



C120-E202-03ENZ2

NTP Operator's Guide

Second Edition



FUJITSU



Introduction

Information systems in recent years have undergone dramatic changes in their configuration from operation centered on the mainframe computer to the distributed system implemented using UNIX servers.

In the distributed system that uses UNIX servers, the clocks must be synchronized among different servers for a number of reasons including individual application server management.

This manual explains the Network Time Protocol (NTP), which is a standard method of synchronizing UNIX servers. This manual also explains how to design and operate the NTP in the Solaris™ Operating System (Solaris OS) to assist the user in designing a distributed system with PRIMEPOWER.

This manual is organized as follows:

1. NTP
Briefly explains NTP.
2. NTP Settings
Describes in detail NTP server and client settings.
3. Application to Actual Operations
Describes issues to be dealt with before NTP is used in actual operations. These issues include time change methods and specifications for NTP in the Solaris(TM) operating system.
4. NTP Status Information
Describes the status information recorded with NTP.

■ Trademarks

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company Limited.

Sun, Sun Microsystems, the Sun logos, Solaris, and all Solaris based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc., in the United States and in certain other countries, and are used under license.

All SPARC trademarks are trademarks or registered trademarks of SPARC International, Inc., in the United States or certain other countries. Products with a SPARC trademark are based on an architecture developed by Sun Microsystems, Inc.

All other company names and product names mentioned herein are the trademarks or registered trademarks of their respective owners.

Request

- The contents of this manual shall not be reproduced in any way or form without the permission of Fujitsu Limited.

All Rights Reserved, Copyright © FUJITSU 2005

Revision History

| Edition | Data | Details |
|---------|------------|---|
| 1 | 2003/10/01 | - |
| 2 | 2005/06/10 | Revision of chapter organization Technical brushup Addition of description for Solaris 10 support |
| 2.1 | 2005/10/06 | Addition of description for SMF support |

| | | |
|-----------|--|-----------|
| 1. | ABOUT NTP | 5 |
| 1.1 | Overview..... | 5 |
| 1.2 | NTP Design Principles | 5 |
| 1.3 | NTP "Stratum" | 6 |
| 1.3.1 | When a server with an accurate reference clock is used | 7 |
| 1.3.2 | When neither a server with an accurate reference clock nor a time-correcting server is used (not recommended)..... | 8 |
| 2. | NTP SETTINGS | 9 |
| 2.1 | Before Configuring Settings..... | 9 |
| 2.1.1 | Checking required patches..... | 9 |
| 2.1.2 | Checking the model | 9 |
| 2.2 | NTP Settings for a Server without an SMC..... | 10 |
| 2.2.1 | Referencing an external NTP server | 10 |
| 2.2.2 | Settings for an NTP server connected to a clockwork..... | 11 |
| 2.3 | NTP Settings for a Server with an SMC..... | 12 |
| 2.3.1 | Design principles | 12 |
| 2.3.2 | SMC settings (when no external NTP server is used) | 12 |
| 2.3.3 | SMC settings (when an external NTP server is used) | 13 |
| 2.3.4 | Partition settings..... | 14 |
| 2.3.5 | Notes on SMC operation..... | 15 |
| 2.3.6 | Notes on installation of a spare SMC | 15 |
| 2.4 | When Configuration of Settings Is Completed..... | 15 |
| 3. | APPLICATION TO ACTUAL OPERATIONS..... | 17 |
| 3.1 | Specifications about Time Changes | 17 |
| 3.1.1 | Specifications for Solaris NTP daemon 3.x..... | 17 |
| 3.1.2 | Interval for time polling with a high-order NTP server | 18 |
| 3.1.3 | Specifications for clock synchronization in Solaris | 18 |
| 3.2 | Reference Server and NTP Clock Synchronization..... | 19 |
| 3.3 | Gradual Clock Synchronization with NTP..... | 19 |
| 3.4 | Time Change Methods in an NTP Operating System..... | 21 |
| 3.4.1 | Without an SMC | 21 |
| 3.4.2 | When an SMC is used..... | 22 |
| 4. | NTP STATUS INFORMATION | 23 |
| 4.1 | ntpq Command | 23 |
| 4.2 | ntptrace Command..... | 24 |
| 4.3 | peerstats File..... | 25 |
| 4.4 | Checking NTP Operation with snoop Data..... | 26 |

1. About NTP

This chapter explains the Network Time Protocol (NTP).

1.1 Overview

NTP is a network protocol used to synchronize a client machine or server machine with another server machine or clock (hardware such as radios, satellite receivers, and modems). For example, the clock of a server that uses the Global Positioning System (GPS), which is synchronized to Coordinated Universal Time (UTC), can be set via a LAN or WAN with a precision on the order of milliseconds or several tens of milliseconds, respectively.

NTP specifications are released by RFC 1305 (NTP Version 3) and others, and Solaris incorporates an NTP implementation in addition to freeware as standard equipment. Though the latest version is Version 4, the NTP implementations for Solaris 2.6 to Solaris 10 are based on Version 3. The NTP implementations for Solaris have been developed based on freeware, but they are only one type of implementation that is not necessarily identical to NTP implementations for other operating systems.

1.2 NTP Design Principles

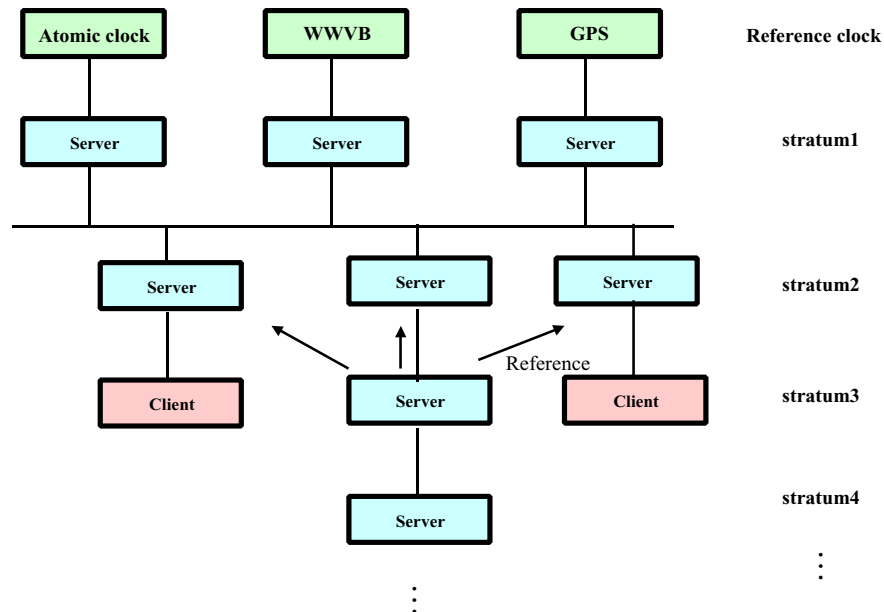
- NTP is a framework used to synchronize the system clocks of the router and host system with a reference clock, which keeps accurate time, at startup and then maintain the time within a specified range. That is, if the NTP server time is changed, the NTP client time is not synchronized.
- NTP assumes use of the GPS or a reference clock (described below) such as a cesium atomic clock with an accuracy on the order of microseconds. In a LAN environment, NTP can maintain the time within a range of several milliseconds. In operation with a WAN or another network that has a longer latency time, NTP can maintain the time within a range of several tens of milliseconds.
- The reference clock is assumed to keep accurate time and basically remain unchanged. (Leap seconds have already been taken into account in the protocol specifications.) Any change of the reference clock from outside is not taken into consideration in the protocol specifications.
- Worldwide, the times in more than 100,000 computer systems under an accurate reference clock (stratum 1) are already being maintained by NTP. In the U.S., the computer systems of financial and securities companies also use an accurate reference clock.
- For NTP operation, reference to an accurate reference clock is assumed. It is desirable to limit references to a local system clock to situations in which an accurate reference clock cannot be referenced because of a network error or hardware failure. In such cases, since the local system depends on the accuracy of the local clock (providing an accuracy within plus or minus 128 ms is recommended), the time is maintained with an accuracy no higher than that of the local clock.

