Challenges in Time Transfer Using the Network Time Protocol (NTP)

Steven E. Sommars, *Nokia*
Naperville, Illinois

BIOGRAPHY

Steven Sommars is a Consulting Member of Technical Staff at Nokia and has concentrated on wireless packet data for over 20 years. He has a PhD in physics from Stony Brook University and is a Senior Member of IEEE.

ABSTRACT

Successful time transfer between a network time protocol (NTP) server and client relies on symmetric network delays and accurate, well-behaved NTP servers. NTP network delays and stratum 1 server performance were monitored for several years from multiple clients located in the United States. Several patterns of asymmetric delay and loss were identified, some pathological. Significant time offsets and other errors were seen in a few NTP stratum 1 servers, including some hosted by national agencies. IoT implications are briefly discussed.

INTRODUCTION

Some wireless radio networks require accurate time or at a minimum good time synchronization. These may utilize embedded GPS, PTP or other technologies. Operation of less time-critical portions (e.g., ~ 1 msec) of the network may use NTP: diagnostic logs, IP latency measurements and so on. After noticing time synchronization errors in my work, I began monitoring several public NTP servers looking for anomalies in the IP transport of NTP messages or in the NTP responses. NTP typically works well with little effort from the user. This paper focuses on the exceptions, where a client receives NTP messages that could cause erroneous time transfer.

Some previous NTP server surveys [1] [2] relied on NTP administrative commands to locate servers. Exhaustive IPV4 address space searches have also been performed. This work instead focused on stratum 1 servers gleaned from public lists and discussion groups such as [3] [4] [5]. The monitored servers are widely used but may not represent an unbiased sample of all public servers.

Time synchronization requirements are application dependent and will not be covered here. This study used simple tools that are sufficient to study some behaviors at the ~msec and higher level. There are significant limitations to what any single remote monitor can observe. For example, if the apparent request and response delay between a GPS-synchronized client and a remote stratum 1 server differ by a constant 10msec out of a total round-trip time of 50 msec, one cannot distinguish between {accurate server, network asymmetry=10 msec} and {server error =10msec, symmetric network}. The accuracy of many NTP stratum 1 servers has been validated to better than 1 microsecond using two-way satellite time transfer (TWSTT) and other techniques.

Monitoring initially focused on United States-based NTP servers but broadened over time. Unexpected network behavior may occur at any time. NTP server timestamp errors also happen, even when operated by respected organizations. NTP servers are identified geographically, not mentioned by name or organization. Potential improvements in stratum 1 servers will be discussed.

Various aspects of NTP time transfer have been reported in PTTI [6] [7] and elsewhere. This study attempts to distinguish client-visible network and server-attributable errors.

Monitoring clients themselves use NTP. However, the results presented here are largely independent of client NTP operation. To the extent possible, this paper describes the information available to clients and not how those clients use NTP. This study was not passive. Where possible, problems were reported to operators and/or manufacturers. In some cases, corrective action was taken. Some commercial NTP servers were no longer supported by the manufacturer but were still in service.

Monitoring Technique

The NTP protocol is defined in [8]. Continuing David Mills' initial work, the *reference* NTP software is maintained through ntp.org [9] and the Network Time Foundation [10]. Many versions of this software have been produced and are in use. There

are also forks such as ntpsec [11] with goals such as security improvement. Independent NTP implementations such as chrony [12], ntimed [13] and openntpd [14] have various goals. Internal details about ntpd (a typical name for the NTP daemon) are difficult to generalize.

NTP monitoring clients were primarily Linux-based but also included FreeBSD and Solaris. The preferred client time synchronization used directly attached GPS + PPS. Where that was not possible, nearby NTP stratum one servers were carefully chosen. The occasional client timing errors for the second configurations are easily seen during the analysis phase.

The number of simultaneous monitoring clients varied over the past seven years but typically ranged from 5-8. Two clients were in residences and used either DSL or cable modem access. Four clients were in commercial hosting centers, all in different states. Clients were also placed at a university and at my workplace. This report describes IPV4 results only. Two clients ran IPV6 but were limited by equipment problems.

As illustrated in Figure 1, NTP clients send NTP mode 3 queries to servers that return NTP mode 4 responses. NTP symmetric and broadcast modes were not studied. The queries and responses were recorded through time-stamped tcpdump packet captures. A packet capture retains all bytes from the mode 3 and 4 NTP packets and allows off-line analysis to evolve. Using packet captures removes most internal client delays from the analysis.

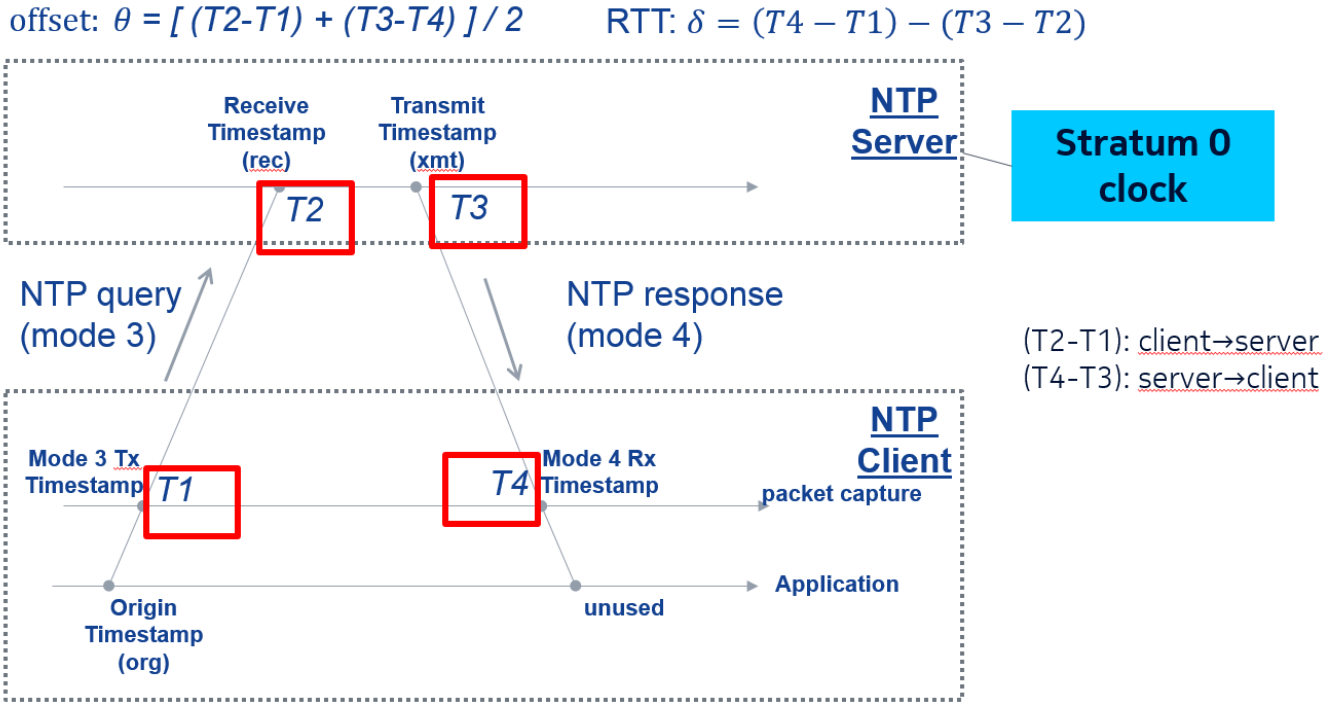

Figure 1 NTP client-server T1, T2, T3, T4. Exchange.

The difference between the application's timestamp generation(org) and T1 taken from the packet capture was typically small, well under 1 msec. The difference between the packet capture timestamp (T4) of the NTP response and the application's timestamp varied with implementation and could be many milliseconds. Because packet captures were collected directly on monitoring clients, the T1 & T4 timestamp accuracy is a concern. In the typical tcpdump implementation, the client's clock is NTP-disciplined. One msec accuracy was the target. Large client timestamp errors, several msec or larger, were sometimes seen but have a well-defined signature: all monitored stratum 1 servers show the same simultaneous shifts in T2-T1 and T4-T3.

The rawstats logging feature of the reference NTP software can be used to collect T1, T2, T3 and T4 instead of packet captures. Though rawstats does not record the entire IP packet, much of the analysis of this paper can be replicated.

Analysis

Several C programs and shell scripts were used to read the packet captures and associate the NTP mode 4 server response with the corresponding NTP mode 3 server request. Each response's T1, T2, T3, T4 timestamps plus other fields from the NTP response were summarized as a single line that can then be graphed. The Chicago-area NTP server shown in Figure 2 is part

of the OWAMP [15] project that measures one-way network delays. This example nicely illustrates a difference between stratum 1 and stratum 2+ operation. When the OWAMP NTP server runs at stratum 1, the request delay (T2-T1) and the response delay (T4-T3) appear as horizontal lines. NTP servers operating at stratum 2+ may exhibit clock drift due to varying network delays. The smoothly increasing (T2-T1) is accompanied by decreases in (T4-T3) and vice versa.

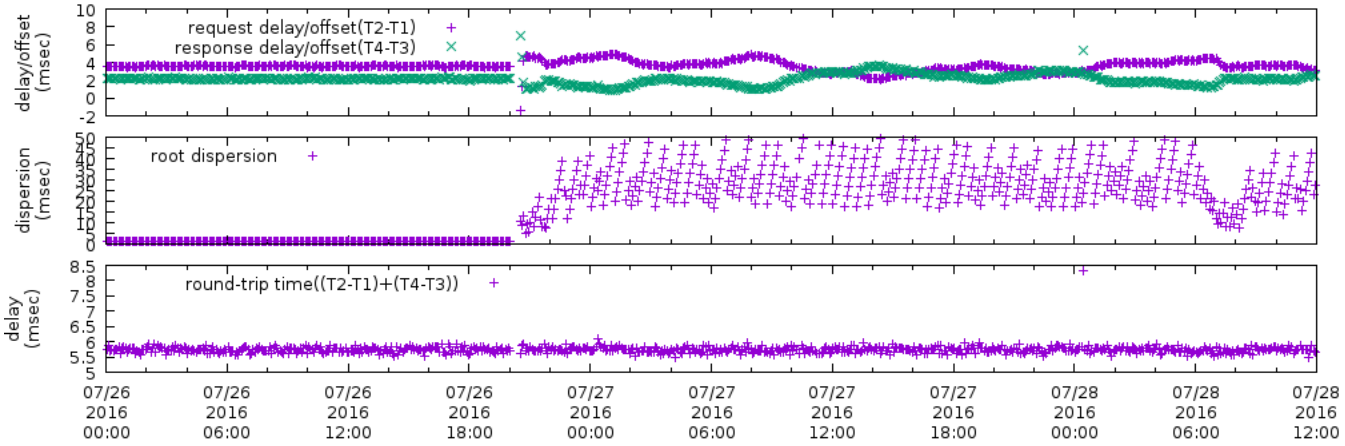

Figure 2 OWAMP NTP server near Chicago. RTT cutoff=15 msec (MYD 57595-57597). Server was initially stratum 1 and then shifted to stratum 2

A wedge scattergram [8] shows the relationship between delay (T2-T1+T4-T3) and offset [(T2-T1)+(T3-T4)]/2 as shown in Figure 3 using the same data as Figure 2.

