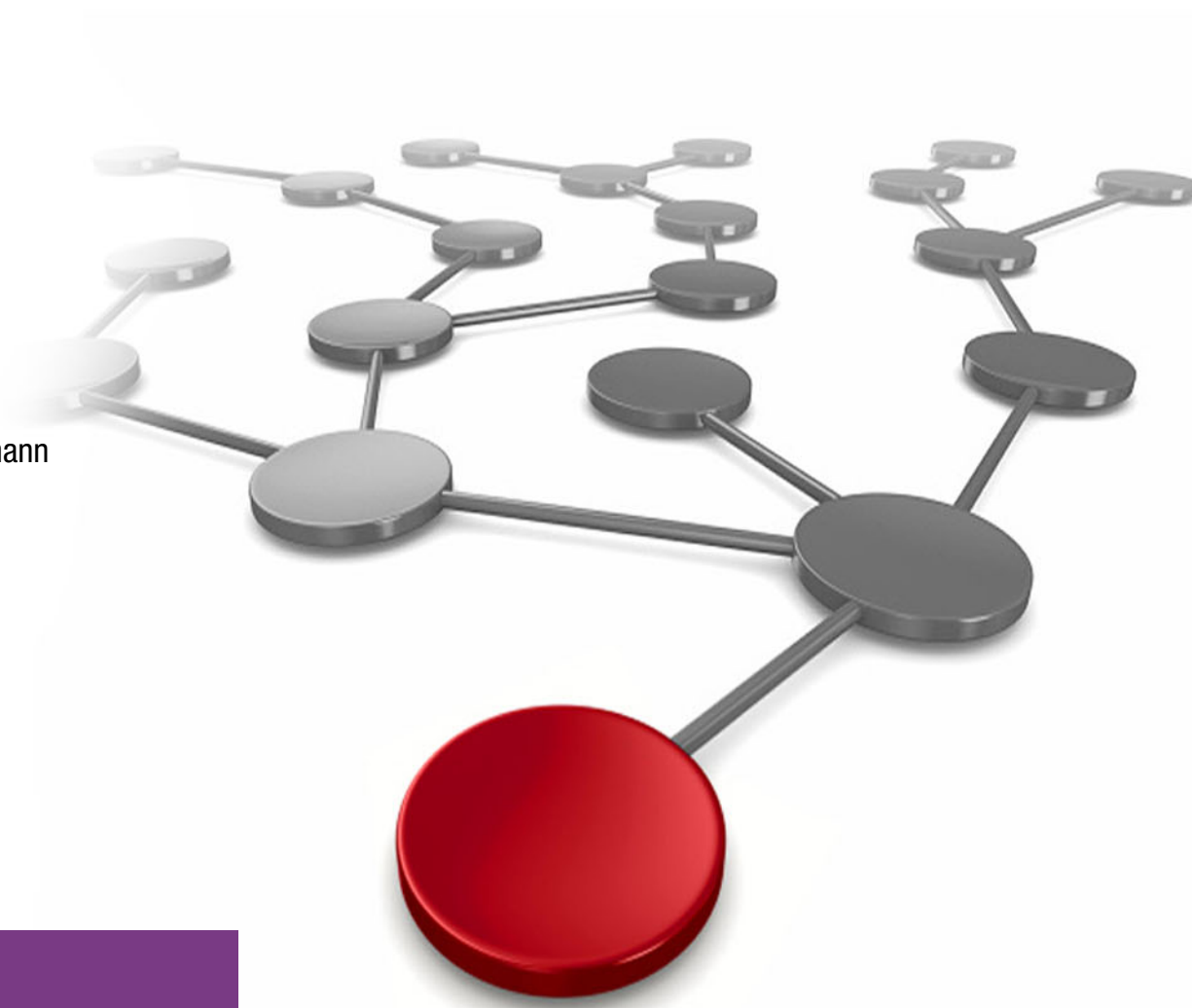


# IBM Z Server Time Protocol Guide

Octavian Lascu  
Franco Pinto  
Gatto Gobehi  
Hans-Peter Eckam  
Jeremy Koch  
Martin Söllig  
Sebastian Zimmermann  
Steve Guendert



IBM Z





IBM Redbooks

**IBM Z Server Time Protocol Guide**

August 2020

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

**First Edition (August 2020)**

This edition applies to IBM Server Time Protocol for IBM Z and covers IBM z15, IBM z14, and IBM z13 server generations.

This document was created or updated on August 3, 2020.

© Copyright International Business Machines Corporation 2020. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

# Contents

<b>Notices</b> .....	vii
Trademarks .....	viii
<b>Preface</b> .....	ix
Authors .....	ix
Comments welcome .....	x
Stay connected to IBM Redbooks .....	xi
<b>Chapter 1. Introduction to Server Time Protocol</b> .....	1
1.1 Introduction to time synchronization .....	2
1.1.1 Insertion of leap seconds .....	2
1.1.2 Time-of-Day (TOD) Clock .....	3
1.1.3 Industry requirements .....	4
1.1.4 Time synchronization in a Parallel Sysplex .....	6
1.2 Overview of Server Time Protocol (STP) .....	7
1.3 STP concepts and terminology .....	9
1.3.1 STP facility .....	9
1.3.2 TOD clock synchronization .....	10
1.3.3 STP Servers .....	11
1.3.4 STP Stratum levels .....	12
1.3.5 Server roles in a CTN .....	14
1.4 Coordinated Timing Network .....	15
1.4.1 CTN ID .....	15
1.5 Coupling Links .....	16
1.6 External Time Source .....	18
1.7 Firmware implementation .....	19
<b>Chapter 2. Planning hardware and software</b> .....	21
2.1 Planning server hardware .....	22
2.1.1 Servers .....	22
2.1.2 Hardware Management Console .....	23
2.1.3 Support Elements -- Connectivity to ETS .....	26
2.2 Planning connectivity .....	27
2.2.1 Planning for coupling links for STP .....	27
2.2.2 Planning for timing-only links .....	32
2.2.3 Considerations for multi-site sysplex .....	34
2.3 External time source .....	35
2.3.1 Precision Time Protocol grandmaster .....	35
2.3.2 NTP server .....	36
2.3.3 NTP server with pulse per second .....	42
2.4 Internal Battery Feature (IBF) .....	46
2.5 Planning z/OS software .....	47
2.5.1 z/OS requirements for STP .....	47
2.5.2 CLOCKxx .....	48
2.6 Planning for z/VM, z/TPF and Linux on IBM Z .....	53
2.6.1 z/VM .....	53
2.6.2 z/TPF .....	56
2.6.3 Linux on IBM Z .....	56
2.7 Special considerations for a single server CTN .....	57

2.7.1	Hardware and ETS planning for single server configuration. . . . .	57
2.7.2	Software planning: z/OS currently running in a sysplex . . . . .	58
2.7.3	Software planning: z/OS systems currently not in a sysplex. . . . .	59
2.8	Migration planning . . . . .	60
2.8.1	CTN types and recovery voting. . . . .	60
2.8.2	Server maintenance and replacement . . . . .	61
2.8.3	N-2 Rule and Coupling Link migration considerations . . . . .	61
2.8.4	Power options (PDU/BPA/IBF) considerations . . . . .	62
	<b>Chapter 3. Configuration and operations . . . . .</b>	<b>63</b>
3.1	Displaying and monitoring the environment . . . . .	64
3.1.1	CPC Detail panel . . . . .	64
3.1.2	The Manage System Time panel . . . . .	64
3.1.3	z/OS display commands. . . . .	67
3.1.4	z/OS messages. . . . .	69
3.2	Configuring the Coordinated Timing Network. . . . .	71
3.2.1	Single Server CTN . . . . .	71
3.2.2	Dual Server CTN. . . . .	78
3.2.3	CTN with three servers - PTS, BTS, and Arbiter. . . . .	80
3.3	Configuring Precision Time Protocol as ETS . . . . .	81
3.3.1	Configuring ETS (PTP w/ PPS). . . . .	83
3.4	Managing the time. . . . .	88
3.4.1	STP time adjustment. . . . .	88
3.4.2	STP offset adjustments. . . . .	91
3.4.3	Changes in local time . . . . .	96
3.4.4	Logical partition time offset . . . . .	99
3.4.5	Parallel Sysplex and multiple time zones . . . . .	100
	<b>Chapter 4. Managing STP configurations . . . . .</b>	<b>105</b>
4.1	Test environment overview . . . . .	106
4.2	Managing CTN configuration - overview . . . . .	106
4.2.1	Changing the name of a CTN (CTN ID) . . . . .	107
4.2.2	Single-server or dual-server auto resume after power-on reset . . . . .	108
4.2.3	Changing the Current Time Server . . . . .	109
4.2.4	Changing the server roles. . . . .	111
4.2.5	Deconfiguring a CTN . . . . .	113
4.3	CTN reconfigurations . . . . .	114
4.3.1	Adding a server to an existing CTN and re-assigning roles . . . . .	114
4.3.2	Splitting a CTN . . . . .	115
4.3.3	Merging Two CTNs. . . . .	118
	<b>Appendix A. IEEE 1588 Precision Time Protocol (PTP) Introduction and Overview</b>	<b>121</b>
	Introduction . . . . .	122
	Background on PTP: History . . . . .	122
	Background on PTP: Technology overview . . . . .	123
	IEEE 1588 PTP scope . . . . .	123
	PTP's underlying technology. . . . .	124
	The concept of PTP Profiles . . . . .	125
	Regulatory Requirements . . . . .	125
	Common Reference and Traceability. . . . .	126
	Overview of Regulations: Past, Present and Future . . . . .	127
	United States: NASD, OATS, FINRA, CAT. . . . .	127
	Europe: MiFID II . . . . .	129
	A comparison of NTP and PTP . . . . .	130

Preparing for PTP .....	130
Conclusion .....	131
<b>Related publications</b> .....	<b>133</b>
IBM Redbooks .....	133
Other publications .....	133
Online resources .....	133
Help from IBM .....	133





# Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

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 grant you any license to these patents. You can send license inquiries, in writing, to:

*IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US*

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 jurisdictions 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 websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

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

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

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.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

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 actual people or business enterprises 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. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

## Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

CICS®	Interconnect®	z/Architecture®
CICSplex®	MVS™	z/OS®
GDPS®	Parallel Sysplex®	z/VM®
IBM®	Redbooks®	z/VSE®
IBM Z®	Redbooks (logo)  ®	z13®
IBM z13®	Resource Link®	z13s®
IBM z13s®	System Storage™	z15™
IBM z14®	System z®	z9®
IBM z15™	System z9®	zEnterprise®

The following terms are trademarks of other companies:

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

UNIX is a registered trademark of The Open Group in the United States and other countries.

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® Z systems. It has been introduced with IBM System z9 generation of IBM Z systems and it provides improved time synchronization in a sysplex or non-sysplex configuration.

This IBM Redbooks® publication is intended for infrastructure architects and system programmers who need to understand the IBM STP functions. Readers are expected to be generally familiar with IBM Z® technology and terminology.

This book provides planning and implementation information for Server Time Protocol functions and associated software support for the IBM z15, z14, and z13 systems generations.

## Authors

This book was produced by a team of specialists from around the world working at IBM Redbooks, Poughkeepsie Center.

**Octavian Lascu** is an IBM Redbooks Project Leader and a Senior IT Consultant for IBM Romania with over 25 years of experience. He specializes in designing, implementing, and supporting complex IT infrastructure environments (systems, storage, and networking), including high availability and disaster recovery solutions and high-performance computing deployments. He has developed materials for and taught over 50 workshops for technical audiences around the world. He is the author of several IBM publications.

**Sebastian Zimmermann** is a **????title???** in **???country???**. He/she has **??** years of experience in **????** field. He/she holds a degree in **????** from **????**. His/her areas of expertise include **??????**. He/she has written extensively on **????????**.

**Martin Söllig** is a Consultant IT Specialist in Germany. He has 28 years of experience working in the IBM Z field. He holds a degree in Mathematics from University of Hamburg. His areas of expertise include z/OS® and IBM Z hardware, specifically in Parallel Sysplex® and GDPS® environments, and also in cryptography on IBM Z.

**Hans-Peter Eckam** is a **????title???** in **???country???**. He/she has **??** years of experience in **????** field. He/she holds a degree in **????** from **????**. His/her areas of expertise include **??????**. He/she has written extensively on **????????**.

**Franco Pinto** is a Client Technical Specialist in IBM Switzerland. He has over 20 years of experience in the mainframe and IBM z/OS fields. His areas of expertise include IBM Z technical pre-sales covering mainframe sizing and installation planning, and providing support on existing and new IBM Z functions.

**Steve Guendert** is a **????title???** in **???country???**. He/she has **??** years of experience in **????** field. He/she holds a degree in **????** from **????**. His/her areas of expertise include **??????**. He/she has written extensively on **????????**.

**Jeremy Koch** is a Systems Programmer in Germany working for Finanz Informatik. He joined Finanz Informatik in 1999. He has 18 years of experience in z/OS®. His areas of expertise

include parallel Sysplex, GDPS, performance, and Z hardware. He is also responsible for Linux on IBM Z as well as for z/VM.

**Gatto Gobehi** is an IBM Z Product Engineer working in IBM Poughkeepsie, NY, USA. He has over 20 years of mainframe engineering and testing experience. He specializes in IBM Z hardware (all areas) and microcode, providing support for IBM Service teams and Customers. He is a customer advocate for several high profile IBM Z Customers.

Thanks to the following people for their contributions to this project:

Mark Maruzzi, Paul Wojciak, Bill Lamastro  
IBM Poughkeepsie, USA

Bob Haimowitz  
IBM Redbooks, Poughkeepsie center

John Houston  
IBM Poughkeepsie, USA

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an IBM Redbooks residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

[ibm.com/redbooks/residencies.html](http://ibm.com/redbooks/residencies.html)

## 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](http://ibm.com/redbooks)

- ▶ Send your comments in an email to:

[redbooks@us.ibm.com](mailto:redbooks@us.ibm.com)

- ▶ Mail your comments to:

IBM Corporation, IBM Redbooks  
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>





# Introduction to Server Time Protocol

This chapter discusses time synchronization and introduces the Server Time Protocol (STP) facility available for IBM Z® systems. The following topics are covered:

- ▶ Introduction to time synchronization
- ▶ Overview of Server Time Protocol (STP)
- ▶ STP concepts and terminology
- ▶ Coordinated timing network
- ▶ Coupling links
- ▶ External Time Source
- ▶ Firmware implementation - overview

## 1.1 Introduction to time synchronization

Historically, the most important requirement for highly accurate time is for navigational purposes. For applications such as very precise navigation and satellite tracking, which must be referenced to the earth's rotation, a time scale that is consistent with the earth's rotation must be used.

In the information technology world, time synchronization has become a critical component for managing the correct order of the events in distributed applications (transaction processing, message logging and serialization), especially for audit and legal purposes.

Today, this time scale is known as Universal Time 1 (UT1). UT1 is computed using astronomical data from observatories around the world. It does not advance at a fixed rate but speeds up and slows down with the earth's rotation rate. While UT1 is actually measured relative to the rotation of the earth with respect to distant stars, it is defined in terms of the length of the mean solar day. This makes it more consistent with civil, or solar, time.

Until 1967, the second was defined on the basis of UT1. Since 1967, the internationally accepted definition of the second has been "9,192,631,770 periods of the radiation corresponding to the transition between two hyperfine levels of the ground state of the cesium-133 atom." In 1967, this definition was already 1000 times more accurate than what could be achieved by astronomical methods, and today it is even more accurate. The atomic definition of the second is primarily aimed at providing an accurate measure of time intervals. At the same time, the need for an accurate time-of-day measure was recognized, leading to the adoption of the following three basic scales of time:

- ▶ International Atomic Time (TAI).  
TAI (*Temps Atomique International*) is based solely on an atomic reference, and its value represents the average of about 200 atomic clocks around the world. Thus, it is more stable than any one clock. It provides an accurate time base that increases at a constant rate with no discontinuities. Its fundamental unit is the second. TAI is not tied to astronomy and has no concept of leap seconds.
- ▶ UT1 (Universal Time) - based on astronomy. Fundamental unit is the average solar day. The old definition of a second was 1/86,400th of this day. No leap seconds.
- ▶ Coordinated Universal Time (UTC - Universal Time, Coordinated or *temps universel coordonné*).  
UTC is derived from TAI and is adjusted to keep reasonably close to UT1. UTC is the official replacement for (and generally equivalent to) the better-known Greenwich Mean Time (GMT).

### 1.1.1 Insertion of leap seconds

In the 1970s the calculation of atomic time replaced the calculation of time based on the irregular rotation of the Earth. Studies of the Earth's dynamics show that the velocity of the Earth's rotation is decreasing, and in consequence a rotational day is longer than a day of 86 400 atomic seconds (24 hours x 3,600 seconds).

When atomic time was adopted, some communities of users - in particular those using celestial navigation - requested that atomic time be synchronized with the rotation of the Earth. To compensate for the Earth's irregular velocity of rotation, the International Telecommunication Union (ITU) defined in 1972 a procedure for adding (or suppressing) a second as necessary, to ensure that the difference between the international time reference and rotational time remained less than 0.9 s. The resulting time scale is Coordinated Universal Time (UTC), the atomic time scale maintained at the Bureau International des



Poids et Mesures (BIPM) with the contribution of 69 national institutes that operate about 400 atomic clocks.

For further information please see the corresponding website of the Bureau International des Poids et Mesures (BIPM):

<https://www.bipm.org/en/bipm-services/timescales/leap-second.html>

The International Earth Rotation and Reference Systems Service (IERS<sup>1</sup>) is responsible for monitoring the Earth's rotation and announces the dates of application of any leap seconds required, usually timed for the end of 30 June or 31 December. Refer to the IERS Bulletin C for the latest information:

<ftp://hpiers.obspm.fr/iers/bul/bulc/bulletinc.dat>

**Note:** The effect of a leap second is the introduction of an irregularity into the UTC time scale, so exact interval measurements are not possible using UTC, unless the leap seconds are included in the calculations. After every positive leap second, the difference between TAI and UTC increases by one second.

### 1.1.2 Time-of-Day (TOD) Clock

The Time-of-Day (TOD) clock was introduced as part of the System/370 architecture to provide a high-resolution measurement of the real time, suitable for the indication of date and time of day. The cycle of the clock is approximately 143 years and wraps on September 18, 2042.

In July 1999, the extended TOD clock facility was announced. The TOD clock was extended by 40 bits on the right. This 104-bit value, along with eight zero bits on the left and a 16-bit programmable field on the right, can be stored by the problem program instruction STORE CLOCK EXTENDED (STCKE).

With proper operating system support, extended TOD clock provides for unique timestamps across the systems in a sysplex<sup>2</sup>. The value of the TOD clock is directly available to application programs by using the instructions STORE CLOCK (STCK), STORE CLOCK FAST (STCKF), and STORE CLOCK EXTENDED (STCKE). These instructions store the value of the clock in a storage location specified by the instruction. Figure 1-1 shows the format of the TOD clock.

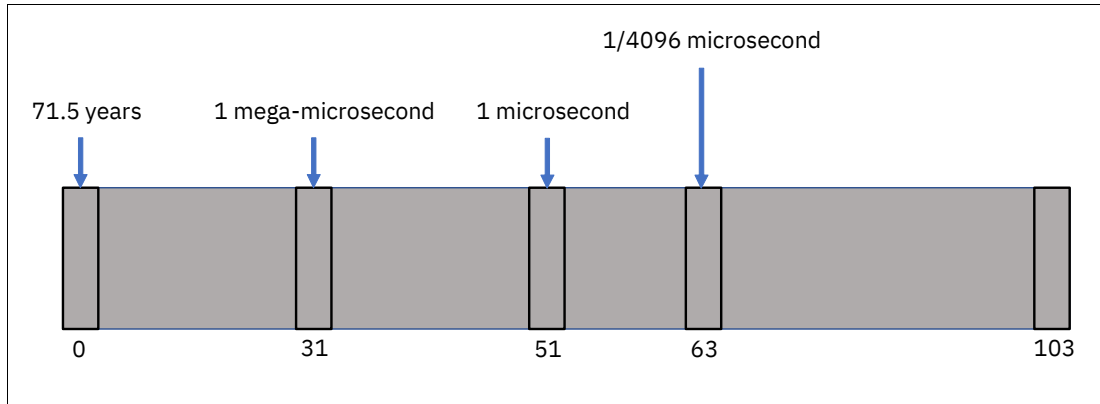


Figure 1-1 TOD format

<sup>1</sup> <https://www.bipm.org/en/worldwide-metrology/liaisons/iers.html>

<sup>2</sup> <https://www.ibm.com/it-infrastructure/z/technologies/parallel-sysplex>

Conceptually, the TOD clock is incremented so that a 1 (one) is added into bit position 51 every microsecond. Actual TOD-clock implementations might not provide a full 104-bit counter but maintain an equivalent stepping rate by incrementing a different bit at such a frequency that the rate of advancing the clock is the same as though a 1 were added in bit position 51 every microsecond.

Figure 1-1 on page 3 shows also the stepping rate (rate at which the bit positions change) for selected TOD clock bit positions. A carry-out of bit 32 of the TOD clock occurs every  $2^{20}$  microseconds (1.048576 seconds). This interval is sometimes called a mega-microsecond.

The use of a binary counter for the time of day, such as the TOD clock, requires the specification of a time origin, or *epoch*. Epoch is the time at which the TOD clock value would have been all zeros. The z/Architecture®, ESA/390, and System/370 architectures established the epoch for the TOD clock as January 1, 1900, 0 a.m. GMT.

**Note:** The TOD clock is architected to be TAI-10 (10 seconds behind TAI). TAI has been introduced in 1961, while UTC has been introduced in 1972, and started off 10 seconds (UTC = TAI -10 seconds). There have been 27 adjustments to the leap second since January 1st, 1972. Thus while the current leap second value is 37 (TAI -UTC = 37 seconds as of January 1st, 2017), for STP, 27 should be used as the leap second value when setting the TOD clock.

### 1.1.3 Industry requirements

In the Financial Industry, where trading occurs across multiple entities and is carried out over network, time synchronization accuracy for business clocks must be in compliance with regulatory mandates established by authorized organizations. Such organizations establish standards and impose regulations for investors' protection. Such organizations include Financial Industry Regulatory Authority in the United States of America and the European Securities and Markets Authority (ESMA, in the European Union).

#### **FINRA<sup>3</sup>**

Regulation changes announced by FINRA significantly reduce the allowable dispersion from universal time of systems processing certain financial trades. The date by which customers must meet these new regulations is staged. For some systems, compliance is required by February 20, 2017. However, it is important that all customers involved in the financial industry carefully evaluate their situation.

All currently supported Z systems machines running Server Time Protocol (STP) already contain the features required to enable the machine to meet the new regulatory timing requirements. However, it is likely that customers will have to purchase new vendor equipment to take advantage of this feature. Pulse Per Second (PPS) is a standard, together with either Precision Time Protocol (PTP), or Network Time Protocol (NTP), that provides precise notifications to connected systems so long as the PPS provider is attached to a compliant time source.

For more information about the regulatory changes, we recommend the following document:

<https://www.finra.org/rules-guidance/notices/16-23>

#### **MiFID II**

The Directive on markets in financial instruments (MiFID II) is an European regulation for financial markets. It has been published in the EU Official Journal on June 12, 2014. The

<sup>3</sup> Financial Industry Regulatory Authority

directive empowers ESMA, the European Securities and Markets Authority to develop regulatory technical standards (RTS) as well as implementing technical standards (ITS). ESMA delivered three sets of technical standards in 2015. The original plan was to put these standards into effect on January 1st, 2017, eventually postponed to January 2018. The standards apply since January 3rd, 2018.

### ***What is RTS 25?***

The Regulatory Technical Standard 25 is a part of the standards developed by ESMA in the context of MiFID 2. RTS 25 defines standards for clock synchronization, acknowledging that this topic has a direct impact in numerous business processes of the financial industry. The ever increasing speed and volume of financial transactions requires more accurate time stamps in business records like order books. Basically, RTS 25 covers two main topics: the time reference to be used and the required level of accuracy for time stamps used in business records.

### ***What are the requirements for clock synchronization as defined in RTS 25?***

- ▶ Article 1 of RTS 25 defines that the time reference used for synchronizing the business clocks used by operators of trading venues shall be Coordinated Universal Time (UTC).
- ▶ Article 2 defines the accuracy level for business clocks of trading venue operators. It specifies requirements which depend on the gateway-to-gateway latency time of a trading system, allowing the maximum divergence from UTC to be either 1 millisecond (gw-to-gw latency of > 1 ms) or 100 microseconds (gw-to-gw latency =< 1 ms). The granularity of the time stamps can be either 1 millisecond or 1 microsecond, also depending on the gateway-to-gateway latency time.

For more information about the regulatory changes, we recommend the following document:

<https://op.europa.eu/en/publication-detail/-/publication/5a8abd26-f1f5-11e3-8cd4-01aa75ed71a1/language-en>

For an overview on how to deploy a Coordinated Timing Network using STP with PPS see Figure 1-2

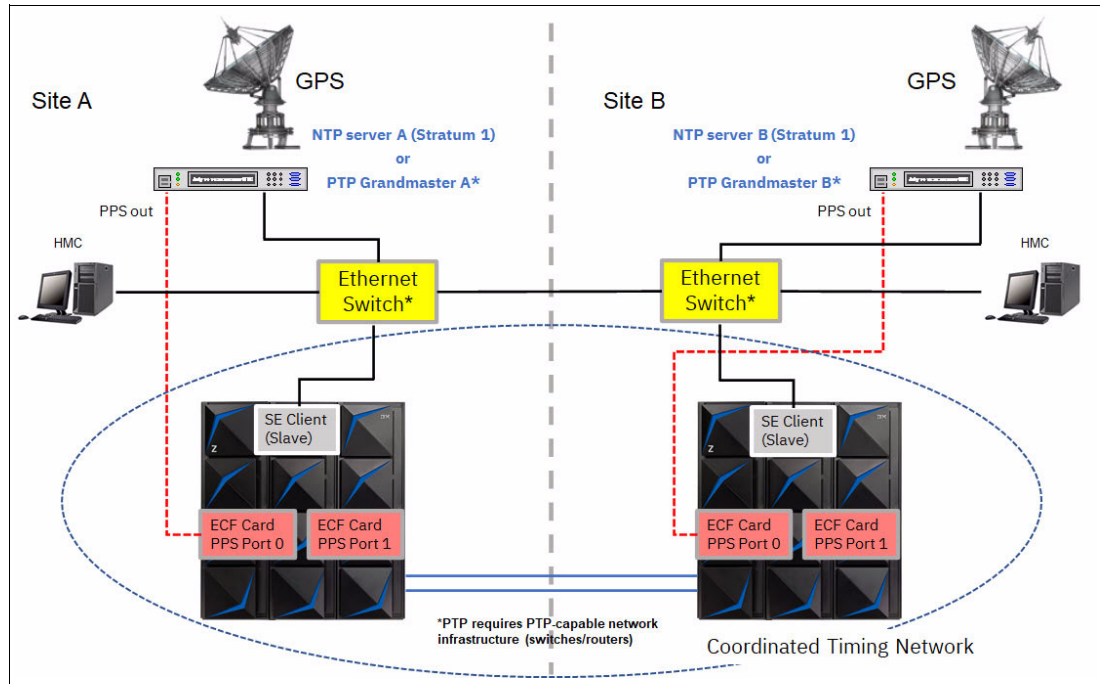


Figure 1-2 Principle of how to deploy STP with PPS<sup>4</sup>

### 1.1.4 Time synchronization in a Parallel Sysplex

In z/Architecture, the STCK and STCKE instructions provide a means by which programs can both establish time-of-day and unambiguously determine the ordering of serialized events, such as updates to a database, a log file, or another data structure. The architecture requires that the TOD clock resolution be sufficient to ensure that every value stored by a STCK or STCKE instruction is unique. Consecutive STCK or STCKE instructions executed, possibly on different CPUs in the same server, must always produce increasing values. Thus, the time stamps can be used to reconstruct, recover, or in many different ways ensure the ordering of these serialized updates to shared data.

Parallel Sysplex® and later on Geographically Dispersed Parallel Sysplex (GDPS®) extended this requirement to the scope of an entire sysplex. Specifically, in a Parallel Sysplex or GDPS, the processes are multisystem processes executing on different servers in the same sysplex. Therefore, time consistency must be maintained across servers. To accommodate this requirement for a Parallel Sysplex or GDPS, the TOD clock architecture was extended by introducing two new major components:

- ▶ External Time Reference (ETR)<sup>5</sup>

The 9037 External Time Reference (ETR) aka the Sysplex Timer was introduced with Parallel Sysplex. An ETR network was comprised on one or two 9037 devices, which synchronized time among themselves. The ETRs provided time services for the IBM Z CPCs.

**Note:** Support for the 9037 ended with the IBM z10 server family

- ▶ Server Time Protocol (STP)<sup>5</sup>

<sup>4</sup> PTP is supported on z15 and newer systems

<sup>5</sup> Thanks to Stephen Guendert, see <https://www.planetmainframe.com/2018/08/brief-history-mainframe-time/>

STP is a message-based protocol, similar to the industry standard Network Time Protocol (NTP). STP allows a collection of IBM Z servers to maintain time synchronization with each other using a time value known as Coordinated Server Time (CST). The network of servers is known as a Coordinated Timing Network (CTN). The mainframe's Hardware Management Console (HMC) plays a critical role with STP CTNs: The HMC can initialize Coordinated Server Time (CST) manually or initialize CST to an external time source. The HMC also sets the time zone, daylight savings time, and leap seconds offsets. It also performs the time adjustments when needed.

### ***TOD-Clock Synchronization Facility***<sup>5</sup>

TOD-Clock Synchronization Facility provides an interface between the operating system (OS) and the designated time server in order to allow the OS to:

- Coordinate setting the local clocks with the Sysplex Time.
- React to losses in synchronization through an external interruption so that data integrity is maintained.

STP supports two external time source options: an external PTP server (master), or an external NTP server. In current implementation, the Support Element encompasses the client code for synchronizing CPC time to the external time source, which provides up to 100 milliseconds (ms) accuracy. For higher time accuracy, both time source options can use a pulse per second (PPS) capable hardware connecting to the PPS ports on the Z CPCs. PPS provides 10 microseconds of accuracy and requires a connection directly to the oscillator. See:

<https://www.planetmainframe.com/2018/08/brief-history-mainframe-time/>

### **The process to achieve time consistency uses the following algorithm:**

1. Server A executes a STCK instruction (time stamp x), which places the clock contents in storage.
2. Server A signals server B.
3. Server B, on receipt of the signal, immediately executes STCK (time stamp y).
4. Server B must then signal Server A, and Server A, upon receipt immediately executes STCK (time stamp z). The three values x, y, and z must be in order. Without this final step (or two) you only test that server B's clock  $\geq$  server A's clock.

For time stamp x and time stamp y to reflect the fact that y is later than x, the two TOD clocks must agree within the time required to send the signal. The consistency required is limited by the time required for signaling between the coupled servers and the time required by the STCK instruction itself. In practical terms, this means that the CPC assigned with the role of CTS is required to provide the TOD clocks to all participating systems with a timestamp so all other CPCs can accomplish steering to get to a synchronized state with each other to within a small number of microseconds, dictated by the fastest possible passing of data from one system to another through a coupling facility (CF) link structure.

## **1.2 Overview of Server Time Protocol (STP)**

Server Time Protocol (STP) is designed to help multiple IBM Z CPCs maintain time synchronization with each other, without the use of a Sysplex Timer. STP uses a message-based protocol in which timekeeping information is passed over externally defined coupling links, such as Coupling Express LR (CE LR) links, Integrated Coupling Adapter Short Range (ICA SR), HCA3-O InfiniBand long reach links and HCA3-O InfiniBand short

reach links. These can be the same links that already are being used in a Parallel Sysplex for coupling facility message communication.

STP is implemented in the Licensed Internal Code (LIC) of IBM Z® servers and CFs for presenting a single view of time to PR/SM™.

**Note:** A time synchronization mechanism, such as Server Time Protocol (STP), is a mandatory requirement for a Parallel Sysplex environment consisting of two or more IBM Z CPCs.

In information technology, two important objectives for survival are:

- ▶ Systems designed to provide continuous availability
- ▶ Near transparent disaster recovery (DR)

Systems designed to provide continuous availability combine the characteristics of high availability and continuous operations to deliver high levels of service around the clock. To attain these objectives, solutions such as GDPS are based on geographical clusters (such as Parallel Sysplex) and remote data mirroring across two or more data centers. An increasing number of enterprises require that the geographical cluster or Parallel Sysplex be dispersed over distances of 100 km or more to mitigate the risk that a single disaster could impact multiple data centers.

An up-to-date time synchronization must provide the following functionality:

- ▶ Improved time synchronization (compared to Sysplex Timer) for System z9® and newer servers in a sysplex or non-sysplex configuration.
- ▶ Scale with distance. Servers exchanging messages over fast, short links require more stringent synchronization than servers exchanging messages over long distances.
- ▶ Scale with server and coupling link technologies. For example, the solution is re-used with appropriate changes when coupling link technologies change in the future.
- ▶ Support a multi-site sysplex of at least 100 km without requiring an intermediate site and it should not preclude going to longer distances in the future.
- ▶ Allow concurrent migration from the ETR network.
- ▶ Allow implementation into an existing Precision Time Protocol or Network Time Protocol (NTP) network, which may or may not utilize a pulse per second (PPS) signal.

The new time synchronization infrastructure developed by IBM for the IBM Z environment is called *Server Time Protocol*.

## Server Time Protocol

Server Time Protocol (STP) is a server-wide facility (FC 1021) that is implemented in the Licensed Internal Code (LIC) starting with the IBM z990 server. STP presents a single view of time to Processor Resource/Systems Manager (PR/SM) and is designed to provide the capability for multiple STP configured servers to maintain time synchronization with each other while providing synchronized time to (supporting) operating systems running across all Z servers participating in the Coordinated Timing Network. It is the follow-on to the External Time Reference. STP is designed to allow events occurring in different servers to be properly sequenced in time.

STP is designed for servers that have been configured in a Parallel Sysplex or a basic sysplex (without a coupling facility), as well as servers that are not in a sysplex but that must maintain time synchronized to an external time reference.

STP is a message-based protocol in which timekeeping information is passed over data links between servers. The timekeeping information is transmitted over externally defined coupling links. Coupling links that can be used to transport STP messages are the no longer available Inter System Channel-3 (ISC-3) links configured in peer mode, Integrated Cluster Bus-3 (ICB-3) links, Integrated Cluster Bus-4 (ICB-4) links, the InfiniBand (IFB) links on servers prior to the z13® server family, as well as the Integrated Coupling Adapter Short Reach (ICA SR) and the Coupling Express Long Reach (CE LR) on the z13/z13s (Driver Level 27 with MCLs), z14 and newer IBM Z server families.

STP provides the following functionality:

- ▶ Allow clock synchronization for the IBM Z server families (beginning with the z990, and z890 servers), without requiring the Sysplex Timer.

**Important:** At the time of this writing (August 2020) the current IBM z15 systems generation support only STP for time synchronization and can coexist in the same Coordinated Timing Network with IBM z14 and IBM z13 systems generations (N-2 generations rule).

- ▶ Support a multi-site timing network of up to 100 km (62 miles)<sup>6</sup> over fiber optic cabling, allowing a Parallel Sysplex to span these distances with no intermediate site requirement.
- ▶ Potentially reduce the cross-site connectivity required for a multi-site Parallel Sysplex.
- ▶ Coexist with an ETR network (older servers only, N-2 generations rule).
- ▶ Allow use of an external time source to set the time to an international time standard, such as Coordinated Universal Time, as well as adjust to the time standard on a periodic basis.
- ▶ Allow setting of local time parameters, such as time zone and daylight saving time (DST).
- ▶ Allow automatic updates of daylight saving time.

**Note:** The IBM z12 EC is the last server to support connections to a Mixed CTN. However, it cannot be configured in the same CTN with any of the following servers: z9 EC, z9 BC, z990 or z890.

Software support for STP is available in all supported z/OS® releases with applicable PTFs.

## 1.3 STP concepts and terminology

This section provides an overview of some of the concepts and terminology related to STP. Understanding this vocabulary will assist you in successfully planning a timing network based on STP.

### 1.3.1 STP facility

STP provides the means by which the TOD clocks in various systems can be synchronized using messages transported over links. STP operates in conjunction with the TOD-clock steering facility, providing a new timing mode, timing states, external interrupts, and machine check conditions.

<sup>6</sup> Extended communications distance equipment (qualified by IBM) must be used, see IBM Resource Link® (IBM ID required).

## TOD-clock steering facility

The TOD-clock steering facility provides a means to change the apparent stepping rate of the TOD clock without changing the physical hardware oscillator that steps the physical clock. This is accomplished by means of a TOD-offset register that is added to the physical clock to produce a logical-TOD-clock value.

TOD-clock steering permits the timing facility control program to adjust the apparent stepping rate of the TOD clock. In normal operation, the update is performed frequently so that the effect, as observed by the program, is indistinguishable from a uniform stepping rate.

The total steering rate is made up of two components:

- ▶ Fine-steering rate

The fine-steering rate is used to correct the inaccuracy in the local oscillator, which is stable over a relatively long period of time. The value normally is less than the specified tolerance of the local oscillator. The change occurs infrequently (on the order of once a day to once a week) and is small.

- ▶ Gross-steering rate

The gross-steering rate is used as a dynamic correction for all other effects, the most predominant being to synchronize time with an external time source or with other clocks in the timing network. The value normally changes frequently (on the order of once per second to once per minute).

## 1.3.2 TOD clock synchronization

This section provides definitions of the timing mode, timing state, and STP clock source state when the STP feature is installed.

### Timing mode

The timing mode specifies the method by which the TOD clock is maintained for purposes of synchronization within a timing network. A TOD clock operates in one of the following timing modes:

- ▶ Local Timing Mode

When the configuration is in local timing mode, the TOD clock has been initialized to a local time and is being stepped at the rate of the local hardware oscillator. The configuration is not part of a synchronized timing network.

- ▶ STP Timing Mode

When the configuration is in STP-timing mode, the TOD clock has been initialized to coordinated server time (CST) and is being stepped at the rate of the local hardware oscillator. In STP timing mode, the TOD clock is steered so as to maintain, or attain, synchronization with CST. To be in STP-timing mode, the configuration must have an STP ID network defined. Refer to 1.4, “Coordinated Timing Network” on page 15, for an explanation of CST.

- ▶ ETR Timing Mode

When the configuration is in ETR-timing mode, the TOD clock has been initialized to the ETR and is being stepped by stepping signals from ETR. To be in ETR-timing mode, the configuration must be part of an ETR network or a mixed CTN.

### Timing states

The timing state indicates the synchronization state of the TOD clock with respect to the timing network reference time. The timing states are:



▶ Synchronized State:

When a configuration is in the synchronized timing state, the TOD clock is in synchronization with the timing-network reference time as defined below:

- If the configuration is in STP timing mode, the configuration is synchronized with coordinated server time (CST).
- If the configuration is in ETR-timing mode, the configuration is synchronized with the ETR

A configuration that has not been initialized or is in local-timing mode is never in the synchronized state.

▶ Unsynchronized State:

When a configuration is in the unsynchronized timing state, the TOD clock is not in synchronization with the timing network reference time as defined below:

- If the configuration is in STP timing mode, the configuration has lost or has not been able to attain synchronization with coordinated server time (CST). The configuration is out of synchronization with CST when the TOD clock differs from CST by an amount that exceeds a model dependent STP-sync-check-threshold value.
- If the configuration is in ETR-timing mode, the configuration has lost synchronization with the ETR.

▶ Stopped State:

When a configuration is in the stopped timing state, either the TOD clock is in the stopped state, or TOD-clock recovery is in progress. After TOD-clock recovery completes, the TOD clock enters either the synchronized or unsynchronized state. This state is only used in ETR mode when transitioning from unsynchronized to synchronized timing state.

### STP clock source state

The STP clock source state indicates whether a usable STP clock source is available. The STP clock source is used to determine the Coordinated Server Time required to be able to synchronize the TOD clock.

▶ Not usable

The *not usable* STP clock source state indicates that a usable STP clock source is not available to STP. When a usable STP clock source is not available, CST cannot be determined.

▶ Usable

The *usable* STP clock source state indicates that a usable STP-clock source is available to STP. When a usable STP-clock source is available, CST has been determined and can be used to synchronize the TOD clock to the STP network.

For further details, please refer to z/Architecture Principles of operations Manual SA22-7832.

### 1.3.3 STP Servers

In this section we discuss STP servers. Throughout the rest of this chapter we discuss STP configurations, since Mixed CTNs or ETR configuration are not available with the most current IBM Z server generations (IBM z13 and newer).

### STP enabled

An STP-enabled server is an STP-capable server that has the STP function enabled. Even after LIC to support STP is installed on a server, the STP function cannot be used until it is enabled.

### STP configured

An STP-configured server is a server that has been configured to be part of a CTN by assigning it a CTN ID. When the STP network ID field is not specified, the server is not configured to be part of a CTN, and is therefore not an STP-configured server. CTN and CTN ID are discussed in the next section.

## 1.3.4 STP Stratum levels

The term *stratum* is found throughout this document and within the Hardware Management Console panels.

**Important:** NTP and its implementations also use the term “stratum”, but the NTP stratum is not related to the STP stratum. STP stratum refers to timing messages distributed within a Coordinated Timing Network, across IBM Z CPCs participating in the respective CTN.

Unless otherwise specified, “stratum” refers to STP stratum.

STP distributes time messages in layers, or *stratums*. The top level (*STP stratum 1*) retrieves time information from the External Time Source (ETS) and distributes time messages to the layer immediately below it (*STP stratum 2*). STP stratum 2 in turn distributes time messages

to STP stratum 3, and STP stratum 3 provides time distribution to STP stratum level 4 CPCs. More layers are conceivably possible, but the current STP implementation is limited to four layers. Figure 1-3 on page 13 illustrates the stratum concept.

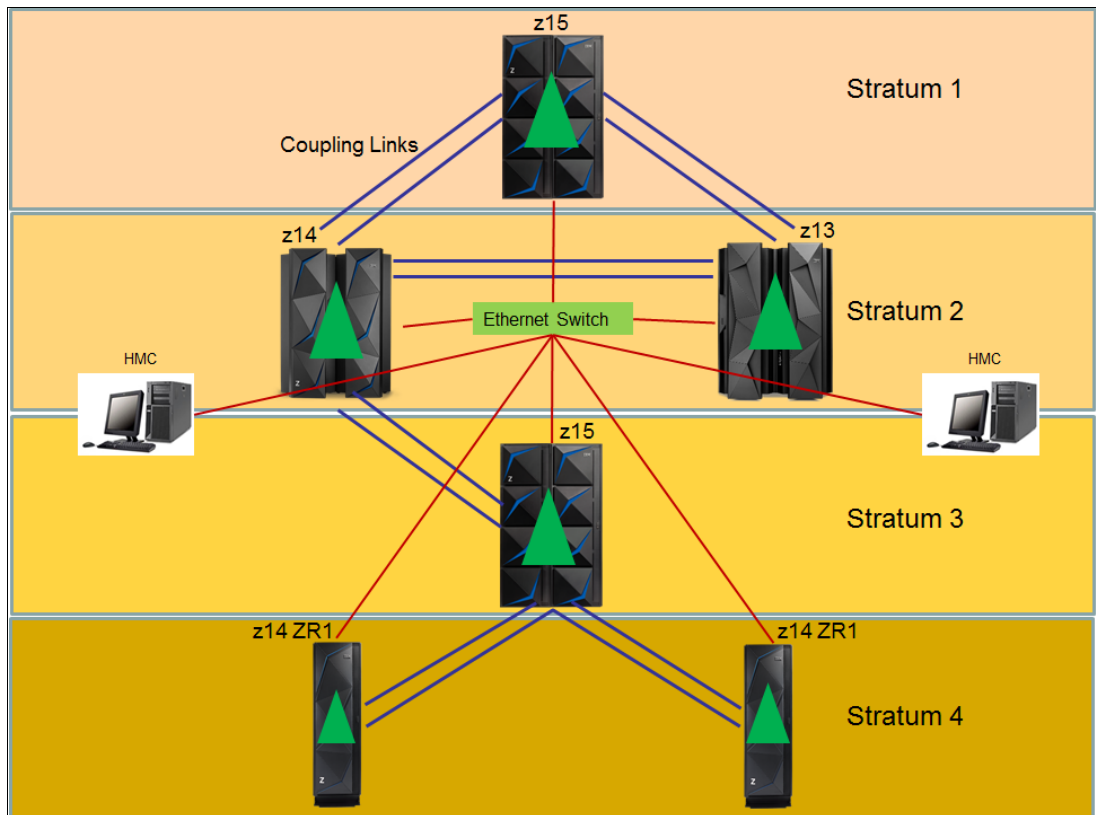


Figure 1-3 STP Stratum levels

### STP Stratum levels (1, 2, 3, 4, and 0)

In a timing network based on STP, stratum levels are used as a means to define the hierarchy of a server in the timing network. A stratum 1 server is the highest level in the hierarchy in the STP network.

The stratum 1 level is determined indirectly in one of the following ways:

- ▶ In an STP-only CTN, the roles of Preferred and Current Time Servers are assigned, and the Current Time Server becomes the active stratum 1 server. Details regarding Preferred and Current Time Servers are described in Section 1.3.5, “Server roles in a CTN” on page 14.

Stratum 2 and stratum 3 levels are determined by how many stratum levels they are away from stratum 1. A server that uses STP messages to synchronize to a stratum 1 server is referred to as a stratum 2 server. Similarly, a server that uses STP messages to synchronize to a stratum 2 server is referred to as a stratum 3 server.

With the introduction of the IBM Z z14 server family, IBM announced support for stratum level 4. Stratum level 4 servers have no direct STP link connections to the stratum level 1 server. The stratum level 4 support was introduced to help alleviate complexity in migration scenarios.

**Note:** This additional stratum level (STP stratum 4) should only be used as a temporary state during CTN reconfiguration (during migration operations). Customers should not run with machines at Stratum level 4 for extended periods of time because of the lower quality of the time synchronization.

### ***STP stratum 0***

The *STP Stratum 0* refers to a server (IBM Z CPC) which is in *unsynchronized* state. In the un-synchronized state the CPC does not receive timing signal (either from an external time source or from another server in a CTN).

## **1.3.5 Server roles in a CTN**

A (STP-only) CTN can have only one active stratum 1 server. The following definitions explain the roles that must be assigned for certain servers in a CTN. CTN configuration is performed through the HMC panels for all participating servers.