1.3 NTP "Stratum"

An NTP network has a stratum structure. Regardless of whether a server is on the Internet or in an intranet, the lower the "stratum" number, the greater the proximity to an accurate clock (UTC). On the Internet, slightly fewer than 100 stratum 1 servers and slightly more than 100 stratum 2 servers are available (public time servers). Stratum 1 servers have reference clocks, but stratum 2 servers have no reference clock and thus reference the stratum 1 servers for the time setting.

In operations where time information is fetched from these servers, additional strata are built such that, for example, file servers are in stratum 3 and other workstations in stratum 4. Stratum 3 servers reference three or more stratum 2 servers to distribute time information to numerous stratum 4 servers.

The term stratum 0 may sometimes be used in ntp.conf. This term refers to the clock itself and does not represent any server stratum.



Before NTP operation, it is necessary to design how the clocks including the reference clock are synchronized among servers. The following sections describe a case with a server that has an accurate reference clock and a case with neither a time-correcting server nor a server that has an accurate reference clock.

1.3.1 When a server with an accurate reference clock is used

Time synchronization is performed based on the reference server clock. In this case, no manual synchronization is required because the reference server itself is in some way or other (atomic clock or GPS) synchronized with world time. However, a correction of about one second may be necessary, such as for adjusting to leap seconds.

First, an accurate reference clock must be provided. The following devices can be used as a reference clock:

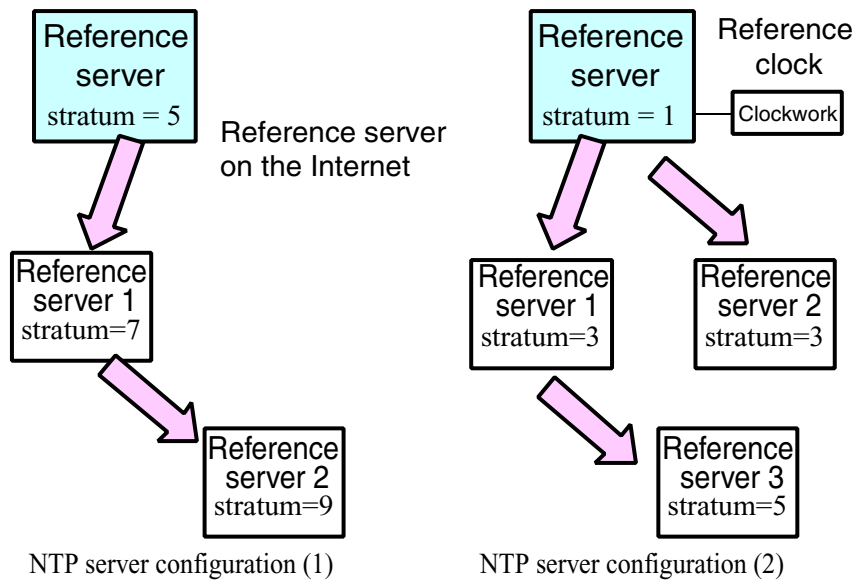
- NTP server using the GPS or a clockwork

One available clockwork (type name: PW007TC1) is a hardware clock shipped by Fujitsu. The clockwork (type name: PW007TC1) synchronizes with NTT time signals (117). Fujitsu recommends a configuration in which an NTP server connected to the clockwork is used for synchronization. Though not shipped by Fujitsu, other devices available on the market can function as an NTP server by synchronizing with time information from the GPS. They can be used as reference servers.

- NTP server on the Internet

This is practical, but accuracy and operation of the server are not necessarily guaranteed. Also, an Internet connection is required.

Configurations are shown below. In NTP server configuration (1), the reference server on the Internet is referenced, and new NTP strata are built for operations. In NTP server configuration (2), the reference server that is synchronized with the reference clock (clockwork (type name: PW007TC1)) is referenced, and new NTP strata are built for operations.



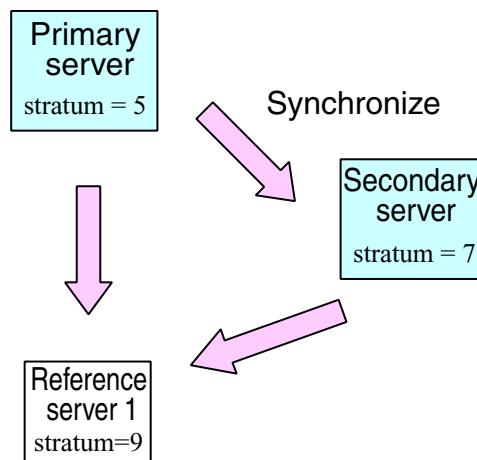
1.3.2 When neither a server with an accurate reference clock nor a time-correcting server is used (not recommended)

Time synchronization of servers is performed based on the clock of at least one other server. In this case, the time must be corrected manually every week or every month because the reference server itself loses synchronization with world time on a daily basis.

However, this correction may cause the following problems. First, if the reference server clock is set back 10 seconds because it was 10 seconds fast, the time in another server that references the reference server clock can be synchronized with difficulty only by drastic correction of the clock of that server. Second, abruptly correcting the clock may cause active applications to malfunction.

Due to the above problems, synchronization of the reference server clock with the precise time requires that all ntp servers be stopped, the reference server clock be corrected, and the clock of each related server then be synchronized. (For details on this procedure, see Chapter 3, "Application to Actual Operations.") Fujitsu does not recommend this method because it is very complicated.

