

ハードウェア SNTTP サーバの開発

鳥山 裕史[†]

町澤 朗彦[†]

岩間 司[†]

金子 明弘[†]

† (独) 情報通信研究機構 〒184-8795 東京都小金井市貫井北町 4-2-1

E-mail: tori@nict.go.jp

あらまし FPGA を用いた、高精度、高スループットな SNTTP サーバを開発したので報告する。正確な外部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、過負荷対策が不要、クラック対策が不要、高精度といった特長がある。

キーワード NTP, 時刻同期, タイムスタンプ, FPGA

Development of A Hardware SNTTP Server

Hiroshi Toriyama[†]

Akihiko Machizawa[†]

Tsukasa Iwama[†]

Akihiro Kaneko[†]

† National Institute of Information and Communications Technology,

4-2-1 Nukuikitamachi, Koganei, Tokyo 184-8795 Japan

E-mail: tori@nict.go.jp

Abstract We have developed a high-precision and high-throughput hardware SNTTP server using a FPGA(Field Programmable Gate Array). This hardware is designed only for a stratum 1 SNTTP server which simply puts timestamps based on an accurate external clock. Our SNTTP server works as fast as the wire-speed of Gigabit Ethernet, and has timestamp accuracy of 8 nano-seconds even in full-traffic. Because of this performance, the server requires no overload protections nor crack protections.

Keyword NTP, Time Synchronization, Time-Stamp, FPGA

1. はじめに

電子商取引や電子行政手続の普及に向けて、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た。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。このような場合、NTP もしくはその簡易版である SNTTP によって時刻情報を取得することが多いが、急増するアクセスにサーバが耐えられなくなるケースも発生している[1]。Internet Draft として、サーバの過負荷を避けるためのガイドライン的な事項を盛り込んだ提案[2]も出されているが、やはり、抜本的には、サーバ能力の強化も必須と考えられる。単体のサーバ能力強化だけでなく、ネットワーク上の多くの場所にサーバが配置されることも、全体としての能力強化、精度向上、信頼性向上のためには重要であり、そのためにも、高性能で、運用コストの低い NTP / SNTTP サーバの開発が望まれている。

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。他の手段で正確さが確保された時計の利用を前提とし、サーバとして必要最少限の機能に限定すれば、単純なステートレスマシンとして SNTTP サーバを構成することができる。

我々のグループでは、高性能 FPGA と高速ネットワークインター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れ替えにより、通過型タイムスタンプ[3]、トラフィック発生装置、パケットキャプチャ装置等として各種実験[4][5]に利用してきた。今回、SNTTP 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、IPv4 および IPv6 のユニキャスト NTP パケットに応答する機能のみ有し、これ以外のパケットには応答しない。

SNTTP サーバ機能以外に、「NTP クライアント補助モード」も実装している。このモードでは、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う。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 SNTTP サーバ、および NTP クライアント補助モードの構成について述べ、性能評価実験の結果を示す。さらに、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。

2. ハードウェア SNTP サーバの構成

2.1. NTP / SNTP 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(RFC1305) / SNTP(RFC2030) 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わせに用いられている。SNTP サーバを、時刻同期ツ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 (最上位階層)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 サーバであっても大きな支障はない。

SNTP サーバを Stratum 1 として用いる場合、参照時計が故障した時に、自動的に Stratum 2 として動作するような構成は取れないが、これは、時計も含め、十分な信頼性を確保する、また、クライアントは複数のサーバを参照する、といったことにより問題が緩和できる。

NTP および 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。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 3 つのタイムスタンプから、自分の時計とサーバ側の時計との差(Offset)を計算する。

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

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、輻輳、サーバおよ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ゆらぐため、これらが誤差となる。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

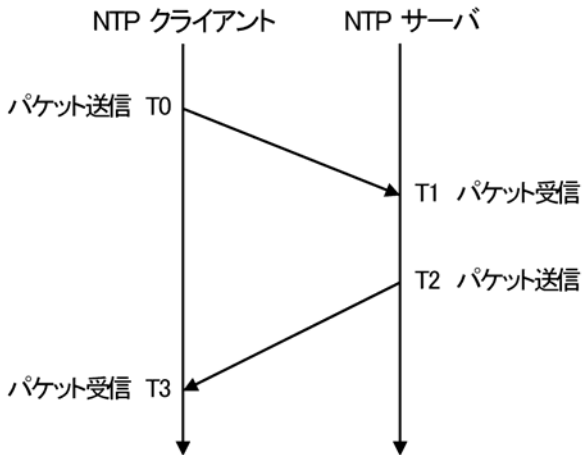


図 1 . NTP による時刻情報取得

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。

2.2. 実装する機能

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。この時、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良い。