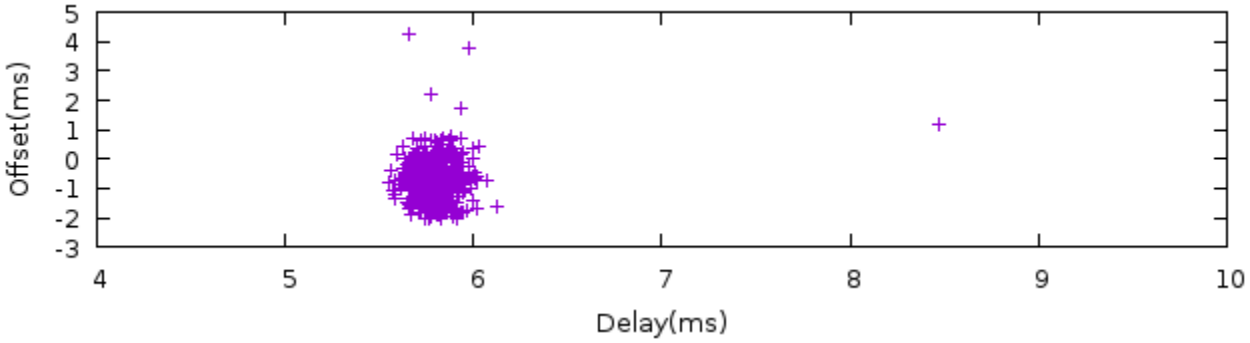

Figure 3 Wedge Scattergram, OWAMP server near Chicago, 2016. Same dataset as previous figure.

The wedge scattergram is less useful in this study than the time series used in Figure 2 because the client's time is believed to be accurate to ~1msec. The request "delay" (T2-T1) represents a combination of network delays and errors in the monitored clock. It is not possible to fully separate network effects from NTP server inaccuracies. However, the transition from stratum 1 to stratum 2+ is readily seen in the top portion of Figure 2 and is also directly reported via the response's NTP stratum. Time series charts are preferred over wedge scattergrams in this paper. For simplicity, typically only the request and response delays are shown. Adapting Mills' terminology, NTP servers returning timestamps and root dispersion inconsistent with the monitoring client are possible *falsetickers*. When the monitored server is assumed to be correct (*truechimer*), the time series is labeled "delay." When the monitored server is believed to be incorrect (falseticker), the time series is labeled "delay/offset." Unless otherwise stated, NTP responses indicating alarm (leap bits set to 11) are not included in the time series.

Discussion / Disclaimers

Many items noted in the following sections are well established parts of NTP folklore, though some specific failures may be little-known. Many portions of the Internet are unrepresented. United States-based monitoring clients are less likely to detect unusual behaviors in international access networks. NTP access from clients using wireless networks was not studied.

Some large government and research organization have their own time scales, often utilizing cesium oscillators. The coupling between the local time scale and public NTP servers may be complex. This study probes the NTP servers, which may not reflect the underlying time scale performance.

Despite the already mentioned limitations and caveats, the two messages of this paper hold: NTP messages may incur unanticipated loss, delay and other impairments when traversing the Internet. NTP servers, even well-run stratum 1, may experience failures.

NETWORK DISCUSSION

Reachability

Theoretically any NTP client can exchange NTP messages with any NTP server if both are on the Internet. The most likely results when a client sends an NTP request (mode 3) are:

- (failure) No response received
- (failure) Administrative response such as an ICMP message indicating server is unreachable
- (failure) NTP response indicating server is reachable, but not providing time (LI bits=11)
- (success) NTP response received

The "no response" result provides little diagnostic information.

Some NTP servers will only respond when the client is specifically authorized (whitelisted). Other NTP servers may implement geoblocks, blocking traffic based on perceived geographic location

NTP has been used in Distributed Denial of Service (DDoS) attacks [16]. UDP port blockage is a common remediation. The NTP message may be silently dropped or may trigger an ICMP error message such as "Administratively prohibited," "Communication administratively filtered," "Host administratively prohibited," or "Protocol unreachable." Messages to the NTP server use UDP destination port 123; the client UDP port is unspecified but is often also 123. Blockage was sometimes seen for each of the follow conditions:

- UDP destination port = 123
- UDP source port ≠ 123
- UDP contents that were much larger than typical NTP mode 3 / 4 messages [by default 48 bytes.].

The first two combine to prevent some monitoring client/NTP server combinations from ever exchanging NTP messages: NTP cannot work between these pairs.

Changes in Asymmetric Delay

Delay from a Chicago monitoring client to a stratum 1 server located in Missouri is shown in Figure 4. Similar delay increases were seen by other monitoring clients, which suggests the delay increase originated near the server. The near-constant response delay strongly suggests that the NTP server itself is operating correctly and that the T2-T1 increase is network related.

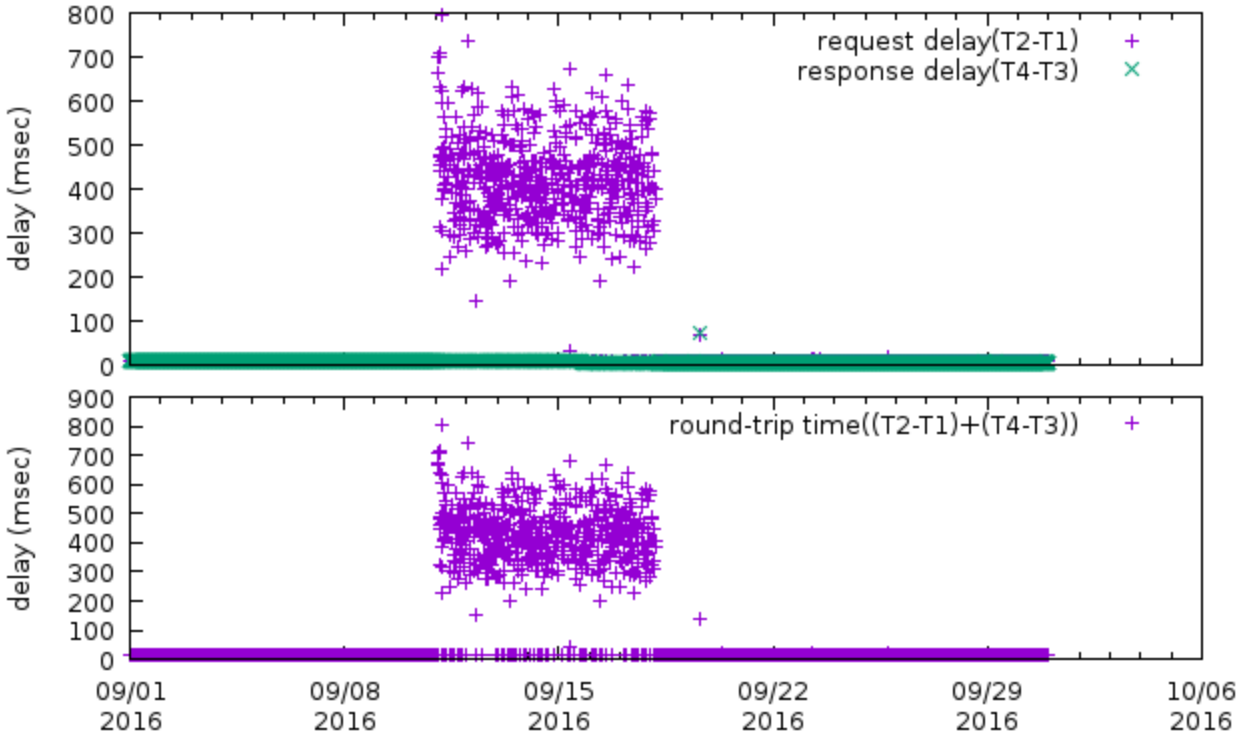

Figure 4 Stratum 1 NTP server in Missouri, (MJD 57632-57661)

Sudden network delay changes are common. While the erratic delays of Figure 4 may be congestion related, Figure 5 shows several sudden client ↔ server delay changes for a stratum 1 server located in Georgia (United States). Monitoring clients were located in Illinois (A, B) and Virginia (C).

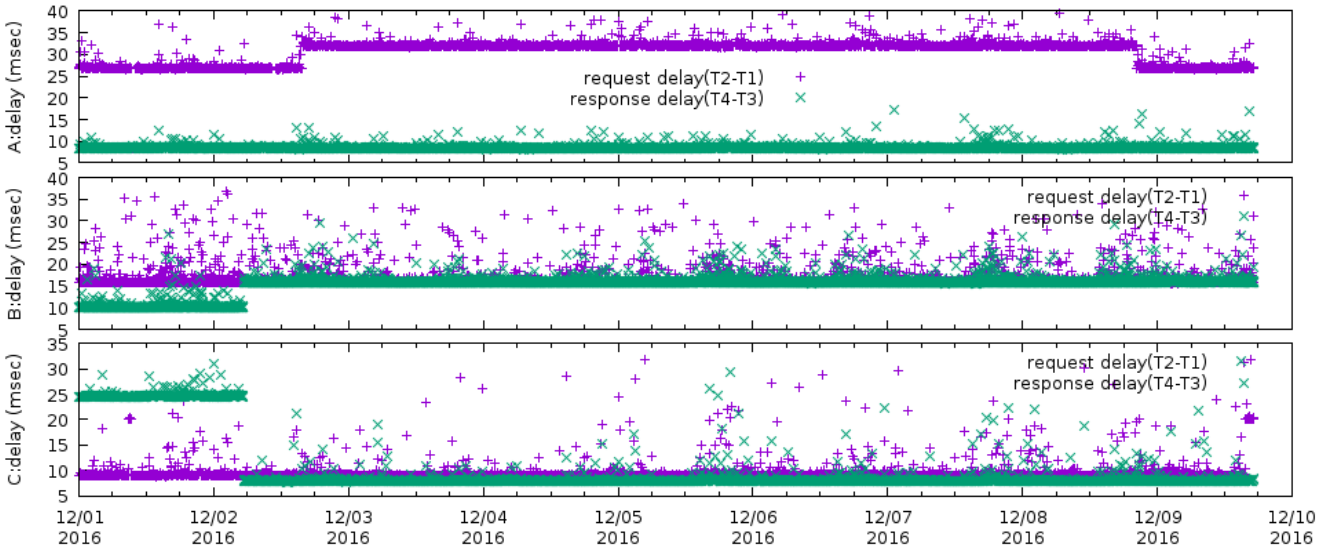

Figure 5 Delays to Georgia(US) Stratum 1 server as seen from 3 monitoring clients (MJD 57723= 57731). Client UDP port=123

There are intervals where some clients see symmetric delays (plus the inevitable "popcorn spikes") and also periods where delays are asymmetric by 15 msec in either direction. Network asymmetry is often unpredictable.

Diurnal Delay Variation

Not surprisingly, some NTP paths show diurnal variations. A stratum 1 NTP server in Hawaii provides an extreme example. Figure 6 shows lowest delays each day beginning at about 06:00 UTC.

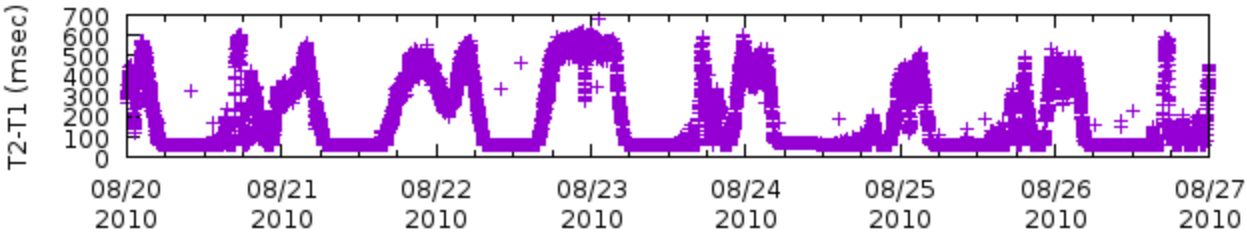

Figure 6 T2-T1 delay between Chicago client and Hawaii stratum 1. (MJD 55428-55434)

The NTP response delay, T4-T3, showed no diurnal variation and was 60-70 msec throughout this period. Other NTP servers showed lower delays on weekends and holidays.

