Redbooks

ibm.com/redbooks

IBM Z Server Time Protocol Guide

Octavian Lascu

Hans-Peter Eckam

Gatto Gobehi

Steve Guendert

Clay Kaiser

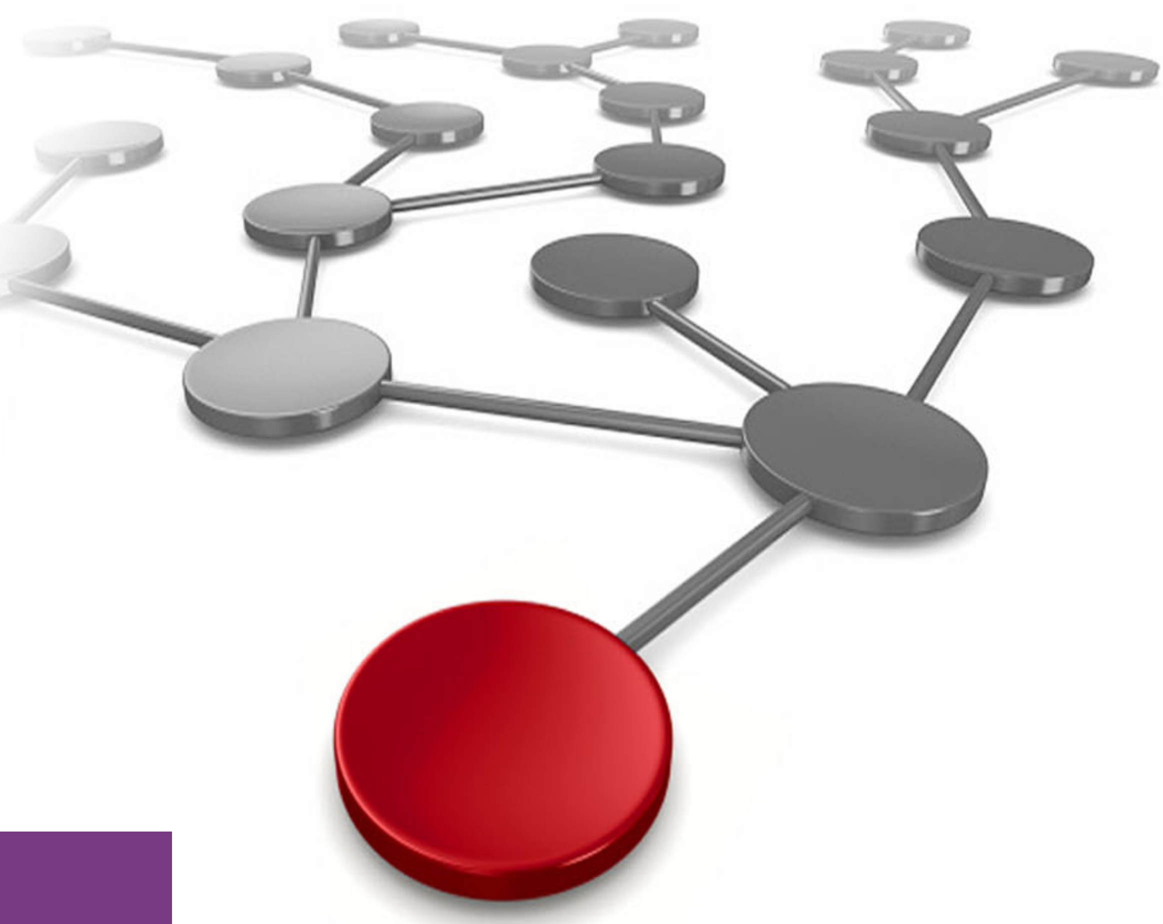
Walter Keller

Jeremy Koch

Franco Pinto

Martin Söllig

Sebastian Zimmermann



IBM Z

Redbooks

IBM Redbooks

IBM Z Server Time Protocol Guide

December 2023

SG24-8480-01

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

Second Edition (December 2023)

This edition applies to IBM Server Time Protocol (STP) for IBM Z platform and covers IBM z16, IBM z15, and IBM z14 server generations.

Contents

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®
CICSPlex®
Db2®
FICON®
GDPS®
IBM®
IBM Z®
IBM z13®
IBM z13s®
IBM z14®
IBM z16™
Parallel Sysplex®
Redbooks®
Redbooks (logo) ®
Resource Link®
z/Architecture®
z/OS®
z/VM®
z/VSE®
z13®
z13s®
z15®
z16™
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® platform. 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 STP functions. Readers are expected to be familiar with IBM Z technology and terminology.

This book provides planning and implementation information for STP functions and associated software support for the IBM z16™, IBM z15®, and IBM z14® platforms.

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 (HADR) solutions and high-performance computing (HPC) 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.

Hans-Peter Eckam is a Support Center Specialist at the IBM Hardware Support Center in Mainz, Germany. Hans-Peter joined IBM in 1986 and has over 25 years of experience supporting IBM Z clients. His areas of expertise include IBM Z Hardware, Sysplex Timer, STP, and connectivity features, such as ESCON, IBM FICON®, and InfiniBand.

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

Steve Guendert is the Brocade Principal Director of Product Management for IBM Z products and technology. He is an industry-recognized worldwide expert on ESCON, IBM FICON, and mainframe I/O, and is a published author of over 50 papers on those subjects. Steve was inducted into the Mainframe Hall of Fame in 2017. He has authored and contributed to several publications, including *Brocade Mainframe Connectivity Solutions* (from Brocade Communications Systems) and *Handbook of Fiber Optic Data Communication, 4th Edition*. Steve has a PhD and an MS in Management Information Systems (MIS), and an MBA. He also holds the distinction of having served on both the SHARE and Computer Measurement Group (CMG) Boards of Directors.

Clay Kaiser is a Firmware Developer for IBM Z Time Synchronization in Poughkeepsie, NY, US. He holds a degree in Electrical Engineering from the University of Illinois Urbana-Champaign. His area of focus includes security and resiliency of time synchronization.

Walter Keller is an IBM z/OS® senior systems engineer with UBS Business Solutions AG. In the last 35 years, he had several roles in the mainframe environment, starting as a mainframe storage manager, IBM Db2® system administrator, and z/OS systems programmer. His areas of expertise include IBM Z hardware, z/OS and its subsystems, IBM Parallel Sysplex®, and storage management.

Jeremy Koch is a Systems Programmer in Germany working for Finanz Informatik. He has more than 20 years of experience in z/OS, IBM z/VM®, and Linux on IBM Z. Jeremy is also designing DR solutions with IBM GDPS®, and working closely with the teams that are responsible for hardware planning and configuration.

Franco Pinto is a senior systems engineer who leads the z/OS and UNIX team at a major bank in Switzerland. He has almost 30 years of experience in the mainframe and IBM z/OS fields. His areas of expertise include IBM Z server management sizing, planning, and supervising deployment of IBM Z.

Martin Söllig is a Consultant IT Specialist in Germany. He has 30 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 Geographically Dispersed Parallel Sysplex (GDPS) environments, and also in cryptography on IBM Z.

Sebastian Zimmermann is a mainframe hardware configurator working for Finanz Informatik, the central IT service provider for the Savings Banks Finance Group in Germany. He holds a degree in computer engineering and has 14 years of experience in IBM Z hardware and its associated components, such as IBM FICON directors, DASD and TAPE systems, and the Hardware Management Console (HMC). His responsibilities also include coordinating mainframe hardware installations and planning the technical infrastructure that is required to integrate a mainframe environment into the overall configuration of a data center.

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

George Kozakos
IBM Australia

Thomas Franklin, Bill Lamastro, Mark Maruzzi, Paul Wojciak
IBM US

Bob Haimowitz and Ewerson Palacio
IBM Redbooks, Poughkeepsie Center

John Houston
IBM Poughkeepsie

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

Comments welcome

Your comments are important to us!

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

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

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, IBM Redbooks
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- 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

1

Introducing the Server Time Protocol

This chapter describes time synchronization and introduces the Server Time Protocol (STP) facility that is available for IBM Z servers.

The following topics are covered in this chapter:

- Introducing time synchronization
- Overview of Server Time Protocol
- STP concepts and terminology
- Coordinated Timing Network
- Coupling links
- External Time Source
- Firmware implementation

1.1 Introducing time synchronization

Historically, the most important requirement for highly accurate time is for navigational purposes. For applications such as 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 by 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 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, which makes it more consistent with civil, or solar, time.

Until 1967, the second was based on 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. Also, the need for an accurate time-of-day (TOD) measure was recognized, which led to the adoption of the following three basic scales of time:

- Temps Atomique International (TAI), which is also known as International Atomic Time. TAI 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 is based on astronomy. Its fundamental unit is the average solar day. The old definition of a second was 1/86,400th of this day. There are no leap seconds.
- Coordinated Universal Time (UTC), also known as Universal Time Coordinated or *temps universel coordonné*. UTC is derived from TAI and adjusted to keep reasonably close to UT1. UTC is the official replacement for (and equivalent to) the better-known Greenwich Mean Time (GMT).

1.1.1 Inserting 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, especially users that use 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) in 1972 defined 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 UTC, the atomic time scale that is maintained at the Bureau International des Poids et Mesures (BIPM) with the contribution of 69 national institutes that operate about 400 atomic clocks.

For more information, see National Institute of Standards and Technology (NIST).

The International Earth Rotation and Reference Systems Service (IERS) is responsible for monitoring the Earth's rotation. It announces the dates of application of any leap seconds that are required, which usually are timed for the end of 30 June or 31 December. For more information, see IERS Bulletin C.

Note: A leap second introduces an irregularity into the UTC time scale, so exact interval measurements are not possible by 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 The TOD clock

The TOD clock was introduced as part of the IBM System/370 architecture to provide a high-resolution measurement of real time that is suitable for the indication of date and time of day. The duration of the TOD clock, beginning with a value of zero and continuing until it wraps around to zero, is approximately 143 years and wraps on 18 September 2042. This duration is referred to as an *epoch*. A single TOD clock is shared by all CPUs in the configuration.

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 (OS) support, the extended TOD clock provides for unique timestamps across the systems in a sysplex.[1] The value of the TOD clock is directly available to application programs by using the instructions STORE CLOCK (STCK), STORE CLOCK FAST (STCKF), and STCKE. These instructions store the value of the clock in a storage location that is specified by the instruction. Figure 1-1 shows the format of the TOD clock.

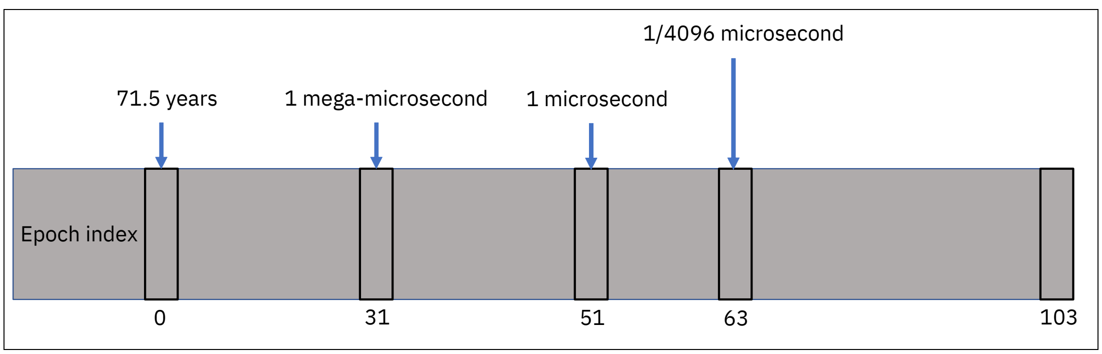


Figure 1-1 TOD format

Conceptually, the TOD clock is incremented so that a 1 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.

[1] For more information, see https://www.ibm.com/it-infrastructure/z/technologies/parallel-sysplex.

Figure 1-1 on page 3 shows also the stepping rate (the 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 usage of a binary counter for the TOD, such as the TOD clock, requires the specification of a time origin, or epoch. The epoch is the time at which the TOD clock value would have been all zeros. The IBM z/Architecture®, IBM ESA/390, and System/370 architectures established the epoch for the TOD clock as 1 January 1900, 0 a.m. GMT.

Note: The TOD clock is TAI-10 (10 seconds behind TAI). TAI was introduced in 1961. UTC was introduced in 1972, and started at TAI -10 seconds. There have been 27 adjustments to the leap second since 1 January 1972. Although the current leap second value is 37 (TAI -UTC = 37 seconds as of 1 January 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 networks, time synchronization accuracy for business clocks must be in compliance with regulatory mandates that are established by authorized organizations. Such organizations establish standards and impose regulations for investors' protection. Such organizations include the Financial Industry Regulatory Authority (FINRA) in the US, and the European Securities and Markets Authority (ESMA) in the European Union (EU).

Financial Industry Regulatory Authority

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

All currently supported IBM Z machines running STP already contain the features that are 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 leverage 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 if the PPS provider is attached to a compliant time source.

For more information about regulatory changes, see FINRA Regulatory Notice 16-23.

Market in Financial Instruments Directive

The Market in Financial Instruments Directive (MiFID II) is a European regulation for financial markets. It was published in the EU Official Journal on 12 June 2014. The directive empowers the ESMA to develop the Regulatory Technical Standard (RTS) and implement technical standards. ESMA delivered three sets of technical standards in 2015. The original plan was to put these standards into effect on1 January 2017, but the plan was y postponed to January 2018. The standards have been in effect since 3 January 2018.

RTS 25

RTS 25 is a part of the standards that are developed by ESMA in the context of MiFID II. RTS 25 defines standards for clock synchronization, which acknowledges that this topic has a direct impact on numerous business processes of the financial industry. The ever-increasing speed and volume of financial transactions require more accurate timestamps 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 timestamps that are used in business records.

Here are the requirements for clock synchronization as defined in RTS 25:

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

For more information about these regulatory changes, see Official Journal of the European Union, L 173, 12 June 2014.

For an overview about how to deploy a Coordinated Timing Network (CTN) by using STP with PPS, see Figure 1-2.

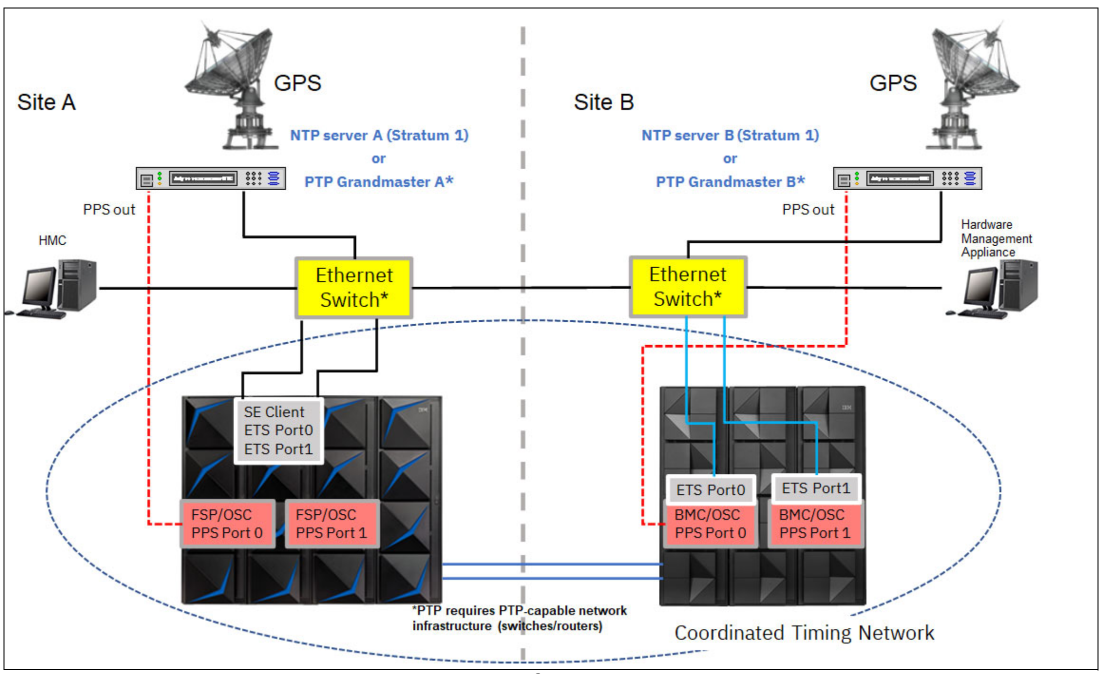


Figure 1-2 Principles of deploying STP with PPS[2]

[2] PTP is supported on IBM z15 and later.

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 TOD 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 ensures that every value that is stored by an STCK or STCKE instruction is unique. Consecutive STCK or STCKE instructions that run, possibly on different CPUs in the same server, must always produce increasing values. Thus, the timestamps can be used to reconstruct, recover, or in many different ways ensure the ordering of these serialized updates to shared data.

Parallel Sysplex and 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 running 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)[3]

  The 9037 ETR (the Sysplex Timer) was introduced with Parallel Sysplex. An ETR network was composed of 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.

- STP[3]

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

TOD-Clock Synchronization Facility[3]

The TOD-Clock Synchronization Facility provides an interface between the OS and the designated time server to allow the OS to accomplish the following tasks:

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

STP supports two ETS options: an external PTP server (master), or an external NTP server. At the time of writing, the client code for synchronizing CPC time to the ETS provides up to 100 milliseconds (ms) accuracy. Enhanced time accuracy can be obtained with the new IBM z16 ETS connectivity direct to CPC (with no Support Element (SE) involvement) without the usage of PPS. In our testing, PTP connectivity through z16 and later provided similar or equivalent accuracy to PPS[4]. However, results can vary based on the complexity of your PTP domain and network infrastructure. NTP connectivity direct to CPC is also enhanced, but should not be used without PPS connectivity if MiFID II compliance is required.

[3] Source: https://www.planetmainframe.com/2018/08/brief-history-mainframe-time/
[4] PPS provides 10 microseconds of accuracy and requires a connection directly to the oscillator card (OSC).

The process to achieve time consistency uses the following algorithm:

1. Server A runs an STCK instruction (timestamp x), which places the clock contents in storage.
2. Server A signals server B.
3. Server B, on receipt of the signal, immediately runs STCK (timestamp y).
4. Then, Server B must signal Server A, and Server A on receipt immediately runs STCK (timestamp z). The three values x, y, and z must be in order. Without this final step (or two), you can test only that server B's clock >= server A's clock.

For timestamp x and timestamp y to reflect the fact that y is later than x, the two TOD clocks must agree within the time that is required to send the signal. The consistency that is required is limited by the time that is required for signaling between the coupled servers and the time that is required by the STCK instruction itself. In practical terms, the CPC that has the role of Current Time Server (CTS) must provide to the TOD clocks of all participating systems with a timestamp so that all other CPCs can get to a synchronized state with each other to within a few microseconds, which is 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 is designed to help multiple IBM Z CPCs maintain time synchronization with each other without using 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 links, Integrated Coupling Adapter Short Range, or on machines earlier than IBM z14 ZR1; HCA3-O InfiniBand long reach links; and HCA3-O InfiniBand short reach links. These links can be the same links that are already being used in a Parallel Sysplex for CF 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 Processor Resource/Systems Manager (PR/SM).

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

In information technology, here are two important objectives for survival:

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

Systems that are designed to provide continuous availability combine the characteristics of high availability (HA) and continuous operations to deliver continuously high levels of service. 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 requires that the geographical cluster or Parallel Sysplex be dispersed over distances of 100 km or more to mitigate the risk that a single disaster might affect multiple data centers.

Up-to-date time synchronization must provide the following functions:

- Improved time synchronization (compared to Sysplex Timer) in a sysplex or non-sysplex configuration.
- Scales with distance. Servers exchanging messages over fast, short links require more stringent synchronization than servers exchanging messages over long distances.
- Scales with server and coupling link technologies. For example, the solution is re-used with appropriate changes when coupling link technologies change in the future.
- Supports a multi-site sysplex of at least 100 km without requiring an intermediate site and not precluding longer distances in the future.
- Allows implementation into an existing PTP or NTP network, which may or may not use a PPS signal.

The time synchronization infrastructure that is developed by IBM for the IBM Z environment is called *STP*.

Server Time Protocol

STP is a server-wide facility (Feature Code 1021) that is implemented in the LIC starting with the IBM z990 server. STP presents a single view of time to PR/SM and is designed to provide multiple STP configured servers to maintain time synchronization with each other while providing synchronized time to (supporting) OSs running across all IBM Z servers participating in the CTN. It is the follow-on to the ETR. STP is designed to allow events occurring in different servers to be properly sequenced in time.

STP is designed for servers that are configured in a Parallel Sysplex or a basic sysplex (without a CF), and for servers that are not in a sysplex but must be time-synchronized to an ETR.

STP is a message-based protocol in which timekeeping information is passed over data links among servers. The timekeeping information is transmitted over externally defined coupling links. Coupling links that can be used to transport STP messages are the Integrated Coupling Adapter Short Reach (ICA SR) and the Coupling Express Long Reach (CE LR) in IBM z14 and later IBM Z server families. Older coupling hardware (such as Inter-System Channel-3 (ISC-3) links configured in peer mode, Integrated Cluster Bus-3 (ICB-3) links, ICB-4 links, the InfiniBand links on servers before IBM z14 server family) are not supported.

STP provides the following functions:

- Allows 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 writing, the IBM z16 servers support only STP for time synchronization, and can coexist in the same CTN with IBM z15 and IBM z14 servers (N-2 generations rule).

Note: At the time of writing, IBM Z servers support STP-only CTN (No ETR or Mixed CTN).

With the introduction of IBM z16 A01 (M/T 3931), servers that are supported in the same CTN are IBM z16, IBM z15, and IBM z14. Although IBM z14 M0x (M/T 3906) servers still support InfiniBand Coupling links (HCA3-O and HCA3-O LR), they cannot be used to connect to IBM z16, IBM z15, or IBM z14 ZR1.

- Supports a multi-site timing network of up to 100 km (62 miles)[5] over fiber optic cabling, allowing a Parallel Sysplex to span these distances with no intermediate site requirement.

- Potentially reduces the cross-site connectivity that is required for a multi-site Parallel Sysplex.
- Allows the usage of an ETS to set the time to an international time standard, such as UTC, and adjust to the time standard on a periodic basis.
- Allows setting local time parameters, such as time zone and DST.
- Allow automatic updates of DST.

Standards-based PTP and NTP[a]: In the future, IBM plans to introduce standards-based PTP and NTP security mechanisms to increase the resiliency of STP and address the time synchronization security concerns tat clients are facing. Future STP implementation is planned to reintroduce the concept of a Mixed CTN, with support for traditional STP and native PTP implementations. Also, the goal of future IBM Z time synchronization development is to enhance the role of IBM Z machines in both traditional on-premises and hybrid cloud PTP and NTP environments that address the many regulatory and security concerns that clients are facing.

a. Statements by IBM regarding its plans, directions, and intent are subject to change or withdrawal without notice at the sole discretion of IBM. Information regarding potential future products is intended to outline general product direction and should not be relied on in making a purchasing decision.

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 that are related to STP. Understanding this vocabulary assists you in successfully planning a timing network that is based on STP.

1.3.1 STP facility

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

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 task is accomplished by using 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.

[5] Extended communications distance equipment (qualified by IBM) must be used. For more information, see IBM Resource Link® (IBMid required).

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. The value normally is less than the specified tolerance of the local oscillator. The change occurs infrequently (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 ETS or with other clocks in the timing network. The value normally changes frequently (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 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 is initialized to a local time and 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 is initialized to CST and stepped at the rate of the local hardware oscillator. In STP timing mode, the TOD clock is steered to maintain, or attain, synchronization with CST. To be in STP timing mode, the configuration must have an STP ID network that is defined. For more information about CST, see 1.4, “Coordinated Timing Network” on page 14.

Timing states

The timing state indicates the synchronization state of the TOD clock regarding the timing network reference time. The timing states are as follows:

- Synchronized state

  When a configuration is in the synchronized timing state, the TOD clock is in synchronization with the CST.

  > **Note:** A configuration that is not initialized or 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. 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.

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 CST that is required 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 is determined and can be used to synchronize the TOD clock to the STP network.

  For more information, see *z/Architecture Principles of Operations*, SA22-7832.

1.3.3 STP servers

This section describes STP servers. Throughout the rest of this chapter, we describe STP configurations because Mixed CTNs or ETR configurations are not available with IBM z14 and later.

STP-enabled server

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

STP-configured server

An STP-configured server is a server that is 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 not an STP-configured server. (CTN and the CTN ID are described in 1.4, “Coordinated Timing Network” on page 14).

1.3.4 STP stratum levels

The term *stratum* is found throughout this document and within the HMC windows.

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 that are distributed within a CTN, across IBM Z CPCs participating in the respective CTN.

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

STP distributes time messages in layers, or *strata*. The top level (*STP stratum 1*) retrieves time information from the ETS and distributes time messages to the layer immediately below it (*STP stratum 2*). STP stratum 2 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 illustrates the stratum concept.

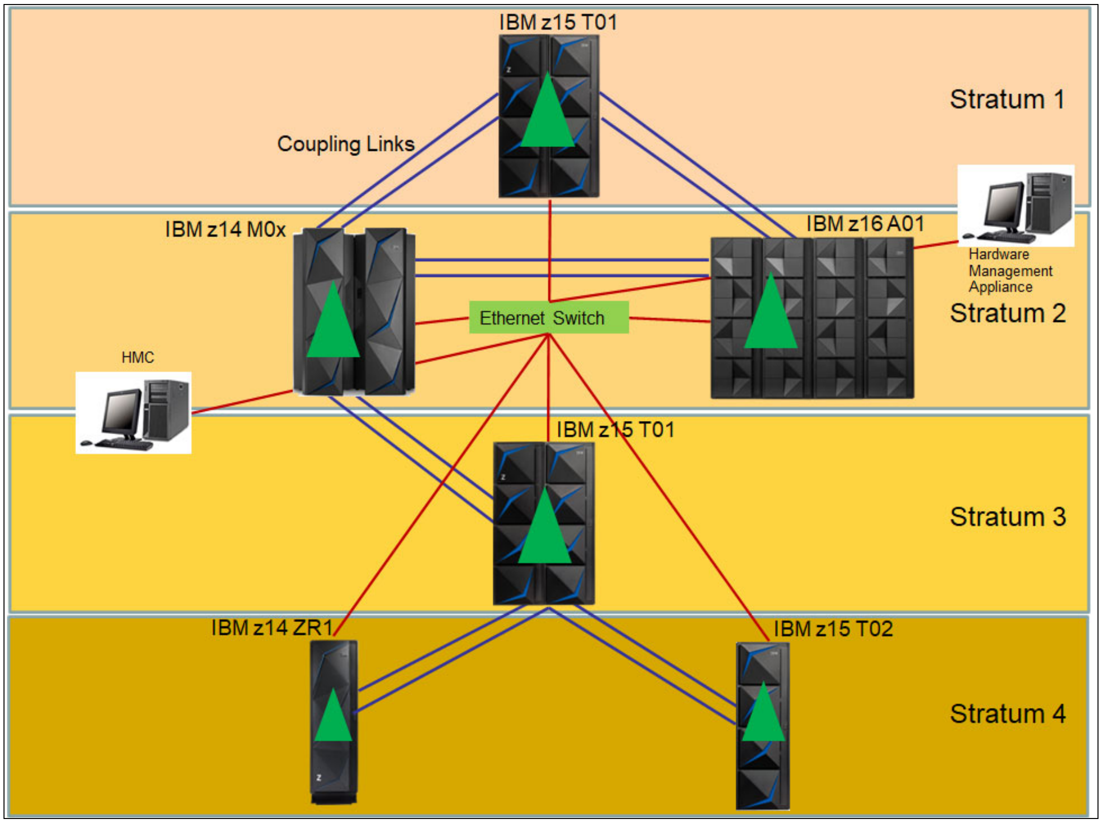


Figure 1-3 STP stratum levels

STP stratum levels (1, 2, 3, 4, and 0)

In a timing network that is based on STP, stratum levels are used 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.

In an STP-only CTN, the roles of Preferred Time Server (PTS) and CTSs are assigned, and the CTS becomes the active stratum 1 server. For more information about PTS and CTSs, see 1.3.5, “Server roles in a CTN” on page 13.

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

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

STP stratum 0

STP stratum 0 refers to a server (IBM Z CPC) that is in an *unsynchronized* state. In the unsynchronized state, the CPC does not receive a timing signal (either from an ETS 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. The CTN configuration is performed through the HMC windows for all participating serves.

> **Important:** IBM z14 is the last server to support the SE **Sysplex Timer** tab (HMC or SE 2.14.1). Later systems (z15 with HMC or SE driver 2.15.0 and later) support only the HMC **Manage System Time** feature (the SE task was discontinued in z15). As a consequence, it is mandatory that the HMC managing the CTN is at the latest level (of the latest server that is available for STP configuration).

Preferred Time Server

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

Backup Time Server

Although optional, it is a best practice to assign a BTS whose role it is to take over as the stratum 1 server if the PTS fails. The BTS is a stratum 2 server that should have connectivity to the PTS and the Arbiter, and to all other stratum 2 servers that are connected to the PTS.

Current Time Server

The CTS is the active (STP) stratum 1 server in a CTN, and it is used to steer the CST to its ETS.

> **Notes:**
>
> - There can be only one active (STP) stratum 1 server in a CTN, and it is the CTS. Only the PTS or the BTS can assume the CTS role and be the active Stratum 1 server.
> - After a recovery situation is resolved and full connectivity is restored, STP sets the CTS to the PTS server if the PTS has connectivity to all systems that are connected to the BTS.

The CTS can be assigned only to either the PTS or the BTS. As a best practice, assign the CTS to the PTS when the configuration is being initialized. Then, if there is a need to reassign the server roles, the CTS (stratum 1) role can be concurrently assigned to the BTS. This action can be part of a planned reconfiguration of the PTS if 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 an additional means for the BTS to determine whether it should take over the CTS role in 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 CTN contains a collection of servers that have their time synchronized. The time is synchronized to a value called CST. The CST represents the time for the entire network of servers (the CTN).

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

With the latest supported servers (IBM z16), a CTN can be configured as an 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 be configured only with STP-capable servers.

Notes:

- Support for the 9037 Sysplex Timer ended with the IBM z10 server family.
- IBM zEC12 and IBM zBC12 are the last IBM Z servers that support being part of a Mixed CTN that involves ETR.
- IBM Z servers starting with IBM z13® support STP timing *only* (no Mixed CTN).
- With IBM z15 and IBM z16, in addition to NTP, PTP (IEEE 1588-V2) was introduced as ETS. Different servers in the CTN may have their ETS as either PTP or NTP. However, one server can have only either PTP or NTP, but not both as ETS.

1.4.1 CTN ID

The CTN ID is an identifier that is used to indicate whether the server is configured to be part of a CTN, and if so, it identifies the CTN. In a STP-only timing network, the CTN ID is also referred as the STP Network ID[6].

The STP network ID (CTN ID) is shown in Figure 1-4 on page 15.

[6] The STP ID consists of 1 - 8 characters, alphanumerics, and symbols.

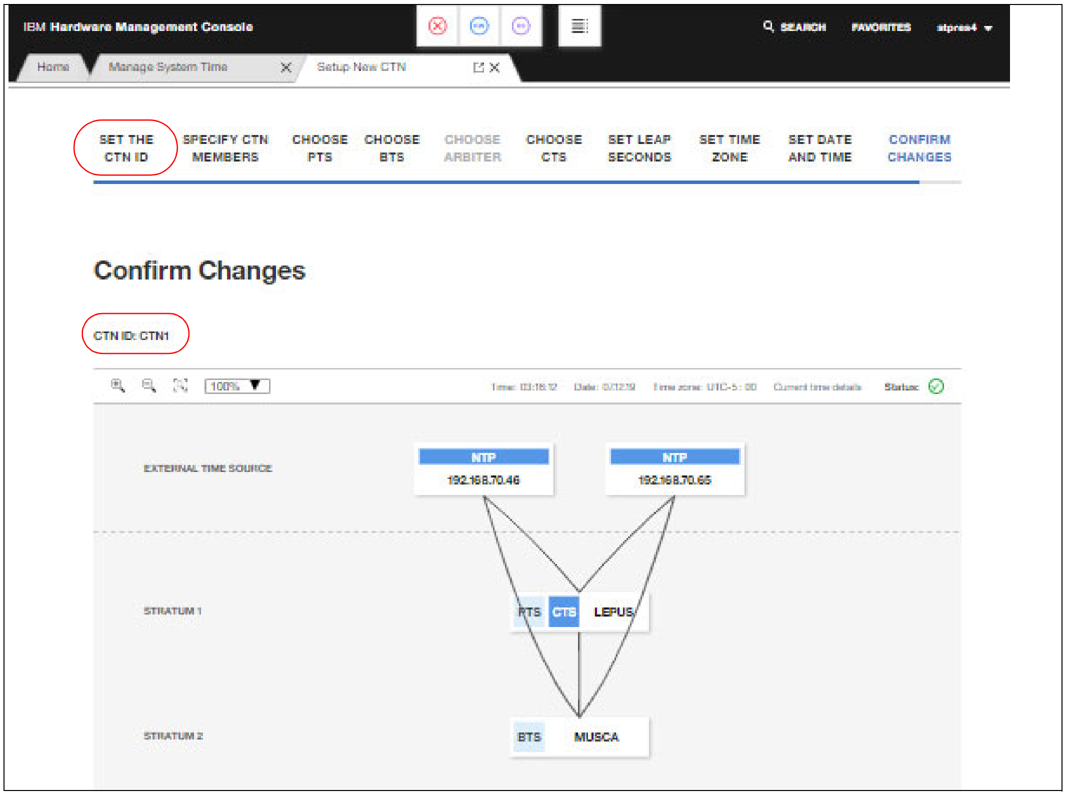


Figure 1-4 Displaying the CTN ID

The HMC is used to enter the CTN ID.

1.5 Coupling links

This section describes the connectivity options that support IBM Parallel Sysplex clustering technology and common time on the IBM Z platform. 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 that are connected to one or more CFs by using high-speed specialized links, which are called *coupling links*, for communication and time-keeping. The CFs 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 CFs. The CF provides critical locking and serialization, data consistency, messaging, and queuing capabilities so that systems in the sysplex can coordinate and share data.

A correctly configured cluster has no single point of failure and can provide users with near-continuous application availability during planned and unplanned outages.

Connectivity options

IBM z16, IBM z15, and IBM z14 ZR1 support three coupling link types:

- ICA SR links (8x optical) connect directly to the CPC drawer and are intended for short distances between CPCs of up to 150 meters.
- CE LR adapters (1x optical) are in the PCIe+ I/O drawers and support unrepeated distances of up to 10 km or up to 100 km over qualified Wavelength Division Multiplexing (WDM) services.
- Internal Coupling (IC[7]) links are for internal links within a CPC.

In addition to these three coupling link types, IBM z14 M0x (M/T 3906) servers also support the following links:

- One InfiniBand link, optical link, long wave.
- Twelve InfiniBand 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 z16 can participate in an IBM Parallel Sysplex cluster with IBM z16, IBM z15 T01, IBM z15 T02, IBM z14 M0x, and IBM z14 ZR1 servers.
>
> However, IBM z16, IBM z15, and IBM z14 ZR1 do not support InfiniBand connectivity, so they can connect to IBM z14 M0x only by using ICA SR or CE LR links.

Figure 1-5 shows the supported coupling link connections for the IBM z16.