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 Ethernet 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 スタックが利用できるものもあるが、処理速度、処理時間のゆらぎの大きさなどの点から、直接 Ethernet フレームを加工する方

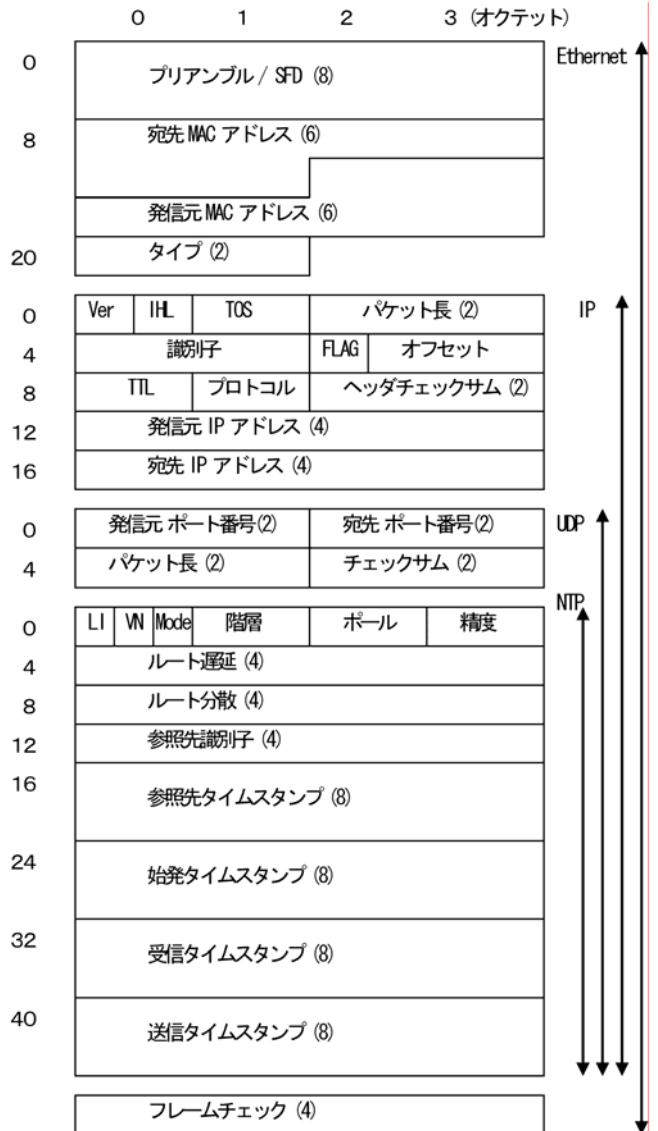


図 2 . NTP メッセージフォーマット

法を選択した。具体的には、MAC アドレス、IP アドレス、ポート番号をそれぞれ、宛先、発信元を入れ替え、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、というのがおお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、キー識別子、メッセージダイジェストなどのオプション規格にも対応していない。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。

SNTP サーバとしての機能は以上のとおりであるが、これ以外に、「NTP クライアント補助モード」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。具体的には、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。このモード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。但し、このモードは NTP 規定外の書き換えを行うので、通常の NTP クライアントでは動作しない。このモードは、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。

2.3. ハードウェア構成

我々のグループでは、高性能 FPGA と高速ネットワークインター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プ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。さらに、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



図 3 . SNTP サーバの外観

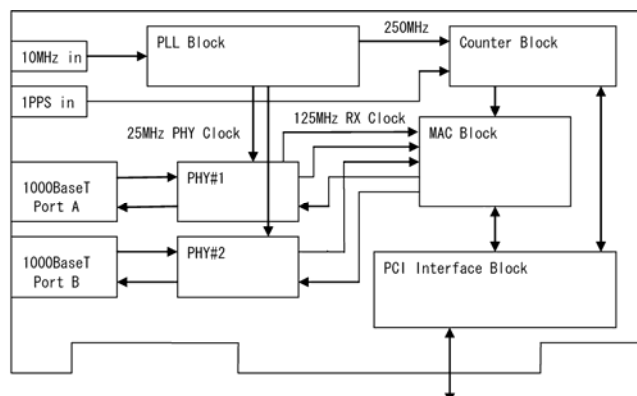


図 4 . SNTP サーバの内部構成

外観を図 3 に示す。

おおまかな内部構造は、図 4 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、PLL Block の一部が 1 個の FPGA で実現されている。1000Base-T 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、クライアント補助モードの場合には 2 ポートを用い、通過型として動作する。

10MHz および 1PPS (秒パルス) 信号源としては、GPS コモンビュー法[6]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、1PPS 信号の立ち上がりでカウンタアップする。秒未満部分のカウントには 10MHz 信号を通過させた 250MHz が入力され、1PPS 信号の立ち上がりからの時間を 4ns きざみでカウントする。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、宛先と発信元の入れ替え、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、1000Base-T 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 1PPS 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。