**Important:** IBM z14® is the last server to support Support Element “Sysplex Timer” task (HMC/SE 2.14.1). Newer systems (z15™, HMC/SE driver 2.15.0 and newer) only support the HMC task “Manage System Time” (the SE task has been discontinued for z15). As a consequence, it is mandatory that the HMC managing the Coordinated Timing Network is at the latest level (of the newest server available for STP configuration).

### **Preferred Time Server (PTS)**

Using the Manage System Time HMC task, a server must be assigned as the preferred server to assume the stratum 1 server role of a CTN. This is the Preferred Time Server. This server should have connectivity to the Backup Time Server and the Arbiter, as well as to all servers that are planned to be stratum 2 servers. The connectivity can be either ICA SR or CE LR links, IFB links (for older servers, ISC-3 links in peer mode, ICB-3 links, and ICB-4 links).

### **Backup Time Server (BTS)**

Although optional, we strongly recommend also assigning a Backup Time Server whose role it is to take over as the stratum 1 server should the PTS fail. The Backup Time Server is a stratum 2 server that should have connectivity to the Preferred Time Server and the Arbiter, as well as to all other stratum 2 servers that are connected to the Preferred Time Server.

### **Current Time Server (CTS)**

The Current Time Server is the active (STP) stratum 1 server in a CTN and is used to steer the Coordinated Server Time (CST) to its External Time Source (ETS).

#### **Notes:**

- ▶ There can only be one active (STP) stratum 1 server in a CTN, and that is the Current Time Server. Only the Preferred Time Server or the Backup Time Server can assume the Current Time Server role and be the active Stratum 1 server.
- ▶ After a recovery situation is resolved and full connectivity is restored, STP will set the CTS to the PTS server, provided the PTS has connectivity to all systems that are connected to the BTS.

The Current Time Server can only be assigned to either the Preferred Time Server or the Backup Time Server. We recommend that the Current Time Server be assigned to the

Preferred Time Server when the configuration is being initialized. Subsequently, if there is a need to reassign the server roles, the Current Time Server (stratum 1) role can be concurrently assigned to the Backup Time Server. This action can be part of a planned reconfiguration of the Preferred Time Server as long as the planned action is not *disruptive*.

### Arbiter

Optionally, the CTN may have a system (IBM Z CPC) with the assigned the role of *Arbiter*. The Arbiter provides additional means for the Backup Time Server to determine whether it should take over the Current Time Server role in the case of unplanned events that affect the CTN. The Arbiter is a stratum 2 server that must have connectivity to both the PTS and the BTS.

## 1.4 Coordinated Timing Network

A Coordinated Timing Network (CTN) contains a collection of servers that have their time synchronized. The time is synchronized to a value called Coordinated Server Time (CST). The CST represents the time for the entire network of servers (CTN).

The servers that make up a CTN are all configured with a common identifier, referred to as the *CTN ID*. Only servers with the same CTN ID are allowed to become members of the CTN. Servers with different CTN IDs can be members of other Coordinated Timing Networks. All servers in a CTN maintain an identical set of time control parameters that are used to coordinate the TOD clocks.

A CTN can be configured as either:

- ▶ STP-only CTN

An STP-only CTN is a timing network in which all servers are configured to be in STP timing mode. It can only be configured with STP-capable servers, and none of the servers can be in ETR timing mode.

- ▶ Mixed CTN

A Mixed CTN allows the coexistence of servers and CFs synchronized in an ETR network, with servers and CFs that are synchronized with CST. The Sysplex Timer provides the timekeeping information in a Mixed CTN.

#### Notes:

- ▶ Support for the 9037 Sysplex Timer has ended with the IBM z10 server family.
- ▶ The IBM zEC12 and zBC12 are the last IBM Z servers that support being part of a Mixed CTN.
- ▶ IBM Z servers starting with z13 support STP timing *only* (no Mixed CTN).

### 1.4.1 CTN ID

The CTN ID is an identifier that is used to indicate whether the server has been configured to be part of a CTN and, if so configured, identifies the CTN.

- ▶ A field that defines the STP network ID<sup>7</sup>

The field has the format STP network ID, as illustrated on Figure 1-4 on page 16

<sup>7</sup> STP ID consist of 1 to 8 characters, alphanumeric and symbols.

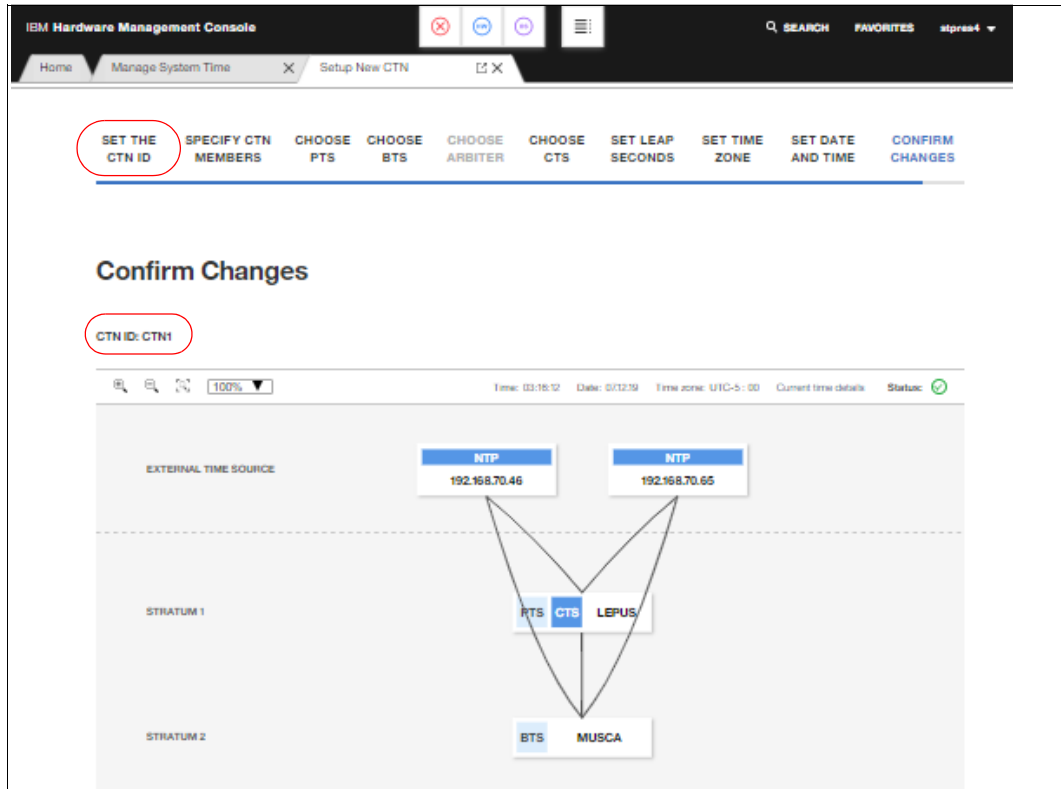


Figure 1-4 Display CTN ID

The Hardware Management Console (HMC) is used to enter the CTN ID.

## 1.5 Coupling Links

This chapter describes the connectivity options that support IBM Parallel Sysplex clustering technology and common time on IBM Z platforms. It covers the following topics:

- ▶ IBM Z Parallel Sysplex
- ▶ Connectivity options

### IBM Z parallel sysplex

Parallel Sysplex brings the power of parallel processing to business-critical applications. A Parallel Sysplex cluster consists of up to 32 IBM z/OS images, connected to one or more coupling facilities (CFs) using high-speed specialized links, called *coupling links*, for communication and time-keeping. The coupling facilities at the heart of the cluster enable high-speed record-level read/write data sharing among the images in a cluster.

Coupling links support communication between z/OS and coupling facilities. The coupling facility provides critical locking/serialization, data consistency, messaging and queueing capabilities that allows the systems in the sysplex to coordinate and share data.

A well configured cluster has no single point of failure and can provide end users with near continuous application availability over planned and unplanned outages.

### Connectivity options

z15 and z14 ZR1 support three coupling link types:

- ▶ Integrated Coupling Adapter Short Reach (ICA SR) links (8x optical) connect directly to the CPC drawer and are intended for short distances between CPCs of up to 150 meters.
- ▶ Coupling Express Long Reach (CE LR) adapters (1x optical) are located in the PCIe+ I/O drawers and support unrepeated distances of up to 10 km or up to 100 km over qualified WDM services.
- ▶ Internal Coupling (IC<sup>8</sup>) links are for internal links within a CPC.

In addition to these three, the IBM z14, z13 and z13s<sup>®</sup> servers also support:

- ▶ 1 x IFB links, optical link, long wave
- ▶ 12 x IFB links, optical links, short wave

**Note:** Parallel Sysplex supports connectivity between systems that differ by up to two generations (n-2). For example, an IBM z15™ can participate in an IBM Parallel Sysplex cluster with z14, z14 ZR1, z13, and z13s systems.

However, the IBM z15 and IBM z14 ZR1 do not support InfiniBand connectivity. These servers (z15 and z14 ZR1) support connectivity the z13 and z13s (Driver 27 with MCLs) using Integrated Coupling Adapter Short Reach (ICA SR) and Coupling Express Long Reach (CE LR) features.

Figure 1-5, shows the supported Coupling Link connections for the z15.

- ▶ Infiniband links are supported between z13, z13s and z14 M0x machines
- ▶ Only ICA SR and CE LR links are supported on z15 and z14 ZR1 machines

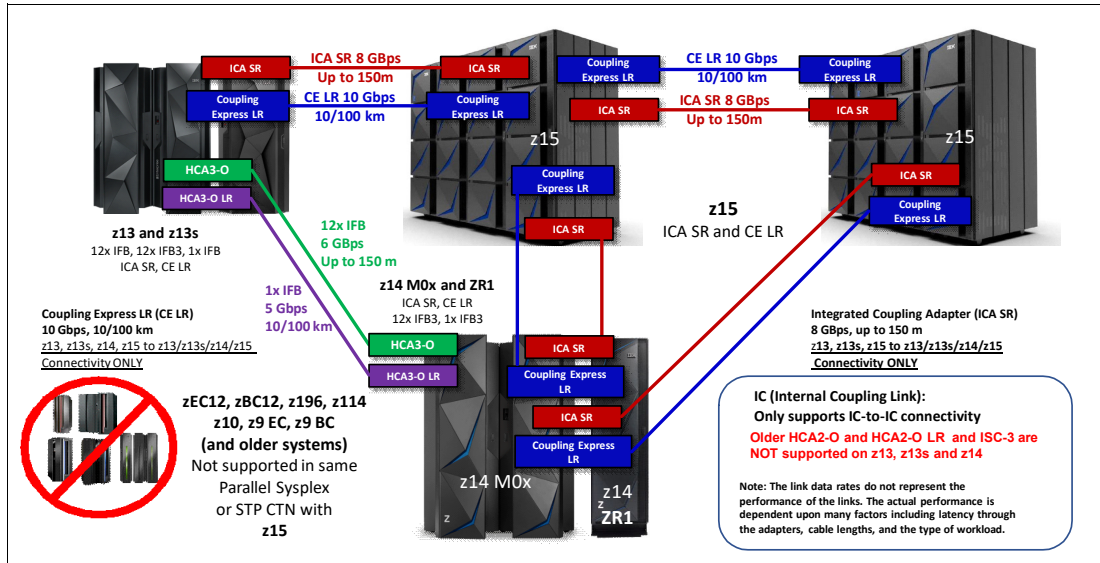


Figure 1-5 Parallel Sysplex connectivity options

- ▶ For further details concerning the supported coupling links, refer to the *IBM Z Connectivity Handbook*, SG24-5444.

<sup>8</sup> Internal Coupling Links (IC) do not carry STP (timing) signals.

## 1.6 External Time Source

**Note:** For an introduction to the Precision Time Protocol and additional information (terminology, implementation overview), see Appendix A, “IEEE 1588 Precision Time Protocol (PTP) Introduction and Overview” on page 121.

The External Time Source (ETS) is used as a reference to which the Coordinated Server Time (CST) is steered (kept in sync.). The current ETS implementation provides time information to the CTN using either:

- ▶ Precision Time Protocol (PTP) or Precision Time Protocol with Pulse Per Second (PPS)

**Note:** Precision Time Protocol is supported on z15 or newer CPCs. It is NOT supported as External Time Source on z14 and earlier CPCs.

- ▶ Network Time Protocol (NTP) or Network Time Protocol with Pulse Per Second (PPS).

### PTP grandmaster clock support

The IBM Z platform can be configured to obtain system time from a PTP Grandmaster (server). The PTP server is configured using the Configure External Time function from the Manage System Time task on the HMC. An ETS must be configured for the Current Time Server (CTS). Redundant connectivity should be configured for ETS availability.

An ETS does not need to be configured for systems that are not the PTS or the BTS. However, if an ETS is configured for one of these servers, its configuration is saved. At a later time, in case the role of the system is changed (CTN reconfiguration from HMC) to either the PTS or BTS, the ETS associated with the CPC can be used.

### PTP Grandmaster with PPS support

STP can use for the External Time Source (ETS) a PTP grandmaster that has a PPS output signal. External time source devices with PPS are available worldwide from several network timing solutions vendors.

The SE ports must be connected to the PTP grandmaster, either directly, or via a network device (i.e. LAN switch) supporting PTP. The PTP ordinary clock code is running on the SE. In addition, the PPS output of the same grandmaster server must be connected to the PPS port on the oscillator (OSC/FSP) card on the Z platform (z15 and later systems). Redundant PPS ports are supported.

The PPS signaling provides a highly accurate event but no time information, combining the two is what provides the system with a highly accurate timestamp.

### NTP server support

The IBM Z platform can be configured to obtain system time from an NTP server. The NTP server address is configured using the Configure External Time function from the Manage System Time task on the HMC. It can also be set on the ETS tab of the System (Sysplex) Time task from the SE (z14 and earlier CPCs only). An ETS must be configured for the Current Time Server (CTS). On systems prior to z15, configuring the Preferred NTP and Secondary NTP servers for the PTS and BTS reduces the risk of an STP-only timing network losing its time source. An ETS does not need to be configured for systems that are not the PTS or BTS. If an ETS is configured, its configuration is saved in case the role of the system changes to the PTS or BTS.



## NTP server with PPS support

STP can use an NTP server that has a PPS output signal as its external time source. This type of external time source device is available worldwide from several vendors that provide network timing solutions.

The NTP output of the NTP server must be connected to the SE LAN because the NTP client is running on the SE. In addition, the PPS output of the same NTP server must be connected to the PPS a port on the oscillator (OSC) card on Z platform.

While NTP provides the time information with relatively low accuracy (~100 milliseconds), the PPS signaling provides a highly accurate event but no time information, combining the two is what provides the system with a highly accurate timestamp.

See Figure 1-2 on page 6 for a schematic drawing of a STP deployment using PPS.

**IEEE 1588 Precision Time Protocol (PTP)<sup>a</sup>:** The initial support implements PTP connectivity via the IBM Z HMC/SE (z15 CPCs only). There is no change to the use of STP CTNs for time coordination, other than the potential to use a PTP-based external time source.

Future implementation is planned to include full connectivity of an external PTP time source directly to the IBM Z CPC, and re-introduction of the concept of a mixed CTN, with support for traditional STP and native PTP implementations. Beyond that, the goal is to enhance the role of IBM Z machines in a PTP environment that addresses the many governmental regulations and security concerns that our clients are facing.

For a brief introduction to PTP, see Appendix A, “IEEE 1588 Precision Time Protocol (PTP) Introduction and Overview” on page 121.

- a. Statements by IBM regarding its plans, directions, and intent are subject to change or withdrawal without notice at the sole discretion of IBM. The development, release, and timing of any future features or functionality described for IBM products remain at the sole discretion of IBM.

## 1.7 Firmware implementation

In this section we describe the implementation of the CPC TOD clock.

### TOD processing

The CPC TOD clocks of all the CPs are automatically set during CPC activation. The time reference used depends on whether Server Time Protocol (STP) is enabled. When STP is enabled, a CPC can participate in a STP CTN. In this case the Current Time Server for the STP CTN provides the time information.

### Server Time Protocol not enabled

During PR/SM initialization, the CPC TOD clocks for each processing unit (PU, or core) are set to the TOD value of the Support Element. Each logical partition (LPAR) starts out with this CPC TOD clock value at the completion of LPAR activation. The operating system running in an LPAR can set a TOD value for itself and this is the only TOD reference it sees. Setting the TOD clock for one logical core in the LPAR sets the TOD clock for all logical cores in that LPAR, but does not affect the logical cores in any other LPAR. The TOD clock value is used for the duration of the LPAR activation, or until a subsequent Set Clock instruction is issued in the LPAR.

## Server Time Protocol enabled

The enablement of STP is supported. Also, during PR/SM initialization, when STP is enabled, the CPC TOD clocks for each PU are set to the TOD value from STP.

The operating system in each LPAR can independently choose whether to synchronize to the current time source for STP, if present. Operating systems in LPARs that do synchronize to STP run with identical TOD values. Operating systems in LPARs that do not synchronize to STP do not need to be aware of the presence of STP and can set their TOD values independently of all other LPARs.

z/OS does not allow you to change the value of the TOD setting when synchronized to STP (STPMODE=YES in the CLOCKxx parmlib member).

The IBM z15 server supports the specification of a logical partition time offset. When all members of a sysplex are in logical partitions on these supported models, the logical partition time offset can be used for:

- ▶ Different local time zone support in multiple sysplexes using the STP Coordinated Timing Network (CTN). Many sysplexes have the requirement to run with a LOCAL=UTC setting in a sysplex (STPMODE=YES) where the time returned from a store clock (STCK) instruction yields local time. To fulfill this requirement, the time initialized for the STP CTN must be local time. With logical partition time offset support, multiple sysplexes can each have their own local time reported to them from a STCK instruction if wanted. For instance, the STP CTN can be set to GMT, one set of sysplex partitions could specify a logical partition time offset of minus 5 hours, and a second set of sysplex partitions could specify a logical partition time offset of minus 6 hours.

External coupling links are also valid to pass time synchronization signals for Server Time Protocol (STP). Therefore the same coupling links can be used to exchange timekeeping information and Coupling Facility messages in a Parallel Sysplex.

Having the TOD clock implemented in the hardware layer provides us with the following advantages:

- ▶ Since all LPARs running on an IBM Z server are fully virtualized, we enable a single LPAR or a cluster of LPARs to run with the exactly the same time derived from the same hardware source, or with a time that is derived from the hardware with a specific offset added.
- ▶ By distributing timing information via highly efficient coupling links, IBM Z servers in the same CTN run on a highly accurate time synchronized within microseconds.



# 2

## Planning hardware and software

This chapter discusses the following planning concepts and considerations:

- ▶ Planning server hardware
- ▶ Planning connectivity
- ▶ External time source
- ▶ Internal Battery Feature (IBF)
- ▶ Planning z/OS software
- ▶ Planning for z/VM, z/TPF and Linux on IBM Z
- ▶ Special considerations for a single server CTN
- ▶ Migration planning

## 2.1 Planning server hardware

STP is a server-wide facility that is implemented in the LIC of all IBM Z servers since z9. Prior to z9 the time reference for a sysplex was provided by an external hardware device called IBM Sysplex Timer, thus the migration from a sysplex receiving the time signal from a sysplex timer, over a sysplex with a subset of servers still connected to a Sysplex Timer (called a Mixed CTN) to a an STP-only CTN had to be carefully planned and has been discussed in deep details in previous STP redbooks (SG24-7280, SG24-7281, and SG24-7380).

The IBM zEC12 and zBC12 CPCs were be the last IBM Z servers to support connections to a Mixed CTN. Moreover, IBM Z servers can coexist in the same sysplex and the same CTN with N-2 generation servers.

For the latest IBM Z server, IBM z15, a Sysplex can only consist of z13, z13s, z14, z14 ZR1 and z15 servers. Older servers like zEC12, zBC12, z196, z114, z10EC, z10BC, z9EC and z9BC servers a cannot be in the same Sysplex and the same CTN with a z15 server. IBM z13® and later servers can only be part of STP-only CTNs.

This section discusses the planning for the servers and the related Hardware Management Console (HMC), Support Element (SE) and microcode, focussing on the following IBM Z CPCs: z13, z13s, z14, z14 ZR1 and z15.

### 2.1.1 Servers

An IBM Z server requires FC 1021 (Server Time Protocol Enablement) to participate in a CTN. The N-2 generation coexistence rule applies both the Coordinated Time Network and IBM Parallel sysplex membership: only servers which are not older than two generations before the newest server in the sysplex and CTN are supported in the same sysplex and CTN.

In a Sysplex with an IBM z15 server, only the following servers are allowed:

- ▶ IBM z15
- ▶ IBM z14
- ▶ IBM z14 ZR1
- ▶ IBM z13
- ▶ IBM z13s®

Older servers like IBM zEnterprise® EC12 (zEC12), IBM zEnterprise BC12 (zBC12), IBM zEnterprise 196 (z196), IBM zEnterprise 114 (z114) or earlier are also capable to run in a sysplex and CTN, but cannot be part of a sysplex or CTN with an IBM z15.

Servers used as stand-alone coupling facilities must be STP enabled and run in a CTN. For stand-alone coupling facilities the same N-2 generations coexistence rule applies. Table 2-1 provides a summary of the servers and the minimum Coupling Facility Control Code (CFCC) level required to be allowed in a sysplex and CTN with a IBM Z z15 server.

Table 2-1 Coupling facility requirements

Server	CFCC level required on the coupling facility
z13 and z13s	CFCC LEVEL 20 and higher
z14 and z14 ZR1	CFCC LEVEL 22 and higher
z15	CFCC LEVEL 24

## 2.1.2 Hardware Management Console

The HMC provides the user interface to manage a CTN.

With the introduction of z14 (HMC 2.14.0, driver level 32) STP management has been implemented as graphical interface on the HMC (as the Manage System Time task). With z15 (HMC 2.15.0, driver level 41) the graphical interface for STP has been further enhanced, and the Sysplex Timer menus on the SE have been discontinued. As such, the HMC graphical interface for STP is the *only* way to manage a CTN that contains an IBM z15 CPC.

### HMC functionality in an CTN

In a CTN, the HMC is used to:

- ▶ Initialize or modify the CTN ID.
- ▶ Initialize the time manually or by using an external time source to keep the Coordinated Server Time (CST) synchronized to the time source provided by the external time source (ETS).
- ▶ Initialize the time zone offset, daylight saving time offset, and leap second offset.
- ▶ Configure the HMC as an NTP server.
- ▶ Configure access to an External Time Source (ETS) so that the CST can be steered to an external time source:
  - To configure the connection from the PTP client (ordinary clock) on the SE to PTP Grandmaster
  - To configure the connection from the PTP client (ordinary clock) on the SE to PTP Grandmaster servers with pulse per second (PPS)
  - To configure the connection from the NTP client on the SE to NTP servers
  - To configure the connection from the NTP client on the SE to NTP servers with pulse per second (PPS)
- ▶ Manage CTN membership (add, remove servers, restrict/unrestrict membership, split and merge CTNs)

**Note:** Split and merge operations are only supported for CTNs with z14 and newer servers.

- ▶ Assign the roles of Preferred, Backup, and Current Time Servers, as well as Arbiter.
- ▶ Adjust time by up to +/- 60 seconds.

#### Notes:

- ▶ Adjustments take approximately 8 hours to steer out one second so steering out 60 seconds would take a little less than 20 days.
- ▶ Even though time adjustments of more than 60 seconds can be performed, these need to be done in 60 second increments.

- ▶ Schedule changes to offsets previously listed (time zone, daylight savings, leap seconds). STP can automatically schedule daylight saving time based on the selected time zone.

- ▶ Monitor the status of the CTN.
- ▶ Monitor the status of the coupling links initialized for STP message exchanges.

The STP HMC GUI is used for the following operation:

- ▶ Setup a new CTN.
- ▶ Configure an External Time Source.
- ▶ Add systems to the CTN.
- ▶ Modify assigned server roles.
- ▶ Remove systems from a CTN.
- ▶ Unconfigure a CTN.
- ▶ Export CTN data (in .xls format).
- ▶ Set CTN member restrictions.

Also these advanced actions can be performed:

- ▶ Join existing CTN.
- ▶ Split to new CTN.
- ▶ View the status of the Pulse Per Second signal (whether or not a CPC is receiving such a signal)
- ▶ Save STP debug data.

### Eligible Hardware Management Consoles

The HMC application level must be equal to or higher than the level of the Support Element consoles that it manages. For example, if a z14 is part of the CTN, HMC application V2.14.0 or later is required. In current CTN configurations including a z15, a the HMC application level 2.15.0 is required.

Historically, at a minimum, management of STP-enabled servers requires HMC application level V2.9.2 with the latest MCL, which has been provided with z9 servers. Some functions have been added or removed during the years in subsequent HMC levels. For example, NTP server function has been made available in HMC application level V2.10.1 and later, while HMC application level V2.12.0 and later no longer support dial-out capability, and HMC application level V2.14.0 provided the new STP graphical interface, which became the only way to define and manage an STP with the removal of the STP panels from the SE with driver 41 (HMC V2.15.0).

Table 2-2 shows the major STP-related changes and the HMC Application levels that introduce them.

Table 2-2 HMC STP-related changes

Change	HMC application level	Supported servers
Support for PTP configuration	V2.15.0	z15 only
Removal of STP panels from the SE	V2.15.0	z15 and earlier N-2 server generations.
CTN Split and Merge	V2.14.1	z14, z14 ZR1 and earlier (N-4) generations
Support for CE LR connections (rolled back to z13, no Going-Away-Signal) Additional Stratum Level 4 Graphical Interface for STP	V2.14.0	z14 and previous (N-4) generations
Removed Mixed CTN support Support for ICA connections HMC STP Panel Enhancements.	V2.13.0 and 2.13.1	z13, z13s and earlier (N-4) generations
HMC NTP authentication Removed dial-out support	V2.12.0 and V2.12.1	zEC12, zBC12 and earlier (N-4) generations

Change	HMC application level	Supported servers
Removed ETR support	V2.11.0 and V2.11.1	z196, z114 and earlier (N-4) generations
NTP server in the HMC	V2.10.1 and V2.10.2	z10 EC, z10 BC and earlier
STP support	V2.9.2	z9 EC, z9BC and earlier

Previously installed HMCs can be upgraded by requesting your IBM Service Representative to order the appropriate ECA for each server that has one or more HMC features that must be upgraded. z15 requires feature code 0062 (HMC Tower), 0063 (HMC Rack Mount) or 0100 (IBM Z Hardware Management Appliance) and HMC Application level V2.15.0.

### **Setting up the HMC for local and remote operation**

The management functions provided by the HMC require that all CPCs that need role assignment in the CTN to be defined objects to any HMC used to manage the CTN. Furthermore, we recommend that all CPCs in the CTN to be defined to any HMC in order to enable CTN reconfigurations (for instance, within a recovery scenario). Even remote HMCs should have all CPCs defined in the CTN.

**Important:** CPCs that are not defined to the HMC used to manage the CTN will not appear in the STP graphical interface (Manage System Time task on the HMC), except when they already have a role assigned (PTS, BTS, or Arbiter). If a CTN reconfiguration is required, it will not be possible to assign a role to the servers not known to the HMC. For role assignment, the HMC needs access to both the CTS and the target server.

There are two methods of defining the server (CPC) objects to the HMC:

- ▶ When the HMC is on the same local LAN as the SE servers, local HMC automatically detects the Support Elements and sets up all the necessary internal configuration information for communication without requiring additional information from the users.
- ▶ When the HMC is on a remote LAN, defining the servers requires manual (user) action.

For details and guidance consult the manual *Hardware Management Console Operations Guide* for the Driver level installed on your HMC. This publication can be found on the IBM Resource Link website at:

<http://www.ibm.com/servers/resourceLink/>

#### **Important:**

- ▶ To define the CPCs to the HMC, the HMC and the SEs must be part of the same HMC domain security.
- ▶ Always follow the IBM guidelines specified in each server documentation to connect HMCs to SEs within a site and between sites.

### **HMC application programming interface (API)**

Specific STP-related attributes have been added to the Defined CPC objects starting with HMC Version 2.10.1. The following STP commands were added in order to manage the CTN or to be able to perform automated recovery actions such as reassigning the PTS, BTS, or Arbiter role in case of a server or site failure:

- ▶ **Swap Current Time Server**
- ▶ **Set STP Configuration**
- ▶ **Change STP-only CTN**

- ▶ **Join STP-only CTN**
- ▶ **Leave STP-only CTN**

For object command details refer to *SNMP Application Programming Interfaces*, SB10-7171.

### **HMC security**

Security for remote HMC connectivity is provided by the HMC user logon procedures (similar to local HMC access). As with a local HMC, all communications between a remote HMC and the SEs it manages is encrypted. Certificates for secure communications are provided and can be replaced if desired. TCP/IP access to the remote HMC is controlled through the HMC internally managed firewall and is limited to HMC-related functions.

**Note:** The HMC Domain Security can be used to isolate servers on a common LAN, or to provide additional security. Individual remote users can be configured to restrict access in the same way that they can be configured on a local HMC.

For a complete description of HMC and Support Element connectivity options refer to the server installation manuals:

- ▶ z15: *8561 Installation Manual for Physical Planning*, GC28-7002
- ▶ z14 ZR1: *3907 Installation Manual for Physical Planning*, GC28-6974
- ▶ z14: *3906 Installation Manual for Physical Planning*, GC28-6965
- ▶ z13s: *2965 Installation Manual for Physical Planning*, GC28-6953
- ▶ z13: *z13 Installation Manual for Physical Planning*, GC28-6938

## **2.1.3 Support Elements -- Connectivity to ETS**

The Support Elements are used for STP management and functionality. The SE's internal storage provides STP initialization, configuration and timing parameters. STP configuration information is mirrored to the alternate SE in case the primary SE encounters a problem.

The STP ID and PPS Port States are stored on the SE and preserved through a power-on reset (POR) and power off/on sequence. Server roles (PTS, BTS, Arbiter) are not preserved across a power-on reset or power off/on sequence of the CTN unless the CTN is configured as a restricted CTN. CTN membership restriction must be configured for preserving STP role information across PORs.

**Note:** Configuration of a CTN that has *been restricted* is described in detail in 4.2.2, "Single-server or dual-server auto resume after power-on reset" on page 108. This is also known as *saving CTN configuration across power-on reset operations*.

Disruptive actions, such as power-on reset, are only allowed for servers without roles, or for a server in a single-server CTN.

The server roles must be reassigned after:

- ▶ The CTS in a single server, unrestricted CTN has been power-on reset or has gone through a power off/on sequence.
- ▶ The PTS and BTS, if assigned, in an unrestricted CTN with two or more servers which experience a power outage at the same time (for example, a data center power outage).
- ▶ A CTN is deconfigured intentionally by the user.

By default, the Support Element data is mirrored automatically each day or when internal code changes are installed through single-step internal code changes.



When the external time source is configured to use a PTP Grandmaster or an NTP server, with or without pulse per second, the client code is on the Support Element. Access to the external time server is automatic and does not have to be scheduled, as described in 2.3, “External time source” on page 35.

## 2.2 Planning connectivity

This section focuses on hardware connectivity requirements, offering planning information and recommendations.

### 2.2.1 Planning for coupling links for STP

STP identifies and maintains a list of all data links on the server that are capable of STP message communication. Coupling peer data links are the only data links considered to be STP-capable data links.

Since the implementation of STP with z9, there is an ongoing evolution of external coupling links which are available to exchange STP messages to keep servers and CFs synchronized:

- ▶ Inter System Channel (ISC) links
- ▶ Integrated Cluster Bus (ICB) links
- ▶ InfiniBand (IFB) links
- ▶ Integrated Coupling Adapter Short Reach (ICA SR) links
- ▶ Coupling Express Long Reach (CE LR) links

There are significant differences between these type of links with respect to the speed of the link, the maximum supported number of links and the maximum supported distance. From an availability point of view, all types should be considered to be equivalent, with the same considerations and recommendations. IBM z15 and z14 ZR1 servers only support CE LR and ICA SR links, while IBM z14, z13 and z13s servers also support HCA-O (InfiniBand) links.

Older coupling link connectivity options such as ISC and ICB will not be discussed in more detail.

Figure 2-1 shows the coupling link connectivity options for a parallel sysplex with a z15.

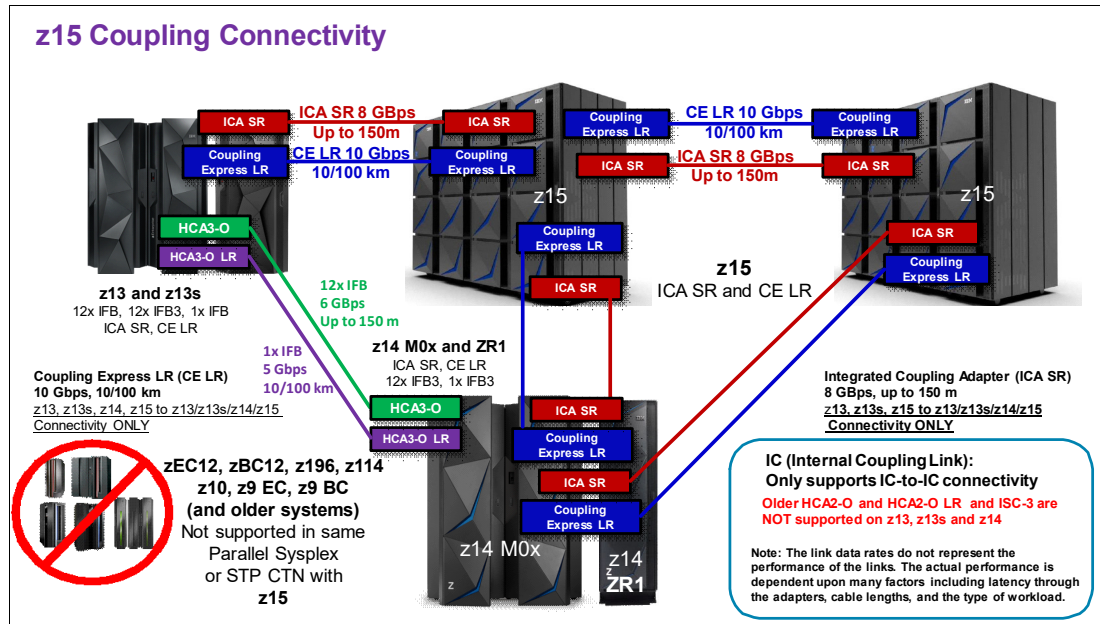


Figure 2-1 Parallel Sysplex connectivity options with z15

## ICA SR links

ICA SR links are high-speed external coupling links available on z15, z14, z14 ZR1, z13 and z13s. ICA SR features are two-port, short-distance coupling adapters.

By feature code, ICA SR coupling links are:

- ▶ FC 0172 - ICA SR
- ▶ FC 0176 - ICA SR1.1

The ICA SR feature (FC 0172) was introduced with the IBM z13. z15 introduces the ICA SR1.1 feature (FC 0176). Both ICA SR and ICA SR1.1 use channel type (CHPID) type CS5. The ICA SR and ICA SR1.1 use PCIe Gen3 technology, with x16 lanes that are bifurcated into x8 lanes for coupling. Both ICA SR and ICA SR1.1 have two ports/links per feature.

The ICA SR and SR1.1 are designed to drive distances up to 150 m and supports a link data rate of 8 GBps. It is designed to support up to four CHPIDs per port and seven subchannels (devices) per CHPID.

More information about planning for configuration of ICA coupling links is available in *IBM Z Connectivity Handbook*, SG24-5444, *IBM z15 Configuration Setup*, SG24-8860, and *IBM Z Planning for Fiber Optic Links*, GA23-1407.

## CE LR links

CE LR links are high-speed external coupling links available on z15, z14, z14 ZR1, z13 and z13s. The CE LR coupling link is hosted in a PCIe I/O drawer and allows the supported IBM Z systems to connect to each other over extended distance (10km/6.2 miles w/o repeaters).

By feature code, CE LR coupling links are:

- ▶ FC 0433 - Coupling Express LR (two ports/links per feature).

CE LR is defined in IOCDs similar to IFB and ICA links, using CHPID type CL5. Even though this feature is a PCIe feature, a PCHID is used instead of an AID to identify the physical card. More information about planning for configuration of CE LR coupling links is available in

*IBM Z Connectivity Handbook, SG24-5444, IBM z15 Configuration Setup, SG24-8860, and IBM Z Planning for Fiber Optic Links, GA23-1407.*

### **InfiniBand (or IFB) links**

IFB links are high-speed external coupling links available on z14 M0x servers (but not on z14 ZR1), z13, z13s and older IBM Z servers. IFB links consist of a Host Channel Adapter (HCA) coupled directly to the processor I/O interface.

By feature code, InfiniBand coupling links available on z14 M0x, z13 and z13s are:

- ▶ FC 0170 - HCA3-0 LR PSIFB 1x (four ports/links per feature).
- ▶ FC 0171 - HCA3-0 PSIFB 12x (two ports/links per feature).

FC 0171 is a 12x (12 lanes of fiber in each direction) InfiniBand-Double Data Rate (IB-DDR) coupling link designed to support a total link data rate of 6 Gbps. FC 0171 is the IFB LR link, which supports a 1x (one lane of fiber in each direction) InfiniBand-Double Data Rate (IB-DDR) or InfiniBand-Single Data Rate coupling link for a distance of up to 10 km and is capable of a data rate of 5 Gbps.

More information about planning for configuration of IFB coupling links is available in *IBM Z Connectivity Handbook, SG24-5444, IBM z14 Configuration Setup, SG24-8460, and IBM Z Planning for Fiber Optic Links, GA23-1407.*

### **Adapter ID assignment and VCHIDs**

An adapter ID (AID) is assigned to every IFB and ICA SR link at installation time. It is unique for the CPC. There is only one AID per feature, so all ports on the feature share the same AID. The adapter ID is:

- ▶ A number between 00 and 3B on z15 T01 (with five CPC drawers)
- ▶ A number between 00 and 17 on z15 T02 (two CPC drawers)
- ▶ A number between 10 and 17 on z14 ZR1
- ▶ A number between 00 and 37 on z14 (with all 4 CPC drawers)
- ▶ A number between 10 and 23 on z13s (with all 2 CPC drawers)
- ▶ A number between 10 and 37 on z13 (with all 4 CPC drawers)

In the I/O configuration program (IOCP or HCD), the AID and port number are used to connect the assigned CHPID to the physical location of the feature.

A physical channel identifier (PCHID) normally has a one-to-one relationship between the identifier and a physical location in the machine, however a PCHID in the range from 0500 to 06FF lacks the one-to-one relationship between the identifier and the physical location, either because they do not have a physical card (like Internal Coupling connections (IC)), or because they are administered through different identifiers (as for IFB and ICA links, with the AIDs). No one-to-one relationship is possible due to the capability to define more than one CHPID for a physical location. Therefore, these are sometimes referred to as Virtual Channel Path Identifiers (VCHIDs).

VCHIDs are assigned automatically by the system and are not defined by you in the IOCDs. A VCHID is also not permanently tied to an AID. Therefore the VCHID assignment can change after a Power-On Reset (POR) if the hardware configuration has changed (for example, if a HCA was added or removed). Due to the automatic assignment of the VCHID at every POR, the client or SSR needs to make sure that the correlation for the channel that they intend to manipulate has not changed. The VCHID that is currently associated with a coupling CHPID can be found by issuing an `MVS™ D CF,CFNM=xxxx` command for the associated CF.

You can also find the association between a CHPID and a VCHID, which is still referred to as a PCHID by the HMC and the SE, by displaying the VCHID details using the graphical STP

interface on the HMC by selecting the connection between two servers in the CTN. For example, in the Figure 2-2 on page 30 the connection between the Stratum 1 server LEPUS and the Stratum 2 server MUSCA is displayed, showing the connection from the LEPUS point of view (using NTP as External Time Source).

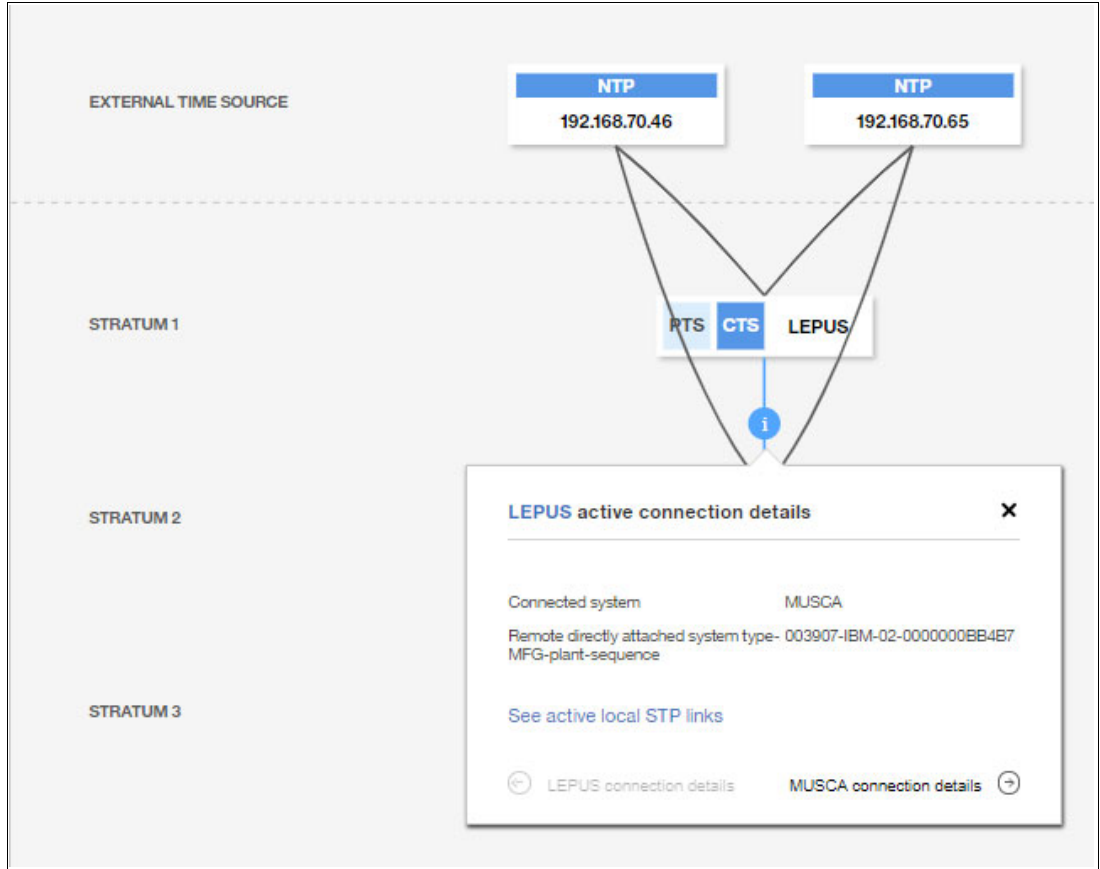


Figure 2-2 Displaying Coupling Link connections in a CTN

By selecting “See active local STP links”, the details for these links are displayed, as shown in Figure 2-3. VCHID 0501 is assigned to Port 02 at AID 0038, VCHID 0503 is assigned to Port 02 at AID 0039.

The screenshot shows a window titled 'LEPUS active local STP links'. It contains a table with the following data:

Channel Type	AID	Port	VCHID
^ Coupling Express LR (1 active links)	0038	02	0501
^ Coupling Express LR (1 active links)	0039	02	0503

Figure 2-3 Displaying Coupling Link connections in a CTN - VCHID information

## Enhanced Drawer Availability (EDA)

The IBM Z servers are designed to allow a single CPC drawer, in a multi-drawer server, to be concurrently removed from the server and reinstalled during an upgrade or repair action. During the process, connectivity to the server I/O resources is provided using a second path from a different CPC drawer. With enhanced drawer availability, and with proper planning to ensure that all resources are still available to run the critical applications in a system with one missing CPC drawer configuration, planned outages can be avoided.

With Server Time Protocol, it is necessary to ensure that coupling links are not impacted during the removal of a CPC drawer. This is especially critical where ICA SR and InfiniBand links are concerned as these are connected directly to the CPC drawers.

Based on the design of redundant I/O interconnect in I/O drawers that allow for connection through multiple paths, CE LR connectivity will not be interrupted. Ensure that all the ICA and IFB links have redundant paths from different books.

This planning should be included as part of the initial installation and any follow-on upgrades that modify the operating environment. The IBM e-Config report can be used to determine the number of CPC drawers, active PUs, memory configuration, and the channel layout.

For more information about Enhanced Drawer Availability refer to:

- ▶ *IBM z15 (8561) Technical Guide*, SG24-8851
- ▶ *IBM z15 (8562) Technical Guide*, SG24-8852
- ▶ *IBM z14 Model ZR1 Technical Guide*, SG24-8651
- ▶ *IBM z14 Technical Guide*, SG24-8451
- ▶ *IBM z13s Technical Guide*, SG24-8294.
- ▶ *IBM z13 Technical Guide*, SG24-8251

### Maximum number of links

The maximum number of attached servers supported by any STP configured server in a CTN is equal to the maximum number of coupling links supported. The number of links that can be installed varies by server type.

**Important:** You must provide at least two coupling links between any two servers that are intended to exchange STP messages for synchronization. This prevents the loss of one link causing the loss of STP communication between the servers.

Table 2-3 lists the maximum number of external coupling links each server type supports in a Parallel Sysplex (z13 and newer servers). Internal coupling channels (ICs) are not used by STP.

Table 2-3 Supported Coupling link options

Type	Description	Feature Code	Link rate <sup>a</sup>	Max unrepeated distance	Maximum number of supported links <sup>b</sup>					
					z15 T01	z15 T02	z14 ZR1	z14	z13s	z13
CE LR	Coupling Express LR	0433	10 Gbps	10 km (6.2 miles)	64	64	32	64	32	64
ICA SR1.1	Integrated Coupling Adapter	0176	8 GBps	150 meters (492 feet)	96*	48	N/A	N/A	N/A	N/A

Type	Description	Feature Code	Link rate <sup>a</sup>	Max unrepeatd distance	Maximum number of supported links <sup>b</sup>					
					z15 T01	z15 T02	z14 ZR1	z14	z13s	z13
ICA SR	Integrated Coupling Adapter	0172	8 GBps	150 meters (492 feet)	96*	48*	16*	80*	16	40*
HCA3-O LR	Infiniband Long Reach (1 x IFB)	0170	5 Gbps	10 km (6.2 miles)	N/A	N/A	N/A	64*	32	64*
HCA3-O	Infiniband (12 x IFB)	0171	6 GBps	150 meters (492 feet)	N/A	N/A	N/A	32*	16	32*
IC	Integrated Coupling Adapter	N/A	Internal speed	N/A	64	64	32	32	32	32

a. The link data rates do not represent the performance of the links. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload.

b. The maximum combined number of supported links depends on the IBM Z model or capacity feature code and the numbers are marked with an asterisk (\*).