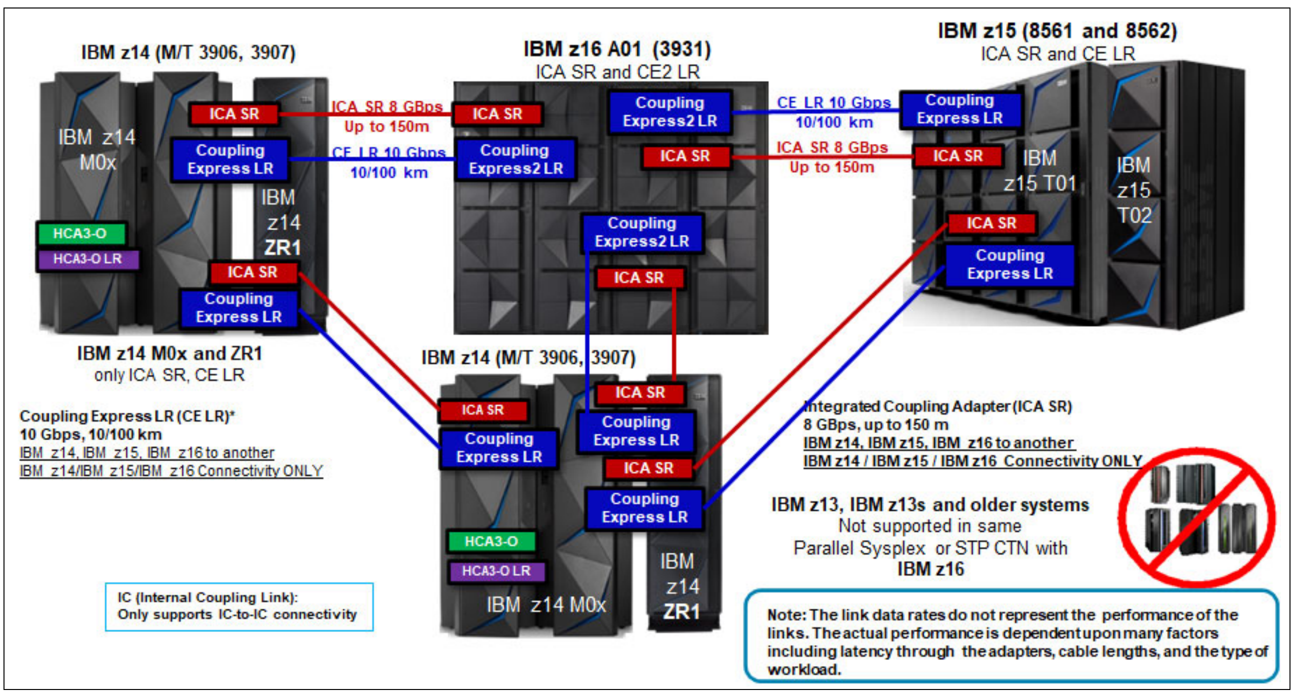


Figure 1-5 Parallel Sysplex connectivity options

InfiniBand links are not supported on IBM z16, IBM z15, and IBM z14 ZR1. The usage of InfiniBand between IBM z14 M0x servers must be planned cheerfully to avoid single points of failure for either the sysplex or the CTN.

[7] IC links do not carry STP (timing) signals.

For more information about supported coupling links, see *IBM Z Connectivity Handbook*, SG24-5444.

1.6 External Time Source

> **Note:** For an introduction to the PTP and more information (terminology and implementation overview), see Appendix A, “IEEE 1588 Precision Time Protocol introduction and overview” on page 137.

The ETS is used as a reference to which the CST is steered (kept in sync.). The current ETS implementation provides time information to the CTN by using either of the following options:

- PTP or PTP with or without PPS

  > **Note:** PTP is supported on IBM z15 or later CPCs. It is *not* supported as ETS on IBM z14 and earlier CPCs.

- NTP or NTP with or without 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 by using the **Configure External Time** function from the **Manage System Time** tab on the HMC. An ETS must be configured for the 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. If the role of the system is changed (a CTN reconfiguration from the HMC) to either the PTS or BTS, the ETS that is associated with the CPC can be used.

PTP Grandmaster with PPS support

STP can use, for the ETS, a PTP Grandmaster that has a PPS output signal. ETS 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 through a network device (such as a local area network (LAN) switch) that supports PTP. The PTP ordinary clock code is running on the SE for IBM z15, or on the CPC directly for IBM z16. In addition, the PPS output of the same Grandmaster clock server must be connected to the PPS port on the oscillator card (OSC) or flexible support processor (FSP) on IBM z15 and later servers. Redundant Grandmaster clocks with PPS ports are supported.

The PPS signaling provides a highly accurate event but no time information. Combining PTP Grandmaster with PPS 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 by using the **Configure External Time** function from the **Manage System Time** tab on the HMC. It can also be set on the **ETS** tab of the **System (Sysplex) Time** tab from the SE (IBM z14 and earlier CPCs only).

An ETS must be configured for the CTS. On systems before IBM 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 ETS. This type of ETS 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 on a port on the OSC on the IBM Z platform.

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

Figure 1-2 on page 5 shows a schematic drawing of an STP deployment that uses PPS.

> **IEEE 1588 PTP:** ETS connectivity is provided through the IBM Z SE for IBM z15 CPCs, or direct to the CPC for IBM z16.
>
> With PTP, there is no change to the usage of STP CTNs for time coordination, other than the potential to use a PTP-based ETS.
>
> Future implementation[a] is planned to re-introduce the concept of a Mixed CTN, with support for traditional STP and native PTP implementations. Also, 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 introduction and overview" on page 137.

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 functions that are described for IBM products remain at the sole discretion of IBM.

1.7 Firmware implementation

This section describes 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 that is used depends on whether STP is enabled. When STP is enabled, a CPC can participate in an STP CTN. In this case, the CTS 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 SE. Each logical partition (LPAR) starts out with this CPC TOD clock value at the completion of LPAR activation. The OS that is running in an LPAR can set a TOD value for itself, which is the only TOD reference that 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 during the LPAR activation, or until a subsequent Set Clock instruction is issued in the LPAR.

Server Time Protocol enabled

Enabling 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 OS in each LPAR can independently choose whether to synchronize to the current time source for STP, if it is present. OSs in LPARs that synchronize to STP run with identical TOD values. OSs 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 and later supports the specification of an LPAR time offset. When all members of a sysplex are in LPARs on these supported models, the LPAR time offset can be used for different local time zone support in multiple sysplexes by using the STP CTN. Many sysplexes must run with a `LOCAL=UTC` setting in a sysplex (`STPMODE=YES`), where the time that is returned from an STCK instruction yields local time. To fulfill this requirement, the time that is initialized for the STP CTN must be local time. With LPAR time offset support, multiple sysplexes can each have their own local time reported to them from an STCK instruction. For example, the STP CTN can be set to GMT, one set of sysplex partitions can specify an LPAR time offset of minus 5 hours, and a second set of sysplex partitions can specify an LPAR time offset of minus 6 hours.

External coupling links are also valid to pass time synchronization signals for STP. Therefore, the same coupling links can be used to exchange timekeeping information and CF messages in a Parallel Sysplex.

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

- Because 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 same time, which is 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 through highly efficient coupling links, IBM Z servers in the same CTN run on a highly accurate time that is synchronized within microseconds.

2

Planning hardware and software

The following topics are covered in this chapter:

- Planning server hardware
- Planning connectivity
- External Time Source
- IBM z16 power failure handling
- Internal Battery Feature
- Planning z/OS software
- Planning for z/VM, z/TPF, and Linux on IBM Z

2.1 Planning server hardware

Server Time Protocol (STP) is a server-wide facility that is implemented in the Licensed Internal Code (LIC) of all IBM Z servers since IBM z9®. Before IBM z9, the time reference for a sysplex was provided by an external hardware device that is called the IBM Sysplex Timer. The migration from a sysplex receiving the time signal from a sysplex timer over a sysplex with a subset of servers that is still connected to a Sysplex Timer (called a Mixed Coordinated Timing Network (CTN)) to an STP-only CTN had to be carefully planned. This topic is described in *Server Time Protocol Planning Guide*, SG24-7280, *Server Time Protocol Implementation Guide*, SG24-7281, and *Server Time Protocol Recovery Guide*, SG24-7380.

2.1.1 Servers

The IBM zEC12 and IBM 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 generation, IBM z16, a sysplex (and CTN) can consist of only the following servers:

- IBM z14 M0x (M/T 3906)
- IBM z14 ZR1 (M/T 3907)
- IBM z15 T01 (M/T 8561)
- IBM z15 T02 (M/T 8562)
- IBM z16 A01 (M/T 3931)
- IBM z16 A02 and IBM z16 AGZ (M/T3932)

Earlier servers like IBM z13, IBM z13s®, IBM zEC12, IBM zBC12 (and earlier) cannot be in the same sysplex and the same CTN with an IBM z16 server. IBM z14 and later servers can be only part of STP-only CTNs.

This section describes the planning for the servers and the related Hardware Management Console (HMC), Support Element (SE), and microcode by focusing on the following IBM Z CPCs: IBM z14 M0x, IBM z14 ZR1, IBM z15 (T01 and T02), and IBM z16 A01.

An IBM Z server requires Server Time Protocol Enablement (Feature Code 1021) to participate in a CTN. The N-2 generation coexistence rule applies to both the CTN and IBM Parallel Sysplex requirement: Only servers that are not earlier than two generations compared to the newest servers in the sysplex and the CTN are supported in the same sysplex and CTN.

Servers that are used as stand-alone Coupling Facilities (CFs) must be STP-enabled and configured in a CTN. For stand-alone CFs, the same N-2 generation coexistence rule applies. Table 2-1 provides a summary of the servers and the minimum Coupling Facility Control Code (CFCC) level that is required to be allowed in a sysplex and CTN with an IBM z16 server.

Table 2-1 Coupling Facility requirements

Server	CFCC level required on the coupling facility
IBM z14 M0x and IBM z14 ZR1	CFCC LEVEL 23
IBM z15	CFCC LEVEL 24
IBM z16	CFCC LEVEL 25

2.1.2 Hardware Management Console

The HMC provides the user interface to manage a CTN.

With the introduction of IBM z14 (HMC 2.14.0 with driver level 32), STP management is implemented as a GUI on the HMC (as the **Manage System Time** tab).

> **Note:** With IBM z16 (HMC 2.16.0 with driver level 51) the GUI for STP is enhanced, and the Sysplex Timer menus on the SE were discontinued. As such, the HMC GUI for STP is the *only* way to manage a CTN that contains an IBM z16 CPC.

HMC functions in a CTN

In a CTN, the HMC is used to complete the following tasks:

- Initialize or modify the CTN ID.
- Initialize the time manually or by using an External Time Source (ETS) to keep the Coordinated Server Time (CST) synchronized to the time source that is provided by the ETS.
- Initialize the time zone offset, Daylight Saving Time (DST) offset, and leap second offset.
- Configure the HMC as a Network Time Protocol (NTP) server.
- Configure access to an ETS so that the CST can be steered to an ETS:
  - To configure the connection from the Precision Time Protocol (PTP) client (ordinary clock) to PTP Grandmaster
  - To configure the connection from the PTP client (ordinary clock) PTP Grandmaster servers with Pulse Per Second (PPS)
  - To configure the connection from the NTP client to NTP servers
  - To configure the connection from the NTP client to NTP servers with PPS

  > **Note:** The PTP client and NTP client are implemented as a CPC Firmware function with IBM z16, so the ETS must be connected (Ethernet connection) to the IBM z16 CPC directly. With IBM z15 and IBM z14, the ETS must be connected (Ethernet connection) to the SE.

- Manage the CTN membership (add or remove servers, restrict or unrestrict membership, and split and merge CTNs).

  > **Note:** Split and merge operations are supported only for CTNs for IBM z14 and later.

- Assign the roles of Preferred, Backup, and Current Time Servers (CTSs), and the Arbiter.
- Adjust the time by up to +/- 60 seconds.

  > **Notes:**
  >
  > - Adjustments take approximately 8 hours to steer out 1 second, so steering out 60 seconds takes a little less than 20 days.
  > - Even though time adjustments of more than 60 seconds can be performed, they must be done in 60-second increments.

- Schedule changes to offsets that were previously listed (time zone, DST, and leap seconds). STP can automatically schedule DST based on the selected time zone.
- Monitor the status of the CTN.
- Monitor the status of the coupling links that are initialized for STP message exchanges.

The STP HMC GUI is used for the following operations:

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

The following advanced actions can also be performed:

- Join an existing CTN.
- Split to a new CTN.
- View the status of the PPS signal (whether a CPC is receiving such a signal).
- Save the STP debug data.
- Set the time server power failover.

Eligible Hardware Management Consoles

The HMC application level must be equal to or higher than the highest level of the SE that it manages. For example, if an IBM z15 is part of the CTN, HMC 2.15.0 or later is required. At the time of writing, CTN configurations include IBM z16 with HMC 2.16.0.

Historically, at a minimum, management of STP-enabled servers requires HMC 2.9.2 with the latest maintenance change level (MCL), which was provided with IBM z9 servers. Some functions were added or removed in later HMC levels. For example, the NTP server function was made available in HMC 2.10.1 and later, HMC 2.12.0 and later no longer support dial-out capability, and HMC 2.14.0 provided the new STP GUI, which became the only way to define and manage an STP with the removal of the STP windows from the SE with driver 41 (HMC 2.15.0).

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
Set the time server power failover.	Version 2.16.0	z16 only
Support for PTP configuration.	Version 2.15.0	IBM z15 and earlier
Removal of STP windows from the SE.	Version 2.15.0	IBM z15 and earlier N-2 server generations
CTN split and merge.	Version 2.14.1	IBM z14, IBM z14 ZR1, and earlier (N-4) generations

Change	HMC application level	Supported servers
Support for Coupling Express Long Reach (CE LR) connections (rolled back to IBM z13, no Going Away Signal (GAS)). Additional Stratum Level 4. GUI for STP.	Version 2.14.0	IBM z14 and previous (N-4) generations
Removed Mixed CTN support. Support for ICA connections. HMC STP window enhancements.	Version 2.13.0 and 2.13.1	IBM z13, IBM z13s, and earlier (N-4) generations
HMC NTP authentication. Removed dial-out support.	Version 2.12.0 and Version 2.12.1	IBM zEC12, IBM zBC12, and earlier (N-4) generations
Removed External Time Reference (ETR) support.	Version 2.11.0 and Version 2.11.1	IBM z196, IBM z114, and earlier (N-4) generations
NTP server in the HMC.	Version 2.10.1 and Version 2.10.2	IBM z10EC, IBM z10BC, and earlier
STP support.	Version 2.9.2	IBM z9EC, IBM z9BC, and earlier

Usually, previously installed HMCs can be upgraded by requesting that your IBM System Services Representative (IBM SSR) order the appropriate ECA for each server that has one or more HMC features that must be upgraded.

The only orderable HMC feature for a new IBM z16 is the Hardware Management Appliance (HMA) (Feature Code 0129). The following existing HMCs can be carried forward within a model conversion:

- HMC Tower (Feature Code 0062) and HMC Rack Mount (Feature Code 0063) carry forward in a conversion from an IBM z15.
- HMC Tower (Feature Code 0082) and HMC Rack Mount (Feature Code 0083) carry forward in a conversion from an IBM z14.

Setting up the HMC for local and remote operation

The management functions that are provided by the HMC require that all CPCs that need role assignment in the CTN to be defined objects to any HMC that is used to manage the CTN. Furthermore, as a best practice, all CPCs in the CTN should be defined to any HMC to enable CTN reconfiguration (for example, within a recovery scenario). Even remote HMCs should have all CPCs defined in the CTN.

> **Important:** CPCs that are not defined to the HMC that is used to manage the CTN will not appear in the STP GUI (**Manage System Time** tab on the HMC), except when they already have a role that is assigned (PTS, Backup Time Server (BTS), or Arbiter). If a CTN reconfiguration is required, it is not possible to assign a role to servers that are not known to the HMC. For role assignment, the HMC needs access to both the Current Time Server (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 area network (LAN) as the SE servers, the local HMC automatically detects the SEs and sets up all the necessary internal configuration information for communication without requiring more information from the users.
- When the HMC is on a remote LAN, defining the servers requires manual (user) action.

For more information and guidance, see the *Hardware Management Console Operations Guide* for the driver level that is installed on your HMC. This publication can be found at the IBM Resource Link website.

> **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 that are specified in each server documentation to connect HMCs to SEs within a site and between sites.

HMC application programming interface

Specific STP-related attributes were added to the Defined CPC objects starting with HMC 2.10.1. The following STP commands were added to manage the CTN or to perform automated recovery actions, such as reassigning the PTS, BTS, or Arbiter role in 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, see *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 communication between a remote HMC and the SEs that it manages is encrypted. Certificates for secure communications are provided and can be replaced if needed. TCP/IP access to the remote HMC is controlled through the HMC internally managed firewall and limited to HMC-related functions.

> **Note:** The HMC Domain Security can be used to isolate servers on a common LAN or provide more 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 SE connectivity options, see the server installation manuals:

- IBM z16 A01: *3931 Installation Manual for Physical Planning*, GC28-7015
- IBM z15 T01: *8561 Installation Manual for Physical Planning*, GC28-7002
- IBM z15 T02: 8562 *Installation Manual for Physical Planning*, GC28-7009
- IBM z14 ZR1: *3907 Installation Manual for Physical Planning*, GC28-6974
- IBM z14 M0x: *3906 Installation Manual for Physical Planning*, GC28-6965

2.1.3 Connectivity to ETS

For IBM z15 and earlier servers, the SEs are used for STP management and functions. The SE‘s internal storage provides STP initialization, configuration, and timing parameters. The STP configuration information is mirrored to the alternative SE in case the primary SE encounters a problem.

With IBM z16, the STP management and functions (including ETS LAN connectivity) moved to the CPC (prior implementations included the SE in the time path), which improves (reduces) dispersion.[1]

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, and Arbiter) are not preserved across a POR or power off/on sequence of the CTN unless the CTN is configured as a restricted CTN. CTN membership restriction must be configured to preserve STP role information across PORs.

Note: The configuration of a CTN that has *been restricted* is described in 4.2.2, “Single- or dual-server auto resume after power-on reset” on page 122. This process is also known as *saving CTN configuration across POR operations*.

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

The server roles must be reassigned after the following situations occur:

- In the CTS in a single server, an unrestricted CTN underwent POR or a power off/on sequence.
- The PTS and BTS (if assigned) in an unrestricted CTN with two or more servers experience a power outage at the same time (for example, a data center power outage).
- A CTN is unconfigured intentionally by the user.

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

When the ETS is configured to use a PTP Grandmaster, or an NTP server with or without PPS, the client code is on the SE for IBM z15. For IBM z16, the client code runs a firmware function in the CPC. 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 36.

2.1.4 Considerations for using PTP as ETS for IBM Z

IBM Z uses the PTP Best Master Clock Algorithm (BMCA) to automatically discover and synchronize to a compatible PTP Grandmaster. A few compatibility considerations apply.

Note: PTP Domain 0 is required for IBM Z.

[1] Dispersion is the maximum difference that is recorded between the time client and the time source (server).

In addition, the IBM Z platform is compatible with PTP Enterprise Profile as defined by the IETF.[2] Read the standard to learn more about the problem statement and the requirements. Generally, the following requirements apply:

- User Datagram Protocol (UDP) connections are required (IPv4 only).
- One-step or two-step mode is allowed.
- End-to-End Delay Mechanism is required.
- Hybrid Mode support is required:
  - Sync messages *must* be multicast event messages (UDP/319).
  - Follow-up messages *must* be multicast general messages (UDP/320).
  - Announce messages *must* be multicast general messages.
  - Delay Request messages *may* be multicast or unicast event messages. IBM Z platform sends unicast Delay Requests.
  - The master clock should be able to respond to multicast or unicast Delay Request messages with corresponding multicast or unicast Delay Response messages.
- Sync, Announce, and Delay Request message rates should be set to once per second.

2.2 Planning connectivity

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

2.2.1 Planning 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 that are considered to be STP-capable data links.

Since the implementation of STP with IBM z9, there is an ongoing evolution of external coupling links that are available to exchange STP messages to keep servers and CFs synchronized (list from the newest to the oldest):

- Integrated Coupling Adapter Short Reach (ICA SR) links
- Coupling Express and Coupling Express2 Long Reach (CE/CE2 LR) links
- InfiniBand links
- Integrated Cluster Bus (ICB) links
- Inter-System Channel (ISC) links

There are significant differences between these types of links regarding 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 z16, IBM z15, and IBM z14 ZR1 servers support only CE LR and ICA SR links, and IBM z14 M0x, IBM z13, and IBM z13s servers also support HCA-O (InfiniBand) links.

Note: If an IBM z16 is part of a CTN or sysplex configuration, only IBM z15, IBM z14 M0x, and IBM z14 ZR1 can be connected to the IBM z16, and only by using ICA SR or CE LR links.

[2] https://datatracker.ietf.org/doc/html/draft-ietf-tictoc-ptp-enterprise-profile

Older coupling link connectivity options such as ISC and ICB are beyond the scope of this book.

Figure 2-1 shows the coupling link connectivity options for a Parallel Sysplex.

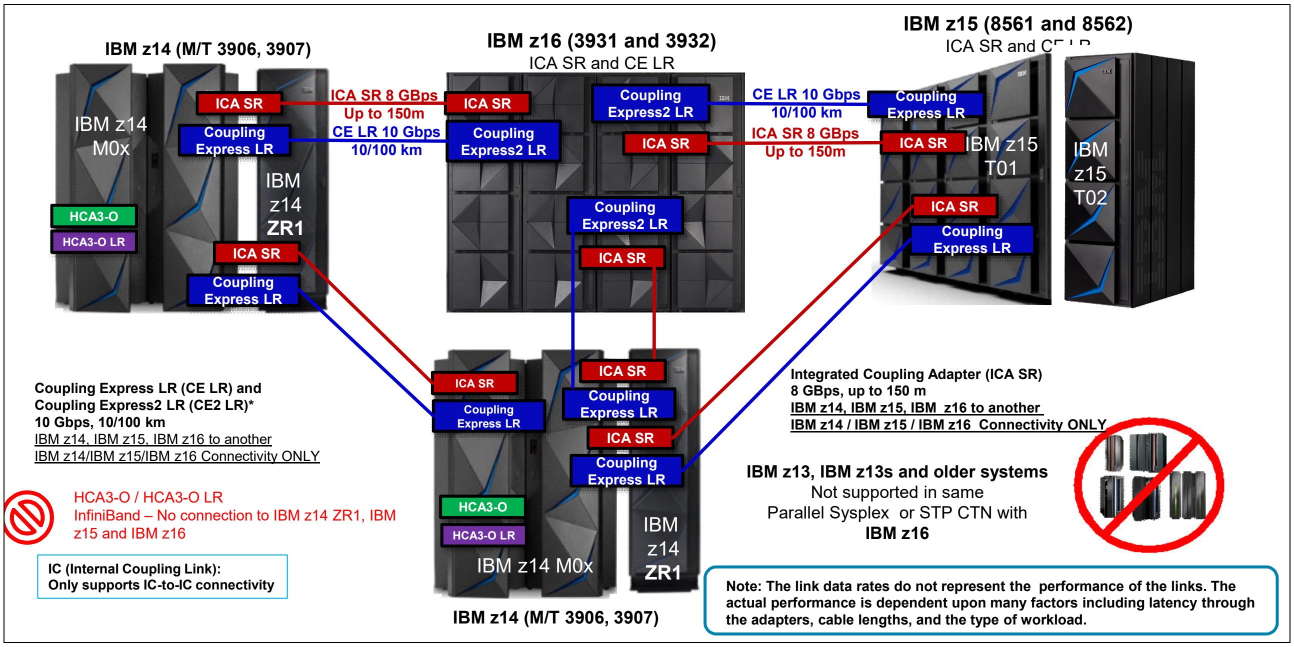


Figure 2-1 Parallel Sysplex connectivity options

ICA SR links

ICA SR links are high-speed external coupling links that are available on IBM z16, IBM z15, IBM z14 M0x, and IBM z14 ZR1. ICA SR features are 2-port, high bandwidth, and short-distance coupling adapters.

Here are the ICA SR coupling links by Feature Code:

- ICA SR (Feature Code 0172)
- ICA SR1.1 (Feature Code 0176)

The ICA SR feature (Feature Code 0172) was introduced with the IBM z13. IBM z15 introduced the ICA SR1.1 feature (Feature Code 0176). Both ICA SR and ICA SR1.1 use channel type (CHPID) type CS5. 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 and links per feature.

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

For more information about planning for the configuration of ICA coupling links, see *IBM Z Connectivity Handbook*, SG24-5444, *IBM z16 Configuration Setup*, SG24-8960, and *IBM Z Planning for Fiber Optic Links,* GA23-1407.

CE LR and CE2 LR links

CE LR links are high-speed external coupling links that are available on IBM z16, IBM z15, IBM z14 M0x, and IBM z14 ZR1. The CE LR coupling links are installed in a PCIe I/O drawer and allow the supported IBM Z servers to connect to each other over extended distance (10 km or 6.2 miles without repeaters).

Here are the CE LR coupling links by Feature Code:

- Coupling Express LR (two ports or links per feature) (Feature Code 0433) for IBM z15, IBM z14 M0x, and IBM z14 ZR1 (no carry forward for IBM z16)
- Coupling Express2 10G LR (two ports per feature) (Feature Code 0434) for IBM z16

CE LR is defined in IOCDS like InfiniBand and ICA links, that is, by using CHPID type CL5. Even though this feature is a PCIe feature, a physical channel identifier (PCHID) is used instead of an adapter ID (AID) to identify the physical card. For more information about planning the configuration of CE LR coupling links, see *IBM Z Connectivity Handbook*, SG24-5444, *IBM z16 Configuration Setup*, SG24-8960, and *IBM Z Planning for Fiber Optic Links,* GA23-1407.

InfiniBand links

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

Here are the InfiniBand coupling links that are available on z14 M0x, z13, and z13s by Feature Code:

- HCA3-0 LR PSIFB 1x (four ports or links per feature) (Feature Code 0170)
- HCA3-0 PSIFB 12x (two ports or links per feature) (Feature Code 0171)

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

For more information about planning for configuration of InfiniBand coupling links, see *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 Virtual Channel Path Identifiers

An AID is assigned to every InfiniBand 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 AID. The AID has the following characteristics:

- A number 00 - 2F on z16 (with all four CPC drawers)
- A number 00 - 3B on z15 T01 (with all five CPC drawers)
- A number 00 - 17 on z15 T02 (two CPC drawers)
- A number 10 - 17 on z14 ZR1
- A number 00 - 37 on z14 (with all four CPC drawers)

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

A PCHID normally has a one-to-one relationship between the identifier and a physical location in the machine, but a PCHID in the range 0500 -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 (IC) connections), or because they are administered through different identifiers (as for InfiniBand 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 IDs 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, so the VCHID assignment can change after a POR if the hardware configuration changed (for example, if an HCA was added or removed). Due to the automatic assignment of the VCHID at every POR, the client or IBM SSR must ensure that the correlation for the channel that they intend to manipulate has not changed. The VCHID that is associated with a coupling CHPID can be found by issuing an MVS `D CF,CFNM=xxxx` command for the associated Coupling Facility (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 by using the STP GUIU on the HMC by selecting the connection between two servers in the CTN. For example, in Figure 2-2, 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 (by using NTP as ETS).

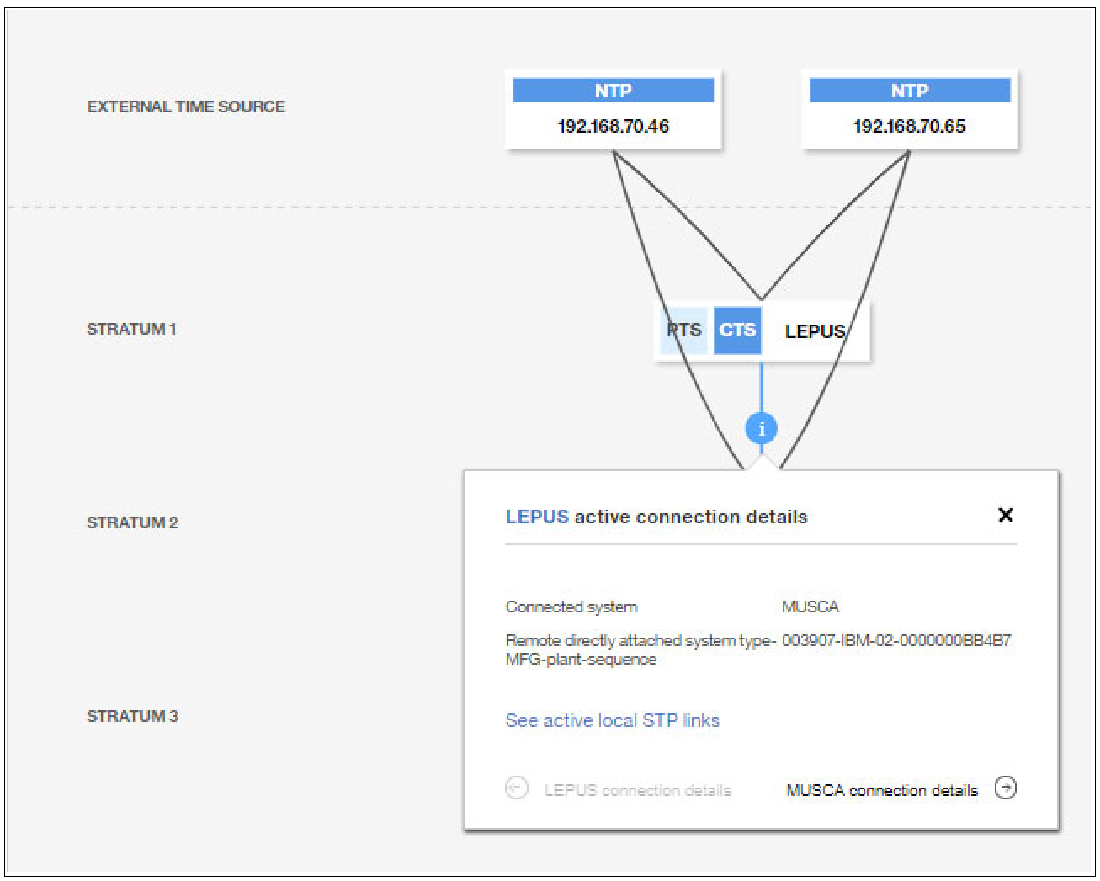


Figure 2-2 Displaying Coupling Link connections in a CTN

To see the details for these links, as shown in Figure 2-3, select **See active local STP links**. VCHID 0501is assigned to Port 02 at AID 0038, and VCHID 0503 is assigned to Port 02 at AID 0039.

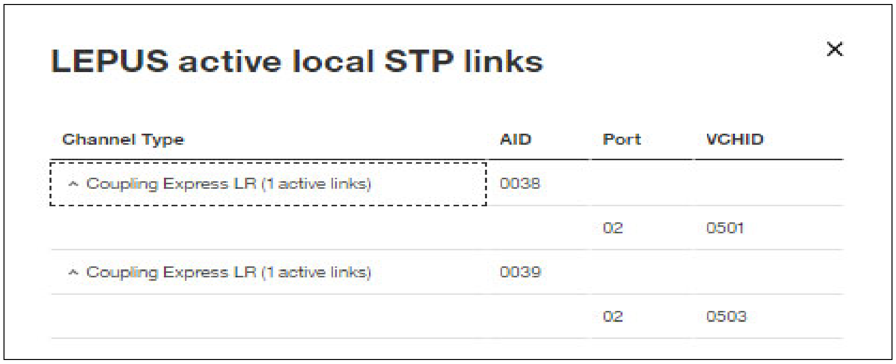


Figure 2-3 Displaying Coupling Link connections in a CTN: VCHID information

Enhanced Drawer Availability

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 by using a second path from a different CPC drawer. With Enhanced Drawer Availability (EDA), 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 STP, it is necessary to ensure that coupling links are not impacted during the removal of a CPC drawer. This task is especially critical where ICA SR and InfiniBand links are concerned because they are connected directly to the CPC drawers.

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

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.

Note: Standard ETS connectivity uses CPC0 and CPC1. ETS cables (Ethernet and PPS) cables might need to be moved to an alternative drawer as part of EDA (for CPC Drawer maintenance).

For more information about EDA, see the following publications:

- *IBM z16 (3931) Technical Guide*, SG24-8951
- *IBM z15 (8561) Technical Guide*, SG24-8851
- *IBM z15 (8562) Technical Guide*, SG24-8852
- *IBM z14 ZR1 Technical Guide*, SG24-8651
- *IBM z14 (3906) Technical Guide*, SG24-8451

Maximum number of links

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

Important: Provide at least two coupling links between any two servers that are intended to exchange STP messages for synchronization to prevent the loss of one link and causing the loss of STP communication between the servers.

Table 2-3 lists the maximum number of external coupling links that each server type supports in a Parallel Sysplex (IBM z14 and later servers). IC channels are not used by STP.

Table 2-3 Supported Coupling link options

Type	Description	Feature Code	Link rate[a]	Max unrepeated distance	Maximum number of supported links[b]				
					IBM z16	IBM z15 T01	IBM z15 T02	IBM z14 ZR1	IBM z14 M0x
CE2 LR	Coupling Express2 LR	0434	10 Gbps	10 km (6.2 miles)	64	N/A	N/A	N/A	N/A
CE LR	Coupling Express LR	0433	10 Gbps	10 km (6.2 miles)	N/A	64	64	32	64
ICA SR1.1	Integrated Coupling Adapter	0176	8 GBps	150 meters (492 feet)	96	96	48	N/A	N/A
ICA SR	Integrated Coupling Adapter	0172	8 GBps	150 meters (492 feet)	N/A	96	48*	16*	80*
HCA3-O LR	InfiniBand Long Reach (1 x InfiniBand)	0170	5 Gbps	10 km (6.2 miles)	N/A	N/A	N/A	N/A	64*
HCA3-O	InfiniBand (12 x InfiniBand)	0171	6 GBps	150 meters (492 feet)	N/A	N/A	N/A	N/A	32*
IC	Integrated Coupling Adapter	N/A	Internal speed	N/A	64	64	64	32	32

a. The link data rates do not represent the performance of the links. The performance depends on many factors, which include 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. 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 by using STP messages when a CF does not exist at either end of the coupling link.

Hardware Configuration Dialog: timing-only links

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 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 CF logical partition (LPAR) can be established by one or more timing-only links.

A timing-only link is established through the CF Channel Path Connectivity List panel in HCD (Figure 2-4). Defining timing-only links is much like defining other coupling links. InfiniBand links are defined as CIB, ICA SR links are defined as CS5, and CE LR links are defined as CL5 by using the normal CF connection dialogs. As with coupling links, timing-only links can be defined as dedicated, reconfigurable, shared, or spanned. The 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 exists between two servers, HCD does not allow defining a timing-only connection between these servers. Also, if a timing-only link exists between two servers, HCD will not allow you to define 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 information about how to set up a timing-only link, see *IBM z16 Configuration Setup*, SG24-8960.

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

Stand-alone coupling facilities and timing-only links

Dynamic I/O reconfiguration supports the migration from timing-only to standard coupling links. However, a stand-alone CF has no z/OS image to facilitate the initiation of a Dynamic I/O reconfiguration. For IBM z14, IBM z14 ZR1, IBM z15, and IBM z16 servers with Driver 36 or later, dynamic I/O changes are supported even for stand-alone CFs by using an activation service that is implemented in firmware (for more information, see *IBM z16 Configuration Setup*, SG24-8960.

Important: As a best practice, use a single IODF for all servers in the installation.

2.2.3 Considerations for a multi-site sysplex

In current business and IT environments, two important objectives for survival are systems that are designed to provide continuous availability and near-transparent disaster recovery (DR). Systems that are designed to deliver continuous availability combine the characteristics of high availability (HA) and near-continuous operations to deliver high levels of service. 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 IBM GDPS.

If a sysplex across two or more sites is configured, it is necessary to synchronize servers at multiple sites. Therefore, you might plan to extend the CE LR[3] links beyond the distance that is supported without repeaters. IBM supports Wavelength Division Multiplexing (WDM) products that are 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[a], transponders[b], and TDM modules[c]) with interface cards that are qualified by IBM for STP.

If both CF 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 optical multiplexers/demultiplexers, optical add-drop modules (OADMs), small form-factor pluggable (SFP) transceivers, dispersion compensation modules, optical amplifiers, and so on.
b. Wavelength-converting transponders: Converts a client-layer signal to a WDM wavelength.
c. Time-Division Multiplexing modules: Used to transmit multiple data links over the same WDM channel.

[3] HCA3-O LR links can also be extended, but they can be used only between IBM z14 M0x servers in a CTN where an IBM z16 is present.

The latest list of qualified WDM vendor products and the links to the corresponding IBM Redbooks publications for each product can be found at IBM Resource Link. They are listed under *Hardware products for servers* on the Library page.

2.2.4 Planning for External Time Source network connectivity

Before IBM z16, the ETS connection was established by using existing LAN connectivity that accessed the HMC and the SE. Starting with IBM z16, the ETS is attached directly to the CPC drawer, which produces higher accuracy compared to an attachment through the SE. For more information, see 2.3, “External Time Source” on page 36.

This approach requires at least one LAN connection from the network that hosts the ETS to the CPC oscillator card (OSC). For redundancy, establish two extra LAN connections that provide access to the PTP or NTP servers.