3. 動作試験

3.1. 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 1PPS 信号および、10MHz 信号を与え、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)

によりアクセスを行い、SNTP サーバ機能が正常に動作していることを確認した。

ntpd のログから、SNTP サーバが十分な安定度を持っていることが確認できたが、同一構内にある、ソフトウェア Stratum 1 サーバとの安定度の差は認められなかった。

3.2. NTP クライアント補助モードでの精密計測

「NTP クライアント補助モード」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 5 は、2 台のサーバからの NTP 返答メッセージから復路の遅延時間に相当するタイムスタンプ差をプロットしたものである。これらの機器は、1 台のストアアンドフォワード型のスイッチングハブを介して接続されており、縦軸の数値には、この通過遅延時間も含まれている。ハードウェア SNTP サーバの計測値はきわめて安定しており、24ns の範囲に収まっている。さらにスイッチングハブを取り除き、直結とした場合には、計測値は 8ns の範囲に収まった(図 6)。1000BaseT の伝送クロックが 125MHz で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。

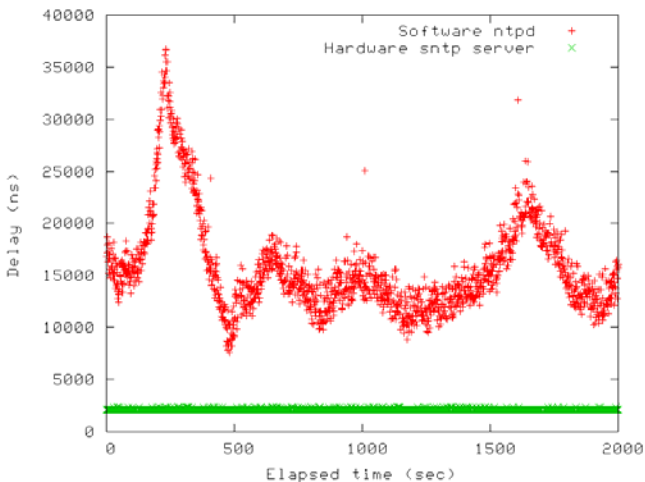


図 5 . 復路タイムスタンプ差の比較

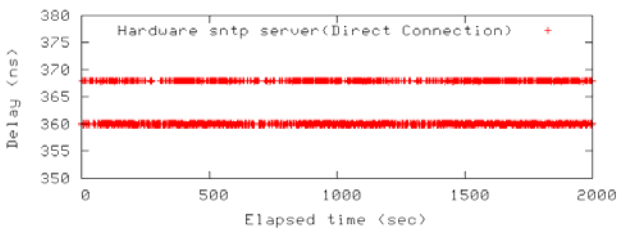


図 6 . 復路タイムスタンプ差 (直結)

4. 本サーバの特長

本 SNTP サーバは、1000BaseT のワイア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 万以上の NTP パケッ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。これらのことから、運用コストの低減が期待できる。

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あろう。

本 SNTP サーバは、IP アドレスを持っておらず、本サーバの MAC アドレス宛に送られるすべての IPv4 および IPv6 のユニキャスト NTP メッセージに応答する。このため、レイア 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 アドレスのサーバとして扱うこともできる。

5. おわりに

NTP サーバとして必要最少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では問題は出ていないが、今後、不正なパケットへの対応、クロックソースが不安定になった場合の挙動など、さまざまな状況でも十分な信頼性が確保できることを確認していく予定である。

我々のグループでは、10 ギガビットイーサネットインターフェースを搭載したハードウェアも開発しており、通過型タイムスタンプとしてワイアスピード動作させている。SNTP サーバの動作はこれより複雑であるが、ワイアスピードを目指して開発を進めたい。

また、「NTP クライアント補助モード」を利用した、高精度な NTP クライアントについても研究開発を進め、手軽な時刻および周波数の供給方法として実用化を目指したい。

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およ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] Dave Plonka, Flawed Routers Flood University of Wisconsin Internet Time Server, NANOG 29, Oct. 2003
- [2] D. Mills, D. Plonka, J. Montgomery, Simple Network Time Protocol (SNTP) Version 4 for IPv4, IPv6 and OSI, Internet Draft draft-mills-sntp-v4-00.txt, Sep. 2003
- [3] 町澤,鳥山,岩間,金子,“通過型高精度 UDP タイムスタンプの開発,” 信学技報, IA2004-25, Jan. 2005
- [4] 鳥山,町澤,岩間,金子,“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
- [5] 岩間,金子,町澤,鳥山,“インターネット環境における遅延時間の統計処理,” 2005 年信学総合大会, B-16-2, Mar. 2005
- [6] 内藤,栗原,“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