The configuration shown below is an example (NTP server configuration (3)). Two NTP servers are required in this configuration. One is the primary server, and the other is the secondary server for operations. The secondary server and client (called the reference server 1 in this example) are synchronized with the clock of the primary server. If the primary server fails, the client is synchronized with the clock of the secondary server.



NTP server configuration (3)

2. NTP Settings

2.1 Before Configuring Settings

2.1.1 Checking required patches

If a time difference equal to or greater than ± 128 ms from an old version of Solaris is detected, the times are forcibly synchronized to eliminate the difference. However, this may change the time drastically, which may cause active applications to malfunction, or set back the time, which results in a system hang.

To prevent these problems, Solaris provides the following patches including corrections that do not forcibly synchronize the time (it is instead synchronized gradually by the `adjtime` function). Apply the appropriate version (or a later one) of the patches listed below. Before applying the patch, be sure to read its README file.

To check its application, execute the `showrev -p` command. (For details, see Section 3.3, "Gradual Clock Synchronization with NTP.")

| | |
|-------------|---|
| Solaris 2.6 | 107298-03 |
| Solaris 7 | 109409-04 |
| Solaris 8 | 109667-07 |
| Solaris 9 | None (already supported by the FCS version) |
| Solaris 10 | None (already supported by the FCS version) |

Note: Even if the clock is to be forcibly synchronized as is done with earlier versions, be sure to apply the appropriate version (or a later one) of the above patches before NTP operation, because earlier versions contain an error.

When the patch has been applied, the following must be added to `ntp.conf` to gradually synchronize the clock:

```
slewalways yes
disable pll
```

- * If a time difference of ± 128 ms or more is detected when a patch has been applied with the correct slew setting, "time reset (slew)" is displayed in a console message (`/var/adm/messages`) as follows:
`xntpd[xxx]: [ID xxxxxx daemon.notice] time reset (slew) 33.483293 s`

2.1.2 Checking the model

Models in the PRIMEPOWER series and GP7000F series can be divided into two groups: one group uses a system management console (SMC (*1)) and the other does not. NTP settings vary depending on the model group. Before starting to configure NTP settings, check the model used.

The following sections each describe the setting configuration procedure for a server with an SMC or a server without an SMC.

- * An SMC is a system console connected via a LAN to the System Control Facility on the main unit to implement main unit status monitoring, log information collection, different types of setting operations, and REMCS center notification.
The PRIMEPOWER models 800, 1000, 2000, 900, 1500, 2500, and HPC2500, and the GPF7000F models 1000 and 2000 use an SMC. For details, see the explanation of system consoles as registered in the IKB.

2.2 NTP Settings for a Server without an SMC

2.2.1 Referencing an external NTP server

If a server is available whose suitability for time synchronization operation is guaranteed, specify that server as "server."

Ideally, three or more NTP servers should be specified. However, if one of the specified NTP servers does not provide the precise time, specifying multiple servers has an adverse effect, because it becomes unclear which server should be used for synchronization. Specifying a single time server is sufficient if that server is accurate and stable.

```
server IP address of external NTP server 1
server IP address of external NTP server 2
server IP address of external NTP server 3
server 127.127.1.0
fudge 127.127.1.0 stratum 9

slewalways yes } To gradually synchronize the time,
disable pll    } apply the appropriate patch, and then
                } add these statements.

enable auth monitor
driftfile /var/ntp/ntp.drift
statsdir /var/ntp/ntpstats/
filegen peerstats file peerstats type day enable
filegen loopstats file loopstats type day enable
filegen clockstats file clockstats type day enable
```

"127.127.1.0" indicates the local clock. As shown below, the same format is used as for the IP addresses of other servers in ntp.conf:

```
127.127.t.u
```

"t" denotes the type of clock. The latest xntpd implements 1 to 33 types of clock by default. Type 1 is a special type that represents Local Clock Driver, which is the kernel clock (local clock). The other types represent high-precision hardware clocks.

Fujitsu offers a clockwork product (type name: PW007TC1, synchronizes with NTT time signals (117)), which is a hardware clock that can be used as a reference clock under Solaris. For details on how to configure NTP settings for this clock, see Section 2.2.2, "Settings for an NTP server connected to the clockwork."

On the fudge line, the local clock is positioned at stratum 9. In this case, xntpd on the local server runs at the stratum level obtained by adding "+1" to the number of the lowest stratum containing external NTP servers. If the stratum number of the external NTP servers is higher than 9, synchronization with the local clock occurs because the local clock is positioned at stratum 9. xntpd thus runs at stratum 10.

On the enable line, the auth and monitor flags are set. The auth flag is required for use of the authentication function, and the monitor flag is required for use of the xntpd monitoring function. (For details, see xntpd(1M).)

The driftfile line specifies the file to which learning results about the precision of the local clock are recorded.

The statsdir line specifies the save destination of the statistical information file specified in filegen. The statistical information file should be specified because it is required for troubleshooting.

2.2.2 Settings for an NTP server connected to a clockwork

To use the clockwork (type name: PW007TC1), the time synchronization application supplied with the clockwork must be installed and activated. For details, see the manuals supplied with the hardware.

In this case, specify the following in `ntp.conf` on the NTP server:

```
server 127.127.1.0

fudge 127.127.1.0 stratum 1

enable auth monitor
driftfile /var/ntp/ntp.drift
statsdir /var/ntp/ntpstats/
filegen peerstats file peerstats type day enable
filegen loopstats file loopstats type day enable
filegen clockstats file clockstats type day enable
```

The local clock (127.127.1.0) runs at stratum 1.

2.3 NTP Settings for a Server with an SMC

2.3.1 Design principles

No partition in the PRIMEPOWER models 800, 1000, 2000, 900, 1500, 2500, and HPC2500 and the GP7000F models 1000 and 2000 has a hardware clock that can be used while the operating system is active. Therefore, NTP must be used to synchronize the time between each partition and the SMC.

Fujitsu recommends referencing an accurate reference server because the clock contained in the SMC is not very accurate. Normally, the SMC is selected as the NTP server, and each partition is set to client mode for operation.

As the SMC is dedicated hardware for use with these models, do not reference the SMC from another server (or client) machine in client mode.

For 24-hour operation, an accurate reference server must be constantly referenced.

2.3.2 SMC settings (when no external NTP server is used)

Configuring the SMC operation settings on the NTP server is very simple. Execute the `partition_setup` command with the `-n` option specified, and all settings will be made automatically. This command sets the following contents in the `/opt/FJSVscsl/etc/ntp.conf` file:

```
server 127.127.1.0 prefer
fudge 127.127.1.0 stratum 5
enable auth monitor
driftfile /var/ntp/ntp.drift
statsdir /var/ntp/ntpstats/
filegen peerstats file peerstats type day enable
filegen loopstats file loopstats type day enable
filegen clockstats file clockstats type day enable
```

The above server line specifies local clock of the SMC.