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 the PPS output option
- An NTP server without or with the PPS output option

The CTS is the only server that adjusts the CST in a CTN by steering it to the time that is obtained from an ETS. In addition, if the CTS has lost connectivity to its ETS servers, the non-CTS server sends its ETS steering information to the CTS, which then steers it to 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. The only information that is conveyed to the client is the NTP stratum level of the NTP server to which it synchronizes its time, and the PTP ordinary clock (running on the client device) synchronizes to the PTP Grandmaster (the PTP Grandmaster is “known” to the client device), which provides increased time accuracy without requiring special hardware (that is, PPS signal support).

Starting with IBM z16, the NTP and PTP servers are connected to the CPC drawer (this change is a major one compared to previous generations where they were connected to the SE). Similar to the PPS signal, the NTP and PTP frames are routed directly to the CPC drawer to provide higher accuracy. In our testing, PTP connectivity through z16 and later provided similar or equivalent accuracy as PPS. However, results might vary based on the complexity of your PTP domain and network infrastructure. NTP connectivity direct to CPC is also enhanced, but should not be used without PPS connectivity if MiFID II compliance is required.

The direct to CPC drawer ETS attachment requires extra tasks to be performed, such as setting the IP addresses, gateway, and static routes (if required) for ETS1 and ETS2 ports in the Customize Network Settings panel on the SE.

2.3.1 Precision Time Protocol Grandmaster

The usage of a PTP Grandmaster as an ETS addresses the requirement of customers who need a highly accurate time source for the IBM Z servers or must use an accurate common time reference across heterogeneous platforms. These customers usually purchase a PTP server (Grandmaster) that obtains the time by using high-precision time technology, such as GPS signals.

PTP requires the appropriate (supporting) infrastructure that can transmit PTP signaling from the PTP Grandmaster to the IBM Z CPC. For more information about the identifying requirements for PTP connectivity, see your PTP Grandmaster and network infrastructure documentation.

Redundant PTP Grandmasters are recommended for ETS availability. In the current implementation, more network connectivity for the IBM Z CPC drawer or SE is required. For more information about configuring an SE network to support PTP, contact your IBM SSR.

For the highest time accuracy, the current IBM Z PTP implementation requires PPS connectivity between the PTP Grandmaster and the IBM Z CPC.

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

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

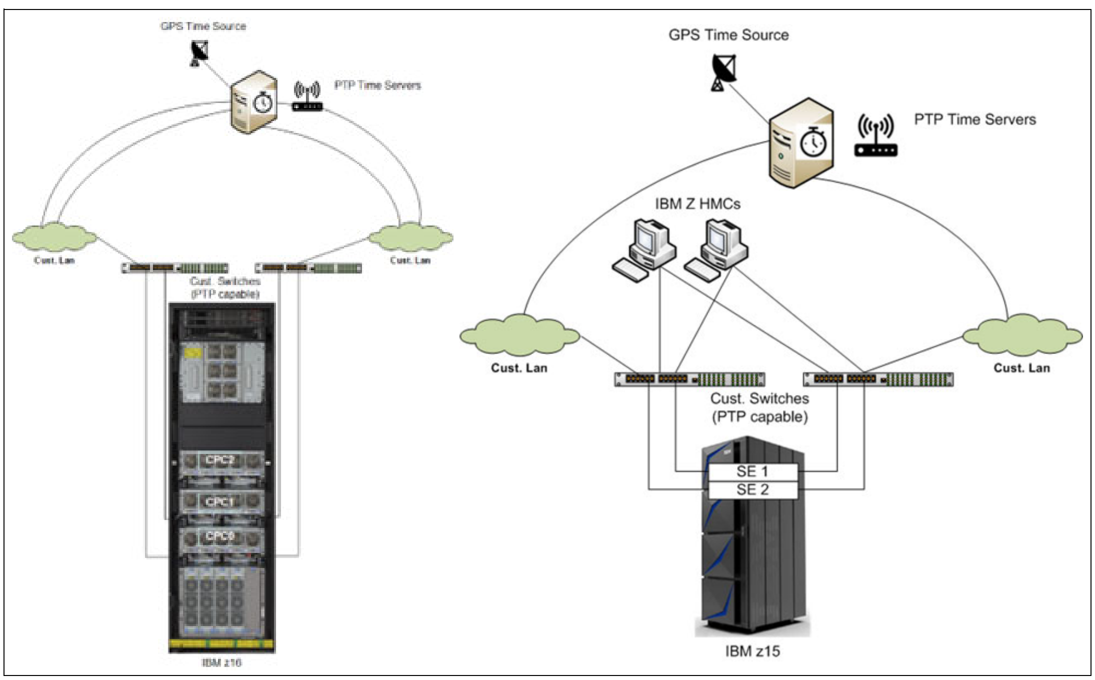


Figure 2-5 PTP connectivity diagram: IBM z16 (left) versus IBM z15

PTP server redundancy

The CTS is the only server that adjusts the CST by steering it to the time that is obtained from an ETS. To provide ETS redundancy, consider connecting two Ethernet connections to the PTP Grandmaster on the CPC drawer or the SE.

If there is a failure that results in PTP signals from the PTS or CTS to become unavailable, the CTN can use the adjustments that are calculated by the BTS to perform the necessary time adjustment steering for the entire CTN. The BTS could also be configured to use NTP, which might change the accuracy of the CTN, depending on the connection topology of the BTS.

For more information about PTP, see Appendix A, “IEEE 1588 Precision Time Protocol introduction and overview” on page 137.

2.3.2 NTP server

The usage of an NTP server as an ETS addresses the requirements of customers who need an accurate time source for the IBM Z servers or an accurate common time reference across heterogeneous platforms. These customers, in most cases, purchase an NTP server that obtains the time by using high-precision time technology, such as GPS signals.

Actual time accuracy, relative to Coordinated Universal Time (UTC), depends on how accurate the NTP server time is regarding UTC. Proper NTP server selection becomes a business decision rather than a technical requirement. STP provides the configuration and alerts based on defined NTP stratum information. Customers can specify an NTP stratum level that 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 that are organized in stratum levels. However, there are significant differences between NTP strata and STP strata, 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 more information about NTP, see the RFCs at the following web pages:

- NTPv3
- NTPv4

For more information about NTP and the NTP Public services project, see the following websites:

- Welcome to the home of the NTP Project
- Network Time Foundation’s NTP Support Wiki

NTP client support with direct CPC attachment

The NTP client on IBM z16 is in a firmware partition that is in the CPC. Additional Ethernet cable connectivity from the CPC drawer to the customer network with access to the configured NTP server provides a more accuracy compared to previous IBM Z server that are using the SE Interface.

NTP client support on the Support Element

The NTP client on IBM z15 and earlier servers can provide accuracy for the CST by accessing an NTP server. An accuracy of 100 milliseconds to the time that is 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 SE.

Planning for NTP client and LAN connectivity

The implementation of the SNTP client in the CPC drawer or on the SE uses the following components:

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

The NTP server can be either an ETS device that is available from several timekeeping device manufacturers, a local NTP server, or an NTP server that is configured on the HMC. The NTP traffic between the SNTP client and the NTP server is not encrypted.

An NTP server on the HMC

As a best practice, configure the NTP servers that are part of ETS so that they are attached directly to the CPC (IBM z16) or the SE (IBM z15 and earlier generations) LAN. These LAN connections are considered, in many configurations, to be a private (dedicated) LAN and should be kept as isolated as possible. Allowing the NTP client that is running on the CPC or 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 LAN is behind a firewall).

Note: For security reasons, we do *not* recommend configuring the SNTP client on the CPC or the SE to directly access NTP servers outside the SNTP client's LAN.

The ability to define an NTP server on the HMC addresses potential security concerns because the HMC is normally attached to both the CPC or SE network and intranet (see Figure 2-6).

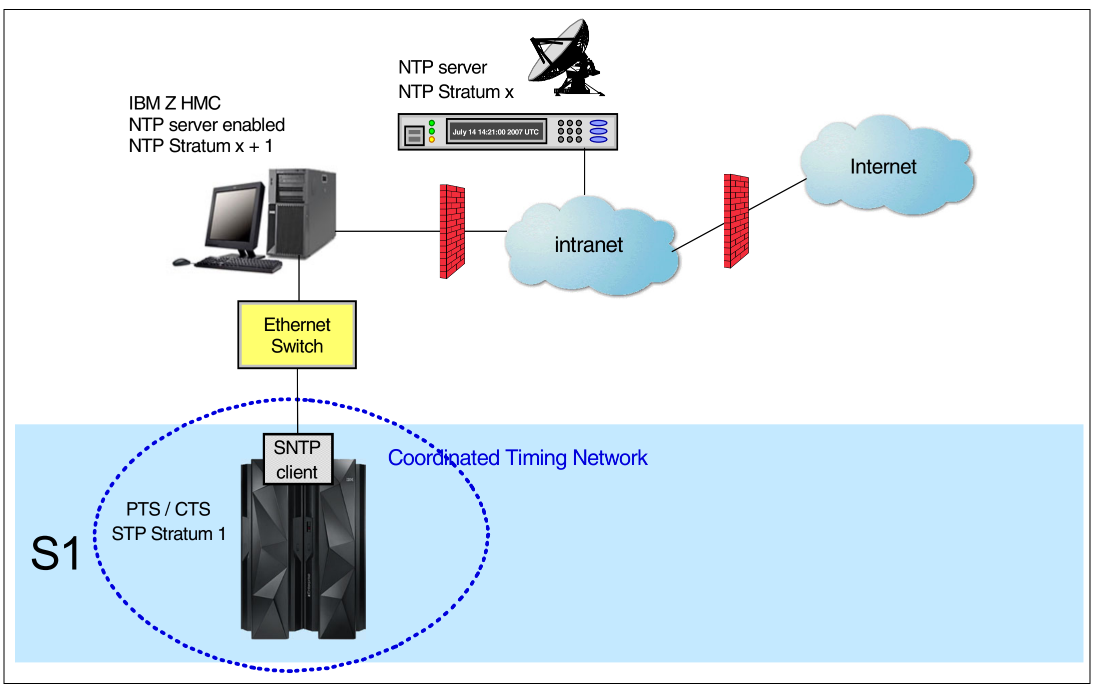


Figure 2-6 HMC acting as an NTP server

The HMC has two LAN ports that are physically isolated: one port for the connection to the HMC, SE, or CPC LAN that can provide an NTP server for the SE or CPC, and the second port for the LAN that is used by the NTP client on the HMC to retrieve time from another NTP time source. So, the NTP server on the HMC can access another NTP server through a separate LAN connection to obtain its time reference.

Note: The NTP server function on the HMC does not provide a PPS output.

HMC NTP broadband authentication support

You can use HMC NTP authentication as of HMC 2.12.0. To use this option on the CPC or 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 use this option:

- Authentication support with a proxy

  Some configurations use a proxy to gain 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 the 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 you use this configuration, you should use HMC authentication to ensure that the timestamps are not tampered with.

- Symmetric key authentication

  With the symmetric key and autokey authentication, the highest level of NTP security is available. HMC 2.12.0 and later provides windows that accept and generate key information to be configured in 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 that are encrypted with a secret key can be decrypted only with the same secret key. Symmetric key authentication supports network address translation (NAT).

  – NTP commands

    NTP command support was added so that you can display the status of remote NTP servers and the current NTP server (HMC).

Planning for an NTP server configuration

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

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

Note: The configuration of network components, such as routers or firewall rules, goes beyond the scope of this document. For more information about security guidelines and recommendations for the HMC or SE network, see *Hardware Management Console Operations Guide* and *Installation Manual for Physical Planning* at IBM Resource Link (select the manual specific to your server).

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

Figure 2-7 on page 41 shows an example that has an NTP server that is attached directly to the CPC or SE LAN. In this case, the NTP server is a single point of failure.

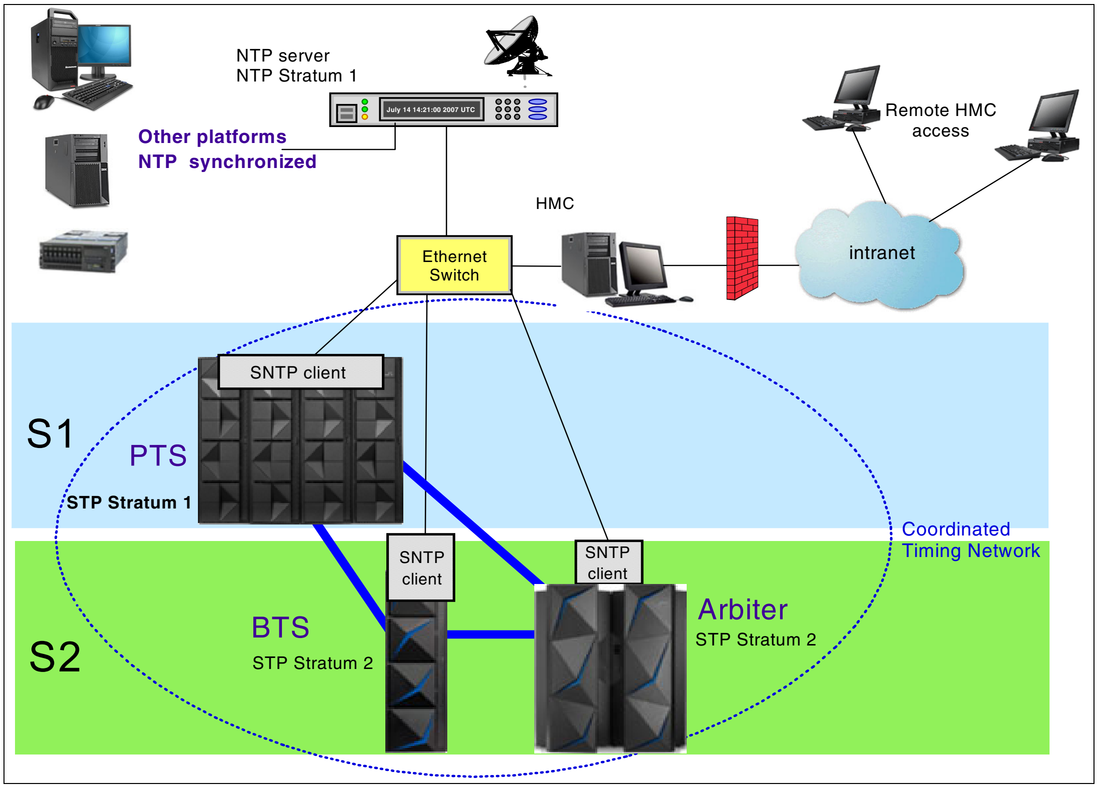


Figure 2-7 IBM Z and other platforms that are configured with an NTP server

One of the NTP output ports of the NTP server is connected directly to the HMC, SE, or CPC LAN, allowing the SNTP client on the CEC or the SE 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, more NTP servers can be installed and configured, as shown in “NTP server redundancy”.

NTP server redundancy

The CTS is the only server that adjusts the CST by steering it to the time that is obtained from an ETS. To provide ETS redundancy, the user should consider configuring two or more NTP servers.[4]

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.

[4] The CPC (IBM z16) or SE (IBM z15 and earlier) can handle up to two NTP servers. However, the HMC can configure 16 NTP servers. Three or more NTP servers are recommended for ETS redundancy for the CTN.

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 that is obtained from the selected NTP server has less accuracy than the non-selected NTP server (or if a discrepancy between the time information that is 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 or CTS role, STP automatically accesses the second NTP server that is configured on the PTS or CTS if the selected NTP server becomes unavailable.
- When NTP servers are configured on the server with the BTS role, the following items are 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 or CTS cannot access any of its NTP servers.

> **Best practice:** 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: 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, or CPC LAN, as in “Example: Single NTP server (typically from a timekeeping device manufacturer)” on page 40. However, this example also introduces using the HMC as an NTP server to provide NTP server redundancy.

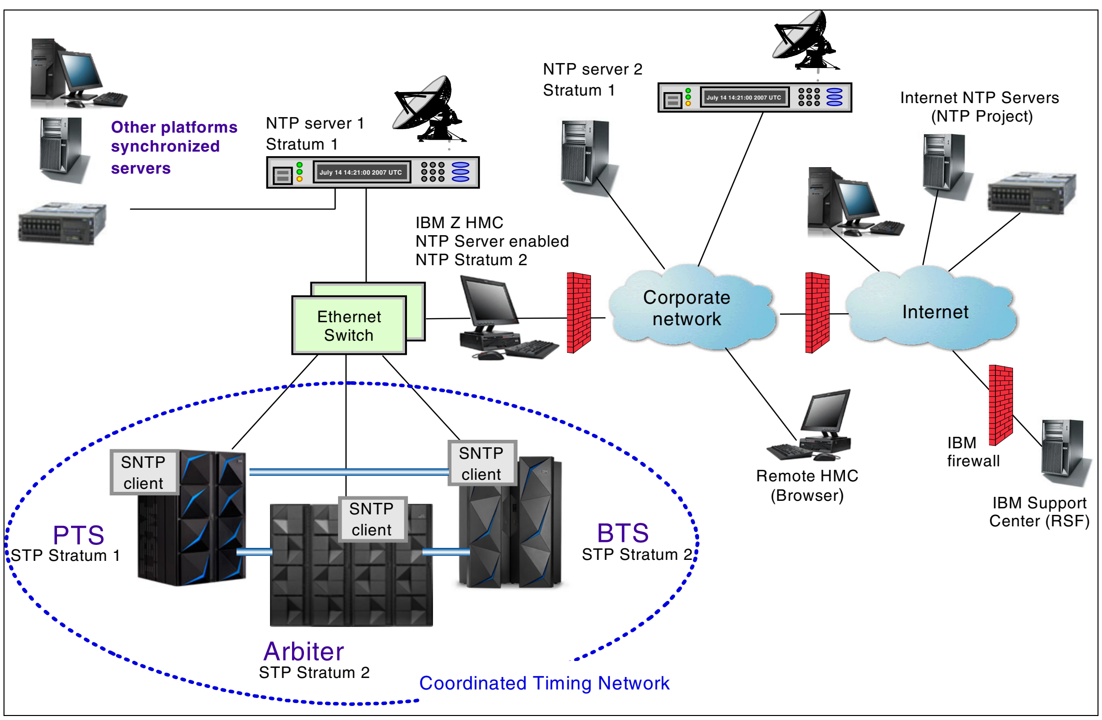


Figure 2-8 IBM Z and other platforms that are configured with NTP servers: Example 2

Because the HMC is normally connected to the SE or CPC LAN, the capability to configure an NTP server on the HMC should be considered. The HMC can connect by using a different network adapter to a remote NTP server that is on the corporate intranet or even on the internet to access its time reference (NTP server 2 in Figure 2-8 on page 42). Such a configuration requires that UDP port 123 is 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 like the HMC (with dual Ethernet adapters) and connected directly to the SE or CPC LAN. This configuration provides separation between the HMC, SE, or CPC LAN and the external networks.

This configuration offers redundancy because the SNTP client on the CEC or the SE allows the 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 “Example: Single NTP server (typically from a timekeeping device manufacturer)” on page 40, browser access to the HMC might be available by using the corporate network through the customer firewall.

Example: NTP server redundancy on PTS and BTS

Figure 2-9 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 “Example: Single NTP server (typically from a timekeeping device manufacturer)” on page 40 and “Example: Using the HMC to provide NTP server redundancy” on page 42 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.

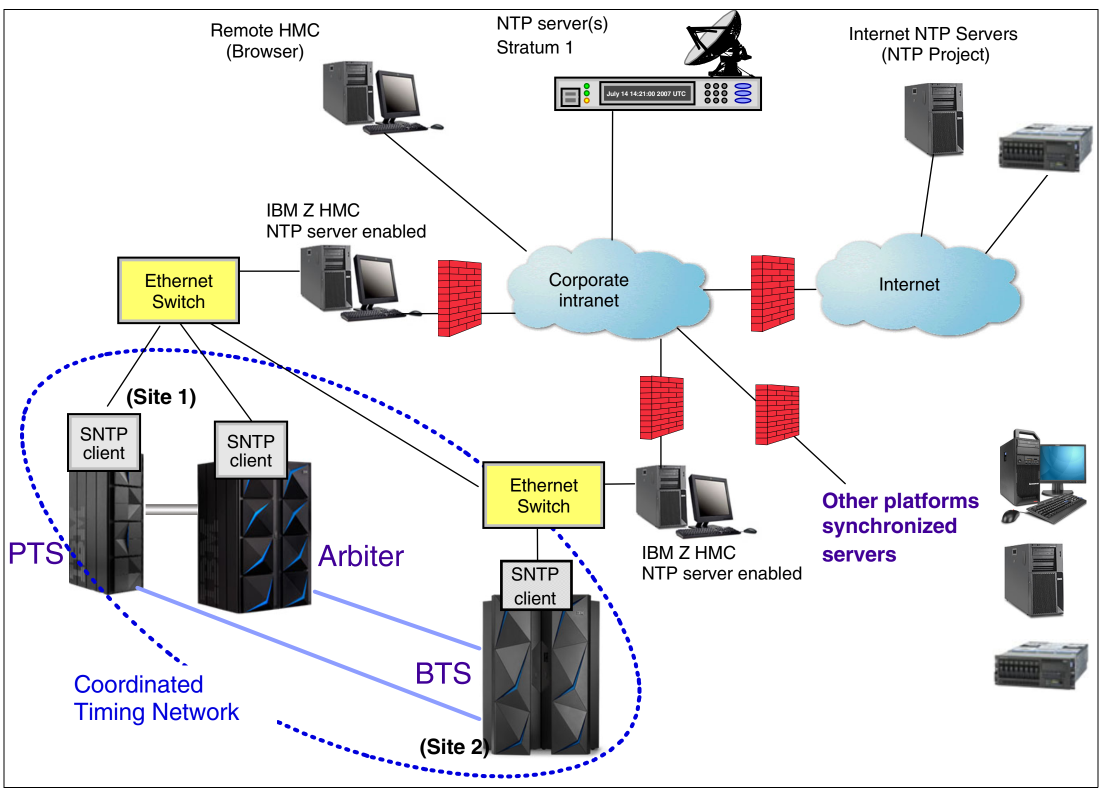


Figure 2-9 IBM Z and other platforms that are configured with NTP servers: Example 3

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

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

The examples that are 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 An NTP server with Pulse Per Second

The time accuracy of an STP-only CTN was 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.

PPS is an electrical signal that precisely indicates the start of a second. STP was 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. Several variables, such as the accuracy of the NTP server to its time source (GPS radio signals, for example) and the cable that is used to connect the PPS signal, determine the ultimate accuracy of STP relative to UTC.

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

Regulation changes that were announced by the Financial Industry Regulatory Authority (FINRA) reduced the allowable dispersion from universal time for systems processing certain financial trades (for more information, see FINRA Regulatory Notice 16-23. PPS is a standard, together with NTP, that provides precise notifications to connected systems if the PPS provider is attached to a compliant time source.

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

Leap seconds are used to synchronize UTC to International Atomic Time (TAI). When a leap second is inserted into UTC (when a leap second is needed, it is added at midnight UTC on either 30 June or 31 December), UTC effectively jumps backward by 1 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 automatically steers to the correct time, but this steering adjustment is made at the rate of 7 hours per second to ensure that no duplicate timestamps are seen.

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 through the STP Adjust Leap Seconds window (see 3.2, “Configuring the Coordinated Timing Network” on page 78). 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 7-hour window that is required for steering to complete. They see message IEA032E, which is issued by the z/OS accuracy monitor, at 60-minute intervals 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 was introduced in 1961, UTC in 1972, UTC 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 timestamps. A spin of 27 seconds at a time will likely cause a system outage, so leap seconds must 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 approach ensures time accuracy across future leap seconds without the problems of getting from a current value of `0` to `27`.

For more information, see the following resources:

- The FINRA requirements that are described in 1.1.3, “Industry requirements” on page 4.
- The `ACCURACY` statement in the CLOCKxx parmlib member in “ACCURACY mmmmm” on page 55.
- *STP recommendations for the FINRA clock synchronization requirements*.

Planning for an NTP server with PPS

To configure an NTP server with PPS, the NTP output of the NTP server must be connected to the CPC or SE LAN, and the PPS output of the same NTP server must be connected to the PPS input that is provided on the OSC of the IBM Z server.

The NTP information is provided through the HMC, SE, and CPC LAN. However, the NTP server propagates its PPS signal through 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 is on the OSC.

At a minimum, one NTP server with a PPS output should be configured at the server that has the CTS role.

An NTP server with PPS redundancy

Each IBM Z server is equipped with two OSC[5] cards, which provide the capability of attaching one or two NTP servers with PPS. Attaching a second NTP server with PPS 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 that is 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 that is obtained from the selected NTP server with PPS has less accuracy than the non-selected NTP server.
- The two NTP servers provide 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 or CTS role, STP automatically accesses the PPS signal from the non-selected NTP server that is configured on the PTS or CTS if the PPS signal from the selected NTP server fails.

  Figure 2-10 on page 47 is an example of this redundancy.

[5] Oscillator (clock) cards. Depending on the server hardware, each server (z13 and later) has redundant OSC cards with PPS input.

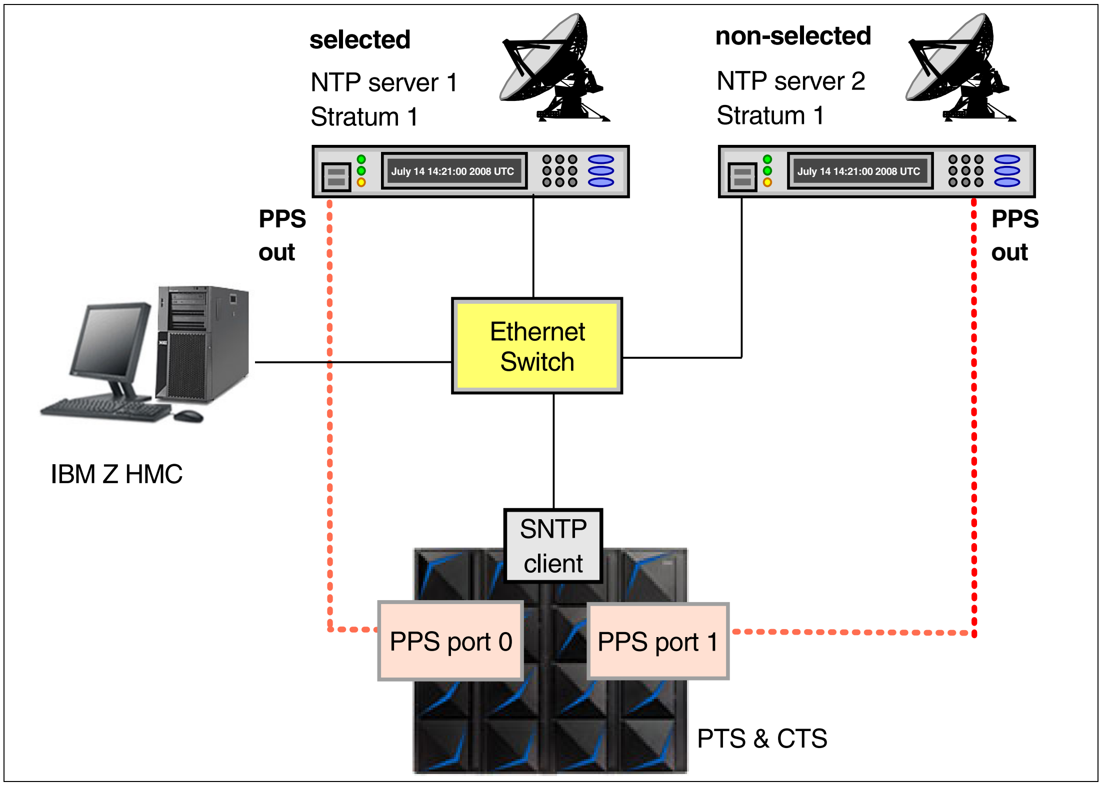


Figure 2-10 IBM Z with NTP PPS

- When NTP servers with PPS are configured on the server with the BTS role, the following items are 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 or CTS does not receive PPS signals from all of its configured NTP servers with PPS.

Figure 2-11 illustrates this type of redundancy.

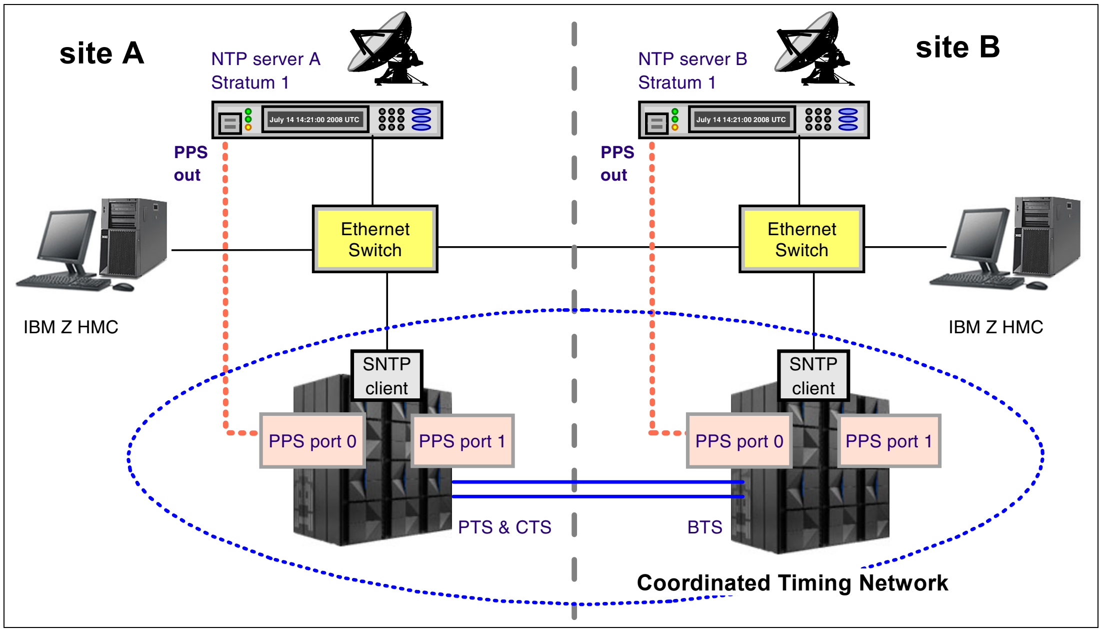


Figure 2-11 NTP servers with PPS on PTS and BTS

> **Best practice:** 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 are configured on PTS or CTS

In this example, two NTP servers with PPS are configured on the PTS or CTS (Figure 2-10 on page 47). The NTP outputs of both NTP servers are connected to the CPC or SE LAN through 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 or CTS is tracking the PPS signal that is received on PPS port 0. If there is a failure that is associated with this PPS signal, STP switches to the PPS signal that is received on PPS port 1 and can continue tracking an ETS to maintain time accuracy.

Example: Different NTP servers with PPS on PTS and BTS

The configuration that is shown in Figure 2-11 allows continuous availability of ETS by configuring a different NTP server at site 1 (NTP server A on the PTS or 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.

During normal operation, the CTN tracks the PPS signals that are received on PPS port 0 of the PTS or 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 or CTS at 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 that is 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 or CTS, the server is able to use the adjustments that are calculated by the BTS to perform the necessary time adjustment steering for the entire CTN.

2.4 IBM z16 power failure handling

A new automated failover option is available starting with IBM z16. As most client data centers are equipped with an uninterruptible power supply (UPS), complete power loss situations are unlikely to occur. The IBM Z focus has moved toward assisting clients by leveraging common power events across their ecosystems. IBM z16 can take automated recovery actions after N-mode is detected on an IBM z16 CPC.

> **Important:** N-mode power provides automated CTN recovery if both PTS and BTS are IBM z16 and N-mode power is enabled on both CPCs.

N-mode power detection is available for both Bulk Power Assembly (BPA) and intelligent Power Distribution Unit (iPDU). N-mode power detection covers a broader scope of power events, such as single wall power loss events and power cord loss events.

This function is a one-time setup, but it can be changed at any time by using the Advance Actions section in the **Manage System Time** tab. After N-mode power detection is enabled, STP can act if any power source detects a power cord or power side failure on the CTS. If N-mode power detection is enabled, any power cord or power side failure on the PTS or CTS sends a signal to the BTS to indicate that an N-mode power event happened on the CTS. If the BTS does not receive a message back within 30 seconds that indicates that the PTS is again fully redundant, the BTS takes over the CTS role. The PTS still is a member of the CTN and remains as the PTS, but it moves into STP stratum level 2.

After the power recovers and the PTS is back to being fully redundant, the CTS role can automatically be returned to the PTS. The timing for the switchback of the CTS role is configured in the **Set time server power failover** tab.

A sample power connectivity configuration for IBM z16 is shown in Figure 2-12.

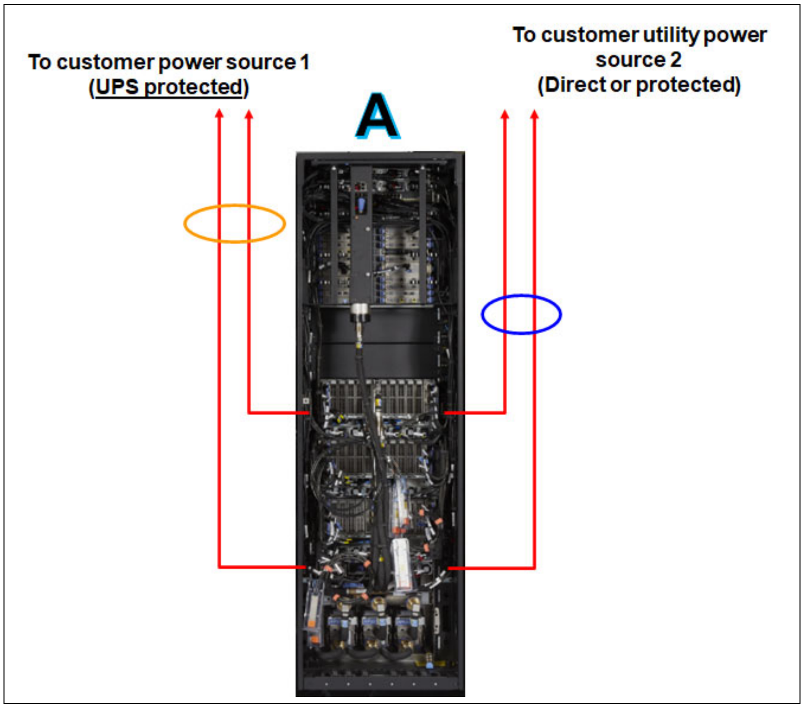


Figure 2-12   IBM z16 power connectivity for enabling N-mode power protection (PDU system)

The configuration of the **Set time server power failover** tab contains the following parameters:

- A checkmark for an automated switch of the CTS to the BTS if there is a power fault detection at the PTS.
- The amount of delay for the switchback of the CTS to the PTS if it is fully redundant.
- Optionally, if an automated switchover occurred, the user can switch the CTS to the PTS immediately. This function is available only if there was a prior automated switch of the CTS to the BTS.

Considerations should be made when moving to a z16 environment that results in a mixed configuration of IBM z16 and earlier generations. Because time server power failover is available only on an IBM z16 server, individual considerations should be made regarding the role assignments. In addition to the global considerations that are described in 2.9, “Migration planning” on page 64, other aspects for the assignment of roles are as follows:

- Assigning the PTS and BTS to z16 CPCs to use the time server power failover concept
- Assigning the PTS role to a server that supports the Internal Battery Feature (IBF) feature

2.5 Internal Battery Feature

Note: IBF is not available for IBM z16.

In some models of IBM Z servers that are equipped with BPAs, which are IBM z13, IBM z13s, IBM z14 M0x, and some models of IBM z15, an IBF can be installed. IBM Z servers that are equipped with a Power Distribution Units (PDUs) (IBM z14 ZR1 and some models of IBM z15) do not support the IBF.

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

The main purpose of an IBF is to keep a IBM Z server running for several minutes when a power outage to the data center occurs *and there is no UPS that is 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 the last IBM Z server to offer an IBF.

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

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

Important: If the CTN in a single data center has three or more servers, as a best practice, assign the Arbiter role, in which case the IBF does not provide any extra benefit for server power outages.

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

2.6 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 as follows:

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

2.6.1 z/OS requirements for STP

STP has been supported by z/OS since z/OS V1R7. Now, all z/OS images, either running in a CTN or unsynchronized, include STP support.

Run an IPL of the z/OS image in the following circumstances:

- After applying z/OS related maintenance (as required).
- After changing the CLOCKxx parameters for systems running with a simulated Sysplex Timer identifier (SIMETRID *nn*).

