Internet Timekeeping Around the Globe[1,2]

David L. Mills[3], Ajit Thyagarjan and Brian C. Huffman
Electrical and Computer Engineering Department
University of Delaware

Abstract

This paper describes a massive survey of Network Time Protocol (NTP) servers and clients in order to discover statistics of time and frequency errors, as well as determine the health and welfare of the NTP synchronization subnet operating in the global Internet. Among the conclusions of the survey are that most NTP clocks are within 21 ms of their synchronization sources and all are within 29 ms on average. However, additional errors up to one-half the roundtrip delay are possible, but relatively rare. There is, however, a disturbing incidence of improperly operating servers and misguided server configurations.

Keywords: computer time, network time, time synchronization

1. Introduction

Timekeeping via the Internet has become a ubiquitous service extending to well over 100,000 public servers and clients in many countries of the world. In addition, an uncounted number of private servers and clients lurk behind the firewalls of many large government and corporate networks. There have been several previous studies [1], [3], [4] which have compiled population and error statistics using time-dissemination protocols such as the Network Time Protocol (NTP) and others. However, only [4] considered reaching out to all hosts and routers in the global Internet. That study indexed the Domain Name System database of eight years ago and found about 21,000 hosts that returned time in one fashion or another. The newer studies used only a small population to evaluate protocol and algorithm improvements over the years since then.

Since the present day Internet has experienced an explosion of numbers, it is no longer acceptable or even possible to survey Internet hosts and routers by indexing a public directory service. However, it is possible to discover them for an organized subnetwork such as used by NTP. This paper presents a comprehensive survey of the NTP servers operating in the global Internet of today. While the search engine used in the survey could find only a fraction of all NTP outposts on the public Internet, it did find 182,538 network paths used by 36,000 servers, 231 of these synchronized to UTC via radio, satellite or modem. Uncounted others, including large numbers of personal workstations operating as broadcast clients, escaped detection. The survey measured the time offsets of each client relative to its server, as well as related variables important to the health and welfare of the *NTP subnet* itself.

A comprehensive description of the NTP architecture, protocol and algorithms is beyond the scope of this paper, but can be found in [2] and citations found there. The NTP timekeeping network consists of a hierarchical tree of primary (stratum 1) time servers at the root. These provide synchronization to secondary servers at increasing stratum levels (hops) to the leaves of the tree. Figure 1 shows the functional organization of a NTP server. The server sends a message to one or more other *peer servers* and expects replies from each of them at some later time. From the four timestamps obtained upon departure and arrival of the request and response messages, the server calculates the clock offset and roundtrip delay relative to each peer. The clock filter algorithm selects the best measurements from each peer and the intersection and clustering algorithms the best

---

1. Sponsored by: DARPA Information Technology Office Contract DABT 63-95-C-0046, NSF Division of Network and Communications Research and Infrastructure Grant NCR 93-01002, Northeastern Center for Electrical Engineering Education Contract A303 276-93, Army Research Laboratories Cooperative Agreement DAA L01-96-2-002, and Digital Equipment Corporation Research Agreement 1417.

2. Reprinted from: Mills, D.L., A. Thyagarajan and B.C. Huffman. Internet timekeeping around the globe. *Proc. Precision Time and Time Interval (PTTI) Applications and Planning Meeting* (Long Beach CA, December 1997), 365-371.

3. Author's address: Electrical Engineering Department, University of Delaware, Newark, DE 19716; Internet mail: mills@udel.edu; URL: www.eecis.udel.edu/~mills.

