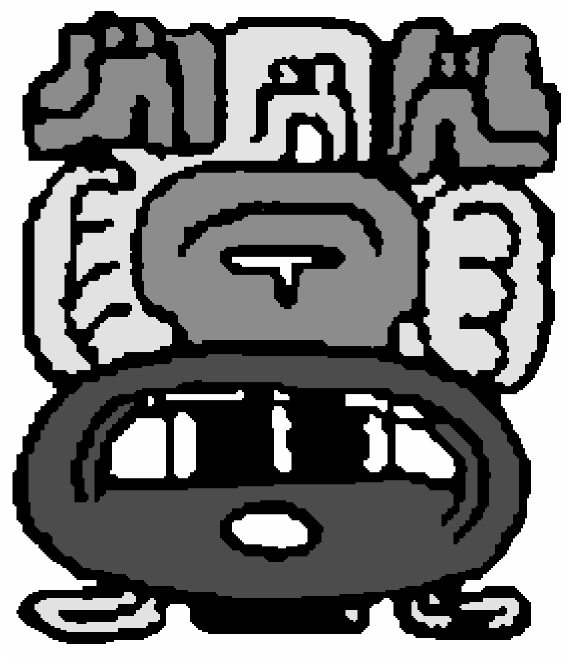
The Network Computer as Precision Timekeeper

David L. Mills
University of Delaware
http://www.eecis.udel.edu/~mills
mills@udel.edu