## 2.2.2 Planning for timing-only links

Timing-only links are coupling links that allow two servers to be synchronized using STP messages when a CF does not exist at either end of the coupling link.

### Hardware Configuration Dialog: timing-only links

Hardware Configuration Dialog (HCD) and Hardware Configuration Manager (HCM) support the definition of timing-only links. The timing-only links bring STP capability into a basic-mode sysplex, or a non-sysplex set of systems that do not currently have coupling links in use the way that a Parallel Sysplex, by definition, would. Also within a Parallel Sysplex, a direct STP connection between servers that both have no coupling facility LPAR, can be established by a timing-only link.

A timing-only link is established through the CF Channel Path Connectivity List panel in HCD. This is shown in Figure 2-4. Defining timing-only links is very much like defining other coupling links. IFB links are defined as CIB, ICA SR links are defined as CS5, and CE LR links are defined as CL5 using the normal CF connection dialogs. As with coupling links, timing-only links can be defined as dedicated, reconfigurable, shared, or spanned. Multiple Image Facility (MIF) can be used to share the timing-only link between z/OS images on a server, but not between CF images on a server.

```

Connect to CF Channel Path

Specify the following values.

Source processor ID . . . . . : ARIES
Source channel subsystem ID . : 2
Source channel path ID . . . . : 80
Source channel path type . . . : CS5

Destination processor ID . . . . . CETUS +
Destination channel subsystem ID . . 2 +
Destination channel path ID . . . . E4 +

Timing-only link . . . . . YES

```

Figure 2-4 HCD: timing-only link definition

The HCD dialog contains an option where the user can specify whether the CF connection is for a timing-only link. The default is no. If the user selects a timing-only link, HCD generates an STP control unit on each side of the connection, and no devices.

**Note:** The addition of timing-only links is fully supported by the Dynamic I/O Reconfiguration process (timing-only links can be added non-disruptively).

For a timing-only link, HCD performs the following checks:

- ▶ The CHPIDs of a timing-only link must belong to different servers. A timing-only link cannot be looped back into the same server.
- ▶ If a CF connection already exists between two servers, HCD does not allow defining a timing-only connection between these servers. Also, if a timing-only link already exists between two servers, HCD will not allow defining a CF connection between these servers. The existing connection type must be broken before the other connection type can be established.

**Note:** CF messages cannot be transferred over timing-only links.

For more details about how to set up a timing-only link, see *IBM z15 Configuration Setup*, SG24-8860, Chapter 9.

Timing-only links are supported between all IBM Z servers. The status of these links can be displayed in the graphical interface for STP in the HMC workplace. It reflects which links are available and initialized.

### ***Standalone coupling facilities and timing-only links***

Dynamic I/O reconfiguration supports the transition from timing-only to standard coupling links. However, a standalone coupling facility has no z/OS image to facilitate the initiation of a Dynamic I/O reconfiguration. So on z13 and z13s servers, a power-on reset is necessary to make the change at the coupling facility end of the link. For z14, z14 ZR1 and z15 servers with Driver 36 or later, dynamic I/O changes are supported even for standalone coupling facilities by using an activation service image (MCS\_1 LPAR), see *IBM z15 Configuration Setup*, SG24-8860, Chapter 9.

**Important:** We recommend that a single IODF be used for all servers in the installation.

## **2.2.3 Considerations for multi-site sysplex**

In current business and IT environments, two important objectives for survival are systems designed to provide continuous availability and near transparent DR. Systems that are designed to deliver continuous availability combine the characteristics of high availability and near continuous operations to deliver high levels of service—targeted at 24x7. To attain high levels of continuous availability and near-transparent DR, the solution should be based on geographical clusters such as a multi-site sysplex and data mirroring. Geographically Dispersed Parallel Sysplex (GDPS) is an industry-leading solution that provides this capability. For more information about GDPS see:

<http://www.ibm.com/it-infrastructure/z/technologies/gdps>

If a sysplex across two or more sites is configured, it is necessary to synchronize servers in multiple sites. Consequently, you might plan to extend the CE LR or HCA3-O LR links beyond the distance supported without repeaters. IBM supports Wavelength Division Multiplexing (WDM) products qualified by IBM for use in multi-site sysplex solutions such as GDPS.

**Attention:** To transmit timing information, STP can choose any defined coupling link that is online between two IBM Z servers. If coupling links are configured over WDM equipment, *all* coupling links must use specific WDM hardware (optical modules<sup>a</sup>, transponders<sup>b</sup>, TDM modules<sup>c</sup>) with interface cards qualified by IBM for STP.

If both coupling facility data and STP timing information must be transmitted between two servers, you cannot select a subset of coupling links to be used only for STP timing information.

- a. Including, but not limited to optical multiplexers/demultiplexers, optical add-drop modules (OADM), small form-factor pluggable (SFP) transceivers, dispersion compensation modules, optical amplifiers, and so forth
- b. Wavelength-converting transponders: converts client layer signal to WDM wavelength
- c. Time-Division Multiplexing modules: used to transmit multiple data links over the same WDM channel

The latest list of qualified WDM vendor products and links to corresponding Redpapers for each product can be found on the Resource Link website at:

<https://www.ibm.com/servers/resourceLink/>

They are listed under Hardware products for servers on the Library page.



## 2.3 External time source

If there are specific requirements to provide accurate time relative to an external time standard for data processing applications, then consider using the ETS function.

In a CTN, the ETS function is provided by the following options:

- ▶ A PTP grandmaster (server) without or with pulse per second (PPS) output option
- ▶ An NTP server without or with a pulse per second (PPS) output option

The Current Time Server (CTS) is the only server that adjusts the Coordinated Server Time (CST) in a CTN by steering it to the time obtained from an external time source. In addition, if the CTS has lost connectivity to its external time source server(s), the non-CTS server will send its external time source steering information to the CTS which will then steer it the appropriate offset.

One of the important differences between NTP and PTP is that NTP synchronizes the client device to the next server as specified in the NTP configuration and the only information conveyed to the client is the NTP stratum level of the NTP server to which it synchronizes its time, while the PTP ordinary clock (running on the client device) synchronizes to the PTP Grandmaster (PTP Grandmaster is “known” to the client device), and thus providing increased time accuracy w/o requiring special hardware (i.e. PPS signal support).

### 2.3.1 Precision Time Protocol grandmaster

The use of a Precision Time Protocol Grandmaster as an external time source addresses the requirement of customers who need a highly accurate time source for the IBM Z servers, or those who need to use an accurate common time reference across heterogeneous platforms. These customers, in most cases, purchase a PTP server (grandmaster) that will obtain the time using high-precision time technology, such as GPS signals.

PTP requires appropriate (supporting) infrastructure capable of transmitting PTP signaling from the PTP grandmaster to the IBM Z CPC. Refer to your PTP grandmaster and network infrastructure documentation for identifying requirements for PTP connectivity.

Redundant PTP grandmasters are recommended for ETS availability. In current implementation, additional network connectivity for the IBM Z Support Element is required. Check with your IBM Service Representative for configuring Support Element network in support of PTP.

For highest time accuracy, current IBM Z PTP implementation requires Pulse Per Second connectivity between the PTP grandmaster and the IBM Z CPC.

**Important:** PTP (IEEE-1588-2) requires specialized HW infrastructure for transmitting timing signals.

An example of PTP connectivity (as ETS) is shown in Figure 2-5 on page 36.

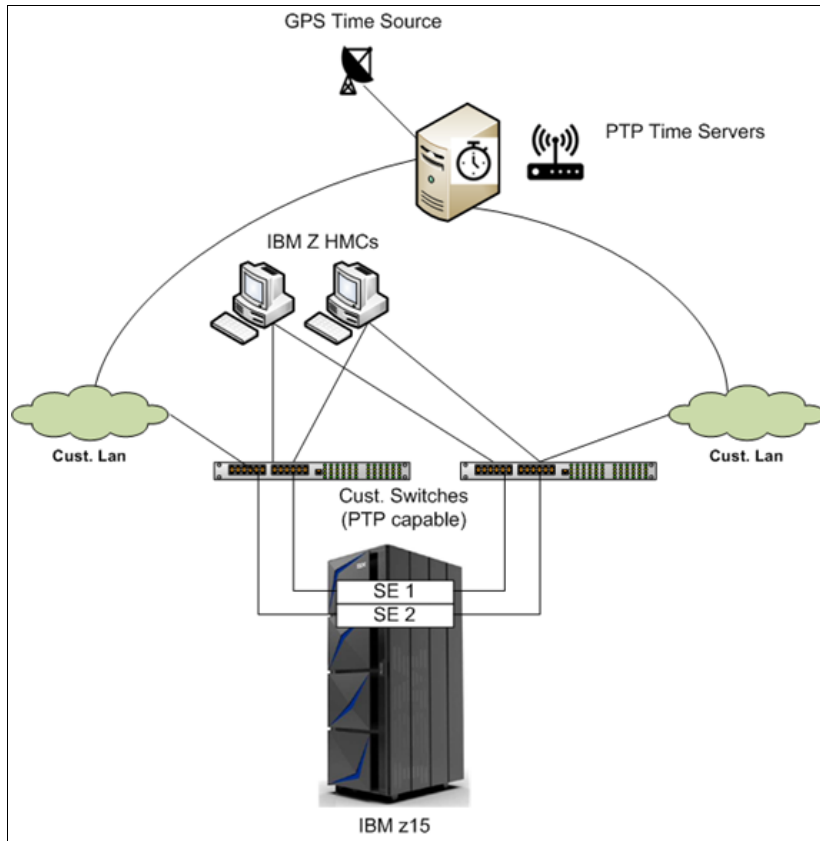


Figure 2-5 PTP connectivity diagram

### 2.3.2 NTP server

The use of a Network Time Protocol server as an external time source addresses the requirement of customers who need an accurate time source for the IBM Z servers, or those who need to use an accurate common time reference across heterogeneous platforms. These customers, in most cases, purchase an NTP server that will obtain the time using high-precision time technology, such as GPS signals.

Actual time accuracy, relative to UTC, depends on how accurate the NTP server time is with respect to UTC. Proper NTP server selection becomes a business decision rather than a technical requirement. STP provides configuration and alerts based on defined NTP stratum information. Customers can specify an NTP stratum level which is unacceptable (the higher the NTP stratum, the lower the time accuracy) for timekeeping information in the respective CTN.

#### Notes:

- ▶ Like STP, NTP is a hierarchical protocol with servers organized in stratum levels. However, there are significant differences between NTP stratum and STP stratum, so these terms cannot be intermixed.
- ▶ Customers can use on-premises NTP servers or can purchase NTP services from a provider over the Internet.

For further information about NTP refer to the RFCs at the following web pages:

- ▶ NTPv3: <http://www.faqs.org/rfcs/rfc1305.html>

- ▶ NTPv4: <http://www.faqs.org/rfcs/rfc5905.html>

For further information about NTP and the NTP Public services project refer to the following websites:

<http://www.ntp.org>  
<http://support.ntp.org>

### **NTP client support on the Support Element**

The NTP client on IBM Z servers can provide accuracy for the Coordinated Server Time by accessing an NTP server. An accuracy of 100 milliseconds to the time provided by the NTP server is possible. Simple Network Time Protocol (SNTP) client support is added to the STP code on the SE to interface with the NTP servers. Access to the NTP server from the CTS is initiated and controlled by the Support Element.

#### ***Planning for NTP client and LAN connectivity***

The implementation of the SNTP client on the Support Element employs:

- ▶ NTP V3 (RFC-1305), NTP V4 (RFC-5905), or SNTP V4 (RFC-4330)
- ▶ IPv4 or IPv6

The NTP server can be either an external time source device available from several timekeeping device manufacturers, a local NTP server, or an NTP server configured on the HMC. The NTP traffic between the SNTP client (running on the Support Element) and the NTP server is not encrypted.

### **NTP server on the HMC**

The NTP server configured as the ETS must be attached directly to the SE LAN. The SE LAN is considered in many configurations to be a private (dedicated) LAN and should be kept as isolated as possible. Allowing the NTP client running on the SE to directly communicate to an NTP server that is on the corporate or the external network (Internet) introduces potential security issues (even if the SE LAN is behind a firewall).

**Note:** For security reasons, we do not recommend configuring the SNTP client on the Support Element to directly access NTP servers outside its local area network.

The ability to define an NTP server on the HMC addresses potential security concerns because the HMC is normally attached directly to the SE LAN. The HMC has two LAN ports that are physically isolated, one port for the connection to the HMC/SE LAN used by NTP client code, and the second port for the LAN used by HMC to access an NTP time server to set its time. So the NTP server on the HMC can access another NTP server via a separate LAN connection to obtain its time reference (Figure 2-6).

**Note:** The NTP server function on the HMC does not provide a pulse per second output.

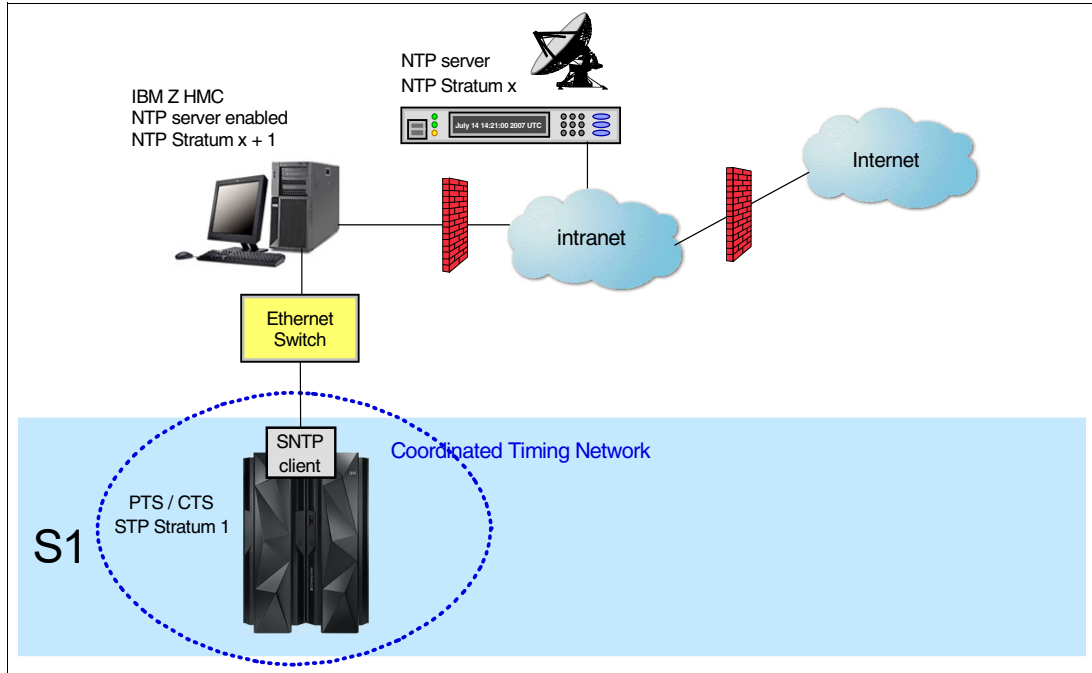


Figure 2-6 HMC acting as NTP server

### HMC NTP broadband authentication support

The possibility to use HMC NTP authentication has been added with HMC level 2.12.0. To exploit this option on the SE, configure the HMC with NTP authentication as the NTP server for the SE. The following scenarios describe the situations in which you could exploit this option.

► **Authentication support with a proxy**

Some configurations use a proxy to access outside the corporate data center. NTP requests are UDP socket packets and cannot pass through the proxy. The proxy must be configured as an NTP server to get to target servers on the web. Authentication can be set up on the proxy to communicate to the target time sources.

► **Authentication support with a firewall**

Some configurations use a firewall. HMC NTP requests can pass through the firewall. When using this configuration, you should use the HMC authentication to ensure un-tampered with time stamps.

► **Symmetric key and autokey authentication**

With the symmetric key and autokey authentication, the highest level of NTP security is available. HMC application level 2.12.0 provides panels that accept and generate key information to be configured into the HMC NTP configuration, and offer the possibility to issue NTP commands.

– Symmetric key (NTPv3-NTPv4) authentication

Symmetric key authentication, as described in RFC-1305 (made available in NTPv3), uses the same key for both encryption and decryption. Users exchanging data keep this key to themselves. Messages encrypted with a secret key can only be decrypted with the same secret key. Symmetric does support network address translation (NAT).

– Autokey (NTPv4) authentication

Autokey authentication, as described in RFC-5906 (made available in NTPv4), uses public key cryptography. The key generation for the HMC NTP is performed by clicking the “Generate Local Host Key” button on the Autokey Configuration panel. Pressing

this button will issue the “ntp-keygen” command to generate the specific key and certificate for this system. Autokey authentication is *not* available with NAT firewall.

- NTP commands  
NTP command support is also added to display the status of remote NTP servers and the current NTP server (HMC).

### Planning for NTP server configuration

There are many possible configuration variations, depending on the type of NTP server equipment used, redundancy and security requirements, and existing LAN topology.

In the examples presented in this document, firewalls are indicated to draw the reader’s attention to the potential network security concerns.

**Note:** Configuration of network components, such as routers or firewall rules, goes beyond the scope of this document. Refer to *Hardware Management Console Operations Guide*, and to the *Installation Manual for Physical Planning* (select the manual specific to your server) for security guidelines and recommendations for the HMC/SE network. For IBM Z hardware documentation, see:

<https://www-01.ibm.com/servers/resourceLink/hom03010.nsf/pages/library?OpenDocument&login>

### Example: Single NTP server (typically from a timekeeping device manufacturer)

Figure 2-7 shows an example that has an NTP server attached directly to the SE LAN. Note that in this case, the NTP server is a single point of failure.

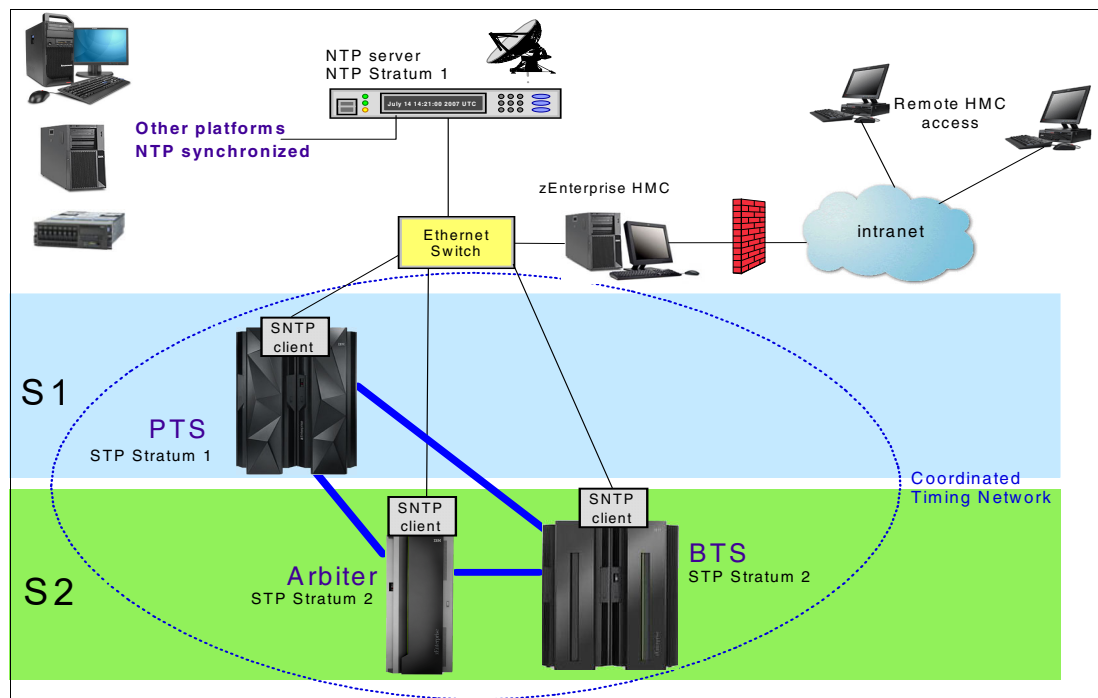


Figure 2-7 IBM Z and other platforms configured with NTP server

One of the NTP output ports of the NTP server is connected directly to the HMC/SE local area network, allowing the SNTP client on the Support Element to obtain the external time from the NTP server. Other NTP output ports of this NTP server can be connected to other platforms that need the same accurate time reference.

In this example, there is no ETS (NTP server, in this case) redundancy. To provide redundancy, additional NTP servers can be installed and configured, as shown in the following section.

### NTP server redundancy

The CTS is the only server that adjusts the CST by steering it to the time obtained from an external time source. To provide ETS redundancy, the user should consider configuring two or more NTP servers.<sup>1</sup>

Up to two NTP servers can be configured on each server in the STP-only CTN. When two NTP servers are configured, the user is responsible for selecting the preferred NTP server. This NTP server is called the *selected* NTP server. The other is called the non-selected NTP server. Normally, the SNTP client uses the time information from the selected NTP server to perform the time adjustment. The SNTP client also compares the quality of both NTP servers and informs the user in case either of the following conditions is detected:

- ▶ The selected NTP server has a stratum level that is lower in the hierarchy than the non-selected NTP server. (NTP stratum 1 server is a better choice than NTP stratum 2.)
- ▶ The time obtained from the selected NTP server has less accuracy than the non-selected NTP server (or if a discrepancy between the time information provided by the two NTP servers exists)

The following planning considerations should be made to provide NTP server redundancy:

- ▶ When two NTP servers are configured on the server that has the PTS/CTS role, STP will automatically access the second NTP server configured on the PTS/CTS if the selected NTP server becomes unavailable.
- ▶ When NTP servers are configured on the server with the BTS role, the following is provided:
  - Access to an NTP server when the BTS becomes the CTS as the result of planned or unplanned recovery
  - Time adjustments to the CTN when the PTS/CTS cannot access any of its NTP servers

**Recommendation:** Configure at least one unique NTP server on both the PTS and the BTS. For redundancy, up to two separate NTP servers can be configured for both the Primary Time Server (PTS) and the Backup Time Server (BTS).

<sup>1</sup> The SE can handle up to two NTP servers. However, the HMC can configure 16 NTP servers. Three or more NTP servers is recommended for ETS redundancy for the CTN.

### Example: Using the HMC to provide NTP server redundancy

A second example is illustrated in Figure 2-8. One of the NTP ports of NTP server 1 is connected directly to the HMC/SE local area network, as in the previous example. However, this example also introduces using the HMC as an NTP server to provide NTP server redundancy.

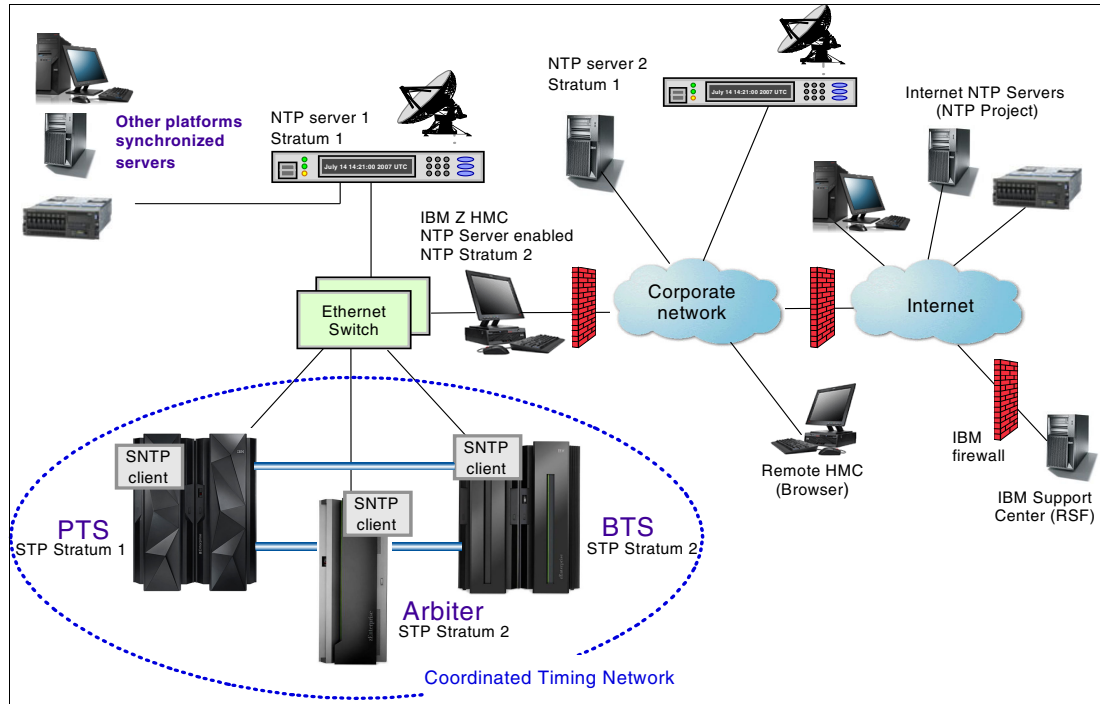


Figure 2-8 IBM Z and other platforms configured with NTP servers: Example 2

Since the HMC is normally connected to the SE LAN, the capability to configure an NTP server on the HMC should be considered. The HMC can connect using a different network adapter to a remote NTP server located on the corporate intranet, or even on the Internet to access its time reference (NTP server 2 in Figure 2-8). Such a configuration requires UDP port 123 being enabled on the customer firewall to allow NTP traffic between the local NTP server (running on HMC) and remote NTP servers.

Alternatively, a separate NTP server (for example, running on UNIX or Linux) can be configured similar to the HMC (with dual Ethernet adapters) and connected directly to the SE LAN. This provides separation between the HMC/SE LAN and the external networks.

This configuration offers redundancy because the SNTP client on the SE allows configuration of two NTP servers. If the preferred NTP server (NTP Server 1 in this example) cannot be accessed, time adjustments can still be made by accessing the NTP server on the HMC.

As in the first example, browser access to the HMC might be available using the corporate network through the customer firewall.

### Example: NTP server redundancy on PTS and BTS

Figure 2-9 on page 42 shows a third example. It involves multiple sites with NTP servers accessing remote NTP servers on the corporate network or the Internet. This example is different from the previous ones because the corporate network is always used as an access path to the time source. In addition, even though the example uses multiple sites, this configuration can be used in a single site as well.

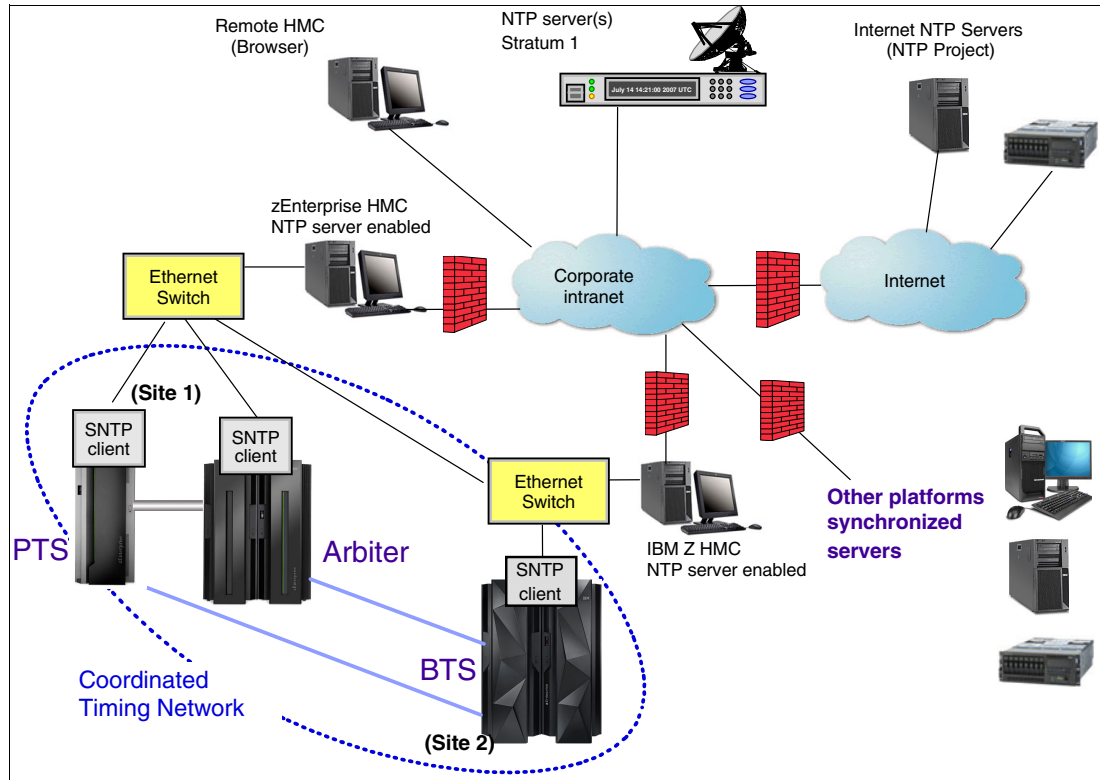


Figure 2-9 IBM Z and other platforms configured with NTP servers: Example 3

This configuration shows that continuous availability of NTP servers can be obtained by configuring a different NTP server in site 1 (on the PTS/CTS) and site 2 (on the BTS). In each site a different HMC is providing the NTP server function. The HMCs in turn can get their time reference from the corporate NTP server 1. The BTS in site 2 accesses the NTP server in site 2 on a periodic basis, calculates time adjustments as required, and propagates them to the PTS in site 1. If the PTS/CTS cannot access the NTP server in site 1, it is able to use the adjustments calculated by the BTS to perform the necessary time adjustment steering for the entire CTN.

This example also requires UDP port 123 to be enabled in the customer firewall for access to NTP servers on the corporate network or on the Internet.

The examples presented illustrate that many variations are possible. The choice of one configuration versus another depends on the security and networking requirements of the enterprise.

### 2.3.3 NTP server with pulse per second

The time accuracy of an STP-only CTN has been improved by adding the capability to configure an NTP server that has a PPS output signal as the ETS device. This type of ETS device is available worldwide from several vendors that provide network timing solutions.

Pulse per second is an electrical signal that very precisely indicates the start of a second. STP has been designed to track to the highly stable, accurate PPS signal from the NTP server and maintain an accuracy of *10 microseconds* as measured at the PPS input of the IBM Z server. A number of variables such as accuracy of the NTP server to its time source (GPS radio signals, for example) and the cable used to connect the PPS signal will determine the ultimate accuracy of STP relative to Coordinated Universal Time (UTC).



In comparison, if STP is configured to use an NTP server without pulse per second, it provides only a time accuracy of *100 milliseconds* to the ETS device.

Regulation changes announced by the Financial Industry Regulatory Authority (FINRA) significantly reduced the allowable dispersion from universal time of systems processing certain financial trades, see <http://www.finra.org/industry/notices/16-23>. PPS is a standard, together with NTP, that provides precise notifications to connected systems so long as the PPS provider is attached to a compliant time source.

**Note:** While NTP together with the PPS standard enables a CTN to operate with a sufficient accuracy compliant to the FINRA regulations, one other aspect has to be considered regarding the FINRA requirements: the **leap seconds**.

Leap seconds are used to synchronize the Coordinated Universal Time (UTC) to the international atomic time (TAI). When a leap second is inserted into UTC (when a leap second is needed, it will generally be added at midnight UTC on either June 30 or December 31), UTC effectively jumps backward by one second. If leap seconds are not specified for the CTN, UTC will be one second ahead of the NTP server that STP is synchronizing to when the leap second event occurs. STP will automatically steer to the correct time, but this steering adjustment is made at the rate of seven hours per second to ensure no duplicate timestamps are seen.

In order to maintain time accuracy when a leap second event occurs, leap seconds must be specified and any leap second adjustment must be scheduled for the appropriate date and time via the STP Adjust Leap Seconds panel, also see 3.2, “Configuring the Coordinated Timing Network” on page 71. STP adds a 61st second to the last minute of the last hour of the day.

Customers who do not specify leap seconds will not meet the FINRA clock synchronization requirement during the seven-hour window required for steering to complete. This will be highlighted by message IEA032E issued by the z/OS accuracy monitor at every 60 minute interval until steering has corrected the time to within the ACCURACY threshold.

Customers that plan to start specifying leap seconds to ensure time accuracy when leap seconds are scheduled have two choices:

- ▶ The leap second value can be updated to the current value (TAI has been introduced in 1961, UTC in 1972, UTC has started 10 seconds apart from TAI, and until 2019, 27 leap seconds have been inserted into UTC) and then kept up to date when new leap seconds are scheduled. As a positive leap second change is a negative UTC change, z/OS spins for the amount of the positive leap second change to avoid duplicate time stamps. A spin of 27 seconds at a time will very likely cause a system outage and so leap seconds need to be added in small increments or while the z/OS systems are down.
- ▶ Alternatively, leap seconds can be added starting from when the next leap second is scheduled. When the next leap second is added, the STP leap second value can be updated from 0 to 1 and then incremented as each subsequent leap second is added. This will ensure time accuracy across future leap seconds without the problem of getting from a current value of 0 to 27.

See also the discussion about the FINRA requirements in 1.1.3, “Industry requirements” on page 4 and about the ACCURACY statement in the CLOCKxx PARMLIB member in “ACCURACY mmmmm” on page 51. See also the White Paper WP102690 „STP recommendations for the FINRA clock synchronization requirements“, <http://www.ibm.com/support/techdocs/atmsastr.nsf/WebIndex/WP102690>.

## Planning for NTP server with PPS

To configure an NTP server with PPS, the NTP output of the NTP server must be connected to the Support Element LAN, and the PPS output of the same NTP server must be connected to the PPS input provided on the External Clock Facility (ECF) card of the IBM Z server.

The NTP information is provided via the HMC/SE local area network. However, the NTP server propagates its pulse per second signal via a coaxial cable. The signals are carried over a copper cable, which is limited in length. The cable length limitation depends on vendor specifications, and certain vendors might offer a fiber optic cable connection as an option. The cable is connected to a coaxial connector, which resides on the ECF card.

As a minimum, one NTP server with a pulse per second output should be configured at the server that has the Current Time Server role.

## NTP server with PPS redundancy

Each IBM Z server has been equipped with two OSC<sup>2</sup> cards, providing the capability of attaching one or two NTP servers with pulse per second. Attaching a second NTP server with pulse per second gives the redundancy that might be required in certain configurations.

When two NTP servers with PPS are configured, the user is responsible for selecting the preferred NTP server. This NTP server is called the selected NTP server with PPS. The other is called the non-selected NTP server with PPS. Normally, STP uses the time information from the selected NTP server to perform the time adjustment. The SNTP client also compares the quality of the NTP data received from both NTP servers and informs the user if one of the following conditions is detected:

- ▶ The selected NTP server with PPS has a stratum level that is lower in the hierarchy than the non-selected NTP server. (NTP stratum 1 server is a better choice than NTP stratum 2.)
- ▶ The time obtained from the selected NTP server with PPS has less accuracy than the non-selected NTP server.
- ▶ The two NTP servers are providing different times.

The following planning considerations should be made to provide NTP servers with PPS redundancy:

- ▶ When two NTP servers with PPS are configured on the server that has the PTS/CTS role, STP will automatically access the PPS signal from the non-selected NTP server configured on the PTS/CTS if the PPS signal from the selected NTP server fails. Figure 2-10 on page 45 is an example of this redundancy.
- ▶ When NTP servers with PPS are configured on the server with the BTS role, the following is provided:
  - Access to an NTP server with PPS when the BTS becomes the CTS as the result of planned or unplanned recovery
  - Time adjustments to the CTN when the PTS/CTS does not receive PPS signals from all of its configured NTP servers with PPS

Figure 2-11 on page 46 illustrates this type of redundancy.

<sup>2</sup> Oscillator (clock) cards. Depending on the server hardware, each sever (z13 and later) have redundant OSC cards with PPS input.

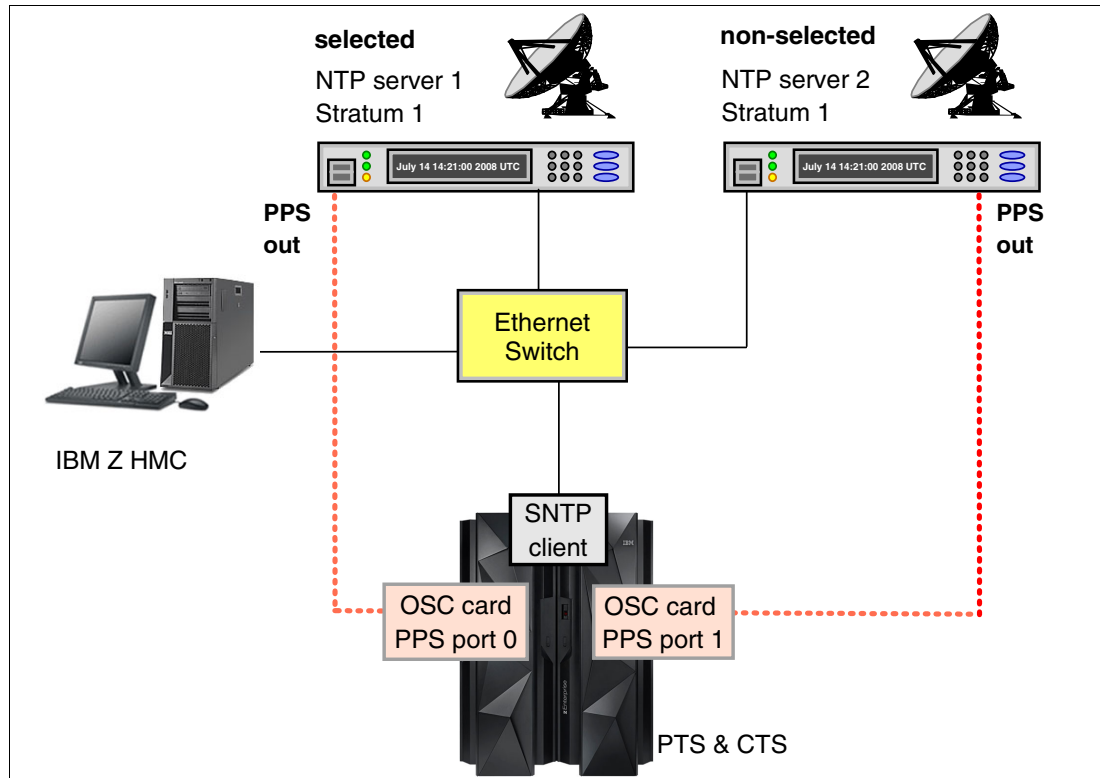


Figure 2-10 IBM Z with NTP pulse per second

**Recommendation:** Configure at least one unique NTP server on both the PTS and the BTS. For redundancy, up to two separate NTP servers can be configured for both the PTS and the BTS.

**Example: Two NTP servers with PPS configured on PTS/CTS**

In this example, two NTP servers with PPS are configured on the PTS/CTS (Figure 2-10). The NTP outputs of both NTP servers are connected to the SE LAN via the Ethernet switch. The PPS output of NTP server 1 is connected to PPS port 0, and the PPS output of NTP server 2 is connected to PPS port 1.

Assume that the PTS/CTS is tracking to the PPS signal received on PPS port 0. If there is a failure associated with this PPS signal, STP will switch to the PPS signal received on PPS port 1 and can continue tracking to an ETS, thus maintaining time accuracy.

**Example: Different NTP servers with PPS on PTS and BTS**

The configuration shown in Figure 2-11 allows continuous availability of ETS by configuring a different NTP server in site 1 (NTP server A on the PTS/CTS) and site 2 (NTP server B on the BTS). The PPS output of each NTP server is connected to PPS port 0 on the respective OSC card.

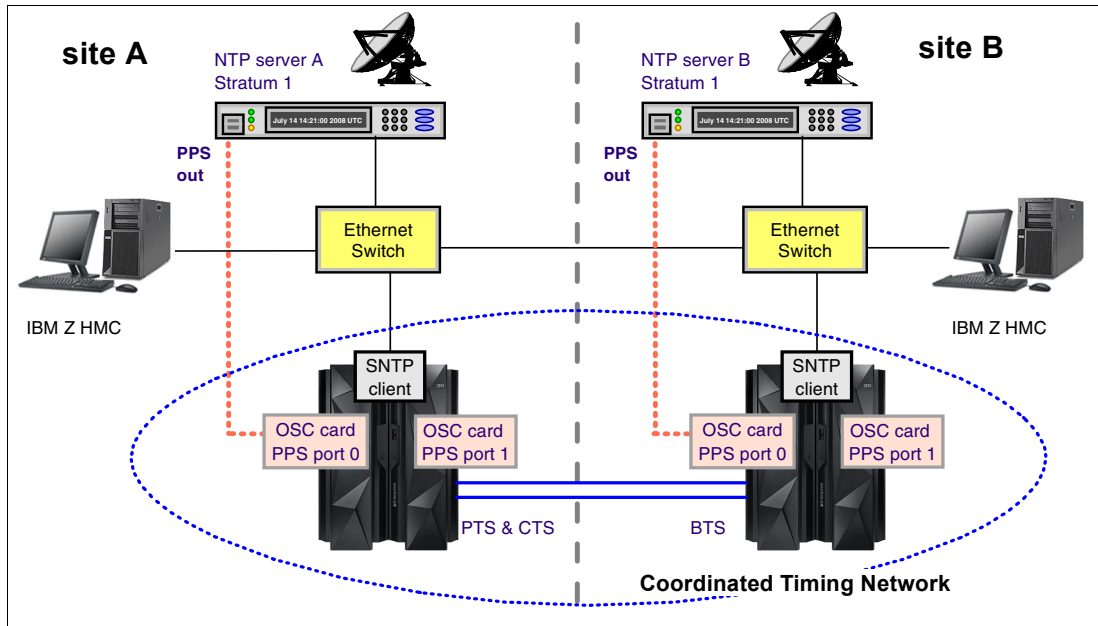


Figure 2-11 NTP servers with pulse per second on PTS and BTS

During normal operation, the CTN will track to the PPS signals received on PPS port 0 of the PTS/CTS. Also, the BTS in site B calculates time adjustments based on NTP data and PPS signals that it receives from NTP server B, and propagates them to the PTS/CTS in site A. The data from the BTS is not used for steering until there is a planned or unplanned recovery action as follows:

- ▶ If the BTS takes over the role of the CTS due to a recovery action, then the STP-only CTN can still continue to track to PPS signals and NTP data received from NTP server B.
- ▶ If there is a failure that results in PPS signals from NTP server A not being received by the PTS/CTS, it is able to use the adjustments calculated by the BTS to perform the necessary time adjustment steering for the entire CTN.

## 2.4 Internal Battery Feature (IBF)

In IBM Z servers that are equipped with Bulk Power Assemblies (BPA), which are z13, z13s, z14 and some models of z15, an Internal Battery Feature (IBF) can be installed. The IBM Z servers equipped with a Power Distribution Units (PDU) (z14 ZR1 and some models of z15) do not support the IBF.

**Note:** With z15 servers two power options, BPA and PDU, are available. BPA models support the IBF, while the PDU models do not support the IBF.

The main purpose of an IBF is to keep an IBM Z server running for several minutes, when a power outage to the data center occurs *and there is no Uninterruptible Power Supply (UPS) available*. With proper procedures, in this time the workload of the server running on IBF can be moved or shut down in a controlled manner.

**Removal of IBF support:** IBM z15 T01 is planned to be the last IBM Z server to offer an Internal Battery Feature (IBF). As client data centers continue to improve power stability and uninterruptible power supply (UPS) coordination, IBM Z continues to innovate to help clients take advantage of common power efficiency and monitoring across their ecosystems. Additional support for data center power planning can be requested through your IBM Sales contact.

If an IBF is installed on an IBM Z server, STP has the capability of receiving notification that external power has failed and the IBF is engaged (CPC running on IBF power). When STP receives this notification from a server that has the role of the PTS/CTS, STP can automatically reassign the CTS role to the BTS, thus automating the recovery action and improving availability.

If the entire CTN is located in a single data center and only has two servers (PTS and BTS), and the IBF is installed on both the PTS and the BTS. This should provide recovery capabilities when the server that is the CTS experiences a power failure.

**Important:** If the CTN in a single data center has three or more servers, we recommend assigning 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.

## 2.5 Planning z/OS software

When a new CTN is configured, there are z/OS considerations that must be planned for. Some key planning considerations are:

- ▶ Apply STP-related maintenance to all z/OS systems that need multisystem time synchronization.
- ▶ Update SYS1.PARMLIB(CLOCKxx) (optional).

### 2.5.1 z/OS requirements for STP

STP is supported by z/OS since z/OS V1R7. So today all z/OS images, either running in a CTN, or unsynchronized, include the STP support.

Please note that a re-IPL of the z/OS image is required:

- ▶ after applying z/OS-related maintenance (as required),
- ▶ after having changed CLOCKxx parameters.

In a Parallel Sysplex this can be done using the *rolling IPL* process, which avoids a sysplex-wide outage.

#### **z/OS maintenance**

Even though today all supported releases of z/OS support STP, additional maintenance might be required. In a sysplex configuration, STP maintenance is required on all z/OS images in the sysplex before the server is STP-configured. For receiving and implementing STP

maintenance in your z/OS, check the Preventive Service Planning (PSP) buckets, STP related PTFs are listed in the PSP buckets for the servers and coupling facilities. To simplify the identification of PTFs for STP, a functional PSP bucket has been created. Please use the Enhanced Preventive Service Planning Tool (EPSPT):

<http://www.software.ibm.com/webapp/set2/psp/srchBroker>

as well as your usual maintenance procedures. For information about SMP/E commands syntax or on receiving and applying SYSMODs, refer to *SMP/E for z/OS User's Guide*, SA23-2270 or *SMP/E for z/OS Reference*, SA23-2276.