Bufferbloat

Figure 6 exhibits symptoms of *bufferbloat* [17], high delays caused by oversized network queues. NTP delays exceeding one second were sometimes seen when uploading unrelated files from a residence using DSL.

Multiple Network Paths

The bimodal paths described in [7] were observed in this study and can often be attributed to Internet network topology. Delays for a cluster of 3 servers in Switzerland were particularly consistent, as shown in Figure 7 for one of those servers. The request delays (T2-T1) fall into clusters of ~62msec and ~75msec.

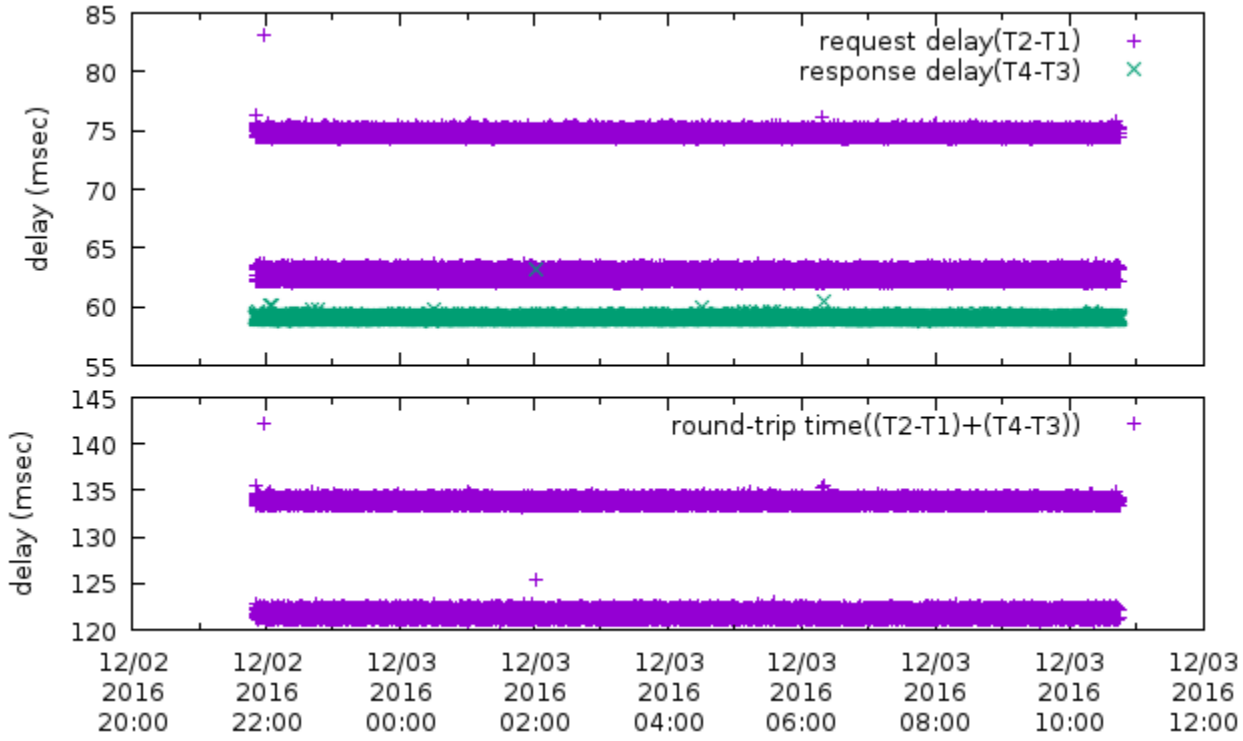

Figure 7 Delays between Chicago client and Switzerland-based server x.y.z.10 (MJD=57724-57725) client UDP ports =62660 – 62665

During this period, NTP requests were sent at higher than normal rates, with permission. The client software polled the three Switzerland-based servers (IP addresses ending in .10, .55 and .56) using client UDP ports of 62660 – 62665 in round-robin. The T2-T1 delays for all three servers are shown in Figure 8; the UDP port number (x-axis) is fuzzed for scatterplot legibility.

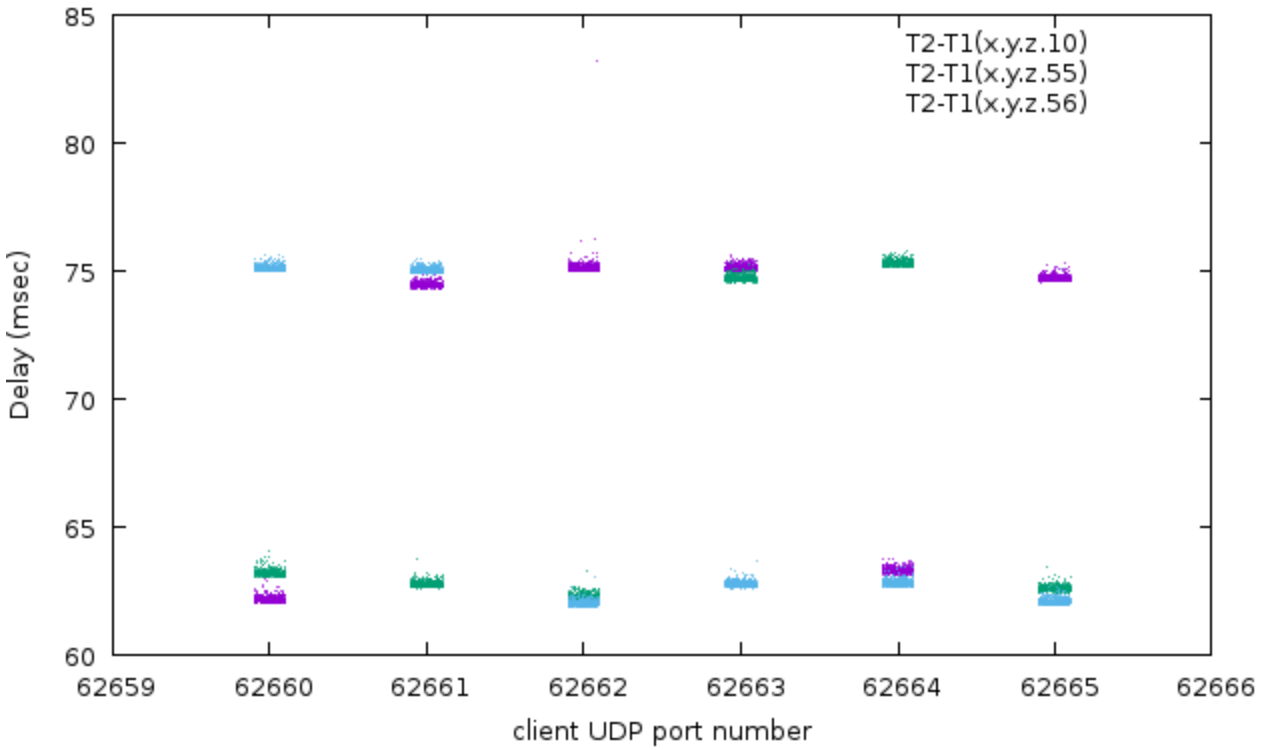

Figure 8 Chicago - Switzerland T2-T1 for three stratum 1 servers, Same time interval as previous figure.

Figure 8 shows a bimodal delay correlated with the client UDP port number / IP address and that the dependency is more complicated than bimodal. Another Chicago monitoring client located on the same campus uses a different ISP and does not show the bimodal behavior of Figure 8 when polling the same three NTP servers. Other client-server combinations showed three or more delay bands.

Such behavior is expected in today's Internet. When a router forwards an IP packet, multiple egress links may be available. Link (path) selection may involve a hash function, e.g.,

Egress_link_index = hash (IP source, IP destination, UDP/TCP source port, UDP/TCP destination port)

Over short time intervals, related traffic tends to follow the same path. Some Internet transport protocols, TCP traffic in particular, are order sensitive and may experience spurious retransmissions and lower performance when packets are reordered during transport.

Different links may have similar delay. Each member of an Ethernet link aggregation group (LAG) typically adds similar delay. At other times, different links may take diverse geographic paths and have measurable delay differences.

Many factors influence the available paths and are largely uncontrollable by an application. Applications may be able to select the NTP client UDP source port(s), but any resulting optimization may be only temporary. NTP monitoring clients in this paper used UDP port 123 in addition to ephemeral ports (typically 32768-65535 for Linux).

Network delays may also be transport protocol dependent. Examples were seen where NTP delay was high yet TCP & ICMP delays were nominal.

Middleboxes

Middleboxes [18] are present throughout the Internet and may provide valuable services such as network address translation (NAT) and network optimizations such as TCP proxies for wireless subscribers. Middleboxes can add challenges in time transfer, particularly when their presence is unsuspected.

Middleboxes can serve as load balancers and distribute incoming client NTP requests between two or more similar local NTP servers. Some operators went further and spread the NTP queries between local NTP servers with different configurations. In a nearby Chicago hosting center, queries toward a stratum 2+ NTP "server" yielded consistent round-trip times of ~4 msec; the server appeared to be located nearby. However, the one-way delays (T4-T3) showed high variation. Looking at the NTP stratum

2 response Reference ID field, which shows the reference IP address, it became clear that multiple NTP servers were in use behind the load balancer, though the exact number was not determined. Figure 9 shows the response delay for two of those NTP servers, both reporting stratum 2. One of the hosting center's stratum 2 NTP servers was synchronized to a remote stratum 1 NTP located in Iowa. Another stratum 2 was synchronized to a stratum 1 NTP located in Maryland. The Washington-area server is believed to suffer from heavy network congestion.

