IBM

Server Time Protocol Recovery Guide

- Server Time Protocol concepts and definitions
- Planning for recovery from various failures
- Restoring STP configuration after a POR

Octavian Lascu
Hans-Peter Eckam
George Kozakos
Paulo Vitor Pereira

Redbooks

ibm.com/redbooks

International Technical Support Organization

Server Time Protocol Recovery Guide

June 2013

SG24-7380-01

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

Second Edition (June 2013)

This edition applies to the Server Time Protocol Facility on IBM zEnterprise EC12 (zEC12), IBM zEnterprise 196 (z196), IBM zEnterprise 114 (z114), IBM System z10™ Enterprise Class (z10 EC), System z10 Business Class (z10 BC), IBM® System z9® Enterprise Class (z9 EC), and System z9 Business Class (z9 BC).

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

ESCON®	Parallel Sysplex®	System z9®
FICON®	Redbooks®	System z®
GDPS®	Redbooks (logo)  ®	z/OS®
Global Technology Services®	S/390®	z10™
IBM®	Sysplex Timer®	z9®
MVS™	System z10®	zEnterprise®

The following terms are trademarks of other companies:

Other company, product, or service names may be trademarks or service marks of others.

Preface

Server Time Protocol (STP) is a server-wide facility that is implemented in the Licensed Internal Code (LIC) of the IBM® zEnterprise® EC12 (zEC12), IBM zEnterprise 196 (z196), IBM zEnterprise 114 (z114), IBM System z10® Enterprise Class (z10 EC), System z10 Business Class (z10 BC), IBM System z9® Enterprise Class (z9 EC), and System z9 Business Class (z9 BC). It provides improved time synchronization in a sysplex or non-sysplex configuration.

This IBM Redbooks® publication will help you plan for and recover from a failure affecting your Mixed or STP-only Coordinated Timing Network. It is intended for technical support personnel requiring information about:

- Recovery concepts and definitions
- Identifying and taking appropriate actions for recovering from a failed component in a Coordinated Timing Network

Readers are expected to be familiar with IBM System z® technology and terminology. For planning information, refer to our companion book, *Server Time Protocol Planning Guide,* SG24-7280, and for implementation details refer to *Server Time Protocol Implementation Guide,* SG24-7281.

Authors

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center.

Octavian Lascu is a project leader at the International Technical Support Organization, Poughkeepsie Center. He writes extensively and teaches IBM classes worldwide on subjects such as cross-platform clustering and High Performance Computing. Before joining the ITSO 11 years ago, Octavian worked in Global Technology Services®, Romania, as senior IT consultant.

Hans-Peter Eckam is a Support Center Specialist in the IBM Hardware Support Center in Mainz, Germany. Hans-Peter joined IBM in 1986 and has 22 years of experience in supporting System z clients. His areas of expertise include System z, Sysplex Timer®, Server Time Protocol, and connectivity such as ESCON®, FICON®, and InfiniBand.

George Kozakos is a Senior IT Specialist with IBM Australia. He has more than 25 years of experience in MVS™ system programming. George's areas of expertise include Server Time Protocol and GDPS®. He holds degrees in Computing Science and Pure Mathematics.

Paulo Vitor Pereira is an IT Specialist in STG Lab Services in São Paulo, Brazil. He joined IBM in 2009 and has six years of experience in System z. His areas of expertise include, z/OS®, WLM, z/BX and System z hardware.

William G White
International Technical Support Organization, Poughkeepsie Center

Donald J Smith, Richard M Hamilton, Steve Cherniak, IBM Poughkeepsie

Gregory K. Hutchison
IBM Gaithersburg

Thanks to the authors of the previous edition of this book:

Authors of the first edition, *Server Time Protocol Recovery Guide*, published in September 2010 were:

Noshir Dhondy, Hans-Peter Eckam, Andy Kilhoffer, Jeremy Koch, Hao Shen, Martin Söllig, Ray Takabe, Octavian Lascu.

Thanks to the ITSO support team:
Robert Haimowiz, Donald R Brennan, Roy P Costa, and Alfred Schwab (editor)
International Technical Support Organization, Poughkeepsie Center

Now you can become a published author, too!

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

Part 1

Concepts and planning

This part introduces STP recovery concepts and definitions, and how they apply to each type of CTN. We also describe recovery in a coordinated timing network, starting with a review of ETR network recovery, followed by recovery for both mixed and STP-only coordinated timing networks, including recovery of an external time source.

1

Concepts and considerations

In this chapter we provide an explanation of the STP terminology. We also present an overview of STP recovery concepts and definitions, and how they apply to each type of CTN.

- Concepts
- Sysplex Failure Management considerations
- Internal Battery Feature considerations
- Multisite considerations

1.1 Concepts

This section provides an overview of STP recovery concepts.

Note: Throughout this document we use the Stratum term for both STP and NTP to identify how far away a server is from its respective time source. Because the STP Stratum level definition is different from NTP Stratum, for clarity, when necessary, we identify the type of Stratum level by prefixing it with STP or NTP (depending on which one is discussed in the respective context).

For detailed definitions of STP and NTP Stratum levels, see Chapter 1 in *STP Planning Guide*, SG24-7280-03.

1.1.1 Sysplex Timer Offline Sequence

When a Sysplex Timer detects a failure, it transmits a special sequence of symbols called the *Offline Sequence* on the CLO links to signal the other Sysplex Timer in an Expanded Availability configuration that it is going offline.

The following recovery rules pertain to the Offline Sequence:

- The primary Sysplex Timer does not depend on receiving the Offline Sequence from the secondary Sysplex Timer. It continues to transmit to the servers whether it receives the Offline Sequence or not.
- If a secondary Sysplex Timer receives an Offline Sequence from the primary Sysplex Timer, it becomes the primary Sysplex Timer and continues to transmit to the servers.
- If all communication between the Sysplex Timers ceases and neither Sysplex Timer receives the Offline Sequence, the primary Sysplex Timer continues to transmit to the servers and the secondary Sysplex Timer discontinues transmission.
- When multiple Clock Link Oscillator (CLO) links are configured between Sysplex Timers, if a failure is detected on a single link, the Offline Sequence is not transmitted by either Sysplex Timer, and both primary and secondary Sysplex Timers continue to transmit to the servers.

These rules ensure that at least one Sysplex Timer continues to transmit for the following types of failures:

- Hardware or software failures detectable in each Sysplex Timer
- Power outage of either Sysplex Timer
- Failure of one or both CLO links
- Failures of extenders used between Sysplex Timers
- A disaster at the data center with the secondary Sysplex Timer

Note: This is not a comprehensive list of Sysplex Timer-related failures.

1.1.2 Synchronization check threshold

In STP timing mode, a server is considered to be in synchronized state if its TOD clock is within the *synchronization check threshold* of Coordinated Server Time (CST).

The synchronization check threshold has been set at 50 microseconds (µs). If the server's TOD clock differs from CST by more than +/- 50µs, it is considered to be unsynchronized. If the server is unable to resynchronize, it becomes a Stratum 0 server.

The synchronization check threshold of 50 µs was selected to accommodate non-Parallel Sysplex configurations where servers need to be synchronized but do not support the Message Time Ordering Facility (MTOF) or an equivalent facility. In such a configuration, the servers might be providing communication over I/O paths and the STP synchronization check threshold must be set to no more than half of the minimum system-to-system data transfer time over I/O paths.

1.1.3 Freewheel interval

The *freewheel interval* is the amount of time a server in a Mixed CTN or STP-only CTN can remain synchronized without receiving timing messages from its time source. When a server loses its time source, it enters freewheel mode and begins recovery processing to try to establish an alternate time source in order to maintain synchronization and remain in the CTN.

If the freewheel interval expires, and the server has not reestablished connectivity to another valid time source within the CTN, then the server switches to an unsynchronized timing state and becomes a Stratum 0.

The duration of the freewheel interval depends on the CTN type:

- In a Mixed CTN, the freewheel interval is approximately 1 second.
- In an STP-only CTN, the freewheel interval is approximately 10 seconds, depending on server type.

STP Stratum 1 freewheel

The freewheel interval has no meaning for a Stratum 1 in an STP-only CTN. The Stratum 1 is always synchronized because it is the time source for the other servers in the CTN.

However, in a Mixed CTN, the Stratum 1 servers are using the Sysplex Timer as the time source. If full ETR connectivity is lost to a server, the freewheel interval applies. The server can either convert to STP timing mode and become a Stratum 2 (or Stratum 3, depending on the available connectivity) or, if no alternate time source can be found, the failure consequences are the same as a server in an ETR network losing all ETR connectivity.

STP Stratum 2 and Stratum 3 freewheel

For Stratum 2 and Stratum 3 servers, the freewheel interval is entered when STP timing messages from the selected time source are not received. If no alternative time source can be found, the server will become unsynchronized and switch to STP Stratum 0.

If an alternative time source *is* available, switching to a different Stratum level may be required in order to receive STP timing messages. A Stratum 2 server that switches to Stratum 3 due to unavailable connectivity will remain synchronized. However, a Stratum 3 server will become unsynchronized if available connectivity requires it to switch to a Stratum 4, since a Stratum 4 is not supported by STP.

1.1.4 Server Offline signal

The *Offline signal (OLS)* is used by a server to indicate to attached control units and servers that the channel is going offline. Offline signals are issued regardless of STP.

Conditions that cause an Offline signal to be sent by a server include:

- Server or LPAR dump
- Server power off
- CHPID taken offline

An Offline signal may not be transmitted for certain failures, such as:

- Channel subsystem failure
- System Assist Processor (SAP) recovery
- Site or server power outage
- Link failure

In a configuration where you have PTS and BTS connected with at least two initialized timing links and without a defined Arbiter, STP will use Offline signals received on initialized timing links between the PTS and the BTS to determine the state of the server at the other end.

With STP Version 2 and above, the following rules apply:

- If the BTS receives OLS on multiple links, including the last link to the CTS within a two-second interval, the BTS will take over as CTS and thus become Stratum 1.
- If the CTS has sent OLS on multiple links, including the last link to the BTS within a two-second interval, the CTS will release its CTS role.
- Console Assisted Recovery (CAR) will then be used to verify the PTS status. Only if CAR indicates that PTS is up and still Stratum 1, the BTS will give up the role as CTS.

If there is only a single initialized timing link between PTS and BTS or if the OLS on the last two links is more than two seconds apart, CAR is used to determine whether BTS can take over (see 1.1.7, “Console-assisted recovery” on page 10).

The Offline signal rules only apply in an STP-only configuration with PTS and BTS assigned and without an Arbiter configured. In a configuration with an Arbiter assigned, the OLS will be ignored and the Arbiter-assisted recovery process will be used instead (see 1.1.6, “Arbiter-assisted recovery” on page 7). Even if the assigned Arbiter is not available, the OLS will still not be used.

Note: The recovery actions described in this chapter assume that the PTS has been assigned the role of the CTS. This is the recommended configuration because only the PTS can perform automatic takeback after a recovery action.

When the BTS has become the CTS, the BTS tasks are performed by the PTS, if available.

1.1.5 Going Away Signal

The *Going Away Signal (GAS)* is a reliable unambiguous signal to indicate that the CPC is about to enter a check stopped state. When a GAS from the CTS is received by the BTS it safely takes over as CTS.

GAS has priority over OLS in a CTN where an Arbiter has not been assigned. The BTS can also use GAS to take over as CTS for CTNs with an Arbiter assigned without communicating with the Arbiter. This is in contrast to OLS where OLS is ignored for CTNs with an Arbiter assigned.

GAS removes the dependency on OLS and CAR in a CTN without an Arbiter assigned and the dependency on BTS to Arbiter communication for CTNs with an Arbiter assigned.

GAS is sent on InfiniBand (IFB) links using HCA3-O to HCA3-O - 12x IFB or 12x IFB3 or HCA3-O LR to HCA3-O LR - 1x IFB for z196 GA2 and later machines.

The current recovery design is still used when GAS is not received by BTS and for other failure types.

1.1.6 Arbiter-assisted recovery

Arbiter-assisted recovery is applicable when both a BTS and an Arbiter are assigned. The BTS does not invoke OLS rules, because the Arbiter provides additional means to determine whether the BTS can take over.

If the BTS loses communication on all of its established paths to the CTS, it attempts to determine the status of the CTS through the Arbiter:

- If both the BTS and the Arbiter cannot communicate with the CTS, then the BTS takes over as the CTS and becomes the Stratum 1.
- If the CTS is still alive after losing communication with both the Arbiter and the BTS, it will switch to Stratum 0 so that the CTN does not end up with two Stratum 1 servers.
- If the Arbiter can communicate with the CTS, then the BTS will not take over, but instead transition to Stratum 3 and get its timing signals from a Stratum 2 server, for example the Arbiter. The Arbiter takes note that the BTS no longer has connectivity to the CTS and, should it subsequently lose contact with the CTS, the Arbiter will inform the BTS accordingly, causing the BTS to proceed with taking over as the CTS.
- If the BTS is unable to communicate with the Arbiter due to a connectivity failure, Console-assisted recovery is invoked by the BTS as an alternate method for determining the status of the CTS.

Note: In a two-site CTN, the location of the Arbiter is critical: it will determine which site remains operational after a loss of communication between the PTS and BTS. Arbiter location considerations are discussed in 2.4.8, “Two-site considerations” on page 39.

Blocking disruptive actions on STP role servers

Disruptive actions such as POR are blocked for the CTS. New function has now been added to block disruptive actions on any of the STP role servers—PTS, BTS, or Arbiter. This prevents a disruptive action causing the CTS to give up the S1 role and go S0. For example, if the PTS has the CTS role and then the Arbiter (or BTS) has a planned or unplanned outage, then a disruptive action on the BTS (or Arbiter) causes the CTS to give up the S1 role and go S0 because it loses communication to both the BTS and Arbiter.

This new function was introduced with the MCL levels shown in Table 1-1.

Table 1-1 Minimum MCL levels for blocking disruptive actions on STP role servers

Driver/Server	MCL	Bundle	Release Date
D86E/z196	N29809.277	45	Sept 8, 2011
D86E/z196	N29802.420	45	Sept 8, 2011
D79F/z10	N24415.078	50	Sept 28, 2011
D79F/z10	N24409.184	50	Sept 28, 2011
D93G/z114 & z196 GA2	Integrated	N/A	Sept 9, 2011

Arbiter-assisted recovery enhancements

The current STP Arbiter-assisted recovery design handles recovery of single failures in an STP-only CTN. Enhancements have now been made to handle planned and unplanned actions that could impact two of the three STP role servers.

This includes safeguarding against the following potential hazards that could result in a CTN-wide failure:

- Planned disruptive actions on the BTS and Arbiter in parallel as part of the same task. Note that disruptive actions on any of the STP role servers will be blocked via the enhancement described in “Blocking disruptive actions on STP role servers” on page 7.
- Unplanned failure of a second STP role server when the STP role for the first unplanned failure is not reassigned or removed.
- Failure to remove the STP role of a server being upgraded to a new machine type. If the STP role is not removed prior to the upgrade, the new node descriptor prevents the server from reassuming the STP role. This puts the CTN in a state equivalent to that when the server with that STP role has an unplanned outage.

Arbiter-assisted recovery is enhanced so that a degraded state is entered when any two of the three STP role servers (PTS, BTS, or Arbiter) agree that they cannot communicate with the third STP role server. A degraded state is entered when:

- The BTS and Arbiter can communicate but cannot communicate with the PTS/CTS. The BTS will take over as the CTS and then Arbiter-assisted recovery is disabled.
- The PTS and BTS can communicate but cannot communicate with the Arbiter .
- The PTS and Arbiter can communicate but cannot communicate with the BTS.

This new function was introduced with the MCL levels shown in Table 1-2.

Table 1-2 MCL levels for enhanced Arbiter-assisted recovery:

Driver/Server	MCL	Bundle	Release Date
D79F/z10	N24406.094	50	Sept 28, 2011
D86E/z196	N29799.110	44	Aug 24, 2011
D93G/z114 & z196 GA2	Integrated	N/A	Sept 9, 2011

Note: The PTS, BTS, and Arbiter servers must all be at the required MCL level for Arbiter-assisted recovery to be disabled.

When Arbiter-assisted recovery is disabled, the Network Configuration panel displays the message shown in Figure 1-1 on page 9.

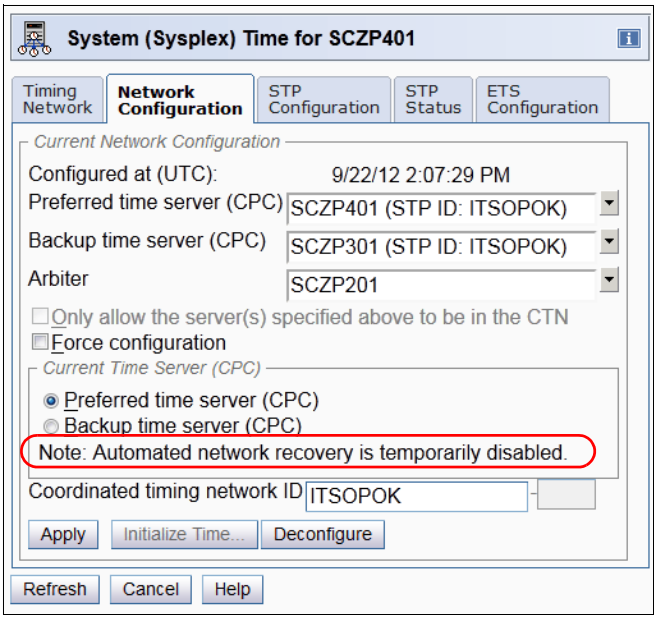

Figure 1-1 Automated network recovery disabled

The hardware message log will explain the reason why recovery has been disabled. For example, when Arbiter-assisted recovery is disabled because both the BTS and Arbiter have lost connectivity to the PTS, the message in Figure 1-2 is issued.

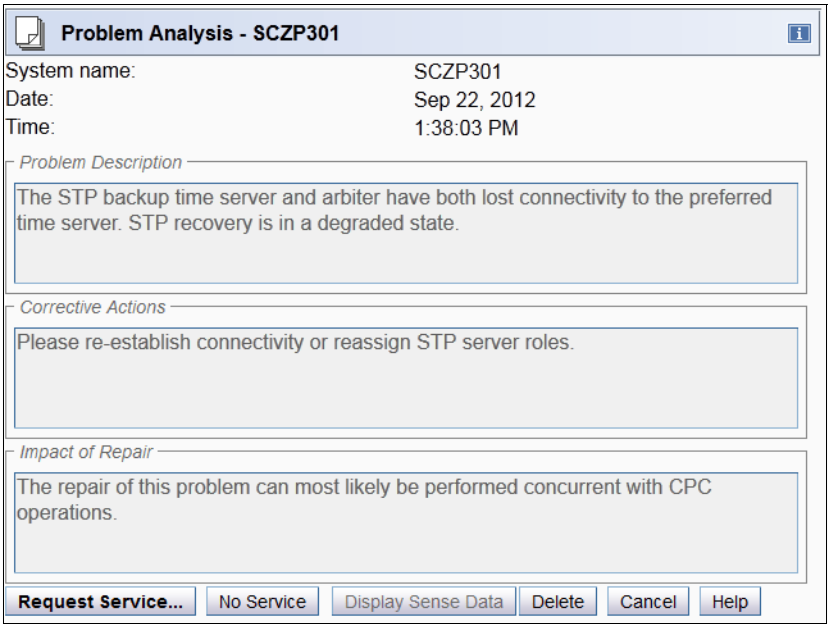

Figure 1-2 Degraded state message

Disabling Arbiter-assisted recovery provides safeguards against:

- Planned disruptive actions done sequentially on the BTS or PTS (whichever is not the CTS) and the Arbiter
- Double failures, unplanned or a combination of planned and unplanned of the BTS or PTS (whichever is not the CTS) and the Arbiter

It does not provide safeguards against planned disruptive actions initiated as part of same HMC task on both the PTS or BTS (whichever is not the CTS) and the Arbiter.

While Arbiter-assisted recovery is disabled:

- The BTS cannot take over as CTS using Arbiter-assisted recovery.
- The CTS will not surrender its role when it loses attachment to the remaining special role server.
- The BTS can still take over as CTS using either Console Assisted Recovery (CAR) or the STP Going Away Signal (GAS) transmitted from the CTS.

Note: After Arbiter-assisted recovery has been disabled, it will not be reenabled until there is full connectivity between the PTS, BTS, and Arbiter.

1.1.7 Console-assisted recovery

Console-assisted recovery uses the HMC and SE LAN in an attempt to determine the status of the PTS (when initiated by the BTS), or the status of the BTS (when initiated by the PTS). Console-assisted recovery helps to determine whether the BTS can take over as CTS, or the PTS can take back its role as a CTS.

Console-assisted recovery in a CTN with BTS

In a CTN that does not have an Arbiter configured, Console-assisted recovery is used to determine whether the BTS should take over the CTS role. The BTS initiates Console-assisted recovery when the BTS has lost communication with the CTS.

If the failure has already been detected through OLS (see 1.1.5, “Going Away Signal” on page 6), the BTS has taken over the CTS role and Console-assisted recovery is used to confirm that the PTS has failed.

If the CTS failure has not yet been recognized through OLS, for example because the failure involved a single link, the BTS will take over if Console-assisted recovery confirms that the CTS has failed.

When the BTS initiates Console-assisted recovery, it sends a command to its Support Element (SE) to determine the state of the CTS by communicating through the HMC.

- If the response indicates that the CTS has failed, the BTS can take over as the new CTS.
- If the response indicates that the status of the CTS is either good or indeterminate, the BTS cannot take over as the new CTS and becomes Stratum 0.

Analysis of the error by Offline signals or Console-assisted recovery is made at the time each process is invoked. In most cases, OLS and Console-assisted recovery are processed almost simultaneously and only one recovery situation is visible to the user. However, if the error conditions change between the time the OLS check is made and Console-assisted recovery is run, the final STP recovery decision will be based on the analysis of conditions at the time Console-assisted recovery is run.

Console-assisted recovery in a CTN with BTS and Arbiter

In a CTN that has an Arbiter configured, Offline signal indications are not used. Console-assisted recovery can be initiated by the BTS or the PTS.

- Console-assisted recovery is initiated by the BTS when it has lost communication with *both* the CTS and the Arbiter and thus cannot initiate the Arbiter-assisted recovery. In this

case, because the Arbiter is configured, although not available, the Offline signals are ignored. Only Console-assisted recovery is used.

The BTS sends a command to its Support Element (SE) to determine the state of the CTS by communicating with the CTS through the HMC.

- If the response indicates that the CTS has failed, the BTS can take over the CTS role.
- If the response indicates that the status of the CTS is either good or indeterminate, the BTS cannot take over as the new CTS.

- Console-assisted recovery is initiated by the PTS when it has lost communication with *both* the BTS and Arbiter.

  The STP-only CTN must only have one CTS, so that the timing integrity of the CTN is not affected. The PTS first surrenders its role as CTS as soon as connectivity to the BTS and Arbiter is lost. Then, it uses Console-assisted recovery in an attempt to determine whether it should take back the CTS role. The PTS sends a command to its Support Element (SE) to determine the state of the BTS by communicating with the BTS through the HMC.

  - If the response indicates that the BTS has failed, the PTS retakes its CTS role.
  - If the response indicates that the BTS has taken over the CTS role, the PTS does not take back the CTS role. The PTS either becomes a Stratum 3 server if a Stratum 2 clock source is available, or becomes unsynchronized.
  - If the response is inconclusive, the PTS does not take back the CTS role. The PTS either becomes a Stratum 3 server if a Stratum 2 clock source is available, or becomes unsynchronized.

Note: The PTS can become a Stratum 3 server before CAR is initiated, if it has a valid clock source to some other server in CTN.

1.1.8 Island condition

An *island* condition occurs when a server detects that one or more servers may be operating as a separate timing network with the same CTN ID, but different PTS, CTS, and Arbiter roles defined.

If an island condition is detected, z/OS issues the following message:

```
IEA391I TWO STP TIMING NETWORKS WITH THE SAME CTN ID HAVE BEEN DETECTED
```

The most common cause of island conditions is a two-site recovery scenario in which connectivity between sites is lost and an incorrect recovery procedure allows both sites to survive, each on its own independent timing network.

1.1.9 Switch to local timing mode

When a server becomes unsynchronized and transitions to a Stratum 0, the resident z/OS systems running in STP timing mode convert to local timing mode.

For z/OS systems running on a server, the impact of switching to a local timing mode depends on the PLEXCFG parameter in IEASYSxx, and ETRMODE or STPMODE specified in CLOCKxx.

- If a system running in either ETR or STP timing mode loses its time source, then:
  - If plexcfg=*xcflocal* or *monoplex*, then the system will continue in LOCAL timing mode. In ETR mode, message `IEA261I NO ETR PORTS ARE USABLE. CPC CONTINUES TO RUN IN`

LOCAL MODE. is issued while in STP mode, message IEA381I THE STP FACILITY IS NOT USABLE. SYSTEM CONTINUES IN LOCAL MODE. is issued.

– If plexcfg=*multisysem* or *any*, the IEA015A is issued in ETR timing mode or IEA394A is issued in STP timing mode. In a Mixed CTN, the S1 z/OS servers issue IEA015A while the S2 z/OS servers issue IEA394A.

- z/OS systems that specify PLEXCFG=MULTISYSTEM *or* PLEXCFG=ANY in IEASYSxx, *and* ETRMODE YES or STPMODE YES in CLOCKxx, issue WTOR message IEA015A or IEA394A to allow operator intervention to resolve the problem before a wait state is loaded.
- z/OS system images that are using a Sysplex timer as their timing source issue WTOR message IEA015A (Example 1-1). The Sysplex timer connectivity needs to be reestablished before a reply with RETRY will be accepted.
- z/OS system images that are using a Stratum 1 or Stratum 2 server as timing source issue WTOR message IEA394A (Example 1-2). Once the CEC is Stratum 1, Stratum 2 or Stratum 3 again, a reply of RETRY will be accepted.

Example 1-1 WTOR message IEA015A

```
IEA015A THIS SYSTEM HAS LOST ALL CONNECTION TO THE SYSPLEX TIMER.

IF THIS EVENT OCCURRED ON SOME, BUT NOT ALL SYSPLEX MEMBERS THE
LIKELY CAUSE IS A LINK FAILURE. TO FIX, ENSURE THAT EACH AFFECTED
SYSTEM HAS AT LEAST ONE CORRECTLY CONNECTED AND FUNCTIONAL LINK.

IF THIS EVENT OCCURRED ON ALL SYSPLEX MEMBERS, THEN THE LIKELY
CAUSE IS A SYSPLEX TIMER FAILURE. TO FIX, REFER TO THE MESSAGE
IEA015A DESCRIPTION IN MVS SYSTEM MESSAGES.

AFTER FIXING THE PROBLEM, REPLY "RETRY" FROM THE SERVICE CONSOLE
(HMC). IF THE PROBLEM WAS NOT CORRECTED, THIS MESSAGE WILL BE
REISSUED AND YOU MAY TRY AGAIN. REPLY "ABORT" TO EXIT MESSAGE
LOOP. PROBABLE RESULT: 0A2-114 WAITSTATE.
```

Example 1-2 WTOR message IEA394A

```
IEA394A THIS SERVER HAS LOST CONNECTION TO ITS SOURCE OF TIME.

IF THIS EVENT OCCURRED ON SOME, BUT NOT ALL NETWORK SERVERS THE
LIKELY CAUSE IS A LINK FAILURE. TO FIX, ENSURE THAT EACH AFFECTED
SERVER HAS AT LEAST ONE CORRECTLY CONNECTED AND FUNCTIONAL LINK.

IF THIS EVENT OCCURRED ON ALL NETWORK SERVERS, THEN THE LIKELY
CAUSE IS A TIMING NETWORK FAILURE. TO FIX, REFER TO THE MESSAGE
IEA394A DESCRIPTION IN MVS SYSTEM MESSAGES.

AFTER FIXING THE PROBLEM, REPLY "RETRY" FROM THE SERVICE CONSOLE
(HMC). IF THE PROBLEM IS NOT CORRECTED, THIS MESSAGE WILL BE
REISSUED AND YOU MAY TRY AGAIN. REPLY "ABORT" TO EXIT MESSAGE
LOOP. PROBABLE RESULT: 0A2-158 WAITSTATE.
```

The WTOR message gives the operator a time window to correct the problem and reply RETRY to reestablish synchronization if the problem has been corrected. The operator can also reply ABORT, thus causing a wait state to be loaded if the problem cannot be rectified.

- If the reply to WTOR message IEA015A is ABORT, wait state 0A2-114 is loaded.
- If the reply to WTOR message IEA394A is ABORT, the wait state depends on the timing mode of the server:

– In ETR timing mode, wait state 0A2-114 is loaded.
– In STP timing mode, wait state 0A2-158 is loaded.

Important: The IEA015A and IEA394A WTOR messages are issued by the Disabled Console Communication Facility (DCCF) *only* to consoles connected to the issuing system. These can be either the sysplex master console, or consoles defined as part of a SYNCDEST group in CONSOLxx.

If no consoles meet these requirements, then DCCF issues the relevant WTOR message to the system console on the HMC. In order to reply to the message via the HMC, the priority message check box must be set.

Note: IEA015A/IEA394A can only be displayed on one console at any time. It is displayed on the first console in the SYNCHDEST console group, and switches to the next console after 125 seconds if no reply is given. It is finally displayed on the HMC where it will remain until a reply is given.

Recovery conditions for a sysplex depend on whether the WTOR messages are issued on all system images in the sysplex or only a partial set, and whether a Sysplex Failure Management (SFM) policy is active (also see 1.2, “Sysplex Failure Management considerations” on page 14).

If the WTOR message is issued by all the z/OS images in the sysplex, you can take your time before replying to WTOR messages IEA0394A or IEA015A, since no z/OS image remains active to monitor the SFM-defined intervals.

Then you have to pay attention to the time interval allowed by the IEA394A and IEA015A messages after the first reply. Once the reply to a WTOR message is answered on the first system image, z/OS will allow a delay of 30 seconds per z/OS image, with a minimum of four minutes, to respond to the last outstanding WTOR message IEA394A or IEA015A. z/OS system images will enter disabled-wait states should you not be able to respond to the IEA394A or IEA015A WTOR message in the allotted time.

If the WTOR message is issued only on a subset of participating sysplex images, the SFM settings specified in the SFM Policy must be considered. The time interval starts at the time the last status update for the system image was made.

1.1.10 External Time Source

In an STP-only CTN, the External Time Source (ETS) function is available using three options:

- Dial out on the HMC
- NTP client support on the Support Element
- NTP client support on the Support Element along with a pulse per second input on the ETR cards

There are no specific recovery actions when the ETS is configured to use a dial-out service. Depending on the ETS configuration, there are two recovery concepts available when NTP with or without PPS is being used:

- NTP Server Availability - Two NTP servers configured for one System z server
- Continuous NTP Server Availability - NTP server configured for both the PTS and the BTS

Note: If STP loses connectivity to all its NTP servers (with or without PPS), all servers in the CTN remain time synchronized. The Coordinated Server Time (CST) may drift away from the NTP time source, until NTP server communication is reestablished.

The following list explains terms that are used in the NTP recovery sections:

- Network Time Protocol (NTP) server

  Provides the capability to keep the Coordinated Server Time (CST) synchronized to the ETS to within 100 milliseconds.

- NTP server with pulse per second (PPS)

  Provides the capability to keep the Coordinated Server Time (CST) synchronized to the ETS to within 10 microseconds. Therefore, a highly stable and accurate PPS output, provided by an NTP server, is utilized.

- Selected and non-selected NTP server

  The user is responsible for selecting the preferred NTP server. The preferred NTP server is called the *selected* NTP server. If two NTP servers have been configured, the second one is called the *non-selected* NTP server.

The following is the order STP selects to steer the CTN when NTP servers with PPS are configured. If one of the items listed is either not configured or has failed, STP will attempt the next item in the list:

1. PPS signals from the selected NTP server configured on the PTS/CTS
2. PPS signals from the non-selected NTP server configured on the PTS/CTS
3. PPS signals from the selected NTP server configured on the BTS
4. PPS signals from the non-selected NTP server configured on the BTS
5. NTP timing information from the selected NTP server configured on the PTS/CTS
6. NTP timing information from the non-selected NTP server configured on the PTS/CTS
7. NTP timing information from the selected NTP server configured on the BTS

Note: In NTP there is also the term stratum. This does not correlate with the STP stratum levels in a CTN.

1.1.11 Coupling channel failure notification

Link failures such as high Bit Error Rates and Interface Control Checks are reported as hardware messages and a Call Home is issued. Coupling channel failure notification is intended to isolate coupling channel failures in the channel subsystem.

1.2 Sysplex Failure Management considerations

The Sysplex Failure Management (SFM) policy allows you to specify failure detection intervals and recovery actions to be initiated in the event of a failure of a system in the sysplex.

Without SFM, when one of the system images in the Parallel Sysplex® fails, you are notified and prompted to take recovery action. You may choose to partition the unresponsive system from the Parallel Sysplex or to take action to try to recover the system. With SFM, an

installation can code a policy to define the recovery actions to be automatically initiated following detection of a Parallel Sysplex failure. Such actions include fencing the failed image that prevents access to shared resources, logical partition deactivation, or dynamic storage reconfiguration.

SFM policy settings can have a significant effect on STP recovery and server availability. SFM makes use of some of the information in the COUPLExx parmlib member. Of particular interest are the INTERVAL, OPNOTIFY, and CLEANUP values. These values reflect recovery-related decisions for the sysplex. To view the current SFM COUPLExx settings, issue the z/OS DISPLAY XCF,COUPLE command, as shown in Figure 1-3.

```
D XCF,COUPLE
IXC357I  15.13.43  DISPLAY XCF 878
SYSTEM SC74 DATA
   INTERVAL   OPNOTIFY     MAXMSG    CLEANUP      RETRY   CLASSLEN
         85         88       2000         30         10        956
```

Figure 1-3 DISPLAY XCF, COUPLE

This output shows that the INTERVAL value is 85. This is the failure detection interval. In a sysplex, each system image periodically updates its own status in the sysplex couple data set, and monitors the status of other system images in the sysplex. If the time elapsed between the last time a system's status was updated and the current time exceeds 85 seconds (the INTERVAL value in Figure 1-3), the system is considered to be in a Status Update Missing (SUM) condition. When this happens, another system in the sysplex will initiate system failure processing according to the SFM policy.

The OPNOTIFY value is the amount of time, following the failure of a system, to update its status before you are notified. XCF on a running system issues WTOR message IXC402D to prompt you after the OPNOTIFY time expires. The OPNOTIFY value must be greater than or equal to the INTERVAL value. When this parameter is omitted, the default is the failure detection interval plus three seconds (88 seconds, as seen in Figure 1-3).

The CLEANUP interval in Figure 1-3 is 30 seconds. This is the amount of time the system waits between notifying sysplex members that this system is terminating and loading a non-restartable wait state. This is the interval permitted for members in the sysplex to perform cleanup processing.

You need to evaluate the trade-off between removing a failing system immediately to prevent disruption to the sysplex, or tolerating a period of non-productivity so that the system and the user have a chance to recover all images in the sysplex.

The INTERVAL, OPNOTIFY, and CLEANUP settings are especially pertinent in recovery scenarios where WTOR message IEA394A (when the STP feature is installed) or IEA015A is issued.

Recovery conditions for a sysplex depend on whether all participating sysplex images are contained within the failed part of the CTN and whether the WTOR messages are issued on all system images in the sysplex or only a partial set.

- If the WTOR message is issued by all z/OS images in the sysplex, you can take your time before replying to WTOR messages IEA0394A or IEA015A, because no z/OS image remains active to monitor the SFM-defined intervals. Then you have to pay attention to the time interval allowed by the IEA394A and IEA015A messages after the first reply.

  Once the reply to a WTOR message is answered on the first system image, z/OS will allow a delay of 30 seconds per z/OS image, with a minimum of four minutes, to respond to the

last outstanding WTOR message IEA394A or IEA015A. z/OS system images will enter the disabled-wait state, should you not be able to respond to the IEA394A or IEA015A WTOR message in the allotted time.

- If the message is issued only on a subset of participating sysplex images, the SFM settings specified in the SFM Policy must be considered. The time interval starts at the time the last status update for the system image was made.

There are other relevant SFM policy settings, such as ISOLATETIME, WEIGHT, and CONNFAIL, which should also be taken into consideration, but are not discussed here.

Note: In GDPS/PPRC configurations running GDPS 3.6 and z/OS 1.11 and later (rolled back to z/OS 1.9 and 1.10 with OA28323 and OA26085) the controlling system is allowed to remain running in an unsynchronized state for a limited amount of time to guarantee a consistent set of secondary PPRC volumes and fulfill its role as the driver of GDPS's recovery actions. For details, see also Timer support for GDPS in *z/OS Version 1 Release 11 Implementation*, SG24-7729.

1.3 Internal Battery Feature considerations

In order to improve STP recovery when power has failed for a single server (PTS or CTS), or when there is a site power outage in a multisite configuration, we suggest that you install the Internal Battery Feature (IBF) on one or more servers in the CTN. If an IBF is installed on your System z server, STP has the capability of receiving notification that client power has failed and that the IBF is engaged. When STP receives this notification from a server that has the role of the PTS or CTS, STP can automatically reassign the role of the CTS to the BTS, thus automating the recovery action and improving availability.

If the entire CTN is located in a single data center, and has only two servers (PTS and BTS), then the recommendation is to install the IBF in both the PTS and the BTS. This should provide recovery capabilities when the server that is the CTS experiences a power failure. If the CTN in a single data center has three or more servers, the recommendation is to assign the Arbiter, in which case the IBF does not provide any additional benefit for server power outages.

If the CTN spans two data centers, we recommend installing the IBF on the servers that will be assigned the roles of PTS, BTS, and Arbiter. This should provide recovery capabilities when the site where the CTS is located experiences a power failure.

If the PTS/CTS has an IBF installed, in case of a power outage it will notify the BTS through the coupling links between the servers (running on IBF power). Because this could be a power glitch, where no immediate action is required, the following recovery action takes place:

- The BTS waits for a period of 30 seconds after notification to take any recovery action.
- If the BTS receives a "normal power" status from the PTS or CTS within 30 seconds, there is no further action.
- If the BTS does not receive a "normal power" status from the PTS or CTS within 30 seconds, the BTS initiates processing to take over the CTS role.
- Automated network recovery is temporarily disabled.
- Because the PTS is still being supplied by the IBF, it stays synchronized as a Stratum 2 server until it powers off.
- When the PTS returns to normal power after the power outage, it resumes the CTS role.

A similar recovery action is performed in case CTN spans multiple sites, and the site that has the PTS or CTS and the Arbiter experiences a power outage and both CTS *and* the Arbiter have the IBF installed.

If the communication between the two sites also fails during this situation (for example, they are connected via DWDM, and the DWDM also fails due to the power outage), IBF will not provide any benefit for STP since the notification of the loss of power cannot be transmitted to the other data center. For a detailed description, refer to 2.4.8, “Two-site considerations” on page 39.

1.4 Multisite considerations

Every installation has a topology, that is, the way servers are arranged and how they communicate with each other. The way that the servers are connected through the actual cables that transmit data, the physical structure of the network, could be called the *physical server topology*. A *logical server topology*, in contrast, is the way that the servers are logically arranged, or the way that the servers behave with one another without regard to the physical interconnection.

In STP, the distinction between the logical and physical design of an STP server configuration is important because both are designed based on two distinct sets of requirements.

The logical STP server design is driven by business requirements. A question to consider might be, based on STP recovery rules, what server role assignments give the system (and business) the best chance of recovering from a failure or disaster. The physical STP server design is driven more by infrastructure and the performance of Parallel Sysplex. Questions to consider might be: how far apart should the servers be placed and what are the cabling requirements?

When planning for STP recovery in multisite server configurations, the logical assignment of server roles plays a greater importance in the survival of the sysplex than the physical placement or distance between servers, given the basic STP recovery rules described in 1.1, “Concepts” on page 4.

2

Recovery scenarios

In this chapter we describe recovery in a Coordinated Timing Network. A review of ETR network recovery is given, followed by recovery for both Mixed and STP-only Coordinated Timing Networks. We also discuss the recovery of an external time source. The following topics are presented:

- Recovery in an ETR network
- Recovery in a Mixed CTN
- Recovery in an STP-only CTN with BTS
- Recovery in an STP-only CTN with BTS and Arbiter
- Restoring STP configuration across PORs in a single- or dual-CEC CTN
- External time source recovery

2.1 Recovery in an ETR network

In this section we describe Sysplex Timer recovery facilities, as well as the reporting and handling of ETR errors by the attached servers and the resident operating systems. Failure recovery scenarios such as link failures, Sysplex Timer failures, and site failures are described in terms of their impact on the z/OS operating system.

2.1.1 Sysplex Timer recovery overview

It is important to understand the recovery design when two Sysplex Timers are configured in an Expanded Availability configuration. The configuration shown in Figure 2-1 illustrates an Expanded Availability configuration with two Sysplex Timers synchronized via CLO links. The Sysplex Timers can be located at different sites in a two-site configuration. Each server has a link to each Sysplex Timer for redundancy.

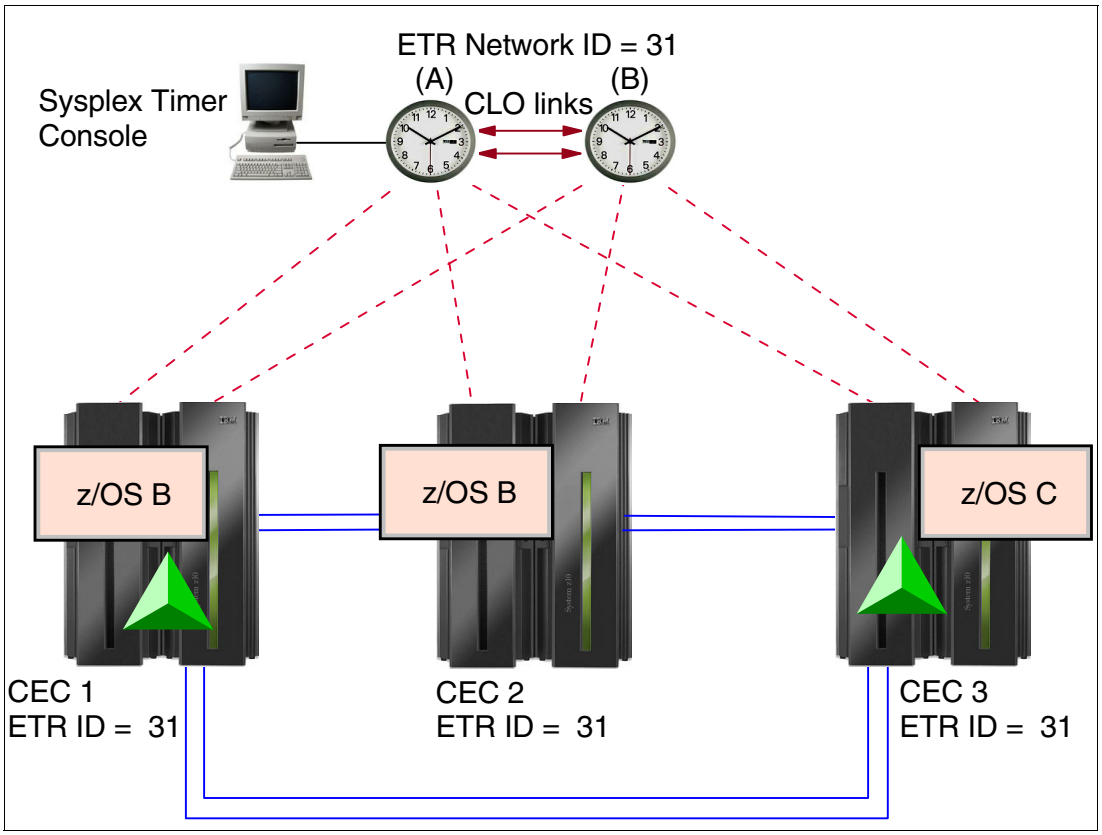

Figure 2-1 ETR network Expanded Availability configuration

In this configuration, both Sysplex Timers are simultaneously transmitting the same time-synchronized data to the attached servers, and no server information is transmitted back to the Sysplex Timers.

The TOD clocks in each server can step to signals received from either server port (0 or 1). For example, the TOD clocks in CEC1 and CEC2 servers could be stepping to signals received from Sysplex Timer (A) while the TOD clocks in server CEC3 could be stepping to signals received from Sysplex Timer (B).

When the server hardware detects the stepping port to be nonoperational, it forces an automatic port switchover, causing the TOD clock to step to signals received from the other

Sysplex Timer, if available. The Sysplex Timers do not switch over, and are unaware of the port change at the server end.

Because the ETR system design allows TODs in each server to step to ETR signals received at either port 0 or port 1, the following Sysplex Timer design rule must be implemented in order to ensure data integrity among the different z/OS images participating in a Parallel Sysplex:

> If at any time two Sysplex Timers in an Expanded Availability configuration lose the capability to synchronize with each other, at least one of the Sysplex Timers must discontinue transmission to the attached servers.

To implement this rule, the Sysplex Timers in an Expanded Availability configuration arbitrate at initialization time to decide the primary and secondary Sysplex Timers. If the Sysplex Timers lose connectivity to each other, only the primary will continue to transmit to the attached servers.

Assigning the primary Sysplex Timer

To ensure that a Sysplex Timer located in a data center is assigned as primary during the arbitration performed at initialization, the following must be implemented:

- If the configuration is using an external time source (ETS):
  - The preferred ETS must be attached to the Sysplex Timer that is to be assigned as the primary Sysplex Timer.
  - When defining the Timer Network to the active console, use the IP address of the Sysplex Timer that will be assigned as the primary Sysplex Timer.
- If the configuration is not using an ETS:
  - When defining the Timer Network to the active console, use the IP address of the Sysplex Timer that will be assigned as the primary Sysplex Timer.

Note: Certain failures after initialization might necessitate a temporary switchover of the primary designation until the failure is repaired.

ETR link failure or server ETR port failure

The server selects which ETR port to use and, if this port loses communication, it will automatically switch to the alternate port. An ETR external interrupt is broadcast to all processors for port availability changes and ETR alert events. If the alternate port is connected to the correct ETR network ID, the system will continue in ETR synchronization mode. Otherwise, the server switches to local TOD stepping mode.

The following message is issued for an ETR port failure:

```
*IEA262I ETR PORT n IS NOT OPERATIONAL
```

After the problem is corrected, the following message is issued:

```
IEA267I ETR PORT n IS NOW AVAILABLE
```

Sysplex Timer port failure

The impact of a Sysplex Timer port failure is similar to an ETR link or server ETR port failure. If access to only one Sysplex Timer port is lost, the following message is issued:

```
*IEA262I ETR PORT n IS NOT OPERATIONAL
```

In this case, the server switches to the alternate ETR port if available. If both Sysplex Timer ports are lost:

- If *plexcfg = xcflocal* or *monoplex*, the following message is issued:

  ```
  *IEA261I NO ETR PORTS ARE USABLE. CPC CONTINUES TO RUN IN LOCAL MODE.
  ```

- If *plexcfg=multisystem*, the IEA015A loss of time source DCCF WTOR is issued instead, and in this situation, z/OS systems running on the server will switch to local timing mode.

Single CLO link failure

The Sysplex Timers maintain synchronization via the remaining active CLO link and continue to transmit time data. There is no impact to the connected servers. However, the following message is issued every eight hours while the problem persists; the reason code indicates which Control Link has failed:

```
*IEA272I ETR SERVICE IS REQUESTED. REASON CODE = nn
```

After the failed link has been recovered, an informational message is issued; the reason code indicates which link is now available:

```
*IEA100I ETR SERVICE INFORMATION IS PROVIDED. REASON CODE = n
```

where `n`:

- RC 77, Control link A failed
- RC 78, Control link B failed

Failure of both CLO links

When communication between Sysplex Timers is lost, the primary unit continues to transmit time data. The secondary unit stops transmitting.

When the ETR network has been implemented as an Expanded Availability configuration where each server has connectivity to both Sysplex Timers, there is no impact because the primary Sysplex Timer will be used for time synchronization. However, in this case the Sysplex Timer is now a single point of failure and the problem needs to be rectified as soon as possible.

Primary Sysplex Timer outage

If the primary Sysplex Timer is removed in a controlled manner, it sends an Offline Sequence to the secondary Sysplex Timer, which becomes the primary Sysplex Timer. No impact should be experienced by any of the attached servers, because they will still retain connectivity to one Sysplex Timer.

However, if the primary Sysplex Timer becomes unavailable without transmitting the Offline Sequence, the secondary Sysplex Timer stops transmitting time data and enters a *dormant* state.

In a multisite configuration, this situation could occur due to failure of the site containing the primary Sysplex Timer. z/OS systems running in a Parallel Sysplex configuration with ETRMODE set to YES in the site containing the secondary Sysplex Timer will issue WTOR message IEA015A. In this scenario, the secondary Sysplex Timer can be recovered and brought back online by reconfiguring the Sysplex Timer configuration from Expanded Availability to Expanded Basic using the procedure documented in *S/390® Time Management and IBM 9037 Sysplex Timer,* SG24-2070. After the secondary Sysplex Timer comes back online and transmits time data to attached servers, you can respond `RETRY` to each WTOR message that was issued.

Also, in case of a total communication failure between sites, the secondary Sysplex Timer will consider that the primary Sysplex Timer has been removed. In this case, the primary Sysplex Timer is still functioning; thus, recovery of the secondary Sysplex Timer should not be performed. In this scenario, the ETR synchronization is lost at the secondary site, and images post a WTOR message. Images at the secondary site will also be partitioned out due to the Sysplex Failure Management (SFM) policy (images running in the primary site) after the failure detection interval (FDI) is reached.

Secondary Sysplex Timer outage

If the secondary Sysplex Timer is removed and an Offline Sequence is received by the primary Sysplex Timer, the primary continues to transmit time data.

This also applies when the primary Sysplex Timer loses connectivity to the secondary Sysplex Timer, whether the Offline Sequence is received or not.

Sysplex Timer synchronization check

Synchronization between the processor TOD clock and the Sysplex Timer is checked every 2^{20} microseconds (1.048576 seconds). If synchronization cannot be performed successfully, a sync-check interrupt is generated. The ETR-sync-check tolerance is model-dependent and is less than the shortest communication delay between any two configurations connecting to the same Sysplex Timer.

If the processor TOD clock can be resynchronized with the Sysplex Timer, recovery is performed. Recovery is successful for multisystem sysplex images if the difference between the processor TOD clock and the Sysplex Timer is less than ETRDELTA.

Recovery for non-multisystem z/OS systems will also be successful, but these may enter local timing mode if the difference between the processor TOD clock and the Sysplex Timer is greater than ETRDELTA (refer to *STP Planning Guide,* SG24-7280 for information on the implications of the ETRDELTA setting).

Recovery from an ETR-sync-check varies, depending on whether the processor TOD clock needs to be moved forward or backward to synchronize with the Sysplex Timer:

- If the processor TOD clock has to be moved forward, the change can proceed immediately.
- If the processor TOD clock has to be moved backward, the CPs will spin until the possibility of duplicated time stamps has passed.

Messages that might be issued to indicate successful recovery from a sync-check include the following:

```
IEA273I TOD CLOCKS DYNAMICALLY ADVANCED TO MAINTAIN ETR SYNCHRONISM.
```

If sync-check recovery is unsuccessful, the operating systems lose access to the Sysplex Timer. For z/OS systems with ETRMODE YES, this results in a switch to local timing mode and issuing of WTOR message IEA015A.

Note: ETR network recovery is covered in detail in *S/390 Time Management and IBM 9037 Sysplex Timer*, SG24-2070.

2.2 Recovery in a Mixed CTN

A Mixed CTN consists of servers that are STP-configured, but the time source is the Sysplex Timer. Since all servers connected to the Sysplex Timer are Stratum 1, there can be multiple Stratum 1 servers in a Mixed CTN. The Stratum 1 servers in a Mixed CTN are in ETR timing mode.

Servers can be indirectly connected to the time source via coupling links as Stratum 2 or Stratum 3 servers. The Stratum 2 and Stratum 3 servers are operating in STP timing mode and are not directly stepping to the Sysplex Timer.

Note: If both Sysplex Timers fail, then the Mixed CTN will no longer have a timing source and all servers will enter local TOD stepping mode.

From the recovery perspective, the Sysplex Timer is the primary time source. Therefore, in a Mixed CTN, the principles of ETR network recovery apply.

Figure 2-2 shows an example of a robust Mixed CTN configuration. It consists of three STP-enabled servers with CTN ID `[ITSOPOK]-[31]`.

There are two Stratum 1 servers (CEC1 and CEC3), and one Stratum 2 server (CEC2). Each server has two coupling links to each of the other two servers. CEC2 has two links to each Stratum 1 server. CEC1 and CEC3 each have links to a Sysplex Timer in an Expanded Availability configuration.

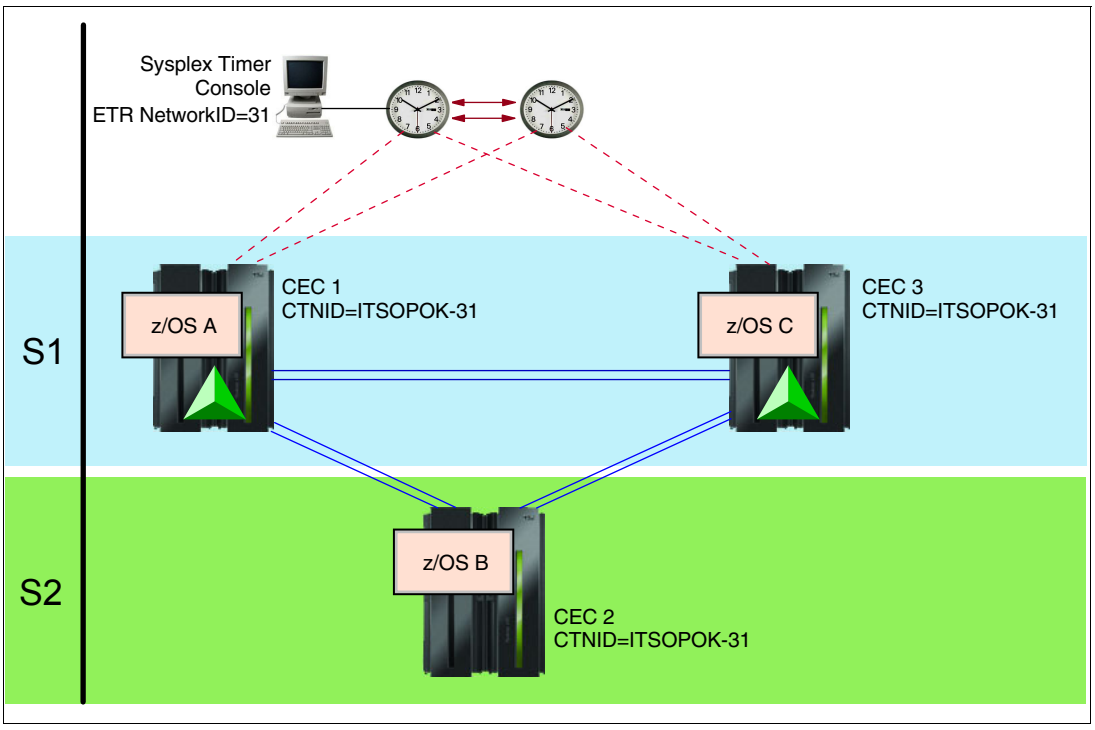

Figure 2-2 Mixed CTN configuration

2.2.1 Single ETR link or port failure

Failure of a single Sysplex Timer link will cause z/OS with ETRMODE set to YES to issue the following message:

```
*IEA262I ETR PORT n IS NOT OPERATIONAL
```

The same message is issued for an ETR port failure, and for a server ETR port failure. If the alternate port on the server is connected to the correct ETR network ID, the system will continue in ETR synchronization mode. Otherwise, full connectivity to the Sysplex Timer is lost and recovery will be as documented in the following section.

2.2.2 Failure of both ETR links or ports

The Sysplex Timer is the time source for a Mixed CTN. The recovery for a failure of both ETR links or ports is similar to recovery in an ETR-only network.

However, complete loss of ETR connectivity of a Stratum 1 server might not result in a server entering local TOD stepping mode (as in an ETR-only network), depending on the configuration: it may be possible for the Stratum 1 server to transition to STP timing mode as a Stratum 2 if required connectivity is in place, and a valid CTN ID is defined.

If STP connectivity is available and a CTN ID has been configured correctly, server transition from ETR to STP, and from Stratum 1 to Stratum 2, has no impact on the resident operating systems or Coupling Facilities. This transition (recovery) applies to z/OS system images that have STPMODE set to YES in CLOCKxx.

As a result, a Mixed CTN has a potentially higher level of recoverability for the Stratum 1 servers than an ETR-only network.

In Figure 2-3 on page 26, if CEC1 stops receiving ETR signals on both its ports, it can become a Stratum 2 server receiving time information via the coupling links from CEC3 Stratum 1. It would also receive STP timing messages from CEC2. However, the STP selection algorithms favor higher stratum levels as time source, so STP messages from a Stratum 1 server would be selected over messages from a Stratum 2.

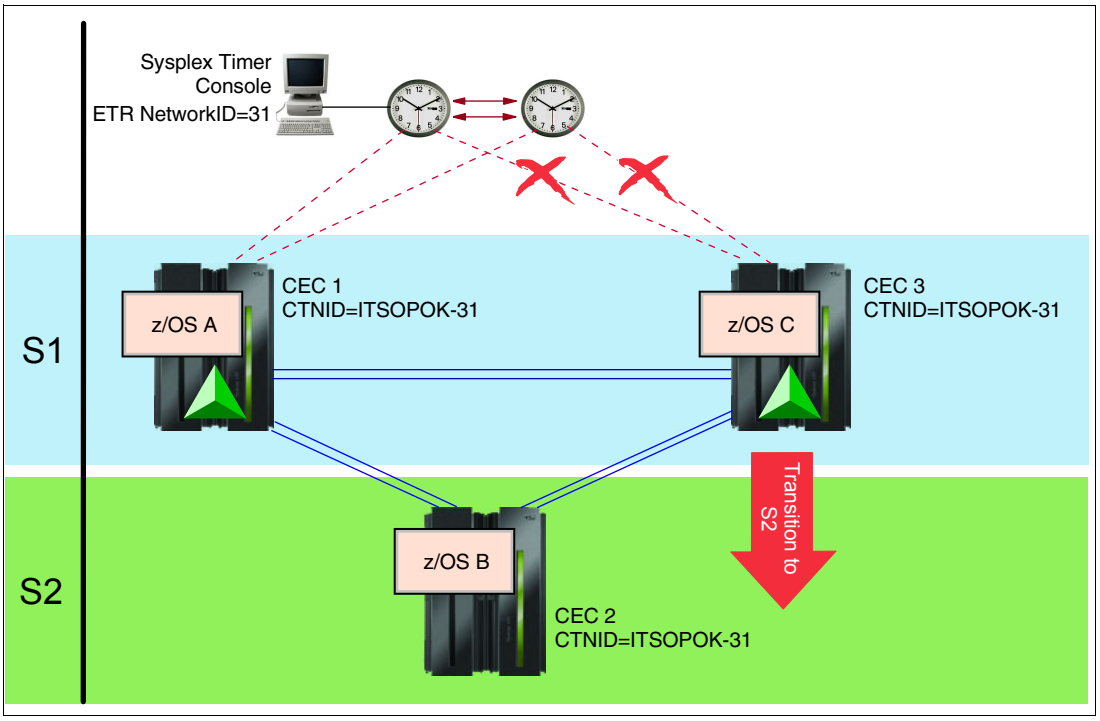

Figure 2-3 Mixed CTN configuration - complete Sysplex Timer link failure

It is important to plan the configuration to provide coupling links to multiple Stratum 1 servers whenever possible. If a server does not have a CF logical partition, timing-only links can be used to provide STP-capable connectivity.

2.2.3 Single coupling link failure

In Figure 2-2, STP messages are being exchanged between the CEC2 Stratum 2 and one of the Stratum 1 servers via the coupling links. There are two coupling links available. However, only one link is selected by the CEC2 for STP timing messages in order to remain synchronized.

If the coupling link being used for synchronization fails, the remaining link is used, if available. In the example used (Figure 2-2), failure of one coupling link between the CEC2 and CEC1 would leave only one link remaining between these servers.

The following z/OS warning message will be issued should this occur:

```
IEA382I THIS SERVER HAS ONLY A SINGLE LINK AVAILABLE FOR TIMING PURPOSES
```

The message is issued when there is only one remaining physical link (PCHID) available for timing purposes between two servers.

2.2.4 Last coupling link failure

In the (unlikely) scenario where the last coupling link between a server and its time source fails, STP timing can recover if there is a coupling link to an alternate server available that does not involve transitioning the server to Stratum 4, which is not supported.

In Figure 2-4, if both coupling links between the CEC2 and the CEC3 fail, there is no impact from an STP timing point of view because CEC2 is still able to use the CEC1 as a valid time source.

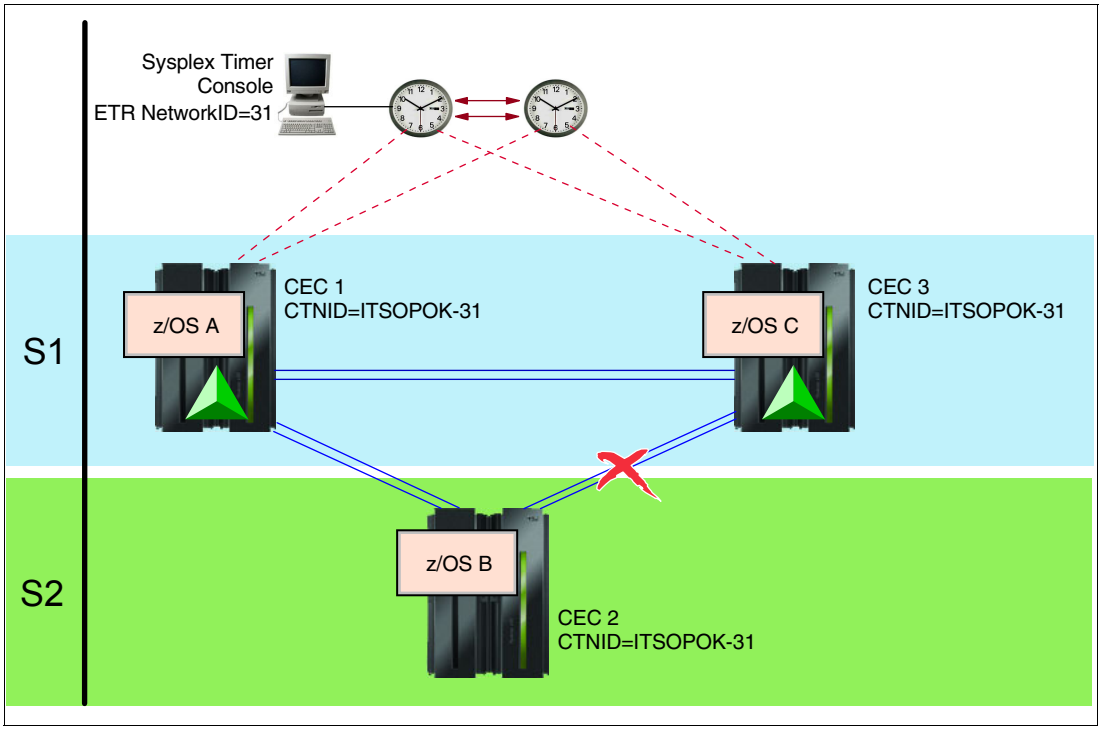

Figure 2-4 Mixed CTN configuration - complete coupling link failure

However, complete failure of the coupling links between servers may have other implications: For example, when the z/OS host's access to a Coupling Facility has been removed and a structure rebuild process is required to recover the inaccessible structures to another Coupling Facility.

2.2.5 Stratum 1 server failure

In a Mixed CTN, at least two Stratum 1 servers are suggested in order to avoid a single point of failure for Stratum 2 servers. Even if a Stratum 2 server is connected to multiple Stratum 1 servers, only one is selected as the time source.

In Figure 2-5 on page 28, CEC2 is a Stratum 2 receiving STP messages from CEC1 and CEC3, and has selected CEC3 as the time source. If CEC3 fails, the CEC2 Stratum 2 selects CEC1 as its time source and continues processing without disruption.

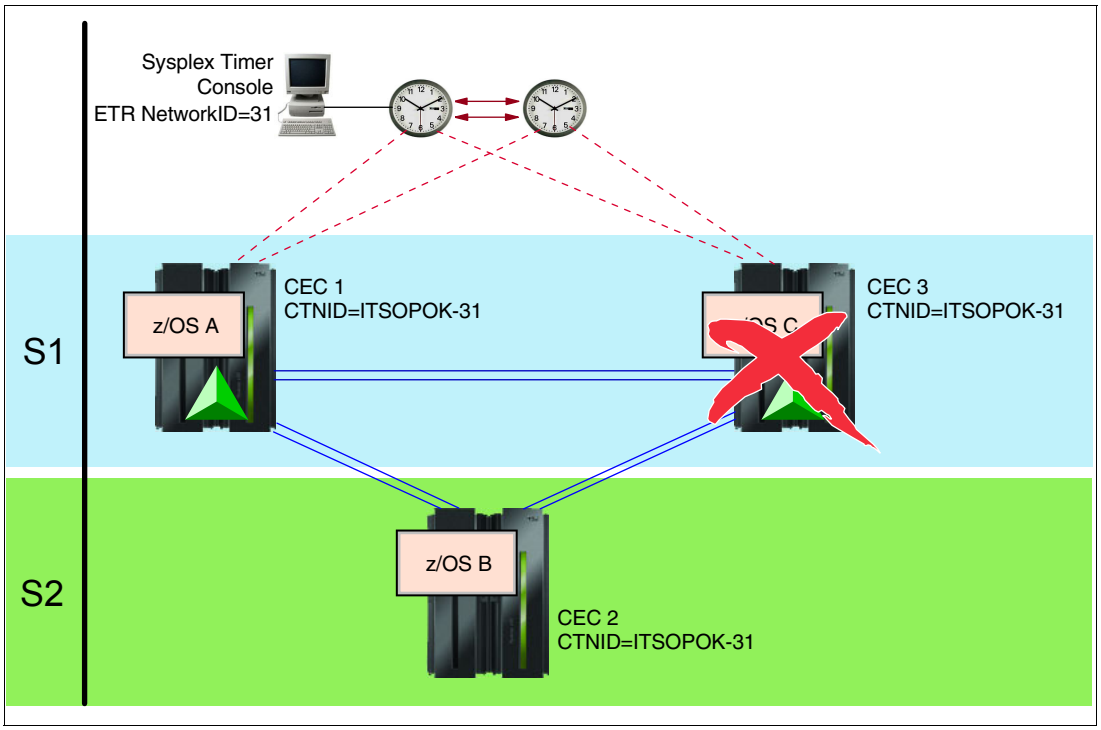

Figure 2-5 Mixed CTN Stratum 1 failure

However, this would leave CEC2 totally dependent on the CEC1 for STP timing messages, and any removal of CEC1 for planned maintenance or possible failure would cause CEC2 to lose its only available timing source.

Suggestion: In a Mixed CTN, note the following:

- There should be at least two Stratum 1 servers connected to Sysplex Timers in an Expanded Availability configuration.
- Each Stratum 2 should be connected to each Stratum 1 with at least two coupling links.

2.2.6 Stratum 2 failure

In Mixed CTN configurations that contain Stratum 3 servers, it is important to have at least two links from each Stratum 3 server to at least two Stratum 2 servers. Connecting two Stratum 3 servers will provide no benefit from an STP timing redundancy perspective because the STP architecture does not allow a server to transition to Stratum 4.

Suggestion: In a Mixed CTN, note the following:

- There should be at least two Stratum 2 servers before any Stratum 3 servers are configured.
- Each Stratum 3 should be connected to each Stratum 2 with at least two coupling links.

2.2.7 Two-site considerations

Figure 2-6 shows a two-site, Mixed CTN configuration. There are two Stratum 1 servers attached to the Sysplex Timers at Site 1. There is one server at Site 2 (100 kilometers away), with coupling links to both Stratum 1 servers in Site 1.

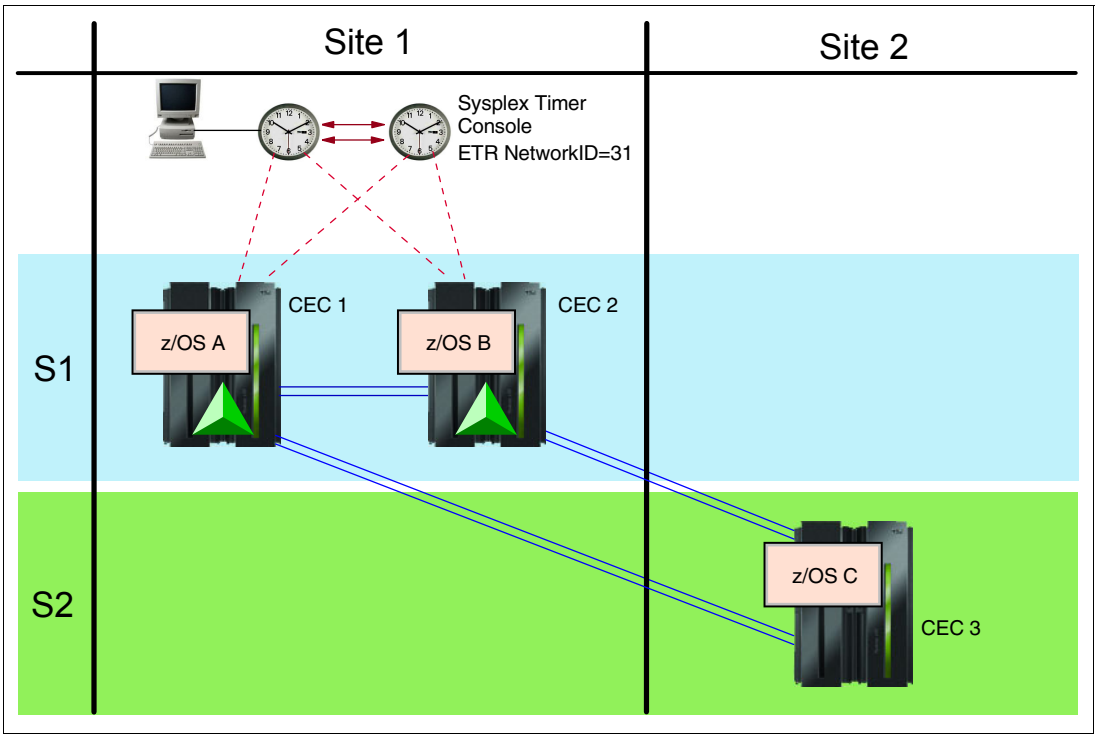

Figure 2-6 Two-site Mixed CTN

The Mixed CTN failure recovery in scenarios previously discussed also applies to a two-site configuration. However, there is an additional recovery scenario involving the complete loss of communication between Site 1 and Site 2.

Site 1 failure

A Site 1 failure causes the server in Site 2 to lose its time source and become unsynchronized at the end of the freewheel interval. z/OS images that have STPMODE set to YES will issue WTOR message IEA394A.

If the time source in Site 1 cannot be recovered, then the operator must reply `ABORT` to the WTOR messages. To recover in such a situation, servers at Site 2 can be reconfigured into an STP-only CTN, if required.

Site 2 failure

A Site 2 failure has no impact on servers located at Site 1.

Loss of communication between Site 1 and Site 2

The Site 1 servers are not affected. However, for the server in Site 2, a complete loss of communication between sites appears like a Site 1 failure and the implications are the same as described previously in “Site 1 failure”.

2.3 Recovery in an STP-only CTN with BTS

Recovery in an STP-only CTN is very different than for a Mixed CTN. The reason for this relates to the assigned roles of PTS, BTS, and in particular Arbiter.

In this section, we discuss recovery scenarios for an STP-only CTN with a single Backup Time Server and no Arbiter assigned. The scenario diagram is shown in Figure 2-7. Note that there are only two servers configured. An Arbiter is recommended for any configuration with more than two servers; therefore, it is not considered in this discussion.

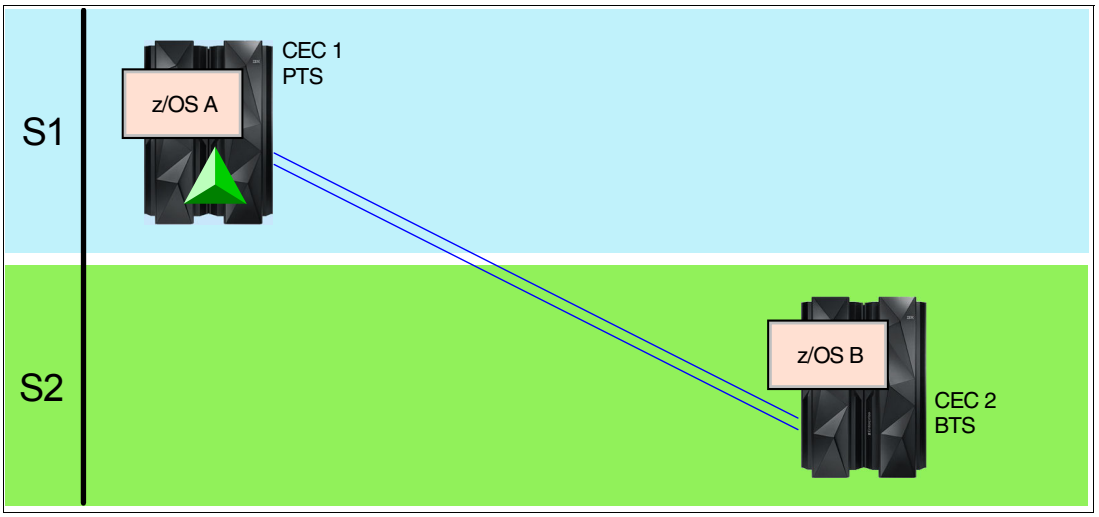

Figure 2-7 STP-only CTN with PTS and BTS assigned

The example configuration has one server configured as Preferred Time Server (PTS) and one as Backup Time Server (BTS). The servers are connected via two coupling links. CEC1 is the PTS and Current Time Server (CTS - Stratum 1), while CEC2 is the BTS (Stratum 2).

Note: The recovery actions described in the following sections assume that the PTS has been assigned the role of Current Time Server. This is the recommended configuration because only the PTS can automatically take back the CTS role after a recovery action.

In cases where the BTS is the Current Time Server, the same tasks described for the BTS are performed by the PTS, if this becomes available.

2.3.1 Current Time Server failure

A failure of the CTS is detected by the BTS when an Offline signal (OLS) is received on all established STP paths to the CTS. The conditions under which an Offline signal is transmitted are described in 1.1.5, “Going Away Signal” on page 6.

Since only one CTS can exist in an STP-only CTN, the PTS has to surrender its CTS role, when OLS is transmitted on all established paths.

When this occurs, the BTS takes over the CTS role and becomes the new Stratum 1 for the CTN. In Figure 2-7, if CEC1 failed with an Offline signal being sent, CEC2 would assume the CTS role and survive. Afterwards, the new CTS uses CAR to assure that the PTS really has surrendered its role as CTS. Only if this check indicates that the PTS is still up and Stratum 1, the BTS surrenders the role as CTS.

If the BTS does not receive the Offline signal on all established links, the loss of STP timing messages causes it to freewheel and initiate Console-assisted recovery as described in 1.1.7, “Console-assisted recovery” on page 10. If Console-assisted recovery is able to determine that the CTS has failed, the BTS takes over the role of CTS, and becomes the new Stratum 1. Otherwise, if Console-assisted recovery is unable to determine the status of the CTS, then synchronization is lost and CEC2 transitions to Stratum 0.

There is one recovery mechanism that overrules the *Server Offline Signal (OLS)* recovery as well as the *Console Assisted Recovery (CAR).* If the CTS enters a *checkstop condition* and the coupling connectivity between the CTS and BTS consists of *Host Communication Adapter 3 cards (HCA3-O)* online on both ends, the CTS sends the Going Away Signal to the BTS. Once the BTS receives the Going Away Signal, the BTS takes over as CTS independent of the results of the Server Offline Signal or the Console Assisted Recovery. The Going Away Signal can be considered the highest recovery mechanism in place, which removes any dependencies from the OLS and CAR recovery. The conditions under which the Going Away Signal is transmitted are described in 1.1.5, “Going Away Signal” on page 6.

2.3.2 Current Time Server power outage: Internal Battery Feature (IBF)

As mentioned in *STP Planning Guide,* SG24-7280, if an Internal Battery Feature (IBF) is installed on the CTS, STP now has the capability of receiving notification that client power has failed and that the IBF is engaged.

The CEC1 server shown in Figure 2-7 on page 30 is the PTS/CTS and has an IBF installed. If CEC1 experiences a power outage and its IBF is engaged, it will notify the CEC2 (BTS) via the coupling links between the servers that it is running on IBF power. The BTS waits for a period of 30 seconds after receiving notification before taking any recovery action, in case the power outage condition is just a transient power outage or glitch (“brown-out”). In case power is restored, the CEC1 server will notify CEC2 that it has returned to “normal power” status.

If the BTS receives a “normal power” status from the PTS within 30 seconds, no further action is taken. However, if the BTS does not receive a “normal power” status from the PTS within 30 seconds, the BTS will initiate takeover processing for the CTS role. As the PTS may still be running on the IBF, it will stay synchronized as a Stratum 2 server until it powers off. When CEC1 returns to normal power after the outage, it automatically resumes the CTS role.

We recommend to also install the IBF in the BTS to provide additional recovery protection when the BTS has the role of CTS.

2.3.3 Backup Time Server failure

Failure of the BTS does not affect the CTN. The PTS (CEC1) will continue to operate as the CTS (STP Stratum 1).

2.3.4 Single coupling link failure

When one of the coupling links selected for timing messages loses communication, another coupling link is selected, if available, and STP timing continues without impact to the CTN.

In the configuration depicted in Figure 2-7 on page 30, if the coupling link used to communicate between CEC1 and CEC2 fails, the remaining link will be used. Because there are two STP links between the servers, only one link will remain available and a warning message indicating a single point of failure is issued:

```
IEA382I THIS SERVER HAS ONLY A SINGLE LINK AVAILABLE FOR TIMING PURPOSES
```

This message is issued when there is only one remaining physical link (PCHID) available for timing purposes between two servers.

2.3.5 Last coupling link failure

Failure of the last coupling link between the CTS and the BTS will appear to each server as loss of communication between the servers. In this case the BTS invokes Console-assisted recovery (CAR). If CAR confirms that the PTS is still CTS and running, or if CAR cannot determine the PTS status, the BTS becomes unsynchronized and transitions to Stratum 0, and the CTS continues to function with no impact. If CAR shows that the PTS no longer acts as CTS, the BTS takes over the role and becomes CTS.

2.3.6 Two-site considerations

Figure 2-8 shows a two-site configuration with the PTS at Site 1 and the BTS at Site 2.

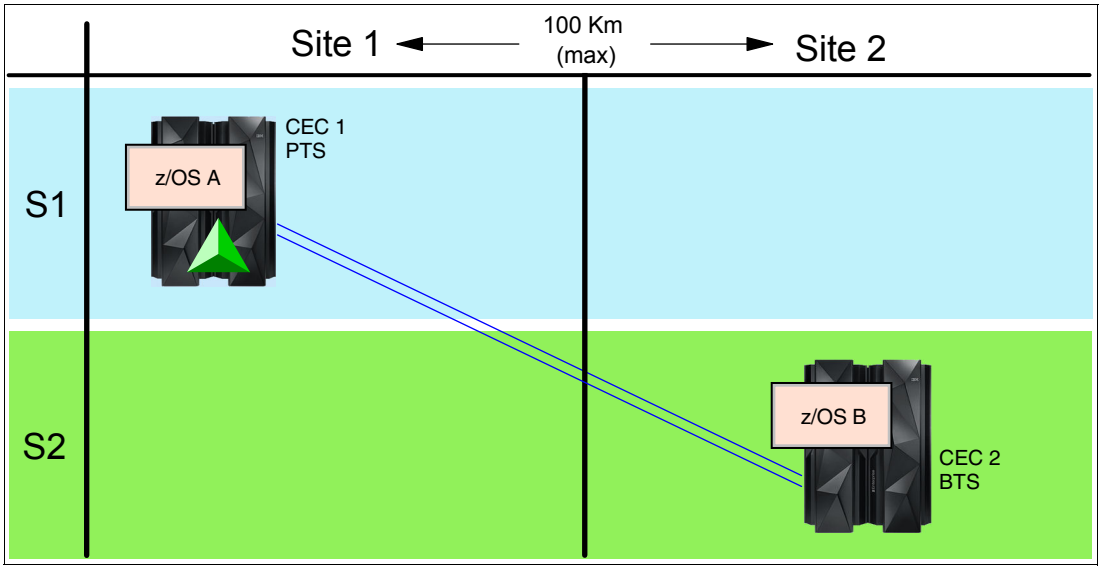

Figure 2-8 Two-site STP-only CTN with PTS and BTS

The server failure scenarios are the same as for the single site configurations previously discussed.

However, in a two-site configuration, a complete site failure can also disable inter-site communication. This makes recovery more complex because the Offline signal may not be transmitted by the CTS nor communicated to the BTS. Also, Console-assisted recovery may not be possible between sites because the HMC network has also failed.

Site 1 failure

If Site 1 fails, the Offline signal will most likely not be received by the BTS, and Console-assisted recovery will find the state of the CTS as indeterminate. The BTS at Site 2 will freewheel until expiration of the freewheel interval, and then switch to STP Stratum 0.

The z/OS systems defined with STPMODE set to YES running on the CEC1 issue WTOR message IEA394A.

Note: For GDPS 3.6 implementations running z/OS 1.11 and later (rolled back to z/OS 1.9 and 1.10 with OA28323 and OA26085), the GDPS controlling system located at Site 2 is allowed to remain unsynchronized for a limited amount of time before it issues message IEA394A via DCCF. This way it will be able to fulfill its role as the driver of GDPS's recovery actions.

Site 1 power outage – CTS at Site 1 with IBF

If the CTN spans two data centers, we suggest to install the IBF on the servers that will be assigned the roles of PTS, BTS, and Arbiter. This provides recovery capabilities when the site where the CTS is located experiences a power failure. In the example shown in Figure 2-8 on page 32, we assume that the IBF is installed at least on CEC1. We suggest that the IBF also be installed on the BTS in order to provide recovery capability when the BTS has the CTS role.

The recovery works the same way as described in 2.3.2, "Current Time Server power outage: Internal Battery Feature (IBF)" on page 31. If Site 1 experiences a power outage, and the IBF is engaged on the CEC1, this notifies the CEC2 (BTS) via the coupling links between the servers that it is running on IBF power. The BTS waits for a period of 30 seconds after notification to take any recovery action, in case the power outage condition is just a transient power outage or glitch. If it is a transient power condition, CEC1 will notify CEC2 that it has now returned to "normal power" status. If the BTS receives a normal power status from the PTS within 30 seconds, there is no further action. However, if the BTS does not receive a normal power status within 30 seconds, the BTS initiates processing to take over the CTS role. Because the PTS might still be running on IBF power, it will stay synchronized as a Stratum 2 server until it powers off. When CEC1 returns to normal power after the outage, it automatically resumes the CTS role.

Note: The timing links between the CTS and BTS need to remain functional until the BTS has taken over the CTS role. If the links fail before the takeover (for example, because the two sites are connected via DWDM that also fails due to the power outage), the failure will be treated as a loss of communication between the two sites.

Installing the IBF on the PTS allows the BTS to successfully take over as the CTS, as compared to the scenario described in "Site 1 failure" on page 32, where the site failure was not a power outage, and the BTS became unsychronized.

Site 2 failure

A Site 2 failure has no impact on the CTS at Site 1. CEC1 will continue in the CTS role.

Loss of communication between sites

For the BTS in Site 2, this looks like a Site 1 failure even though the CTS is still active in Site 1. Cross-site communication must be restored before replying `RETRY` to the WTOR message IEA394A. Replying `ABORT` to WTOR message IEA394A will result in z/OS systems on the BTS (CEC2) loading a non-restartable wait state 0A2-158. This also looks like a Site 2 failure to the Site 1 server. CEC1 at Site 1 is not affected.

2.4 Recovery in an STP-only CTN with BTS and Arbiter

In this section we discuss recovery scenarios for an STP-only CTN with an Arbiter assigned.

The diagram in Figure 2-9 is used in this scenario. A PTS, BTS, and Arbiter are assigned. Each server has coupling links connected to the other two servers. CEC1 is the Stratum 1, and CEC2 and CEC3 are both Stratum 2.

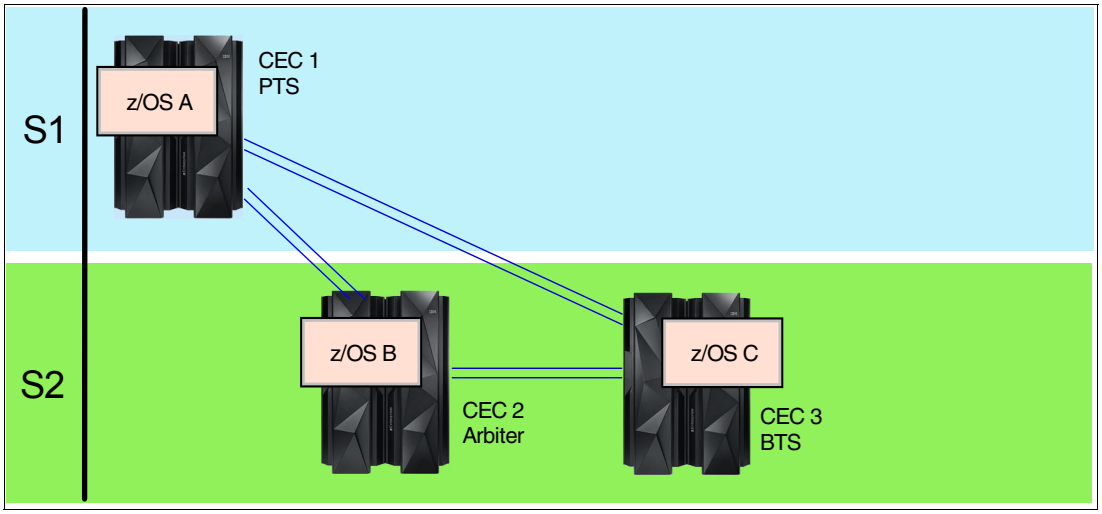

Figure 2-9 STP-only CTN with PTS, BTS, and Arbiter assigned

Configuration of an Arbiter is suggested whenever there are three or more servers in the STP-only CTN. With an Arbiter available, the Arbiter-assisted recovery process becomes the key mechanism for the BTS to determine whether it should take over the role of CTS when communication to the CTS is lost. Refer to 1.1.6, “Arbiter-assisted recovery” on page 7. The only recovery mechanism that overrules the Arbiter-assisted recovery is the Going Away Signal (only available when using HCA3-O cards on both ends of the link). For details refer to 1.1.5, “Going Away Signal” on page 6.

When an Arbiter is configured:

- The Offline signal recovery process is not invoked.
- The BTS only invokes Console-assisted recovery if it cannot communicate with the Arbiter.
- The PTS will also invoke Console-assisted recovery when it loses communication with both the BTS and the Arbiter.
- The Going away signal (initiated by the CTS once it enters a checkstop condition) received at the BTS has highest priority and overrules the Arbiter-assisted recovery as well as the Console-assisted recovery.

Note: The Going Away Signal recovery mechanism is only available if the PTS to BTS connectivity consists of at least one online InfiniBand link based on the Host Communication Adapter 3 (HCA3-O) card on both ends of the link.

Including an Arbiter in the configuration provides a more robust recovery mechanism, because Arbiter-assisted recovery can handle recovery scenarios when an Offline signal cannot be transmitted, such as a channel subsystem failure.

2.4.1 Current Time Server failure

If the CTS fails, the BTS communicates with the Arbiter to determine whether the Arbiter has also lost communication with the CTS. If the Arbiter has also lost communication with the

CTS, the BTS assumes the CTS role. Since only one CTS can exist in an STP-only CTN, the PTS has to surrender its role as the CTS.

There is one recovery mechanism that overrules the Arbiter-assisted recovery. If the CTS enters a checkstop condition and the coupling connectivity between the CTS and BTS consists of Host Communication Adapter 3 cards (HCA3-O) on both ends, the CTS sends the Going Away Signal to the BTS. Once the BTS receives the Going Away Signal, the BTS takes over as CTS independent of the results of the Arbiter-assisted recovery. The Going Away Signal can be considered the highest recovery mechanism in place, which removes any dependencies from the Arbiter-assisted recovery and the CAR recovery. The conditions under which the Going Away Signal is transmitted are described in 1.1.5, “Going Away Signal” on page 6.

Important: Having only an Arbiter and Current Time Server available provides no STP time synchronization recovery protection for the CTN.

A CTN without an operational PTS cannot perform a CTS switch, and the Arbiter-assisted recovery is therefore temporarily disabled. The CTN is considered to be in a partially degraded status.

The CTN should be reconfigured as soon as possible to ensure that at least a Current Time Server and Backup Time Server are available. If this is not performed and the BTS that took over the role of CTS also fails, the entire CTN would become unsynchronized.

In the configuration shown in Figure 2-9 on page 34, CEC3 assumes the role of CTS and becomes a Stratum 1 server when CEC1 fails. This leaves the CTN without a Backup Time Server until CEC1 returns to the configuration. If extended unavailability of the PTS is expected and there are additional servers in the CTN, it is suggested that the Arbiter assignment be removed and CEC2 be assigned the PTS role.

Note: If the CTN in a single data center has three or more servers, the recommendation is to assign the Arbiter, in which case the IBF does not provide any additional benefit for server power outages. Arbiter-assisted recovery is used as described previously for a power outage that affects the CTS (CEC1 in Figure 2-9 on page 34).

2.4.2 Backup Time Server failure

If the BTS fails, z/OS systems on BTS post WTOR messages. No further recovery actions are taken. The CTN continues without a Backup Time Server and is therefore subject to a single point of failure until the BTS becomes operational, or the CTN is reconfigured to define a new BTS.

