ハードウェアSNTPサーバの開発

鳥山　裕史[†]　町澤　朗彦[†]　岩間　司[†]

Development of a Hardware SNTP Server

Hiroshi TORIYAMA[†], Akihiko MACHIZAWA[†], and Tsukasa IWAMA[†]

あらまし　パソコン，情報家電機器等の普及に伴い，手軽に時刻情報を得たいという需要が増加している．インターネットでの時刻情報取得はNTPまたはSNTPによるのが一般的であるが，今後の需要に見合ったサーバ能力及び時刻供給精度を確保していくためには，高性能なサーバの開発が不可欠である．本論文では，筆者らの開発した高精度，高スループットなハードウェアSNTPサーバについて述べ，また，その性能を評価する．本サーバは，正確な基準時計の使用を前提としたStratum 1専用機であり，処理速度はGbEワイヤスピード，タイムスタンプ精度は8ナノ秒の性能を有する．本サーバには，このような高精度，高スループットという点以外に，過負荷対策が不要，クラック対策が不要といった特長もあり，これによりサーバ運用コスト低減も期待できる．

キーワード　NTP，時刻同期，タイムスタンプ，FPGA

1. まえがき

電子商取引や電子行政手続の普及に向け，時刻を安全かつ正確に把握することが，ますます重要となってきている．このような用途では，時刻情報取得に多少の費用がかかっても，「信頼できる」ことが重視される．一方，パソコン，家庭用ブロードバンドルータ，その他情報家電機器の普及に伴い，手軽に時刻情報を得たいという需要も急増している．このような目的では，NTP [1], [2] 若しくはその簡易版であるSNTP [3] によって時刻情報を取得する方法が一般的であり，実際，一般ユーザ向けに公開されているサーバへのアクセスも急増している [4]．サーバ側では，サーバ機材を増設し，DNSラウンドロビンまたは負荷分散装置によって負荷の増加に対応しているケースも多いが，このような方法では，アクセス増に伴う機器コスト，管理コスト等の増大が問題となる．NTPは，本来，サーバの階層化によって負荷集中を避けるように考慮されているが，中間階層のサーバであっても，正式なサービスとして運用する場合には，相応の運用コストが発生し，そのサーバ台数に応じた総コストは必ずしも低くない．multicastを用いれば，サーバ負荷の問題は生じないが，現状の家庭用インターネット環境ではmulticastが利用できないケースも多い．このようなことから，少ない台数でアクセス増に対応でき，かつ，機器コスト，管理コストの低いサーバの開発が必要であると考えられる．

このような実際の需要増以外にも，バグ等のトラブルでNTPサーバにアクセスが集中するケースもある [5]．もちろん，このような不適切なクライアントは修正されるべきであるが，今後も，同様の可能性は存在し，高スループットのサーバであれば，そのような場合でもサービスが継続できる可能性が高い．

NTP以外の時刻供給方法としてIEEE1588 [6] があり，高精度な時刻供給デバイス等も開発されているが，これらは主にLANなど近距離での利用を前提としたものであり，インターネット経由でのパソコン時計合せなどには，あまり利用されていない．

NTPサーバとしては，ntpdが広く用いられている．これは複雑なアルゴリズムを実装したソフトウェアであるが，その複雑さのほとんどは，複数のサーバからの時刻情報を処理し，ローカルな仮想時計を適切に保つ部分に由来している．この複雑な処理を省略したものが，SNTPサーバ，またはクライアントであり，特に，SNTPクライアントは，パソコンのOSや，家庭用ルータなどに組み込まれ，広く用いられている．

[†] 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

SNTP サーバは，時刻源の選択機能をもたないが，信頼できる唯一の時刻源を利用する Stratum 1（最上位階層）サーバとして利用する場合，フルスペックの NTP サーバと比べ，基本的な機能に差はなく，メッセージフォーマットも同一であるため，クライアントからは，どちらであるかの区別はつかない．もちろん，クライアント側が ntpd であれば，サーバ選択機能は問題なく動作する．

公開 Stratum 1 サーバへのアクセスのほとんどは，ユニキャストのクライアントリクエストであると考えられ，この機能のみ実装したサーバでも，ほとんどのニーズに対応することができる．このように，SNTP を更に単純化することにより，実装の単純化，高速化が期待できる．また，FPGA などによるハードウェア化も十分に可能となり，これにより，飛躍的な高速化，高精度化が期待できる．