As the times in the SMC and partitions are checked when `xntpd` on the SMC starts, `xntpd` may be prevented from starting, depending on the check result. For details, see Section 2.3.5, "Notes on SMC operation."

Note)

- For Solaris 9 or before

The location of "ntp.conf" and "ntp" start script used on SMC are changed from Solaris standard.

- Solaris 10

Service Management Method is changed to SMF (Service Management Facility) on Solaris 10.

SMC uses its specific configuration file so that System Console Software installer makes the following changes.

If `/etc/inet/ntp.conf` exists, it renames the file name to `/etc/inet/ntp.conf.orig`.

Then it makes a symbolic link from `/etc/inet/ntp.conf` to `/opt/FJSVscsl/etc/ntp.conf`.

2.3.3 SMC settings (when an external NTP server is used)

Execute the `partition_setup` command to create the `/opt/FJSPVscsl/etc/ntp.conf` file as described in Section 2.3.2, "SMC settings (when no external NTP server is used)." To use external NTP servers, modify `ntp.conf` as follows:

```
server IP address of external NTP server 1
server IP address of external NTP server 2
server IP address of external NTP server 3
server 127.127.1.0
fudge 127.127.1.0 stratum 5

slewalways yes      } To gradually synchronize the time,
disable pll         } apply the appropriate patch, and then
                    } add these statements.

enable auth monitor
driftfile /var/ntp/ntp.drift
statsdir /var/ntp/ntpstats/
filegen peerstats file peerstats type day enable
filegen loopstats file loopstats type day enable
filegen clockstats file clockstats type day enable
```

The above server lines specify the external NTP servers. Ideally, three or more NTP servers should be specified, but in the worst case, none of the time servers provides the precise time. (In this case, it is unclear which of the time servers should be used for synchronization.) Specifying a single time server is sufficient if that server is accurate and stable.

`xntpd` on the SMC runs at the stratum level obtained by adding "+1" to the number of the lowest stratum containing external NTP servers. If the stratum number of the external NTP servers is higher than 5, synchronization with the local clock occurs because the local clock is positioned at stratum 5. (In this event, the stratum number specified by `fudge` has to be increased until it is higher than that of the external NTP servers.)

On the `enable` line, the `auth` and `monitor` flags are set. The `auth` flag is required for use of the authentication function, and the `monitor` flag is required for use of the `xntpd` monitoring function. (For details, see `xntpd(1M)`.)

The `driftfile` line specifies the file to which learning results about the precision of the local clock are recorded.

The `statsdir` line specifies the save destination of the statistical information file specified in `filegen`. The statistical information file should be specified because it is required for troubleshooting.

Please start the system additionally in the single user mode at the time of the NTP server at the time of SMC when there are time of the NTP server and a time difference when the system is started up. For the detailed procedure, see Section 3.4.2, "With an SMC."

Because the times in the SMC and partitions are checked when `xntpd` on the SMC starts, `xntpd` may be prevented from starting, depending on the check result. For details, see Section 2.3.5, "Notes on SMC operation."

Note)

- For Solaris 9 or before

The location of "`ntp.conf`" and "`ntp`" start script used on SMC are changed from Solaris standard, the `ntpdate` command is not issued on the SMC, and `xntpd` is started.

- Solaris 10

Service Management Method is changed to SMF (Service Management Facility) on Solaris 10.

SMC uses its specific configuration file so that System Console Software installer makes the following changes.

If "`/etc/inet/ntp.conf`" exists, it renames the file name to "`/etc/inet/ntp.conf.orig`".

Then it makes a symbolic link from "`/etc/inet/ntp.conf`" to "`/opt/FJSPVscsl/etc/ntp.conf`".

2.3.4 Partition settings

Configuring partition settings is also simple. Install Solaris in a partition by using the `partition_setup` command, and the `/etc/inet/ntp.conf` file is then automatically created. Make the following modifications as necessary:

```
server IP address of the SMC
server 127.127.1.0
fudge 127.127.1.0 stratum 9

slewalways yes } To gradually synchronize the time,
disable pll    } apply the appropriate patch, and then
                } add these statements.

enable auth monitor
driftfile /var/ntp/ntp.drift
statsdir /var/ntp/ntpstats/
filegen peerstats file peerstats type day enable
filegen loopstats file loopstats type day enable
filegen clockstats file clockstats type day enable
```

Be sure to define the SMC as the primary NTP server for the partition.

Each partition in the GP7000F and PRIMEPOWER models 800, 1000, 2000, 900, 1500, 2500, and HPC2500 references the SMC clock at the firmware level when activated. Therefore, if a partition is synchronized with an external NTP server other than the SMC, the time becomes unstable during activation (loss of synchronization between the firmware clock and NTP internal clock).

Because the times in the SMC and partition are checked when `xntpd` on the SMC starts, `xntpd` may be prevented from starting, depending on the check result. For details, see Section 2.3.5, "Notes on SMC operation."

2.3.5 Notes on SMC operation

At the start of NTP operation, the SMC checks the SMC time and the time of each partition through the daytime port of the partition, and it compares the times. To guarantee normal NTP operation, xntpd on the SMC is prevented from starting if:

- A partition time is faster than the SMC time by two seconds or more.
- The SMC time is faster than a partition time by 1,000 seconds or more.

If either of the above conditions is met, synchronize the SMC clock with the partition clock correctly. For the detailed procedure, see Section 3.4.2, "When an SMC is used."

If the daytime port of a partition is disabled, the SMC cannot check and compare the SMC time with the partition time, and normal NTP operation may be disabled. If a daytime port is disabled, check and compare the SMC time with the partition time to ensure that neither of the above xntpd start restriction conditions is met.

Note that the daytime port of a partition in which Solaris 10 is installed is disabled by default and the SMC thus cannot check the time in the partition. When starting NTP operation, check the time difference from the SMC. To enable a daytime port, execute the following command in the partition:

```
# /usr/sbin/inetadm -e svc:/network/daytime:stream <RETURN>
```

2.3.6 Notes on installation of a spare SMC

Installing a spare SMC is becoming more common in preparation for any hardware failure in the main SMC. However, this has produced an undesirable condition. A definition error or error related to human error can lead to a time difference between the main SMC and spare SMC. This can cause a reset of the time and result in abnormal termination of high-order applications.

Detailed examples of NTP definitions are provided in Appendix D, "Installation of a Spare System Console," in the System Console Software User's Guide. Following the associated explanations, use NTP correctly.