A CTN without an operational BTS cannot perform a CTS switch, and the Arbiter-assisted recovery is therefore temporarily disabled. The CTN is considered to be in a partially degraded status.

The CTN should be reconfigured as soon as possible to ensure that at least a Current Time Server and Backup Time Server are available. In a configuration where another server is available, this should be assigned the BTS role. In the configuration illustrated in Figure 2-9 on page 34, we suggest that the Arbiter assignment be removed and CEC2 be assigned the BTS role.

2.4.3 Arbiter failure

If the Arbiter fails, no further recovery actions are taken. The CTN will continue to operate with only a CTS and a BTS. Since a CTN without an operational Arbiter cannot get assistance from the Arbiter in the case the PTS to BTS communication gets lost, the Arbiter-assisted recovery gets temporarily disabled. The CTN is considered to be in a partially degraded status.

However, the Arbiter is still known to the configuration, even though it is no longer available, causing subsequent recovery actions to ignore the Offline signal rules. This leaves only Console-assisted recovery available for the CEC1 and CEC3, should communication links between the two servers be lost. Also, since the Arbiter has been lost, the Arbiter-assisted recovery is being temporarily disabled, since a subsequent BTS failure should not depict the PTS to be isolated, further causing it to become unsynchronized. The CTN is considered to be in a partially degraded status.

If the Arbiter is expected to be nonoperational for an extended period of time, it is suggested that the Arbiter assignment be removed, so that the BTS can use the OLS rules.

2.4.4 Single coupling link failure

Failure of a single STP link between any of the servers in the configuration will not have any impact on the CTN, because a redundant link is available. If only one link is left, the following message is displayed:

```
IEA382I THIS SERVER HAS ONLY A SINGLE LINK AVAILABLE FOR TIMING PURPOSES.
```

2.4.5 All coupling links fail between the CTS and BTS

Figure 2-10 shows a scenario where there is a complete coupling link failure between the PTS and BTS.

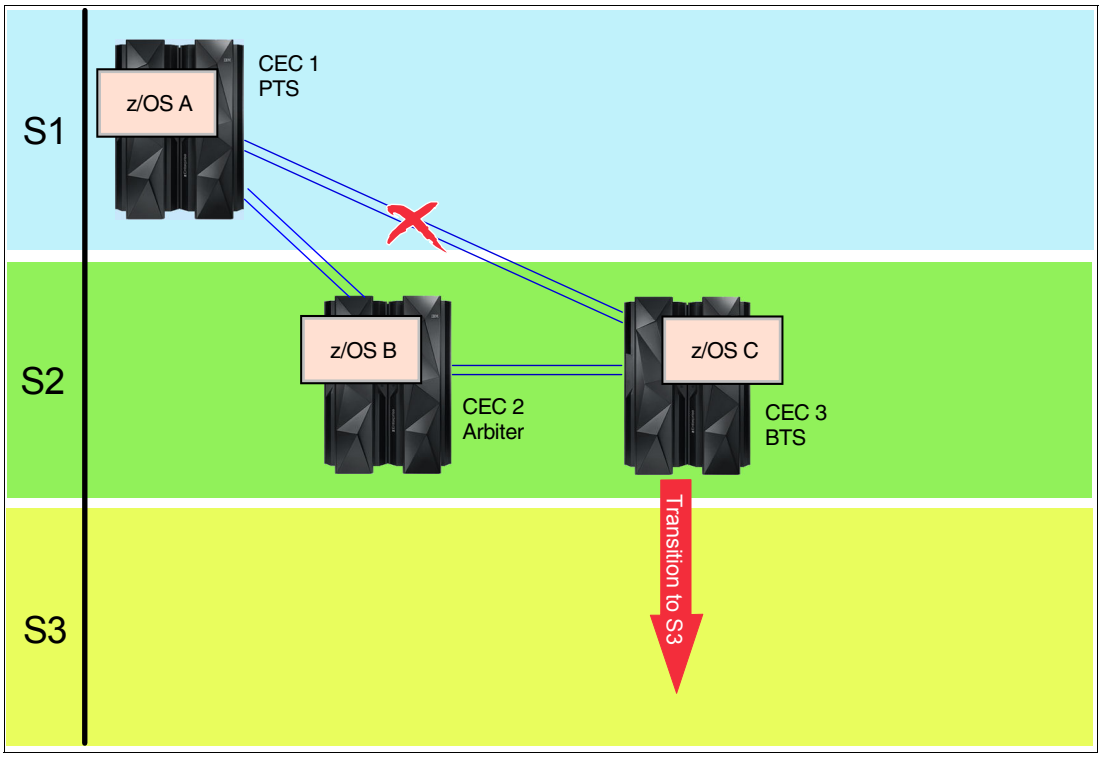

Figure 2-10 STP-only CTN - complete link failure between PTS/CTS and BTS

The BTS will cease to receive STP timing messages from the CTS and will contact the Arbiter to determine whether the Arbiter can communicate with the CTS. Because the Arbiter is still in communication with the CTS, the BTS cannot take over the CTS role, and therefore must transition to a Stratum 3 in order to remain in the STP-only CTN.

Even though the BTS no longer has direct connectivity with the CTS, it is still able to function as BTS. The Arbiter takes note that the BTS no longer has connectivity to the CTS and, should it subsequently lose contact with the CTS, the Arbiter will inform the BTS accordingly, causing the BTS to proceed with taking over as the CTS.

Alternatively, if connectivity is subsequently reinstated between the CTS and the BTS, the BTS will transition back to a Stratum 2 and the Arbiter will be informed of its availability.

In this example, CEC3 would transition from a Stratum 2 to a Stratum 3 while the connectivity to CEC1 is unavailable. However, the CEC3 would still be able to function as BTS through communication with the Arbiter.

2.4.6 Last coupling link failure between the CTS and Arbiter

It is possible to lose all connectivity between the CTS and the Arbiter. This is depicted in Figure 2-11.

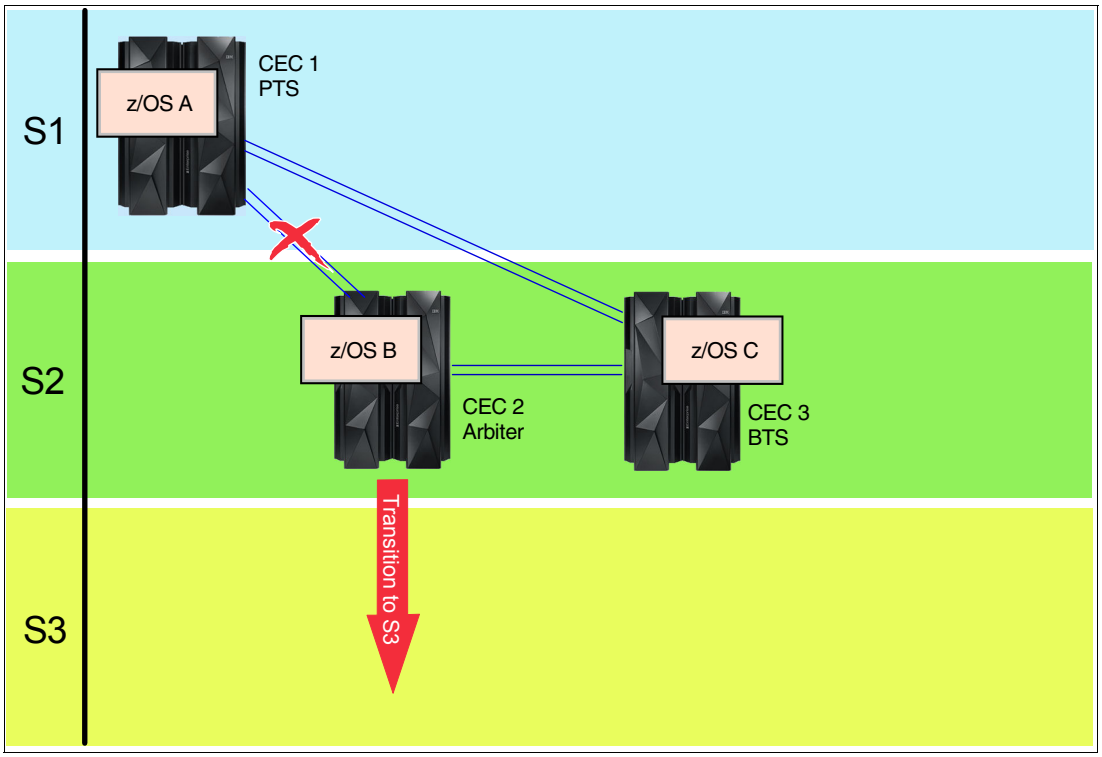

Figure 2-11 STP-only CTN - complete link failure between PTS/CTS and Arbiter

Because the Arbiter no longer has direct connectivity to the CTS, it needs to transition to a Stratum 3 to continue to receive STP timing signals and remain in the STP-only CTN.

In this state the Arbiter is still able to communicate with the BTS and therefore is still available for the BTS to use, should the BTS subsequently lose contact with the CTS.

When communication is reinstated between the CTS and the Arbiter, the Arbiter will transition back to a Stratum 2 due to the clock selection algorithms placing priority on higher stratum levels for STP link selection.

2.4.7 All coupling links fail between the BTS and Arbiter

For completeness, we also describe recovery when BTS loses communications with the Arbiter, as shown in Figure 2-12 on page 39.

A loss of the coupling connectivity between BTS and Arbiter has no impact on the CTN, as long as no other failure occurs. But this makes the Arbiter unavailable for the BTS to initiate Arbiter-assisted recovery should communication with the CTS be lost, requiring the BTS to use Console-assisted recovery to determine CTS status.

Should a subsequent CTS failure occur, and the Console-assisted recovery results are inconclusive, the BTS will be unable to take over the CTS role. The BTS and Arbiter will become unsynchronized and transition to Stratum 0. All z/OS systems resident on the BTS

and the Arbiter will issue a WTOR IEA394A message or continue operating in local timing mode (see 1.1.9, “Switch to local timing mode” on page 11).

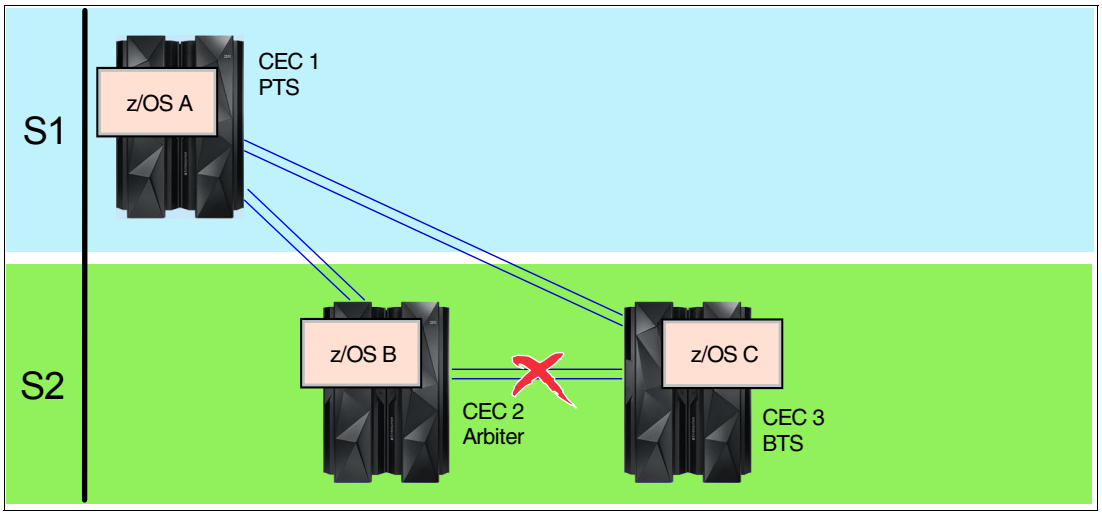

Figure 2-12 STP-only CTN - complete link failure between BTS and Arbiter

2.4.8 Two-site considerations

When the STP-only CTN spans multiple sites, the location of the Arbiter becomes critical.

Suggestion: In a non-GDPS implementation the suggestion is to have CTS and Arbiter at Site 1.

In a GDPS implementation in an STP-only CTN the suggestion for the Arbiter configuration depends on the freeze policy settings:

- If your freeze policy setting is FREEZE=GO or FREEZE=SWAP,GO, the advice is to configure both the PTS and the Arbiter in the same site as the primary disks. With z/OS 1.11 and later (rolled back to z/OS 1.9 and 1.10) the GDPS controlling system will be allowed to remain running in an unsynchronized state for a limited amount of time. In the event of a Site 1 failure, the GDPS controlling system in Site 2 will be able to guarantee a consistent set of secondary PPRC volumes and fulfill its role as the driver of GDPS recovery actions.
- If the freeze policy is FREEZE=STOP or FREEZE=SWAP,STOP, configure both the BTS and the Arbiter at the same site as the secondary disks.

Site 1 failure - CTS and Arbiter at Site 1, BTS at Site 2

This is the suggested configuration, shown in Figure 2-13 on page 40.

Both the CTS and Arbiter are lost, because they are located in Site 1. The systems located in Site 2 survive the failure, but have lost the STP time source and therefore the BTS needs to go through a recovery scenario to determine if it is able to assume the role of CTS.

The recovery involves the following steps:

1. The BTS loses all communication with the CTS.
2. The BTS tries to communicate with the Arbiter to establish whether the Arbiter can communicate with the CTS.

3. The BTS cannot communicate with the Arbiter and initiates Console-assisted recovery to attempt to determine the state of the CTS.
4. The results of Console-assisted recovery are most likely indeterminate, so the BTS cannot take over the CTS role.
5. The BTS eventually becomes unsynchronized and changes to Stratum 0, as well as all other CECs on Site 2. z/OS systems resident at Site 2 with STPMODE set to YES issue the WTOR IEA394A message.
6. Manual action via the HMC is required to reassign the role of CTS to the BTS before replying `RETRY` to WTOR IEA394A messages.

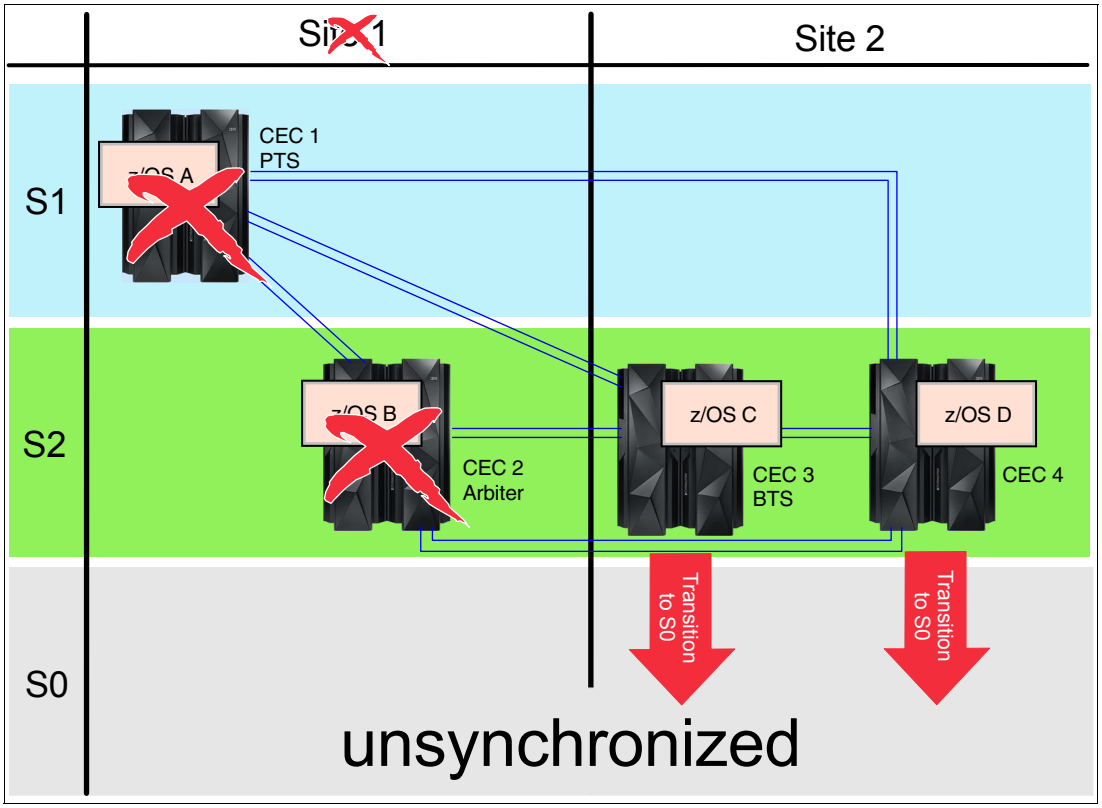

Figure 2-13 STP-only CTN - Site 1 failure with CTS and Arbiter at Site 1, BTS at Site 2

Site 1 power outage: CTS and Arbiter at Site 1 with IBF

If the CTN spans two data centers, it is suggested to install the IBF on the servers that will be assigned the roles of PTS, BTS, and Arbiter. This should provide recovery capabilities when the site where the CTS is located experiences a power failure. In the example shown in Figure 2-14 on page 41, assume that the IBF is installed at least on CEC1 (PTS/CTS) and CEC2 (Arbiter). It is suggested that the IBF also be installed on the BTS in order to provide recovery capability when the BTS has the role of the CTS.

If Site 1 has a power outage, and the IBF is engaged on both CEC1 and CEC2, they will notify CEC3 (BTS) via the coupling links between the servers that they are running on IBF power. The BTS waits for a period of 30 seconds after notification to take any recovery action, in case the power outage condition is just a transient power outage or glitch. In the case that this is indeed a transient power condition, CEC1 and CEC2 will notify CEC3 that they have now returned to “normal power” status. If the BTS receives a normal power status from the PTS and Arbiter within 30 seconds, there is no further action.

However, if the BTS does not receive a normal power status within 30 seconds, then it will initiate processing to take over the CTS role and the PTS and Arbiter will stay synchronized as Stratum 2 servers until they power off. If the BTS got the IBF notification from both the CTS *and* the Arbiter, automated network recovery will also be temporarily disabled. This is important because otherwise the BTS would surrender its role as CTS once it loses communication to the PTS and Arbiter when they power off. Automated recovery will also be temporary disabled if the Arbiter does not have an IBF installed. In this case the Arbiter loose power immediately when the power outage occurs and both server PTS as well as BTS identify they both can not see the Arbiter anymore. As a consequence the automated network recovery gets disabled, since Arbiter assisted recovery can not be performed without the Arbiter.

When the PTS returns to normal power after the outage, it automatically resumes the CTS role.

Installing the IBF on the PTS and Arbiter allows the BTS to successfully take over as the CTS, as compared to the scenario described in “Site 1 failure - CTS and Arbiter at Site 1, BTS at Site 2” on page 39, where the site failure was not a power outage, and the BTS and CEC4 became unsychronized.

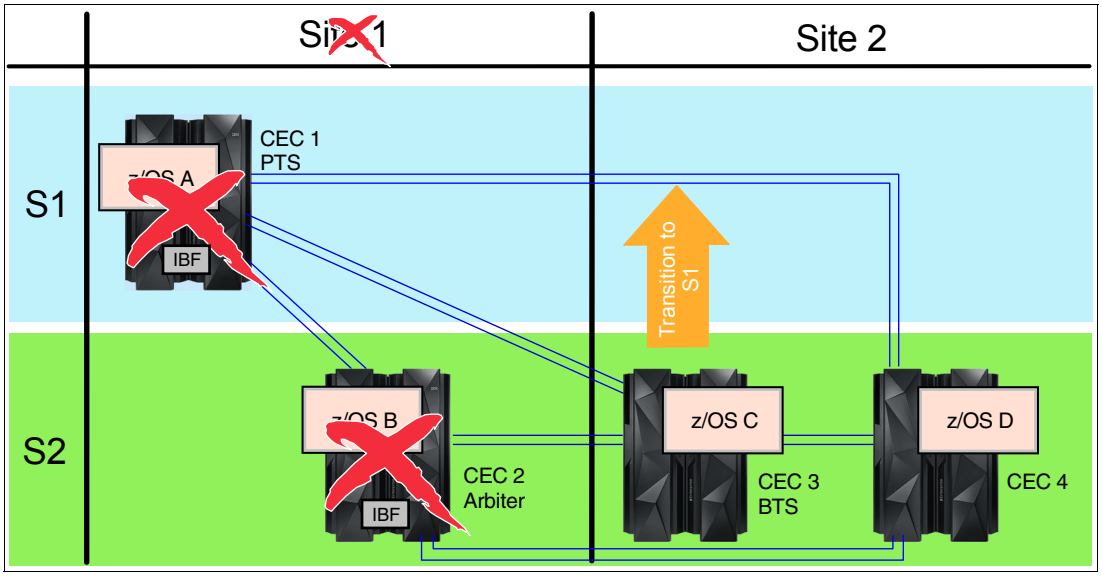

Figure 2-14 STP-only CTN - Site 1 power outage with CTS and Arbiter at Site 1, IBF installed, BTS at Site 2

Note: The timing links from the BTS on Site 2 to the CTS and Arbiter need to remain functional until the BTS has taken over the CTS role and automated recovery is disabled. In case the timing links fail (for example, because they are connected via DWDM, which also fails due to the power outage) before the BTS has taken over, the failure will be treated as a loss of communication between the two sites.

Site 2 failure: CTS and Arbiter at Site 1, BTS at Site 2

The CTS and the Arbiter both survive this failure because they are both located at Site 1. This is shown in Figure 2-15 on page 42.

The BTS and CEC4 are lost because they are located at Site 2. Both the CTS and Arbiter survive and are able to remain synchronized without further recovery processing being required.

The recovery processing is as follows:

1. The CTS loses all communication with the BTS.
2. The CTS tries to communicate with the Arbiter, which is successful. The Arbiter confirms that the BTS is no longer reachable.
3. The PTS retains the role of CTS and all servers in Site 1 remain synchronized.
4. The automated network recovery gets disabled since the PTS and the Arbiter both lose communication to the BTS.

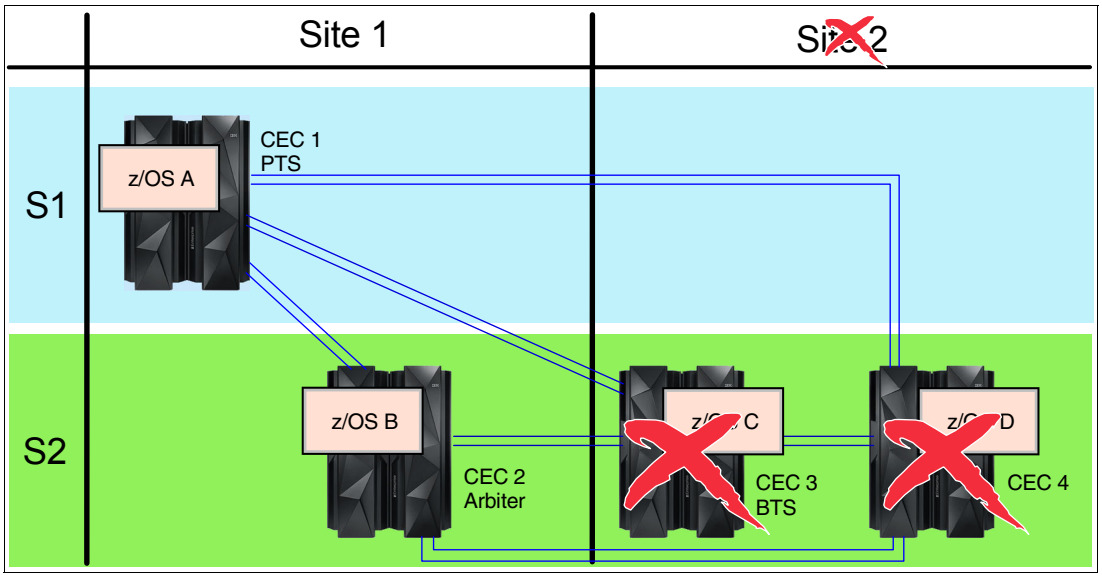

Figure 2-15 STP-only CTN - Site 2 failure with CTS and Arbiter at Site 1, BTS at Site 2

Site 1 failure: CTS at Site 1, BTS and Arbiter at Site 2

Important: The following configuration is generally *not* suggested and is discussed only to highlight the risks associated with having the Arbiter at a different site from the CTS. Although a Site 1 failure will result in z/OS systems at Site 2 not being affected, a Site 2 failure *will* impact z/OS systems at Site 1, which is unacceptable for most installations.

This configuration is only recommended in a GDPS environment with a freeze policy of FREEZE=STOP or FREEZE=SWAP,STOP.

The BTS and the Arbiter both survive this failure because they are both located at Site 2, as shown in Figure 2-16 on page 43.

The CTS and CEC2 are lost because they are located at Site 1. Both the BTS and Arbiter survive and are able to remain synchronized without further recovery processing being required.

The recovery processing is as follows:

1. The BTS loses all communication with the CTS.
2. The BTS tries to communicate with the Arbiter to establish whether the Arbiter can communicate with the CTS.
3. Both the BTS and Arbiter have lost communication with the CTS, so the BTS assumes the CTS role and all servers at Site 2 remain synchronized.

4. The automated network recovery gets disabled since the BTS and the Arbiter both lost communication to the PTS.

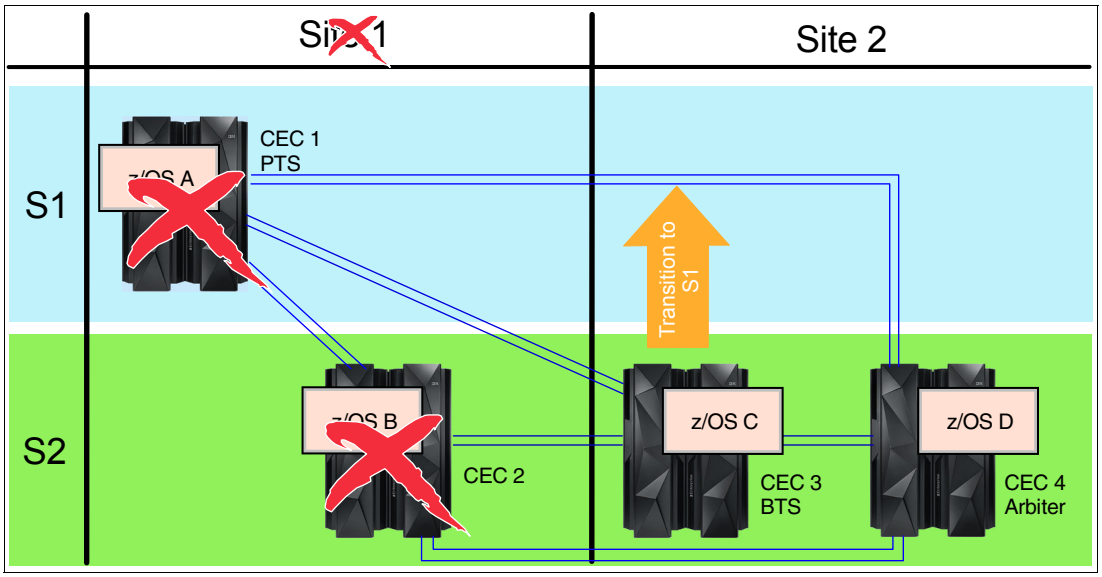

Figure 2-16 STP-only CTN - Site 1 failure with CTS at Site 1, BTS and Arbiter at Site 2

Site 2 failure: CTS at Site 1, BTS and Arbiter at Site 2

Important: Placing both the BTS and Arbiter at Site 2 *will* impact Site 1 systems during a Site 2 outage. Using this configuration is highly discouraged and should not be implemented unless there are specific reasons for doing so (for example, GDPS with FREEZE=STOP or FREEZE=SWAP,STOP) and the consequences for your Site 1 systems are understood.

This particular configuration is very dangerous because a Site 2 outage impacts Site 1 systems, which is unacceptable in most client environments; refer to Figure 2-17 on page 44 for details. In this scenario, the CTS survives the Site 2 failure because it is located at Site 1. However, because the CTS has lost contact with both the BTS and Arbiter, it enters recovery processing—which eventually leads to loss of the CTS role and all Site 1 servers becoming unsynchronized.

The recovery processing for a Site 2 failure where the CTS is located at Site 1 and the BTS and Arbiter are located at Site 2 consists of the following steps:

1. The CTS loses all communication with both the Arbiter and BTS.
2. The PTS surrenders the role of CTS.
3. The PTS initiates Console-assisted recovery to attempt to determine the state of the BTS.
4. The results of Console-assisted recovery are most likely indeterminate so the PTS cannot resume the CTS role.
5. The PTS eventually becomes unsynchronized and changes to Stratum 0. CEC 2 also loses its time source and transitions to Stratum 0. z/OS systems resident at Site 1 with STPMODE set to YES issue a WTOR IEA394A message.
6. Manual action via the HMC is required to reassign the role of CTS to the PTS before replying `RETRY` to the WTOR IEA394A message.

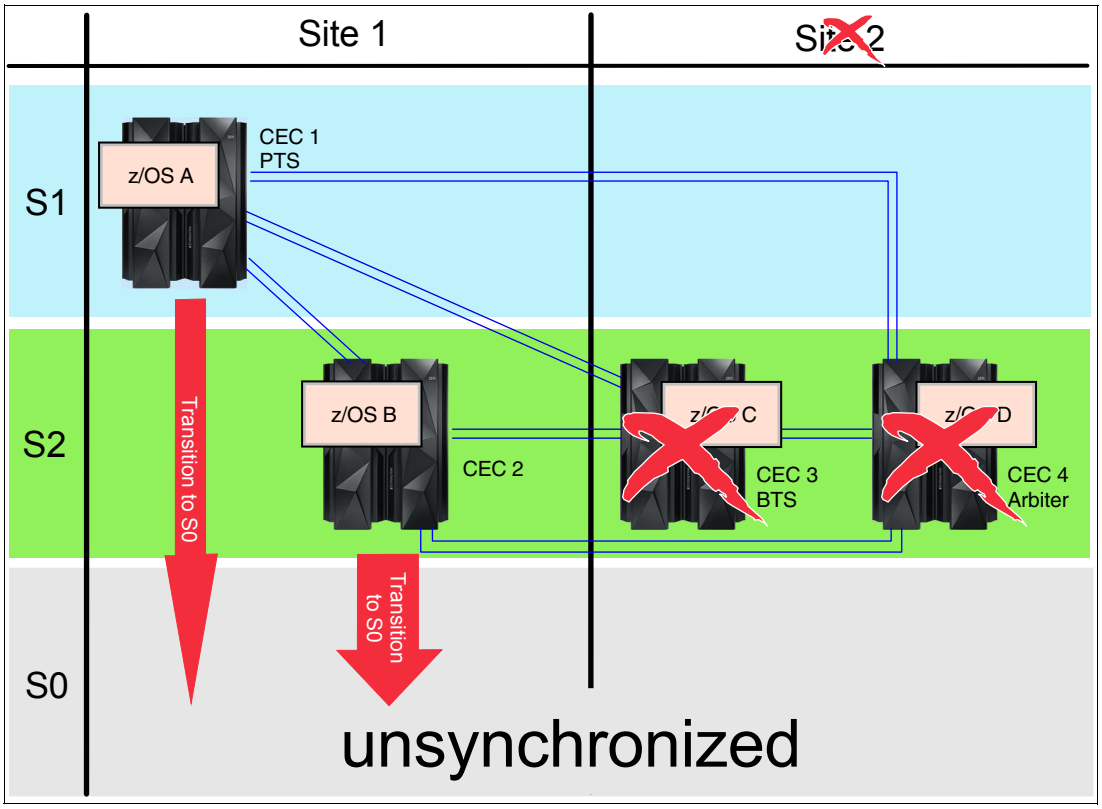

Figure 2-17 STP-only CTN - Site 2 failure with CTS at Site 1, BTS and Arbiter at Site 2

Site 2 power outage: CTS, BTS, and Arbiter with IBF

In the example shown in Figure 2-18, assume that the IBF is installed on CEC1 (PTS/CTS), CEC3 (BTS), and CEC4 (Arbiter).

If Site 2 has a power outage, and the IBF is engaged on both CEC3 and CEC4, they will notify CEC1 (PTS/CTS) via the coupling links between the servers that they are running on IBF power. The PTS waits for a period of 30 seconds after notification to take any recovery action, in case the power outage condition is just a transient power outage or glitch. In the case that this is indeed a transient power condition, CEC3 and CEC4 will notify the CEC1 that they have now returned to “normal power” status. If the PTS/CTS receive a normal power status from the BTS and Arbiter within 30 seconds, there is no further action.

However, if the PTS/CTS does not receive a normal power status within 30 seconds, and it has received the IBF notification from both the BTS *and* the Arbiter, automated network recovery will be disabled. Thus the PTS will not surrender the CTS role when both the BTS and Arbiter get powered off eventually (IBF cannot keep the server up) and the PTS cannot communicate with both of them. This is different than the case described in “Site 2 failure: CTS at Site 1, BTS and Arbiter at Site 2” where the site failure was not a power outage.

When CEC3 and CEC4 return to normal power after the outage, they automatically join the CTN as Stratum 2 servers.

Note: The timing links from the CTS to the Arbiter and BTS need to remain functional until the 30-second interval has expired so that the automated network recovery is disabled. In case the timing links fail before that (for example, say the two sites are connected via DWDM, which also fails due to the power failure), the failure will be treated as a loss of communication between the two sites.

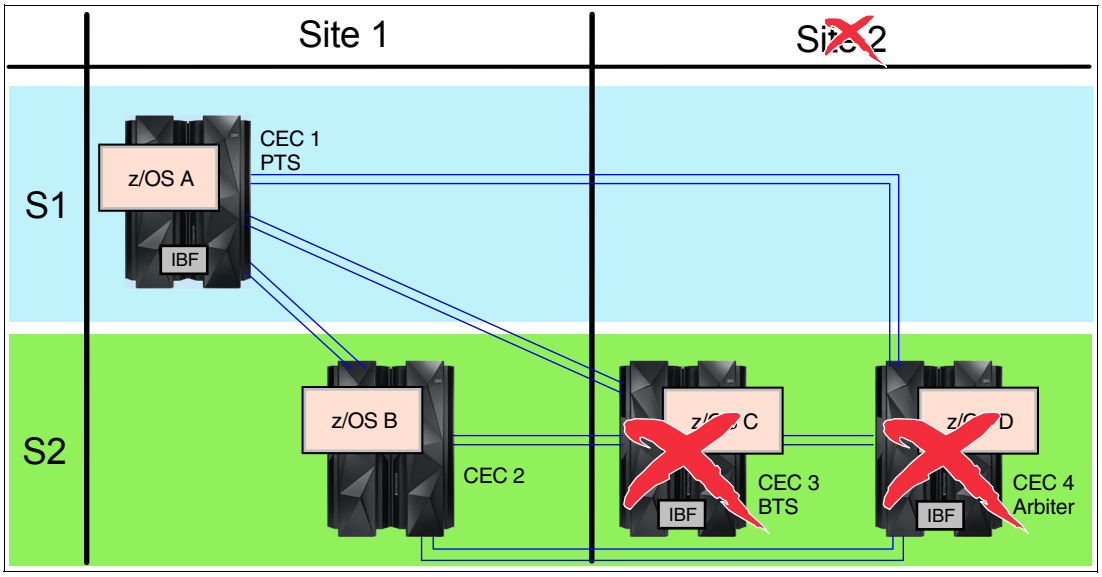

Figure 2-18 STP-only CTN - Site 2 power outage with CTS at Site 1, BTS and Arbiter at Site 2, IBF installed

Loss of communication between sites

For Site 2 servers this looks like a Site 1 failure, and for Site 1 servers, this looks like a Site 2 failure. For both sites, recovery processing is dependent upon where the Arbiter is located.

2.5 Restoring STP configuration across PORs in a single- or dual-CEC CTN

STP provides an option to save the configuration across PORs for STP-only CTNs with one or two servers. The selection of this option is made via the checkbox for **Only allow the servers specified above to be in the CTN** in the Network Configuration tab (see Figure 2-19 on page 46). When **Only allow the servers specified above to be in the CTN** has been selected, the CTN's timing and configuration settings are saved so that they will not need to be reentered after a loss of power or a power-on reset (POR) of the servers.

Note: For brevity's sake, the capability implemented by selecting **Only allow the server(s) specified above to be in the CTN** is referred to as the "save configuration" feature and CTNs for which this capability is selected are referred to as "bounded" CTNs.

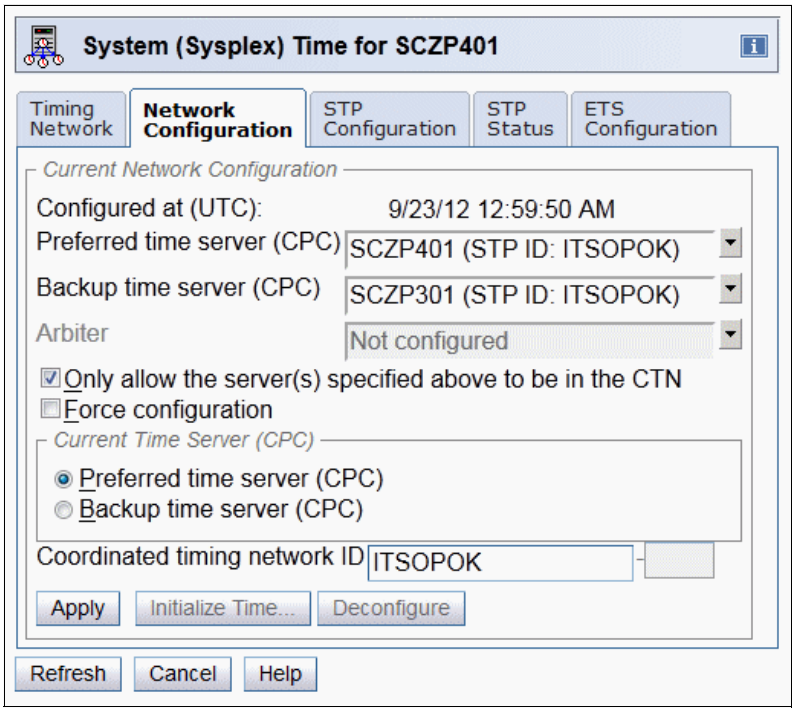

Figure 2-19 Dual CEC - saving configuration across POR

2.5.1 Single-server CTN

In this section we discuss recovery scenarios for a single-server STP-only CTN that uses the “save the STP configuration” feature. The scenario diagram is shown in Figure 2-20. Note that only one server is configured.

Figure 2-20 Single CEC configuration - STP Stratum 1

The example configuration has one server configured as Preferred Time Server (PTS) and Current Time Server (CTS - Stratum 1).

Prerequisites

For single-server bounded CTNs, the CTN must contain only a single server and be operating in STP-only mode. The configuration data saved includes the server’s role as the PTS/CTS reflected in the Network Configuration tab (see Figure 2-21 on page 47).

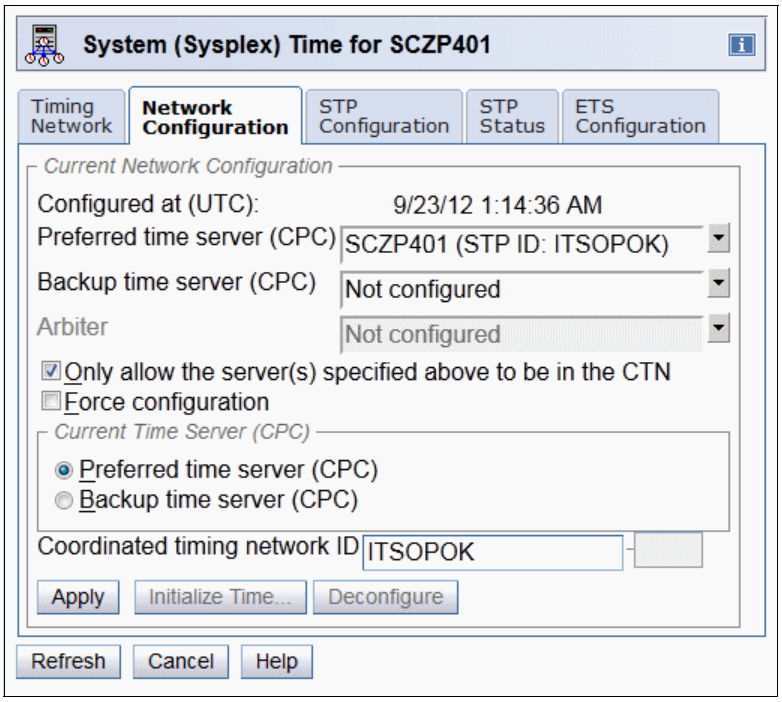

Figure 2-21 Configuring a single CEC bounded CTN - Network Configuration tab

Note: If the “save configuration” feature is used to limit the CTN to a single server CTN, no other server can join this CTN unless this option is deselected in the Network Configuration window of the Current Time Server of the existing CTN.

This restriction can be removed concurrently at any time by deselecting **Only allow the server(s) specified above to be in the CTN**.

Restoring the CTN configuration after power off/on sequence

When power is restored, the server will resume its role as PTS/CTS using the timing configuration information previously provided. Due to the simplicity of the CTN configuration, there are no considerations to impact the recovery process.

2.5.2 Dual server CTN

In this section we discuss planning for recovery for a dual server STP-only CTN. The scenario diagram is shown in Figure 2-22 on page 48. Note that only two servers are configured.

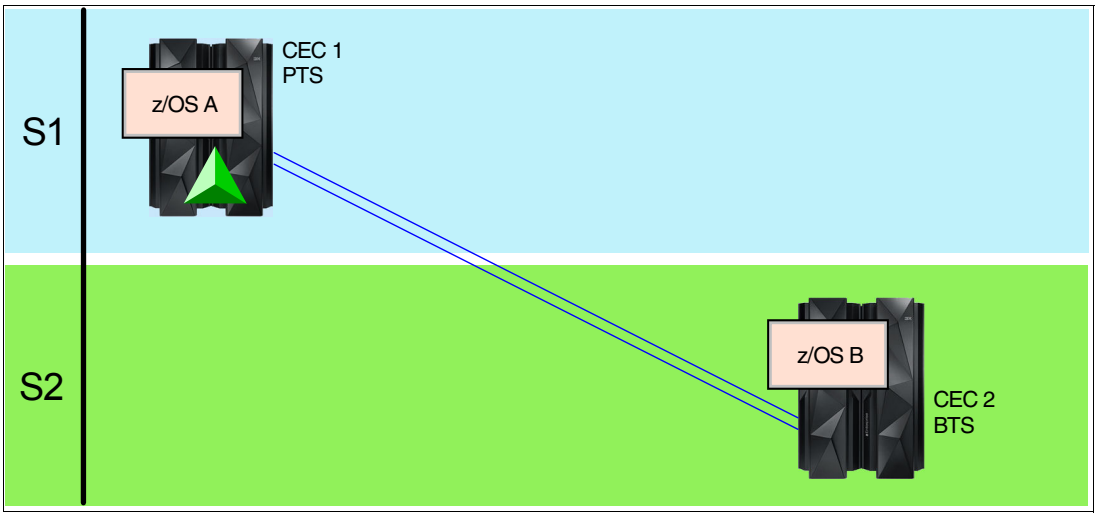

Figure 2-22 STP-only CTN with PTS and BTS assigned

The example configuration has one server configured as Preferred Time Server (PTS) and one as Backup Time Server (BTS). The servers are connected via six coupling links. CEC1 is the PTS and Current Time Server (CTS - Stratum 1), while CEC2 is the BTS (Stratum 2).

Note: The recovery actions described in this section assume that a BTS has been assigned and the PTS has been assigned the role of Current Time Server (CTS). This is the suggested configuration because the PTS/CTS is a single point of failure and only the PTS can automatically retake the CTS role after a recovery action.

Prerequisites

For dual server bounded CTNs, the CTN must contain two servers and be operating in STP-only mode.

The configuration data saved includes the servers' roles as the PTS/CTS and BTS, reflected in the Network Configuration tab (see Figure 2-23 on page 49).

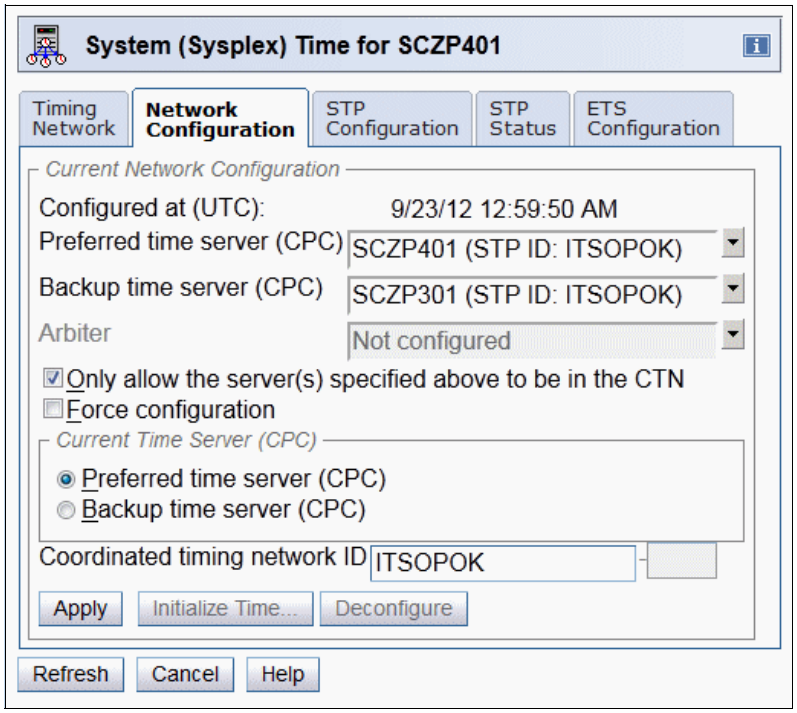

Figure 2-23 Dual CEC CTN configuration - Network Configuration tab

Note: If the “save configuration” feature is used to limit the CTN to a dual server CTN, a number of restrictions apply:

- No other server can join the CTN.
- The CTN ID cannot be changed.
- No server role assignment changes can be made.
- No reverse migration from STP-only to a mixed CTN.

These restrictions can be removed concurrently at any time by deselecting **Only allow the server(s) specified above to be in the CTN** from the CTS, as long as the two servers can communicate via coupling links.

Restoring the CTN configuration after a power off/on sequence

When power is restored to both the PTS and the BTS, the servers will normally resume their roles as PTS/CTS and BTS, using the timing configuration information previously provided. However, with two servers, there are some considerations that could impact the configuration restore process.

All scenarios in this chapter assume that the PTS has been assigned the role of CTS when the CTN was configured.

When one server loses power

If one server with IBF loses power, it notifies the nonfailing server that it is operating on IBF power and, if power is not restored within 30 seconds, the nonfailing server assumes or continues the CTS role. When power is restored to the failed CEC, the two servers revert back to their original roles.

If a server without IBF loses power, the recovery process depends upon whether OLS signals are received by the surviving server on all established STP paths within a period of two

seconds. If OLS signals are received, the nonfailing server assumes or continues the CTS role. When power is restored to the failed CEC, the two servers revert back to their original roles.

If a server without IBF loses power and OLS signals are not received by the surviving server, the recovery process depends upon whether LAN connectivity is available between the servers so that Console Assisted Recovery (CAR) initiated by the nonfailing server can confirm the loss of the failed server. If LAN connectivity is available and CAR confirms that the failed server is no longer operating as a time server, the nonfailing server assumes or continues the CTS role. When power is restored to the failed CEC, the two servers revert to their original roles.

If a server without IBF loses power, OLS signals are not received by the surviving server, and either LAN connectivity is not available between the servers or CAR is unable to confirm the loss of the failed server, the recovery process depends upon which time server failed. If the BTS failed, the PTS continues to operate as the CTS. When power is restored to the BTS, it resumes its role as Stratum 2. If the CTS server failed, the BTS transitions to Stratum 0. The BTS has to transition to Stratum 0 because it cannot confirm that the CTS is inoperative and cannot risk creating two separate time sources when there could be tasks operating on the two servers that require a common time source.

If the BTS transitions to Stratum 0, a manual recovery process can be utilized if the PTS can be definitively determined to be inoperative, perhaps by physical examination. If the PTS can be determined to be inoperative, **Assume Current Time Server** can be selected on the BTS network configuration panel (see Figure 2-24 on page 51) to cause the BTS to assume the role of current time server. A more detailed explanation of this option is provided in “Forcing a server to assume the CTS role” on page 52. If a definitive evaluation of the condition of the PTS is not possible, the BTS must remain at Stratum 0 for the reasons previously stated. Regardless of whether the BTS assumes the role of the CTS, when power is restored to the PTS and connectivity is reestablished between the PTS and BTS, the two servers revert back to their original roles.

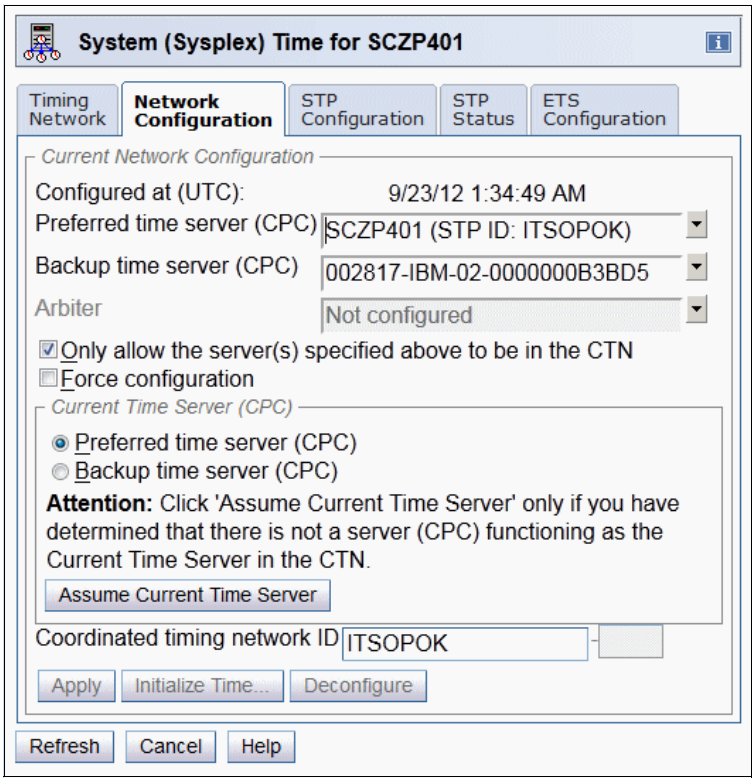

Figure 2-24 Assume the CTS role in a dual CEC CTN

When both servers lose power

When power is lost to both servers, the recovery process depends on whether coupling link or LAN connectivity is reestablished between the servers. Both servers remain at Stratum 0 until coupling link or LAN connectivity is reestablished between the servers. If the servers are able to communicate via coupling link, they resume their roles as PTS/CTS and BTS. If the servers cannot communicate via coupling link but can communicate via LAN, Console Assisted Recovery (CAR) is initiated.

Console Assisted Recovery initiated by the PTS

If the PTS determines that the BTS has failed or is a Stratum 0, the PTS becomes the Stratum 1. When coupling link connectivity is restored, the BTS will become a Stratum 2 server.

If the PTS determines that the BTS is Stratum 1, the PTS remains at Stratum 0. When coupling link connectivity is restored, the PTS retakes its role as Stratum 1 and the BTS becomes a Stratum 2 server.

If the PTS cannot determine the status of the BTS, the PTS remains at Stratum 0. When coupling link connectivity is restored, the servers resume their roles as PTS/CTS and BTS.

Console Assisted Recovery initiated by the BTS

If the BTS determines that the PTS has failed, the BTS becomes the Stratum 1. When coupling link connectivity is restored, the PTS will rejoin the CTN as Stratum 2 and then retake its role as CTS and become a Stratum 1 server.

If the BTS determines that the PTS is Stratum 1, the BTS remains at Stratum 0. When coupling link connectivity is restored, the BTS joins the CTN as Stratum 2.

If the BTS cannot determine the status of the PTS, the BTS remains at Stratum 0. When coupling link connectivity is restored, the servers resume their roles as PTS/CTS and BTS.

When CAR fails

In both the CAR from the PTS perspective scenarios and the CAR from the BTS perspective scenarios, there is a scenario in which CAR is unable to assign the role of CTS to either server, resulting in an outage for both servers until link path connectivity can be reestablished between the servers. In these situations, a complete outage can be avoided by forcing one of the servers to assume the CTS role.

Forcing a server to assume the CTS role

As previously outlined, there are situations in which one or both servers are available but an automated assignment of the CTS role is not possible due to a lack of communication between the servers. In these situations, if the status of the servers can be determined manually, it is possible to force one of the servers to assume the CTS role without permanently reconfiguring the CTN. While this process can be viewed from the perspective of either the PTS or the BTS, we review the process from the perspective of the PTS.

Scenario and initial configuration

The scenario that follows is a typical example of a situation for which the “Assume CTS role” capability was designed. In this scenario, two servers, SCZP301 and SCZP401, are configured in an STP-only CTN as the PTS and BTS, respectively. The two servers are connected by coupling links, as can be seen in the STP Status tab (see Figure 2-25).

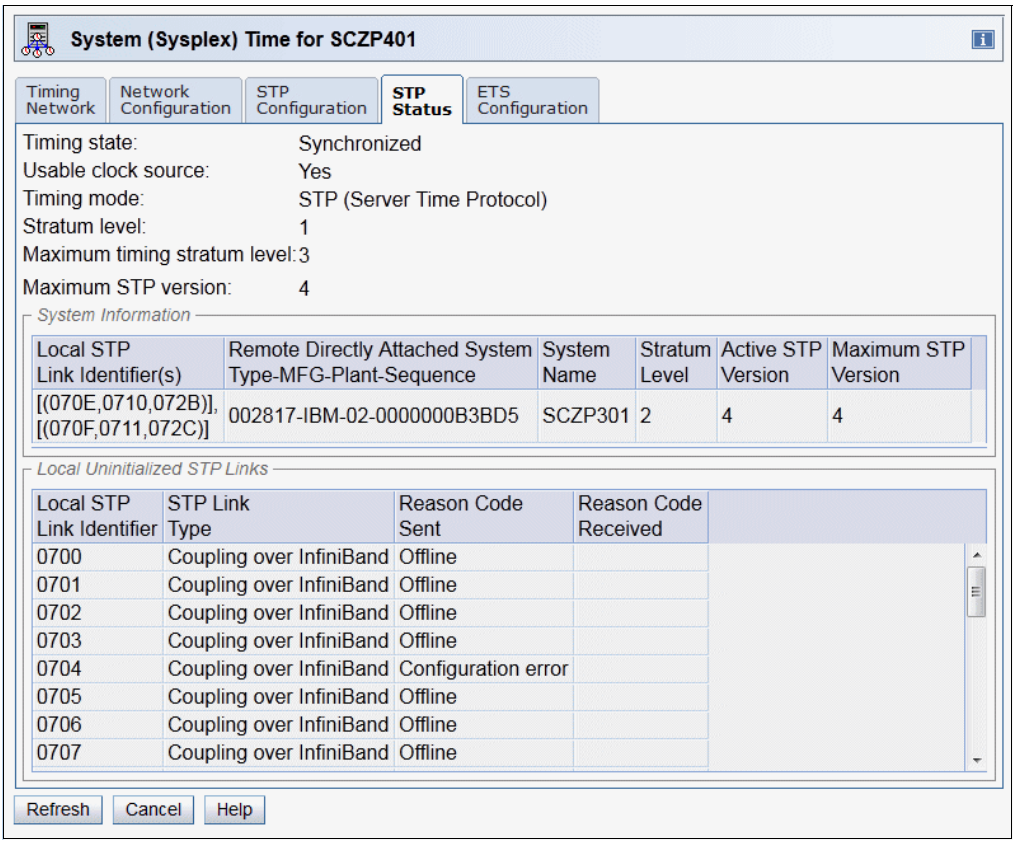

Figure 2-25 Dual CEC - STP Status tab

Neither server has the Internal Battery Feature and the STP configuration has been saved by selecting **Only allow the server(s) specified above to be in the CTN** in the Network Configuration tab (see Figure 2-26).

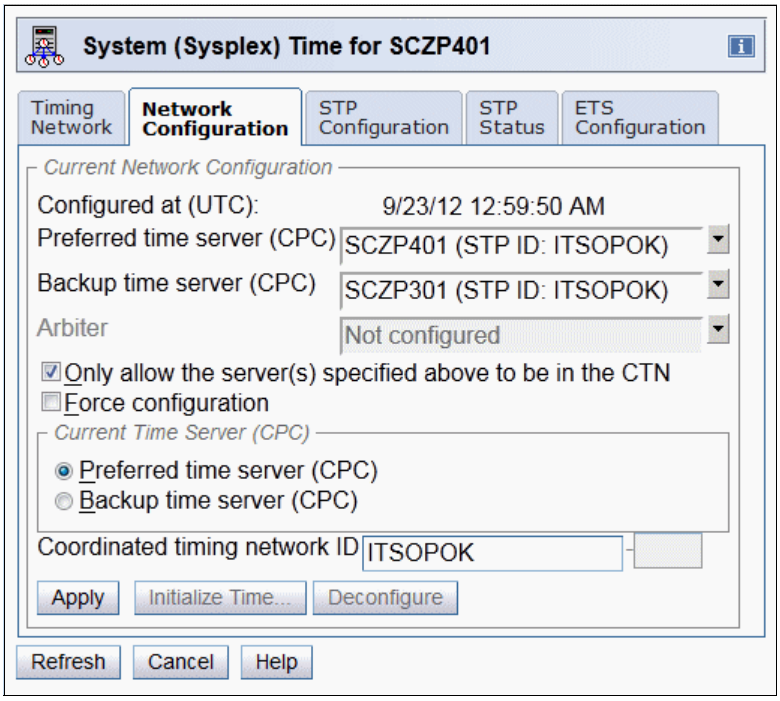

Figure 2-26 Saving STP configuration across PORs - dual CEC CTN

Problem description

After a power outage involving both servers, server SCZP401 is powered on but server SCZP301 remains powered off due to an environmental issue. Since server SCZP301 has completely lost power, there is no communication on the LAN path to its Support Element since it is also powered off. Because it is unknown when power will be restored to Server SCZP301, the more pressing concern is to get one of the servers working so that production workload processing can be resumed. As is typically the case, the most important workload is processed on the PTS server, SCZP401 in this case.

The *unsynchronized* status and the problem with coupling link connectivity is evidenced in the STP Status tab for server SCZP401 (see Figure 2-27 on page 54).

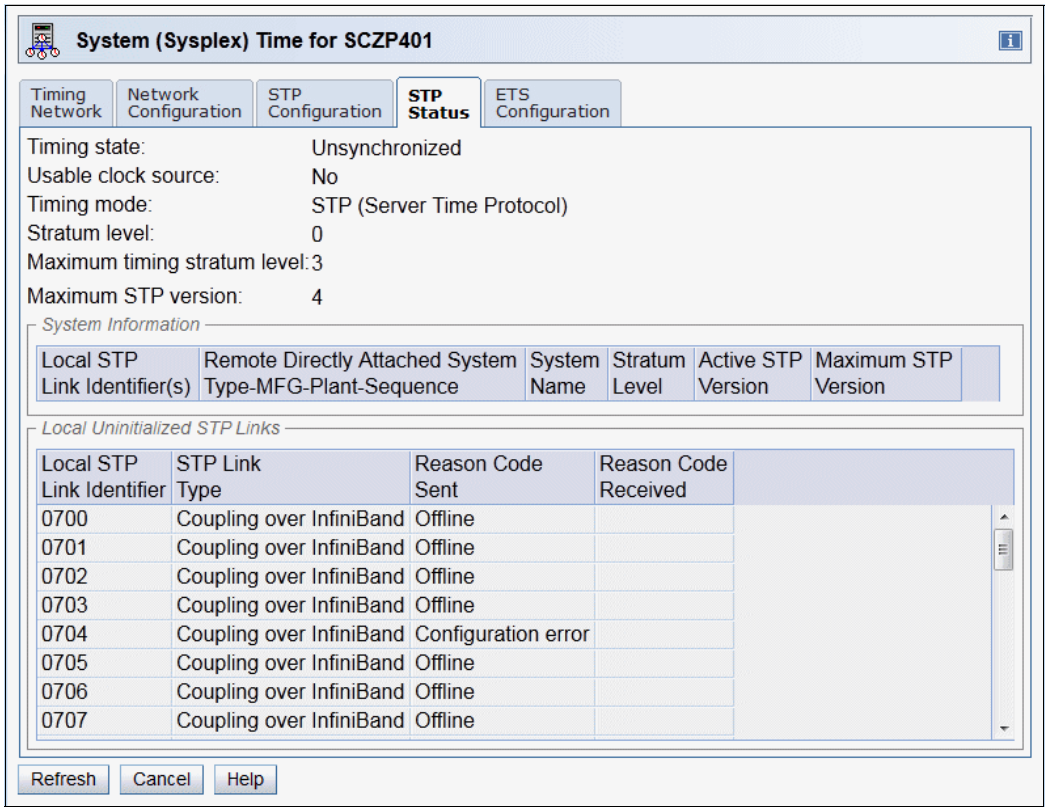

Figure 2-27 STP Status tab - SCZP401 unsynchronized (STP Stratum 0)

Additional confirmation of the communication problem to server SCZP301 can be found by looking at the Network Configuration tab for server SCZP401 (see Figure 2-28 on page 55). The BTS field shows the node descriptor instead of the CPC name. This indicates that the HMC cannot see server SCZP301 on the LAN path, which can be related to the power off condition of its Support Element.

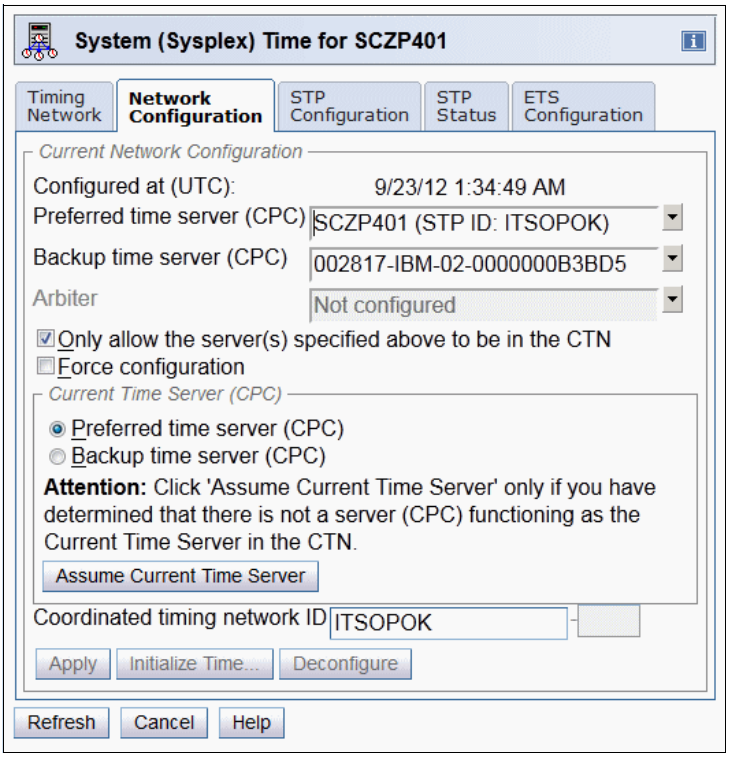

Figure 2-28 SCZP401 Network Configuration tab: no coupling link to SCZP301

User recovery actions

The PTS server, SCZP401, needs to assume the CTS role so that production workload processing can continue while the cause of the coupling link connectivity problem is being investigated. To do this, select **Assume Current Time Server** in the Network Configuration tab (see Figure 2-29 on page 56) and then select **Apply**.

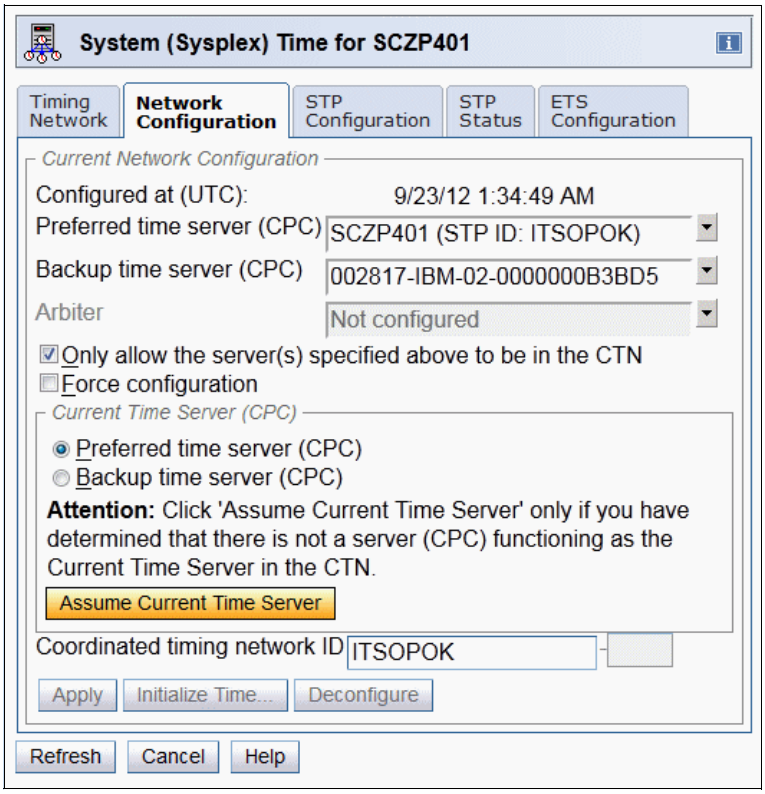

Figure 2-29 SCZP401 assuming CTS role

Confirm the Assume Current Time Server role by answering **Yes**, as shown in Figure 2-30.

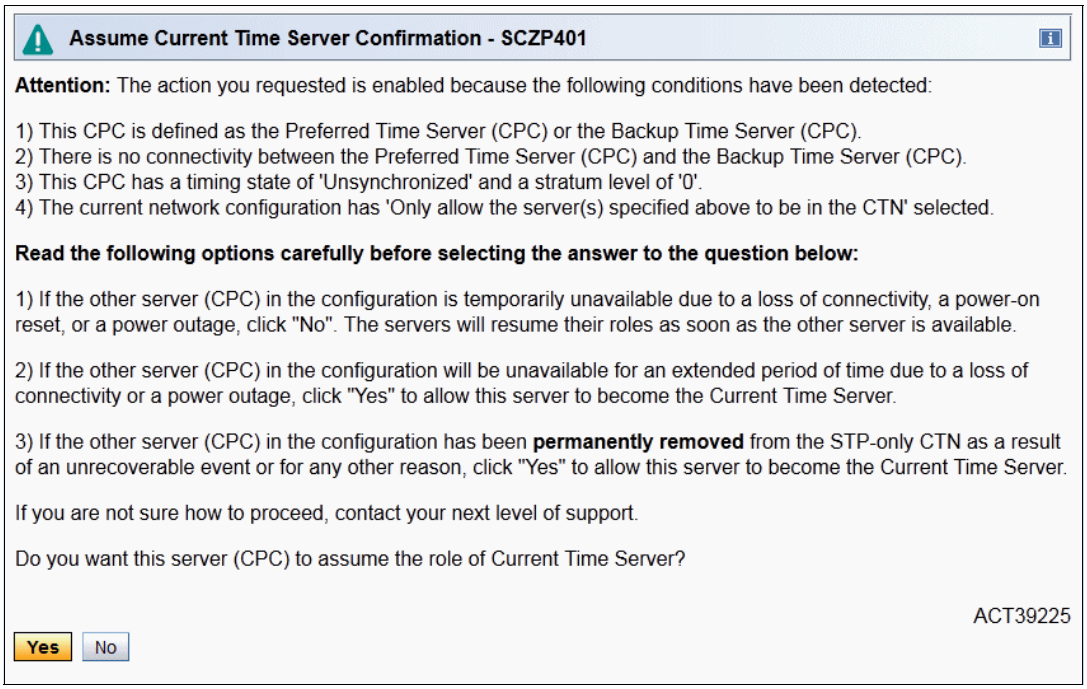

Figure 2-30 Assume Current Time Server message ACT39225

The Assume Current Time Server function is completed (Figure 2-31 on page 57). Select **OK**.

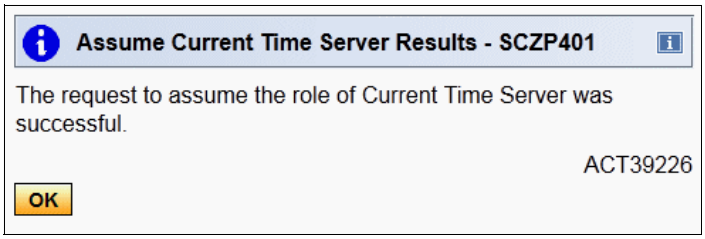

Figure 2-31 Assume CTS successful confirmation message ACT39226

Recovered configuration

After the PTS SCZP401 has assumed the CTS role, production processing can resume. As can be seen in the STP Status tab in Figure 2-32, SCZP401 is now synchronized but still has no coupling link connectivity to the BTS.

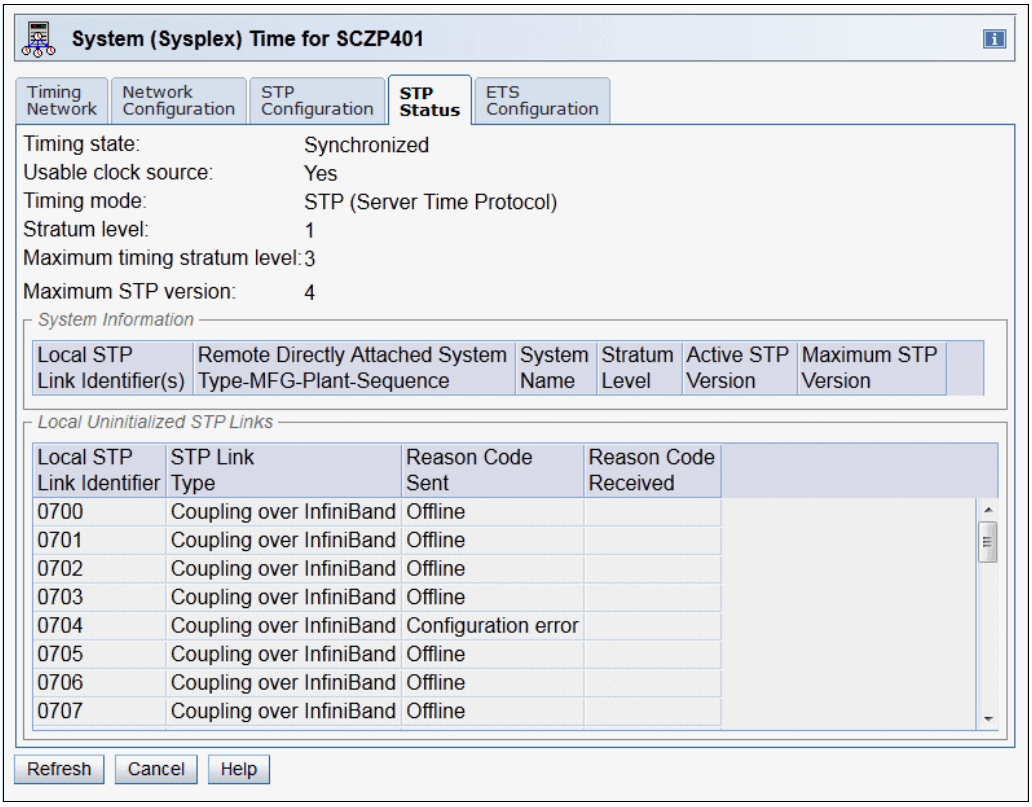

Figure 2-32 SCZP401 synchronized but no link to SCZP301

In addition, the option to Assume Current Time Server is no longer available in the Network Configuration tab because the CTS role has been assigned to the PTS (see Figure 2-33 on page 58).

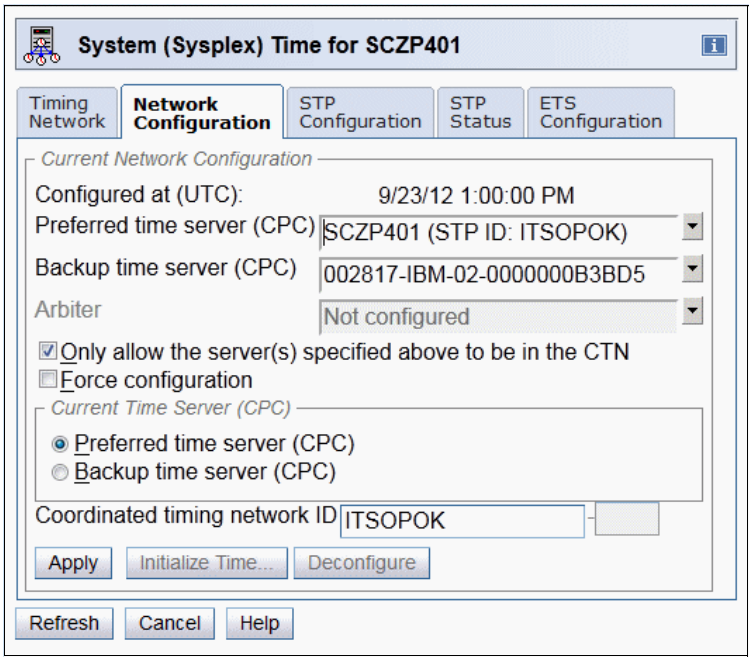

Figure 2-33 Assume Current Time Server button not present

Now that the production workload is once again being processed on the PTS, time is available to investigate the power problem that prevents server SCZP301 to power on. Once the environment problem has been corrected, corrective action can be taken to bring the second server back into the CTN because the BTS and regular workload processing can be resumed.

Forcing a CTN to delete the saved configuration

Deleting the saved configuration is accomplished by deselecting **Only allow the server(s) specified above to be in the CTN** in the Network Configuration tab and then selecting **Apply**. Since the configuration data is stored on both servers and this information needs to be in synch, deleting the saved configuration requires that the two servers have coupling link connectivity. However, circumstances can arise in which coupling link connectivity is not available and the loss of connectivity is permanent. In this case, the saved configuration no longer applies and can significantly impact the operation of the servers in the CTN (see “Forcing a server to assume the CTS role” on page 52).

One potential cause for this situation is a permanent removal of one of the servers prior to deleting the saved configuration. In this circumstance, the removed server will never rejoin the CTN, so the configuration needs to be changed to reflect this reality to avoid continually impacting the operation of the remaining server. In Figure 2-34 on page 59, server SCZP301 (only the Node ID is shown since the server is completely powered off) is no longer available and needs to be removed from the network configuration.

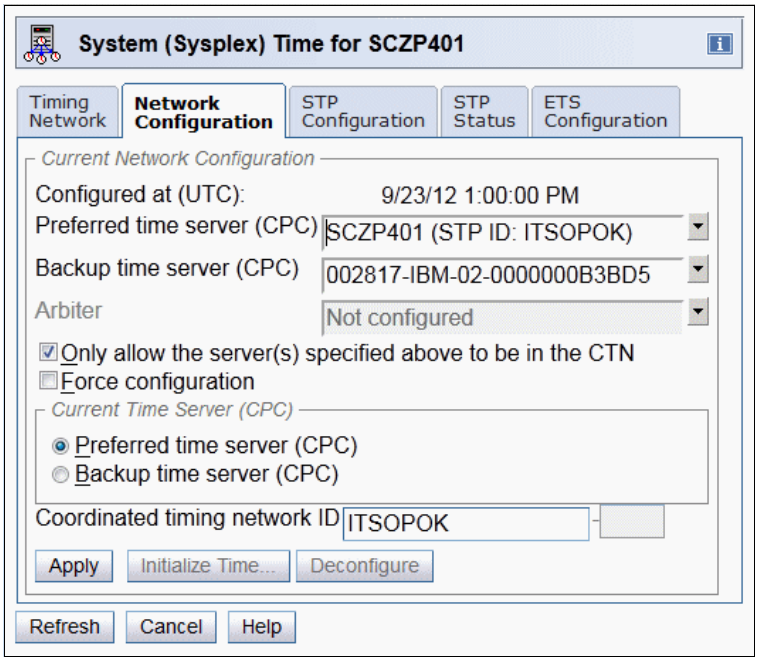

Figure 2-34 Unavailable server before deletion

The first step in correcting the configuration is to force the deselection of **Only allow the server(s) specified above to be in the CTN** by additionally selecting the **Force configuration** option and then selecting **Apply** (see Figure 2-35).

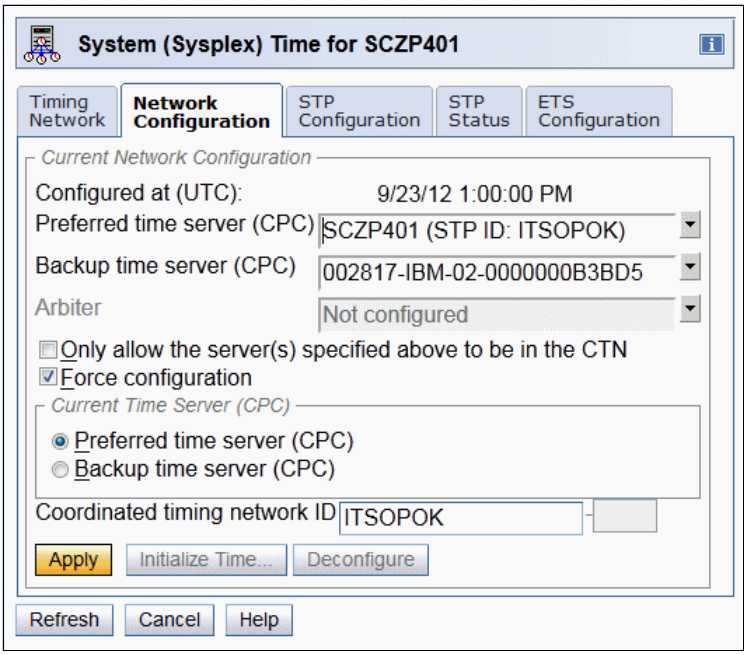

Figure 2-35 Deleting unavailable server

A confirmation message appears. Select **Yes** so that the operation can complete. Now that **Only allow the server(s) specified above to be in the CTN** is no longer selected, the server which is no longer available can be removed from the configuration. After the configuration has been corrected, the **Only allow the server(s) specified above to be in the CTN** option

should once again be selected and then select **Apply** to save the revised configuration (see Figure 2-36).

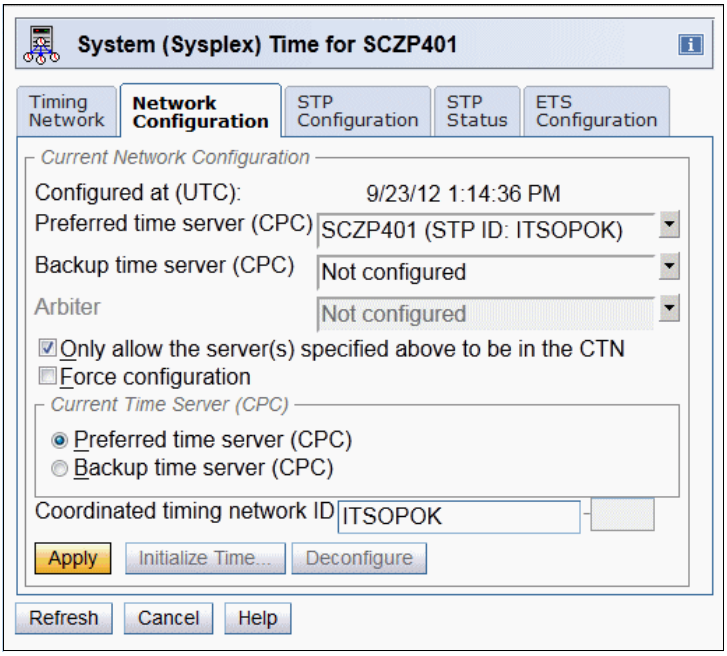

Figure 2-36 Saving the STP configuration for single server

2.5.3 Considerations for MES upgrade

Care must be taken when applying an MES upgrade that changes the machine type. Consider a situation in which server SCZP401 is a zEC12 and server SCZP301 is a z196 that is being upgraded to a zEC12 via an MES. The initial configuration can be seen in Figure 2-37 on page 61. All the configuration changes being shown need to be performed from server SCZP401 since this server is the designated Current Time Server. In this example the CTS does not change and therefore all changes are being done from the CTS. If the MES would target server SCZP401, the configuration would also require a PTS/CTS role change which would change the designated CTS, in which case the configuration would have to be applied from server SCZP301 since it would become the new PTS/CTS.

The following configuration change must be accomplished while the two servers have coupling link connectivity or the **Force Configuration** option will be required to make the change (see “Forcing a CTN to delete the saved configuration” on page 58).

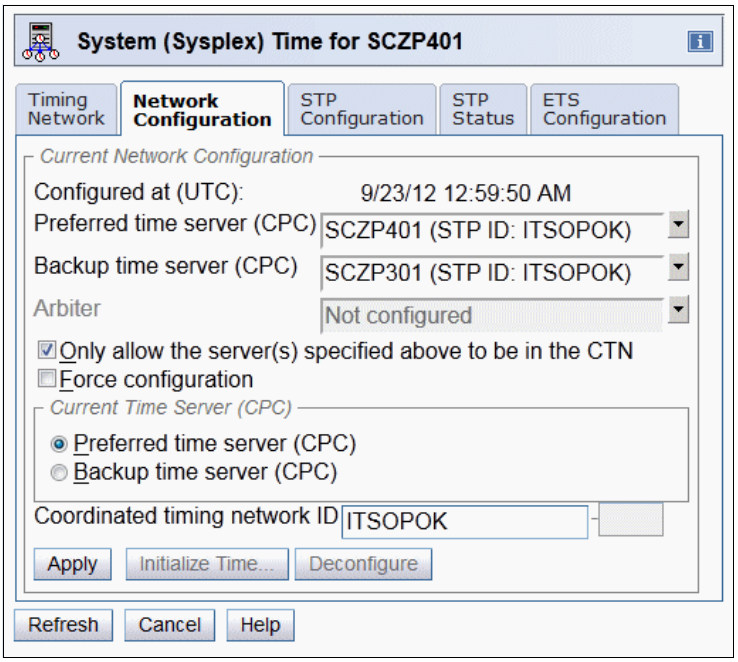

Figure 2-37 Initial configuration (before MES)

Prior to applying such an MES, the saved configuration must be deleted by deselecting **Only allow the server(s) specified above to be in the CTN** and then selecting **Apply** in the Network Configuration tab (see Figure 2-38).

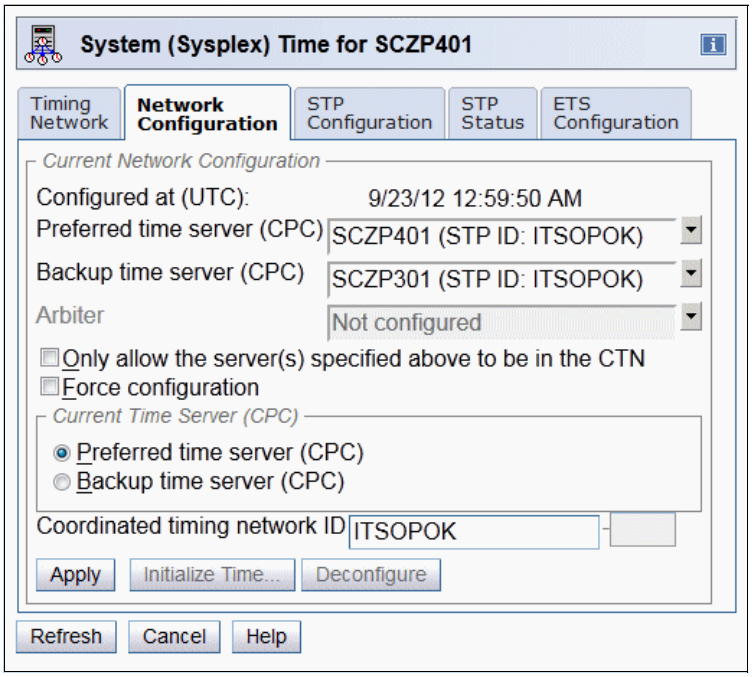

Figure 2-38 Allowing other servers to join the CTN

Then remove the machine being upgraded from the CTN configuration (see Figure 2-39 on page 62) by setting the BTS to **Not configured**.

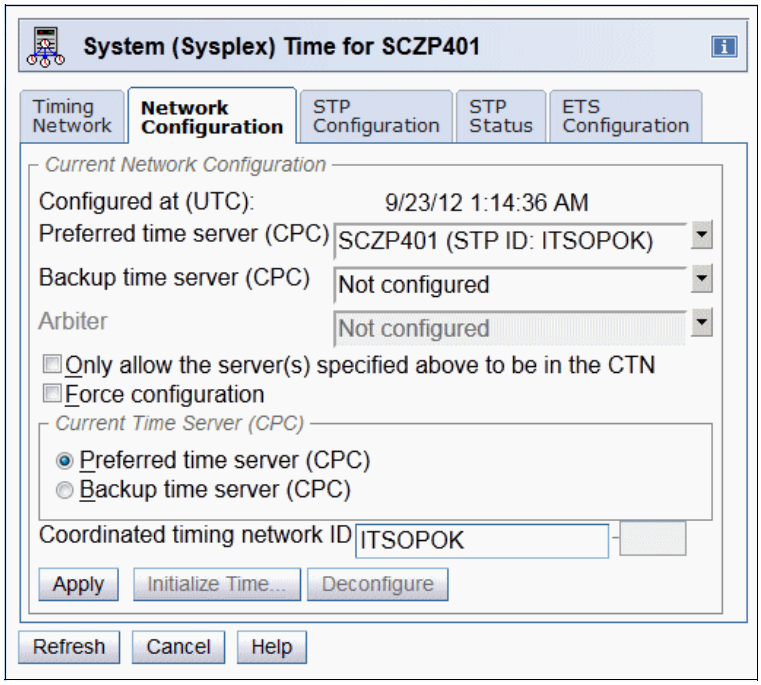

Figure 2-39 Removing the CEC to be upgraded

You can ensure that the configuration and timing values are not lost due to an unanticipated outage of server SCZP401 while the MES is being applied to the SCZP301 server by once again selecting **Only allow the server(s) specified above to be in the CTN** and then selecting **Apply** in the Network Configuration tab, as shown in Figure 2-40.

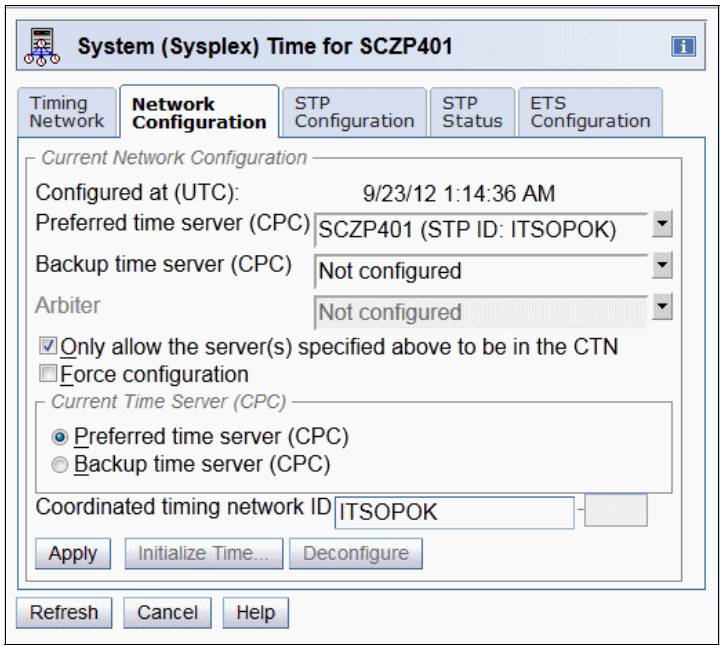

Figure 2-40 Saving the configuration for a single CEC

Once the MES of SCZP301 is complete and the two servers have reestablished coupling link connectivity, the following steps will have server SCZP301 back into the CTN and reestablish the original CTN configuration:

1. Deselect **Only allow the server(s) specified above to be in the CTN** and select **Apply.** This will have server SCZP301 join the CTN.
2. Configure server SCZP301 to be the BTS by selecting it from the BTS pull-down menu and select **Apply**.
3. Select **Only allow the server(s) specified above to be in the CTN** and select **Apply**. This will finally save the CTN configuration values. Note that the Network Configuration tab (see Figure 2-41) is now identical to that shown in the original configuration.

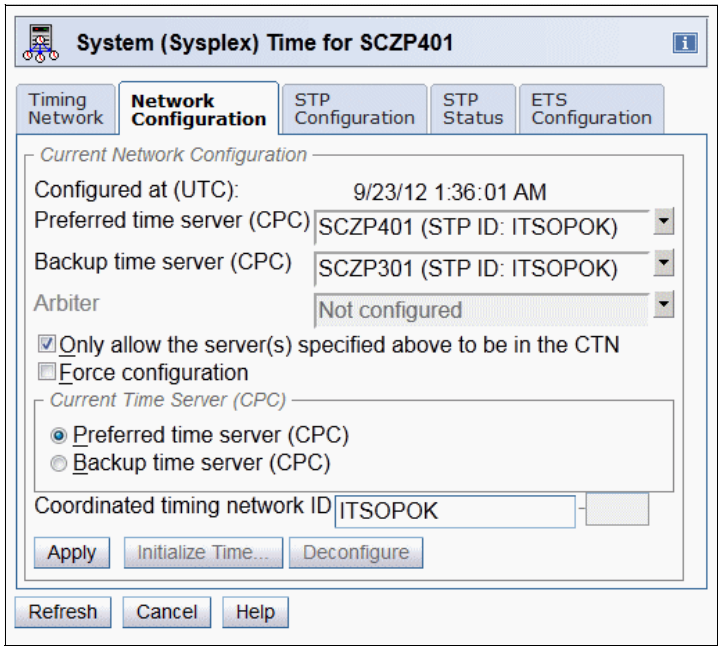

Figure 2-41 Saving the configuration after MES

2.6 External time source recovery

External time source (ETS) in an STP-only CTN can be provided by either:

- Using dial-out on the HMC (only available for HMC versions 2.11.1 and before)
- Using an NTP server (LAN connection)
- Using an NTP server with a pulse per second output option (LAN connection and coaxial cable to the PPS port of an ETR card)

Important: The ETS function provides time accuracy relative to an external time standard. Regardless of the ETS option selected, failures associated with ETS do not affect the capability of servers in a CTN to stay synchronized with each other. As long as the timing state of the servers remains synchronized, z/OS images that depend on synchronization are not affected. The only effect of unsuccessful recovery of an ETS failure is that the CTN will slowly drift away from ETS time.