筆者らのグループでは，高性能 FPGA と高速ネットワークインタフェースを備えた汎用性の高いハードウェアを開発し，ファームウェアの入換えにより，通過型タイムスタンパ [7]，トラヒック発生装置，パケットキャプチャ装置等として各種実験 [8], [9] に利用してきた．今回，SNTP サーバとして動作するファームウェアを開発し，その動作確認を行った．このサーバは，クライアントからの IPv4 または IPv6 のユニキャスト NTP メッセージに対し，サーバ応答を返す機能のみを有し，これ以外のパケットには応答しない．

この SNTP サーバ機能以外に，「NTP クライアント補助機能」も実装している．これは，NTP クライアントの直前に挿入された本ハードウェアが，通過する NTP パケットに正確なタイムスタンプを書き込む動作を行い，これにより，高精度な NTP クライアントを構成する，または，NTP サーバの安定度を正確に測定することが可能となる．

本論文では，上記の SNTP サーバ，及び NTP クライアント補助機能の構成について述べ，性能評価実験の結果を示す．更に，本開発品の利用方法の例と，それに対する検討結果を示す．本論文の構成は，以下のとおりである．まず **2.** で NTP による時刻同期方法について述べ，**3.** で本サーバの構成，**4.** で，実際に動作させた結果を示し，**5.** で，実験結果等を元に，本サーバの応用等について考察する．

2. NTP による時刻同期

インターネットプロトコルによる時刻同期には，NTP または SNTP が広く用いられている．SNTP は，NTP に規定されている，複数のサーバからの時刻情報を処理し，ローカルな仮想時計を適切に保つための複雑な機構などを省略したもので，パソコンやネットワーク機器の簡易な時計合せに用いられている．SNTP サーバを，時刻同期トリーの中間階層に用いるのは適切ではないが，正確さが確保された時計とともに用いられる Stratum 1 サーバに用いる場合，他の NTP サーバを参照する必要がないため，SNTP サーバであっても支障はない．SNTP サーバでは，基準時計が故障したときに，自動的に Stratum 2 として動作するような構成をとることはできないが，正確な基準時計からの時刻供給では，Stratum 2 として動作を続けるより，停止する方が適切である場合が多いと考えられる．同時に停止する確率が低い状態で，複数の Stratum 1 サーバを運用し，クライアントがこれらの選択を行うことにより，可用性を確保することができる．

NTP 及び SNTP で用いられるメッセージフォーマットは同一であり，サーバ，クライアントとも，相手がどちらであるかの区別はつかない．以下，本論文では，メッセージフォーマットや，両者を区別する必要がない場合は NTP と表記する．

クライアントがサーバから時刻を取得する場合，図 1 に示すように，まず，クライアント側の時計を基準にしたタイムスタンプ T0 を含むメッセージがサーバに送られる．サーバにこのメッセージが届くと，サーバ側の時計を基準としたタイムスタンプ T1 を付加し，内部処理の後，タイムスタンプ T2 を付加したメッセージを返す．クライアント側では，返送メッセージを受信した時刻 T3 と，メッセージに含まれる三つのタイ

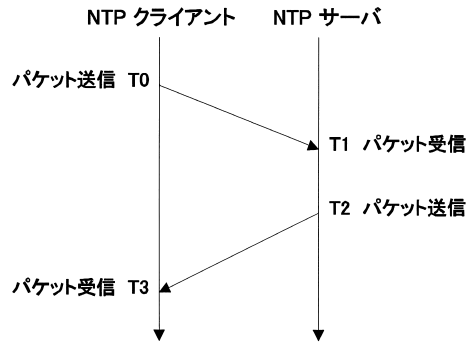

図 1 NTP による時刻比較
Fig. 1 Time comparison using NTP.