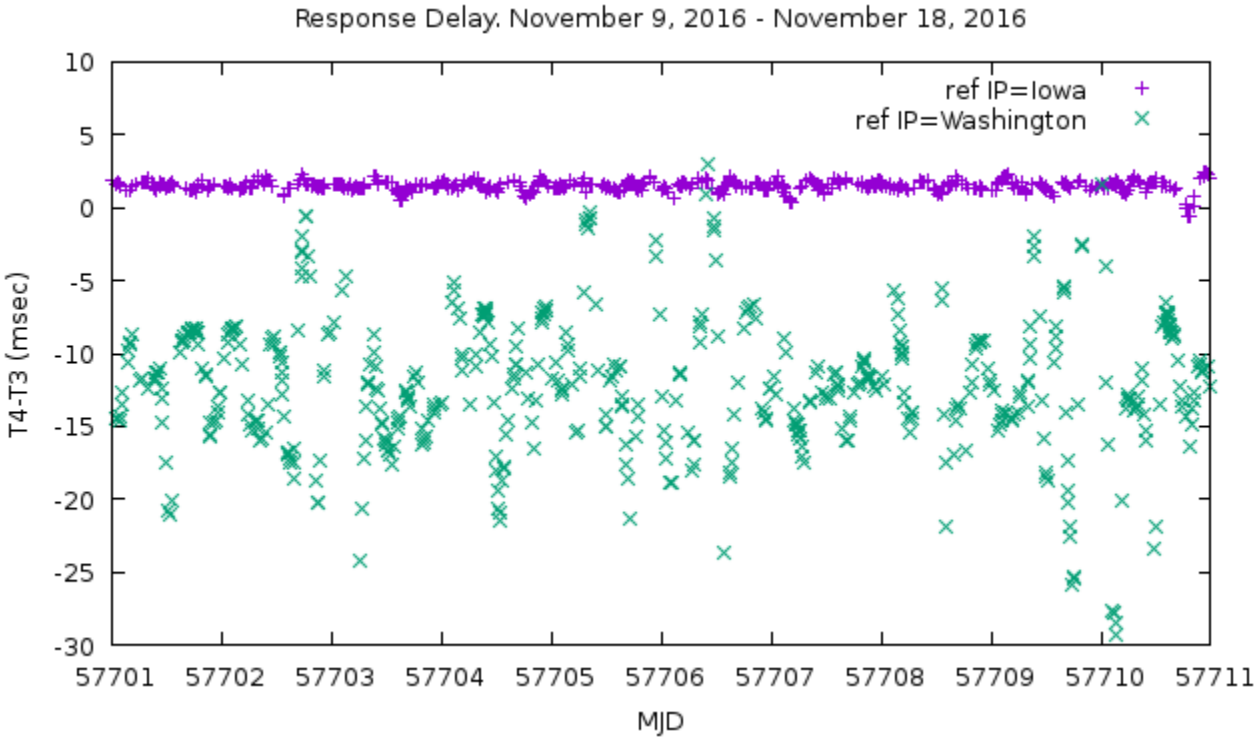

Figure 9 stratum 2+ NTP server at Chicago hosting center

Figure 10 gives a simplified diagram of the inferred hosting center configuration.

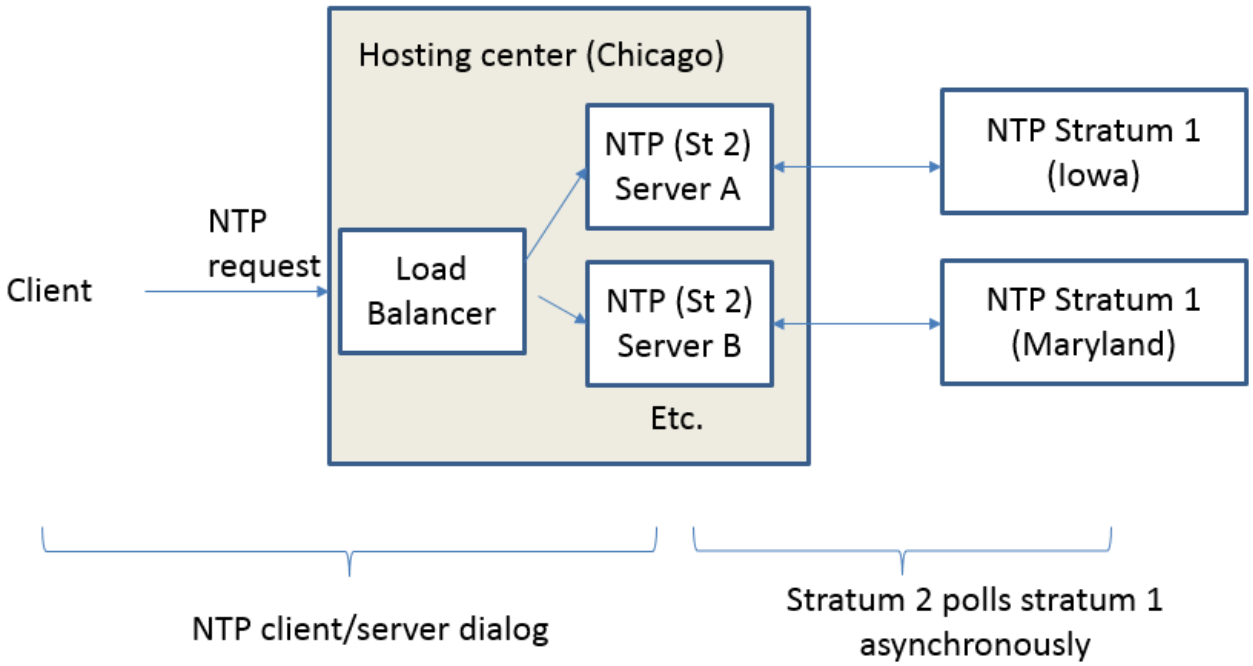

Figure 10 Inferred Hosting Center Configuration

When load balancers are used at stratum 1 NTP server sites (load balancer and NTP servers collocated), systematic differences in timestamps are sometimes seen.

Load balancers can also be bottlenecks. In 2013 reachability slowly worsened over the course of several months for one heavily used North American laboratory site. Bypassing the site's load balancer immediately improved reachability from ~45% to ~100%. Though a detailed failure analysis was not available, I will offer a potential explanation. A load balancer may forward a request to a specific server based upon load-related information and store (in RAM) each client's UDP source port and IP source address so that future requests from that client can be sent to the same server. When servicing many clients [19], the memory required to store the client information may become excessive. Load balancers servicing millions of IoT devices may require special configuration.

Occasionally NTP responses with no corresponding NTP requests were detected. Some of these were triggered in response to NTP administrative commands [20]. Many other isolated NTP responses are believed to be load balancer-related: the NTP response matched expectations except that the IP source address was incorrect and was associated with another collocated host. Here is one example.

```
2012-11-19 15:35:39.702 client -> X.Y.41.209 NTP Version 3, client
        Transmit Timestamp: Nov 19, 2012 15:35:39.701654 UTC
2012-11-19 15:35:39.902 X.Y.41.40 -> client  NTP Version 3, server
        Origin Timestamp:   Nov 19, 2012 15:35:39.701654 UTC
```

This behavior was seen by all monitoring clients for a small number of NTP server sites. One hopes that an IP addressing error would be rejected by carefully written NTP client software: matching an outgoing NTP request to an incoming response should utilize IP addresses and UDP port numbers, in addition to T1.

Queuing Delays

The average daily response fraction for a group of heavily used stratum 1 NTP servers located in Maryland as seen from a monitoring client located in Virginia 50km away is shown in Figure 11. T4-T3 was typically under 5 msec throughout this period.

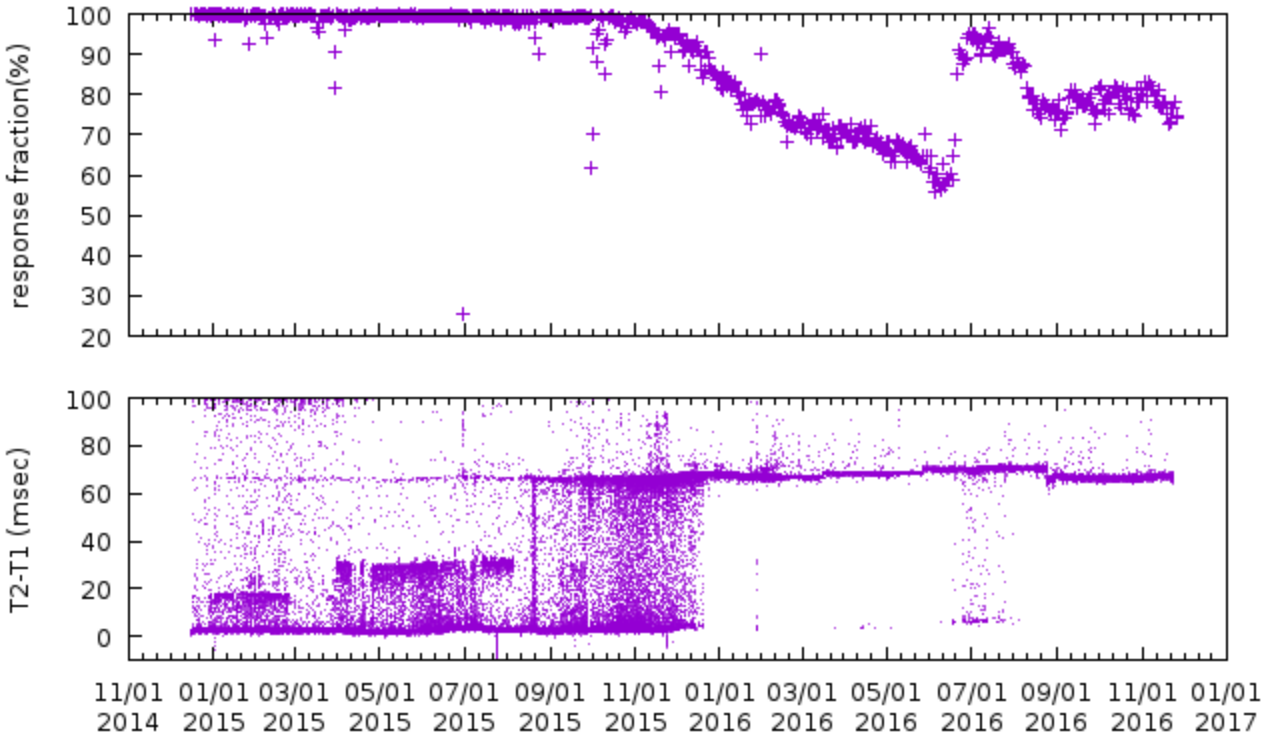

Figure 11 Maryland-based NTP server as seen from monitoring client located in Virginia. MJD 57006-57716 [T2-T1<100msec shown]

Similar behavior was seen from all monitoring clients. Though the precise queuing logic is not known, a ~60msec queue in the client → server direction is apparent.

Excessive Delays

Queuing delays can be excessive.

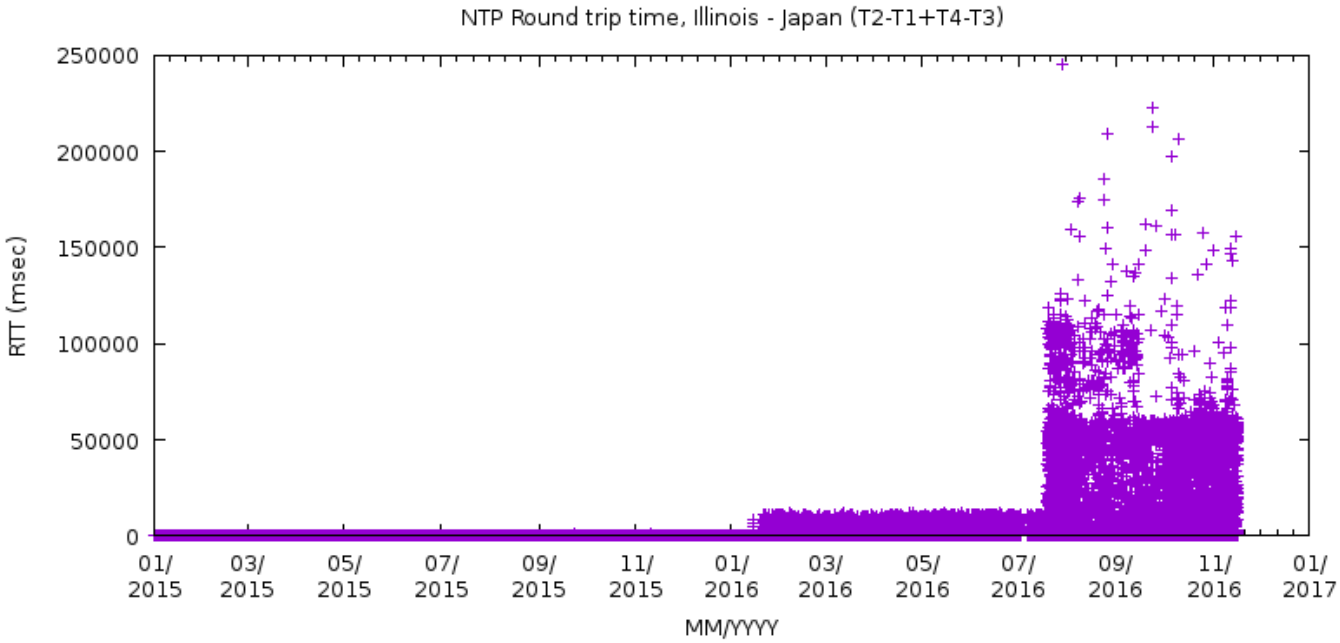

Figure 12 NTP RTT, Illinois(Client) – Japan(Server) MJD 57023 - 57709

NTP round-trip times from Illinois to one NTP server in Japan are shown in Figure 12. Starting in January 2016, delays grew from a typical maximum of ~500msec to as much as 10 seconds and further increased to 60+ seconds in July 2016. The high delays are in the reply direction (T4-T3, Japan→Illnois). The scatterplot above exaggerates the high RTT queries; 80% of the

post July 2016 queries had RTT under 1 second. Manual pings run at times of high delays between the client and host did not show the same delays, nor did traceroutes. There are hints that the high delays may be network operator specific.