## 2.5.2 CLOCKxx

The z/OS system data set SYS1.PARMLIB contains many members that are used during IPL to determine how the z/OS system should be configured. One member is CLOCKxx, where xx is a 2-character suffix. The CLOCKxx member performs the following functions:

- ▶ Provides the means to specify that z/OS should make use of STP and its time zone parameters. For this the statements STPMODE, STPZONE and TIMEZONE are used.
- ▶ Provides the means of specifying that z/OS should make use of the Sysplex Timer<sup>3</sup> and its time zone parameters. For this the statements ETRMODE, ETRZONE and TIMEZONE are used.
- ▶ Allows you to specify the difference between the local time and UTC if you do not want to use the time zone used by ETR or STP. For this the statements ETRDELTA and TIMEDELTA are used.
- ▶ Specifies how much time deviation for the TOD clock from the external time source is accepted. For this the statements ACCURACY and ACCMONINTV are used.

For STP support the following CLOCKxx statements have been added in z/OS in z/OS 1.7:

- ▶ STPMODE
- ▶ STPZONE
- ▶ TIMEDELTA

An External Time Reference (ETR) is a 9037 Sysplex Timer. Although today the IBM Z servers do not support any connectivity to a 9037 Sysplex Timer, the statements regarding the connection to an active ETR are still supported in the CLOCKxx member.

---

<sup>3</sup> Sysplex timer is not available for z196 and later generations, also, zEC12/zBC12 is the last server generation to support Mixed CTN. z13 and later support STP-only CTN.

**Notes:** If the z/OS system image is running on a server that is in ETR timing mode, and both STPMODE and ETRMODE have been specified as YES, z/OS uses ETRMODE YES.

If the z/OS system image is running on a server that is in STP timing mode, and both STPMODE and ETRMODE have been specified as YES, z/OS uses STPMODE YES.

For migration planning from an ETR network to a Mixed CTN, or from a Mixed CTN to an STP-only CTN, it has been important that both ETRMODE YES and STPMODE YES were specified prior to the migration for systems that are running on servers that are in ETR timing mode. The ETR parameters could be removed after the migration to STP has been completed and the Sysplex Timers have been removed.

However, on IBM Z servers used today, which do not support any ETR connectivity and therefor cannot run in ETR mode, it does not matter if ETRMODE statement is specified, as long as STPMODE YES is specified.

For this reason the statements regarding ETR will not be discussed in detail.

In a multi-system environment, the objective is to share one SYS1.PARMLIB concatenation with as many z/OS images as possible. This simplifies system management by helping to present a single system image to system programmers so that changes must be made in one place rather than in every system image in the sysplex.

For more information about the CLOCKxx member see *z/OS MVS Setting Up a Sysplex*, SA23-1399.

## Syntax format of CLOCKxx

Example 2-1 shows the CLOCKxx parameters format.

*Example 2-1 Syntax format of CLOCKxx parameters and values*

---

```

OPERATOR {PROMPT } | {NOPROMPT}
TIMEZONE d.hh.mm.ss
ETRMODE {YES} | {NO }
ETRDELTA nn
ETRZONE {YES} | {NO }
SIMETRID nn
STPMODE {YES} | {NO }
TIMEDELTA nn
STPZONE {YES} | {NO }
ACCURACY mmmm
ACCMONINTV {mmss | mss}

```

---

## IBM supplied default for CLOCKxx

The IBM supplied default PARMLIB member of SYS1.PARMLIB is CLOCK00. If the STP-related maintenance is installed, CLOCK00 contains the parameters shown in Example 2-2.

*Example 2-2 Default parameters and values for CLOCKxx*

---

```

OPERATOR NOPROMPT
TIMEZONE W.00.00.00
STPMODE YES
STPZONE YES
TIMEDELTA 10

```

ACCURACY 0

---

**Statements and parameters for CLOCKxx**

This section identifies the CLOCKxx statements and provides a short description of each statement. For a detailed description of the statements in the CLOCKxx member see the *MVS Initialization and Tuning Reference*, SA23-1380.

**OPERATOR {PROMPT | NOPROMPT}**

This specifies whether the operator is to be prompted to set the TOD clock during system image initialization. The default NOPROMPT specifies that the system image is not to prompt the operator during TOD initialization unless the TOD clock is not set.

OPERATOR PROMPT and SIMETRID are mutually exclusive keywords. If both are specified, the system image rejects the CLOCKxx member during system image initialization and issues a message to prompt the operator for specifying a valid CLOCKxx member, or for using the default values (by pressing the Enter key). Otherwise, the operator must re-IPL the system image.

**TIMEZONE d.hh.mm.ss**

This specifies the difference between the local time and the UTC. If STPMODE YES and STPZONE YES are specified and the server is being synchronized using STP messages, the system image ignores the TIMEZONE parameter.

The format is *d.hh.mm.ss*, where *d* specifies the direction from UTC (E for East and W for West), and *hh.mm.ss* specifies the number of hours (*hh*) minutes (*mm*) and seconds (*ss*) that the local time differs from UTC. The default value is 00.00.00.

**ETRMODE {YES | NO}**

This specifies whether z/OS should use ETR timing mode. As IBM Z servers do not support ETR mode any more, this statement should be removed. However, if the z/OS system image is running on a server that is in STP timing mode, and both STPMODE and ETRMODE have been specified as YES, z/OS uses STPMODE YES.

**ETRDELTA nn**

This is the ETR version of the TIMEDELTA statement. ETRDELTA should be replaced by TIMEDELTA.

**ETRZONE {YES | NO}**

This is the ETR version of the STPZone statement. ETRZOE should be replaced by STPZONE.

**SIMETRID nn**

This specifies the simulated Sysplex Timer identifier. SIMETRID allows z/OS images running on the same server, in native mode in logical partitions, or as VM guests, to participate in a multisystem sysplex when no real STP is available. In these environments, the z/OS TOD clocks are synchronized by PR/SM or the z/VM® host.

Do not use SIMETRID to coordinate z/OS images running on different servers. Instead, run in STP timing mode with STPMODE YES.

**STPMODE {YES | NO}**

This specifies whether z/OS should use STP timing mode. The default value is YES, which specifies that z/OS is to use STP timing mode if STP is configured. When NO is specified, local z/OS timing mode will be used.



**TIMEDELTA nn**

This indicates the greatest difference, after IPL, between the system's TOD and the Coordinated Server Time (the STP stratum 1 server's time value), by which the system will adjust its TOD, when necessary, to match the stratum 1 server's TOD. The default value is 10 (seconds).

**Note:** Selecting a TIMEDELTA value of 10 seconds does not mean that the processor TOD and STP CST can be out of synchronization by as much as 10 seconds. The two are synchronized to a tolerance of a few microseconds and the value of TIMEDELTA has no effect on the synchronization tolerance between the processor TOD and CST.

If the difference between the system's TOD and the Coordinated Server Time (CST) exceeds the TIMEDELTA, the result is:

- ▶ If the system is in STP synchronization mode and is part of a multisystem sysplex, the WTOR message IEA394A is issued.
- ▶ If the system is not part of a multisystem sysplex, processing continues, but STP is not used for the remainder of that IPL.

If a value of 0 seconds is selected for TIMEDELTA, no deviation between the processor TOD clock and the Stratum 1 Server's TOD can be corrected by z/OS, so no TOD adjustment is possible. When a synch check is recognized in the TIMEDELTA 0 case, one of the two previously described actions will result.

**STPZONE {YES | NO}**

This specifies whether the system image is to get the time zone constant from the STP. The time zone constant specifies the difference between the local time and the UTC. The default is YES, when also STPMODE is set to YES.

**ACCURACY mmmmm**

The z/OS TOD clock accuracy monitor was introduced with z/OS V2.1 to monitor the accuracy to which a CTN is synchronized with an external time source (ETS). The ACCURACY statement in the CLOCKxx PARMLIB member specifies a threshold for the time deviation between the UTC time of the CTN and an external time source. The accuracy monitor checks the deviation and issues a message if found to exceed the specified ACCURACY threshold.

The z/OS TOD clock accuracy monitor is enabled by specifying a non-zero value between 0 and 60000 milliseconds (60 seconds) for the ACCURACY parameter in the CLOCKxx PARMLIB member. The default value for ACCURACY is zero, meaning that the accuracy monitor is disabled. When CLOCKxx is processed at IPL time, and a non-zero value is specified for the ACCURACY parameter for a system running on an STP network that is synchronizing to an external time source, the ACCURACY function is activated. This results in the following IPL-time message:

```
IEA034I THE TOD CLOCK ACCURACY MONITOR IS ACTIVE
```

The monitor will then check the time deviation every 60 minutes unless the ACCMONINTV keyword is specified with a user-defined timing interval. If the TOD clock exceeds +/- the ACCURACY value, the following message is issued and then reissued every 60 minutes until the condition is corrected:

```
IEA032E TOD CLOCK ACCURACY LIMITS MAY HAVE BEEN EXCEEDED
```

When the time difference is corrected the following message is issued:

## IEA033I THE TOD CLOCK IS NOW WITHIN SPECIFIED ACCURACY BOUNDS

**Note:** IBM recommends that ACCURACY be set to 20 for customers that need to meet the FINRA clock synchronization requirement. This will cause IEA032E to be issued when the steering adjustment exceeds 20 milliseconds.

To help customers meet the FINRA clock synchronization audit requirement, new function have been added to the z/OS accuracy monitor via APAR OA51786. A message is issued at a user specified interval, with a default of every 60 minutes, giving details of the current status from the PRT correction steering information block.

See also the discussion about the FINRA requirements in 1.1.3, “Industry requirements” on page 4 and in 2.3.3, “NTP server with pulse per second” on page 42.

**ACCMONINTV mmss | mss**

Sets a timing interval when the ACCURACY function is to be processed The default value is 6000 (60 minutes).

**General rules and recommendations**

**Note:** The synchronization mode for the server and the synchronization mode for a z/OS image running on the server are usually the same but can be different. If the server is in ETR timing mode, a z/OS image could be running in local, ETR, or SIMETR mode. If the server is in STP timing mode, a z/OS image could be running in local, STP, or SIMETR mode.

The following general rules apply to the creation of CLOCKxx:

- ▶ The following combinations of parameters are not valid and will be rejected:
  - ETRMODE=NO and ETRZONE=YES
  - STPMODE=NO and STPZONE=YES

**Recommendation:** STPMODE YES and STPZONE YES are the defaults with z/OS. Letting the STP settings default gives the most configuration flexibility.

- ▶ If both ETRDELTA and TIMEDELTA are specified in the same CLOCKxx member, z/OS will use the second entry, whichever one that is, and flag the first entry as an error. Thus, ETRDELTA *should be removed and replaced with* TIMEDELTA.
- ▶ The CLOCKxx member for a system image that is a member of a multisystem sysplex must contain the specification STPMODE YES. The system image then uses the Sysplex Timer or STP timing signals to synchronize itself with the other members of the sysplex. The system image uses a synchronized time stamp to provide appropriate sequencing and serialization of events within the sysplex.

A usual and reasonable CLOCKxx member for a z/OS system running in a Parallel Sysplex is shown in Example 2-3.

*Example 2-3 Usual example values for CLOCKxx*

---

```
OPERATOR NOPROMPT
STPMODE YES
STPZONE YES
TIMEDELTA 10
```

---

## 2.6 Planning for z/VM, z/TPF and Linux on IBM Z

On IBM Z servers, also other operating systems than z/OS are supported. This section describes the STP support of z/VM, z/TPF and Linux on IBM Z.

### 2.6.1 z/VM

z/VM exploits the Server Time Protocol (STP) facility to generate time stamps for guest and system DASD write I/O operations, allowing these I/O operations to be synchronized with those of other systems. This support allows data used by z/VM and its guests to be replicated asynchronously over long distances by IBM System Storage™ z/OS Global Mirror (formerly known as Extended Remote Copy, or XRC). For example, this allows z/VM to participate in a Geographically Dispersed Parallel Sysplex Metro/z/OS Global Mirror (GDPS/MzGM) environment.

This baseline STP support is available in z/VM since version 6.2, while versions 6.1 and 5.4, required the PTFs for APARs VM64814 and VM64816, and provides the following functions:

- ▶ Synchronizes z/VM TOD clock with STP server at IPL. Before the introduction of baseline STP support it was up to z/VM, under operator control, to change the TOD clock during IPL by the operator prompt for date and time, entering YES to the Change TOD clock {YES|NO} prompt during IPL. With STP support enabled in z/VM, no prompts are received any more to change the TOD clock; instead, STP is automatically used to initialize the TOD clock with the CST and set the same time zone used by the CTN, displaying this message:

```
HCP986I "TOD Clock Synchronized via STP."
```

- ▶ Maintains a delta value of TOD changes over the lifetime of the z/VM IPL.
- ▶ Supports STP time zone management.
- ▶ CPC must be either a member of an STP-only CTN, or a stratum 2 or higher member of a mixed-CTN.

To enable this support, the following FEATURES statements for SYSTEM CONFIG have been added:

- ▶ **STP\_Timestamping:** Timestamps will be added to write channel programs issued to all DASD devices that have the XRC LIC installed.
- ▶ **STP\_TIMEZone/STP\_TZ:** System time zone will be derived from the STP server.
- ▶ **XRC\_OPTIONal:** System behaves differently when STP is suspended. Instead of deferring all I/O that is to be timestamped until STP sync is restored, it stops timestamping but continues issuing I/O.
- ▶ **XRC\_TEST:** Only allowed on 2nd-level z/VM. This will enable STP\_Timestamping without STP availability. Manually-specified TOD value is used for timestamping, intended for vendor test support.

For instance, Example 2-4 show the FEATURES section of a SYSTEM CONFIG that has STP\_TZ enabled only.

*Example 2-4 z/VM SYSTEM CONFIG with STP enabled*

```

/*****
/*
/*           Features Statement           */
/*****
Features ,

```

```

Disable ,                /* Disable the following features */
  Set_Privclass ,       /* Disallow SET PRIVCLASS command */
  Auto_Warm_IPL ,      /* Prompt at IPL always */
  Clear_TDisk ,        /* Don't clear TDisks at IPL time */
Enable,
  STP_TZ ,             /* timezone from STP on CPC */
Retrieve ,             /* Retrieve options */
  Default 20 ,         /* Default.... default is 20 */
  Maximum 255 ,       /* Maximum.... default is 255 */
MaxUsers noLimit ,    /* No limit on number of users */
Passwords_on_Cmnds ,  /* What commands allow passwords? */
  Autolog yes ,       /* ... AUTOLOG does */
  Link yes ,          /* ... LINK does */
  Logon yes ,         /* ... and LOGON does, too */
Vdisk Userlim 144000 blocks /* Maximum vdisk allowed per user */

```

---

For more details about the FEATURES statements, refer to *z/VM: CP Planning and Administration*, SC24-6271.

### STP state flow in z/VM

Depending on what FEATURES statements have been used and on the STP status, you might have different results:

- ▶ During IPL:
  - If STP\_Timestamping, STP\_TZ, or both are specified: Perform activation by querying STP and setting the TOD clock to match the STP TOD value.
  - If STP\_TZ is specified: Set the system time zone to match the STP time zone.
  - If either of these fails, STP will enter SUSPENDED state. Otherwise, STP activation completes successfully and STP is considered ACTIVE.
- ▶ When STP is ACTIVE:
  - If STP\_Timestamping is specified: I/O to XRC-capable DASD will be timestamped (when required) and I/O to non-XRC-capable DASD will be unchanged.
- ▶ When STP is SUSPENDED:
  - If STP\_Timestamping is specified and XRC\_OPT is not specified: I/O to XRC-capable DASD that must be timestamped will be deferred until STP becomes ACTIVE again, while I/O that does not need to be timestamped and I/O to non-XRC-capable will continue to be issued.
  - If STP\_Timestamping is specified and XRC\_OPT is also specified: I/O to all DASD will be issued without a timestamp until STP becomes ACTIVE again.

Some events in the system might cause STP state changes:

- ▶ From ACTIVE to SUSPENDED: Occurs when an STP machine check is received that informs CP that the TOD value must be synchronized.
- ▶ From SUSPENDED to ACTIVE: In response to machine check/external interrupts received, CP will attempt to re-synchronize with the STP server. A successful resync will *not* change the system TOD value, but will update the delta between the system time and the STP TOD value and will update the system time zone (if STP\_TZ is enabled).
- ▶ From ACTIVE to ACTIVE: External interrupts might be received that require CP to query STP time zone information (for example, the time zone was changed via the HMC). This does not cause STP to go SUSPENDED, but will cause the system time zone to change.

## CP commands

With the baseline STP support to provide accurate XRC timestamping associated with VM guests and host I/O operations, some CP commands were added or updated:

- ▶ **QUERY STP:** This command has been added, and is used to display the Server Time Protocol status. Example 2-5 shows a possible response.

*Example 2-5 Query STP information in z/VM*

---

```
Q STP
Server Time Protocol synchronization activated.
Ready; T=0.01/0.01 17:30:45
```

---

- ▶ **SET TIMEZONE:** This command is used to change the system's time zone ID and time zone offset. It has been updated to display the response shown in Example 2-6 if z/VM is configured to use STP time zone.

*Example 2-6 Setting the time zone*

---

```
SET TIMEZONE EST
HCPTZN987E SET TIMEZONE not valid - STP timezone in use
Ready(00987); T=0.01/0.01 17:28:55
```

---

- ▶ **QUERY TIMEZONES:** This command is used to display the list of active and inactive time zone definitions on your system. It has been updated to show the STP time zones when z/VM is configured to use it.

For a full description of possible responses, refer to *z/VM: CP Commands and Utilities Reference*, SC24-6268.

## z/VM IOCP

If z/VM is used to create IOCP decks for the z/OS logical partitions (not z/OS Guest machines), z/VM can be used to create timing-only links. To define timing-only links, simply define the peer channel (CFP or CBP) with a control unit of UNIT=STP, and no devices. See 2.2.2, "Planning for timing-only links" on page 32.

## VM guests

A z/OS sysplex can be established among z/OS images running as guests under z/VM, but the guests still cannot take advantage of STP or the Sysplex Timer. Only their write I/O to an XRC-enabled storage will have the STP timestamps if STP\_Timestamping is specified. z/OS sysplex guests use z/VM timer services and must specify SIMETRID in the parmlib CLOCKxx member to synchronize their virtual TOD clocks.

To be able to change the date and time of a guest system running under z/VM, an OPTION TODENABLE statement must be added in the directory entry for the guest virtual machine. Then change the system date and time for the guest system using the procedures of the guest system. TODENABLE must not be specified for guest machines that will be specifying SIMETRID. This support requires that all of the TODs for the system be synchronized, and if TODENABLE is specified, the operator can subsequently change the guest TOD clock.

The three z/OS guest images in Figure 2-12 belong to the same sysplex. The guest images must use SIMETRID in their CLOCKxx parmlib members. The TOD clock for a VM guest must not be altered from the value set by LPAR during logical partition activation, either by prior systems IPLed in the partition or z/VM during its initialization.

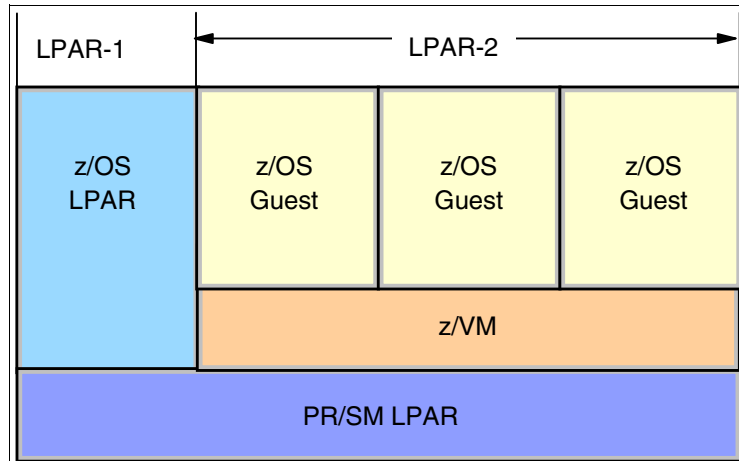


Figure 2-12 A sysplex within a logical partition and VM guests

## 2.6.2 z/TPF

Server Time Protocol (STP) is supported in z/TPF with APAR PJ36831. PJ36831 was released on PUT 07. Information about STP support in z/TPF can be found in the TPF Information Center:

<http://publib.boulder.ibm.com/infocenter/tpfhelp/current/index.jsp>

For more information about z/TPF, refer to the TPF web site at:

<http://www.ibm.com/tpf>

Or contact [tpfqa@us.ibm.com](mailto:tpfqa@us.ibm.com).

## 2.6.3 Linux on IBM Z

Your Linux instance might be part of an extended remote copy (XRC) setup that requires synchronization of the Linux time-of-day (TOD) clock with a timing network.

Linux on Z supports STP based TOD synchronization. This support can be enabled:

- ▶ at boot time by passing the **stp=on** parameter to the Linux kernel, or
- ▶ at runtime by modifying **sysfs**:

At runtime, the TOD synchronization using STP can be turned on:

```
# echo 1 > /sys/devices/system/stp/online
```

Or can be turned off:

```
# echo 0 > /sys/devices/system/stp/online
```

For more details, consult your Linux on Z distribution documentation.

## 2.7 Special considerations for a single server CTN

You might plan a single-server CTN for one of the following reasons:

- ▶ You have a requirement to coordinate time accuracy across IBM Z and other platforms of your enterprise. As discussed in “NTP server” on page 36, you can use the ETS option of using an NTP server to accomplish this.
- ▶ You have a single server and want to implement a sysplex made up of multiple members. Even though this can be accomplished without using STP and coding SIMETRID nn for the CLOCKxx member of each z/OS image, there are reasons, mentioned in the following paragraphs, that you might want to implement this using STP and coding STPMODE YES in the CLOCKxx member.

You might have both z/OS systems as well as other operating systems that do not support STP running on the single server. In the following text, OS-xyz applies to any operating system that does not support STP, for example z/VSE®, as well as z/OS systems that do not have STPMODE YES coded in the CLOCKxx member. All active LPARs in the single server CTN can benefit from the time accuracy provided by using the ETS function. All LPARs (OS-xyz included) can still maintain time within 100 ms of the External Time Source.

There are a few instances that will cause a larger difference between the CEC time and the ETS time that OS-xyz running on a STP configured server will not be able to handle. These limitations do not apply to z/OS images that have STPMODE YES coded in the CLOCKxx member. For details, refer to:

<http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/FLASH10631>

**Note:** The sync check limitation discussed in the referenced document does not apply to a single server CTN, since a stratum 1 server (PTS/CTS) is the source of time and cannot get sync checks.

Regardless of the reason you are planning a single server CTN, the hardware and ETS planning will be the same. The software planning in each case might be unique and will determine whether you can do a concurrent implementation or not.

### 2.7.1 Hardware and ETS planning for single server configuration

Hardware planning includes installing an IBM Z server. There are no considerations for coupling links as this will be a single-server CTN (no external timing links required).

For planning your ETS configuration refer to 2.3, “External time source” on page 35.

**Note:** Since you are implementing a single server CTN, your server will be the Primary Time Server and Current Time Server (PTS/CTS). There is no Backup Time Server (BTS) or Arbiter in this configuration. However, all rules for providing ETS redundancy to the PTS/CTS apply here.

It is highly recommended that you save the STP configuration by restricting CTN membership as shown in Figure 2-13 on page 58.

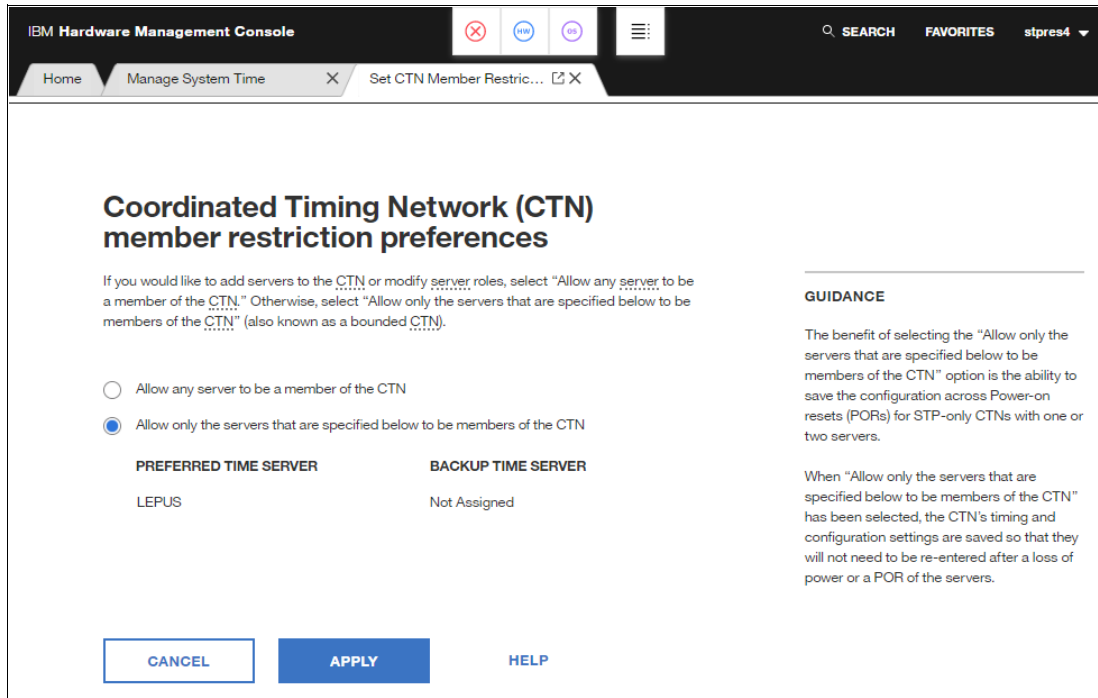


Figure 2-13 Enabling configuration save for single CPC CTN

**Note:** After a power off/on sequence, the TOD value in the Current Time Server is initialized from the TOD values stored in its Support Element. If you have previously configured an external time source, STP code does not perform an *initialize time* using the ETS, but automatically accesses the configured ETS and starts adjusting the time so as to maintain time accuracy. STP allows time adjustments of up to +/- 60 seconds.

## 2.7.2 Software planning: z/OS currently running in a sysplex

In this section, the assumption is that you have running z/OS systems in a sysplex with SIMETRID coded in the CLOCKxx member prior to configuring a CTN.

The main software planning consideration is whether you should continue to run with SIMETRID or whether you should update the CLOCKxx member to use STPMODE YES and STPZONE YES or NO.

Continuing to run with SIMETRID will allow you to configure the CTN concurrently without stopping production. This is because in SIMETRID simulation mode, STP interrupts and machine checks are not applicable.

For configuring the CTN, refer to 3.2.1, "Single Server CTN" on page 71. When the procedure gets you to the initialize time step for a new CTN, the TOD value can be changed since the z/OS systems are not running in STP synchronization mode. Appropriate adjustments to the LPAR offset of each z/OS image running with SIMETRID will be made by PR/SM code. There are no WTOR messages posted to the z/OS Console.



As mentioned in 2.5, “Planning z/OS software” on page 47, if STP is available when z/OS is not being run as a VM guest, IBM recommends using appropriate CLOCKxx statements instead of SIMETRID<sup>4</sup> because of the following advantages:

- ▶ If you update the CLOCKxx member to use STPMODE YES and STPZONE YES, then Daylight Savings Time (DST) changes can be handled without using manual procedures you might be using today to change the time zone, such as using the SET TIMEZONE command.
- ▶ z/OS messages will be posted to the z/OS console when z/OS detects certain conditions with STP such as messages related to ETS.
- ▶ If Leap Seconds are important to you, z/OS will recognize them when you specify Leap seconds during the Initialize Time procedure as well as later when Leap seconds are adjusted. See “Leap second offset adjustment” on page 92.

**Attention:** In order to change from z/OS systems running with SIMETRID to STPMODE YES, you must plan to take an outage of the z/OS systems to activate the STP mode. Once the CTN is configured, you can shut down and re-IPL the z/OS systems using the updated CLOCKxx member.

### 2.7.3 Software planning: z/OS systems currently not in a sysplex

In this section, the assumption is that you have running z/OS systems not in a sysplex prior to configuring a CTN.

You might already be running with the defaults in CLOCKxx member, namely STPMODE YES and STPZONE YES. You might be disregarding the messages z/OS posts at IPL time and z/OS has been completing its IPL in local timing mode. Since the z/OS systems are currently running in local mode, you can configure an CTN without planning to take an outage and shutting down the z/OS systems.

After the CTN is configured (see “Configuring the Coordinated Timing Network” on page 71), plan a *rolling re-IPL* of your z/OS systems using either the defaults in CLOCKxx or updating CLOCKxx to use STPMODE YES and STPZONE YES or NO.

As each z/OS system is re-IPLed, it will now start operating in STP synchronization mode, allowing you to get the advantages stated in the previous section.

**Note:** Once the z/OS systems are running in STPMODE YES, you are now enabled to follow the procedures to set up a sysplex as well.

<sup>4</sup> When a synch check occurs, LPAR takes an offset equal to the synch check and generates an STP machine check. z/OS systems running in STP timing mode would process the sync check by clearing the offset. However, SIMETRID systems are not enabled for STP sync checks and will not get an STP sync check machine check. This means the LPAR will have an offset (aka an EPOCH) which will not be cleared. The logical TOD for the SIMETRID LPARs will drift away from the STP time of the CEC as each sync check occurs. With OA26800 (see <http://www-01.ibm.com/support/docview.wss?uid=isg1OA26800>), the LPAR offset is not cleared at IPL and so there is no mismatch when systems are re-IPLed due to the LPAR EPOCH. However, when an LPAR is reactivated, the LPAR EPOCH is cleared. Now when z/OS is IPLed in that reactivated LPAR it will not be able to join the sysplex due to the LPAR EPOCH mismatch as the running system has a non-zero LPAR EPOCH, while the system being IPLed has a zero EPOCH. This means if one LPAR in the sysplex needs to be reactivated, then all LPARs in the sysplex must be reactivated to ensure there is no mismatch of LPAR EPOCHs.

## 2.8 Migration planning

Over time, to an existing CTN structural changes have to be performed. The server technology changes, new servers have to be added to the CTN or others have to be removed, and also the coupling connectivity technology changes. In this section several aspects are discussed that have to be taken into account when migrations to the actual server and connectivity hardware are performed.

### 2.8.1 CTN types and recovery voting

In a CTN there are different roles a server can have, depending on the number of servers in a CTN, as described in chapter 1.3.5, “Server roles in a CTN” on page 14. These roles are essential for the availability of the time source in the CTN, for the servers obtaining a role perform a voting who has to take over the role of the Current Time Sever (CTS) in a case of recovery scenario.

- ▶ Only one server in the CTN (single server CTN):  
In a single server CTN there is only one server available to perform the role of the CTS. This server is defined as the PTS, and if this server is not available, there is no CTS. The CTN should be configured as a restricted CTN, for which the option **Allow only the servers that are specified below to be members of the CTN** has to be selected in the CTN member restriction preferences, to avoid losing the CTN definitions after the POR of the single server. This is described in 2.1.3, “Support Elements -- Connectivity to ETS”.
- ▶ Two servers in the CTN:  
In a two-server CTN the servers have the roles of PTS and BTS. Normally, the PTS performs the CTS task, and the BTS only takes over the CTS task when it can determine that the PTS is not available. During an unplanned outage, console assisted recovery is used to determine that the BTS should assume the role of CTS. Console assisted recovery is a path, not a destination.
- ▶ Three or more servers in the CTN:  
In a CTN with three or more servers, all three roles PTS, BTS and Arbiter should be assigned. When the connectivity between the PTS and the BTS is lost, a majority voting with the Arbiter is performed to determine if PTS or BTS have to obtain the role of the CTS. This recovery is called Arbiter Assisted Recovery (CAR is not used).

With the IBM zEnterprise 196 GA2 announcement, the recovery voting for STP in a CTN consisting of three or more servers, with the roles of PTS, BTS and arbiter assigned, has been enhanced.

- ▶ A server obtaining a role in a CTN cannot accidentally be removed from the CTN in a planned disruptive action any more, unless its role has been shifted to another server. If a planned disruptive action is attempted on the PTS, BTS or arbiter, STP SE code blocks that action until the role is reassigned to another server in CTN or the role has been removed from the CTN definition. This protects from accidentally causing a *Degraded Triad State* (which means that either PTS, BTS or Arbiter have no connectivity to the other two servers with roles in the CTN).
- ▶ If a Degraded Triad State is entered due to an unplanned outage, which means that when any two of the special role servers (PTS, BTS, Arbiter) agree they cannot communicate with the third special role server, the normal voting is disabled. As a consequence, the recovery mechanisms used in a two-server CTN can be used, when also the connection between the two surviving servers with roles breaks.

This is described in detail in the White Paper WP102037 „STP Enhancement - Recovery Voting“:

<http://www.ibm.com/support/techdocs//atsmastr.nsf/WebIndex/WP102037>

## 2.8.2 Server maintenance and replacement

As described in the previous section, the PTS, BTS and Arbiter roles should be assigned when three or more servers are available. This role assignment should also be assured when a disruptive maintenance to a server is performed or when a server is being replaced.

Since 2012 (with z196 GA2) an active CTN role has to be shifted to another server, before the server having the role assigned can be deactivated or any other disruptive action on the server can be performed. When no other server is available for taking over the role, the CTN definitions have to be altered. If the relocation of the server role has not been performed before the planned disruptive action, the SE code blocks the action.

So we recommend to plan carefully to shift the CTN roles between servers before any disruptive actions to any server holding a CTN role.

## 2.8.3 N-2 Rule and Coupling Link migration considerations

As already discussed in 2.1, “Planning server hardware” on page 22, there is a N-2 connectivity and coexistence rule for servers running in the same sysplex, applicable also to servers running in the same CTN. If an IBM z15 server is present in the CTN, only the following servers are allowed in the same CTN:

- ▶ IBM z15
- ▶ IBM z14
- ▶ IBM z14 ZR1
- ▶ IBM z13
- ▶ IBM z13s

Note that IBM z15 and z14 ZR1 only support the following coupling connectivity:

- ▶ Integrated Coupling Adapter Short Reach (ICA SR) links
- ▶ Coupling Express Long Reach (CE LR) links,

This implies that all servers that connect to a z15 or z14 ZR1 must use these two coupling connectivity options. It is possible to also run z13 servers that only have InfiniBand (IFB) links installed in the same CTN together with z14 ZR1 or z15 servers, but then there cannot be any direct coupling connectivity between these servers.

We recommend considering this CTN layout only for short-term migration scenarios. For permanent use all servers in the CTN should be equipped with the same coupling link technology, respectively ICA and CE LR links.

When migrating from HCA3-O LR links to CE LR links, the same fibre cabling infrastructure can be used, however CE LR cannot connect to HCA3-O LR, and CE LR does not support the STP Going Away Signal (GAS).

The GOSIG is a reliable unambiguous signal send by a HCA3-O or ICA SR link to indicate that the server on which the feature is running is about to enter a failed (check stopped) state. When the GAS sent by the CTS and is received by the BTS, the BTS can safely take over as the CTS without relying on the previously mentioned recovery methods (Offline Signal (OLS) in a two-server CTN or the Arbiter voting in a CTN with three or more servers). However, the GAS can only be sent by coupling links that are installed directly on the CPC drawer fanouts, and since the CE LR card is located in the I/O drawer, the Going Away Signal is not supported by CE LR.

On the other hand, coupling links installed directly on the CPC drawer fanouts go offline if the CPC drawer is removed or maintained. Because CE LR cards are located in the I/O drawers, they benefit from the Enhanced Drawer Availability (EDA) and the Redundant I/O Interconnect® (RII), which ensure that an I/O card does not lose connectivity when a CPC drawer is removed or maintained in a IBM Z server with more than one CPC drawer.

For more details about coupling links connectivity options see *IBM z15 (8561) Technical Guide*, SG24-8851.

## 2.8.4 Power options (PDU/BPA/IBF) considerations

As already discussed in 2.1, “Planning server hardware” on page 22, the IBM z15 servers are offered with two power options: Power Distribution Unit (PDU) or Bulk Power Assembly (BPA). IBM z14 ZR1 servers are only available with the PDU power option. A server with the PDU power option does not support the Internal Battery Feature (IBF). See also 2.4, “Internal Battery Feature (IBF)” on page 46.

With an IBF installed in the CTS, STP has the capability of receiving notification that external power has failed and that the IBF is engaged. Then STP can automatically reassign the role of the CTS to the BTS. If the customer has been using this signal for any recovery automation purposes, this has to be considered before assigning any active CTN role to a server with PDU (without IBF).

For more information about the power options and the IBF see *IBM z15 (8561) Technical Guide*, SG24-8851.



# 3

## Configuration and operations

This chapter discusses considerations when planning for the operational management of the Coordinated Timing Network (CTN). This chapter covers the following topics:

- ▶ Displaying and monitoring the environment
- ▶ Configuring the Coordinated Timing Network
- ▶ Managing the time

## 3.1 Displaying and monitoring the environment

This section discusses the facilities available for monitoring your STP CTN.

### 3.1.1 CPC Detail panel

The CPC Details panel on the HMC is the first panel to look at for timing information about a server. Select the server you want to work with and then click on “System Details”, as shown in Figure 3-1.

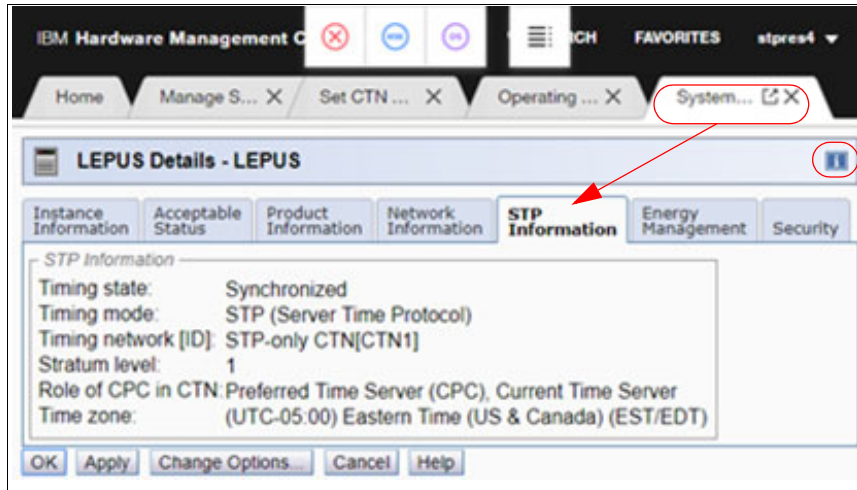


Figure 3-1 System Details - STP Information

If your server has the STP Feature Code (FC 1021) installed, there will be a tab entitled STP Information. This is the easiest way to check whether your servers have the STP feature code installed. The number of fields shown on the STP Information tab varies depending on the type of timing network. This is a display-only panel. It does not allow any modification of the timing network. This menu can be used to observe the current STP status for the selected CPC. Note that there is no refresh of the information on this panel (only the status at the time the menu was accessed is displayed).

The left Context-Sensitive Help area is a function that provides brief descriptions of the action buttons available on each individual panel. The buttons can be sequentially highlighted by pressing the Tab key and thus getting the information for all of them. The help function itself is activated or deactivated by clicking the “i” button in the upper right corner of the opened windows (see Figure 3-1).

**Note:** Ensure that the HMC that you intend to use as an operational focal point for CTN management is at the latest driver level with the latest maintenance change levels (MCLs).

### 3.1.2 The Manage System Time panel

To display and manage the STP configuration, you can access the Manage System Time task for the CPCs known to the HMC.

Figure 3-2 on page 65 shows the Manage System Time task tab.

The screenshot shows the IBM Hardware Management Console (HMC) interface for the 'Manage System Time' task. The top navigation bar includes 'Home' and 'Manage System Time - L...'. The main content area is divided into several sections:

- CTN ID:** Shows 'CTN1' with a dropdown arrow.
- Membership:** Displays 'Unrestricted CTN'.
- STP ACTIONS:** A list of actions including 'Add systems to CTN', 'Configure External Time Source', 'Deconfigure CTN', 'Export CTN data (.xls)', 'Modify assigned server roles', 'Remove systems from CTN', 'Set CTN member restriction', and 'Setup new CTN'. There is also an 'Advanced Actions' section with 'Join existing CTN', 'Save STP debug data', and 'Split to new CTN'.
- Graphical Topology View:** A diagram showing the network topology. At the top, two NTP servers (192.168.70.46 and 192.168.70.65) are connected to a central node containing PTS, CTS, and LEPUS. Below this, two nodes labeled BTS and MUSCA are connected to the central node. The diagram is organized into stratum levels: STRATUM 1 (PTS, CTS, LEPUS) and STRATUM 2 (BTS, MUSCA).

Red callouts are present: #1 points to the CTN ID and Membership section; #2 points to the graphical topology view; #3 points to the top status bar (Time: 6:13:19 AM, Date: 12/7/19, Time zone: UTC-5:00, Current time details); and #4 points to the STP ACTIONS list.

Figure 3-2 The Manage System Time task tab on the HMC

The “Manage System Time” task provides an overview of your CTN. Here you can view details and make changes. This menu is accessible only through the HMC. The “Manage System Time” menu contains the following areas:

1. Global CTN details (see #1 in Figure 3-2)

You can use the selection menu at the small black arrow next to the CTN ID to select the CTN you want to work with. Furthermore, you can see which members belong to the CTN and in which STP stratum and which role they are. It is also easy to see if the selected CTN is restricted or not.

2. Graphical topology view (see #2 in Figure 3-2 on page 65)

Here your CTN is shown with the servers, time sources and connections among each other. You can also see in which Stratum your server is located. By selecting a connection or an object further details can be retrieved:

- When selecting a server, the role within the CTN and the timing status is displayed as shown in Figure 3-3 on page 66 (This view can be zoomed in or out or scrolled if not all CPCs can be seen at one time):

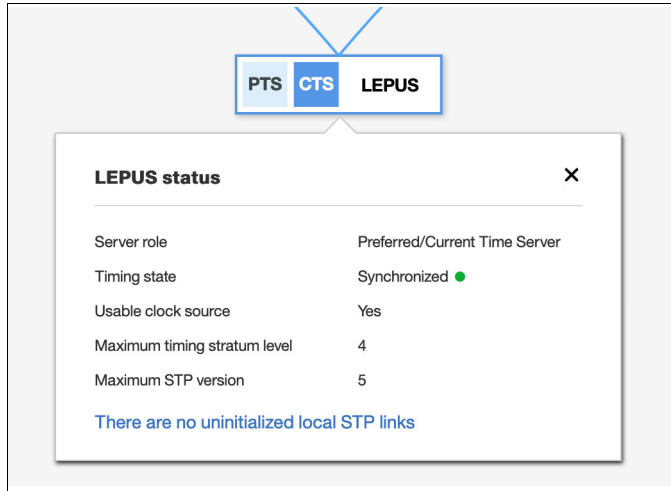


Figure 3-3 System details - STP menus

- When selecting a link among two servers, details about the connection are displayed by clicking on the "See active local STP links" and these are displayed (indicated by the red arrow) in another window, as shown in Figure 3-4:

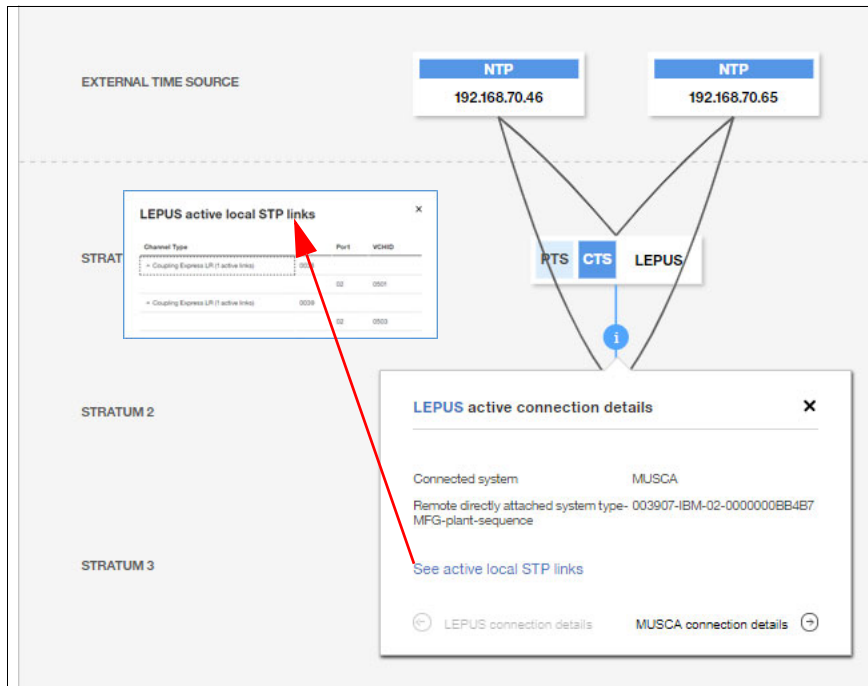


Figure 3-4 Timing links information

- When selecting a connection to the NTP server, details about this connection are displayed as shown in Figure 3-5 on page 67



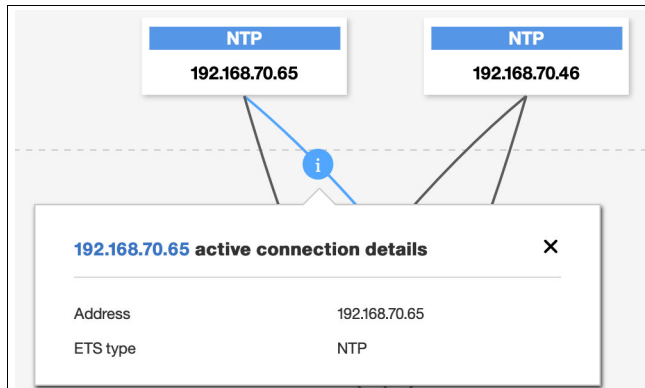


Figure 3-5 ETS link details

3. Topology toolbar (see #3 in Figure 3-2 on page 65)

You can find details about your CTN by clicking “Current Time details”, as shown in Figure 3-6. You’ll also find menus for adjusting the time, setting time zone offset, and adjusting leap second offset. Further information about the tasks can be found in 3.4, “Managing the time” on page 88.