ムスタンプから，自分の時計とサーバ側の時計との差（Offset）を計算する．

$$\mathrm{Offset} = ((T1 - T0) - (T3 - T2))/2$$

この計算式では，往路と復路の伝送遅延時間が同じことを仮定しているが，実際には，通信回線の非対称性，ふくそう，サーバ及びクライアントのプロセススケジューリングなどの要因で，それぞれが独立して揺らぐため，これらが誤差となる．フルスペックの NTP クライアントでは，複数のサーバ，またはピアを用い，それぞれから取得した時刻情報を時系列的に処理することにより，これらの誤差の影響を小さくしている．

NTP の仕様には，このようなサーバ，クライアントモード以外に，broadcast モード，コントロールメッセージ，認証オプション等が含まれているが，公開 Stratum 1 サーバへのアクセスのほとんどは，ユニキャストによる認証なしのクライアントリクエストであり，この機能のみ実装したサーバでも，現状のほとんどのニーズに対応することができる．このように SNTP を更に単純化することにより，FPGA などによるハードウェア化も十分に可能となる．ソフトウェアでも，機能を絞った実装とすれば高速化が期待できるが，一般に時刻取得関数はコストが高く，タスクスケジューリングの影響で時刻精度を上げにくいことなどから，十分な時刻精度を高スループットに処理するためにはハードウェアによる実装が有利である．専用ハードウェアによる方法は，汎用サーバ上での ntpd などのソフトウェア利用に比べ，コスト面では不利になる可能性があるが，過負荷対策・クラッキング対策が不要なことから，これらの対策機器コスト，運用コストを含めると，むしろコスト面でも有利になる場合も多いと考えられる．

更に，NTP / SNTP version 4 [2], [3] では，サーバの状態通知，アクセスコントロール等に用いられる KoD（Kiss-o'-Death）パケットの仕様，サーバ過負荷を避けるためにクライアントが守るべき事項，などが追加されている．後者としては，ポーリング間隔の制限，サーバアドレス解決に DNS を用いること，KoD によるアクセス拒否に対応する機能を備えるべきであること，などが列挙されている．

3. ハードウェア SNTP サーバの構成

この章では，本サーバに実装する SNTP サーバ機能，及びクライアント補助機能について述べ，そのハードウェア構成の概要を説明する．

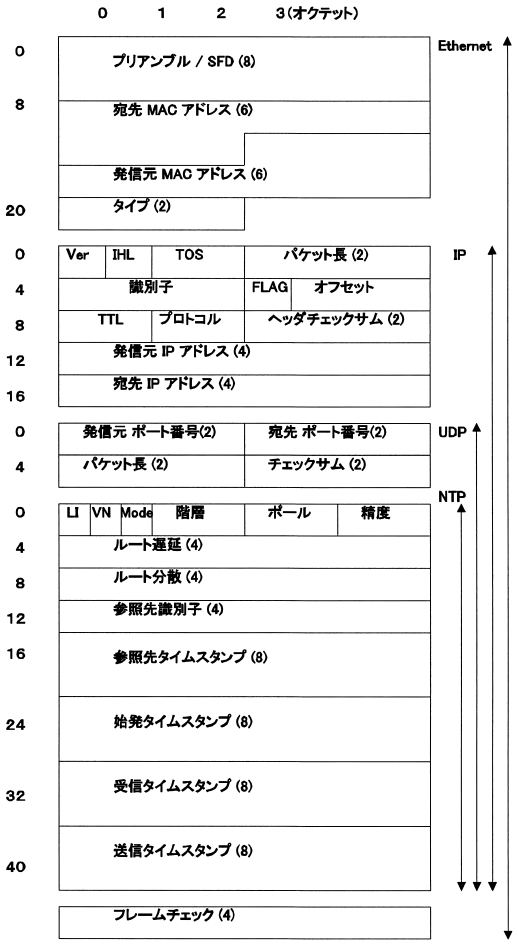

図 2　NTP メッセージフォーマット
Fig. 2　NTP message format.

3.1 SNTP サーバ

SNTP サーバの動作は単純で，基本的には，クライアントからの NTP メッセージに入っている送信タイムスタンプを始発タイムスタンプの位置に移し，サーバ時計を基準とした受信タイムスタンプ，送信タイムスタンプを付加し，送信元に送り返すだけである．このとき，サーバの状態を示すフラグ類，参照先の情報などを付加するが，Stratum 1 サーバの場合，そのほとんどは固定値でよい．

一般に，このようなサーバを構成する場合，汎用の IP スタック上に構築するが，上記のような単純な動作であれば，受け取ったイーサネットフレーム（図 2）に加工を施し，それを返送するような構成でも容易に実現できる．FPGA でも，汎用の IP スタックが利用