Because there are no specific recovery actions when the ETS is configured to use a dial-out service, we only discuss recovery actions in an STP-only CTN using NTP servers or NTP servers with PPS as the external time source.

In this section we provide an overview of external time source recovery based on the most common ETS redundancy configurations, introduced in *Server Time Protocol Planning Guide,* SG24-7280. The recovery process depends on the implemented ETS option and the redundancy configured for that option. The STP design provides continuous availability of ETS, while maintaining the special roles of PTS and BTS you have assigned. If the PTS/CTS is not able to access the time information from its configured NTP servers, STP is able to steer the CTN using the calculated time adjustments from the BTS.

With z/OS 1.11 and later (rolled back to z/OS 1.9 and 1.10) the following z/OS message will show up indicating status changes with respect to the external time source for the CTN:

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = nn
```

The alert code `nn` uniquely identifies the STP alert condition.

2.6.1 ETS recovery using NTP servers

The Current Time Server (CTS) is the only server that adjusts the Coordinated Server Time (CST) by steering it to the time obtained from an external time source (ETS). To provide ETS redundancy, you should consider configuring two or more NTP servers, as described in *Server Time Protocol Planning Guide,* SG24-7280. Up to two NTP servers can be configured on each server in the STP-only CTN. When two NTP servers are configured, you are responsible for selecting the preferred NTP server, which is called the selected NTP server, while the other is called the non-selected NTP server.

Configured NTP servers on the PTS/CTS are accessed once every 10 minutes by their SNTP client. Once every hour, assuming a successful access of the selected NTP server, the SNTP client sends a CST adjustment to the STP facility. Normally, the SNTP client on the CTS uses the time information from the selected NTP server to perform the time adjustment. The time information from the non-selected NTP server is only used when there is a failure associated with accessing time information from the selected NTP server.

Configured NTP servers on the BTS are also accessed once every 10 minutes. The BTS calculates a value for time adjustment based on this access, and communicates the information to the PTS over the coupling links. If the PTS/CTS cannot access both its configured NTP servers, it will switch over to using the timing information sent from the BTS to steer the STP-only CTN.

Note: NTP client code runs on every SE of the zEC12, z196, z114, z10, and z9 servers that are in the STP-only CTN. However, only the NTP client of the PTS/CTS and BTS access the NTP servers configured to them; the NTP clients on the SEs of the other servers in the CTN do not.

Figure 2-42 on page 65 shows possible failures that can affect recovery when configured with an NTP server. They include:

- Loss of LAN connectivity between the Support Element and the NTP server
- Complete NTP server failure or bad NTP data from the NTP server

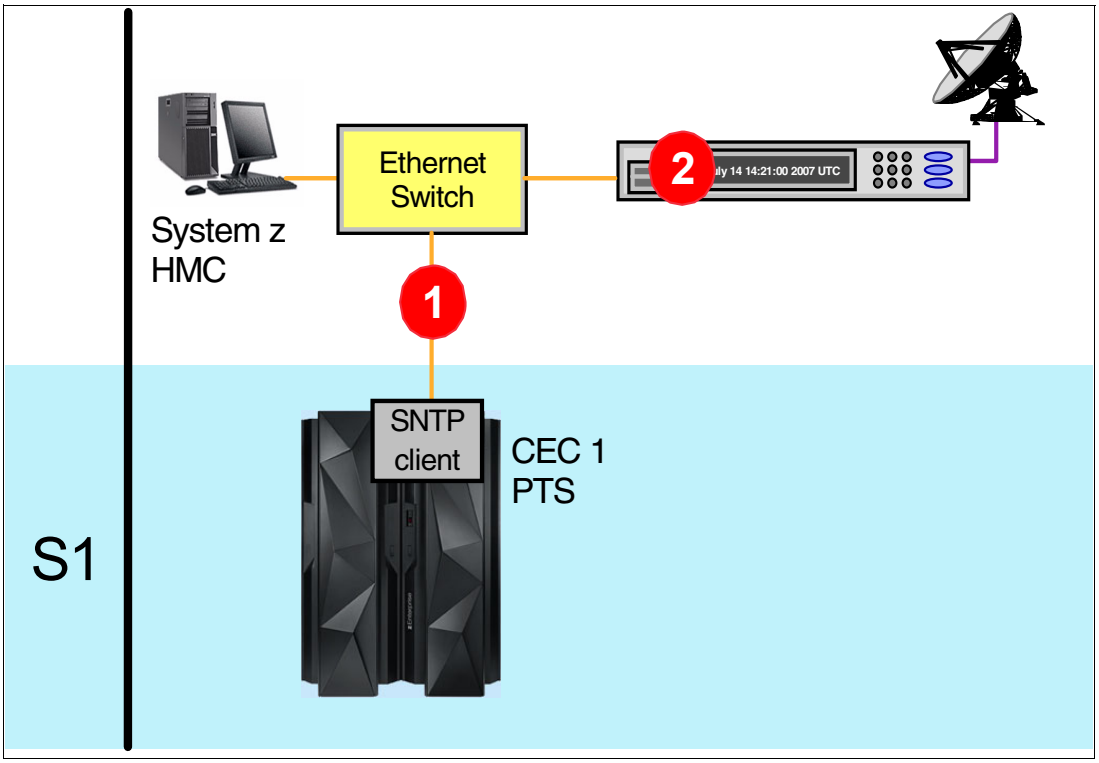

Figure 2-42 Possible failures affecting the NTP server

The following is the order in which a recovery action results in selecting a different NTP server than the selected NTP server on the PTS/CTS:

1. When the SNTP client has two unsuccessful attempts (two hours) at sending a CST adjustment value to the STP facility, based on valid NTP data from the selected NTP server, it will switch to sending timing adjustment information based on the nonselected NTP server.
2. When the SNTP client has two unsuccessful attempts (two hours) at sending a CST adjustment value to the STP facility, based on valid NTP data from the nonselected NTP server, STP will steer the CTN using the calculation from the BTS. The BTS information could be based on:
   - A selected NTP server at the BTS, if valid data can be accessed
   - A nonselected NTP server, if valid data cannot be accessed from the selected NTP server
3. When STP is not able to switch to any operational NTP server, the automatic base steering described in *Server Time Protocol Planning Guide,* SG24-7280 can continue, if the CTN was previously steered to an NTP Stratum 1 External Time Source. Base steering allows STP to compensate for the drift characteristics of the oscillator, thereby maintaining relatively good time accuracy at the Current Time Server, even if an ETS is not available.

Scenario 1: Redundant NTP servers on the PTS/CTS

In this section we describe ETS recovery scenarios for a CTN using redundant NTP servers configured on the PTS/CTS only.

Figure 2-43 on page 66 illustrates a configuration of two NTP servers configured on the PTS/CTS. NTP server 1 is the selected NTP server, NTP server 2 is the nonselected NTP

server. In this configuration, NTP server 2 is a Stratum 2 NTP server configured on the HMC. The HMC gets its time information from the Stratum 1 NTP server on the corporate network.

Under normal circumstances, only the time information provided by NTP server 1 is used for time adjustments of the CTN.

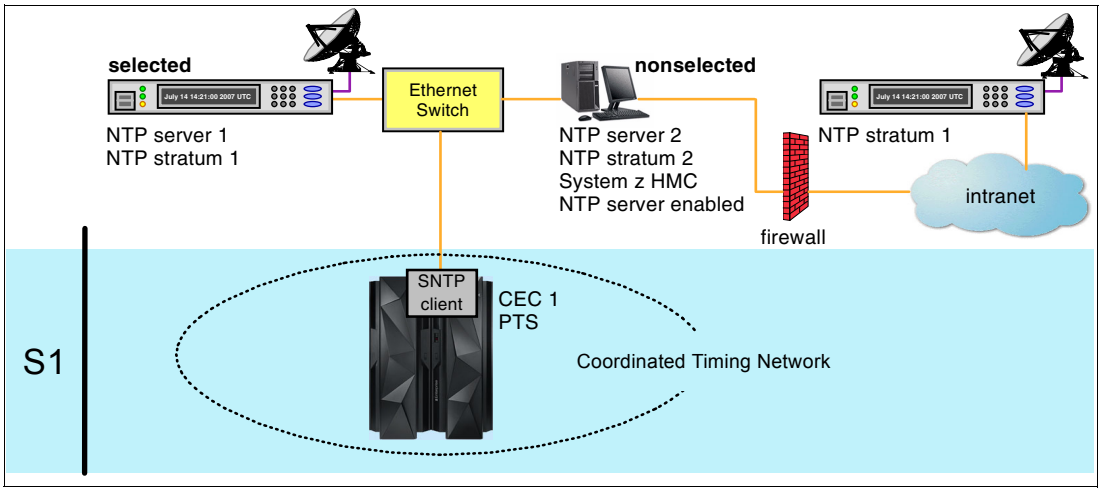

Figure 2-43 NTP server redundancy on PTS/CTS

Problem awareness

Refer to Figure 2-42 on page 65 for failures discussed in this section. If the failure is an NTP server 1 failure shown as (2), a hardware message is generated on the PTS/CTS after two hours indicating NTP server 1 could not be accessed. Also, the following message is posted:

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 0A
```

If the failure is a LAN access failure shown as (1), both NTP servers will not be accessible and therefore two hardware messages will be generated after two hours, one for NTP server 1 and one for NTP server 2. Also, the following z/OS message is issued:

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 06
```

A hardware message is also posted when changes in NTP stratum level or source ID are detected on either the selected or nonselected NTP server. The corresponding STP alert codes for message IEA031I are: 12 and 23 for the PTS, and 22 and 23 for the BTS.

Recovery

Refer to Figure 2-42 on page 65 for failures referred to in this section. If the selected NTP server becomes unavailable, but the nonselected NTP server is still available (failure 2), the SNTP client on the PTS/CTS will use the nonselected NTP server as its ETS, and will continue steering the CTN according to timing information received from NTP server 2.

However, in this configuration, the network connection between the SE and the Ethernet switch is a single point of failure for both NTP servers. So if the failure is a LAN failure, there is no recovery possible, and the CTN continues to use automatic base steering, as described in *Server Time Protocol Planning Guide,* SG24-7280.

For more detailed recovery information and the actions that must be taken in response to various failures, see the ETS recovery information in *“External Time Source recovery” on page 191*.

Scenario 2: Redundant NTP servers on PTS and BTS

In this section, we describe ETS recovery scenarios for a CTN using redundant NTP servers configured on the PTS/CTS and the BTS.

Figure 2-44 illustrates a configuration with two NTP servers. NTP server 1 is the selected NTP server on the PTS/CTS and nonselected NTP server on the BTS. NTP server 2 is the selected NTP server on the BTS and nonselected NTP server on the PTS.

Under normal circumstances, only the time information provided by NTP server 1 on the PTS is used for time adjustments of the CTN. The BTS also periodically accesses NTP server 2, calculates a value for time adjustment based on this access, and communicates the information to the PTS/CTS over the coupling links. The PTS/CTS uses this data to steer the CTN only if it cannot access NTP server 1 or NTP server 2.

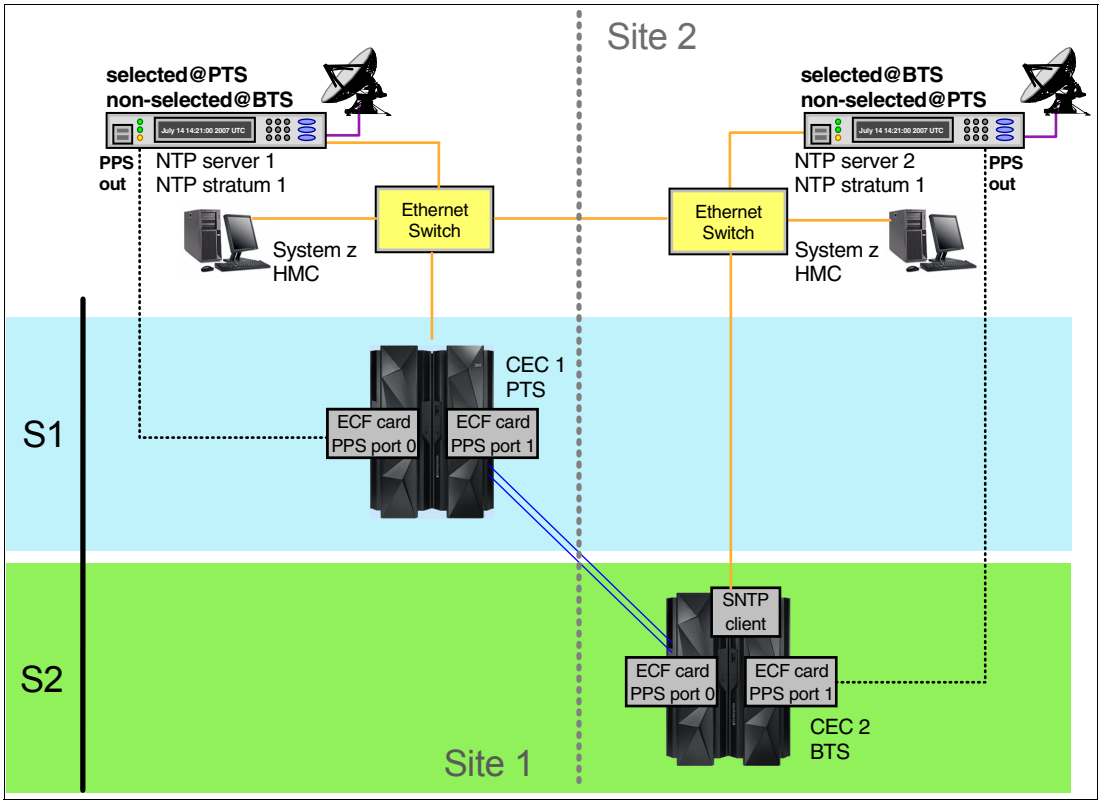

Figure 2-44 NTP server redundancy on PTS and BTS

Problem awareness

Refer to Figure 2-42 on page 65 for failures discussed in this section. If the PTS/CTS experiences failure 2, a hardware message is generated on the PTS/CTS after two hours, indicating NTP server 1 could not be accessed. If it experiences failure 1, a hardware message is generated that NTP servers 1 and 2 could not be accessed.

If the BTS experiences failure 2, a hardware message is generated on the BTS after two hours, indicating NTP server 2 could not be accessed. If it experiences failure 1, a hardware message is generated that NTP server 1 and 2 could not be accessed.

The following message is issued if no usable NTP-server can be accessed:

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 06
```

Recovery

If the PTS/CTS is not able to access NTP server 1 for two hours, the non-selected NTP server takes over the ETS role. If the PTS encounters a LAN failure and cannot reach any of the two NTP servers, it will start using the time adjustment information sent by the BTS approximately an hour later to steer the CTN.

If the BTS is not able to access NTP server 2 or NTP server 1 for two hours, then there is no recovery action. However, the problem should be corrected as soon as possible to maintain ETS redundancy.

Note: All configured NTP servers (on PTS/CTS and BTS) are accessed independent of one another. However, not every NTP server failure causes STP to switch to an alternate NTP server. This happens only if the NTP server currently being used to steer the CTN is affected.

For more detailed recovery information and the actions that must be taken in response to various failures, see the ETS recovery information in *“External Time Source recovery” on page 191*.

Compared to Scenario 1: Redundant NTP servers on the PTS/CTS, this configuration provides an additional degree of continuous availability of NTP servers, and is suitable for a single-site implementation and also for a dual-site implementation, with PTS and BTS at different sites.

To provide even more redundancy, you might also consider configuring an additional NTP server on the HMC, as shown in Figure 2-45 on page 69. The NTP server on the HMC is the nonselected NTP server at the PTS/CTS. If the selected NTP server fails at the PTS/CTS, the nonselected NTP server takes over the ETS role and provides the time information.

In case both NTP servers at Site 1 are not accessible for a certain period of time (for example because of LAN problems), the time adjustment information sent by the BTS will be used as mentioned previously.

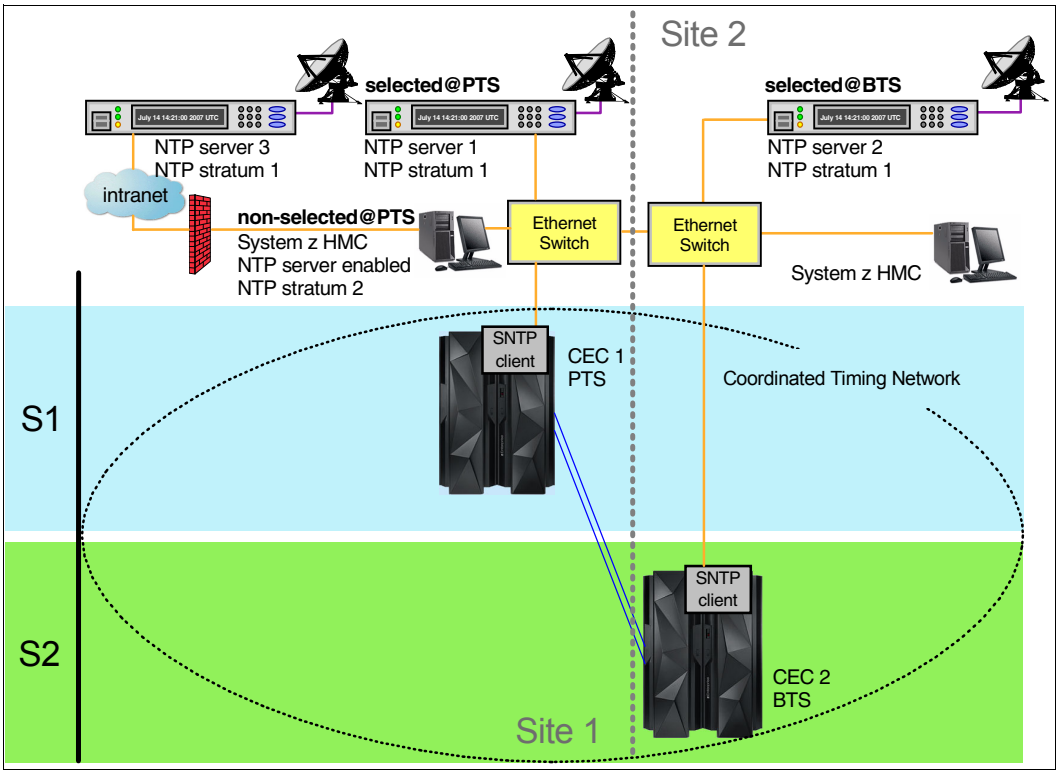

Figure 2-45 Continuous NTP server availability - enhanced configuration

2.6.2 ETS recovery using NTP servers with the PPS option

To provide ETS redundancy, you should consider configuring two or more NTP servers with PPS, as described in *Server Time Protocol Planning Guide,* SG24-7280. Up to two NTP servers with PPS can be configured on each server in the STP-only CTN. When two NTP servers with PPS are configured, you are responsible for selecting the preferred NTP server with PPS. This NTP server with PPS is called the selected NTP server with PPS, the other is called the nonselected NTP server with PPS. As stated in *Server Time Protocol Planning Guide*, SG24-7280, the PPS output of the same NTP server has to be connected to the PPS input provided on the External Time Reference (ETR) card of the System z10 or System z9 server, or to the PPS ports on the ECF/FSP card of the System zEC12, z196 and z114.

Note: To improve readability, for the remainder of this section an NTP server with PPS is simply referred to as an NTP server. This should not be confused with the option discussed in 2.6.1, “ETS recovery using NTP servers” on page 64 where the NTP server is being used without the PPS option.

Configured NTP servers on the PTS/CTS are accessed once a minute by the SNTP client. Once every 10 minutes, assuming successful access of both NTP servers, the SNTP client sends time adjustment information based on both NTP servers to the STP facility. Normally, the STP facility on the PTS/CTS uses the time information in conjunction with the PPS signal from the selected NTP server to perform the time adjustment. The time information or PPS signal from the nonselected NTP server is only used when there is a failure associated with accessing time information or PPS signals from the selected NTP server.

Configured NTP servers on the BTS are also accessed once a minute by the SNTP client, and time adjustment information based on both NTP servers sent to the STP facility on the

BTS. Normally, the STP facility on the BTS uses the time information in conjunction with the PPS signal from the selected NTP server to calculate a time adjustment. The BTS then communicates this information to the PTS over the coupling links. The time information or PPS signal from the nonselected NTP server on the BTS is only used for this calculation when there is a failure associated with accessing time information or PPS signals from the selected NTP server. If the PTS/CTS cannot access both its configured NTP servers, it will switch over to using the timing information sent from the BTS to steer the STP-only CTN.

Possible failures that can affect an NTP server with PPS include (see Figure 2-46):

- Loss of LAN connectivity between the Support Element and the NTP server or bad NTP data; in this case, good PPS signals are still received by the PPS port on the ETR card.
- PPS signal not received by the PPS port on the ETR card; in this case, valid NTP data from the NTP server is still available over the LAN.
- Complete NTP server failure affecting both the NTP data and the PPS output of the NTP server.

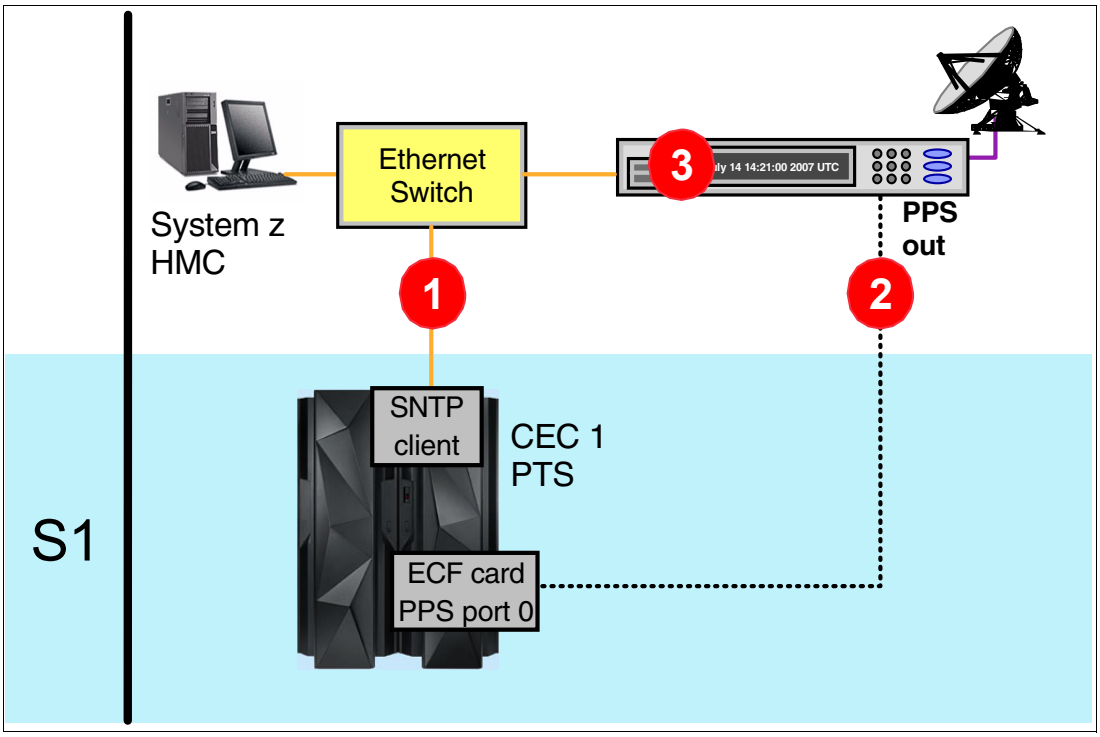

Figure 2-46 Possible failures affecting the NTP server with PPS

Assuming two NTP servers with PPS are configured on both the PTS/CTS and the BTS, and STP is tracking to the PPS signals received by the PPS port associated with the selected NTP server on the PTS/CTS:

- If there is a failure type (1), STP continues using the PPS signals received on the PPS port of the selected NTP server on the PTS/CTS.
- If there is a failure type (2) or (3), STP switches to using NTP data and PPS signals from the nonselected NTP server on the PTS/CTS.

Assuming only one NTP server with PPS is configured on the PTS/CTS and the BTS, and STP is tracking to the PPS signals received by the PPS port associated with the selected NTP server on the PTS/CTS:

- If there is a failure type (1), STP continues using the PPS signals received on the PPS port of the selected NTP server on the PTS/CTS.
- If there is a failure type (2) or (3), STP switches to using the time adjustment information received from the BTS, which could be based on:
  - A selected NTP server at the BTS, if valid data and PPS signals can be accessed
  - A nonselected NTP server, for failure type (2) or (3) on the BTS

For failures regarding the PPS signal, message IEA031I is posted with a reason code indicating the alert condition.

It should be noted that regardless of the specific redundancy provided by an NTP server with PPS configuration:

- If PPS signals are not received by any of the configured NTP servers on the PTS/CTS and the BTS, but valid NTP data is available, then STP continues using the NTP data for steering the CTN, following the same recovery flow described in 2.6.1, “ETS recovery using NTP servers” on page 64.
- When STP is not able to switch to any operational NTP server, the automatic base steering described in *Server Time Protocol Planning Guide,* SG24-7280, can continue. Base steering allows STP to compensate for the drift characteristics of the oscillator, thereby maintaining relatively good time accuracy at the Current Time Server, even if an ETS is not available.

Scenario 1: Redundant NTP servers with PPS on the PTS/CTS

In this section, we describe ETS recovery scenarios for a CTN using redundant NTP servers with PPS configured on the PTS/CTS only.

In the configuration shown in Figure 2-47 on page 72, NTP server 1 and NTP server 2 are both configured on the PTS/CTS. The NTP output of both NTP servers is connected to the SE LAN. The PPS output of NTP server 1 is connected to PPS port 0 of the ECF/FSP card, and the PPS output of NTP server 2 is connected to PPS port 1. NTP server 1 is the selected NTP server, and NTP server 2 is the nonselected NTP server.

In normal operation, the STP facility on the PTS/CTS uses the time information in conjunction with the PPS signal from the selected NTP server to perform the time adjustment and steer the CTN. Once the CTN is tracking to the PPS signal, the SE LAN is no longer a single point of failure because STP uses the PPS signal to steer the time in the CTN.

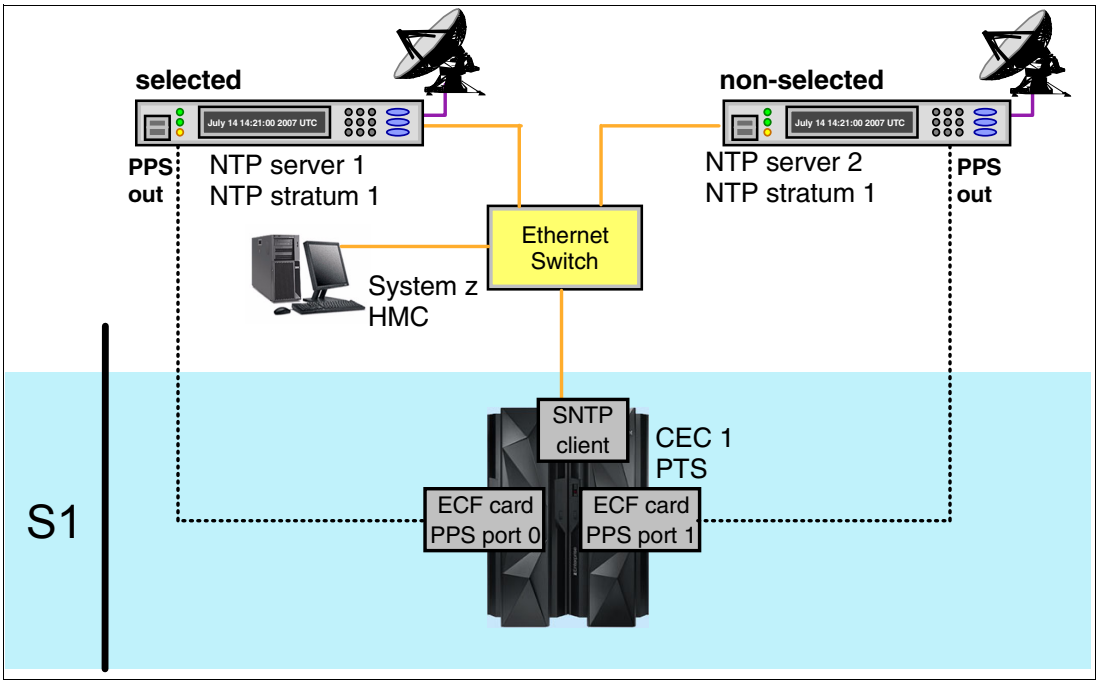

Figure 2-47 NTP server with PPS redundancy on PTS/CTS

Problem awareness

Refer to Figure 2-46 on page 70 for failures discussed in this section. If the failure is a LAN access failure shown as (1), it will most likely impact access to both NTP server 1 and 2. Two hardware messages will be generated after twenty minutes—one for NTP server 1 and one for NTP server 2.

If the failure is associated with just the PPS signal from NTP server1, shown as (2), or an NTP server 1 failure shown as (3), a hardware message is posted after approximately two minutes, indicating NTP server 1 failure.

Message IEA031I will pop up on the z/OS images with STP alert codes 83, 84, and 81 indicating the ETS error conditions.

Recovery

If NTP server 1 is not accessible by the SNTP client on the SE (failure 1), but the PPS signal is still received on PPS port 0, no recovery is required because STP will continue to steer the CTN using the PPS signals from NTP server 1.

For failures 2 and 3 on NTP server 1, STP switches to using the time information and the PPS signals from the nonselected server, NTP server 2.

Important: In normal operation, if the PTS/CTS continues to receive a PPS signal from the selected PPS port of the NTP server, there is no recovery action, even if the SE does not receive valid NTP data from the selected NTP server over the LAN.

For more detailed recovery information and the actions that must be taken in response to various failures, see the ETS recovery information in Chapter 6, “External Time Source recovery” on page 191.

Scenario 2: Redundant NTP servers with PPS on PTS and BTS

In this section, we describe ETS recovery scenarios for a CTN using redundant NTP servers configured on the PTS/CTS and the BTS.

Figure 2-48 illustrates a configuration with two NTP servers. NTP server 1 is the selected NTP server on the PTS/CTS; its PPS output is connected to PPS port 0 of the ECF/FSP card on the PTS/CTS. NTP server 2 is the selected NTP server on the BTS; its PPS output is connected to PPS port 1 of the ECF/FSP card on the BTS. Due to PPS cable length limitations, NTP servers must be located in the vicinity of their associated server.

Under normal circumstances, only the time information provided by NTP server 1 and its PPS signals are used for time adjustments of the CTN. The BTS also periodically accesses NTP server 2, and calculates a value for time adjustment based on this access and the PPS signals it receives. The time adjustment is communicated to the PTS/CTS over the coupling links. The PTS/CTS uses this data to steer the CTN only if it does not receive PPS signals from NTP server 1.

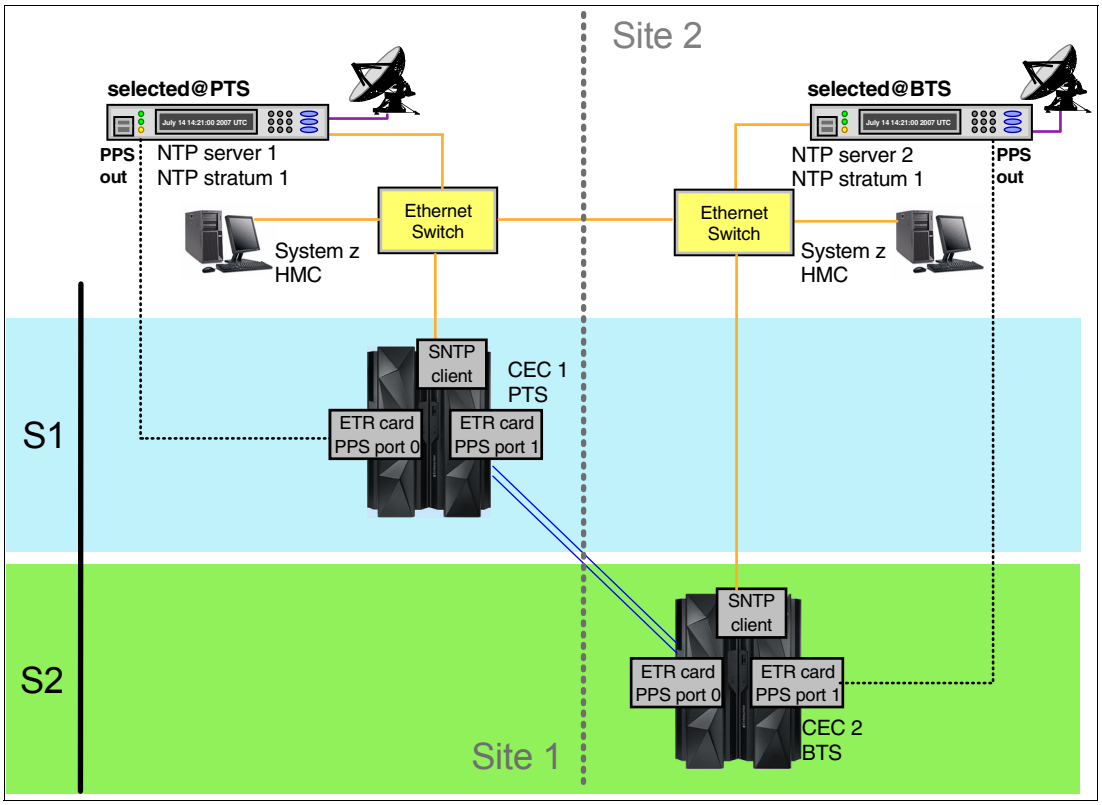

Figure 2-48 Two NTP servers, one on PTS, one on BTS, both using PPS

The scenario presented in this section illustrates a dual-site configuration, but it can also be used for a single-site implementation.

Problem awareness

Refer to Figure 2-46 on page 70 for failures discussed in this section. If the PTS/CTS experiences a LAN access failure shown as (1), a hardware message is generated after 20 minutes for NTP server 1. Similarly, if the BTS experiences a LAN access failure, a hardware message is generated after 20 minutes for NTP server 2.

If the failure is associated with just the PPS signal from NTP server 1, shown as (2), or an NTP server 1 failure shown as (3), a hardware message is posted on the PTS/CTS after

approximately two minutes, indicating NTP server 1 failure. In addition to that, z/OS message IEA031I with STP alert codes 83 and 82 is issued. For similar failures associated with NTP server 2, a hardware message is posted on the BTS after approximately two minutes.

Recovery

If NTP server 1 is not accessible (failure 1), but the PPS signal is still received on PPS port 0, no recovery is required because STP will continue to steer the CTN using the PPS signals from NTP server 1.

For failure scenarios 2 and 3 on NTP server 1, the PTS/CTS starts using the time adjustment information received from the BTS, which is based on NTP server 2 and its PPS signals.

Important: In normal operation, if the PTS/CTS continues to receive PPS signals from the selected PPS port of the NTP server, there is no recovery action, even if the SE does not receive valid NTP data from the selected NTP server over the LAN.

For failure scenarios 1, 2, or 3 on NTP server 2, no recovery is required.

For more detailed recovery information and the actions that must be taken in response to various failures, see the ETS recovery information in Chapter 6, “External Time Source recovery” on page 191.

2.6.3 ETS recovery summary

Here are some points to remember when planning and implementing ETS redundancy. The term “NTP server” is used generically here to refer to both ETS options (NTP server and NTP server with PPS) unless specifically noted that a point applies only to the NTP server with the PPS option.

- Failures associated with ETS and possible recovery actions do not affect the capability of servers in a CTN to stay synchronized with each other. z/OS systems that depend on synchronization are not impacted if the ETS is not available. The only effect of unsuccessful recovery for an ETS failure is that the CTN will slowly drift away from ETS time.
- Up to two NTP servers can be configured on each server in the STP-only CTN. When two NTP servers are configured, you are responsible for selecting the preferred NTP server. This NTP server is called the selected NTP server, while the other is called the nonselected NTP server.
- NTP client code runs on every SE of the zEC12, z196, z114, z10, and z9 servers that are in the STP-only CTN. However, only the NTP client of the PTS/CTS and BTS access the NTP servers configured to them; the NTP clients on the SEs of the other servers in the CTN do not.
- The Current Time Server (CTS) is the only server that adjusts the Coordinated Server Time (CST) by steering it to the time obtained from an external time source (ETS). Either the PTS or the BTS can be the CTS.
- We suggest to configure at least one unique NTP server or NTP server with PPS on the PTS and the BTS. Configuring an NTP server on the BTS provides two benefits:
  - Access to an NTP server when the BTS becomes the CTS as the result of planned or unplanned recovery
  - Time adjustments to an NTP server when the PTS/CTS cannot access any of its NTP servers

- All configured NTP servers (on PTS/CTS and BTS) are accessed independent of each other. However, not every NTP server failure or PPS failure causes STP to switch to an alternate NTP server. This happens only if the NTP server or PPS failure currently being used to steer the CTN is affected.
- Hardware messages are posted for the failure condition even if there was no recovery action that resulted in switching to an alternate NTP server.
- z/OS message IEA031I is issued on all systems when changes with respect to the ETS occur, even when there is no recovery action necessary.
- If an NTP server with PPS is configured on the PTS/CTS, the CTN can continue to track to the PPS signals even if there are failures that result in no valid NTP data being received by the SE.
- The following is the order STP selects to steer the CTN when NTP servers with PPS are configured. If one of the items listed is either not configured or has failed, STP will attempt the next item in the list:
  a. PPS signals from the selected NTP server configured on the PTS/CTS
  b. PPS signals from the nonselected NTP server configured on the PTS/CTS
  c. PPS signals from the selected NTP server configured on the BTS
  d. PPS signals from the nonselected NTP server configured on the BTS
  e. NTP timing information from the selected NTP server configured on the PTS/CTS
  f. NTP timing information from the nonselected NTP server configured on the PTS/CTS
  g. NTP timing information from the selected NTP server configured on the BTS
  h. NTP timing information from the nonselected NTP server configured on the BTS
- Multisite CTN configurations do not have any specific ETS redundancy considerations, other than the general suggestion to configure an NTP server both on the PTS and the BTS. Normally the BTS is in a different site than the PTS. When the PTS and BTS are in different sites, and each site has its own unique NTP server, it is suggested that the NTP server in the other site is configured as the nonselected NTP server. In other words, the PTS/CTS should have the NTP server located in the site with the BTS as its nonselected NTP server, and the BTS should have the NTP server located at the site with the PTS/CTS as its nonselected NTP server.
- The CTS assignment does not change as a consequence of an ETS failure. When STP is not able to switch to any operational NTP server, the CTN continues to use the adjustment value from the last successful NTP server query for steering. In addition, the automatic base steering (see *Server Time Protocol Planning Guide,* SG24-7280) can continue. This allows STP to compensate for the drift characteristics of the oscillator and to maintain time accuracy at the Current Time Server.

- The ETS redundancy configurations presented in this section are examples and not necessarily the only possible ones. Various other combinations can be implemented:
  - You can have two NTP servers on both the PTS and the BTS.
  - You can have an NTP server with PPS on the PTS/CTS and only an NTP server (without PPS) on the BTS, as shown in Figure 2-49.

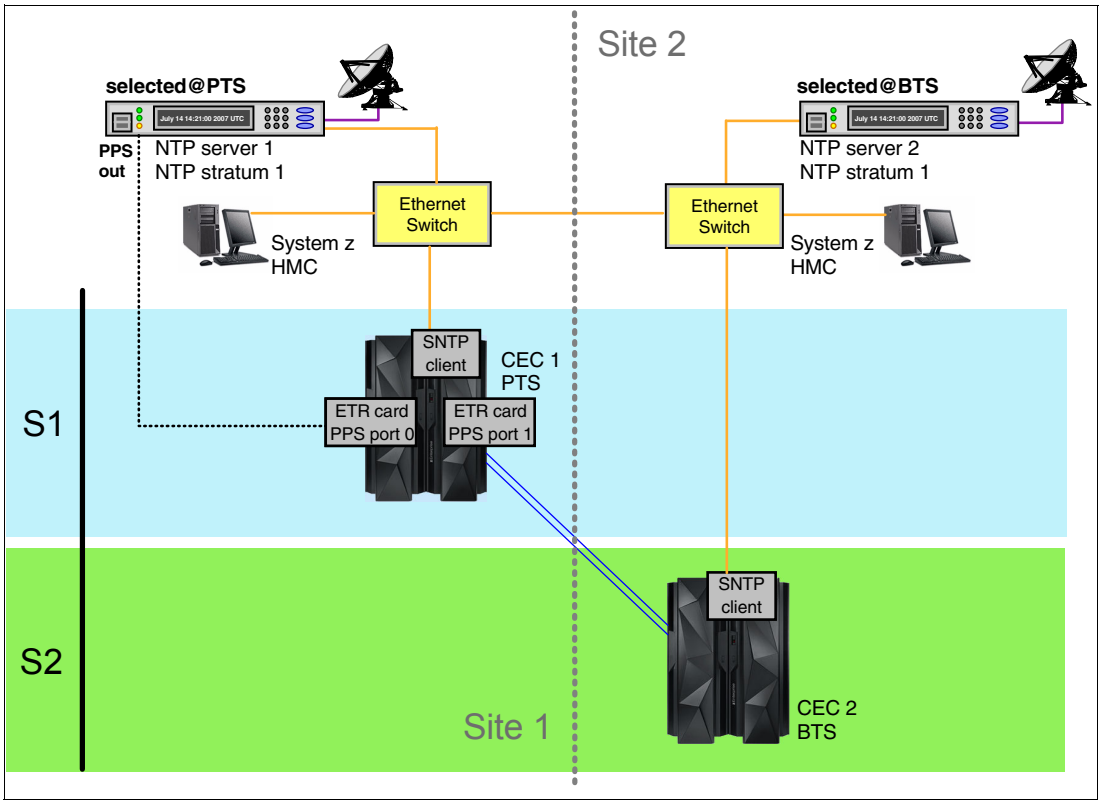

Figure 2-49 Dual-site ETS redundancy with mixed NTP servers

Part 2

Implementation

The chapters in this part describe STP recovery for different CTN-related failures. Each failure scenario is discussed in terms of:

Description A general description of the recovery scenario and sample configuration.

Problem awareness The external symptoms that alert you to the problem. It presents the z/OS messages that are issued, and describes the indications that are seen at the System (Sysplex) Time tabs on the Hardware Management Console.

Problem determination The tools and procedures available to assist in identifying the cause of the problem.

User actions The actions required to recover from, or to minimize the impact of, the CTN-related failure. The appropriate user actions vary depending on the CTN configuration.

Throughout this chapter, most of the examples described are topology variations using the base configuration illustrated in Figure 2-50.

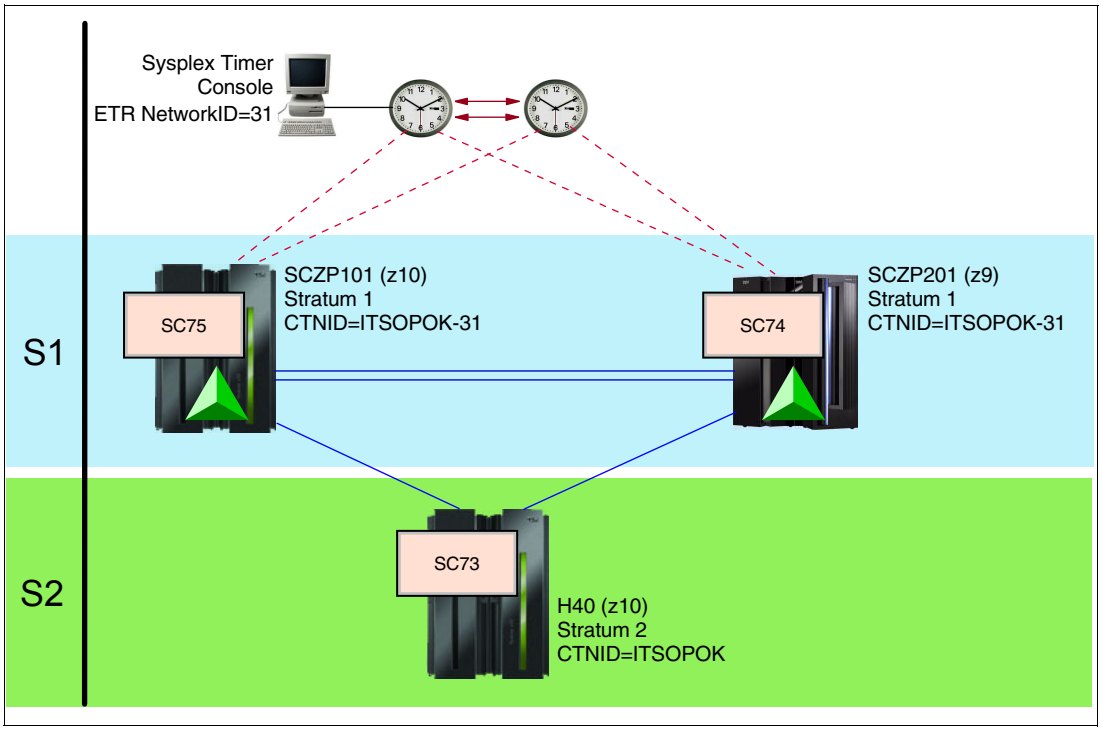

Figure 2-50 Base test configuration

Where:

- All servers are configured in one Coordinated Timing Network.
- The HMC has Ethernet connections to Support Elements for all servers in the timing network.

One Parallel Sysplex with three z/OS images and two coupling facility logical partitions spans the CTN. All z/OS system images are Version 1.11 with STP maintenance applied, and STPMODE YES is the default in the active CLOCKxx member.

3

Recovery in a Mixed CTN

A Mixed CTN consists of servers that are STP-configured, but the time source is the Sysplex Timer. Because all servers connected to the Sysplex Timer are Stratum 1, there can be multiple Stratum 1 servers in a Mixed CTN. These servers are in ETR timing mode. Servers can be indirectly synchronized to the time source as Stratum 2 or Stratum 3 servers. The Stratum 2 and Stratum 3 servers are operating in STP timing mode and are not directly connected to the Sysplex Timer.

We discuss the following topics in this chapter:

- Overview
- Coupling link failure, single link
- STP Stratum 1 failure
- Two-site: Site 1 failure
- Two-site: Sysplex Timers failures
- Two-site: Site 2 failure
- Two-site: loss of communication

3.1 Overview

In this section, recovery scenarios for a Mixed CTN are discussed. All examples assume that the latest MCL maintenance level has been applied and the maximum STP version available on servers that are attached in STP Timing Mode is STP Version 4.

For a CTN configuration with more than two servers, we suggest to attach at least two servers to the Sysplex Timer in ETR Timing Mode to avoid a single point of failure.

3.2 Stratum 1 loss of ETR synchronization

In a Mixed CTN, recovery of the failure of one ETR link or port is similar to that in an ETR network. However, recovery is different if both links fail, because it is now possible to recover from the failure of both ETR links or ports. This scenario is shown in Figure 3-1.

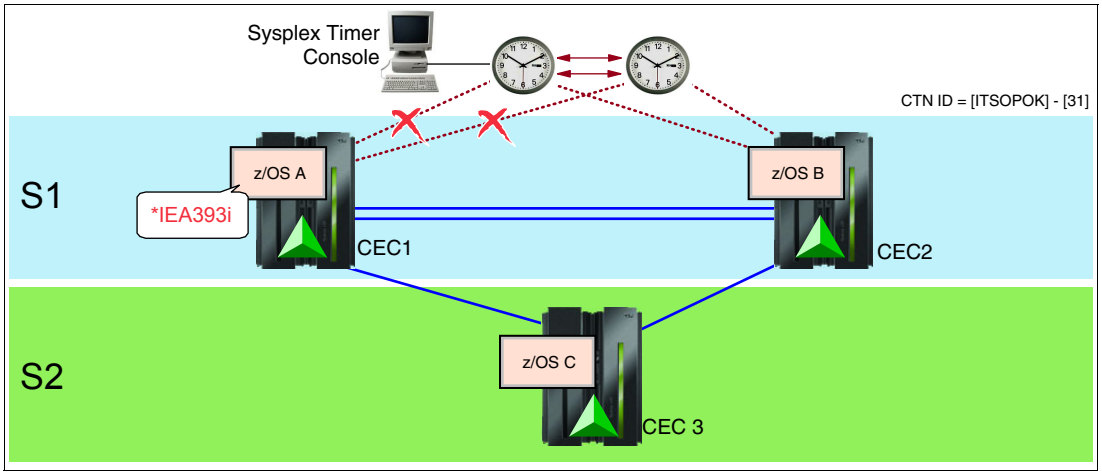

Figure 3-1 Failure of both ETR links or ports

A server selects which ETR port to use for synchronization and, if communication is lost on the selected port, the server automatically switches to the alternate port.

Failure of a single ETR link or port causes z/OS images with ETRMODE YES to issue message IEA393I, as shown in Figure 3-2 on page 81. The same message is issued for a Sysplex Timer port failure, for a server ETR port failure, or for a link failure.

If the alternate port is connected to the correct ETR Network ID, the system will continue in ETR synchronization mode. Otherwise, full connectivity to the Sysplex Timer is lost.

In a Mixed CTN, complete loss of connectivity between a Stratum 1 server and its Sysplex Timers may not result in a server entering local TOD stepping mode, as would be the case in an ETR network. If sufficient connectivity is in place and a valid CTN ID is defined, the Stratum 1 server will transition to STP timing mode as a Stratum 2. Such a transition has no impact on the resident z/OS system images or Coupling Facilities.

Figure 3-1 describes a failure of both ETR links or ports. If SCZP101 stops receiving ETR signals on both of its ports, it will become a Stratum 2 server, receiving time information through the coupling links from the Stratum 1 SCZP201. It can also receive STP timing messages from the server H40. However, because the algorithms favor selection of a Stratum 1 over a Stratum 2, it is likely that STP messages from SCZP201 will be selected.

3.2.1 Problem awareness

The key highlights of this scenario are:

- Message IEA393I appears when an ETR link or port fails:

  ```
  IEA393I ETR PORT n IS NOT OPERATIONAL. THIS MAY BE A CTN CONFIGURATION CHANGE
  ```

  If both ETR ports or links fail, z/OS message IEA393I is displayed twice, once for each port, as shown in Figure 3-2.

```
*IEA393I ETR PORT 0 IS NOT OPERATIONAL. THIS MAY BE A CTN CONFIGURATION CHANGE.
*IEA393I ETR PORT 1 IS NOT OPERATIONAL. THIS MAY BE A CTN CONFIGURATION CHANGE.
```

Figure 3-2 Dual ETR link or port failure: message IEA393I

- SCZP101 transitions from Stratum 1 to Stratum 2. The transition has no impact on z/OS system image SC75 or on the Coupling Facility partition.

3.2.2 Problem determination

The status of the Sysplex Timer connections should be determined using the DISPLAY ETR command, and the ETR Status and STP Status tabs of the System (Sysplex) Time task on the HMC.

In a single ETR link or port failure, there should be a single OPERATIONAL and ENABLED server ETR port, and a single NONOPERATIONAL and ENABLED server ETR port for each operating system. Figure 3-3 shows the response from the DISPLAY ETR command.

```
DISPLAY ETR
IEA282I 18.29.03 TIMING STATUS 050
SYNCHRONIZATION MODE = ETR
  CPC PORT 0       ACTIVE ==>   CPC PORT 1
  NONOPERATIONAL                OPERATIONAL
  ENABLED                       ENABLED
                                ETR NET ID=31
                                ETR PORT=01
                                ETR ID=00

ETR DATA COULD NOT BE OBTAINED FOR CPC PORT 0
THIS SERVER IS PART OF TIMING NETWORK ITSOPOK-31
```

Figure 3-3 Single ETR link or port failure: DISPLAY ETR and Port 0 not operational

In the case of a dual ETR link or port failure, there will be two NONOPERATIONAL and ENABLED server ETR ports for each operating system. As shown in Figure 3-1 on page 80, SCZP101 stops receiving ETR signals on both its ports and becomes a Stratum 2 server receiving time information through the coupling links from SCZP201 Stratum 1.

Message IEA380I is displayed on z/OS system image SC75 indicating that SCZP101 is now receiving time information using STP:

```
IEA380I THIS SYSTEM IS NOW OPERATING IN STP TIMING MODE.
```

The ETR Status Word section of the ETR Status tab contains information describing the state for each ETR port. The state of a port describes its ability to communicate with its host system. In Figure 3-4, the state of Port 0 and Port 1 in a dual ETR link or port failure is “Loss of light”, which indicates that the ports are not receiving an optical signal from the Sysplex Timer.

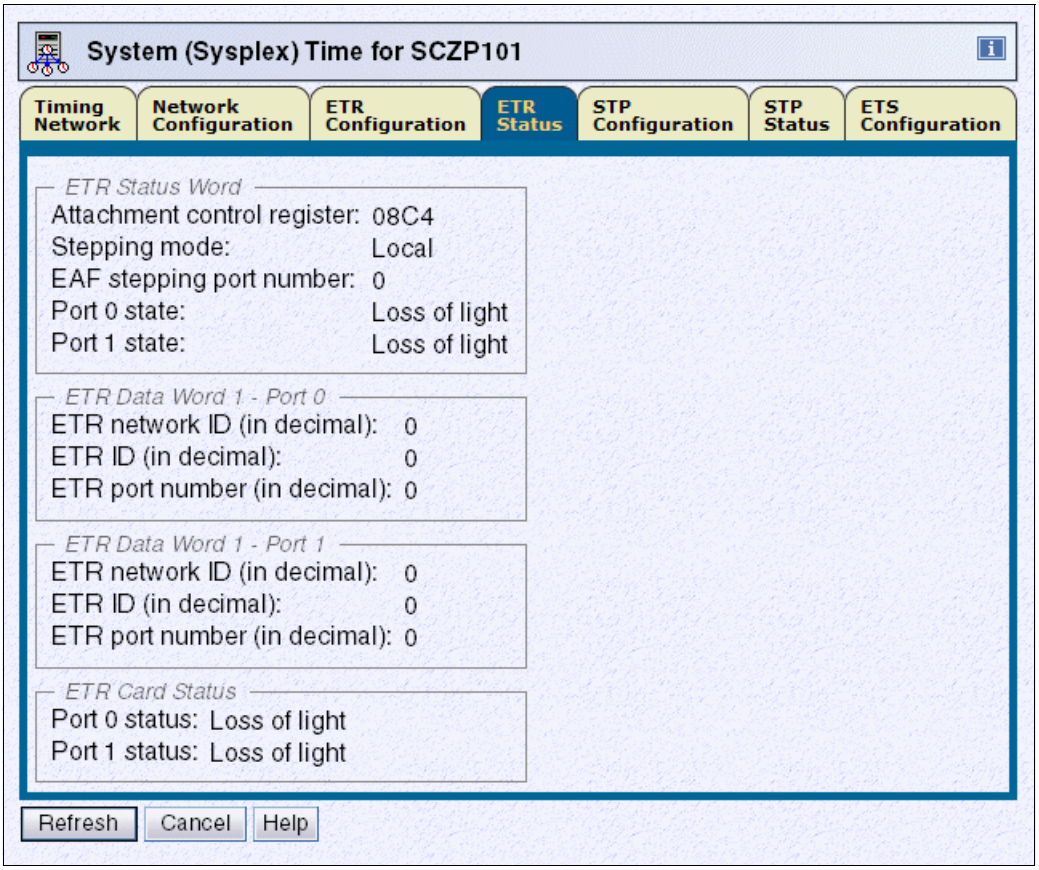

Figure 3-4 DUAL ETR link or port failure: ETR Status tab

In Figure 3-1 on page 80, SCZP101 was originally a Stratum 1 server using ETR timing mode. Figure 3-5 on page 83 shows the STP Status tab of SCZP101:

- Timing mode is now STP.
- Stratum level has changed to 2.

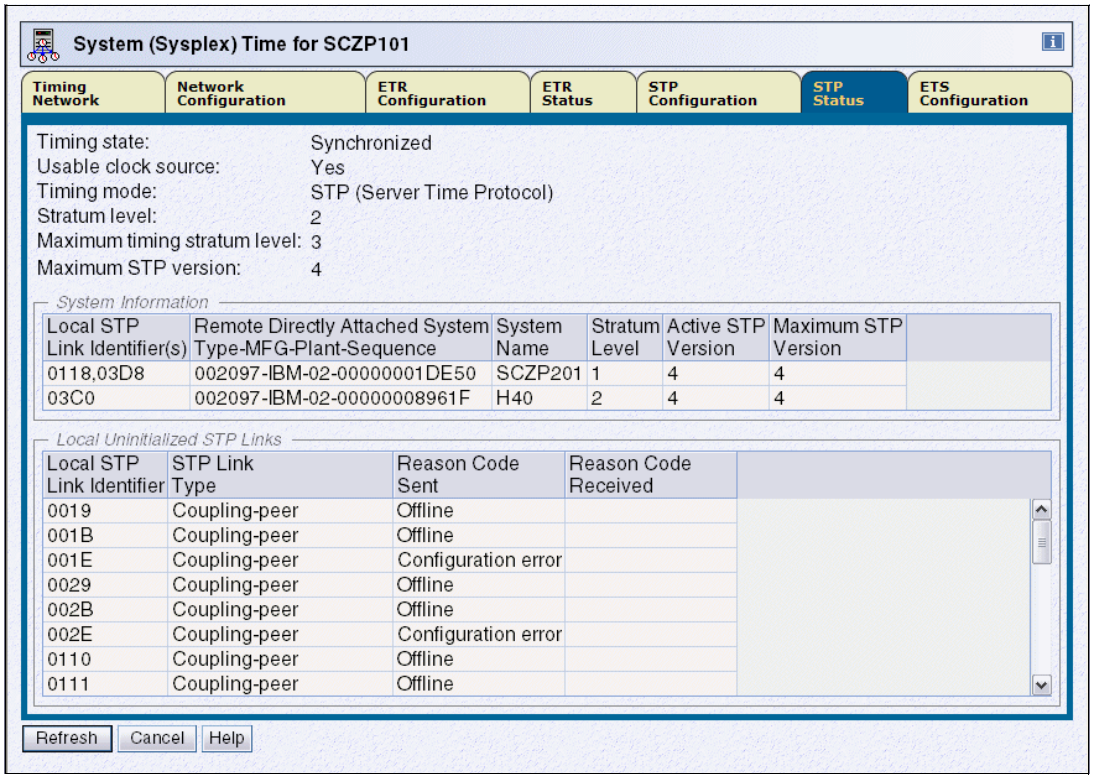

Figure 3-5 STP Status tab: after dual port or link failure

In a Mixed CTN configuration, potential causes of loss of ETR port availability (single or both ports) for a Stratum 1 include:

- ETR port or Sysplex Timer port or link failure
- Failure of the ports in the CLO card in either Sysplex Timer unit
- Failure of both CLO links between the Sysplex Timers
  - Failure of repeater hardware when extended distances are involved
  - Loss of power to repeater hardware

3.2.3 User actions

Loss of one ETR link or port communication does not have any impact on z/OS system images if the availability recommendations have been followed. In general, all attached system images should have an enabled connection to each Sysplex Timer unit in an Expanded Availability configuration.

When the link or port communication is restored, z/OS message IEA267I is issued, once for each server ETR port that has been restored. For example, when both failing ETR links or ports are restored, message IEA267I is displayed twice, as shown in Figure 3-6.

```
*IEA267I ETR PORT 0 IS NOW AVAILABLE
*IEA267I ETR PORT 1 IS NOW AVAILABLE
```

Figure 3-6 ETR ports available after communication is re-established

The status of all ETR connections should also be verified using the DISPLAY ETR command at the z/OS console. All ports should be both OPERATIONAL and ENABLED, as shown in Figure 3-7.

```
DISPLAY ETR
IEA282I 19.01.21 TIMING STATUS 079
SYNCHRONIZATION MODE = ETR
  CPC PORT 0       ACTIVE ==>   CPC PORT 1
  OPERATIONAL                   OPERATIONAL
  ENABLED                       ENABLED
  ETR NET ID=31                 ETR NET ID=31
  ETR PORT=01                   ETR PORT=01
  ETR ID=01                     ETR ID=00
  THIS SERVER IS PART OF TIMING NETWORK ITSOPOK-31
```

Figure 3-7 Single ETR link or port failure: DISPLAY ETR after recovery and user action

As shown in Figure 3-8, the ETR Status tab shows that the recovery actions were successful, because the state of both Port 0 and Port 1 is now Operational.

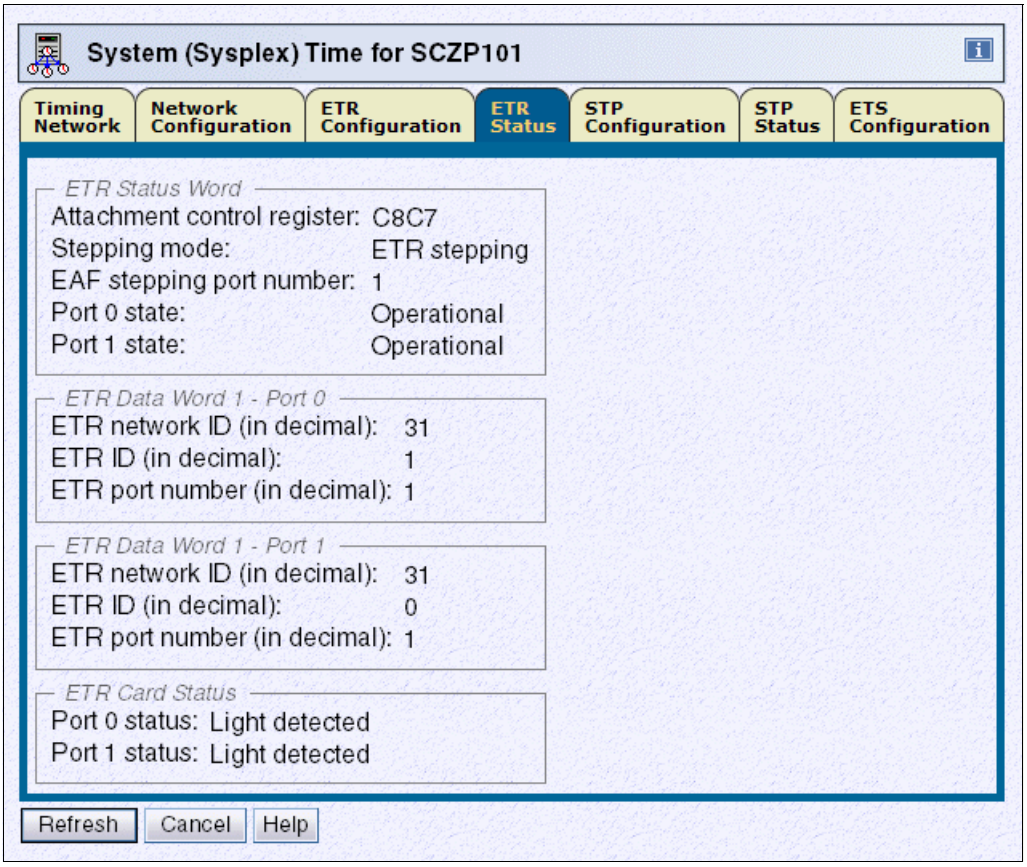

Figure 3-8 ETR Status tab: after recovery and user action

As shown in Figure 3-9 on page 85, the STP Status tab indicates that with ETR connectivity restored, SCZP101 has:

- Timing mode returned to ETR.
- Stratum level returned to 1.

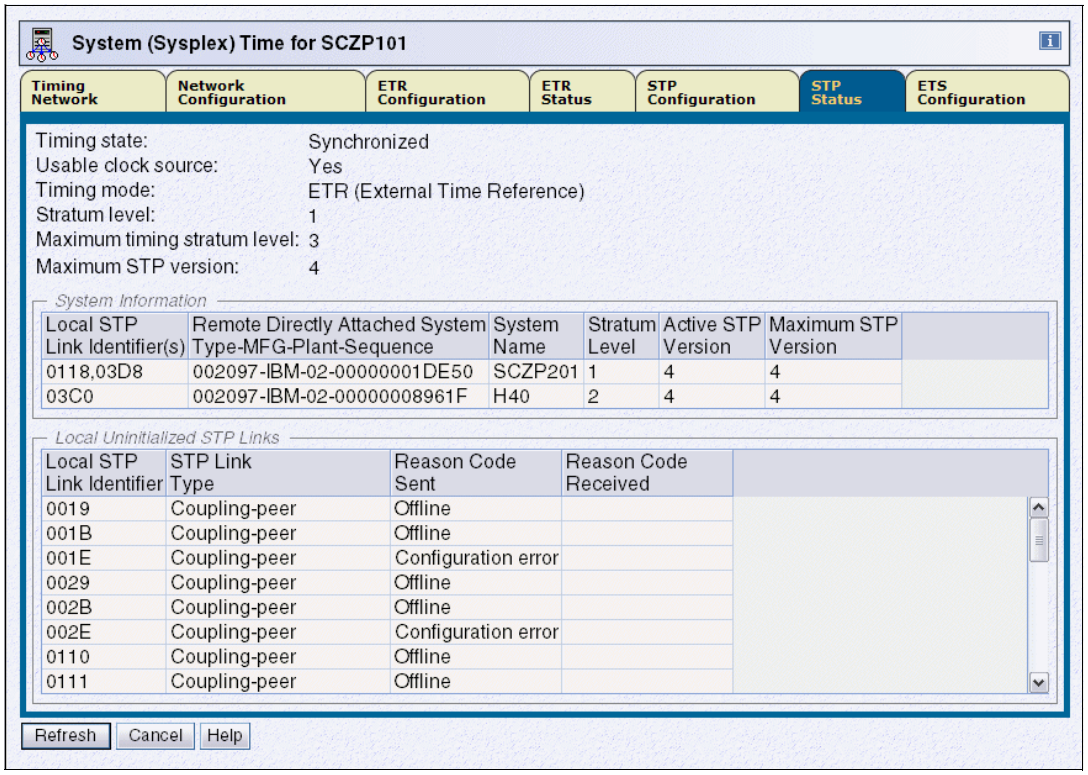

Figure 3-9 STP Status tab: after recovery and user action

3.3 Coupling link failure, single link

Coupling links are used to exchange STP messages between servers. In a Mixed or STP-only CTN, only one link between any two servers is selected for STP messages. If this particular link fails, one of the other remaining links is used with no impact on the CTN.

Note: Only one coupling link is used at a given time to exchange STP messages between any two servers. However, when multiple links are available, STP may change from one link to another.

In Figure 3-10 on page 86, there are multiple coupling links available between Stratum 1 servers SCZP101 and SCZP201. The links are not used for synchronization purposes because the time source for both servers is the Sysplex Timer. STP messages exchanged on these links are used to keep a history of timing packets to allow instant transition to Stratum 2 (if needed), should a failure occur.

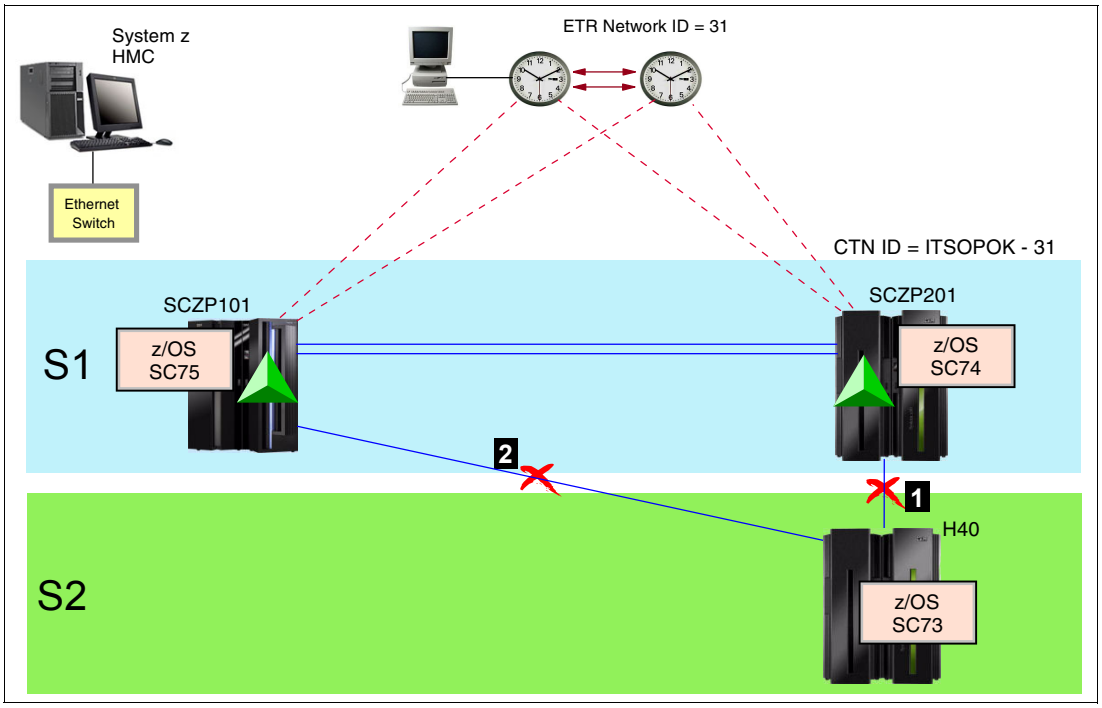

Figure 3-10 Coupling link failure

STP messages are also being exchanged between H40, Stratum 2 and both SCZP101 and SCZP201 Stratum 1, through single coupling links.

When a single coupling link fails among multiple coupling links, there will be no indication of a problem, because one of the remaining links will be selected. However, if the failure of a coupling link leaves a server with only a single remaining link to use as a timing source, z/OS information message IEA382I is issued on z/OS images that reside on server H40.

In this scenario, a failure of coupling link 1 is described (see Figure 3-10). The key highlights of this scenario are:

- Message IEA382I appears when coupling link 1 fails:

  `IEA382I THIS SERVER HAS ONLY A SINGLE LINK AVAILABLE FOR TIMING PURPOSES`

- H40 remains as Stratum 2 after the failure.

3.3.1 Problem awareness

In Figure 3-10, if any of the links between the two Stratum 1 servers fail, there will be no indication of an STP problem because these links are not being used for time synchronization.

- Coupling link 1 failure

  Assuming that Stratum 2 H40 is using Stratum 1 SCZP201 as its time source, when coupling link 1 fails, H40 has an alternate time source to SCZP101 available through link 2. There is no impact from an STP timing point of view because H40 is still able to use SCZP101 as a valid time source. However, message IEA382I is displayed because H40 now has only one timing link to SCZP101:

  `IEA382I THIS SERVER HAS ONLY A SINGLE LINK AVAILABLE FOR TIMING PURPOSES`

This message is issued when there is only one remaining physical link (PCHID) available for timing purposes.

- Coupling link 2 failure

  Similarly, if coupling link 2 fails, H40 loses connectivity to its time source. There is no impact from an STP timing point of view because H40 is still able to use SCZP201 as a valid time source; however, message IEA382I is displayed because H40 only has one timing link to SCZP201:

  ```
  IEA382I THIS SERVER HAS ONLY A SINGLE LINK AVAILABLE FOR TIMING PURPOSES
  ```

- Coupling links 1 and 2 failure

  If both coupling links 1 and 2 fail, H40 loses connectivity to all sources of timing. In this case, H40 will become unsynchronized and drop to Stratum 0. z/OS image SC73 on H40 will issue WTOR message IEA394A.

3.3.2 Problem determination

The status of the coupling links should be determined using the DISPLAY ETR command and the STP Status tab.

As shown in Figure 3-11, issuing the DISPLAY ETR command from SC73 shows that H40 is a Stratum 2 server in STP timing mode with only one link available for timing purposes.

```
RO SC73,D ETR
IEA386I 12.33.47 TIMING STATUS 259
SYNCHRONIZATION MODE = STP
THIS SERVER IS A STRATUM 2
CTN ID = ITSOPOK-31
NUMBER OF USABLE TIMING LINKS =  1
THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS
```

Figure 3-11 Coupling link failure: DISPLAY ETR

In Figure 3-12, the System Information section of the STP Status tab for H40 indicates that H40 is synchronized to server SCZP101 through a single coupling link identified by PCHID 0110.

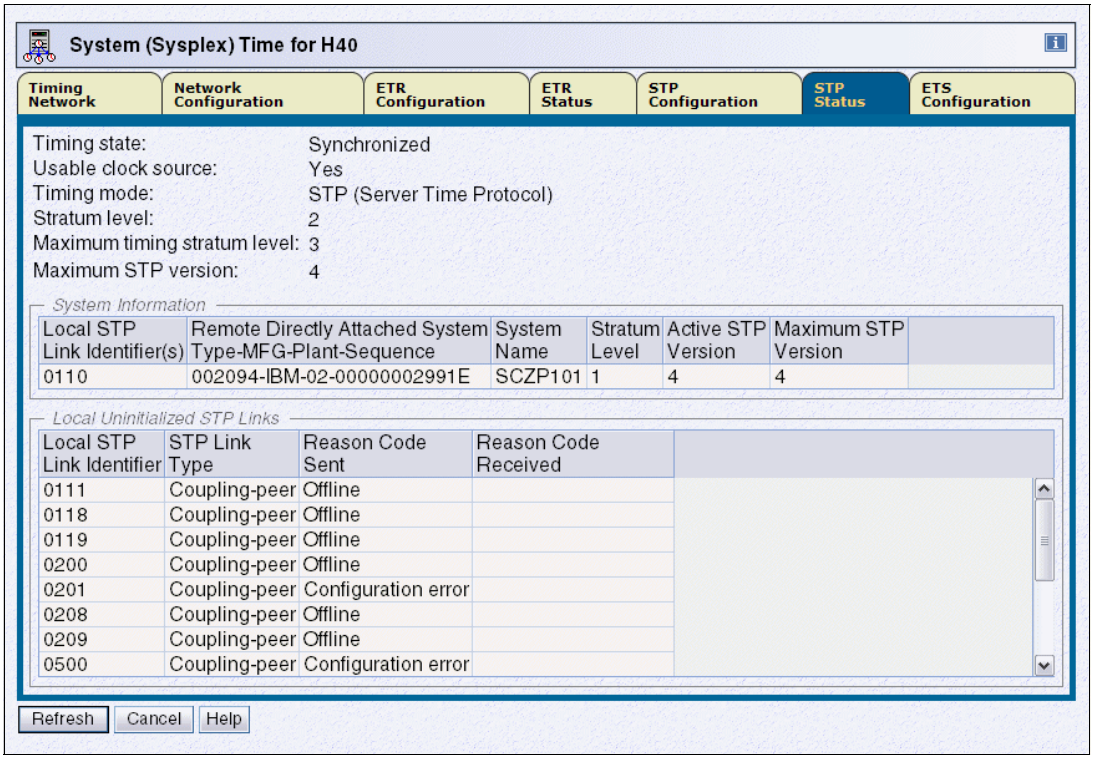

Figure 3-12 Coupling link failure: STP Status tab

3.3.3 User actions

You have to identify the coupling link that has failed and take the appropriate actions to restore link connectivity. When link connectivity has been restored, the number of timing links can be verified with the DISPLAY ETR command and checking the STP Status tab.

Figure 3-13 shows that after recovery H40 has two coupling links to two separate sources of time:

- The `NUMBER OF USABLE TIMING LINKS` is now 2.
- The indication `THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS` is no longer present.

```
IEA386I 12.58.25 TIMING STATUS 150
SYNCHRONIZATION MODE = STP
THIS SERVER IS A STRATUM 2
CTN ID = ITSOPOK-31
NUMBER OF USABLE TIMING LINKS =  2
```

Figure 3-13 Coupling link failure: DISPLAY ETR from H40 after recovery and user action

In Figure 3-14 on page 89, the System Information section of the STP Status tab also indicates that H40 has two coupling links available for timing purposes, one to SCZP101 through PCHID 0110 and the other to SCZP201 through PCHID 0111.

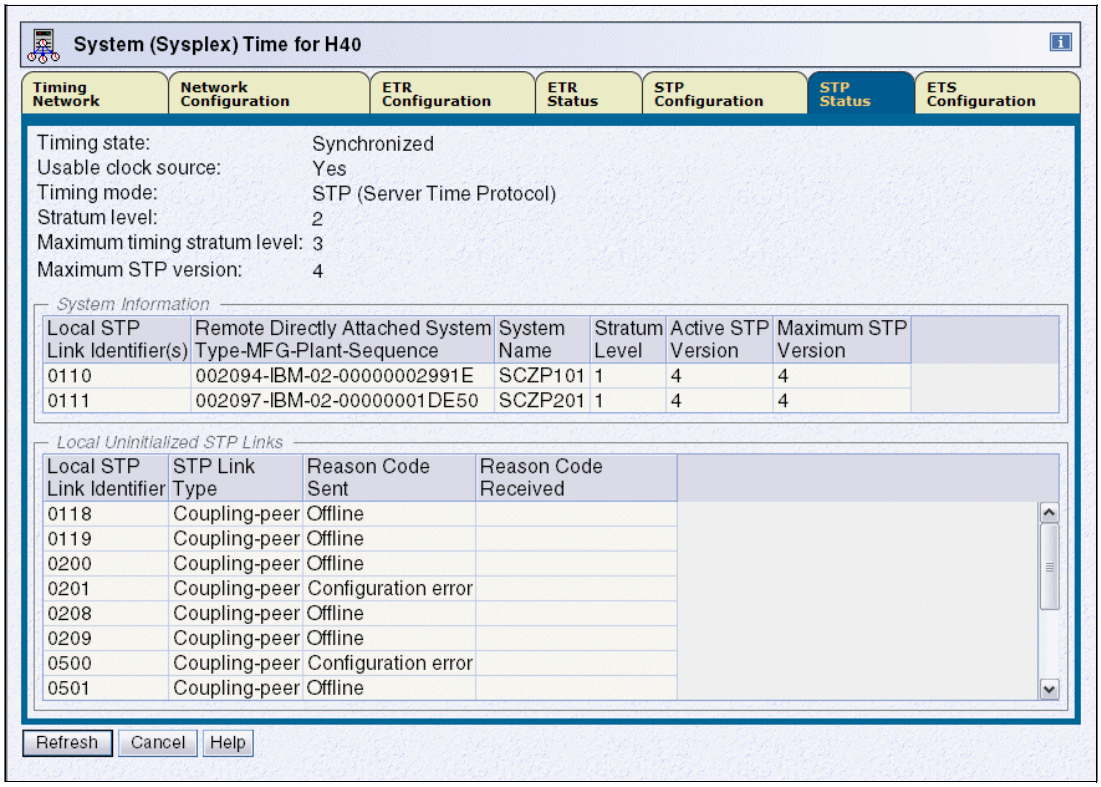

Figure 3-14 Last coupling link failure: STP Status tab after recovery and user action

3.4 STP Stratum 1 failure

In a Mixed CTN, at least two Stratum 1 servers are recommended, if possible, in order to avoid a single point of failure for Stratum 2 servers. Where a Stratum 2 server is connected to multiple Stratum 1 servers, only one is selected as the time source.

In Figure 3-15, H40 is a Stratum 2 receiving STP messages from both SCZP101 and SCZP201 Stratum 1 servers, and has selected SCZP101 as the time source. If SCZP101 fails, H40 Stratum 2 selects SCZP201 Stratum 1 as its time source and continues processing without disruption.

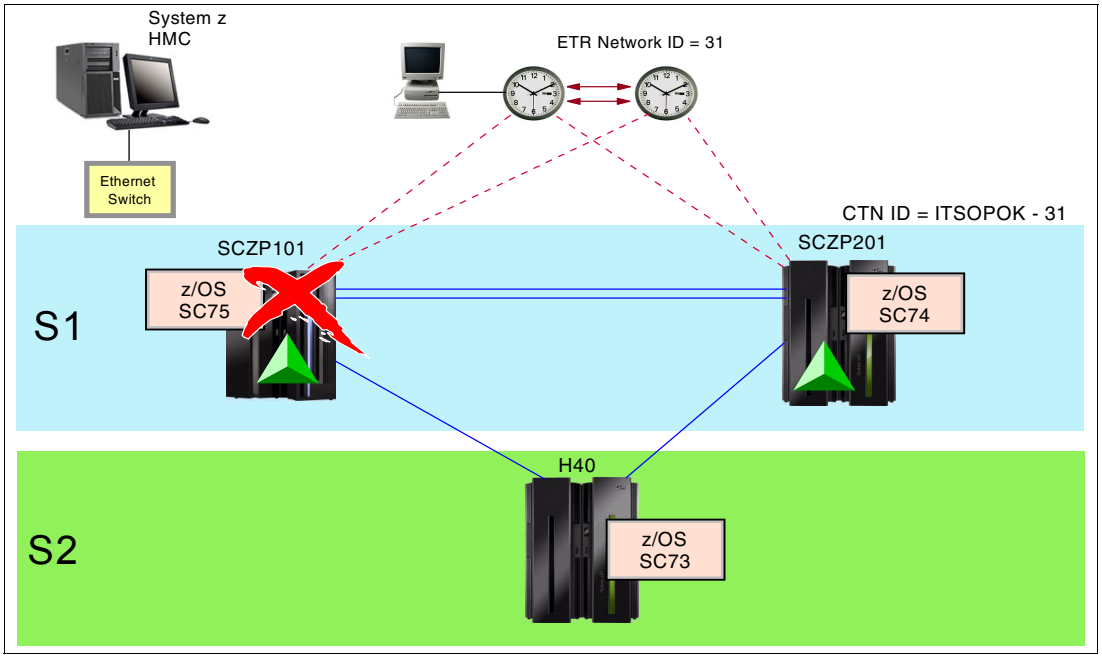

Figure 3-15 Configuration: Stratum 1 failure

However, this would leave H40 totally dependent upon SCZP201 for STP timing messages, and any removal of SCZP201 for planned maintenance or a possible failure would cause H40 to lose its only available timing source.

The key highlights of this scenario are:

- SCZP101 Stratum 1 fails.
- H40 receives timing from SCZP201 and remains at Stratum 2.

3.4.1 Problem awareness

z/OS message IXC101I is issued, indicating that the server in this scenario (SCZP101) has failed:

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC75 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
```

An active SFM system isolation policy automatically starts system partitioning. This is highlighted by an IXC101I message a few seconds after detection of the failure. Partitioning triggers the cleanup of resources for the failing system.

As shown in Figure 3-16, if system isolation fails, then SFM issues WTOR message IXC102 after the XCF CLEANUP time has elapsed.