For a system with `STPMODE YES` and `PLEXCFG=MONOPLEX`, the system switches from LOCAL timing mode to STP timing mode when the CPC is STP-synchronized.

z/OS maintenance

Even though all supported releases of z/OS support STP, more 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 CFs. To simplify the identification of PTFs for STP, a functional PSP bucket was created. Use the Enhanced Preventative Service Planning Tool (EPSPT) and your usual maintenance procedures. For more information about SMP/E commands' syntax or on receiving and applying SYSMODs, see *SMP/E for z/OS User's Guide*, SA23-2270 or *SMP/E for z/OS Reference*, SA23-2276.

2.6.2 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 CLOCK*xx*, where *xx* is a 2-character suffix. The CLOCKxx member performs the following functions:

- Provides the means to specify that z/OS should use STP and its time zone parameters. The statements **`STPMODE`**, **`STPZONE`**, and **`TIMEZONE`** are used.
- Provides the means of specifying that z/OS should use the Sysplex Timer[6] and its time zone parameters (**`TIMEZONE`**).
- Allows you to specify the difference between the local time and UTC if you do not want to use the time zone that is used by ETR or STP. The statements **`ETRDELTA`** and **`TIMEDELTA`** are used.
- Specifies how much time deviation for the time-of-day (TOD) clock from the ETS is acceptable. The statements **`ACCURACY`** and **`ACCMONINTV`** are used.

[6] The sysplex timer is not available for z196 and later generations. Also, zEC12 and zBC12 are the last server generation to support Mixed CTN. z13 and later support STP-only CTN.

For STP support, the following CLOCKxx statements were added in z/OS in z/OS 1.7:

- **STPMODE**
- **STPZONE**
- **TIMEDELTA**

An ETR is a 9037 Sysplex Timer. Although 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.

> **Notes:** If the z/OS system image is running on a server that is in ETR timing mode and both **STPMODE** and **ETRMODE** are 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** are 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, `ETRMODE YES` and `STPMODE YES` were specified before 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 was completed and the Sysplex Timers were removed.
>
> Now, on IBM Z servers, which do not support any ETR connectivity and cannot run in ETR mode, it does not matter if the **ETRMODE** statement is specified if `STPMODE YES` is specified.
>
> For this reason, statements regarding ETR are not described in detail in this publication.

In a multi-system environment, the objective is to share one SYS1.Parmlib concatenation with as many z/OS images as possible. This approach 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.

The syntax format of CLOCKxx

Example 2-1 shows the CLOCKxx parameters format.

Example 2-1 The syntax format of the 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 mmmmm
ACCMONINTV {mmss | mss}
```

IBM supplied default for CLOCKxx

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

Example 2-2 Default parameters and values for CLOCK00

```
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 *MVS Initialization and Tuning Reference*, SA23-1380.

OPERATOR {PROMPT | NOPROMPT}

This statement specifies whether the operator should 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 keywords are specified, the system image rejects the CLOCKxx member during system image initialization and issues a message to prompt the operator for a valid CLOCKxx member, or to use the default values (by pressing the Enter key). Otherwise, the operator must perform an IPL of the system image.

TIMEZONE d.hh.mm.ss

This statement specifies the difference between the local time and the UTC. If **STPMODE YES** and **STPZONE YES** are specified and the server is being synchronized by 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 statement specifies whether z/OS should use ETR timing mode. Because IBM Z servers no longer support ETR mode, 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** are specified as `YES`, z/OS uses **STPMODE YES**.

ETRDELTA nn

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

ETRZONE {YES | NO}

This statement is the ETR version of the **STPZONE** statement. **ETRZONE** should be replaced by **STPZONE**.

SIMETRID nn

This statement specifies the simulated Sysplex Timer identifier. **SIMETRID** allows z/OS images running on the same server, in native mode in LPARs, 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 Processor Resource/Systems Manager (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 statement specifies whether z/OS should use STP timing mode. The default value is `YES`, which specifies that z/OS should use STP timing mode if STP is configured. When `NO` is specified, the local z/OS timing mode is used.

TIMEDELTA nn

This statement indicates the greatest difference between the system's TOD and the CST (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). **TIMEDELTA** is used at IPL for the sync clock and also when a sync check occurs.

> **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 CST exceeds **TIMEDELTA**, the following results occur:

- If the system is in STP synchronization mode and part of a multisystem sysplex, the WTOR message IEA394A is issued.
- If the system is not part of a multisystem sysplex, processing continues, and the system switches to the LOCAL timing mode.

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 synchronization check is recognized in the `TIMEDELTA 0` case, one of the two previously described results.

STPZONE {YES | NO}

This statement specifies whether the system image should 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 **STPMODE** is set to `YES`.

ACCURACY mmmmm

The z/OS TOD clock accuracy monitor was introduced with z/OS 2.1 to monitor the accuracy to which a CTN is synchronized with an 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 ETS. The accuracy monitor checks the deviation and issues a message if the deviation exceeds the specified **ACCURACY** threshold.

The z/OS TOD clock accuracy monitor is enabled by specifying a nonzero value 0 - 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 nonzero value is specified for the **ACCURACY** parameter for a system running on an STP network that is synchronizing to an ETS, the **ACCURACY** function is activated, which results in the following IPL-time message:

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

Then, the monitor checks 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:** As a best practice, see **ACCURACY** to `20` for customers that must meet the FINRA clock synchronization requirement. This setting causes IEA032E to be issued when the steering adjustment exceeds 20 milliseconds.
>
> To help customers meet the FINRA clock synchronization audit requirement, a new function was added to the z/OS accuracy monitor by APAR OA51786. A message is issued at a user-specified interval, with a default of every 60 minutes, giving details of the status from the PRT correction steering information block.
>
> For more information about FINRA requirements, see 1.1.3, "Industry requirements" on page 4 and 2.3.3, "An NTP server with Pulse Per Second" on page 44.

ACCMONINTV mmss | mss

This statement sets a timing interval when the **ACCURACY** function is 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 they can be different. If the server is in ETR timing mode, a z/OS image might be running in local, ETR, or SIMETR mode. If the server is in STP timing mode, a z/OS image might 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`

  > **Best practice: STPMODE YES** and **STPZONE YES** are the defaults with z/OS. Leaving the STP settings as the default gives the most configuration flexibility.

- If both **ETRDELTA** and **TIMEDELTA** are specified in the same CLOCKxx member, z/OS uses the second entry, and flags 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**. Then, the system image uses the Sysplex Timer or STP timing signals to synchronize itself with the other members of the sysplex. The system image uses a synchronized timestamp 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
```

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

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

2.7.1 z/VM

z/VM uses the STP facility to generate timestamps for guest and system DASD write I/O operations so that these I/O operations are synchronized with the I/O operations of other systems. This support allows data that is 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 (XRC)). For example, this configuration allows z/VM to participate in a GDPS Metro z/OS Global Mirror (GDPS/MzGM) environment.

This baseline STP support is available in z/VM since Version 6.2[7] and provides the following functions:

- Synchronizes the z/VM TOD clock with the STP server at IPL. Before the introduction of baseline STP support, z/VM, under operator control, changed the TOD clock during IPL through the operator prompt for date and time. The operator entered `YES` to the Change TOD clock {YES|NO} prompt during IPL. With STP support enabled in z/VM, no prompts are received to change the TOD clock; instead, STP automatically initializes the TOD clock with the CST and sets the same time zone that is used by the CTN, and displays 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.
- The CPC must be either a member of an STP-only CTN, or a stratum 2 or higher member of a Mixed CTN.

[7] At the time of this writing, z/VM 7.1 and later are the z/VM versions in support. Older z/VM versions (z/VM 6.4 and older) are not supported.

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

- **STP_Timestamping**: Timestamps are added to write channel programs that are issued to all DASD devices that have the XRC LIC installed.
- **STP_TIMEZone/STP_TZ**: The system time zone is derived from the STP server.
- **XRC_OPTional**: The 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 second-level z/VM. This statement enables **STP_Timestamping** without STP availability. A manually specified TOD value is used for timestamping, and is intended for vendor test support.

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_Cmds ,            /* 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 information about the **FEATURES** statements, see *z/VM: CP Planning and Administration*, SC24-6271.

STP state flow in z/VM

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

- During IPL:
  - If **STP_Timestamping**, **STP_TZ**, or both are specified: Perform an 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 approaches fails, STP enters the **SUSPENDED** state. Otherwise, STP activation completes successfully and STP is considered ACTIVE.

- When STP is `ACTIVE`.

  If **`STP_Timestamping`** is specified, the 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: The I/O to XRC-capable DASD that must be timestamped will be deferred until STP becomes `ACTIVE` again. The I/O that does not need to be timestamped, and I/O to the non-XRC-capable DASD 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 the Control Program (CP) that the TOD value must be synchronized.
- From `SUSPENDED` to `ACTIVE`: In response to machine checks or external interrupts that are received, the CP attempts to resynchronize 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 the system time zone (if **`STP_TZ`** is enabled).
- From `ACTIVE` to `ACTIVE`: External interrupts might be received that require the CP to query the STP time zone information (for example, the time zone was changed through the HMC). This action does not cause STP to enter the `SUSPENDED` state, but it causes the system time zone to change.

CP commands

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

- **`QUERY STP`**: This command displays the STP 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 changes the system's time zone ID and time zone offset. It was updated to display the response that is 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 displays the list of active and inactive time zone definitions on your system. It was updated to show the STP time zones when z/VM is configured to use it.

For a full description of possible responses of these CP commands, see *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 LPARs (not z/OS guest machines), z/VM can be used to create timing-only links. To define timing-only links, define the peer channel (CFP or CBP) with a control unit of `UNIT=STP`, and no devices. For more information, see 2.2.2, “Planning for timing-only links” on page 34.

VM guests

A z/OS sysplex can be established among z/OS images running as guests under z/VM, but the guests still cannot leverage 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 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 by 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 the TODs for the system are synchronized, and if **`TODENABLE`** is specified, the operator can change the guest TOD clock later.

The three z/OS guest images in Figure 2-13 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 that is set by the LPAR during the LPAR activation, either by systems that underwent IPL in the partition or z/VM during its initialization.

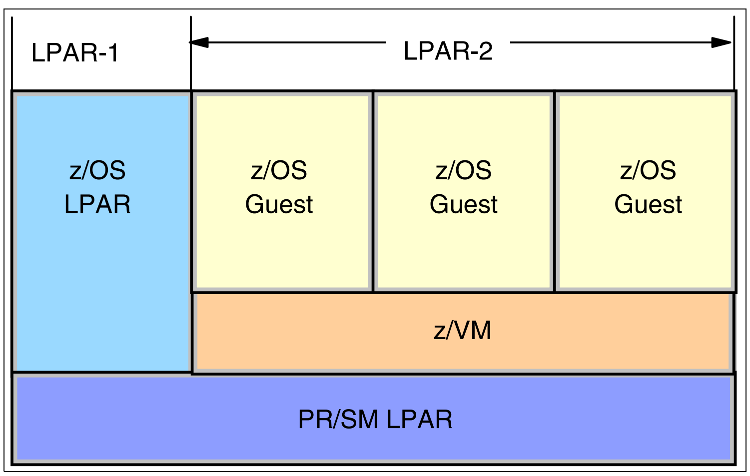


Figure 2-13 A sysplex within a logical partition and VM guests

2.7.2 z/TPF

STP is supported in z/TPF with APAR PJ36831. PJ36831 was released on z/Transaction Processing Facility Product Update (z/TPF PUT) 07.[8] For more information about STP support in z/TPF, see Accessing the TPF Product Information Center.

For more information about z/TPF, see the TPF website or contact tpfqa@us.ibm.com.

[8] z/TPF PUT is a maintenance update that provides small programming enhancements (SPEs) and maintenance updates for the z/TPF.

2.7.3 Linux on IBM Z

Your Linux instance might be part of an XRC setup that requires synchronization of the Linux TOD clock with a timing network.

Linux on IBM Z supports STP-based TOD synchronization. This support can be enabled at the following times:

- At boot time by passing the **`stp=on`** parameter to the Linux kernel.
- St run time by modifying **`sysfs`**:

  At run time, TOD synchronization that uses STP can be turned on:

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

  It can also be turned off:

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

For more information, consult your Linux on IBM Z distribution documentation.

2.8 Special considerations for a single-server CTN

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

- You must coordinate time accuracy across IBM Z and other platforms of your enterprise. As describe in 2.3, "External Time Source" on page 36, you can use the ETS option of using a PTP or an NTP server to accomplish this task.
- You have a single server and want to implement a sysplex that is made up of multiple members. Even though this task can be accomplished without using STP and coding **`SIMETRID`** *nn* for the CLOCKxx member of each z/OS image, there are reasons that you might want to implement the sysplex by using STP and coding **`STPMODE YES`** in the CLOCKxx member.

  You might have both z/OS systems and other OSs that do not support STP running on the single server. In the following text, OS-xyz applies to any OS that does not support STP, for example, IBM z/VSE®, and 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 that is provided by using the ETS function. All LPARs (OS-xyz included) can still maintain time within 100 ms of the ETS.

  There are a few instances that cause a larger difference between the CEC time and the ETS time that OS-xyz running on an 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 more information, see STP - when the operating system doesn't support it!

  Note: The sync check limitation that is described in see STP - when the operating system doesn't support it does not apply to a single-server CTN because a stratum 1 server (PTS or CTS) is the source of time and cannot obtain sync checks.

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

2.8.1 Hardware and ETS planning for single-server configuration

Hardware planning includes installing an IBM Z server. There are no considerations for coupling links because no external timing links are required.

For more information about planning your ETS configuration, see 2.3, “External Time Source” on page 36.

Note: Because you are implementing a single-server CTN, your server will be the PTS and CTS. There is no BTS or Arbiter in this configuration. However, all rules for providing ETS redundancy to the PTS and CTS apply here.

As a best practice, save the STP configuration by restricting CTN membership, as shown in Figure 2-14 on page 62. This approach avoids z/OS issuing `IEA389I THIS STP NETWORK HAS NO SERVER TO ACT AS BACKUP/ARBITER` at IPL time and whenever a link availability change or configuration change occurs

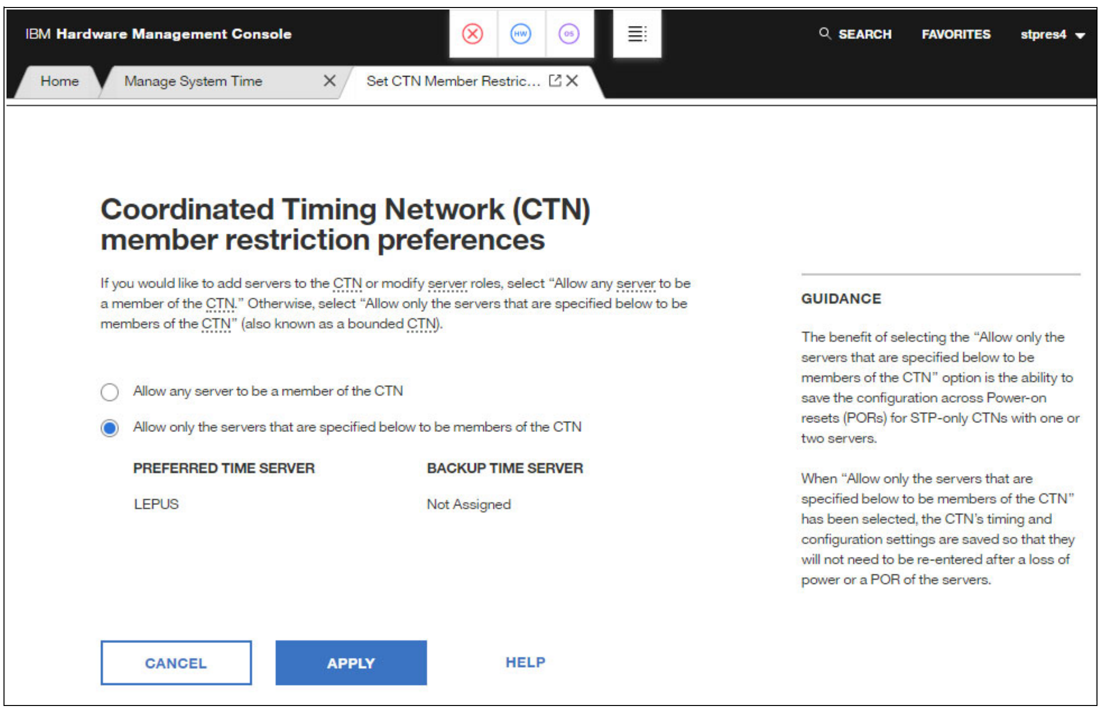


Figure 2-14 Enabling a configuration save for a single-CPC CTN

Note: After a power off/on sequence, the TOD value in the CTS is initialized from the TOD value that is stored in its SE. If you previously configured an ETS, the STP code does not perform an *initialize time* by using the ETS, but automatically accesses the configured ETS and starts adjusting the time to maintain time accuracy. STP allows time adjustments of up to +/- 60 seconds.

2.8.2 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 before configuring a CTN.

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

Continuing to run with **SIMETRID** allows you to configure the CTN concurrently without stopping production because in **SIMETRID** simulation mode, STP interrupts and machine checks are not applicable. Continuing to run on a CPC that is configured for STP exposes problems due to unprocessed sync checks. For more information, see Implementing and using SIMETR with STP.

For more information about configuring the CTN, see 3.2.1, “Single-server CTN” on page 78. When the procedure reaches the initialize time step for a new CTN, the TOD value can be changed because 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 the PR/SM code. There are no WTOR messages that are posted to the z/OS Console.

As described in 2.6, “Planning z/OS software” on page 52, if STP is available when z/OS is not running as a VM guest, use appropriate CLOCKxx statements instead of **SIMETRID** because of the following advantages.

Note: When a sync check occurs, LPAR takes an offset equal to the sync check and generates an STP machine check. z/OS systems running in STP timing mode process the sync check by clearing the offset.

However, **SIMETRID** systems are not enabled for STP sync checks and do not get an STP sync check machine check. The LPAR will have an offset (an EPOCH) that will not be cleared. The logical TOD for the **SIMETRID** LPARs drifts away from the STP time of the CPC as each sync check occurs.

With currently supported z/OS versions, the LPAR offset is not cleared at IPL, so there is no mismatch when systems undergo an IPL due to the LPAR EPOCH. However, when an LPAR is reactivated, the LPAR EPOCH is cleared. Now, when z/OS undergoes IPL in that reactivated LPAR, it cannot join the sysplex due to the LPAR EPOCH mismatch because the running system has a nonzero LPAR EPOCH, and the system undergoing IPL has a zero EPOCH. If one LPAR in the sysplex must be reactivated, then all LPARs in the sysplex must be reactivated to ensure that there is no mismatch of LPAR EPOCHs.

- If you update the CLOCKxx member to use **STPMODE YES** and **STPZONE YES**, then DST changes can be handled without using manual procedures that you might be using to change the time zone, such as using the **SET TIMEZONE** command.
- z/OS messages are posted to the z/OS console when z/OS detects certain conditions with STP, such as messages that are related to ETS.
- If leap seconds are important to you, z/OS recognizes them when you specify leap seconds during the initialize time procedure, and later when leap seconds are adjusted. For more information, see “Leap second offset adjustment” on page 104.

Attention: To change from z/OS systems running with `SIMETRID` to `STPMODE YES`, you must plan to an outage of the z/OS systems to activate the STP mode. After the CTN is configured, you can shut down and perform an IPL of the z/OS systems by using the updated CLOCKxx member.

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

In this section, the assumption is that you have running z/OS systems that are not in a sysplex before you configure a CTN.

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

After the CTN is configured (see 3.2, "Configuring the Coordinated Timing Network" on page 78), for non-sysplex systems, the system switches dynamically to STP if `STPMODE=YES` is specified.

Note: When the z/OS systems are running in `STPMODE YES`, follow the procedures to set up a sysplex.

2.9 Migration planning

Over time, structural changes must be performed on an existing CTN. The server technology changes; new servers must be added to the CTN or others must be removed; and the coupling connectivity technology changes. In this section, several aspects are described that must be accounted for when migrations to the server and connectivity hardware are performed.

2.9.1 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 1.3.5, "Server roles in a CTN" on page 13. These roles are essential for ensuring the availability of the time source in the CTN because the servers with the necessary roles vote for which server takes over the role of the CTS in a recovery scenario.

- Only one server in the CTN (single-server CTN)

  In a single-server CTN, there is only one server that is available with 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** must be selected in the CTN member restriction preferences to avoid losing the CTN definitions after the POR of the server. This process is described in 2.1.3, "Connectivity to ETS" on page 27.

- 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 takes over the CTS task only 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 there 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 vote with the Arbiter is performed to determine whether the PTS or BTS can take on the role of the CTS. This recovery is called Arbiter Assisted Recovery (Console Assisted Recovery (CAR) is not used).

With the IBM zEnterprise® 196 GA2 announcement, recovery voting for an STP in a CTN with three or more servers with the roles of PTS, BTS, and Arbiter assigned is enhanced:

- A server obtaining a role in a CTN cannot accidentally be removed from the CTN in a planned disruptive action anymore, unless its role has been shifted to another server. If a planned disruptive action is attempted on the PTS, BTS, or Arbiter, the 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 block prevents you 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 occurs because of an unplanned outage, which means that when any two of the special role servers (PTS, BTS, or Arbiter) agree that they cannot communicate with the third special role server, normal voting is disabled. Therefore, the recovery mechanisms that are used in a two-server CTN can be used in this situation and when the connection between the two surviving servers with roles breaks.

For more information about recovery voting, see *STP Enhancement - Recovery Voting*.

2.9.2 Server maintenance and replacement

As described in 2.9.1, “CTN types and recovery voting” on page 64, the PTS, BTS, and Arbiter roles should be assigned when three or more servers are available. This role assignment should also be used when disruptive maintenance to a server is performed or when a server is being replaced.

Since 2012 (with z196 GA2), an active CTN role must be shifted to another server before the server with the assigned role 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 must be altered. If the relocation of the server role has not been performed before the planned disruptive action, the SE code blocks the action.

As a best practice, shift the CTN roles between servers before taking any disruptive actions against any server holding a CTN role.

2.9.3 The N-2 rule and Coupling Link migration considerations

As described in 2.1, “Planning server hardware” on page 22, there is an N-2 connectivity and coexistence rule for servers running in the same sysplex that is also applicable to servers running in the same CTN. If an IBM z16 server is present in the CTN, only the following servers are allowed in the same CTN:

- IBM z16 A01
- IBM z15 T01
- IBM z15 T02
- IBM z14 M0x
- IBM z14 ZR1

IBM z16, IBM z15, and IBM z14 ZR1 support only the following coupling connectivity:

- ICA SR links
- CE LR links

All servers that connect to an IBM z16, IBM z15, or IBM z14 ZR1 must use these two coupling connectivity options. You may also run IBM z14 M0x servers that have InfiniBand links installed, but there cannot be any direct InfiniBand coupling connectivity between these servers and IBM z16, IBM z15, and IBM z14 ZR1.

> **Attention:** Using InfiniBand links between IBM z14 M0x servers in a configuration with IBM z16 or IBM z15 must be carefully planned to avoid single points of failure.

As a best practice, use this CTN layout only for short-term migration scenarios. For permanent usage, all servers in the CTN should be equipped with the same coupling link technology, that is, ICA and CE LR links.

When migrating from HCA3-O LR links to CE LR links, the same fiber cabling infrastructure can be used, but CE LR cannot connect to HCA3-O LR, and CE LR does not support the STP GAS.

> **Note:** IBM z16 introduces the new Coupling Express2 LR features (Feature Code 0434). These features are link-compatible with previous CE LR features (Feature Code 0433).

The GOSIG is a reliable, unambiguous signal that is sent 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 is sent by the CTS and received by the BTS, the BTS can safely take over as the CTS without relying on the previously mentioned recovery methods, that is, an offline signal in a two-server CTN, or the Arbiter voting in a CTN with three or more servers. However, the GAS can be sent only by coupling links that are installed directly on the CPC drawer fanouts. Because the CE LR card is in the I/O drawer, the GAS is not supported by CE LR.

Conversely, coupling links that are installed directly on the CPC drawer fanouts go offline if the CPC drawer is removed or maintained. Because CE LR cards are in the I/O drawers, they benefit from the EDA and the 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 information about coupling link connectivity options, see *IBM z16 (3931) Technical Guide*, SG24-8951, and *IBM z15 (8561) Technical Guide*, SG24-8851.

2.9.4 Power option considerations: PDU, BPA, or IBF

As described in 2.1, “Planning server hardware” on page 22, the IBM z16 and IBM z15 servers are offered with two power options: PDU or BPA. IBM z14 ZR1 servers are available only with the PDU power option. A server with the PDU power option does not support the IBF, as described in 2.5, “Internal Battery Feature” on page 51.

Note: IBM z16 servers do not support the IBF feature. N-mode power is provided for handling CTN reconfiguration for CPC power failures. For more information, see 2.4, “IBM z16 power failure handling” on page 49.

With an IBF that is installed in the CTS, STP can receive notifications 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 was using this signal for any recovery automation purposes, you must plan for this situation before assigning any active CTN role to a server with a PDU (or without the IBF).

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

3

Configuration and operations

This chapter describes considerations for planning the operational management of the Coordinated Timing Network (CTN).

The following topics are covered in this chapter:

- Displaying and monitoring the environment
- Configuring the Coordinated Timing Network
- Configuring Precision Time Protocol as ETS
- Time server power failure failover (IBM z16 only)
- Managing the time

3.1 Displaying and monitoring the environment

This section describes the facilities that are available to monitor your Server Time Protocol (STP) CTN.

3.1.1 CPC Details window

The CPC Details window on the Hardware Management Console (HMC) is the first window to look at for timing information about a server. Select the server that you want to work with and then click **System Details**, as shown in Figure 3-1.

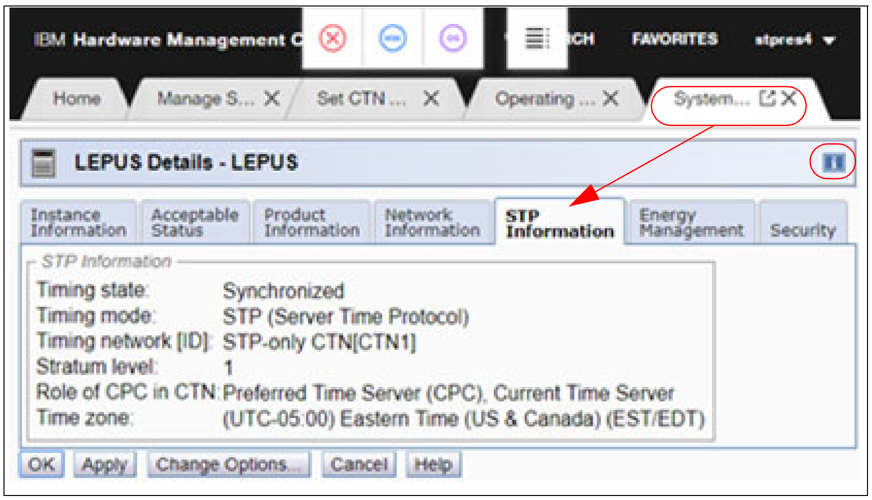


Figure 3-1 System Details: STP Information

If your server has an STP feature (Feature Code 1021) installed, there is a tab that is named **STP Information**. Use this tab to check whether your servers have the STP feature code installed. The number of fields that is shown on the **STP Information** tab varies depending on the type of timing network. This tab is a display-only. You cannot modify the timing network from this tab. Use this tab to observe the STP status for the selected CPC. There is no refresh of the information on this tab (only the status at the time that the menu was accessed is displayed).

The context-sensitive help area is a function that provides brief descriptions of the action buttons that are available on each individual tab. The buttons can be sequentially highlighted by pressing the Tab key. The help function itself is activated or deactivated by clicking **i** in the upper right of the opened window (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 window

To display and manage the STP configuration, use the **Manage System Time** tab to view and manage the CPCs that are known to the HMC.

Figure 3-2 on page 71 shows the **Manage System Time** tab.

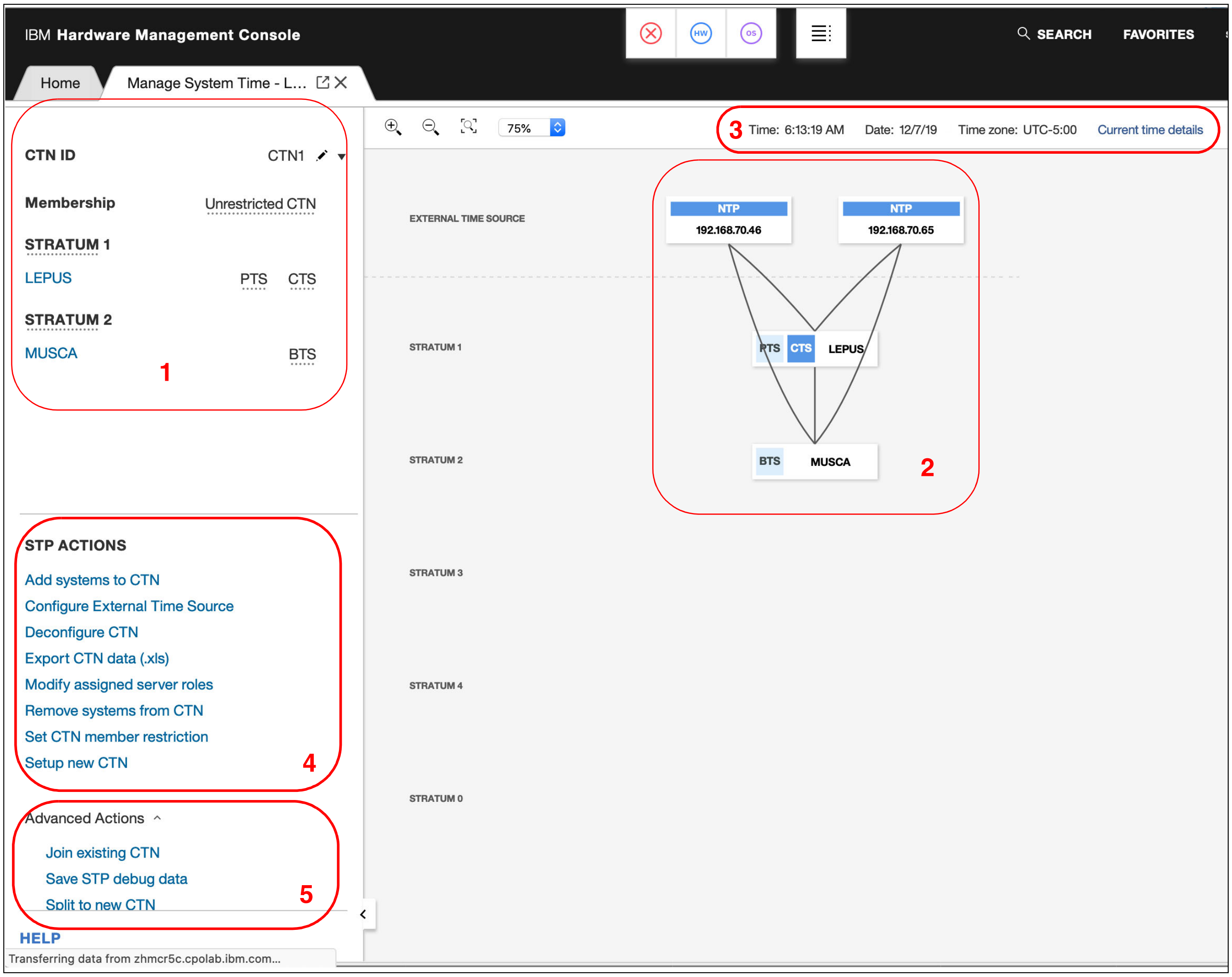


Figure 3-2 The Manage System Time tab on the HMC

The **Manage System Time** tab provides an overview of your CTN. Here you can view details and make changes. This tab is accessible only through the HMC. The **Manage System Time** tab has the following information:

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

   You can use the selection menu next to the small black arrow next to the CTN ID to select the CTN with which you want to work. Furthermore, you can see which members belong to the CTN and in which STP stratum and which roles they are. You also can see whether the selected CTN is restricted.

2. Graphical topology view (see #2 in Figure 3-2 on page 71)

   Here, your CTN is shown with its servers, time sources, and connections. You can also see in which stratum that your server is located. By selecting a connection or an object, you can see more details:

   – When selecting a server, the role within the CTN and the timing status is displayed, as shown in Figure 3-3 (This view can be zoomed in or out or scrolled.

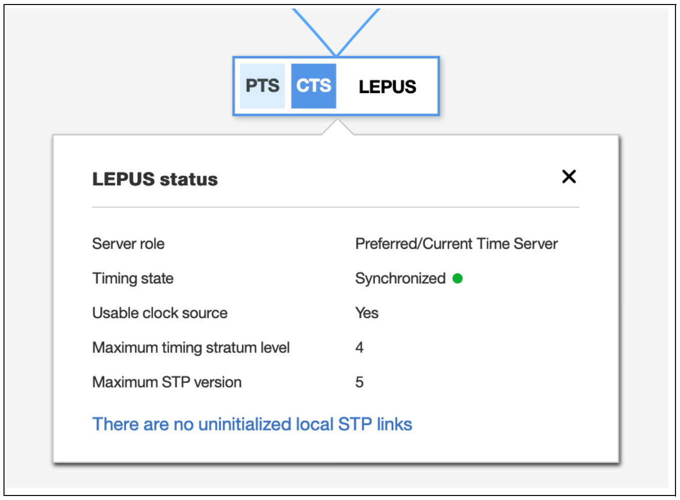


Figure 3-3 System details: STP menus

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

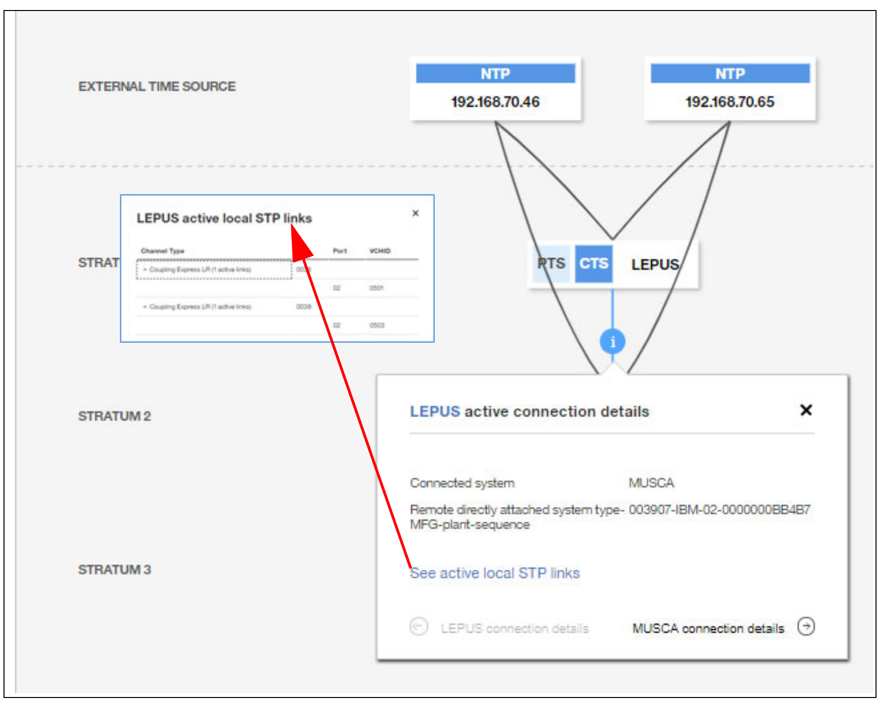


Figure 3-4 Timing links information

– When selecting a connection to the Network Time Protocol (NTP) server, details about this connection are displayed, as shown in Figure 3-5.

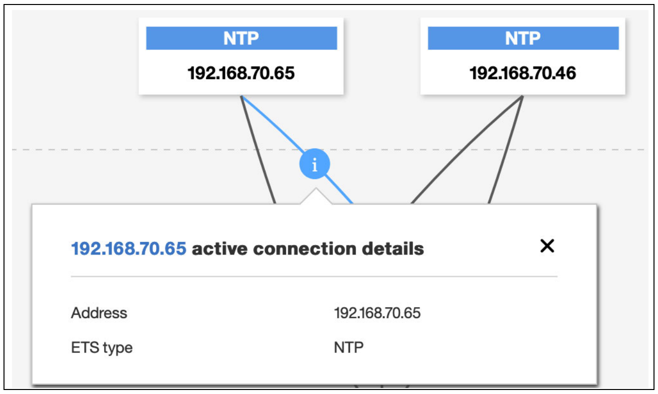


Figure 3-5 External Time Source link details

3. Topology toolbar (see #3 in Figure 3-2 on page 71)

   You can find details about your CTN by clicking **Current Time details**, as shown in Figure 3-6. You can also find menus for adjusting the time, setting the time zone offset, and adjusting the leap second offset. For more information, see 3.5, “Managing the time” on page 100.

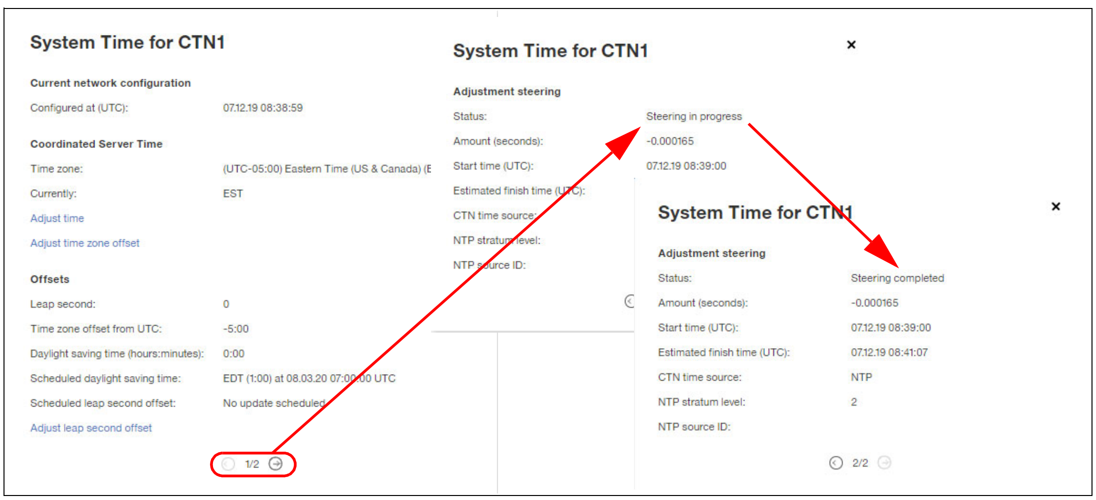


Figure 3-6 Current Time details

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

4. STP Actions area (see #4 in Figure 3-2 on page 71)

   With these menus, you can perform actions for your CTN by following the provided workflow. For example, you can add systems to the CTN or configure the External Time Source (ETS). For more information, see Chapter 4, “Managing Server Time Protocol configurations” on page 119.

5. Advanced Actions (see #5 in Figure 3-2 on page 71)

   A user can perform certain tasks on the Pulse Per Second (PPS) signal for troubleshooting by selecting **Advance Actions** → **Control Pulse per second (PPS) signal**. You can perform these tests on either PPS port of the oscillator (see Figure 3-7):

   – Detect a PPS signal from the ETS.

   – Perform an internal diagnostic on the PPS port.

   – Fence the PPS port by using the License Internal Code.

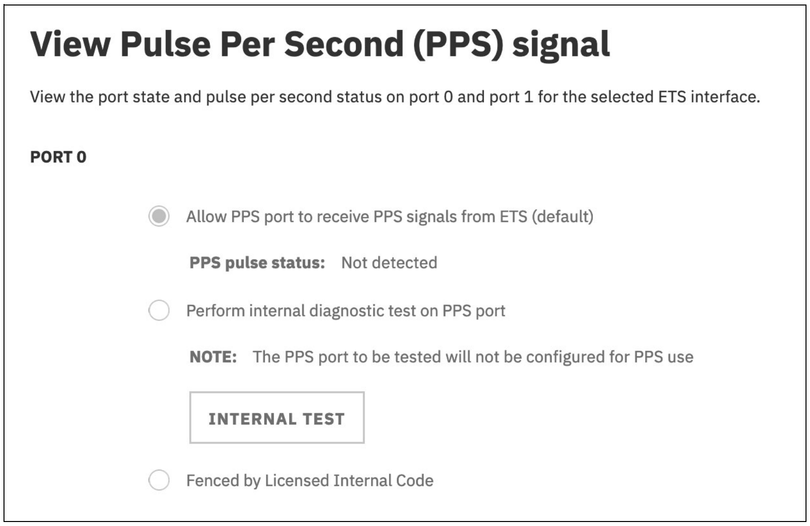


Figure 3-7 Advanced Actions: Pulse Per Second (PPS) signal

3.1.3 z/OS DISPLAY commands

There are several 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 using.

DISPLAY ETR

Before STP, the **DISPLAY ETR** command was used to display the synchronization mode and the status of the External Time Reference (ETR) ports that were seen by z/OS. With STP support, the command itself has not changed. However, the output was updated to display STP-related information where applicable.

In a CTN, the time-of-day (TOD) clock is being steered to the time that is provided by the Current Time Server (CTS). No reference to a Sysplex Timer is displayed.

Example 3-1 shows the DISPLAY ETR command in an STP-only CTN.

Example 3-1 The 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
```

The stratum level, CTN ID, and node ID information for the CTS is displayed, as are several extra, optional lines that show a different output depending on the CTN topology. For example, if this server was assigned as a Preferred Time Server (PTS), a Backup Time Server (BTS), or an Arbiter Server role, you see 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 which timing mode is used. The following modes are available:

- CDTS: Console dial-up time service
- CNTP: NTP without PPS
- CNTX: NTP without PPS from a secondary server
- PPSN: NTP with PPS
- PPSX: NTP with PPS from a secondary server

For more information about the output message IEA386I, see “z/OS MVS System Messages” for your z/OS version at the IBM z/OS Internet Library.