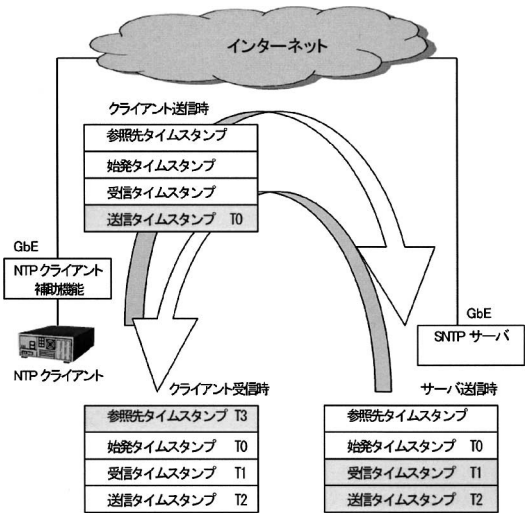

図 3　NTP クライアント補助機能
Fig. 3　NTP client assistance function.

できるものもあるが，処理速度，処理時間の揺らぎの大きさなどの点から，直接イーサネットフレームを加工する方法を選択した．具体的には，MAC アドレス，IP アドレス，ポート番号について，それぞれあて先，発信元を入れ換え，必要なタイムスタンプを付加し，チェックサムを再計算する，というのが大まかな処理の流れとなる．図 2 は，IPv4 の場合を示しているが，処理内容は IPv6 の場合でも同様である．

今回の実装では，対応するのはユニキャストのみであり，メッセージダイジェストなどのオプション規格には対応していない．ARP 機能も実装していないが，これは，NTP メッセージに ARP 解決待ちが発生すると大きな計時誤差の原因となることから，もともと，スタティックに解決すべきだとの判断に基づいている．

3.2　クライアント補助機能

本サーバには，SNTP サーバ機能以外に，「NTP クライアント補助機能」も実装した．これは，NTP クライアントの直前に通過型のネットワークデバイスとして挿入し，通過する NTP パケットに正確なタイムスタンプを書き込む動作を行う．具体的な動作は，図 3 のとおりで，クライアントが送出した NTP パケットの送信タイムスタンプを上書きし，復路のパケットの通過時には，その時点のタイムスタンプを参照先タイムスタンプの位置に書き込む．NTP メッセージフォーマットには，クライアント受信時のタイムスタンプを書き込むフィールドがないため，重要度の低い，このフィールドを流用している．この機能を利用することにより，高精度な NTP クライアントを構成することが可能となる．ただし，NTP 規定外の書換えを行うので，通常の NTP クライアントでは動作しない．この機能は，NTP / SNTP サーバの安定度等を精密に測定する用途にも利用することができる．

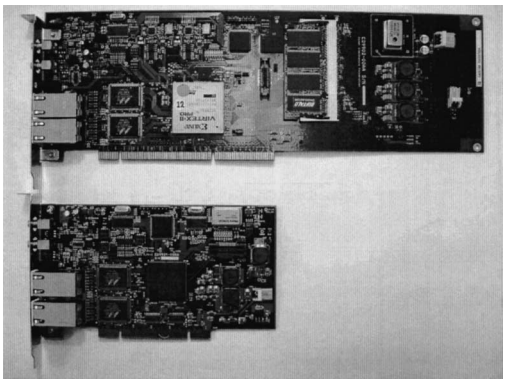

図 4　ハードウェア SNTP サーバ
Fig. 4　Hardware SNTP servers.

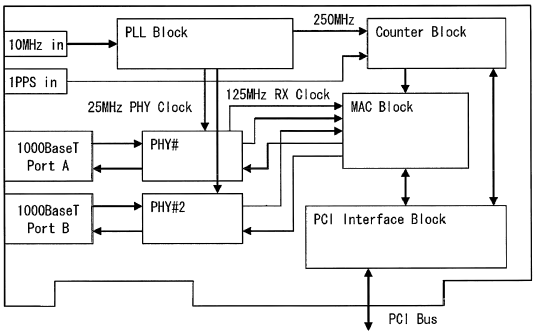

図 5　SNTP サーバの内部構成
Fig. 5　Block diagram of hardware SNTP server.