```
IXC102A XCF IS WAITING FOR SYSTEM SC75 DEACTIVATION. REPLY DOWN
WHEN MVS ON SC75 HAS BEEN SYSTEM RESET
```

Figure 3-16 z/OS message IXC102A: after failure

Note: Before replying DOWN to IXC102A or IXC402D, perform a hardware SYSTEM RESET on the z/OS system being removed. This is necessary to ensure that this z/OS system can no longer perform any I/O operations, and that it releases any outstanding I/O reserve. A SYSTEM RESET ensures that other systems continue to have access to the data sets on the shared devices.

3.4.2 Problem determination

Figure 3-17 shows the STP Status tab of SCZP201 after Stratum 1 SCZP101 has failed. SCZP201 has connectivity through a single link to H40.

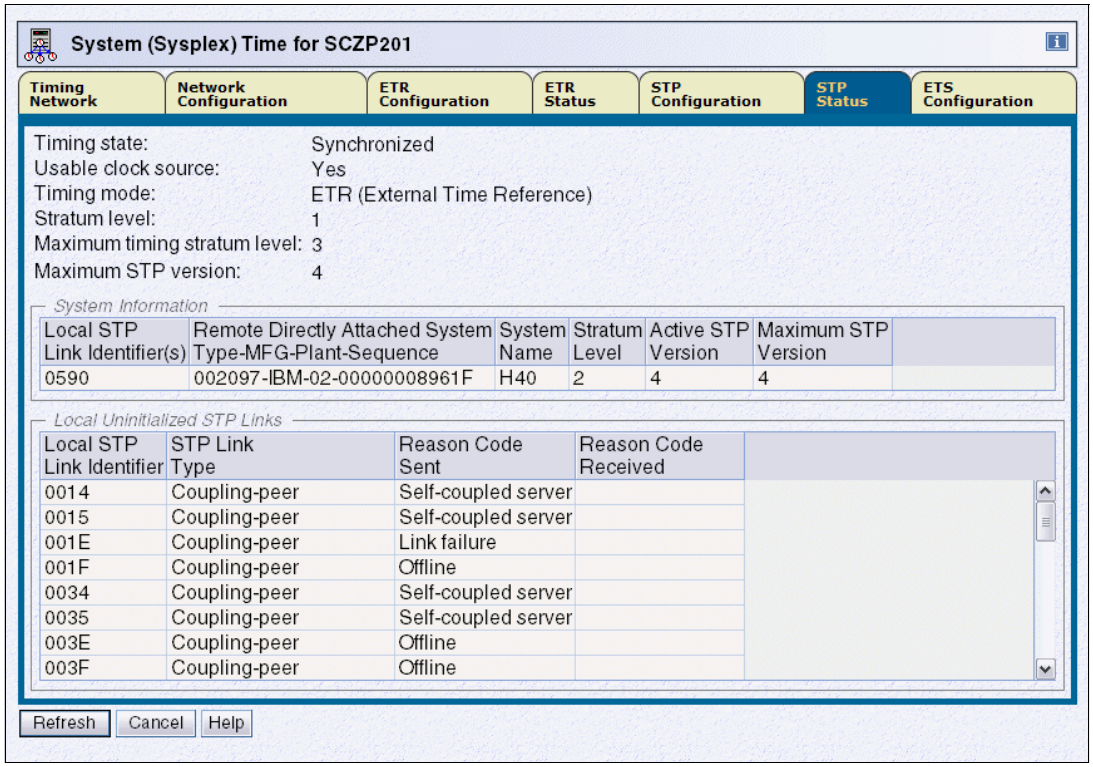

Figure 3-17 STP Status tab: SCZP201 after SCZP101 failure

3.4.3 User actions

Recover and re-IPL SCZP101 using installation recovery procedures. In Figure 3-18, the System Information section of the STP Status tab for SCZP201 shows that SCZP201 is Stratum 1 and has connectivity to both Stratum 1 SCZP101 and Stratum 2 H40 after successful recovery.

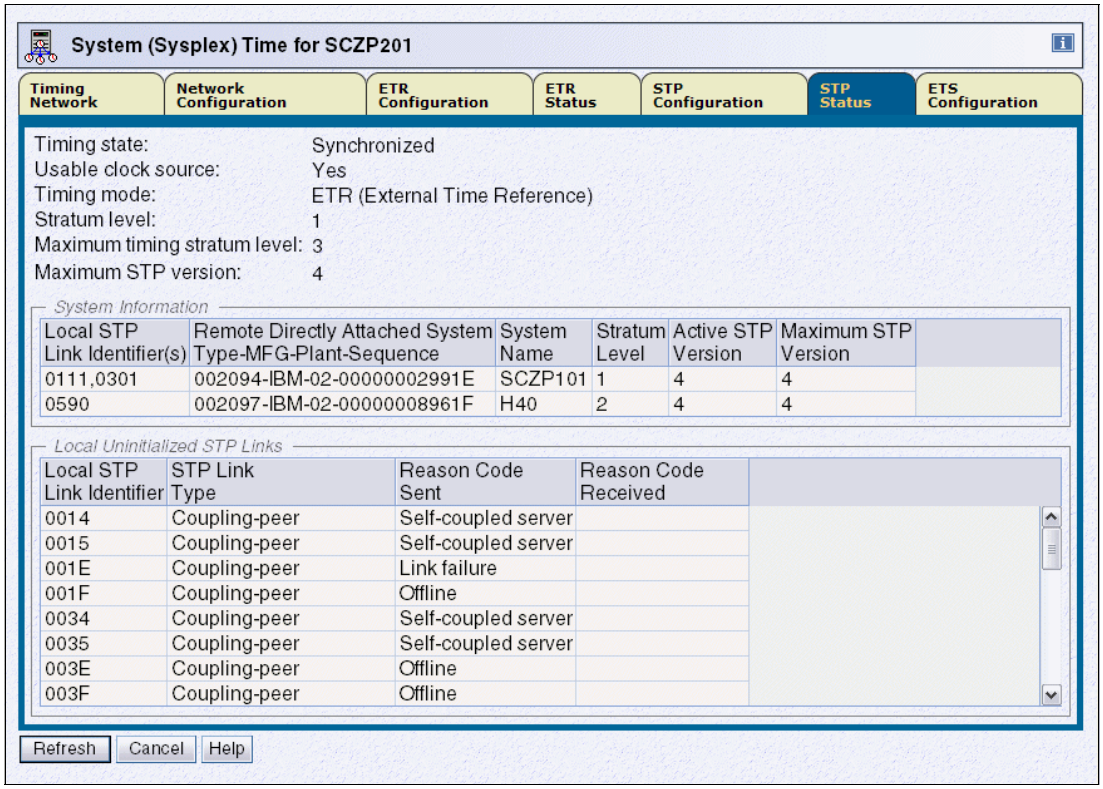

Figure 3-18 STP Status tab: SCZP201 after recovery and user action

3.5 Two-site: Site 1 failure

Complete failure at Site 1 causes the servers at Site 2 to lose the time source and become unsynchronized. Figure 3-19 shows a two-site Mixed CTN configuration. Stratum 1 SCZP101 is at Site 1, while Site 2 consists of STP Stratum 2 H40 and STP Stratum 3 SCZP201. In this scenario, Site 1 completely fails.

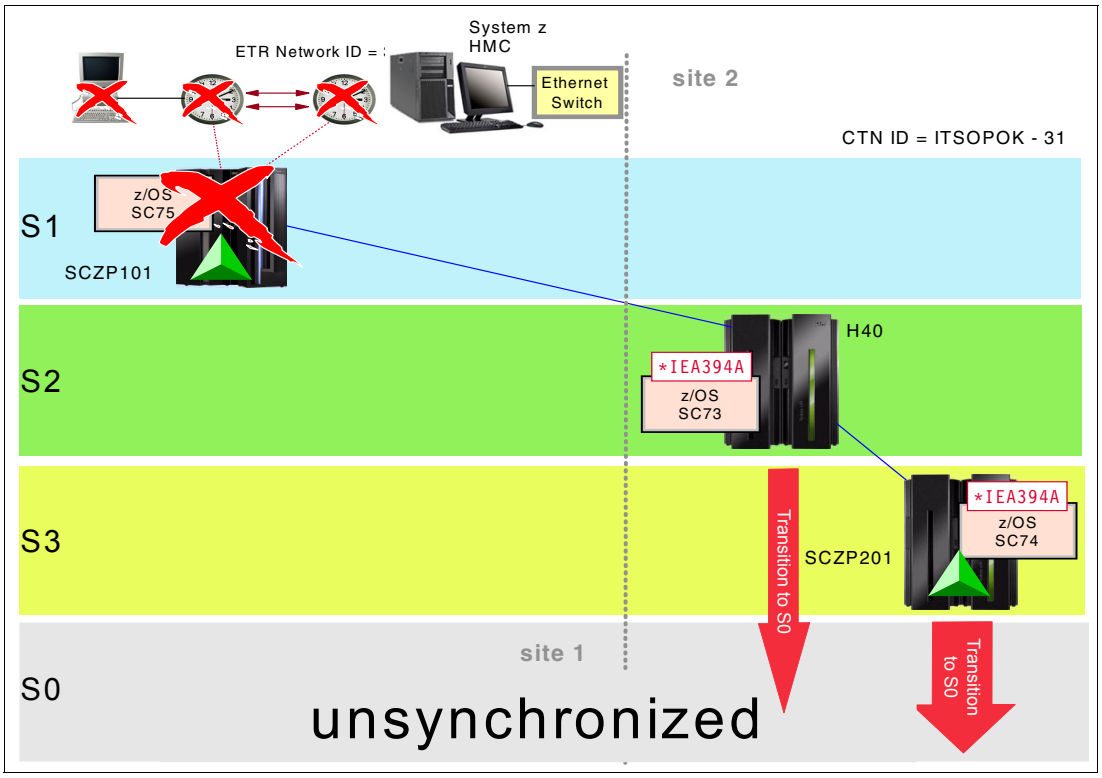

Figure 3-19 Configuration: Site 1 failure

After the time source at Site 1 completely fails, all servers at Site 2 become unsynchronized. z/OS images on SCZP201 and H40 that have ETRMODE YES or STPMODE YES will issue WTOR message IEA394A.

3.5.1 Problem awareness

When Site 1 fails, z/OS system images on servers at Site 2 issue WTOR message IEA394A, as shown in Figure 3-20.

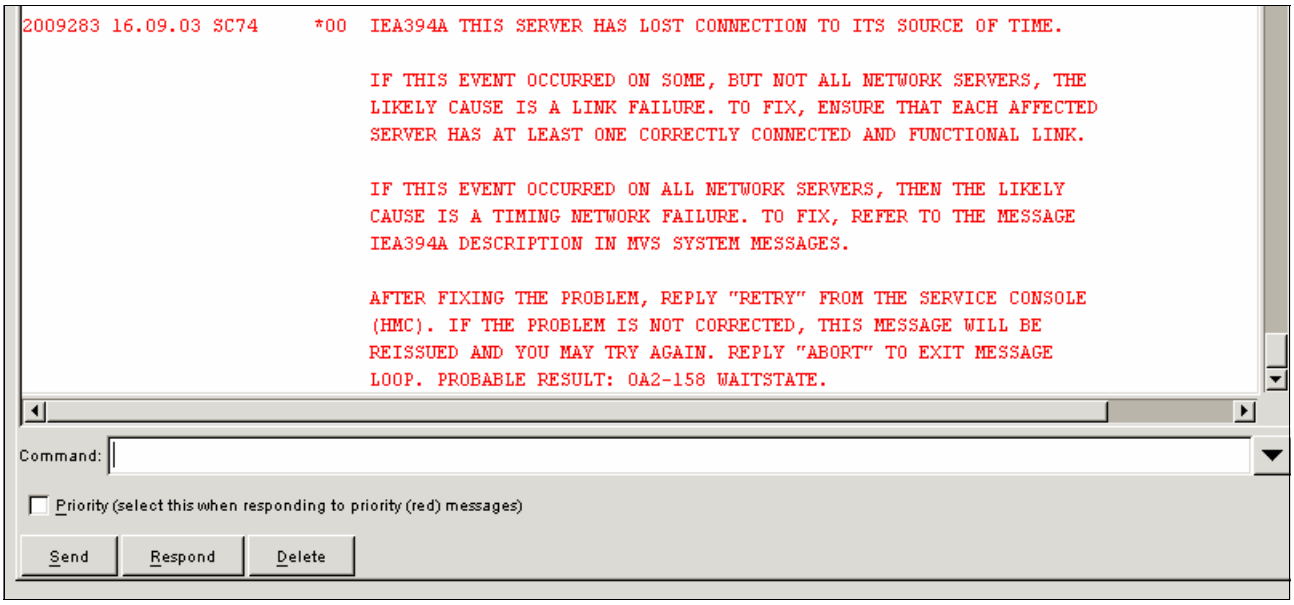

Figure 3-20 WTOR message IEA394A on z/OS image SC74

3.5.2 Problem determination

All servers at Site 2 have lost the time source and become unsynchronized. In Figure 3-21 on page 95, the STP Status tab for both servers at Site 2 will show that:

- Timing State is *unsynchronized*.
- Usable clock source is *no*.
- Stratum level is *0*.
- There is still a local STP link identifier between the two servers at Site 2, from H40 to SCZP201. As expected, the stratum level for SCZP201 is also 0.

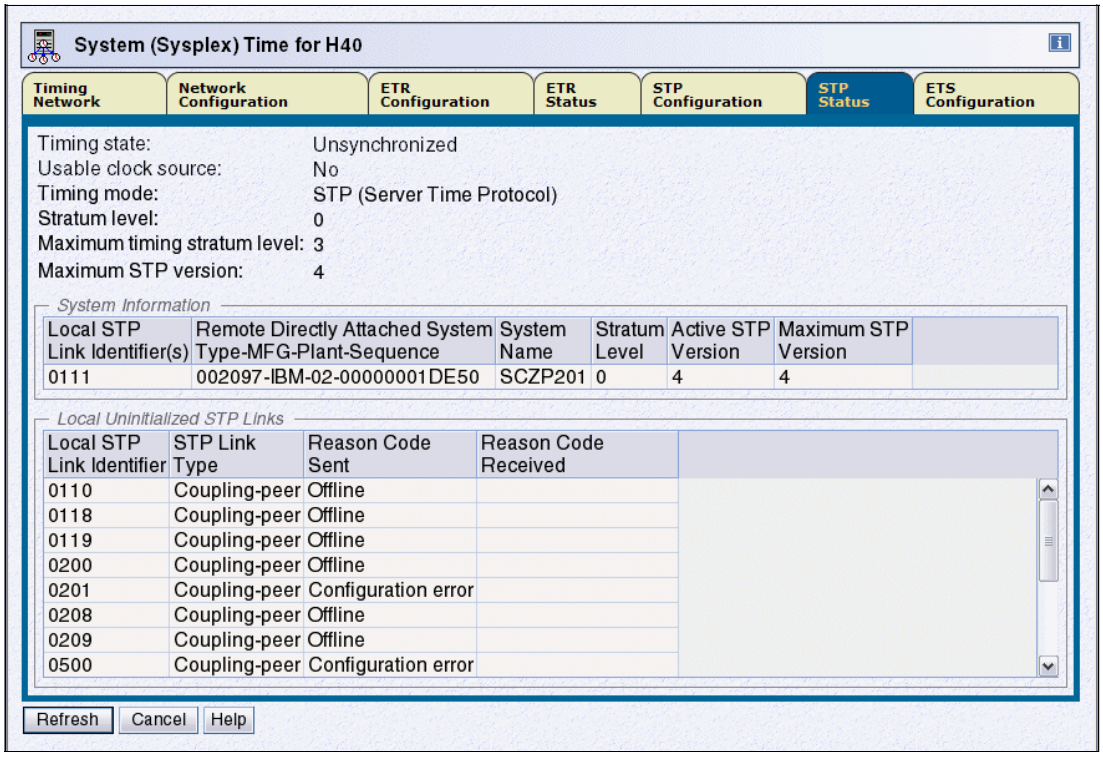

Figure 3-21 STP Status tab: H40 after failure

3.5.3 User actions

In this case, you can either recover Site 1 (including the Sysplex Timer) and reply RETRY to the WTOR messages, or you can decide to reconfigure the servers at Site 2 into an STP-only CTN and reply RETRY to the WTOR messages. In this example it is shown how the CTN is being reconfigured assuming Site 1 remains unavailable.

To reconfigure the CTN and retry the systems at Site 2:

1. Because the CTN will be reconfigured from a Mixed to an STP-only CTN, the ETR network ID must be *manually* removed on servers that will participate in the STP-only CTN (SCZP201 and H40 in this example).

   Figure 3-22 on page 96 shows the ETR Configuration tab for SCZP201. Clear the ETR network ID and click **Apply**. The CTN configuration message ACT37363 needs to be confirmed as shown in Figure 3-23 on page 96. For server H40 the ETR ID does not need to be manually removed because this will be automatically done when reconfiguring into an STP-only CTN using this server as Preferred and Current Time Server.

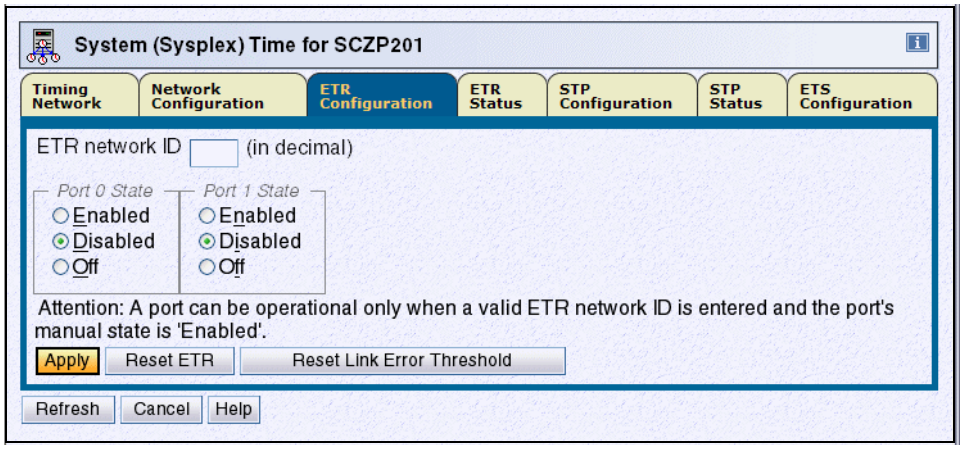

Figure 3-22 ETR Configuration tab: SCZP201

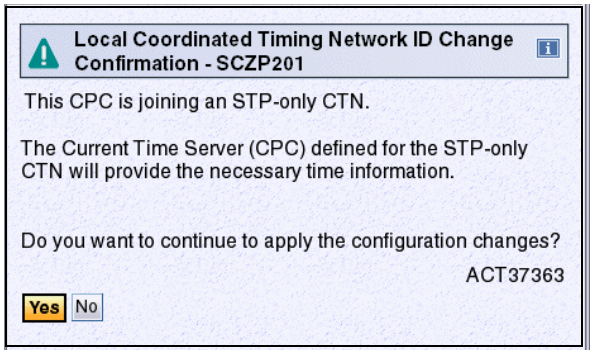

Figure 3-23 SCZP201 reconfiguration: Message ACT37363

2. Reconfigure the servers at Site 2 in an STP-only CTN. Figure 3-24 on page 97 shows the setup of the Network Configuration tab for H40. The CTN reconfiguration needs to be performed in two phases because at this point in time the CTN ID between H40 and SCZP201 does not match. Once H40 has been assigned as CTS for the STP-only CTN, its ETR network ID will automatically disappear. Furthermore, the CTN ID will match the one from SCZP201.
   - Since H40 is being chosen to become the CTS, the reconfiguration must be done from its Network Configuration tab. Initially, all fields in the Current Network Configuration section display Not Configured.
   - The Initialize Time button is not enabled because the time configuration for H40 has not changed, and thus the time does not need to be initialized.
   - In the first phase H40 is configured as the Preferred Time Server, as shown in Figure 3-24 on page 97. SCZP201 is visible in the server menu but does not have an STP ID displayed because it does not match the CTN ID of H40.
   - Click the **Force configuration** check box. This is mandatory because there is no CTS configured yet.
   - Click **Apply** to reconfigure the CTN to STP-only. The regular Network Configuration Change Confirmation panel ACT37348 needs to be answered because the Force Configuration option was used.
   - In the second phase SCZP201 is configured as the Backup Time Server using the Network Configuration tab of H40, as shown in Figure 3-24 on page 97. Since H40 has been reconfigured to a STP-only CTN, its CTN ID now matches the CTN ID of

SCZP201. As a consequence it is shown in the server menu including the STP ID, ITSOPOK in this case. This configuration change can be done without the force configuration option because there is already a CTS assigned in this Timing Network.

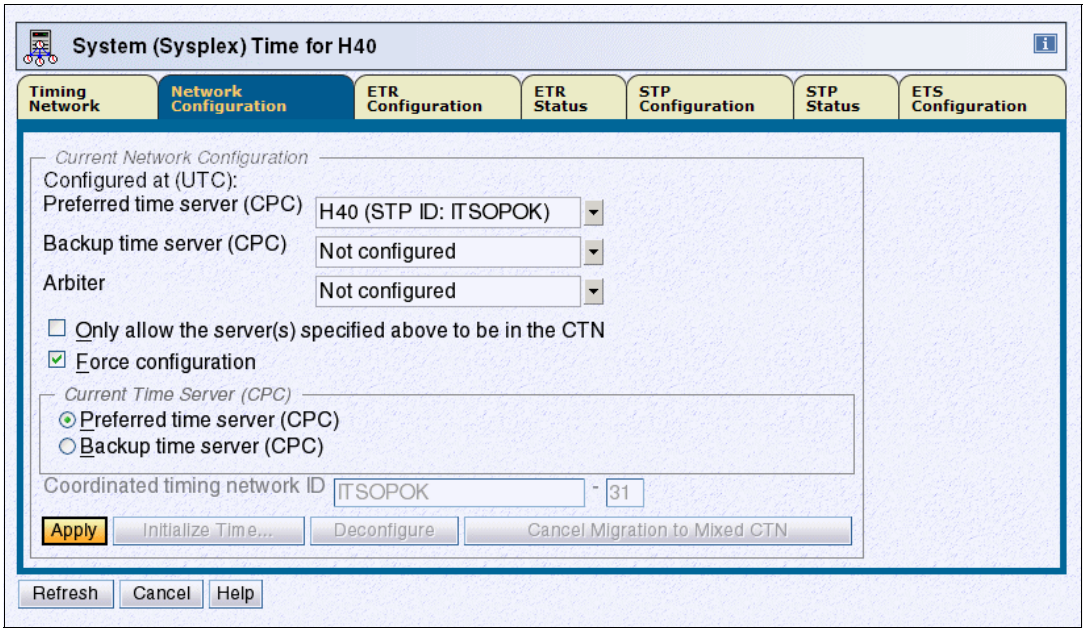

Figure 3-24 Network Configuration tab: H40 - assign PTS

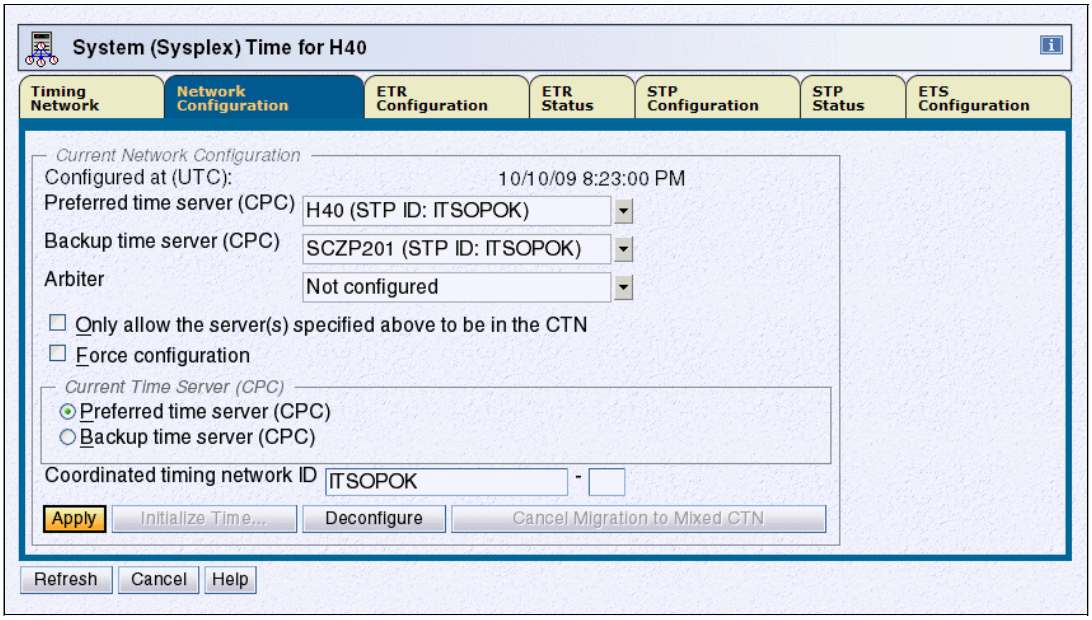

Figure 3-25 Network Configuration tab: H40 - assign BTS

Figure 3-26 on page 98 shows the STP status tab for H40. The CTN recovers with H40 as Stratum 1 and SCZP201 as Stratum 2.

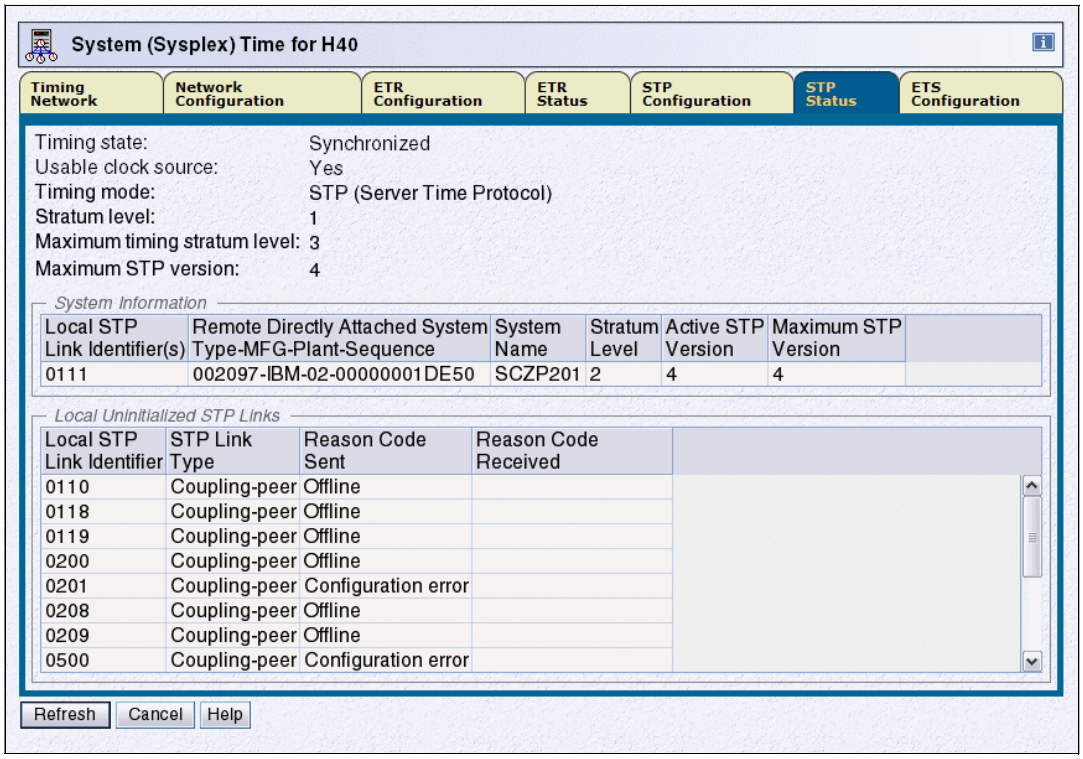

Figure 3-26 STP status tab: H40 after recovery and user action

3. After the CTN reconfiguration is successful, reply RETRY to each outstanding WTOR message IEA394A for z/OS images on H40 and SCZP201. If the reply is entered from the HMC, the priority message check box must be set.

4. When the failure at Site 1 has been recovered, and connectivity between SCZP101 and the other servers has been re-established, SCZP101 can be reconfigured to join the STP-only CTN. This is done from the ETR Configuration tab on SCZP101 by:

   – Disabling the ETR ports on the ETR configuration tab
   – *Manually* removing (nullifying) the ETR network ID

   When this is done, SCZP101 will automatically join the STP-only CTN and synchronize to the Coordinated Server Time provided by H40, as shown on the STP Status tab of H40 in Figure 3-27. z/OS and Coupling Facility images on SCZP101 can then be activated and IPLed to rejoin the sysplex.

   Furthermore, if the decision has been taken to keep this CTN in an STP-only CTN, the final steps to complete the STP-only CTN migration should be taken:

   – Time Zone offset adjustment.
   – Optionally add an External Time Source to the CTN.

   For details, refer to *Server Time Protocol Implementation Guide*, SG24-7281.

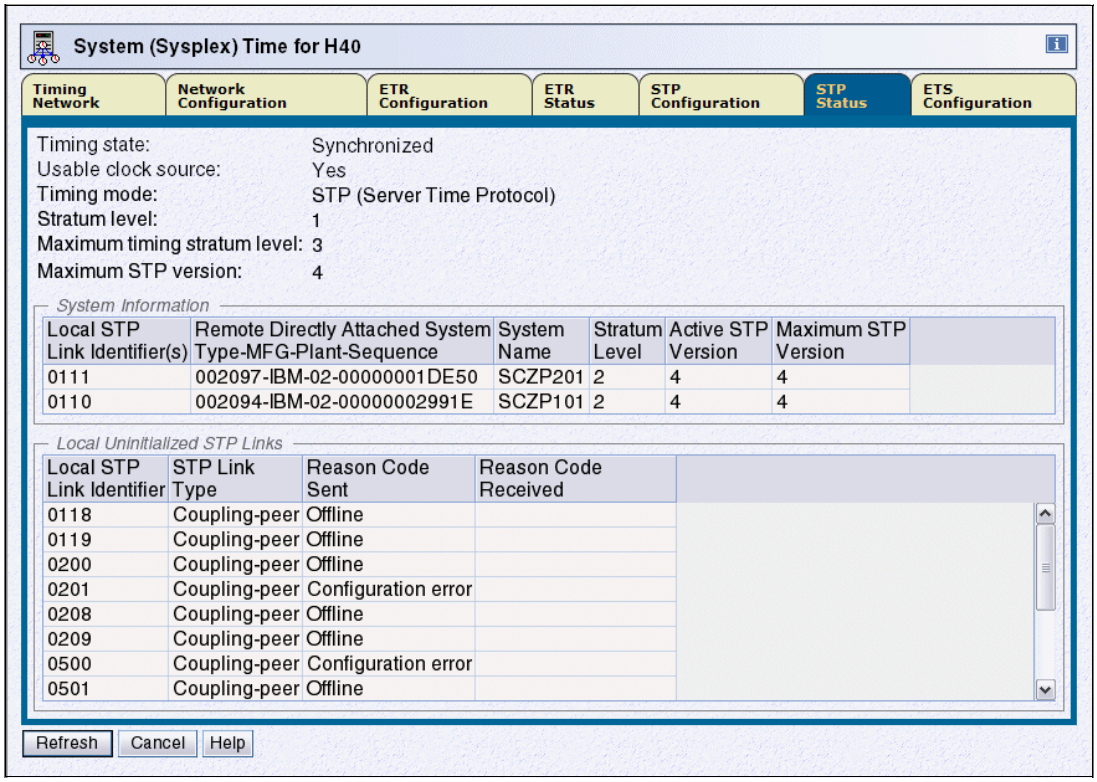

Figure 3-27 STP Status tab: H40 after recovery and user action

Optionally, after recovery of Sysplex Timers at Site 1 is complete, the CTN can be returned to its original state, as shown in Figure 3-19 on page 93, by:

- Reconfiguring SCZP101 as the PTS and CTS.
- Initiating a migration from STP-only to Mixed CTN from SCZP101. To perform the migration, the CTN ID should be restored to its initial value by specifying the ETR network ID [31] on the Network Configuration tab.

  The migration will enable the ETR ports on SCZP101 and the server will synchronize to the Sysplex Timer at the end of the migration.

3.6 Two-site: Sysplex Timers failures

Failure of both Sysplex Timers at Site 1 causes the servers at both sites to lose their time source and become unsynchronized. On SCZP101, z/OS images that have ETRMODE YES issue WTOR message IEA015A. On SCZP201 all z/OS images that have STPMODE YES and operate in STP Timing Mode issue WTOR message IEA394A. This configuration is similar to the one described in 3.5, “Two-site: Site 1 failure” on page 93. However, in this scenario, only connectivity to the Sysplex Timers located at Site 1 is lost.

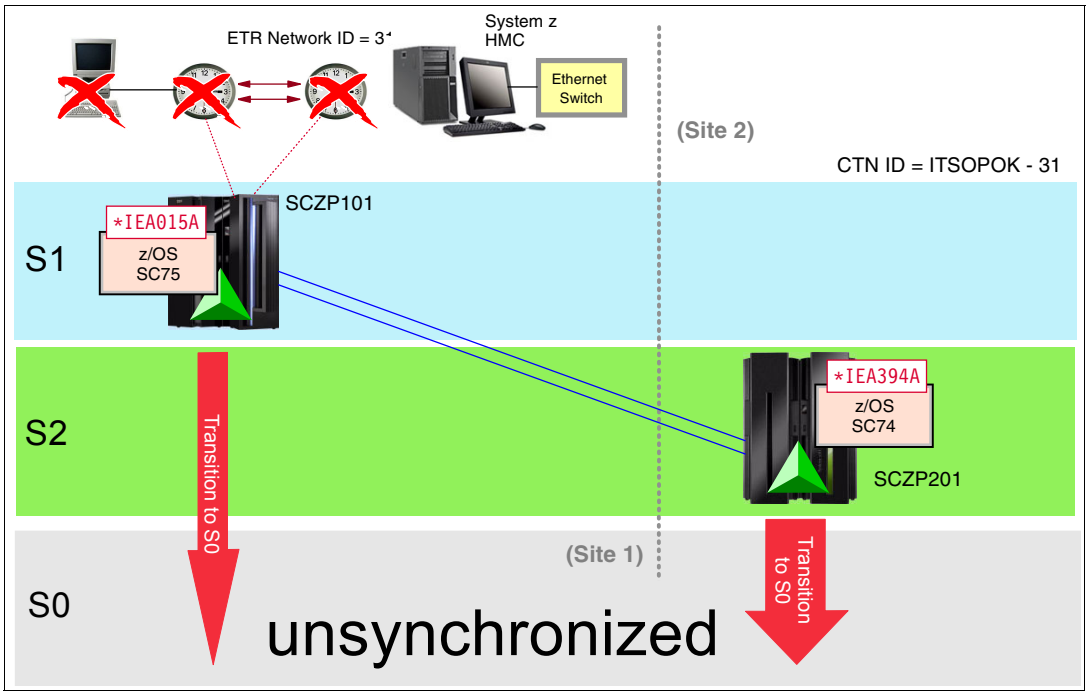

Figure 3-28 Sysplex Timers failure (single site or two sites)

The key highlights of this scenario are:

- Sysplex Timers located at Site 1 fail.
- All servers at Site 1 and Site 2 lose time synchronization. All z/OS images that have ETRMODE YES and operate in ETR Timing Mode issue WTOR message IEA015A. All z/OS images that have STPMODE YES and operate in STP Timing Mode issue WTOR message IEA394A.

The scenario presented here does not follow the recommendation to have two Stratum 1 servers in a Mixed CTN configuration. However, we recognize that there are cases when the availability rules must be temporarily ignored.

- In a two-site scenario, because the site hosting the second server (SCZP201 in this example) is beyond the 40 km distance supported between two Sysplex Timer units in an Expanded Availability configuration.
- In a single site, installations with only two servers installed may need to test STP timing mode on one server before migrating to an STP-only CTN.

The recovery described for this scenario applies to a single- or dual-site situation.

3.6.1 Problem awareness

When ETR connectivity fails, all z/OS images on servers at Site 1 issue WTOR message IEA015A, and at Site 2 issue WTOR message IEA394A. Figure 3-29 on page 101 shows z/OS system image SC75 on SCZP101.

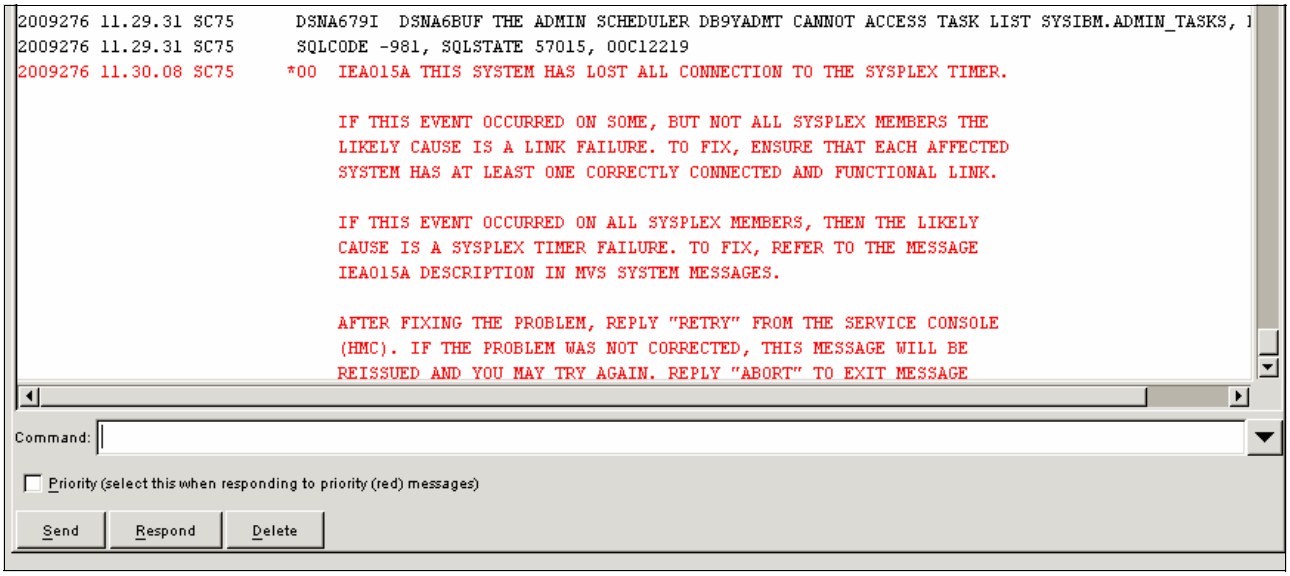

Figure 3-29 z/OS image SC75: WTOR message IEA015A

3.6.2 Problem determination

All servers at Site 1 and Site 2 have lost the time source and are unsynchronized. In Figure 3-30, the STP Status tab for both servers at Site 1 and Site 2 shows that:

- Timing State is unsynchronized.
- Usable clock source is No.
- Stratum level is 0.

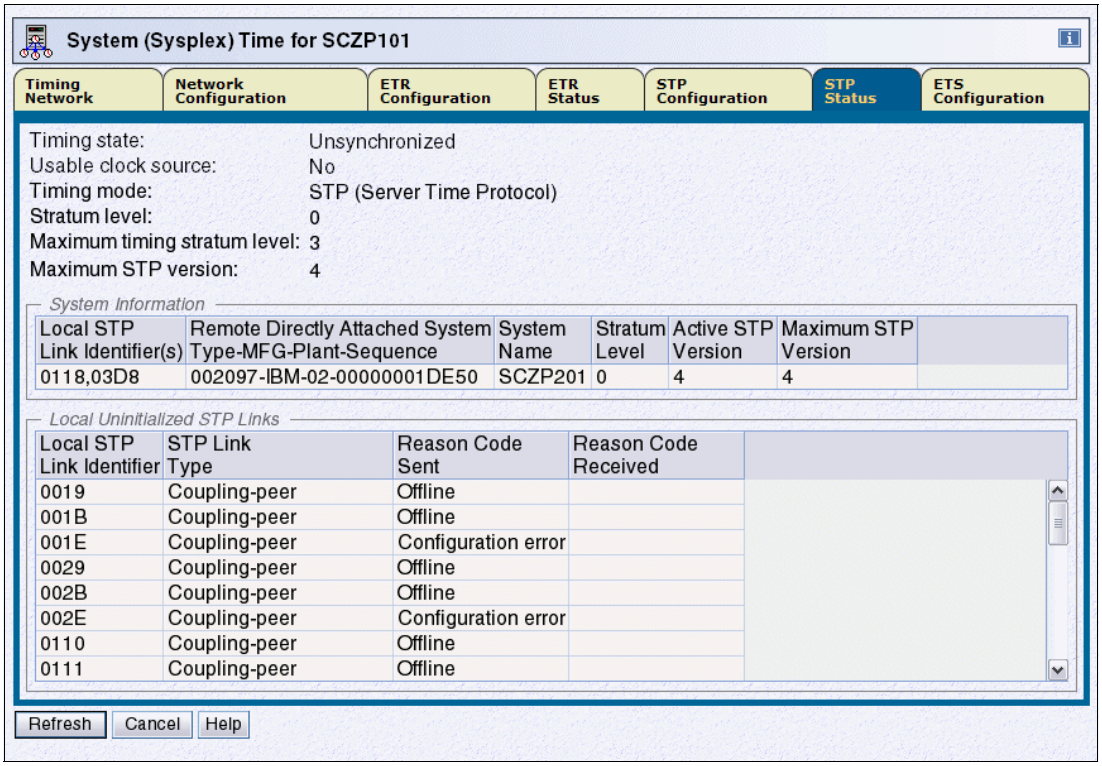

Figure 3-30 STP Status tab: SCZP101 after failure

3.6.3 User actions

In this case, you can either correct the Sysplex Timer problem and reply RETRY to the WTOR messages, or you can decide to reconfigure the servers at both sites into an STP-only CTN and reply RETRY to the WTOR messages. In this example we show how the CTN is being reconfigured assuming the Sysplex Timer remains unavailable.

To reconfigure the CTN and retry the z/OS systems:

1. Because the CTN will be reconfigured from a Mixed to an STP-only CTN, the ETR network ID must be manually removed on servers that will participate in the STP-only CTN (SCZP101 and SCZP201 in this example).

   As the target configuration is an STP-only CTN, z/OS images that are in WTOR recovery coming from ETR timing mode (WTOR IEA015A) need to be re-IPLed because the required ETR timing mode cannot be restored in this case.

   Figure 3-31 shows the ETR Configuration tab for SCZP201. Clear the ETR network ID and click **Apply**. The CTN configuration message ACT37363 needs to be confirmed, as shown in Figure 3-31. This step needs to be done for every server except the one that was previously attached to the Sysplex Timer. In this case we do not clear the ETR network ID for server SCZP101. This server will be used to reconfigure the CTN into an STP-only CTN, shown in the next steps.

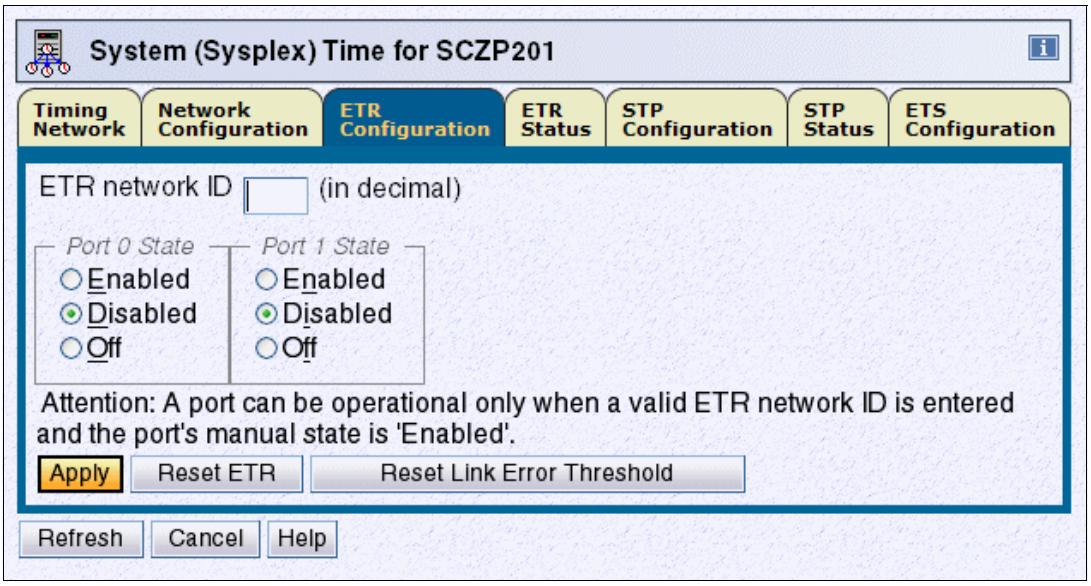

Figure 3-31 ETR Configuration tab: SCZP201

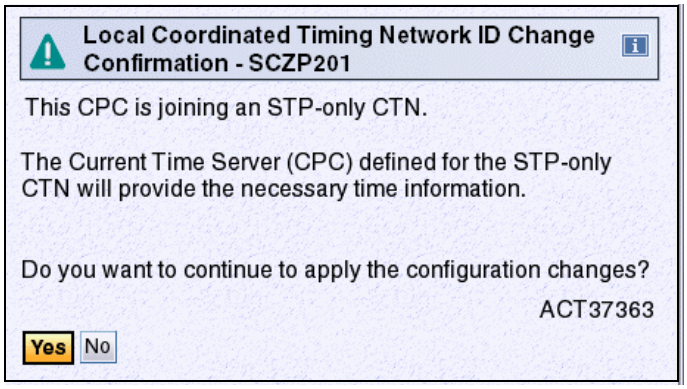

Figure 3-32 SCZP201 reconfiguration: Message ACT37393

2. Reconfigure the two servers in an STP-only CTN. Figure 3-33 on page 104 shows the setup of the Network Configuration tab for SCZP101. The CTN reconfiguration needs to be performed in two phases because at this point in time the CTN ID between SCZP101 and SCZP201 does not match. Once SCZP101 has been assigned as CTS for the STP-only CTN, its ETR network ID automatically disappears. Furthermore, the CTN ID will match the one from SCZP201, which was previously manually modified.
   - Because SCZP101 is being chosen to become the CTS, the reconfiguration must be done from its Network Configuration tab. Initially, all fields in the Current Network Configuration section display Not Configured.
   - The only button that is enabled is the Apply button. Because SCZP101 is already synchronized, the time need not be initialized.
   - In the first phase SCZP101 is configured as the Preferred Time Server, as shown in Figure 3-33 on page 104. SCZP201 is visible in the server menu but does not have an STP ID displayed because it does not match the CTN ID of SCZP101.
   - Click the **Force configuration** check box. This is mandatory because there is no CTS configured yet.
   - Click **Apply** to reconfigure the CTN to STP-only. The regular Network Configuration Change Confirmation panel ACT37348 needs to be answered because the Force Configuration option was used.
   - In the second phase, SCZP201 is configured as the Backup Time Server using the Network Configuration tab of SCZP101, as shown in Figure 3-34. Because SCZP101 has been reconfigured to a STP-only CTN, its CTN ID now matches the CTN ID of SCZP201. As a consequence it is shown in the server menu including the STP ID, ITSOPOK in this case. This configuration change can be done without the force configuration option because there is already a CTS assigned in this Timing Network.

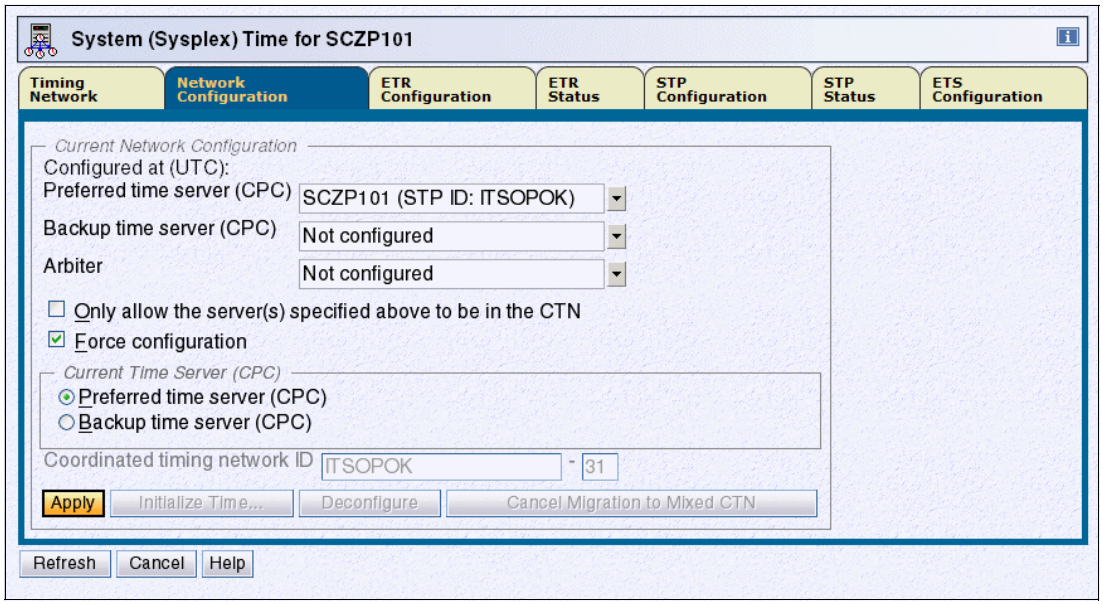

Figure 3-33 Network Configuration tab: SCZP101 - assign PTS

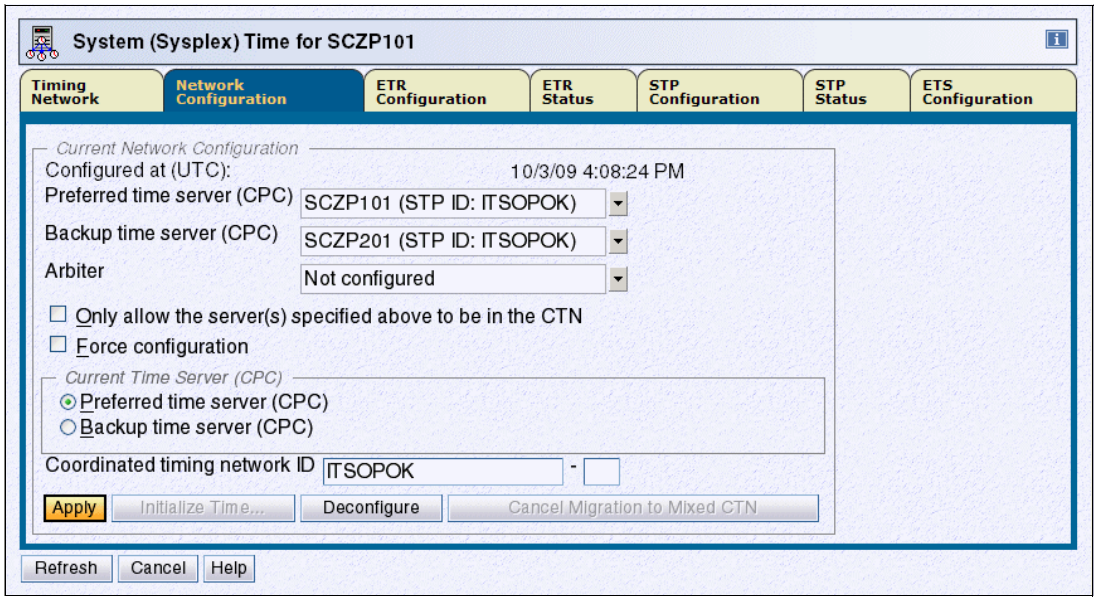

Figure 3-34 Network Configuration tab: SCZP101 - assign BTS

Figure 3-35 on page 105 shows the STP status tab for SCZP101. The CTN recovers with SCZP101 as Stratum 1 and SCZP201 as Stratum 2.

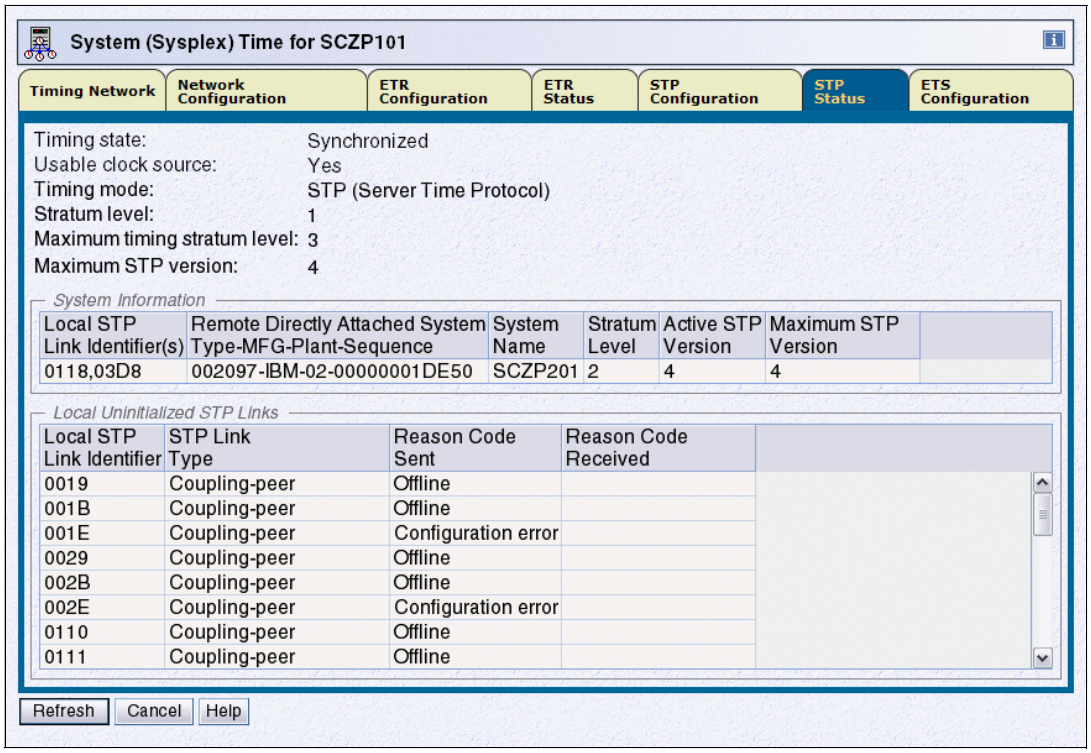

Figure 3-35 STP Status tab: SCZP101 after recovery and user action

3. When the CTN reconfiguration has been successful, reply RETRY to the outstanding WTOR message IEA394A for z/OS image SC74. If the reply is entered from the HMC, the priority message check box must be set. Because SC75 posted IEA015I, it cannot be recovered using the RETRY option because its Timing Mode changed while it was in the WTOR recovery. Answer the outstanding WTOR message IEA015A for image SC75 with ABORT, perform a hardware SYSTEM RESET, and follow the normal recovery procedure to re-IPL image SC75.

   Furthermore, if the decision has been taken to keep this CTN in an STP-only CTN, the final steps to complete the STP-only CTN migration should be taken:

   – Time Zone offset adjustment.
   – Optionally add an External Time Source to the CTN.

   For details, refer to the *Server Time Protocol Implementation Guide*, SG24-7281.

4. Optionally, after recovery of Sysplex Timers at Site 1 is complete, the CTN can be returned to its original state (shown in Figure 3-28 on page 100) by initiating a migration from STP-only to Mixed CTN from SCZP101. To perform the migration, the CTN ID should be restored to its initial value by specifying the ETR network ID [31] on the Network Configuration tab.

   The migration will enable the ETR ports on SCZP101, and the server will synchronize to the Sysplex Timer at the end of the migration.

3.7 Two-site: Site 2 failure

In Figure 3-36, SCZP101 is located at Site 1, and servers H40 and SCZP201 are located at Site 2. In this scenario, Site 2 fails. The key highlights of this scenario are:

- Failure of Site 2 has no impact on Site 1 servers.
- Both servers at Site 2 fail.

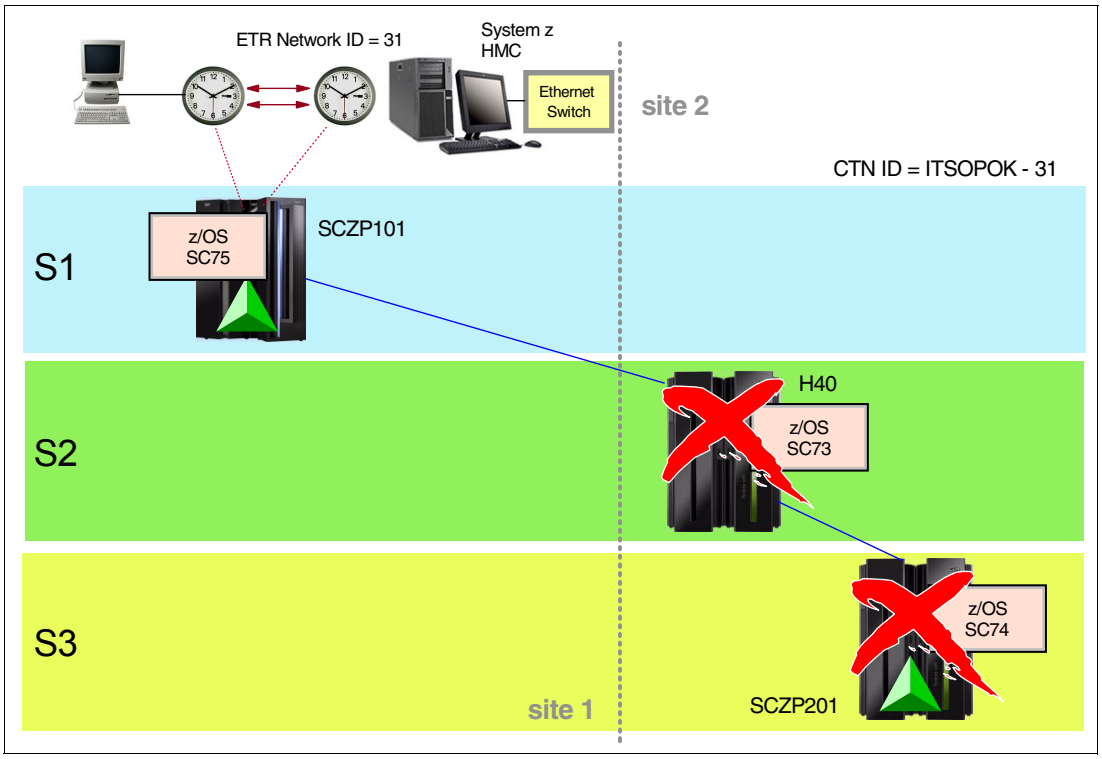

Figure 3-36 Configuration: Site 2 failure

3.7.1 Problem awareness

As shown in Figure 3-37, when an SFM policy is *not* active and the Parallel Sysplex becomes aware of the system failure, message IXC402D is issued, indicating that system images SC73 and SC74 (which reside on failing servers) are not operative.

```
*017 IXC402D SC73 LAST OPERATIVE AT 09:11:35. REPLY DOWN AFTER SYSTEM
                  RESET, OR INTERVAL=SSSSS TO SET A REPROMPT TIME.
*018 IXC402D SC74 LAST OPERATIVE AT 09:11:35. REPLY DOWN AFTER SYSTEM
                  RESET, OR INTERVAL=SSSSS TO SET A REPROMPT TIME.
```

Figure 3-37 z/OS message IXC402D

With the WTOR message IXC402D, you are requested to reply DOWN when the system is confirmed as shut down with a SYSTEM RESET. The INTERVAL option allows you to specify a period of time for system operation recovery. If the system has not recovered within this period, message IXC402D is issued again. The INTERVAL reply can be in the range 0 to 86400 seconds (24 hours). Once the SYSTEM RESET has been performed, reply DOWN to IXC402D.

3.7.2 Problem determination

Use the DISPLAY XCF command to determine the sysplex status. In Figure 3-38, the DISPLAY XCF command is issued from Site 1 LPAR SC75. With Site 2 completely down, only SC75 is active in the sysplex.

```
D XCF,SYSPLEX,ALL
IXC335I  09.26.21  DISPLAY XCF 765
SYSTEM   TYPE SERIAL LPAR STATUS TIME         SYSTEM STATUS
SC75     2094 991E   1C   10/10/2009 09:26:21 ACTIVE         TM=ETR
```

Figure 3-38 DISPLAY XCF: SYSPLEX, ALL

Examine the STP Status tab to determine the status of the CTN. As shown in Figure 3-39, after Site 2 fails, server SCZP101 has no STP connectivity to any other server.

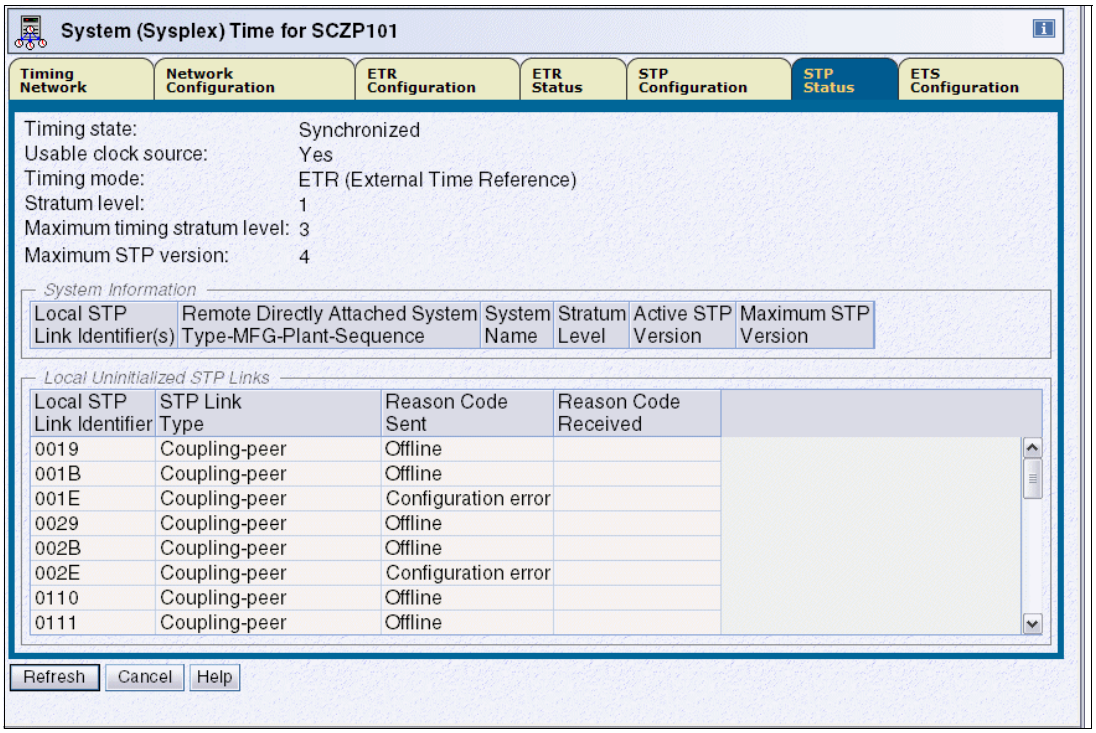

Figure 3-39 Site 2 failure: STP Status tab: Site 1 survives

3.7.3 User actions

A Site 2 failure has no impact on servers located at Site 1. Re-IPL servers H40 and SCZP201 at Site 2 using installation recovery procedures. When both H40 and SCZP201 are operational, they will rejoin the CTN. Examine the STP Status tab to verify that the CTN has successfully returned to its original state.

Figure 3-40 shows the STP Status tab of Site 1 Stratum 1 SCZP101. It has connectivity to Site 2 Stratum 2 H40.

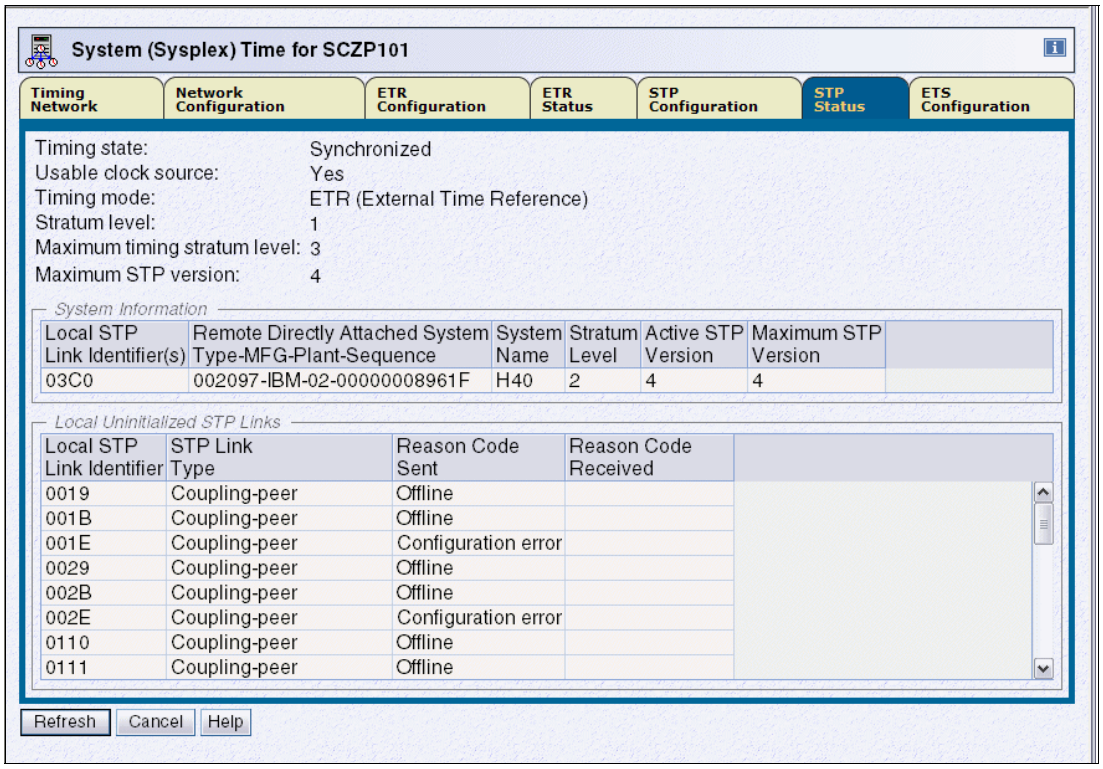

Figure 3-40 SCZP101 STP Status tab: Site 2 restored

Figure 3-41 shows the STP Status tab of Site 2 Stratum 2 H40. It has connectivity to both Site 1 Stratum 1 SCZP101 and Site 2 Stratum 3 SCZP201.

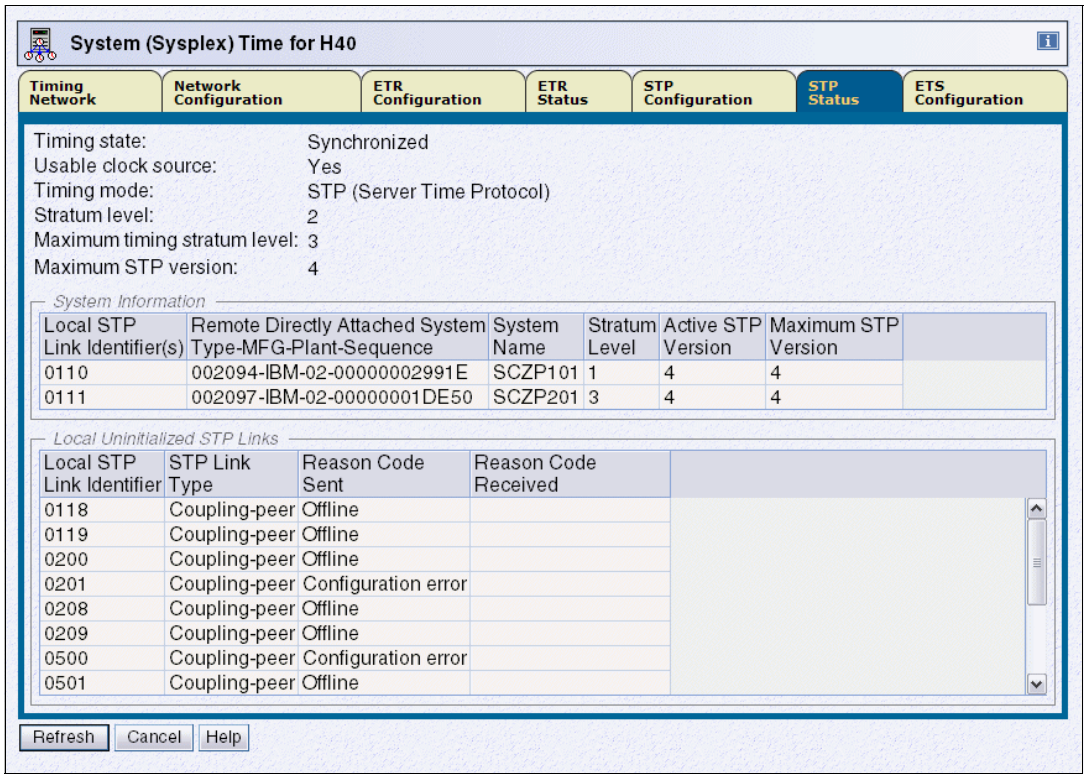

Figure 3-41 H40 STP Status tab: Site 2 restored

The DISPLAY XCF command issued from Stratum 1 SCZP101 shows that the CTN has returned to its original state, as shown in Figure 3-42.

```
D XCF,SYSPLEX,ALL
IXC335I  10.22.41 DISPLAY XCF 616
SYSTEM   TYPE SERIAL LPAR STATUS TIME         SYSTEM STATUS
SC75     2094 991E   1C   10/10/2009 10:22:41 ACTIVE         TM=ETR
SC74     2097 DE50   2C   10/10/2009 10:22:41 ACTIVE         TM=STP
SC73     2097 961F   03   10/10/2009 10:22:41 ACTIVE         TM=STP
```

Figure 3-42 DISPLAY XCF: after recovery and user action

3.8 Two-site: loss of communication

In Figure 3-43, server SCZP101 is located at Site 1, and servers SCZP201 and H40 are located at Site 2. In this scenario, there is a communication failure between Site 1 Stratum 1 SCZP101 and Site 2 Stratum 2 H40. The communication failure does not impact the Site 1 servers. However, for the servers at Site 2, a complete loss of communication between sites appears like a Site 1 failure, and the implications are the same as described in 3.5, “Two-site: Site 1 failure” on page 93. The key highlights of this scenario are:

- The communication failure does not impact the servers at Site 1.
- All servers at Site 2 become unsynchronized, drop to Stratum 0, and z/OS images that have ETRMODE YES or STPMODE YES issue WTOR message IEA394A.

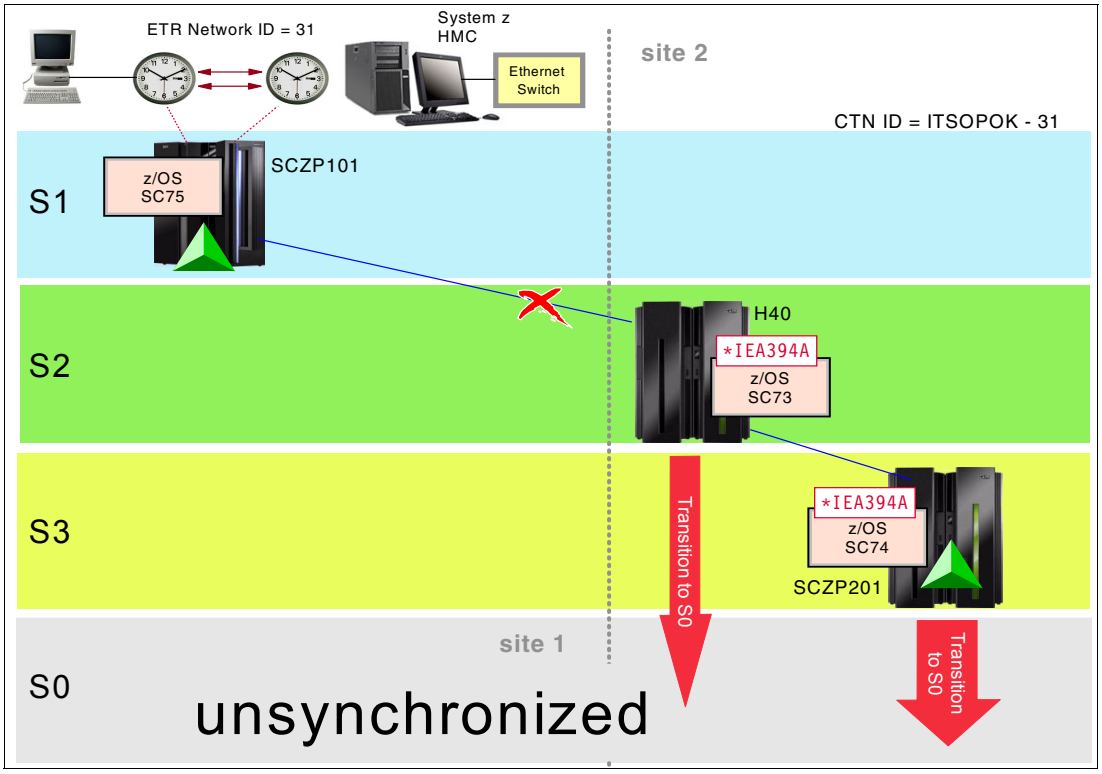

Figure 3-43 Mixed CTN: two-site loss of communication

3.8.1 Problem awareness

When the link between Site 1 and Site 2 fails, z/OS images on servers at Site 2 issue WTOR message IEA394A, as shown in Figure 3-44 on page 111.

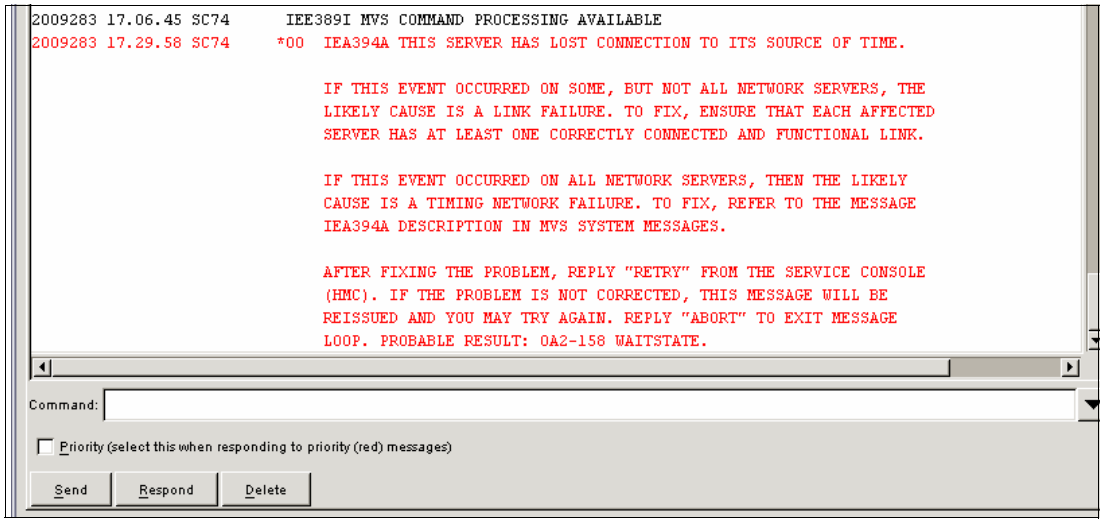

Figure 3-44 WTOR message IEA394A on image SC74

This is one scenario where you need to be aware of what has been specified in the SFM policy, if one is active. The z/OS system images that are still up (SC75 in this example) will see the sysplex images with the IEA394A message as not responding and will take actions as specified in the SFM policy after the INTERVAL time has expired.

3.8.2 Problem determination

In Figure 3-45, the STP Status of Site 1 Stratum 1 SCZP101 shows that SCZP101 has no STP connectivity to any other server.

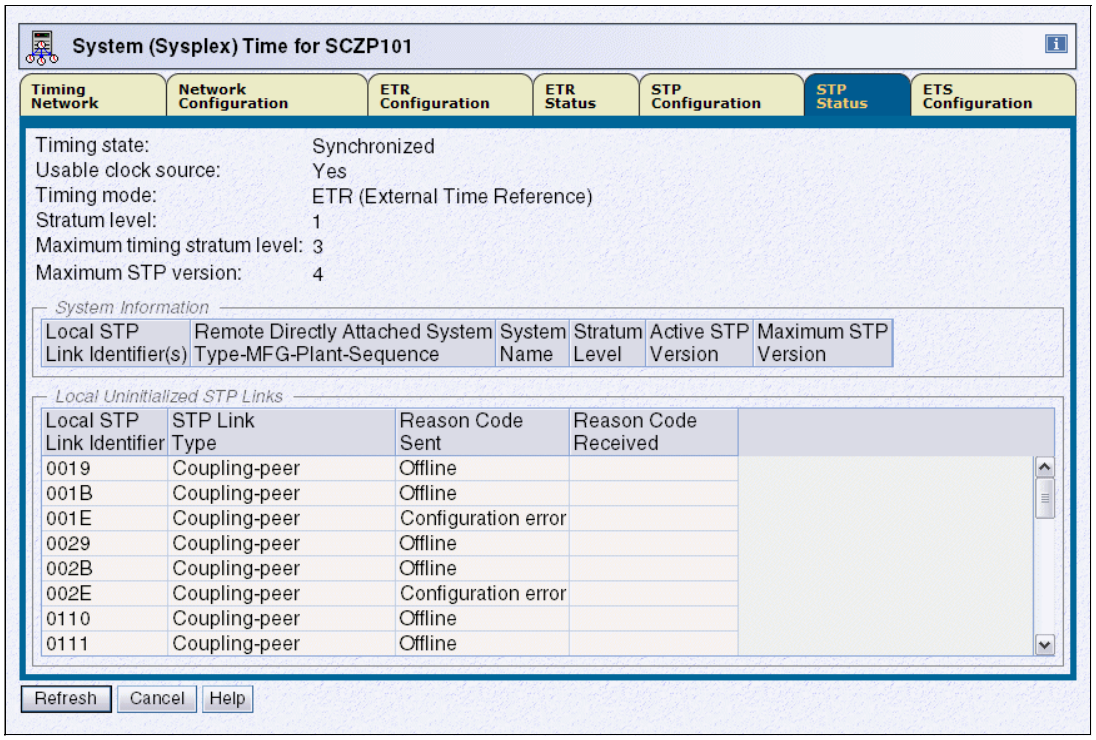

Figure 3-45 STP Status tab: SCZP101 after failure

All servers at Site 2 have lost their time source and become unsynchronized. In Figure 3-46, the STP Status tab for H40 at Site 2 shows that:

- Timing State is unsynchronized.
- Usable clock source is No.
- Stratum level is 0.

Servers at Site 2 H40 and SCZP201 have connectivity to each other but not to SCZP101 at Site 1.

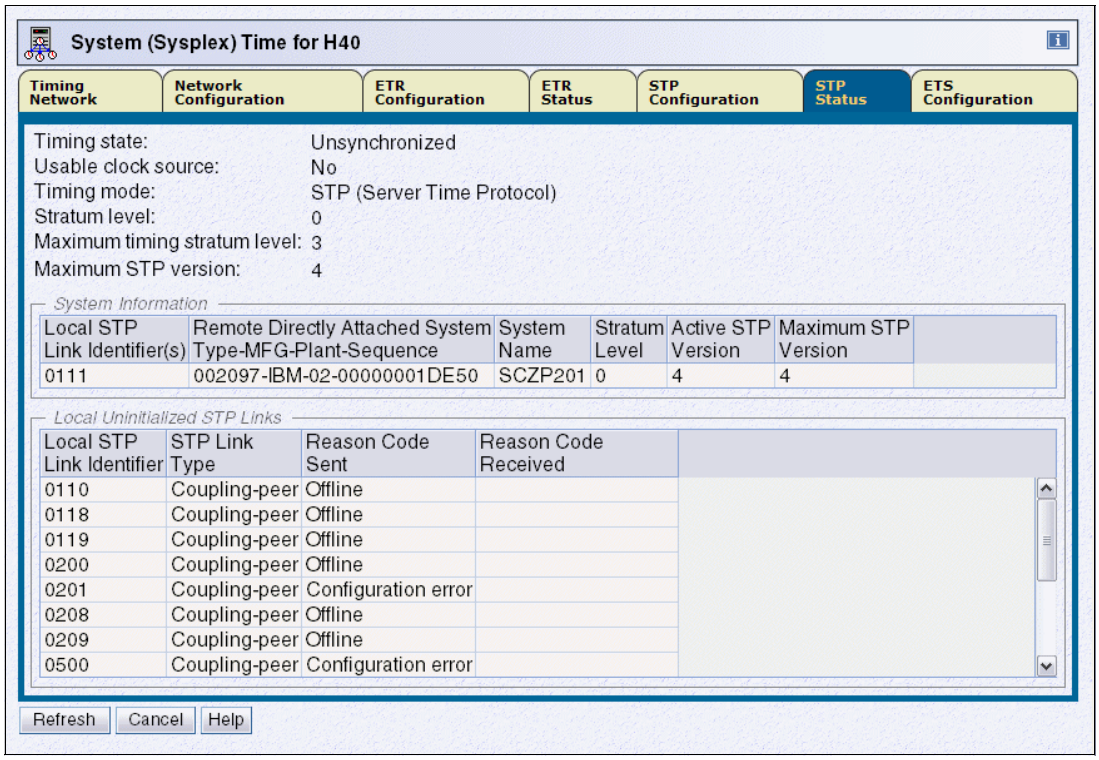

Figure 3-46 STP Status tab: H40 after failure

3.8.3 User actions

If possible, restore communication between Site 1 and Site 2.

- If communication between sites has been restored, use the Operating System messages window, click the **Priority** check box, and respond RETRY to outstanding WTOR message IEA394A.
- If connectivity cannot be restored, you have the option of responding ABORT. The consequences are that z/OS images will go to the *disabled wait* state.

Figure 3-47 shows the STP Status tab of H40 after communication has been restored between Site 1 and Site 2. H40 has returned to Stratum 2, and has connectivity to both Stratum 1 SCZP101 and Stratum 3 SCZP201. All servers on both sites have reverted to their original stratum levels.

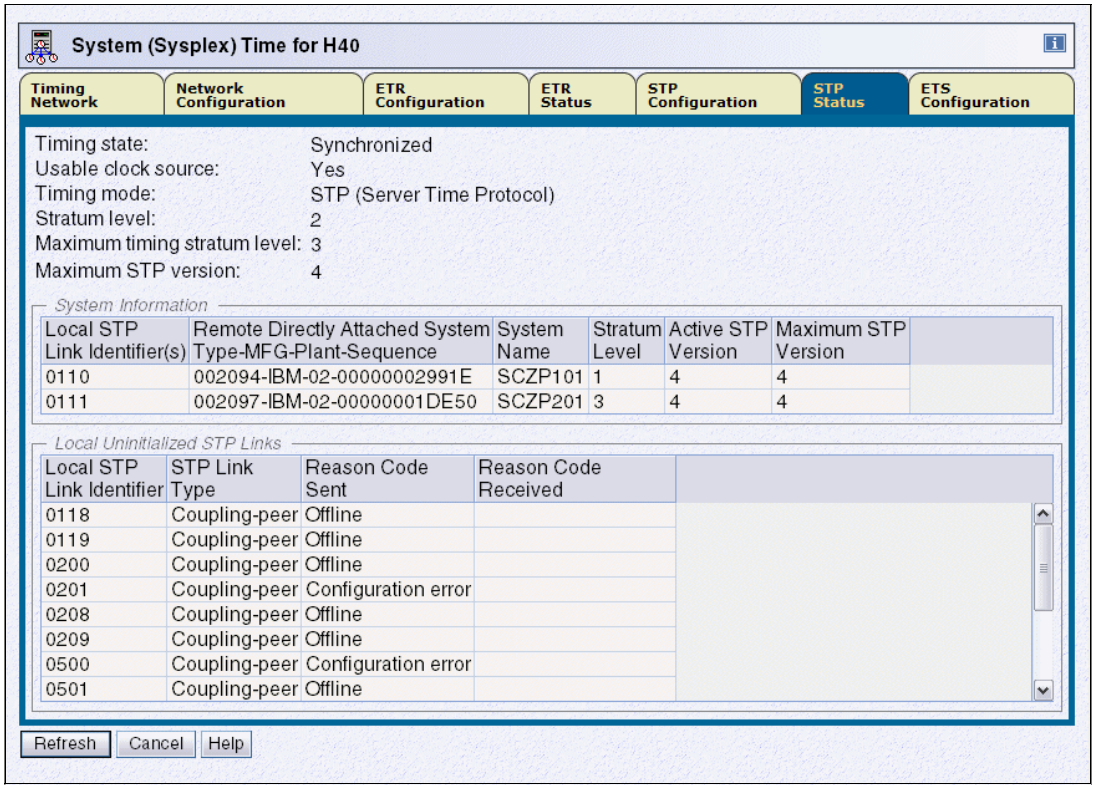

Figure 3-47 STP Status tab: H40 after recovery and user action

4

Recovery in an STP-only CTN with BTS

Recovery in an STP-only CTN is different from a Mixed CTN. The reason for this relates to the assigned roles of PTS, BTS, and (in particular) Arbiter. The following scenarios are described:

- Current Time Server failure - No Going Away Signal (GAS) or OLS received
- Current Time Server failure - Going Away Signal (GAS) or OLS received
- Current Time Server power outage - Internal Battery Feature (IBF)
- Coupling links failure - No Going Away Signal (GAS) or OLS received
- Coupling links failure - Going Away Signal (GAS) or OLS received

4.1 Recovery scenarios overview

In this section, recovery scenarios for an STP-only CTN with only a Backup Time Server and no Arbiter assigned are discussed. All examples in this section assume that the latest MCL maintenance has been applied and the maximum STP version available on both the PTS and BTS is STP Version 4.

For a CTN configuration with more than two servers, we strongly advise having an Arbiter configured. See Chapter 5, “Recovery in an STP-only CTN with BTS and Arbiter” on page 141 for details about that process.

The display of the server names within the Network Configuration and the STP Status panel varies on the status of the HMC-to-SE communication. In most scenarios shown in this chapter it is assumed there is still communication between the HMC and the SE of the Server that has failed. As a consequence, the SE names are shown in the Network Configuration panel and in the STP Status panel. In cases where a server is completely lost (typically when a power drop occurs), the server will be shown with its Node ID, because the HMC cannot communicate with the Support Element at that point in time.

4.2 Current Time Server failure: no Going Away Signal (GAS) or OLS received

A failure of the CTS is detected by the BTS when a Going Away Signal (GAS) is received or if there are no links between the PTS and BTS that support GAS, when Offline signals are received on all established STP paths from the CTS within a time interval of two seconds. The conditions under which the Going Away Signal and the Offline signal are transmitted are described in 1.1.5, “Going Away Signal” on page 6 and 1.1.4, “Server Offline signal” on page 5.

When the condition is detected, the BTS initiates a takeover of the CTS role to become the new Stratum 1 for the CTN. Because only one CTS can exist in an STP-only CTN, the PTS has to surrender its role of the CTS when OLSs are transmitted on all established paths.

However, in the scenario shown in Figure 4-1, note the following points:

- The BTS does not receive a Going Away Signal or an Offline signal on all established links.
- We assume that Console-assisted recovery does not return conclusive information about the state of the CTS.

The BTS cannot make the decision to take over as the CTS. Time synchronization is lost on the BTS, and SCZP301 transitions to Stratum 0.

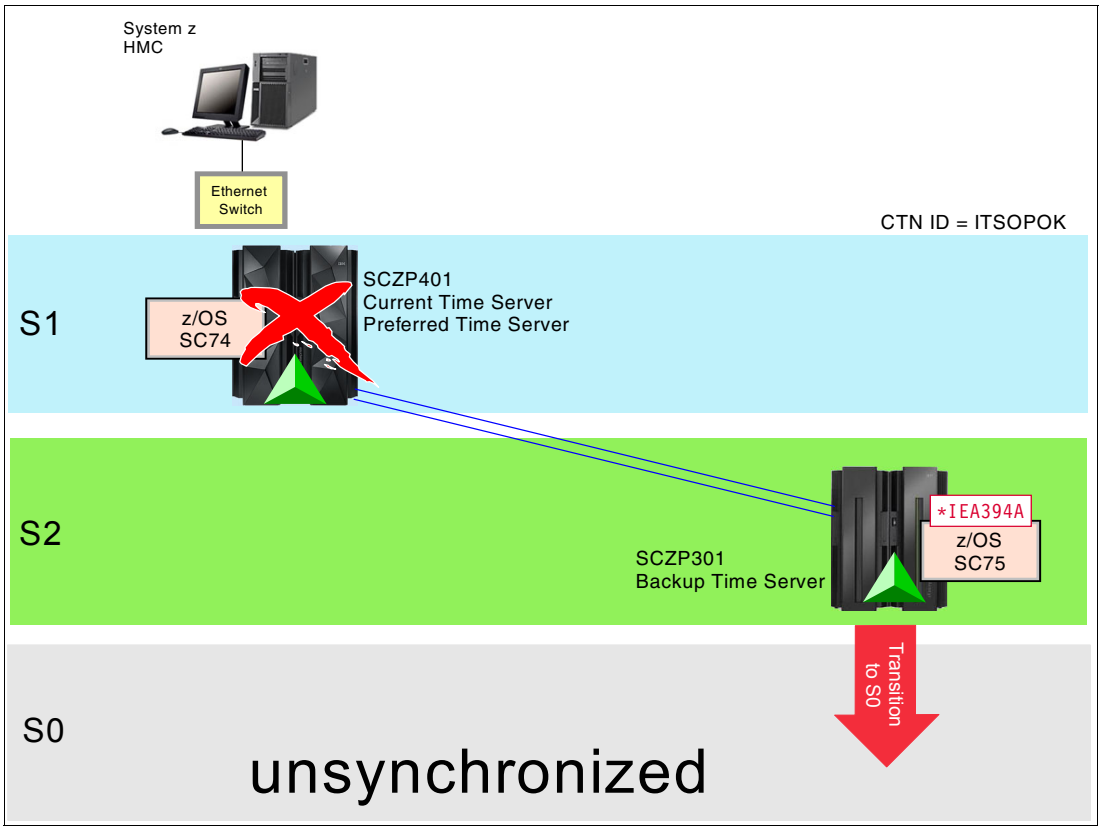

Figure 4-1 STP-only CTN: CTS failure and no Going Away Signal or OLS received

The key highlights of this scenario are:

- The CTS fails.
- The BTS does not receive a Going Away Signal.
- The BTS does not receive Offline signals on multiple links from the CTS.
- The BTS no longer receives timing messages from the CTS.
- The BTS invokes Console-assisted recovery. Because Console-assisted recovery does not confirm the failure of the CTS, the BTS becomes a Stratum 0 server.

  WTOR message IEA394A is issued on all z/OS system images on the BTS.

4.2.1 Problem awareness

As shown in Figure 4-2 on page 118, when SCZP401 fails, z/OS system images on SCZP301 that lose their time source post WTOR message IEA394A.

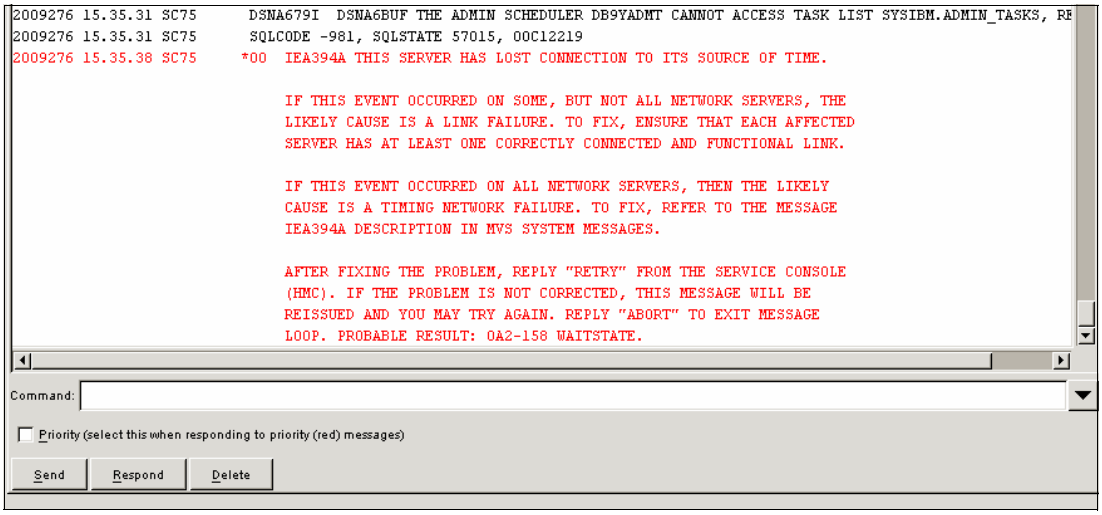

Figure 4-2 SCZP301 WTOR message IEA394A

4.2.2 Problem determination

Figure 4-3 shows the STP Status tab of the BTS SCZP301 after the CTS failure.

- The timing state is unsynchronized.
- Usable clock source is no.
- Stratum level is 0.
- SCZP301 has no STP connectivity to any other server.

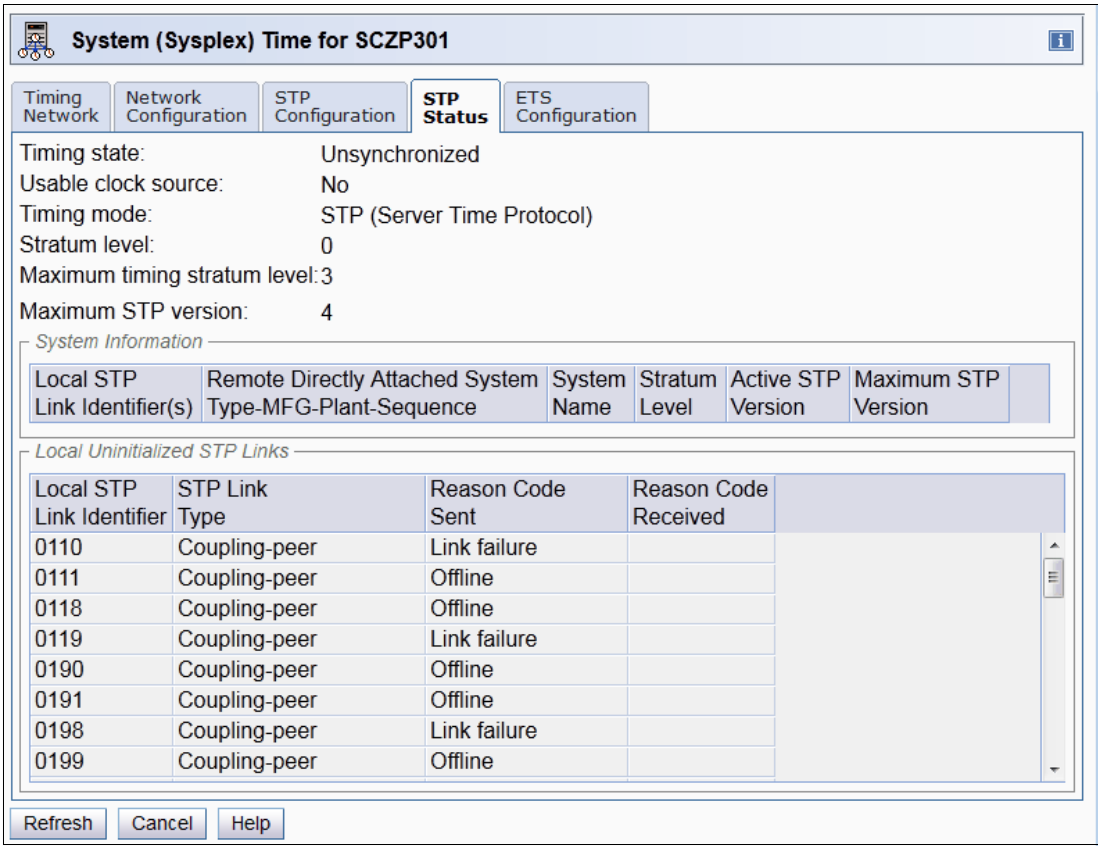

Figure 4-3 STP Status tab: SCZP301 after failure

4.2.3 User actions

If the PTS cannot be recovered, the BTS may be reassigned as the CTS. Note the following recovery steps:

1. Reassign the BTS as CTS and reconfigure the CTN.
2. On the Operating System messages window for z/OS images on SCZP301, click the **Priority** check box and reply `RETRY` to WTOR message IEA394A.

Because SCZP301 is being chosen as the CTS, the reconfiguration must be performed from the Network Configuration tab on SCZP301. Figure 4-4 shows the setup of the Network Configuration tab of SCZP301. Initially, all fields in the Current Network Configuration section are set to `Not configured`.

- SCZP301 is being configured as the Backup Time Server, which was its initial role.
- SCZP401 is selected as the Preferred Time Server. Note that the entry does not show any STP ID field, because SCZP401 is currently not available. However, this ensures that when SCZP401 is repaired and rejoins the CTN, it will automatically recognize its role and provide redundancy.
- The Initialize Time button is enabled. Because SCZP301 is already synchronized, the time does not need to be initialized.
- Because this is an STP-only CTN reconfiguration after a CTS failure, the **Force configuration** box must be clicked to skip the CTS connectivity checks.
- Click **Apply** to reconfigure the CTN with SCZP301 as the new CTS.

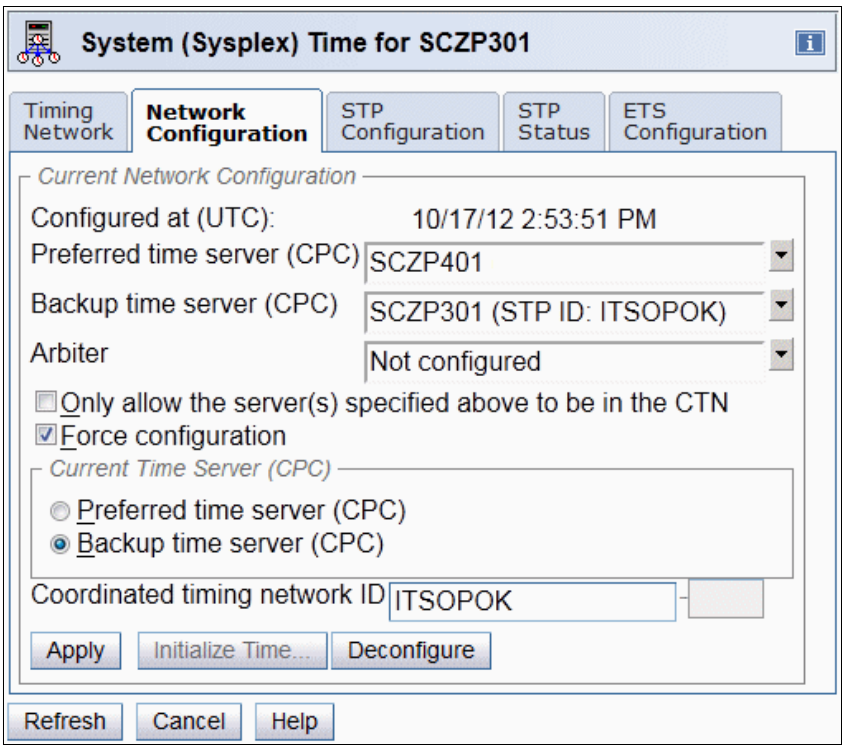

Figure 4-4 Network Configuration tab: SCZP301 apply

Because **Force configuration** was selected, message ACT37348 is displayed, as shown in Figure 4-5. Click **Yes** to continue.

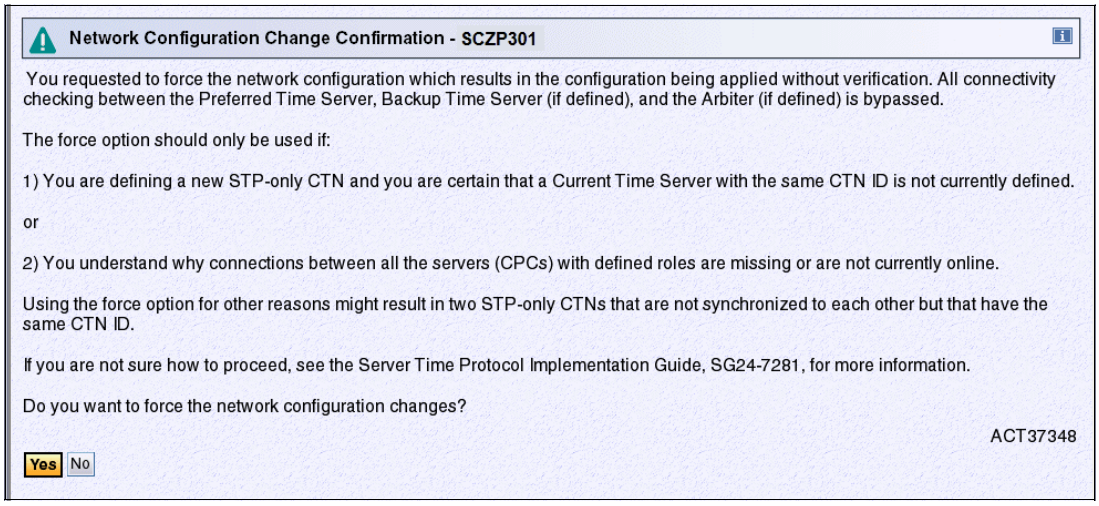

Figure 4-5 Message ACT37348

As shown in the STP Status tab of the CTS SCZP301 in Figure 4-6, after successful configuration, SCZP301 transitions to Stratum 1.

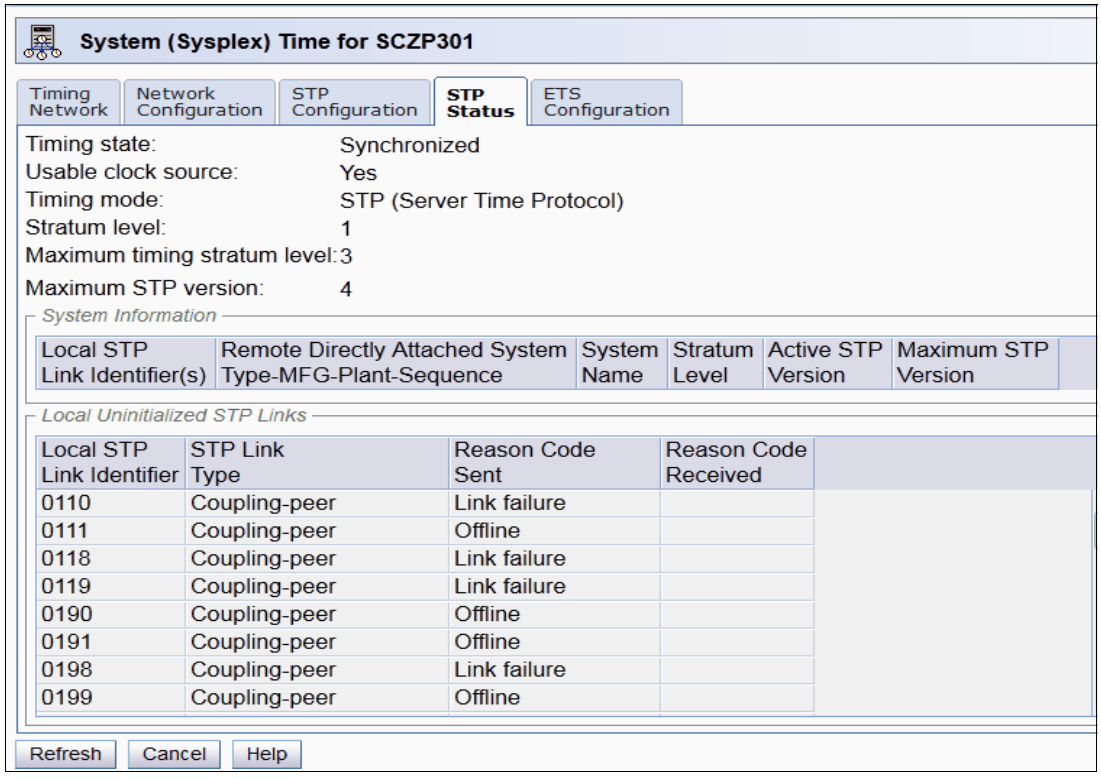

Figure 4-6 STP status tab: SCZP301 after recovery and user action

Figure 4-7 shows the Network Configuration tab for SCZP301.

- The Current Network Configuration section shows SCZP301 being defined as the BTS.
- The Current Time Server (CPC) section shows that the BTS (SCZP301) is now the CTS.

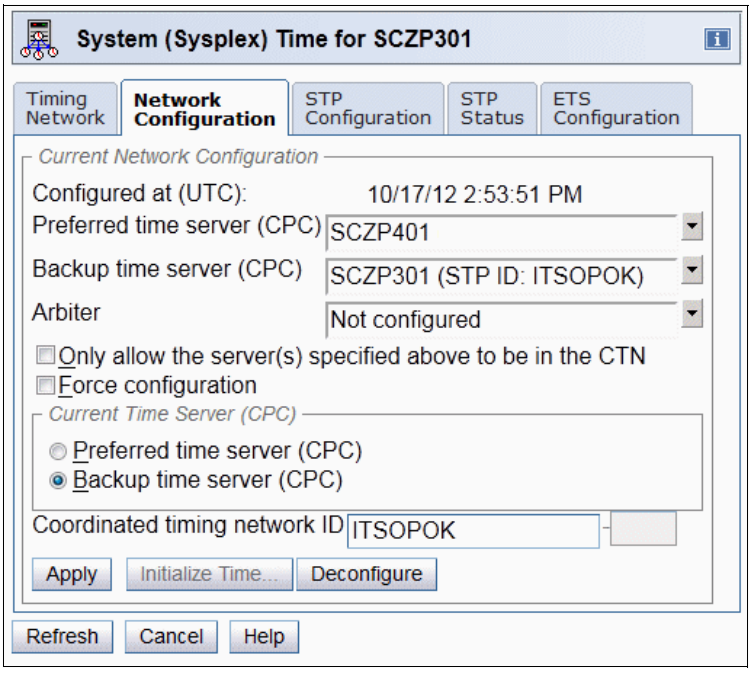

Figure 4-7 Network Configuration tab: SCZP301 after recovery and user action

4.3 Current Time Server failure - Going Away Signal (GAS) or OLS received

The configuration for this scenario is shown in Figure 4-8. The difference from the previous scenario is that the BTS successfully receives either the Going Away Signal or the Offline signals issued from the CTS on all the links within a 2-second time period. As a result, the PTS gives up its CTS role and the BTS takes over as the CTS and transitions to Stratum 1. Console assisted recovery is later invoked by the BTS to verify that the PTS has indeed failed.

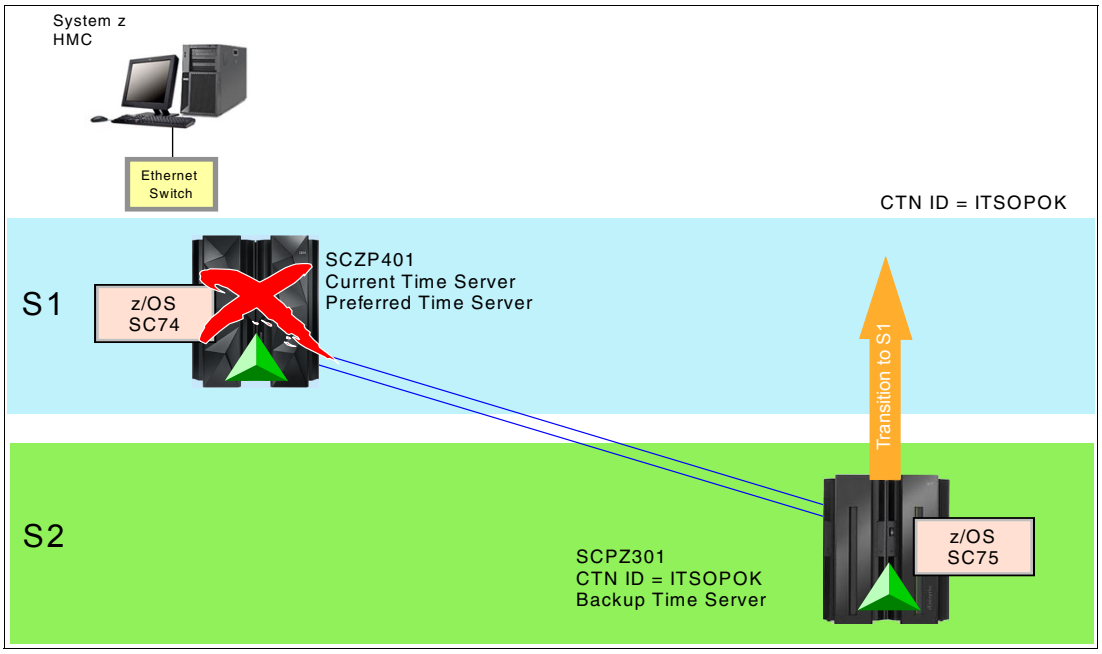

Figure 4-8 STP-only CTN: CTS failure and OLS received

These are the key highlights of this scenario:

- The PTS fails, issues a Going Away Signal if it has the required link type or Offline signals on multiple links to the BTS, and releases its CTS role.
- If the BTS receives the Going Away Signal or Offline signals on multiple links from the PTS within two seconds, it takes over the CTS role. The BTS invokes Console-assisted recovery to confirm the PTS failure.

4.3.1 Problem awareness

z/OS messages are issued indicating that z/OS image SC74 on SCZP401has failed:

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC74 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
```

An active SFM system isolation policy automatically starts system partitioning. This is highlighted by an IXC101I message only a few seconds after detection of the failure. Partitioning triggers the cleanup of resources for the failing system.

