Network Time Protocol (NTP): Overview and Configuration

Stanislav Shalunov ⟨shalunov@internet2.edu⟩

Performance Workshop, Atlanta, 2005-04-22

DRAFT

Talk overview

- NTP Overview
  - Purpose and utility
  - Strata, the loosely coupled network
  - Limitations, need for symmetry
  - Influence of temperature
  - Network use
  - Use of cryptography
  - Implementations of NTP
- Installing ntpd
  - OS selection
  - Obtaining, compiling, and installing

Talk overview, continued

- Configuring ntpd
  - Configuration file location and format
  - The (too) simple way: {broad,multi}castclient
  - Selecting the stratum; considerations for stratum 1
  - Number of servers, their strata, location, and options
  - Access control
- Monitoring ntpd
  - ntpd -p (through the network)
  - ntptime (local status check)
  - statistics (local log files)

NTP Overview

- Purpose of NTP
- Utility of NTP
- Packet Exchange of NTP
- NTP: Strata
- NTP: the network
- NTP Limitations
- Topological symmetry
- Time and Temperature
- Network Use by NTP
- Cryptographic Authentication
- NTP: Implementations

Purpose of NTP

- Synchronize clocks of a server constellation to UTC
- Does not work very well for synchronizing to *some* time
- Must be the correct time
- Attempt to keep time strictly monotonically increasing
- Minimize offset (difference from the true time)
- Minimize skew (difference of time change rate from the true rate)
- Sometimes these goals contradict each other
- Compromise sensibly

Utility of NTP

- Timestamps (e.g., in log files) make sense
- Timestamps can be compared to those made on other machines
- Tools that depend on time work in distributed environments
  - `make` on an NFS-mounted filesystem
  - Kerberos
  - SecurID tokens
- Time stepping (which breaks all of the above) is avoided
- **One-way delay can be measured**

Packet Exchange of NTP

- UDP port 123
- Uses its own binary format (RFC 1305, RFC 2030)
- No public online version of the NTPv4 spec at this time
- Packets sent

  $\longrightarrow$ Client sends request with a timestamp

  $\longleftarrow$ Server returns a packet with three timestamps

  1. Echo of the client timestamp
  2. When did I receive the request?
  3. When did I send the response?
- Client can estimate the offset between it and the server
- Client can have several servers
- Client chooses **one** server to synchronize to
- Minimizes the offset and skew with a feedback loop

NTP: Strata

- Each NTP node has a *stratum*
- Stratum is an integer between 0 and 16, inclusively
- Stratum 0 means a physical clock, **never a computer**
  - Cesium or rubidium oscillator
  - GPS or CDMA receiver
- Stratum 16 is reserved for devices that are not synchronized
- The stratum of any NTP-synchronized device is the stratum of the device it is synchronized to plus 1, e.g.:
  - GPS receiver: stratum 0
  - Computer connected to it by a serial line: stratum 1
  - Client that gets the time from that computer: stratum 2

NTP: the network

- NTP servers form a loosely coupled network
- Each node decides which server to use for synchronization based on complex selection algorithm (voting-like)
- Sanity checks
- Error estimates
- Resiliency to clock failure
- Resiliency to network failure
- Enforces the global nature of NTP

NTP Limitations

- Needs external servers to work well, even with a local clock
- Can produce systematic errors with asymmetric paths
- Can have problems with asymmetric congestion

Topological symmetry

- Assume asymmetric paths between client and server
  - Packet needs t_1 to get from client to server
  - t_2 from server to client
- The offset that the client will see is the true offset plus $\frac{t_1 - t_2}{2}$
- No way around this: the extra offset shows up as an error in the server clock
- If there are multiple servers, not all is lost
- But the paths must be diverse, or all will have the same error

