Network Working Group
Request for Comments 1128

D. L. Mills
University of Delaware
October 1989

Measured Performance of the Network Time Protocol in the Internet System

Status of this Memo

This paper describes a series of experiments involving over 100,000 hosts of the Internet system and located in the U.S., Europe and the Pacific. The experiments are designed to evaluate the availability, accuracy and reliability of international standard time distribution using the DARPA/NSF Internet and the Network Time Protocol (NTP), which is specified as an Internet Standard in RFC-1119. NTP is designed specifically for use in a large, diverse internet system operating at speeds from mundane to lightwave. In NTP a distributed subnet of time servers operating in a self-organizing, hierarchical, master-slave configuration exchange precision timestamps in order to synchronize subnet clocks to each other and national time standards via wire or radio.

The experiments are designed to locate Internet hosts and gateways that provide time by one of three time distribution protocols and evaluate the accuracy of their indications. For those hosts that support NTP, the experiments determine the distribution of errors and other statistics over paths spanning major portions of the globe. Finally, the experiments evaluate the accuracy and reliability of precision timekeeping using NTP and typical Internet paths involving DARPA, NSFNET and other agency networks. The experiments demonstrate that timekeeping accuracy throughout most portions of the Internet can be ordinarily maintained to within a few tens of milliseconds, even in cases of failure or disruption of clocks, time servers or networks.

This memo does not specify a standard. Distribution of this memo is unlimited.

Keywords: internet clock synchronization, standard time distribution, Internet protocol, timekeeping experiments.

Table of Contents

List of Figures

List of Tables

1. Introduction

How do hosts and gateways in a large, dispersed networking community know what time it is? How accurate are their clocks? In a 1988 survey involving 5,722 hosts and gateways of the Internet system [MIL88a], 1158 provided their local time via the network. Sixty percent of the replies had errors greater than one minute, while ten percent had errors greater than 13 minutes. A few had errors as much as two years. Most host clocks are set by eyeball-and-wristwatch to within a minute or two and rarely checked after that. Many of these are maintained by some sort of battery-backed clock/calender device using a room-temperature quartz oscillator that may drift as much as a second per day and can go for weeks between manual corrections. For many applications, especially distributed internet applications, much greater accuracy and reliability is required.

The Network Time Protocol (NTP) is designed to distribute standard time using the hosts and gateways of the Internet system. The Internet consists of over 100,000 hosts on over 800 packet-switching networks interconnected by a comparable number of gateways. While the Internet backbone networks and gateways are engineered and managed for good service, operating speeds and service reliabilities vary considerably throughout the regional and campus networks of the system. This places severe demands on NTP, which must deliver accurate and reliable standard time in spite of component failures, service disruptions and possibly mis-engineered implementations.

NTP and its forebears were developed and tested on PDP11 computers and the Fuzzball operating system, which was designed specifically for timekeeping precisions of a millisecond or better [MIL88b]. An implementation of NTP as a Unix 4.3bsd system daemon called *ntpd* was built by Michael Petry and Louis Mamakos at the University of Maryland. A special-purpose hardware/software implementation of NTP was built be Dennis Ferguson at the University of Toronto. Over a dozen NTP primary time servers are synchronized by radio or satellite to national time standards in the U.S. and Canada. About half of these are connected directly to international backbone networks and are intended for ubiquitous access, while the remainder are connected to regional networks and intended for regional and local access. It is estimated that there are well over 2000 secondary servers in North America, Europe and the Pacific synchronized by NTP directly or indirectly to these primary servers.

This paper describes several large scale experiments designed to evaluate the availability, accuracy and reliability of standard time distribution using NTP and the hosts and gateways of the Internet. The first is designed to locate hosts that support at least one of three time protocols specified for use in the Internet, including NTP. Since Internet hosts are not centrally administered and network time is not a required service in the TCP/IP protocol suite, experimental determination is the only practical way to estimate the penetration of time service in the Internet. The remaining experiments use only NTP and are designed to assess the nominals and extremes in various types of errors that occur in regular system operation, including those due to the network paths between the servers, the radio propagation path to the source of synchronization and the radio clock itself.

This paper does not describe in detail the architecture or protocols of NTP, nor does it present the rationale for the particular choice of synchronization method and statistical processing algorithms.

Further information on the background, model and algorithms can be found in [MIL89a], while details of the latest NTP protocol specification can be found in [MIL89b].

1.1. Standard Time and Frequency Dissemination

In order that atomic and civil time can be coordinated throughout the world, national administrations operate primary time and frequency standards and maintain Coordinated Universal Time (UTC) by observing various radio broadcasts and through occasional use of portable atomic clocks. A primary frequency standard is an oscillator that can maintain extremely precise frequency relative to a physical phenomenon, such as a transition in the orbital states of an electron. Presently available atomic oscillators are based on the transitions of the hydrogen, cesium and rubidium atoms and are capable of maintaining UTC frequency to 10^{-13} and time to 100 ns when operated in multiple ensembles at various national standards laboratories.

The U.S. National Institute of Standards and Technology (NIST - formerly National Bureau of Standards) operates three radio services for the dissemination of standard time and frequency information [NBS79]. One of these uses high-frequency (HF or CCIR band 7) transmissions from Fort Collins, CO (WWV), and Kauai, HI (WWVH). Signal propagation is usually by reflection from the upper ionospheric layers, which vary in height and composition throughout the day and season and result in unpredictable delay variations at the receiver (see Section 3.2). While these services and those operated by the National Research Council of Canada (CHU) and other countries can be received over large areas of the world, reliable frequency comparisons can be made only to the order of 10^{-7} and time accuracies are limited to the order of a millisecond [BLA74].

A second service operated by NIST is the low-frequency (LF or CCIR band 5) transmissions from Boulder, CO (WWVB), which can be received over the continental U.S. and adjacent coastal areas. Signal propagation is via the lower ionospheric layers, which are relatively stable and have predictable diurnal variations in height. With appropriate receiving and averaging techniques and corrections for diurnal and seasonal propagation effects, frequency comparisons to within 10^{-11} are possible and time accuracies of from a few to 50 microseconds can be obtained [BLA74]. However, there is only one station and it operates at modest power levels.