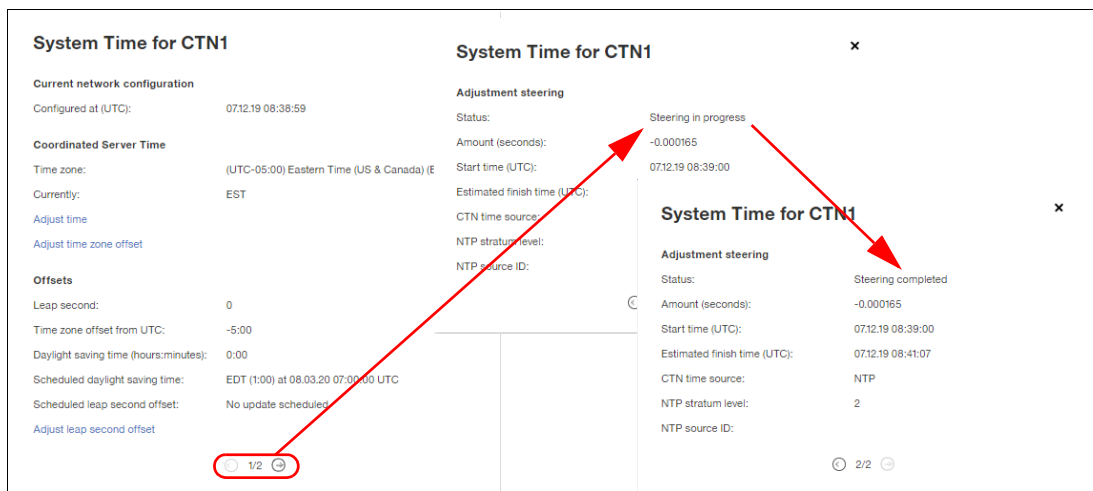


Figure 3-6 Current time details

By clicking on the second page (following the arrows in Figure 3-6) you can find out information about current time status and steering information.

4. STP Actions area (see #4 in Figure 3-2 on page 65)

The menus provided allow you to perform actions for your CTN following the provided workflow. For example, you can add systems to the CTN or configure the external time source. Detailed information is provided in Chapter 4, “Managing STP configurations” on page 105.

### 3.1.3 z/OS display commands

There are a number of z/OS commands that display information about the timing mode or the status of the CTN as viewed by z/OS. The output from these commands can change, depending on the CTN type and the timing mode that the server is currently using.

## DISPLAY ETR

Prior to STP, the DISPLAY ETR command was used to display the synchronization mode and the status of the ETR ports as seen by z/OS. With STP support, the command itself has not changed. However, the output has been updated to display STP-related information where applicable.

In a CTN the TOD clock is being steered to the time provided by the Current Time Server (CTS). No reference to a Sysplex Timer is displayed. See Example 3-1.

### Example 3-1 z/OS DISPLAY ETR command in an STP-only CTN

---

```
D ETR
IEA386I 106.01.17 TIMING STATUS 103
SYNCHRONIZATION MODE = STP
  THIS SERVER IS A STRATUM 1
  CTN ID = CTN1          ETS ID = CNTP
  THE STRATUM 1 NODE ID = 003907.ZR1.IBM.02.0000000BB4B7
  THIS IS THE PREFERRED TIME SERVER
```

---

Stratum Level, CTN ID, Node ID information for the CTS is displayed in addition to a number of extra optional lines that show a different output depending on the CTN topology. For example, if this server has been assigned as a Preferred Time Server, a Backup Time Server or an Arbiter Server role, and the number of usable timing links.

With z/OS 2.4 the command was extended by the ETS ID. Now you can see in z/OS wutch timing mode is used. The following modes are available:

- ▶ CDTS - Console dial up time service
- ▶ CNTP - NTP (Network Timing Protocol) without PPS (Pulse Per Second)
- ▶ CNTX - NTP without PPS from a secondary server
- ▶ PPSN - NTP with PPS
- ▶ PPSX - NTP with PPS from a secondary server

Refer to “z/OS MVS System Messages” for your z/OS version for further information on the output message *IEA386I*. See the IBM z/OS Internet library at:

<https://www-01.ibm.com/servers/resourceLink/svc00100.nsf/pages/zosInternetLibrary?OpenDocument>

## DISPLAY XCF,SYSPLEX,ALL

The DISPLAY XCF,SYSPLEX,ALL command displays the system status and the last recorded system status monitor time stamp for each system in the sysplex, as shown in Example 3-2

### Example 3-2 z/OS DISPLAY XCF,SYSPLEX,ALL command

---

```
D XCF,S,ALL
IXC337I 06.01.28 DISPLAY XCF 106
SYSPLEX PLEX60          MODE: MONOPLEX

SYSTEM SC60            STATUS: ACTIVE
                       TIMING: STP CTNID: CTN1
                       STATUS TIME: 12/07/2019 06:01:28.277410
                       JOIN TIME: 12/07/2019 06:00:02.938771
                       SYSTEM NUMBER: 0100015A
                       SYSTEM IDENTIFIER: B4B73907 0100015A
                       SYSTEM TYPE: 3907 SERIAL: B4B7 LPAR: 01
```

.....

---

## DISPLAY CF

The DISPLAY CF command does not directly provide information regarding the CTN type or timing mode of the server. However, the output does display the CF Request Time Ordering status.

**Note:** CF Request Time Ordering is also referred to as the Message Time Ordering Facility (MTOF).

The DISPLAY CF command can be used to verify whether CF Request Time Ordering is required and enabled, as shown in Example 3-3.

### Example 3-3 z/OD DISPLAY CF command

---

```
D CF
IXL150I 14.55.46 DISPLAY CF 975          COUPLING FACILITY 003907
.IBM.02.000000088888
                PARTITION: 2C  CPCID: 00
                CONTROL UNIT ID: FFFC
NAMED CF7B          COUPLING FACILITY SPACE
UTILIZATION . . . . .
    CFCC RELEASE 22.00, SERVICE LEVEL 02.09
    BUILT ON 11/30/2019 AT 14:18:00
    COUPLING FACILITY HAS 1 SHARED AND 0 DEDICATED PROCESSORS
    DYNAMIC CF DISPATCHING: ON
```

### CF REQUEST TIME ORDERING: REQUIRED AND ENABLED

.....

---

STP requires that each coupling facility within a Parallel Sysplex is enabled for CF Request Time Ordering before any server within the Parallel Sysplex can be defined STP CTN.

Refer to “z/OS MVS System Messages” for your z/OS version for further information on the output message *IXL150I*. See the IBM z/OS Internet library at:

<https://www-01.ibm.com/servers/resourceLink/svc00100.nsf/pages/zosInternetLibrary?OpenDocument>

There are two time-related CFCC commands, which are further described in “Coupling facility commands” on page 98.

## 3.1.4 z/OS messages

Ensure that your operations staff reviews all STP-related messages and plan for which ones they would like automation (or console operator staff) to take action on. Note that some messages will be issued on every member of the sysplex with STPMODE YES specified, which might cause automation to take multiple or redundant actions.

To improve the delivery of important information to the operator and to better integrate with system automation tools, z/OS provides STP-related messages in addition to the hardware messages posted on the HMC.

**Note:** These messages will be posted by every z/OS sysplex member using STPMODE YES in the CTN. The operations staff or automation tools can monitor these messages even if there are stand-alone CFs in the CTN assigned the special roles of PTS and BTS.

To address the CTS change, a new z/OS message was created:

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

Where ccccc is BACKUP or PREFERRED. This informational message does not require any action, but ensures that the operational staff responsible for STP is aware of the change.

**Note:** The IEA031I STP Alert messages will be replaced by IEA037I but will still be displayed in z/OS. Therefore, your automation should only report on the IEA037I messages.

```
IEA037I STP ALERT RECEIVED. STP ALERT CODE = nn, REASON = rr
```

Alert Codes:

- 01 -> Alert related to the external time source
- 02 -> Alert related to the state of the PPS ports
- 04 -> Alert related to the hardware

Please refer to “z/OS MVS System Messages” published for your z/OS version for detailed Reason Codes and further information on the alert Code Message.

Refer to “z/OS MVS System Messages” for your z/OS version for further information on the output message *IEA037I*. See the IBM z/OS Internet library at:

<https://www-01.ibm.com/servers/resourceLink/svc00100.nsf/pages/zosInternetLibrary?OpenDocument>

## **z/OS messages context**

z/OS messages might help you identify various configuration issues and status changes for the servers in your CTN. However, planning the roles of your servers and changing your CTN configuration considering the following aspects:

Although stand-alone CFs will typically provide best connectivity to other servers requiring time synchronization. Stand-alone CFs do not produce z/OS messages for operations or interception by automation routines. For example, the following information messages at IPL or interrupt time might not be displayed:

```
IEA380I THIS SYSTEM IS NOW OPERATING IN STP TIMING MODE.
IEA381I THE STP FACILITY IS NOT USABLE. SYSTEM CONTINUES IN LOCAL TIMING MODE
```

If the condition being raised relates to connectivity between two servers, the information might be available to a z/OS system image at the other end of the link. However, if both ends of the link are CF partitions, no warning message is available to the user.

There are several IEAxxx and IXCxxx messages that report current and changed timing status. As an example, the following message reports the result of a successful rename of an CTN:

```
IXC438I COORDINATED TIMING INFORMATION HAS BEEN UPDATED
FOR SYSTEM: SC60
PREVIOUS CTNID: CTN2
```

CURRENT CTNID: CTLSOMRS

In general, there are no z/OS messages that are posted only on the PTS, BTS, or Arbiter.

- ▶ Certain messages do not appear on the CTS since it is the time source:

```
IEA382I THIS SERVER HAS ONLY A SINGLE LINK AVAILABLE FOR TIMING PURPOSES.
IEA383I THIS SERVER RECEIVES TIMING SIGNALS FROM ONLY ONE OTHER NETWORK NODE.
IEA390I TOD CLOCKS DYNAMICALLY ADJUSTED TO MAINTAIN STP SYNCHRONISM.
```

- ▶ The following message might not appear on certain special role servers:

```
IEA388I THIS SERVER HAS NO CONNECTION TO THE nnnnnnnnnn
Where nnnnnnnnnnn = 'PREFERRED 'I 'BACKUP 'I 'ARBITER '
```

For example, the following message never appears on a z/OS system running on the BTS:

```
IEA388I THIS SERVER HAS NO CONNECTION TO THE BACKUP
```

### 3.2 Configuring the Coordinated Timing Network

This section describes how to configure a CTN.

#### 3.2.1 Single Server CTN

The sequence of steps in configuring a Coordinated Timing Network with only a single server consists of the following steps:

1. Set the CTN ID
2. Specify CTN Members
3. Choose PTS
4. Choose CTS
5. Set Leap Seconds
6. Set Time Zone
7. Configure External Time Source
8. Set Date and Time
9. Confirm Changes

We will create a Single Server CTN from scratch with the CPC named "LEPUS" which is not a member of any CTN at this point in time (see Figure 3-7 on page 71).

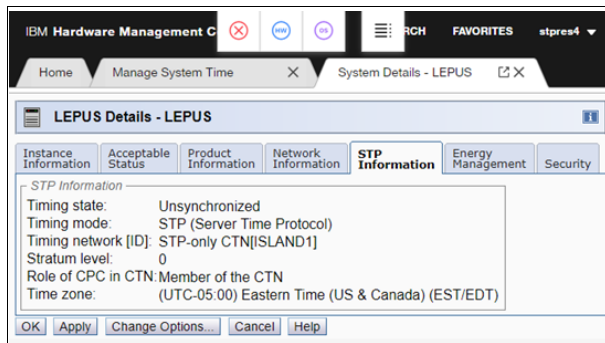


Figure 3-7 Server Details - STP information for LEPUS

The process of creating a new CTN is implemented in the workflow 'Setup new CTN' in the 'Manage System Time'-Panel on the HMC.

1. Create a new CTN (Note, that if there are no servers available, this operation is not allowed). See Figure 3-8.

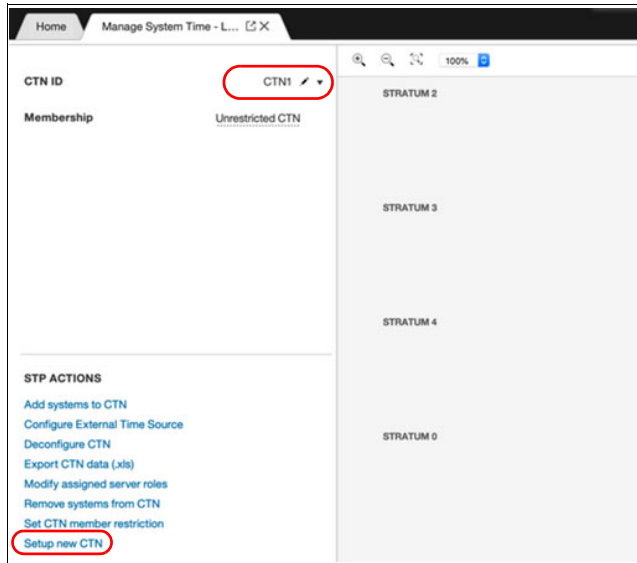


Figure 3-8 Create new CTN

The first step is to set the new CTN ID. Select the “Setup new CTN” task, then you type in the name of your new CTN and continue with 'Next' (Figure 3-9).

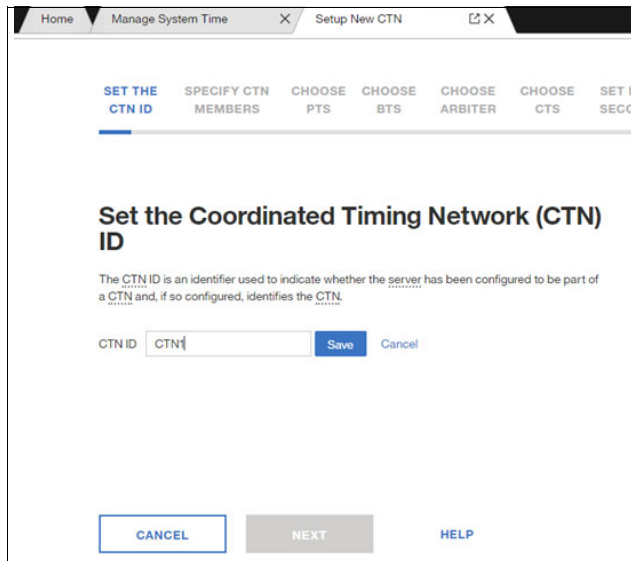


Figure 3-9 Setting the CTN ID

2. Specify CTN Members

In this step you can select the systems that shall become members of the new CTN. Only systems that are currently not a member of any active CTN are shown here. Once all members are selected continue with 'Next'. (Figure 3-10 on page 73).

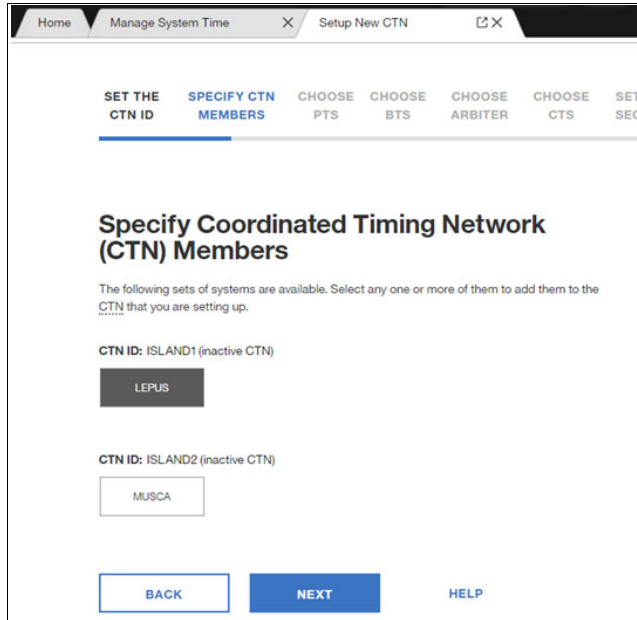


Figure 3-10 Specify LEPUS as a CTN member

3. Choose the PTS

In this step you select the server that will become the PTS. All systems that are available for role assignment are being displayed. In this example we select LEPUS as PTS and click on 'Next' (see Figure 3-11 on page 73)

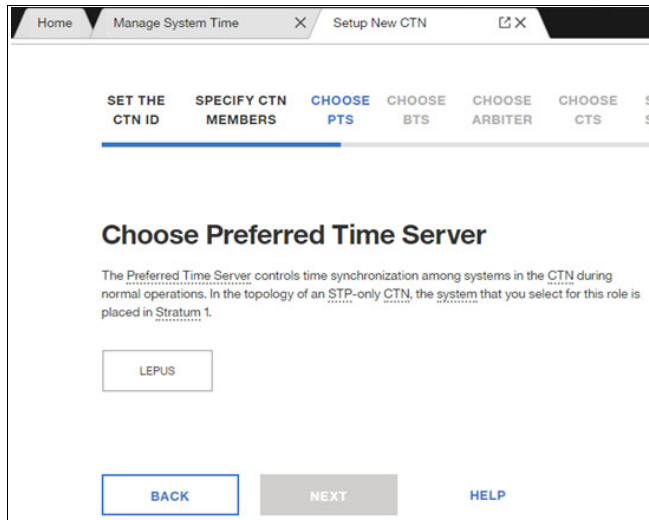


Figure 3-11 Choosing the Preferred Time Server

4. Choose the CTS

Here you select the system that is going to be the Current Time Server and so will become Stratum 1. The CTS adjusts the Coordinated Server Time by steering it to the time that is being obtained from an external time source. Since we only have a single server in our configuration the PTS will become the CTS, as shown in Figure 3-12 on page 74.

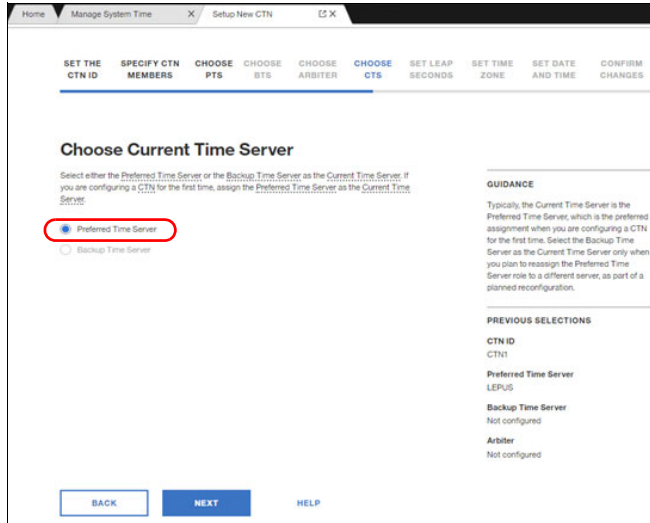


Figure 3-12 Choosing the Current Time Server

5. Set the Leap Seconds

In this window you can specify the leap seconds for the CTN. If you do not have the requirement to add leap seconds you can leave this value at 0. You can add leap seconds to your CTN at a later point in time without interruption. See also -> Pointer to Leap Second Section. Enter the desired amount of leap seconds and Click on 'Next' (see Figure 3-13 on page 74).

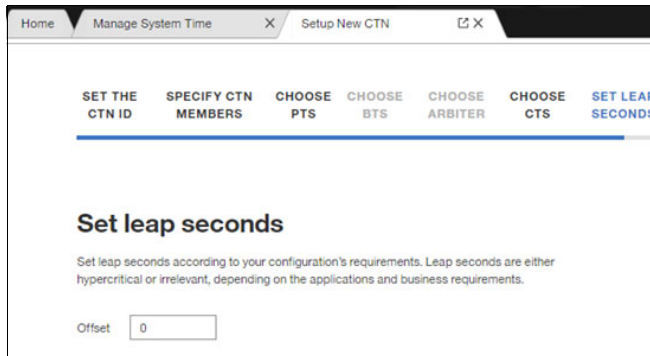


Figure 3-13 Setting the Leap Seconds

6. Set the Time Zone

Here you can select a time zone from the dropdown panel or create your own time zone by selecting one of the user defined timezones (UD1 to UD5).

If the systems in your CTN span multiple time zones see 3.4.5, "Parallel Sysplex and multiple time zones" on page 100.

In the section Clock adjustment for daylight saving time you can specify your DST offset and whether to automatically adjust for DST as shown in Figure 3-14 on page 75, or to set DST, or standard time manually.



Figure 3-14 Setting the Daylight Savings Time

## 7. Set the Date and Time

This window lets you define the CTN time. The most common option will be using an external time source, but you can also manually define a date and time or specify a delta by which the time will be modified.

We choose to use the external time source. If no external time source is configured clicking on 'Next' will lead you to the 'Configure External Time Source' option. If an external time source has been configured before, you can manually go to this option by clicking on 'Configure External Time Source' (see Figure 3-15 on page 75).

Once you finish configuring the external time source you will automatically continue with the 'Setup new CTN' task.



Figure 3-15 Moving to configuring the External Time Source

## 8. Configure External Time Source

- a. As a first step you select the system to modify its external time source (Figure 3-16 on page 76).

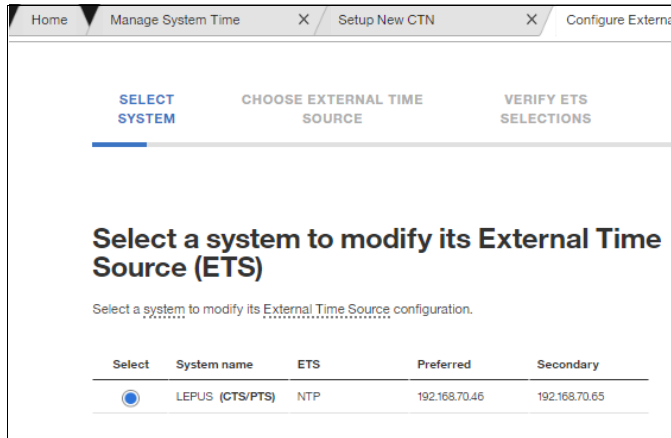


Figure 3-16 Selecting the system for the ETS

- b. The next step is to choose from the ETS options. You can select NTP or NTP with PPS for enhanced accuracy or you can even select to not use an external time source (Figure 3-17).

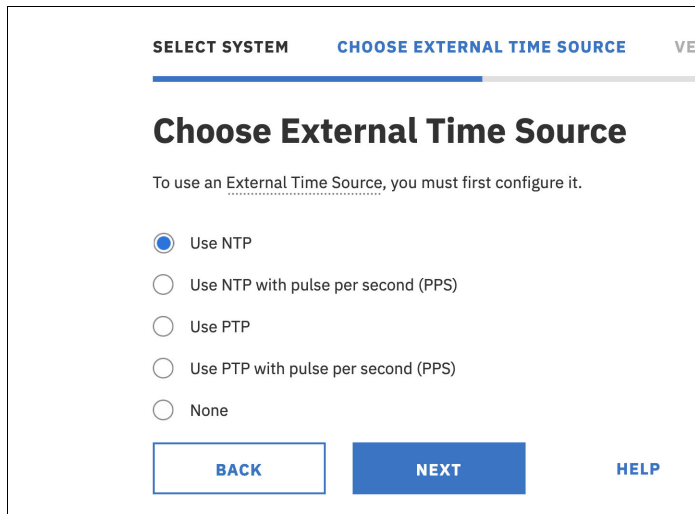


Figure 3-17 Choosing NTP as ETS

- c. In the following step you can enter IP addresses or domain names of NTP servers. 'Test Connectivity' will check if the NTP servers can be reached and display the results in the Stratum, Source and Status fields (see Figure 3-18 on page 77).

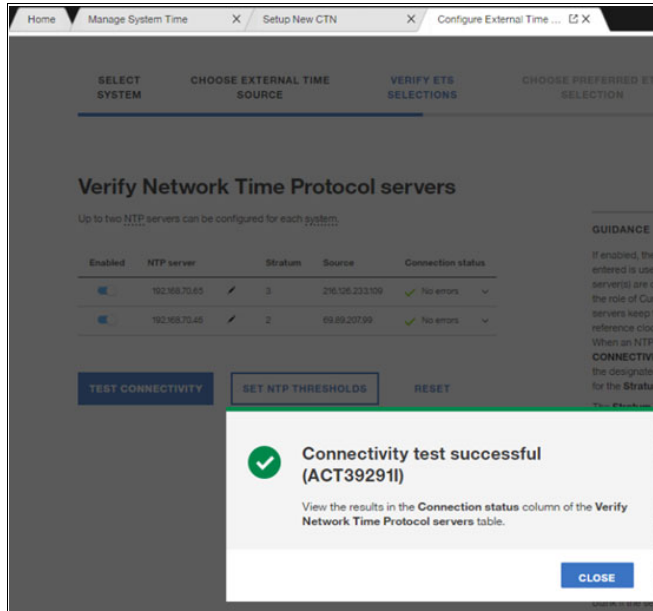


Figure 3-18 Testing NTP Connectivity

- d. After this you specify which of the NTP-servers is the preferred one. Usually you should select the one with the lower Stratum level since that one provides better accuracy. Before you apply the new settings you get an overview of all settings in the 'Confirm External Time Source configuration' window. The new settings and the old settings are displayed next to each other so you can easily see the changes that will be performed. If everything looks as desired, click on 'Apply' (see Figure 3-19 on page 77)

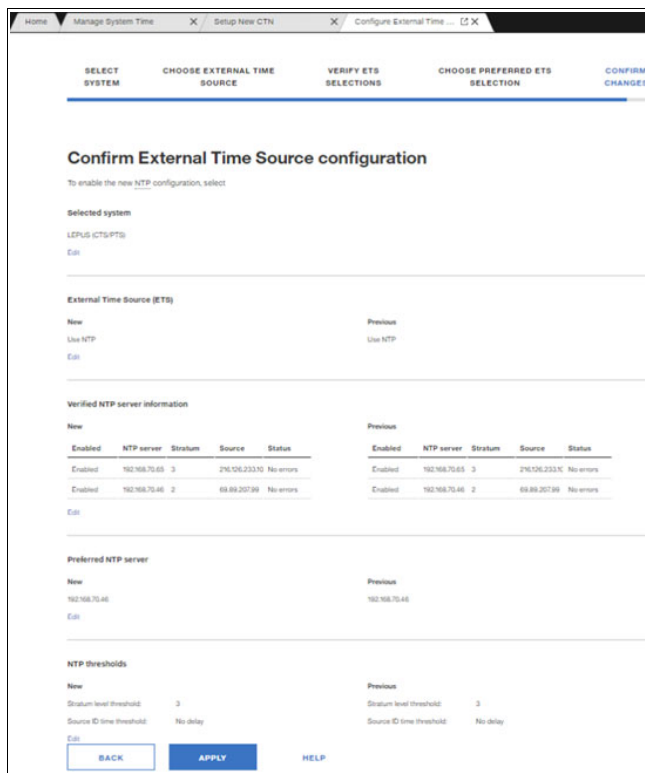


Figure 3-19 ETS configuration - confirmation of changes

9. CTN Configuration - Confirm Changes

Once the ETS configuration has ended you are lead back to the 'Setup New CTN' workflow. The last panel is 'Confirm Changes' where you get a graphic of what your CTN will look like (Figure 3-20).

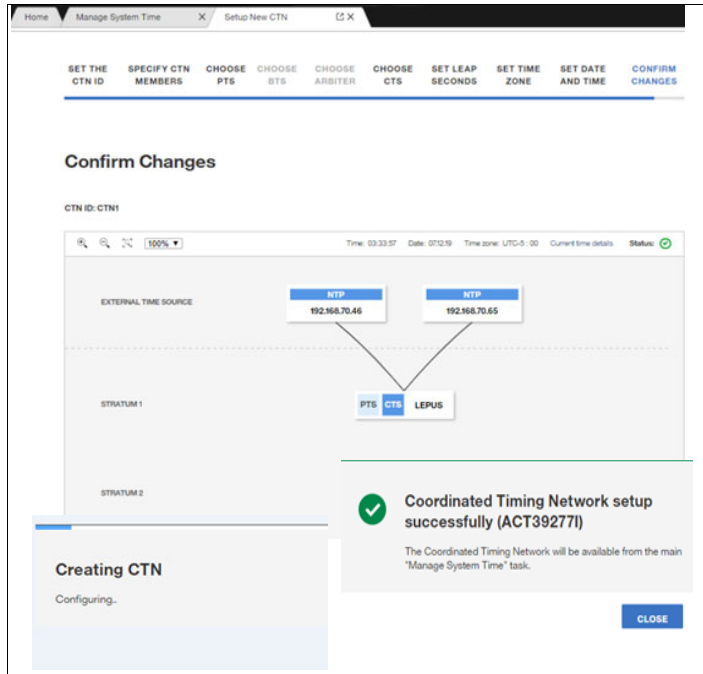


Figure 3-20 CTN Configuration - final step

Once the task has finished you can see your new CTN in the 'Manage System Time' task on the HMC, as shown in Figure 3-21 on page 78.

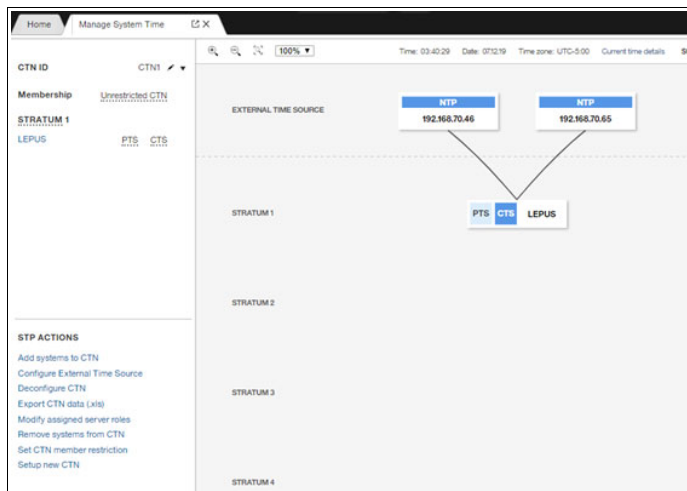


Figure 3-21 CTN Configuration - Result

3.2.2 Dual Server CTN

The sequence of steps in configuring a Coordinated Timing Network with PTS and BTS consists of the following steps:

1. Set the CTN ID

- 2. Specify CTN Members
- 3. Choose PTS
- 4. Choose BTS
- 5. Choose CTS
- 6. Set Leap Seconds
- 7. Set Time Zone
- 8. Configure External Time Source
- 9. Set Date and Time
- 10. Confirm Changes

The process is identical to configuring a Single Server CTN (see 3.2.1, “Single Server CTN” on page 71) except for the additional step 4 (Choose BTS, after selecting the PTS - see Figure 3-22).

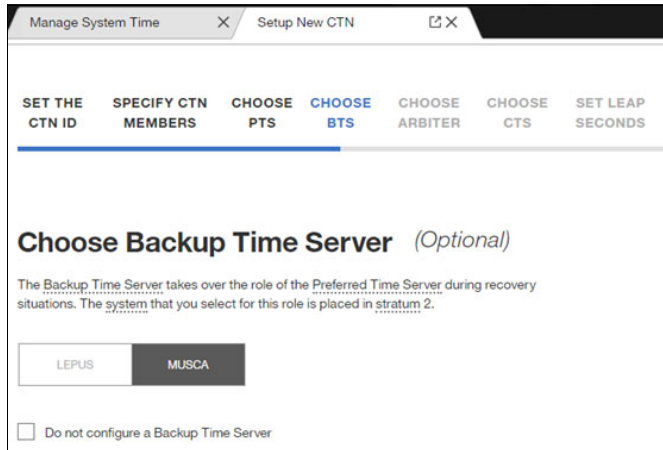


Figure 3-22 Choosing Backup Time Server

The window displays the systems that are eligible for BTS role assignment. The systems that have already been assigned the PTS role is displayed as unavailable (in this case LEPUS). You can also choose not to assign a BTS. Before committing the changes you will be presented the 'Confirm Changes' window where you can see what your CTN will look like, then you can apply the configuration. The configuration is shown in Figure 3-23 on page 79.

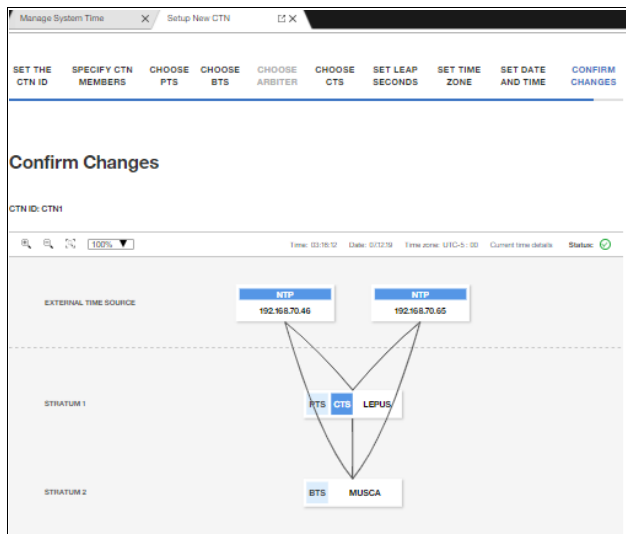


Figure 3-23 Dual server CTN - configured CTN

**Important:** At least one external time source needs to be configured for both CPCs that will assume the roles of PTS and BTS. It is recommended to set up dual ETS for each of the two CPCs, preferably, the same external time sources (primary/alternate) for both servers.

### 3.2.3 CTN with three servers - PTS, BTS, and Arbiter

The sequence of steps in configuring a Coordinated Timing Network with PTS, BTS and Arbiter consists of the following steps:

1. Set the CTN ID
2. Specify CTN Members
3. Choose PTS
4. Choose BTS
5. Choose Arbiter
6. Choose CTS
7. Set Leap Seconds
8. Set Time Zone
9. Configure External Time Source
10. Set Date and Time
11. Confirm Changes

The process is identical to configuring a Dual Server CTN (see 3.2.2, “Dual Server CTN” on page 78) except for the additional step 5 (Choose Arbiter). Step 5 will become available if a BTS is selected (see Figure 3-24).

The servers to assume the arbiter role must have coupling connectivity to both the PTS and the BTS. Also, there is no need to configure an external time source to this CPC.

Figure 3-24 Choosing the Arbiter (three or more servers in a CTN)

When all changes have been applied you can see the layout of the new CTN in the Manage System Time tab, as shown in Figure 3-25 on page 81.

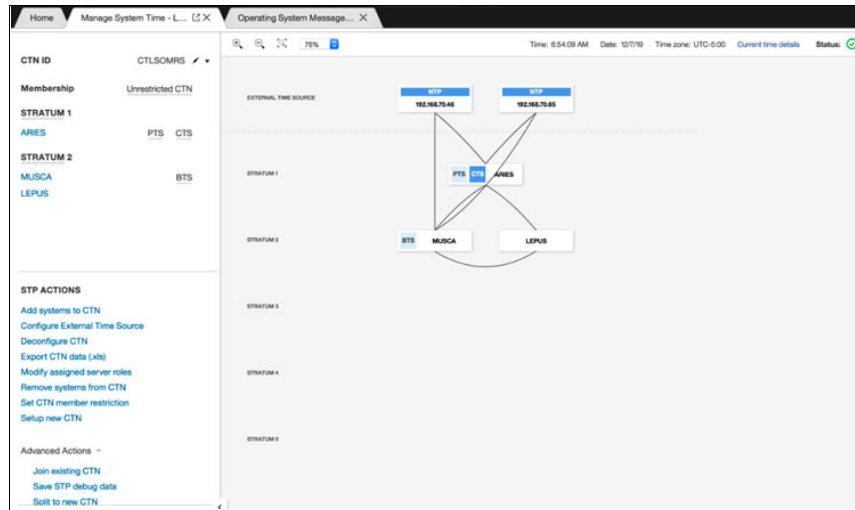


Figure 3-25 Three server CTN - final configuration

### 3.3 Configuring Precision Time Protocol as ETS

In this section we introduce the procedure used to configure Precision Time Protocol (PTP) as external Time Source (ETS) for one of the CPC in the CTN.

**Important:** This section assumes that the IBM Z Support Element (SE) has been properly connected to (PTP supporting and enabled) network infrastructure with at least one Ethernet port. The physical infrastructure, from PTP Grandmaster and any network elements used to connect the Support Element ethernet ports em3 and/or em4, must be configured to ensure PTP signaling communication as well as IP routing between the SE and the PTP Grandmaster.

The scenario is provided as a guide to migrate the ETS for a CPC (member of a CTN) from using NTP to using PTP as ETS. The scenario starts with a three-CPC CTN that has the PTS using PTP as ETS and the BTS using NTP as ETS. The initial configuration is shown in Figure 3-26 on page 82.

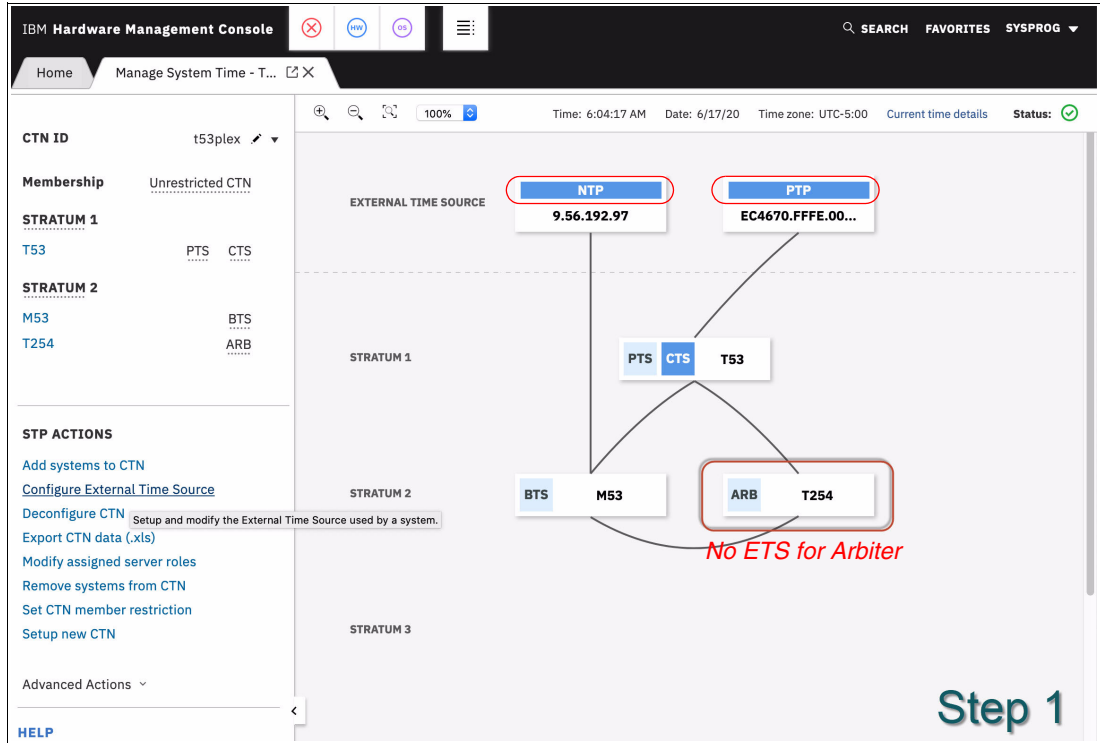


Figure 3-26 Initial configuration - three server CTN w/ mixed ETS

This is a short description of the scenario:

Initial status: The servers in the “t53plex” are configured as follows:

- T53 - PTS and CTS, uses PTP w/ PPS as ETS
- M53 - BTS, uses NTP w/ PPS) as ETS
- T254 - Arbiter - has no ETS.

The following steps (high level) are used:

- ▶ The ETS infrastructure is prepared for T254 and appropriate Ethernet ports on the SEs are connected to a PTP capable switch. Figure 3-27 shows the ports of the z15 Support Element which can be configured for PTP connectivity.

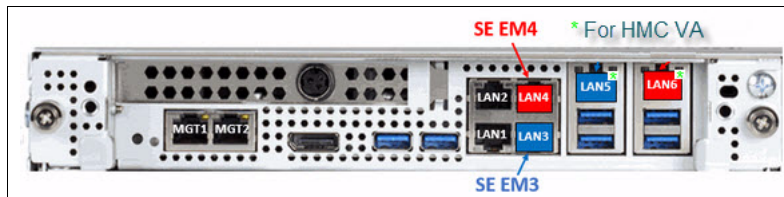


Figure 3-27 z15 Support Element ports used for PTP (em3, em4)

**Note:** Currently, only z15 CPCs support PTP as ETS. For current implementation, connectivity to ETS is done via the support element.

- ▶ We configure the ETS for the Arbiter (T254)
- ▶ We re-assign server roles in the CTN (T254 becomes BTS, while M53 becomes the Arbiter)



We describe in detail only the ETS assignment for T254 (PTP w/ PPS). The CTN roles reassignment is described in 4.2.4, “Changing the server roles” on page 111.

### 3.3.1 Configuring ETS (PTP w/ PPS)

1. We start by selecting the “Configure External Time Source” (Figure 3-26 on page 82). The server selection menu is opened, as shown in Figure 3-28.

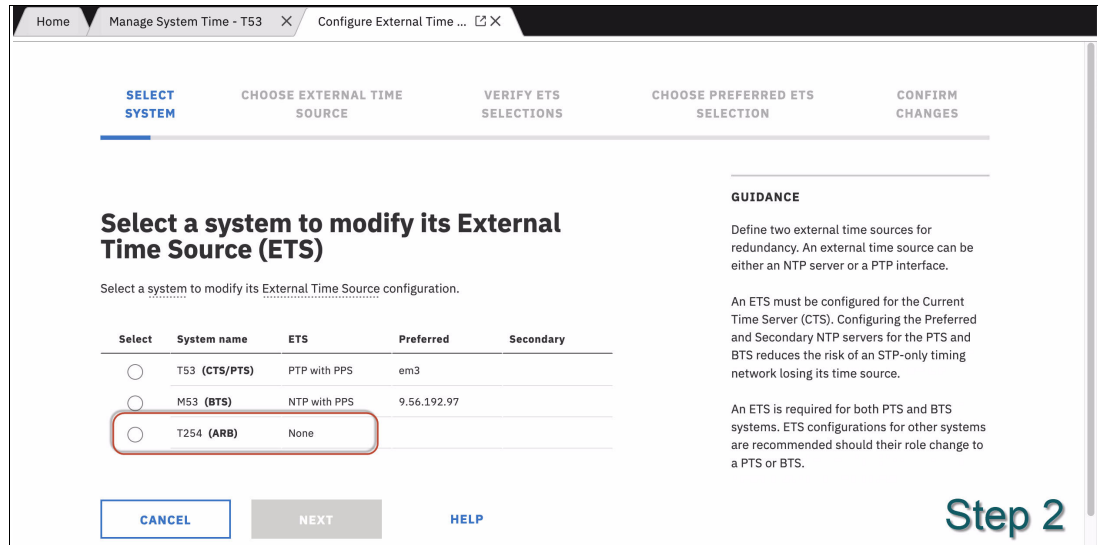


Figure 3-28 System selection

2. Next, we select the system - T254, as shown in Figure 3-29, and click “Next”.

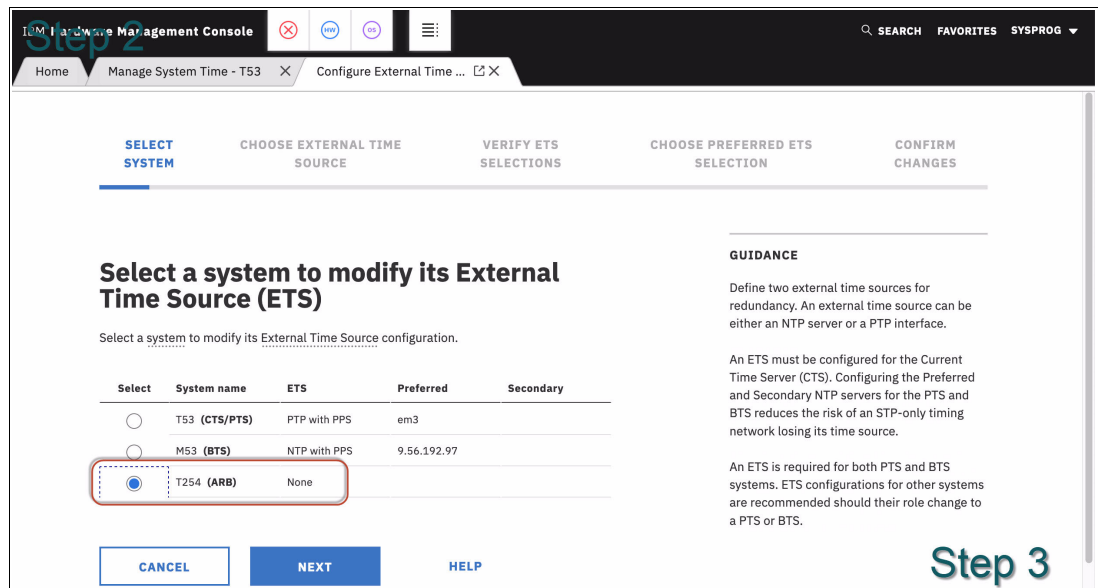


Figure 3-29 T254 - system selected

3. For the External Time Source type, we select “Use PTP with pulse per second (PPS)”, as shown in Figure 3-30, then click “Next”.

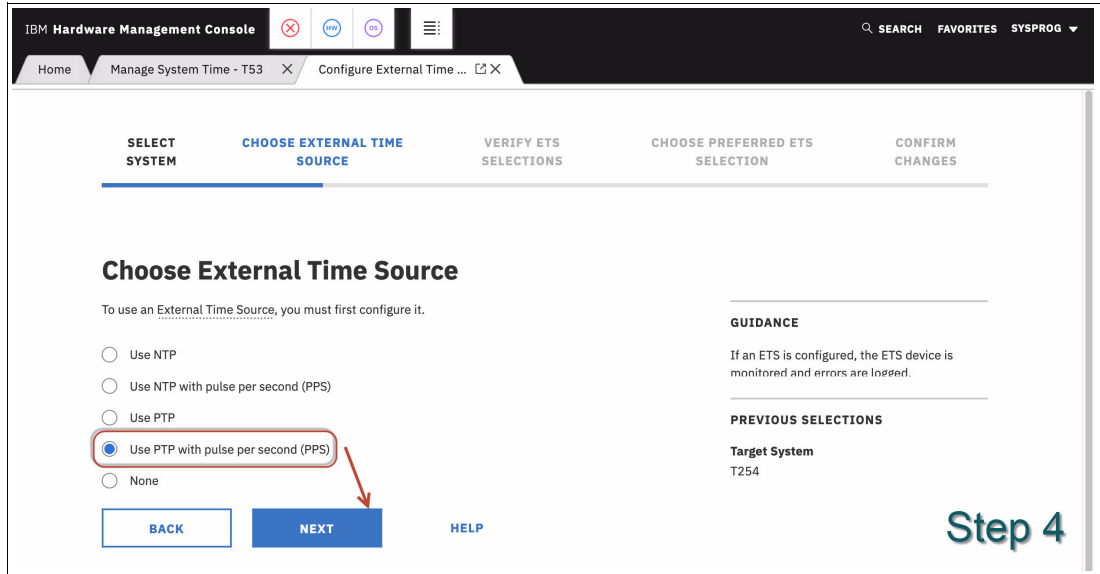


Figure 3-30 ETS type selection

- At this time there is no specified Ethernet interface for accessing the ETS from the Support Element, so we proceed further to SE interface selection. The SE Ethernet interface selection menu is shown, as in Figure 3-31. For Ethernet ports naming, see also Figure 3-27 on page 82.