Time and Temperature

- Oscillator frequency depends on temperature
- Typical correspondence: 1PPM (part per million) of clock rate for 1°C
- NTP can resolve rate differences of .001PPM
- For comparison: the temperature inside a modern computer will vary by 10°C depending on CPU load
- NTP could notice human movement around the host (!), and certainly open windows or A/C failures
- NTP will compensate, but best servers sit in constant-temperature machine rooms

Network Use by NTP

- A pair of small packets every 64–1024 seconds for each server
- 8 times more if `burst` keyword is used
- For vast majority of sites, nothing to worry about
- Less than 0.01% of Abilene traffic in bytes (circa 5 GB/day for the entire R&E world)
- For comparison: DNS is circa 0.15% of Abilene traffic in bytes

Cryptographic Authentication

- Not present in NTPv3
- Relatively new in NTPv4
- Not needed for the vast majority of cases
- No cryptography will help against an attacker who delays packets
- Only makes sense for time distribution within organization
- Only the shared keys part is likely to work
- Does your ultimate time source (e.g., GPS receiver) get its time through a cryptographically authenticated channel?
- Current best advice: do not bother

NTP: Implementations

- xntpd (shipped with some ancient operating systems): NTPv3; University of Delaware
- ntpd: the current state of the art; University of Delaware
- No complete NTP implementations for Windows (SNTP only, numerous)
- If you were interested enough to listen this far, you want ntpd

Installing ntpd

- OS selection for ntpd
- Obtaining ntpd
- Compiling/installing ntpd

OS selection for ntpd

- David Mills is the author of the NTP standards, of xntpd, and ntpd
- He seems to prefer FreeBSD

OS selection for ntpd, continued

"I have confronted Linux many, many times with broken this and that and concluded I will not attempt to examine, fix or judge anything Linux. You are on your own." —David L. Mills, 2004-10-20, <cl5odl$c6b$1@dewey.udel.edu>

"There have been reports that Linux is unfriendly at clock frequencies other than 100 Hz." —David L. Mills, 2004-05-07, <409BD15F.6D6AC8BD@udel.edu>

"I suspect the Linux team has not noticed that the adjtime() and ntp_adjtime() syscalls needs to recalibrate the tick adjustment value so to achieve something like 500 PPM as with 100-Hz clocks. [...] Unless Linux surmounts that challenge, NTP is a lose." —David L. Mills, 2004-03-23, <40607A18.51C1B2E2@udel.edu>

OS selection for ntpd, continued

"I say this in the urgent agenda to avoid Linux recursive complexity and keep FreeBSD clean and clear of messing fingers." —David L. Mills, 2004-02-16, <4030DEF4.BA19AE09@udel.edu>

"[Y]ou did say Linux. All of our tests are in other systems and they work as intended." —David L. Mills, 2003-05-31, <3ED80509.7675F0A2@udel.edu>

"When you said 'Linux' the red light came on." —David L. Mills, 2003-05-11, <3EBDCD7C.14A814CA@udel.edu>

"I am even more astonished if anybody considers Linux NTP anywhere near a conformant implementation of anything." —David L. Mills, 2003-04-10, <3E959842.473D0113@udel.edu>

OS selection for ntpd, continued

"The TAI [code] works perfectly in FreeBSD with ntpd and ntptime and Autokey. Linux experience may be different." —David L. Mills, 2003-03-24, <3E7F370D.EC4D8756@udel.edu>

Und so weiter, und so fort, again and again *ad infinitum*.