The third service operated by NIST uses ultra-high frequency (UHF or CCIR band 9) transmissions from the Geosynchronous Orbiting Environmental Satellite (GOES). There is some speculation on the continued operation of GOES, especially if the LORAN-C [FRA82] and Global Positioning System (GPS) [BES82] radiopositioning systems operated by other U.S. agencies continue to evolve as expected. While the OMEGA [VAS78] radionavigation system operated by the U.S. Navy and other countries can in principle provide worldwide frequency and time distribution, this system is unlikely to long survive the operational deployment of GPS.

1.2. The Network Time Protocol

An accurate, reliable time distribution protocol must provide the following:

1. The primary time reference source(s) must be synchronized to national standards by wire, radio or portable clock. The system of time servers and clients must deliver continuous local time based on UTC, even when leap seconds are inserted in the UTC timescale.
2. The time servers must provide accurate and precise time, even with relatively large statistical delays on the transmission paths. This requires careful design of the data smoothing and deglitching algorithms, as well as an extremely stable local clock oscillator and synchronization mechanism.
3. The synchronization subnet must be reliable and survivable, even under unstable conditions and where connectivity may be lost for periods up to days. This requires redundant time servers and diverse transmission paths, as well as a dynamically reconfigurable subnet architecture.
4. The synchronization protocol must operate continuously and provide update information at rates sufficient to compensate for the expected wander of the room-temperature crystal oscillators used in ordinary computer systems. It must operate efficiently with large numbers of time servers and clients in continuous-polled and procedure-call modes and in multicast and point-to-point configurations.
5. The system must operate with a spectrum of systems ranging from personal workstations to supercomputers, but make minimal demands on the operating system and supporting services. Time-server software and especially client software must be easily installed and configured.

In NTP one or more primary time servers synchronize directly to external reference sources such as radio clocks. Secondary time servers synchronize to the primary servers and others in the configured subnet using NTP. Subnet peers calculate clock offset and delay between them using timestamps with 200 picosecond resolution exchanged at intervals of up to about 17 minutes. As explained in [MIL89a], the protocol uses a distributed Bellman-Ford algorithm [BER87] to construct minimum-weight spanning trees within the subnet based on hierarchical level and total synchronization path delay to the root.

Besides NTP, there are several protocols designed to distribute time in local-area networks, including the DAYTIME protocol [POS83a], TIME Protocol [POS83b], ICMP Timestamp message [DAR81] and IP Timestamp option [SU81]. The DCN routing protocol incorporates time synchronization directly into the routing protocol using algorithms similar to NTP [MIL83]. The Unix 4.3bsd time daemon *timed* uses a single master-time daemon to measure offsets of a number of slave hosts and send periodic corrections to them [GUS85]. However, these protocols do not include engineered algorithms to compensate for the effects of statistical delay variations encountered on wide-area networks and are unsuitable for precision time distribution throughout the Internet.

1.3. Determining Time and Frequency

In this paper to synchronize frequency means to adjust the clocks in the network to run at the same frequency, to synchronize time means to set the clocks so that all agree at a particular epoch with respect to UTC, as provided by national standards, and to synchronize clocks means to synchronize

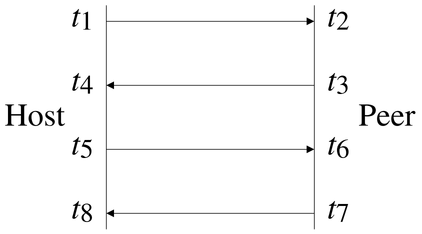

Figure 1. Calculating Delay and Offset

them in both frequency and time. A clock synchronization subnet operates by measuring clock offsets between the various peers in the subnet and so is vulnerable to statistical delay variations on the various transmission paths between them. In the Internet the paths involved can have wide variations in delay and reliability, while the routing algorithms can select landline or satellite paths, public network or dedicated links or even suspend service without prior notice.

In statistically noisy subnets accurate time synchronization requires carefully engineered filtering and selection algorithms and the use of redundant resources and diverse transmission paths, while accurate frequency synchronization requires finely tuned oscillator tracking loops and multiple phase comparisons over relatively long periods of time. For instance, while only a few comparisons are usually adequate to resolve local time for an Internet host to within a few tens of milliseconds, dozens of measurements over many hours are required to resolve frequency to within a few tens of milliseconds per day.

In NTP the roundtrip delay and clock offset relative to a peer time server are calculated as follows. Number the times of sending and receiving NTP messages as shown in Figure 1 and let i be an even integer. Then t_{i-3}, t_{i-2}, t_{i-1} and t_i are the contents of the four most recent timestamps as above. The roundtrip delay d_i and clock offset c_i of the receiving host relative to the sending peer is:

$$d_i = (t_i - t_{i-3}) - (t_{i-1} - t_{i-2}) \ ,$$

$$c_i = \frac{(t_{i-2} - t_{i-3}) + (t_{i-1} - t_i)}{2} \ .$$

This method amounts to a continuously sampled, returnable-time system, which is used in some digital telephone networks [MIT80]. Among the advantages are that the order and timing of the messages are unimportant and that reliable delivery is not required. Obviously, the accuracies achievable depend upon the statistical properties of the outbound and inbound data paths. Further analysis and experimental results bearing on this issue can be found in [COL88], [MIL85a] and [MIL85b].

The computed offsets are first filtered to reduce stochastic noise and then evaluated to determine the most accurate and reliable selection among usually several peers. The filtered offsets from the selected peer are used to adjust the phase and frequency of the host clock, which is implemented by an adaptive-parameter, first-order, phase-locked oscillator. The loop parameters are carefully engineered to provide low frequency errors, so that the host clock will retain good accuracy during subnet outages of a day or more.