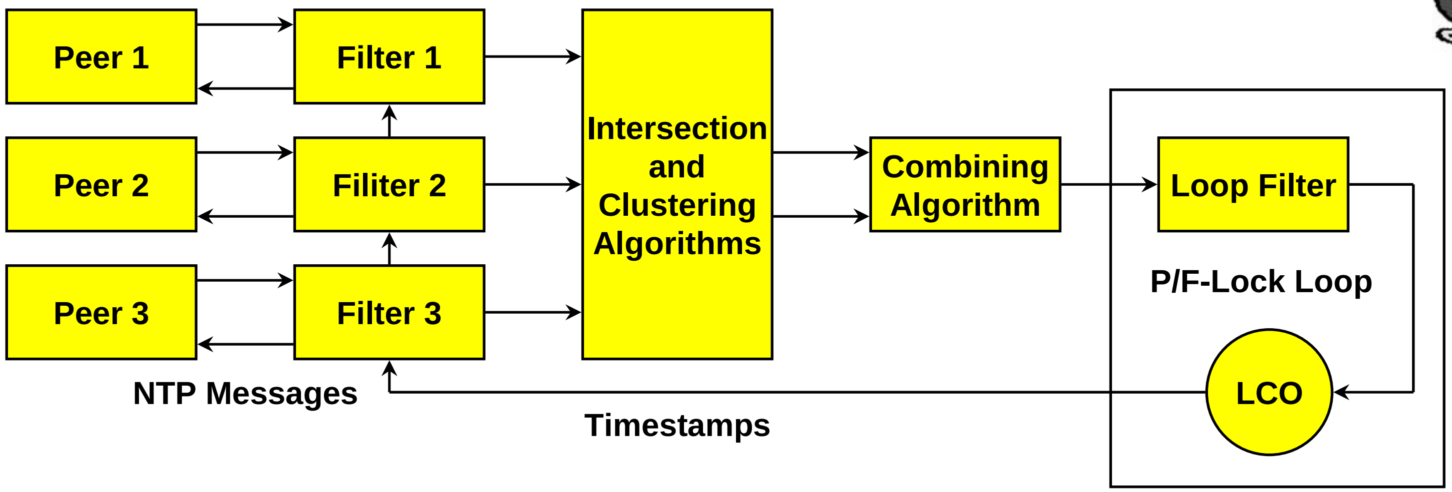


alautun, Maya glyph

Introduction

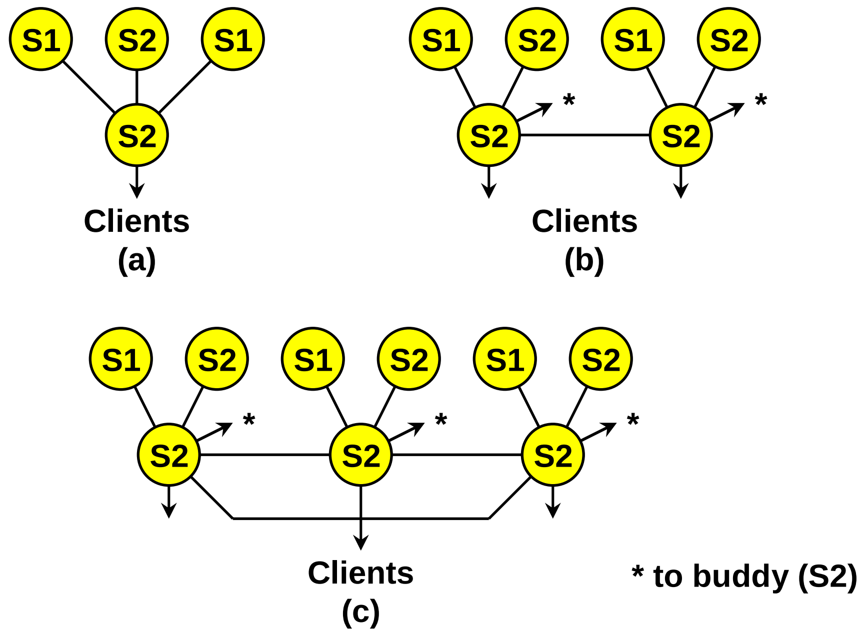
- Network Time Protocol (NTP) synchronizes clocks of hosts and routers in the Internet
- Provides submillisecond accuracy on LANs, low tens of milliseconds on WANs
- Primary (stratum 1) servers synchronize to UTC via radio, satellite and modem; secondary (stratum 2, ...) servers and clients synchronize via hierarchical subnet
- Reliability assured by redundant servers and diverse network paths
- Engineered algorithms used to reduce jitter, mitigate multiple sources and avoid improperly operating servers
- Unix NTP daemon ported to almost every workstation and server platform available today - from PCs to Crays
- Well over 100,000 NTP peers deployed in the Internet and its tributaries all over the world

NTP architecture



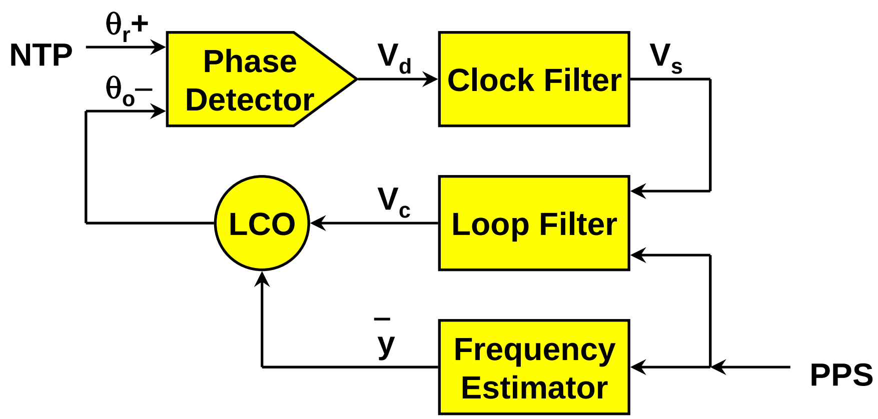
- Multiple synchronization peers for redundancy and diversity
- Data filters select best from a window of eight clock offset samples
- Intersection and clustering algorithms pick best subset of peers and discard outlyers
- Combining algorithm computes weighted average of offsets for best accuracy
- Loop filter and local clock oscillator (LCO) implement hybrid phase/frequency-lock feedback loop to minimize jitter and wander