Also note that no partition has a hardware clock. Therefore, when a partition is rebooted, the SMC time is obtained, and system startup is based on that time, unlike the NTP framework. If the same stratum level is mistakenly specified for both the main and spare SMCs, any time difference between the two SMCs may result in a time discrepancy. Be very careful regarding this point.

2.4 When Configuration of Settings Is Completed

When configuration of all settings on every server and client is completed, restart the system or execute the command shown below as a superuser on all servers and clients to start NTP operation. When configuration of NTP settings has been completed, NTP operation starts automatically when the system is restarted. Therefore, the following operation need not be repeated each time the system is restarted.

```
# /etc/init.d/xntpd start <RETURN>
```

Note: Using the SMC, restart the system or enter the "/etc/init.d/FJSVscsl restart" command to start NTP.

If settings are changed during operation, enter the following command before the above command:

```
# /etc/init.d/xntpd stop <RETURN>
```

The NTP start method has been changed in Solaris 10.

- 1) Check the NTP status.

```
# /usr/bin/svcs -a | grep ntp <RETURN>
disabled          januar_29  svc:/network/ntp:default
Inactive state
```

- 2) Enter the following command to start the NTP script:

```
# /usr/sbin/svcadm enable svc:/network/ntp:default <RETURN>
Start (Specify "disable" to stop it.)
```

3) Check the status to determine whether NTP is active.

```
# /usr/bin/svcs -a |grep ntp <RETURN>
online          January _29  svc:/network/ntp:default
Active state
```

Subsequently, NTP operation starts automatically when the system is restarted.

If NTP is in the active state (online) in step 1), enter the following command to apply the settings. In this event, skip steps 2) and 3).

```
# /usr/sbin/svcadm restart svc:/network/ntp:default <RETURN>
```

Note: Part of the start script has been changed in Solaris 10.

Solaris 9 and earlier versions support the -w option (wait for synchronization with the server) in the ntpdate command executed in the start script. If synchronization fails because of an inactive NTP server or network error, retry processing is executed to wait for the start of the xntpd daemon. However, the -w option has been eliminated in Solaris 10, which starts the xntpd daemon without retry processing even if synchronization with the NTP server fails for the above reasons. That is, the xntpd daemon starts without time correction by ntpdate, and clock synchronization may thus take a long time depending on the settings, if there is a substantial time difference with the server.

When starting the system, check whether the NTP server is active and stable. To use the same specifications as those for Solaris 9 or earlier, edit the start script (/lib/svc/method/xntp), and add the -w option.

Execute the ntpq command, and check whether NTP is actually operating.

```
# ntpq -np
      remote           refid      st t when poll reach  delay  offset  disp
=====
+192.168.246.9  192.168.246.16  7 u   61   64  377    0.67  -103.51  0.72
*192.168.246.16 127.127.1.0     6 u   60   64  377    0.67  -102.50  0.56
127.127.1.0    127.127.1.0     9 l   59   64  377    0.00   0.000 10.01
```

This example shows that synchronization occurs with the NTP server whose IP address is 192.168.246.16 with a time difference of -102.5 ms. Another reference server, 192.168.246.9, is synchronized with 192.168.246.16 with a time difference of about 1 millisecond (103.51-102.5). (For details on '+' and '+', see Section 4.1, "ntpq Command.")

Note: The first edition recommends a peer specification for primary and secondary server operation without a cluster external clock (external NTP server). However, Solaris 9 has related problems, such as the ntpdate command not ending at startup, xntpd not starting, and difficulty in achieving clock synchronization. Therefore, even for clusters, make settings for a hierarchical structure without making a peer specification.

3. Application to Actual Operations

3.1 Specifications about Time Changes

Actual operation may sometimes require that a clock be manually synchronized by applying NTP in an environment without an accurate reference clock. This section explains time-change-related specifications for NTP and Solaris.

3.1.1 Specifications for Solaris NTP daemon 3.x

If a closed NTP operation within an organization is executed without referencing NTP servers on the Internet and the NTP primary server in the organization does not have any reference clock, periodic manual clock synchronization is required. However, NTP was originally a protocol intended for synchronizing a clock with the precise time (UTC), and the case of a "manual change of the server time" is thus not assumed (*1). If necessary, therefore, change the server time manually after stopping NTP.

*1 5.3. Step Phase Adjustments of RFC-1305 (NTP V3) states that synchronization is required immediately if a detected shift from the server is 128 ms or more and the elapsed time after the last synchronization with the server is 900 seconds or more. However, there is no detailed specification of operations related to time changes and a greater part of the operations may well depend on actual implementation. The following table lists time-change operations of NTP Version 3.4y for Solaris 7, NTP Version 3.5.93e for Solaris 8 and 9, and NTP Version 3.5.93e+sun03/08/29 for Solaris 10.

Table 3.1 Clock synchronization methods and their effects

| Server status | Client status | Operation | Effect |
|-----------------|-----------------|---|--|
| xntpd operating | xntpd operating | Changes the time with the server (change of 128 ms or less) | The time in the client is corrected gradually. |
| xntpd operating | xntpd operating | Changes the time with the server (change of more than 128 ms) | After the time difference with the server exceeds 128 ms, it is monitored for 900 s. If a time difference of 128 ms or more is still detected, the NTP internal time is forcibly changed after resynchronization (64 seconds x 5). (If the appropriate patch described in Section 2.1 has been applied, the local clock is not set back. (*1)) |

* If the appropriate patch described in Section 2.1 has not been applied, the local clock may be set back. For details, see Section 3.3, "Gradual Clock Synchronization with NTP."

If a time difference of 1,000 seconds (about 17 minutes) or more from the reference server is detected during xntpd activation or clock correction, the NTP process on the client stops.

The three NTP thresholds in the above table are summarized as follows:

- (1) Step threshold (± 128 ms): Upper and lower limits of time differences between the system and reference server (Basically, the time difference must be kept within this range.)
- (2) Stepout threshold (900 s): Period in which any time difference of ± 128 ms or more from the offset value of the reference server is ignored
- (3) Panic threshold ($\pm 1,000$ s): If the time difference between the system and reference server is 1,000 seconds or more, clock synchronization with NTP is abandoned (xntpd is exited).

3.1.2 Interval for time polling with a high-order NTP server

If a time difference of 128 ms or more is detected in a change of the interval at which NTP in Solaris polls a high-order NTP server, the following action is taken:

- a) The client first checks the reference server time at a 64-second interval, and the interval is then gradually extended. The maximum interval is 1,024 seconds. The initial and maximum values can be specified as power-of-two values. However, the interval cannot be a constant. As long as the time difference between the reference server and NTP client is within 128 ms, the time is gradually corrected. The time is corrected either by increasing the integrated value if the client clock is slower or by decreasing the value if the clock is faster.
- b) If the time difference between the server and client exceeds 128 ms, the client monitors the server time for 900 seconds and then checks the time difference. If the time difference is still 128 ms or more, the local server time is corrected by time polling at least five times (the time taken for correction depends on the polling interval and other conditions). If the appropriate patch has been applied with the slew option specified, the xntp internal clock is corrected in steps but the system clock (local clock) is never set back. (If the patch has not been applied or the slew option was not specified, the clock may be set back.) For details, see Section 3.3, "Gradual Clock Synchronization with NTP."