DISPLAY XCF,SYSPLEX,ALL

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

Example 3-2 The 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 The z/OS 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 (CF) within a Parallel Sysplex is enabled for CF Request Time Ordering before any server within the Parallel Sysplex can be defined an STP CTN.

For more information about the output message IXL150I, see “z/OS MVS System Messages” for your z/OS version at the IBM z/OS Internet Library.

There are two time-related CFCC commands, which are described in “Coupling Facility commands” on page 110.

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 act on. Some messages are issued on every member of the sysplex with **STPMODE YES** specified, which might cause the automation process to take multiple or redundant actions.

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

> **Note:** These messages are posted by every z/OS sysplex member that uses **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 that are 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 *cccccc*`

Where cccccc is `BACKUP` or `PREFERRED`. This informational message does not require any action, but ensures that the operational staff that is responsible for STP are 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 report only on the IEA037I messages.

The IEA0371 message informs you that an STP alert was received:

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

The alert codes are as follows:

01	The alert is related to the ETS.
02	The alert is related to the state of the PPS ports.
04	The alert is related to the hardware.

For more information about Reason Codes and the alert Code Message, see the “z/OS MVS System Messages” resource that is published for your z/OS version. For more information about the output message *IEA037I*, see the “z/OS MVS System Messages” resource for your z/OS version. Both resources can be found at the IBM z/OS Internet Library.

z/OS messages context

z/OS messages might help you identify various configuration issues and status changes for the servers in your CTN. However, wh8ile planning the roles of your servers and changing your CTN configuration, you must consider the following aspects.

Stand-alone CFs typically provide the best connectivity to other servers that require 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 that is 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 statuses. For example, the following message reports the result of a successful rename of a 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. However, you might see the following scenarios:

- Certain messages do not appear on the CTS because 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`*

  nnnnnnnnnnn is `'PREFERRED '| 'BACKUP '| '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

Note: In this example, we use NTP as the ETS. A Precision Time Protocol (PTP) as ETS configuration example is described in 3.3, “Configuring Precision Time Protocol as ETS” on page 89.

The sequence of steps in configuring a CTN with only a single server consists of the following steps:

1. Set the CTN ID.
2. Specify the CTN members.
3. Choose the PTS.
4. Choose the CTS.
5. Set the leap seconds.
6. Set the time zone.
7. Set the date and time.
8. Configure the ETS.
9. Confirm the changes.

In this example, we create a single-server CTN from scratch with a CPC that is named “LEPUS”, which is not a member of any CTN (see Figure 3-8).

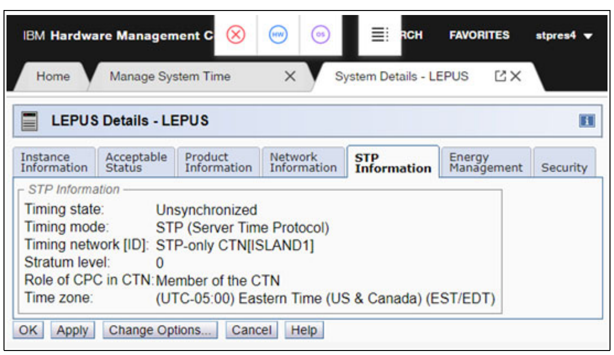


Figure 3-8 Server details: STP information for LEPUS

The process of creating a CTN is part of the workflow “Set up new CTN” in the Manage System Time window on the HMC.

Complete the following steps:

1. Create a CTN, as shown in Figure 3-9 on page 79. (If there are no servers that are available, this operation is not allowed.)

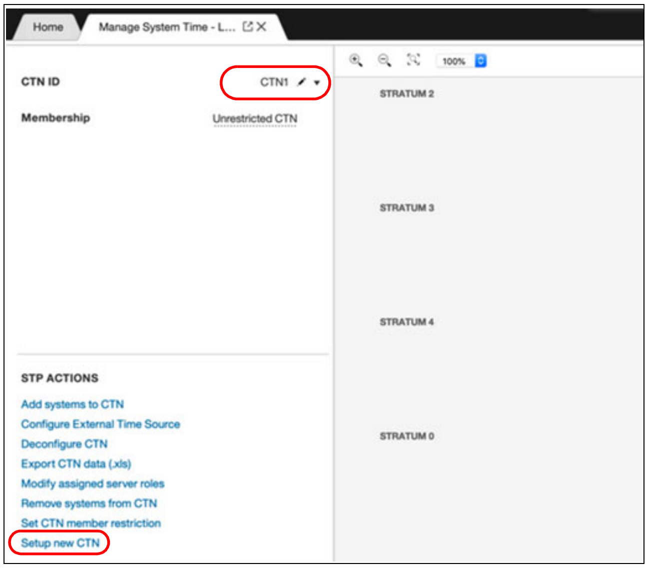


Figure 3-9 Creating a CTN

2. Set the CTN ID by selecting **Setup new CTN** and entering the name of your new CTN. Click **Next** (Figure 3-10).

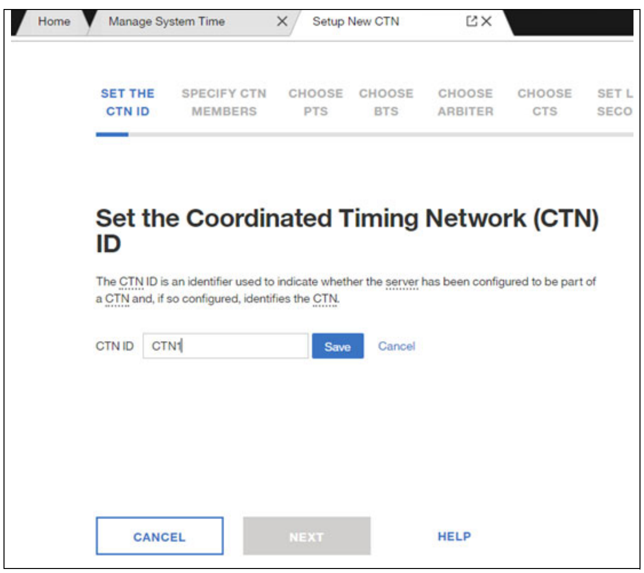


Figure 3-10 Setting the CTN ID

3. Specify the CTN members.

   Select the systems that will become members of the new CTN. Only systems that are not a member of any active CTN are shown here. When all the correct members are selected, click **Next** (Figure 3-11).

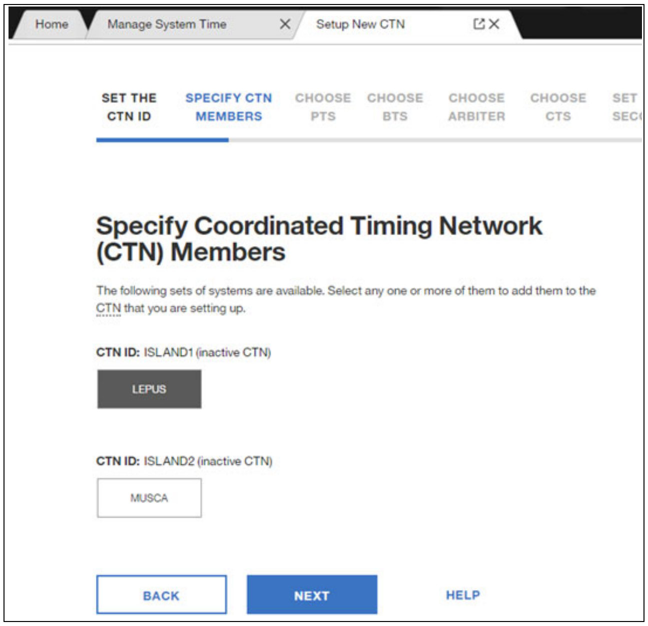


Figure 3-11 Specifying LEPUS as a CTN member

4. Choose the PTS.

   Select the server that will become the PTS. All systems that are available for role assignment are displayed. In this example, we select **LEPUS** as the PTS and click **Next** (see Figure 3-12).

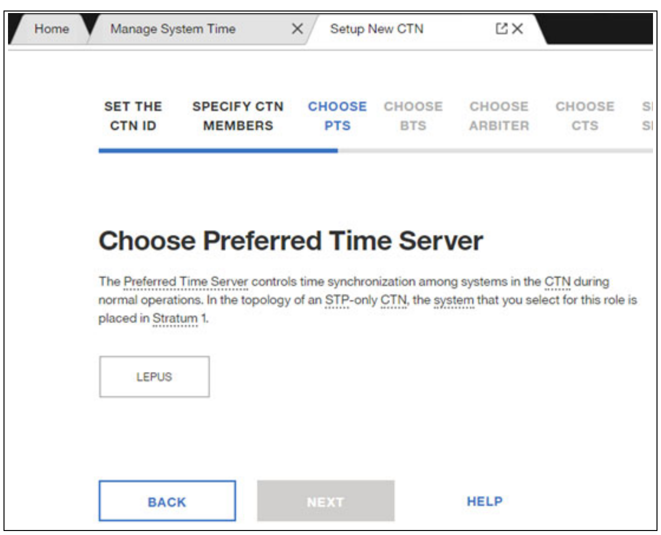


Figure 3-12 Choosing the Preferred Time Server

5. Choose the CTS.

   Select the system that is going to be the CTS and stratum 1. The CTS adjusts the Coordinated Server Time (CST) by steering it to the time that is obtained from an ETS. Because we have only a single server in our configuration, the PTS becomes the CTS, as shown in Figure 3-13.

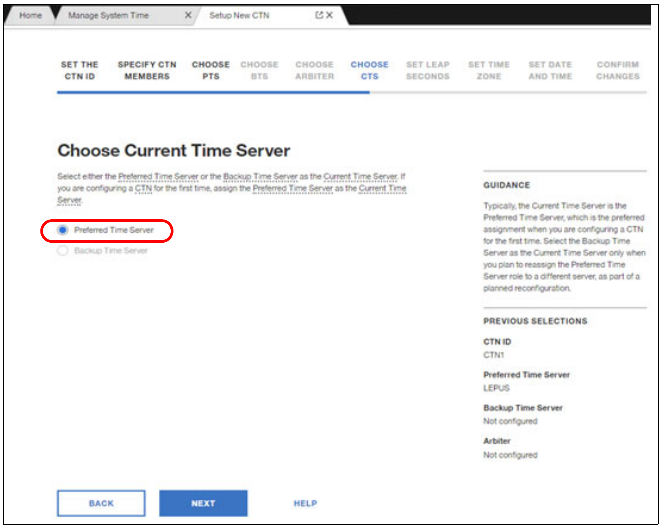


Figure 3-13 Choosing the Current Time Server

6. Set leap seconds.

   Specify the leap seconds for the CTN. If you do not need to add leap seconds, you can leave this value at 0. You can add leap seconds to your CTN at a later point without interruption.

   Enter the number of leap seconds based on the year and click **Next** (see Figure 3-14).

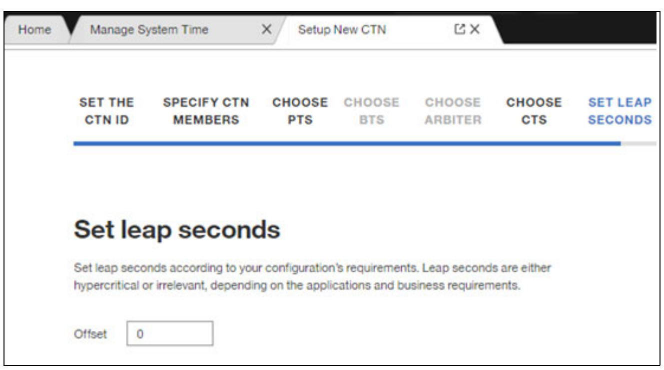


Figure 3-14 Setting leap seconds

7. Set the time zone.

   Select a time zone from the drop-down menu or create your own time zone by selecting one of the user-defined time zones (UD1 - UD5).

   If the systems in your CTN span multiple time zones, see 3.5.5, “Parallel Sysplex and multiple time zones” on page 113.

In the section "Clock adjustment for Daylight Saving Time (DST)", you can specify your DST offset and whether to automatically adjust for DST, as shown in Figure 3-15, or to set DST or standard time manually.

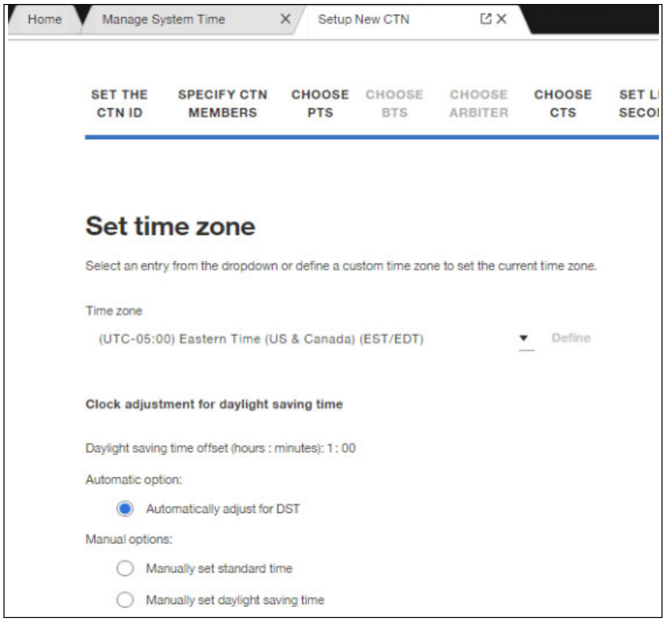


Figure 3-15 Setting the Daylight Saving Time

8. Set the date and time.

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

   In this example, we choose to use the ETS. If no ETS is configured, click **Next** to open the **Configure External Time Source** option. If an ETS was configured before, you can manually go to this option by clicking **Configure External Time Source** (see Figure 3-16).

   When you finish configuring the ETS, you automatically see the **Setup new CTN** tab.

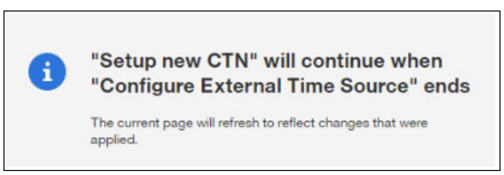


Figure 3-16 Moving to configuring the External Time Source

9. Configure the ETS:

   a. Select the system to modify its ETS (Figure 3-17 on page 83).

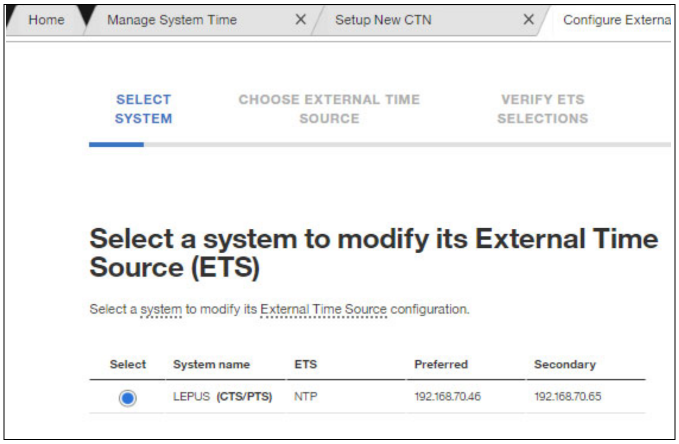


Figure 3-17 Selecting the system for the ETS

b. Choose from the ETS options. You can select NTP, PTP, NTP with PPS, or PTP with PPS for enhanced accuracy, or you can select to not use an ETS (Figure 3-18).

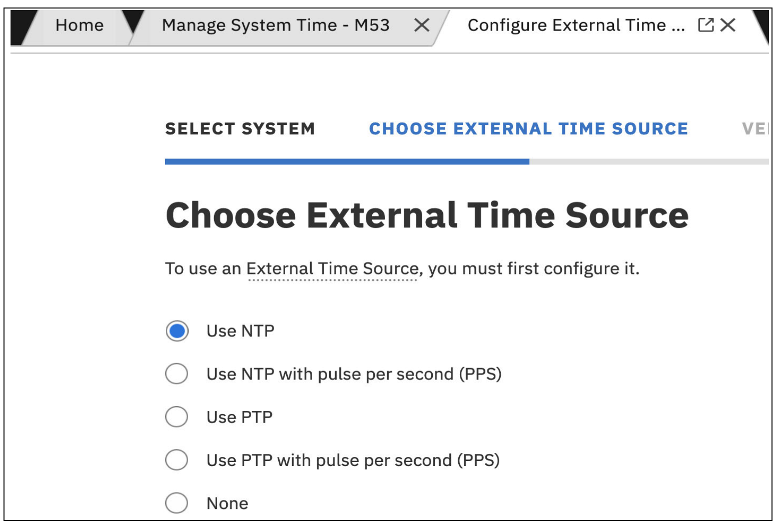


Figure 3-18 Choosing an NTP as the ETS

c. Enter the IP addresses or domain names of the NTP servers. Click **Test Connectivity**, which checks whether the NTP servers can be reached and displays the results in the Stratum, Source, and Status fields (see Figure 3-19).

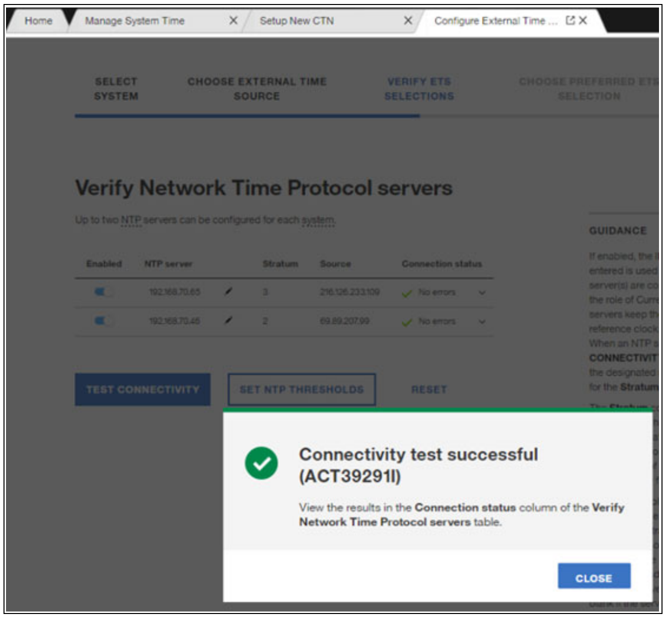


Figure 3-19 Testing NTP connectivity

d. Specify which of the NTP servers is the preferred one. Usually, you should select the one with the lower stratum level because that one provides better accuracy. Before you apply the new settings, you see an overview of all the settings in the Confirm External Time Source configuration window. The new settings and the old settings are displayed next to each other so that you can see the changes that will be performed. If everything looks correct, click **Apply** (see Figure 3-20 on page 85).

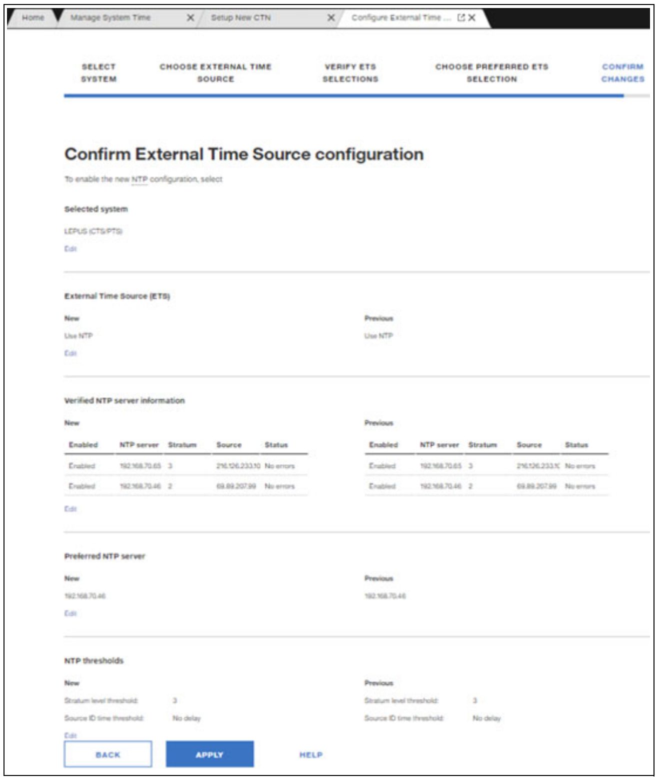


Figure 3-20 ETS configuration: Confirmation of changes

10. CTN configuration - Confirm Changes.

    When the ETS configuration completes, the Setup New CTN workflow appears. The last window is Confirm Changes, where you see a graphic of what your CTN will look like (Figure 3-21).

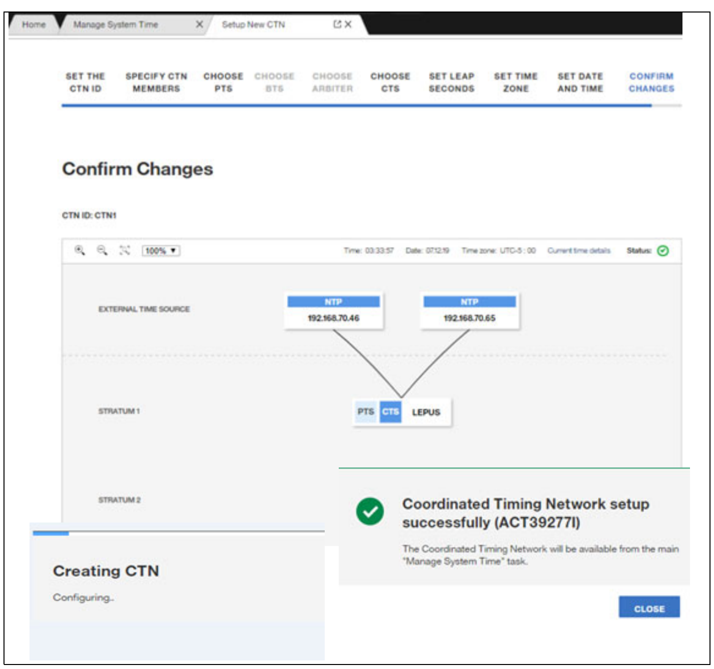


Figure 3-21 CTN configuration: Final step

When the task finishes, you can see your new CTN in the **Manage System Time** tab on the HMC, as shown in Figure 3-22.

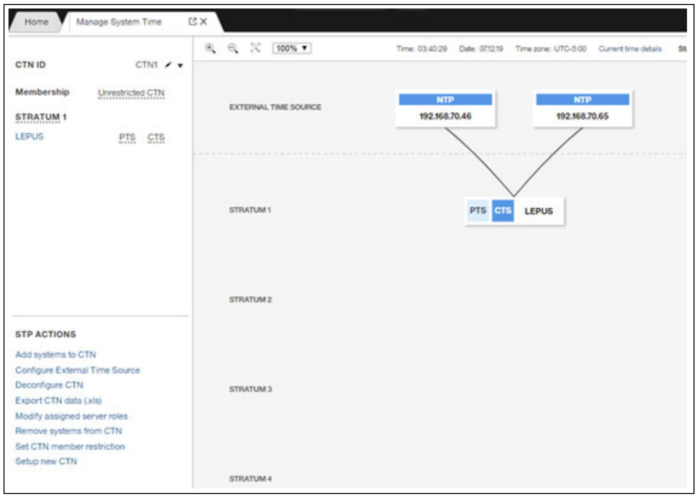


Figure 3-22 CTN configuration result

3.2.2 Dual-server CTN

Note: In our example, we use an NTP as the ETS. A PTP as the ETS configuration example is described in 3.3, “Configuring Precision Time Protocol as ETS” on page 89.

To configure a CTN with a PTS and a BTS, complete the following steps:

1. Set the CTN ID.
2. Specify the CTN members.
3. Choose a PTS.
4. Choose a BTS.
5. Choose a CTS.
6. Set the leap seconds.
7. Set the time zone.
8. Configure the ETS.
9. Set the date and time.
10. Confirm the changes.

The process is identical to configuring a single-server CTN (see 3.2.1, “Single-server CTN” on page 78) except for the additional step 4 (see Figure 3-23 on page 87).

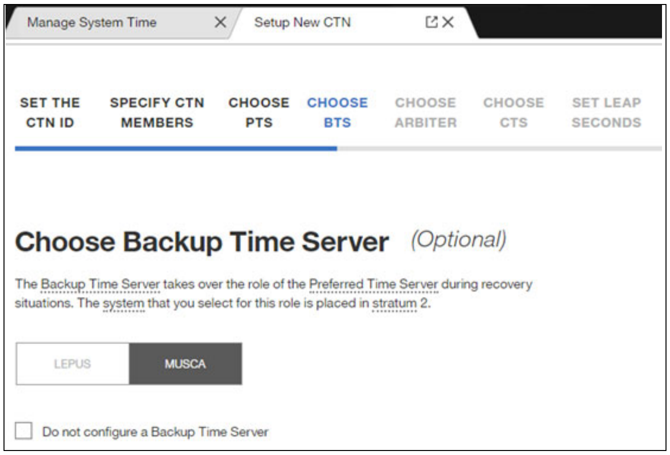


Figure 3-23 Choosing the Backup Time Server

The window shows the systems that are eligible for BTS role assignment. The systems that already are assigned the PTS role are shown as unavailable (in this case, the LEPUS system). You can also choose not to assign a BTS. Before committing the changes, you are prompted by the Confirm Changes window, where you can see what your CTN will look like, and then you can apply the configuration. The configuration is shown in Figure 3-24.

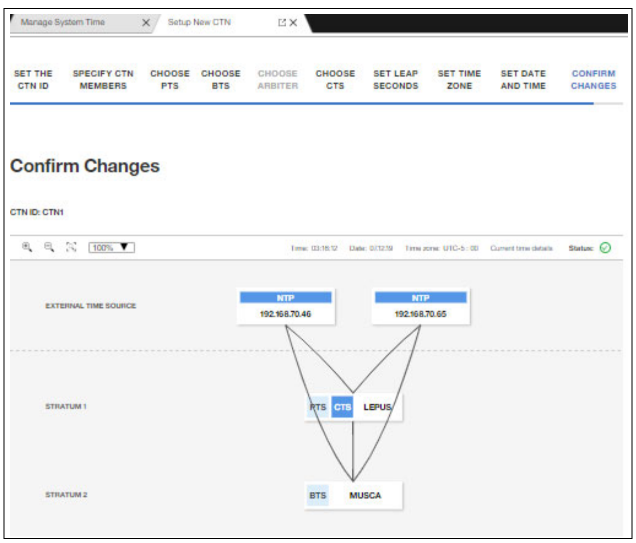


Figure 3-24 Dual-server CTN: Configured CTN

Important: At least one ETS must be configured for both CPCs, which assume the roles of PTS and BTS. As a best practice, set up dual ETSs for each of the two CPCs, preferably the same ETSs (primary and alternative) for both servers.

3.2.3 CTN with three servers: PTS, BTS, and Arbiter

Note: In our example, we use an NTP as the ETS. A PTP as the ETS configuration example is described in 3.3, "Configuring Precision Time Protocol as ETS" on page 89.

To configure a CTN with a PTS, a BTS, and an Arbiter, complete the following steps:

1. Set the CTN ID.
2. Specify the CTN members.
3. Choose the PTS.
4. Choose the BTS.
5. Choose the Arbiter.
6. Choose the CTS.
7. Set leap seconds.
8. Set the time zone.
9. Configure the ETS.
10. Set the date and time.
11. Confirm the changes.

The process is identical to configuring a dual-server CTN (see 3.2.2, "Dual-server CTN" on page 86) except for the additional step 5. Step 5 becomes available if a BTS is selected (see Figure 3-25).

Servers that assume the Arbiter role must have coupling connectivity to both the PTS and the BTS. Also, there is no need to configure an ETS to this CPC.

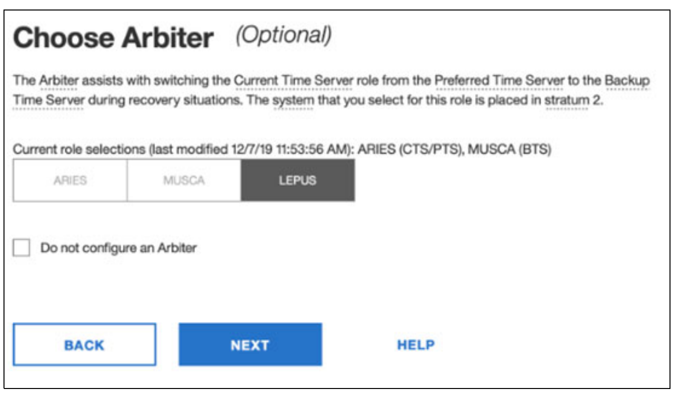


Figure 3-25 Choosing the Arbiter (three or more servers in a CTN)

When all changes are applied, you can see the layout of the new CTN in the **Manage System Time** tab, as shown in Figure 3-26 on page 89.

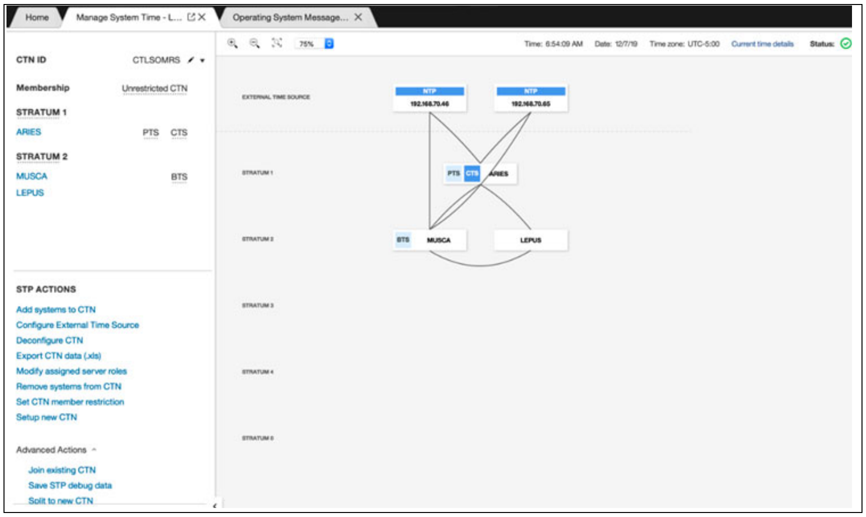


Figure 3-26 Three-server CTN: Final configuration

3.3 Configuring Precision Time Protocol as ETS

In this section, we introduce the procedure that is used to configure PTP as ETS for one of the CPCs in the CTN.

At the time of writing, the servers that support PTP as ETS are IBM z16 and IBM z15. Although the configuration process is similar (as are the functions), due to different ETS connectivity, there are slight differences in the menus for the PTP interface selection.

The following example is based on an IBM z15 that is being configured for PTP as ETS. Later in this section, we also provide a view of an IBM z16 PTP interface in Figure 3-32 on page 93.

Important: This section assumes that the IBM Z CPC has been properly connected to a PTP supporting and enabled network infrastructure:

- Connected to the Support Element (SE) dedicated ETS interfaces (named em3 and em4) for IBM z15.
- Connected directly to the oscillator card (OSC) on the CPC drawers for IBM z16 (the interfaces are named[a] ETS1 and ETS2).

The physical infrastructure, from the PTP Grandmaster and any network elements that are used to connect the CPC, must be configured to ensure PTP signaling communication and IP routing (if enabled).

