固定成分の推定値 $\hat{\xi}$ は以下の計算式によって求められる。

$$\begin{aligned} \hat{\xi} = & m_outgoing + network_delay \\ & + buffering_time + s_incoming \\ & + s_interaction \end{aligned} \tag{23}$$

$$buffering_time = S/L \tag{24}$$

$$network_delay = \sum_{k=1}^H (a_k S + b_k) + \sum_{k=0}^H 5.2l_k \tag{25}$$

ただし、 k 段目のスイッチングハブのフレームサイズ遅延特性 a_k 及び b_k 、また、 $m_outgoing$ 及び $(s_incoming + s_interaction)$ は、機器のデータシートなどにより与えられる。また、 S は時刻同期パケットのフレーム長、 L はスレーブのリンク速度、 H はスイッチングハブの段数、そして、 l_k は、 k 段目と $k+1$ 段目のスイッチングハブを接続しているケーブル長である。

また、処理遅延の標準偏差は、 k 段目のスイッチングハブの処理遅延の分散を $\sigma_{\psi_k}^2$ とすれば、分散の加法性を用いて次式で与えられる。

$$\sigma_{\psi}^2 = \sum_{k=1}^H \sigma_{\psi_k}^2 \tag{26}$$

5. 性能評価及び考察

本章では、提案方式の性能について評価する。まず、5.1 で耐キューイング性能について調べ、次に、5.2 でクロックレートの推定精度について、提案方式と線

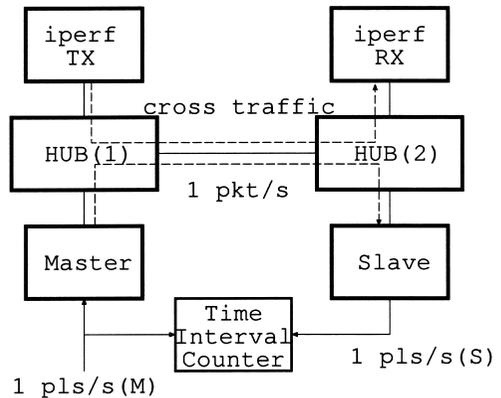


図9 性能評価システム構成図
Fig.9 Performance evaluation system.

形計画法と比較する。更に、5.3 で初期特性について NTP と比較する。5.4 で最後に、遅延の固定成分の影響について考察する。

5.1 耐キューイング精度

図9に示す構成で提案方式の性能を評価した。マスタは外部の原子時計からの1 pls/s(M)で駆動し、2台のスイッチングハブを介してスレーブの同期を行い、スレーブからのパルス出力1 pls/s(S)と時刻源の1 pls/s(M)をインタバルカウンタにて時刻差を計測した。スレーブは、自時計の毎正秒にパルスを出力することにより、マスタとスレーブの時刻差を計測することが可能となる。一方、クロストラヒックは、iperf (Ver. 1.1.1)^{注1)}のUDPモードを用いて、1500バイト長のIPv4/UDPパケットを所定レートで発生させ、ハブ(1)で、キューイングが生じる。

なお、本論文では、マスタとしてPUTSを用いている。PUTSは、通過型タイムスタンプとして開発されているが、付加機能として、任意のビットパターンのイーサネットフレームを外部1 pls/s信号に同期して送出することが可能であり、PUTSでは $m_outgoing = 260$ ns一定である。また、本論文では、拡張デジタル出力ボード(インタフェース社PCI-2703)によってパルス出力を行い、実測した結果、 $s_incoming + s_interaction = 5025$ nsを得た。また、スレーブは、ソフトウェアタイムスタンプの精度を高めるために、割込みハンドラでプロセッササイクルカウンタを取得している[19]。スレーブの諸元を表4に示す。また、ハブ(1)及びハブ(2)は、ともに

(注1): <http://dast.nlanr.net/projects/Iperf/>

表 4 測定に用いた PC の諸元
Table 4 Specification of the PC.