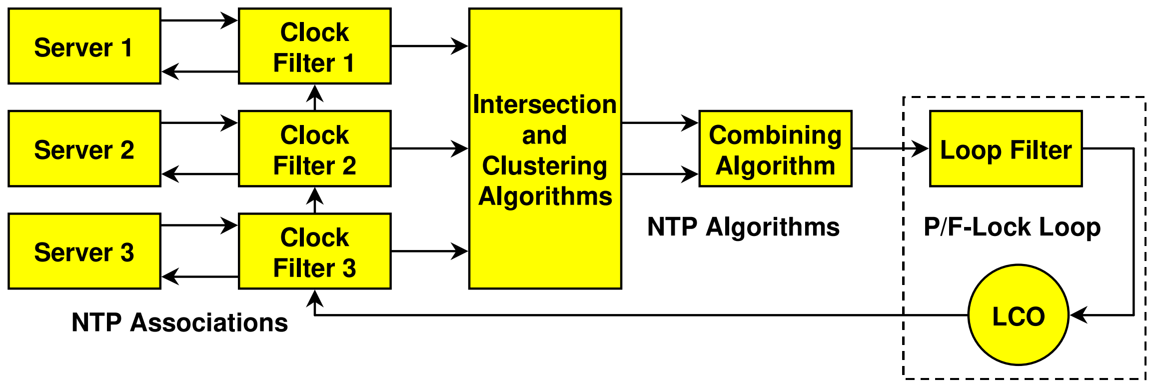

Figure 1. NTP Server Functional Organization

combination of several peere operating simultaneously. The actual local clock oscillator (LCO) time and frequency are disciplined using an adaptive parameter, hybrid phase/frequency-lock loop.

2. Survey Methodology

The survey which generated the bulk of the results presented in this paper was performed over some months in late 1995. The goal was to find as many NTP servers operating in the Internet as practical and determine their performance with respect to time and frequency errors, as well as the robustness and stability of the NTP subnet itself. The results serve as a baseline calibration of the subnet performance, as well as useful insights into subnet configuration and security management strategies.

The search engine used a variety of techniques, ranging from indexing available public databases to exploring the Internet NTP subnet in real time with available monitoring protocols and tools. In the following, *subnet* without the qualifier NTP means any subtree of the global NTP subnet tree. The strategy was to identify as many servers as possible from public lists, such as the database of public servers maintained at the NTP web page *www.eecis.udel.edu/~ntp*. There are currently 69 public primary (stratum 1) servers and 105 public secondary (stratum 2) servers listed in that database. In the initial round, these servers were queried using NTP monitoring and control protocols to identify their client servers. In the next round, those servers were queried to identify their client servers and so on. The survey continued for many rounds until no new servers could be found. Occasionally, the identity of a new server was discovered by other means and the survey continued for new rounds as necessary.

It was vital in the planning and execution of the survey to avoid network and server congestion due to the query messages themselves. Therefore, the intervals between messages were randomized over moderately long intervals in the order of minutes. Due to the volume of data involved, individual rounds required up to days to complete and the entire survey required many rounds.

The data for each server consists of a spreadsheet, where each line of the spreadsheet corresponds to a protocol instantiation, called an *association* between the server and its peer server, as shown in Figure 1. The peer is either another server or a primary reference clock, usually a GPS receiver. Typical server configurations include several associations, which run simultaneously in order to provide redundancy and diversity. The association data includes the endpoint Internet addresses, stratum, mode, reachability status, relative clock offset, roundtrip delay and error estimate called *dispersion*.

Each association operates in a designated mode, depending on whether the server can synchronize the peer or whether the peer can synchronize the server. Symmetric modes are used when either peer can synchronize the other and where mutual backup is necessary in multiple-peer subnets. Client/server modes are used in subnets of modest size and where the highest accuracy is required. Broadcast modes are used in large subnets where the delays are small and the accuracy requirements not demanding.

The data analysis phase of the project considered 230,774 associations, but discarded 818 with misconfigured servers, 3,673 with broadcast servers, 25,640 with unreachable servers, 17,195 with invalid stratum and 1,293 duplicates, leaving 180,520 representing valid measurements of clock offset and roundtrip delay. In these, 36,478 distinct Internet addresses were found, each identifying a NTP server. It should be noted that these servers represent only a fraction of all servers in the NTP subnet. Clients operating in broadcast mode were not found, nor were many others on private networks or behind firewalls. Since these statistics were

Stratum	Servers	Associations	Maximum	Top Ten	Mean
1	220	11,223	652	421	51.0
2	4,438	49,164	356	276	11.1
3	6,591	106,825	558	377	16.2
4	2,254	14,221	398	262	6.3
5	317	9,90	68	25	3.1
6-14	60	115	na	na	na
Total	13,880	182,538	652	421	13.2

Figure 2. Association Population by Stratum

determined, a few additional valid associations were found, bringing the total to 182,538 associations and 38,722 servers.