3.3　ハードウェア構成

筆者らのグループでは，高性能 FPGA と高速ネットワークインタフェースを備えた汎用性の高いハードウェアを開発し，通過型タイムスタンパ等として各種実験に利用してきた．今回の SNTP サーバも，このハードウェアを用い，専用のファームウェアを開発することにより実現した．更に，SNTP 専用機として不要な機能を削除し，小型化したものも試作した．これらの外観を図 4 に示す．

大まかな内部構造は，図 5 のようになっており，このうち，Counter Block，MAC Block，PCI Interface Block のほとんどと，PLL Block の一部が 1 個の

FPGA で実現されている．1000BaseT インタフェースを 2 個備えており，SNTP サーバとして稼動する場合には Port A のみを用い，クライアント補助機能の場合には 2 ポートを用い，通過型として動作する．

10 MHz 及び 1 PPS（秒パルス）信号源としては，GPS コモンビュー法 [10] によって正確に校正されているセシウム時計など，高精度，高安定なものを使用することを前提としている．秒カウンタの初期値は，PCI バス経由で設定され，1 PPS 信号の立上りでカウントアップする．秒未満部分のカウンタには 10 MHz 信号を逓倍した 250 MHz が入力され，1 PPS 信号の立上りからの時間を 4 ns 刻みでカウントする．NTP のタイムスタンプ形式は，2 進固定小数点であるので，このカウンタ値に定数

$4 \times 10^{-9}/2^{-32}$

を乗じた数値をタイムスタンプの秒小数部（32 ビット）としている．

このタイムスタンプ演算，あて先と発信元の入換え，その他の情報の付加などの処理は，すべて FPGA 上でパイプライン処理され，1000BaseT の最大レートでの入力に対応することができる．フレーム受信から返答フレーム送信までの遅延は固定値 832 ns である．

ボード上には，小型の OCXO（恒温槽付水晶発振器）が搭載されており，これとの比較により 1 PPS 入力の異常が検知された場合，SNTP サーバとしての機能を自動停止させることができる．また，簡易な NTP クライアントを構成する際のローカルクロックとして利用することも可能である．

4. 動作試験

この章では，一般的な NTP クライアントを用い，正常に時刻取得ができるかの動作確認試験，NTP クライアント補助機能を用いた時刻精度評価実験，及び，公開試験運用について述べる．

4.1 ソフトウェアクライアントからの利用

本 SNTP サーバに，日本標準時に同期した 1 PPS 信号及び 10 MHz 信号を与え，図 6 の実験用ネットワーク上で稼動させた．これに対し，

・WindowsXP 標準装備の SNTP クライアント，
・ntpd での server 指定（IPv4），
・ntpd での server 指定（IPv6）

からアクセスを行い，SNTP サーバ機能が正常に動作していることを確認した．

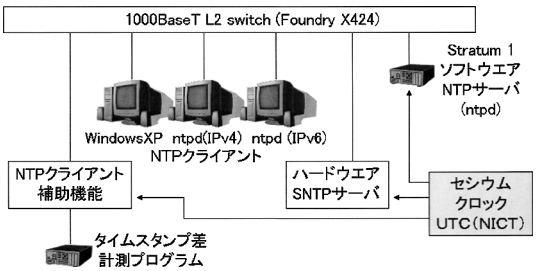

図 6　実験構成図
Fig. 6　Experiment configuration.

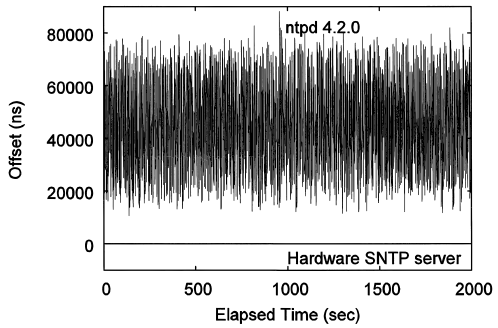

図 7　時刻比較結果
Fig. 7　Measured clock difference.