Mother board	Supermicro P4DPR-iG2
CPU	Intel Xeon 2.4 GHz
Network controller	ADMtek AN983(B) drv tulip-0.9.14
OS	linux 2.4.18

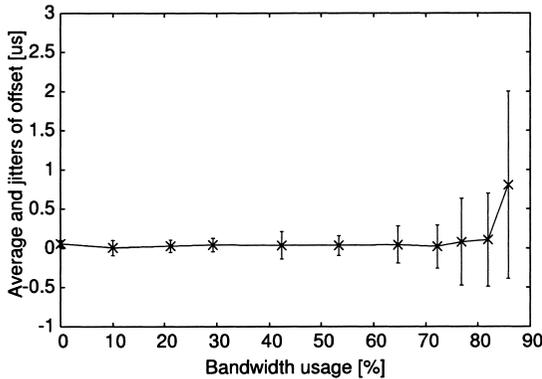


図 10 クロストラフィックによる同期精度の低下
Fig. 10 Synchronization accuracy deteriorated by cross traffic.

表 1 で、最も通過遅延の小さいスイッチングハブ (C) を用いている。

クロストラフィックの量を変化させて、同期精度を計測した結果を図 10 に示す。図 10 の横軸はクロストラフィックのデータ量を回線使用率で表している。また、縦軸は、スレーブの毎正秒パルス出力と規準となる原子時計の毎正秒パルスとの時間差 (時刻オフセットに相当) の平均を示している。更に、ジッタの標準偏差をグラフ中にエラーバーで示した。クロストラフィックの回線使用率 80% 以下ではマイクロ秒以下で同期していることが確認された。しかも、クロストラフィックの回線使用率 50% 以下では 0.1 マイクロ秒以下で同期している。一方、ネットワークの回線使用率が 80% を超えると、急激に同期精度が劣化してしまう。

回線使用率が 80% を超えると、急激に同期精度が劣化してしまう理由は、クロックレート推定の誤差が大きくなるために、最終的な時刻オフセットの推定値の精度が劣化すると考えられる。提案方式では、式 (20) のように、 N 個の標本ペアが用いられるが、パケットがキューイングされずにハブを通過する確率は、回線の空き確率であるので、回線使用率を θ とすれば、 y_{i-k} と y_{i-N-k} の両パケットともキューイングされていないペアの個数の平均 p は、

$$p = N(1 - \theta)^2 \quad (27)$$

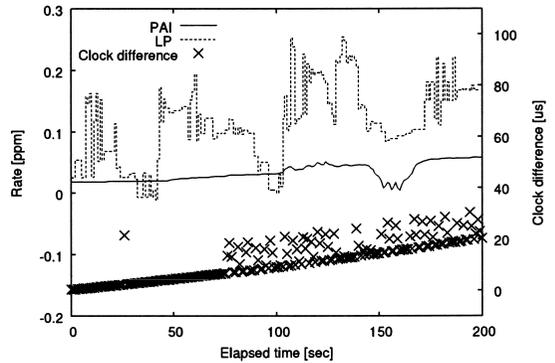


図 11 クロックレート推定性能の比較
Fig. 11 Comparison of clock rate estimation.

で与えられる。 $p < 1$ のとき、 y_{i-k} と y_{i-N-k} の両パケットともキューイングされていないペアはほとんど存在せず、クロックレートの推定は困難となる。このような回線使用率 θ_1 は標本数 N の関数となる。

$$\theta_1 > (1 - 1/\sqrt{N}) \quad (28)$$

ここで、 $N = 32$ とすると、 $\theta_1 > 82.3\%$ となる。一方、図 10 では、正に、82% 付近から急激に精度が劣化している。更に、クロックレート推定の誤差が大きくなると、式 (21) による時刻オフセットのドリフトを十分に補償することができず、最終的な時刻オフセットの推定の精度が劣化すると考えられる。

また、クロストラフィックがバースト的である場合、回線使用率が瞬間的に 80% を超える場合もあるが、影響は小さいと考えられる。これは、上で述べたように、本方式の精度がキューイングされずに通過した同期パケットの割合に依存していると考えられるが、クロストラフィックと時刻同期パケットの到着は独立過程であり、時刻同期パケットがキューイングされずにハブを通過する確率は、瞬間的なキュー長ではなく、平均的な回線使用率に依存すると考えられるためである。

