Using NTP on the Sun Fire™ 15K / 12K Servers

Jason Beloro, Enterprise Systems Products Engineering

Sun BluePrints™ OnLine—June 2003

http://www.sun.com/blueprints

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95045 U.S.A.
(650) 960-1300

Part No. 817-2979-10
Revision 04, 6/2/03
Edition: June 2003

Please
Recycle

Using NTP on the Sun Fire 15K/12K Servers

The Sun Fire™ 15K/12K server system controller (SC) is a SPARCengine® CP1500 board, which is a member of Sun's family of platforms based on CompactPCI technology. One of the key components of the CP1500 board is the NVRAM device that contains a time-of-day (TOD) clock and non-volatile memory that stores the MAC address of the CP1500 board. The NVRAM has its own lithium battery to operate the clock and to retain the contents of the NVRAM during power-off situations.

The real-time clock circuitry has an accuracy of plus or minus one second per day. Despite this accuracy, both field and internal sources have noted that the SCs are losing from 1.4 to 8 seconds per hour. As a result, several bugs have been opened. The solution to prevent these time skews is to use the Network Time Protocol (NTP), which is designed to synchronize computer clocks over a network.

This article addresses the time skew issues for the Sun Fire 12K/F15K server and explains how the system controllers and domains can be configured as NTP clients to external servers. A sample configuration is also provided.

NTP Concerns

The *System Management Services (SMS) 1.1 Administrator Guide* and the *System Management Services (SMS) 1.2 Administrator Guide for the Sun Fire™ 15K/12K Systems* contain the following statements:

1. If NTP makes a large time correction on the SC, the `setdate`(1M) command will not store the correct time offsets.

2. The main SC propagates the time to the spare SC by file propagation. This process causes the time to be off on the spare SC even if NTP is in use.

In response to these statements, Enterprise Systems Products (ESP) Engineering has made the following observations:

1. Many of the assumptions that the offset will not be correct are based on the ideas that the offset was set during a skewed SC time period or that domains will reboot and receive the time through the OpenBoot™ PROM server daemon, `osd`(1M), during a skewed SC time period. ESP Engineering does not believe that these examples are valid.

   If NTP is in use on the SC and the `setdate`(1M) command is used to set the time offset for a domain, the correct domain offset will be established from a correct SC time source. The corrections made by NTP are towards the correct time, not away from the correct time. Therefore, if the SC time is maintained by NTP, the offset should continue to be correct.

   If the SC and the domains are synchronized using NTP, then using the `setdate`(1M) command is not necessary. The `setdate`(1M) command would only be necessary if a user desires a special case for a non-standard SC time or domain time. Otherwise, the SC times and domain times will be correct because the same NTP server will align the SC times and domain times.

   Adjustments can be made locally in the Solaris™ Operating Environment (OE) at any time by setting the time zone. Setting the time zone does not change the offset established by the `setdate`(1M) command. However, if NTP is used on the domain, NTP nullifies the offset when it makes corrections and synchronizes the time with the NTP servers. If an offset is needed, NTP should not be running on the domain; however, that does not prevent NTP from running on the SCs.

2. With NTP on the SCs, the time on the spare SC is corrected. The time is propagated to the spare SC from the main SC. The act of reading, packaging, and restoring the time takes a significant amount of time. If the main SC is busy, a significant time error on the spare SC can be introduced. Without NTP on the spare SC, a failover could cause the time to be off by more than a second. With NTP, the NTP daemon, `xntpd`(1M), would constantly be making corrections to the time. Thus, the amount of time error on the spare SC is decreased.

3. The `setdate`(1M) command is already inaccurate because it requires the time to be set manually. In addition, it is only accurate to the second, and even this level of accuracy is questionable. Thus, the accuracy is significantly better with NTP running on the SCs than it is with a manual time adjustment.

Recommended NTP Configuration

ESP Engineering strongly recommends that the main SC, the spare SC, and each bootable domain on the platform be configured as NTP clients to the same reliable NTP servers that are external to the Sun Fire 15K/12K server. Make sure that the System Management Software patches are current and that the latest Solaris OE recommended patch cluster is installed on the SCs and domains.

If the Solaris 8 OE is installed on the SCs, make sure that the Kernel Update patch level is at KU-13 (108528-13) or higher. For the latest revision of the KU patches, check the SunSolve℠ program web site at: `http://sunsolve.sun.com`

Also, for the Solaris 8 OE, you *must* have the following NTP patch installed:

- Patch-ID: 109667-04
- Keywords: security, `xntpd`, NTP, system clock, `SIGPOLL`, `SLEWALWAYS` overflow
- Synopsis: SunOS 5.8: `/usr/lib/inet/xntpd` and `/usr/sbin/ntpdate patch`
- Date: Oct/15/2001

Caution – For the Solaris 8 OE, do not use Patch-IDs 109667-01, 02, or 03. Using one of these patches could prevent NTP from fixing the time drift.