3.1.3 Specifications for clock synchronization in Solaris

Changing the time in Solaris is prohibited while the Windows system is running and multi-user mode is set (see the main page about the date command). If the time is actually set back, the X Windows system may hang.

To change the system date/time for test purposes, be sure to use the date command in single-user mode.

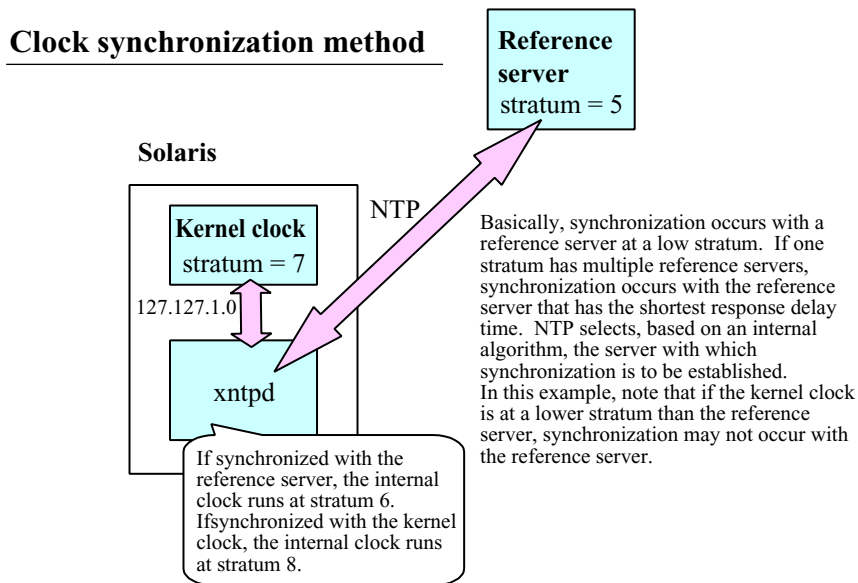
Caution: Changing /etc/system

Do not specify the following system parameter because it would have an adverse effect on the system clock:

```
set dosyncodr=0
```

3.2 Reference Server and NTP Clock Synchronization

The NTP daemon `xntpd` has an internal NTP clock, and it synchronizes the times of the local clock and reference server with the logical structure shown below.



3.3 Gradual Clock Synchronization with NTP

One of the patches listed below must be applied to prevent forced clock synchronization that eliminates a time difference (aggregated in R04111). (The `adjtime` function is used for gradual clock synchronization.)

| | |
|-------------|-----------|
| Solaris 2.6 | 107298-03 |
| Solaris 7 | 109409-04 |
| Solaris 8 | 109667-07 |
| Solaris 9 | None |
| Solaris 10 | None |

* If no patch has been applied, "time reset (slew)" is displayed.

When the patch has been applied (with the `slew` option specified), the following must be added to `ntp.conf` to synchronize the clock gradually:

```
slewalways yes  
disable pll
```

Before applying the patch, be sure to read its README file.

In verification results, a significant difference was evident in the clock synchronization speed when the slew option is specified with VL of xntpd.

The following table lists the operation verification results after application of the above patch with the slew option specified.

Time synchronization speed when the slew option is specified

| OS version | xntpd version | Patch ID | Time taken for a one-second correction |
|-------------|-------------------------------|-----------|--|
| Solaris 2.6 | xntpd-3.4y | 107298-03 | About 16 or 17 seconds |
| Solaris 7 | xntpd-3.4y | 109409-04 | About 16 or 17 seconds |
| Solaris 8 | xntpd-3.5.93e | 109667-04 | About 2,000 sec |
| Solaris 9 | xntpd-3.5.93e | - | About 2,000 sec |
| Solaris 10 | xntpd-3.5.93e+sun 03/08/29 | - | About 2,000 sec |

Solaris 8, 9, and 10 require quite a long time to correct a time difference of one second. This is because NTP has an internal value representing the "maximum time for time adjustments of less than one second," and it was changed for Solaris 8 and later (xntpd-3.5.93e).

In synchronization with an accurate reference server, the time difference from the reference server seldom exceeds plus or minus 128 ms. However, if an inaccurate Solaris clock is manually corrected for operation, clock synchronization may take quite a long time. In this event, use one of the following methods to prevent that situation:

- Stop xntpd, synchronize the clock by using `ntpdate -B`, and then reactivate xntpd. (Solaris 8 and later)
- If the time difference from the reference server is large, synchronize the clock to a certain degree in single-user mode, and then start the system.
- Provide an accurate reference server.

Note: Operation specifications for xntp when the slew option is specified

xntpd has internal time information (it does not have a clock). If the slew option is specified, the internal time of xntpd is synchronized in steps. (The offset value displayed for the `ntpq` command and in `peerstats` seems to indicate synchronization in steps.) However, the specifications require that the system clock be synchronized gradually by the `ajtime` function.

3.4 Time Change Methods in an NTP Operating System

3.4.1 Without an SMC

Basically, the clock must be synchronized on the client and server machines using NTP in single-user mode. However, this method cannot be applied to an environment for 24-hour operation.

The `-a` option of the `date` command can be used on a system in which NTP client operation cannot be stopped. In the `-a` option, specify the relative time (seconds) used for the change. The clock is synchronized by quickening or retarding the clock. Concretely, the time can be corrected by following the procedure below.

- 1) Stop `xntpd` on each client.

```
# /etc/init.d/xntpd stop
```

Note: `xntpd` under Solaris 10 can be stopped as follows:

```
# /usr/sbin/svcadm disable svc:/network/ntp:default
```

- 2) A time difference from the NTP server can be eliminated by simultaneous execution of `date -a` on each client. (In the following example, the clock is set back by seven seconds.)

```
# date -a -7
```

- 3) Wait until the time difference from the NTP server is eliminated on all clients (*1). No client may be in the stopped state at this point. The server time can be fetched, for example, by using the daytime protocol as follows (the daytime protocol uses port number 13):

```
# telnet NTP_server_IP_address 13
```

Note: For enhanced security, Solaris 10 is initialized to prevent unnecessary services from starting. By default on an NTP server running under Solaris 10, the daytime protocol cannot be used. To enable the daytime protocol, enter the following command on the server:

```
# /usr/sbin/inetadm -e svc:/network/daytime:stream
```

- 4) Restart `xntpd`.

```
# /etc/init.d/xntpd start
```

Note1: For details on how to start `xntpd` under Solaris 10, see Section 2.4, "When Configuration of Settings Is Completed."