5.2 クロックレート推定性能

従来方式の中では、最もクロックレート推定性能が高いと考えられている線形計画法 (LP) と提案方式 (PAI) によるクロックレート推定値を比較した結果を図 11 に示す。参照データ点数は、線形計画法で 64 点の窓とし、1 秒ずつ窓をずらした。同様に、提案方式では $N = 32$ とし、参照データ点数をそろえている。なお、評価には図 1 と同じデータを用いており、右縦軸は時刻同期をしない場合の、マスタとスレーブの時計をネットワークを介して比較した差 (clock

difference) を示している。実験システムの構成は前節と同じく、図 9 に示す構成を用いて、クロストラヒックのない状態で十分同期してから計測を開始し、70 秒経過後から、クロストラヒックとして iperf により、1500 バイト長の IPv4/UDP パケットを、回線使用率 50% の一定レートで流している。また、線形計画法は GNU 線形計画法キット^(注2)のシンプレックス法を利用した。

図 11 より、提案方式の推定値の方が安定していることが分かる。これは、高精度時刻同期では、処理遅延の変動が無視できなくなるため、線形計画法で仮定している“ノイズは非負”が成り立たないためと考えられる。また、線形計画法の解は制約条件の張る凸多面体の頂点を移動するため、推定値は離散値となり、図 11 のようにステップ状となる。更に、評価データは、途中からキューイング遅延を含んでおり、提案方式の推定精度は若干低下しているが、線形計画法の変動よりは小さい。

5.3 初期特性

NTP の初期特性と比較する。ntp (Ver. 4.2.0)^(注3)を用いて、図 9 と同様の構成を用いた。ただし、クロストラヒックのない状態で計測した。また、NTP のポーリング間隔を 16 秒固定とし、ntp サーバには PUTS を用いた。PUTS の SNTP サーバ精度は 8 ナノ秒である [20]。実験結果を図 12 に示す。横軸は同期開始からの経過時間を示しており、縦軸は、スレーブの毎正秒パルス出力と規準となる原子時計の毎正秒パルスとの時間差を示している。

NTP では、PLL/FLL のフィードバック係数が小さいため、初期化に 1 時間ほど要する (図 12)。一方、提案方式は、プログラム起動後、数秒でマイクロ秒以下の精度に達する。

5.4 固定成分の影響

遅延の固定成分 ξ の及ぼす影響について考える。式 (21) で、固定成分 ξ を独立して扱わず、オフセットと固定遅延成分の和 $O + \xi$ を一体化して、仮想的なオフセットとみなした場合、 ξ 相当分の誤差を含むことになる。イーサネットの最小フレーム長は 84 バイトであるため、リンク速度が 1 Gbit/s の場合、表 1 で測定したスイッチングハブでは 1 段で最低、 $84 [\text{byte}] \times 8 [\text{bit}] \times a [\text{ns/bit}] + b [\text{ns}] = 1408 [\text{ns}]$ の遅延が生じ、ケーブル遅延も 100 m で 520 ns 生じる。したがって、少なくとも $2 \mu\text{s}$ 程度の誤差を含むことになり、本論文で対象とするマイクロ秒以下の高精度同

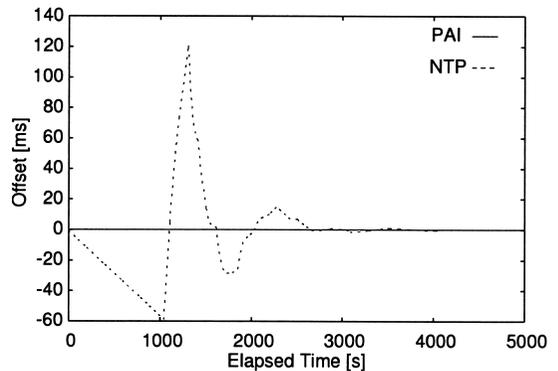


図 12 同期初期特性
Fig. 12 Initial character of synchronization.

期では ξ をオフセットと一体化して考えることはできない。しかし、もし、必要とする精度が数 μs 程度ならギガビットイーサネットで、数十 μs 精度程度ならファストイーサネットでも、 ξ を考慮せずに簡便にオフセットを推定することが可能である。

6. む す び

LAN 上での高精度時刻同期方式を開発した。正確に一定間隔でブロードキャストされる時刻同期パケットの到着間隔を利用することにより、パケット交換ネットワークで生じるキューイング遅延の影響を除去し、高精度に同期できることが確認された。回線使用率 80% 程度のクロストラヒックが存在する場合でも、マイクロ秒以下の精度で同期可能である。また、本方式では、スレーブは整数の減算とビットシフト演算及びメジアンフィルタのみで実現され、初期同期時間も短いため、センサネット等の計算能力の小さなスレーブにも実装可能である。しかも、本方式は、ネットワーク全体の同期システムを置き換える必要はなく、NTP などの他システムとの混在が可能である。本方式では、マスタを GPS に同期させることにより、同期系全体を広域同期を得ることができる。