In the above totals, there were 3,673 servers operating in broadcast modes, or about one in ten. The number of clients of these servers is unknown, since broadcast clients ordinarily do not exchange messages with their servers. In fact, most broadcast servers have tens to hundreds of clients, so the actual client population is considerably underestimated. The 1,293 duplicates represent symmetric associations, where each of two peers agree to back up the other. Ordinarily, the offsets measured by symmetric peers have nearly the same magnitude, but opposite sign, and the delays are nearly the same, but this was not a factor in the analysis.

Figure 2 shows the number of associations by stratum. Among the 182,538 associations, there are 13,880 different servers with an overall average of 13.2 associations per server. The maximum number of associations per server is 652, but this understates the true value, since the maximum number the server monitoring cache can hold is near this number. From other data discussed later, higher values to 734 have been observed on one server which is certainly not the most heavily loaded. The fact that at most stratum levels the maximum and the top ten have about the same number of associations and the overall average is much lower suggests the population is dominated by a relatively small number of heavily loaded servers and the remainder relatively lightly loaded. About half the number of servers operate at stratum 3, which usually function as department servers dependent on redundant campus servers operating at stratum 2. The servers shown operating at stratum 6 and above are very likely broken or unstable.

3. Performance Indicators

The three most significant factors that affect timekeeping accuracy are network jitter, oscillator frequency stability and asymmetric network delays. The influence of

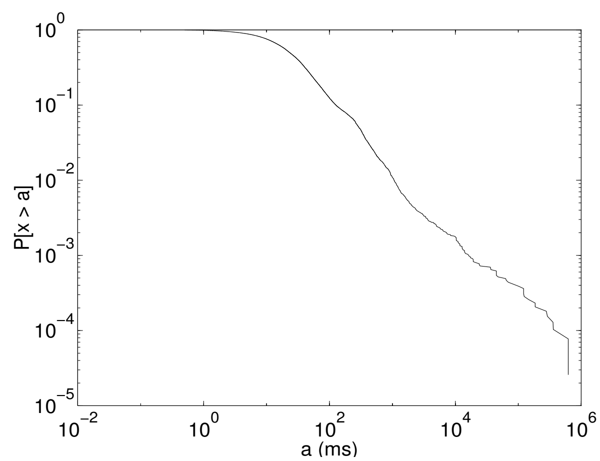

Figure 3. Time Offsets

these three factors varies widely in the NTP subnet, depending on the particular network path and particular oscillator implementation. The survey results allow each of these factors to be studied, evaluated and displayed as graphs and tables and statistics derived from them. However, the survey was conducted over a period of months, where the particular network conditions were not completely stationary. For this reason, the results should be regarded as an approximation to a true snapshot.

3.1 Time Offsets

The most important indicator of timekeeping performance is the time offset measured by the server relative to the peer server for each association. Figure 3 shows the cumulative distribution function (CDF) of time offsets for 38,722 servers, where the median is 23.3 ms, mean 234 ms and maximum 686 ms. However, 3,833 servers with offsets greater than 128 ms were not synchronized by NTP at the time of the survey. When these servers are excluded, the median is 20.1 ms and mean 28.7 ms. Note that these results represent the global population at large, whereas clients on typical LANs have offsets generally less than a millisecond [3].