ETS connectivity: Physical network (local area network (LAN)) connectivity from the SE (IBM z15) or directly from the CPC drawer (IBM z16) to the client network infrastructure is performed by an IBM System Services Representative (IBM SSR) during system installation.

PPS signal connectivity is similar for IBM z15 and IBM z16.

For more information, see the following publications:

- *IBM 8561 Installation Manual for Physical Planning*, GC28-7002
- *IBM 8562 Installation Manual for Physical Planning*, GC28-7011
- *IBM 3931 Installation Manual for Physical Planning*, GC28-7015

a. The numbering and naming of the interfaces might differ for your configuration.

The scenario in this section is provided as a guide to migrate the ETS for a CPC (in this case, a IBM z15 member of a CTN) from using NTP to using PTP as ETS. The scenario starts with a 3-CPC CTN that has the PTS using PTP as ETS and the BTS using NTP as ETS. The initial configuration is shown in Figure 3-27.

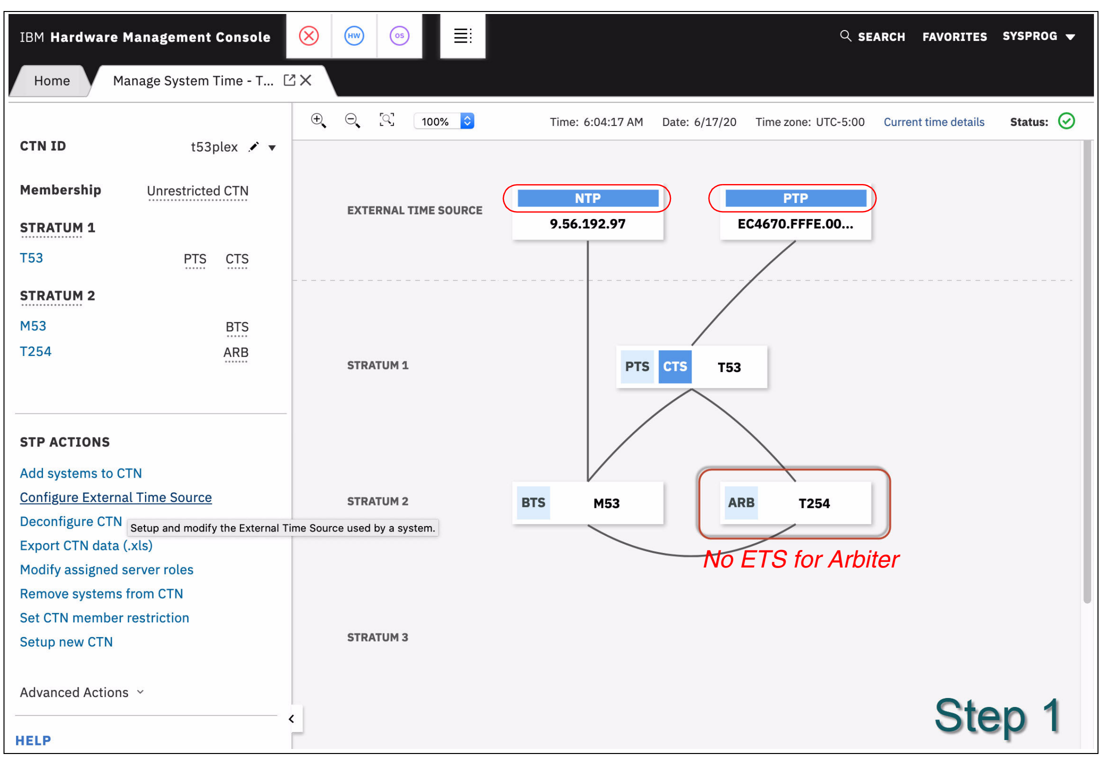


Figure 3-27 Initial configuration: Three-server CTN with a mixed ETS

Here is a short description of the scenario:

- Initial status: The servers in the “t53plex” are configured as follows:
  - T53 (IBM z15 T01): PTS and CTS. Uses PTP with PPS as ETS.
  - M53 (IBM z14 M0x): BTS. Uses NTP with PPS) as ETS.
  - T254 (IBM z15 T02): Arbiter. It has no ETS.
- The following high-level steps are performed:
  - The ETS infrastructure is prepared for T254, and the appropriate Ethernet ports on the SEs are connected to a PTP-capable switch.

    Note: At the time of writing, IBM z16 and IBM z15 CPCs support PTP as ETS.

  - We configure the ETS for the Arbiter (T254).
  - We reassign server roles in the CTN (T254 becomes BTS, and M53 becomes the Arbiter).

We describe in detail only the ETS assignment for T254 (PTP w with PPS). The CTN roles’ reassignment is described in 4.2.4, “Changing the server roles” on page 125.

3.3.1 Configuring ETS for an IBM z15 (PTP with PPS)

To configure ETS for an IBM z15 (PTP with PPS), complete the following steps:

1. Click **Configure External Time Source** (Figure 3-27 on page 90). The Select System window opens, as shown in Figure 3-28.

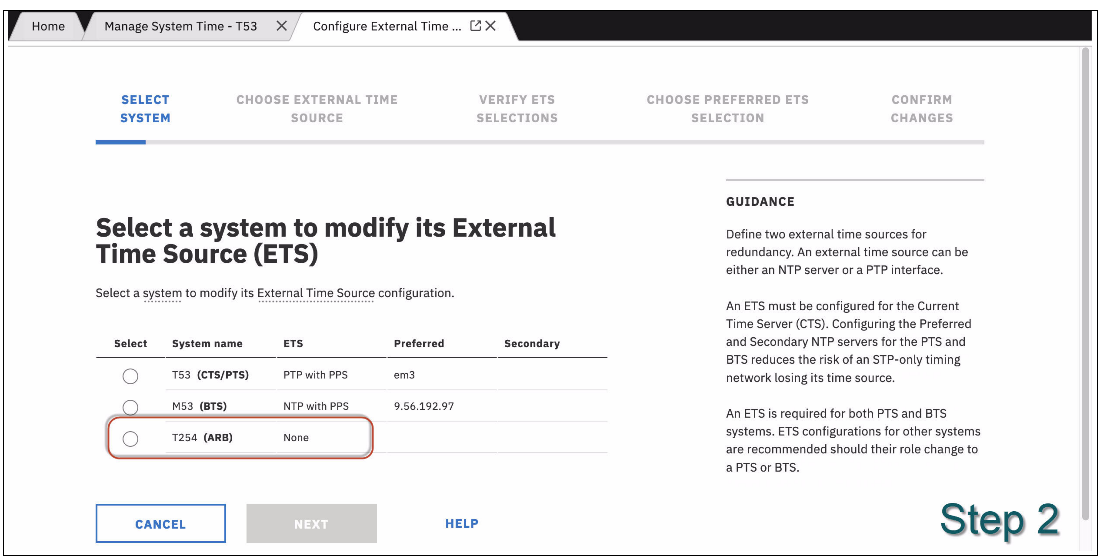


Figure 3-28 System selection

2. Select the system T254, as shown in Figure 3-29. Click **Next**.

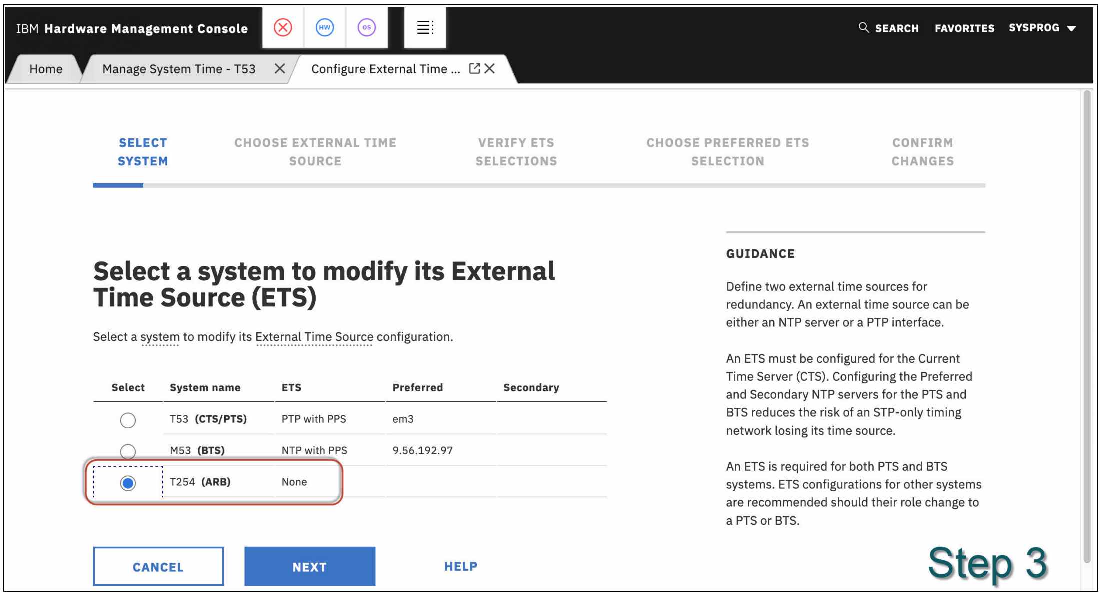


Figure 3-29 Selecting the T254 system

3. For the ETS type, select **Use PTP with pulse per second (PPS)**, as shown in Figure 3-30. Click **Next**.

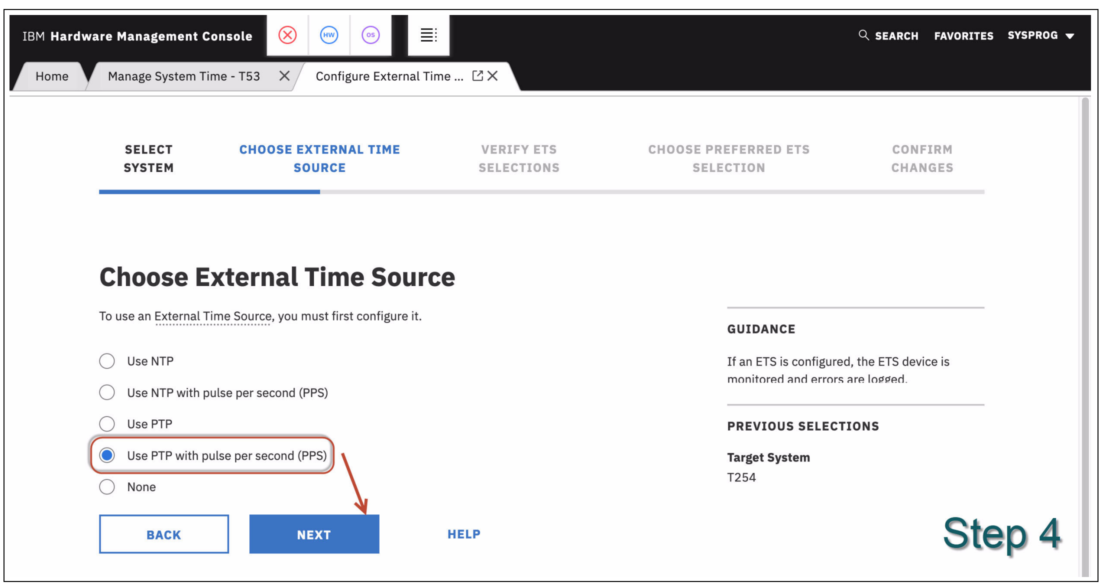


Figure 3-30 Selecting the ETS type

4. Now, there is no specified Ethernet interface for accessing the ETS from the SE, so proceed to the SE interface selection (IBM z15). The SE Ethernet interface selection menu is shown in Figure 3-31 on page 93. For Ethernet ports for IBM z16, see Figure 3-32 on page 93.

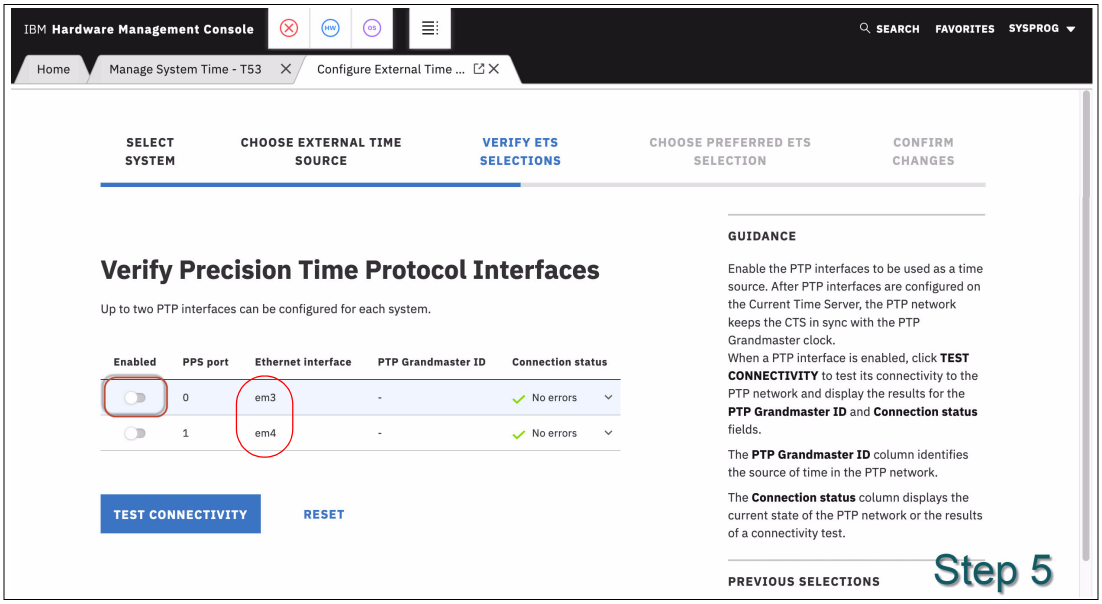


Figure 3-31 SE Ethernet interface selection (IBM z15)

Figure 3-32 shows ETS interfaces for an IBM z16.

> **Note:** IBM z16 ETS interfaces (on the CPC) use a different naming convention than the SE ETS interfaces that are used for IBM z15.

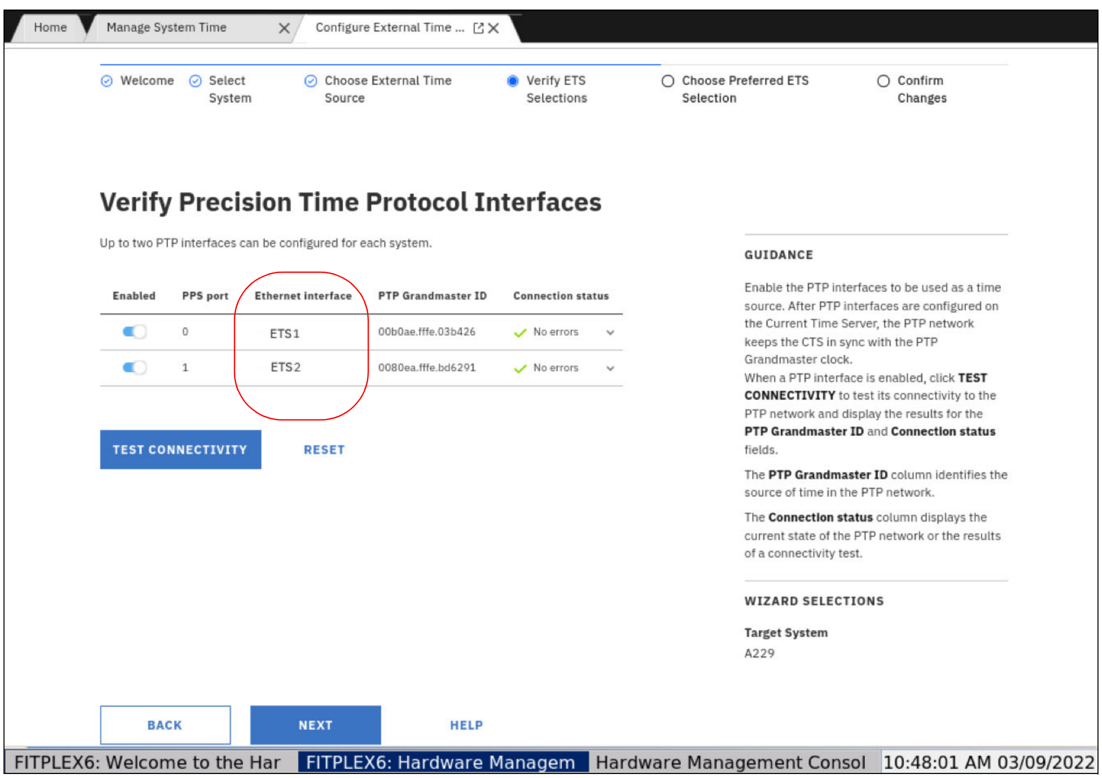


Figure 3-32 ETS interfaces for an IBM z16 A01 (M/T 3931)

5. Because we know that the em3 interface is connected, select **em3**. Use the SE interface to test the connectivity (after the interface is selected), as shown in Figure 3-33.

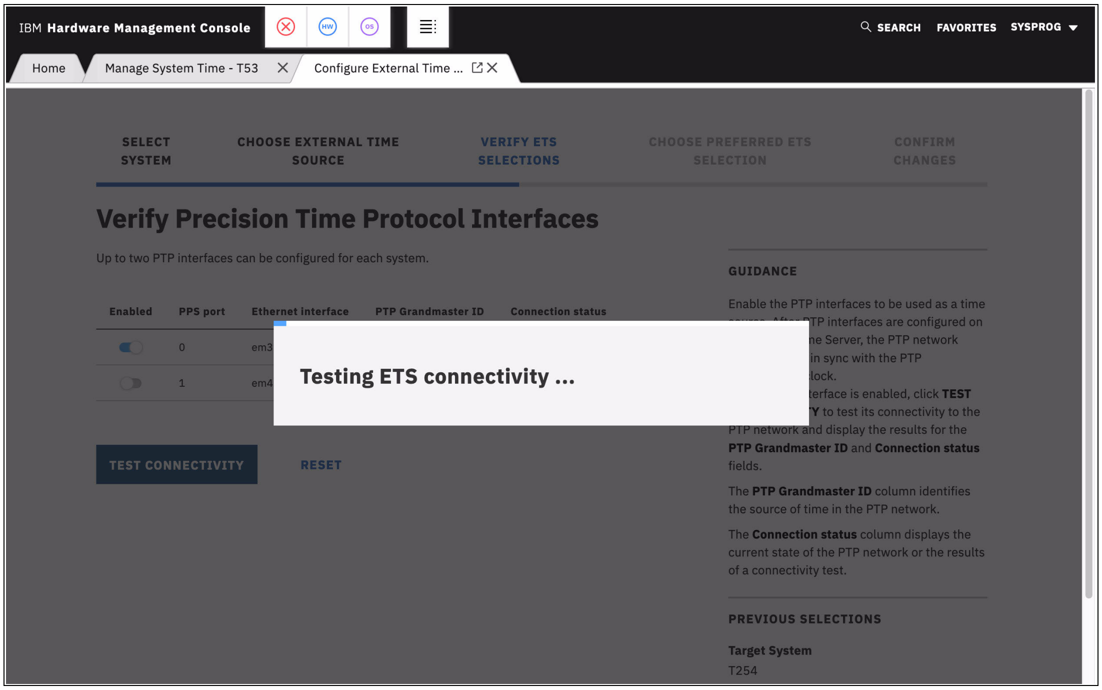


Figure 3-33 PTP source connectivity test

6. When the test is successful, the message that is shown in Figure 3-34 appears. Click **Close**.

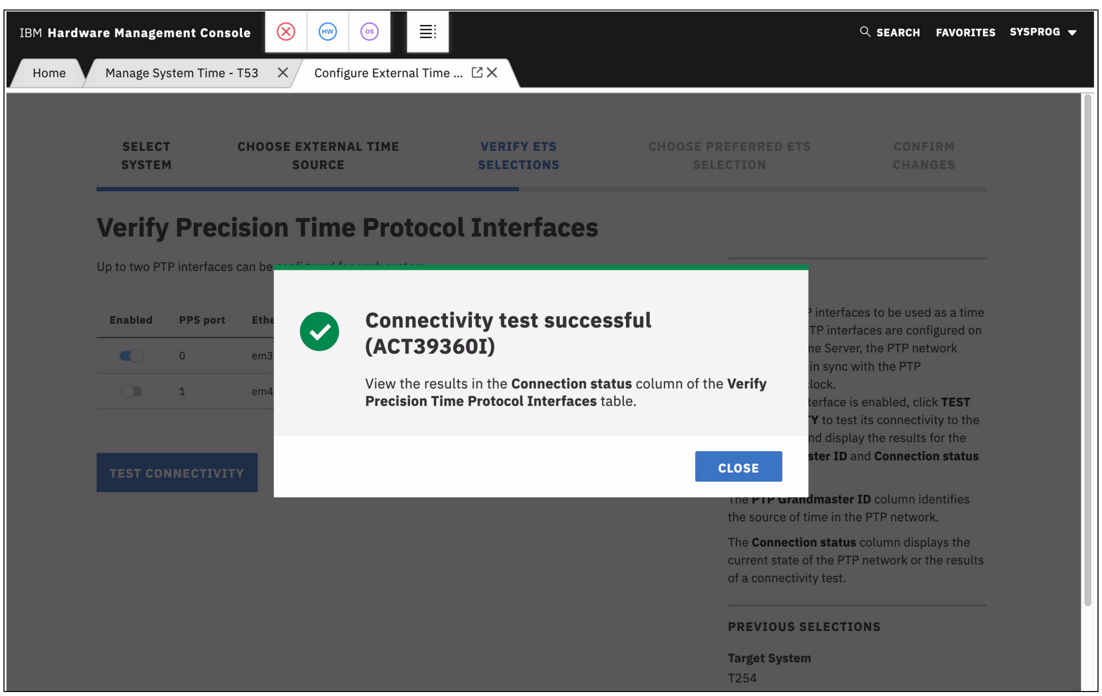


Figure 3-34 ETS connectivity test successful

7. The configured interface is shown in Figure 3-35 on page 95. The PTP Grandmaster is discovered and its name is populated automatically. Click **Next**.

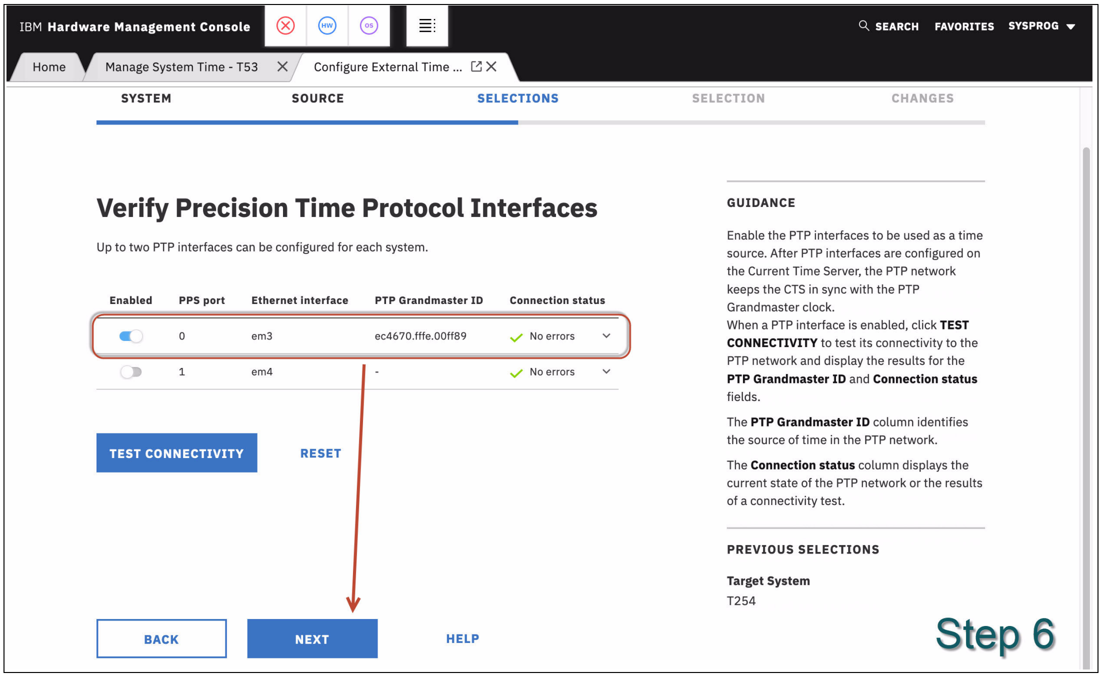


Figure 3-35 Verifying the interfaces for PTP

8. After the Ethernet interfaces for PTP access are configured, select the preferred interface (em3), as shown in Figure 3-36. Click **Next**.

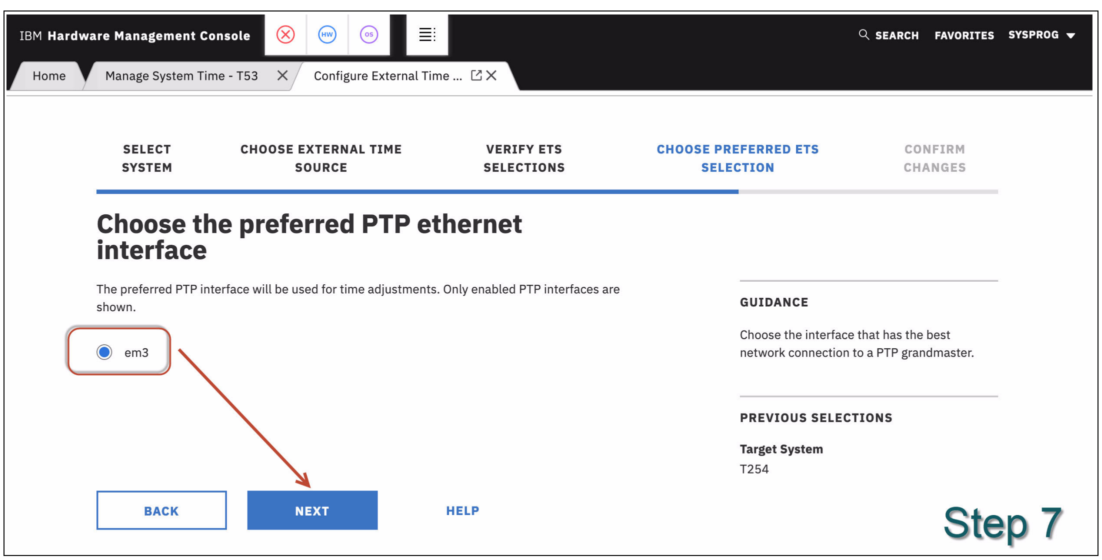


Figure 3-36 Selecting the preferred interface

9. The Confirm External Time Source configuration window opens, as shown in Figure 3-37. To commit the changes, click **Next**.

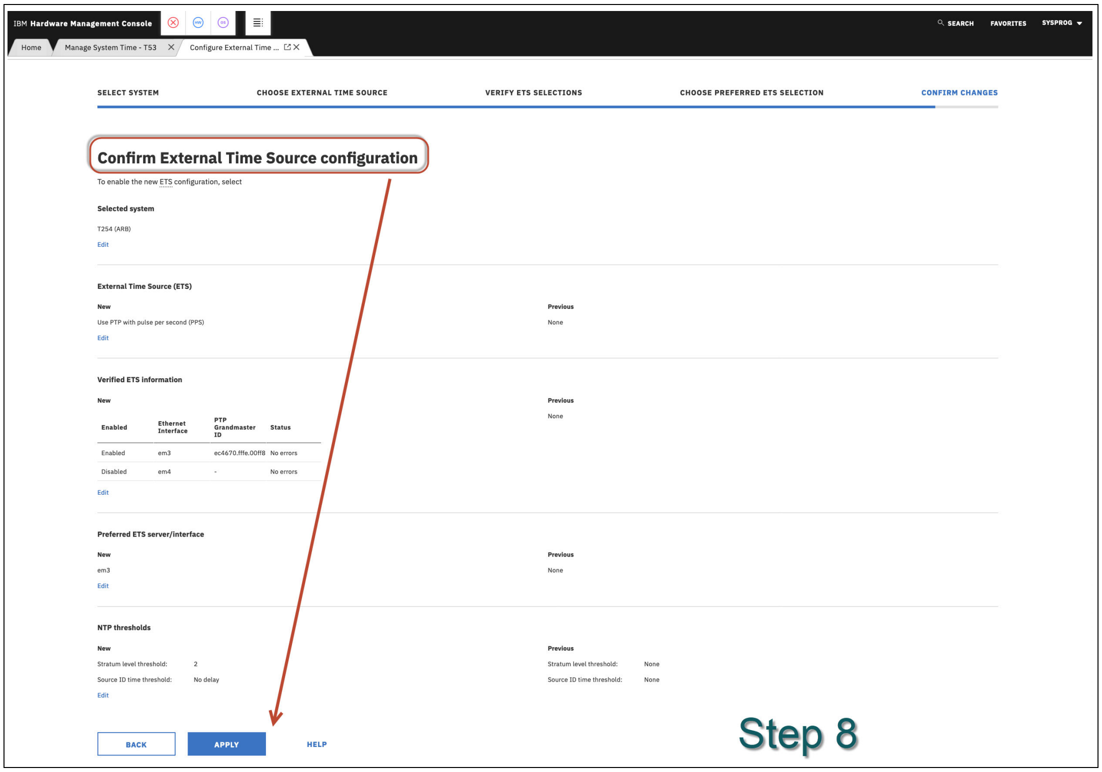


Figure 3-37 Confirmation menu

10. Figure 3-38 shows a success configuration message. Click **Close**.

> **Note:** Make sure that you save the data that you entered into the alternative SE, as noted in Figure 3-38.

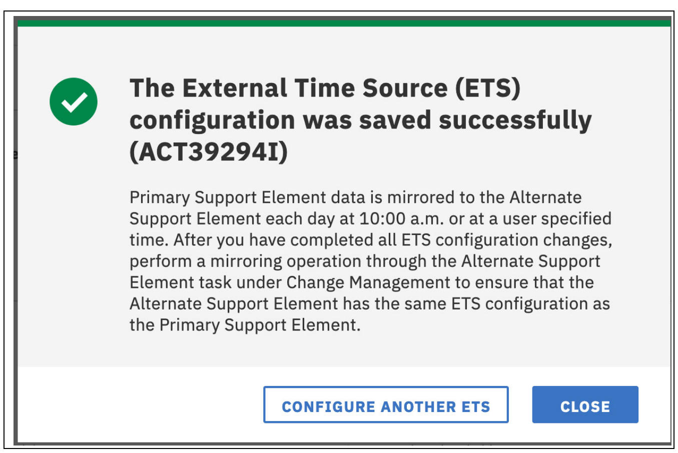


Figure 3-38 Success message for ETS configuration

11. The CTN configuration has not changed. The Arbiter (T254) has access to an ETS and is prepared to assume the PTS role or BTS role.

12. Swap the server roles in the “t53plex” CTN (as described in 4.2.4, “Changing the server roles” on page 125). The final configuration is shown in Figure 3-39.

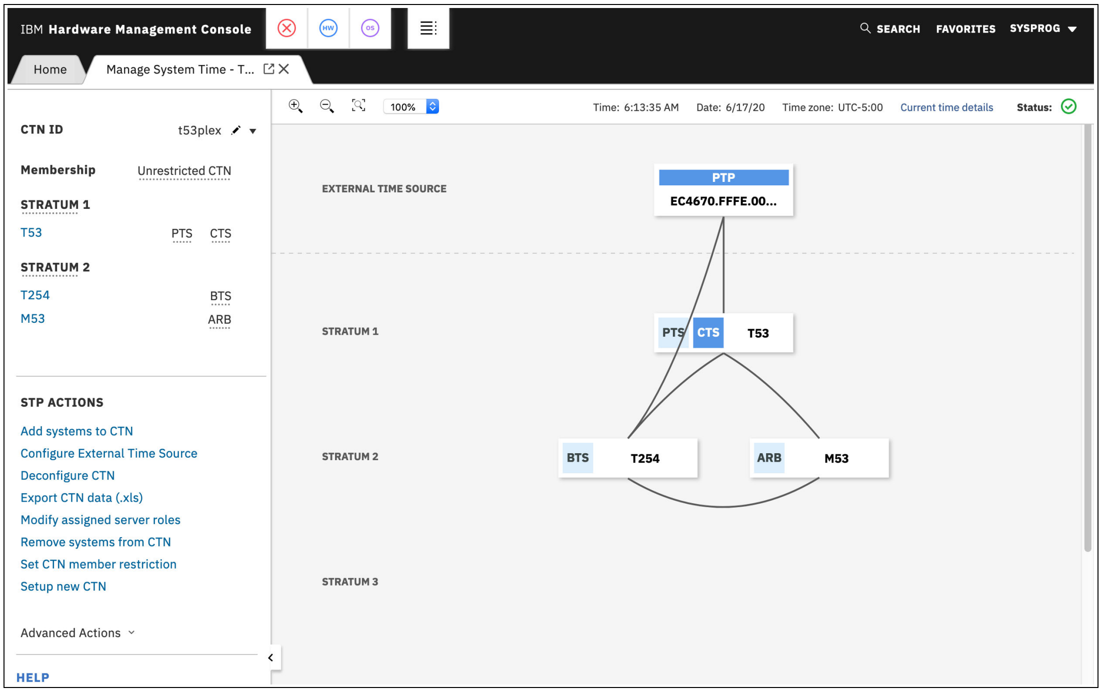


Figure 3-39 PTS and BTS using PTP as ETS

3.4 Time server power failure failover (IBM z16 only)

N-mode power considerations were introduced in 2.4, “IBM z16 power failure handling” on page 49. In a CTN configuration where both the PTS and the BTS are IBM z16 CPCs, it is now possible to configure CTS failover to the BTS based on an N-mode power signal that is available for IBM z16.

An N-mode power signal assumes that the IBM z16 CPCs (both the PTS and the BTS) have at least one side of their power feed that is protected by an uninterruptible power supply (UPS). An example of such connectivity (a Power Distribution Unit (PDU)-based system) is shown in Figure 2-12 on page 50. Redundant power connectivity is provided for both PDU and Bulk Power Assembly (BPA) systems.

N-mode provides a decision mechanism for CTS role to fail over to the BTS in a situation when there is an imminent power failure of the CTS CPC, thus avoiding an unscheduled CTN reconfiguration. Other decision mechanisms for CTS failover are Console Assisted Recovery and Arbiter.

Setting the time server power failover feature (optional, and available only if both the PTS and the BTS are IBM z16 CPCs) should be done for both the PTS and the BTS.

To enable the time server power failover feature in your CTN, complete the following steps:

1. Access the menu that is shown in Figure 3-40.

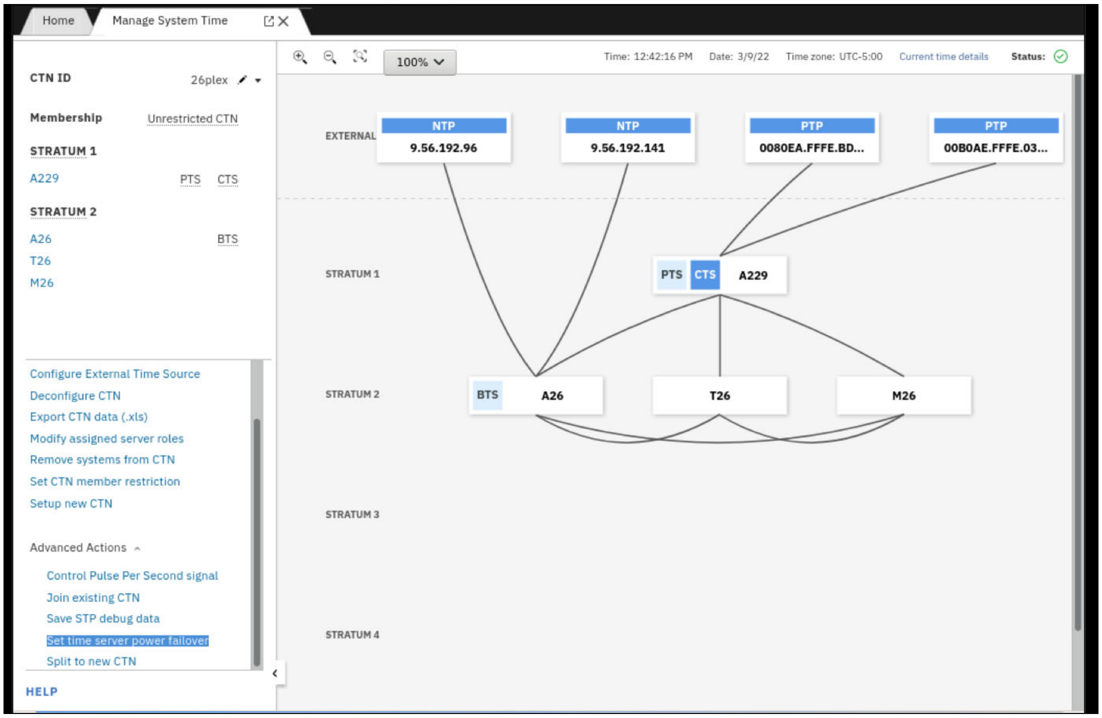


Figure 3-40 Set Time Server Power failover: Initial configuration

2. The menu directs you to change the failback timer (after the N-mode switch of the CTS to the BTS occurred) for the BTS to surrender the CTS role back to the PTS (Figure 3-41).