今後は、ユビキタス環境に向けて、無線 LAN での高精度時刻同期に展開する予定であるが、チャンネル競合及び電波干渉などが問題となる。また、単独の GPS では大気の影響によって、100 ナノ秒以下の精度での時刻同期は困難であるため、光ファイバを用いた広域高精度時刻同期の研究も進める予定で

(注2): <http://www.gnu.org/software/glpk/glpk.html>

(注3): <http://www.ntp.org/>

ある .

文 献

- [1] S. Johannessen, "Time synchronization in a local area network," IEEE Control Syst. Mag., vol.24, no.4, pp.61–69, 2004.
- [2] "IEEE1588: IEEE standard for a precision clock synchronization protocol for networked measurement and control systems," IEEE Standard, 1588, 2002.
- [3] S. Rodrigues and A. Pietilainen, "IEEE-1588 telecommunications applications," Proc. of the 2005 IEEE 1588 Conference, Zurich, 2005.
- [4] D.L. Mills, "Network time protocol (version 3)," RFC 1305, IETF, 1992.
- [5] D.L. Mills, "Precision synchronization of computer network clocks," ACM Comput. Commun. Rev., vol.24, no.2, pp.28–43, 1994.
- [6] K.C. Claffy, G.C. Polyzos, and H.-W. Braun, "Measurement considerations for assessing unidirectional latencies," Internetworking: Research and Experience, vol.4, no.3, pp.121–132, 1993.
- [7] D.L. Mills, "Adaptive hybrid clock discipline algorithm for the network time protocol," IEEE/ACM Trans. Netw., vol.6, no.5, pp.505–514, 1998.
- [8] J. Levine, "Time synchronization over the internet using an adaptive frequency-locked loop," IEEE Trans. Ultrason., Ferroelectr. Freq. Control, vol.46, no.4, pp.888–896, 1999.
- [9] P.F. Driessen, "DPLL bit synchronizer with rapid acquisition using adaptive kalman filtering techniques," IEEE Trans. Commun. Technol., vol.42, no.9, pp.2673–2675, 1994.
- [10] A. Bletsas, "Evaluation of kalman filtering for network time keeping," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol.52, no.9, pp.1452–1460, 2005.
- [11] V. Paxson, "On calibrating measurements of packet transit times," Proc. ACM SIGMETRICS 1998, pp.11–21, 1998.
- [12] S.B. Moon, P. Skelly, and D.F. Towsley, "Estimaion and removal of clock skew from network measurements," Proc. IEEE INFOCOM '99, pp.227–234, New York, 1999.
- [13] 北口善明, 町澤朗彦, 鶴 正人, 尾家祐二, 箱崎勝也, "パケット自己消滅法によるネットワーク時刻同期精度の向上," 情処学論, vol.46, no.4, pp.1017–1024, 2005.
- [14] H. Kopetz and G. Bauer, "The time-triggered architecture," Proc. IEEE, vol.91, no.1, pp.112–126, 2003.
- [15] 福富 実, "時刻同期方式," 特許 2705665, 1995.
- [16] 添田 喬, 中溝高好, 大松 繁, 信号処理の基礎と応用, 日新出版, 1979.
- [17] 北口善明, 町澤朗彦, 中川普一, 西村道明, 日迫 彰, 箱崎勝也, "PC における時刻精度の精密計測とその評価," 信学技報, NS2003-160, 2003.
- [18] 町澤朗彦, 鳥山裕史, 岩間 司, 金子明弘, "通過型高精度 UDP タイムスタンプの開発," 信学論 (B), vol.J88-B,

no.10, pp.2002–2011, Oct. 2005.

- [19] 町澤朗彦, 北口善明, "割込みハンドラと高精度 PC によるソフトウェアタイムスタンプの精度改善," 信学論 (B), vol.J87-B, no.10, pp.1678–1685, Oct. 2004.
- [20] 鳥山裕史, 町澤朗彦, 岩間 司, 金子明弘, "ハードウェア SNTP サーバの開発," 信学技報, CQ2005-12, 2005.
- [21] C. Williamson, "Internet traffic measurement," IEEE Internet Computing, vol.5, no.7, pp.70–74, 2001.