Configuring NTP

To set up the SCs and domains as NTP clients, edit the `ntp.conf` file. The `xntpd`(1M) daemon reads the `ntp.conf` file at startup time. The configuration must contain a minimum of three time servers (with independent time sources), as recommended by the NTP protocol designers.

Using a text editor, edit the `ntp.conf` file by inserting the following lines for the main SC, the spare SC, and the bootable domains. Be sure to add the `prefer` argument to the host that is preferred for synchronization.

```
# /etc/inet/ntp.conf
server NTP_Server1 prefer
server NTP_Server2
server NTP_Server3
```

The drift file should also be added to record the frequency offset of the local clock oscillator. The drift file is read at startup to set the initial frequency offset. Add the following line to the `ntp.conf` file:

```
driftfile /var/ntp/ntp.drift
```

Statistics and file generation should also be used. Be sure to indicate the directory path where statistics files should be created. In the following example, `peerstats` are generated to record peer statistics, and `loopstats` are generated to record loop filter statistics. Finally, `clockstats` are generated to record clock driver statistics.

To set up statistics and file generation, add the following lines to the `ntp.conf` file:

```
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
```

Note – You should set up a `cron`(1M) job to remove the existing statistics files on a regular basis.

The following is a full example of an `ntp.conf` file:

```
# /etc/inet/ntp.conf
server clock.via.net prefer
server navobs1.usnogps.navy.mil
server ntp.ucsd.edu
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
```

After you have edited the `ntp.conf` file, you must stop and restart the NTP daemon, as in the following example:

```
sc:# /etc/init.d/xntpd stop
sc:# /etc/init.d/xntpd start
```

Note – NTP is an insecure protocol. If you are running a secure environment, you should take additional steps to implement the appropriate level of security. Refer to the Sun BluePrints™ articles cited in the “References” section.

Other NTP Configuration Options

Other NTP configuration options, such as reference clock support (the `fudge` command) and address restriction, were not tested for the recommended configuration in this article. Almost all of the supported commands and options of `xntpd`(1M) can still be used for configuring NTP on the SCs and domains. However, they are not necessary for the minimum recommended configuration in this article.

Caution – Do not use the `SLEWALWAYS` option on the Sun Fire 15K/12K server. Using this option could prevent NTP from fixing the time drift.

Consult `xntpd`(1M) for the full list of configuration options. For public NTP servers, check the Public NTP Server list on: `http://www.ntp.org`

Acknowledgements

Special thanks is given to those who assisted in creating this document, especially the following people for their expertise and help: Dan Anderson, Rob Bartlett, Darin Carlson, David Fan, Jacob Lowman, Tony Ho Nguyen

References

- Deeths, David and Glen Brunette. “Using NTP to Control and Synchronize System Clocks - Part I: Introduction to NTP.” Sun BluePrints OnLine, July 2001. `http://www.sun.com/blueprints/online.html`
- ———. “Using NTP to Control and Synchronize System Clocks - Part II: Basic NTP Administration and Architecture.” Sun BluePrints OnLine, August 2001. `http://www.sun.com/blueprints/online.html`

- ———. “Using NTP to Control and Synchronize System Clocks - Part III: NTP Monitoring and Troubleshooting.” Sun BluePrints Online, September 2001. `http://www.sun.com/blueprints/online.html`
- Network Time Protocol Project. “NTP: The Network Time Protocol.” `http://www.ntp.org`
- Sun Microsystems. *SPARCengine™ CP1500 360MHz/440MHz Technical Reference and Installation Manual.* `http://www.sun.com/oem/products/manuals/806-2104-03.pdf`
- ———. “Configuring NTP,” *System Management Services (SMS) 1.2 Administrator Guide for the Sun Fire™ 15K/12K Systems.* `http://www.sun.com/products-n-solutions/hardware/docs/Servers/High-End_Servers/Sun_Fire_15K/SW_FW_Documentation/SMS/index.html`
- SunSolve program web site. “The emulated TOD clock on the SC should be NTP aware / accessible” (Sun BugID 4510625). `http://sunsolve.sun.com`
- ———. “Starcat SC suffers drifts because `sys_tick_freq` is set incorrectly on boot” (Sun BugID 4656724). `http://sunsolve.sun.com`

About the Author

Jason Beloro is an engineer in the Enterprise Systems Products group. He has been at Sun for six years, working on various software projects on the Sun Enterprise™ 10000 server and now the Sun Fire 15K server. Jason has a Bachelor of Science in Computer Engineering from the University of California, San Diego.

Ordering Sun Documents

The SunDocs℠ program provides more than 250 manuals from Sun Microsystems, Inc. If you live in the United States, Canada, Europe, or Japan, you can purchase documentation sets or individual manuals through this program.

Accessing Sun Documentation Online

The `docs.sun.com` web site enables you to access Sun technical documentation online. You can browse the `docs.sun.com` archive or search for a specific book title or subject. The URL is `http://docs.sun.com/`

To reference Sun BluePrints OnLine articles, visit the Sun BluePrints OnLine web site at: `http://www.sun.com/blueprints/online.html`