NTP configurations



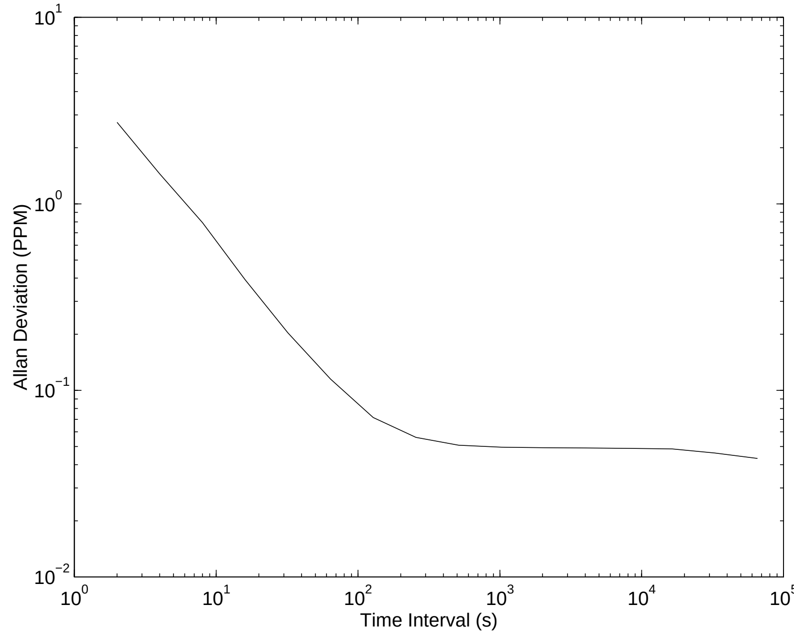
- (a) Workstations use multicast mode with multiple department servers
- (b) Department servers use client/server modes with multiple campus servers and symmetric modes with each other
- (c) Campus servers use client/server modes with up to six different external primary servers and symmetric modes with each other and external secondary (buddy) servers

Improved NTP local clock model



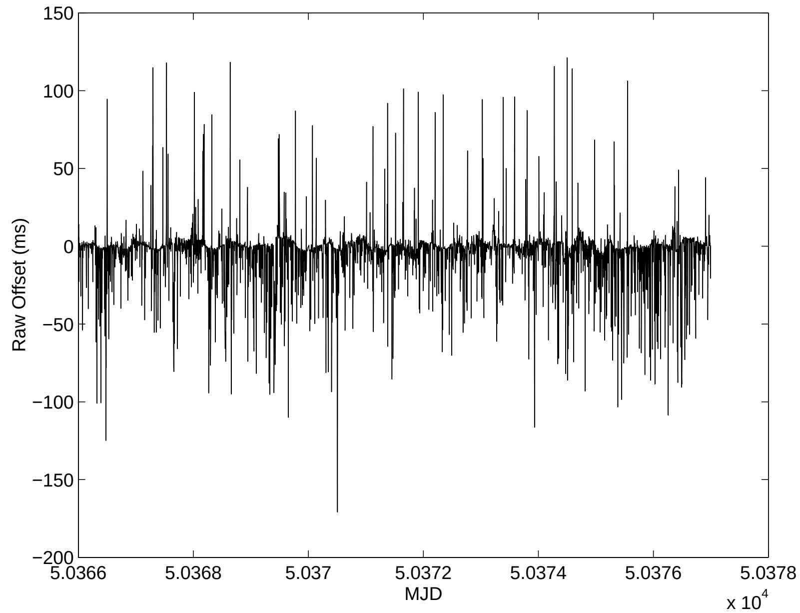
- Type II, adaptive-parameter, hybrid phase/frequency-lock loop estimates LCO phase and frequency
- Phase signal $V_d = \theta_r - \theta_o$ between NTP and local clock oscillator (LCO) computed from timestamps, then filtered to produce control signal V_c
- Auxilliary frequeny-lock loop disciplines LCO frequency *y* to pulse-per-second (PPS) signal, when available
- Loop parameters automatically optimized for update intervals from 16 s to 16,384 s

Optimizing the local clock parameters



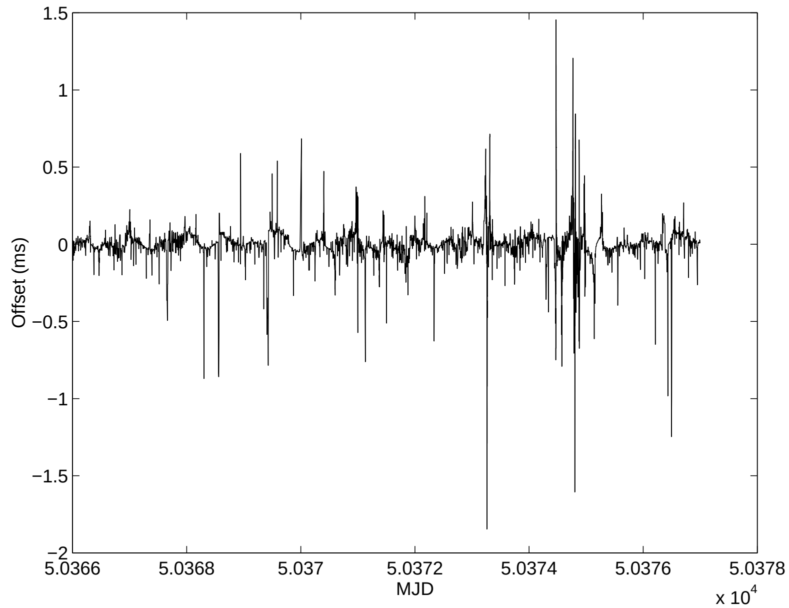
- Allan deviation shows computer clock stability calculated from free-running oscillator offsets measured over 11 days
  - Characteristic to the left shows white phase variations
  - Characteristic to the right shows flicker frequency variations
- Inflection represents best update interval - about 1000 s
  - Phase-lock mode optimum below 1000 s
  - Frequency-lock mode optimum above 1000 s