Note2 To synchronize the clock gradually in Solaris 8, use the `-B` option of the `ntpdate` command. However, if the time difference is too large, the synchronization takes quite a long time. Therefore, start the synchronization after setting a time that is as close to the server time as possible (see Section 3.3, "Gradual Clock Synchronization with NTP").

Since the `-B` option of the `ntpdate` command cannot be used in Solaris 7 or earlier, synchronization is necessary in order to decrease the time difference to less than 0.5 seconds. (If the time difference from the server is less than 0.5 seconds immediately after `xntpd` activation, the time is gradually corrected by the `ntpdate` command using `adjtime`.) However, such synchronization is actually very difficult to implement. Therefore, apply a patch (see Section 3.3, "Gradual Clock Synchronization with NTP") before starting operation so that the time is gradually corrected, comment out the execution locations of the `ntpdate` command in `/etc/init.d/xntpd`, and then execute the script shown in step 4). Also, since quite a long time will be required in this case if the time difference is large, start the synchronization after setting a time that is as close to the server time as possible.

3.4.2 When an SMC is used

The SMC works as an NTP server for each partition. If a partition is stopped, the SMC time changed, and the partition restarted, the partition time is also changed. This enables partitions to operate with the correct time.

To change the SMC time, follow the procedure below. SMC root authority is required for these operations. The SMC time need not be adjusted in an operation mode where the SMC references a server that keeps accurate time.

- 1) Stop all partitions.
- 2) Start the SMC in single-user mode.

```
% su - <RETURN>
# /usr/sbin/shutdown -g0 -y -i0 <RETURN>

{?} ok boot -s <RETURN>
```

- 3) Enter the date command to set the correct SMC time.

```
# date 10080045 <RETURN>
```

- 4) Reboot the SMC.

```
# /usr/sbin/shutdown -g0 -y -i6 <RETURN>
```

- 5) Start each partition by using SMC machine management.

4. NTP Status Information

This section explains NTP status information.

4.1 ntpq Command

Executing the ntpq command (-np option) displays the information shown below. For details on the command option, see ntpq (1M).

```
# ntpq -np
      remote          refid      st t when poll reach  delay  offset  disp
=====
+192.168.246.9  192.168.246.16  7 u  61  64  377   0.67 -103.51  0.72
*192.168.246.16 127.127.1.0     6 u  60  64  377   0.67 -102.50  0.56
 127.127.1.0    127.127.1.0    9 l  59  64  377   0.00  0.000  10.01
```

The clock runs slow with regard to server 192.168.246.16, indicated by the * mark, with a time difference of -102.5 ms. If the time difference from the server becomes larger, "*" is not displayed.

The offset value displayed by the ntpq command shows a time difference between the xntpd internal time of the local server and the xntpd internal time of the reference server.

The following table explains the items displayed by the ntpq command, starting from the left side.

| | |
|---------------|--|
| remote | Host name or IP address of the referenced NTP server |
| refid | Host name/IP address of the high-level server referenced by the server displayed in "remote" |
| st | Stratum number of the server displayed in "remote" |
| t | Server type u : unicast l : local m : multicast b : broadcast |
| when | Elapsed time after synchronization (seconds) |
| poll | Synchronization interval (seconds) |
| reach | Reachable register (octal number) |
| delay | Delay time (milliseconds) |
| offset | Time difference from the NTP server (milliseconds) |
| disp | Dispersion (time dispersion) (milliseconds) |

The following table explains the symbols displayed in front of remote items.

| | |
|--------------|---|
| Blank | Removed from the synchronization candidate list because of a high stratum number (low accuracy) or unreachability |
| x | Removed from the synchronization candidate list (has an inaccurate clock as determined by an algorithm) |
| . | Removed from the synchronization candidate list (culled from the end of the candidate list) |
| - | Removed from the synchronization candidate list (by the clustering algorithm) |
| + | On the synchronization candidate list |
| # | Currently synchronized but distant server |

| | |
|---|-------------------------------------|
| * | Server currently being synchronized |
| o | Synchronized (pps signal in use) |

4.2 ntptrace Command

Executing the ntptrace command displays the following information:

```
# ntptrace
localhost: stratum 2, offset 0.000066, synch distance 0.02495
techntp: stratum 1, offset 0.000088, synch distance 0.00000, refid 'GPS'
```

The displayed items, starting from the left side, have the following meanings:

1. Host name of the reference server
2. Stratum number of the reference server
3. Time difference between the system clock of the local machine and the xntpd internal time of the reference server (seconds)
4. Synchronizing distance
5. Reference clock ID (displayed only for stratum1)
 - * The offset value displayed for the ntptrace command is the time difference between the system clock of the local machine and the system clock of the reference server. For example, if the option (slew option) for gradual clock synchronization is selected, this value may differ from the offset value displayed for the ntpq command.

4.3 peerstats File

The `/var/ntp/ntpstats/` directory (or the directory specified by `statsdir` in `ntpd.conf`) contains daily time difference information such as `peerstats.20010124`. For example, it contains the following record:

(Example) 51933 27.688 10.34.150.1 9614 -0.000903 0.00523 0.00035

Each field is explained in the paragraph on `peerstats` in `xntpd (1M)`.

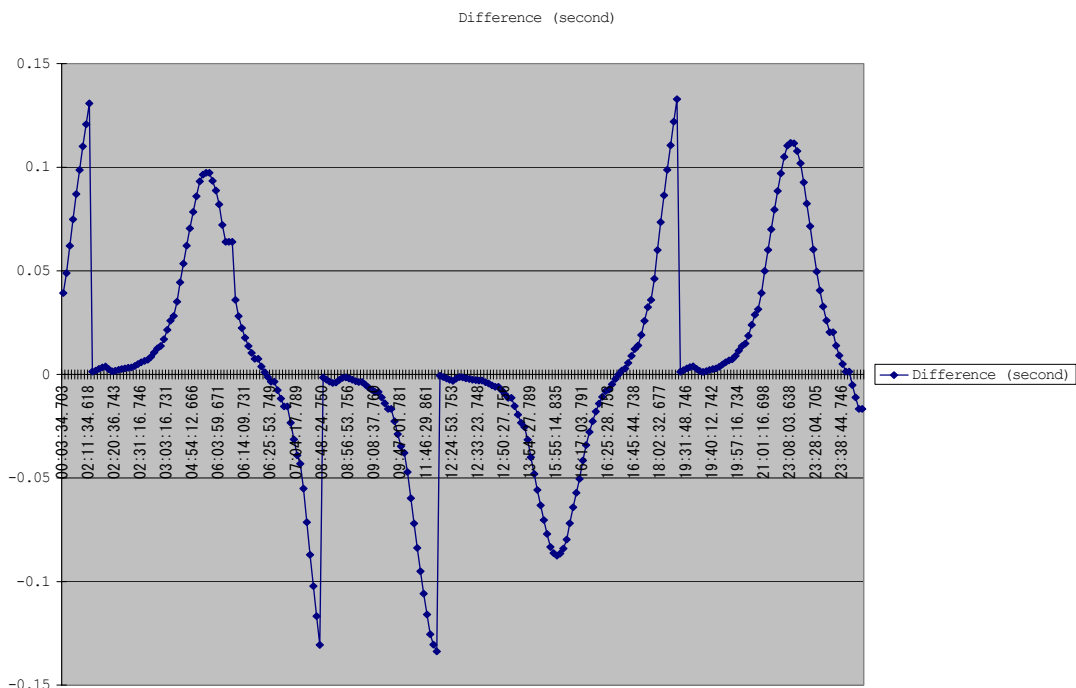
The fields are arranged in the following sequence and format:

1. Julian Date date
2. UTC time (unit: seconds)
3. Remote IP address (Note: 127.127.1.0 is the local clock.)
4. peerstatus in hexadecimal notation
5. Offset in seconds (-> time difference from the reference clock)
6. Delay in seconds
7. Dispersion in seconds

The date and time (seconds) are UTS times. (+9 hours (+9*60*60) must be added for conversion to Japanese standard time.)

The above example shows a time difference of -0.903 ms from the server 10.34.150.1.

Analysis of changes in time differences in this file can verify the amount of the time difference from the reference server during operation. The following figure shows an example of changes in the time difference from the reference server.



Note: The offset value in `peerstats` is the time difference between the `xntpd` internal clock and the `xntpd` internal clock of the reference server. The value does not necessarily match the time difference between the system clock of the operating system and the reference server clock. The time difference between the system clock of the operating system and the reference server clock can be examined with the `ntptrace` command.

4.4 Checking NTP Operation with snoop Data

Almost all information about NTP operation (e.g., the time difference from the reference server and the synchronous state) can be referenced with the `ntpq` command, `ntptrace` command, and `peerstats` file. However, to check whether a response from the NTP server is delayed in Server-Client mode operation or to reference details about response packets, an analysis of snoop data is required.

An example of NTP snoop data is shown below. Since NTP uses UDP port 123, only NTP data can be retrieved.

```
522  1.91667 10.34.151.57 -> netst11      NTP symmetric active (Mon Jun  3
12:13:03 2002)
550  0.41989 10.34.197.112 -> netst11      NTP client (Mon Jun  3 12:11:54
2002)
551  0.00037      netst11 -> 10.34.197.112 NTP server (Mon Jun  3 12:13:09
2002)
552  0.06355      netst11 -> 10.34.184.26 NTP client (Mon Jun  3 12:13:09
2002).....(1)
553  0.01175 10.34.184.26 -> netst11      NTP server (Mon Jun  3 12:13:09
2002) .....(2)
```

- (1) Operation when "server" is specified in `ntp.conf` (client mode). This packet is used to make a time inquiry to the server.
- (2) Response from the NTP server to the inquiry of (1)

Note: Responses are returned from the NTP server only in Server-Client mode.

Detailed snoop (-v option specified) data is shown below (NTP part only).

```
NTP:
NTP: Leap = 0x0 (OK)
NTP: Version = 3
NTP: Mode    = 4 (server)
NTP: Stratum = 1 (primary reference)
NTP: Poll   = 6
NTP: Precision = 242 seconds
NTP: Synchronizing distance = 0x0000.0000 (0.000000)
NTP: Synchronizing dispersion = 0x0000.0000 (0.000000)
NTP: Reference clock = GPS
NTP: Reference time = 0xc0a55d3c.00000000 (Mon Jun  3 12:13:00 2002)
NTP: Originate time = 0xc0a55d45.4d01d000 (Mon Jun  3 12:13:09 2002)
NTP: Receive time = 0xc0a55d45.51a60d27 (Mon Jun  3 12:13:09 2002)
NTP: Transmit time = 0xc0a55d45.526d47e4 (Mon Jun  3 12:13:09 2002)
```

Details on the packet format are shown below.

To check for the time synchronization state in Server-Client mode among different kinds of information, reference the Leap value in a response packet from the NTP server.

A check of whether a response packet is delayed on the NTP server can be performed by examining the difference between Receive Time and Transmit Time.

[Details]

The packet format is explained below using displayed results of snoop data. Though the names of some fields are different from those used in RFC1305, the field contents are the same.

Leap : 2bit

| | | | |
|------|---------------------------------------|----|------------------------|
| 0x00 | no warning(OK) | -> | Synchronous state |
| 0x01 | last minute has 61 seconds | } | Used for a leap second |
| 0x02 | last minute has 59 seconds | | |
| 0x11 | alarm condition(clock unsynchronized) | | |

-> This value is set immediately after server startup or at the time that synchronization is lost.

Version: NTP Version (3 bits)

| | |
|---|-----------|
| 2 | Version 2 |
| 3 | Version 3 |

Mode (3 bits)

| | |
|---|--|
| 0 | reserved |
| 1 | symmetric active |
| 2 | symmetric passive |
| 3 | client |
| 4 | server |
| 5 | broadcast |
| 6 | reserved for NTP control message (Refer to RFC1305 Appendix B) |
| 7 | reserved for private use |

Stratum number (8 bits)

| | |
|-------|---------------------|
| 0 | unspecified |
| 1 | primary reference |
| 2-255 | secondary reference |

Polling interval (8 bits)

This value is 2 to the x-th power (unit: seconds). In the above example, the value is 2 to the 6th power. Thus, the polling interval is 64 seconds.

Precision of the local clock (8 bits)

This value is 2 to the x-th power (unit: seconds).

Synchronizing distance : (32bit)

This 32-bit fixed-point number represents the round-trip delay time to the primary or reference source.

Synchronizing dispersion : (32bit)

This 32-bit fixed-point number represents the maximum error with regard to the primary or reference source.

Reference Clock : Reference clock identifier (32 bits)

This 32-bit code identifies a specific reference clock. For details, see RFC1305.

Reference Time : (64bit)

This time in the 64-bit timestamp format is the time at which the local clock was last set or corrected.

Originate Time : (64bit)

This time in the 64-bit timestamp format is the time at which the client made a request to the server.

Receive Time : (64bit)

This time in the 64-bit timestamp format is the time at which the server received a request from the client.

Transmit Time : (64bit)

This time in the 64-bit timestamp format is the time at which the server returned a response to the client.