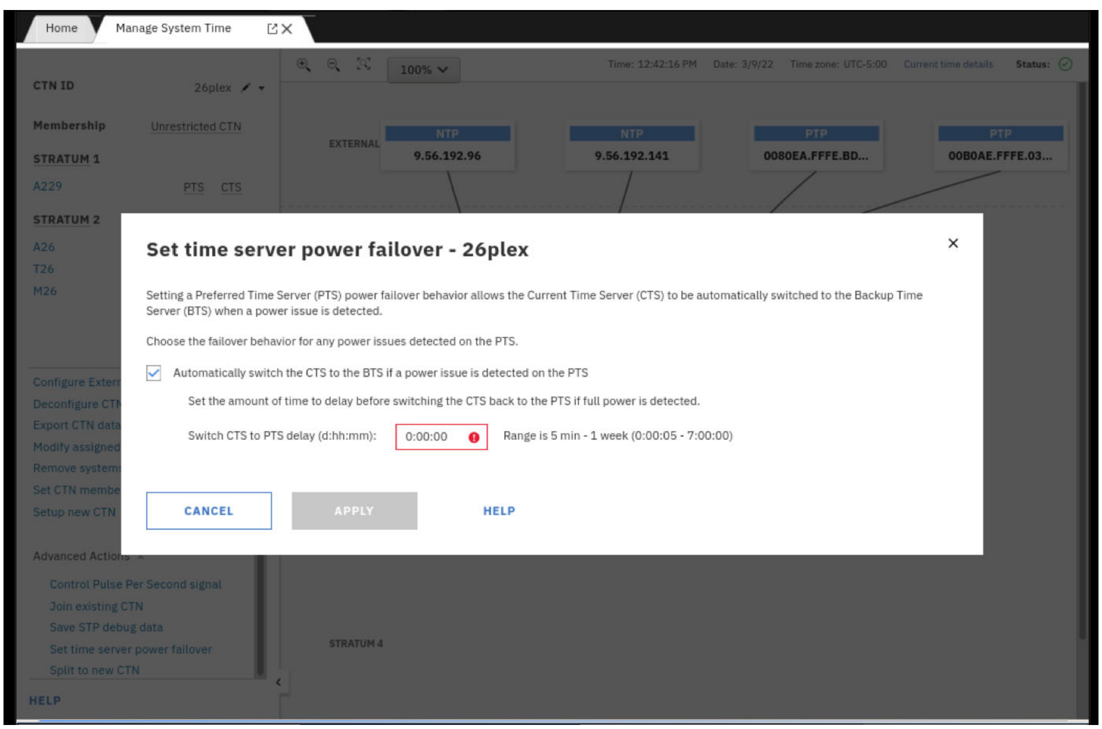


Figure 3-41 BTS timer for surrendering the CTS role (failback)

3. Set the failback timer (minimum 5 minutes), as shown in Figure 3-42.

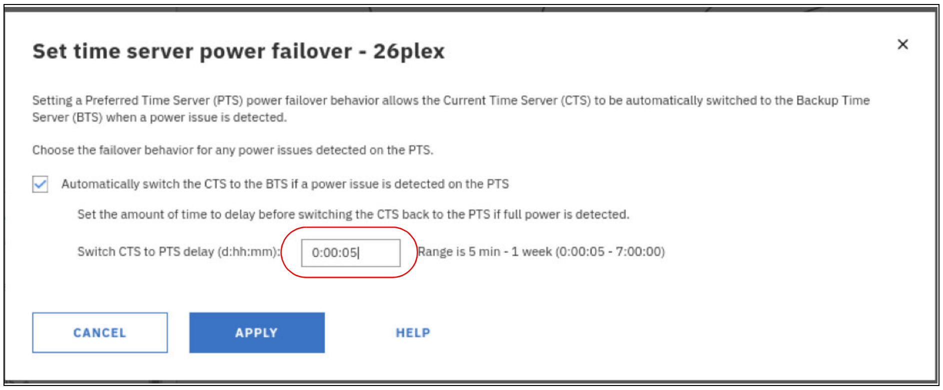


Figure 3-42 Failback timer (the BTS to the PTS after an N-mode switch on the PTS)

4. If the BTS is not an IBM z16, a warning message appears, as shown in Figure 3-43. Configuring the PTS (CTS) failback timer fails if both the PTS and the BTS are not IBM z16 servers.

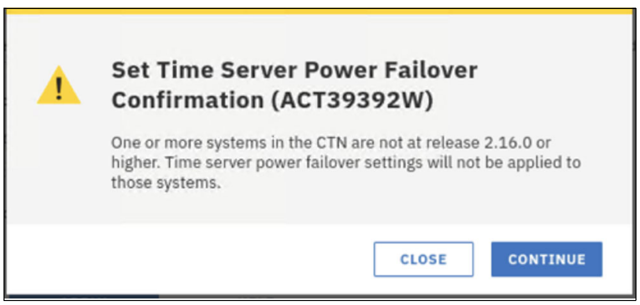


Figure 3-43 System-level warning for time server power failover capability

5. If both servers are IBM z16, a confirmation message appears (Figure 3-44).

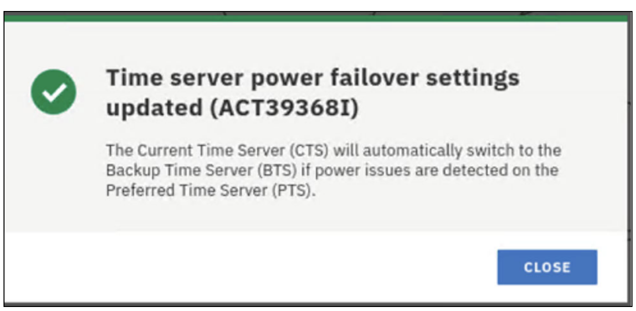


Figure 3-44 Confirmation message: N-mode power failover for the PTS enabled

Manual failback for time server power failure

When the PTS surrenders the CTS role to the BTS because of a potential power failure (for N-mode power systems), you can manually initiate CTS failback to the PTS, as shown in Figure 3-45, by using the same time server power failover menu (see Figure 3-40 on page 98).

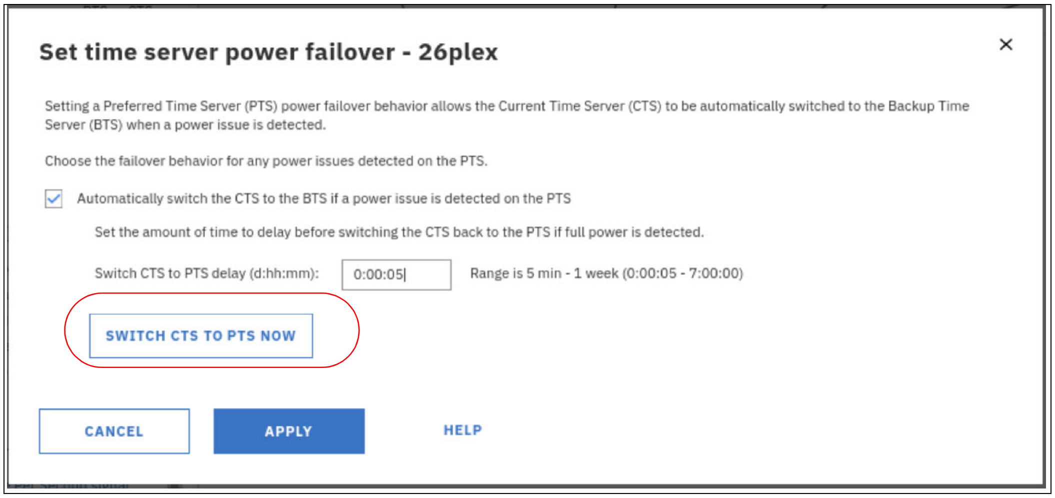


Figure 3-45 Manual failback option

3.5 Managing the time

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

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

3.5.1 STP time adjustment

Without regular adjustment, the time on the CTS drifts from its initial setting because of the frequency drift of the oscillator that is used to step the TOD clock. This situation might not meet time accuracy requirements. Therefore, adjustments might need to be made regularly, either manually or by referencing an ETS.

To do a manual adjustment of time, click **Adjust time** in the Current time details window (over the graphical topology view) in the **Manage System Time** tab, as shown in Figure 3-46 on page 101.

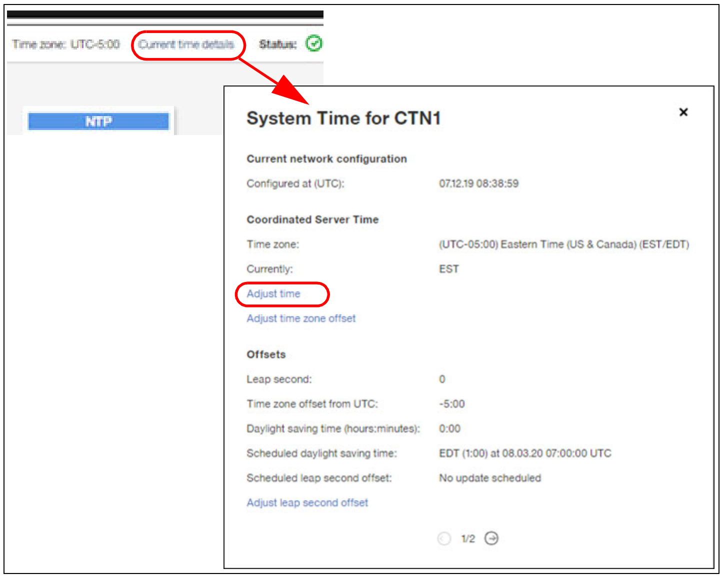


Figure 3-46 Adjusting the time

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 click **Apply**, the CTN time is adjusted to the ETS time, as shown in Figure 3-47.

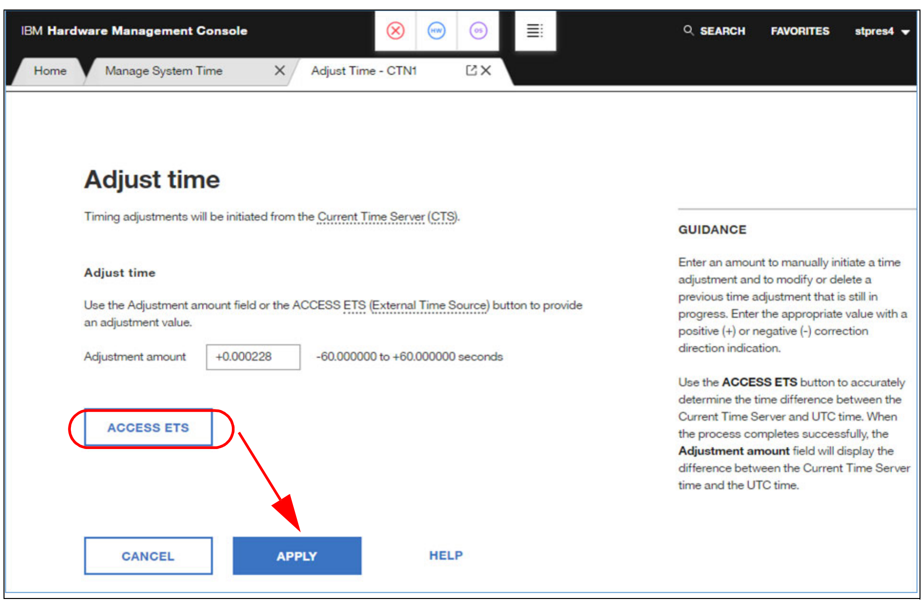


Figure 3-47 Applying the time adjustment

The following adjustments are possible:

- Adjustment steering
- Base steering

Adjustment steering

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

The offset that is specified is gradually incorporated into the standard timing messages in small enough increments or decrements that the operating systems (OSs), 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 1 second for every 7 hours.

Base steering

There is another time-steering method that is built in to the STP facility that works in a similar way to adjustment steering. This method is an automatic function that does not require user intervention. This method is known as base steering, which is performed at the CTS and requires an ETS.

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

Base steering is performed only if the server is an NTP stratum 1. The stratum level of the NTP server can be seen in the Manage System Time window by clicking the NTP server, as shown in Figure 3-48.

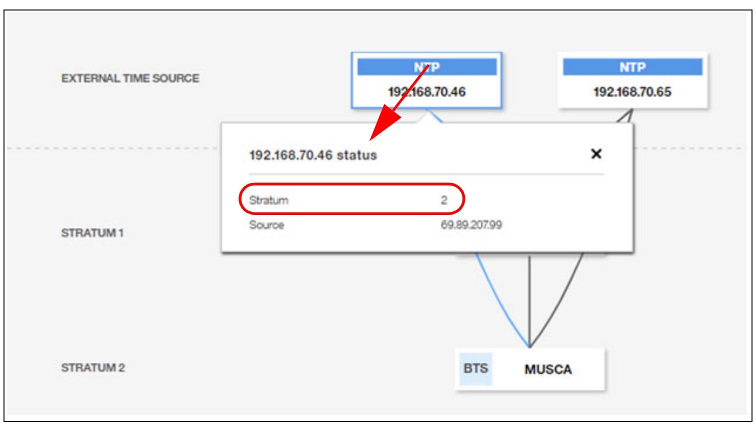


Figure 3-48 NTP stratum information

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

Table 3-1 STP time steering

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

Important: After a CTN is configured, the only way to set the date and time is to unconfigure the CTN and then go through the process to initialize the time again. For more information, see Chapter 4, “Managing Server Time Protocol configurations” on page 119.

The **Adjust Time** button is not available when the CTN time source is NTP with PPS. Because the STP uses PPS input every second, manual adjustment is disabled.

3.5.2 STP offset adjustments

The STP timing message includes the following items:

- CST
- Leap second offset
- Time zone offset
- DST offset

The values in this list are transmitted from the CTS to all the servers in the CTN. How the z/OS system image uses these values depends on options that are specified in the TIME macro with options that are specified in CLOCKxx at IPL.

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

Table 3-2 TIME macro options

Option	TIME macro with ZONE=LT	TIME macro with ZONE=UTC	TIME macro with STORE CLOCK
Include TOD in the result.	Yes	Yes	Yes
Include leap second offset.	Yes	Yes	No
Include a time zone offset.	Yes	No	No

In addition, the parameters that are 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

Option	STPMODE=NO STPZONE=NO	STPMODE=YES STPZONE=NO	STPMODE=YES STPZONE=YES
Step TOD to CTS.	No[a]	Yes	Yes
Include a time zone offset from the CTS.	No	No	Yes
Include time zone offset from CLOCKxx.	Yes	Yes	No
Allow local time adjustment through z/OS `SET` commands.	Yes	Yes	No

a. The base TOD steering affects all LPARs regardless whether they operate in STP mode or not because it changes the hardware TOD gradually without notifying the OSs. The STP mode corrects the difference when the local hardware TOD drifts and differs from Stratum 1 (the CTS or CST), which results in an STP sync check. The STP sync check is presented only to LPARs running in STP mode. Several other STP machine checks and external interruptions are also presented only to those LPARs running in STP mode.

Leap second offset adjustment

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

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

Although leap second adjustments usually are positive, it is theoretically possible to have a negative leap second adjustment.

> **Important:** Leap seconds are automatically built into the UTC time that is obtained from an ETS. Any leap second offset that is defined is accounted for when calculating the delta between the CTS and the time that is received from the ETS to prevent double accounting.
>
> If the ETS is used to incorporate an extra leap second into the CTS TOD, then the requirement that leap second adjustments should occur simultaneously worldwide has been breached. This situation occurs because TOD adjustments through the ETS are implemented over a period through 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 30 June or 31 December. However, the offset adjustment might be scheduled for a particular date, depending on your requirements.

A positive leap second is inserted after UTC 23:59:59, as UTC 23:59:60, and then UTC 00:00:00 occurs 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 (at the time of writing, it has not occurred) and is implemented by excluding UTC 23:59:59.

When the leap second offset occurs, z/OS is interrupted 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 window is used to apply the leap second offsets when they become available (see Figure 3-49). To do this task, click **Adjust Leap Second offset** from the Current time details window (over the graphical topology view) in the **Manage System Time** tab.

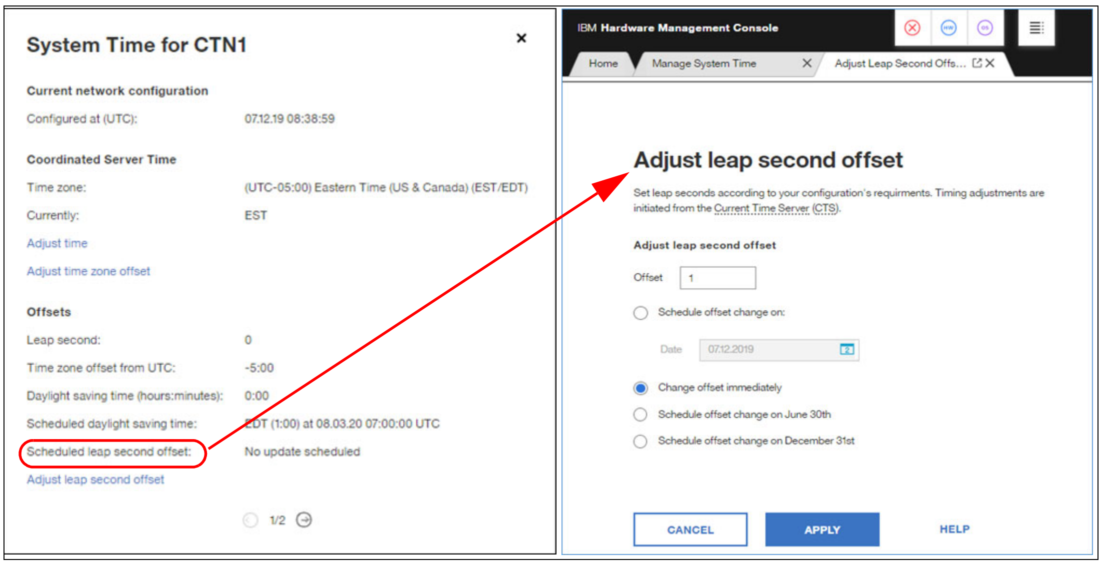


Figure 3-49 Leap seconds adjustment

Attention: The value of “1” for the leap seconds is only for display purposes of this scenario. Use a correct value for the leap second offset (currently, the value for leap seconds is 27).

When the Adjust Leap Second Offset window opens, 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 leap second (negative or positive), you must add or subtract this leap second from the value that is displayed in the window.

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

A negative leap second appears to any application that 1 second 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 window.

Important: Leap second adjustments occur simultaneously worldwide (regardless of the time zone) when UTC 23:59:59 occurs. STP and the OS take care of the adjustment (no user intervention or interaction is needed).

Time zone offset adjustment

In the Current time details window (over the graphical topology view) in the **Manage System Time** tab, you can see the set time zone offset.

You can change the time zone by using the **Adjust Time Zone Offset** tab. 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 3-50).

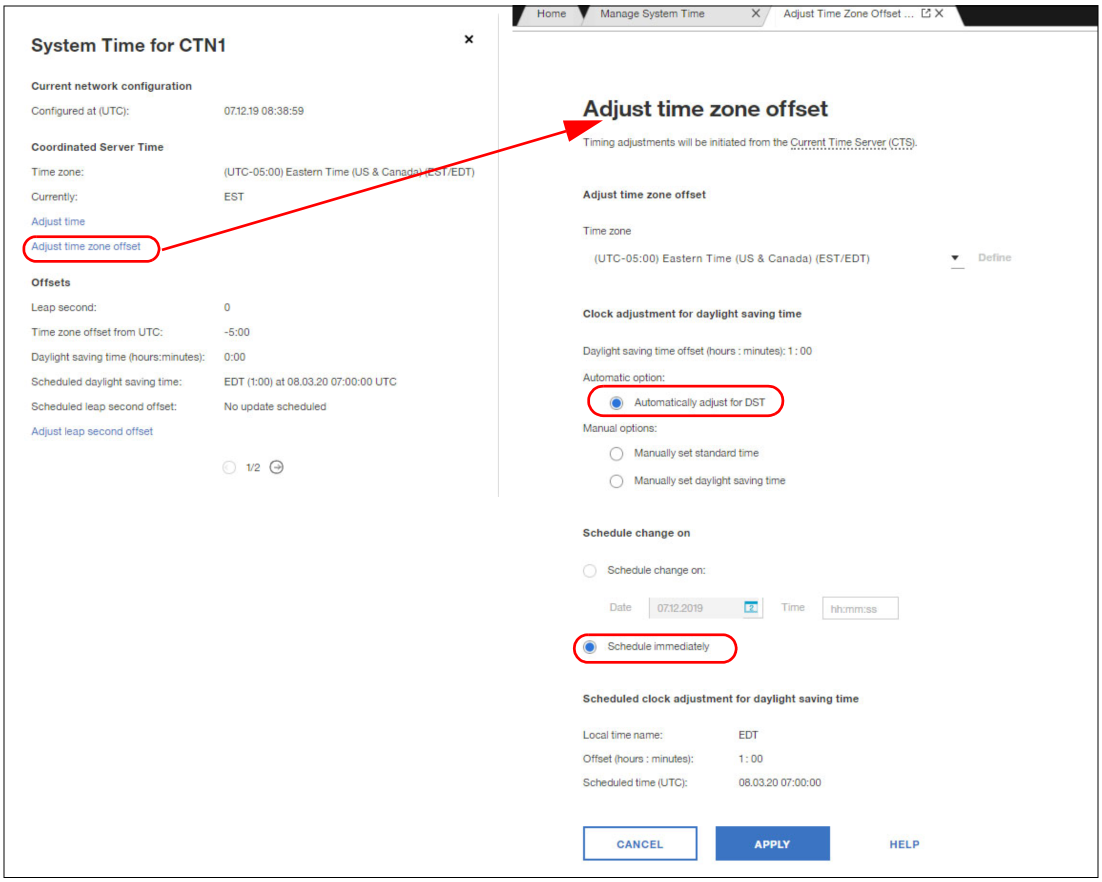


Figure 3-50 Time zone adjustment

Supported time zone offsets

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

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

In addition, each entry can optionally have a time zone algorithm that is 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** button can be used to activate automatic adjustment for the site.

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

The Scheduled Clock Adjustment for daylight saving section is displayed if an adjustment is 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 offset entry for the time zone meets the requirements. However, you can check this situation indirectly by applying the time zone offset entry for the time zone with “automatically adjust”, and then review fields in the Schedule Clock Adjustment for daylight saving section for correctness in relation to the next time zone adjustment.

Daylight saving changes that are implemented by a government authority

An example of daylight saving changes that are implemented by a government authority is the US Energy Policy Act of 2005 that changed DST in the US in 2007.

For an STP CTN, which has the option of automatically scheduling DST, support for the new DST start and end dates is available if the server has the correct level of microcode.

Future DST changes that are implemented by any governmental authority might require MCL updates. Update the server MCLs as needed.

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 - UD5) can be used to define the time zone. If a user-defined time zone entry is selected, the **Define** button is enabled (Figure 3-51).

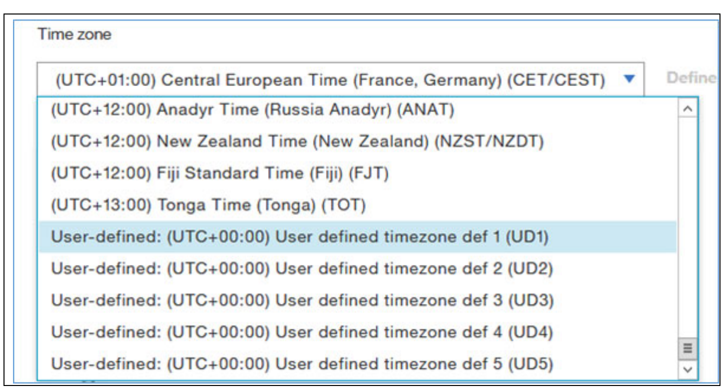


Figure 3-51 User-defined time zone entries

The Description (maximum of 80 characters) and Standard time name fields (maximum of 4 characters) must be entered, or an error message appears when you click **OK**. The standard time name is an abbreviation that is displayed on various windows to differentiate standard time from daylight saving.

The UTC offset must be entered in +/- hours and minutes, and it ranges -14 to +14 hours.

Also, if the time zone is subject to daylight saving adjustments, then the daylight saving name and daylight saving offset must be specified. Optionally, algorithms for DST start and DST end can be defined to support automatic clock adjustment by selecting the "Define adjustment of clock for daylight saving" option. The algorithm is saved when you click **OK**, but it is not sent to STP until you click **Apply** on the Adjust Time Zone Offset window.

3.5.3 Changes in local time

z/OS allows the user to obtain either STORE CLOCK (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`** option in the CLOCKxx member.

> **Note:** In z/OS, **`ETRMODE`** can be set in the CLOCKxx member, which requires a connection to a Sysplex Timer. In terms of hardware, connecting the mainframe server to a Sysplex Timer is no longer possible since IBM z196. Therefore, **`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`**.

The **`TIMEZONE`** parameter in the CLOCKxx member is used to set the time zone offset at IPL, and several z/OS **`SET`** commands can be used to dynamically adjust the offset when required. Similarly, the CF supports the concept of time zone offset and allows dynamic modification of the time zone offset through a command.

z/OS commands

On a z/OS system, the local date and time can be modified dynamically. The ability to do this task depends on what options are 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 through z/OS command		
	Local time ZONE=LT	**UTC time ZONE=UTC**	**STCK time STCK**
`STPMODE=NO, STPZONE=NO`	Yes	Yes[a]	Yes[a]
`STPMODE=YES, STPZONE=NO`	Yes	No	No
`STPMODE=YES, STPZONE=YES`	No[b]	No	No

a. UTC time cannot be changed by a z/OS command. It can be modified during the IPL processing response to the following message:

```
IEA888A   [UTC DATE=yyyy.ddd,CLOCK=hh.mm.ss] LOCAL
        DATE=yyyy.ddd,CLOCK=hh.mm.ss
```

REPLY U, OR UTC/LOCAL TIME 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 shown in Example 3-4.

Example 3-4 The 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 the DST offset, if any) that is applicable to the time zone.

SET DATE

This command is used to change the local date. Here is the syntax:

```
SET DATE=yyyy.ddd
```

The command has the following restrictions:

- *yyyy* is the year. It must be in the range 1900 - 2042. The value that is specified must consist of four digits and must be within 70 years of the UTC date, or 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. Here is the syntax:

```
SET CLOCK=hh.mm.ss
```

This command is used with the **SET DATE** command to set a maximum value of 23.53.47 on 2042.260. This command does not update the server's TOD clock or the logical TOD of the LPAR on which this z/OS image operates. The change that is made by this command is effective during 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 resets the time zone offset to the value that was read in from the CLOCKxx member during IPL, therefore changing the local date and time. Here is the syntax:

```
SET RESET
```

This command annuls all previous **SET DATE**, **SET CLOCK**, and **SET TIMEZONE** commands, and reestablishes the relationship between the local date and time and the UTC date and time + time zone offset.

SET TIMEZONE

This command changes the time zone offset to a different value from what was specified at IPL through the **TIMEZONE** parameter in CLOCKxx. This command automatically adjusts the local date and time. Here is the syntax:

```
SET TIMEZONE={W|E}.hh[.mm]
```

The time zone offset direction is West (W) or East (E). West is the default. The value for *hh* must be 00 - 15, and the value for *mm* must be 00 - 59.

The DST changes can be handled manually by using the **SET CLOCK** command instead of STP doing them automatically. When you use the manual method, there is always a degree of error because the difference between the local time and UTC time does not exactly match the time zone offset that would have been achieved by updating the **TIMEZONE** statement in CLOCKxx and then running an IPL.

The new z/OS **SET TIMEZONE** command overcomes this problem by applying the correct offset value to the CVTTZ field, which causes an exact time zone offset to be applied.

> **Tip:** If the **SET CLOCK** command is used to change local time for DST offset purposes, then use the **SET TIMEZONE** command instead for better accuracy.
>
> If the time zone offset for DST is changed dynamically by 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 does not regress at the next IPL.

Coupling Facility commands

In a Parallel Sysplex environment, CFs require time awareness to support CF request time ordering. The server TOD is used for this purpose.

CFs also support the concept of time zone offset, which is used only for timestamping messages that are displayed on the console.

There is no CFCC command that is available to display time. However, all messages that appear on the CF console include a timestamp 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 timestamp 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 that is used by the CF. Here is the syntax:

```
DISPLAY TIMEZone
```

This command produces a single line indicating how many hours and minutes the current time zone is east or west of GMT (Example 3-5).

Example 3-5 The 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 to change the time zone offset, if needed. Here is the syntax is:

```
TIMEZone {0|hh|hh:mm|:mm} {East|West}
```

Use the command that is shown in Example 3-6 to adjust the local time as shown in messages on the CF console for the onset and removal of DST.