Raw Data Offsets



- Data for path between U Delaware NTP server and USNO NTP server navobs.wustl.edu in St. Louis
- Both servers synchronized to GPS receivers
- Delay budget includes 38.1 ms propagation delay plus 26.3 ms queueing delay
- Errors include .05 ms nonreciprocal delay + 4.8 ms mean offset + 19.5 ms RMS

Processed Data Offsets

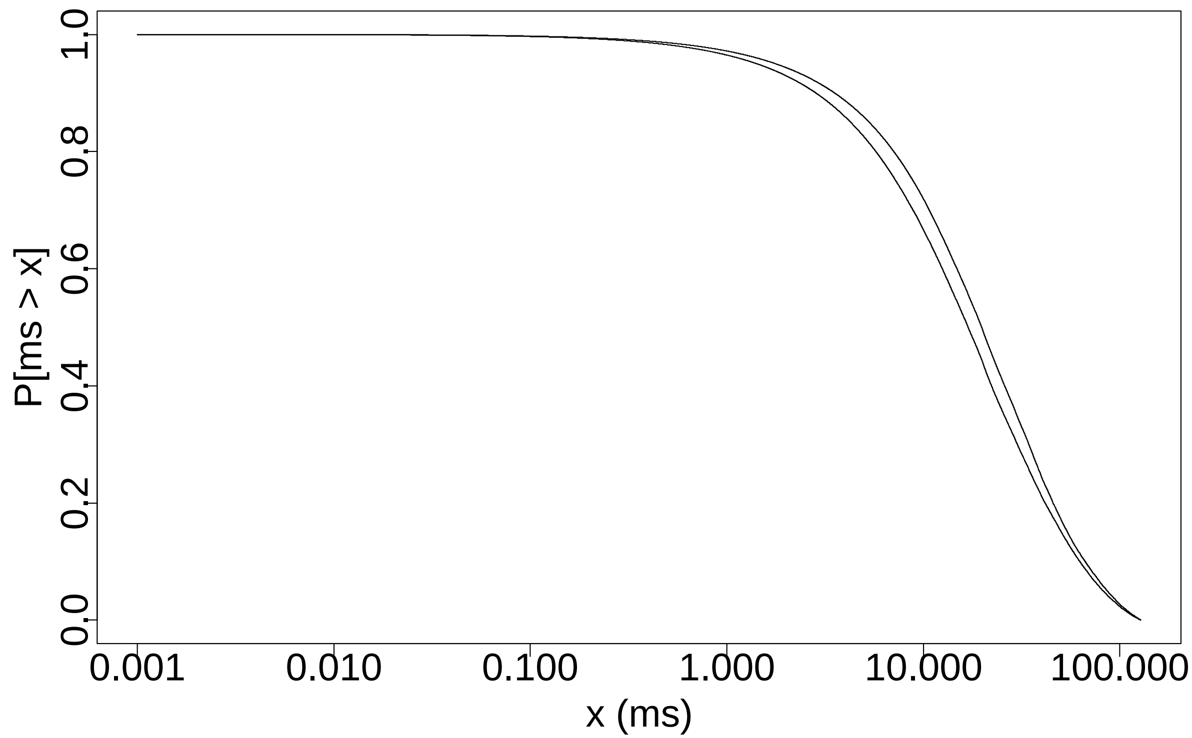


- Data for path between U Delaware NTP server and USNO NTP server navobs.wustl.edu in St. Louis
- Raw data processed by clock filter and PLL/FLL
- Errors include .05 ms nonreciprocal delay +.003 ms mean offset + 0.15 RMS
- Note spikes due various causes - should be eliminated with PPS discipline