However:

"Most of you know my feelings about Linux, and I am on record as not allowing Linux anywhere near here, but the Linux system we don't have runs 2.6.4 kernel and now running latest ntpd 4.2.0. With virtually no customization (NFS and NIS and ssh) and vanilla ntp.conf it runs like a charm with or without the kernel [NTP support] enabled." —David L. Mills, 2004-04-01, <406C6164.57C8A7BB@udel.edu>

OS selection for ntpd, continued

- If convenient, use FreeBSD in favor of Linux
- If you want stratum 1, FreeBSD is strongly recommended (Linux needs non-standard kernel mods, PPSkit, to run well as stratum 1)
- Linux should work, but distributions can have surprises
- Other Unix systems might work (including Mac OS X)
- Don't even *think* about Windows

Hardware requirements for NTP

- A CPU and a network interface
- Memory and hard disk optional
- An old 486 will handle the load of a department
- Temperature stability (lack of advanced power management and HLT instruction) more important than the rest
- Anything, really
- (But need a time source and good interrupt handling ability for stratum 1)

Obtaining ntpd

- On FreeBSD, use the system binaries
- On other OSes, especially Linux
  - Uninstall the system ntpd
  - Grab the latest production tarball from ntp.org

Compiling/installing ntpd

- On FreeBSD, do nothing
- On other OSes
  - Ensure OpenSSL is installed, if you want NTP cryptography
  - tar xzf ntp-4.2.0.tar.gz
  - cd ntp-4.2.0
  - ./configure
  - gmake
  - su
  - gmake install
  - (Change 4.2.0 to the appropriate version number, of course.)

Enabling ntpd

- On FreeBSD, insert the line "`ntpd_enable="YES"`" into /etc/rc.conf
- On Linux, the official way depends on the distribution, but it's probably easiest to add the ntpd command to the end of `/etc/rc.local`

Configuring ntpd

- Configuration file location and format
- {broad,multi}castclient
- Selecting the stratum
- Stratum 1 servers
- Selecting the number of servers
- Server strata
- Server location
- Useful server options
- Access Control

Configuration file location and format

- `/etc/ntp.conf`
- One configuration statement per line
- Most important configuration statements
  - `driftfile` ⟨*filename*⟩
    - Important to include, not on by default
    - On FreeBSD, drift file is normally `/var/db/ntp.drift`
    - On Linux, varies, but there's `/etc/ntp/drift`
    - Historic location is `/etc/ntp.drift`
  - `server` ⟨*ntp.example.edu*⟩ [⟨*options*⟩]
- These two statements are enough to make a working server

{broad,multi}castclient

- A lazy way to configure
- `broadcastclient` will listen to broadcasts
- `multicastclient` ⟨*IP address*⟩ will listen to multicast on 224.0.1.1 (or another)
- Can be used for end-user machines
- **Do not use on machines that will run OWAMP**

Selecting the stratum

- Stratum 1 if you need maximum possible accuracy (single microseconds)
- Stratum 2 is suitable for machines that serve time to others or are used in measurements (accuracy can be better than 1 ms)
- Stratum 3 is a reasonable end-user stratum
- Most measurement nodes will probably use NTP stratum 2
  - Best accuracy bang for the buck
  - OK accuracy for most needs
  - Can be set up on most machines with no new hardware

Selecting the stratum, continued

- The best measurement nodes will use stratum 1
  - Need a stratum 0 time source
  - Oscillators too expensive for most
  - CDMA receivers: cheap, work anywhere a CDMA cell phone works
  - GPS receivers: best accuracy bang for the buck, but require an antenna that can see the sky
- Leave stratum 3 for your laptop and (maybe) your mail server

Stratum 1 servers

- Best accuracy
- Special hardware
- Extra work in configuration
- Should really run FreeBSD
- Use pulse per second (PPS) mode if you can at all

Selecting the number of servers

- Lazy people select one server (perhaps pool.ntp.org)
- Don't be lazy! Time can be OK with one server, but error estimate is worthless
- The minimum number of servers for reasonable error estimate is 3
- Anything in excess of 6 servers is likely to result in too many switches from server to server
- Prepare for losing connectivity to some servers (at least no single server failure...)
- Reasonable numbers:
  - 4 well-selected servers
  - 5 servers selected with, perhaps, not the same degree of care