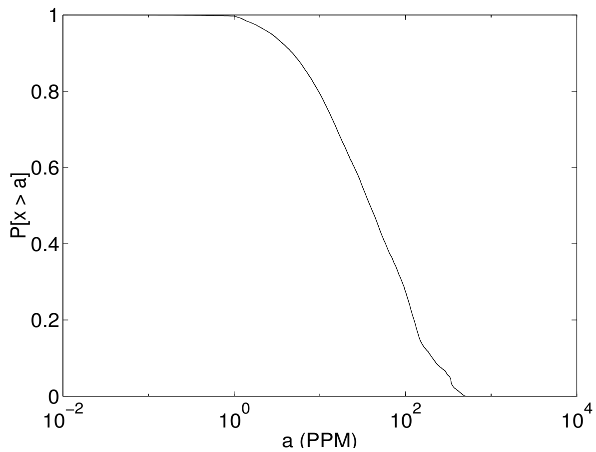

Figure 4. Systematic Frequency Offsets

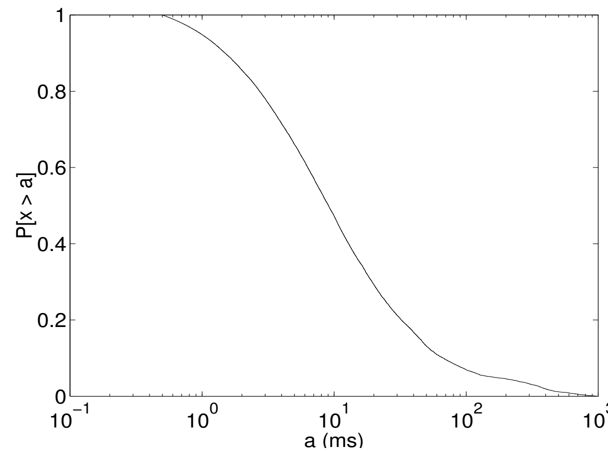

Figure 5. Oscillator Phase Errors

According to Figure 2, of the total population of 38,722 servers, 220 primary servers at the root of the NTP subnet tree provide synchronization to 13,660 secondary servers and these provide synchronization to the 24,842 remaining servers at the leaves of the tree. In the present NTP design, each server on the path from the leaves to the root independently disciplines its clock relative to the servers at the next lower stratum, in order to provide a stable, low-jitter reference for local applications. Thus, the figure shows the offset of a client relative to its server, not the offset relative to the primary servers at the root. In principle, the protocol could determine this by accumulating the offsets on each hop along the path to the primary servers. However, this statistic would result in large jitter and be unsuitable if used to discipline the system clock. In the interest of an economical packet header, it was not included in the NTP protocol design.

3.2 Clock Discipline Errors

The NTP clock discipline algorithm operates to minimize both phase and frequency errors. To the extent it makes good predictions based on measured time offsets produced by the protocol, phase errors due to systematic frequency offsets can be minimized. However, in order to provide millisecond accuracies with measurement intervals up to 1024 s, it is necessary to estimate the current frequency to less than one PPM, even when the systematic frequency offset is hundreds of PPM. The systematic frequency offsets and oscillator phase errors of 19,873 servers were determined from a special survey. Figure 4 shows the CDF for these servers less 593 outlyers with offsets greater than 500 PPM, from which the median is 38.6 PPM and mean 78.1 PPM. Frequency offsets greater than 500 PPM are not credible, since that is the limit of the discipline loop capture range.

While most clock oscillators are moderately well behaved, there is a significant tail on the distribution. The results demonstrate that the NTP clock discipline must have a very large capture range and an adaptive time constant. In the present design, the capture range is 500 PPM and the time constant varies with the measurement interval from 64 s to 1024 s.

Oscillator phase errors are due both to frequency estimate errors and jitter in the network. Figure 5 shows the CDF for these errors for the 19,873 servers in the special survey, less 131 outlyers with errors over 1 s, from which the median is 9.1 ms and mean 37.0 ms. Errors over 1 s can occur only if the peer is defective and the server is synchronized to another peer. Errors over 128 ms can occur only during initial synchronization after reboot or infrequently due to network or peer server disruptions. If these servers are excluded from the population, the median is 8.0 ms and mean 16.7 ms.