Reference clock sources

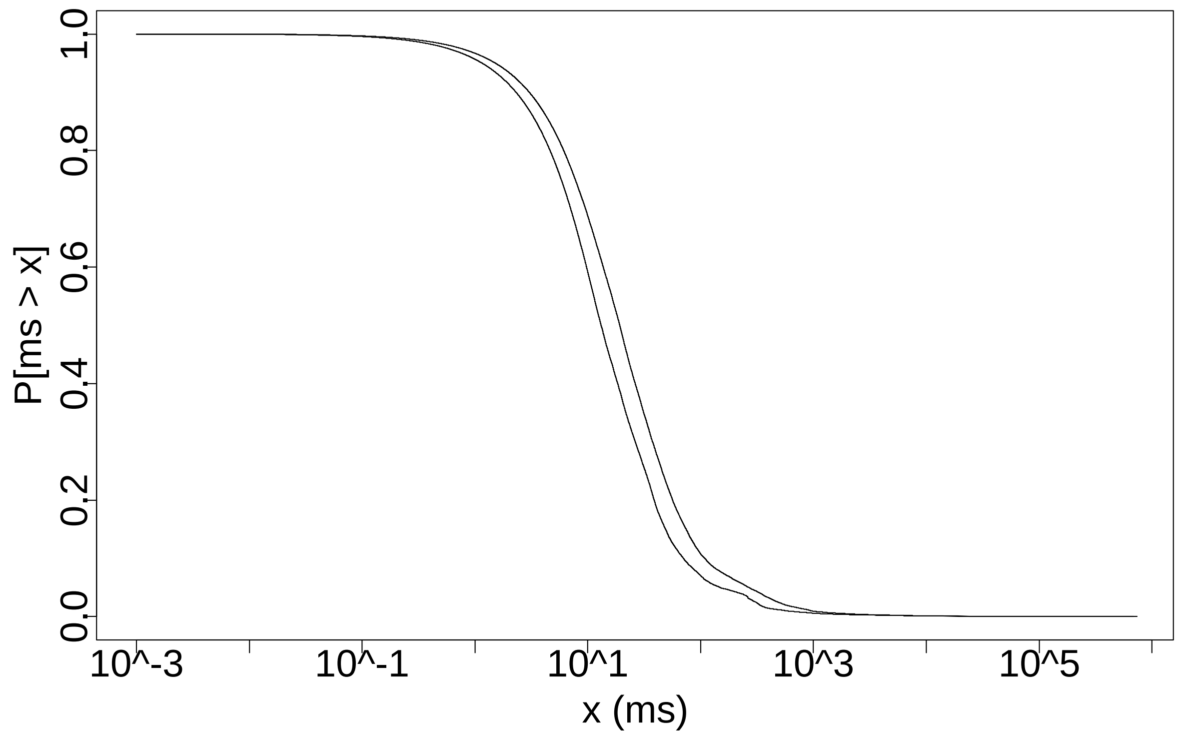
- In a survey of 36,479 peers, found 1,733 primary and backup external reference sources
- 231 radio/satellite/modem primary sources
  - 47 GPS satellite (worldwide), GOES satellite (western hemisphere)
  - 57 WWVB radio (US)
  - 17 WWV radio (US)
  - 63 DCF77 radio (Europe)
  - 6 MSF radio (UK)
  - 5 CHU radio (Canada)
  - 7 modem time service (NIST and USNO (US), PTB (Germany), NPL (UK))
  - 25 other (precision PPS sources, etc.)
- 1,502 local clock backup sources (used only if all other sources fail)
- For some reason or other, 88 of the 1,733 sources appeared down at the time of the survey