If system isolation fails, then SFM issues WTOR message IXC102A after the XCF CLEANUP time has elapsed:

```
IXC102A XCF IS WAITING FOR SYSTEM SC74 DEACTIVATION. REPLY DOWN WHEN MVS ON SC74
HAS BEEN SYSTEM RESET
```

Note: Before replying `DOWN` to IXC102A or IXC402D, you must perform a hardware SYSTEM RESET on the z/OS system being removed. This is necessary to ensure that the system releases any outstanding I/O reserves. A SYSTEM RESET ensures that other systems continue to have access to the data sets on the shared devices.

Because the Current Time Server was switched from the Preferred Time Server to the Backup Time Server when the PTS failed, z/OS (Version 1.11 or higher) issues message

IEA395I on all z/OS images that are members of this CTN and use STPMODE=YES in the CLOCKxx member:

```
IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE BACKUP
```

4.3.2 Problem determination

Figure 4-9 on page 123 shows the Network Configuration tab for the BTS SCZP301:

- The Preferred Time Server field of the Current Network Configuration section shows SCZP401 being defined as the PTS, but there is no STP ID associated with it (as compared with the Backup Time Server field). This indicates that the server Support Element is still visible to the HMC, but SCZP301 no longer has STP connectivity to SCZP401.
- The Current Time Server (CPC) section shows that the BTS SCZP301 is now the CTS.

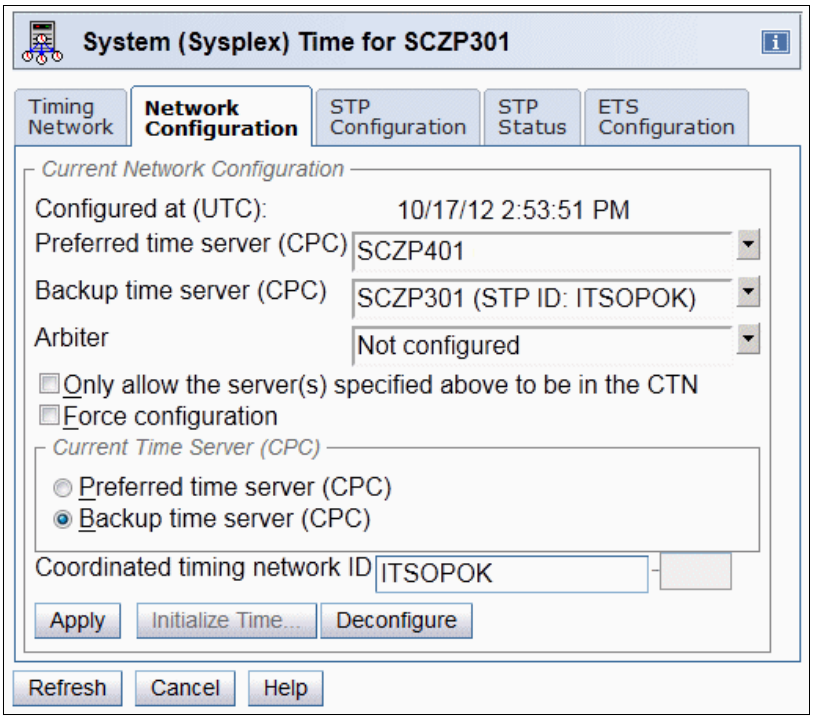

Figure 4-9 Network Configuration tab: SCZP301 after failure

Figure 4-10 shows the STP Status tab of SCZP301 after SCZP401 has failed. SCZP301 has transitioned to Stratum 1 and has no STP connectivity to SCZP401.

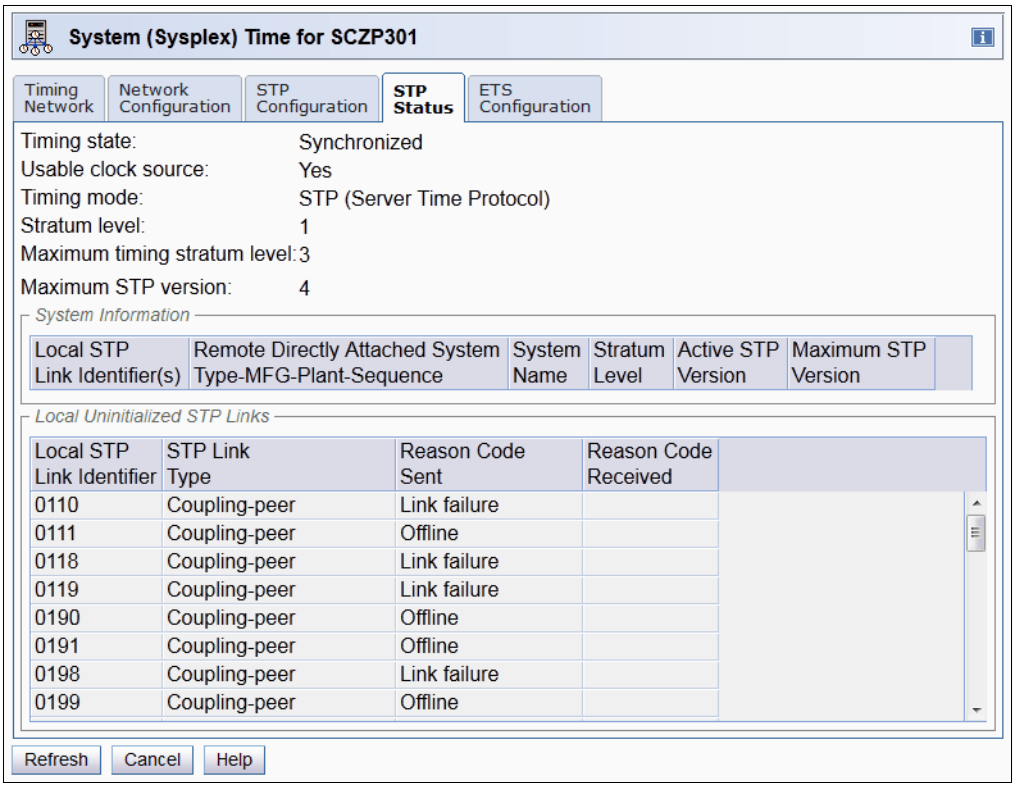

Figure 4-10 STP Status tab: SCZP301 after failure

4.3.3 User actions

No user actions are required for the CTN. The BTS SCZP301 takes over the role of CTS. SCZP401 remains unavailable until the repair action is completed.

When SCZP401 is restarted, it joins the CTN. Because the STP-only CTN was not reconfigured by reassigning roles during the outage, SCZP401 automatically takes back its original role as CTS. As a consequence, z/OS (Version 1.11 or higher) issues message IEA395I on all z/OS images that are members of this CTN and uses STPMODE=YES in the CLOCKxx member:

```
IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE PREFERRED
```

Figure 4-11 shows the Network Configuration tab of SCZP301 after SCZP401 has been recovered:

- The Preferred Time Server field of the Current Network Configuration section shows SCZP401 being the PTS with an STP ID. This indicates that SCZP301 now has connectivity to SCZP401.
- The Current Time Server (CPC) section shows that SCZP401 is the CTS.

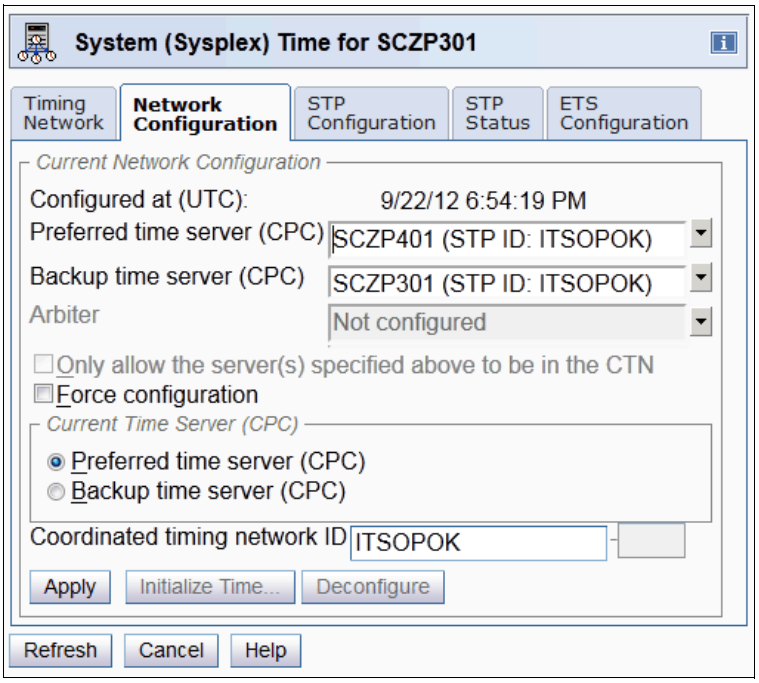

Figure 4-11 Network Configuration tab: SCZP301 after recovery

4.4 Current Time Server power outage - Internal Battery Feature (IBF)

The configuration is shown in Figure 4-12 on page 126. In this scenario, the CTS is hit by a power outage, but not affected immediately because it does have the Internal Battery Feature (IBF) installed. The PTS/CTS notifies the BTS through the coupling links that it is running on IBF power. If the BTS does not get a “normal power” status within the next 30 seconds from the PTS/CTS, it initiates processing to take over the CTS role. For details about the IBF recovery, refer to 1.3, “Internal Battery Feature considerations” on page 16.

In this scenario, we assume that the power outage exists for more than 30 seconds, so the BTS takes over as Stratum 1. Furthermore, the scenario shows that the PTS fails because the IBF ran out of power. The time the IBF can support the PTS depends on the hardware configuration of the PTS.

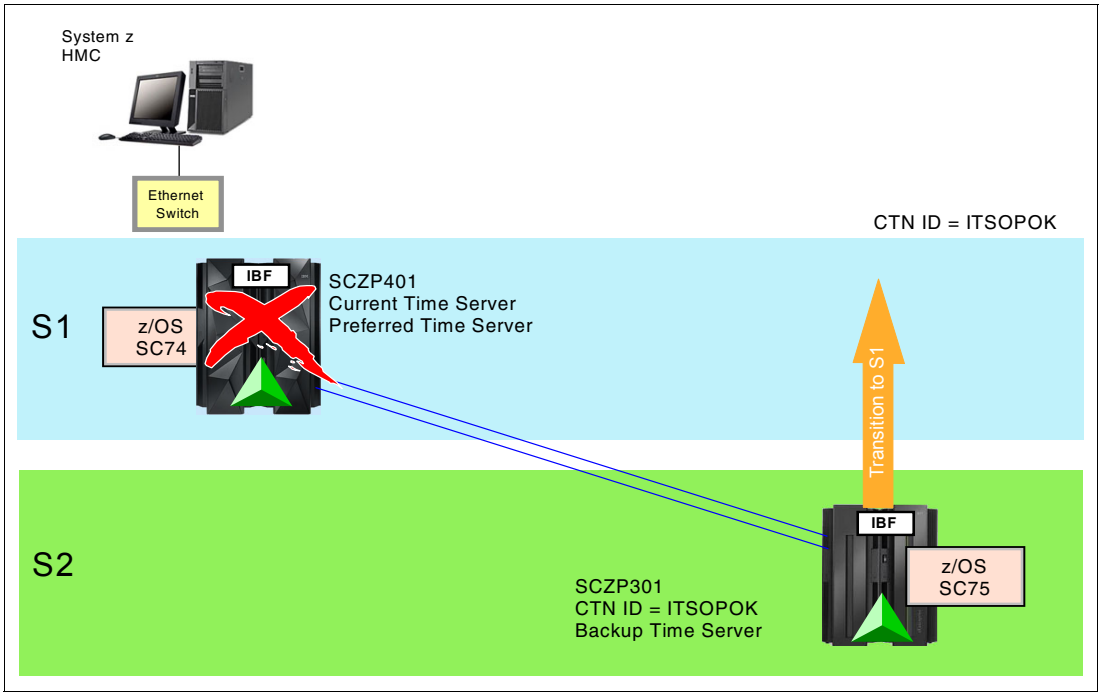

Figure 4-12 STP-only CTN: CTS failure with Internal Battery Feature installed in the PTS and BTS

The key highlights of this scenario are:

- The PTS/CTS is hit by a power outage, but the PTS power is supplied by the IBF of the PTS.
- The PTS/CTS sends a message through the coupling links to the BTS indicating that it is running on IBF power.
- The BTS takes over the CTS role and becomes Stratum 1 if this condition exists for more then 30 seconds, while the PTS transitions to Stratum 2.
- The PTS is lost when its IBF runs out of power.

4.4.1 Problem awareness

A Hardware Message is posted on the HMC indicating that there is a power problem at the PTS. However, there is no impact at this point because the PTS stays on battery power for a certain period of time.

Because the Current Time Server switched from the Preferred Time Server to the Backup Time Server after 30 seconds, z/OS (Version 1.11 or higher) issues message IEA395I on all z/OS images that are members of this CTN and use STPMODE=YES in the CLOCKxx member:

```
IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE BACKUP
```

After the IBF has been drained, the PTS and its LPARs are lost. As a consequence, z/OS messages are issued indicating that z/OS image SC74 on SCZP401 has failed:

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC74 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
```

An active SFM system isolation policy automatically starts system partitioning. This is highlighted by an IXC101I message only a few seconds after detection of the failure. Partitioning triggers the cleanup of resources for the failing system. If system isolation fails, then SFM issues WTOR message IXC102A after the XCF CLEANUP time has elapsed:

```
IXC102A XCF IS WAITING FOR SYSTEM SC74 DEACTIVATION. REPLY DOWN WHEN MVS ON SC74
HAS BEEN SYSTEM RESET
```

4.4.2 Problem determination

If the BTS has not received a “normal power” message for more than 30 seconds, the BTS takes over as CTS. The Network Configuration tab in Figure 4-13 shows the following information:

- The Current Time Server has switched to the Backup Time Server.
- Automated network recovery is temporarily disabled.

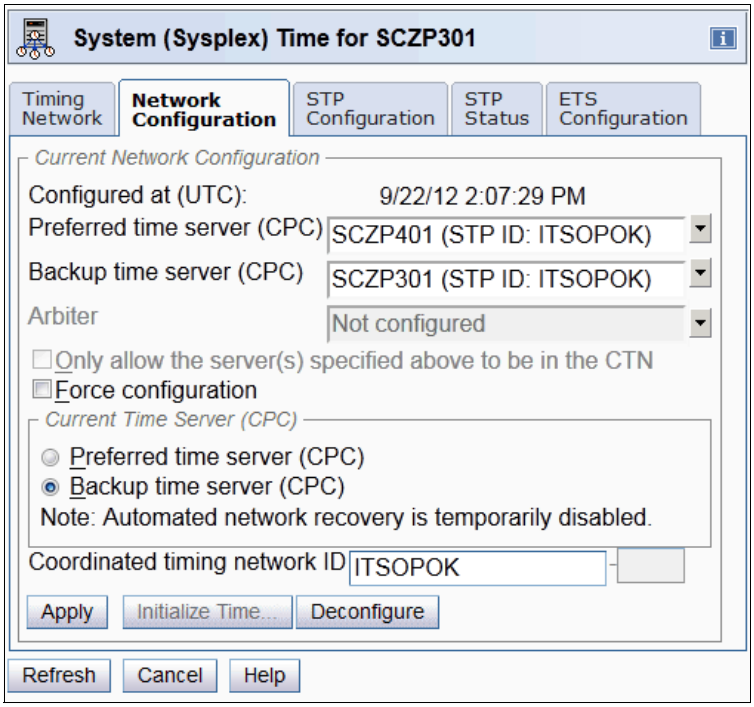

Figure 4-13 Network Configuration tab: Server SCZP301 - 30 seconds after the power outage

In addition, the STP status tab shows server SCZP301 being Stratum 1. Furthermore, after the IBF has been drained, server SCZP401 is powered down. After SCZP401 is lost, the Network Configuration window on SCZP301 still shows SCZP401 being the PTS, but its subsequent STPID is missing. This indicates that the CTN has lost communication to server SCZP401. The STP status tab shows that SCZP301 is Stratum 1 and there is no more connectivity to server SCZP301, as shown in Figure 4-14 on page 128.

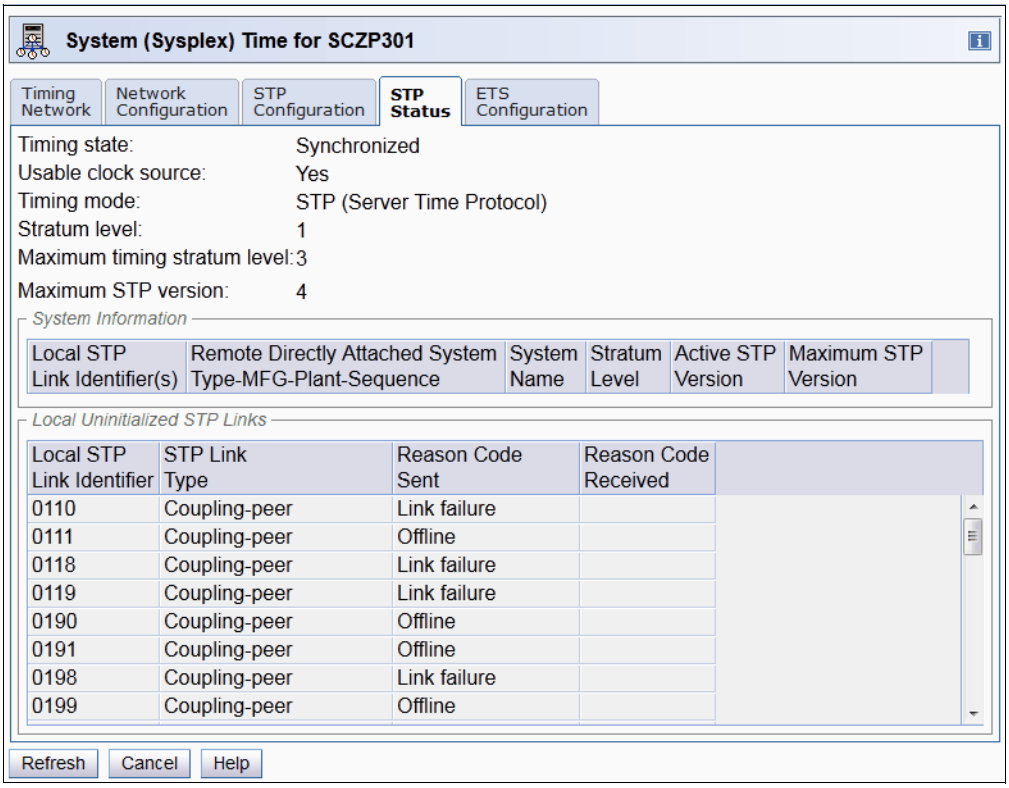

Figure 4-14 STP status tab: SCZP301 (BTS) after the IBF of SCZP401 is discharged

4.4.3 User actions

No user actions are required for the CTN. The BTS (SCZP301 in this scenario) took over the role of the CTS. SCZP401 remains unavailable until its power has been restored.

After SCZP401 has been powered up again, it joins the CTN. Because the STP-only CTN was not reconfigured by reassigning roles during the outage, SCZP401 will automatically take back its original role as CTS. As a consequence, z/OS (Version 1.11 or higher) issues message IEA395I on all z/OS images that are members of this CTN and use STPMODE=YES in the CLOCKxx member:

```
IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE PREFERRED
```

Figure 4-15 shows the Network configuration tab of SCZP301 after SCZP401 has been recovered:

- The Current Time Server (CPC) section shows that SCZP401 is the CTS.
- Automated network recovery is no longer disabled.

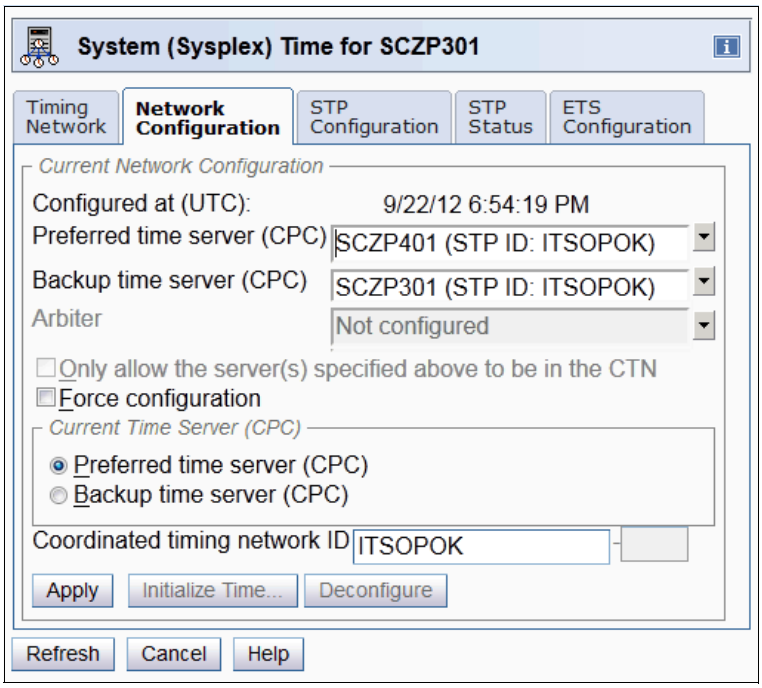

Figure 4-15 Network configuration tab: SCZP301 after recovery

4.5 Backup Time Server failure

Figure 4-16 shows an STP-only CTN configuration, the PTS and CTS are SCZP401, and the BTS is SCZP301. In this scenario, the BTS SCZP301 fails.

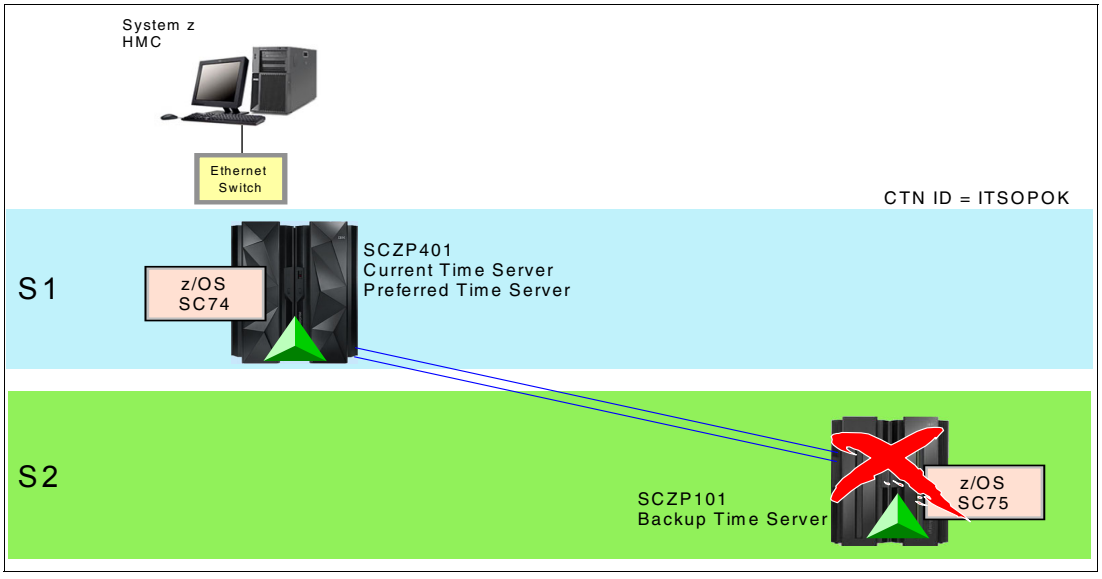

Figure 4-16 STP-only CTN: BTS failure

The BTS failure has no impact on the PTS.

4.5.1 Problem awareness

z/OS messages are issued indicating that z/OS images on server SCZP301 have failed:

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC75 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
```

An active SFM system isolation policy automatically starts system partitioning. This is highlighted by an IXC101I message only a few seconds after detection of the failure. Partitioning triggers the cleanup of resources for the failing system. If system isolation fails, then SFM will issue WTOR message IXC102A after the XCF CLEANUP time has elapsed:

```
IXC102A XCF IS WAITING FOR SYSTEM SC75 DEACTIVATION. REPLY DOWN WHEN MVS ON SC75
HAS BEEN SYSTEM RESET
```

Note: Before replying `DOWN` to IXC102A or IXC402D, the user must perform a hardware SYSTEM RESET on the z/OS system being removed. This is necessary to ensure that the system releases any outstanding I/O reserves. A SYSTEM RESET ensures that other systems continue to have access to the data sets on the shared devices.

4.5.2 Problem determination

As shown in Figure 4-17, issuing the **`DISPLAY ETR`** command on a z/OS image on SCZP401 shows that the PTS has no link to the Backup Time Server.

```
DISPLAY ETR
IEA386I 16.27.22 TIMING STATUS 099
SYNCHRONIZATION MODE = STP
THIS SERVER IS A STRATUM 1
CTN ID = ITSOPOK
THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
THIS IS THE PREFERRED TIME SERVER
THIS SERVER HAS NO LINK TO THE BACKUP TIME SERVER
THIS STP NETWORK HAS NO SERVER TO ACT AS ARBITER
```

Figure 4-17 DISPLAY ETR

The **`DISPLAY XCF`** command in Figure 4-18 shows that SC75 is no longer a part of the sysplex.

```
DISPLAY XCF,SYSPLEX
IXC336I  16.28.40  DISPLAY XCF 104
SYSPLEX PLEX75
SYSTEM   TYPE SERIAL LPAR STATUS TIME         SYSTEM STATUS
SC74     2827 B8D7   01   10/01/2012 13:33:11 ACTIVE         TM=STP
```

Figure 4-18 DISPLAY XCF: SYSPLEX

Depending on the cause of the server failure, the System (Sysplex) Time task might or might not be available for SCZP301.

Figure 4-19 on page 131 shows the Network Configuration tab for SCZP401. The Current Network Configuration section shows SCZP301 being defined as the BTS, but there is no

STP ID associated with it (unlike the Preferred Time Server field). This indicates that the Support Element for SCZP301 is still visible to the HMC, but SCZP301 is not available or has no STP connectivity to SCZP401, so SCZP301 is not functional as a BTS.

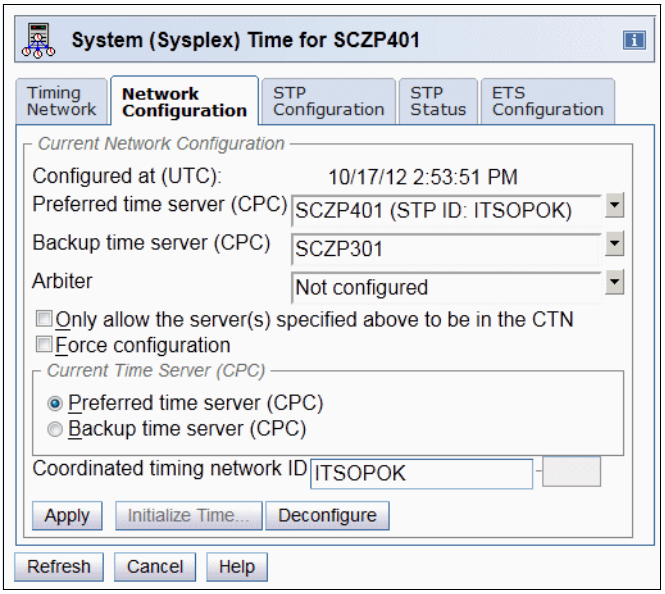

Figure 4-19 STP-only CTN: Network Configuration tab, SCZP401 view

The STP Status tab on SCZP401 displays that there is no connectivity to another system (including SCZP301), as shown in Figure 4-20.

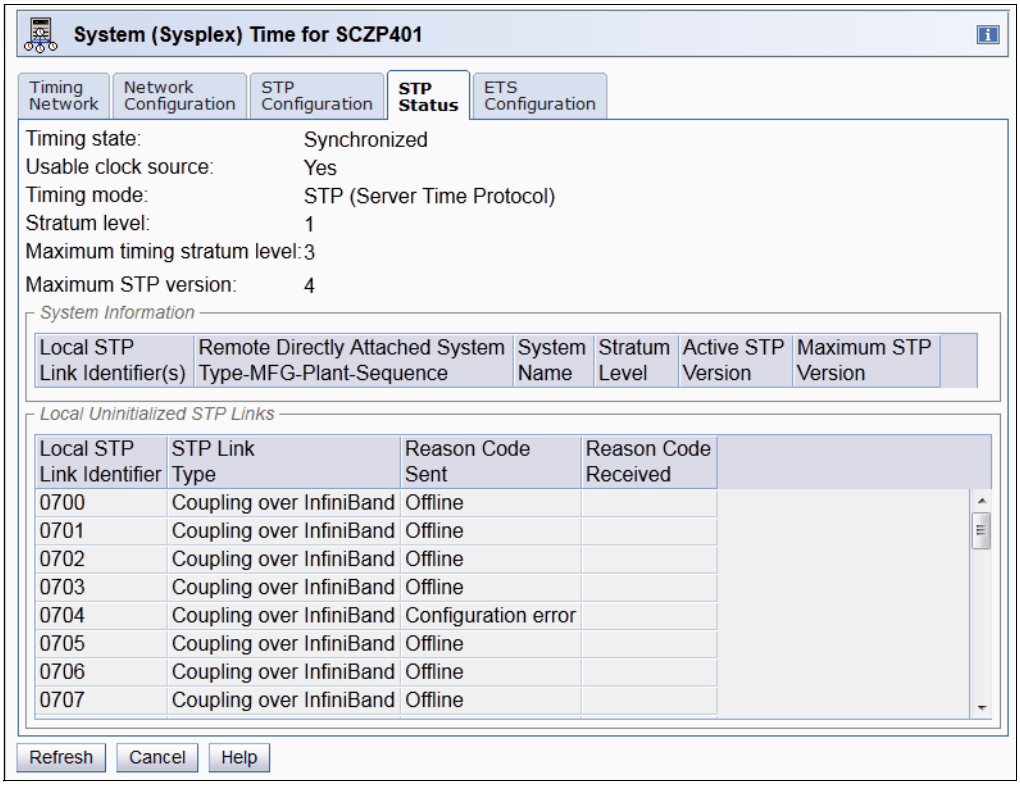

Figure 4-20 STP-only CTN: STP Status tab, SCZP401 view

4.5.3 User actions

Failure of the BTS does not affect the PTS (because it is also the CTS). SCZP401 will continue as the CTS.

1. Repair SCZP301 using installation recovery procedures.

   After the repair action, SCZP301 joins the CTN and regains its role as BTS. In Figure 4-21, the Network Configuration tab shows SCZP301 defined as the BTS in the Current Network Configuration section.

2. Re-IPL z/OS image SC75 on SCZP301.

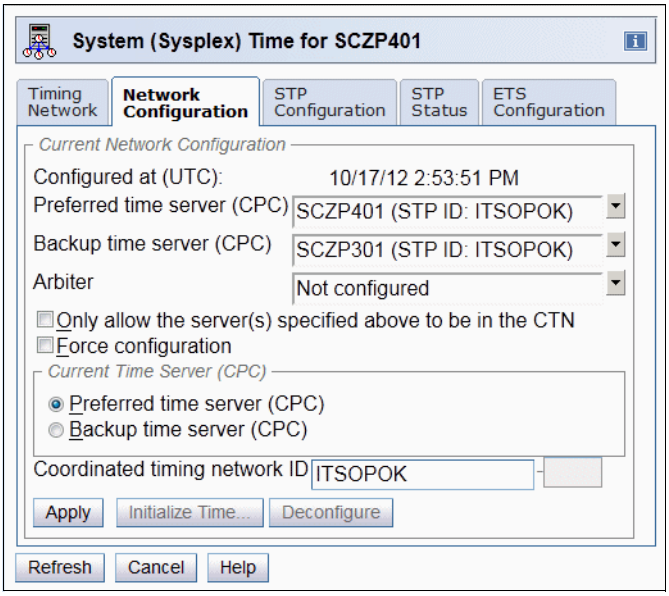

Figure 4-21 Network configuration tab after SCZP401 repair action

On the STP Status tab for SCZP401 (Figure 4-22 on page 133), the local STP link identifiers show that STP connectivity between SCZP401 and SCZP301 has been reestablished.

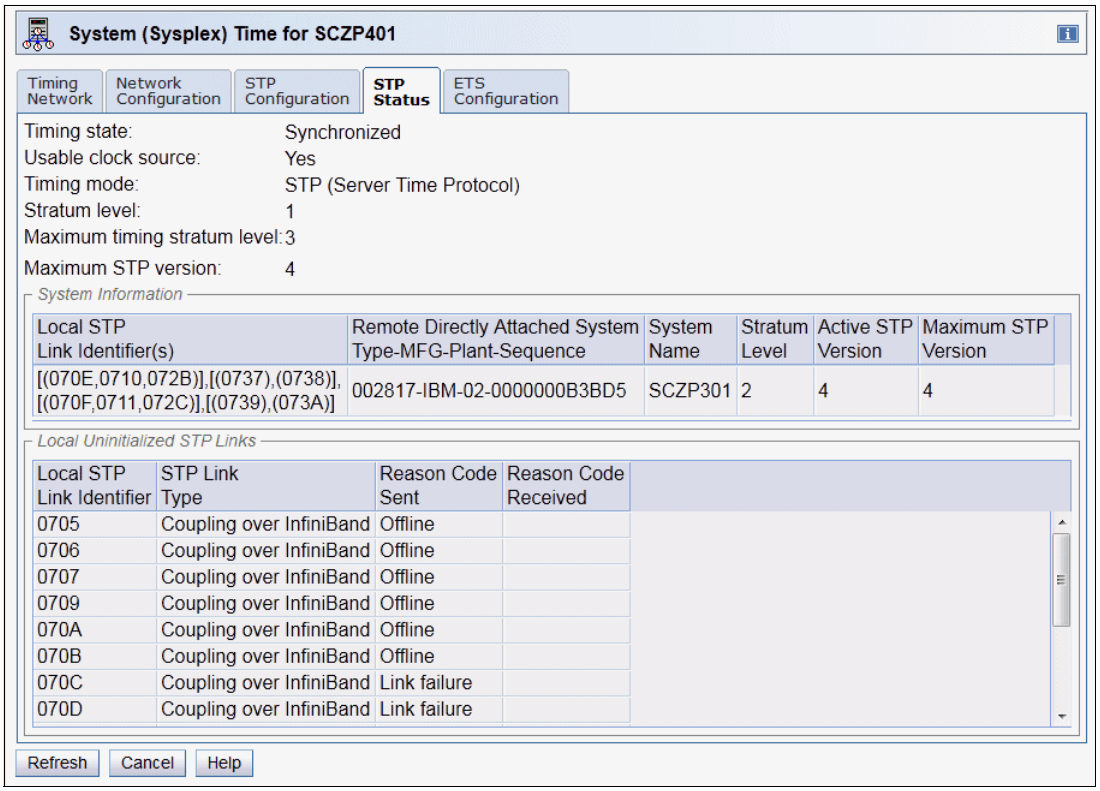

Figure 4-22 STP Status tab after SCZP401 repair action

4.6 Coupling links failure: No Going Away Signal (GAS) or OLS received

When multiple coupling links are configured, and not all coupling links fail, there is no indication of an STP problem because one of the remaining links is used for STP messaging. However, if a failure leaves a server with a single link to use as a timing source, z/OS information message IEA382I is issued on the z/OS images:

```
IEA382I THIS SERVER HAS ONLY A SINGLE LINK AVAILABLE FOR TIMING PURPOSES
```

When a coupling link fails between CTS and BTS, this appears to each server as loss of communication between the servers. Recovery is different from the server failure scenarios, because each server is still functional.

In this scenario, a Going Away Signal is not issued, and Offline signals (OLS) are not transmitted by the CTS or the failure involves only the last path in a two-second interval. This might be because only one coupling link was configured between the PTS and the BTS (not recommended), or because other links have been previously removed or have failed. The following actions occur:

- The BTS no longer receives STP timing messages from the CTS.
- A Going Away Signal is not received by the BTS,
- OLS has not been received on multiple links in the 2-second interval, and the BTS does not recognize the Offline signal for a single link.
- The BTS invokes Console-assisted recovery, which confirms that the CTS is still up and running.

- The BTS becomes unsynchronized and transitions to Stratum 0.
- The CTS continues to function without impact.

In the configuration shown in Figure 4-23, the last coupling link between the CTS SCZP401 and the BTS SCZP301 fails. SCZP401 remains as the CTS and continues to function, while SCZP301 becomes unsynchronized (Stratum 0). The key highlights of this scenario are:

- The BTS does not receive Offline signals from the CTS. The BTS no longer receives timing from CTS, and it becomes unsynchronized and goes to Stratum 0.
- WTOR message IEA394A is issued on all z/OS system images on the BTS.
- The CTS remains active and continues to function without impact.

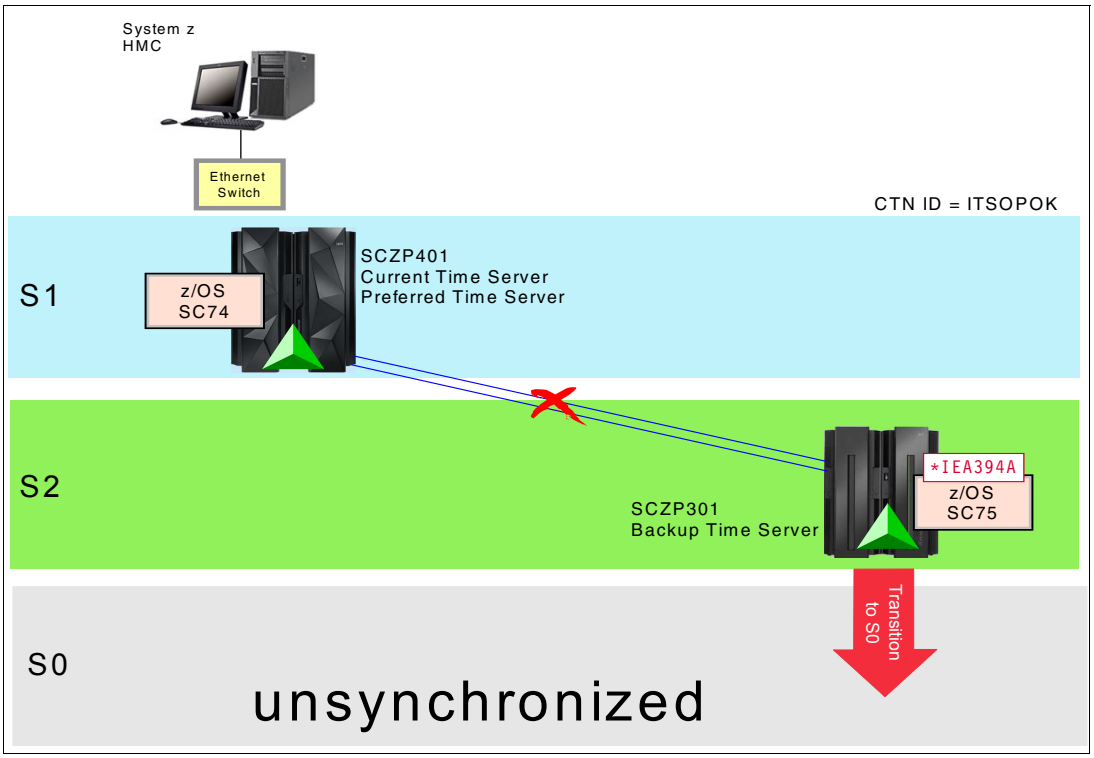

Figure 4-23 STP-only CTN: last coupling link failure

4.6.1 Problem awareness

As shown in Figure 4-24 on page 135, every system on the BTS SCZP301 that loses its time source posts WTOR message IEA394A.

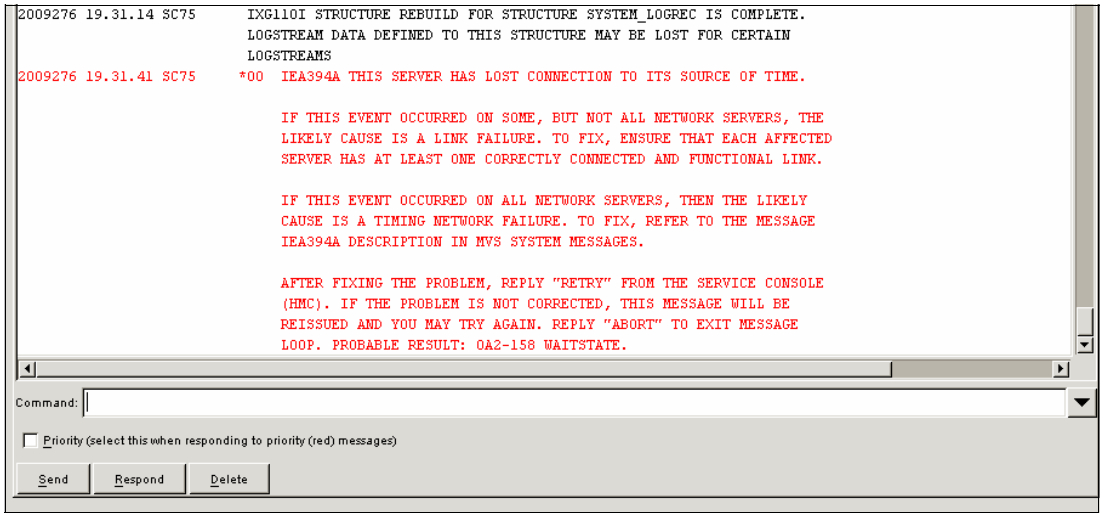

Figure 4-24 z/OS WTOR message IEA394A: SCZP301

4.6.2 Problem determination

Figure 4-25 shows the STP Status tab of the CTS SCZP401 after the link failure. With the failure of the link to the BTS SCZP301, SCZP401 has no connectivity to any other servers.

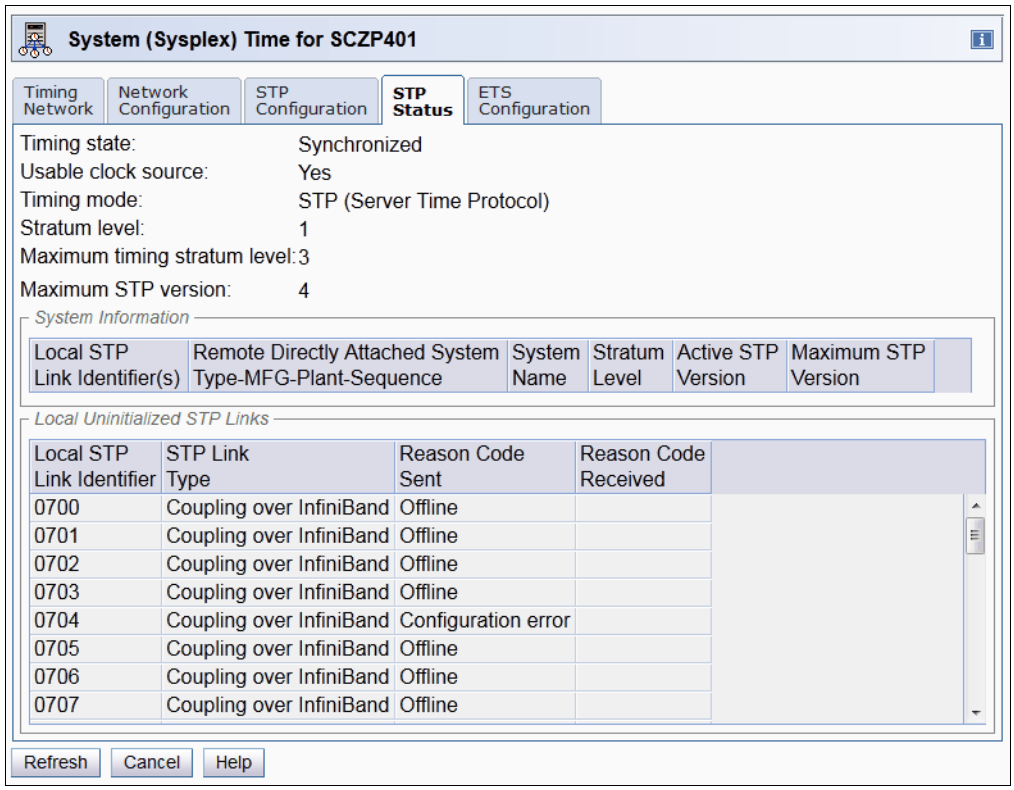

Figure 4-25 STP Status tab: SCZP401

4.6.3 User actions

The CTS SCZP401 continues to function. To avoid an outage (if the CTS SCZP401 fails), take these actions:

1. Repair the failing links using installation recovery procedures.

   The BTS SCZP301 will subsequently rejoin the CTN.

2. On the Operating System messages window for z/OS image SC75 on SCZP301, click the Priority check box and reply `RETRY` to WTOR message IEA394A.

4.7 Coupling links failure: Going Away Signal (GAS) or OLS received

When multiple coupling links are configured and not all coupling links fail, there is no indication of an STP problem because one of the remaining links will be selected. If the failure involves the last link between the PTS and the BTS, then the BTS loses its timing source.

In this example, the configuration shown in Figure 4-26 is similar to the previous scenario shown in 4.6, “Coupling links failure: No Going Away Signal (GAS) or OLS received” on page 133. The difference is that the failure affects multiple links in a 2-second interval and the BTS will recognize the OLS on multiple links from the CTS, or a Going Away Signal is received. The following actions occur:

- The CTS sends a Going Away Signal or OLS on all paths to the BTS and gives up its CTS role. WTOR message IEA394A is issued on all z/OS system images on SCZP401.
- When receiving the Going Away Signal or OLS on multiple paths from the PTS, the BTS takes over the CTS role. The BTS also invokes Console-assisted recovery. In this example, Console-assisted recovery confirms that SCZP401 is stratum 0 and only SCZP301 is assigned the CTS role.

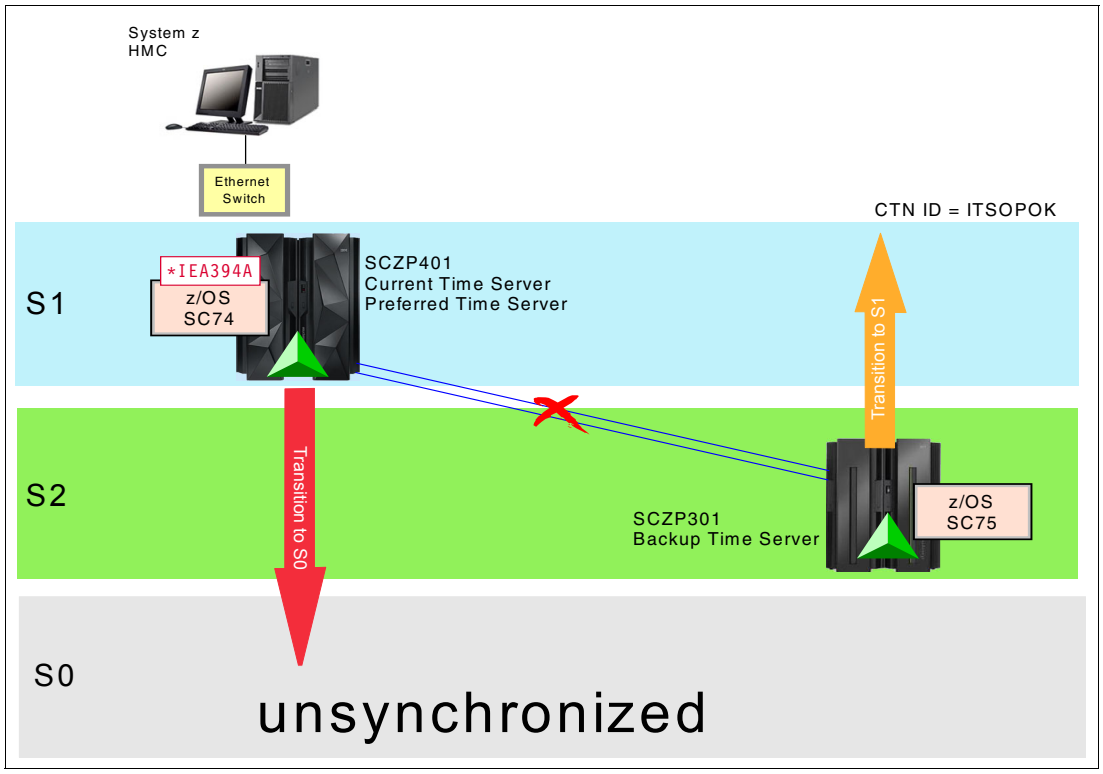

Figure 4-26 STP-only CTN: multiple coupling links failure

4.7.1 Problem awareness

On SCZP401, every z/OS image that has ETRMODE YES or STPMODE YES issues WTOR message IEA394A, as shown in Figure 4-27.

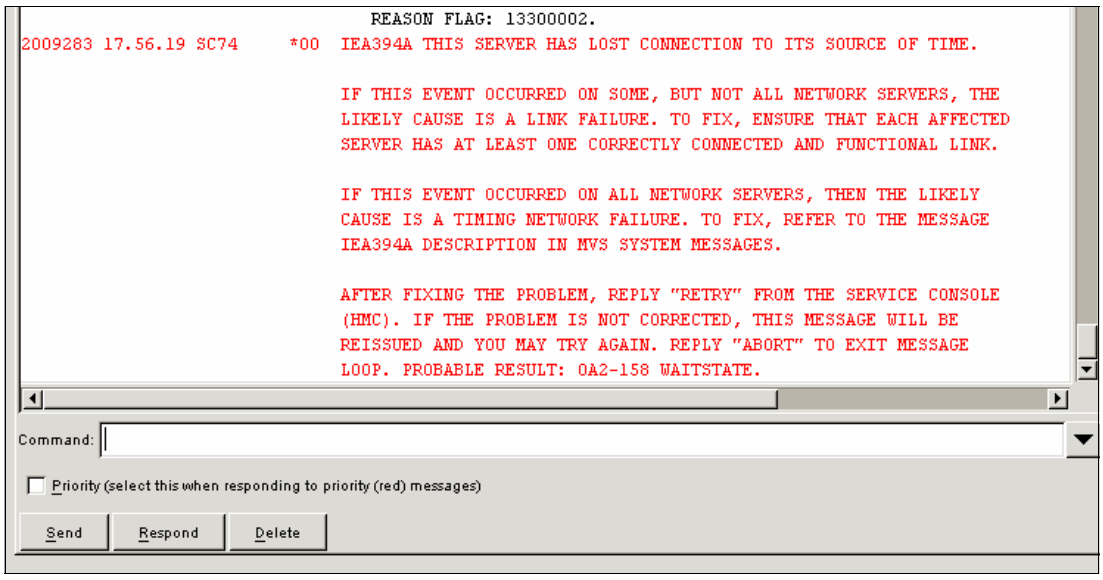

Figure 4-27 z/OS WTOR message IEA394A: SCZP401

Because the Current Time Server was switched from the Preferred Time Server to the Backup Time Server when the PTS failed, z/OS (Version 1.11 or higher) issues message

IEA395I on all z/OS images that are members of this CTN and use STPMODE=YES in the CLOCKxx member (SC75 in this case):

```
IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE BACKUP
```

4.7.2 Problem determination

In this scenario, identifying the problem requires examination of the STP Configuration tabs for the CTS and the BTS. In this situation, they look different. Figure 4-28 shows the Network Configuration tab of the CTS SCZP401 after the last coupling link has failed:

- All fields in the Current Network Configuration section are set to `Not configured`. This implies that SCZP401 is not synchronized.
- The Initialize Time button is enabled, while all other buttons are disabled.

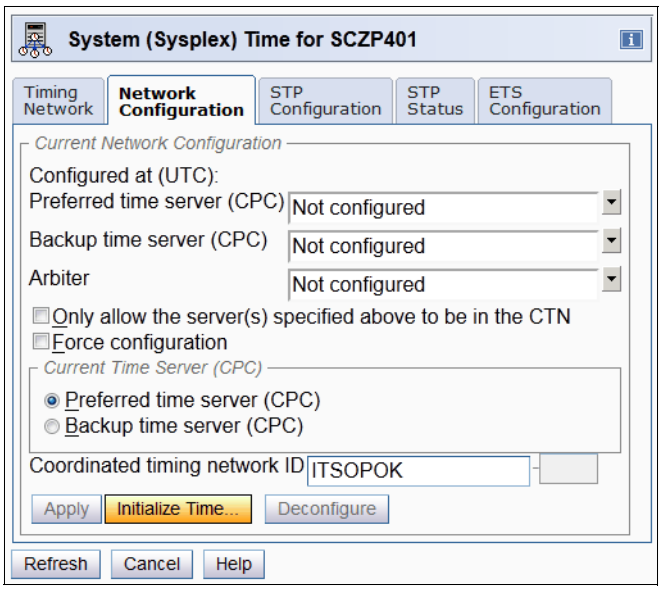

Figure 4-28 Network Configuration tab: SCZP401 after failure

Figure 4-29 shows the Network Configuration tab for the BTS SCZP301:

- The Preferred Time Server field of the Current Network Configuration section shows SCZP401 defined as the PTS, but there is no STP ID associated with it (as compared to the Backup Time Server field). This indicates that SCZP401 has no connectivity to SCZP301.
- The Current Time Server (CPC) section shows that the BTS SCZP301 is now the CTS.

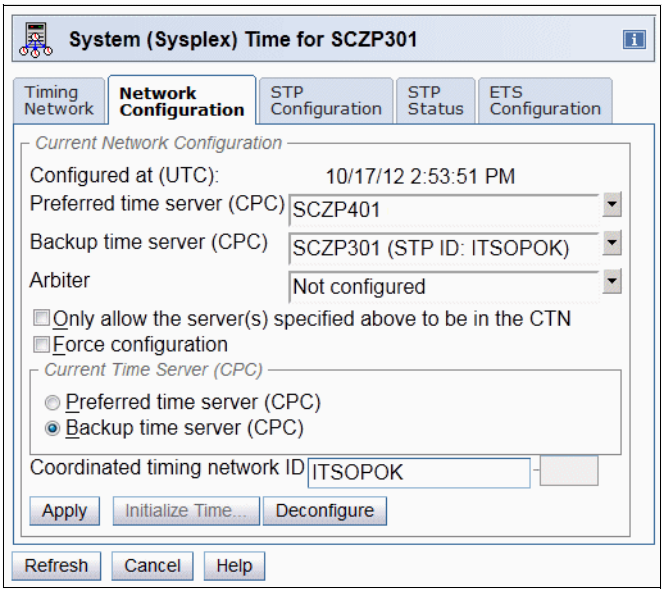

Figure 4-29 Network Configuration tab: SCZP301 after failure

Figure 4-30 shows the STP Status tab of SCZP301 after the last coupling link has failed. SCZP301 has transitioned to Stratum 1 and has no STP connectivity to any other servers. Coupled with the information from Figure 4-28 on page 138, you can determine that SCZP401 is operational, but there is a loss of communication between the two servers.

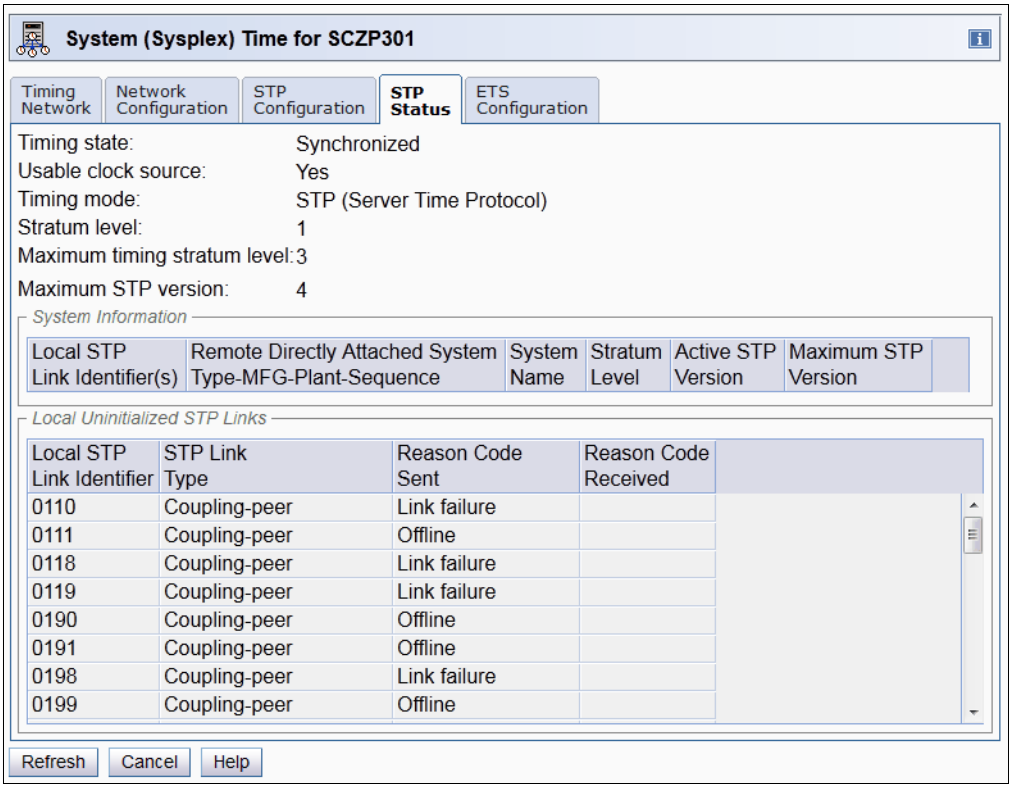

Figure 4-30 STP Status tab: SCZP301 after failure

4.7.3 User actions

The BTS SCZP301 takes over the role of CTS. To avoid an outage if the BTS SCZP301 fails:

1. Repair the failing links using installation recovery procedures.
   - SCZP401 will subsequently rejoin the CTN in its original role as PTS.
   - The BTS SCZP301 transitions back to Stratum 2.

   Figure 4-31 on page 140 shows the Network Configuration tab for SCZP401 after repair actions. The Preferred Time Server field of the Current Network Configuration section shows SCZP401 being defined as the PTS, and there is now an STP ID associated with it. This indicates that SCZP401 has connectivity to SCZP301.

   The Current Time Server (CPC) section shows that SCZP401 is now the PTS and CTS. As a consequence, z/OS (Version 1.11 or higher) issues message IEA395I on all z/OS images that are members of this CTN and use STPMODE=YES in the CLOCKxx member (SC75 in this case):

   ```
   IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE PREFERRED
   ```

2. Re-IPL z/OS system image SC74 on SCZP401.

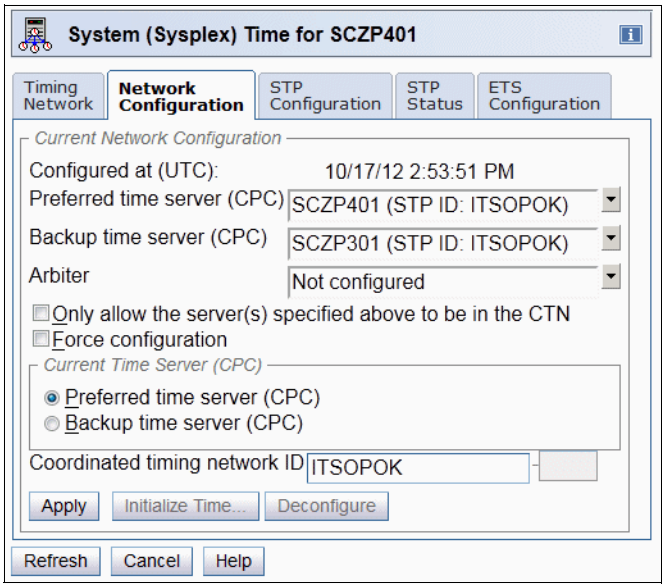

Figure 4-31 Network Configuration tab: SCZP401 after repair actions

5

Recovery in an STP-only CTN with BTS and Arbiter

In this chapter we discuss recovery scenarios for an STP-only CTN with an Arbiter-assigned configuration. The following possible failures are discussed:

- Current Time Server failure
- Backup Time Server failure
- Arbiter failure
- Multiple coupling links failure between the CTS and the BTS
- Last coupling link failure between the CTS and the Arbiter
- Last coupling link failure between the BTS and the Arbiter
- Two-site: Site 1 failure: CTS and Arbiter at Site 1, BTS at Site 2
- Two-site: Site 1 power outage - CTS and Arbiter at Site 1 with IBF
- Two-site: Site 2 failure: CTS and Arbiter at Site 1, BTS at Site 2
- Two-site: Site 2 power outage - CTS, BTS, and Arbiter with IBF

5.1 Overview of using an Arbiter

Configuring an Arbiter is strongly advised whenever there are three or more servers in the STP-only CTN. With an Arbiter available, Arbiter-assisted recovery becomes the primary mechanism for the BTS to determine whether it should take over the role of the CTS when communication with the PTS is lost. Refer to 1.1.6, “Arbiter-assisted recovery” on page 7 for more information.

Note: If the links between the PTS and BTS support the going away signal (GAS), then the going away signal is used by the BTS to take the CTS role without communicating with the Arbiter.

Note the following points when an Arbiter is configured:

- The Offline signal recovery process is not invoked.
- The going away signal is used without communicating with the Arbiter.
- Arbiter-assisted recovery is disabled when the PTS, or the BTS, or the Arbiter loses communication with the other two.
- The BTS invokes Console-assisted recovery only if it cannot communicate with the Arbiter.
- The PTS also invokes Console-assisted recovery when it loses communication with both the BTS and the Arbiter.

Including an Arbiter in the configuration provides a more robust recovery mechanism, because Arbiter-assisted recovery can handle recovery scenarios when an Offline signal might not be transmitted, such as during a channel subsystem failure.

The displays of the server names in the Network Configuration panel and the STP Status panel vary regarding the status of the HMC-to-SE communication. In most scenarios shown in this chapter it is assumed there is still communication between the HMC and the SE of the server that has failed. As a consequence, the SE names are shown in both the Network Configuration panel and in the STP Status panel. In cases where a server becomes completely lost (typically when a power drop occurs), the server will be shown with its Node ID because the HMC cannot communicate with the Support Element at that point in time.

5.2 Current Time Server failure

The recovery mechanism used by the BTS to take the CTS role when the PTS/CTS fails when an Arbiter is configured depends on whether any of the links between the PTS and BTS support the going away signal. If a going away signal is received by the BTS, then the BTS will take the CTS role without communicating with the Arbiter. Otherwise, the BTS will communicate with the Arbiter to determine whether it can take the CTS role.

Figure 5-1 shows a CTN configuration with a BTS and an Arbiter, where server SCZP301 is the PTS/CTS. If CTS fails, then if a going away signal is received, the BTS will assume the CTS role. Otherwise, the BTS communicates with the Arbiter to determine whether the Arbiter has also lost communication with the CTS. If the Arbiter has also lost communication with the CTS, the BTS will assume the CTS role. Because only one CTS can exist in an STP-only CTN, the PTS must surrender its role as the CTS. As a result, the BTS takes over as CTS and transitions to Stratum 1.

Note: There is no indication regarding whether the BTS took the CTS role when the PTS failed through the going away signal or by communicating with the Arbiter.

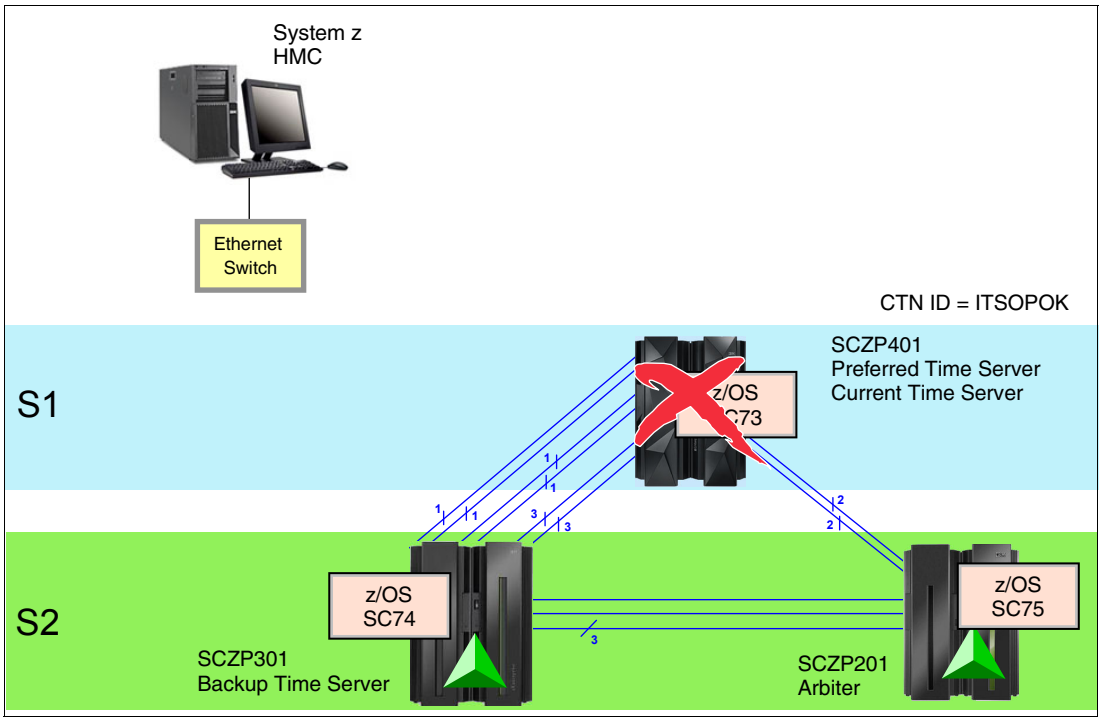

Figure 5-1 STP-only CTN with BTS and Arbiter: CTS failure

5.2.1 Problem awareness

z/OS messages are issued on z/OS images running on the BTS and Arbiter, indicating that the z/OS image on SCZ401 has failed. An active SFM system isolation policy automatically starts system partitioning. This is highlighted by an IXC101I message that is sent only a few seconds after the failure detect interval expires. Partitioning triggers the cleanup of resources for the failing system. If system isolation fails, then SFM will issue WTOR message IXC102A after the XCF CLEANUP time has elapsed:

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC73 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE

IXC102A XCF IS WAITING FOR SYSTEM SC73 DEACTIVATION. REPLY DOWN
WHEN MVS ON SC73 HAS BEEN SYSTEM RESET
```

Note: Before replying `DOWN` to IXC102A or IXC402D, you must perform a hardware SYSTEM RESET on the z/OS system being removed. This is necessary to ensure that this z/OS system no longer performs I/O operations and releases any outstanding I/O reserves. A SYSTEM RESET ensures that other systems continue to have access to the data sets on the shared devices.

Because the Current Time Server was switched from the Preferred Time Server to the Backup Time Server when the PTS failed, z/OS (Version 1.11 or higher) issues message IEA395I on all z/OS images that are members of this CTN and use STPMODE=YES in the CLOCKxx member:

```
IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE BACKUP
```

5.2.2 Problem determination

Figure 5-2 shows the result of running the z/OS **DISPLAY ETR** command after the CTS SCZP401 has failed.

Note: The output from SC74 indicates that SCZP301 maintains its role as BTS. However, it has also transitioned to Stratum 1 and now assumes the role of Current Time Server.

SCZP201 maintains its role as Stratum 2 Arbiter. Both servers do not have connectivity to SCZP401.

```
RO SC73,D ETR
IEA386I 10.33.26 TIMING STATUS 055
SYNCHRONIZATION MODE = STP
THIS SERVER IS A STRATUM 1
CTN ID = ITSOPOK
THE STRATUM 1 NODE ID = 002817.M32.IBM.02.0000000B3BD5
THIS IS THE BACKUP TIME SERVER
THIS SERVER HAS NO LINK TO THE PREFERRED TIME SERVER

RO SC75,D ETR
IEA386I 10.33.26 TIMING STATUS 077
SYNCHRONIZATION MODE = STP
THIS SERVER IS A STRATUM 2
CTN ID = ITSOPOK
THE STRATUM 1 NODE ID = 002817.M32.IBM.02.0000000B3BD5
THIS IS THE ARBITER SERVER
THIS SERVER HAS NO LINK TO THE PREFERRED TIME SERVER
NUMBER OF USABLE TIMING LINKS = 5
THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS
```

Figure 5-2 DISPLAY ETR: after failure

Figure 5-3 shows the Network Configuration tab for the BTS SCZP301. The following items are displayed:

- The Preferred Time Server field of the Current Network Configuration section shows server SCZP401 being defined as PTS, but there is no STP ID associated with it, as compared to the Backup Time Server and Arbiter fields. This indicates that SCZP401 has no connectivity to SCZP301.
- The Current Time Server (CPC) section shows that the BTS (SCZP301) is now the CTS.

The Current Time Server (CPC) section includes the insert.

Note that automated network recovery is temporarily disabled.

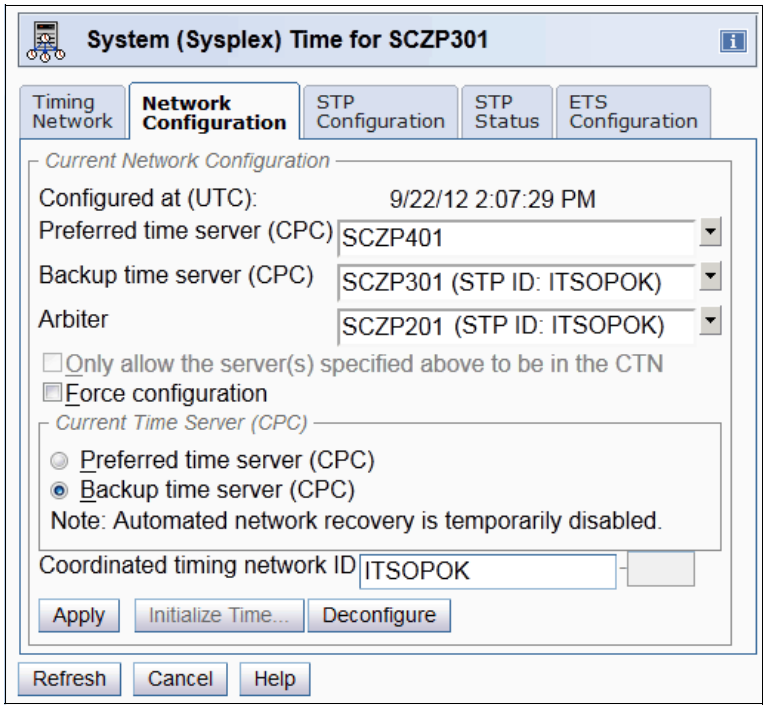

Figure 5-3 Network Configuration tab: SCZP401 after failure

Figure 5-4 shows the STP Status tab for the BTS/CTS SCZP301. SCZP301 has transitioned to Stratum 1 and only has connectivity to the Arbiter SCZP201.

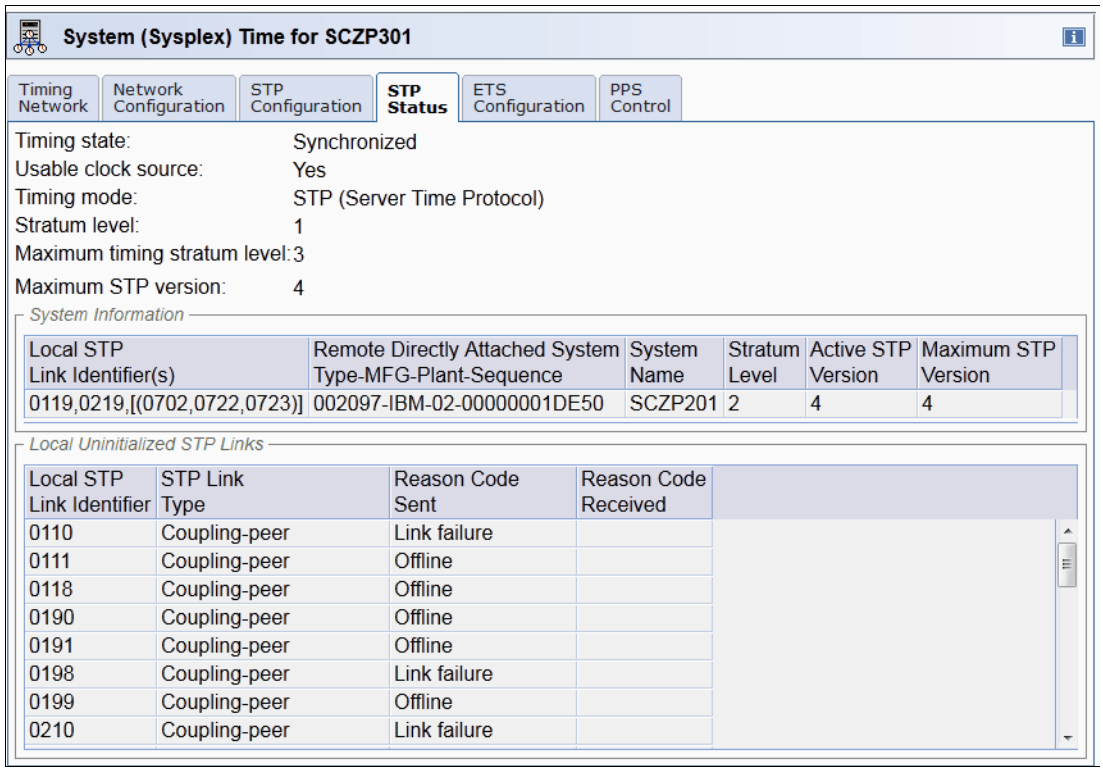

Figure 5-4 STP Status tab: SCZP401 after PTS failure

5.2.3 User actions

When the problem is fixed and server SCZP401 becomes operational, it rejoins the CTN.

Because the STP-only CTN was not reconfigured by reassigning roles during the failure, the PTS automatically takes back its original role as CTS. As a consequence, z/OS (Version 1.11 or higher) issues message IEA395I on all z/OS images that are members of this CTN and that use STPMODE=YES in the CLOCKxx member:

```
IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE PREFERRED
```

Examine the STP Status tabs from various servers to verify that the CTN has successfully returned to its original state.

Figure 5-5 on page 146 shows the Network Configuration tab for the PTS after it has become operational. Several items are displayed:

- The Preferred Time Server field of the Current Network Configuration section shows SCZP401 being defined as the PTS. SCZP301 in the Backup Time Server field and SCZP201 in the Arbiter field have an associated STP ID. This indicates that SCZP401 has connectivity to SCZP301 and SCZP201.
- The Current Time Server (CPC) section shows that the PTS (SCZP401) is now the CTS.

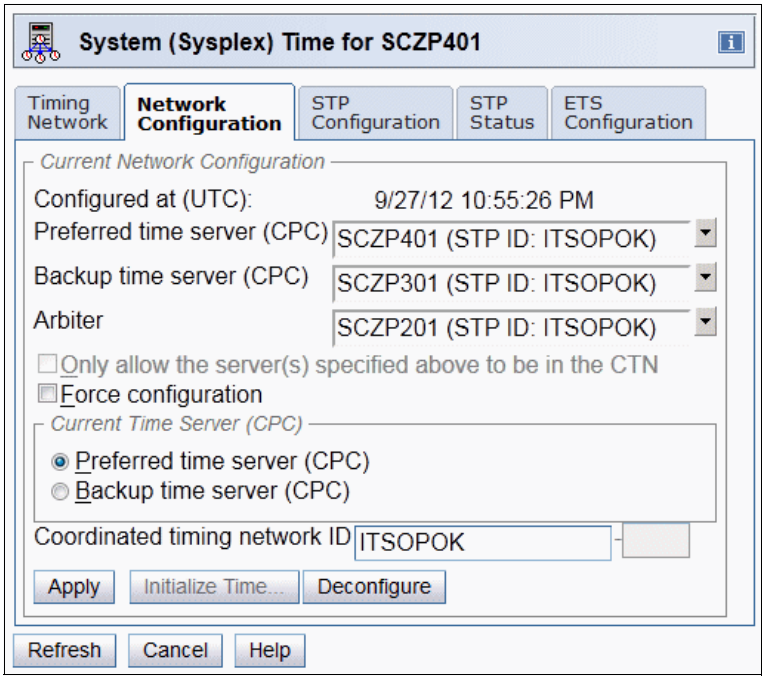

Figure 5-5 Network Configuration tab: SCZP401after recovery and user action

In Figure 5-6, the CTS SCZP401 returns to its original Stratum 1 level and has connectivity to both the Arbiter SCZP201 and BTS SCZP301 servers. With the PTS back in the CTN, SCZP301 transitions back to Stratum 2.

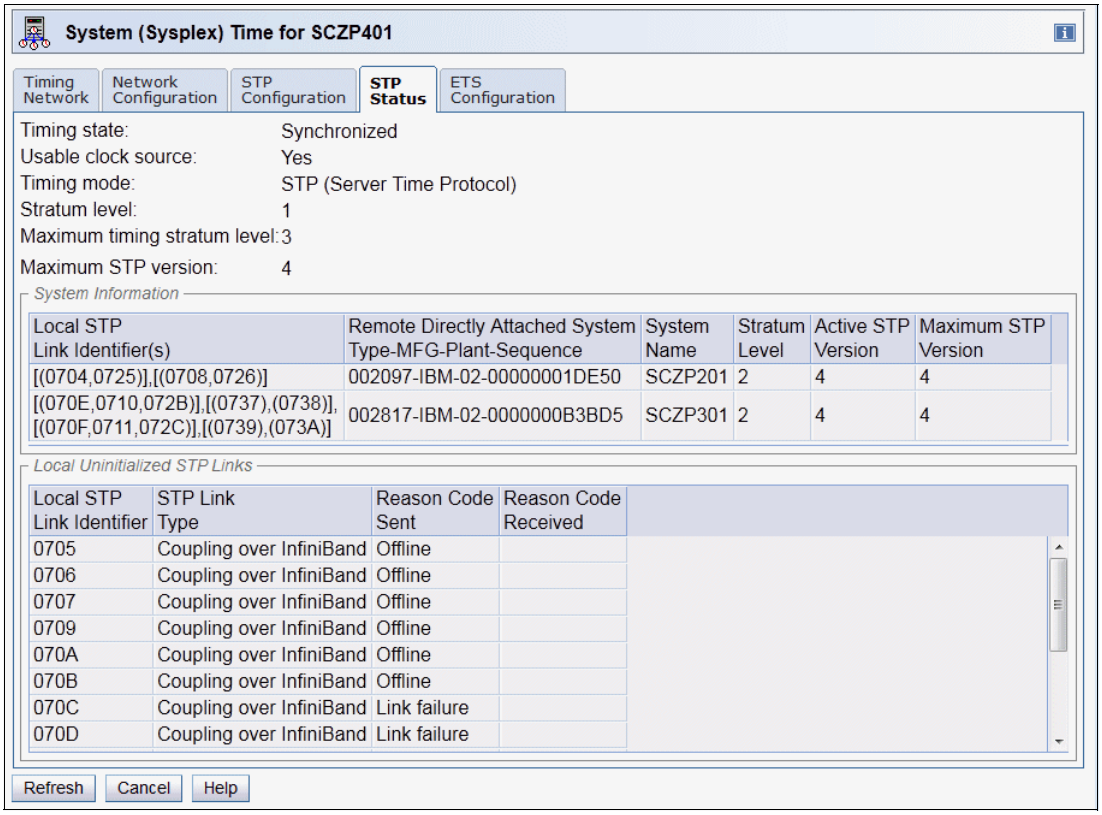

Figure 5-6 STP Status tab: Server SCZP401 after recovery and user action

Re-IPL z/OS images on SCZP401 using the installation recovery procedures.

5.3 Backup Time Server failure

In this scenario, the BTS fails. The CTN continues to operate without a BTS.

The CTN continues without a server performing the Backup Time Server role and is subject to a single point of failure until BTS is operational or another server is configured as BTS, as shown in Figure 5-7.

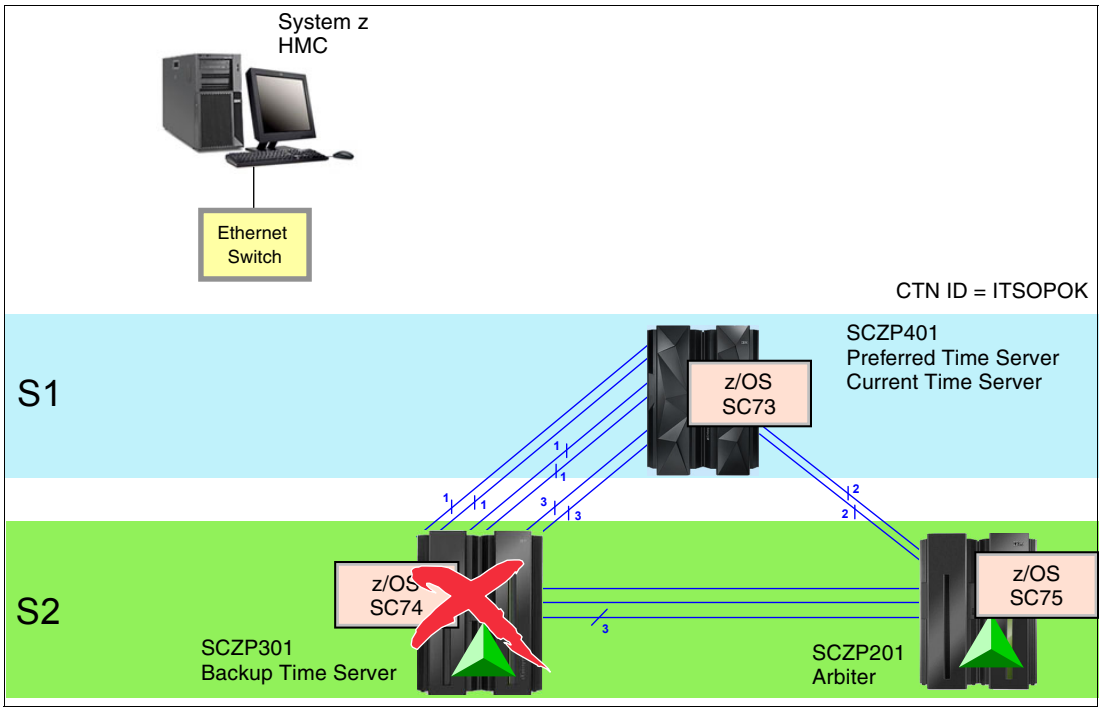

Figure 5-7 STP-only CTN with BTS and Arbiter: BTS failure

5.3.1 Problem awareness

z/OS messages similar to those shown in Figure 5-8 are issued, indicating that the BTS SCZP301 has failed.

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC74 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
```

Figure 5-8 z/OS message IXC101I: after failure

An active SFM system isolation policy automatically starts system partitioning. In Figure 5-8, this is highlighted by an IXC101I message only a few seconds after failure detection interval expires. Partitioning triggers the cleanup of resources for the failing system.