While high delays on trans-Pacific routes are not a practical concern for time synchronization, similar delays were occasionally seen for a few continental United States servers.

Security
A client using NTP may expose itself to security attacks, including incoming and outgoing denial of service (DoS), attacks against NTP administration mechanisms and other security probes. The default NTP message exchange is also subject to falsified NTP responses by means of a man-in-the-middle attack. No examples of intentional timestamp alteration were detected. A situation was seen where NTP requests intended for a government operated stratum 1 server were redirected by the network operator for unknown reasons to a local stratum 2 server.

Multiple Responses
IP packets may be duplicated on the Internet; NTP is no exception. Occasionally a single mode 3 NTP request resulted in multiple mode 4 responses. Duplication on the client → server path will result in multiple responses each having different T2, T3. Duplication on the server → client direction will result in multiple responses sharing the same T2, T3. In extreme cases, a single request triggered hundreds/thousands of responses. Multiple responses are not common but should be anticipated by client software.

Attempts to Calibrate Delay Asymmetry
Attempts to characterize and calibrate network asymmetry are problematic. Equipment failures, system maintenance, and changes in network usage patterns may cause unanticipated delay changes. Recalibrated network delays may not be available in a timely manner. Perhaps network asymmetry calibrations should be limited to special cases where private networks are in use or where specific service level agreements (SLAs) exist.

NTP SERVER DISCUSSION
Many stratum 1 NTP servers were observed to send occasional responses that were erroneous in some aspect. Often the errors were temporary, lasting for seconds or minutes. A few stratum 1 servers provided chronically inaccurate NTP timestamps.

A stratum 1 server should return accurate T2, T3 values when possible. If not, it should set the alarm bits. The server should also return correct/meaningful values for other fields, especially for the root dispersion. This designation is somewhat subjective: potential errors at the monitoring client must be considered. As the client-server network delay increases, the ability to detect small NTP server timestamp errors decreases. Causes of NTP stratum 1 errors included:

- Leap second related
- Operation with undisciplined clock or free-running PPS
- Administrative error or inaction
- Reset related
- Firmware errors
- GPS rollover
- Software errors
- GPS antenna, power supply, cesium and other failures.

Due to network access restrictions, organizations may be unable to directly monitor their public NTP servers and thus delay problem detection.

NTP Server Error Heuristics
Heuristics suggesting errors in stratum 1 servers can be used, i.e., causality violations and complementary changes in request and response delays. Accurate client timestamps are required.

If (T2-T1) or (T4-T3) is smaller than Delay_minimum, where

Delay_mininum = minimum_propagation_delay – (ClientRootDispersion + ServerRootDispersion)

causality is violated and the server timestamp may be incorrect. If the NTP server location is known, the minimum_propagation_delay can be estimated conservatively using the great circle distance and the speed of light.

The right side of Figure 2 shows characteristic complementary changes in the request and response delays. Smooth changes are typical of server clock drift and highly atypical of network effects. Seeing 1) roughly constant round-trip-times, 2) smooth

complementary changes in request (T2-T1) and response (T4-T3) delays and 3) no drift in the client's clock is a good heuristic indication of server clock drift. If the drift magnitude exceeds the server's root dispersion, that server may be delivering erroneous timestamps.

Detecting the same problematic NTP server using multiple monitoring clients reduces the risk of false positives.

Well-behaved Server Initialization

Figure 13 shows the request and response delay/offset for an NTP server located in Africa reporting reference ID = GPS. The NTP server reports loss of synchronization (leap indication=3) at 07:24. The computed request and response delays are consistent with the increased root dispersion. Three hours elapsed before the server returned to its pre-initialization condition.

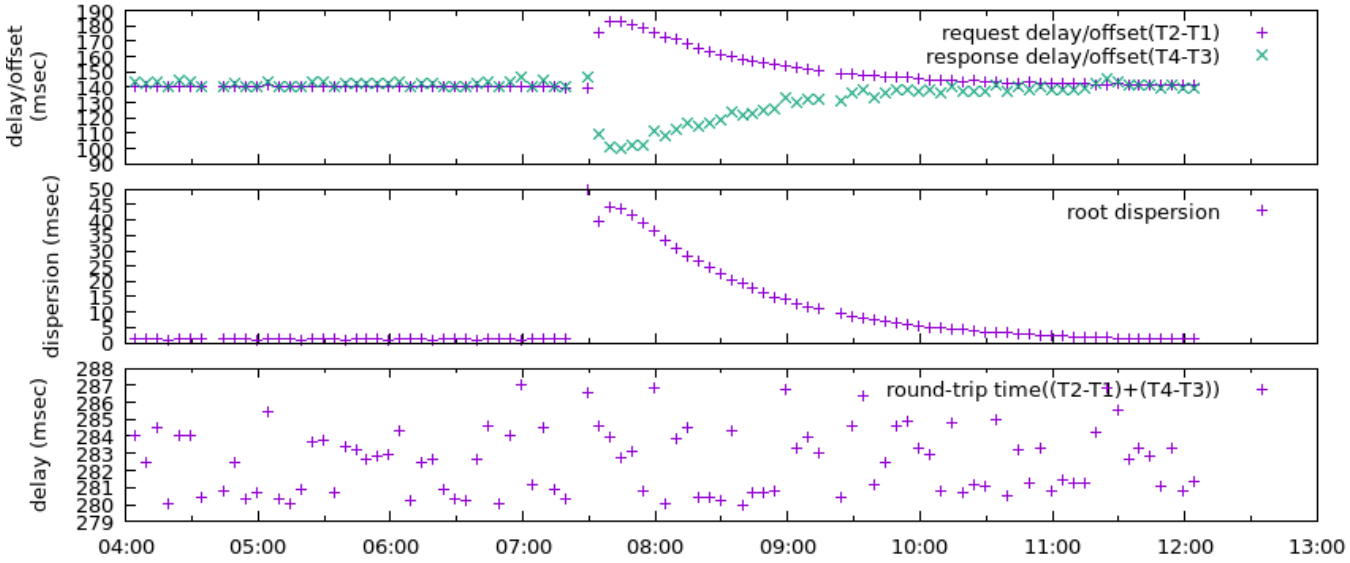

Figure 13 2016/12/23 (MJD 57745) Stratum 1 Server (Africa) initialization

The term "initialization" is used loosely here. Without additional information, it is not possible to distinguish NTP server resets, reboots or recovery from loss of GPS/PPS.

Server Initialization Errors

Occasionally one finds NTP stratum 1 servers with timestamps offset by an integral number of seconds. Some of these errors are correlated with leap-second processing. Others occur during initialization after NTP server reset. On November 10, 2016 (MJD 57702) an NTP stratum 1 server in Florida returned these values:

```
LI Stratum   Root  Ref                    T2-T1
             Disp  ID                    (msec)
3     0   0.000000 INIT 16:48:06.570597 Not meaningful
0     1   0.000000 GPS. 16:50:15.171854 2030.95
0     1   0.000000 GPS. 16:52:32.964334 8.08692
0     1   0.000000 GPS. 16:54:59.089189 32.0828
```

The server timestamp was briefly off by 2 seconds. Unfortunately, it also consistently sets the NTP response root dispersion to 0. Based on timestamp analysis, the root dispersion should be at least 50 msec for this server.

Backup Synchronization (Flywheel)

Figure 14 shows an NTP server running at first with reference ID "GPS", which changed to "FLY" on November 4, 2016; both periods reported NTP stratum 1 operation, no alarm. This pattern suggests that the GPS signal was lost and that the NTP server used a backup oscillator during flywheel operation. After 13 days on November 17, 2016, the root dispersion reached 17 seconds, a growth rate of 15000 PPM, and then the server set the LI bits to 11 (alarm). The actual time offset grew at a rate of ~15ppm. The NTP server is overly pessimistic but is not a falseticker.

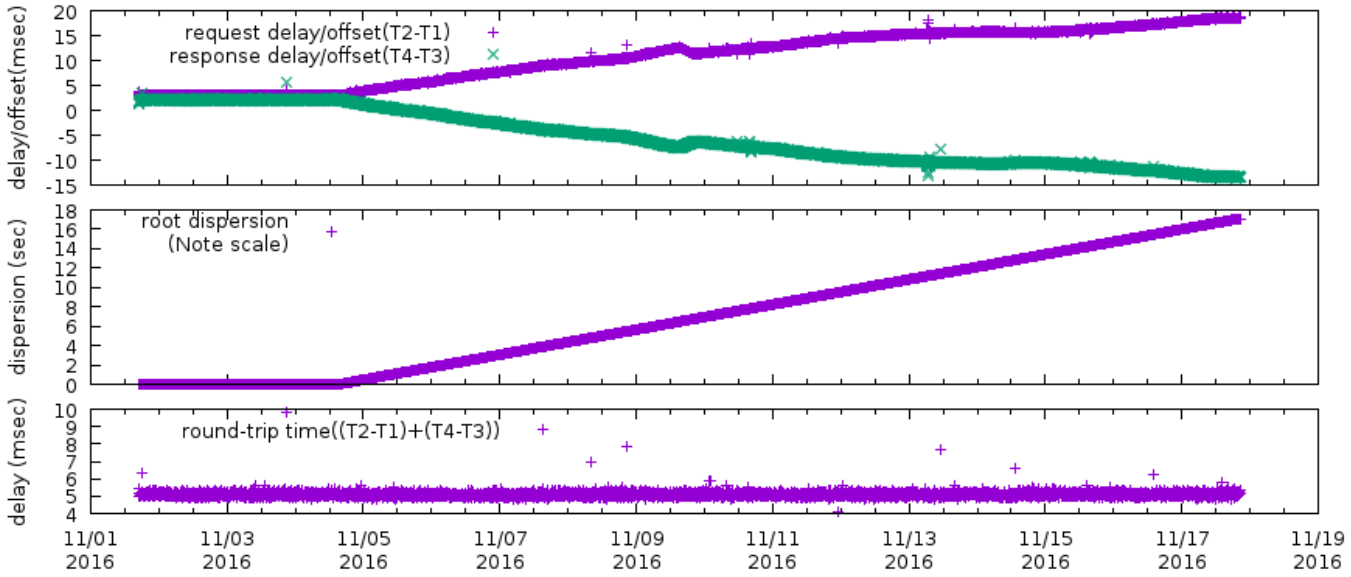

Figure 14 Illinois NTP stratum 1operating on flywheel (MJD 57693-57709). RTT cutoff=10msec

Root dispersion grew too slowly, or not at all, for some servers operating on flywheel. From January 6-29, 2014, the NTP server shown in Figure 15 operated on flywheel; the reference ID was GPS at other times.