Peer clock offsets - filtered/unfiltered data



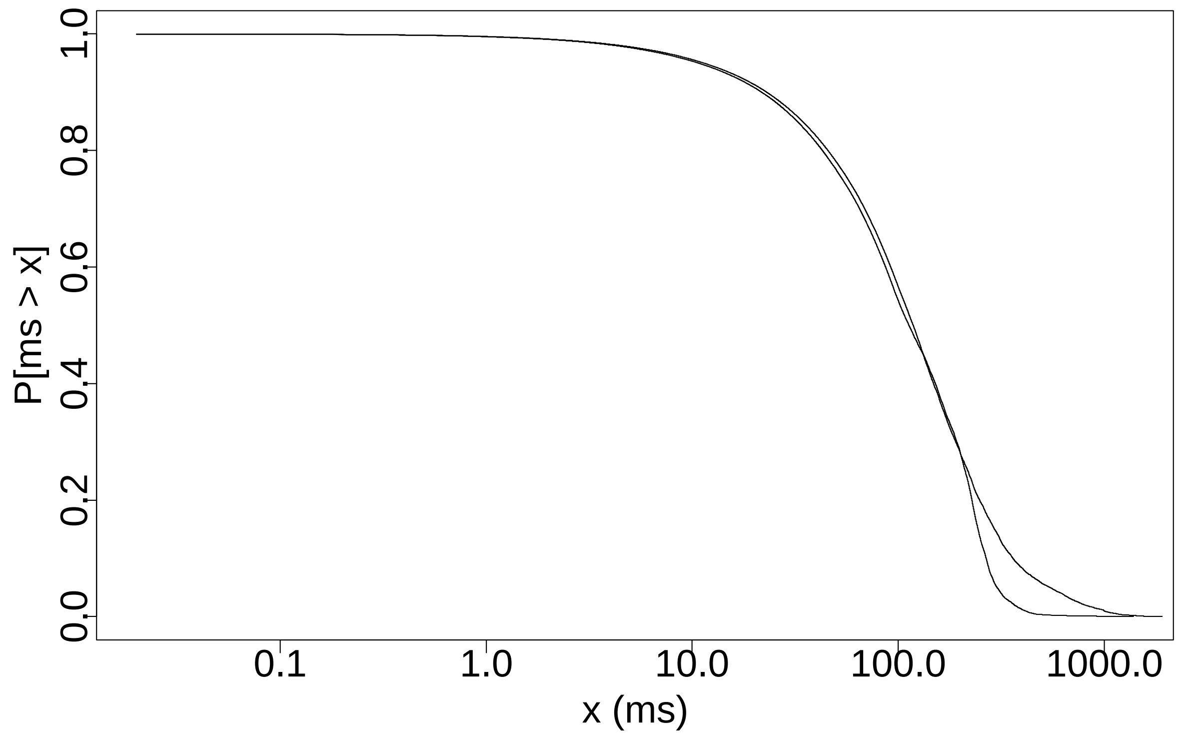
- Cumulative distribution function of peer-peer absolute clock offsets
  - 182,538 peers used by 34,679 clients; 94,489 peers survived intersection and clustering algorithms.
  - Upper curve: unfiltered (median 23 ms, mean 231 ms, max 686 s)
  - Lower curve: filtered (median 19 ms, mean 148 ms, max 686 s)

Peer clock offsets-same/different subnets



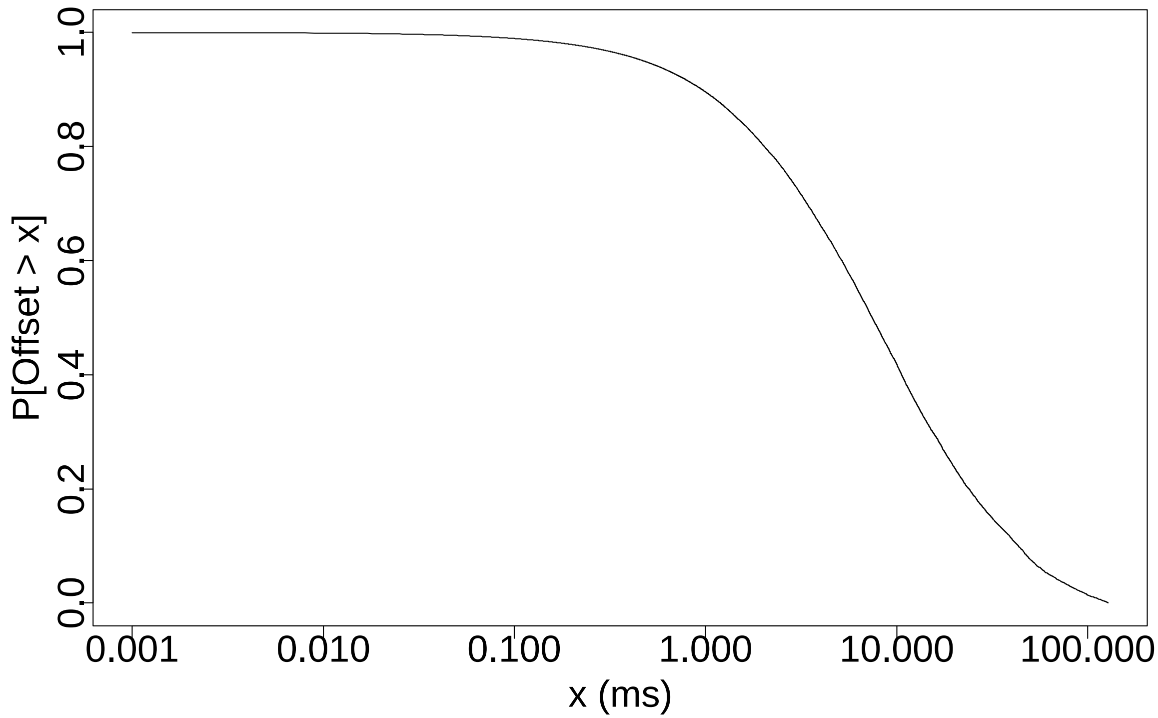
- Cumulative distribution function of peer-peer absolute clock offsets
  - 182,538 peers used by 34,679 clients, 85,730 on the same subnet, 96,808 on a different subnet.
  - Upper curve: different subnet (median 19 ms, mean 161 ms, max 621 s)
  - Lower curve: same subnet (median 13 ms, mean 188 ms, max 686 s)