Protocol	Valid	Timeout	Error	Unknown
ICMP	11533	61343	265	532
TIME	8441	1400	2293	na
NTP	784	713	6956	na

Table 1. Time Responses by Protocol

2. Discovering Internet Timetellers

An experiment designed to discover Internet time-serving hosts and evaluate the quality of their indications was conducted over a nine-day interval in August 1989. This experiment is an update of previous experiments conducted in 1985 [MIL85b] and early 1988 [MIL88a]. It involved sending time-request messages in each of three time distribution protocols, ICMP Timestamp, TIME and NTP, to every Internet address that could reasonably be associated with a working host. Previously, lists of such addresses were derived from the Internet host table maintained by the Network Information Center (NIC) at SRI International and widely known as the file "hosts.txt". The NIC host table as of 10 August 1989 contained 6382 distinct host and gateway addresses.

With the proliferation of the Internet domain-name system used to resolve host addresses from host names [MOC87], the NIC host table has become increasingly inadequate as a discovery vehicle for working host addresses. In a comprehensive survey of the domain-name system, Mark Lotter of SRI International recently compiled a revised host table of 137,484 entries. Each entry includes two lists, one containing the Internet addresses of a single host or gateway and the other containing its associated domain names. For the experiment this 9.4-megabyte table was sorted by address and extraneous information deleted, such as entries containing missing or invalid addresses, to produce a control file of 112,370 entries.

The experiment itself was conducted with the aid of the control file and a specially constructed experiment program written for the Fuzzball operating system [MIL88b]. The data were collected using experiment hosts located at the University of Delaware and connected to the University of Delaware campus network and SURA regional network. The experiment program reads each entry from the control file in turn and sends time-request messages to the first Internet address found. If no reply is received after one second, the program tries again. If no reply is received after an additional second, the program abandons the attempt and moves to the next entry in the control file. The program accumulates error messages and sample data for up to eight samples in each of the three time protocols. It abandons a host upon receipt of an ICMP error message [DAR81] and abandons further hosts on the same network upon receipt of an ICMP net-unreachable message. Using this procedure, attempts were made to read the clock for 107,799 distinct host addresses.

In the first series of experiments the clock offsets were measured for each of the three time protocols relative to the experiment host local clock, which is synchronized via radio to NBS standards to within a few milliseconds (see below). The maximum, minimum and mean offset for up to eight replies in each protocol was computed and written to a statistics file. It can happen that more than one reply is received for a single time-request message if the roundtrip interval is longer than one

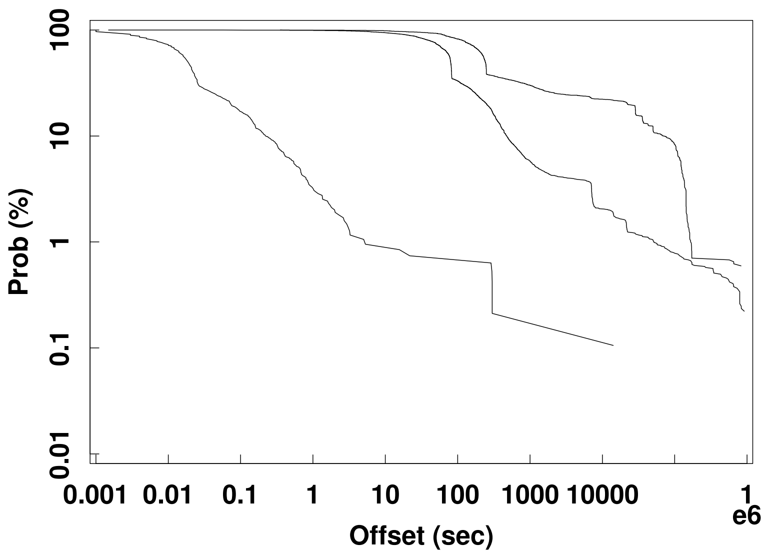

Figure 2. Error Distribution by Protocol

second, in which case only the first reply is counted. The inherent design of all three protocols is resistant to message duplication.

The resulting raw statistics file was then cleaned to remove miscellaneous and unrelated commentary, including occasional duplicates resulting from multiple occurrences of the same name with different addresses. The resulting file contains valid responses, ICMP error messages of various kinds, timeout messages and other error indications. In the tabulation shown in Table 1 the timeout column shows the number of occasions when no reply was received, while the error column shows the error messages received, including ICMP time-exceeded, ICMP host-unreachable and ICMP port-unreachable messages. The unknown column tabulates occurrences of a specially marked ICMP Timestamp reply that indicates the host supports the protocol, but does not have a synchronized time-of-day clock.

In summary, of the 107,799 host addresses surveyed, 94,260 resulted in some kind of entry in the statistics file. Of these 20,758 hosts (22%) were successful in returning an apparently valid indication. Note that there may be more than one attempt to read a host clock and that some clocks were read using more than one protocol.

Then, the list of 20,758 apparently valid responses was processed to remove all entries except the first for each address and protocol. In addition, if a host replied to an NTP request, all other entries for that host were deleted, while, if a host did not reply to an NTP request, but did for a TIME request, all other entries for that host were deleted. This results in a list of 8455 hosts which provided an apparently valid time indication, 3694 for ICMP Timestamp, 7666 for TIME and 789 for NTP.

In order to discover as many NTP hosts as possible, the NTP synchronization subnet operating in the Internet was searched starting from the known primary servers using the Unix NTP-monitoring program *ntpdc* and its Fuzzball counterpart *netspy*. This search, together with the 789 hosts discovered using the domain-name system and additional information gathered by other means, resulted in a total of about 990 NTP hosts. These hosts were then surveyed again, while keeping track of ancillary information to determine whether the implementation and configuration were correct. This resulted in a list of 946 hosts apparently synchronized to the NTP subnet and operating correctly.

2.1. Evaluation of Timekeeping Quality

In evaluating the quality of standard time distribution it is important to understand the effects of errors on the applications using the service. For many applications the maximum error under all conditions is more important than the mean error under controlled conditions. In these applications conventional statistics such as mean and variance are insufficient. A useful statistic has been found to be the error distribution plotted on log-log axes and showing the probability $P(x>a)$ that a sample x from the population exceeds the value a on the x axis. Figure 2 shows the error distributions for each of the three time protocols included in the survey. The top line in Figure 2 is for ICMP Timestamp, the next down is for TIME and the bottom is for NTP.