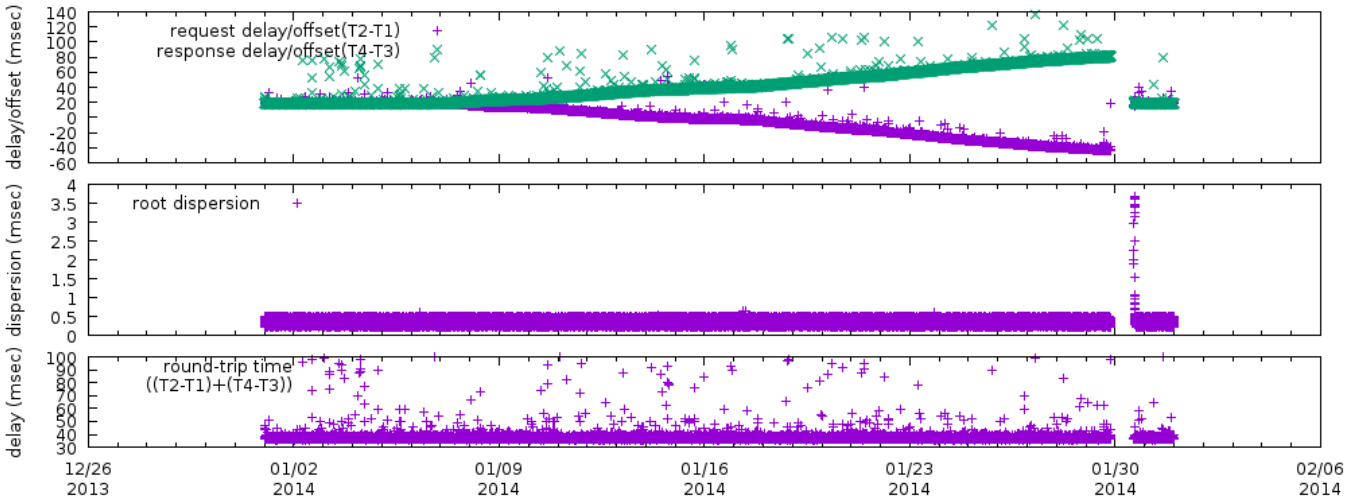

Figure 15 Northeastern United States NTP server operating on flywheel (MJD 56658-56688) RTT cutoff=100msec

It is unclear how a careful client should cope with servers delivering "backup" time. Taking different actions based on the reference ID field seems risky. Some ntpd versions limit the maximum age (e.g., NTP_MAXAGE) of the NTP response's reference timestamp. If no recent updates have been made the NTP response is not used. Unfortunately, NTP servers do not use this field consistently. The server of Figure 15 continued to update the reference timestamp while operating in flywheel mode.

NTP Stratum 1 Servers with Single Time Source

Stratum 1 servers with multiple independent clocks/NTP peers may be able to detect time inconsistencies and set the alarm flag or increase the root dispersion. Servers with a single time source have limited ability to detect malfunctioning clocks.

Less Common Failures

Figure 16 shows the delays seen for an NTP stratum 1 (ref ID=TRUE) server located in California.

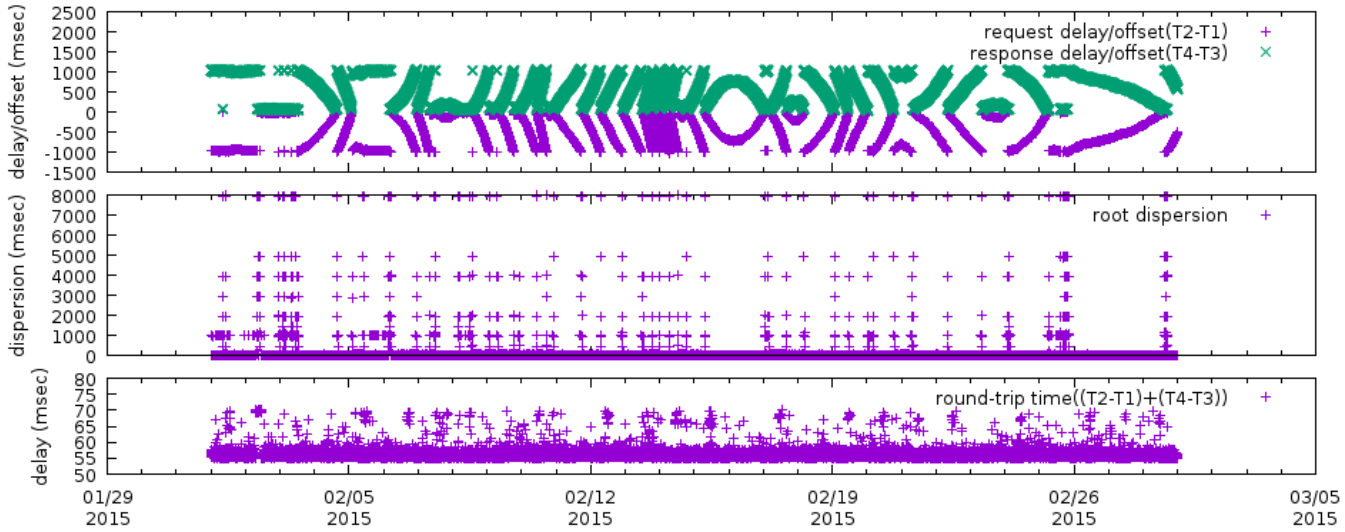

Figure 16 California stratum 1 server. (MJD 57054 - 57071). RTT cutoff=70msec

There are periods when the time offset is large, nearly one second, but the root dispersion is only 5-30 msec. The cause of underlying failure is not known; perhaps the PPS is free-running.

Leap Seconds

Much has been written concerning the motivation for and challenges of leap seconds [21]. David Malone [22] [23] has monitored NTP server leap second behavior for many years. This study observed similar problems, incorrect setting of leap bits and "off by one second" errors, for a number of NTP servers. Figure 17 shows an example of a "false leap second" seen at one European standards agency. Although none was scheduled, the NTP leap indicator was set starting at 00:00 UTC on March 31, 2015. When the date became April 1, 2015, the stratum 1 server executed a leap second procedure: the server's timestamp was erroneously decremented by one second. The timestamp error continued for about 20 minutes. During the first few minutes of this interval, the root dispersion was under 1 msec, but then it increased to over one second.

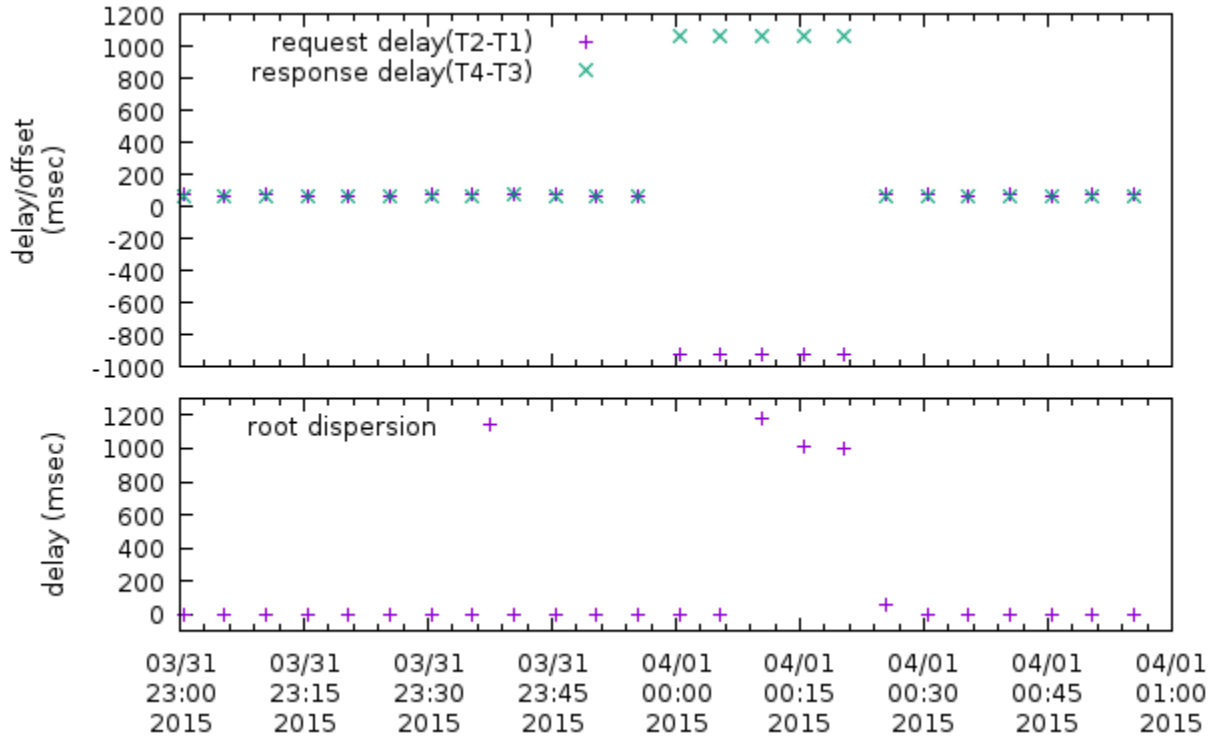

Figure 17 Chicago client monitoring European stratum 1 (MJD 57112-57113) during false leap second

Leap second smearing is supported by Google [24], the reference NTP software and others. On June 30, 2015, at least one NTP pool server showed the behavior of Figure 18, which is characteristic of a "Google smear." The deviations from a purely linear smear have been reported by Martin Burnicki [25] .

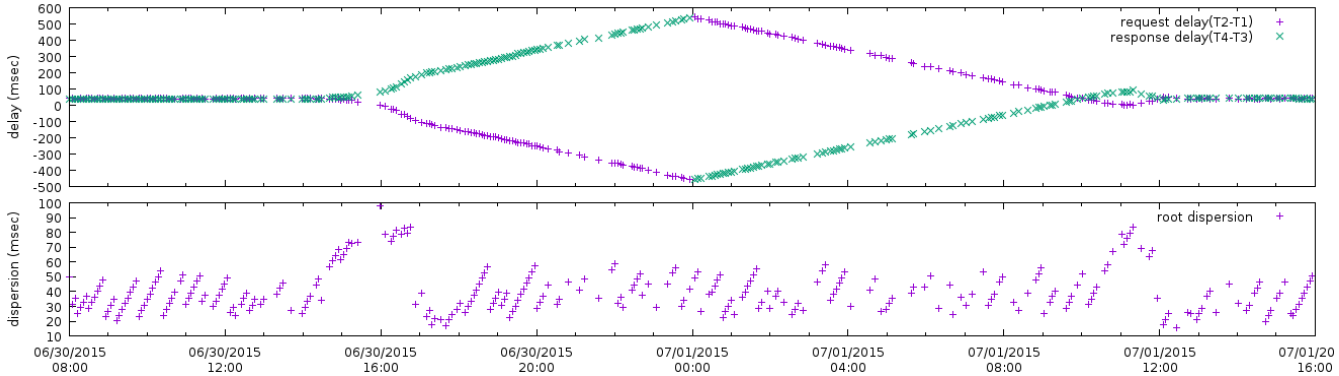

Figure 18 NTP pool server located in Europe (MJD 57203-57204). Server executed "Google smear"

The December 31, 2016 leap second showed the same smearing signature for at least 7 NTP pool servers. Over 10% of the NTP pool servers exhibited probable leap-second-related errors of at least 1 second.