Peer roundtrip delays



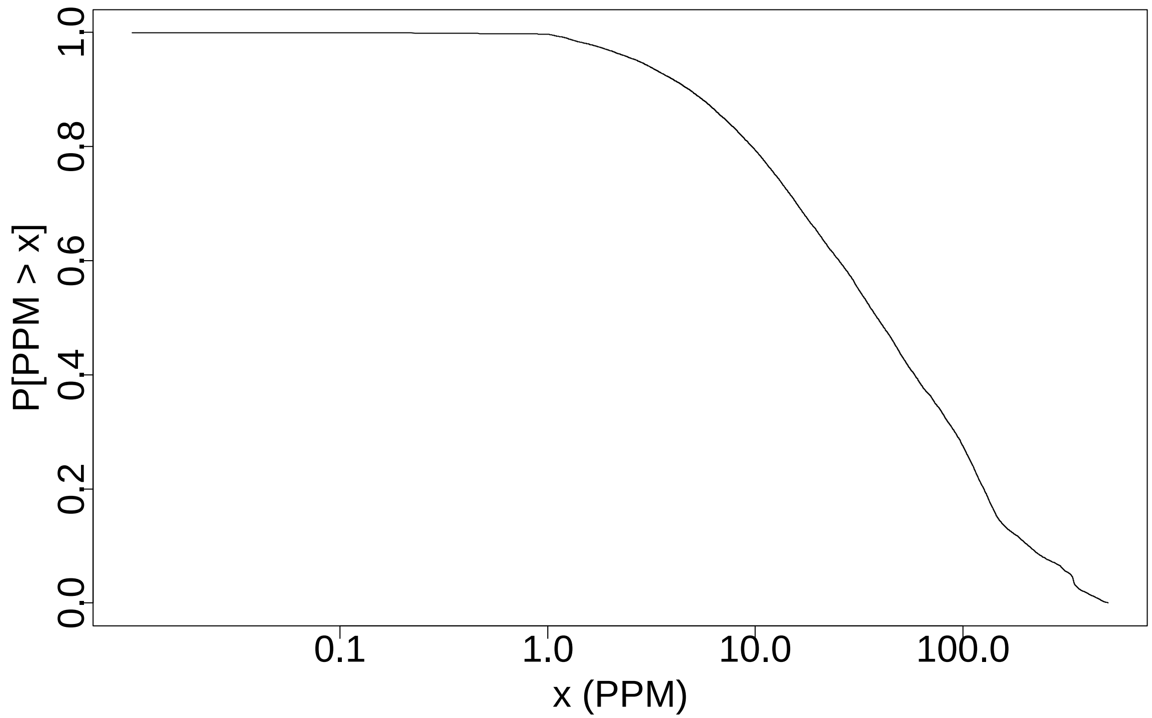
- Cumulative distribution function of peer-peer absolute roundtrip delays
  - 182,538 samples excludes measurements where synchronization distance exceeds 1 s. since by specification these cannot synchronize the local clock
  - Upper curve: different subnets (median 118 ms, mean 173 ms, max 1.91 s)
  - Lower curve: same subnet (median 113 ms, mean 137 ms, max 1.40 s)