The graphs shown in Figure 2 suggest several tentative conclusions. First, the time accuracy of the various hosts varies dramatically over at least nine decades from milliseconds to over 11 days. To be sure, not many hosts showed very large errors and there is cause to believe these hosts either were never synchronized or were operating improperly. In the case of NTP, for example, which is designed expressly for time synchronization, eight hosts showed errors above ten seconds, a value considered barely credible for a host synchronized by NTP in the Internet. It is very likely that some or all of these hosts, representing about one percent of the total NTP population, were using an old NTP implementation with known bugs. On the other hand, one percent of the ICMP Timestamp hosts show errors greater than a day, while one percent of TIME hosts show errors greater than a few hours. Clearly, at least on some machines for the latter two protocols, time is not considered a cherished service.

At the other end of the scale, Figure 2 suggests that at least 30 percent of the hosts in all three protocols make some attempt to maintain accurate time to about 30 ms with NTP, a minute with TIME and a couple of minutes with ICMP Timestamp. Between this regime and the one-percent regime the accuracies deteriorate; however, in general, NTP hosts maintain time about a thousand times more accurate than either of the other two protocols.

2.2. Discussion

While this experiment was designed to assess the ubiquity and quality of Internet time service, several interesting facts have emerged. The experiment identified almost 100,000 hosts in the composite domain-name database; however, there is every suspicion that the domain-name survey did not capture the entire collection of hosts identifiable by that means. There may be hidden servers, servers that were down during the survey and network errors of various kinds that occurred during the survey. In addition, there were a few instances where the same address appeared in the domain-name database with different names, which would tend to slightly overestimate the number of distinct hosts.

Years of experience with the Internet have demonstrated the utility of ICMP error messages as useful indicators of network problems. One of the goals in system design is to avoid "black holes," where an attempt to reach a host results in no response at all. The success, or lack thereof, in attaining this goal is apparent in the data above. Of 94,260 attempts to reach a distinct host, 63,456 resulted in no response at all, or about a 67% failure rate. Some of the failures are probably due to the fact that not all hosts support the ICMP Timestamp message or even ICMP messages at all. A conforming host should return an ICMP error message in case of inability to deliver an ICMP echo, timestamp or information-request message; however, this detail of the ICMP specification is not always implemented.

There is reason to believe that many, perhaps most hosts are connected to local nets, most often Ethernets. While it is in principle possible to determine (from address-resolution failures) that a particular host is unreachable, there are few if any gateways that do that. In fact, some gateway implementations drop the first datagram arriving for a host that requires address resolution. This is one of the factors that drove the experiment design to make two attempts for each address rather than one.

The design of the experiment suppresses attempts using TIME if ICMP Timestamp fails and suppresses attempts using NTP if TIME fails. The intent of this design is to avoid network overhead for attempts unlikely to succeed. In the TCP/IP implementations derived from the various recent Berkeley Unix distributions, the standard configuration includes support for ICMP Timestamp and TIME, but not NTP. Some configurations elect not to support either TIME or NTP; however, ICMP Timestamp is supported by the kernel, so is highly likely to be available, even if the others are not.

From the design of the experiment, it would be expected that only those hosts that respond to ICMP Timestamp would be surveyed for TIME. In fact, there are somewhat more hosts surveyed than this. This may be due to any of several factors, including the fact that some hosts are represented more than once in the control file with different names and may behave differently due to network errors on subsequent attempts. In the case of ICMP Timestamp attempts, most of the errors were due to miscellaneous ICMP host-unreachable and time-exceeded messages, but the incidence is well down in the statistical noise. On the other hand, in the case of TIME most of the errors were due to ICMP port-unreachable messages as expected. There seems little justification for the surprisingly high level of timeouts in these cases, since implementations capable of returning ICMP Timestamp messages would ordinarily be capable of returning ICMP port-unreachable messages in case the user-datagram protocol module itself or TIME or NTP was not implemented.

3. Survey of NTP Hosts

The above experiment was designed to assess the performance of all time servers that could be found in the Internet, regardless of protocol, system management discipline or protocol conformance. In another experiment a number of NTP primary time servers was surveyed. Primary servers are synchronized by radio or satellite to NBS standards and located at or near points of entry to national and international backbone networks. Since they are monitored and maintained on a regular basis, their performance can be taken as representative of a managed system.

				Unfiltered		Filtered	
Host	Synch	Hops	Samples	Offset	Delay	Offset	Delay
FORD	GOES	10	8097	2	190	2	178
ISI	WWVB	12	2214	-12	269.5	-9	220
MIT	WWV	11	991	-8	178	-9	163
NCAR	WWVB	8	1563	6	231	6	216
UIUC	WWVB	7	3986	-9	198	-10	177
UMD	WWVB	5	16105	1	60	1	52

Table 2. Time Responses by Category

The experiment operated over a two-week period in August 1989 using paths between six primary servers on the east coast, west coast and midwest. All measurements were made from an experiment host located at the University of Delaware. All of the paths involve links operating at 1.5 Mbps or higher, although there are up to a dozen links on some paths and some lower speed links are in use. Samples of roundtrip delay and clock offset were collected at intervals from one to seventeen minutes on all six paths and the data recorded in files for later analysis.

Table 2 shows the results of the survey, which involved about 33,000 samples. For each server the name, synchronization source, number of network hops and number of samples are shown. The offset and delay columns show the sample medians in milliseconds. The unfiltered columns show the raw data, while the filtered columns show the data after processing by the clock-filter algorithm used in NTP and described below. Note that the number of samples collected depend on whether the server is selected for clock synchronization as determined by the clock-selection algorithm. As in previous surveys of this type, statistics based on the sample median yield more accurate results than those based on the sample mean. However, results for the trimmed mean (also called Fault-Tolerant Average [KOP87]) with 25% of the samples removed are within a millisecond of the values shown in Table 2. The reduced delay for the filtered data is an artifact of the filtering algorithm.