On January 31, 2017 (during PTTI 2017), over 110 public NTP servers incorrectly advertised leap=1. Thirty of these executed a leap second procedure similar to that shown in Figure 17. Past behavior suggests that the false leap second problems may persist for months.

Stale T2

In some versions of the reference NTP software, the T2 timestamp is generated when ntpd begins processing one or more NTP requests queued in the operating system. Additional NTP queries could arrive before the queue was emptied, causing a stale value of T2 to be used by ntpd. The client's computed value of T2-T1 is therefore too low and could even be negative, as seen in Figure 19 for an NTP server with RefID=CDMA from 2011. The response delay, T4-T3, showed no anomalies.

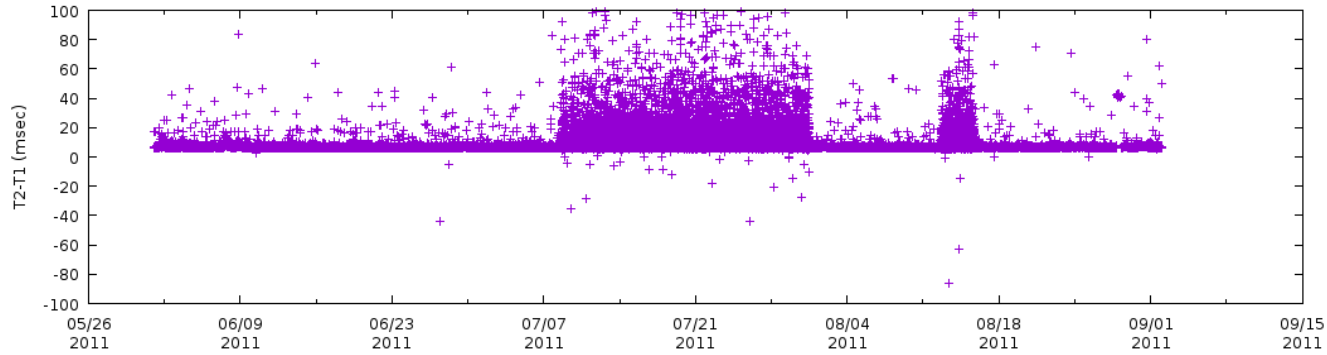

Figure 19 Stale T2, NTP stratum 1 server located in Indiana. MJD 55713-55805. Note "negative" delays

Vendor NTP server software updates greatly reduced the negative delays shown in Figure 19. Operating system support of per-packet time timestamps such as SO_TIMESTAMP has made this problem less common, although it is still seen on some public NTP servers today.

Reference ID

The stratum 1 NTP server reference ID field may be set to any 4-character ASCII string, though standards exist. IANA maintains a list [26] of NTP Reference IDs, all 4 character strings. For example, GPS is assigned the 4 characters 'G' 'P' 'S' '\0'. During monitoring, 17 other strings with variations on GPS were seen. The most common reference IDs observed were GPS, PPS and CDMA. Some reference IDs have clear interpretations: USNO, PTB, or WWVB. References IDs such as 635 and 0x7f are less obvious.

The reference NTP software allows the server administrator to configure stratum 1 operation while using a local, undisciplined clock source. Common choices include LOCL, LCL, CLK, FREE, NONE and FLY. While operating without a

true stratum 0 clock, the NTP server may have degraded accuracy. Should a stratum 1 NTP response listing a local oscillator as the reference ID be discounted as "potentially inaccurate"? The server's root dispersion may or may not reflect the current uncertainty. Without additional information, an NTP client may not be able to distinguish commercial NTP servers using OXCO or Rb oscillators during holdover from less expensive NTP servers that have simply lost their PPS.

NTP servers using PPS and those that rely solely on NMEA sentence arrival time may both use the reference ID `GPS`. (NMEA timings may have significant jitter.)

Unless the semantics are somehow known, it is risky to make use of the reference ID in time transfer.

Off by N-years

On Nov 19, 2012 some responses from a group of North American NTP stratum one servers were in error by 12 years.

Request (T1 from NTP client)	Reply (T2 from server)
Nov 19, 2012 20:47:52.110249	Nov 19, 2012 20:47:52.132228
Nov 19, 2012 20:49:29.259841	Nov 19, 2012 20:49:29.277597
Nov 19, 2012 20:51:06.665791	Nov 19, 2012 20:51:06.684298
Nov 19, 2012 20:52:40.029127	Nov 19, 2012 20:52:40.046487
Nov 19, 2012 20:54:15.354737	Nov 19, **2000** 20:54:15.342575
Nov 19, 2012 20:55:48.475327	Nov 19, **2000** 20:55:48.442512
Nov 19, 2012 20:57:21.298888	Nov 19, 2012 20:57:21.317553
Nov 19, 2012 20:58:56.491482	Nov 19, 2012 20:58:56.509978
Nov 19, 2012 21:00:29.092069	Nov 19, **2000** 21:00:28.987584
Nov 19, 2012 21:02:02.291681	Nov 19, 2012 21:02:02.310593

Etc.

This pattern continued for about one hour. The failure was later noted on the organization's website.
Other multi-year errors were attributed to GPS rollover [27] and resulted in 1024 week timestamp offsets.

NTP Servers Installation and Administration

Inexpensive GNSS synchronized stratum 1 NTP servers may be marketed as "plug and play appliances." Install an antenna, connect a few cables and NTP service is available. Ongoing maintenance may be neglected.

- Hardware degradation or failure should be noticed by an operator and corrected.
- Network connectivity may be lost.
- Software updates may not be installed. Without timely software updates, the NTP server may deliver incorrect time or be vulnerable to security attacks. Some NTP servers became unsupported during this study and later experienced failures such as GPS rollover that caused them to deliver incorrect time. Such units should be removed from service.

It took two days for a failure at a large East coast stratum 1 site to be corrected; see Figure 20. During this period, the server's root dispersion was always 0.

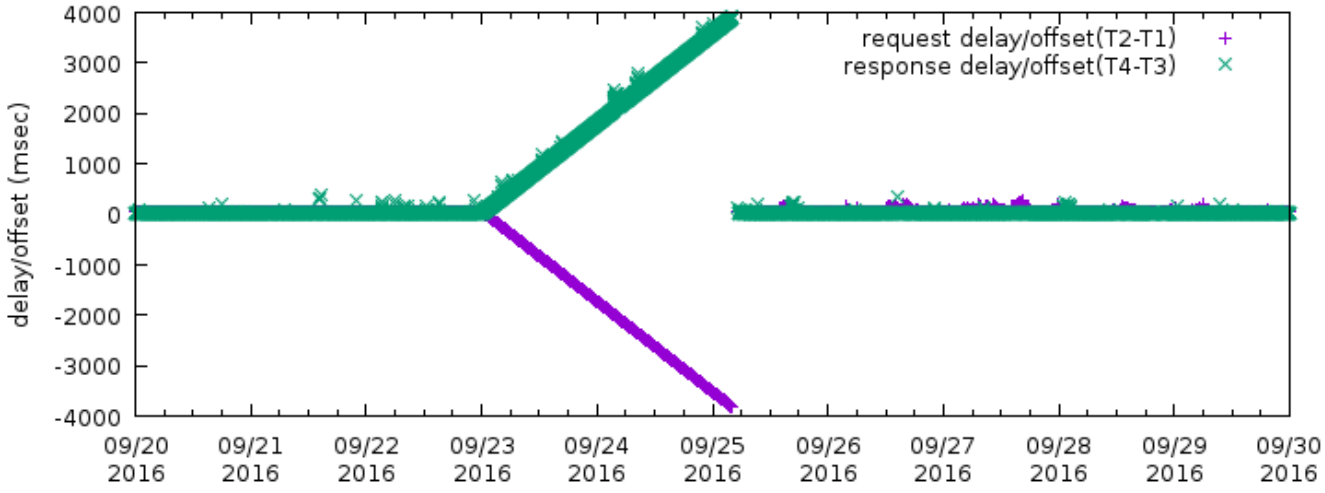

Figure 20 East coast stratum 1 monitored from Midwest. MJD 57651-57660

Two stratum 1 servers in Georgia have had time offsets of one second for several years, possibly due to misconfigured leap second adjustments.

As with other essential network nodes, NTP stratum 1 servers should be administered and constantly monitored.

NTP Server Internal Delays
The NTP server internal delay (T3-T2) is typically insignificant, though there are exceptions.

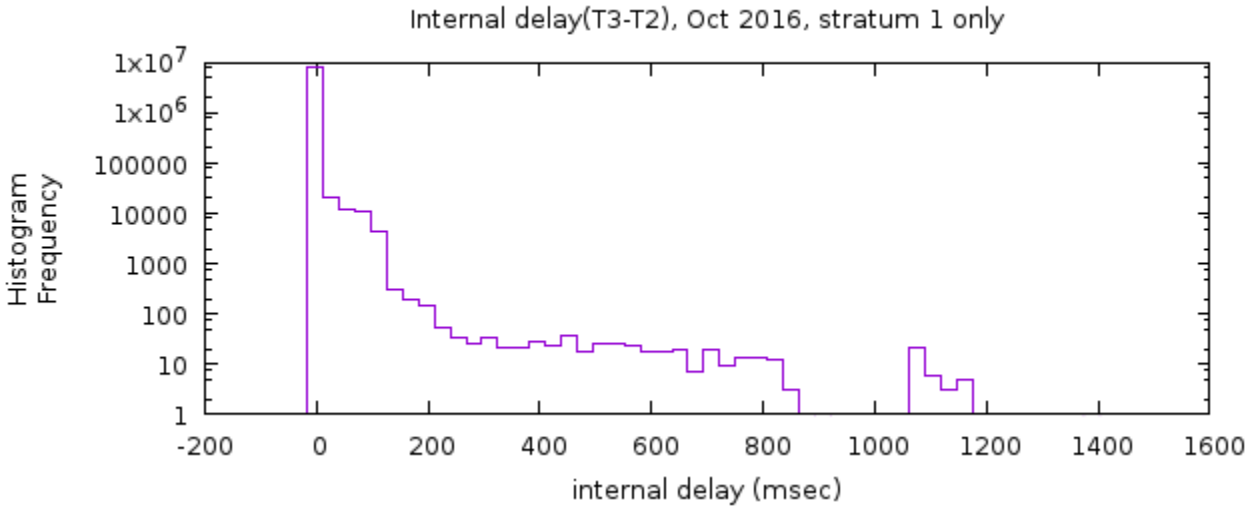

Figure 21 Internal NTP server delay, stratum 1 servers, October 2016 (-100ms < delay < 1500ms)

Sixty of the monitored Stratum 1 servers had one or more NTP responses reporting internal delay > 100 msec. Servers at one European agency occasionally reported slightly negative internal delays (T3<T2).

An infrequent error seen on some stratum 1+ servers has (T3-T2) ~ -1 second. Here are several examples from one server. Times are given in seconds since the Unix Epoch (Jan 1, 1970):

T1	T2	T3	T4	date	MJD
1393240104.969479715	1393240105.999999059	1393240105.000190634	1393240105.030632	2014/02/24	56712
1394374372.967450999	1394374373.999999140	1394374373.000149419	1394374373.032062	2014/03/09	56725
1417388204.967745999	1417388205.999999574	1417388205.000245543	1417388205.022712	2014/11/30	56991
1432805931.969056999	1432805932.999999061	1432805932.000154833	1432805932.026865	2015/05/28	57170

In each case T2 is too large by about one second and falls just prior to a one-second boundary. This particular error is believed to be caused by obsolete versions of the reference NTP software incorrectly “fuzzing” the timestamp low-order-bits.