Server strata

- Only servers of stratum $n+1$ (where n is my stratum) matter
- Important to select servers of the same strata in all cases
- Use all stratum 1 peers for stratum 1 configurations (unless you can have multiple physical time sources)
- Use all stratum 1 servers for stratum 2 configurations
- Use all stratum 2 servers for stratum 3 configurations
  - But don't do stratum 3 configurations for measurement
- **Do not mix strata**

Server location

- Best server is your server: get a GPS
- Second-best server is your provider's server: demand it
  - GigaPoPs could use the closer of ntp-e.abilene.ucaid.edu and ntp-w.abilene.ucaid.edu
- Third-best server is your neighbor's server: ask around
- If you still haven't found 4 or 5 servers, there's the public server list at http://ntp.isc.org/bin/view/Servers/StratumOneTime
- Hint: servers operated by NIST are well-administered and tend not to go away (you pay for them with your tax dollars, might as well use them)

Useful server options

- The one I find most useful: `iburst`
  - Send more packets in the beginning
  - Synchronize faster
  - Why isn't this the default?
- `burst`: more packets, better accuracy
- `tinker huffpuff 7200`: can be a big help in case of asymmetric congestion (but not asymmetric topology, of course)

Access control

- `restrict` keyword
- Takes IP address ranges
- Some people like to restrict access so that only the servers can talk to them
- Usually an overkill: theft of NTP service is uncommon
- PITA: servers change IP addresses, you need to change the config file

Monitoring ntpd

- ntpq -p
- ntptime
- statistics {loop,peer,clock,raw}stats

ntpq -p

- Outputs a two-line header and then one line per server, with 11 fields
- 1st char: '*' for the host we sync to; '+' for OK; '−' and 'x' for bad; ' ' for unreachable; 'o' for PPS sync source
- hostname
- Where does it get time from?
- Its stratum
- Type of peer ('u' is for 'unicast')
- Seconds ago we heard from it
- How often do we currently poll, in seconds
- Octal reachability mask, 377 means A-OK
- Delay (round-trip) to the server, in milliseconds
- Offset in milliseconds
- Jitter in milliseconds

`ntpq -p` example output

```
$ ntpq -p mail.internet2.edu
     remote           refid      st t when poll reach   delay   offset  jitter
==============================================================================
-ntp0.usno.navy. .USNO.           1 u  651 1024  333   78.347   18.023   0.124
+ntp-e.abilene.u nms4-nycm.abile  2 u  238 1024  377   34.944    0.734   0.143
+ntp-w.abilene.u nms4-snva.abile  2 u   77 1024  377   55.667    0.402   0.137
 207.75.164.255  0.0.0.0         16 u    -   64    0    0.000    0.000 4000.00
*gps1.tns.its.ps .GPS.           1 u  152 1024  377   52.301    0.445   0.040
```

Strata mixed, unreachable server, poorly reachable server...

ntptime

- The simplest way to look at NTP state
- Is it synchronized?
- Maximum error (not very useful)
- Estimated error
- Offset from target
- Frequency skew in PPM

statistics {loop,peer,clock,raw}stats

- A way to really understand what is going on
- This is where open window detection comes in
- Outside the scope of this presentation

Resources

- man ntpd
- man ntp.conf
- /usr/share/doc/ntp/ (complete HTML documentation)
- e2epi.internet2.edu/owamp/details.html#NTP
- twiki.ntp.org/bin/view/Support/SelectingOffsiteNTPServers
- Newsgroup: comp.protocols.time.ntp (free tier 4 support: David Mills might answer your question here—unless you're using Linux, of course)

Summary of minimum working configuration

- Decide on stratum 2
- Install latest FreeBSD release on an old box
- `ntpd_enable="YES"` in `/etc/rc.conf`
- Select four stratum 1 servers
- Edit `/etc/ntp.conf` so that it contains one `driftfile` line and four `server` lines
- Type "`ntpd`" as root
- Wait a day or so for the clock to settle
- Try to keep the temperature constant

Questions?