As shown in Figure 5-9, if system isolation fails, then SFM will issue WTOR message IXC102A after the XCF CLEANUP time has elapsed.

```
IXC102A XCF IS WAITING FOR SYSTEM SC74 DEACTIVATION. REPLY DOWN WHEN MVS ON
SC74 HAS BEEN SYSTEM RESET
```

Figure 5-9 z/OS message IXC102A: after failure

Note: Before replying `DOWN` to IXC102A or IXC402D, you must perform a hardware SYSTEM RESET on the z/OS system being removed. This is necessary to ensure that this z/OS system releases any outstanding I/O reserves. A SYSTEM RESET ensures that other systems continue to have access to the data sets on the shared devices.

5.3.2 Problem determination

Figure 5-10 shows the Network Configuration tab for the CTS SCZP401. The Backup Time Server field of the Current Network Configuration section shows SCZP301 being defined as the BTS, but there is no STP ID associated with it (as compared to the Preferred Time Server and Arbiter fields). This indicates that SCZP301 has no connectivity to SCZP401.

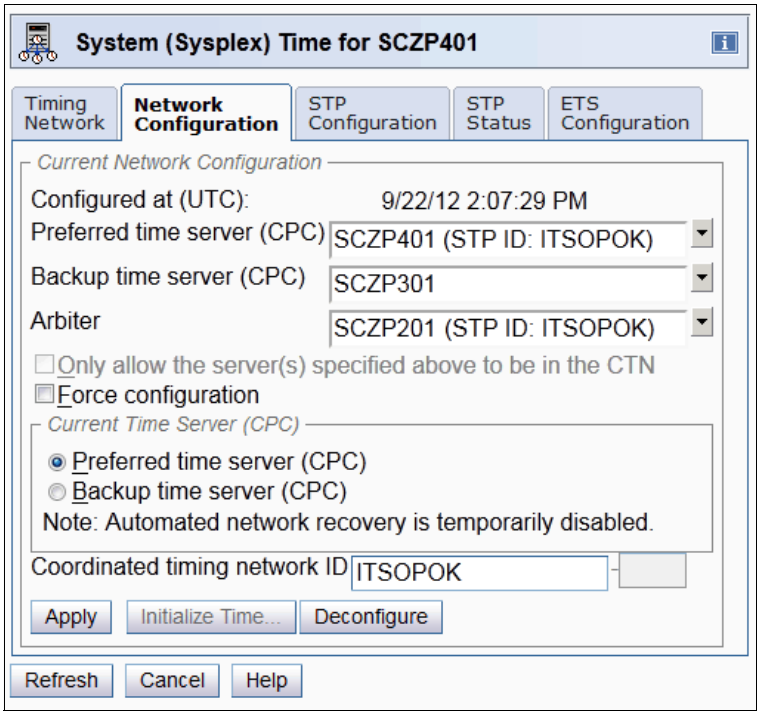

Figure 5-10 Network Configuration tab: PTS SCZP401 after failure

Figure 5-11 shows the STP Status tab for the CTS SCZP401, and that SCZP401 only has connectivity to the Arbiter SCZP201.

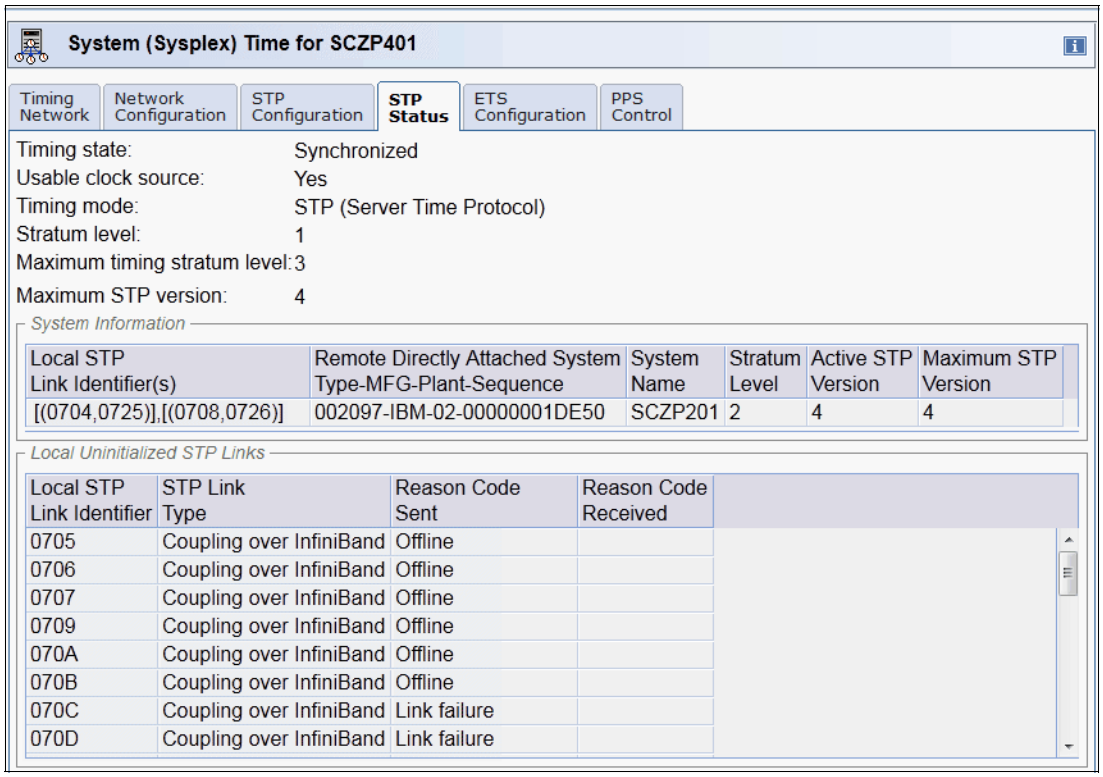

Figure 5-11 STP Status tab: PTS SCZP301 after failure

5.3.3 User actions

Reconfigure the CTN as soon as possible to ensure that at least a Current Time Server and Backup Time Server are available. In a configuration where another server is available, assigned it the BTS role.

Important: Given the configuration shown in Figure 5-7 on page 148, a BTS failure introduces a single point of failure into the environment because a subsequent failure of the PTS will bring down the entire CTN. To avoid this, give a high priority to restoring the BTS to an operational state.

In Figure 5-12, the z/OS DISPLAY XCF command shows that SC74 has rejoined the sysplex.

```
D XCF,SYSPLEX
IXC336I  13.33.11  DISPLAY XCF 441
SYSPLEX PLEX75
SYSTEM   TYPE SERIAL LPAR STATUS TIME         SYSTEM STATUS
SC73     2827 B8D7   01   10/01/2012 13:33:11 ACTIVE         TM=STP
SC74     2817 3BD5   05   10/01/2012 13:33:08 ACTIVE         TM=STP
SC75     2097 DE50   03   10/01/2012 13:31:11 ACTIVE         TM=STP
```

Figure 5-12 DISPLAY XCF: after recovery and user action

Figure 5-13 shows the Network Configuration tab for the CTS SCZP401 after the BTS SCZP301 as successfully recovered. Note that with SCZP301 operational, the Backup Time Server field of the Current Network Configuration section shows SCZP301 being defined as the BTS, and there is now an STP ID associated with it. This indicates that SCZP301 is a member of the CTN.

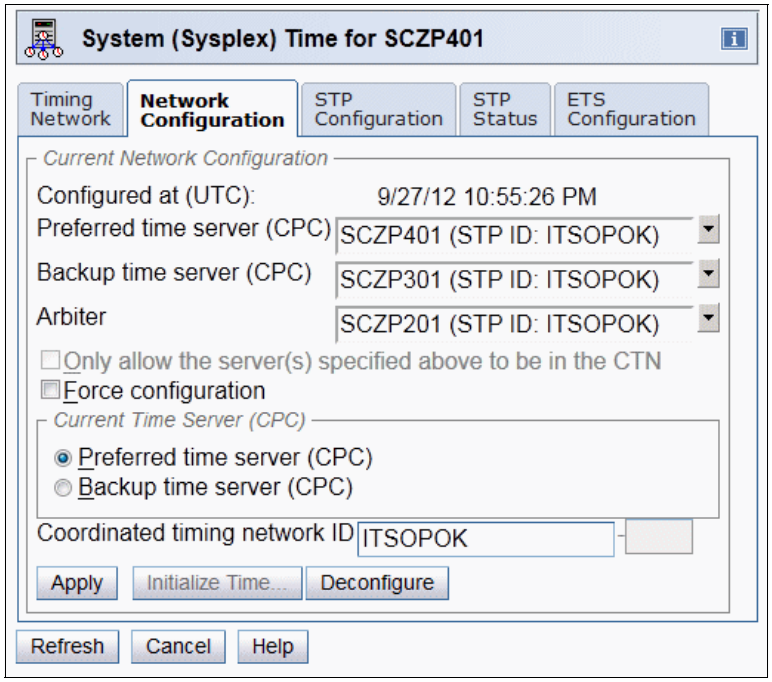

Figure 5-13 Network Configuration tab: SCZP401 after recovery and user action

Figure 5-14 shows the STP Status tab for the CTS SCZP401. After recovery, SCZP401 has connectivity to the BTS SCZP301 and the Arbiter SCZP201.

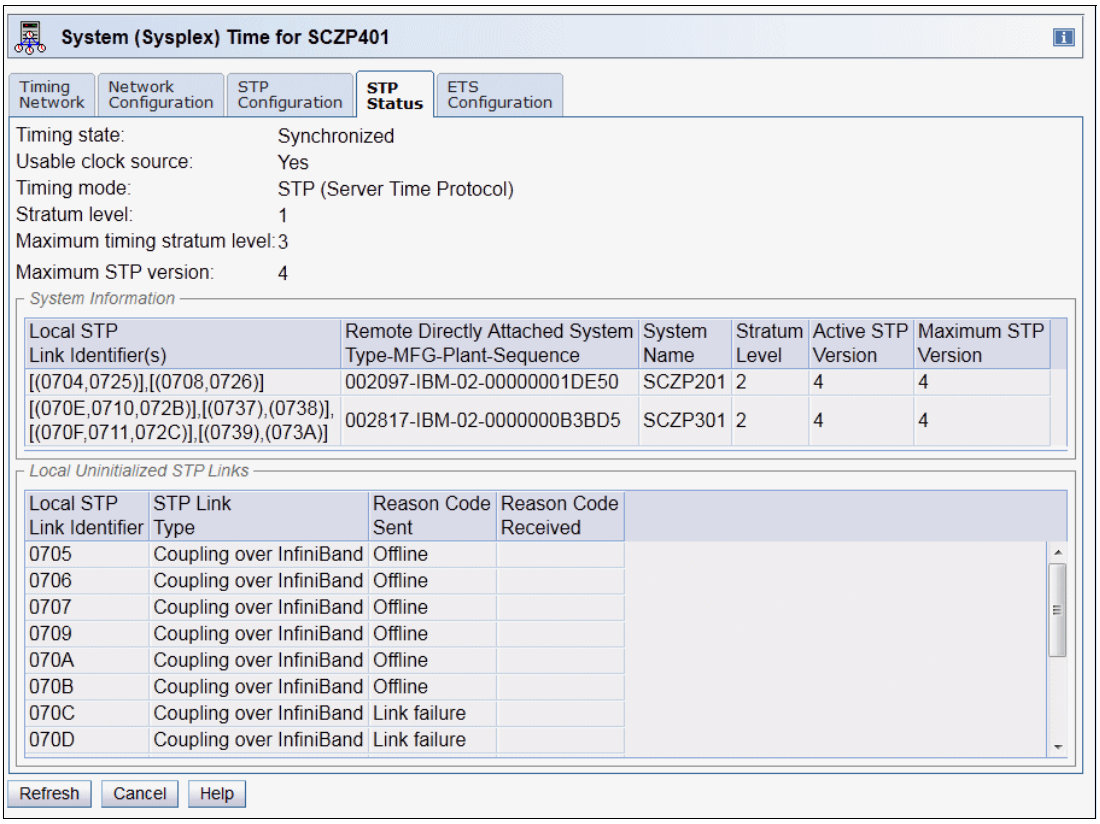

Figure 5-14 STP Status tab: SCZP401 after recovery and user action

5.4 Arbiter failure

Figure 5-15 shows a configuration where the Arbiter, SCZP201, has failed.

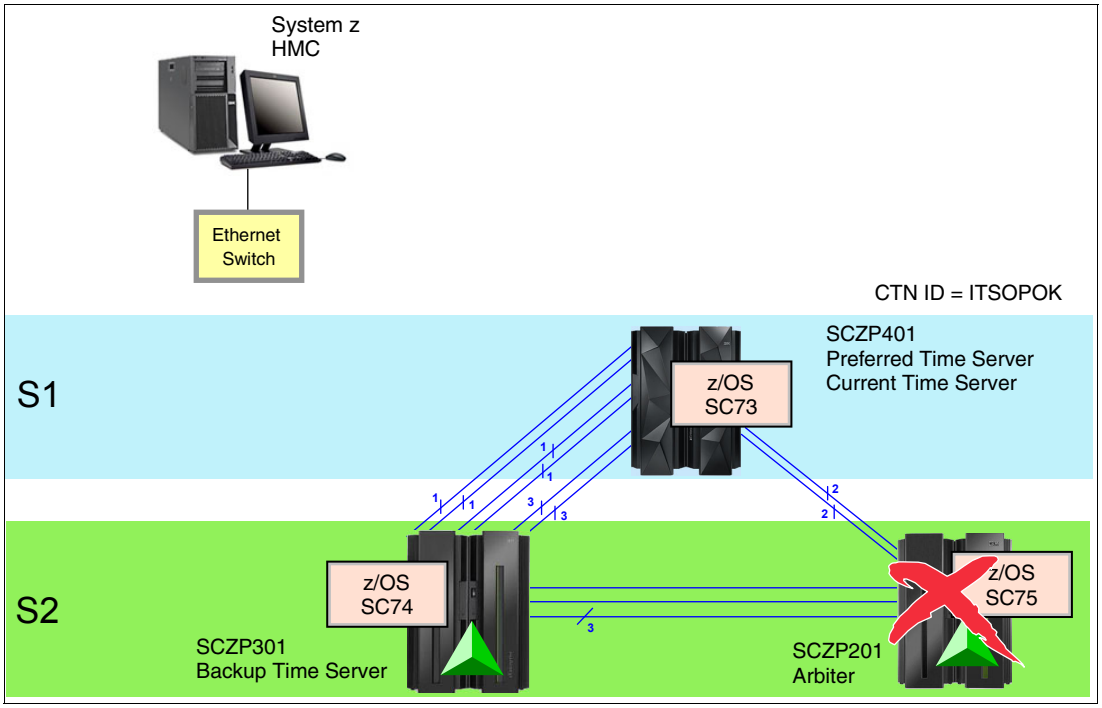

Figure 5-15 STP-only CTN: Arbiter failure

If the Arbiter fails, no further recovery actions are taken. The CTN will continue to operate with only a CTS and a BTS. However, the Arbiter is still known to the configuration, even though it is no longer available, thereby causing subsequent recovery actions to ignore the Offline signal rules.

5.4.1 Problem awareness

z/OS messages will be issued indicating that the Arbiter SCZP201 has failed. An active SFM system isolation policy automatically starts system partitioning. This is highlighted by an IXC101I message that is issued only a few seconds after detection of the failure. Partitioning triggers the cleanup of resources for the failing system:

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC75 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE

IXC102A XCF IS WAITING FOR SYSTEM SC75 DEACTIVATION. REPLY DOWN WHEN MVS ON SC75
HAS BEEN SYSTEM RESET
```

If system isolation fails, then SFM issues WTOR message IXC102A after the XCF CLEANUP time has elapsed.

Note: Before replying `DOWN` to IXC102A or IXC402D, you must perform a hardware SYSTEM RESET on the z/OS system being removed. This is necessary to ensure that this z/OS system releases any outstanding I/O reserves. A SYSTEM RESET ensures that other systems continue to have access to the data sets on the shared devices.

5.4.2 Problem determination

In Figure 5-16, the z/OS **DISPLAY ETR** command shows that neither the CTS SCZP401 nor the BTS SCZP301 have connectivity to the Arbiter.

```
RO SC73,D ETR
IEA386I 13.46.46 TIMING STATUS 613
SYNCHRONIZATION MODE = STP
THIS SERVER IS A STRATUM 1
CTN ID = ITSOPOK
THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
THIS IS THE PREFERRED TIME SERVER
THIS SERVER HAS NO LINK TO THE ARBITER SERVER

RO SC74,D ETR
IEA386I 13.46.46 TIMING STATUS 190
SYNCHRONIZATION MODE = STP
THIS SERVER IS A STRATUM 2
CTN ID = ITSOPOK
THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
THIS IS THE BACKUP TIME SERVER
THIS SERVER HAS NO LINK TO THE ARBITER SERVER
NUMBER OF USABLE TIMING LINKS = 10
```

Figure 5-16 DISPLAY ETR: after failure

Figure 5-17 shows the Network Configuration tab for the CTS SCZP401. The Arbiter field of the Current Network Configuration section shows SCZP201 being defined as the Arbiter, but there is no STP ID associated with it (as compared to the Preferred Time Server and Backup Time Server fields). This indicates that the server Support Element is still visible to the HMC, but SCZP401 has no connectivity to SCZP201.

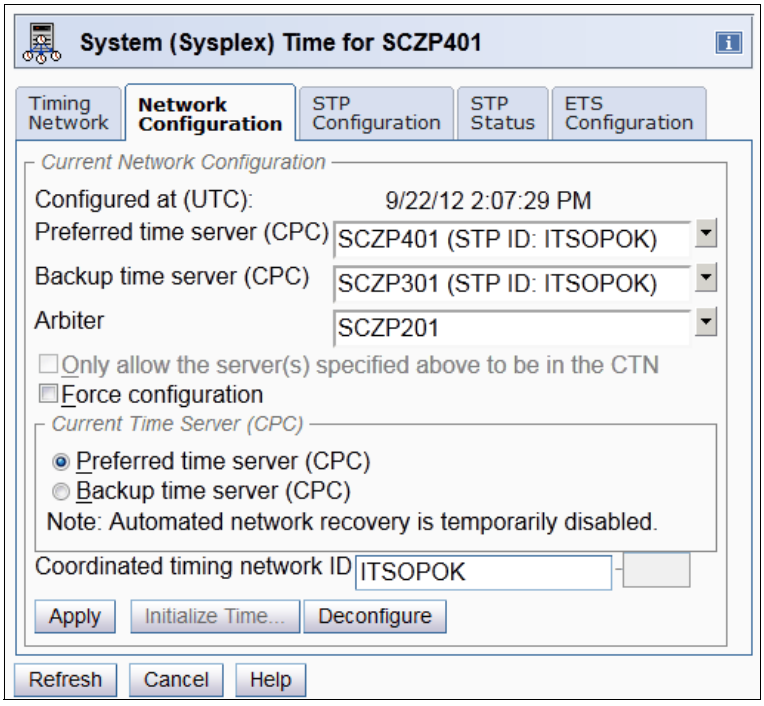

Figure 5-17 Network Configuration tab: SCZP201 after failure

Figure 5-18 shows the STP Status tab for the CTS SCZP401, and that SCZP401 only has connectivity to the BTS SCZP301.

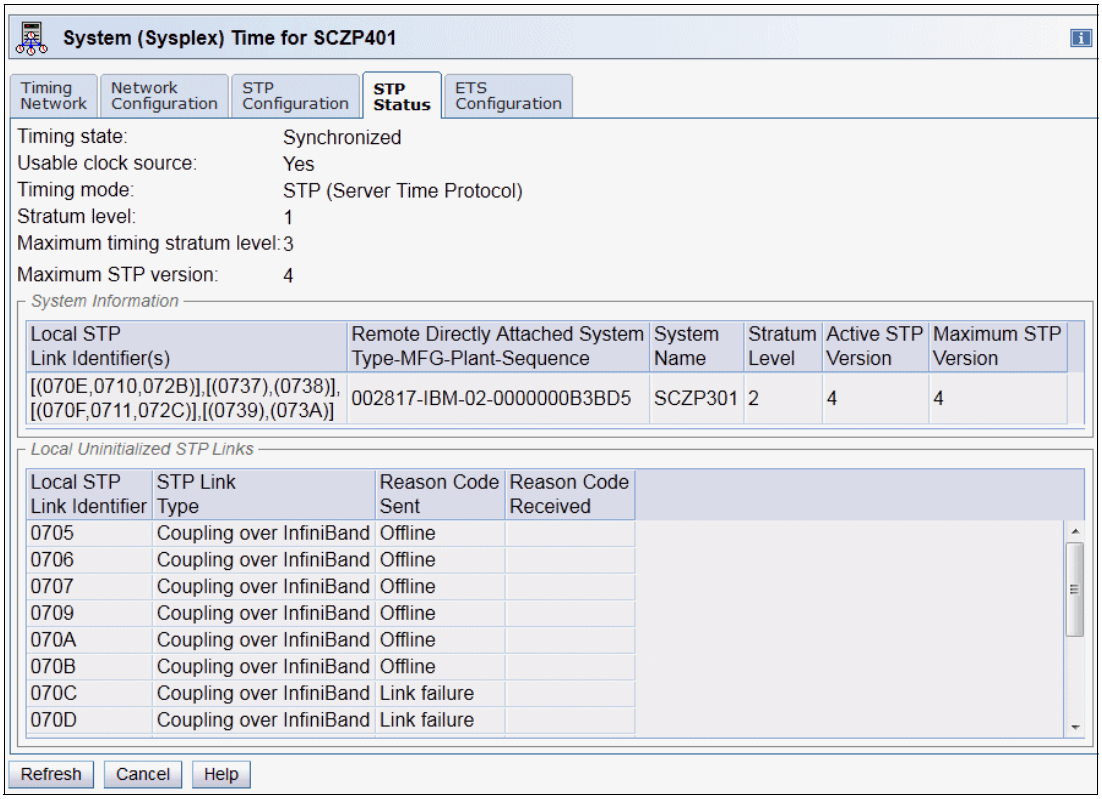

Figure 5-18 STP Status tab: SCZP401 after SWCZP201 failure

5.4.3 User actions

The CTN continues to operate with only a CTS and a BTS. If another server is available, it is advisable to configure it to take over the role of Arbiter.

Figure 5-19 shows the output from the z/OS DISPLAY XCF command. Even though the Arbiter is nonoperational, the CTN continues to operate.

```
D XCF,SYSPLEX
IXC336I  14.35.15 DISPLAY XCF 441
SYSPLEX PLEX75
SYSTEM   TYPE SERIAL LPAR STATUS TIME         SYSTEM STATUS
SC73     2827 B8D7   01   10/01/2012 14:35:15 ACTIVE         TM=STP
SC74     2817 3BD5   05   10/01/2012 14:35:02 ACTIVE         TM=STP
```

Figure 5-19 DISPLAY XCF: Arbiter down

If the Arbiter is expected to be nonoperational for an extended period of time and no other server can be configured as a new Arbiter, we strongly advise that the Arbiter assignment be removed so that the BTS can use the OLS rules.

Figure 5-20 shows the Network Configuration tab for the CTS SCZP401 after the Arbiter SCZP201 was successfully recovered. Note that with SCZP201operational, the Arbiter field of the Current Network Configuration section shows SCZP201 defined as the Arbiter; there is now an STP ID associated with it. This indicates that SCZP201 is a member of the CTN.

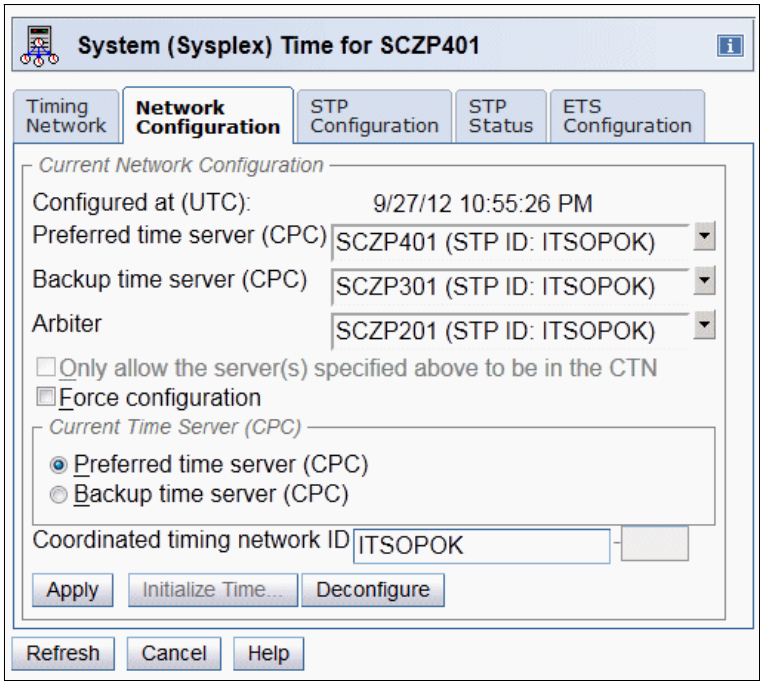

Figure 5-20 Network Configuration tab: SCZP401 after recovery and user action

Figure 5-21 shows the STP Status tab for the CTS SCZP401. After recovery, SCZP401 has connectivity to the BTS SCZP301 and the Arbiter SCZP201.

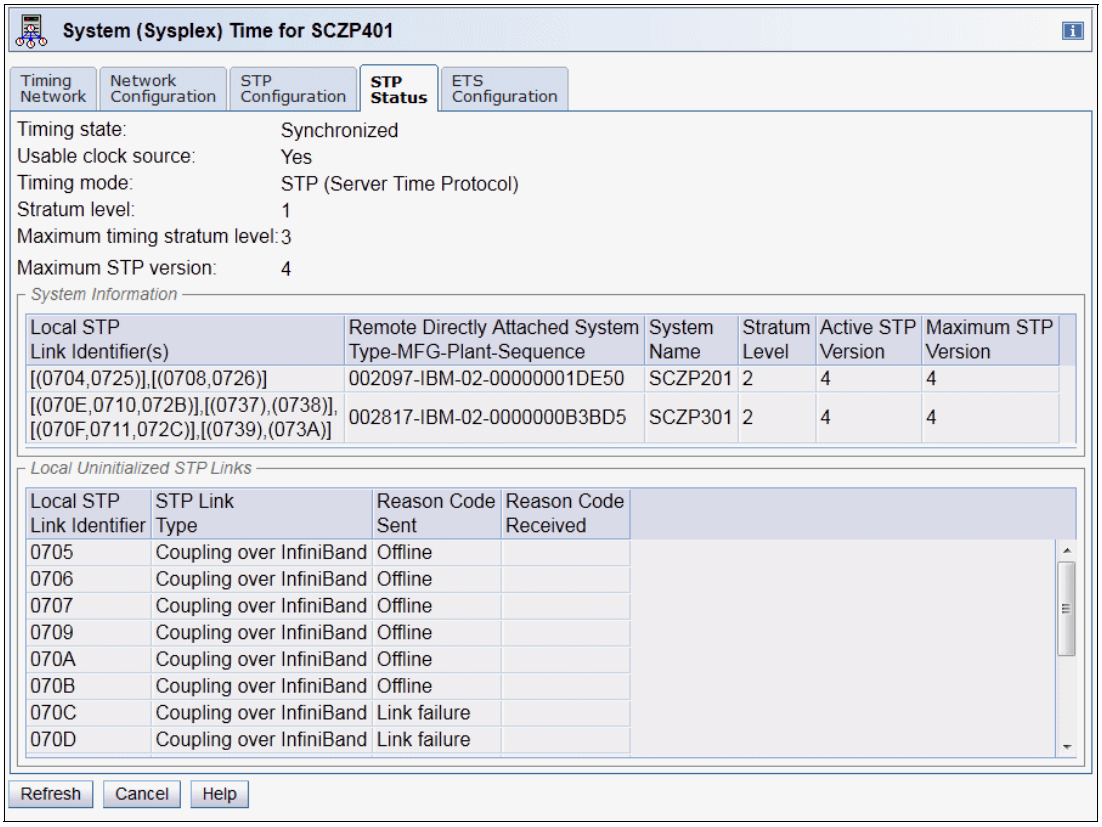

Figure 5-21 STP Status Configuration tab: SCZP401 after recovery and user action

After performing a repair action on SCZP201 and re-IPLing system image SC73, the **DISPLAY XCF** command in Figure 5-22 shows that the sysplex is back to its original state.

```
D XCF,SYSPLEX
IXC336I  16.45.22 DISPLAY XCF 441
SYSPLEX PLEX75
SYSTEM   TYPE SERIAL LPAR STATUS TIME         SYSTEM STATUS
SC73     2827 B8D7   01   10/01/2012 16:45:22 ACTIVE         TM=STP
SC74     2817 3BD5   05   10/01/2012 16:45:19 ACTIVE         TM=STP
SC75     2097 DE50   03   10/01/2012 16:45:22 ACTIVE         TM=STP
```

Figure 5-22 DISPLAY XCF: after Arbiter recovery

5.5 Multiple coupling links failure between the CTS and the BTS

When there are multiple coupling links between any two servers, only one coupling link is used at a time to exchange STP messages. If this particular link fails, one of the remaining links will be used with no impact on the CTN. There will be no indication of an STP timing problem because one of the remaining links will be automatically selected. In this case, STP recovery is transparent to operations.

In the following scenario, only the failure of the last coupling link between any two servers is considered. Figure 5-23 shows a scenario where the last coupling link between the PTS and the BTS fails.

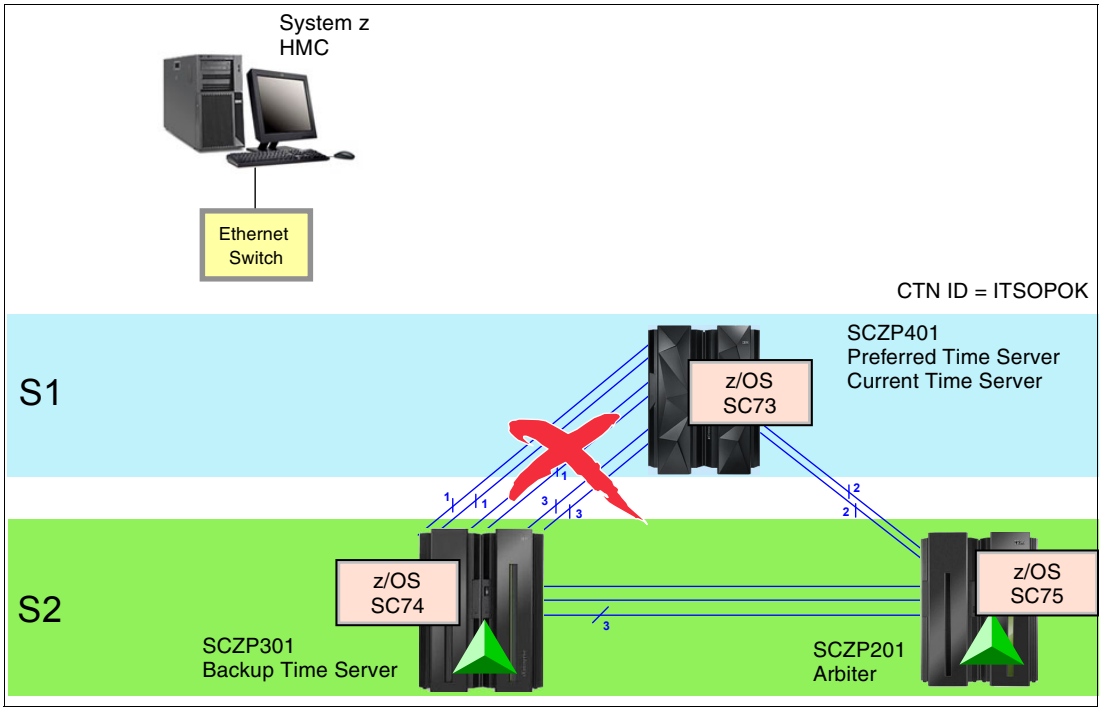

Figure 5-23 STP-only CTN: multiple link failure between CTS and BTS

BTS will cease to receive STP timing messages from the CTS and will contact the Arbiter to determine whether the Arbiter can communicate with the CTS. Because the Arbiter is still in communication with the CTS, BTS will not take over the CTS role, and therefore will transition to Stratum 3 to remain in the STP-only CTN.

Even though the BTS no longer has direct connectivity with the CTS, it is still able to function in the BTS role. The Arbiter takes note that the BTS no longer has connectivity to the CTS and, if it subsequently loses contact with the CTS, the Arbiter will inform the BTS accordingly, causing the BTS to proceed with taking over as the CTS.

In Figure 5-23, when the coupling link between SCZP401 and SCZP301 fails, there is no impact from an STP timing point of view because SCZP301 is still able to use the Arbiter SCZP201 as a valid time source.

The key highlights of this scenario are:

- The BTS no longer receives messages from CTS, and contacts the Arbiter to determine whether the Arbiter can communicate with CTS.
- The Arbiter still has communication, and the BTS will not take over the CTS role.
- The BTS transitions from Stratum 2 to Stratum 3.

5.5.1 Problem awareness

Message IXL158I is issued if the link failure also affects connectivity to a Coupling Facility partition in a sysplex:

```
IXL158I PATH 54 IS NOW NOT-OPERATIONAL TO CUID: FFFA 778
```

```
COUPLING FACILITY 002827.IBM.02.00000000B8D7
PARTITION: 0F CPCID: 00
```

This message indicates that the CHPID that was in use by the system to communicate with the identified Coupling Facility has become nonoperational. The cause is either a path failure or loss of the Coupling Facility. The system discontinues using the nonoperational path.

5.5.2 Problem determination

Figure 5-24 displays the Network Configuration tab seen from SCZP401, which shows that SCZP401 has connectivity to the Arbiter SCZP201. The BTS is still assigned to SCZP301, but there is no STP ID in the BTS entry, which is an indication that there is no STP connectivity between SCZP401 and SCZP301.

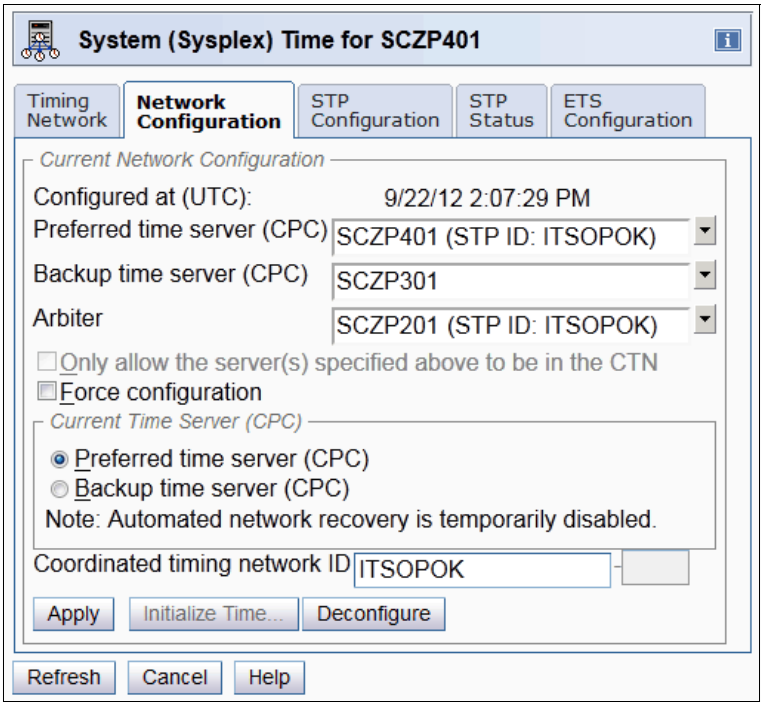

Figure 5-24 STP Configuration tab: SCZP401 after failure

Figure 5-25 displays the STP Status tab of the BTS. It shows that SCZP301 has connectivity to the Arbiter SCZP201, but has no connectivity to SCZP401. SCZP301 has transitioned from Stratum 2 to Stratum 3. However, its timing state is still *synchronized*, because it has communication with the Arbiter SCZP201.

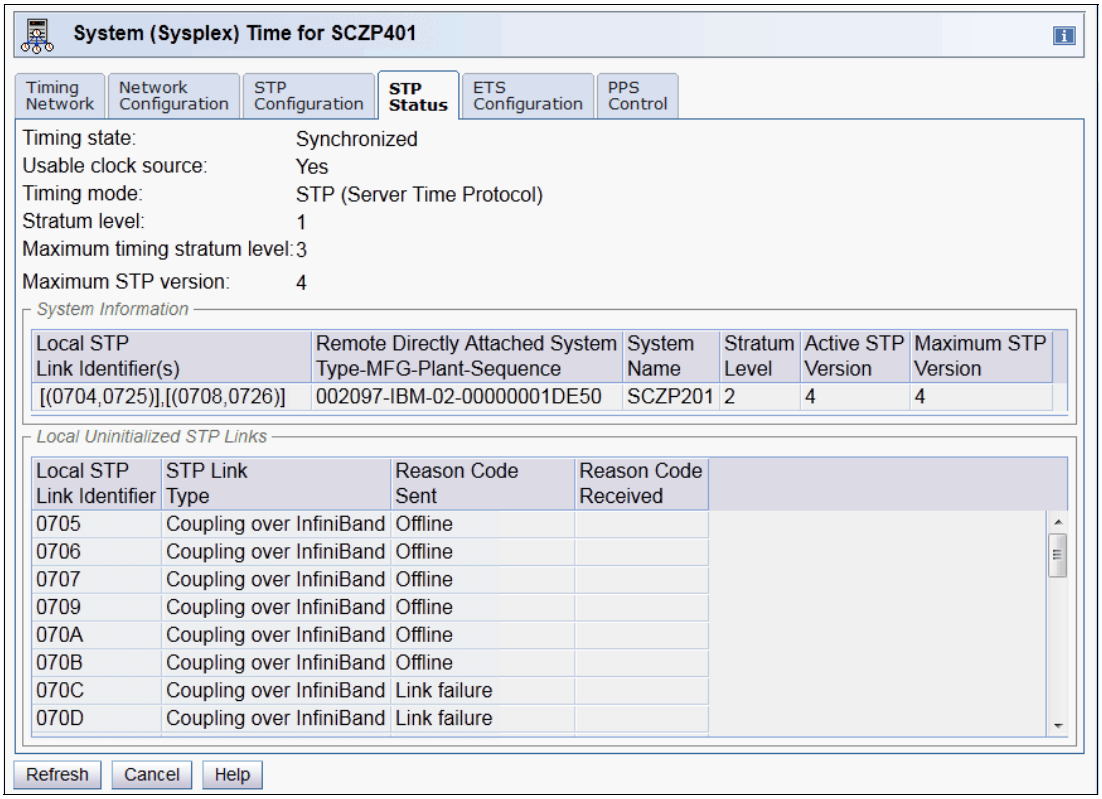

Figure 5-25 STP Status tab: BTS SCZP301 after failure

The same information is confirmed by the responses to the **DISPLAY ETR** command from the z/OS system images on each server, as shown in Figure 5-26 on page 161.

- On SC73, hosted on the PTS and CTS, the following messages are seen:
  - `THIS SERVER HAS NO LINK TO THE BACKUP TIME SERVER`
- On SC74, hosted on the BTS, the following messages are seen:
  - `THIS SERVER HAS NO LINK TO THE PREFERRED TIME SERVER`
  - `THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS`

```
SC73 RESPONSES -----------------------------------------------------
 IEA386I 15.47.23 TIMING STATUS 188
 SYNCHRONIZATION MODE = STP
   THIS SERVER IS A STRATUM 1
   CTN ID = ITSOPOK
   THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
   THIS IS THE PREFERRED TIME SERVER
   THIS SERVER HAS NO LINK TO THE BACKUP TIME SERVER
SC74     RESPONSES -----------------------------------------------------
 IEA386I 15.47.23 TIMING STATUS 134
 SYNCHRONIZATION MODE = STP
   THIS SERVER IS A STRATUM 3
   CTN ID = ITSOPOK
   THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
   THIS IS THE BACKUP TIME SERVER
   THIS SERVER HAS NO LINK TO THE PREFERRED TIME SERVER
   NUMBER OF USABLE TIMING LINKS = 5
   THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS
SC75     RESPONSES -----------------------------------------------------
 IEA386I 15.47.23 TIMING STATUS 791
 SYNCHRONIZATION MODE = STP
   THIS SERVER IS A STRATUM 2
   CTN ID = ITSOPOK
   THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
   THIS IS THE ARBITER SERVER
   NUMBER OF USABLE TIMING LINKS = 4
   THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS
```

Figure 5-26 DISPLAY ETR command output

5.5.3 User actions

No user action is required on the CTN. Identify the coupling link that has failed and take the appropriate action to repair it.

When connectivity is subsequently restored between the CTS and the BTS, the BTS will transition back to Stratum 2 and the Arbiter will be informed of its availability. The BTS SCZP301 will return to Stratum 2, and will have connectivity to both the CTS SCZP401 and the Arbiter SCZP201.

When link connectivity has been restored, it can be verified by examining the HMC. Figure 5-27 shows the Network Configuration tab seen from the CTS SCZP401. The servers have maintained their original roles, and the CTS is connected to both the BTS and the Arbiter.

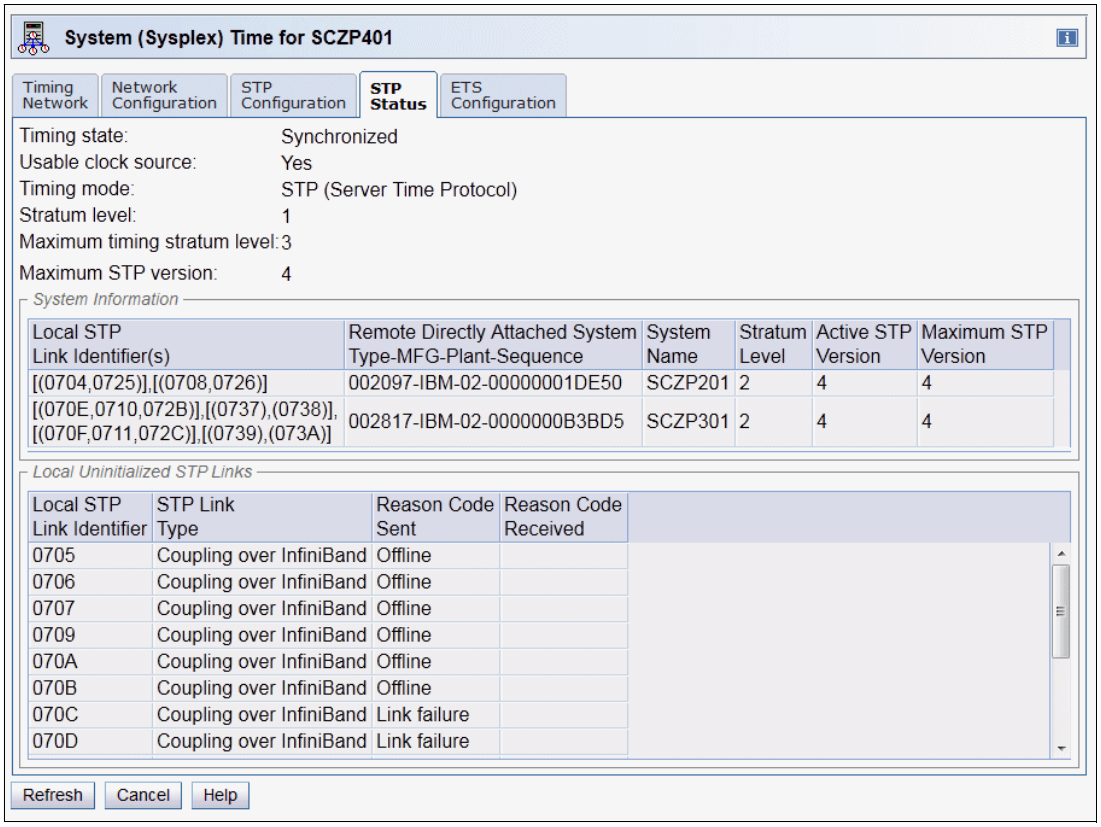

Figure 5-27 Network Configuration tab: SCZP401 after recovery and user action

5.6 Last coupling link failure between the CTS and the Arbiter

It is possible to lose all connectivity between the CTS and the Arbiter, as shown in Figure 5-28 on page 163.

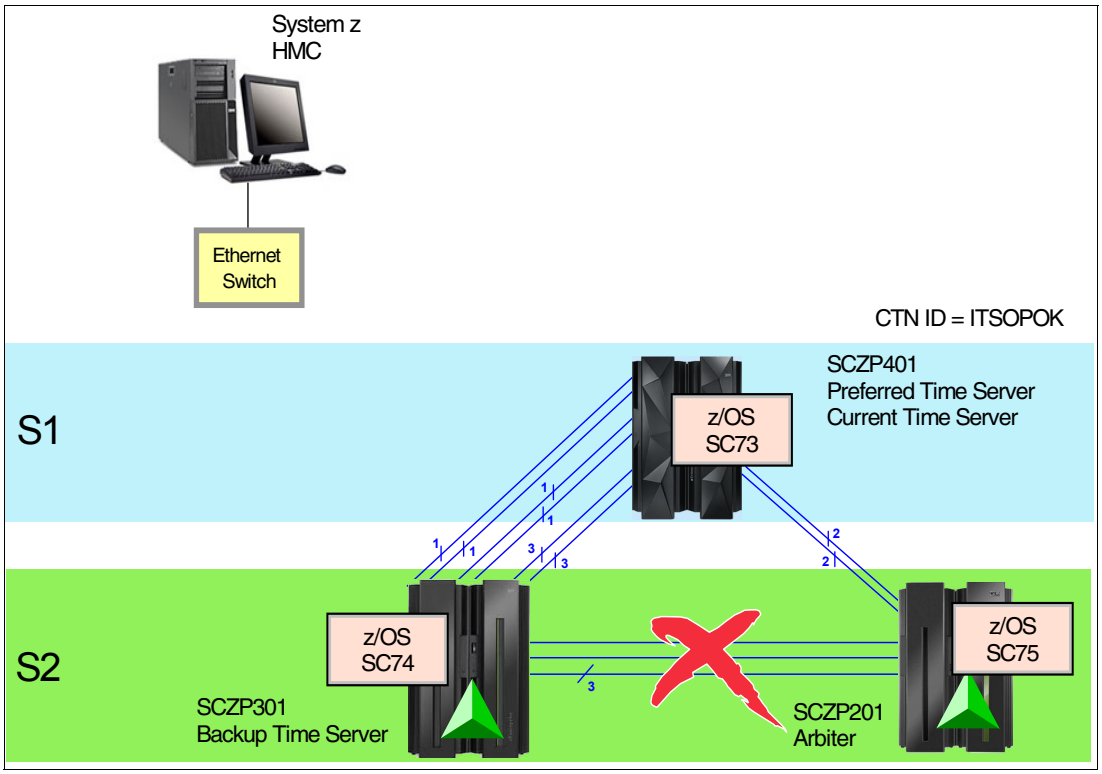

Figure 5-28 STP-only CTN; last link failure between CTS and Arbiter

Because the Arbiter no longer has direct connectivity to the CTS, it transitions to Stratum 3 to continue to receive STP timing signals.

In this state, the Arbiter is still able to communicate with the BTS, and therefore is still available for the BTS to use if the BTS subsequently loses contact with the CTS.

In Figure 5-28, when the coupling link between the CTS SCZP401 and the Arbiter SCZP201 fails, there is no impact on STP timing because the Arbiter SCZP201 is still able to use the BTS SCZP301 as a valid time source. However, message IEA383I is displayed because the Arbiter SCZP201 only has one network node to receive its timing signals from the BTS SCZP301. The key highlight of this scenario is that the Arbiter no longer has connectivity to the CTS and transitions to Stratum 3.

5.6.1 Problem awareness

When the last coupling link fails, message IXL158I is issued because the link failure also involves connectivity to a Coupling Facility partition:

```
IXL158I PATH 54 IS NOW NOT-OPERATIONAL TO CUID: FFD5 778
COUPLING FACILITY 002097.IBM.02.00000001DE50
PARTITION: 0D  CPCID: 00
```

This message indicates that the CHPID that was in use by the system to communicate with the identified Coupling Facility has become nonoperational. The cause is either a path failure or loss of the Coupling Facility. The system discontinues using the nonoperational path.

5.6.2 Problem determination

Figure 5-29 displays the Network Configuration tab seen from SCZP401, which shows that SCZP201 has connectivity to the BTS. The Arbiter is still assigned to SCZP201, but it has no STP ID.

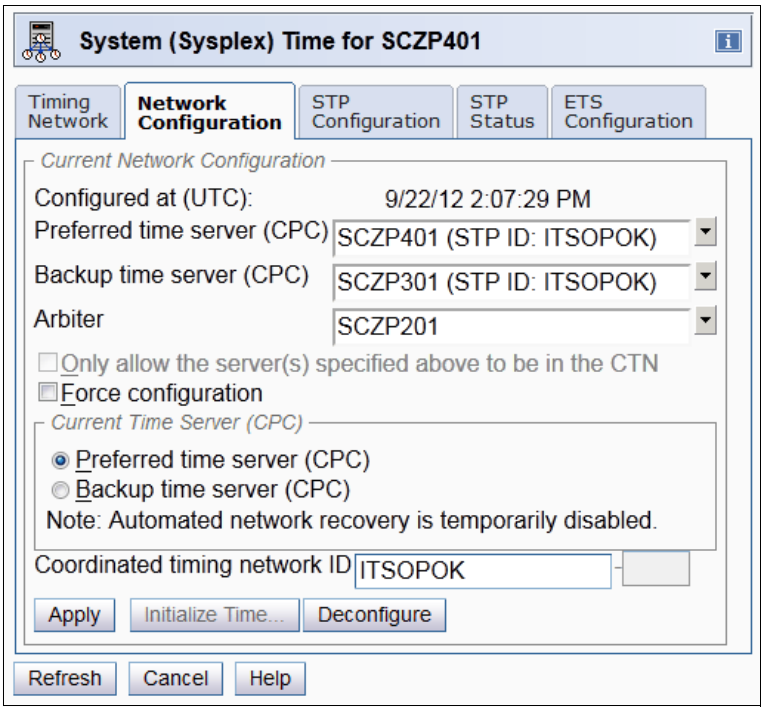

Figure 5-29 STP Configuration tab: SCZP201 after failure

The status of the coupling links can be determined using the STP Status tab. Figure 5-30 on page 165 shows the STP Status tab for SCZP401. It shows that SCZP401 has connectivity to SCZP301, but not to SCZP201.

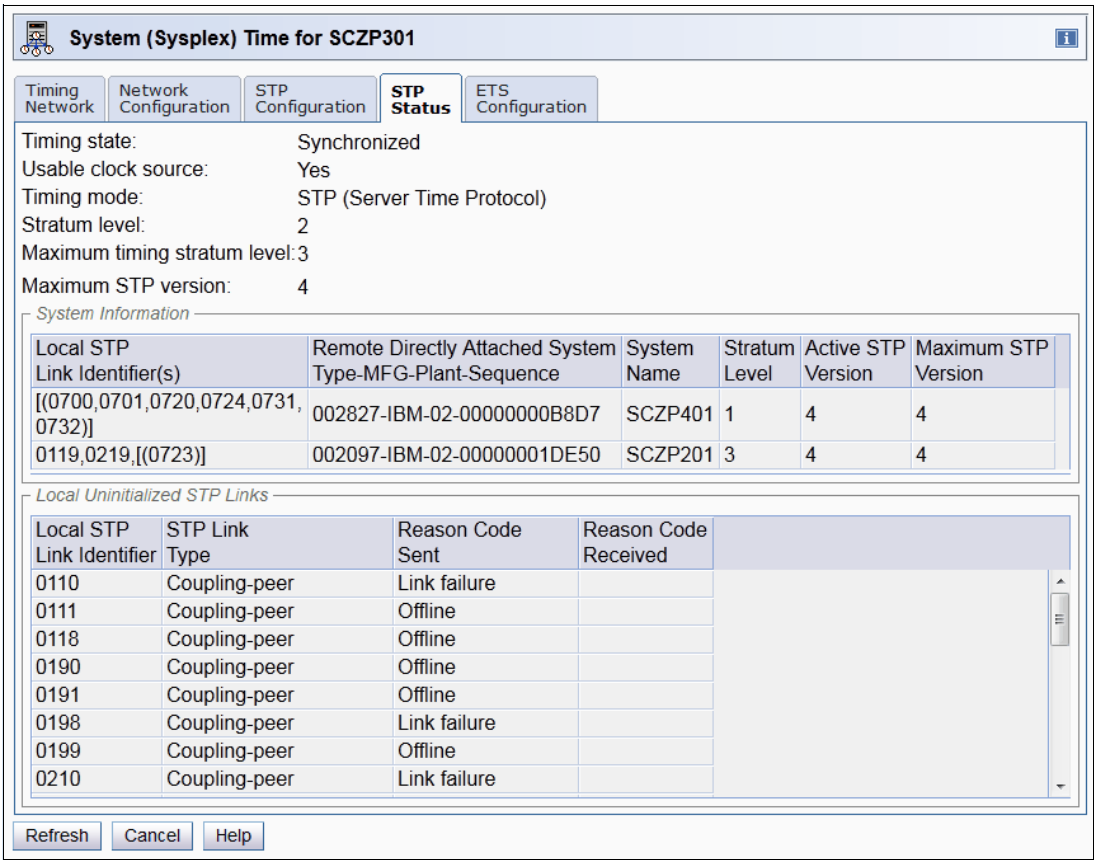

Figure 5-30 STP Status tab: SCZP301 after failure

An observation of the STP Status tab for SCZP301 shows that SCZP301 has connectivity to SCZP401 and to the Arbiter SCZP201. It also shows that SCZP201 is now at Stratum 3. SCZP201 transitions from Stratum 2 to Stratum 3 when connectivity to SCZP401 becomes unavailable.

Similar information is available by running the **`DISPLAY ETR`** command from the z/OS system images on each server, as shown in Figure 5-31 on page 166.

- On SC73, hosted on the PTS and CTS, the following messages are seen:
  - `THIS IS THE PREFERRED TIME SERVER`
  - `THIS SERVER HAS NO LINK TO THE ARBITER SERVER`
- On SC75, hosted on the Arbiter, the following messages are seen:
  - `THIS SERVER IS A STRATUM 3`
  - `THIS SERVER HAS NO LINK TO THE PREFERRED TIME SERVER`
  - `THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS`

```
SC73 RESPONSES ------------------------------------------------
 IEA386I 16.34.18 TIMING STATUS 188
 SYNCHRONIZATION MODE = STP
   THIS SERVER IS A STRATUM 1
   CTN ID = ITSOPOK
   THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
   THIS IS THE PREFERRED TIME SERVER
   THIS SERVER HAS NO LINK TO THE ARBITER
SC74 RESPONSES ------------------------------------------------
 IEA386I 16.34.18 TIMING STATUS 134
 SYNCHRONIZATION MODE = STP
   THIS SERVER IS A STRATUM 2
   CTN ID = ITSOPOK
   THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
   THIS IS THE BACKUP TIME SERVER
   NUMBER OF USABLE TIMING LINKS = 10
SC75 RESPONSES ------------------------------------------------
 IEA386I 16.34.18 TIMING STATUS 791
 SYNCHRONIZATION MODE = STP
   THIS SERVER IS A STRATUM 3
   CTN ID = ITSOPOK
   THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
   THIS IS THE ARBITER SERVER
   THIS SERVER HAS NO LINK TO THE PREFERRED TIME SERVER
   NUMBER OF USABLE TIMING LINKS =  5
   THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS
```

Figure 5-31 DISPLAY ETR command output

5.6.3 User actions

Identify the coupling link that has failed and take the appropriate actions to restore link connectivity, such as configuring the CHPID online. When link connectivity has been restored, it can be verified by examining the HMC.

Figure 5-32 shows the Network Configuration tab for SCZP201. The CTN is operational, and the servers have maintained their original roles.

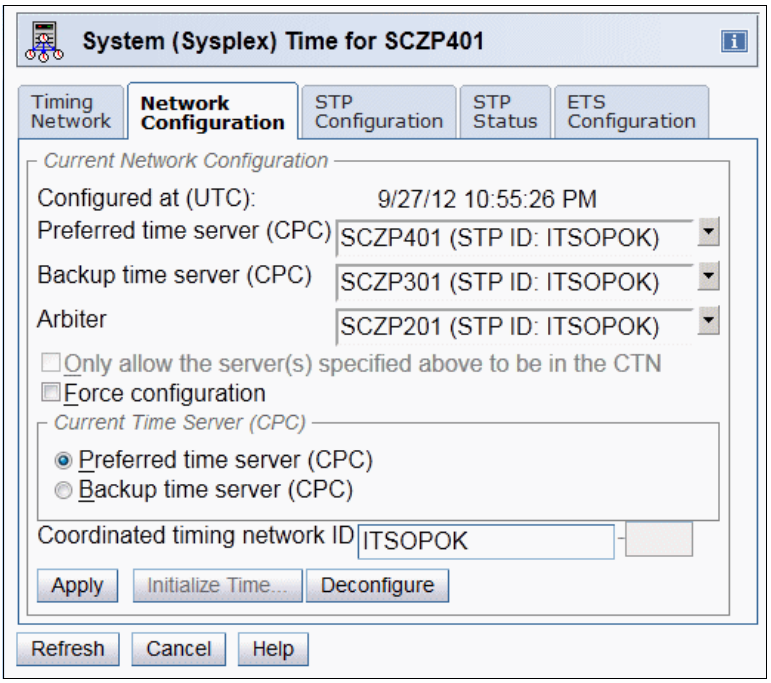

Figure 5-32 Network Configuration tab: SCZP401 after repair action

When communication is reinstated between the CTS and the Arbiter, the Arbiter transitions back to Stratum 2. After the coupling link problem has been corrected, the STP Status tabs for SCZP201 show connectivity to SCZP401 and SCZP301.

5.7 Last coupling link failure between the BTS and the Arbiter

In the following scenario, only the failure of the last coupling link between any two servers is considered. Figure 5-33 on page 168 shows an example where the last coupling link between the BTS and the Arbiter fails. When the coupling link between the Arbiter SCZP201 and the BTS SCZP301 fails, there is no impact from an STP timing point of view because the BTS SCZP301 is still able to use the CTS SCZP401 as a valid time source.

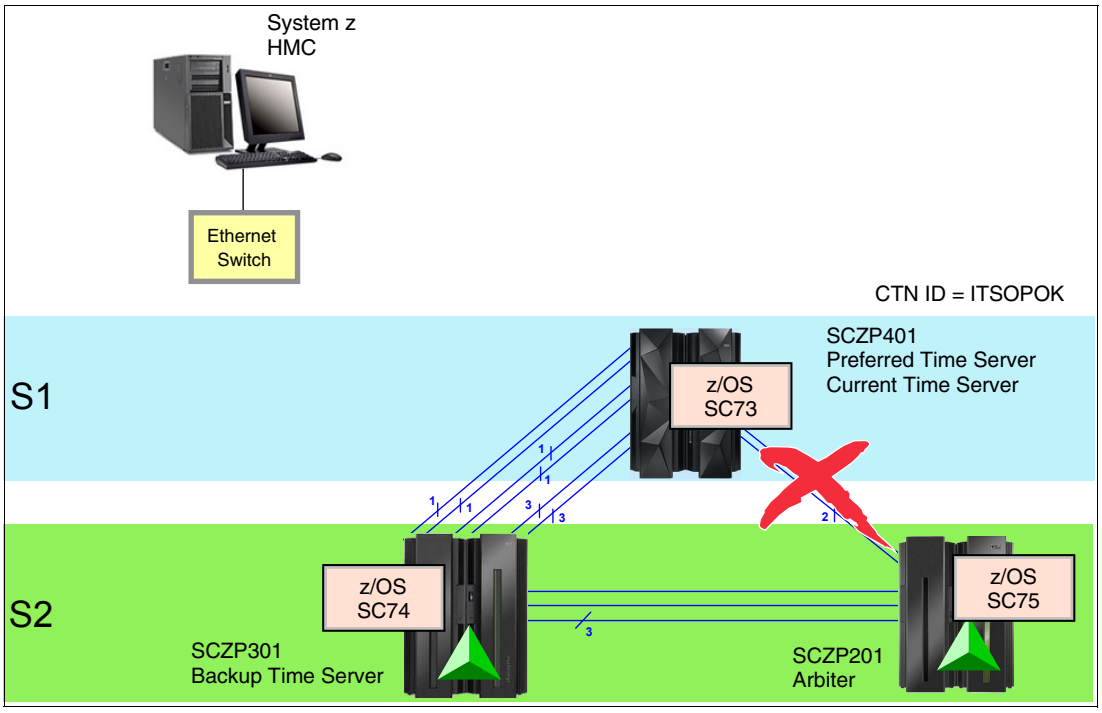

Figure 5-33 STP-only CTN, BTS and Arbiter, last link failure between BTS and Arbiter

At this stage, the CTN is unaffected but a single point of failure has been created. The Arbiter can no longer fulfill its mission. If a CTS failure occurs later, the BTS will be unable to initiate Arbiter-assisted recovery and might become unsynchronized.

The key highlights of this scenario are:

- Both the BTS and the Arbiter remain at Stratum 2.
- There is no impact from a time synchronization point of view because the BTS is still able to use the CTS as a valid time source.
- A single point of failure has been created. The BTS will no longer be able to initiate Arbiter-assisted recovery if the need occurs.

5.7.1 Problem awareness

When the last coupling link fails, message IXL158I is issued:

```
IXL158I PATH 53 IS NOW NOT-OPERATIONAL TO CUID: FFD5 838
COUPLING FACILITY 002097.IBM.02.00000001DE50
PARTITION: 0D  CPCID: 00
```

This message indicates that the CHPID that was in use by the system to communicate with the identified Coupling Facility has become nonoperational. The cause is either a path failure or loss of the Coupling Facility. The system discontinues using the nonoperational path.

5.7.2 Problem determination

Figure 5-34 displays the Network Configuration tab seen from SCZP301. It shows that SCZP301 has connectivity to the PTS. The Arbiter is still assigned to SCZP201, but there is no STP ID in its entry, which is an indication that there is no STP connectivity between SCZP301 and SCZP201.

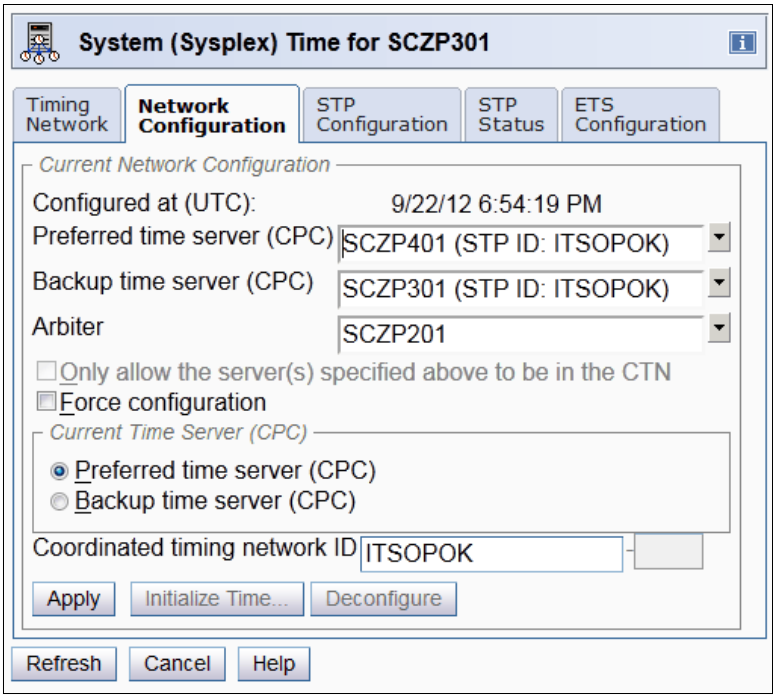

Figure 5-34 Network Configuration tab: SCZP301 after failure

Both the Arbiter, SCZP201, and the BTS, SCZP301, remain at Stratum 2 after the link failure. SCZP301 can continue to function as the BTS through communication with SCZP401.

The status of the coupling links can be determined using the STP Status tab. Figure 5-35 shows the STP Status tab of the BTS SCZP301. It shows that SCZP301 has connectivity to SCZP401, but has no connectivity to the Arbiter SCZP201.

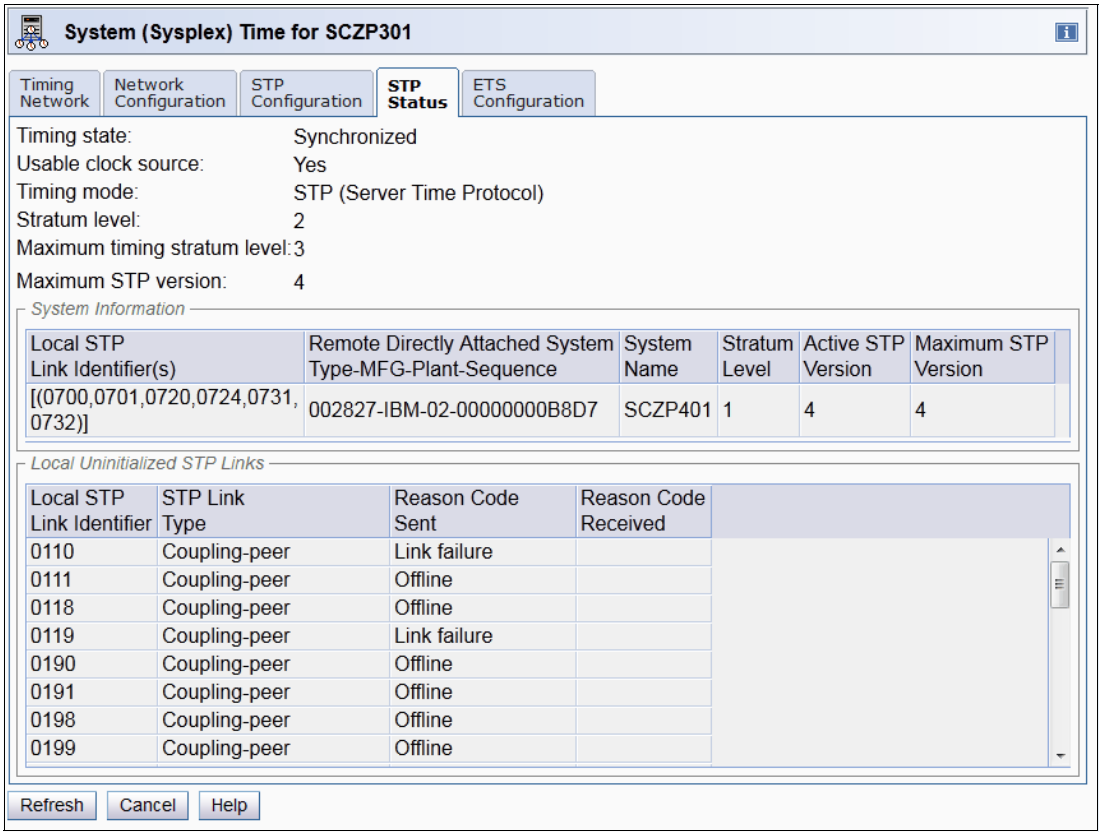

Figure 5-35 STP Status tab: BTS SCZP301 after failure

The Arbiter, SCZP201, remains at Stratum 2 after the link failure. However, it cannot continue to fulfill its role as an Arbiter because it has lost communication with the BTS SCZP301.

The information is also available by running the **DISPLAY ETR** command from the z/OS system images on each server, as shown in Figure 5-36.

- On SC74, hosted on the BTS, the following messages are seen:
  - THIS SERVER IS A STRATUM 2
  - THIS IS THE BACKUP TIME SERVER
  - THIS SERVER HAS NO LINK TO THE ARBITER SERVER
  - THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS

- On SC75, hosted on the Arbiter, the following messages are seen:
  - THIS SERVER IS A STRATUM 2
  - THIS SERVER HAS NO LINK TO THE BACKUP TIME SERVER
  - THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS

```
SC73 RESPONSES ----------------------------------------------------
 IEA386I 18.03.24 TIMING STATUS 346
 SYNCHRONIZATION MODE = STP
   THIS SERVER IS A STRATUM 1
   CTN ID = ITSOPOK
   THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
   THIS IS THE PREFERRED TIME SERVER
SC74 RESPONSES ----------------------------------------------------
 IEA386I 18.03.24 TIMING STATUS 262
 SYNCHRONIZATION MODE = STP
   THIS SERVER IS A STRATUM 2
   CTN ID = ITSOPOK
   THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
   THIS IS THE BACKUP TIME SERVER
   THIS SERVER HAS NO LINK TO THE ARBITER SERVER
   NUMBER OF USABLE TIMING LINKS =  10
   THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS
SC75 RESPONSES ----------------------------------------------------
 IEA386I 18.03.24 TIMING STATUS 943
 SYNCHRONIZATION MODE = STP
   THIS SERVER IS A STRATUM 2
   CTN ID = ITSOPOK
   THE STRATUM 1 NODE ID = 002827.H43.IBM.02.00000000B8D7
   THIS IS THE ARBITER SERVER
   THIS SERVER HAS NO LINK TO THE BACKUP TIME SERVER
   NUMBER OF USABLE TIMING LINKS = 4
   THIS SERVER HAS ONLY A SINGLE SOURCE OF TIMING SIGNALS
```

Figure 5-36 DISPLAY ETR command output

5.7.3 User actions

Identify the coupling link that has failed and take the appropriate actions to restore link connectivity, such as configuring the CHPID online.

When link connectivity has been restored, it can be verified from the System (Sysplex) Time task windows on the HMC.

Figure 5-37 shows the Network Configuration tab from SCZP401. It shows that SCZP301 now has connectivity to both the PTS and the Arbiter. All servers have kept their original roles.

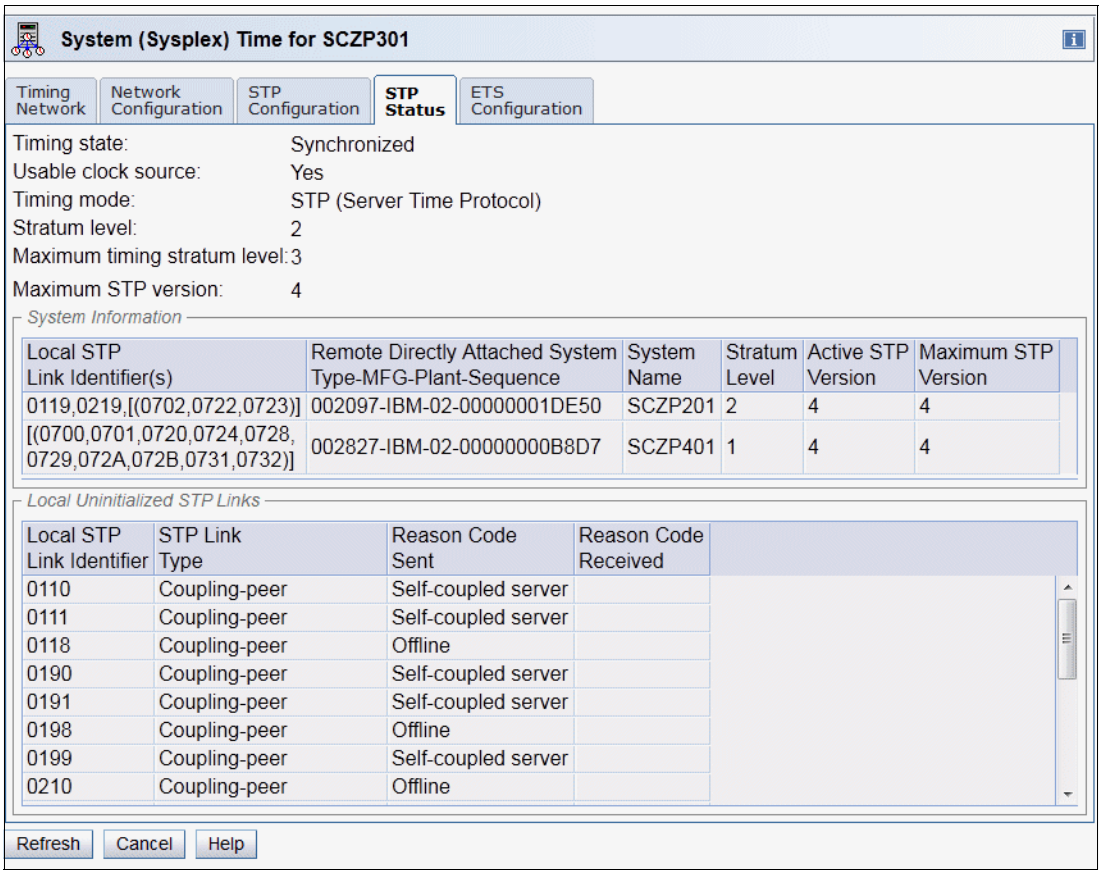

Figure 5-37 Network Configuration tab: PTS SCZP301 after recovery and user action

5.8 Two-site: Site 1 failure, CTS and Arbiter at Site 1, BTS at Site 2

Figure 5-38 on page 173 shows the recommended two-site configuration. In this scenario, both the CTS and the Arbiter are lost because they are located at Site 1, which has failed. The BTS located at Site 2 survives the failure, but has lost the STP time source, and therefore needs to determine whether it is able to assume the role of CTS. The key highlights of this scenario are listed here:

- Site 1 is lost. Both the CTS and the Arbiter are down.
- The BTS loses communication with the CTS.
- The BTS tries to communicate with the Arbiter to determine the status of the CTS, but communication with the Arbiter is also unavailable.
- The BTS becomes unsynchronized and transitions to Stratum 0.

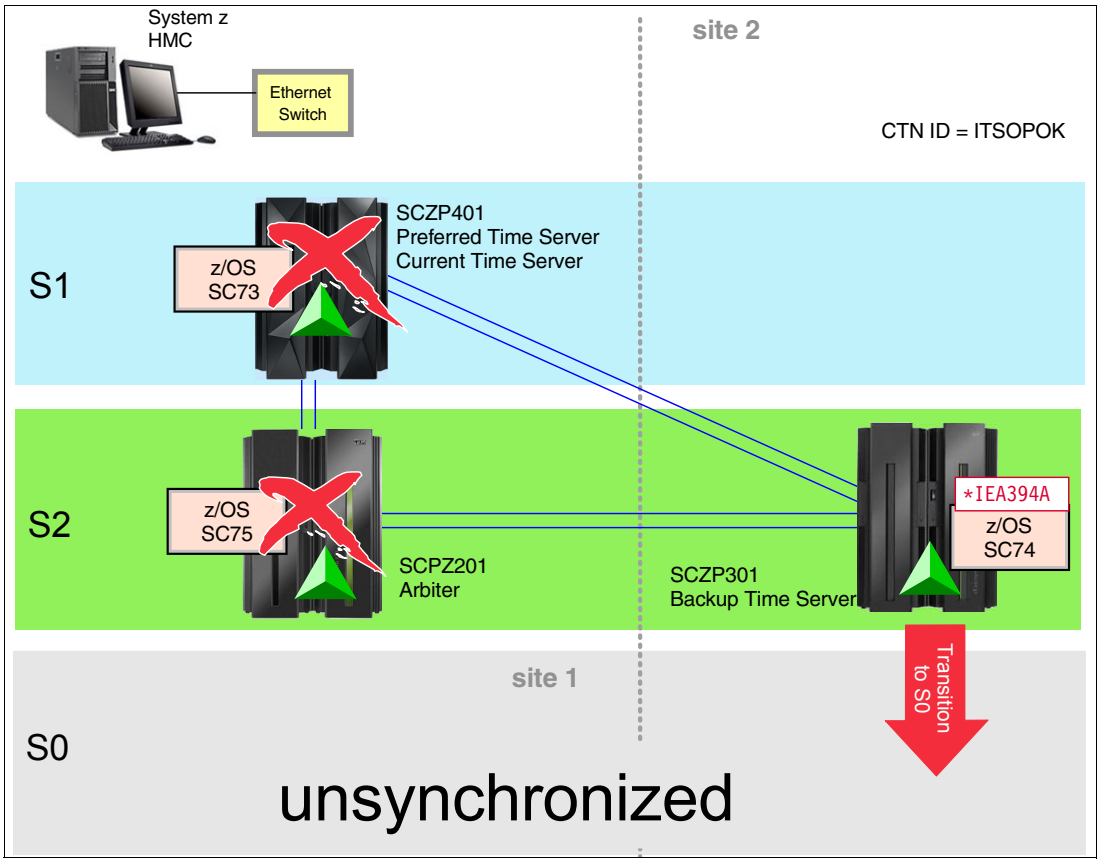

Figure 5-38 STP-only CTN, BTS and Arbiter CTN, Site 1 failure

Note: When the STP-only CTN spans multiple sites, the location of the Arbiter is critical. Refer to *Server Time Protocol Planning Guide,* SG24-7280, to understand the consequences of locating the Arbiter at Site 2.

5.8.1 Problem awareness

As shown in Figure 5-39, every z/OS system image on server SCZP301 that loses its time source will post WTOR message IEA394A.

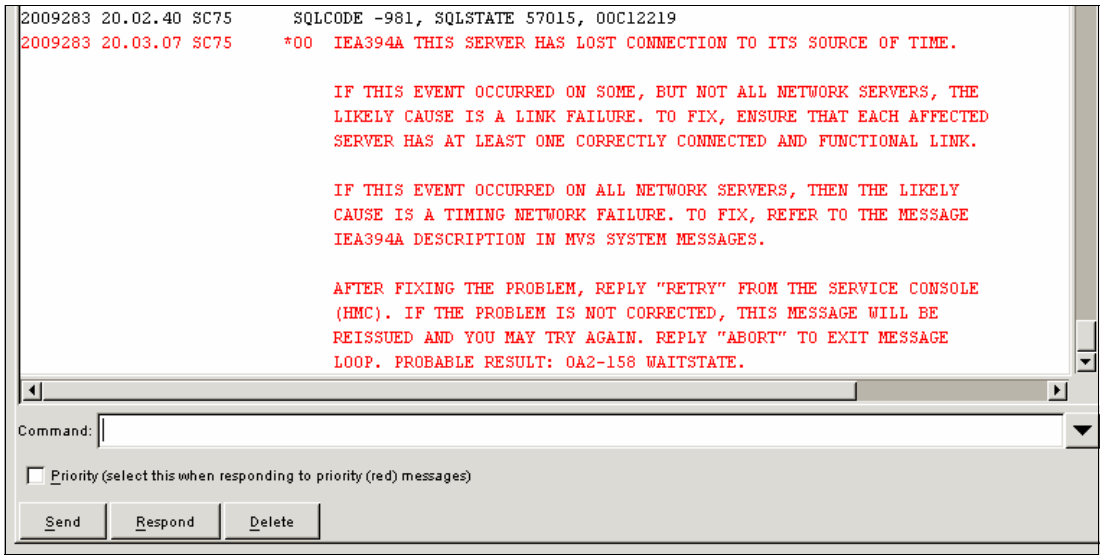

Figure 5-39 WTOR message IEA394A: z/OS image SC74

5.8.2 Problem determination

Figure 5-40 shows the STP Status tab of SCZP301 after Site 1 has failed.

- Timing state is unsynchronized.
- Usable clock source is no.
- Stratum level is 0.
- SCZP301 has no local STP link identifiers to any other servers.

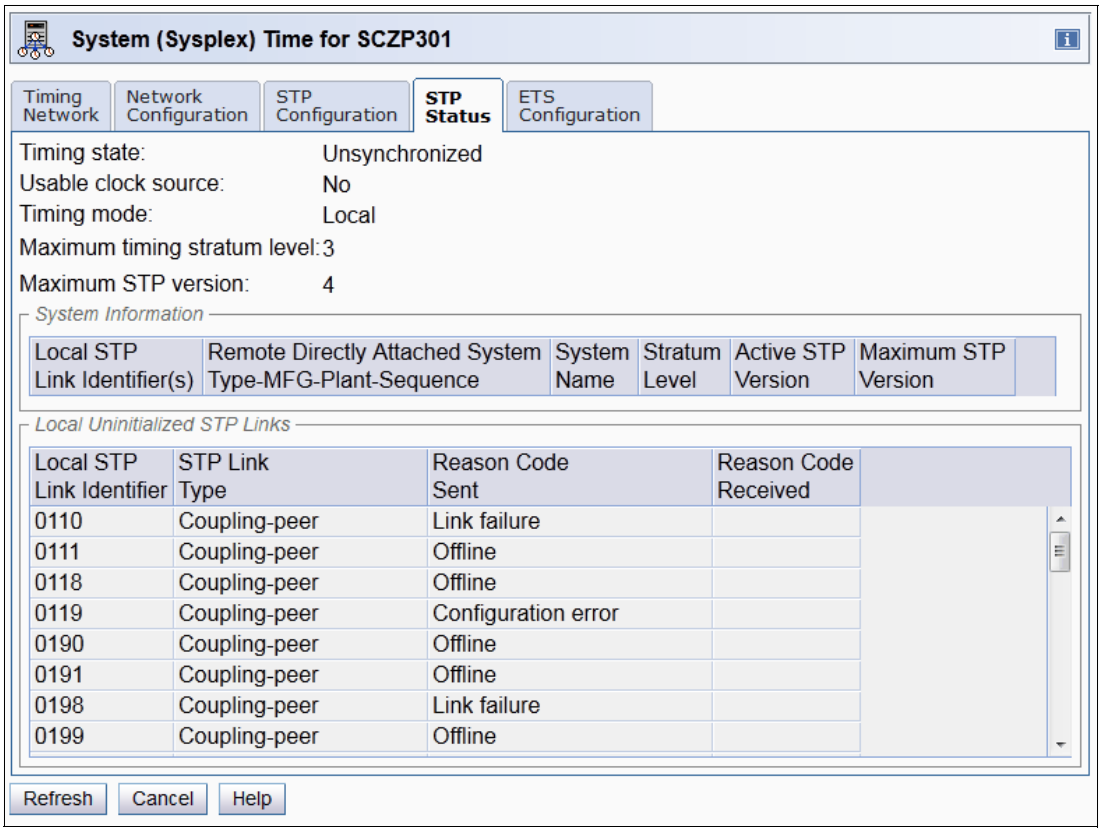

Figure 5-40 STP Status tab: SCZP301 after failure

5.8.3 User actions

If Site 1 cannot be recovered, the BTS may be reassigned as the PTS. The recovery process is identical to that described in Chapter 4.2, “Current Time Server failure: no Going Away Signal (GAS) or OLS received” on page 116.

Figure 5-41 on page 176 shows the Network Configuration tab of the BTS SCZP301 after Site 1 has failed. Initially, all fields in the Current Network Configuration section display `Not configured`. The Initialize Time button is enabled. However, there is no need to reinitialize the time.

To reconfigure the BTS as the PTS, follow these steps:

1. Set SCZP301 as the Preferred Time Server. The Coordinated Timing Network fields are left unchanged.
2. Click the **Force configuration** check box. This is mandatory because there is no CTS configured yet.

3. Click **Apply** to reconfigure SCZP301 as the PTS and CTS.

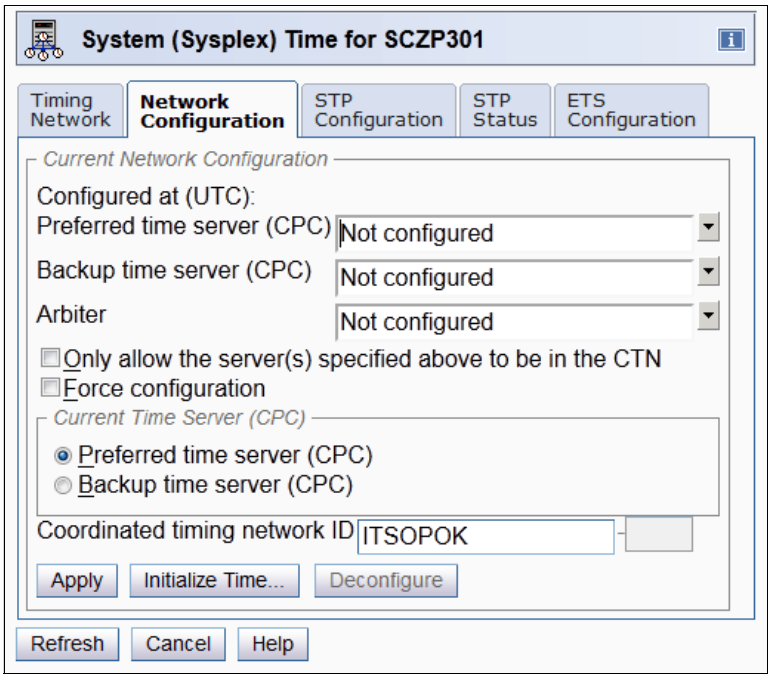

Figure 5-41 Network Configuration tab: SCZP301 after failure

Figure 5-42 shows the Network Configuration tab of SCZP201 after it has been reassigned as the PTS.

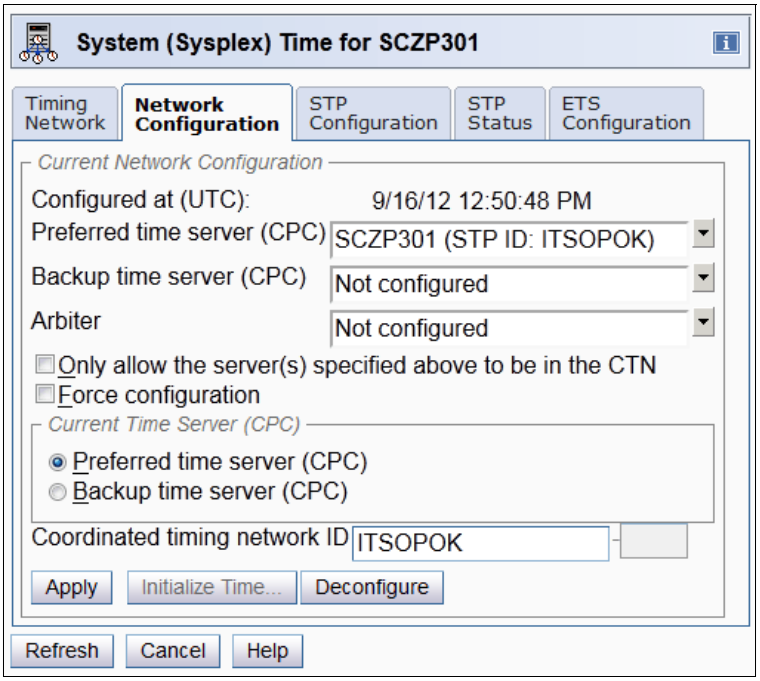

Figure 5-42 Network Configuration tab: SCZP301 after CTN reconfiguration

As shown in the STP Status tab in Figure 5-43, after reconfiguration, SCZP301 is Stratum 1.

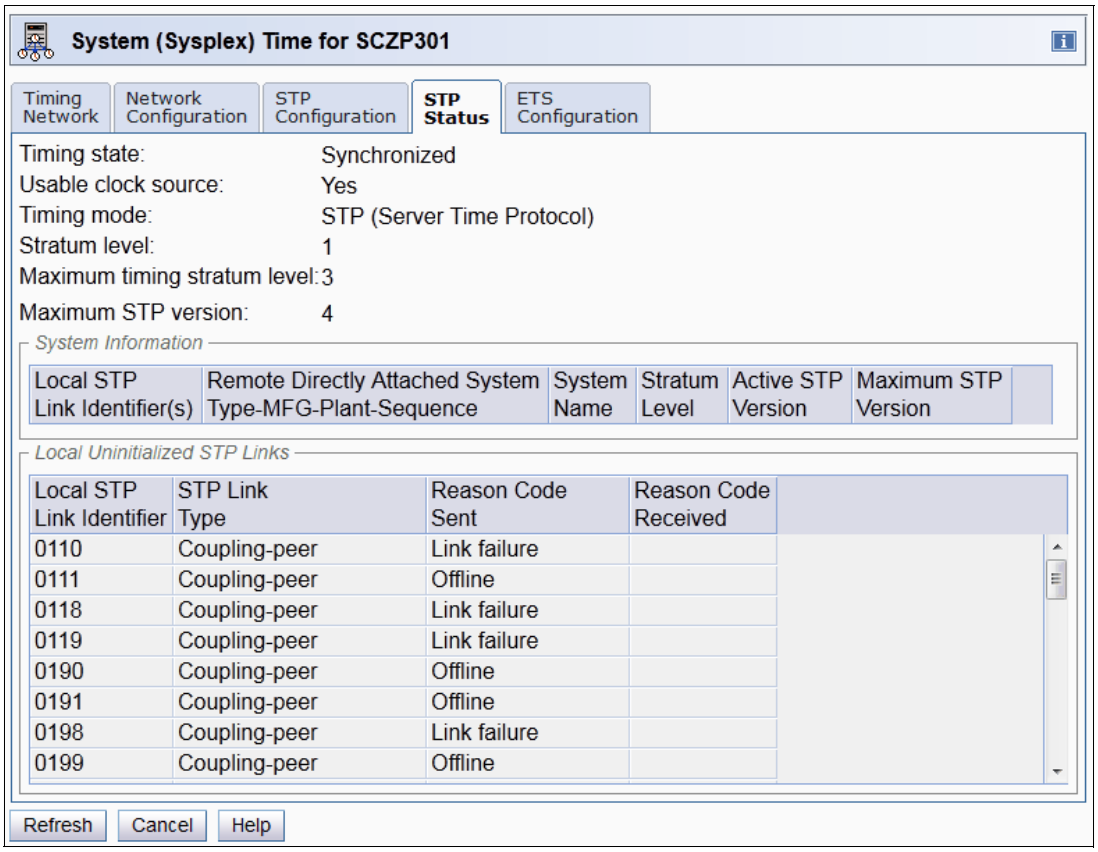

Figure 5-43 STP Status tab: SCZP301 after CTN reconfiguration

To resume processing on system image SC74 on SCZP301, click the **Priority** check box on the Operating System messages window and reply `RETRY` to WTOR message IEA394A. There is nothing that needs to be done other than replying to IEA394A after the server has regained STP synchronization.

5.9 Two-site: Site 1 power outage, CTS and Arbiter at Site 1 with IBF

Figure 5-44 on page 178 shows the recommended two-site configuration. In this scenario, Site 1 is affected by a power outage. Both servers, the PTS and the Arbiter, are not immediately affected, because they both have the Internal Battery Feature (IBF) installed. They both notify the BTS through the coupling links that they are running on IBF power. If the BTS does not get a “normal power” status within the next 30 seconds from the PTS or CTS, it initiates processing to take over the CTS role. For details about the IBF recovery and considerations, refer to 1.3, “Internal Battery Feature considerations” on page 16.

In our scenario, we assume that the power outage exists for more than 30 seconds, so the BTS takes over as Stratum 1. Furthermore, the scenario shows that the PTS and Arbiter fail because their IBF runs out of power. The time frame that each IBF can maintain the PTS and the Arbiter depends on the hardware configuration of each server.

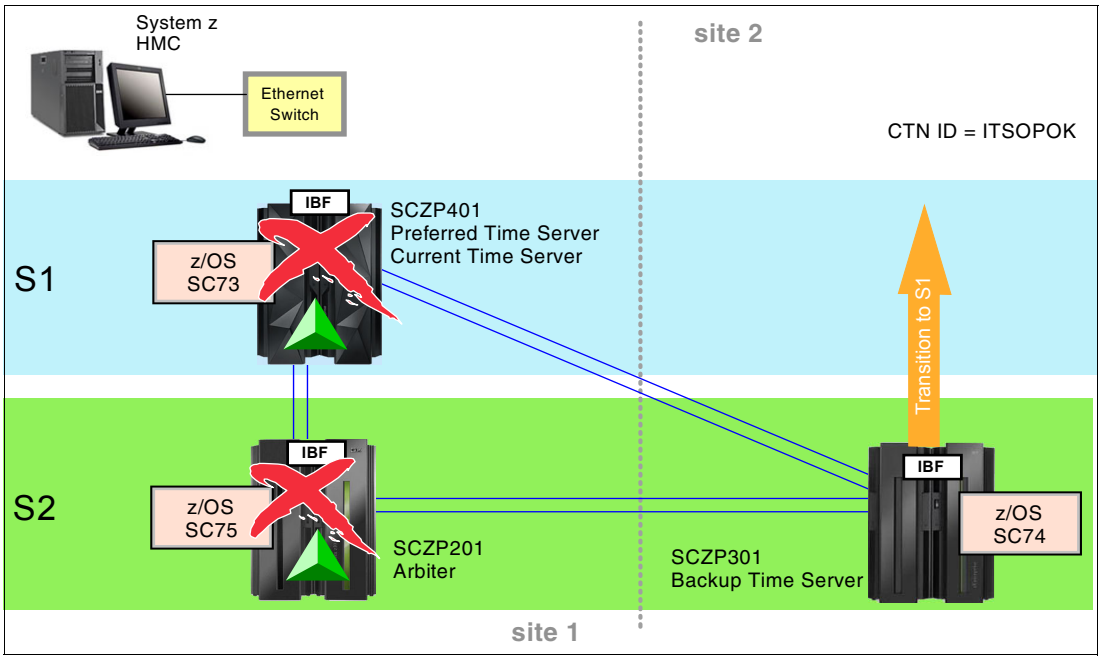

Figure 5-44 STP-only CTN, BTS and Arbiter CTN, site failure with Internal Battery Feature

The key highlights of this scenario are listed here:

- Site 1 experiences a power outage. Both servers, the PTS and the Arbiter, are powered by their IBF.
- The PTS and CTS, and the Arbiter, send a message through coupling links to the BTS indicating that they are running on IBF power.
- The BTS takes over the CTS role and becomes Stratum 1 because this condition exists for more than 30 seconds, while the PTS transitions to Stratum 2.
- The PTS and the Arbiter are lost when their IBFs run out of power.

5.9.1 Problem awareness

A Hardware Message is posted on the HMC indicating there is a power problem at the PTS and the Arbiter. However, there is no impact at this time because the PTS and the Arbiter stay on battery power for a certain period of time.

Because the Current Time Server switched from the Preferred Time Server to the Backup Time Server after 30 seconds, z/OS (Version 1.11 or higher) issues message IEA395I on all z/OS images that are members of this CTN and use STPMODE=YES in the CLOCKxx member:

```
IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE BACKUP
```

After the IBF has been drained, both servers, SCPZ301 and SCZP201, power down and their LPARs are lost. Both servers' images, SC75 and SC73, are lost immediately and get partitioned out right after the power outage:

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC75 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC73 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
```

An active SFM system isolation policy automatically starts system partitioning. This is highlighted by an IXC101I message only a few seconds after detection of the failure. Partitioning triggers the cleanup of resources for the failing system. If system isolation fails, then SFM issues WTOR message IXC102A after the XCF CLEANUP time has elapsed:

```
IXC102A XCF IS WAITING FOR SYSTEM SC75 DEACTIVATION. REPLY DOWN WHEN MVS ON SC75
HAS BEEN SYSTEM RESET
IXC102A XCF IS WAITING FOR SYSTEM SC73 DEACTIVATION. REPLY DOWN WHEN MVS ON SC73
HAS BEEN SYSTEM RESET
```

5.9.2 Problem determination

If BTS has not received a “normal power” message for more than 30 seconds, it takes over as CTS. The network configuration tab in Figure 5-45 shows the following information:

- The Current Time Server has switched to the Backup Time Server.
- Automated network recovery is temporarily disabled.