Experimental NTP Servers
After inquiry, some public NTP servers were found to be running experimental software/hardware. It is not obvious how a client knows that the queried server is being tested.

Root Dispersion
In addition to previously mentioned issues, transient root dispersion errors can occur. The underlined portion of Figure 22 shows server timestamps that are off by 3 seconds yet the root dispersion is only 1 msec.

```
New Zealand server.  2016/11/26 MJD=57718
Stratum=1 reported, leap=00 (no alarm)
T1(client)         T2-T1      T4-T3      Root Disp
  sec                (ms)       (ms)         (ms)
1480178618.447978  97.652    103.259       1.098
1480178917.704253  97.594    114.349       1.022
1480179218.242489  97.927    118.234    6937.805
1480179517.828767 3098.4  -2900.37        1.159
1480179817.867029  3098.1  -2903.04     3000.289
1480180118.020221  98.140     98.945       1.038
1480180418.301428  97.945    118.678       1.221
1480180718.774723  97.948    121.819       1.159
```

Figure 22 Root dispersion variation.

The reference NTP software includes several mechanisms that help clients minimize the impact of one-shot errors

Systemic Problems
Some NTP servers have systemic problems. Figure 23 shows the request and response delay for three NTP servers operated by the same North American organization. RTT cutoffs were used to highlight the server clock drift.

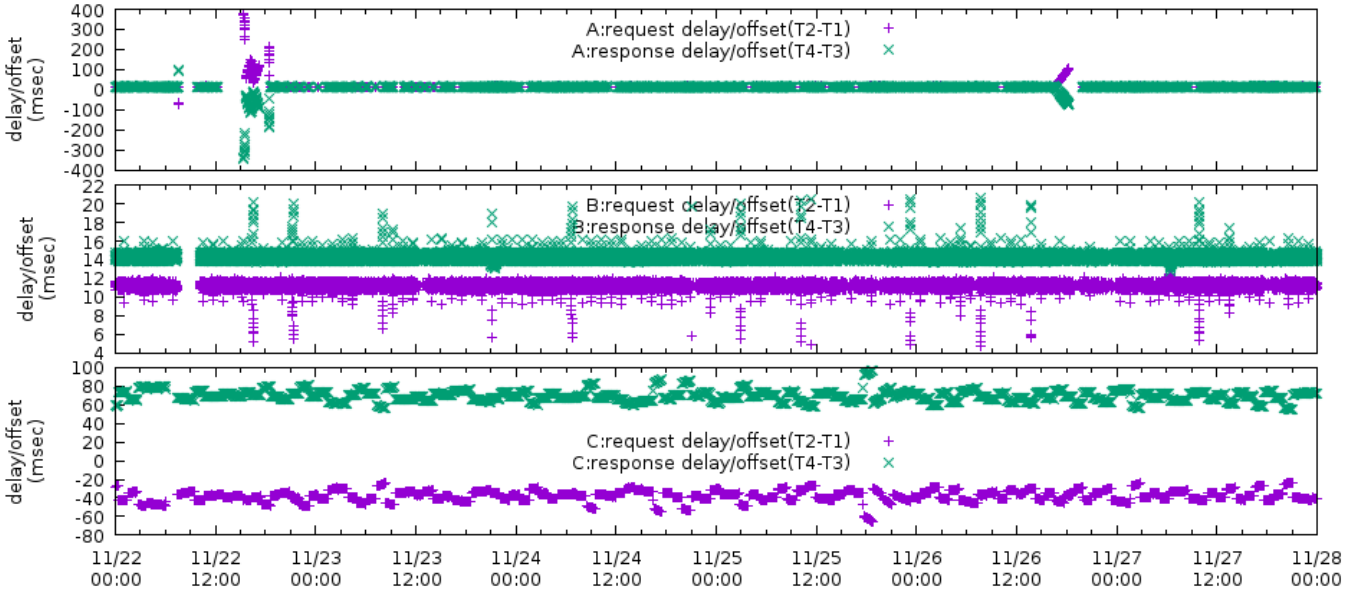

Figure 23 Three North American Stratum 1 servers in 2016. RTT cutoff=32/26/33 msec. (MJD= 57714-57719)

Server A showed time errors approaching 400msec at times. Server B showed frequent small timestamp spikes. Server C showed jitter/wander plus systematic offsets of ~40-50 msec. Reported root dispersions for the three servers were 0, 0 and 9-10 msec for servers A, B and C respectively. The responsible organization has been notified.

INTERNET OF THINGS (IOT)
Many IoT devices needing accurate time will not contain GNSS or other radios. Use of network-based two-way client/server time transfer using protocols such as NTP will continue. Developers should understand the need for accurate time within their application. They should also understand that some public NTP sources run with minimum oversight and may deliver inaccurate timestamps.

CONCLUSION
There are some points to remember when using NTP for time transfer:

- Wired network delays are often unpredictable with varying asymmetry.
  Network asymmetry attributable error is bounded by RTT/2; use of nearby servers may reduce the potential error.
- NTP clients have little/no ability to query NTP servers beyond the basic timestamp exchange.
  Established NTP filtering techniques may help clients deal with transients.
- Even well-maintained NTP servers may occasionally return erroneous timestamps. A few servers may be chronic falsetickers. Using multiple, independent NTP servers is a possible safeguard but may not be foolproof.
- The root dispersion reported by stratum 1+ servers may be optimistic or simply set to 0.
- Clustered NTP servers may have common flaws.
- Use of NTP adds security concerns, both for the device and for others. Critical applications may require use of authenticated time transfers.

Applications with demanding time synchronization requirements may benefit from local, well-administered NTP servers that include stratum 0 clocks or make use of nearby (cloud edge) stratum 1 servers. In an IoT era, local servers may be essential to prevent overloading the small number of public stratum 1 servers.

ACKNOWLEDGEMENTS

The author thanks Steve Myers, Jeff Sommars and Tim Sommars for their help in data collection.

Appendix: NTP Response Message

The 48-byte NTP response message contains these required fields. [8]

Field	Size	Comments
Leap Indicator(LI)	2 bits	00=no leap warning. 01=leap second insertion 10=leap second deletion 11=clock unsynchronized
Version	3 bits	Most recent protocol version = 4
Mode	3 bits	3=NTP client query 4=NTP server response
Stratum	8 bits	0-16
Poll	8 bits	Suggested polling interval in $\log_2$ sec
Precision	8 bits	Precision of system clock in $\log_2$ sec
Root Delay	4 bytes	RTT to reference clock
Root Dispersion	4 bytes	Total dispersion to the reference clock
Reference ID	4 bytes	Identifies server or reference clock
Reference Timestamp	8 bytes	Time when the system clock was last set or corrected
Origin Timestamp (org)	8 bytes	T1 (returned from Mode 3 query)
Receive Timestamp (rec)	8 bytes	T2 (NTP request arrival time)
Transmit Timestamp (xmt)	8 bytes	T3 (NTP response transmission time)

REFERENCES

[1] N. Minar, "A Survey of the NTP Network," 1999. [Online]. Available: alumni.media.mit.edu/~nelson/research/ntp-survey99.

[2] C. D. Murta and P. R. Torres, "NTP Survey - 2005," [Online]. Available: http://www.ntpsurvey.arauc.br/.

[3] T. Yates, "The NTP pool system," [Online]. Available: https://lwn.net/Articles/701222/.

[4] "ntp.org mailing lists," [Online]. Available: http://lists.ntp.org/listinfo.

[5] BIPM, "Time Dissemination Services," 2016. [Online]. Available: ftp://ftp2.bipm.org/pub/tai/scale/TIMESERVICES/timeservices.pdf.

[6] M. A. Lombardi, "International Comparisons of Network Time Protocol Servers," *Proc. 2014 PTTI Mtg.*

[7] D. Matsakis, "Network Time Protocol (NTP) Accuracy As Seen. By The Users," *Proc 2014 PTTI Mtg.*

[8] D. Mills and etal, "RFC 5905. Network Time Protocol Version 4: Protocol and Algorithms Specification," June 2010. [Online]. Available: https://www.ietf.org/rfc/rfc5905.txt.

[9] "NTP: The Network Time Protocol," [Online]. Available: http://www.ntp.org/.

[10] "NTP Project," [Online]. Available: http://nwtime.org/projects/ntp/.

[11] "Welcome to NTPsec," [Online]. Available: NTP: The Network Time Protocol.

[12] "Chrony," [Online]. Available: https://chrony.tuxfamily.org/.

[13] "Ntimed Project," [Online]. Available: http://nwtime.org/projects/ntimed/.

[14] "OpenNTPD," [Online]. Available: http://www.openntpd.org/.

[15] S. Shalunov, B. Teitelbaum, A. Karp, J. Boote and M. Zekauskas, "RFC 4656, A One-way Active Measurement Protocol (OWAMP)," September 2006. [Online]. Available: https://tools.ietf.org/html/rfc4656.

[16] J. Czyz, M. Kallitsis, M. Gharaibeh, C. Papadopoulos, M. Bailey and M. Karir, "Taming the 800 Pound Gorilla: The Rise and Decline of NTP DDoS Attacks," *Proceedings of the 2014 Conference on Internet Measurement Conference*, pp. 435-448, 2014.

[17] J. Gettys, "Bufferbloat: Dark Buffers in the Internet," *IEEE Internet Computing*, 2011.

[18] B. Carpenter and S. Brim, "RFC 3234: Middleboxes: Taxonomy and Issues," 2002.

[19] J. A. Sherman and J. Levine, "Usage Analysis of the NIST Internet Time Service," *Journal of Research of the National Institute of Standards and Technology*, vol. 121, pp. 33-46, 2016.

[20] D. Mills, "ntpq - standard NTP query program," [Online]. Available: https://www.eecis.udel.edu/~mills/ntp/html/ntpq.html.

[21] S. Allen, "The Future of Leap Seconds," [Online]. Available: https://www.ucolick.org/~sla/leapsecs/onlinebib.html.

[22] D. Malone, "Various Time stuff," [Online]. Available: http://www.maths.tcd.ie/~dwmalone/time/.

[23] D. Malone, "The Leap Second Behaviour of NTP Servers," in *Traffic Monitoring and Analysis workshop*, Louvain La Neuve, Belgium, 2016.

[24] M. Shields, "Making every (leap) second count with our new public NTP servers," [Online]. Available: https://cloudplatform.googleblog.com/2016/11/making-every-leap-second-count-with-our-new-public-NTP-servers.html?m=1.

[25] M. Burnicki, "NTP Leap Smearing Test Results," [Online]. Available: https://www.meinberg.de/download/burnicki/ntp_leap_smearing_test_results.pdf.

[26] IANA, "Network Time Protocol (NTP) Parameters," [Online]. Available: https://www.iana.org/assignments/ntp-parameters/ntp-parameters.xhtml.

[27] "Two NTP servers with GPS date calculation errors," 3 May 2015. [Online]. Available: http://lists.ntp.org/pipermail/hackers/2015-May/006866.html.

[28] D. L. Mills, Computer Network Time Synchronization: The Network Time Protocol on Earth and in Space, Second Edition, CRC Press, 2010.

[29] D. Reilly, H. Stenn and D. Sibold, "Network Time Protocol Best Current Practices.," [Online]. Available: https://datatracker.ietf.org/doc/draft-ietf-ntp-bcp/.