The residual offset errors apparent in Figure 2 can be traced to subtle assymetries in path routing and network/gateway configurations. If these can be calibrated, perhaps using a portable atomic clock, reliable time transfer over the Internet should be possible to a few milliseconds if measurements are made over periods in the order of weeks. Assuming individual time offset measurements can be made with confidence (see below) to this order, frequency transfer over the Internet can be determined to about 10^{-7} in a day and to less than 10^{-8} in two weeks.

3.1. Errors Due to Statistical Delays

In order to more completely assess the accuracy and reliability that clocks can be synchronized using NTP and the Internet, the paths illustrated in Table 2 were carefully measured in several surveys conducted over a period of 18 months. Each survey used up to six time servers and lasted up to two weeks. A typical survey involves the path between experiment hosts at the University of Delaware and USC Information Sciences Institute, located near Los Angeles, over a complex path of up to twelve network hops involving NSFNET, ARPANET and several other nets and gateways. This

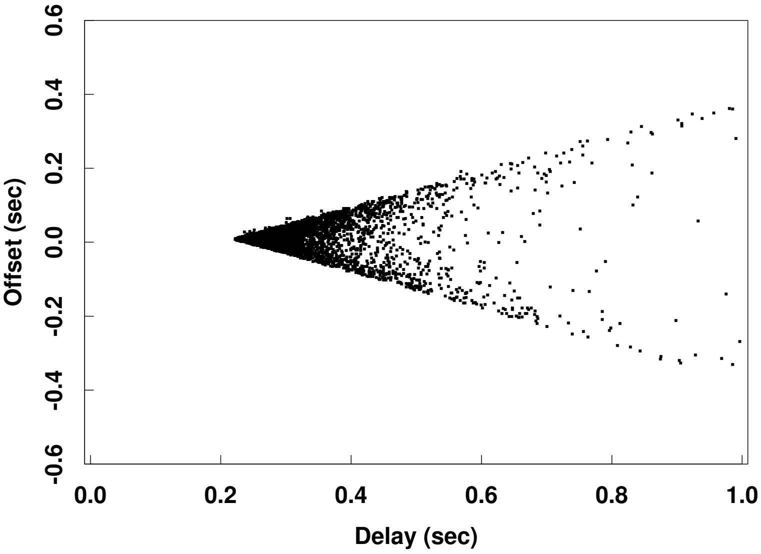

Figure 3. Offset vs Delay

path was purposely selected as among the poorer performing ones in order to determine how well clocks can be synchronized under nonideal conditions.

A number of algorithms for deglitching and filtering time-offset data are summarized in [MIL89a], including trimmed-mean and various kinds of nonlinear filters. The NTP Version 1 protocol used a type of median filter in which a window consisting of the last n sample offsets is continuously updated and the median sample selected as the estimate. In this algorithm the outlyer (sample furthest from the median) is discarded and the process repeated until only a single sample offset is left, which is then produced as the offset estimate. It was used in the Fuzzball and Unix implementations for about two years until the end of 1987.

Experiments during the development of NTP Version 2 have produced an algorithm which provides higher accuracy together with a lower computational burden. The key to the new algorithm became evident through an examination of scatter diagrams plotting clock offset versus roundtrip delay. Without making any assumptions about the distributions of queueing and transmission delays on either direction along the path, but assuming the intrinsic frequency offsets of the host and peer clocks are relatively small, let d_0 and c_0 represent the delay and offset when no other traffic is present on the path and so represents the true values. The problem is to accurately estimate d_0 and c_0 from a sample population of d_i and c_i collected under typical conditions over a relatively long period.

Figure 3 shows a typical scatter diagram for the path under study, in which the points (d_i,c_i) are concentrated near the apex of a wedge defined by lines extending from the apex with slopes +0.5 and -0.5, corresponding to the locus of points as the delay in one direction increases while the delay in the other direction does not. From these data it is obvious that good estimators for (d_0,c_0) are points near the apex and that the best offset samples occur at the lower delays. Therefore, an appropriate technique is simply to select from the *n* most recent samples the sample with lowest