Local clock phase offsets



- Histogram of local clock absolute phase offsets
  - 19,873 Internet peers surveyed running NTP Version 2 and 3
  - 530 offsets equal to zero deleted as probably unsynchronized
  - 664 offsets greater than 128 ms deleted as probably unsynchronized
  - Remaining 18,679 offsets: median 7.45 ms, mean 15.87 ms

Local clock frequency offsets



- Histogram of local clock absolute frequency offsets
  - 19,873 Internet peers surveyed running NTP Version 2 and 3
  - 396 offsets equal to zero deleted as probably spurious (self synchronized)
  - 593 offsets greater than 500 PPM deleted as probably unsynchronized
  - Remaining 18,884 offsets: median 38.6 PPM, mean 78.1 PPM

Performance of typical NTP servers in the global Internet

NTP Server	Location	Days	Mean	RMS	Max	>1	>5	>10	>50
Austron GPS	DCnet	91	0.0	0.012	1.000	0	0	0	0
rackety	DCnet	95	0.066	0.053	2.054	11	0	0	0
mizbeaver	DCnet	17	0.150	0.171	1.141	2	0	0	0
churchy	DCnet	42	0.185	0.227	3.150	15	0	0	0
pogo	DCNet	88	0.091	0.057	1.588	8	0	0	0
beauregard	DCnet	187	0.016	0.108	2.688	30	0	0	0
umd1	U Maryland	78	4.266	2.669	35.89	29	29	28	0
swifty	Australia	84	2.364	56.70	3944	27	27	27	13
ntps1	Germany	70	0.810	10.86	490.9	12	12	12	6
time_a	NIST Boulder	85	1.511	1.686	80.56	28	19	11	2
fuzz	San Diego	77	3.889	2.632	47.59	27	27	23	0
la	Los Angeles	83	0.650	0.771	17.84	28	8	3	0
enss136	NSFnet WashDC	88	0.657	1.203	32.65	38	23	10	0

- Table shows number days surveyed, mean absolute offset, RMS and maximum absolute error and number of days on which the maximum error exceeded 1, 5, 10 and 50 ms at least once
- Servers represent LANs, domestic WANs and worldwide Internet
- Results show all causes, including software upgrades and reboots

A day in the life of a busy NTP server

- NTP primary (stratum 1) server rackety is a Sun IPC running SunOS 4.1.3 and supporting 734 clients scattered all over the world
- This machine supports NFS, NTP, RIP, IGMP and a mess of printers, radio clocks and an 8-port serial multiplexor
- The mean input packat rate is 6.4 packets/second, which corresponds to a mean poll interval of 157 seconds for each client
- Each input packet generates an average of 0.64 output packets and requires a total of 2.4 ms of CPU time for the input/output transaction
- In total, the NTP service requires 1.54% of the available CPU time and generates 10.5, 608-bit packets per second, or 0.41% of a T1 line
- The conclusion drawn is that even a slow machine can support substantial numbers of clients with no significant degradation on other network services

The Sun never sets on NTP

- NTP is argueably the longest running, continuously operating, ubiquitously available protocol in the Internet
- USNO and NIST, as well as equivalents in other countries, provide multiple NTP primary servers directly synchronized to national standard cesium clock ensembles and GPS
- Over 230 Internet primary servers in Australia, Canada, Chile, France, Germany, Isreal, Italy, Holland, Japan, Norway, Sweden, Switzerland, UK, and US
- Over 100,000 Internet secondary servers and clients all over the world
- National and regional service providers BBN, MCI, Sprint, Alternet, etc.
- Agencies and organizations: US Weather Service, US Treasury Service, IRS, PBS, Merrill Lynch, Citicorp, GTE, Sun, DEC, HP, etc.
- Several private networks are reported to have over 10,000 NTP servers and clients; one (GTE) reports in the order of 30,000 NTP-equipped workstations and PCs

NTP online resources

- Internet (Draft) Standard RFC-1305 Version 3
  – Simple NTP (SNTP) RFC-1769
  – Designated SAFEnet standard (Navy)
  – Under consideration in ANSI, ITU, POSIX
- NTP web page http://www.eecis.udel.edu/~ntp
  – NTP Version 3 release notes and HTML documentation
  – List of public NTP time servers (primary and secondary)
  – NTP newsgroup and FAQ compendium
  – Tutorials, hints and bibliography
- NTP Version 3 implementation and documentation for Unix, VMS and Windows
  – Ported to over two dozen architectures and operating systems
  – Utility programs for remote monitoring, control and performance evaluation