付 録

式 (11) の根拠

まず, 1 段のスイッチングハブで生じる平均キューイング時間を, クロストラヒックのパケット長分布から推定する. インターネットを流れているトラヒックのパケット長は, 長短に偏在していることが知られており [21], これは, データパケットの多くが MTU サイズとなる一方, コントロールパケットの多くが TCP ヘッダのみであり, その割合は, 長いパケットが約半数, 短いパケットが約 40% となり, 残りの 10% ほどはランダムなパケット長である. 本論文では, インターネットではなく, LAN を対象としているが, パケット長の分布を同様と仮定する. さて, ここでは, 平均キューイング時間がスイッチングハブなどの処理遅延の変動よりも十分大きいことを示せばよい. 更にモデルを簡略化し, ランダムな長さのパケットを無視し, 長短パケットの割合を 50%/50% とする. この場合, 短パケットの割合が増えるため, 推定される平均キューイング時間は短くなる. また, キュー長も最長 1 と仮定する. キューに複数のパケットがたまっている場合よりも, やはり, 推定される平均キューイング時間は短くなる.

さて, 時刻同期パケットがキューイングされたときに, 回線を占有しているパケットが長パケットである確率は, $P_l = L_l / (L_l + L_s)$ である. ここで, L_l, L_s は, それぞれ, 長パケット及び短パケットのフレーム長である. イーサネットのプリアンブルや FCS (Frame Check Sum) などのオーバヘッドを考慮すると, 長パケットのフレーム長 (L_l) は $L_l = 1538$ バイトであり, リンク速度を 1 Gbit/s とすれば, ポート占有時間 (T_l) は $T_l = 12304$ ns である. また, 短パケットのフレーム長及びポート占有時間は, それぞれ, $L_s = 84$ バイト及び $T_s = 672$ ns である. これらの値を代入すれば,

$$P_l = L_l / (L_l + L_s) = 94.8\% \quad (\text{A-1})$$

である. 同様に, 短パケットが回線を占有している確

率は次のようになる．

$$P_s = L_s / (L_l + L_s) = 5.2\% \quad (\text{A}\cdot 2)$$

一定長パケットによるキューイング時間の分布は一様分布となり，その平均は（ポート占有時間/2）である[18]．したがって，平均キューイング時間 $\bar{\eta}$ は少なくとも，

$$\bar{\eta} = P_l * T_l / 2 + P_s * T_s / 2 = 5.85 \text{ us} \quad (\text{A}\cdot 3)$$

と推定される．

一方，表 1 より，個々のスイッチングハブの処理遅延の標準偏差は 100 ns 以下であるため，

$$\bar{\eta} \gg \sigma_\psi \quad (\text{A}\cdot 4)$$

と考えることができる．この関係は，スイッチングハブの段数が増えても維持される．なんとなれば，平均キューイング時間は加算的であるが，標準偏差は 2 乗で加算的（分散の加法性による）であることによる．

（平成 18 年 1 月 18 日受付，5 月 17 日再受付）



鳥山 裕史（正員）

昭 56 名工大・情報卒．昭 58 名大大学院情報工学専攻博士課程前期課程了．同年郵政省電波研究所（現情報通信研究機構）入所．平 2～5 ATR 通信システム研究所．平 5～6 ドイツテレコム研究所客員研究員．画像符号化，情報通信などの研究に従事．



町澤 朗彦（正員）

昭 59 上智大・理工・電気電子卒．同年郵政省電波研究所（現情報通信研究機構）入所．平 6 科学技術庁に出向し，IMnet 立上げに参与．平 8～11 Univ. Canterbury 客員研究員．平 15 JGN2 立上げに参与．画像の高速率符号化，視覚情報処理，計算機ネットワーク，時刻同期の研究に従事．



岩間 司（正員）

昭 58 山梨大・工・電子卒．昭 60 東工大大学院修士課程了．同年郵政省電波研究所（現情報通信研究機構）入所．以来，電波伝搬特性解析，移動通信のセル構成，標準時，時刻認証基盤技術の研究に従事．現在，電磁波計測部門タイムアプリケーショングループ主任研究員．平 2 本会篠原記念学術奨励賞受賞．IEEE 会員．