3.3 Asymmetric Delays

The NTP protocol design does not permit an absolute measurement of time difference between servers, unless an external common-view means is available, as in [1]. As a result, clock offset errors cannot be distinguished from asymmetric network delays. There is evidence to suggest that errors due this cause are not uncommon; however, by design the errors due to asymmetric delays are bounded by half the roundtrip delay [2], although errors this large are rare. Figure 6 shows the CDF for the delays measured between the 38,722 servers used in the time offset survey. The results show median delay 118 ms, mean 186 ms and maximum a whopping 1.9 s. As with other survey data analyzed, the tails on the distributions can be very long. From anecdotal evidence, the larger values due to the widely varying incidence of congestion in the Internet of modern times.

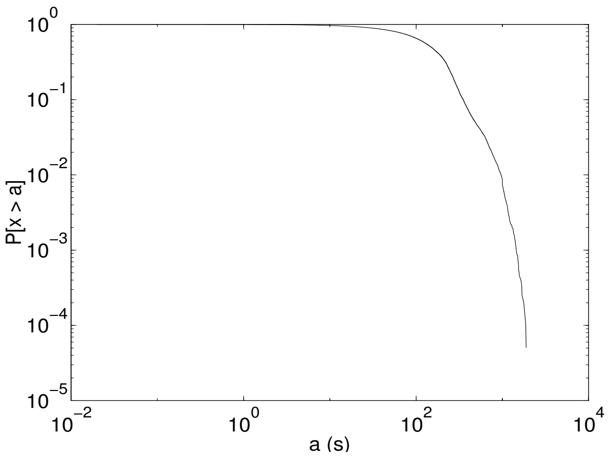

Figure 6. Roundtrip Delays

Count	Status
34,384	Selected for synchronization
319	Miscellaneous checks
60,105	Potential candidates
6,359	Rejected by clustering algorithm
68,966	Standby servers
1,854	Rejected by intersection algorithm
10,551	Peer unreachable
182,538	Total

Figure 7. Data Mitigation

From anecdotal evidence and previously reported results [1], [3], the delay asymmetry is seldom more than one-tenth the roundtrip delay. For LANs with delays in the low milliseconds, the asymmetric errors are generally negligible. Even on long paths spanning the Atlantic where propagation delays dominate, the errors are small compared to those due to network jitter. However, in cases where one direction uses cable and the other satellite, for example, or where two peers are connected to different providers with very different network grids, the errors can be substantial.

3.4 Data Mitigation

A good idea of the general health of the NTP subnet can be determined from the status code maintained for each association. Figure 7 shows a breakdown of the 182,538 associations considered previously. The table is best interpreted starting from the bottom. At each step upward from the bottom, the indicated number of associations are discarded for the reason given on the same line. The line under the heading line shows the number of associations actually used to discipline the system clock. Since there is only one such association per server, this number represents the actual number of servers synchronized by NTP. Since there were 38,722 servers resulting from the time offset survey, the remaining 4,338 servers may be unsychronized, about 11 percent of the total. Caution is advised when making such conclusions, since the survey errors due to changing network conditions could well account for the differences.

For the remaining lines of Figure 7, a few notes will clarify the operations involved. The intersection algorithm detects and discards associations which show a time interval outside the correct time interval determined for the majority of associations. Standby associations are those deemed correct, but not among the lowest ten in order of increasing synchronization distance. When more than ten associations are admitted to the clustering algorithm, excessive *clockhopping* can occur among the peers, which can result in excessive system clock jitter. The relatively high number of standby associations was traced to a provider using a very misguided configuration where each of about 100 NTP servers maintained associations with all of the others.

The clustering algorithm repeatedly discards outlyer associations until either a minimum number of three remain or the residual sample variance cannot be reduced by further discards. The remaining population, including the top three lines in the table, represent candidate associations considered the pick of the litter. Selection among the candidates to determine the ones actually used to discipline the system clock is determined on the basis of synchronization distance. The combining algorithm computes the offset provided to the clock discipline as a average of the available candidates weighted by this distance. Thus, the average number of associations used to synchronize the system clock is 1.75.