4.2 NTP クライアント補助機能での精密計測

「NTP クライアント補助機能」を用い，本 SNTP サーバと ntpd によるソフトウェア NTP サーバの比較を行った．図 6 の実験ネットワークに設置した NTP クライアントから，SNTP サーバ，及び比較のためのソフトウェアサーバに向けて毎秒 1 回の NTP パケットを送出し，その時刻差測定値をプロットしたのが，図 7 である．ソフトウェアサーバは，ntpd を使用し，PPS ドライバによって，シリアルポートに入力された日本標準時信号にロックしている．更に，このサーバは，高精度時刻 PC [11], [12] と同様の改造が施されており，CPU 及び周辺チップのクロックは，日本標準時の 10 MHz にロックした状態で稼動している．これは，ソフトウェア NTP サーバとしては十分安定したものといえるが，60 マイクロ秒程度のジッタが観測されている．これに対し，ハードウェア SNTP サーバの測定値は極めて安定しており，図 7 では，1 本の直線になっている．この縦軸を拡大したのが図 8 で，ジッタは，おおむね 40 ナノ秒程度に収まっているのが分かる．更に L2 スイッチを取り除き，ケーブルのみによる直結とした場合には，測定値は 8 ナノ秒の範囲に収まった．1000BaseT の伝送クロックが 125 MHz で

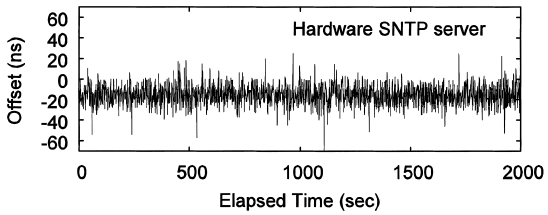

図 8　時刻比較結果（拡大図）
Fig. 8　Measured clock difference (enlarged).

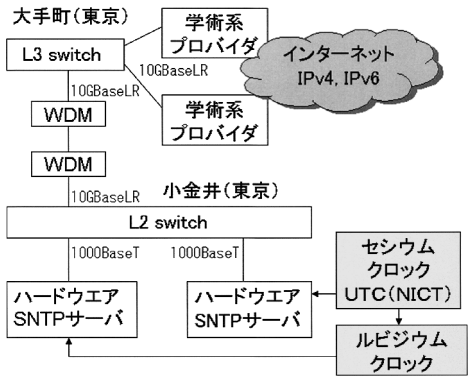

図 9　公開試験運用
Fig. 9　Experimental service.

あることを考えると，ほぼ限界の精度に達しているといえる．

4.3　公開試験運用

ハードウェア SNTP サーバの機能に問題がないか，また，長期にわたって運用上の問題が出ないかを確認するため，インターネットから，IPv4，IPv6 のいずれでも自由にアクセスできる状態で公開試験運用を行っている．サーバを収容するスイッチから学術プロバイダまで 10 Gbit/s 回線を確保し，ハードウェア SNTP サーバも最大 10 台稼動できる状態になっているが，現在，外部に公開しているサーバは 2 台である（図 9）．現在までのところ，うるう秒対応，IPv6 対応なども含め問題は報告されていない．

5.　考　　察

本 SNTP サーバは，1000BaseT のワイヤスピードで動作するため，サーバのための負荷対策が不要，外部からのパケットにより変更される部分がないので，特段のクラック対策も不要である．また，毎秒 100 万程度の NTP リクエストに対応できる能力があるため，多くのアクセスが集中するような環境でも，さほど多くない設置台数での対応が可能である．文献 [4] で示されている PC ベース NTP サーバのスループットは，毎秒 1000 リクエスト程度，現在市販されている NTP サーバ専用機でも毎秒 1 万リクエスト程度であり，本 SNTP サーバとは大きな差がある．複数台のサーバを負荷分散装置で集約する方法と比べても，負荷分散装置のインタフェースが GbE である限り，何台のサーバを集約しても，本サーバのスループットを超えることはできない．汎用 PC サーバは安価であるので，ある程度の台数を設置しても，機器コスト自体は大きくないが，設置スペース，発熱量などは台数に応じて増大し，これらが運用コストを押し上げる．本サーバの場合，サーバカードを収容する PC を含めても，19 インチラック 1 U サイズに収まり，発熱量も 100 W 以下であるため，ファシリティ面での問題は少ない．