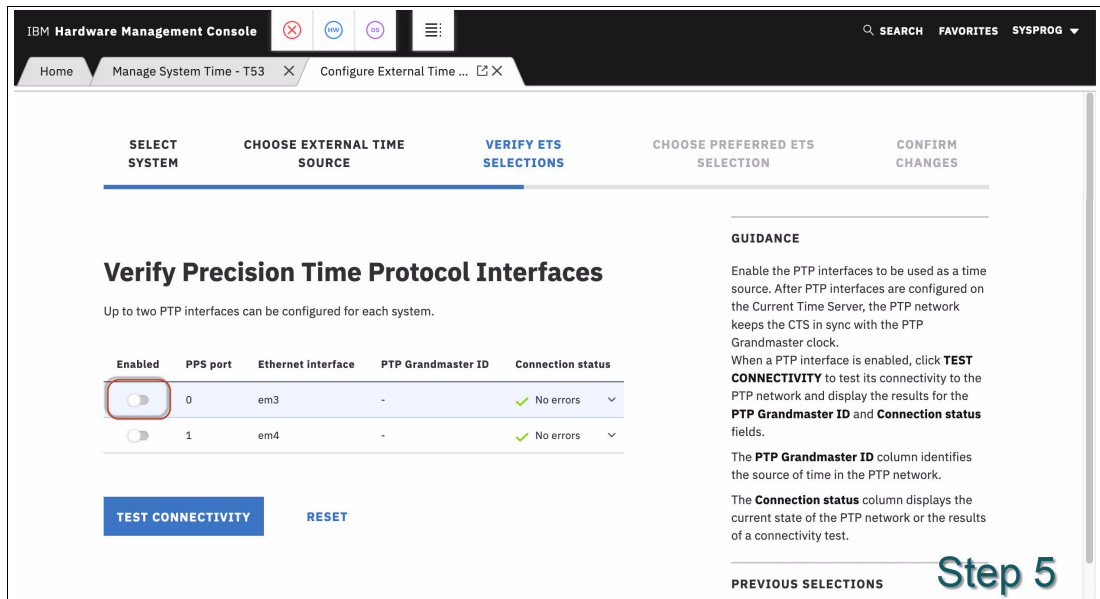


Figure 3-31 SE Ethernet interface selection

- Since we know that “em3” interface has been connected, we select “em3”. SE test the connectivity (once the interface has been selected), as shown in Figure 3-32 on page 85.

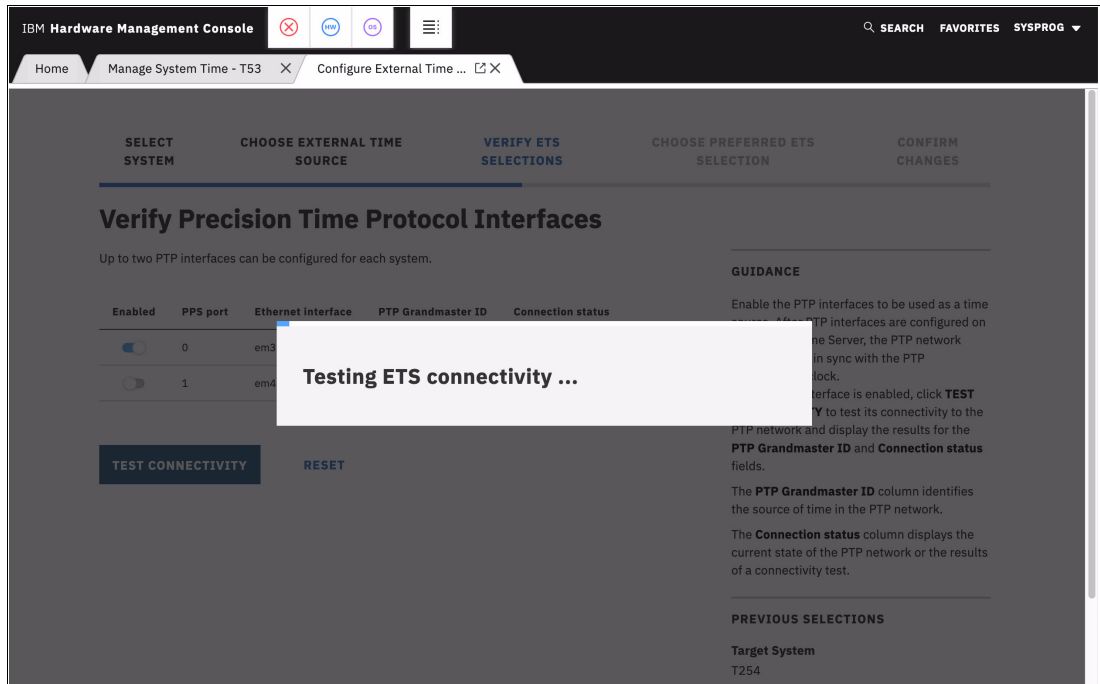


Figure 3-32 PTP source connectivity test

6. Once the test has been successful, the message shown in Figure 3-33 is displayed. Click “Close” to proceed to next step.

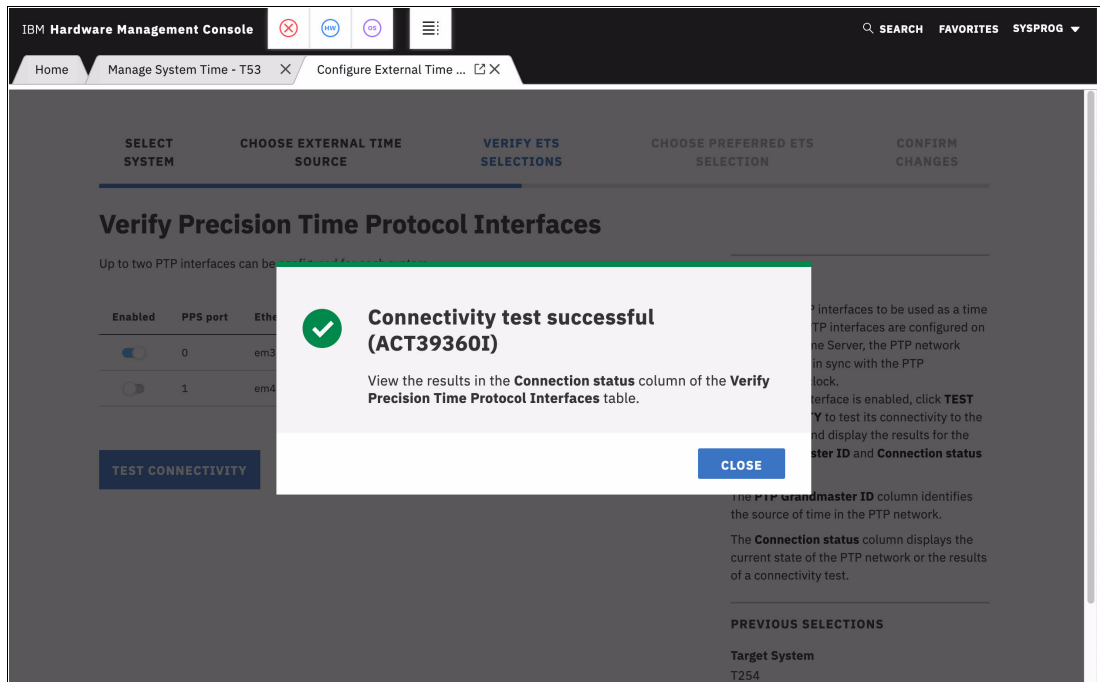


Figure 3-33 ETS Connectivity test success

7. The configured interface is shown in Figure 3-34 on page 86. The PTP Grandmaster is discovered and its name is populated automatically. Click “Next” to proceed to next step.

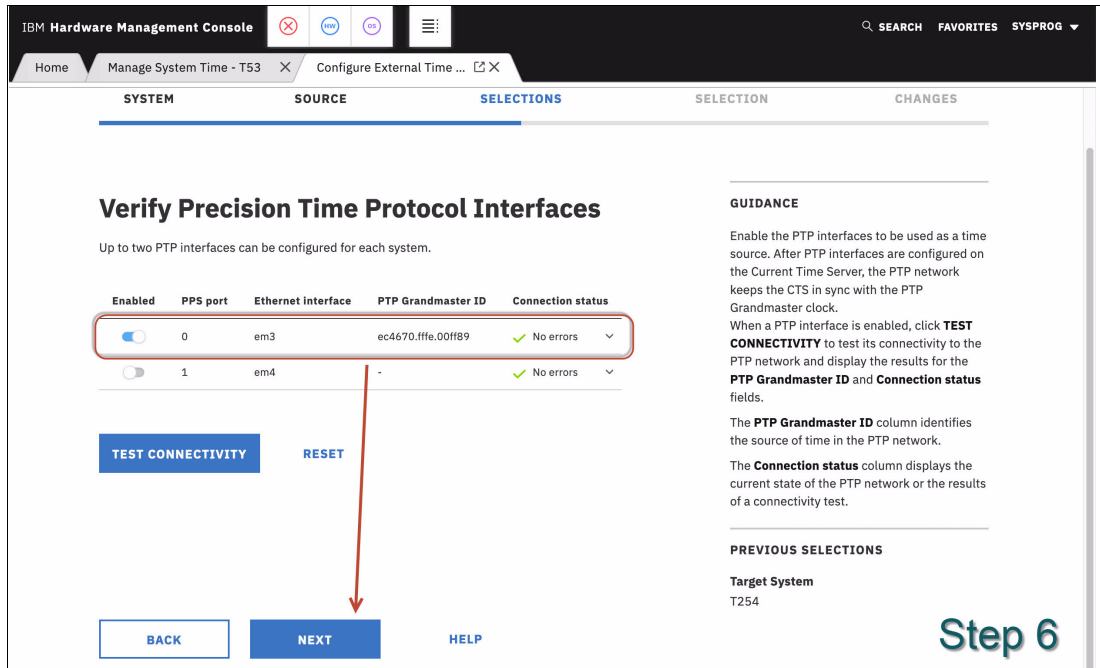


Figure 3-34 Verify interfaces for PTP

- 8. Once the Ethernet interfaces for PTP access have been configured, we select the preferred interface (em3), as shown in Figure 3-35 and click “Next”.

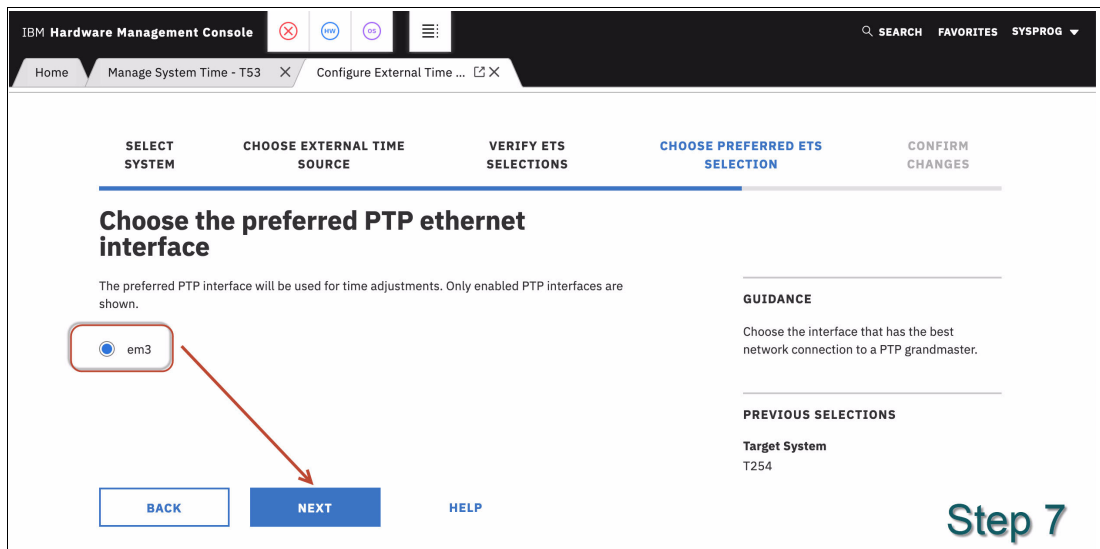


Figure 3-35 Preferred interface selection

- 9. Configuration verification - confirmation menu is show, as in Figure 3-36 on page 87We click “Next” to commit changes.

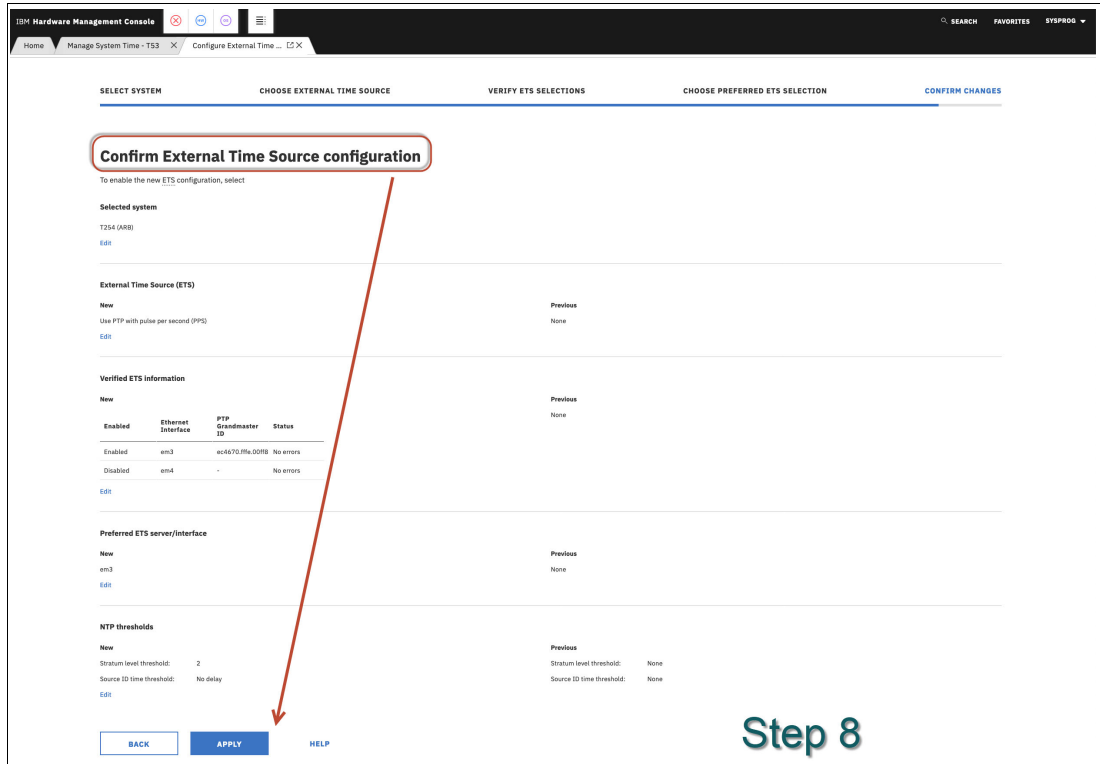


Figure 3-36 Confirmation menu

10. Successful configuration message is displayed in Figure 3-37. Click “Close”.

**Note:** Make sure you save the data you have just entered to the alternate Support Element, as the message in Figure 3-37 indicates.

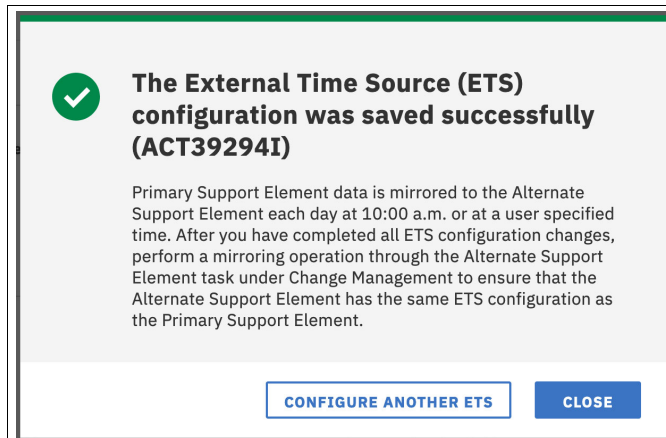


Figure 3-37 Success message for ETS configuration

11. At this time the CTN configuration has not changed. The Arbiter (T254) has access to an external time source and is prepared for assuming PTS or BTS role.

12. We swap the server roles in the “t53plex” CTN (procedure is described in 4.2.4, “Changing the server roles” on page 111. The final configuration is shown in Figure 3-38 on page 88.

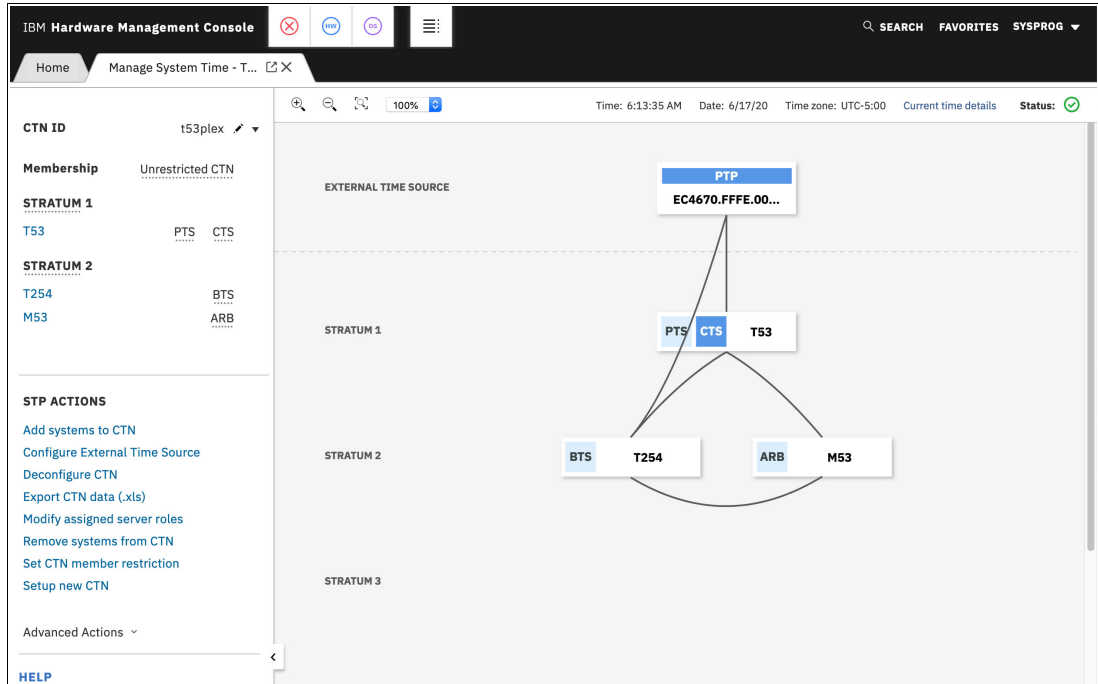


Figure 3-38 PTS and BTS use PTP as ETS

## 3.4 Managing the time

This section discusses the following time modifications within a CTN and the operational requirements:

- ▶ STP time adjustment
- ▶ STP offset adjustments
- ▶ Changes in local time
- ▶ Logical partition time offset
- ▶ Parallel Sysplex and multiple time zones

### 3.4.1 STP time adjustment

Without regular adjustment, the time on the Current Time Server drifts from its initial setting because of the frequency drift of the oscillator used to step the TOD clock. This might not meet time accuracy requirements. Therefore, adjustments might need to be made on a regular basis.

This can be done manually or by referencing an external time source.

Manual Adjustments of time can be performed by clicking Adjust time from the “Current time details” Panel (over the graphical topology view) in the Manage System Time task, as shown in Figure 3-39 on page 89.

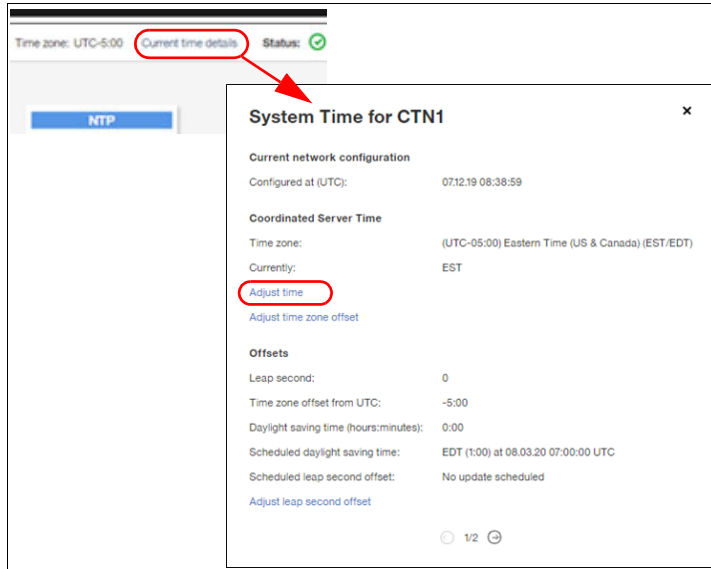


Figure 3-39 Adjust time

In the Adjust Time window, if you click “Access ETS”, the deviation of the CTN time from the ETS time is calculated and displayed. If you now confirm the menu by clicking “Apply”, the CTN time is successively adjusted to the ETS time, as shown in Figure 3-40.

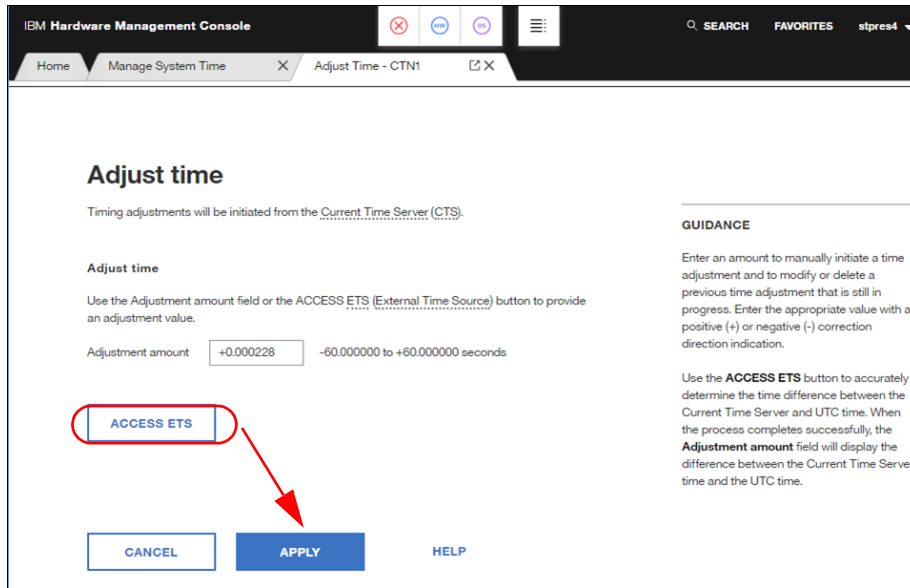


Figure 3-40 Applying time adjustment

The following adjustments are possible:

### Adjustment steering

STP supports adjustment steering, which allows the time at the Current Time Server to be changed by up to +/- 60 seconds. Adjustments greater than 60 seconds can be implemented in multiple increments of +/- 60 seconds. This can take considerable elapsed time to achieve.

The offset specified is gradually incorporated into the standard timing messages in small enough increments or decrements that the operating systems, subsystems, and applications are unaware that time is speeding up or slowing down.

The input of the offset to be steered out is done either manually or through the ETS.

**Important:** In a CTN, the adjustment steering rate is approximately one second every 7 hours.

**Base steering**

There is another time-steering method built into the STP facility that works in a similar fashion to the adjustment steering. This is an automatic function requiring no user control. This is known as base steering, which is performed at the Current Time Server and requires an ETS.

By comparing the time obtained from subsequent accesses of a configured ETS with the corresponding Coordinated Time Server values, STP can compute the amount of drift that has occurred between the events. This represents the inherent inaccuracy of the Current Time Server oscillator over time. With this information, STP can automatically introduce a compensation offset into the STP timing messages by additional steering to counter the drift. As a result, the all servers in the CTN self-correct over time so that the offset returned from future ETS accesses approaches zero as greater accuracy is achieved.

Base steering is only performed if the server is an NTP stratum 1. The stratum level of the NTP server can be seen in the Manage System Time Panel by clicking on the NTP Server, shown in Figure 3-41:

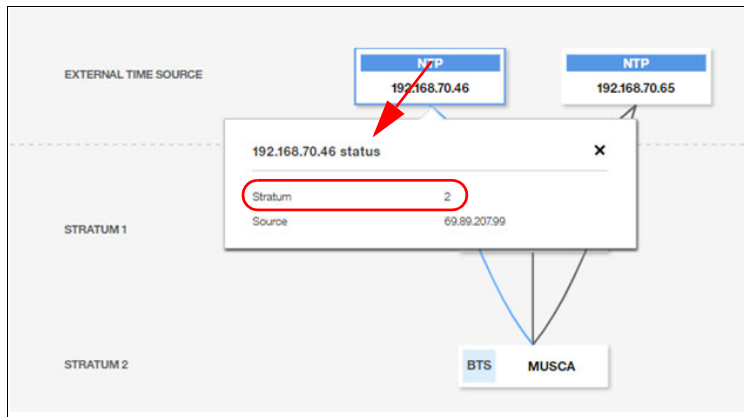


Figure 3-41 NTP stratum information

Table 3-1 provides a summary of STP clock adjustment functions.

Table 3-1 STP time steering

	NTP server access adjustment	Manual adjustment	Manual set
Adjustment steering	Yes	Yes	No
Base steering	Yes, after multiple ETS accesses	No	No



**Important:** Once a CTN is configured, the only way to set the date and time is to deconfigure the CTN and then go through the process to initialize the time again. Refer to Chapter 4, “Managing STP configurations” on page 105 for details.

The Adjust Time button is not available when the CTN time source is NTP with pulse per second. Since STP utilizes PPS input every second, manual adjustment is disabled.

### 3.4.2 STP offset adjustments

The STP timing message includes the following:

- ▶ Coordinated Server Time
- ▶ Leap second offset
- ▶ Time zone offset
- ▶ Daylight saving time offset

The values in the previous list are transmitted from the Current Time Server to all the servers in the CTN. How the z/OS system image uses these values is dependent on options specified in the TIME macro in combination with options specified in CLOCKxx at IPL.

Different time results can be received depending on the options specified in the TIME macro, as shown in Table 3-2.

Table 3-2 TIME macro options

	TIME macro with ZONE=LT	TIME macro with ZONE=UTC	TIME macro with STCK
Include TOD in result	Yes	Yes	Yes
Include leap second offset	Yes	Yes	No
Include time zone offset	Yes	No	No

In addition, the parameters specified in the CLOCKxx member at IPL determine where these values are obtained.

Table 3-3 shows comparable details for z/OS systems on a server in STP timing mode.

Table 3-3 z/OS system on a server in STP timing mode

	STPMODE=NO STPZONE=NO	STPMODE=YES STPZONE=NO	STPMODE=YES STPZONE=YES
Step TOD to Current Time Server.	No <sup>a</sup>	Yes	Yes
Include time zone offset from Current Time Server.	No	No	Yes
Include time zone offset from CLOCKxx.	Yes	Yes	No
Allow local time adjustment via z/OS SET commands.	Yes	Yes	No

- a. The base TOD steering affects all logical partitions regardless whether they operate in STP mode or not because it changes the H/W TOD gradually without notifying it to operating systems. The STP mode makes the difference when the local H/W TOD drifts and differs from Stratum 1 (CTS/CST) resulting in an STP synch check. The STP synch check is to be presented only to logical partitions running in STP mode. Several other STP machine checks and external interruptions are also to be presented only to those logical partitions running in STP mode.

### Leap second offset adjustment

Adjustments are only necessary (if leap seconds are used) if the applications require precise synchronization accuracy to UTC. Examples of such specific requirements might be legal or contractual requirements for time stamps to be within tolerance of UTC Time, or if time stamps are used for time-dependent banking, scientific, or navigational purposes.

Periodically, the International Earth Rotation and Reference Systems Service in Paris advises that a leap second adjustment, which might be either positive or negative, is required to be introduced into civil time. For the most current leap seconds information refer to (search for “leap seconds”):

<https://www.iers.org/IERS/>

Each of the leap seconds introduced so far has been positive. However, it is theoretically possible to have a negative leap second adjustment.

**Important:** Leap seconds are automatically built into UTC time obtained from an external time source. Any leap second offset that is defined is taken into account when calculating the delta between the Current Time Server and the time received from the ETS. This is required to prevent double accounting.

If the ETS is used to incorporate an additional leap second into the Current Time Server TOD, then be aware that the requirement that leap second adjustments should occur at the same time worldwide has been breached. This is because TOD adjustments via the ETS are implemented over a period of time via adjustment steering rather than immediately.

If leap seconds are used, the adjust leap seconds facility should be used to ensure that the new offset is applied as a single adjustment at the correct time.

Typically, adjustments to the leap second offset are scheduled to be applied on either June 30th or December 31st. However, the offset adjustment might be scheduled for a particular date depending upon the requirements.

A positive leap second is inserted after UTC 23:59:59, as UTC 23:59:60, then UTC 00:00:00 would occur as normal. Note that 23:59:60 will never be made visible to z/OS users of the TOD clock. A negative leap second is potentially possible (however, it has not occurred to date) and is implemented by excluding UTC 23:59:59.

When the leap second offset occurs, z/OS is interrupted just as it is for DST offset changes if STPMODE YES was specified in the CLOCKxx member of SYS1.PARMLIB. The offset value is updated if STPZONE YES was also specified.

The Adjust Leap Second panel is used to apply the leap second offsets when they become available (see Figure 3-42 on page 93). Clicking Adjust Leap Second offset from the “Current time details” Panel (over the graphical topology view) in the Manage System Time Task.

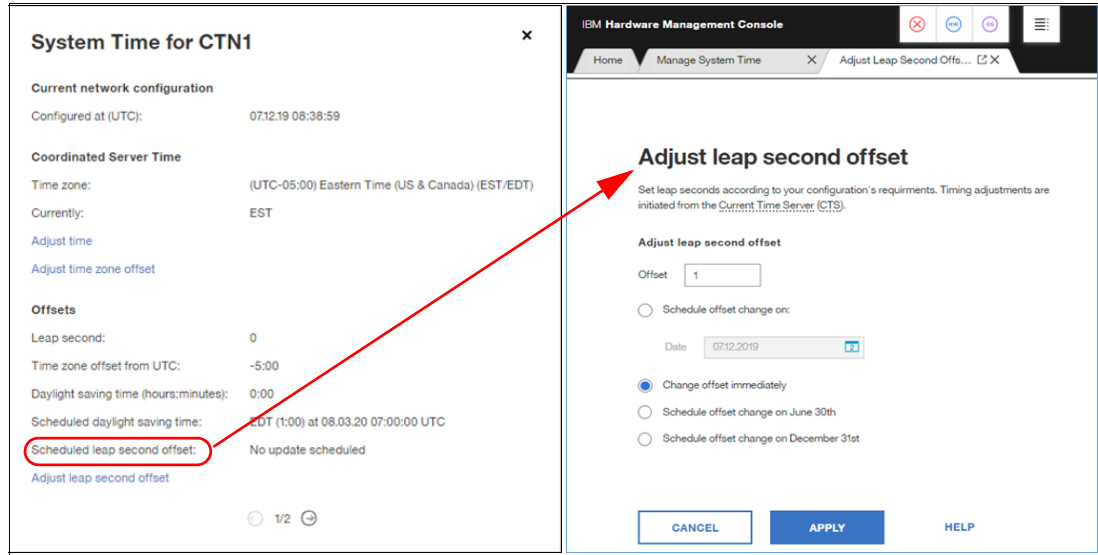


Figure 3-42 Leap seconds adjustment

**Attention:** The value of “1” for the leap seconds is just for display purposes of this scenario. Please use a correct value for the leap second offset (currently the value for leap seconds is 27).

When the Adjust Leap Second Offset panel is displayed, the current leap second offset in effect for the STP CTN value is displayed in the Offset box. This value is either inherited from the Sysplex Timer during the migration process from an ETR network or set during the Initialize Time function.

**Important:** If you want to add a new leap second (negative or positive) you have to add or subtract this leap second from the value displayed in the panel.

While z/OS spins for positive leap seconds, there is no need for a spin during a negative leap second.

A negative leap second would just appear to any application that 1 second had jumped past without them seeing it. A large change in leap seconds can cause a system outage.

You can also set the time at which the leap seconds should be implemented in the panel.

**Important:** Leap second adjustments occur at the same time worldwide (regardless of the time zone), when UTC 23:59:59 will occur, and it will be taken care of by STP and the operating system (no user intervention or interaction).

### Time zone offset adjustment

In the “Current time details” panel (over the graphical topology view) in the Manage System Time Task you can see the currently set time zone offset.

With the task “Adjust Time Zone Offset” you can change the time zone. Select one of the supported time zones that are provided by default or use a user-defined time zone to meet the requirements (see Figure 3-43 on page 94).

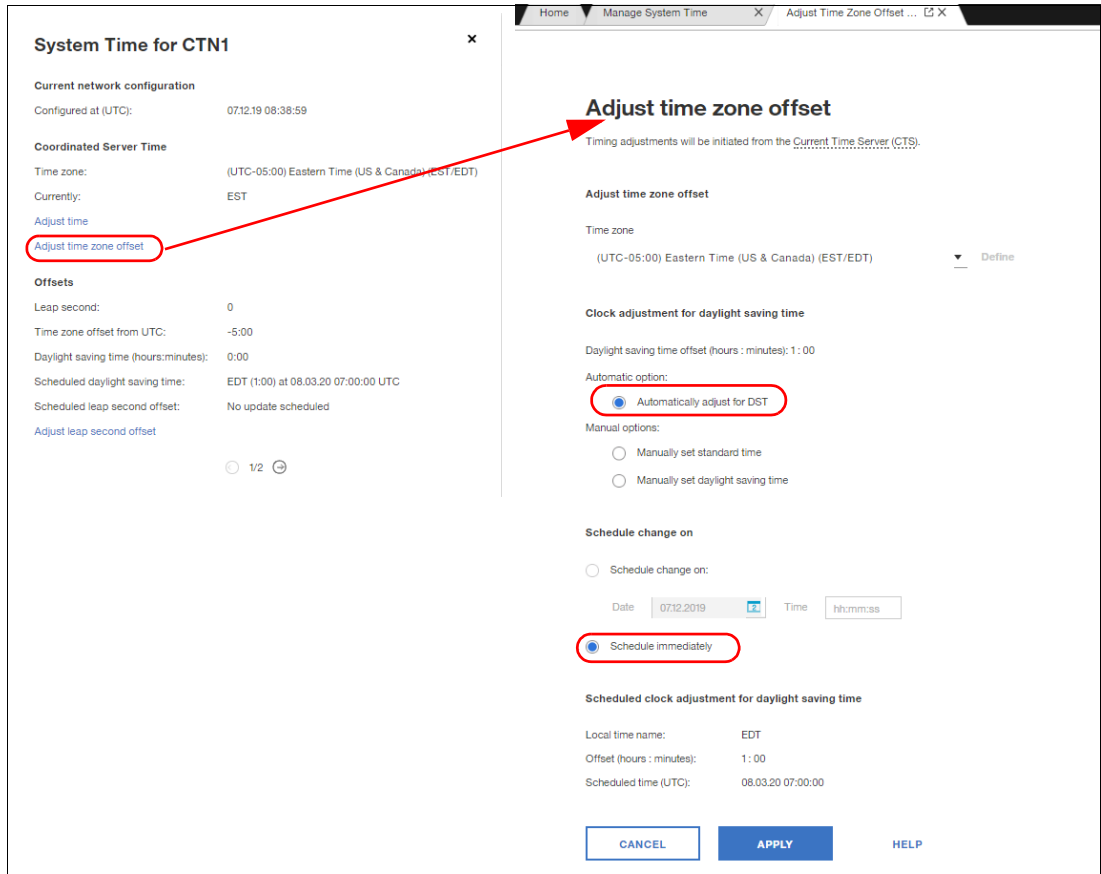


Figure 3-43 Timezone adjustment

### Supported time zone offsets

A number of supported time zone entries are provided by default. Each of these entries has a defined offset from UTC.

**Note:** The time zone on the Support Element is independent of the time zone set for the CTN. The same is also true for the HMC. It is possible to have a different time zone at each Support Element, the HMC, and for the CTN.

In addition, each entry can optionally have a time zone algorithm defined that contains the following daylight saving information:

- ▶ Daylight saving offset
- ▶ Daylight saving automatic adjustment information (optional):
  - Daylight saving date and time start algorithm
  - Daylight saving date and time end algorithm

If the selected time zone supports automatic adjustment by providing a time zone algorithm with the necessary start and end information, then the Automatically adjust radio button can be used to activate automatic adjustment for the site.

Alternatively, the time zone might not support automatic adjustment, or is handled manually. In this case select either Set standard time or Set daylight saving time to indicate whether a daylight saving period is active when the selected time zone is activated.

The Scheduled Clock Adjustment for Daylight Saving Time section is displayed if an adjustment has been scheduled either automatically as the result of an Adjust Time Zone offset algorithm or manually through the schedule change on facility.

**Tip:** There is no direct way to review the time zone algorithm that might optionally be part of a supported time zone entry. Therefore, there is no direct way to verify that the daylight saving time offset entry for the time zone will meet the requirements. However, this can be done indirectly by applying the time zone offset entry for the time zone with automatically adjust and then reviewing fields in the Schedule Clock Adjustment for Daylight Saving Time section for applicability correctness in relation to the next time zone adjustment.

### ***DST changes implemented by a government authority***

An example of this is the USA Energy Policy Act of 2005 that changed daylight saving time in the USA in 2007.

For an STP CTN, which has the option of automatically scheduling DST, support for the new daylight saving time start and end dates are available provided that the server has the correct level of microcode.

Future DST changes implemented by any governmental authority might require MCL updates. Make sure to keep the server MCLs as up to date as possible.

### ***User-defined time zone offsets***

If a supported time zone entry that meets the requirements cannot be found, then one of the five user-defined time zones (that is, UD1 to UD5) can be used to define the desired time zone. If a user-defined time zone entry is selected, the Define button is enabled (Figure 3-44 on page 95).

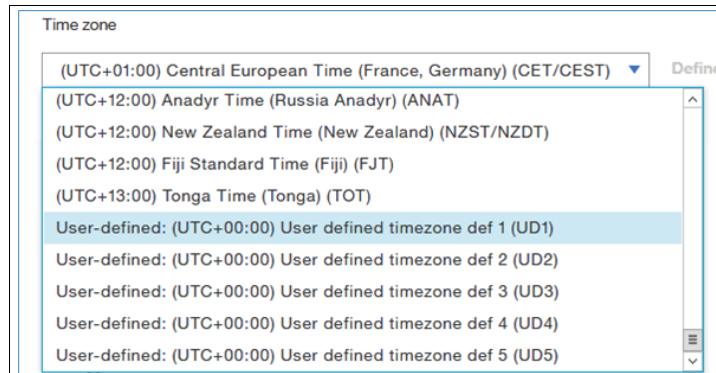


Figure 3-44 User-defined timezone entries

The Description (maximum of 80 characters) and Standard time name fields (maximum of 4 characters) must be entered, otherwise an error message appears when **OK** is clicked. The standard time name is an abbreviation displayed on various panels to differentiate standard time from daylight saving time.

The UTC offset must be entered in +/- hours and minutes and ranges from -14 to +14 hours.

Also, if the time zone is subject to daylight saving adjustments, then the daylight saving time name and daylight saving offset must be specified. Optionally, algorithms for daylight saving time start and daylight saving time end can be defined to support automatic clock adjustment by selecting the Define adjustment of clock for daylight saving time option. The algorithm is saved when **OK** is clicked, but it is not sent to STP until **Apply** is clicked on the Adjust Time Zone Offset panel.

### 3.4.3 Changes in local time

z/OS allows the user to obtain either STCK time, UTC time, or local time depending on their requirements. The difference between UTC time and local time is usually the time zone offset under normal circumstances. The time zone offset can be managed at the z/OS level by specifying the STPZONE=NO options in the CLOCKxx member.

**Note:** In z/OS, the ETRMODE can be set in the CLOCKxx member, which requires a connection to a Sysplex timer. In terms of hardware, a connection of the mainframe server to a Sysplex Timer is no longer possible since the IBM z196. Therefore, the ETRMODE should no longer be set. If the server runs in STP timing mode, and both STPMODE and ETRMODE were specified as YES, z/OS uses STPMODE YES.

When this is done, the TIMEZONE parameter in the CLOCKxx member is used to set the time zone offset at IPL, and a number of z/OS SET commands can be used to dynamically adjust the offset when required. Similarly, the coupling facility supports the concept of time zone offset and allows dynamic modification of the time zone offset via a command.

#### z/OS commands

On a z/OS system, local date and time can be modified dynamically. The ability to do this is dependent on what options have been specified in the CLOCKxx member at IPL, as shown in Table 3-4.

Table 3-4 z/OS time adjustment through command cross-reference

	Adjust time via z/OS command		
	Local time ZONE=LT	UTC time ZONE=UTC	STCK time STCK
STPMODE=NO, STPZONE=NO	Yes	Yes <sup>a</sup>	Yes <sup>a</sup>
STPMODE=YES, STPZONE=NO	Yes	No	No
STPMODE=YES, STPZONE=YES	No <sup>b</sup>	No	No

a. UTC time cannot be changed by a z/OS command. It can be modified during the IPL processing response to message.

```
IEA888A [UTC DATE=yyyy.ddd,CLOCK=hh.mm.ss] LOCAL
DATE=yyyy.ddd,CLOCK=hh.mm.ss
```

REPLY U, OR UTC/LOCAL TIME which is issued at IPL when OPERATOR PROMPT is specified in CLOCKxx with STPMODE NO and SIMETRID not specified.

b. Any attempt to change the local time or date when the server is operating in STP mode generates message IEA279I: IEA279I ALL CLOCK RELATED SET COMMANDS ARE IGNORED WHEN IN STP MODE.

#### DISPLAY TIME

This command can be used to display the local time and date, and the UTC time and date, as seen in Example 3-4.

Example 3-4 z/OS DISPLAY TIME command

```
D T
IEE136I LOCAL: TIME=hh.mm.ss DATE=yyyy.ddd UTC: TIME=hh.mm.ss DATE=yyyy.ddd
```

Under normal circumstances the difference between local time and UTC time is the time zone offset (incorporating daylight saving time offset, if any) applicable to the time zone.

### **SET DATE**

This command is used to change the local date. The syntax is:

```
SET DATE=yyyy.ddd
```

The command has the following restrictions:

- *yyyy* is the year. It must be in the range 1900-2042. The value specified must consist of four digits and must be within 70 years of the UTC date, otherwise the SET command is ignored.
- *ddd* is the day. It must be in the range 001-366 and meet leap year restrictions.

The maximum date that can be specified is 2042.260.

### **SET CLOCK**

This command is used to change the local time. The syntax is:

```
SET CLOCK=hh.mm.ss
```

This command is used in conjunction with the SET DATE command to set a maximum value of 23.53.47 on 2042.260. The server's TOD clock is not updated by this command, nor is the logical TOD of the logical partition this z/OS image is operating on. The change made by this command is effective for the duration of this IPL only.

Also, z/OS does not change the date when the new time implies a change of date, so either use the DATE parameter or wait for time to pass midnight if the new time is for tomorrow.

### **SET RESET**

This command causes the time zone offset to be reset to the value that was read in from the CLOCKxx member during IPL, causing the local date and time to be changed accordingly.

The syntax is:

```
SET RESET
```

This annuls all previous SET DATE, SET CLOCK, and SET TIMEZONE commands and reestablishes the relationship local date and time = UTC date and time + time zone offset.

### **SET TIMEZONE**

This command can be used to change the time zone offset to a different value from that specified at IPL via the TIMEZONE parameter in CLOCKxx. This automatically adjusts local date and time accordingly. The syntax is:

```
SET TIMEZONE={W|E}.hh[.mm]
```

The time zone offset direction is West (W) or East (E). West is defaulted if not specified. The value for hh must be between 00 and 15 and the value for mm must be between 00 and 59.

The daylight saving time changes can be handled manually using the SET CLOCK command rather than having it done automatically via Server Time Protocol (STP). Using this method there is always a degree of error because the difference between the local time and UTC time will not exactly match the time zone offset that would have been achieved by updating the TIMEZONE statement in CLOCKxx and IPLing.

The new z/OS SET TIMEZONE command overcomes this problem by applying the correct offset value in the CVTTZ field, causing an exact time zone offset to be applied.

**Tip:** If the SET CLOCK command is used to change local time for daylight saving offset purposes, then use the SET TIMEZONE command instead for better accuracy.

Remember that if the time zone offset for daylight saving purposes is changed dynamically using either the z/OS SET CLOCK or the SET TIMEZONE commands, the TIMEZONE statement in CLOCKxx must be updated so that the new offset is not regressed at the next IPL.

## Coupling facility commands

In a Parallel Sysplex environment, coupling facilities require time awareness to support CF request time ordering. The server TOD is used for this purpose.

Coupling facilities also support the concept of time zone offset, which is used only for the purpose of timestamping messages that are displayed on the console.

There is no CFCC command available to display time. However, all messages that appear on the CF console include a time stamp in local time format, which is the server TOD with the time zone offset applied. Therefore, the current local date and time at the CF console can be indirectly determined by entering any command (valid or invalid) and reviewing the time stamp in the resulting response.

Because the CF supports a local time format that incorporates the time zone offset, it also provides methods to both display the current time zone offset setting and to change it if required.