3.5 Primary Servers

The accuracy and stability of the global NTP subnet depends on the reference clocks used to synchronize the primary servers. There are many different means used to do this, including radio, satellite, telephone modem and cesium clock ensemble. Figure 8 shows the population characteristics of the 231 primary reference clocks found in the survey. Since this survey was completed, the national time dissemination services of France, Chile and Italy have become available and additional time servers operated by NIST and USNO have come online, making the total operating today significantly larger. The surprisingly large number of DCF77 reference clocks is undoubtedly due to the very inexpensive receivers available for this signal in Europe. In the interval since this survey was completed, and as inexpensive

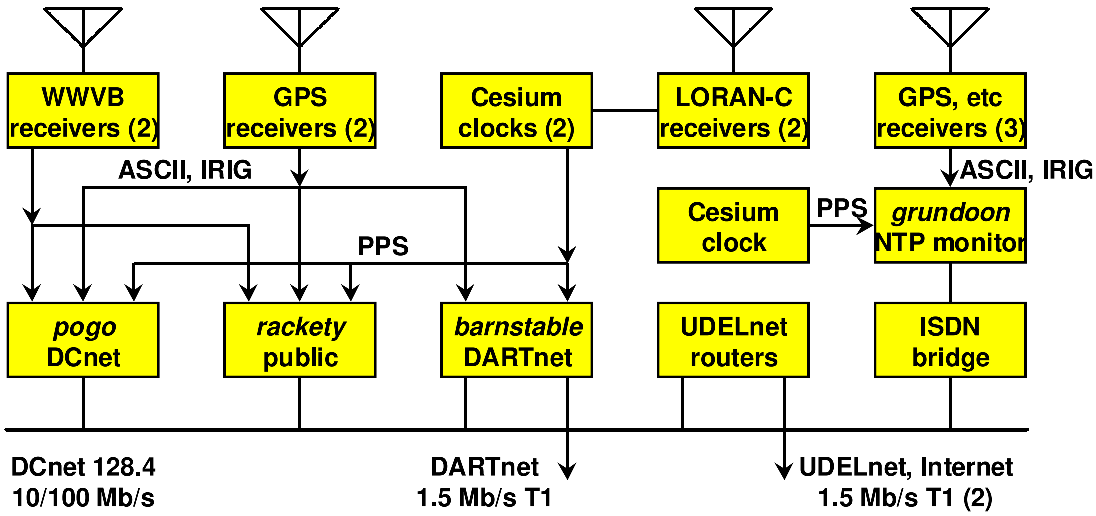

Figure 9. DCN Network Time Servers

Count	Service
47	satellite (GPS, GOES)
57	WWVB radio (US)
17	WWV radio (US)
63	DCF77 radio (Europe)
6	MSF radio (UK)
5	CHU radio (Canada)
7	modem (NIST, USNO, PTB, NPL)
25	other (cesium clock PPS, etc.)
231	Total

Figure 8. Primary Reference Clocks

GPS receivers have come on the market, new servers are using GPS.

In addition to the 231 primary servers found, some 88 servers were found where the reference clock had either failed or had been misconfigured. In the NTP design, reference clocks are treated as ordinary peer servers, so many primary servers maintain associations for other servers as backup. When the radio fails, the other associations kick in and the server continues to operate at a higher stratum.

In some configurations, a special reference clock driver provides backup when all other sources fail. The driver simulates a reference clock operating at some higher stratum, usually 3. If all sources of lower stratum fail, the free-running server system clock becomes the NTP reference and the clients depending on the server synchronize to it. At the time of the survey, 1,502 servers were found to use this feature.

4. A Day in the Live of a Busy NTP Server

While some NTP servers are dedicated timekeepers, most serve other functions as well, including those of routers, file servers and ordinary workstations. It is of some concern to determine how much the resource requirements of the timekeeping function impact on these other functions. Of principle concern are the processing and network overhead resulting from client requests to a busy time server.