高速で，かつ経路変動のない回線が利用できる範囲では，マイクロ秒以下の精度での時刻同期が期待できる．GPS 受信機も低廉化しているが，アンテナ工事等が難しい場合には，IP による時刻同期の方が手軽なケースも多い．データセンタ，地下街，トンネル内などに設置した機器間の同期をとるような場合でも，もともと機器制御用に IP ネットワークが用意されていることが多く，この場合，本 SNTP サーバの利用により，新たな配線を設けることなく必要な時刻同期精度を得られる可能性がある．

本 SNTP サーバは，IP アドレスをもっておらず，本サーバの MAC アドレスあてに送られるすべての IPv4 及び IPv6 のユニキャスト NTP メッセージに応答する．このため，レイヤ 2 インターネットエクスチェンジのような形態で，複数のインターネットプロバイダが SNTP サーバを共用し，それぞれ自 AS（自律システム）内の IP アドレスをもったサーバとして扱うこともできる．

6.　む　す　び

SNTP サーバとして必要最小限の機能のみに絞り，FPGA で構成することにより，高精度，高スループットなサーバを開発した．これまでの試験運用の範囲では，うるう秒対応，基準時計信号が不安定になった場合の挙動なども含め，問題は生じていない．今後，パブリックな Stratum 1 サーバの定常運用に向けて，準備を進めたいと考えている．

本サーバには KoD 関係の機能は搭載されていない．このサーバ自体が過負荷となることはないが，アクセス拒否の通知機能等は，不要なトラヒックの発生を避

けるために有効と考えられる．この機能の実装についても，今後，検討を進めたい．

また，筆者らのグループでは，10 ギガビットイーサネットインタフェースを搭載したハードウェアも開発しており，既に，通過型タイムスタンパとしてはワイヤスピードで動作させている [7]．SNTP サーバの動作はこれより複雑であるが，ワイヤスピードを目指して開発を進めたい．

謝辞　コーダ電子（株）野間氏と佐武氏には，ファームウェア実装及びハードウェア小型化に関し多大なる助言を頂いた．ここに感謝する．

文　献

[1] D.L. Mills, “Network time protocol (version 3),” RFC 1305, IETF, 1992.

[2] D.L. Mills, “Network time protocol version 4 core protocol specification,” Electrical Engineering Technical Report 06-01-02 University of Delaware, Jan. 2006.

[3] D.L. Mills, “Simple network time protocol (SNTP) version 4 for IPv4, IPv6 and OSI,” RFC 4330, IETF, Jan. 2006.

[4] J. Levine, M.A. Lombardi, and A.N. Novick, NIST Computer Time Services: Internet Time Service (ITS), Automated Computer Time Service (ACTS), and time.gov Web Sites, NIST Special Publication, pp.250–259, May 2002.

[5] D. Plonka, “Flawed routers flood university of Wisconsin Internet time server,” NANOG 29, Oct. 2003.

[6] IEEE1588: Standard for a precision Clock Synchronization Protocol for Networked Measurement and Control Systems, IEEE, 2002.

[7] 町澤朗彦，鳥山裕史，岩間　司，金子明弘，“通過型高精度 UDP タイムスタンパの開発，” 信学論（B），vol.J88-B, no.10, pp.2002–2011, Oct. 2005.

[8] 鳥山裕史，町澤朗彦，岩間　司，金子明弘，“高速インターネット環境におけるパケット遅延時間の精密測定，” 信学技報，IA2004-24, Jan. 2005.

[9] 岩間　司，金子明弘，町澤朗彦，鳥山裕史，“インターネット環境における遅延時間の統計処理，” 2005 信学総大，B-16-2, March 2005.

[10] 内藤隆光，栗原則幸，“電子時刻認証システム開発—遠隔地の標準時校正手法の確立，” 通信総合研究所第 105 回研究発表会，Nov. 2003.

[11] H. Okazawa, A. Machizawa, S. Nakagawa, Y. Kitaguchi, T. Asami, and A. Ito, “Advanced NTP synchronization device for Internet monitoring tools,” Proc. INET2001, June 2001.

[12] 北口善明，町澤朗彦，箱崎勝也，中川晋一，“高精度時刻 PC による片道遅延時間によるネットワーク帯域推定手法，” 信学論（B），vol.J87-B, no.10, pp.1696–1703, Oct. 2004.

（平成 18 年 1 月 20 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．画像の高能率符号化，視覚情報処理，計算機ネットワークの研究に従事．日本認知科学会会員．

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，光・時空標準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．