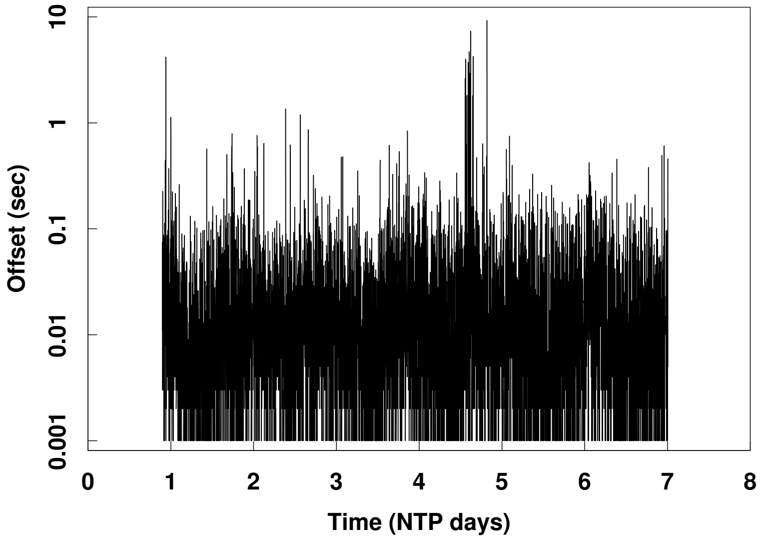

Figure 4. Raw Offsets

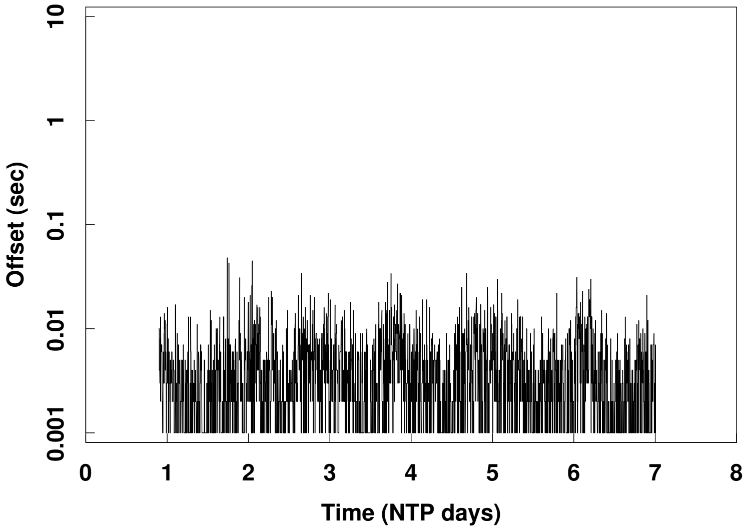

Figure 5. Filtered Offsets

delay and use its associated offset as the estimate. This is the basis of the NTP Version 2 algorithm described in detail in [MIL89b].

Figure 4 shows the raw time-offset series for the path under study over a six-day interval, in which occasional errors up to several seconds are apparent. Figure 5 shows the time-offset series filtered by the NTP Version 2 algorithm, in which large errors have been dramatically reduced and may even reveal a subtle diurnal traffic variation over this path. Finally, the overall performance of the

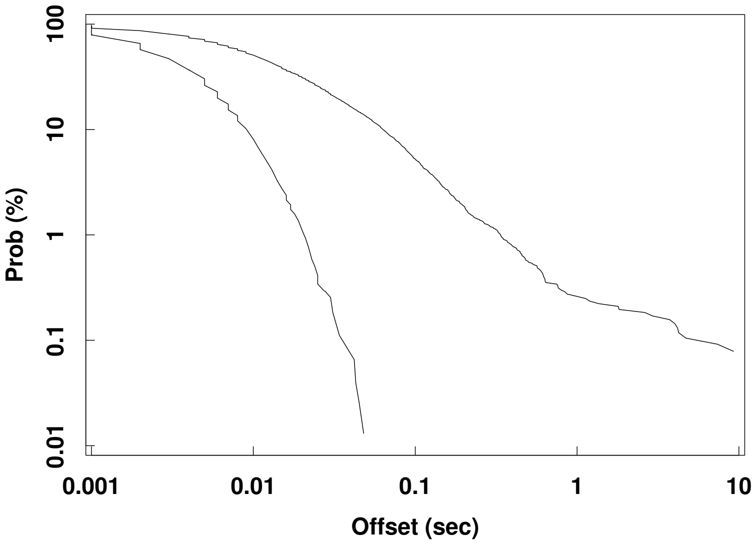

Figure 6. Error Distribution for NTP hosts

path is apparent from the error distributions shown in Figure 6. The upper line shows the distribution for the raw data, while the lower line shows the filtered data. The significant facts apparent from the latter line are that the median error over all samples was only a few milliseconds, while the maximum error was no more than 50 ms.

3.2. Errors due to Radio Propagation Effects

In order to assess the accuracy of the above results, it is necessary to consider the inherent accuracy and precision of the primary synchronization paths and radio clocks themselves. The radio clocks used in the surveys are all capable of resolution to within a millisecond and are potentially accurate to within a millisecond or two relative to the propagation medium. However, the absolute accuracy depends on knowledge of the radio propagation path to the source of standard time and frequency. In the absence of calibration by portable atomic clock, the conventional method is to estimate the propagation delay along the great-circle path between the known geographic coordinates of the transmitter and receiver. However, this can result in errors as great as two milliseconds when compared to the actual oblique ray path. Additional errors can be introduced by unpredictable latencies in the radio clocks, operating system, hardware and in the protocol software (e.g., encryption delays) for NTP itself.

An evaluation of the performance of the synchronization accuracy of the NTP primary servers relative to national standards in principle requires calibration by a portable atomic clock; however, it is possible to estimate the accuracy by means of a detailed analysis of the radio propagation path itself. In the case of the NBS WWVB service on 60 kHz, the variations in path delay are relatively well understood and limited to the order of 50 microseconds [BLA74]. In the case of the NBS GOES service the accuracy is limited by the ability to accurately estimate the distance along the line-of-sight path to the satellite orbit and the ability to maintain accurate stationkeeping in geosynchronous

Locations				Lat	Long
University of Delaware (Newark)				30.68N	75.73W
CHU Radio (Ottawa)				45.30N	75.75W
SSN: 184		Date: 1 Sept 89		Dist: 625 km	
UTC	MUF	Path Delay (ms)			
(hours)	(MHz)	2.5 MHz	5 MHz	10 MHz	15 MHz
0	14.3	5.1j2	5.1j2	5.1j2	
2	11.6	4.0n2	4.0n2	2.7n1	
4	9.7	4.0n2	4.0n2		
6	8.3	4.0n2	4.0n2		
8	7.3	4.0n2	4.0n2		
10	9.2	5.1j2	5.1j2		
12	13.3		5.1j2m	5.1j2	
14	15.4		5.1j2	5.1j2	3.2j1
16	16.5		5.1j2	5.1j2m	3.2j1
18	17.0		5.1j2	5.1j2m	3.2j1
20	16.8		5.1j2	5.1j2	3.2j1
22	16.0		5.1j2m	5.1j2	3.2j1

Table 3. Radio Propagation Delay

orbit. In principle, the estimation errors for either of these services is small compared to the accuracy expected of Internet timestamps generated with NTP.

However, in the case of the NBS WWV/H and NRC CHU services, which operate on HF frequencies from 2.5 through 20 MHz, radio propagation is determined by the upper ionospheric layers, which vary in height throughout the day and night, and by the geometric ray path determined by the maximum usable frequency (MUF) and other factors, which also vary throughout the day, night, season and phase of the 11-year sunspot cycle.