### **DISPLAY TIMEZONE**

Use the CFCC DISPLAY TIMEZONE command to display the current time zone offset being used by the coupling facility. The syntax is:

```
DISPLAY TIMEZone
```

This produces a single line indicating how many hours and minutes the current time zone is east or west of Greenwich Mean Time (Example 3-5 Example 3-18).

*Example 3-5 CFCC DISPLAY TIMEZONE command*

---

```
2006272 11:06:47 => display timezone
2006272 11:06:47 CF0271I Timezone is 04:00 West of Greenwich Mean Time
```

---

### **TIMEZONE**

The CFCC supports a command allowing the time zone offset to be changed, if this is a requirement. The syntax is:

```
TIMEZone {0|hh|hh:mm|:mm} {East|West}
```

Use the command shown in Example 3-6 to adjust the local time as displayed in messages on the coupling facility console for the onset and removal of daylight saving time.

*Example 3-6 CFCC adjusting timezone*

---

```
2006272 11:17:31 => timezone 05:00 west
2006272 11:17:31 CF0271I Timezone is 05:00 West of Greenwich Mean Time
```

---

### **Coupling facility implications at daylight saving time changes**

When a CF image partition is activated and the server is using Server Time Protocol source, the CFCC uses only one of the following time offset options:



- ▶ The logical partition time offset specified in the image profile The TIMEZ offset
- ▶ The TIMEZ offset overrides the logical partition time offset. Use the TIMEZ command for DST changes as described in the document at the following website:

<http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD103077>

### 3.4.4 Logical partition time offset

IBM Z servers support a function called logical partition time offset. Logical partition time offset3 support provides for the optional specification of a fixed time offset (specified in days, hours, and quarter hours) for each logical partition activation profile. The offset, if specified, is applied to the time that a logical partition receives from Server Time Protocol.

It is sometimes necessary to run multiple Parallel Sysplexes with different local times and run with the time set to TOD Clock LOCAL. This causes the results returned in the store clock (STCK) instruction to reflect local time. With logical partition time offset support, logical partitions on each server in a Parallel Sysplex that must do this can specify an identical time offset that shifts time in the logical partition sysplex members to the desired local time. The logical partition time offset value, if specified, is applied to the time value that a logical partition receives from the time source. The time zone offset and DST offset are independent of this parameter.

Remaining logical partitions on the servers can continue to participate in current date production Parallel Sysplexes utilizing the same STP messages with the time provided by STP. This function is supported by all supported releases of z/OS.

Open the image profile of the image to be modified. From the General tab, select Logical Partition Time Offset from the Clock Type Assignment area (Figure 3-45 on page 99).

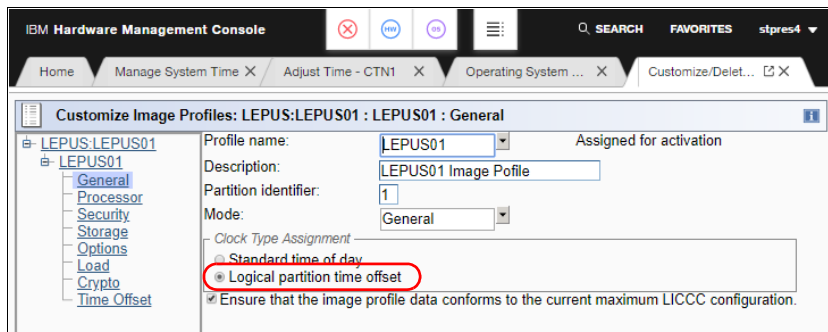


Figure 3-45 Customize Image Profiles - General

A new selection becomes available on the left side of the panel, called *Time Offset*. Use the Time Offset window to set the offset, as seen in Figure 3-46 on page 100.

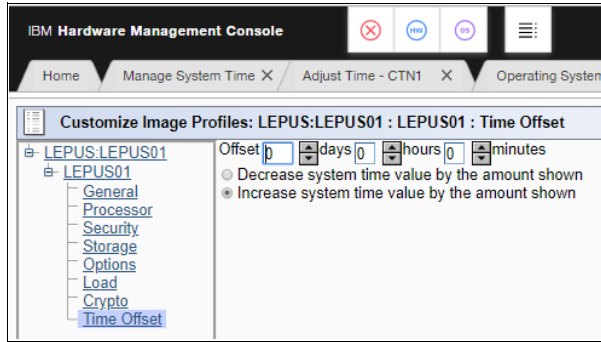


Figure 3-46 Customize Image Profiles - Time offset

The settings available on the Time Offset window have the following meanings:

► **Offset**

Type or spin to the number of days, hours, and minutes to be set as the offset from the time of day supplied by its time source. The offset can be in the following range:

- 0 to 999 days
- 0 to 23 hours
- 0, 15, 30, or 45 minutes

► **Decrease system time value by the amount shown.**

Select this choice to set the logical partition's clock back from the time of day supplied by its time source by the number of days, hours, and minutes in the offset. Use this setting to provide a local time zone west of UTC.

► **Increase system time value by the amount shown.**

Select this choice to set the logical partition's clock ahead of the time of day supplied by its time source by the number of days, hours, and minutes in the offset. Use this setting to provide a local time zone east of UTC or a date and time in the future.

### 3.4.5 Parallel Sysplex and multiple time zones

Some Parallel Sysplex installations have a requirement where different z/OS sysplex members are IPLed in different time zones. This is accomplished by:

- Having all sysplex members on a common time source with STPMODE YES
- Not using STP to obtain time zone information, by specifying STPZONE NO and TIMEZONE d.hh.mm.ss.
- Alternatively, one group of sysplex members in a common time zone could use STP as the time zone source, using STPZONE YES, while other sysplex members do not.

This section discusses how to create and manage the multiple time zone environment.

#### Create the multiple time zone environment

The example in this section reflects a consolidation of multiple sysplexes into a single sysplex. This in itself is not remarkable. What is unusual is the fact that the sysplex is supporting applications intended to operate in images that support different time zones.

The following issues have to be addressed when creating the environment:

- IBM software considerations
- Vendor software considerations
- Staggered time zone setup considerations

### ***IBM software considerations***

Producing a sysplex with staggered time zone offsets introduces many peculiarities into the environment that not every installation or department is happy with. However, there might be business reasons why this configuration is favored, such as sysplex consolidation efforts or business consolidations. Some, but not all, of the peculiarities are:

- ▶ **SMF considerations**

SMF records contain local time. If SMF records from two systems are used for accounting purposes, the data is not accurate. For example, assume that a job is submitted on logical partition 1A, getting a job reader start time in one time zone, and the job executes on another image, getting job start and end times in the other time zone.

- ▶ **RMF considerations**

RMF displays and reports are suspect for the same reasons cited under SMF.

- ▶ **OPERLOG considerations**

Although the records are ordered by UTC (GMT) time stamps, they are displayed with local time offsets. For example, in a merged OPERLOG, messages from the two systems described appear to be out of sequence.

- ▶ **JES2 time-sensitive command considerations**

Some JES2 commands, such as \$TA and \$PJQ, have time-sensitive parameters that would cause actions to be taken at different times. For example, the commands would execute N hours earlier on one system than the other.

The system will function properly, but the external view will complicate the role of some people in the enterprise.

### ***Vendor software considerations***

Since a sysplex allows work to be routed to any machine in the configuration, programs asking for local time can get a variety of answers depending on the server where the work was routed. To help in this situation there is software available that will provide an altered date and time to users, user applications, subsystems software, CICS® regions, and even specific transactions or terminal IDs.

### ***Configuration considerations***

To set up the sysplex with different time zones, there are considerations for z/OS and middleware. In-depth details are beyond the scope of this book, but some examples are mentioned here:

- ▶ **CICS complex (CICSplex)**

Although originally a full-function CICS ran in a single address space (region) in the z/OS environment, most CICS users are more accustomed to running multiple, interconnected CICS regions. Using the CICS Multi-region Operation (MRO) Intercommunication Facility, users can combine CICS regions into a complex of subsystems.

Using multi-region operations, CICS functions can be separated into individual regions, with the different types of CICS regions being classified as resource managers. With the latest enhancements to MRO, these CICS resource manager regions can reside in one or more z/OS images.

- ▶ **CICSplex® system manager**

If there are servers in different geographical locations, are there connections between those processors, or are they managed as separate entities, each with its own workload? If these separate units exist in the enterprise, it is likely that there is a need to define

multiple CICSplexes, and so manage the enterprise CICS systems as though they belonged to more than one enterprise.

Much of CICSplex SM's activity is time-dependent. For example, it can be specified that a monitor definition or an analysis definition is to be active during a particular time period. CICSplex SM does not require every instance in a single CICSplex to be running in the same time zone, and so must be able to accommodate any time zone differences between entities. Therefore:

- ▶ Whenever a time period definition (using the CICSplex SM PERIODEF view) is created, a time zone must be specified in the definition. For example, a time period definition called MORNING can be created for the hours 08:00 through 11:59 Eastern standard time.
- ▶ A time zone must be specified for each CICSplex SM address space (CMAS) in its data repository initialization job, EYU9XDUT. A permanent change to the CICSplex SM address space time zone value can be made, even while the CMAS is running, via the CICSplex SM user interface.
- ▶ A time zone must be established for each managed CICS system. When a CICS system is defined in a CICSplex SM, the time zone in which the system is running can be specified. Alternatively, if no time zone is specified in the CICS system definition, the CICS system is assumed to be running in the time zone that is the default for the CICSplex SM address space to which it connects. Allow the time zone of a managed CICS system to default to that of its CICSplex SM address space. The time zone of a managed CICS system can be altered subsequently while the CICS system is running. Any change made in this way lasts for the lifetime of the CICS system, or until it is next changed, whichever is sooner.
- ▶ A time zone must be specified for every CICSplex when it is first defined. This time zone is used by the CICSplex SM monitor function to determine the actual time at which the monitor interval for the CICSplex expires. The CICSplex time zone can be altered via the CICSplex SM user interface.

Time zones are specified using single-character codes in the range B through Z. For example, code S represents Mountain Standard Time, code T represents Central Standard

Time, and code C represents Eastern Europe Time. A complete list of the codes can be found in CICSplex SM Setup and Administration - Volume 1. CICSplex SM allows offsets (known as time zone adjustments) in the range 0 through 59 minutes to be specified to accommodate regions that are not running in the standard time zones. Also, daylight saving time can be specified.

Because multiple CICSplex SM entities require a time zone to be specified, there is obvious potential for conflicting time zones to be specified. For example, it is possible that a CMAS and a CICS instance in the same CICSplex could be in different time zones. CICSplex SM always honors the time zone of the CICS-managed application system. Suppose the following conditions exist:

- ▶ The time period definition time zone is S.
- ▶ The CICSplex SM address space time zone is B.
- ▶ The CICS-managed application system time zone is C.

Time zone C is used by the CICS-managed application system, and the CMAS makes any necessary adjustments between time zones B, C, and S to ensure that the time zone is honored.

## Manage the multiple time zone environment

**Note:** In this section, it might be preferable to use the new z/OS SET TIMEZONE command instead of the z/OS SET CLOCK command. Refer to "SET TIMEZONE" on page 97 for more information.

Since each sysplex member is in a different time zone, it is necessary to stagger Daylight Saving Time changes. Daylight saving time changes could be accomplished individually in each sysplex member using the SET CLOCK command from the z/OS console.

The cities used in the examples that follow include Sydney in the southern hemisphere (during summertime) and London, Atlanta, and Los Angeles in the northern hemisphere (using standard time).

**Operational considerations**

In Figure 3-47, each of the three cities accurately reflects the number of hours from the Greenwich meridian when Daylight Saving Time would be in effect in the southern hemisphere. That is, Sydney's TIMEZONE parameter is shown while Daylight Saving Time is in effect. London and Atlanta TIMEZONE parameters are shown with standard time in effect. Each hemisphere is changing the offset in different directions. When the southern hemisphere is changing from Daylight Saving Time to standard time (back), the northern hemisphere is changing from standard time to Daylight Saving Time (a forward change).

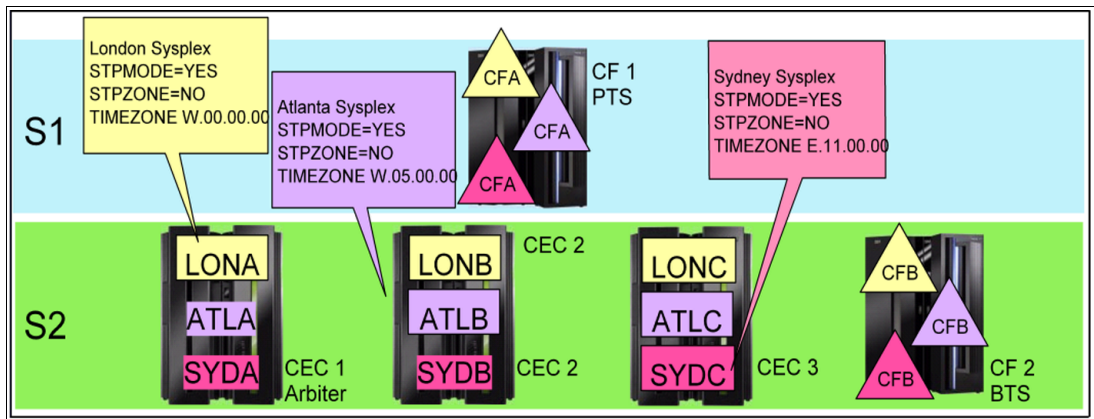


Figure 3-47 Staggered offset changes: one sysplex supporting multiple time zones

When planning for Daylight Saving Time changes in different countries, remember that not all countries agree on which dates the switch will occur. In fact, each country's government might change the Daylight Saving Time switch dates from year to year. A country that has multiple time zones might have a different change date in each zone. In our example, however, assume that all countries are going either to Daylight Saving Time (London and Atlanta) or to standard time (Sydney) on the same arbitrary day.

Just to keep things interesting, assume that this sysplex is physically located in Los Angeles, California. Any z/OS images representing Pacific Standard Time in the USA are not depicted in Figure 3-47. However, assume that those images contain STPZONE YES and that STP is being used to obtain Daylight Saving Time offsets.

**Note:** The world time zones shown in Table 3-5 represent only the local standard time offset from the Greenwich meridian and do not include Daylight Saving Time offsets in either hemisphere.

Table 3-5 Local standard time offsets

Sydney	London	Atlanta	Los Angeles
02:00 AM	04:00 PM	11:00 AM	08:00 AM
12:00 PM	02:00 AM	09:00 PM	06:00 PM

Sydney	London	Atlanta	Los Angeles
05:00 PM	07:00 AM	02:00 AM	11:00 PM
08:00 PM	10:00 AM	05:00 AM	02:00 AM

Which sysplex members (cities) change their offsets first? The answer is Sydney: SYDA, SYDB, and SYDC. Since a new day will dawn in Sydney prior to in London and Atlanta, Sydney changes first. The operators in Los Angeles must be at Sydney's z/OS console to enter the SET TIMEZONE command at the correct instant.

What is the local time in Los Angeles when Sydney's DST offset changes? Remember that Sydney is on DST in the southern hemisphere. Assuming that Sydney wants the offset changed at 02:00 a.m. Sydney time, the Los Angeles operator must enter the command at 08:00 a.m. After the offset change, Sydney's local time will be changed from 02:00 a.m. to 01:00 a.m. (back one hour).

Several hours later, the same thing must be done for the London sysplex members. Again, assuming that London will change the offset at 02:00 a.m. London time, the Los Angeles operator must enter the command at 06:00 p.m. The only difference is that the northern hemisphere is changing the offset in the opposite direction from Sydney. London is going from standard time to Daylight Saving Time (forward).

Sometime later, it is Atlanta's turn. Atlanta is also going from standard time to Daylight Saving Time. The Los Angeles operator must enter the SET CLOCK command on Atlanta's z/OS console at 11:00 p.m.

Finally, the offset for Los Angeles can be changed through the Server Time Protocol, Adjust time zone offset panel. These images have STPZONE YES in their CLOCKxx members.

**Note:** Each customer situation is different. Certain ones are much more complicated, and others less so. Planning and excellent communications with the user community are essential when Daylight Saving Time schedules is a prime concern.

For some countries, time zones are not aligned on hourly divisions. For example, the time zone might be E.07.30.00 or W.08.45.00.

After a SET CLOCK command has been entered, change the TIMEZONE parameter in the CLOCKxx member in each affected image to reflect the new DST offset. The next IPL uses the new TIMEZONE value, and z/OS recognizes the correct local time for that image.



# Managing STP configurations

In this chapter we describe standard tasks. With the workflow available since HMC 2.14.1 Driver 36 most of the tasks are guided and pretty straight forward. The “Managed System Time” option is also no longer available on the Support Element (SE).

In this chapter we describe the following configuration changes:

- ▶ Changing the CTN ID
- ▶ Single Server or dual Server auto resume after Power-on Reset
- ▶ Changing the current time server
- ▶ Changing server roles
- ▶ Deconfiguring a CTN
- ▶ Adding a server to an existing CTN and reassigning server roles
- ▶ Splitting a CTN
- ▶ Merging two CTNs

## 4.1 Test environment overview

Figure 4-1 depicts the test environment used for the scenarios described in this chapter.

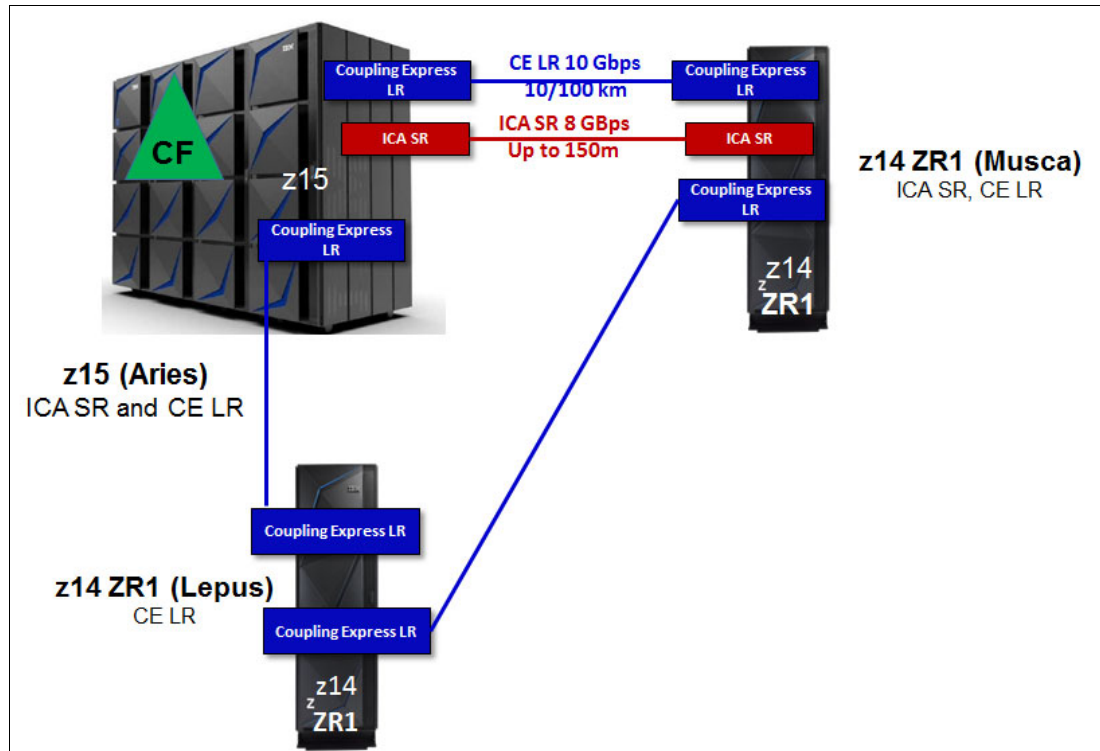


Figure 4-1 Test environment used for this chapter

The three CPCs available are:

- ▶ ARIES (z15)
- ▶ MUSCA (z14 ZR1)
- ▶ LEPUS (z14 ZR1)

## 4.2 Managing CTN configuration - overview

All the functions described in this chapter are performed using the “Manage System Time” HMC task. The task is invoked from the IBM Z HMC after selecting a CPC and clicking on “Configuration -> Manage System Time”.



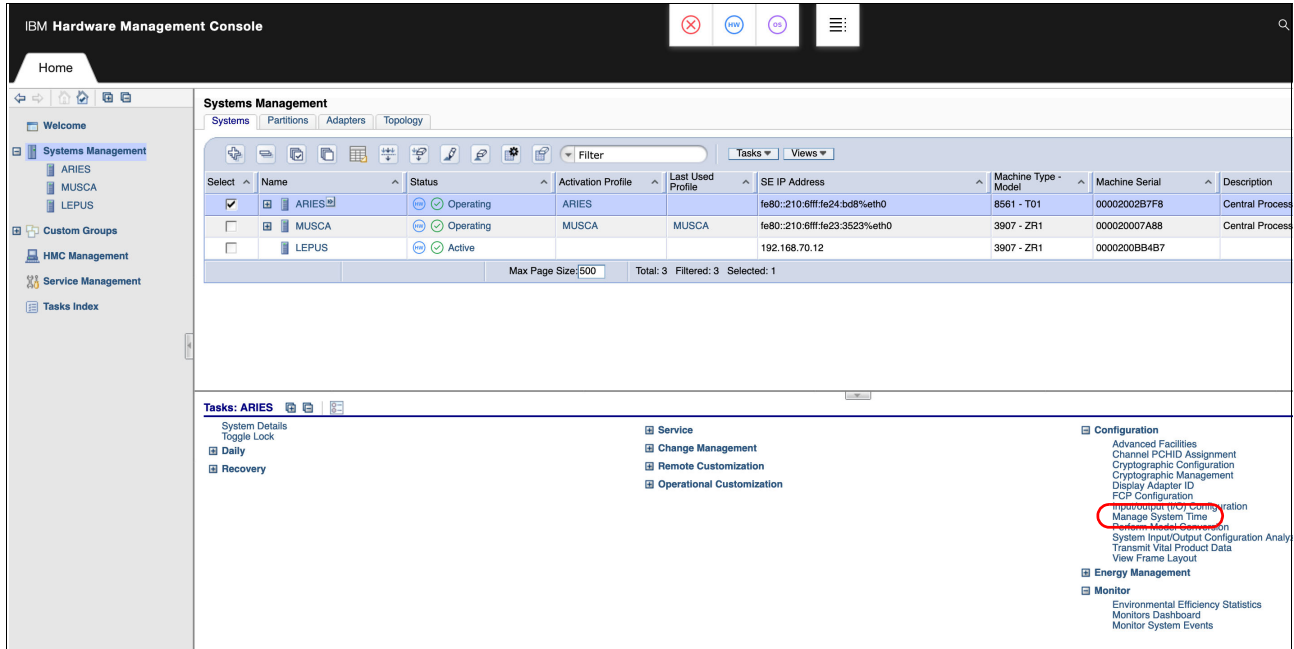


Figure 4-2 The Manage System Time Task - HMC

### 4.2.1 Changing the name of a CTN (CTN ID)

The name of a CTN can be changed in the “Manage System Time” task. The change is non-disruptive to the systems operation. On the upper left corner of the Manage System Time tab you can see the name of the currently select CTN. By clicking on the pen-symbol you can modify the name, as shown in Figure 4-3 on page 107.

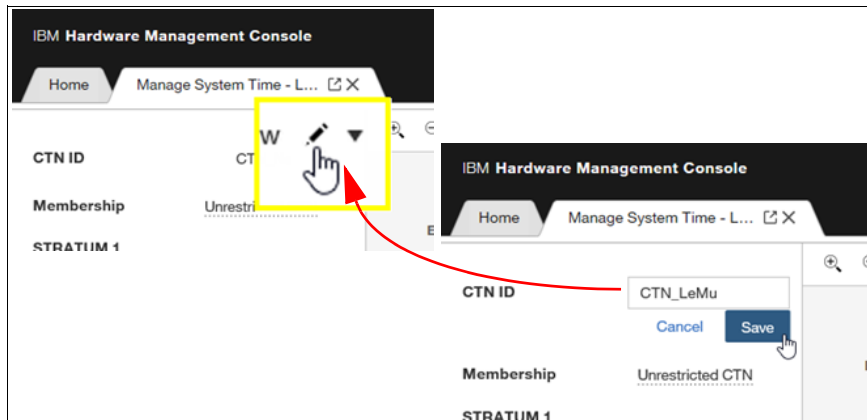


Figure 4-3 Changing the CTN ID (name)

A confirmation dialog is displayed, as shown in Figure 4-4 on page 108.

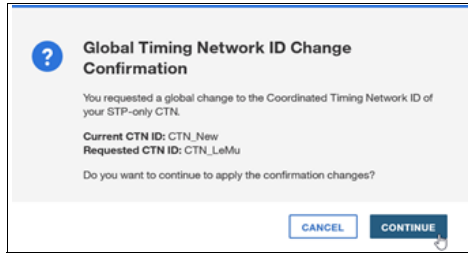


Figure 4-4 Confirmation dialog

Verify the changes and click “Continue”. The rename operation success message is displayed, as in Figure 4-5.

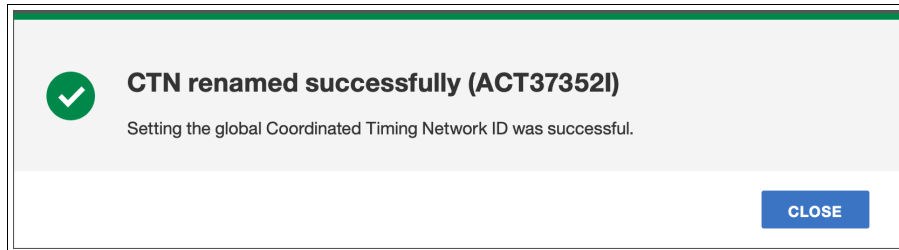


Figure 4-5 CTN ID change success

### 4.2.2 Single-server or dual-server auto resume after power-on reset

If the CTN consists of one or two servers it can be marked as “restricted”. This is shown in Figure 4-6.

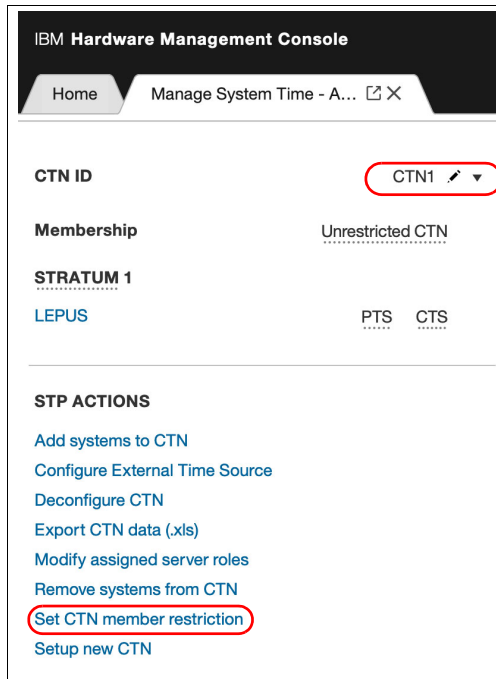


Figure 4-6 Restricting a CTN

A dialog is shown as in Figure 4-7 on page 109. Select “Allow only the servers that are specified below to be members of the CTN” and click on “Apply”.

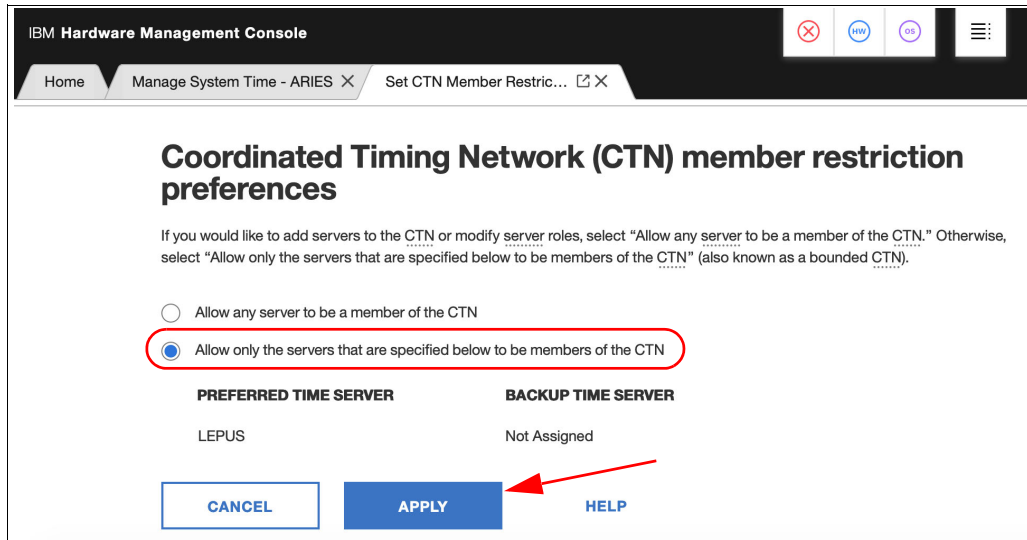


Figure 4-7 Applying CTN restriction

#### Important:

With this option set, no additional server can participate in this CTN.

During a recovery scenario, users should not attempt to reconfigure a restricted CTN UNLESS all servers are available in the CTN.

This saves the CTN configuration for handling a Power-on Reset (POR) of all servers in the CTN. Because the CTN settings and roles are saved, the CTN is automatically recovered after the POR.

If the CTN is marked as unrestricted, the CTN has to be recovered manually after the POR (configuration from scratch). In this scenario, a single server CTN is recovered. A dual server CTN can be recovered in a similar way.

After a POR the CPCs defined in a unrestricted CTN are all stratum 0. To recover, the “Setup new CTN” dialog from the “Manage System Time” task can be used. The configuration steps are described in 3.2, “Configuring the Coordinated Timing Network” on page 71.

### 4.2.3 Changing the Current Time Server

Changing the Current Time Server from the Preferred or Backup Time Server is yet another non-disruptive modification. To perform this change, select the CTN ID you want to modify then click on “Modify assigned server roles”, as in Figure 4-8 on page 110:

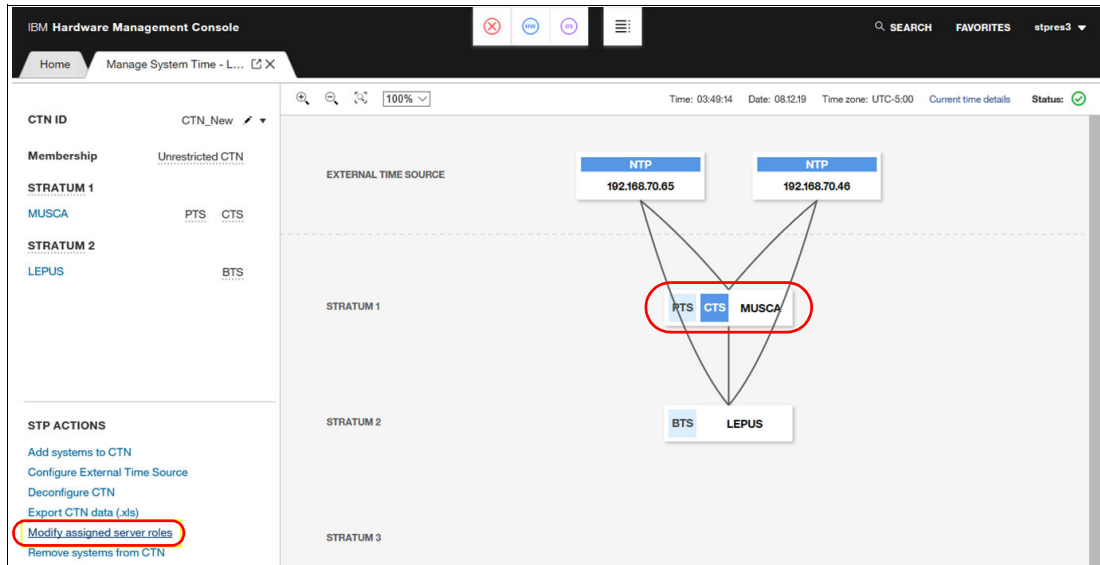


Figure 4-8 Modify assigned server roles

Figure 4-9 shows the “Modify Assigned Server Roles” dialog:

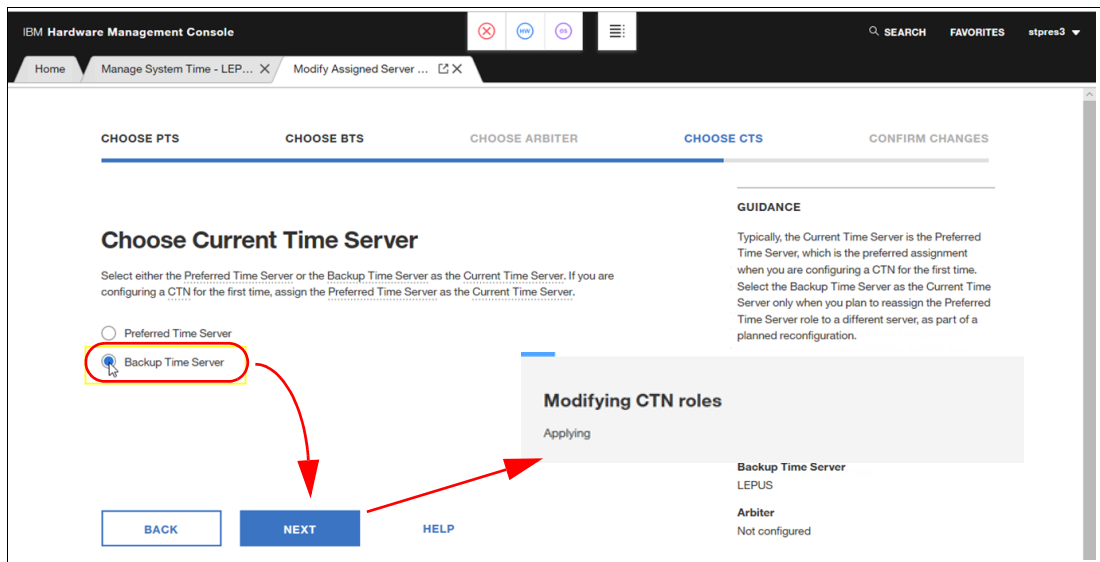


Figure 4-9 Modify server roles

Confirm your selection by clicking on NEXT. The changes will then be performed. A confirmation pop-up window is displayed during this period.

When the modification is completed the new configuration is displayed. Note that the CPC named “LEPUS” is now the CTS, as shown in Figure 4-10 on page 111.

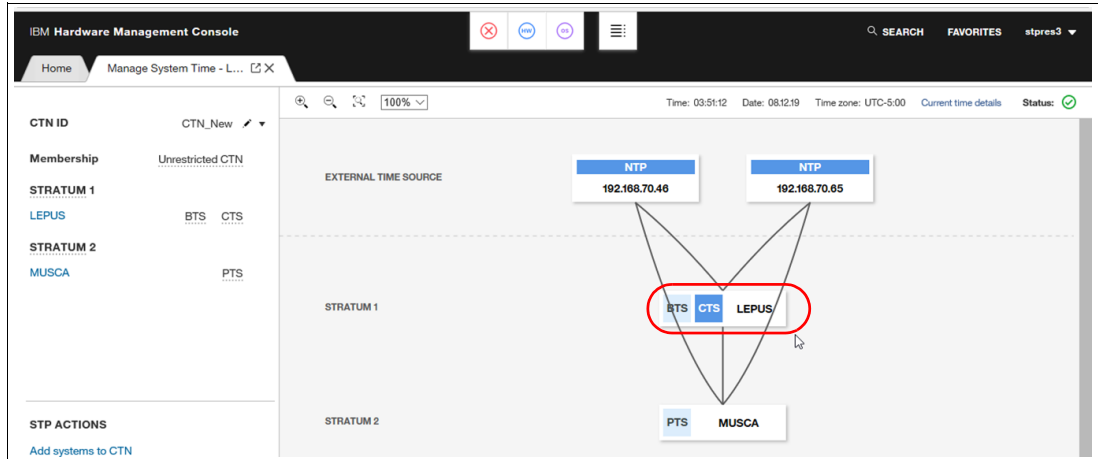


Figure 4-10 CTS role assigned to BTS

### 4.2.4 Changing the server roles

Changing server roles may be necessary if you plan maintenance on a system with a role in a CTN. This HMC task is also non-disruptive. In this example we swap the role of PTS and BTS and we want the new PTS to be the CTS.

We select the CTN ID we want to modify and click on “Modify assigned server roles” under “STP ACTIONS”, as shown in Figure 4-11 on page 111.

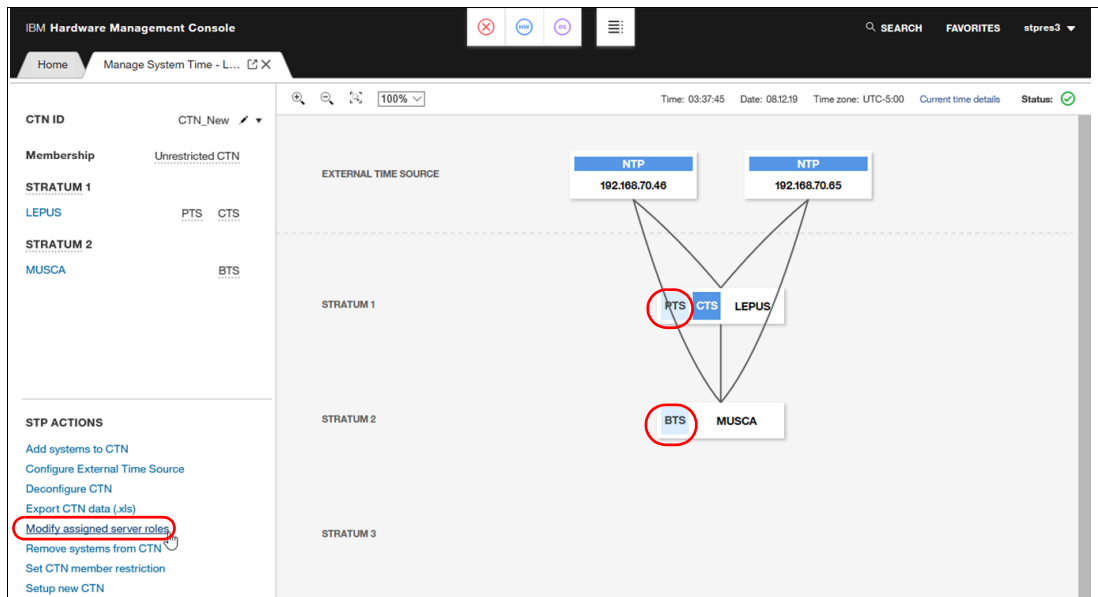


Figure 4-11 Modifying server roles

We select the server which should become the new Preferred Time Server PTS. The role should switch from LEPUS to MUSCA so we click on MUSCA and confirm by clicking on the NEXT button, as shown in Figure 4-12 on page 112.

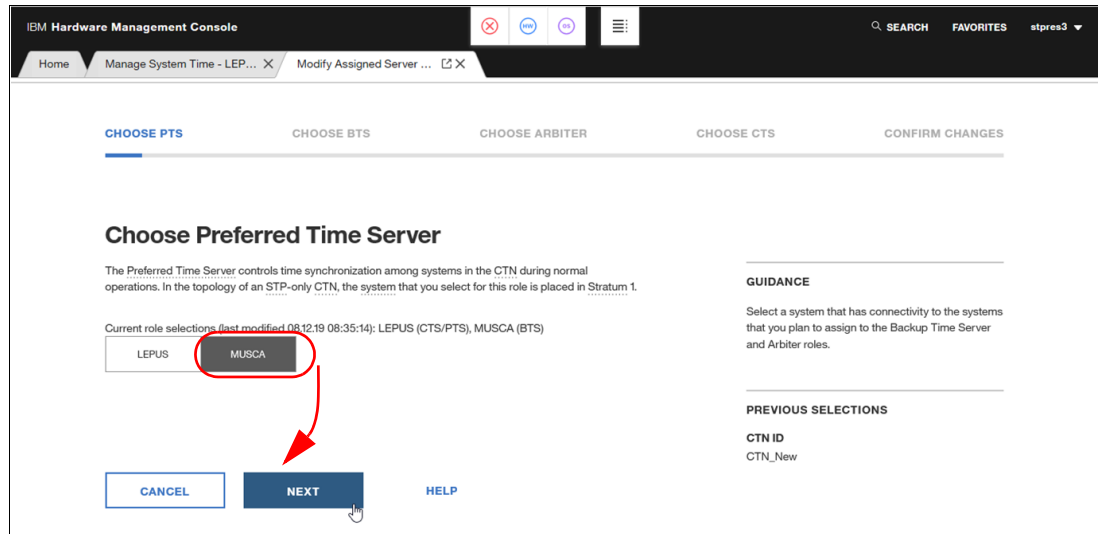


Figure 4-12 Selecting the new PTS

On the next window we choose the server that has to perform as the Backup Time Server (optional). Since MUSCA is already selected as PTS, only LEPUS can be chosen in this example.

Note that to perform maintenance on a server this must not play a role in the CTN, thus, to run without a BTS we can check the “Do not configure a Backup Time Server” checkbox. However, in this example we select LEPUS as BTS and then click NEXT, as shown in Figure 4-13 (we just swap the server roles).

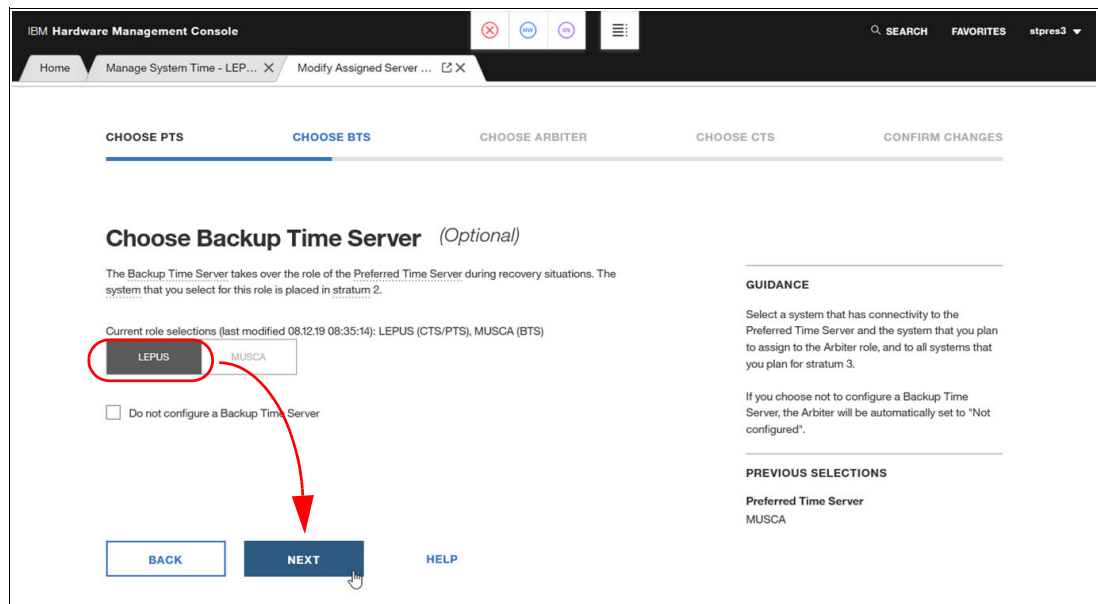


Figure 4-13 Selecting the BTS

Then we define the CTS. We choose “MUSCA” (the new PTS) as our CTS, thus we check “Preferred Time Server” and click NEXT, as shown in Figure 4-14 on page 113.

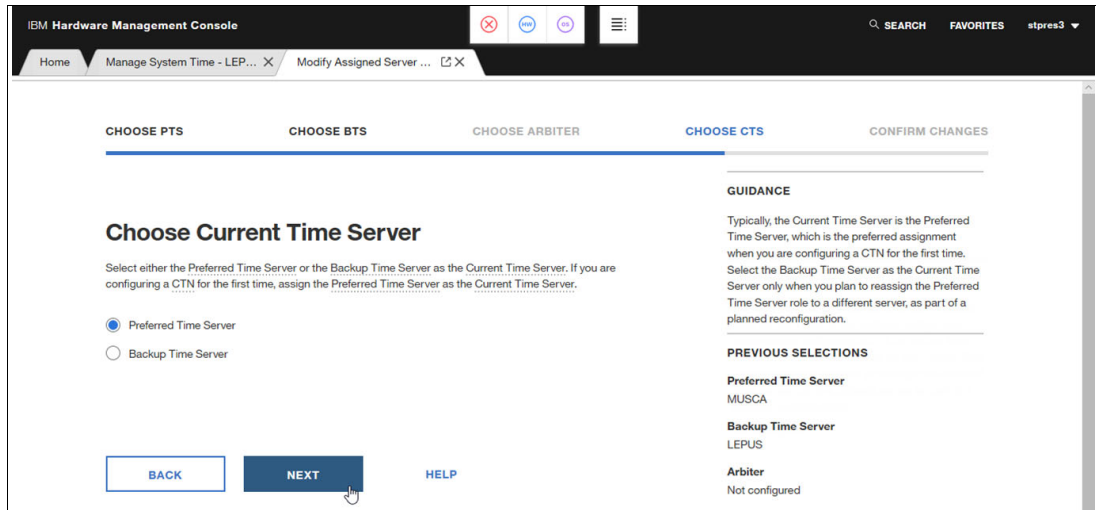


Figure 4-14 Selecting the CTS

The next window shows how the new configuration will look like. If the proposed configuration matches the expectation, we confirm by clicking on “NEXT”. After the changes were successfully applied, the message ACT39295I is displayed and the new roles are active.

## 4.2.5 Deconfiguring a CTN

If a CTN is no longer needed, it can be deconfigured. All roles from the servers in the CTN are removed. In this example a single server CTN will be deconfigured, leaving LEPUS in unsynchronized state (STP stratum 0). We select the CTN ID to be deconfigured, then we choose “Deconfigure CTN” under STP ACTIONS, as shown in Figure 4-15 on page 113.

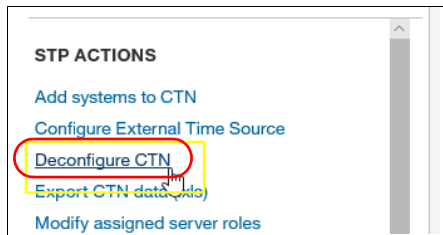


Figure 4-15 Initiating CTN deconfiguration

**Note:** If the CTN is restricted, it must be switched to unrestricted before deconfiguring the CTN. Also server role changes can be performed only if the CTN is unrestricted.