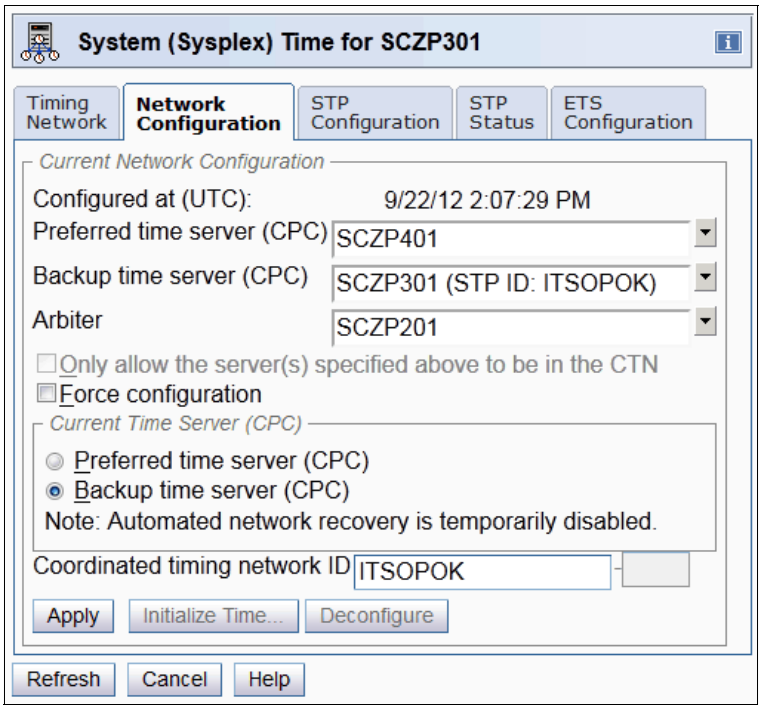

Figure 5-45 Network Configuration tab: SCZP301 30 seconds after the power outage

In addition, the STP status tab shows that server SCZP301 is Stratum 1. Furthermore, after the IBFs of server SCZP401 and SCZP201 have been discharged, both servers are powered down. After SCZP401 and SCZP201 are lost, the Network Configuration window in SCZP301 still shows SCZP401 as the PTS and SCZP201 as the Arbiter, but their subsequent STPID is missing. This indicates the CTN lost communication to server SCZP401 and SCZP201. The STP status tab shows SCZP301 as Stratum 1 and no connectivity to either server SCZP401 or server SCZP201, as displayed in Figure 5-46 on page 180.

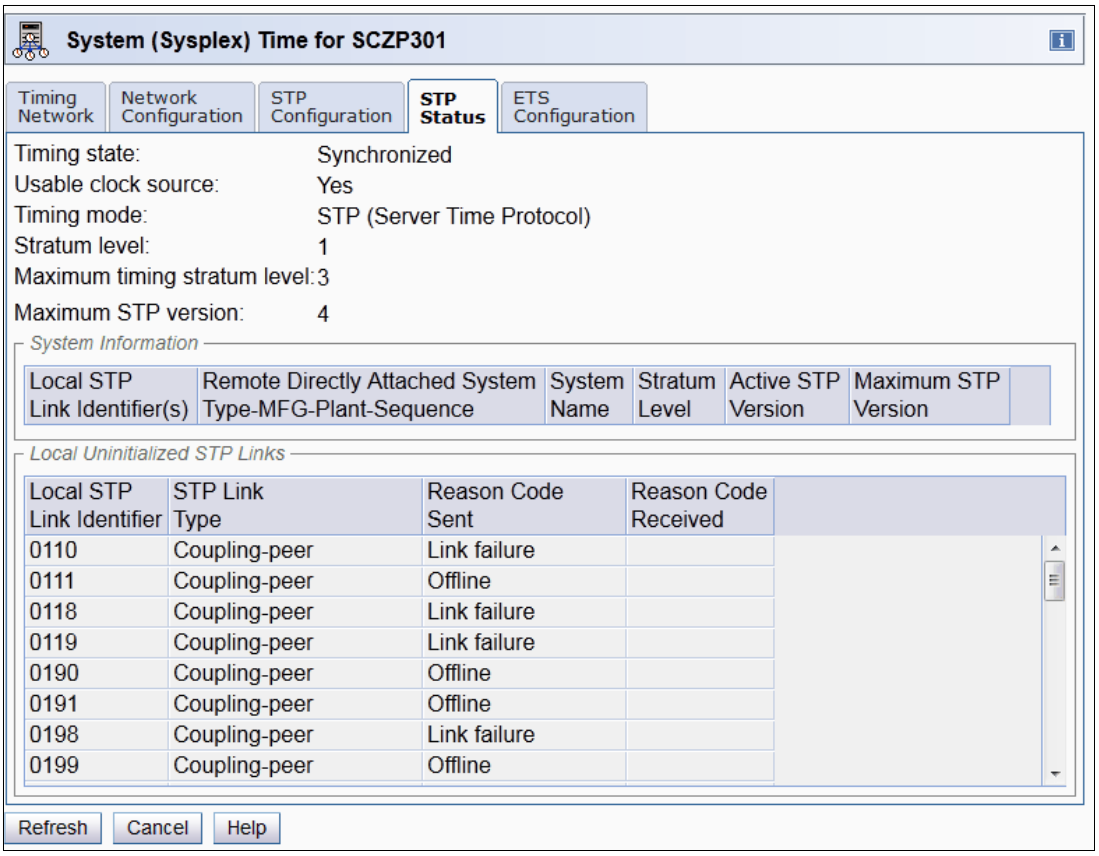

Figure 5-46 STP status tab: SCZP301 (BTS) after SCZP401 and SCZP201 IBFs are discharged

5.9.3 User actions

No user actions are required for the CTN. The BTS SCZP301 in this scenario took over the role of the CTS. SCZP401 and SCZP201 remain unavailable until their power has been restored.

After SCZP401 and SCZP201 have been powered up again, they will join the CTN. Because the STP-only CTN was not reconfigured by reassigning roles during the outage, SCZP401 will automatically take back its original role as CTS, and SCZP201 becomes the Arbiter assigned before the power outage. As a consequence, z/OS (Version 1.11 or higher) issues message IEA395I on all z/OS images that are members of this CTN and that use STPMODE=YES in the CLOCKxx member:

```
IEA395I THE CURRENT TIME SERVER HAS CHANGED TO THE PREFERED
```

Figure 5-47 shows the Network configuration tab of SCZP301 after both server SCZP401 and SCZP201 have been recovered.

- The Current Time Server (CPC) section shows that SCZP401 is the CTS.
- Automated network recovery is not disabled.

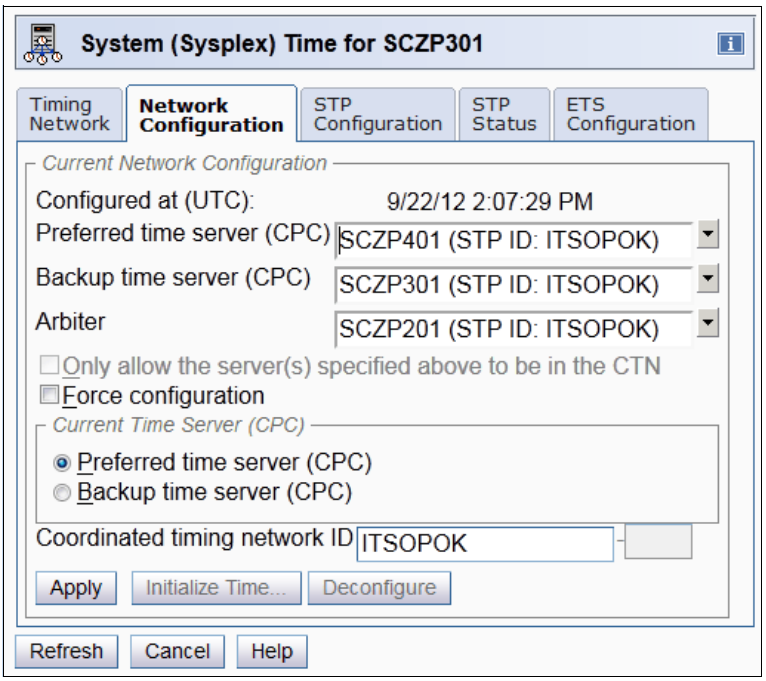

Figure 5-47 Network configuration tab: SCZP301 after recovery

5.10 Two-site: Site 2 failure, CTS and Arbiter at Site 1, BTS at Site 2

In this scenario (see Figure 5-48 on page 182), both the CTS and the Arbiter survive this failure because they are located at Site 1. Only the BTS is lost, because it is located at Site 2. Both the CTS and the Arbiter survive and are able to remain synchronized without the need for further recovery processing. The key highlights of this scenario are listed here:

- Site 2 has failed. Consequently, the BTS is down or not operational.
- The CTS loses all communication with the BTS. It communicates successfully with the Arbiter.
- The CTS retains its role and servers at Site 1 remain synchronized.

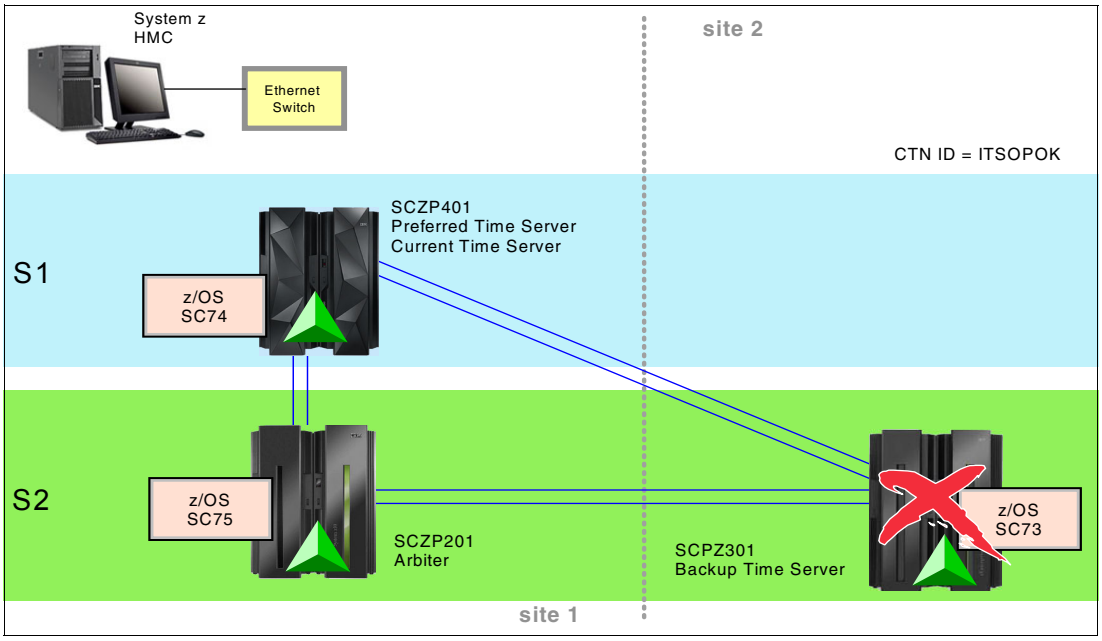

Figure 5-48 STP-only CTN, BTS and Arbiter configuration, Site 2 failure

5.10.1 Problem awareness

z/OS messages are issued, indicating that the system images on the BTS (SCPZ301) have failed:

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC73 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
```

An active SFM system isolation policy automatically starts system partitioning. This is highlighted by an IXC101I message only a few seconds after detection of the failure. Partitioning triggers the cleanup of resources for the failing system. If system isolation fails, then SFM will issue WTOR message IXC102A after the XCF CLEANUP time has elapsed:

```
IXC102A XCF IS WAITING FOR SYSTEM SC73 DEACTIVATION. REPLY DOWN
WHEN MVS ON SC73 HAS BEEN SYSTEM RESET
```

Note: Before replying `DOWN` to IXC102A or IXC402D, you must perform a hardware SYSTEM RESET on the z/OS system being removed. This is necessary to ensure that this z/OS system releases any outstanding I/O reserves. A SYSTEM RESET ensures that other systems continue to have access to the data sets on the shared devices.

5.10.2 Problem determination

Figure 5-49 shows the STP Status tab for the CTS SCZP401. The CTS SCZP401 has connectivity to the Arbiter SCZP201, but no connectivity to the BTS SCZP301 because of the site failure.

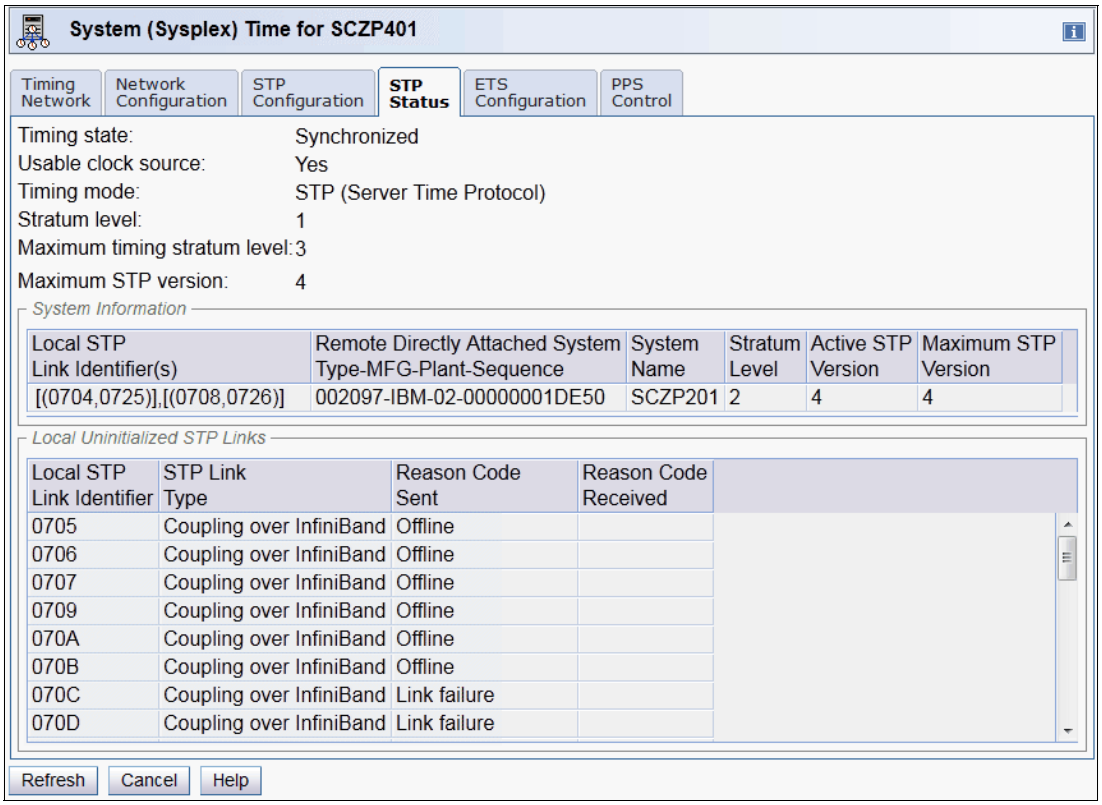

Figure 5-49 STP Status tab: PTS SCZP201 after failure

Figure 5-50 on page 184 shows the Network Configuration tab for the CTS SCZP401.

- The Current Time Server (CPC) section shows that the PTS SCZP401 is the CTS.
- If the Support Element for SCZP301 is still visible to the HMC, the Backup Time Server field still indicates SCZP301 as the BTS, but with no STP ID associated with it. If the server Support Element is unavailable, the field displays the CPC node descriptor.

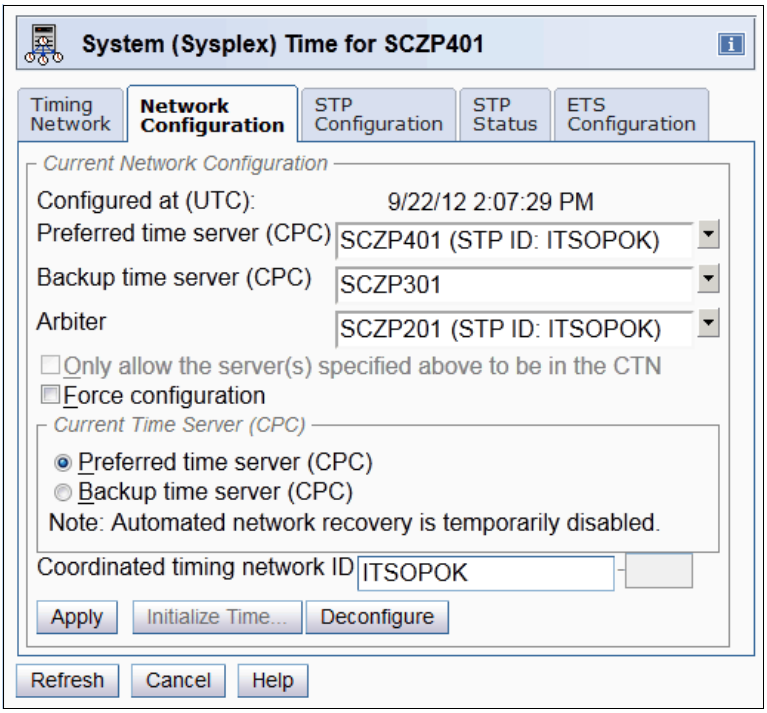

Figure 5-50 Network Configuration tab: PTS SCZP201 after failure

5.10.3 User actions

The CTS and the Arbiter both survive this failure because they are located at Site 1.

Restore Site 2 and re-IPL SCZP301 using installation recovery procedures.

When SCZP301 is operational, it rejoins the CTN. Examine the STP Status tabs from various servers to verify that the CTN has successfully returned to its original state.

Figure 5-51 shows the Network Configuration for SCZP401 after server SCZP301 has become operational.

- The Preferred Time Server field of the Current Network Configuration section shows SCZP401 being defined as the PTS.
- The Backup Time Server field of the Current Network Configuration section shows SCZP301 being defined as the BTS. The STP ID value is shown.
- The Current Time Server (CPC) section shows that the PTS (SCZP401) is the CTS.

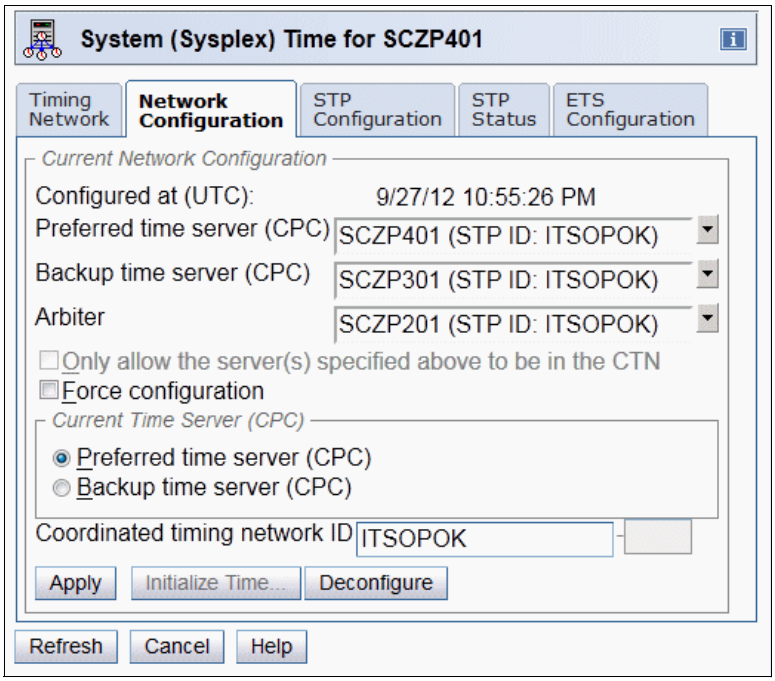

Figure 5-51 Network Configuration tab: PTS SCZP401 after recovery and user action

As shown in Figure 5-52, the CTS SCZP401 has connectivity to both the Arbiter SCZP201 and the BTS SCZP301.

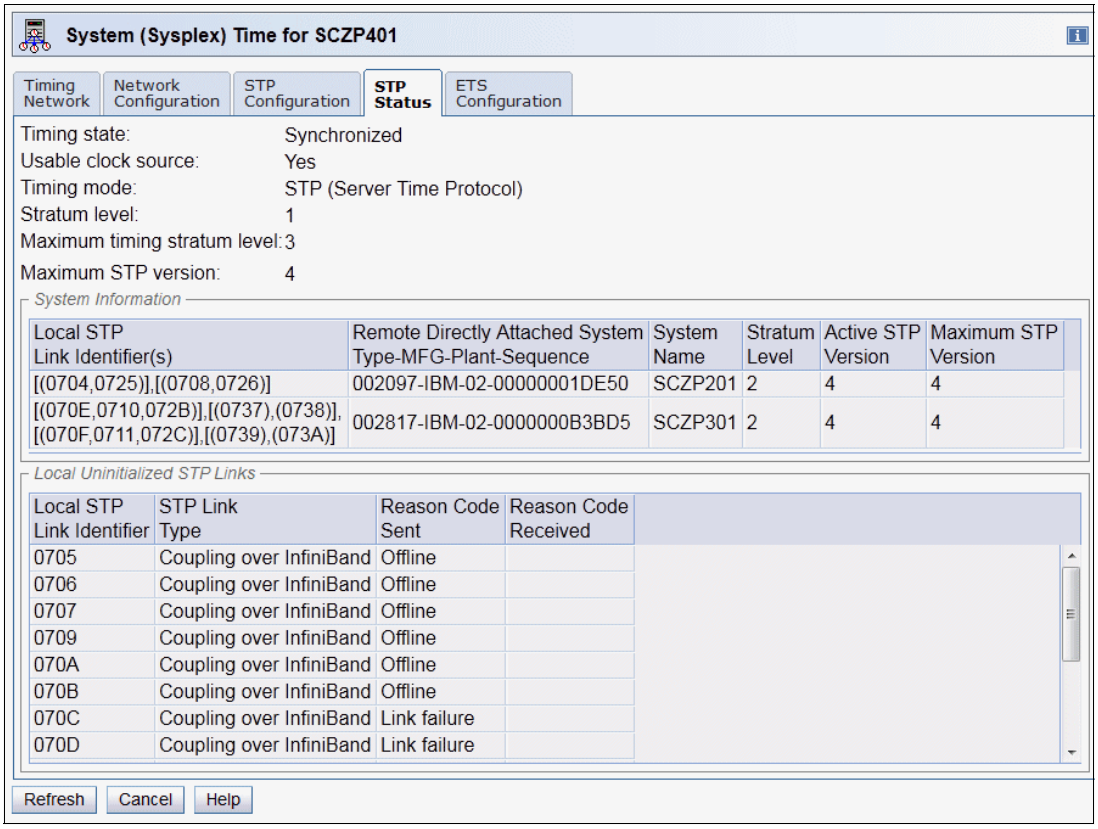

Figure 5-52 STP Status tab: PTS SCZP401 after recovery and user action

5.11 Two-site: Site 2 power outage, CTS, BTS, and Arbiter with IBF

In this example, all three servers have the Internal Battery Feature (IBF) installed, as shown in Figure 5-53. The PTS is located at Site 1 and the BTS and the Arbiter are located at Site 2. Usually this configuration is not recommended, because a Site 2 failure impacts systems at Site 1. However, the Internal Battery Feature installed in all three servers adds a different aspect to this configuration and the recovery behavior.

In this scenario, Site 2 is affected by a power outage. Both servers (BTS and the Arbiter) are not immediately affected, because they both have the Internal Battery Feature (IBF) installed. They both notify the PTS/CTS through the coupling links that they are running on IBF power. If the PTS/CTS does not receive a “normal power” status within the next 30 seconds from the BTS and Arbiter, it remains Stratum 1 and disables the automated recovery.

In this scenario, we assume the power outage exists for more than 30 seconds, so the PTS/CTS remains as Stratum 1 and disables the automated recovery. Furthermore, the scenario shows that the BTS and Arbiter have failed because their IBFs run out of power. The time each IBF can hold up the BTS and the Arbiter depends on the hardware configuration of each server.

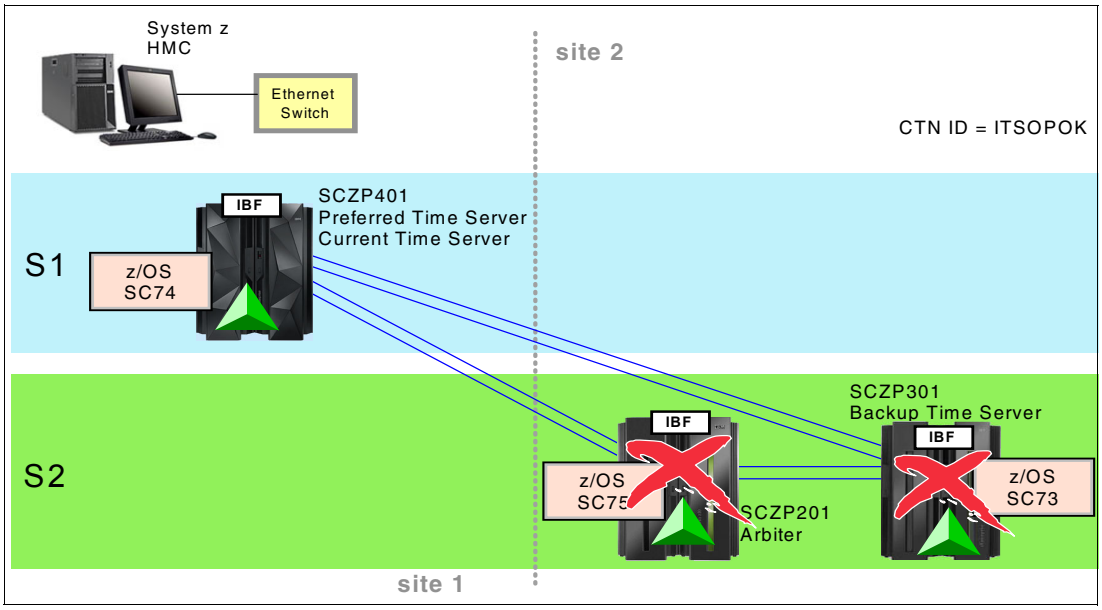

Figure 5-53 STP-only CTN, BTS and Arbiter CTN, site failure with Internal Battery Feature

The key highlights of this scenario are:

- Site 2 is hit by a power outage. Both servers, the BTS and the Arbiter, are maintained by their IBF.
- The BTS and Arbiter send a message through coupling links to the PTS/CTS indicating that they are running on IBF power.
- The PTS/CTS remains as Stratum 1 and disables automated network recovery, because this condition exists for more than 30 seconds.
- The BTS and the Arbiter are also lost after their IBFs run out of power.

5.11.1 Problem awareness

A Hardware Message is posted on the HMC, indicating that there is a power problem at the BTS and the Arbiter. However, there is no impact at this point in time because the BTS and the Arbiter stay on battery power for a certain period of time. After the power outage exceeds this time frame, they both power down and their LPARs are unavailable. Both servers' images, SC73 and SC75, are lost immediately and get partitioned out right after the power outage:

```
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC73 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
IXC101I SYSPLEX PARTITIONING IN PROGRESS FOR SC75 REQUESTED BY XCFAS.
REASON: SYSTEM ENTERED WAIT STATE
```

An active SFM system isolation policy automatically starts system partitioning. This is highlighted by an IXC101I message only a few seconds after detection of the failure. Partitioning triggers the cleanup of resources for the failing system. If system isolation fails, then SFM issues WTOR message IXC102A after the XCF CLEANUP time has elapsed:

```
IXC102A XCF IS WAITING FOR SYSTEM SC73 DEACTIVATION. REPLY DOWN WHEN MVS ON SC73
HAS BEEN SYSTEM RESET
IXC102A XCF IS WAITING FOR SYSTEM SC75 DEACTIVATION. REPLY DOWN WHEN MVS ON SC75
HAS BEEN SYSTEM RESET
```

5.11.2 Problem determination

After the PTS has not received a “normal power” message for more than 30 seconds, it disables automated network recovery and remains at Stratum 1. The network configuration tab (Figure 5-54) shows the following information:

- The PTS continued to play the CTS role.
- Automated network recovery is temporarily disabled.

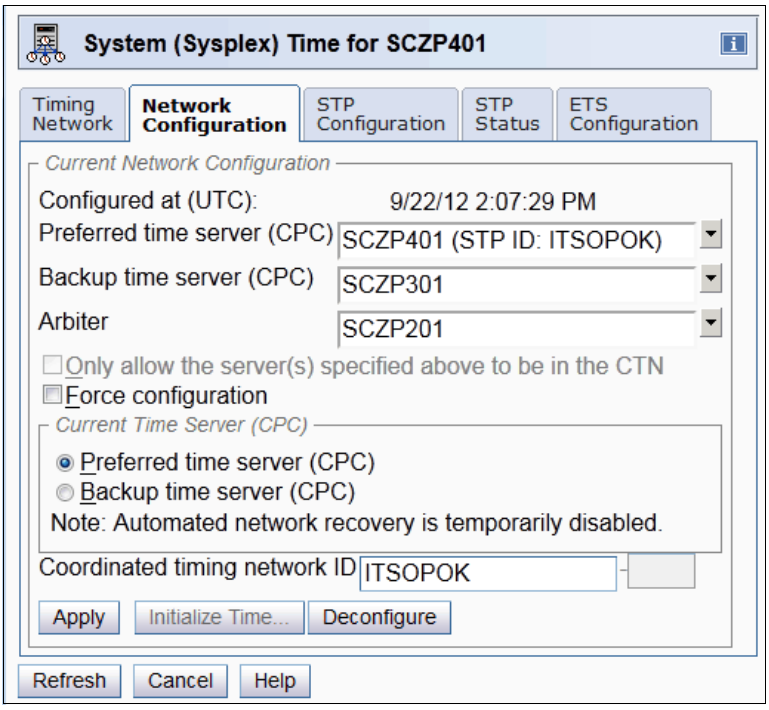

Figure 5-54 Network Configuration tab: SCZP401 30 seconds after the power outage

After the IBFs of SCZP301 and SCZP201 are drained, both servers are powered down. As a consequence, the STP Status tab shows that there is no connectivity to either SCZP301 nor SCZP201, as shown in Figure 5-55 on page 189. In addition, the Network Configuration window still shows both servers with their assigned role (BTS and Arbiter), but the STPID is missing because the CTN lost connectivity to them.

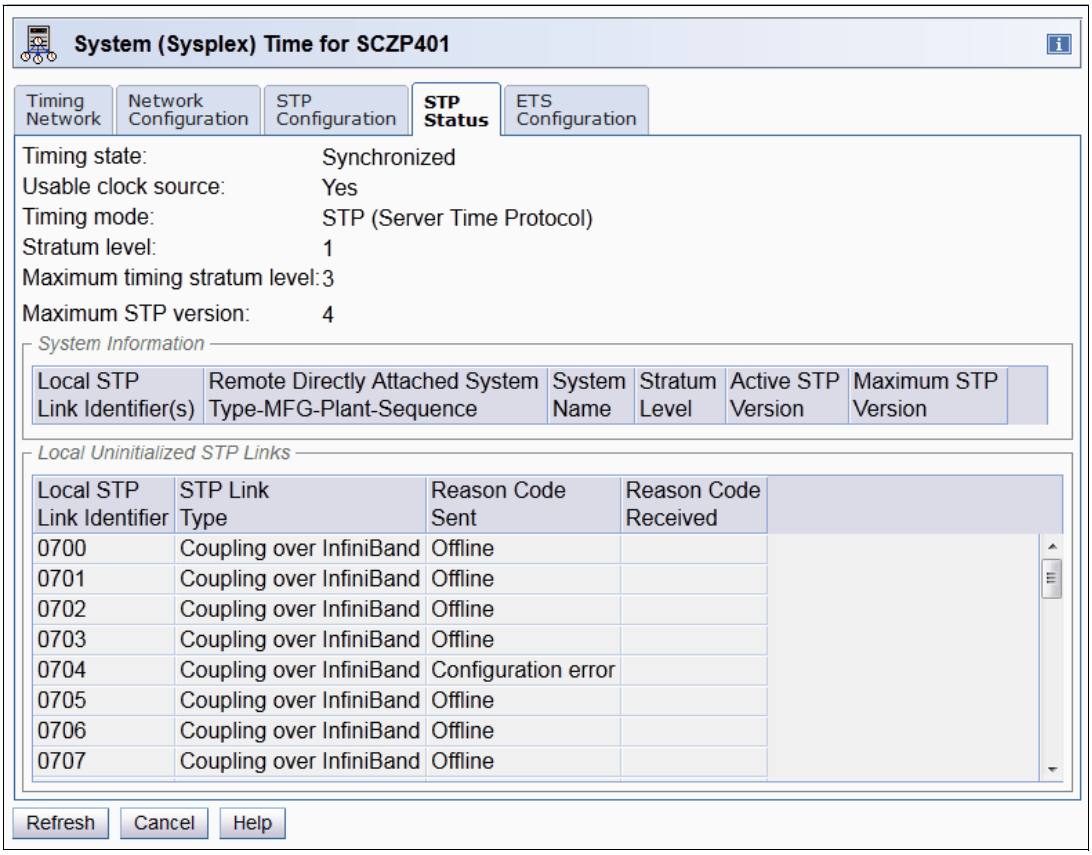

Figure 5-55 STP status tab: SCZP401 after SCZP301 and SCZP201 IBFs are discharged

5.11.3 User actions

No user actions are required for the CTN. The PTS SCZP401 in this scenario maintained the CTS role. SCZP301 and SCZP201 remain unavailable until their power has been restored.

After SCZP301 and SCZP201 have been powered up again, they will join the CTN. Because the STP-only CTN was not reconfigured by reassigning roles during the outage, SCZP301 becomes the BTS and SCZP201 becomes the Arbiter, as they were assigned before the power outage.

Figure 5-56 on page 190 shows the Network configuration tab of SCZP401 after both server SCZP301 and SCZP201 have been recovered.

- Both servers SCPZ301 and SCZP201 are displayed including an STP ID. This indicates server SCZP401 has coupling connectivity to both servers.
- The Current Time Server (CPC) section shows that SCZP401 is still the CTS.
- Automated network recovery is no longer disabled.

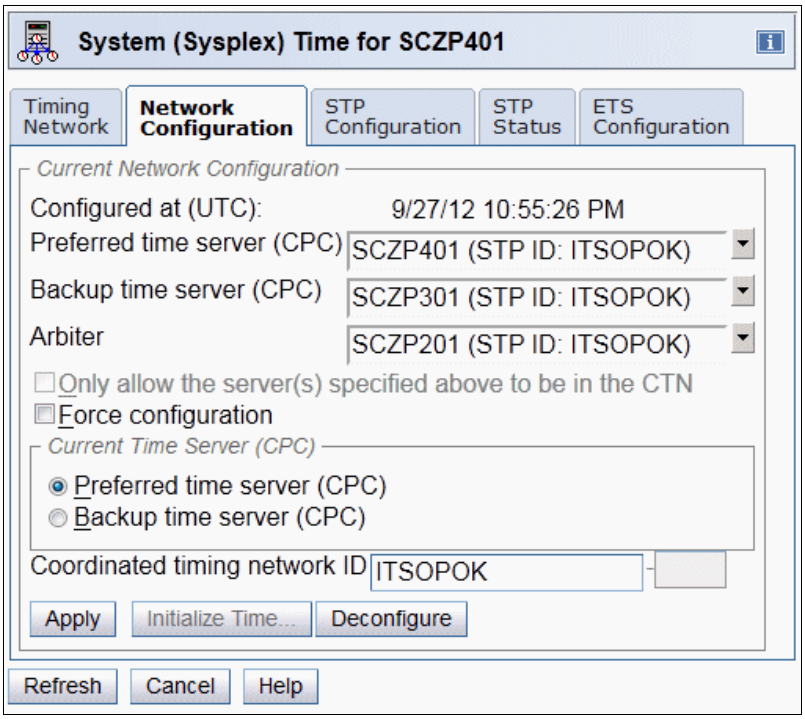

Figure 5-56 Network configuration tab: SCZP401 after recovery

6

External Time Source recovery

In this chapter we describe CTN actions to recover an external time source (ETS). External time source in an STP-only CTN can be provided by one of the following methods:

- Using dial-out on the HMC[1]
- Using an NTP server (LAN connection)[2]
- Using an NTP server with a pulse per second output option (LAN connection and coaxial cable to the PPS port of an ETR card)

We also describe the recovery actions needed to synchronize the CTN to an ETS when the time difference is greater than the 60-second threshold preventing STP from steering to the ETS.

Important: The ETS function provides time accuracy relative to an external time source. Regardless of the ETS option selected, failures associated with ETS do not affect the capability of servers in a CTN to stay synchronized with each other. If the timing state of the servers remains synchronized, z/OS images that depend on synchronization are not affected.

The only effect of unsuccessful recovery for an ETS failure is that the CTN will slowly drift away from ETS time.

Because there are no specific recovery actions when the ETS is configured to use a dial-out service, we only discuss recovery actions in an STP-only CTN using NTP servers or NTP servers with PPS as the external time source.

The recovery process depends on the implemented ETS option and the redundancy configured for that option. The STP design provides continuous availability of ETS, while maintaining the special roles of PTS and BTS you have assigned. If the PTS/CTS is not able to access the time information from its configured NTP servers, STP is able to steer the CTN using the calculated time adjustments from the BTS.

[1] Dial-out support is not available for HMC V2.12.0 or later.
[2] Stating with HMC V2.12.0, HMC NTP authentication support is available.

6.1 ETS recovery using NTP servers

The Current Time Server (CTS) is the only server that adjusts the Coordinated Server Time (CST) by steering it to the time obtained from an external time source. To provide external time source redundancy, consider configuring two or more NTP servers. Up to two NTP servers can be configured on each server in the STP-only CTN. When two NTP servers are configured, you are responsible for selecting the preferred NTP server. This NTP server is called the *selected NTP server*. The other server is called the *non-selected NTP server*.

Configured NTP servers on the PTS/CTS are accessed once every 10 minutes by the Simple Network Time Protocol (SNTP) client. Once every hour, assuming there is a successful access of the selected NTP server, the SNTP client sends a CST adjustment to the STP facility. Normally, the SNTP client on the CTS uses the time information from the selected NTP server to perform the time adjustment. The time information from the non-selected NTP server is only used when there is a failure associated with accessing time information from the selected NTP server.

Configured NTP servers on the BTS are also accessed once every 10 minutes. The BTS calculates a value for time adjustment based on this access, and communicates the information to the PTS over the coupling links. If the PTS/CTS cannot access both its configured NTP servers, it will switch over to using the timing information sent from the BTS to steer the STP-only CTN.

Server access: NTP client code runs on every System z server's Support Element (SE). However, only the NTP client of the PTS/CTS and BTS access the NTP servers configured to them; the NTP clients on the SEs of the other servers in the CTN do not.

Figure 6-1 shows possible failures that can affect recovery when configured with an NTP server. They include:

- Loss of LAN connectivity between the Support Element and the NTP server
- Complete NTP server failure or bad NTP data from the NTP server

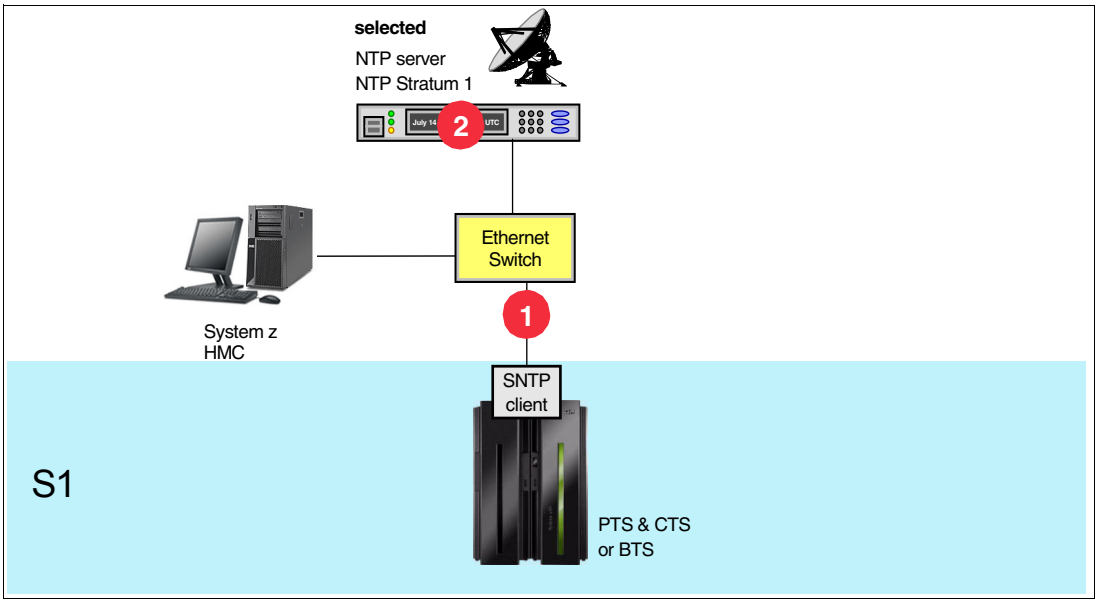

Figure 6-1 Possible failures affecting the NTP server

The following list shows the order in which a recovery action results in selecting a different NTP server than the selected NTP server on the PTS/CTS:

1. When the SNTP client has two unsuccessful attempts (two hours) at sending a CST adjustment value to the STP facility, based on valid NTP data from the selected NTP server, it will switch to sending timing adjustment information based on the non-selected NTP server.
2. When the SNTP client has two unsuccessful attempts (two hours) at sending a CST adjustment value to the STP facility, based on valid NTP data from the non-selected NTP server, STP will steer the CTN using the calculation from the BTS. The BTS information can be based on:
   - Selected NTP server at the BTS, if valid data can be accessed.
   - Non-selected NTP server, if valid data cannot be accessed from the selected NTP server.

When STP is unable to switch to any operational NTP server, the automatic base steering can continue. Base steering allows STP to compensate for the drift characteristics of the oscillator, thereby maintaining relatively good time accuracy at the Current Time Server, even if an ETS is not available.

6.1.1 Redundant NTP servers on the PTS/CTS

The example shown in Figure 6-2 on page 194 depicts a configuration with two NTP servers configured on the PTS/CTS:

- NTP server 1 is an NTP Stratum 2 that has been defined on an HMC and is the selected time source (selected@PTS/CTS).
- NTP server 2 is an NTP Stratum 2 that has been defined on an HMC connected to the same Ethernet switch as the NTP server 1 and the SE, and is the non-selected time source (non-selected@PTS/CTS).

Failure description

The selected NTP server on the PTS/CTS becomes unavailable (see failure type 2 in Figure 6-1 on page 192), as shown in Figure 6-2 on page 194.

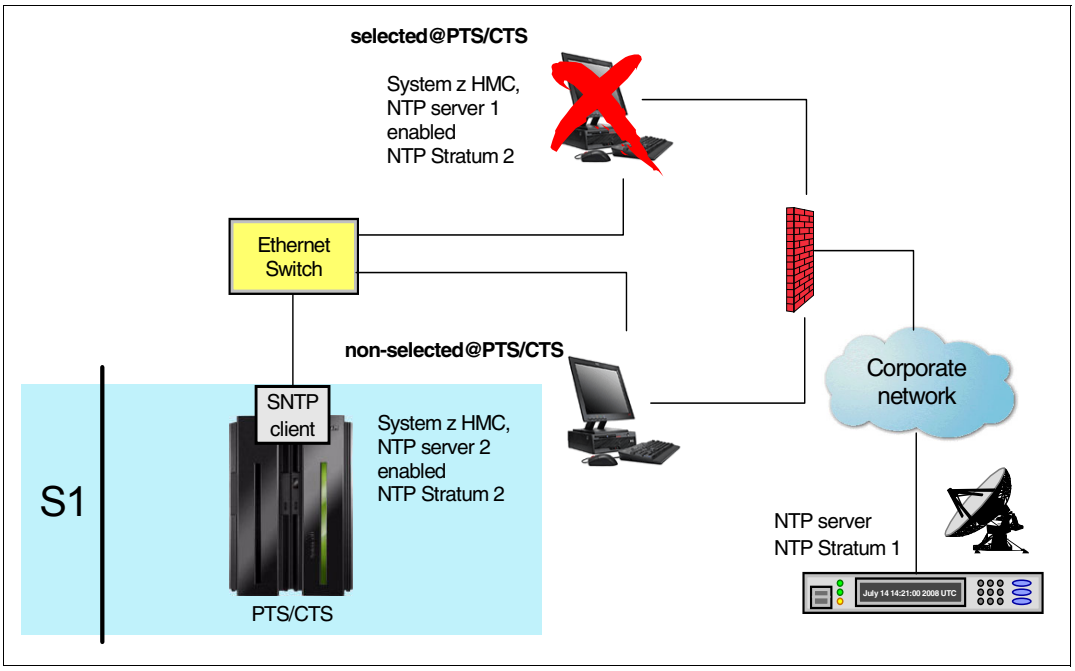

Figure 6-2 NTP server on PTS/CTS failure

Problem awareness

When the SNTP client has two unsuccessful attempts (one to two hours) at sending a CST adjustment value to the STP facility, based on valid NTP data from the selected NTP server, a hardware message is issued for the PTS as shown in Figure 6-3.

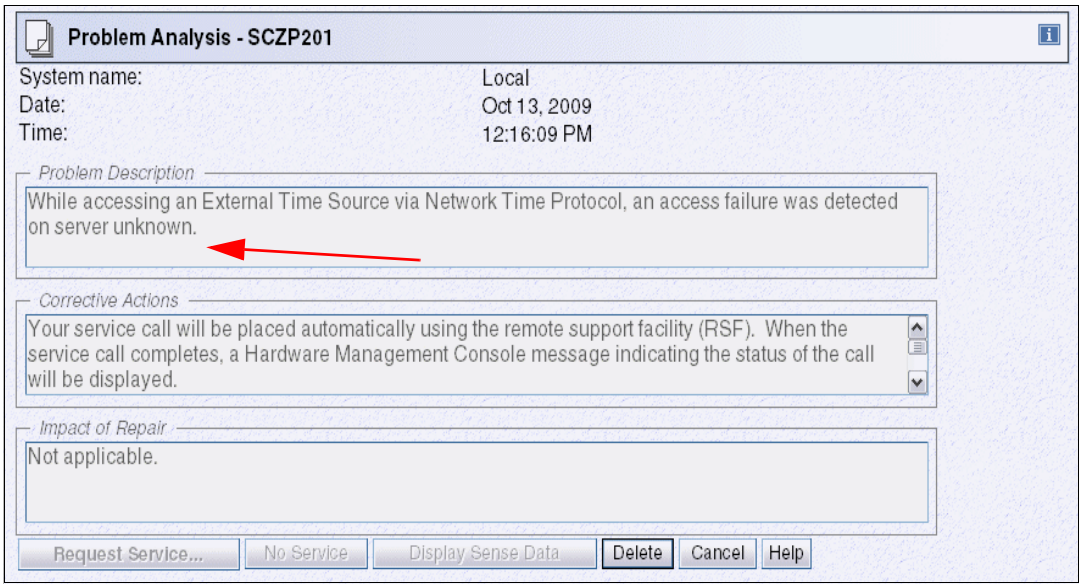

Figure 6-3 Hardware message: access failure

Also, for any STP timing mode z/OS image running at z/OS V1.11 (roll back to z/OS V1.10 and z/OS V1.9) on any server (z9 and above) within CTN, a z/OS message is sent to the MCS console and a hardcopy log (Example 6-1 on page 195).

Example 6-1 Message IEA031I: NTP server switch to non-preferred NTP server

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 0A
```

Problem determination

Within the System (Sysplex) Time window, click the **ETS Configuration** tab for the PTS/CTS. The status for the selected NTP server will display `Timeout failure` (Figure 6-4).

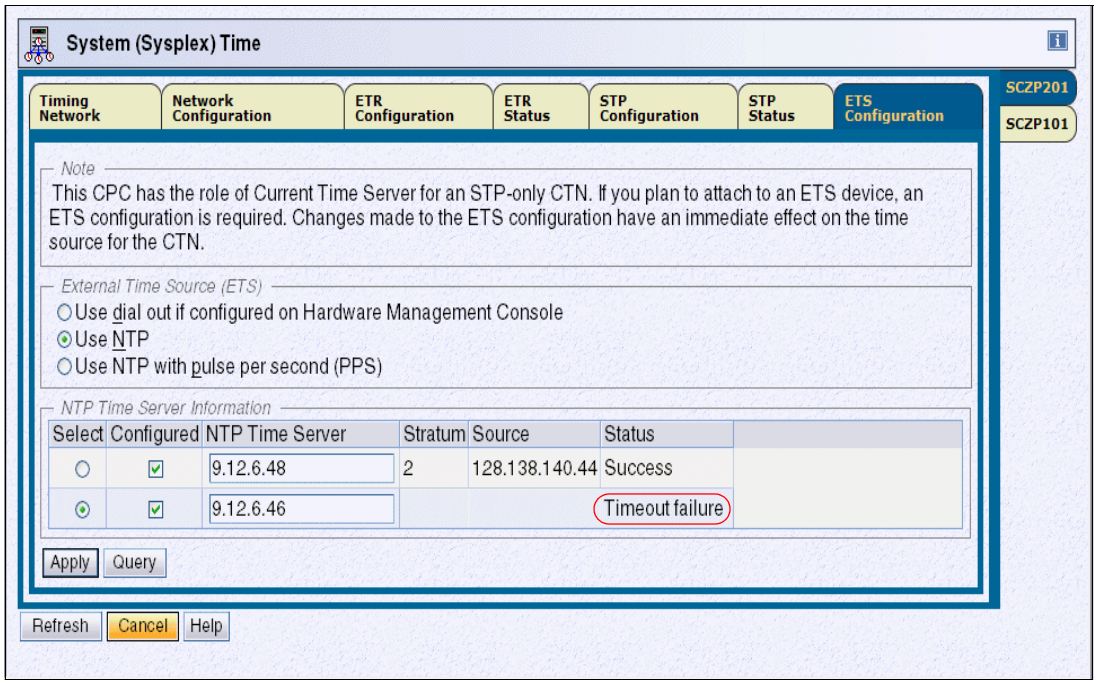

Figure 6-4 ETS configuration tab: Timeout failure

User actions and recovery

Check the NTP server for a failure condition. If it does not have a failure indication, then perform network problem determination.

If the non-selected NTP server is available, after approximately two hours, it is used to steer the CTN.

In this configuration, the network connection between the SE and the Ethernet switch is a single point of failure for both NTP servers. So if the failure is a LAN failure, no recovery is possible and the CTN continues to use automatic base steering.

6.1.2 Redundant NTP servers on PTS and BTS

Figure 6-5 on page 196 shows a configuration with two NTP servers:

- NTP server 1 on the HMC is the selected NTP server on the PTS/CTS.
- NTP server 2 on the HMC is the selected NTP server on the BTS.

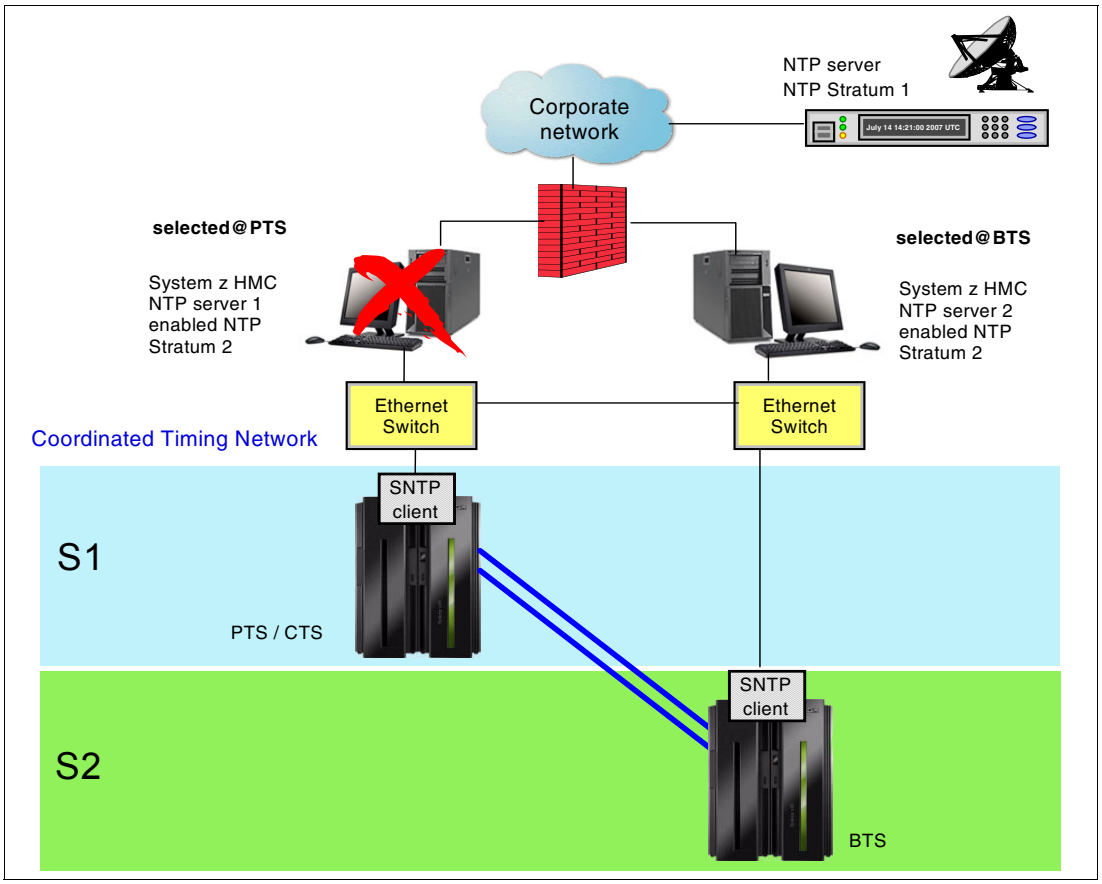

Figure 6-5 Recovery example 2: NTP without PPS

Failure description

The selected NTP server on the PTS/CTS becomes unavailable (see failure type 2 in Figure 6-1 on page 192).

Problem awareness

A hardware message is issued for the PTS after one to two hours (Figure 6-6).

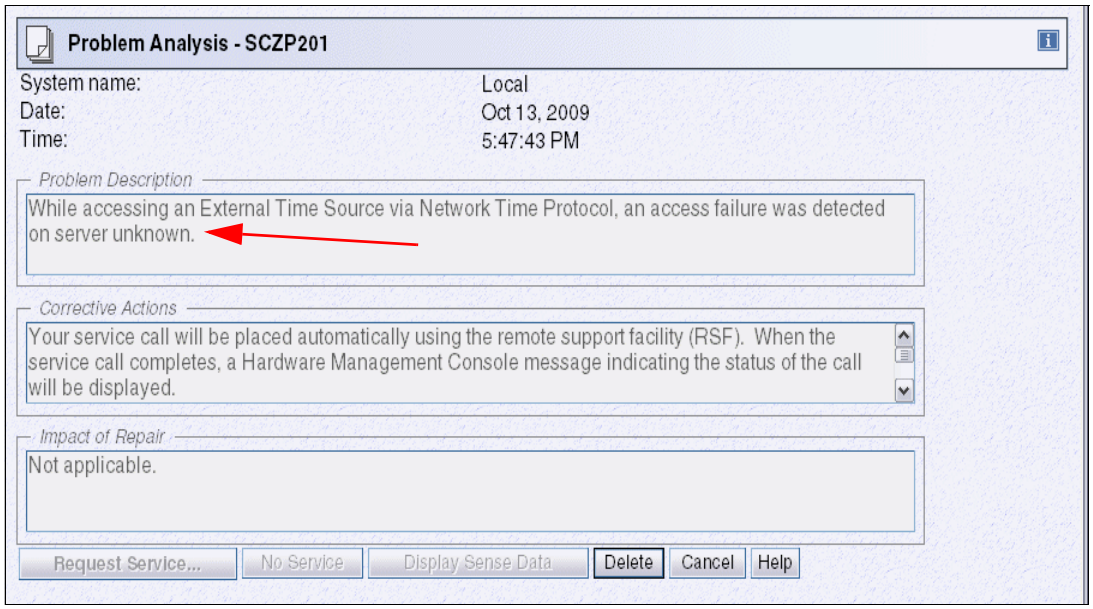

Figure 6-6 Hardware message: access failure

Also, for any STP timing mode z/OS image running at z/OS V1.11 (roll back to z/OS V1.10 and z/OS V1.9) on any server (z9 and above) within CTN, a z/OS message is sent to the MCS console and a hardcopy log (Example 6-2).

Example 6-2 Message IEA031I: NTP server failure

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 06
```

Problem determination

In the System (Sysplex) Time window, click the **ETS Configuration** tab for the PTS/CTS. The status for the selected NTP server will display `Timeout failure` (Figure 6-7 on page 198).

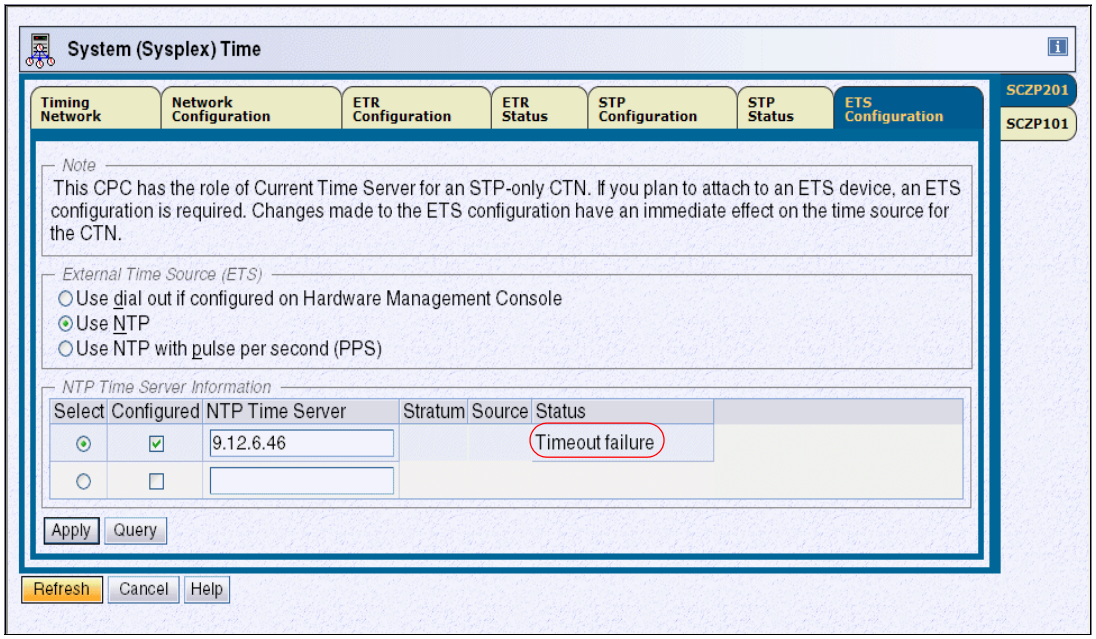

Figure 6-7 ETS configuration tab: Timeout failure

In the System (Sysplex) Time window, click the **Timing Network** tab for the PTS/CTS. The CTN time source will display NTP (Backup Time Server), as shown in Figure 6-8.

Continuous NTP Server Availability state consideration: The Access External Time Source option available in the Adjust Time panel always accesses the ETS configured on the CTS. As a consequence, this option will fail with a timeout message when in the Continuous NTP Server Availability state.

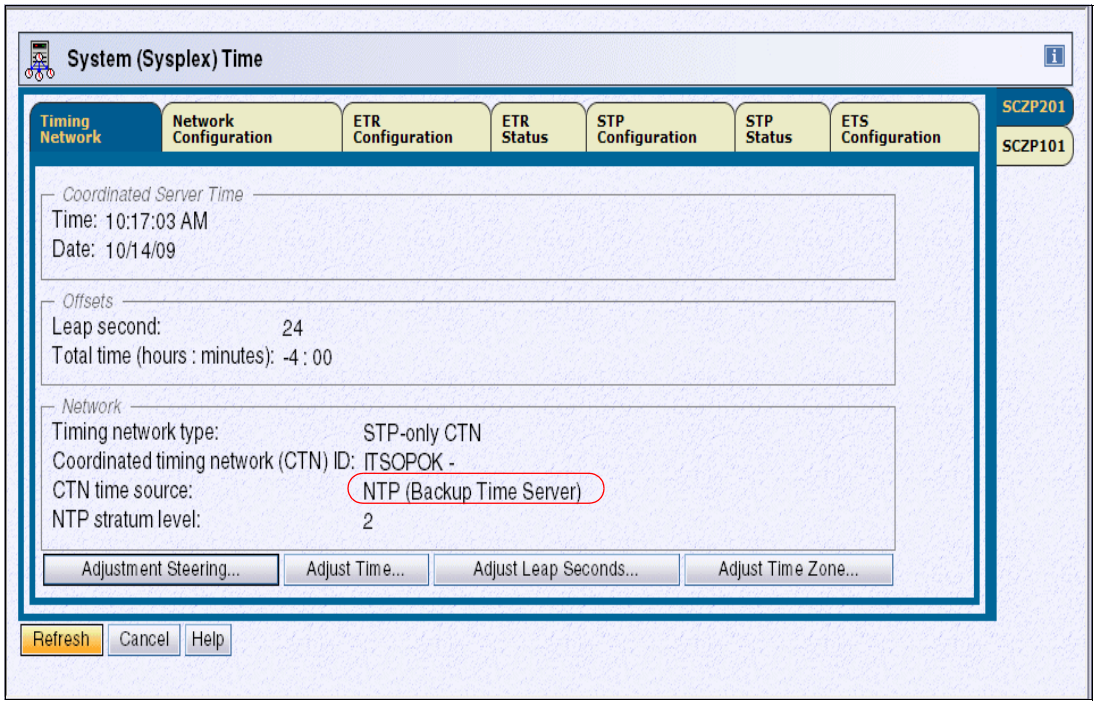

Figure 6-8 Timing Network tab: CTN time source

User actions and recovery

Check the related NTP server for a failure condition. If the NTP server does not have a failure indication, perform network problem determination.

If the PTS/CTS is unable to access NTP server 1 for two hours, it will start using the time adjustment information sent by the BTS approximately an hour later to steer the CTN.

If the BTS is unable to access the NTP server 2 for two hours, then there is no recovery action. However, correct the problem as soon as possible to maintain ETS redundancy.

For additional scenarios, refer to 2.6, “External time source recovery” on page 63.

6.2 ETS recovery using NTP servers with PPS

To provide ETS redundancy, consider configuring two or more NTP servers. Up to two NTP servers with PPS can be configured on each server in the STP-only CTN. When two NTP servers with PPS are configured, you are responsible for selecting the preferred NTP server with PPS. This NTP server with PPS is called the *selected NTP server with PPS*. The other server is called the *non-selected NTP server with PPS*. The PPS output of the same NTP server has to be connected to the PPS input provided on the External Time Reference (ETR) card of the System z10 or System z9 server.

Terminology consideration: For the remainder of this section, an NTP server with PPS might also be referred to simply as an “NTP server” to improve readability.

Do not confuse this with the option discussed in 6.1, “ETS recovery using NTP servers” on page 192, where the NTP server is being used without the PPS option.

Configured NTP servers on the PTS/CTS are accessed once a minute by the SNTP client. Once every 10 minutes, assuming a successful access of both NTP servers, the SNTP client sends time adjustment information based on both NTP servers to the STP facility. Normally, the STP facility on the PTS/CTS uses the time information in conjunction with the PPS signal from the selected NTP server to perform the time adjustment. The time information or PPS signal from the non-selected NTP server is only used when there is a failure associated with accessing time information or PPS signals from the selected NTP server.

Configured NTP servers on the BTS are also accessed once a minute by the SNTP client, and time adjustment information based on both NTP servers is sent to the STP facility on the BTS every 10 minutes. Normally, the STP facility on the BTS uses the time information in conjunction with the PPS signal from the selected NTP server to calculate a time adjustment. The BTS then communicates this information to the PTS over the coupling links. The time information or PPS signal from the non-selected NTP server on the BTS is only used for this calculation when there is a failure associated with accessing time information or PPS signals from the selected NTP server. If the PTS/CTS cannot access both its configured NTP servers, it switches over to using the timing information sent from the BTS to steer the STP-only CTN.

Listed here are several possible failures that can affect an NTP server with PPS:

- Loss of LAN connectivity between the Support Element and the NTP server or bad NTP data; in this case, good PPS signals are still received by the PPS port on the ETR card.
- PPS signal not received by the PPS port on the ETR card; in this case, valid NTP data from the NTP server is still available over the LAN.
- Complete NTP server failure affecting both the NTP data and the PPS output of the NTP server.

Figure 6-9 illustrates these possible failures.

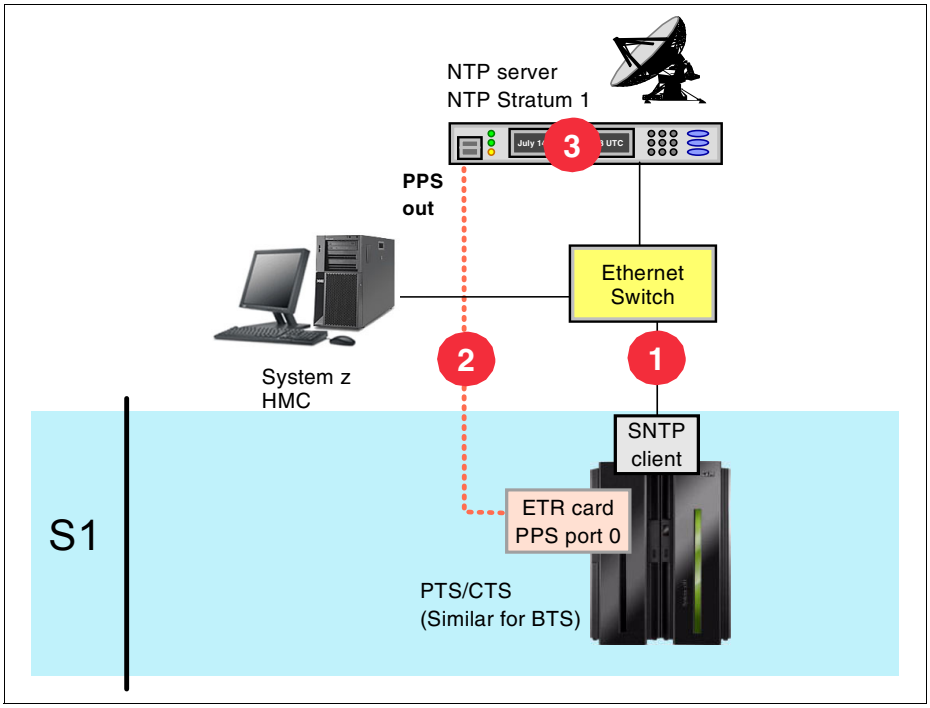

Figure 6-9 Possible failures affecting the NTP server with PPS

Assuming two NTP servers with PPS are configured on both the PTS/CTS and the BTS, and that STP is tracking the PPS signals received by the PPS port associated with the selected NTP server on the PTS/CTS, then note the following outcomes:

- If there is a failure type (1), STP continues using the PPS signals received on the PPS port of the selected NTP server on the PTS/CTS.
- If there is a failure type (2) or (3), STP switches to using NTP data and PPS signals from the non-selected NTP server on the PTS/CTS.

Assuming only one NTP server with PPS is configured on the PTS/CTS and the BTS, and that STP is tracking the PPS signals received by the PPS port associated with the selected NTP server on the PTS/CTS, then note the following outcomes:

- If there is a failure type (1), STP continues using the PPS signals received on the PPS port of the selected NTP server on the PTS/CTS.
- If there is a failure type (2) or (3), STP switches to using the time adjustment information received from the BTS. The time adjustment information from the BTS could be based on the selected NTP server at the BTS, if valid data and PPS signals can be accessed.

Regardless of the specific redundancy provided by an NTP server with PPS configuration, note the following points:

- If PPS signals are not received by any of the configured NTP servers on the PTS/CTS and the BTS, but valid NTP data is available, then STP continues using the NTP data for steering the CTN, following the same recovery flow described in 6.1, “ETS recovery using NTP servers” on page 192.
- When STP is unable to switch to any operational NTP server, the automatic base steering can continue. Base steering allows STP to compensate for the drift characteristics of the oscillator, thereby maintaining relatively good time accuracy at the Current Time Server, even if an ETS is not available.

6.2.1 Redundant NTP servers with PPS on PTS/CTS

The diagram in Figure 6-10 on page 202 shows a configuration with two NTP servers with a PPS output that have been configured on the PTS/CTS. One NTP server is the selected NTP server and its PPS cable has been connected to PPS port 0 at the ETR card. The second NTP server is the non-selected NTP server and its PPS cable has been connected to PPS port 1 at the ETR card.

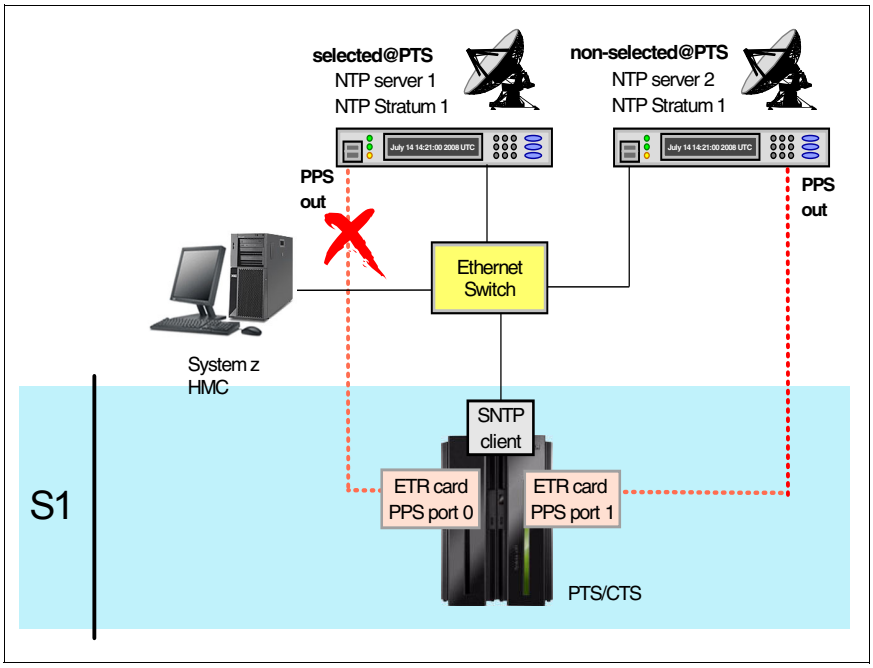

Figure 6-10 Recovery example 1: NTP with PPS

Failure description

The PPS signal of the selected NTP server on the PTS/CTS becomes unavailable.

Problem awareness

A hardware message (shown in Figure 6-11) is issued for the PTS after two minutes.

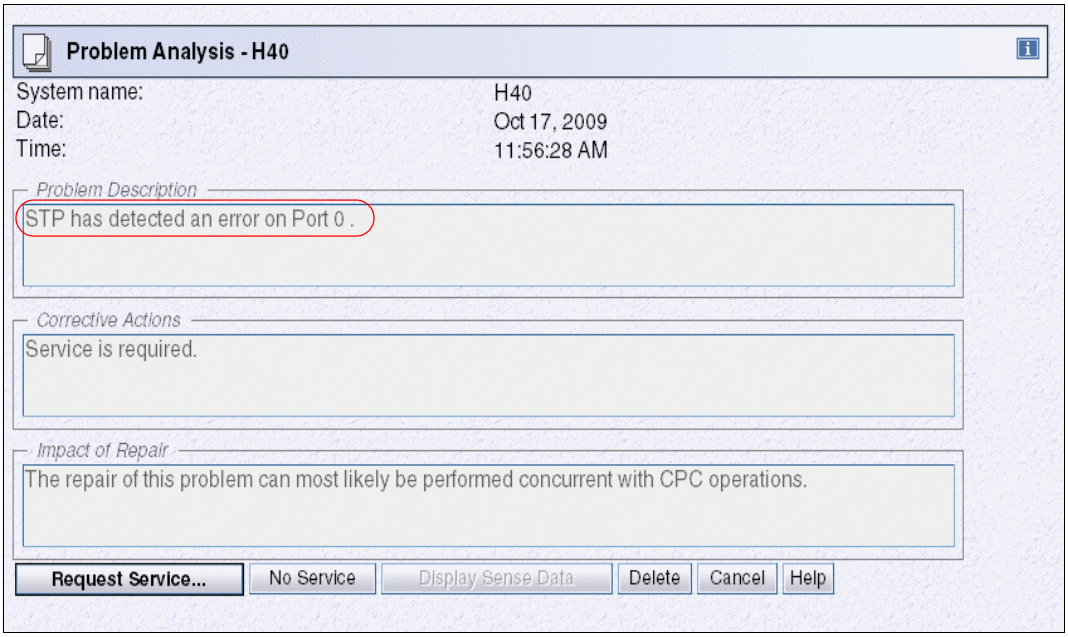

Figure 6-11 Hardware message: PPS signal missing

For any STP timing mode z/OS image running at z/OS V1.11 (roll back to z/OS V1.10 and z/OS V1.9) on any server (z9 and above) within the CTN, when PPS signals of the selected NTP server on PTS/CTS become unavailable, z/OS messages indicating the PPS status

change are sent to the MCS console and to a hardcopy log. Example 6-3 shows the message when the NTP server providing the PPS signal has been changed.

Example 6-3 Message IEA031I: PRT source ID change

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 83
```

Example 6-4 shows the message when the PPS signal is not available at all.

Example 6-4 Message IEA031I: No PPS signal

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 84
```

After two minutes, another z/OS message is sent to indicate that the PPS signal from the non-preferred NTP server of PTS/CTS is now used by STP, as shown in Example 6-5.

Example 6-5 Message IEA031I: Switch to non-preferred Pulse Per Second (PPS) port

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 81
```

Problem determination

In the System (Sysplex) Time window, click the **ETS Configuration** tab for the PTS/CTS. The port status for PPS port 0 will display `No PPS signal` and PPS port 1 will display `Tracking to PPS signal` (Figure 6-12).

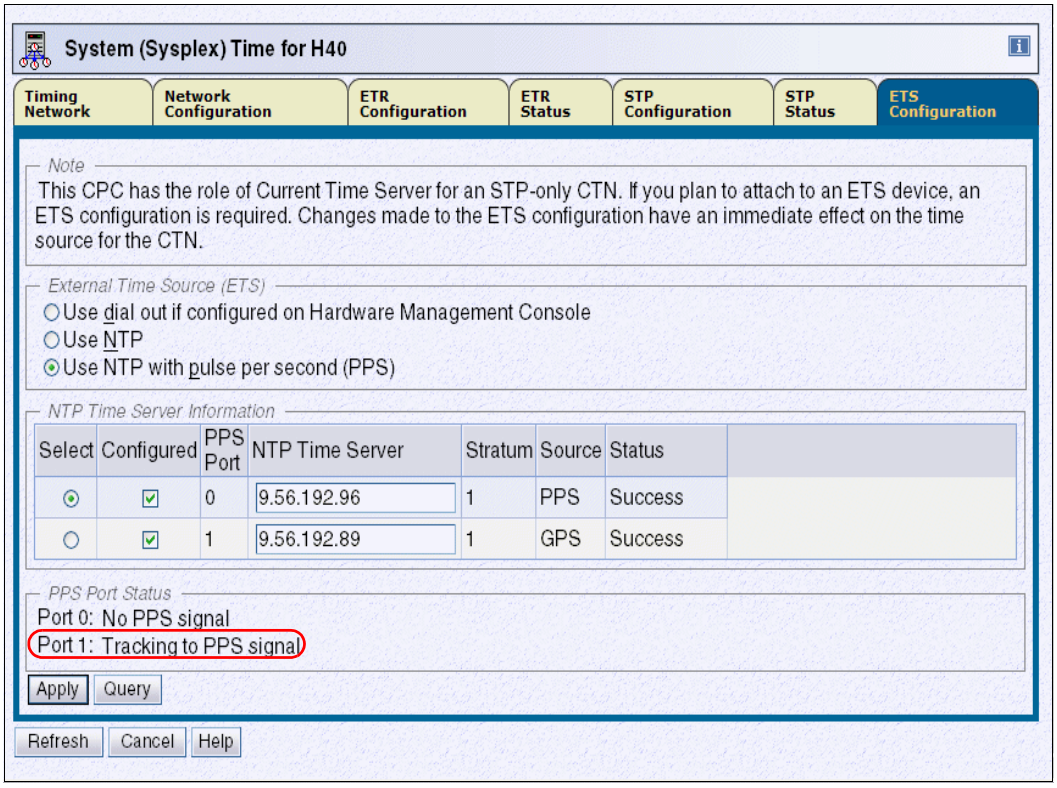

Figure 6-12 ETS configuration tab: No PPS signal

You can check the Console Event log. To view the console log from the HMC application, first select the PTS/CTS server and then follow these steps:

1. Select **Recovery** → **Single Object Operations** to log on to the Support Element.

2. From the SE workplace, select **Service Management** → **View Console Events**. Table 6-1 shows the events that have occurred.

Table 6-1 View Console Events

Date	Time	Console Event
10/17/2009	11:54:28.180	PPS port 0 is not receiving pulse per second (PPS) signals.
10/17/2009	11:54:28.140	Pulse per second (PPS) signals from PPS port 1 are being used.

User actions and recovery

Check the related NTP server for a failure condition. If the NTP server does not have a failure indication, check PPS connectivity.

If NTP server 1 is not accessible by the SNTP client on the SE (failure 1 in Figure 6-9 on page 200), but the PPS signal is still received on PPS port 0, no recovery is required because STP will continue to steer the CTN using the PPS signals from NTP server 1.

For failures 2 and 3 on NTP server 1, STP will switch to using the time information and the PPS signals from the non-selected server, NTP server 2.

Important: In normal operation, if the PTS/CTS continues to receive the PPS signal from the selected PPS port of the NTP server, there is no recovery action, even if the SE does not receive valid NTP data from the selected NTP server over the LAN.

6.2.2 Redundant NTP servers with PPS on PTS and BTS

The diagram in Figure 6-13 on page 205 presents a configuration where an NTP server has been defined and selected on the PTS/CTS. The appendant PPS cable has been connected to the ETR card PPS port 0 at the PTS/CTS. Another NTP server has been defined and selected on the BTS. The appendant PPS cable has been connected to PPS port 0 at the ETR card of the BTS.