Timekeeper *rackety*, one of the most experienced timekeepers in the business, is a well known primary time server with clients all over the world. The configuration of this server in the context of the other timekeeping gear in our laboratory is shown in Figure 9. Together with other timekeepers *pogo* and *barnstable*, this Sun Microsystems SPARCstation IPC functions as a file server. print server and router as well, although in this machine these functions have low demand. An Austron 2201A GPS receiver functions as the primary reference clock and a Spectracom WWVB receiver as backup. The operating system kernel is modified to use the pulse-per-second (PPS) signal from a Hewlett Packard 5061A cesium oscillator to discipline the system clock generally within a few tens of microseconds. While implemented expressly for the SunOS operating system, these kernel modifications are also in current Digital Unix and Sun Solaris operating systems.

In order to calibrate the resource demands of the timekeeping function, ordinary Unix commands were used to measure the CPU time consumed by the NTP daemon over about a day, which came to 1.54 percent. A network sniffer program captured all of the NTP packets for *rackety* that flew the DCnet wires during that day and saved them for later analysis. The 734 clients sent an average of 6.4 packets per second, which corresponds to a mean interarrival interval of 157 s for each client. This is curious, since after an initial training period, most NTP Version 3 clients should increase the poll interval in steps to 1024 s. The fact the interarrival interval is so small suggests there are still a large number of previous versions in use.

Each input packet generated on average 0.64 output packets and required a total of 2.4 ms of CPU time for the request/response transaction. It is not known how many of the 734 clients were using cryptographic authentication; but, if so, the time to generate the request and response cryptosums on an IPC is about 300 μs, but only about 30 μs on an UltraSPARC 170.

Rackety is connected via a 10-Mb/s Ethernet, routers and 1.544 Mb/s T1 circuits to Internet backbone networks as shown in Figure 9. While the T1 circuits are not dedicated to NTP traffic, the other customers are few. Averaged over the day, the 76-octet NTP packet rate was measured at 10.5 packets per second. Exclusive of overhead bits, this is.064 percent of the Ethernet capacity and a maximum of 0.21 percent on each direction of the T1 capacity. The conclusion to be drawn here is that the resources required for NTP are minimal, even when hundreds of clients are involved and even with a slow processor.

5. Conclusions

The results of the survey demonstrate that most NTP clocks are within 21 ms of their servers and all are within 29 ms on average. The oscillator phase error may add a few milliseconds to these figures, but in most cases can be neglected. There may be additional errors due to asymmetrical paths, but this is bounded by one-half the roundtrip delay. While it is not possible using the NTP protocol itself to estimate the errors when multiple-hop paths are involved, a reasonable estimate may be the number of hops times these figures. It is clear from the graphs that the error distributions can have surprisingly large tails well above these figures, in some cases due to network congestion and in other cases due to improperly operating servers or misguided choice of servers.

The roundtrip delay statistic has proven very useful, not only to assess timekeeping performance, but in the general study of the worldwide Internet performance. The fact that NTP servers determine this statistic on a regular basis suggests the use of NTP as a network monitor with a coupling to the monitoring infrastructure using the Simple Network Monitoring Protocol (SNMP). In fact, this has been done on an experimental basis and described in a report now in preparation.

Perhaps the most disturbing results of the survey is the significant fraction of servers showing improper operating parameters or incorrectly configured servers. In many cases, this is due to benign neglect on the part of the operator, since timekeeping is not usually regarded as a revenue service, and providing NTP service to the community is a volunteer sport. This problem has been addressed in the design of the next version of the NTP protocol in the form of features to discover servers and configure the NTP subnet automatically and without the chance of operator error.

5.1 References

Note: The following publications are available in PostScript format at *www.eecis.udel.edu/~mills*.

1. Mills, D.L. The network computer as precision timekeeper. Proc. PTTI Meeting (December 1996).

2. Mills, D.L. Improved algorithms for synchronizing computer network clocks. IEEE/ACM Trans. Networks (June 1995).

3. Mills, D.L. Precision synchronization of computer network clocks. ACM Computer Communications Review (April, 1994).

4. Mills, D.L. On the accuracy and stability of clocks synchronized by the Network Time Protocol, ACM Computer Communications Review (January, 1990).