The next pop-up window appears stating that the “Deconfigure CTN” task is disruptive. Enter the password of the HMC user and click on APPLY. After the progress pop-up disappeared, message ACT39285I confirms the successful execution of the “Deconfigure CTN” task. LEPUS is now unsynchronized (STP Stratum 0).

**Attention:** Any sysplex operations will be suspended and any z/OS systems running in STP mode will go into a *disabled wait* state if CTN deconfiguration is performed. This is a HIGHLY DISRUPTIVE action.

## 4.3 CTN reconfigurations

In this section we provide examples of CTN reconfiguration operations:

- ▶ Adding a server to a CTN and re-assigning server roles
- ▶ Splitting a CTN
- ▶ Merging two CTNs

### 4.3.1 Adding a server to an existing CTN and re-assigning roles

In this sample we add the System MUSCA to the existing CTN *CTN\_New*. MUSCA is part of the inactive CTN *CTN3* and is therefore ready to be added to *CTN\_New*.

**Note:** A server must be part of a CTN (active or inactive) to be added to an existing CTN.

We select the CTN ID *CTN\_New* to be the receiving CTN. Then we click on “Add systems to CTN” under STP ACTIONS. A window with all available Systems and their corresponding CTNs is shown. In our case only MUSCA CPC is available. We select MUSCA by clicking on its name and then continue (NEXT), as shown in Figure 4-16.

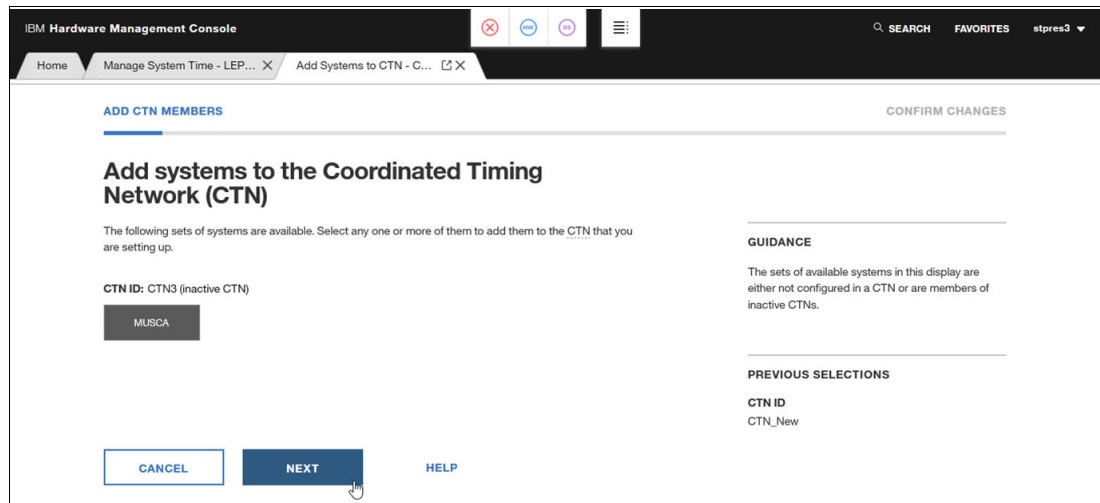


Figure 4-16 Add a server to an existing CTN - selecting the server

A window which shows the target configuration will then be shown. We accept this by clicking APPLY. The next window provides an overview of the system(s) that will join.

The message shown in Figure 4-17 on page 115 informs that the CTN ID for the server selected to join the existing CTN will change to match the target CTN ID.



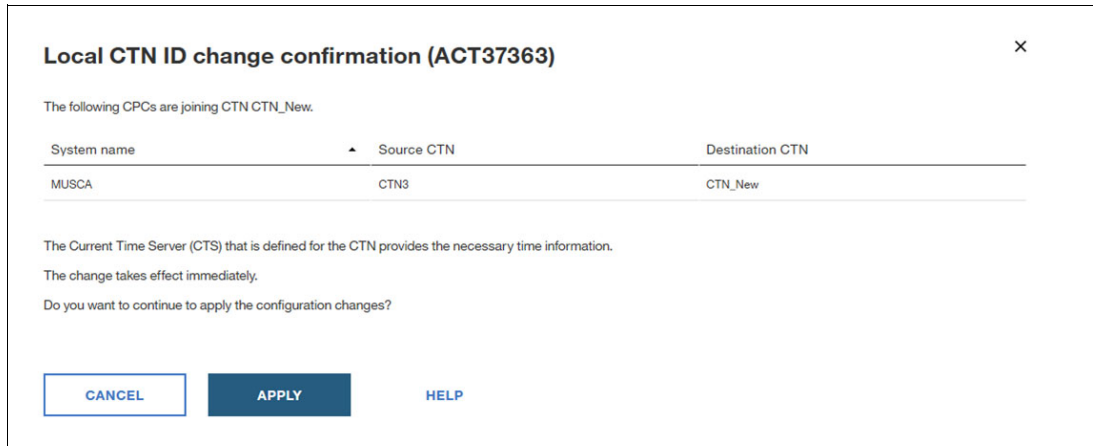


Figure 4-17 CTN ID change confirmation

We click on APPLY and the changes are executed. The new configuration will be active immediately. A pop-up window confirms the activation and message ACT39275I “Systems added to CTN successfully” is displayed.

### 4.3.2 Splitting a CTN

**Important:** Split and merge operations are only available for z14 (Driver 36) and newer servers.

Splitting a CTN into two distinct CTNs can be performed using the “Split to new CTN” task under STP ACTIONS/Advanced Actions menu. The systems to be split off the existing CTN must NOT have a role assigned. All these roles must be removed by executing the “Modify assigned server roles” task, which was already described earlier in this chapter.

In this example we split MUSCA and LEPUS off the CTN with the ID “CTLSOMRS” into the newly created CTN ID “CTN\_Split”. LEPUS will also become the PTS and CTS of “CTN\_Split”. ARIES will remain PTS and CTS of the existing CTN “CTLSOMRS”.

The current configuration of CTLSOMRS is shown in Figure 4-18 on page 115:

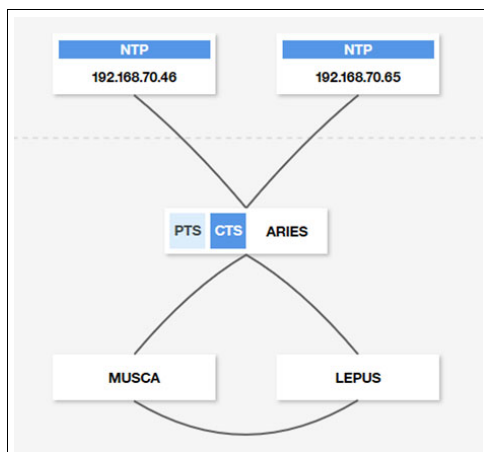


Figure 4-18 Starting point of “Split CTN” scenario

**Attention:** It is very important that you do not split a running sysplex, therefore a warning message appears as in Figure 4-19. Make sure you have properly reconfigured your sysplex workloads to avoid splitting sysplexes!

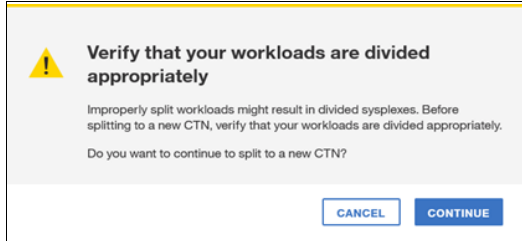


Figure 4-19 Warning message

If everything has been double-checked, click CONTINUE. On the next Window we specify the name of the (new) CTN ID to create. In our case the new CTN ID is named "CTN\_Split", as shown in Figure 4-20.

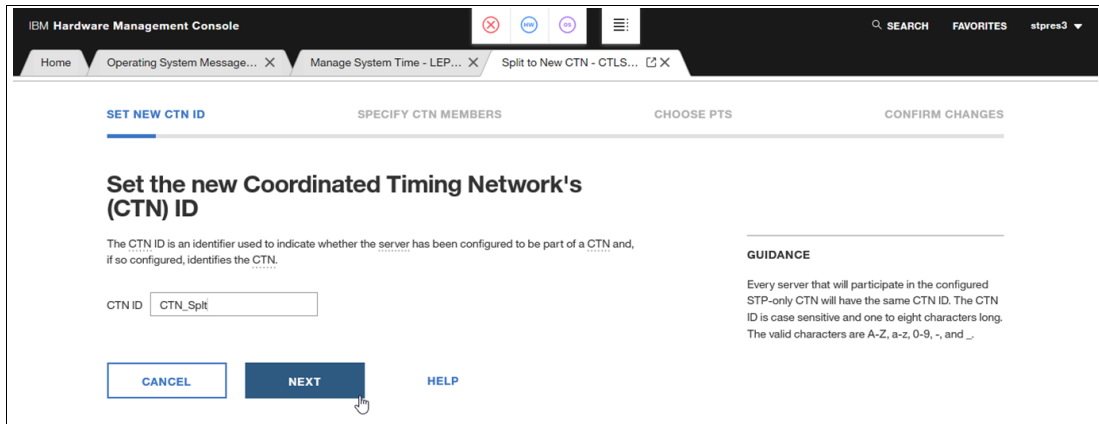


Figure 4-20 New CTN ID

The next window lists all members that are eligible to be split into the new CTN. If a system does not appear, it may still have a role assigned. Remove the role from the server you desire and start over.

We choose to split MUSCA and LEPUS from CTLSOMRS, as shown in Figure 4-21 on page 117.

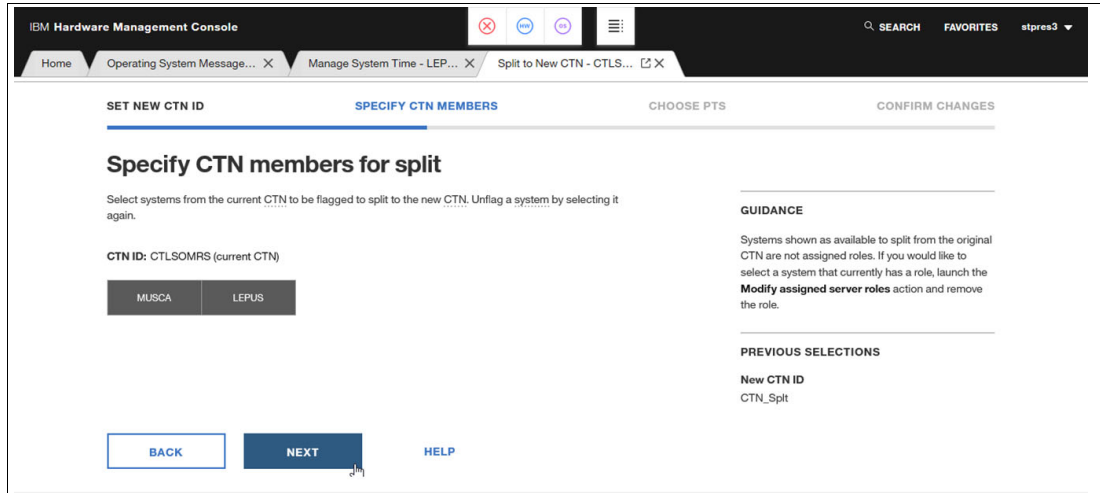


Figure 4-21 Servers to be split into the new CTN

**Attention:** If there are systems detected running in a same named sysplex which would be separated by the split, a warning message appears (see Figure 4-22 on page 117). Take this very seriously and double check if you really want to split!

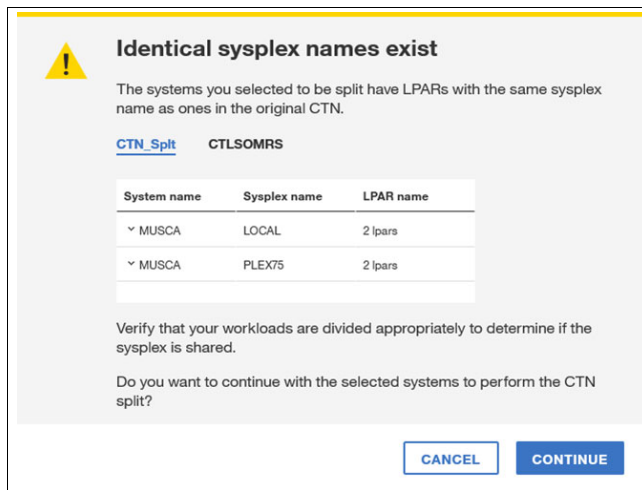


Figure 4-22 Final warning message before splitting the CTN

If everything has been checked, confirm this by clicking on CONTINUE.

On the next window we can choose which system should be the Preferred Time Server. In our case we select LEPUS, as shown in Figure 4-23 on page 118:

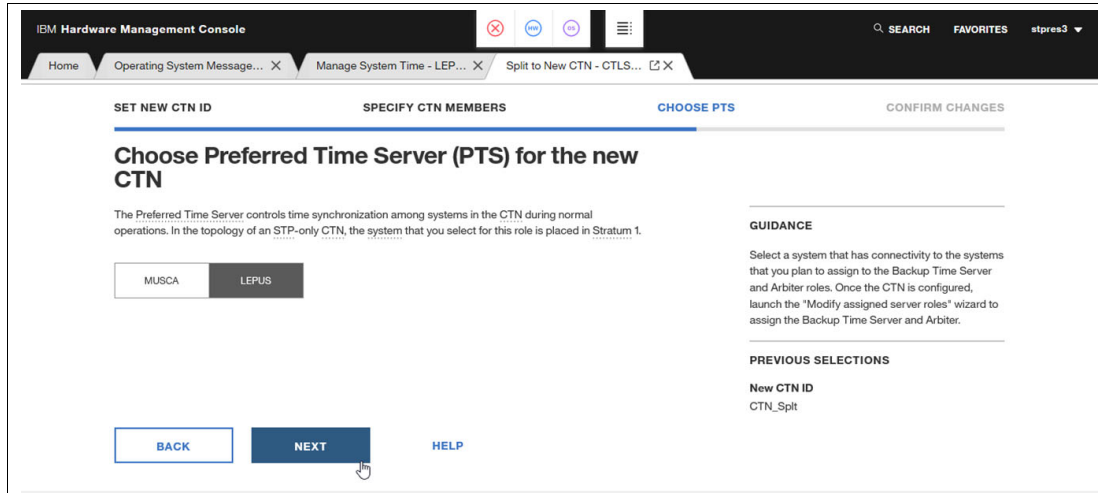


Figure 4-23 Server roles for the new CTN

On the next window we verify both CTNs. If both CTNs match expectations, we confirm by clicking on APPLY and the changes will be performed immediately. After a short period, the successful execution is confirmed with message ACT39329I “The CTN split task completed successfully”. This concludes the split and you now have two valid CTNs.

### 4.3.3 Merging Two CTNs

To merge two CTNs into one, we select the “Join existing CTN” task under STP ACTIONS/Advanced Actions. We make sure that we select first the CTN ID to be joined into an existing CTN. The CTN selected first to be joined now will no longer exist after the successful execution of the merge CTN operation.

**Important:** This procedure could take a considerable period of time if the CST of the two systems are not very close to each other.

In this example we merge *CTN\_Join* into *CTLSOMRS*. That means the systems that are defined in *CTN\_Join* will become part of the CTN with the ID “CTLSOMRS”, while “*CTN\_Join*” will no longer exist.

CTLSOMRS is a single server CTN with the member ARIES (PTS/CTS). *CTN\_Join* is a two server CTN with the members, MUSCA and LEPUS.

We first select *CTN\_join* as the current CTN ID. Then we start the merge CTN operation by clicking on “Join existing CTN”. A new tab opens in which we select the CTN to join, as shown in Figure 4-24.

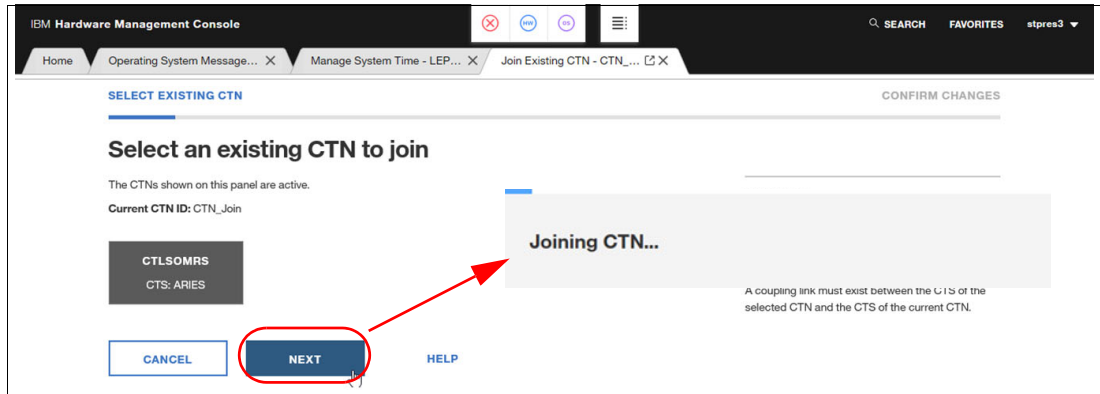


Figure 4-24 Initiating the merge operation

We select CTLSOMRS and click NEXT. Note that the Current CTN ID is *CTN\_Join*. On the next window we verify the result. If they meet expectations, we click on APPLY. The join process is started and a progress pop-up is displayed (see Figure 4-24 on page 119).

When the join process is finished another pop-up window appears which confirms the successful execution (see Figure 4-25). CTLSOMRS is now the current CTN ID and *CTN\_Join* does no longer exist.

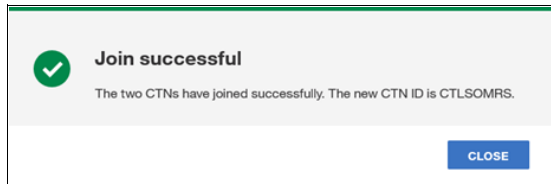
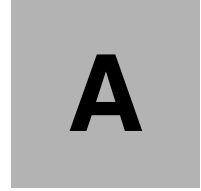


Figure 4-25 Successful join message





# IEEE 1588 Precision Time Protocol (PTP) Introduction and Overview

This appendix provides an introduction and overview of Precision Time Protocol (PTP). The chapter will provide a background on PTP technology and the relevant standards, summarize the regulatory requirements in industry that led to IBM moving forward with PTP, offer a brief comparison with Network Time Protocol (NTP), and discuss things clients can do to prepare their data center for transitioning to using PTP.

## Introduction

As announced in the September 2019 IBM z15 product announcement's statement of general direction, IBM has introduced IEEE 1588 Precision Time Protocol (PTP) as an external time source for IBM Z Server Time Protocol (STP) running in an IBM Z Coordinated Timing Network (CTN). The initial implementation provides PTP connectivity via the IBM Z Support Element (SE) and management interface of the External Time Source via the IBM Z Hardware Management Console (HMC) appliance. For the current implementation (as announced on April 14, 2020, there is no change to the use of STP CTNs for time coordination, other than the potential to use a PTP-based external time source. Future implementation<sup>1</sup> is planned to include full connectivity of an external PTP time source directly to the IBM Z CPC, and re-introduction of the concept of a mixed CTN, with support for traditional STP and native PTP implementations. Beyond that, the goal is to enhance the role of IBM Z machines in a PTP environment that addresses the many governmental regulations and security concerns that clients are facing.

## Background on PTP: History

The Network Time Protocol (NTP) was introduced in 1985. During the 1990s, public discussion started on standardizing the technology and techniques for synchronizing clocks in devices not associated with NTP, examples of such devices being those typically being used in measurement and control applications. At the time, these discussions were taking place between engineers developing systems associated with industrial automation and the IEEE 1451 family of standards. There was enough interest in starting a new, separate standardization activity on clock synchronization to warrant formation of a committee dedicated to this purpose. This committee held its initial meeting in April 2001 and was sponsored by the Institute of Electrical and Electronics Engineers (IEEE) Technical Committee on Sensor Technology of the Instrumentation and Measurement Society (IEEE IMS). A formal application was submitted and approved by the IEEE on July 18, 2001. The committee produced a draft of the standard, which went through the IEEE standards approval process in Spring 2002. The IEEE Standards Board Review Committee completed its approval process in September 2002, and the IEEE 1588-2002 standard was published in November 2002. Much of the standard was based on early prototypes built at Agilent Technologies under the leadership of John Eidson between 1990 and 1998. As conceived, PTP is capable of near nanosecond precision while taking advantage of a network infrastructure that is similar to what NTP uses.

The IEEE 1588-2002 standard with additional features and improved performance was later revised and became the IEEE 1588-2008 standard, more commonly known as IEEE 1588 Version 2 (V2). Versions 1 and 2 are not compatible, i.e., it is not possible to have a mix of V1 and V2 devices in the same timing network.

In June 2013, and IEEE Standards Association Project Authorization Request (PAR) was approved to revise IEEE 1588-2008. During the time from 2008-2013, additional industries with use cases beyond those associated with the early PTP implementations emerged. The stated purpose of the revision was to reflect the common needs of the various industries using the standard to meet the requirements of more security in a precision clock synchronization protocol, and that is also suitable for higher speed applications in the sub-nanosecond range. It also was decided that it needed to support both IPV4 and IPV6.

---

<sup>1</sup> Statements by IBM regarding its plans, directions, and intent are subject to change or withdrawal without notice at the sole discretion of IBM. The development, release, and timing of any future features or functionality described for IBM products remain at the sole discretion of IBM.



The revision was approved by the IEEE Standards Association in November 2019 and will officially be known as IEEE V2.1, or also as IEEE 1588-2020. It is scheduled to be formally published in early 2020.

## Background on PTP: Technology overview

The IEEE 1588 standard provides a scope for the standard. This scope summarizes the standard, its goals and technology quite well. The latest scope (from IEEE V2.1) is summarized in the following section. Please keep in mind that this is the overall scope, it is not IBM Z (or any other platform) specific. The information is presented here to help you understand the basic idea behind PTP.

### IEEE 1588 PTP scope

The Precision Time Protocol (PTP) standard enables accurate and precise synchronization of the real-time clocks of devices in networked distributed systems. The protocol is applicable to systems where devices communicate via networks, including Ethernet. The standard allows multicast communication, unicast communication or both. The standard specifies requirements for mapping the protocol to specific network implementations and defines such mappings, including User Datagram Protocol (UDP)/Internet Protocol (IP versions 4 and 6), and layer-2 IEEE 802.3 Ethernet.

PTP enables heterogeneous systems that include clocks of various inherent precision, resolution, and stability to synchronize to a grandmaster clock. The protocol supports synchronization in the sub-microsecond range with minimal network bandwidth and local clock computing resources. The protocol enhances support for synchronization to better than 1 nanosecond. The protocol specifies how corrections for path asymmetry are made, if the asymmetry values are known. The grandmaster can be synchronized to a source of time external to the system, if time traceable to international standards or other source of time is required. The protocol provides information for devices to compute Coordinated Universal Time (UTC) from the protocol distributed time, if the grandmaster is traceable to international standards and is able to access pending leap second changes. Options are also provided to allow end devices to compute other time scales from the protocol distributed time scale.

PTP defines timing domains in which system timing is consistent. The protocol establishes the timing topology. The default behavior of the protocol allows simple systems to be installed and operated without requiring the administrative attention of users to determine the system timing topology.

The PTP standard defines all needed data types, message formats, required computations, internal states, the behavior of devices with respect to transmitting, receiving, and processing protocol communications. It also provides for the management of protocol artifacts in devices and defines formal mechanisms for message extensions, the requirements for profiles that allow customization for specific application domains, and finally defines conformance requirements.

Optional specifications are provided for protocol security. The PTP standard documents conditions under which V 2.1 is backward compatible with IEEE 1588-2008.

## PTP's underlying technology

This section will introduce the basic terminology and components of PTP, including the various types of clocks, the synchronization methodology, and the basics of the delay mechanisms used in the synchronization process.

At a high level, PTP is conceptually a two-stage process that relies on protocol specific message traffic to perform its synchronization functions. PTP defines a number of protocol messages and classifies each as event messages, or general messages. Management and signaling messages fall into the general message category.

The first stage of the PTP process is the use of an algorithm, known as the Best Master Clock algorithm (BMCA), to organize all clocks in a given PTP domain into a hierarchy. PTP clocks in the domain are self-organized into a hierarchy in which the PTP Grandmaster clock with the highest priority and best quality is at the top of the pyramid, and ordinary clocks are at the bottom of the pyramid. Only certain types of PTP clocks (ordinary clocks and boundary clocks) participate in the BMCA's self-organizing hierarchy. The second stage PTP process consists of the passing of the PTP message traffic to perform time synchronization.

The PTP synchronization methodology is unique, and it is the heart of the IEEE 1588 standard(s). PTP defines a hierarchy based on criteria that describe a combination of a device's timekeeping capability and the traceability of its time source. The master will serve as a reference for one or more ordinary devices. The process of selecting the top master clock (the PTP grandmaster) from the list of participating devices in the hierarchy is defined in PTP's Best Master Clock Algorithm (BMCA). The BMCA is a key part of the "simple, administration free" aspects of PTP's objectives as a protocol. PTP devices will transmit "Announce" messages at configurable intervals. The devices in the PTP domain will process these announce messages according to the BMCA and select the grandmaster. Attributes contained in the "Announce" message that play a role in the BMC selection include:

1. Device's time source (Atomic clock, GPS, free-running oscillator)
2. The device clock ID
3. The device's priority (determined and configurable by administrators).

The IEEE 1588 PTP standard(s) define five device types: ordinary clocks, boundary clocks, end-to-end transparent clocks, peer-to-peer transparent clocks, and management nodes:

- ▶ A PTP *ordinary clock* is a device that either serves time or synchronizes to time and communicates on the network through a single PTP port. An ordinary clock is known as a PTP grandmaster clock if it is serving time to the entire PTP network/domain (making it the ultimate source of time for all other devices in the network/PTP domain). An ordinary clock synchronizes to another clock which is capable of serving time.
- ▶ A PTP *boundary clock* is a multiport device that synchronizes to the reference time on one of its ports and serves time on one or more ports (one port is an ordinary port, driven by the upper level master, while the remainder are master ports). Therefore, boundary clocks can be said to terminate and then start the PTP time distribution. Boundary clock functionality is usually built into PTP network aware components such as switches, bridges, and routers.
- ▶ A PTP *end-to-end transparent clock* is a multiport network device that measures the length of time a PTP message spends within the device as it is routed from the ingress port to an egress port. This measurement eliminates variations in message delays and asymmetry that the device might introduce in the transfer of PTP messages. End-to-end transparent clocks are typically PTP aware network switches.
- ▶ A PTP *peer-to-peer transparent clock* is a multiport device that measures the link delay of each port and adds that information and the device residence time to PTP messages

traversing the device. These clocks are especially suited for networks with redundant paths. In such networks, the PTP message will always contain the actual delay it experiences on the network, regardless of the path taken.

- ▶ Our final component is a *PTP management node*. These are simply a network connected device used to configure and monitor PTP devices.

## The concept of PTP Profiles

One of the goals of the IEEE 1588 PTP standards are to make it possible to set up and run a PTP network with minimal device settings and administrative overhead. The concept of PTP profiles makes this goal possible. PTP profiles allow organizations or industry groups to specify a subset of options, features, and default values for protocol attributes that will meet the performance requirements of applications and eliminate or minimize device configuration settings. Through the use of PTP profiles, organizations can specify selections of attribute values and optional features of the PTP standard that, when using the same transport protocol, work together and achieve a performance that meets the requirements of particular applications.

PTP profiles make PTP better suited for particular applications and/or industry use cases while simultaneously adhering to the more general requirements of the IEEE 1588 PTP standard. Profiles themselves are standardized and defined by a recognized standards development organization (SDO) that has jurisdiction over a particular industry. Examples of such SDOs include the IEC, IEEE, OETF, ANSI, and the ITU.

PTP profiles not only change some aspects of the standard, but they also further extend the standard via mandating some of the optional features in the standard:

1. A PTP profile may define its own BMC algorithm, configuration, monitoring, and management mechanism.
2. A PTP profile may also define use of a specific transport mechanism, use of multi-cast or unicast, and whether end-to-end or peer-to-peer path delay mechanisms will be used.
3. PTP profiles will also define any IEEE 1588 optional components that are required, permitted, or prohibited.
4. PTP profiles may also define completely new transport mechanisms such as optical transport networks (OTNs) and data types.

The IEEE 1588 PTP standards define two default profiles, one for the delay request-response mechanism and the other for the peer delay mechanism. PTP profiles provide tremendous flexibility in "morphing" PTP to the needs of almost any application or use case. This flexibility has proven highly useful to the energy and telecommunications industries. As additional profiles are developed, it will likely prove invaluable to the financial and banking industry as well.

## Regulatory Requirements

This section provides a solid background on time synchronization regulations. This section is important, because understanding this topic, especially the most recent regulatory changes, will make it clear why IBM decided PTP was important technology for our IBM Z clients.

The world's financial markets, even though very dependent on the IBM mainframe, effectively operate as high-performance distributed systems where the time stamp of any trade can have a huge influence on the financial fortunes of millions of investors both large and small. The

integrity and trust needed for these systems requires highly accurate and traceable time synchronization.

Accurate and trusted timing plays a critical role in financial markets. Since the inception of the direct electronic trading of financial instruments in the 1990s, the speed of financial market transactions has increased at an exponential rate, mirroring the advances made in CPU processing and network performance. The widespread usage of electronic trading platforms and automated stock exchanges that began in the late 1990s has fundamentally changed the way financial market transactions take place. Market makers are no longer individuals working telephones or waiting in a physical line to place orders. Computers now are automatically executing trades, both buying from and selling to each other based on software algorithms. All this automation led to a substantial increase in the number of exchanges competing against each other, while simultaneously the amount of time required to execute transactions and the “spread size” decreased. High Frequency Trading, or HFT became a widespread practice. Today, according to Financial Review, half of all transactions in today's stock markets are the result of HFT.

The rapid expansion of computer-based trading with its highly sophisticated algorithms has increased the need for synchronization of trading systems and traceability to a common reference time scale to help prevent irregularities and to aid forensics. The practice of HFT makes it essential for all stock exchanges, trading platforms, and associated IT hardware to be able to document that their time stamps are accurate. Institutions that interact with financial exchanges require high accuracy time for several reasons. One reason is that accurate time is necessary for institutions to control their own trading traffic. Accurate time is also required to settle disagreements and prevent fraud. The time stamps collected by each institution must be examined when there are disagreements or errors in processing trades. The most famous example of this occurred during the Flash Crash of 2010.

## Common Reference and Traceability

Author Lee Segall once said “It is possible to own too much. A man with one watch knows what time it is, a man with two watches is never quite sure.” Applying this to our discussion: Time stamps created by different networks or systems can only be meaningfully compared if they are based on the same reference. For time and computer systems, the global reference is Coordinated Universal Time (UTC). UTC is the time scale that underpins the Global Positioning System (GPS) and broadcasts time signals and precise time services. The Bureau International de Poids et Mesures (BIPM) in Sevres, France is responsible for calculating UTC. The BIPM collaborates internationally with over 70 timing institutes and laboratories around the world who contribute data from over 400 atomic clocks to the BIPM. UTC is a post-processed virtual time scale that does not produce a physical signal. UTC is calculated/obtained by taking a weighted average of all the submitted clock data from these institutes and laboratories.

The timing laboratories that contribute to UTC operate local time scales that do produce a physical signal and these local realizations of UTC can be used as measurement references for real-time applications such as providing time to financial markets. BIPM refers to these local UTC time scales as UTC(k), where (k) is the acronym of the specific laboratory. BIPM publishes a monthly document called Circular T, that lists the UTC-UTC(k) time differences for each UTC contributing laboratory. This time difference value, along with its associated measurement uncertainty, is provided for each contributing laboratory every five days. Circular T allows each of the participating laboratories to know the recent difference of its local UTC time scale with respect to UTC, therefore establishing traceability to UTC and the SI second. (As a quick review, the SI second is the basic unit of time in the International System of Units (SI)). SI is defined as the duration of a specified number of cycles of the

microwave radiation corresponding to a transition within Cesium 133 atoms.) The U.S. National Institute of Standards and Technology (NIST) maintains a local realization of UTC(NIST) at its laboratories in Boulder, CO. The UTC(NIST) time scale is kept in close agreement with UTC. Over a ten-year period from 2006-2015, the average UTC-UTC(NIST) was 4.2 nanoseconds.

The concept of traceability for measurements of time of day refers to a continuous chain of comparisons extending from a time comparison, time stamp, or clock synchronization performed by a user, back through the distribution of one of the UTC(K) time scales, and so on to the reference time scale UTC. Each of these comparisons in the chain is carried out with some inherent level of inaccuracy (uncertainty of the measurement). The uncertainties of each link in the overall chain can then be combined to give the total uncertainty of the time signal at the point where it is available to the user. Understanding the dissemination chain and calculating this uncertainty is essential to meeting requirements of being able to demonstrate traceability. Each link in the chain must be documented, along with its associated uncertainty evaluation, and the resulting total uncertainty of the timestamp or time output must be determined and recorded.

There are two additional requirements for demonstrating time traceability to a common reference. One is for the timing equipment to be calibrated so that its unknown internal delays do not bias its time output. The second is for the equipment to be monitored continuously so that any fault or anomaly can be detected, and the time output not used until the equipment is once again working correctly. The calibration evidence and monitoring results should be archived so that the status of the timing equipment at any given point in time can be verified later.

## Overview of Regulations: Past, Present and Future

Prior to the advent of High Frequency Trading (HFT) and distributed electronic trading platforms, many of the clocks maintained for financial transactions were mechanical devices that physically stamped the time (ink) onto paper documents used to record the transactions. These clocks were seldom synchronized to the granularity expected today. By the late 1990s it became obvious that the performance of these clocks was far too limited for financial markets, and synchronization requirements were implemented in the U.S. and in Europe. Since there are some substantial differences, let's examine the U.S. and EU requirements in separate following sub-sections.

### United States: NASD, OATS, FINRA, CAT

In August 1996, the U.S. Securities and Exchange Commission (SEC) issued a report alleging that the National Association of Securities Dealers (NASD) and NASDAQ did not always act in the best interest of customers. This report led to a financial settlement between the SEC and NASD that also included new regulations. To comply with the new regulations, the NASD issued a series of rules that were approved by the SEC in March 1998 and went into effect in August 1998. One rule was NASD OATS (Order Audit Trail System) Rule 6953, entitled "Synchronization of Member Business Clocks". Rule 6953 required computer systems and mechanical clocks to be synchronized every business day before the stock market opened to ensure that event time stamps are accurate. It also required the synchronization to be within three seconds of the NIST atomic clock, which while not particularly stringent, did finally force the removal of old clocks that could not display seconds. The NYSE adopted similar requirements in 2003.

The US based Financial Industry Regulatory Authority (FINRA) is a not for profit entity that is responsible for overseeing US brokerage firms and works closely with the SEC. In 2008, FINRA issued new requirements for stock market time synchronization. These new requirements, contained in FINRA OATS Rule 7430 "Synchronization of Member Business Clocks", superseded Rule 6953 for the NASDAQ (2008) and for the NYSE (2011). Similar to the previous requirements of Rule 6953, Rule 7430 lists NIST time as the official reference for U.S. stock market transactions. Rule 7340 made the synchronization requirement more stringent by a factor of three, now requiring synchronization within one second of NIST time. Rule 7340 also required that this one second synchronization be maintained at all times markets are open.

In relatively short order, it was realized that with the increasingly widespread use of automated trading platforms and HFT, Rule 7340's requirements were not stringent enough. In late 2014, FINRA sent out FINRA Regulatory Notice 14-47 for comments. Notice 14-47 proposed further tightening the synchronization requirements by an additional factor of 20, to within 50 milliseconds (ms) of NIST time. This dramatic change led to significant debate, delaying approval, in the form of FINRA Rule 4590, until April 2016. The SEC subsequently issued its Regulatory Notice 16-23 in July 2016. Under the rule change, effective August 15, 2016, computer clocks that are used to record events in NMS securities and OTC equity securities must be synchronized to within a 50-millisecond drift tolerance of the NIST atomic clock. The tolerance includes:

1. The difference between the NIST standard and a time provider's clock
2. Transmission delay from the source; and
3. The amount of drift in the member's clock

Firms were given six months from the effective date, until February 20, 2017, to meet the new synchronization standard for clock systems that capture time in milliseconds. Firms were given 18 months from the effective date, until February 19, 2018, to meet the new synchronization standard for clock systems that do not capture time in milliseconds. Firms must also document and maintain their clock synchronization procedures. Among other requirements, members must keep a log of the times when they synchronize their clocks and the results of the synchronization process. This log should include notice of any time the clock drifts more than the tolerance specified in Rule 4590. This log should be maintained for the period and accessibility specified in SEC Rule 17a-4(b), and it should be maintained and preserved for the required period in paper format or in a format permitted under *SEC Rule 17a-4*.

This documentation/log requirement stemmed in part from the 2012 SEC Rule 613, which required FINRA and U.S. based stock exchanges to establish a consolidated audit trail (CAT) that would enable regulators to better monitor and analyze trading activity. The CAT rules required FINRA and the stock exchanges to collect and accurately identify every order for all stocks and stock options across all U.S. markets. The rules also required the sending of complete documentation about orders to a central repository by 8 a.m. Eastern Time the day following a trade. CAT requires time stamps of millisecond resolution at five places in the audit trail:

1. The time of order origination
2. The time when the order is routed
3. The time when the order is received
4. The time when the order was modified and/or cancelled
5. The time when the order was executed

## Europe: MiFID II

The European Securities and Markets Authority (ESMA) is the closest equivalent in Europe to the SEC in the U.S. ESMA has the goal of ensuring fair and equitable financial markets and protecting investors. ESMA is empowered by the Market in Financial Instruments Directive (MiFID) to draft regulatory technical standards for European financial markets and firms. MiFID has been applicable across the European Union (EU) since November 2007. ESMA began the process of updating MiFID with a new Directive, MiFID II, in 2011 and MiFID II was formally adopted by the EU Parliament in June 2014, with the new requirements applicable January 3, 2018. The MiFID II clock synchronization requirements are more stringent than the latest U.S. requirements discussed earlier. The specific section of MiFID II that details these new clock synchronization requirements is Commission Delegated Regulation (EU) 2017/574 and is better known as Regulatory Technical Standard (RTS)-25. There are four articles to RTS-25 which are summarized below.

### **Article 1: Reference Time**

Article 1 states that business clocks that provide the timestamp for any reportable event should be coordinated to UTC, using either a link to one of the laboratories maintaining a UTC(k) realization of UTC, or the time signals disseminated by GPS or other satellite system. If using a satellite system, any offset from UTC must be accounted for and removed from the timestamp.

### **Article 2: Level of Accuracy for operators of trading venues**

Article 2 describes the level of accuracy, i.e. the maximum divergence from UTC, that should be achieved by the operators of financial trading venues, taking into account the gateway to gateway latency of their trading systems. This is summarized in Table A-1.

*Table A-1 Level of accuracy for operators of trading venues*

Gateway-to-gateway latency	Max. divergence from UTC	Timestamp granularity
> 1 millisecond	1 millisecond	1 millisecond or better
<= 1 millisecond	100 microseconds	1 microsecond or better

### **Article 3: Level of Accuracy for members or participants of a trading venue**

Article 3 defines the level of accuracy that apply to business clocks of members or participants of financial trading venues. These are summarized in Table A-2.

*Table A-2 Level of accuracy for business clocks*

Type of trading activity	Max. divergence from UTC	Timestamp granularity
Algorithmic HFT	100 microseconds	1 microsecond or better
Voice trading systems	1 second	1 second or better
Non-algorithmic, human intervention	1 second	1 second or better
Concluding negotiated transactions	1 second	1 second or better
Other	1 millisecond	1 millisecond or better

### **Article 4: Compliance with the maximum divergence requirements**

Article 4 specifies that operators of trading venues and their members or participants shall establish a system of traceability to UTC. They shall be able to demonstrate traceability to UTC by documenting the system design, functioning and specifications. They shall be able to identify the exact point at which a timestamp is applied and demonstrate that the point within

the system where the timestamp is applied remains consistent. The traceability system should be reviewed at least once per year to ensure compliance with the regulations.

## A comparison of NTP and PTP

PTP aims for much tighter synchronization than typically is achievable with NTP, with PTP capable of synchronizing devices within nanoseconds of each other over a common networking infrastructure. PTP is capable of even picosecond level synchronization via implementation of the high accuracy optional component of the latest IEEE 1588 standard along with ITU-T's Synchronous Ethernet technology.

For many use cases/industries, NTP is “good enough”. If you need the tighter synchronization such as required in the financial/banking industry, PTP is likely a better option for you. As these regulations become increasingly strict, we likely will see increased adoption of PTP in the industries subjected to these regulations.

The key reason why PTP is capable of much tighter, more accurate time synchronization is due to its use of hardware timestamping. NTP has extra latency and less accuracy because it uses software-based timestamping and all timestamp requests in NTP must wait for the local operating system. Also, unlike NTP, PTP devices will account for device latency by accounting for the amount of time that synchronization messages spend in each device.

PTP's BMC algorithm allows for PTP to more easily adapt to changing conditions/failure scenarios than NTP. The PTP BMC algorithm ensures PTP devices always have the highest quality time reference, even in failure and degraded operation scenarios. Finally, NTP requires all devices to be configured to reference a predetermined set of time servers prior to use, and its performance suffers when messages have to traverse network elements such as switches.

## Preparing for PTP

How do you prepare your environment for implementing PTP? The process is very similar to what you would do prior to introducing any other new technology into your data center: evaluate what you have, what you need, your budget, and take care of the gaps.

In terms of evaluating what you have, you need to look at several things. First of all, look at the time servers that are connected to the outside world (such as to an antenna receiving GPS time) and are providing your IBM Z Sysplex's Coordinated Timing Network (CTN) with its Server Time Protocol (STP) external time reference. Most likely today, your external time reference is NTP. Hopefully, your time server can also function as a PTP grandmaster. It likely can, but you will want to verify this, and ask the vendor what, if any code upgrades are needed for PTP support.

Secondly, the networking equipment in your time synchronization network will need to support PTP, specifically at a minimum you will need it to function as a PTP transparent clock. The good news here is that with PTP being a standard since 2002, it is likely that the switches/routers in your data center network, especially in the mainframe environment, are “PTP compliant” and you may only need to enable PTP in the management software, or worst case, do a code/OS update.

Third, you may want to use PTP for other server environments in the data center, especially for servers that play a role with applications impacted by the aforementioned regulations. In this case, you need to look at the server hardware, such as NICs, to see if they support PTP.



Does the server OS support PTP, and if so, do you have the driver levels required? What about the applications themselves?

Finally, it should go without say but make certain you have all of the IBM Z requirements met as well.

Once you have a complete understanding of what you have, you can look at the gaps and determine what you need, develop a budget, and address the gaps. Going forward, for new network, server, and other hardware purchases you will want to make certain that you are buying equipment that will support PTP. The same holds true for software and OS updates.

## Conclusion

This chapter explained the basics of PTP, including the history, an overview of the technology, the regulations that led to IBM to support PTP on IBM z, and some advice on how to prepare for implementing PTP. IBM Z started to investigate PTP in 2012-2013. These investigative/research activities ramped up in 2017 as clients started to inquire about PTP. The client interest in PTP led to IBM joining, actively participating, and now leading in PTP related standards development organizations (SDOs). This in turn led to IBM developing support for PTP with IBM Z CTNs and the announced statement of direction for z15 and future IBM Z machines.



# Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

## IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- ▶ *IBM z15 Technical Introduction*, SG24-8850
- ▶ *IBM z15 (8561) Technical Guide*, SG24-8851
- ▶ *IBM z15 (8561) Technical Guide*, SG24-8852
- ▶ *IBM Z Connectivity Handbook*, SG24-5444
- ▶ *IBM z14 (M/T 3907) Technical Guide*, SG24-8451
- ▶ *IBM z14 ZR1 Technical Guide*, SG24-8651
- ▶ *IBM z14 (M/T 3906) Technical Introduction*, SG24-8450
- ▶ *IBM z14 ZR1 Technical Introduction*, SG24-8550

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

[ibm.com/redbooks](http://ibm.com/redbooks)

## Other publications

These publications are also relevant as further information sources:

- ▶ Hardware Management Console (HMC) Operations Guide Version 2.15.0.  
See IBM [Resource Link](#) (requires IBM ID authentication).
- ▶ IBM Hardware Management Console Help

## Online resources

The IBM Resource Link for documentation and tools website is also relevant as another information source:

<http://www.ibm.com/servers/resourceLink>

## Help from IBM

IBM Support and downloads

[ibm.com/support](http://ibm.com/support)

IBM Global Services

[ibm.com/services](https://ibm.com/services)

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(->Hide)>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine.fm still open and **File>Import>Formats** the Conditional Text Settings (ONLY!) to the book files.

Draft Document for Review August 3, 2020 8:31 pm

8480spine.fm 135



# IBM Z Server Time Protocol Guide

SG24-8480-00  
ISBN DocISBN



(1.5" spine)  
1.5" <-> 1.998"  
789 <-> 1051 pages



Redbooks

# IBM Z Server Time Protocol Guide

SG24-8480-00  
ISBN DocISBN



(1.0" spine)  
0.875" <-> 1.498"  
460 <-> 788 pages

Redbooks

# IBM Z Server Time Protocol Guide

SG24-8480-00  
ISBN DocISBN



(0.5" spine)  
0.475" <-> 0.873"  
250 <-> 459 pages

Redbooks

# IBM Z Server Time Protocol Guide

(0.2" spine)  
0.17" <-> 0.473"  
90 <-> 249 pages

(0.1" spine)  
0.1" <-> 0.169"  
53 <-> 89 pages

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(->Hide)>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine:fm still open and **File>Import>Formats** the Conditional Text Settings (ONLY!) to the book files.

Draft Document for Review August 3, 2020 8:31 pm

8480spine.fm 136



# IBM Z Server Time Protocol Guide

SG24-8480-00

ISBN DocISBN

(2.5" spine)  
2.5" <-> mmm.n"  
1315 <-> mmm pages



# IBM Z Server Time Protocol Guide

SG24-8480-00

ISBN DocISBN



(2.0" spine)  
2.0" <-> 2.498"  
1052 <-> 1314 pages





SG24-8480-00

ISBN DocISBN

Printed in U.S.A.

Get connected