In an effort to calibrate how these effects affect the limiting accuracy of the NTP primary servers using WWV/H and CHU services, existing computer programs were used to determine the MUF and propagation geometry for typical ionospheric conditions forecast for September 1989. The results are shown in Table 3 by two-hour intervals throughout a 24-hour period for the path between the University of Delaware (Newark), and CHU (Ottawa). Each line of the table shows UTC (hour), MUF (MHz) and delay (ms) for forecast frequencies of 2.5, 5, 10 and 15 MHz. In case no propagation path is likely, the delay entry is left blank. The delay itself is followed by a code indicating whether the path is entirely in sunlight (j), in darkness (n) or mixed (x) and the number of hops. A symbol (m) indicates two or more geometric paths are likely with similar amplitudes, which may result in multipath fading and unstable indications.

From Table 3 it can be seen that the delay decreases as the ionospheric layers fall during the night (to about 250 km) and rise during the day (to about 350 km). The delay also changes when the number of hops and thus the oblique ray geometry changes. The maximum delay variation for this particular path is from 2.7 to 5.1 ms, a variation of 2.4 ms. While this may be an extreme case (a forecast path to Hawaii varies from 8.6 to 9.7 ms), the results demonstrate that the ultimate accuracy of HF-radio derived NTP time may depend on the ability to accurately estimate the propagation path variations or to confine observations to the same time each day.

3.3. Errors due to Radio Clock Phase Noise

Precision timekeeping at the NTP primary servers requires an exceptionally stable local oscillator reference in order to deliver accurate time when the radio propagation medium, transmitter or radio clock has failed. Furthermore, the oscillator must maintain accurate frequency in case the radio clock has excessive phase noise or experiences propagation anomaly, such as can happen when a WWV/H radio clock changes frequency. For instance, in order to maintain time to within a few milliseconds for a day without outside synchronization, the local oscillator frequency must be accurate to within 10^{-7} or better.

Accuracies like this usually require a relatively expensive oven-compensated quartz oscillator, which is not a standard component in most computer systems. Accordingly, the NTP host-clock model involves an adaptive-parameter control loop which continuously corrects phase and frequency variations of the reference oscillator relative to the indications received from the radio clock. The loop parameters are chosen to match the characteristics of uncompensated board-mounted crystals used in most computing equipment, where frequency can vary several parts in 10^6 as the result of normal room temperature changes. Commercial radio clocks typically have similar oscillators and control loops, although none are known with the adaptive-parameter design used in NTP.

A problem occurs when the reference oscillator in the radio clock itself becomes destabilized due to propagation path disturbance or, in case of WWV/H clocks, when the path fails and a frequency change is made. Sometimes this can result in temporary frequency surges, which the reference oscillator in the primary server will attempt to follow. If synchronization with the radio transmitter is lost following a surge, the primary server will ordinarily continue at its last estimated frequency, which can lead to gross time errors if allowed to continue. In order to cope with this problem, NTP continuously evaluates the quality of the time indications received from the radio clock and modulates their affect on the frequency estimate. If the quality estimate is low, the effect on the frequency estimate is reduced; while, if the estimate is high, the effect is increased. The goal is to maintain the best frequency estimate possible in the face of widely varying quality in indications received from the radio clock.

The final experiment reported in this paper involves an assessment of how well the NTP estimation algorithm behaves under typical propagation conditions with a commercial WWV/H radio clock. In order to separate the effects of host-clock wander from the effects due to the radio clock itself, the host clock was derived from a precision oven-compensated quartz oscillator with rated stability of $\pm 5 \times 10^{-9}$ per day and aging rate of 1×10^{-9} per day. The oscillator was set to within an estimated accuracy of $\pm 1 \times 10^{-8}$ relative to the 20-MHz WWV transmission under good propagation conditions near midday at the midpoint of the propagation path. The estimated offset and frequency ordinarily

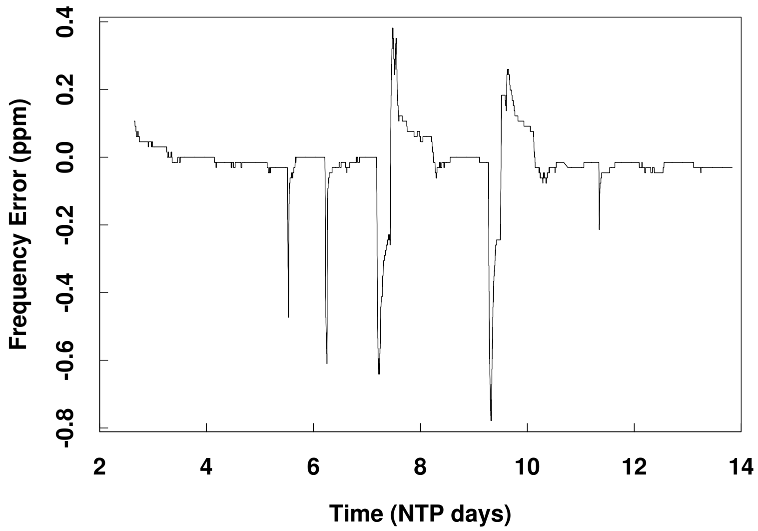

Figure 7. Radio Clock Frequency Error

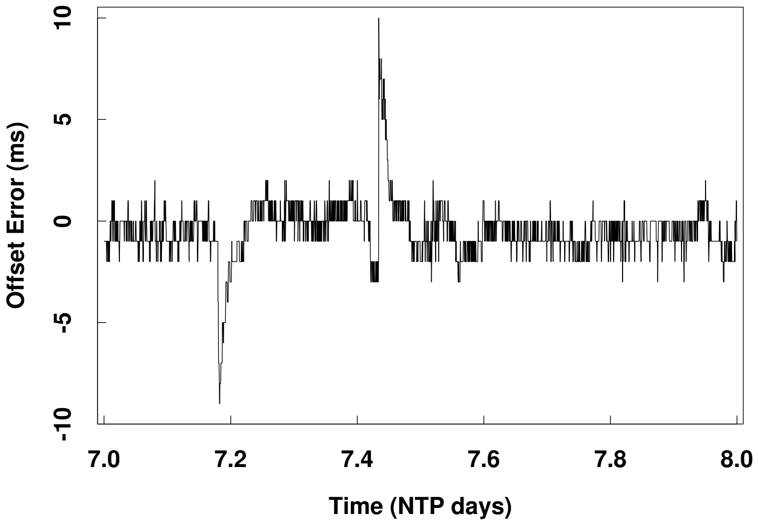

Figure 8. Radio Clock Offset Error

produced by the NTP local clock algorithm was then recorded at 30-second intervals for a period of about weeks.