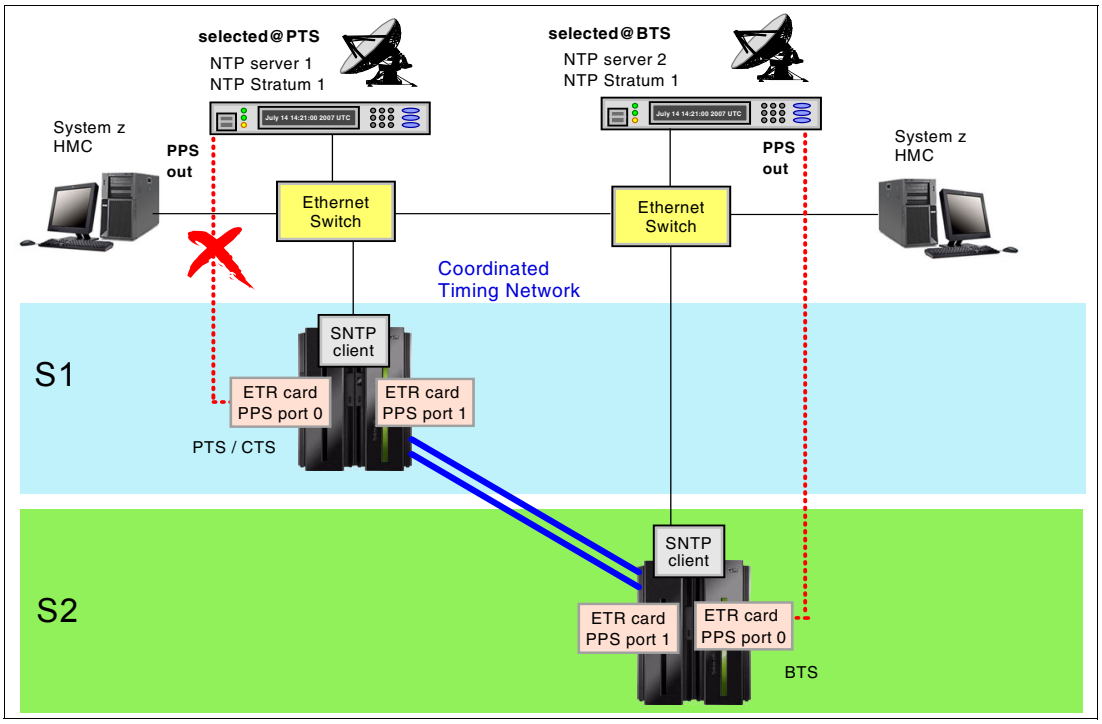

Figure 6-13 Recovery example 2: NTP with PPS

Failure description

The PPS signal of the selected NTP server on the PTS/CTS becomes unavailable.

Problem awareness

A hardware message (shown in Figure 6-14 on page 206) is issued for the PTS after two minutes.

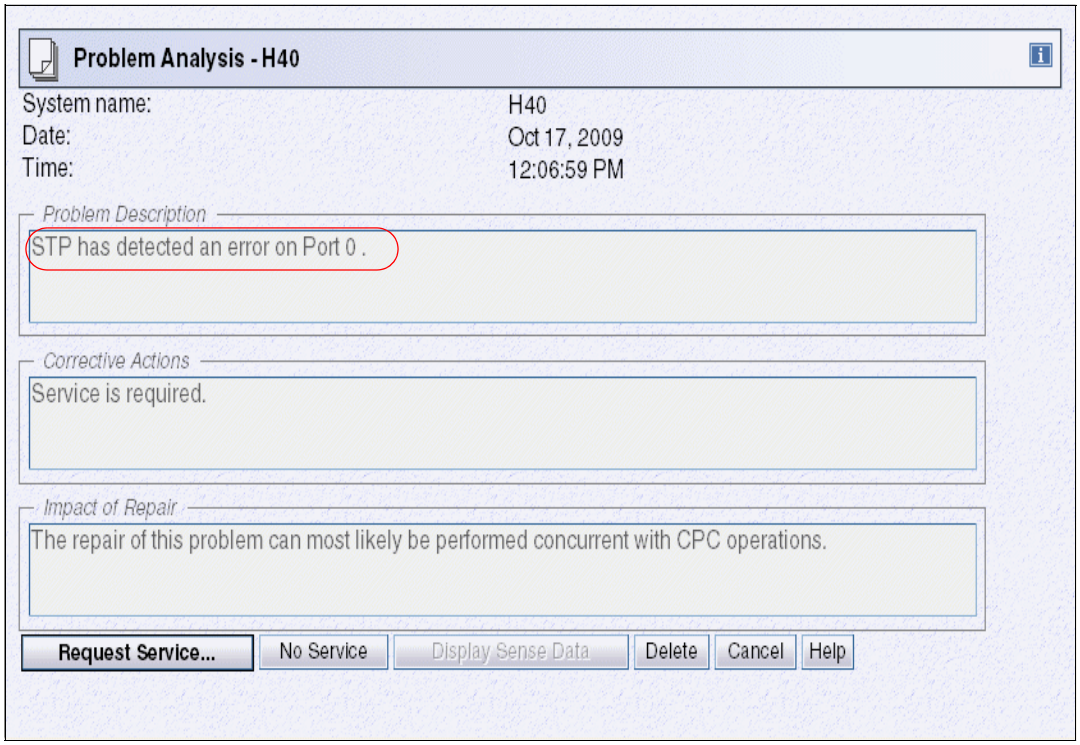

Figure 6-14 Hardware message: PPS signal missing

For any STP timing mode z/OS image running at z/OS V1.11 (roll back to z/OS V1.10 and z/OS V1.9) on any server (z9 and above) within CTN, when the PPS signal of the selected NTP server on PTS/CTS becomes unavailable, z/OS messages indicating a PPS status change are sent to the MCS console and to a hardcopy log.

Example 6-6 shows the message when the NTP server providing the PPS signal has been changed.

Example 6-6 Message IEA031I: PRT source ID change

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 83
```

Example 6-7 shows the message that signals that the original PPS source has been restored.

Example 6-7 Message IEA031I: Switch to preferred PPS port

```
IEA031I STP ALERT RECEIVED.  STP ALERT CODE = 82
```

The alert shown in Figure 6-15 on page 207 is sent by BTS to indicate that the PPS signal tracking is switched from the selected PPS port of PTS/CTS to the selected PPS port of BTS.

Problem determination

In the System (Sysplex) Time window, click the **ETS Configuration** tab for the PTS/CTS. The port status for PPS port 0 will display `No PPS signal` (Figure 6-15 on page 207).

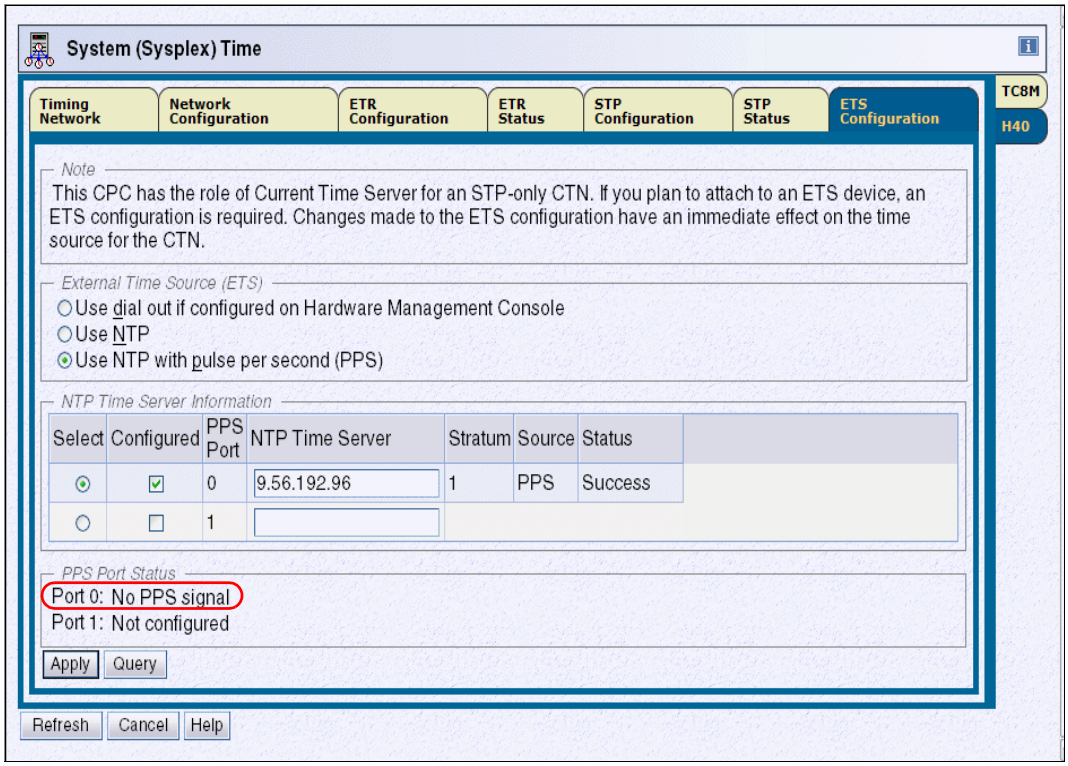

Figure 6-15 ETS configuration tab for PTS: No PPS signal

In the System (Sysplex) Time window, click the **ETS Configuration** tab for the BTS. The port status for PPS port 0 will display `Tracking to PPS signal` (Figure 6-16 on page 208).

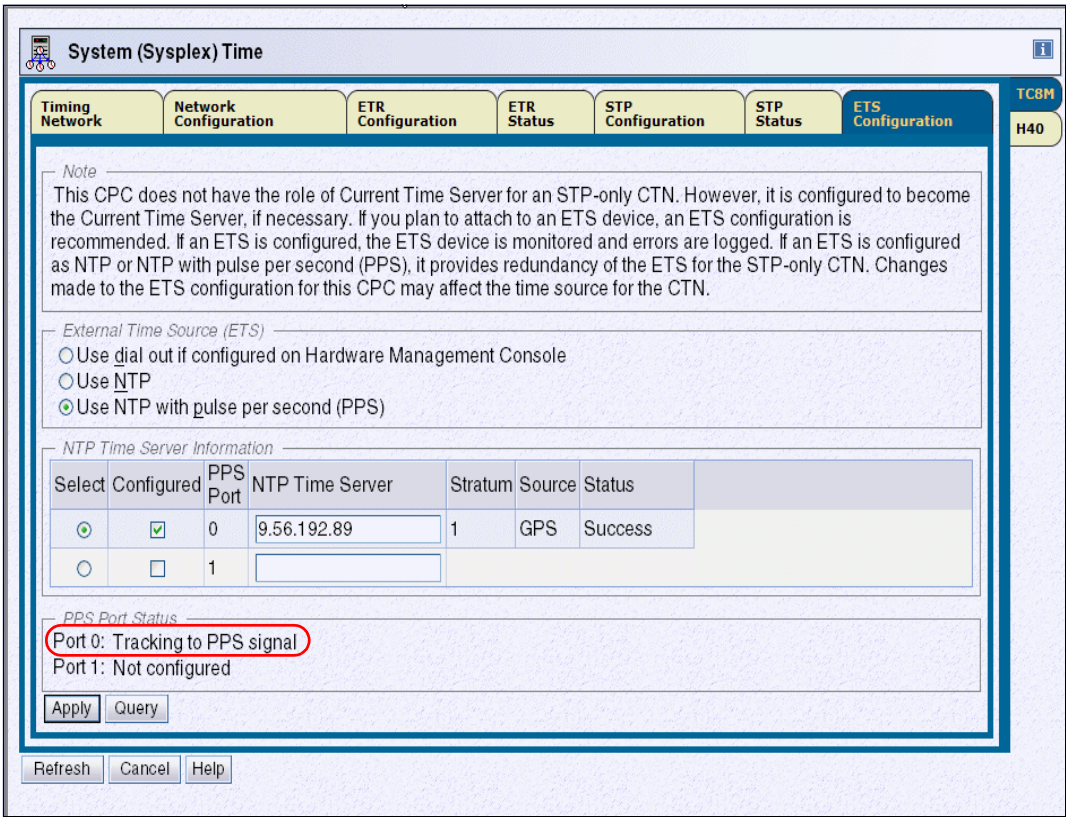

Figure 6-16 ETS configuration tab for BTS: Tracking to PPS signal

In the System (Sysplex) Time window, click the **Timing Network** tab for the PTS/CTS; the CTN time source will display `NTP with pulse per second (Backup Time Server)`, as shown in Figure 6-17 on page 209).

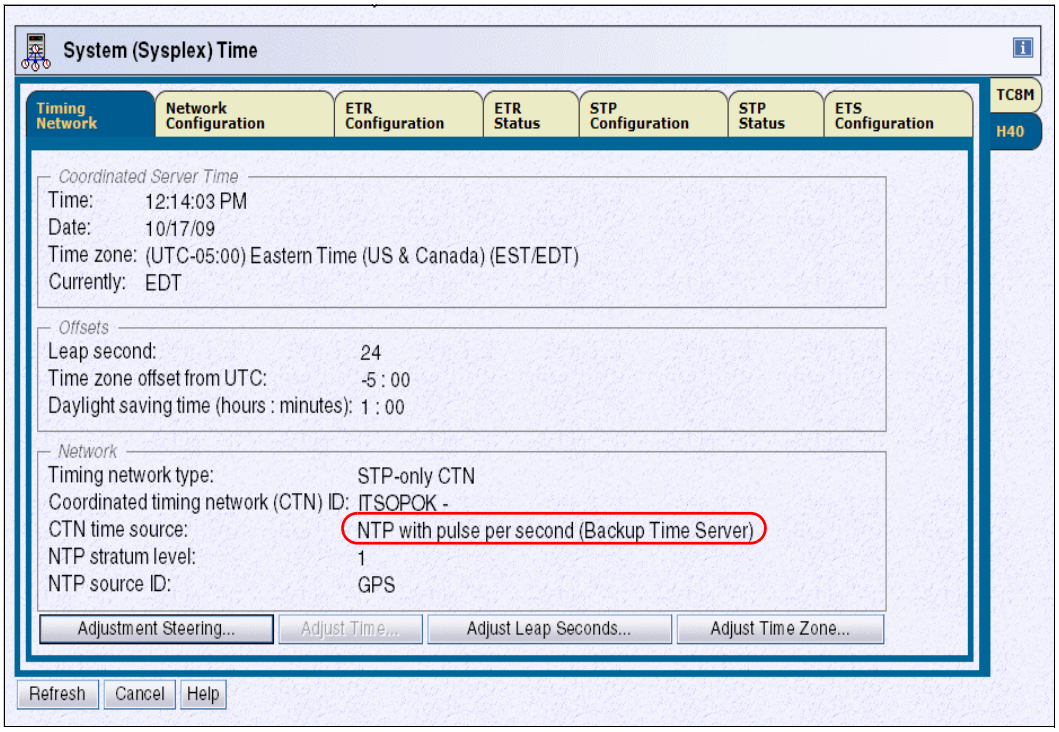

Figure 6-17 Timing Network tab: NTP with pulse per second (Backup Time Server)

You can check the Console Event log. To view the console log from the HMC application, first select the PTS/CTS server and then follow these steps:

1. Select **Recovery** → **Single Object Operations** to log on to the Support Element.
2. From the SE workplace, select **Service Management** → **View Console Events**.

Table 6-2 View Console Events

Date	Time	Console Event
10/17/2009	12:04:59.590	PPS port 0 is not receiving pulse per second (PPS) signals.
10/17/2009	12:04:59.280	Adjustments to Coordinated Server Time for the STP-only CTN are being made using information from the backup system.

User actions and recovery

Check the related NTP server for a failure condition. If the NTP server does not have a failure indication, check PPS connectivity.

If NTP server 1 is not accessible (failure 1 in Figure 6-9 on page 200), but the PPS signal is still received on PPS port 0, no recovery is required because STP will continue to steer the CTN using the PPS signals from NTP server 1.

For failure scenarios 2 and 3 on NTP server 1, the PTS/CTS will start using the time adjustment information received from the BTS, which is based on NTP server 2 and its PPS signals.

Important: In normal operation, if the PTS/CTS continues to receive the PPS signal from the PPS port of the selected NTP server, there is no recovery action, even if the SE does not receive valid NTP data from the selected NTP server over the LAN.

For failure scenarios 1, 2, or 3 on NTP server 2, the relevant hardware message and z/OS alerts will be sent, but no recovery is required,

6.3 Synchronizing the CTN to an ETS when time difference is greater than 60-second threshold

When an NTP server has been configured, STP will steer Coordinated Server Time (CST) to the external time source (ETS) accessed by the NTP server providing the time difference is not greater than 60 seconds. If the 60-second threshold is exceeded, the status of the NTP time server on the ETS configuration panel is `CPC/NTP time difference > 60 seconds`.

This condition can occur when an NTP server is configured for the first time for an existing CTN. It is discussed in the *Server Time Protocol Implementation Guide,* SG24-7281, in 5.6.1 “Adding an ETS to an existing CTN.” However, it can also occur if there is an NTP failure or connectivity problem for an extended period of time during which the CST drifts away from the ETS time.

Important: Problems in accessing an ETS have no effect on the STP-only CTN other than that the STP UTC time is no longer steered to the ETS. The only effect is that the STP UTC time is no longer accurate. The condition will be alerted through message IEA031I with the appropriate alert code.

Similarly, resolving an NTP server problem so that STP begins steering to an ETS is transparent to z/OS, other than the IEA031I messages indicating the change of NTP server status.

There are two preferred procedures to correct CST to the NTP server time. Which procedure to use depends on whether a sysplex-wide outage can be scheduled, and the time frame within which the STP UTC time must be corrected.

There is also a third procedure that can be used to correct the STP UTC time immediately without any disruption to the CTN through the use of leap seconds.

The three procedures are:

- Using STP time adjustments to get CST within 60 seconds and then allowing steering to synchronize to the ETS
- Initializing CST to the correct time
- Applying leap seconds to correct the time difference

These procedures are explained in more detail in the following sections.

Using STP time adjustments

This is the suggested approach because it does not require any disruption to the CTN and is transparent to z/OS. Time adjustments of +/- 60 seconds can be made using the “Adjust Time...” button on the Timing Network panel through the CTS server.

Time adjustments in 60-second increments can be made until STP UTC time is within 60 seconds of the ETS. When the time difference is 60 seconds or less, STP will begin steering to the ETS. The steering rate is 7 hours for each second of adjustment. When a time adjustment is made, the expected duration to steer the adjustment is displayed.

After CST is within 60 seconds of the ETS, steering will commence and the “Adjustment Steering” button on the Timing Network panel displays the expected duration to steer to the ETS.

Initializing CST to the correct time

If the expected duration for time adjustments and steering cannot be tolerated, then the STP UTC time can be corrected by deconfiguring the CTN and initializing with the correct time. This is disruptive to the CTN and requires a sysplex-wide outage.

Follow these steps to initialize CST to the correct time:

1. Shut down all z/OS and CF LPARs that are using STP timing mode and are in a MULTISYSTEM sysplex. These are systems that specify STPMODE YES in CLOCKxx and PLEXCFG=MULTISYSTEM in IEASYSxx.

   Systems that are running with LOCAL timing mode, or that specify SIMETRID in CLOCKxx, or that specify PLEXCFG=XCFLOCAL or MONOPLEX in IEASYSxx, are not affected.

2. Deconfigure the CTN through the “Deconfigure” button on the Timing Network panel through the CTS server.

3. Initialize the time and configure the CTN as described in *Server Time Protocol Implementation Guide,* SG24-7281, in 5.4.4 “Time initialization, 5.4.5 Completing time initialization” and 5.4.6 “Defining server roles.”

Note: You can choose not to shut down z/OS and CF LPARs that are using STP timing mode and are in a MULTISYSTEM sysplex. The z/OS systems will issue the loss of time source synchronous WTOR IEA394A when the CTN is deconfigured. Then, when the CTN is reconfigured, you can reply RETRY to IEA394A.

For more details about responding to IEA394A, see 1.1.9, “Switch to local timing mode” on page 11 and 1.2, “Sysplex Failure Management considerations” on page 14.

Applying leap seconds to correct the time difference

If you cannot take a sysplex-wide outage and have a requirement to correct the time difference immediately, then leap seconds can be used. For details about the use of leap seconds, refer to white paper 102081 for more information:

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102081

Important: Although leap seconds are being used here for a different purpose than intended, the use of leap seconds is fully supported.

However, use this procedure under the direction of IBM Support.

As an ETS supplies UTC time, STP steers CST to synchronize STP UTC time with the ETS. UTC time is calculated by adjusting CST by the number of leap seconds specified through the following equation:

```
STP UTC time = CST time - leap seconds
```

Making a leap second change affects STP UTC time but does not affect CST time. A negative leap second adjustment results in a forward jump in UTC time. A positive leap second adjustment results in a backward jump in UTC time.

- When a negative leap second adjustment is made, z/OS processes the change with immediate effect, resulting in a forward jump in UTC time by the amount of the leap second adjustment.
- When a positive leap second adjustment is made, z/OS must spin for the amount of the leap second adjustment to prevent duplicate UTC time stamps.

Important: To prevent duplicate UTC time stamps, make positive leap second adjustments in 2-second increments to avoid system impact due to the spin.

After the leap second adjustment or adjustments have been made and the STP UTC time is synchronized to the ETS, the leap seconds can be removed. This can be performed nondisruptively, assuming a 1-second difference between STP UTC time and the true UTC time reflected by the NTP server can be tolerated.

The leap second adjustment can be changed in 1-second increments. Every time a 1-second change is made to leap seconds, STP UTC time will be 1-second different to the NTP server time.

STP will take 7 hours to steer the 1-second adjustment. When steering completes for the 1-second adjustment, then the next 1-second adjustment can be made until the leap seconds have been removed.

Note: For every leap second adjustment, the following message will be issued on all z/OS systems:

```
IEA392I STP TIME OFFSET CHANGES HAVE OCCURRED.
```

Example steps in correcting a time difference between CST and the ETS time

In this example the STP UTC time is ahead of the ETS, resulting in a status of `CPC/NTP time difference > 60 seconds` for the NTP Time Server on the ETS panel (Figure 6-18).

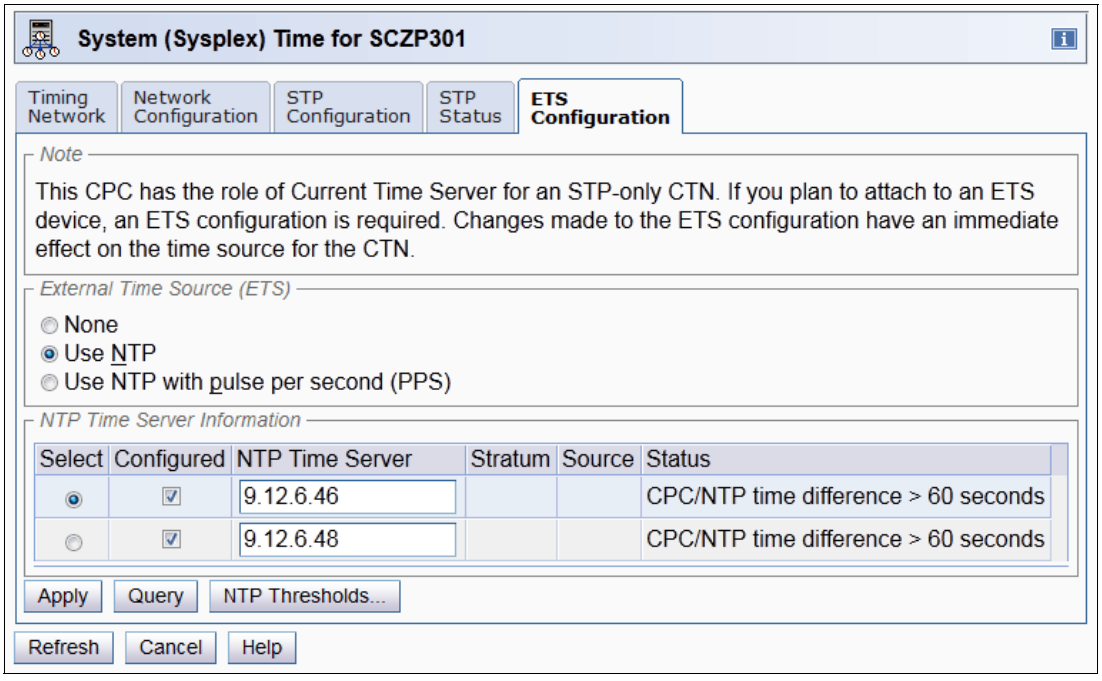

Figure 6-18 NTP Time Server status

To determine the actual time difference between STP UTC time and the ETS accessed by the NTP Time Server, select the “Adjust Time” button on the Network Timing panel and then the “Access External Time Source” button (Figure 6-19).

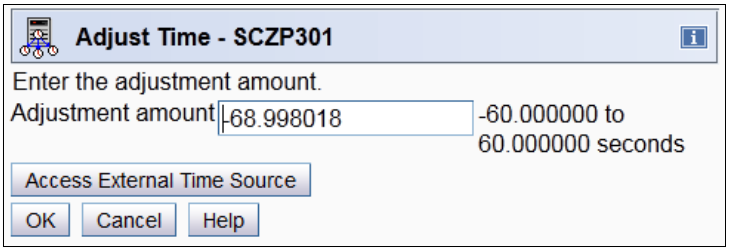

Figure 6-19 Adjust Time through Access External Time Source

This shows that the difference between STP UTC time and the ETS is -68.998018 seconds. A backward change in UTC time needs to be made through a positive leap second adjustment of 69 seconds. The positive leap second adjustment is made in 2-second increments (Figure 6-20).

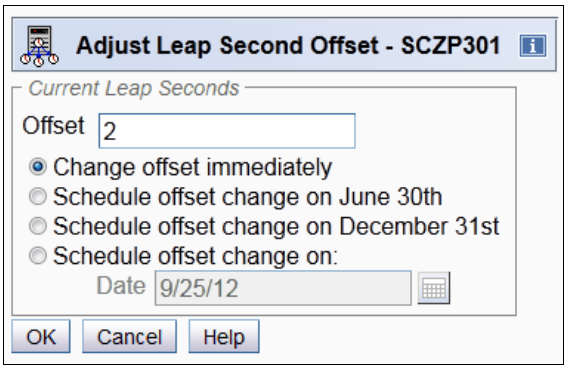

Figure 6-20 Adjust Leap Second Offset

After adjusting leap seconds to 10 through 2-second increments, the status of the NTP Time Server is `Success`, as shown in Figure 6-21 and Figure 6-22.

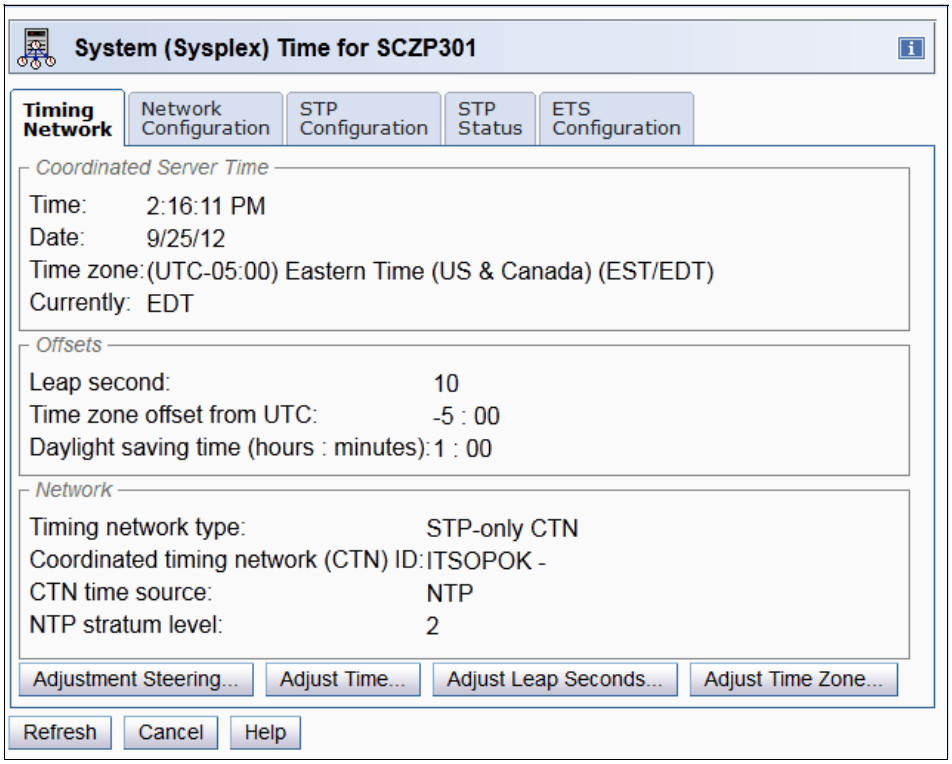

Figure 6-21 Timing Network panel

Figure 6-22 shows the status.

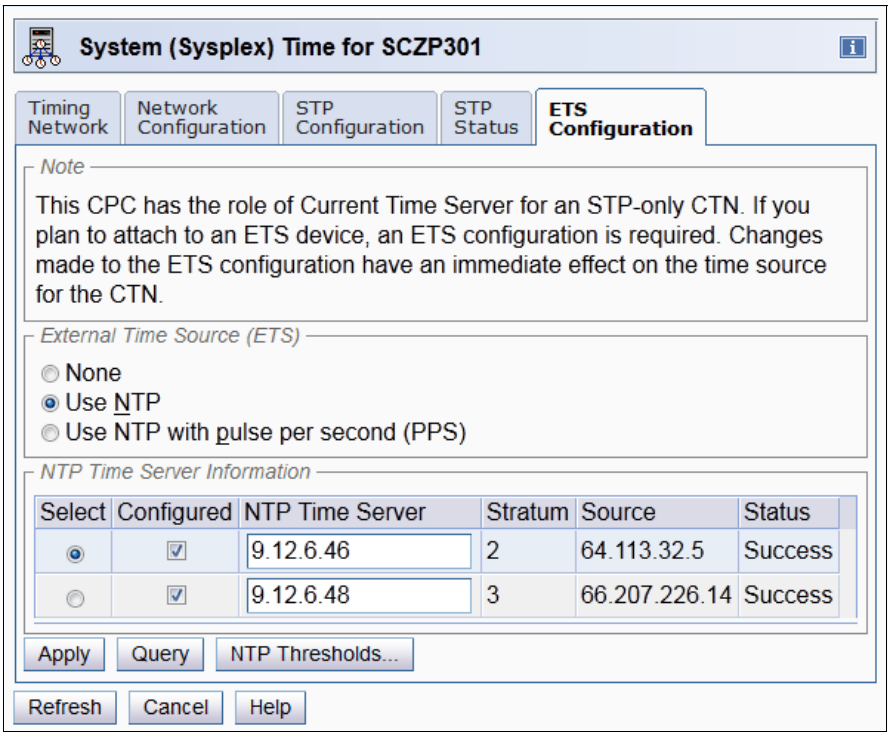

Figure 6-22 NTP Time Server status

After adjusting leap seconds to 69 in 2-second increments, STP UTC time is synchronized to the ETS, as shown in Figure 6-23 and Figure 6-24.

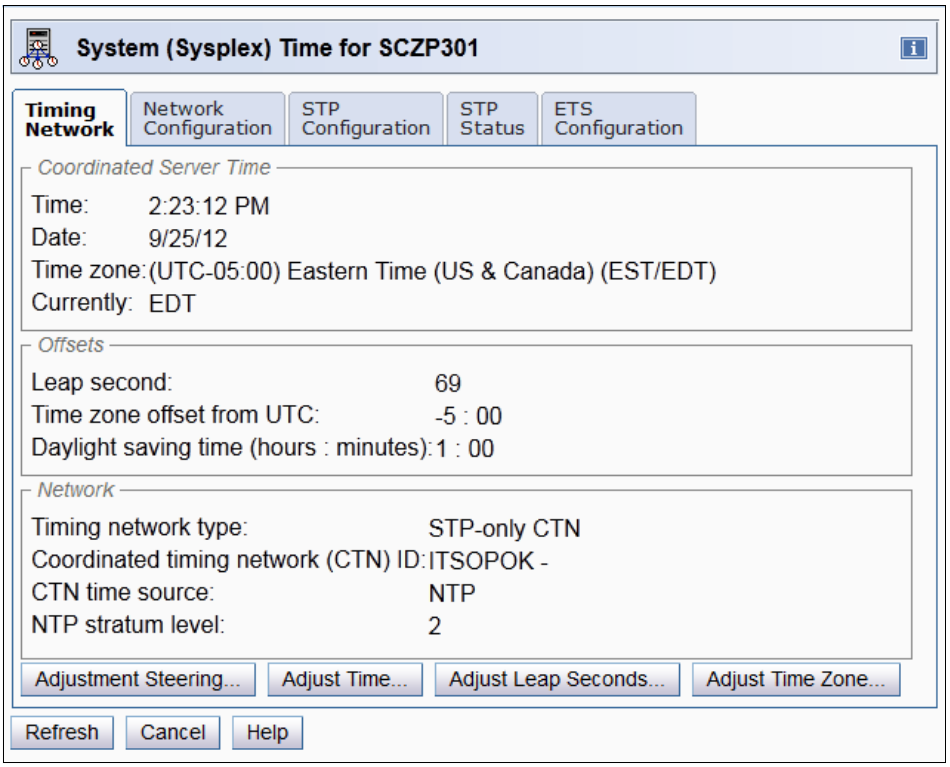

Figure 6-23 Timing Network panel

Figure 6-24 shows the status.

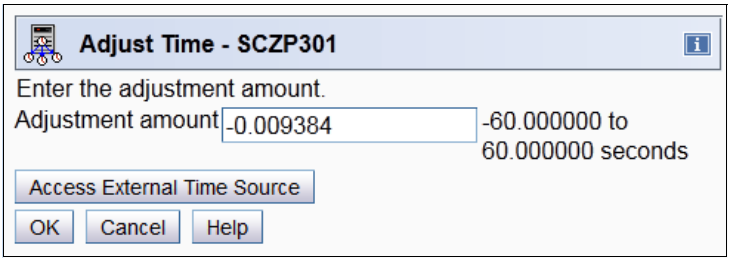

Figure 6-24 Access External Time Source

It is advisable to remove the 69 leap seconds in 1-second increments. As each leap second is removed, the STP UTC time will be 1 second different to the ETS. This 1-second difference will be steered out in 7 hours and then the next 1-second adjustment can be made until the added leap seconds are removed. These steps are shown in Figure 6-25, Figure 6-26, and Figure 6-27.

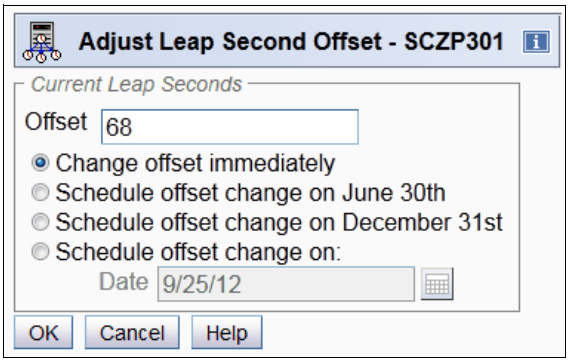

Figure 6-25 Adjust Leap Second Offset

Figure 6-26 shows the next step.

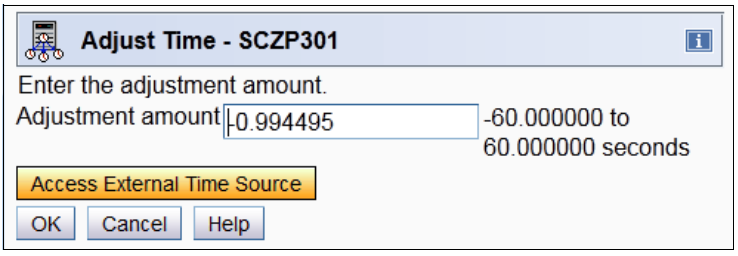

Figure 6-26 Access External Time Source

Figure 6-27 shows the Time Adjustment.

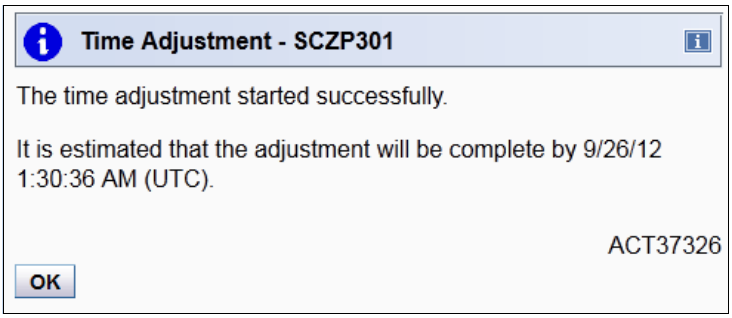

Figure 6-27 Time Adjustment

7

Restoring STP configuration across PORs in a single or dual CEC CTN

In this chapter we explain how you can restore STP configuration across PORs in a single CEC CTN or a dual CEC CTN.

7.1 Saving the configuration across power-on reset

STP provides an option to save the configuration across power-on resets (PORs) for STP-only CTNs with one or two servers.You select this option through the check box for **Only allow the server(s) specified above to be in the CTN** in the Network Configuration tab (Figure 7-1 on page 218).

When this check box is selected, the CTN's timing and configuration settings are saved so that they will not need to be reentered after a loss of power or a power-on reset of the servers.

"Bounded" CTNs: For brevity's sake, the capability implemented by selecting **Only allow the server(s) specified above to be in the CTN** is referred to as the "save configuration" feature. CTNs for which this capability is selected can also be referred to as "bounded" CTNs.

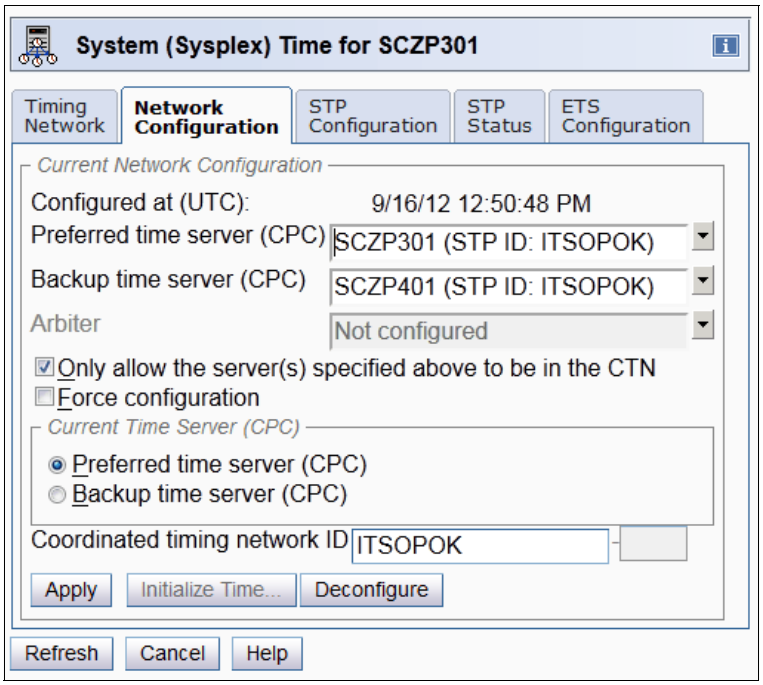

Figure 7-1 Dual CEC - saving configuration across POR

7.2 Single Server CTN

In this section, we discuss recovery scenarios for a single server STP-only CTN. The scenario diagram is shown in Figure 7-2. Note that only one server is configured.

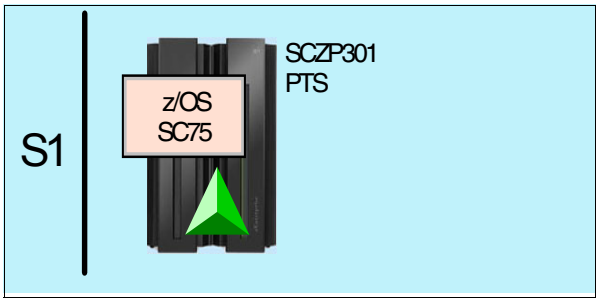

Figure 7-2 Single CEC configuration - STP Stratum 1

The example configuration has one server configured as Preferred Time Server (PTS) and Current Time Server (CTS - Stratum 1).

7.2.1 Prerequisites for single-server bounded CTN

For single-server bounded CTNs, the CTN must contain only a single server and be operating in STP-only mode. This feature, saving STP configuration across PORs, applies to all currently supported servers.

To enable this feature, the CTN must be configured with the check box **Only allow the server(s) specified above to be in the CTN** selected (Figure 7-3).

The configuration data saved includes both the server's roles as the PTS and CTS reflected in the Network Configuration tab.

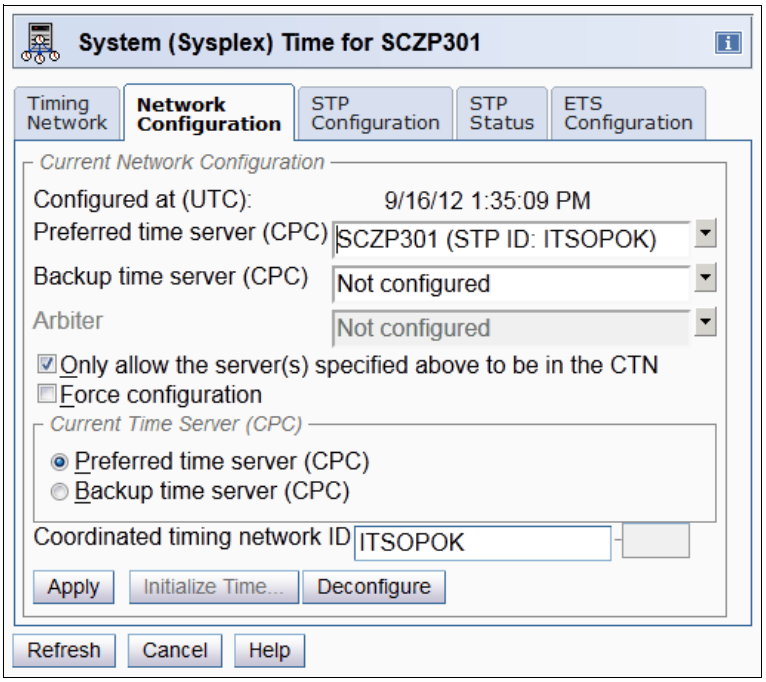

Figure 7-3 Configuring a single CEC "bounded" CTN - Network Configuration tab

Deselection consideration: If the “save configuration” feature is used to limit the CTN to a single-server CTN, no other server can join this CTN unless this option is deselected in the Network Configuration tab of the Current Time Server of the existing CTN.

This restriction can be removed concurrently at any time by deselecting **Only allow the server(s) specified above to be in the CTN**.

7.2.2 Restoring STP configuration after power off and power on sequence

When power is restored, the server resumes its role as PTS/CTS using the timing configuration information previously provided. Due to the simplicity of the CTN configuration, there are no considerations to impact the STP configuration restore process.

7.3 Dual-server CTN

In this section, we discuss STP configuration restore scenarios for a dual-server STP-only CTN. Note that only two servers are configured in the scenario diagram shown in Figure 7-4.

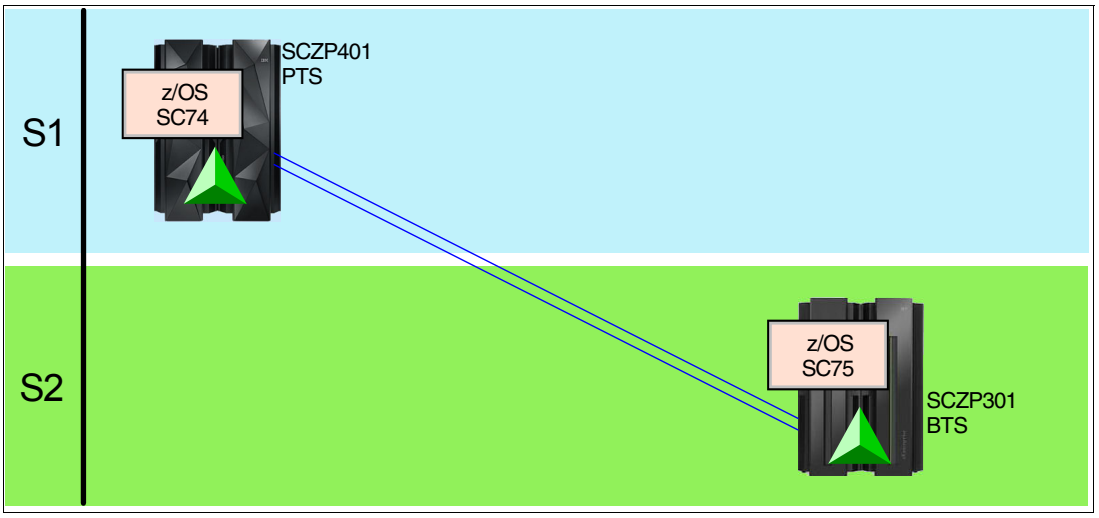

Figure 7-4 STP-only CTN with PTS and BTS assigned

The example configuration has one server configured as Preferred Time Server (PTS) and one as Backup Time Server (BTS). The servers are connected through two coupling links. SCZP401 is the PTS and Current Time Server (CTS - Stratum 1). SCZP301 is the BTS (Stratum 2).

Assumption: The actions described in this section assume that a BTS has been assigned, and that the PTS has been assigned the role of Current Time Server (CTS). This is the preferred configuration because the PTS/CTS is a single point of failure and only the PTS can automatically take back the CTS role after the power on and power off sequence is complete and the STP configuration is restored.

7.3.1 Prerequisites for dual-server bounded CTN

For dual-server bounded CTNs, the CTN must contain two servers and be operating in STP-only mode. The servers must be z9s or newer. It is important that you install the latest MCLs to make sure configuration restore works as described in this section.

Unsupported configuration: A zEC12 and z9 cannot be part of the same CTN; this configuration is not supported.

To enable this feature, the CTN must be configured with the check box **Only allow the server(s) specified above to be in the CTN** selected (Figure 7-5). The configuration data saved includes the servers' roles as the PTS/CTS and BTS, as reflected in the Network Configuration tab.

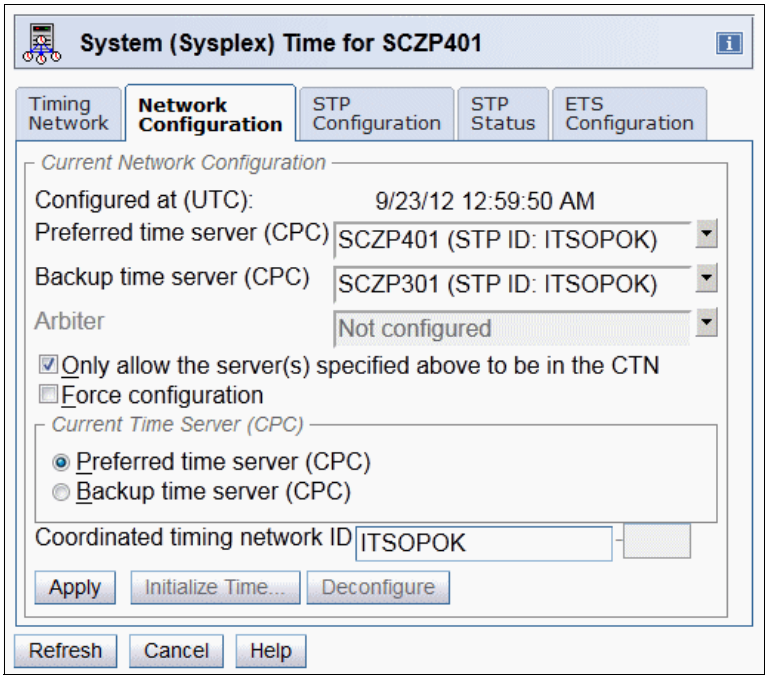

Figure 7-5 Dual CEC bounded CTN configuration - Network Configuration tab

Restrictions: If the "save configuration" feature is used to limit the CTN to a dual-server CTN, a number of restrictions apply:

- No other server can join the CTN.
- The CTN ID cannot be changed.
- No server role assignment changes can be made.
- No reverse migration from STP-only to a mixed CTN.

These restrictions can be removed concurrently at any time by deselecting **Only allow the server(s) specified above to be in the CTN** from the CTS, if the two servers can communicate through coupling links.

7.3.2 Restoring the CTN configuration after power off and power on sequence

When power is restored to both the PTS and the BTS, the servers will normally resume their roles as PTS/CTS and BTS, using the timing configuration information previously provided. However, with two servers, there are considerations that can impact the configuration restore process.

All scenarios in this chapter assume that the PTS has been assigned the role of CTS when the CTN was configured.

One server loses power

One of the reasons for implementing a “bounded” CTN is to shorten system recovery time for one- or two-CEC configurations running in STP-only timing mode. However, there are considerations to keep in mind when you operate such a configuration.

Here we discuss two possible situations when a server can lose power: planned outages and unplanned outages.

Planned outage (maintenance)

- In the Network Configuration Tab, from the server that will remain activated, deselect the **Only allow the server(s) specified above to be in the CTN** check box.
- Ensure the CTS is assigned to the PTS, because there will only be a PTS in the CTN after the deactivation.
- Remove the PTS or BTS role from the server that will be deactivated, because you cannot perform disruptive actions to a server that has one of them. If you are removing the PTS role from the server that will be deactivated, assign it to the BTS because you cannot have a CTN without a PTS.
- Deactivate (power off) the server.

The activated server will remain as Stratum 1 even after the deactivated server has its power restored. When it happens, the server that was deactivated now joins as Stratum 2 and you must reassign the roles to the same configuration they had before, and then select the **Only allow the Server(s) specified above to be in the CTN** check box (Figure 7-7 on page 225).

Unplanned outage

For an unplanned outage, recovery depends on which CEC experiences the outage. Refer to Chapter 4, “Recovery in an STP-only CTN with BTS” on page 115, for more information about this topic.

Both servers lose power

When power is lost to both servers, the process depends on whether coupling link or LAN connectivity is reestablished between the servers when the power is restored. Both servers remain at Stratum 0 until coupling link or LAN connectivity is reestablished between them.

If the servers are able to communicate through any coupling link, they resume their original roles as PTS/CTS and BTS. If they cannot communicate through any coupling link but can communicate through LAN, then Console Assisted Recovery (CAR) is initiated.

Console Assisted Recovery initiated by the PTS

If the PTS determines that the BTS has failed or is a Stratum 0, the PTS becomes the Stratum 1. When coupling link connectivity is restored, the BTS becomes a Stratum 2 server.

If the PTS determines that the BTS is Stratum 1, the PTS remains at Stratum 0. When coupling link connectivity is restored, the PTS retakes its role as Stratum 1 and the BTS becomes a Stratum 2 server.

If the PTS cannot determine the status of the BTS, the PTS remains at stratum 0. When coupling link connectivity is restored, the servers resume their roles as PTS/CTS and BTS.

Console Assisted Recovery initiated by the BTS

If the BTS determines that the PTS has failed, the BTS becomes the Stratum 1. When coupling link connectivity is restored, the PTS rejoins the CTN as Stratum 2 and then retakes its role as CTS and becomes a Stratum 1 server.

If the BTS determines that the PTS is Stratum 1, the BTS remains at Stratum 0. When coupling link connectivity is restored, the BTS joins the CTN as Stratum 2.

If the BTS cannot determine the status of the PTS, the BTS remains at Stratum 0. When coupling link connectivity is restored, the servers resume their roles as PTS/CTS and BTS.

7.3.3 Forcing a server to “Assume CTS role”

When STP configuration cannot be restored through Console Assisted Recovery from either PTS or BTS, this can result in an outage for both servers until link path connectivity can be reestablished between the servers. In these situations, you can force one of the servers to assume the CTS role.

In these situations, if the status of the servers can be determined manually, it is possible to force one of the servers to assume the CTS role without permanently reconfiguring the CTN. Although this process can be viewed from the perspective of either the PTS or the BTS, we review the process from the perspective of the PTS.

Scenario and initial configuration

The following scenario is a typical example of a situation for which the “Assume CTS role” capability was designed. In this scenario, two servers named SCZP401 and SCZP301 are configured in an STP-only CTN as the PTS and BTS, respectively. The two servers are connected by coupling links as can be seen in the STP Status tab (Figure 7-6).

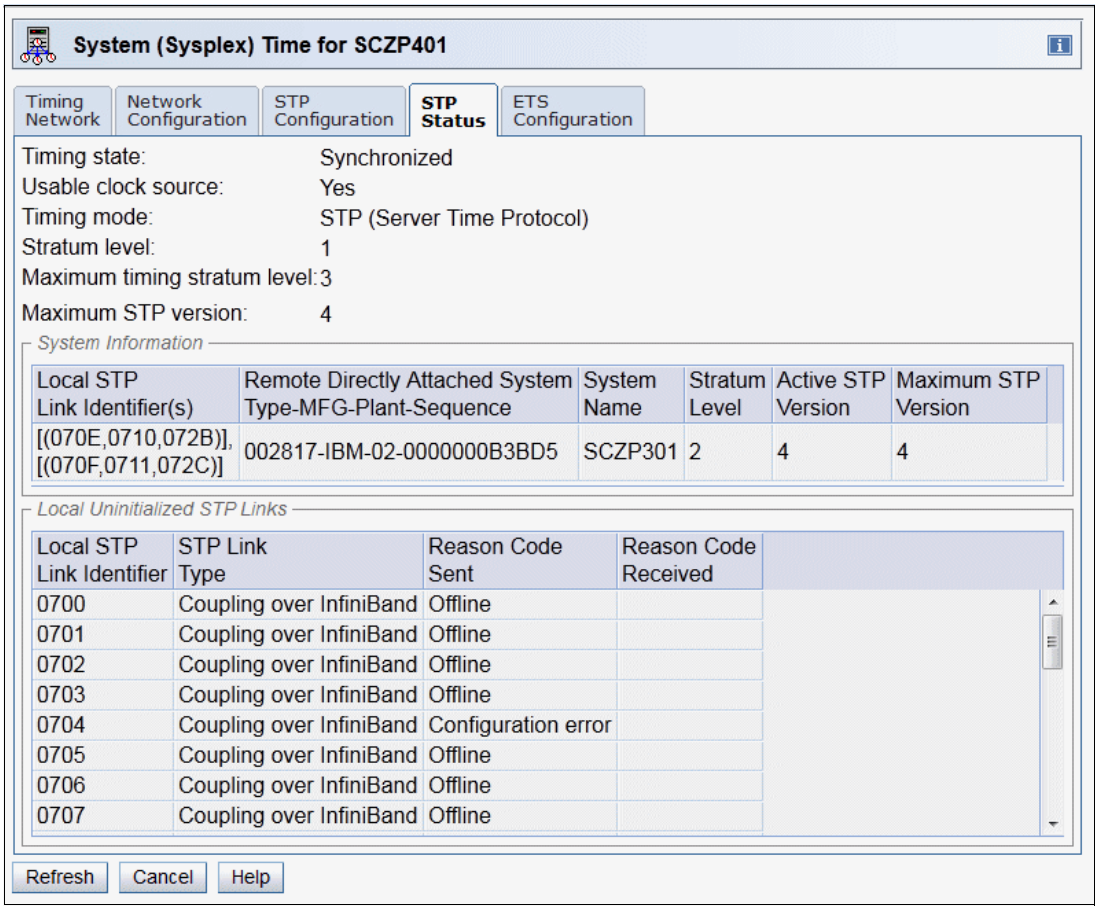

Figure 7-6 Dual CEC - STP Status tab

If there is no LAN communication between the servers, they are unable to resume their timing roles. The STP configuration has been saved by selecting **Only allow the server(s) specified above to be in the CTN** in the Network Configuration tab (Figure 7-7 on page 225).

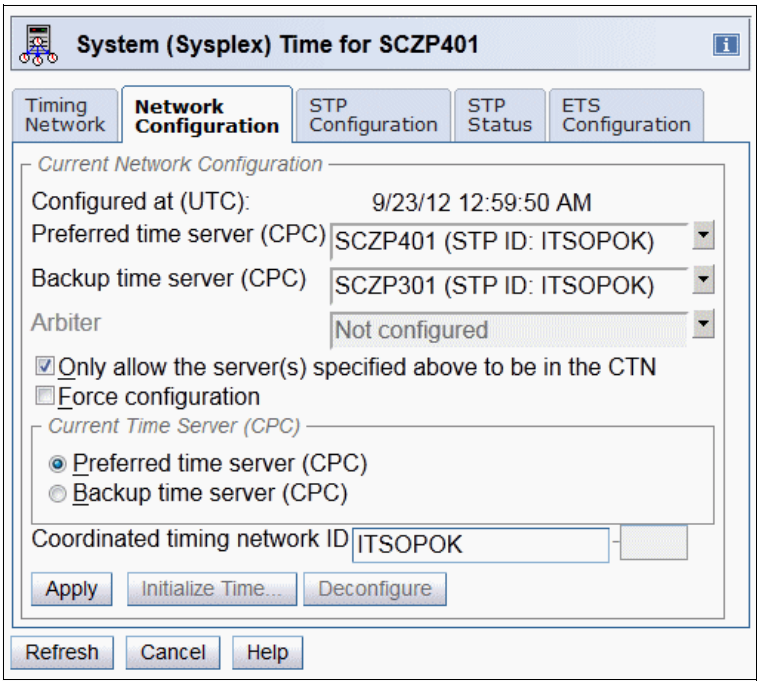

Figure 7-7 Saving STP configuration across PORs - dual CEC CTN

Problem description

After a power outage involving both servers, they are powered on and operational, but in an unsynchronized state (Stratum 0). The servers are unable to resume their timing roles because coupling link connectivity has not been established and there is no LAN connectivity between them; see “Both servers lose power” on page 222.

Later investigation reveals that this is due to a problem in the I/O configurations used by SCZP301 during the activation process. However, the cause of the problem is not immediately apparent and the more pressing concern is to get one of the servers working so that production workload processing can be resumed. As is typically the case, the most important workload is processed on the PTS server, SCZP401 in this case.

The *unsynchronized* status and the problem with the (missing) coupling link connectivity is evidenced in the STP Status tab for server SCZP401 (Figure 7-10 on page 228).

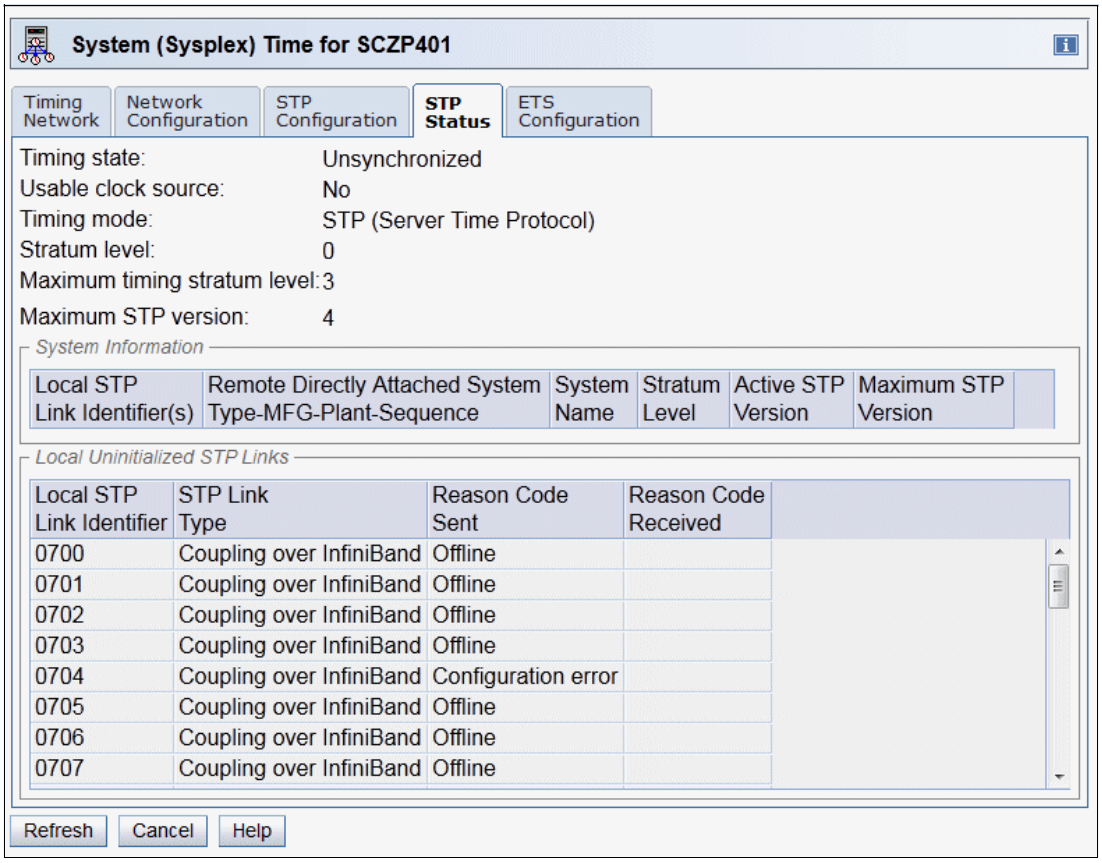

Figure 7-8 STP Status tab - SCZP401 unsynchronized (STP Stratum 0)

Additional confirmation about the coupling link problem is evidenced by the lack of the server name SCZP301 and an STP ID, which is replaced by the Type-MFG-Plant-Sequence number in the “Backup time server” field in the Network Configuration tab for SCZP401 (Figure 7-9 on page 227).

User recovery actions

The PTS server, SCZP401, needs to assume the CTS role so that production workload processing can continue while the cause of the coupling link connectivity problem is being investigated. To do this, click **Assume Current Time Server** in the Network Configuration shown in Figure 7-9 on page 227, and then click **Yes** in the Assume Current Time Server Confirmation window.

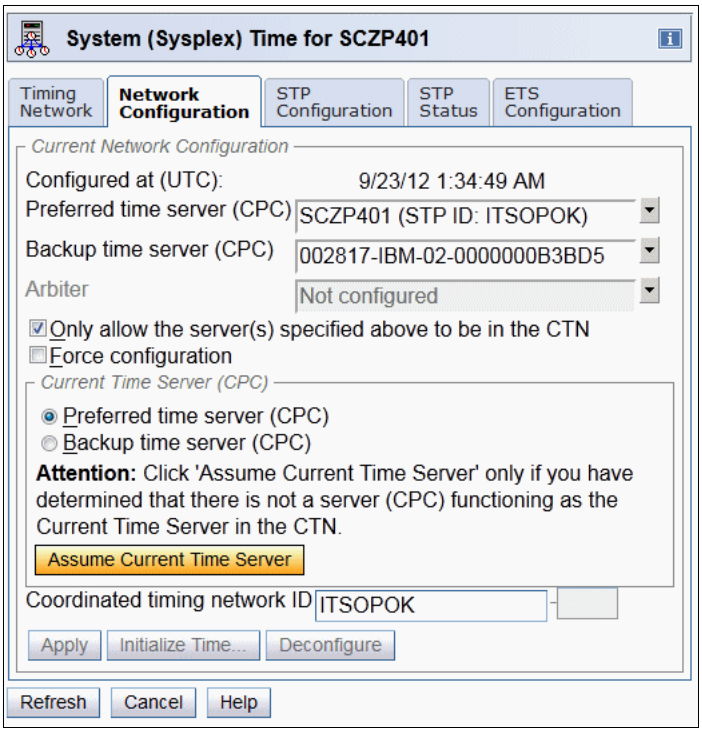

Figure 7-9 SCZP401 assuming CTS role

Recovered configuration

After the PTS (SCZP401) has assumed the CTS role, production processing can resume. As shown in the STP Status tab in Figure 7-10 on page 228, SCZP401 is now synchronized but still has no coupling link connectivity to the BTS.

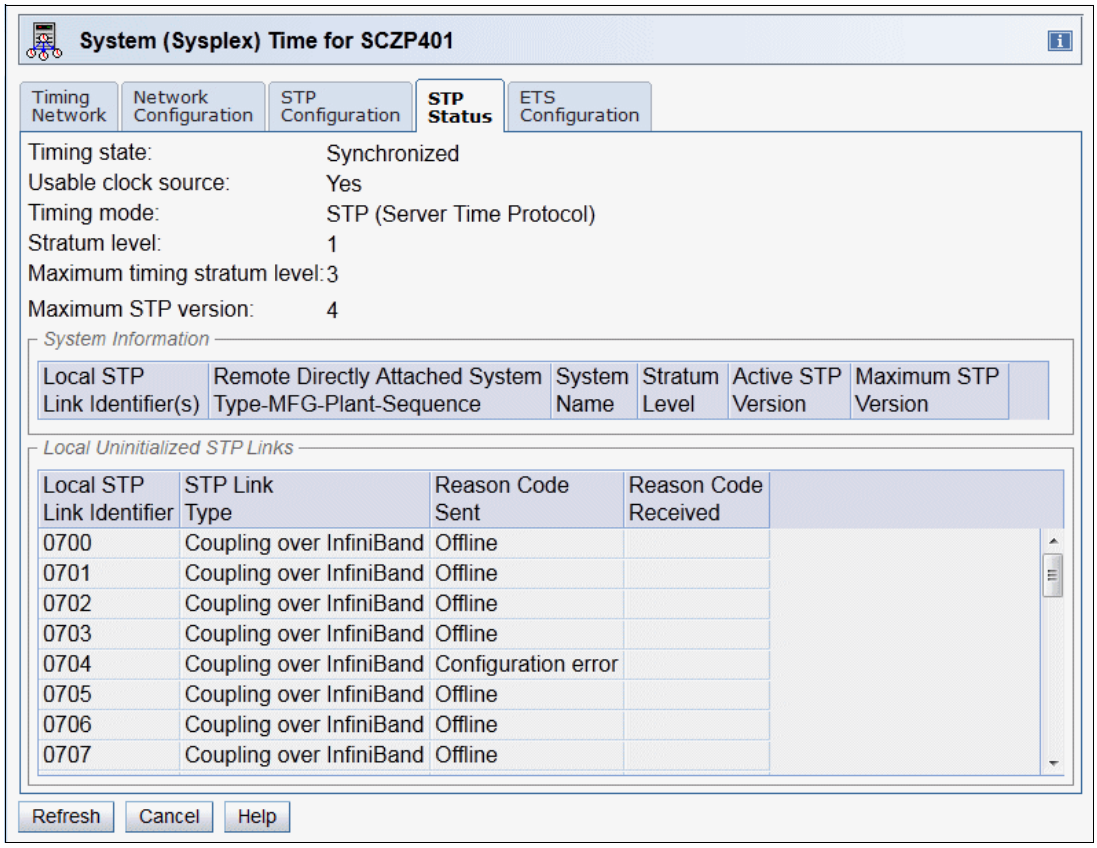

Figure 7-10 SCZP401 synchronized but no link to SCZP301

In addition, the option to Assume Current Time Server is no longer available in the Network Configuration tab because the CTS role has been assigned to the PTS (Figure 7-11).

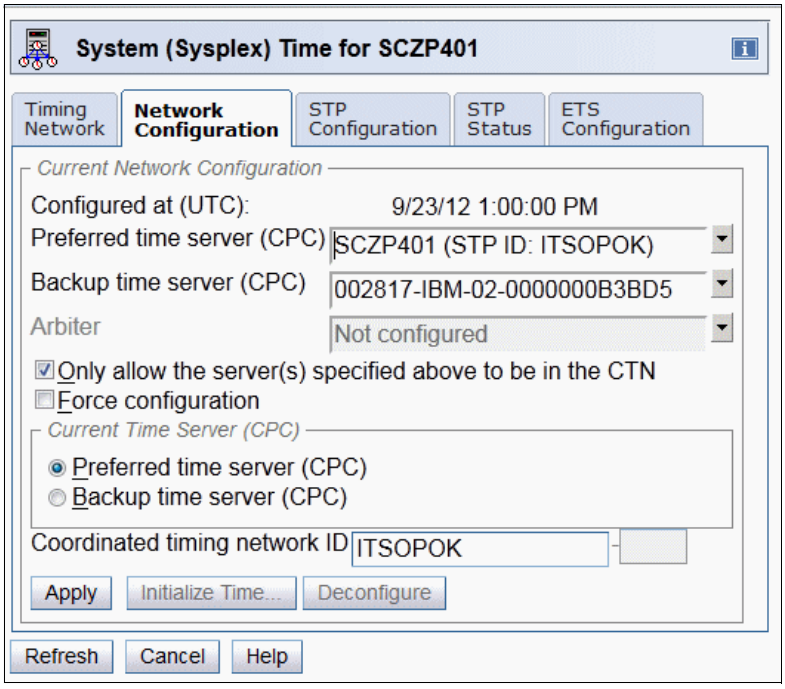

Figure 7-11 Assume Current Time Server button not present

Now that the production workload is again being processed on the PTS, you have time available to investigate the loss of coupling link connectivity. After the cause has been established, corrective action can be taken to bring the second server, SCZP301, back into the CTN as the BTS and regular workload processing can be resumed.

7.3.4 Forcing a CTN to delete the saved configuration

Deleting the saved configuration is accomplished by deselecting **Only allow the server(s) specified above to be in the CTN** in the Network Configuration tab and then selecting **Apply**. Because the configuration data is stored on both servers and this information needs to be in synch, deleting the saved configuration requires that the two servers have coupling link connectivity. However, circumstances can arise in which coupling link connectivity is not available and the lost of connectivity is permanent. In this case, the saved configuration no longer applies and can significantly impact the operation of the servers in the CTN (see 7.3.3, “Forcing a server to “Assume CTS role”” on page 223).

One potential cause for this situation is a permanent removal of one of the servers prior to deleting the saved configuration. In this circumstance, the removed server will never rejoin the CTN, so the configuration needs to be changed to reflect this reality to avoid continually impacting the operation of the remaining server.

In Figure 7-12, server SCZP301 (represented as Type-MFG-Plant-Sequence) is no longer available and needs to be removed from the network configuration.

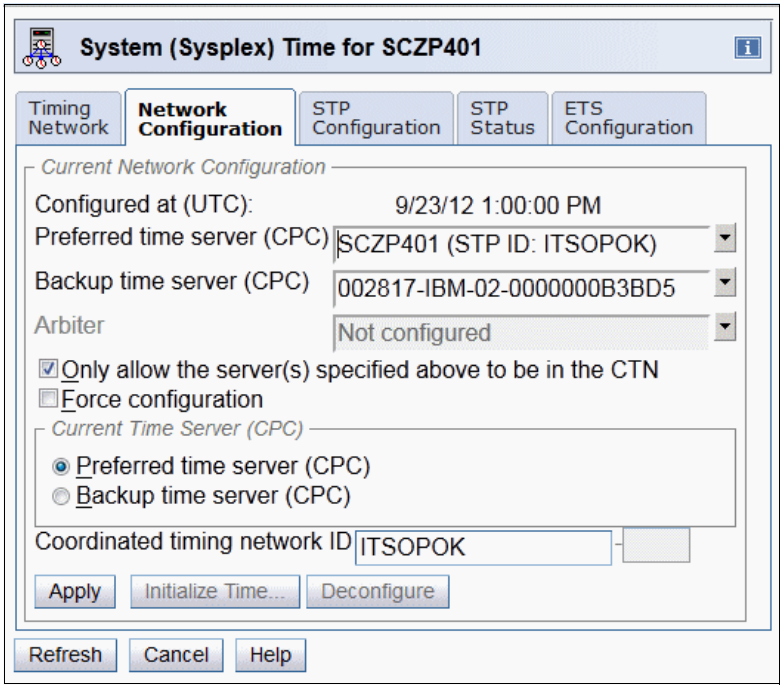

Figure 7-12 Unavailable server (SCZP301) before deletion

The first step in correcting the configuration is force the deselection of **Only allow the server(s) specified above to be in the CTN** by additionally selecting the **Force configuration** option and then selecting **Apply** (Figure 7-13 on page 230).

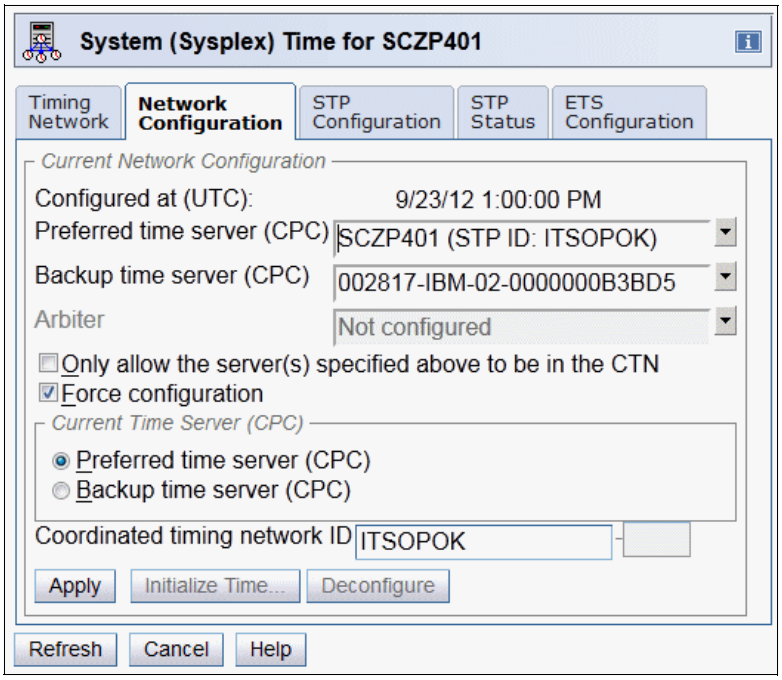

Figure 7-13 Deleting unavailable server SCZP301

A confirmation message appears. Select **Yes** so that the operation can complete. Now that **Only allow the server(s) specified above to be in the CTN** is no longer selected, the server that is no longer available can be removed from the configuration.

After the configuration has been corrected select the **Only allow the server(s) specified above to be in the CTN** option again. Then select **Apply** to save the revised configuration (Figure 7-14).

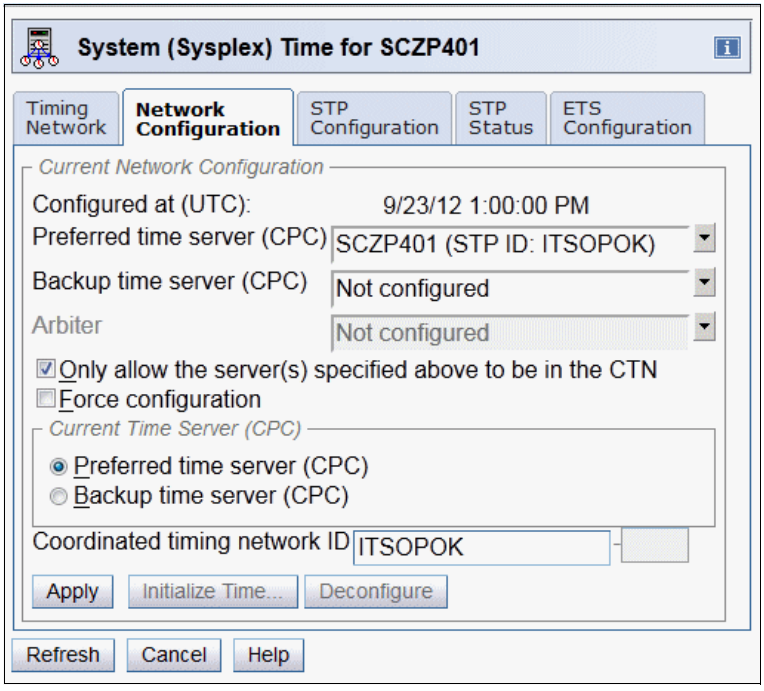

Figure 7-14 Saving STP configuration for a single server

7.4 Considerations for an MES upgrade

Use care when applying an MES upgrade that changes the machine type. For example, consider a situation in which server SCZP401 is a zEC12 and server SCZP301 is a z196 that is being upgraded to a zEC12 through an MES. Figure 7-15 shows the initial configuration.

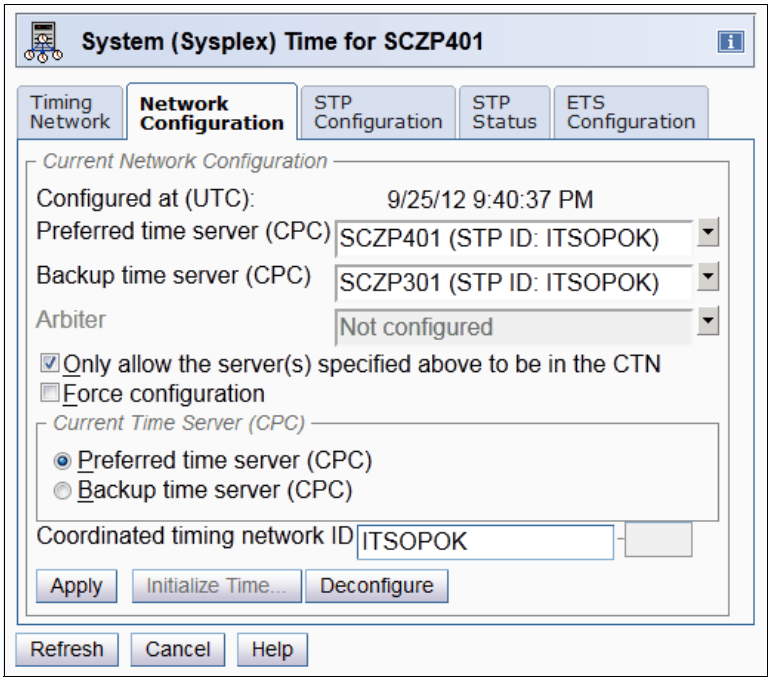

Figure 7-15 Initial configuration (before MES)

Prior to applying such an MES, the saved configuration must be deleted by deselecting **Only allow the server(s) specified above to be in the CTN** and then selecting **Apply** in the Network Configuration tab (Figure 7-16 on page 232).

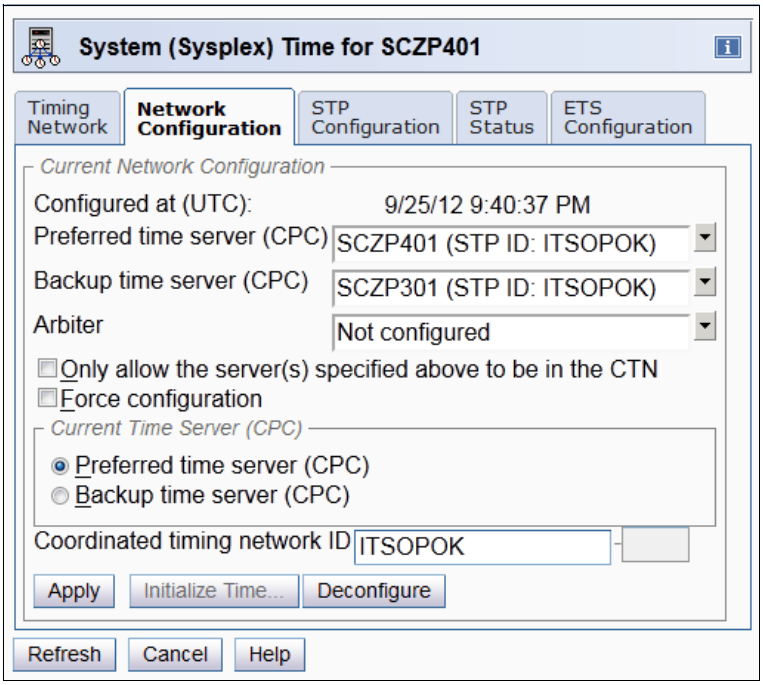

Figure 7-16 Allowing other servers to join the CTN

Then remove the machine being upgraded from the CTN configuration (Figure 7-17).

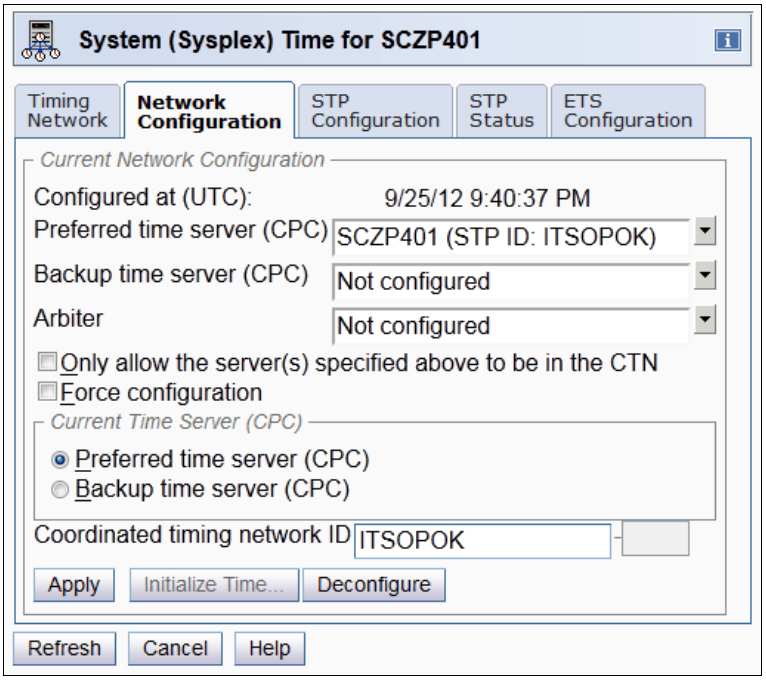

Figure 7-17 Removing the CEC to be upgraded

This must be accomplished while the two servers have coupling link connectivity or the **Force Configuration** option will be required to make the change (see “Forcing a CTN to delete the saved configuration” on page 229).

You can ensure that the configuration and timing values are not lost due to an unanticipated outage of the SCZP401 server while the MES is being applied to the SCZP301 server by

again selecting **Only allow the server(s) specified above to be in the CTN** and then selecting **Apply** in the Network Configuration tab (Figure 7-18).

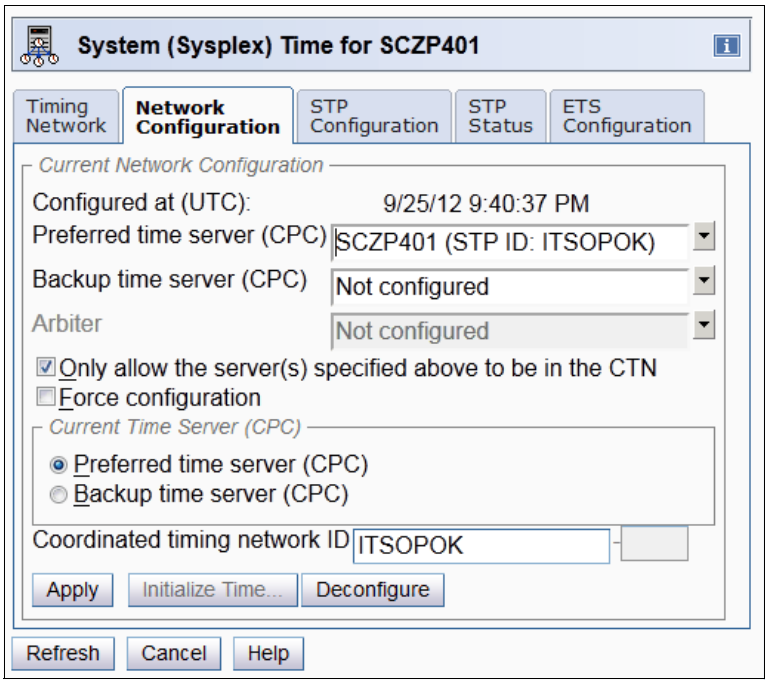

Figure 7-18 Saving configuration for single CEC

After the MES of SCZP301 is complete and the two servers have reestablished coupling link connectivity, **Only allow the server(s) specified above to be in the CTN** should be deselected and SCZP301 added back into the CTN configuration as the BTS.

When the reconfiguration is complete, **Only allow the server(s) specified above to be in the CTN** should once again be selected to save the CTN configuration values. Note that the Network Configuration tab (Figure 7-19 on page 234) is now identical to that shown in the original configuration.

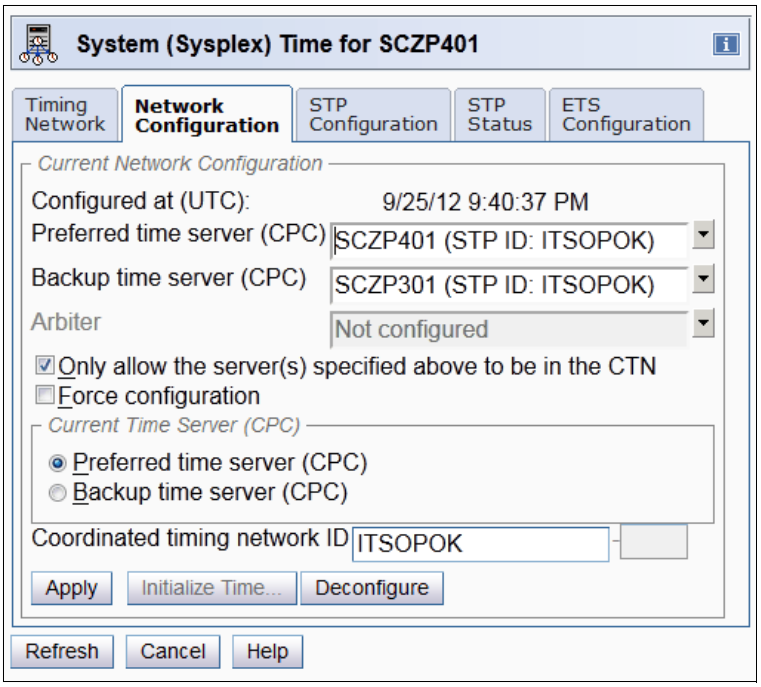

Figure 7-19 Saving configuration after MES

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

For information about ordering these publications, see “How to get Redbooks” on page 237. Note that some of the documents referenced here may be available in softcopy only.

- *Server Time Protocol Planning Guide,* SG24-7280
- *Server Time Protocol Implementation Guide,* SG24-7281
- *Getting Started with InfiniBand on System z10 and System z9,* SG24-7539
- *IBM zEnterprise EC12 Technical Guide*, SG24-8049
- *IBM zEnterprise EC12 Technical Introduction*, SG24-8050
- *IBM zEnterprise 196 Configuration Setup*, SG24-7834
- *IBM zEnterprise EC12 Configuration Setup*, SG24-8034
- *IBM zEnterprise 196 Technical Guide*, SG24-7833
- *IBM System z9 Enterprise Class Technical Guide*, SG24-7124
- *IBM System z9 Business Class Technical Introduction*, SG24-7241
- *IBM System z10 Enterprise Class Configuration Setup,* SG24-7571
- *IBM System z10 Enterprise Class Technical Guide,* SG24-7516
- *IBM System z10 Enterprise Class Technical Introduction,* SG24-7515
- *IBM System z10 Business Class Technical Overview,* SG24-7632
- *IBM System z Connectivity Handbook*, SG24-5444
- *S/390 Time Management and IBM 9037 Sysplex Timer*, SG24-2070

Other publications

These publications are also relevant as further information sources:

- *z/Architecture Principles of Operation*, SA22-7832
- *zEnterprise EC12 Installation Manual for Physical Planning*, GC28-6914
- *zEnterprise 196 Installation Manual for Physical Planning*, GC28-6897
- *zEnterprise 114 Installation Manual for Physical Planning*, GC28-6907
- *IBM System z10 Business Class Installation Manual for Physical Planning,* GC28-6875
- *System z10 Enterprise Class Installation Manual for Physical Planning,* GC28-6864
- *System z10 Processor Resource/Systems Manager Planning Guide,* SB10-7153

- *IBM System z10 Enterprise Class System Overview,* SA22-1084
- *IBM System z10 Business Class System Overview,* SA22-1085
- *IBM System z10 Enterprise Class Service Guide,* GC28-6866
- *System z9 Enterprise Class System Overview*, SA22-6833
- *System z9 Business Class System Overview*, SA22-1083
- *System z9 EC Installation Manual for Physical Planning*, GC28-6844
- *System z9 BC Installation Manual for Physical Planning*, GC28-6855
- *System z9 Processor Resource/Systems Manager Planning Guide,* SB10-7041
- *System z9 Support Element Operations Guide Version 2.9.1,* SC28-6858
- *zSeries 990 Installation Manual for Physical Planning*, GC28-6824
- *zSeries 890 Installation Manual for Physical Planning*, GC28-6828
- *zSeries 890 and 990 Processor Resource/Systems Manager Planning Guide*, SB10-7036
- *z890 and z990 Support Element Operations Guide Version 1.8.2,* SC28-6831
- *Hardware Management Console Operations Guide Version 2.12.0*, SC28-6919
- *zEnterprise System Support Element Operations Guide Version 2.12.0*, SC28-6920
- *System z CHPID Mapping Tool User's Guide*, GC28-6900
- *Hardware Management Console Operations Guide Version 2.10.1,* SC28-6873
- *Hardware Management Console Operations Guide Version 2.10.0,* SC28-6867
- *Support Element Operations Guide V2.10.1,* SC28-6879
- *Support Element Operations Guide V2.10.0,* SC28-6868
- *Planning for the 9037 Model 2*, SA22-7233
- *Model 5900 and 4900 ESCON Server Installation and User's Guide*, GA22-1082
- *z/OS MVS Setting Up a Sysplex*, SA22-7625
- *IOCP User's Guide,* SB10-7037
- *Stand-Alone Input/Output Configuration Program User's Guide,* SB10-7152
- *Planning for Fiber Optic Links,* GA23-0367
- *Common Information Model (CIM) Management Interfaces,* SB10-7154

Online resources

These web sites and URLs are also relevant as further information sources:

- IBM Resource Link

  http://www.ibm.com/servers/resourcelink/
- Server Time protocol Web page

  http://www-03.ibm.com/servers/eserver/zseries/pso/stp.html
- GDPS

  http://www.ibm.com/systems/z/gdps/

How to get Redbooks

You can search for, view, or download Redbooks, Redpapers, Technotes, draft publications and Additional materials, as well as order hardcopy Redbooks publications, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

L

M

N

O

P

R

S

T

U

W

Z

IBM Redbooks

Server Time Protocol Recovery Guide

(0.2”spine)
0.17”<->0.473”
90<->249 pages

Server Time Protocol Recovery Guide

Server Time Protocol concepts and definitions

Planning for recovery from various failures

Restoring STP configuration after a POR

Server Time Protocol (STP) is a server-wide facility that is implemented in the Licensed Internal Code (LIC) of the IBM zEnterprise EC12 (zEC12), IBM zEnterprise 196 (z196), IBM zEnterprise 114 (z114), IBM System z10 Enterprise Class (z10 EC), System z10 Business Class (z10 BC), IBM System z9 Enterprise Class (z9 EC), and System z9 Business Class (z9 BC). It provides improved time synchronization in a sysplex or non-sysplex configuration.

This IBM Redbooks publication will help you plan for and recover from a failure affecting your Mixed or STP-only Coordinated Timing Network. It is intended for technical support personnel requiring information about:

- Recovery concepts and definitions
- Identifying and taking appropriate actions for recovering from a failed component in a Coordinated Timing Network

Readers are expected to be familiar with IBM System z technology and terminology. For planning information, refer to our companion book, Server Time Protocol Planning Guide, SG24-7280, and for implementation details refer to Server Time Protocol Implementation Guide, SG24-7281.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-7380-01 ISBN 073843809X