The results of the experiment are shown in Figure 7 and Figure 8. Figure 7 shows the estimated frequency error for the entire period and reveals generally good behavior, except for occasional periods where apparent phase hits cause the control loop to surge. The times of these surges are near times when the path MUF between the transmitter and receiver are changing rapidly (see Table 3) and the receiver must change operating frequency to match. An explanation for the surges is evident in Figure 8, which shows the measured offsets during an interval including a typical surge. The figure shows a negative phase excursion of about 10 ms near the time the MUF would ordinarily fall in the evening and a similar positive excursion near the time the MUF would ordinarily rise in the morning. Since these excursions are far beyond those expected due to ionospheric behavior alone, the results may point to a defect in the radio clock design itself.

4. Conclusions

Over the years it has become something of a challenge to discover and implement architectures, algorithms and protocols which deliver precision time in a statistically rambunctious Internet. In perspective, for the ultimate accuracy in frequency and time transfer, navigation systems such as LORAN-C, OMEGA and GPS, augmented by portable atomic clocks, are the preferred method. On the other hand, it is of some interest to identify the limitations and estimate the magnitude of the timekeeping errors using NTP and typical Internet hosts and network paths. This paper has identified some of what are believed to be the major limitations in accuracy and measured their effects in large-scale experiments involving major portions of the Internet.

The results demonstrated in this paper suggest several improvements that can be made in subsequent versions of the protocol and hardware/software implementations, such as improved radio-clock designs, improved timebase hardware, at least at the primary servers, improved frequency-estimation algorithms and more diligent monitoring of the synchronization subnet. When a sufficient number of these improvements mature, NTP Version 3 may appear.

5. References

[BER87] Bertsekas, D., and R. Gallager. *Data Networks.* Prentice-Hall, Englewood Cliffs, NJ, 1987.

[BES82] Beser, J., and B.W. Parkinson. The application of NAVSTAR differential GPS in the civilian community. *Navigation 29, 2* (Summer 1982).

[BLA74] Blair, B.E. (Ed.). *Time and Frequency Theory and Fundamentals*. National Bureau of Standards Monograph 140, U.S. Department of Commerce, 1974.

[COL88] Cole, R., and C. Foxcroft. An experiment in clock synchronisation. *The Computer Journal 31*, 6 (1988), 496-502.

[DAR81] Defense Advanced Research Projects Agency. Internet Control Message Protocol. DARPA Network Working Group Report RFC-792, USC Information Sciences Institute, September 1981.

[FRA82] Frank, R.L. History of LORAN-C. *Navigation 29, 1* (Spring 1982).

[GUS85] Gusella, R., and S. Zatti. The Berkeley UNIX 4.3BSD time synchronization protocol: protocol specification. Technical Report UCB/CSD 85/250, University of California, Berkeley, June 1985.

[KOP87] Kopetz, H., and W. Ochsenreiter. Clock synchronization in distributed real-time systems. *IEEE Trans. Computers C-36,* 8 (August 1987), 933-939.

[MIL83] Mills, D.L. DCN local-network protocols. DARPA Network Working Group Report RFC-891, M/A-COM Linkabit, December 1983.

[MIL85a] Mills, D.L. Algorithms for synchronizing network clocks. DARPA Network Working Group Report RFC-956, M/A-COM Linkabit, September 1985.

[MIL85b] Mills, D.L. Experiments in network clock synchronization. DARPA Network Working Group Report RFC-957, M/A-COM Linkabit, September 1985.

[MIL88a] Mills, D.L. Network Time Protocol (version 1) specification and implementation. DARPA Network Working Group Report RFC-1059, University of Delaware, July 1988.

[MIL88b] Mills, D.L. The fuzzball. *Proc. ACM SIGCOMM 88 Symposium* (Palo Alto, CA, August 1988), 115-122.

[MIL89a] Mills, D.L. Internet time synchronization: the Network Time Protocol. DARPA Network Working Group Report RFC-1129, University of Delaware, October 1989.

[MIL89b] Mills, D.L. Network Time Protocol (version 2) specification and implementation. DARPA Network Working Group Report RFC-1119, University of Delaware, September 1989.

[MIT80] Mitra, D. Network synchronization: analysis of a hybrid of master-slave and mutual synchronization. *IEEE Trans. Communications COM-28,* 8 (August 1980), 1245-1259.

[MOC87] Mockapetris, P. Domain names - concepts and facilities. DARPA Network Working Group Report RFC-1034, USC Information Sciences Institute, November 1987.

[NBS79] *Time and Frequency Dissemination Services.* NBS Special Publication 432, U.S. Department of Commerce, 1979.

[POS83a] Postel, J. Daytime protocol. DARPA Network Working Group Report RFC-867, USC Information Sciences Institute, May 1983.

[POS83b] Postel, J. Time protocol. DARPA Network Working Group Report RFC-868, USC Information Sciences Institute, May 1983.

[SU81] Su, Z. A specification of the Internet protocol (IP) timestamp option. DARPA Network Working Group Report RFC-781. SRI International, May 1981.

[VAS78] Vass, E.R. OMEGA navigation system: present status and plans 1977-1980. *Navigation 25, 1* (Spring 1978).

Security considerations

Not applicable.

Author's address

David L. Mills
Electrical Engineering Department
University of Delaware
Newark, DE 19716
Phone (302) 451-8247
EMail mills@udel.edu