Example 3-6 The CFCC command to adjust the time zone

```
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 with Daylight Saving Time changes

When a CF image partition is activated and the server is using an STP source, the CFCC uses only one of the following time offset options:

- The LPAR time offset that is specified in the image profile. The **TIMEZ** offset overrides the LPAR time offset.
- Use the **TIMEZ** command for DST changes, as described in Coupling Facility - TIMEZ Command during Daylight Saving Time Changes.

3.5.4 Logical partition time offset

IBM Z servers support a function that is called *LPAR time offset*. LPAR time offset supports the optional specification of a fixed time offset (specified in days, hours, and quarter hours) for each LPAR activation profile. The offset, if specified, is applied to the time that an LPAR receives from STP.

Sometimes, you must run multiple Parallel Sysplexes with different local times and run with the time set to TOD Clock LOCAL. This situation causes the results that are returned in the STCK instruction to reflect local time. With LPAR time offset support, LPARs on each server in a Parallel Sysplex can specify an identical time offset that shifts time in the LPAR sysplex members to the local time. The LPAR time offset value, if specified, is applied to the time value that an LPAR receives from the time source. The time zone offset and DST offset are independent of this parameter.

The remaining LPARs on the servers can continue to participate in current date production Parallel Sysplexes that use the same STP messages with the time that is provided by STP. This function is supported by all supported releases of z/OS.

To set the LPAR time offset, complete the following steps:

1. 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 3-52).

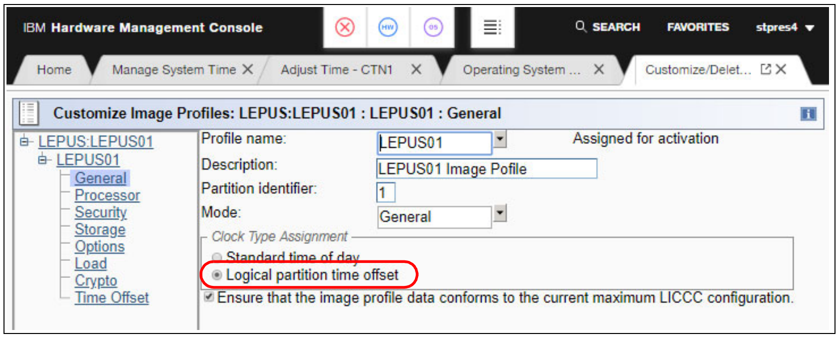


Figure 3-52 Customize Image Profiles: General

2. A new selection that is called **Time Offset** becomes available at the left. Use the Time Offset window to set the offset, as shown in Figure 3-53.

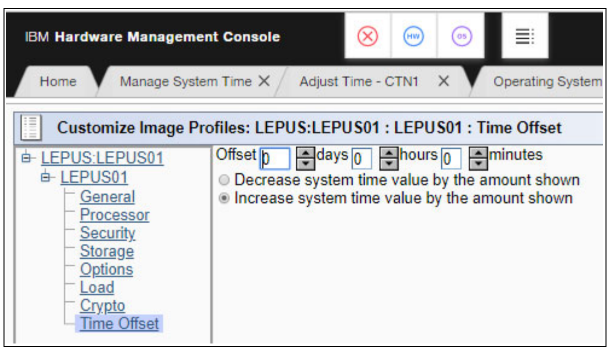


Figure 3-53 Customize Image Profiles: Time offset

The settings that are 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 that is supplied by its time source. The offset can be in the following range:

- 0 - 999 days
- 0 - 23 hours
- 0, 15, 30, or 45 minutes

– **Decrease system time value by the amount shown**

Select this choice to set the LPAR's clock back from the time of day that is 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 LPAR's clock ahead of the time of day that is 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.

3.5.5 Parallel Sysplex and multiple time zones

Some Parallel Sysplex installations have a requirement where different z/OS sysplex members undergo IPL in different time zones. You can meet this requirement by using the following means:

- 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 by using **STPZONE YES** while other sysplex members do not.

This section describes how to create and manage a multiple time zone environment.

Creating the multiple time zone environment

The example in this section reflects a consolidation of multiple sysplexes into a single sysplex. Although this action in itself is not remarkable, what is unusual is the fact that the sysplex supports applications that are intended to operate in images that support different time zones.

The following issues must be addressed when creating the environment:

- IBM software considerations
- Vendor software considerations
- Configuration considerations

IBM software considerations

Producing a sysplex with staggered time zone offsets introduces many peculiarities into the environment that are not acceptable to everyone. However, there might be business reasons why this configuration is favored, such as sysplex consolidation efforts or business consolidations. Here are some of the peculiarities:

- 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 LPAR 1A, obtains a job reader start time in one time zone, and runs on another image, with job start and end times in the other time zone.

- RMF considerations

  RMF displays and reports are suspect for the same reasons that are cited under SMF.

- OPERLOG considerations

  Although the records are ordered by UTC (Greenwich Mean Time (GMT)) timestamps, they are displayed with local time offsets. For example, in a merged OPERLOG, messages from the two systems that are 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 cause actions to be taken at different times. For example, the commands run *N* hours earlier on one system than the other.

The system functions properly, but the external view complicates the role of some people in the enterprise.

Vendor software considerations

Because a sysplex allows work to be routed to any machine in the configuration, programs asking for local time can get various answers depending on the server where the work was routed. To help in this situation, there is software that is available that provides an altered date and time to users, user applications, subsystems software, IBM 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 (IBM 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 MROs, CICS functions can be separated into individual regions, with the different types of CICS regions classified as resource managers. With the latest enhancements to MRO, these CICS resource manager regions can be 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 you must define multiple CICSplexes, and so manage the enterprise CICS systems as though they belonged to more than one enterprise.

Much of the CICSPlex System Manager activity is time-dependent. For example, you can specify that a monitor definition or an analysis definition is active during a particular period. CICSPlex System Manager does not require every instance in a single CICSplex to be running in the same time zone, and so it must be able to accommodate any time zone differences between entities:

- Whenever a time period definition (by using the CICSPlex System Manager PERIODEF view) is created, a time zone must be specified in the definition. For example, a time period definition that is called MORNING can be created for the hours 08:00 - 11:59 eastern standard time.
- A time zone must be specified for each CICSPlex System Manager address space (CMAS) in its data repository initialization job, which is called EYU9XDUT. A permanent change to the CMAS time zone value can be made, even while the CMAS is running, by using the CICSPlex System Manager user interface.
- A time zone must be established for each managed CICS system. When a CICS system is defined in a CICSPlex System Manager, 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 CMAS to which it connects. Allow the time zone of a managed CICS system to default to that of its CMAS. The time zone of a managed CICS system can be altered later while the CICS system is running. Any change that is 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 System Manager monitor function to determine the actual time at which the monitor interval for the CICSplex expires. The CICSplex time zone can be altered by using the CICSPlex System Manager user interface.

Time zones are specified by using single-character codes B - Z. For example, code S represents Mountain Standard Time, and 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*, SC33-0784-02. CICSPlex System Manager allows offsets (known as time zone adjustments) in the range 0 - 59 minutes to be specified to accommodate regions that are not running in standard time zones. Also, DST can be specified.

Because multiple CICSPlex System Manager 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 might be in different time zones. CICSPlex System Manager always acknowledges the time zone of the CMAS. Suppose that the following conditions exist:

- The time period definition time zone is S.
- The CMAS time zone is B.
- The CMAS time zone is C.

Time zone C is used by the CMAS, and the CMAS makes any necessary adjustments between time zones B, C, and S to ensure that the time zone is acknowledged.

Managing a 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. For more information, see “SET TIMEZONE” on page 110.

Because each sysplex member in our example is in a different time zone, you must stagger DST changes. DST changes can be accomplished individually in each sysplex member by using the **SET CLOCK** command from the z/OS console.

The cities that are used in the examples that follow include Sydney in the southern hemisphere (during summertime) and London, Atlanta, and Los Angeles in the northern hemisphere (by using standard time).

Operational considerations

In Figure 3-54, each of the three cities accurately reflect the number of hours from the Greenwich meridian when DST would be in effect in the southern hemisphere. Sydney's **TIMEZONE** parameter is shown while DST is in effect. The London and Atlanta **TIMEZONE** parameters are shown while standard time is in effect. Each hemisphere is changing the offset in different directions. When the southern hemisphere is changing from DST to standard time (back), the northern hemisphere is changing from standard time to DST (a forward change).

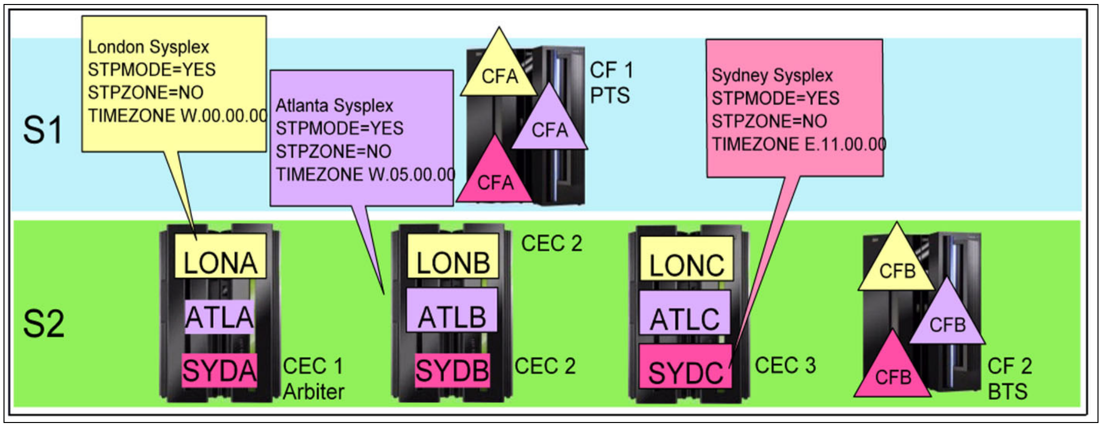


Figure 3-54 Staggered offset changes: one sysplex supporting multiple time zones

When planning for DST changes in different countries, not all countries agree on which dates that the switch occurs. In fact, each country's government might change the DST switch dates from year to year. A country that has multiple time zones might have a different change dates in each zone. In our example, assume that all countries are going either to DST (London and Atlanta) or to standard time (Sydney) on the same arbitrary day.

Also, assume that this sysplex is in Los Angeles, California. Any z/OS images representing Pacific Standard Time in the US are not shown in Figure 3-54. However, assume that those images contain **STPZONE YES** and that STP is being used to obtain DST offsets.

> **Note:** The world time zones that are shown in Table 3-5 represent only the local standard time offset from the Greenwich meridian and do not include DST 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
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. Because a new day dawns in Sydney before 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? Sydney is on DST in the southern hemisphere. Assuming that Sydney wants the offset changed at 02:00 AM. Sydney time, the Los Angeles operator must enter the command at 08:00 AM. After the offset change, Sydney's local time will be changed from 02:00 AM. to 01:00 AM. (back one hour).

Several hours later, the same thing must be done for the London sysplex members. Again, assuming that London changes the offset at 02:00 AM. London time, the Los Angeles operator must enter the command at 06:00 PM 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 DST (forward).

Sometime later, it is Atlanta's turn. Atlanta is also going from standard time to DST. The Los Angeles operator must enter the **SET CLOCK** command on Atlanta's z/OS console at 11:00 PM.

Finally, the offset for Los Angeles can be changed through the STP in the Adjust time zone offset window. These images have **STPZONE YES** in their CLOCKxx members.

> **Note:** Each customer situation is different. Certain ones are more complicated, and others less so. Planning and excellent communications with the user community are essential when DST schedules are 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 is 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.

4

Managing Server Time Protocol configurations

In this chapter, we describe several standard tasks for maintaining your Server Time Protocol (STP) environment. With the workflow that has been available since Hardware Management Console (HMC) 2.14.1 Driver 36, most of the tasks are guided.

> **Note:** The **Manage System (Sysplex) Time** option is no longer available on the Support Element (SE).

The following topics are covered in this chapter:

- Test environment overview
- Managing the Coordinated Timing Network configuration: Overview
- CTN reconfiguration operations

4.1 Test environment overview

Figure 4-1 shows the test environment that is used for the scenarios that are described in this chapter.

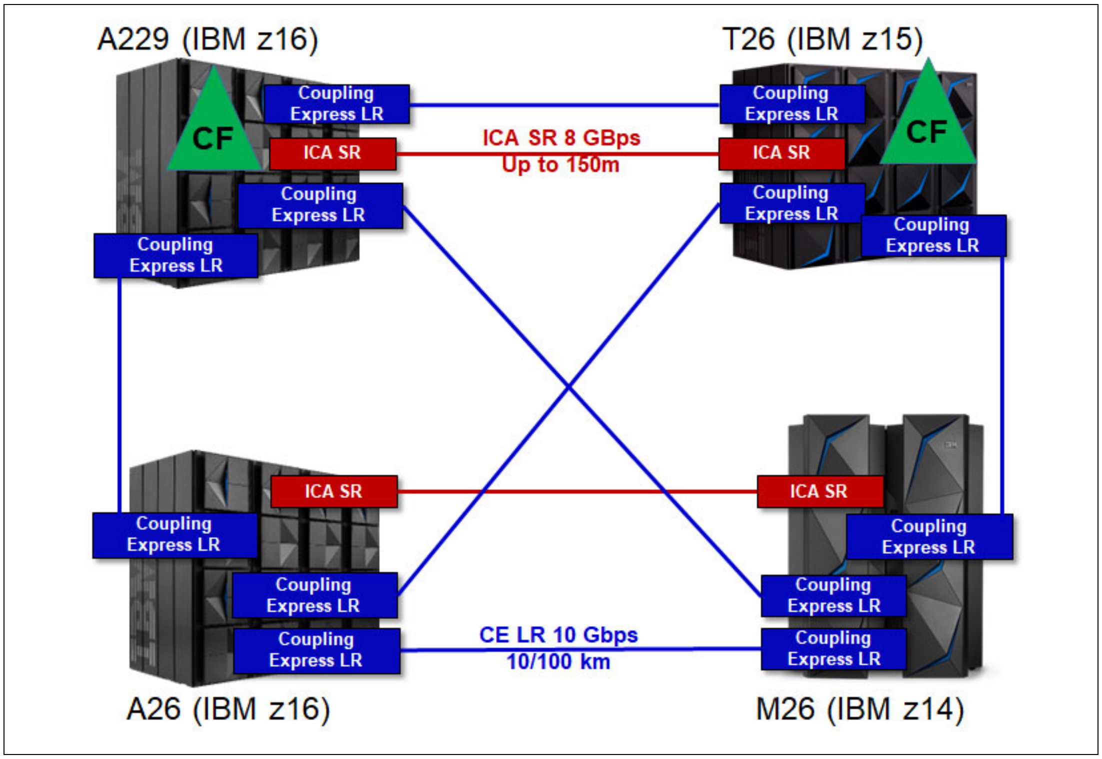


Figure 4-1 Test environment that is used for this chapter

Here are the four CPCs that are available:

- A229 (IBM z16)
- A26 (IBM z16)
- T26 (IBM z15)
- M26 (IBM z14)

4.2 Managing the Coordinated Timing Network configuration: Overview

All the functions that are described in this chapter are performed by using the “Manage System Time” tab on the HMC. To open this tab, go to the IBM Z HMC and select a CPC. Click **Systems Management**, and then select the CPC. Select **Configuration** → **Manage System Time** (see Figure 4-2 on page 121).

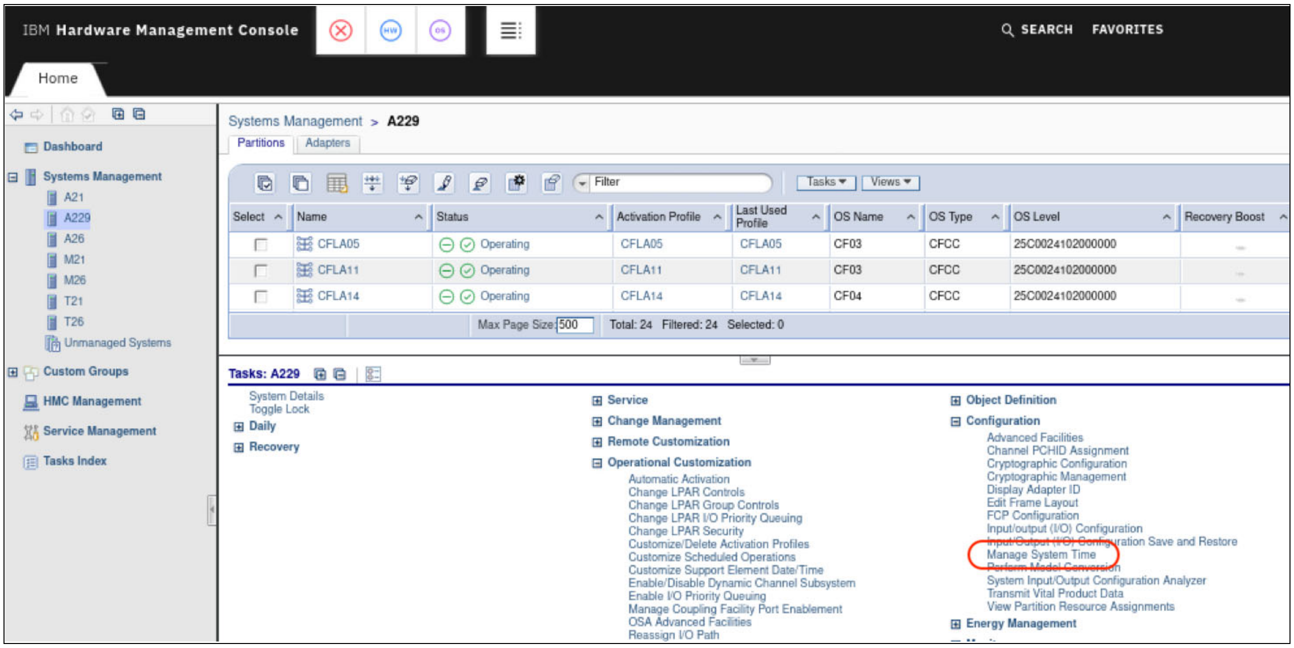


Figure 4-2 The Manage System Time tab

4.2.1 Changing the name (ID) of a Coordinated Timing Network

The name of a Coordinated Timing Network (CTN) can be changed in the **Manage System Time** tab. The change is non-disruptive to the systems operation. At the upper left of the **Manage System Time** tab, you can see the name of the selected CTN. To modify the name, click the pen symbol, as shown in Figure 4-3.

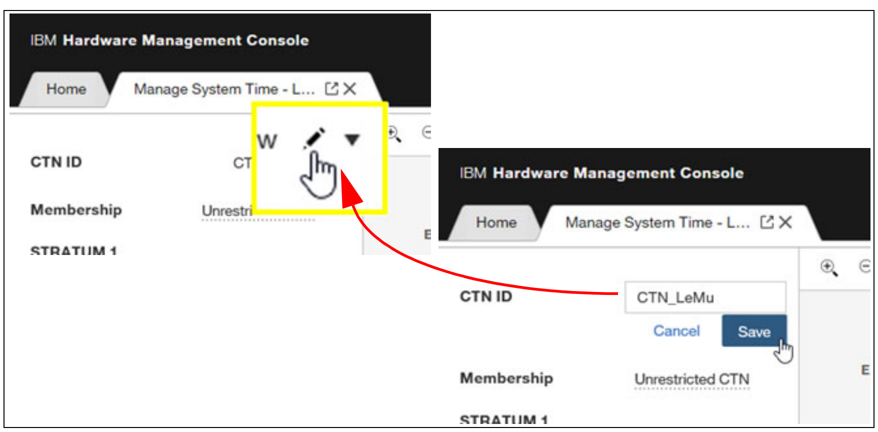


Figure 4-3 Changing the CTN ID (name)

A confirmation opens, as shown in Figure 4-4.

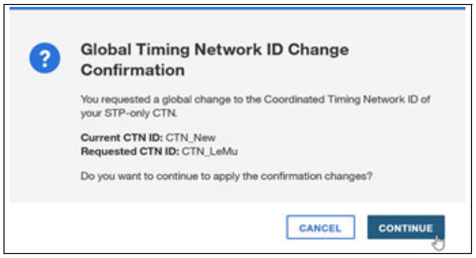


Figure 4-4 Confirmation dialog

Verify the changes and click **CONTINUE**. The rename operation success message opens, as shown in Figure 4-5.

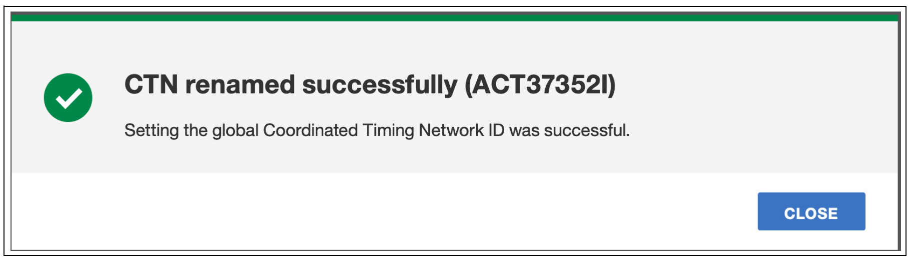


Figure 4-5 CTN ID change success

4.2.2 Single- or dual-server auto resume after power-on reset

If the CTN consists of one or two servers, it can be marked as “restricted”. Marking the CTN as “restricted” 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, as shown in Figure 4-6.

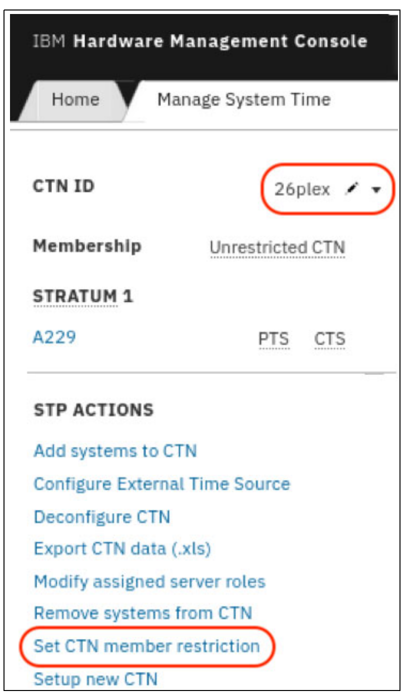


Figure 4-6 Restricting a CTN

Attention: Here are considerations for future CTN join, merge, and split operations:

- If either of the two CTNs that are being joined or merged is membership-restricted, the join operation fails and might lead to an outage. Therefore, to have a successful join or merge operation, neither of the two CTNs may be restricted.
- With a restricted or bounded CTN, you cannot add or remove a CPC from the CTN.
- A restricted CTN cannot be split.

In Figure 4-7, select **Allow only the servers that are specified below to be members of the CTN** and click **APPLY**.

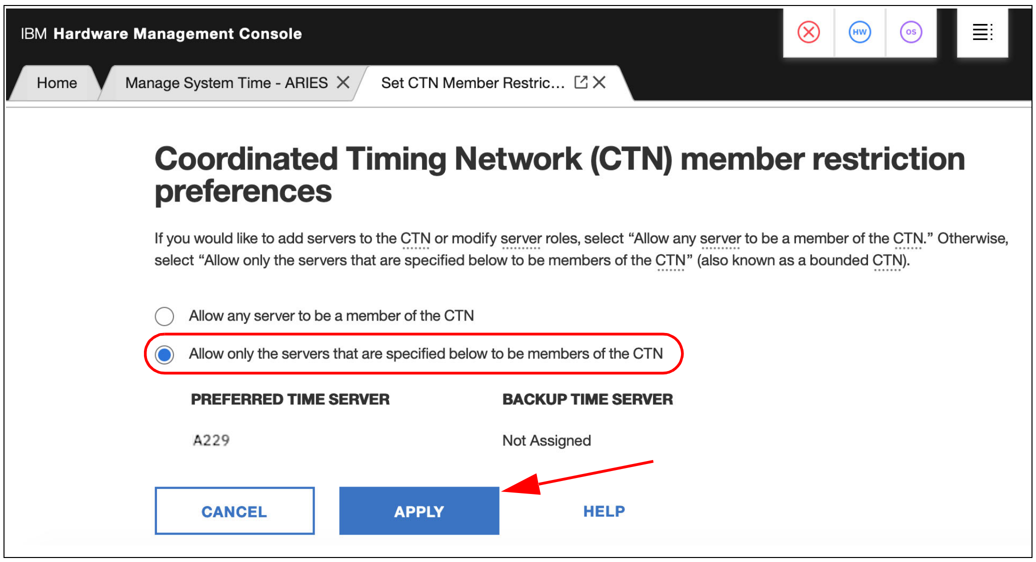


Figure 4-7 Applying a CTN restriction

Important: With the CTN membership restriction option set, no additional servers 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.

If the CTN is marked as unrestricted, the CTN must 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.

Important: A CTN with three or more CPCs cannot be membership-restricted. If a site is down and then power is restored, the (previously) unrestricted CTN must be reconfigured (the configuration is not saved across PORs).

After a POR, the CPCs that are defined in an unrestricted CTN are all stratum 0. To recover the strata, click the **Manage System Time** tab and open the "Setup new CTN" dialog. The configuration steps are described in 3.2, "Configuring the Coordinated Timing Network" on page 78.

4.2.3 Changing the Current Time Server

Changing the Current Time Server (CTS) from the Preferred Time Server (PTS) or Backup Time Server (BTS) is non-disruptive. To perform this change, complete the following steps:

1. Select the CTN ID that you want to modify, and then click **Modify assigned server roles**, as shown in Figure 4-8.

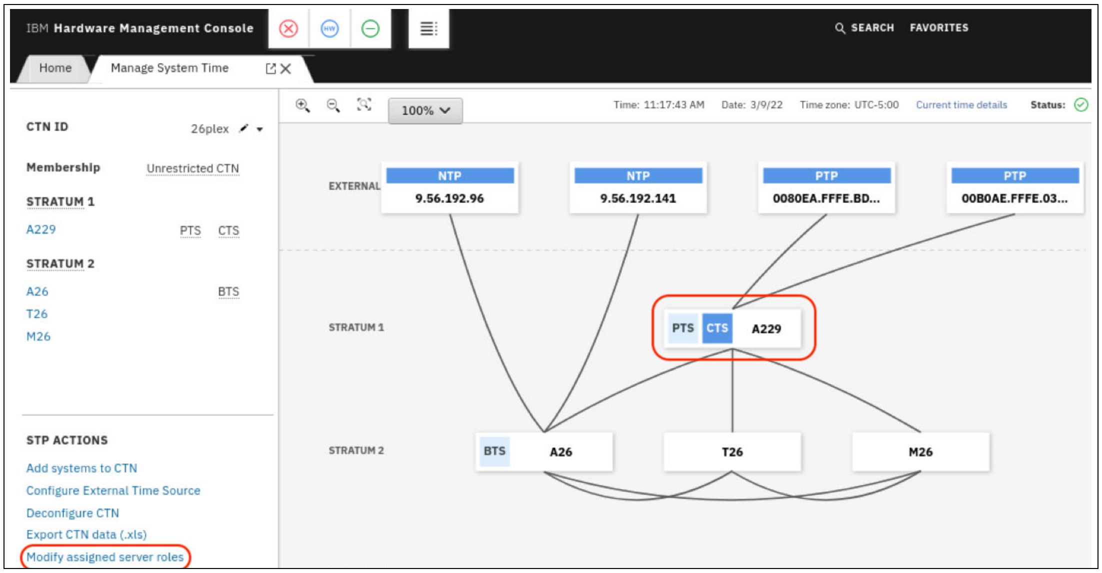


Figure 4-8 Modify assigned server roles menu

Figure 4-9 shows the Modify Assigned Server Roles dialog.

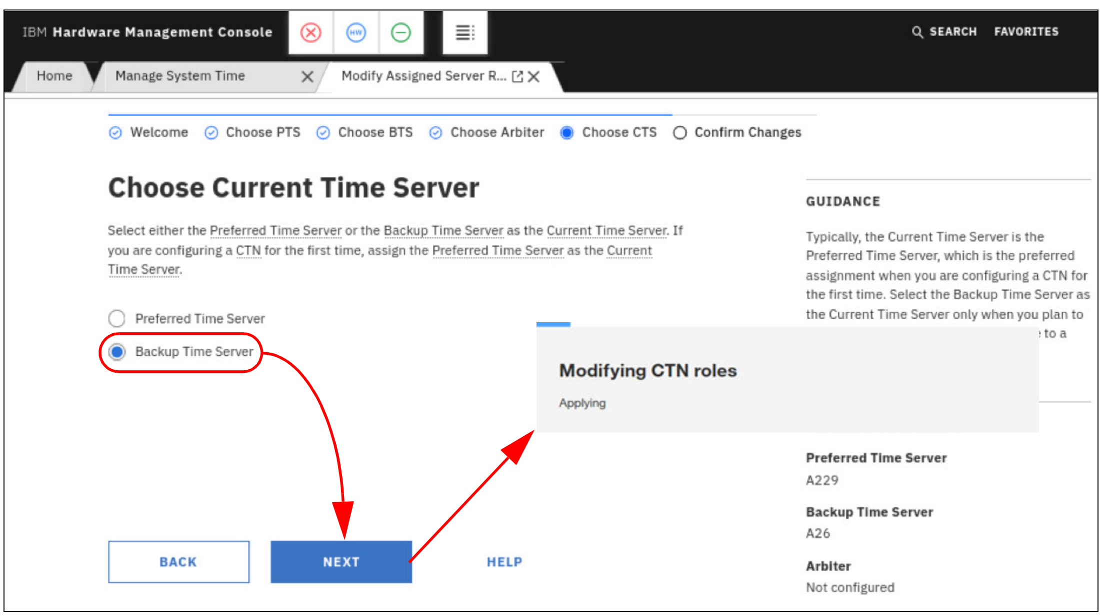


Figure 4-9 Modifying server roles

2. Confirm your selection by clicking **NEXT**. The changes are performed. A confirmation window opens.

   When the modification completes, the new configuration is displayed. The CPC that is named A26 is now the CTS, as shown in Figure 4-10 on page 125.

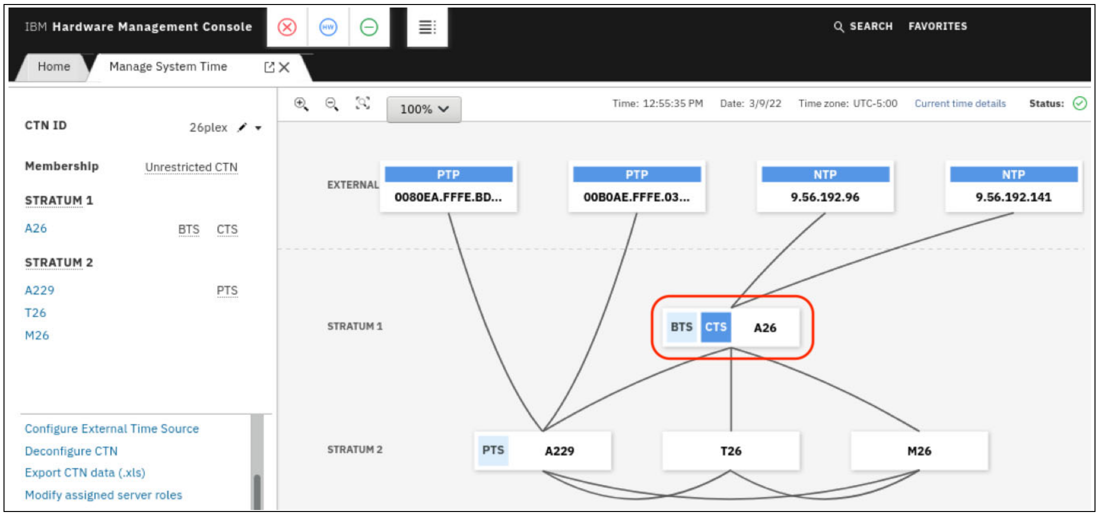


Figure 4-10 CTS role that is assigned to the BTS

4.2.4 Changing the server roles

Changing the server roles might be necessary if you plan to perform maintenance on a system with a role in a CTN. This HMC task is non-disruptive. In the following example, we swap the role of PTS and BTS, and we want the new PTS to be the CTS.

Note: To perform maintenance on a server, *it must NOT play* a role in the CTN.

Complete the following steps:

1. Select the CTN ID that you want to modify and click **Modify assigned server roles** under STP ACTIONS, as shown in Figure 4-11.

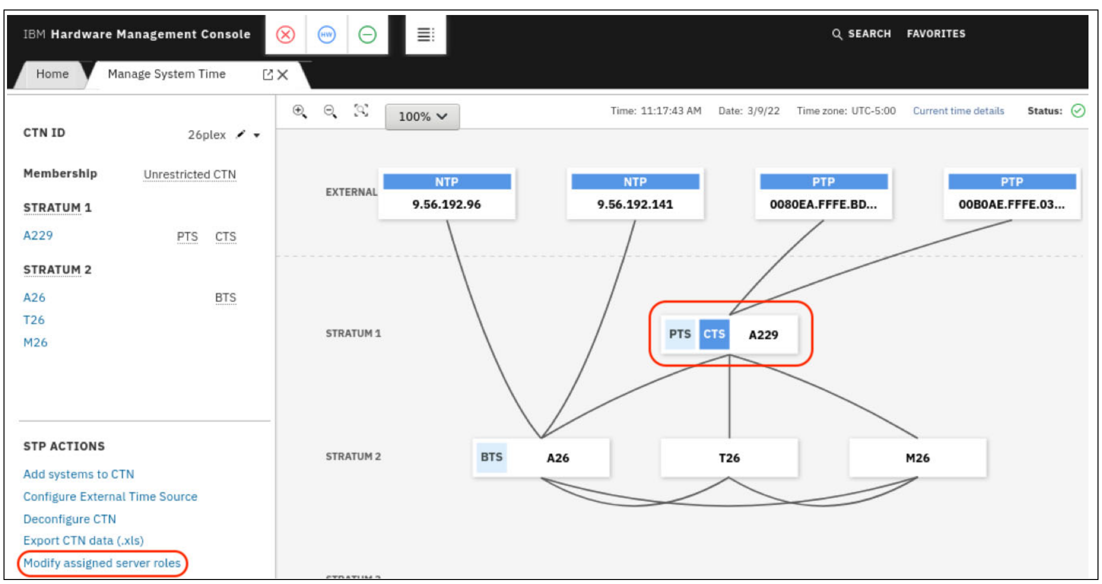


Figure 4-11 Modifying server roles

2. Select the server that you want to become the new Preferred Time Server (PTS). In our example, the role should switch from A229 to A26, so we select **A26** and confirm by clicking **NEXT**, as shown in Figure 4-12.

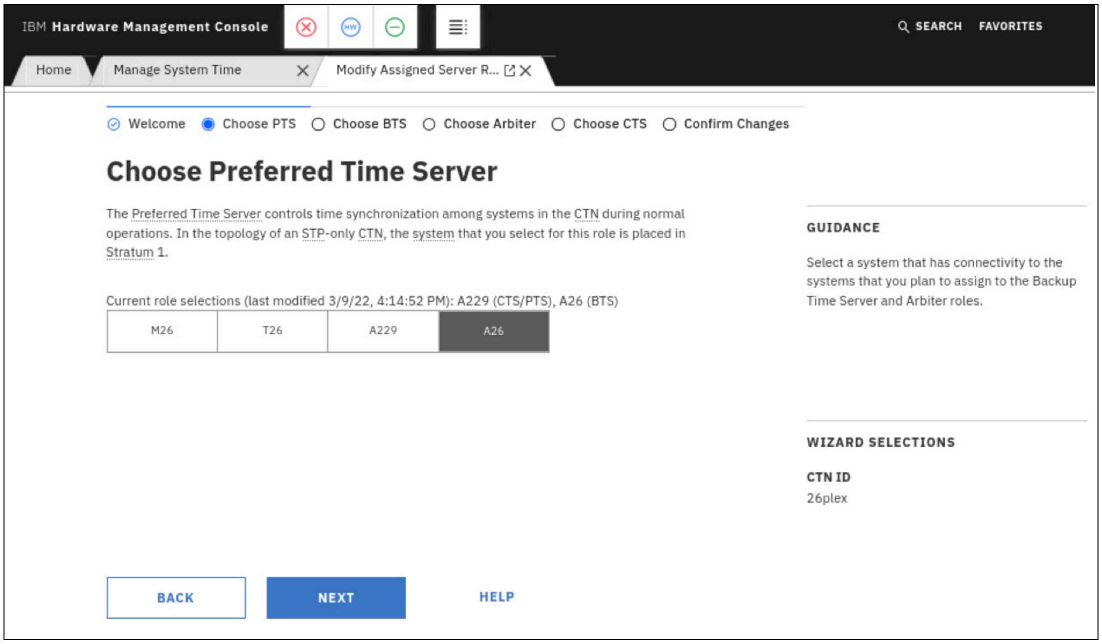


Figure 4-12 Selecting the new PTS

3. Select **A229** as the BTS and then click **NEXT**, as shown in Figure 4-13 (in our example, we swap the server roles).

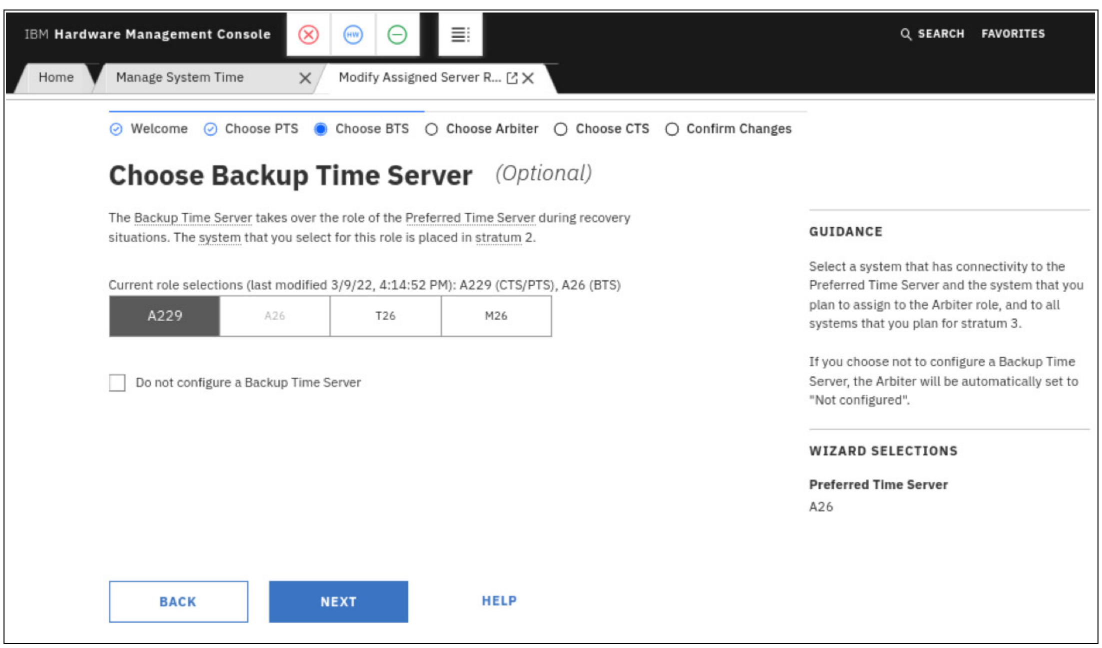


Figure 4-13 Selecting the BTS

4. Define the CTS. In our example, we chose A26 (the new PTS) as our CTS, so we select **Preferred Time Server** and click **NEXT**, as shown in Figure 4-14 on page 127.

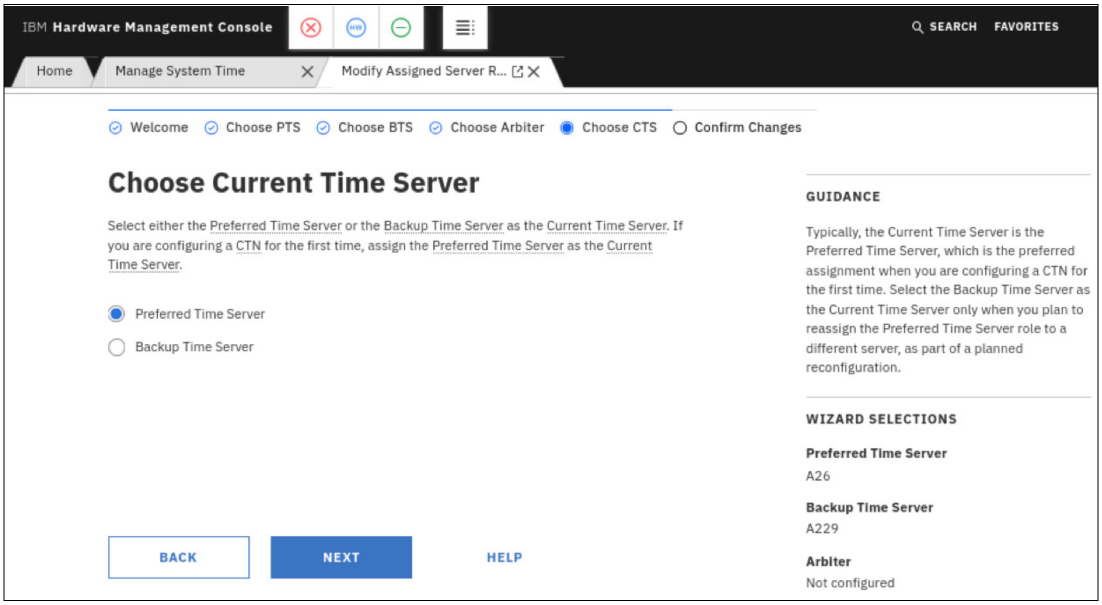


Figure 4-14 Selecting the CTS

5. When you click **NEXT**, the window that opens shows how the new configuration looks. If the proposed configuration matches the expectation, confirm by clicking **NEXT**. After the changes are successfully applied, the message ACT39295I is displayed and the new roles become 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 is deconfigured, which leaves the CPC in an unsynchronized state (STP stratum 0).

Complete the following steps:

1. Select the CTN ID to be deconfigured, and then click **Deconfigure CTN** under STP ACTIONS, as shown in Figure 4-15.

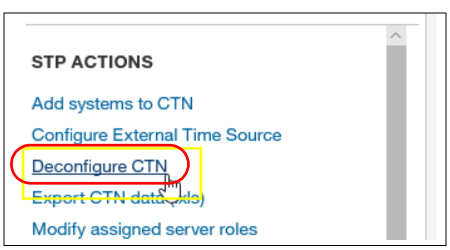


Figure 4-15 Initiating the CTN deconfiguration

> **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.

2. The next window opens and states that the **Deconfigure CTN** task is disruptive. Enter the password of the HMC user and click **APPLY**. After the progress window disappears, message ACT39285I confirms the successful execution of the **Deconfigure CTN** task. The CPC 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 action is *highly disruptive*.

4.3 CTN reconfiguration operations

In this section, we provide the following examples of CTN reconfiguration operations:

- Adding a server to a CTN and reassigning server roles
- Splitting a CTN
- Merging two CTNs

4.3.1 Adding a server to an existing CTN and reassigning roles

In this example, we add the system T26 to the existing CTN that is named 26plex. T26 is part of the inactive CTN t26plex.

Note: A server must be part of a CTN (active or inactive) to be added to an existing CTN.

Complete the following steps:

1. Select the CTN ID 26plex as the receiving CTN. Then, click **Add systems to CTN** under STP ACTIONS. A window with all the available systems and their corresponding CTNs opens.

   In our example, only the T26 CPC is available. Select **T26** and then continue by clicking **NEXT**, as shown in Figure 4-16.

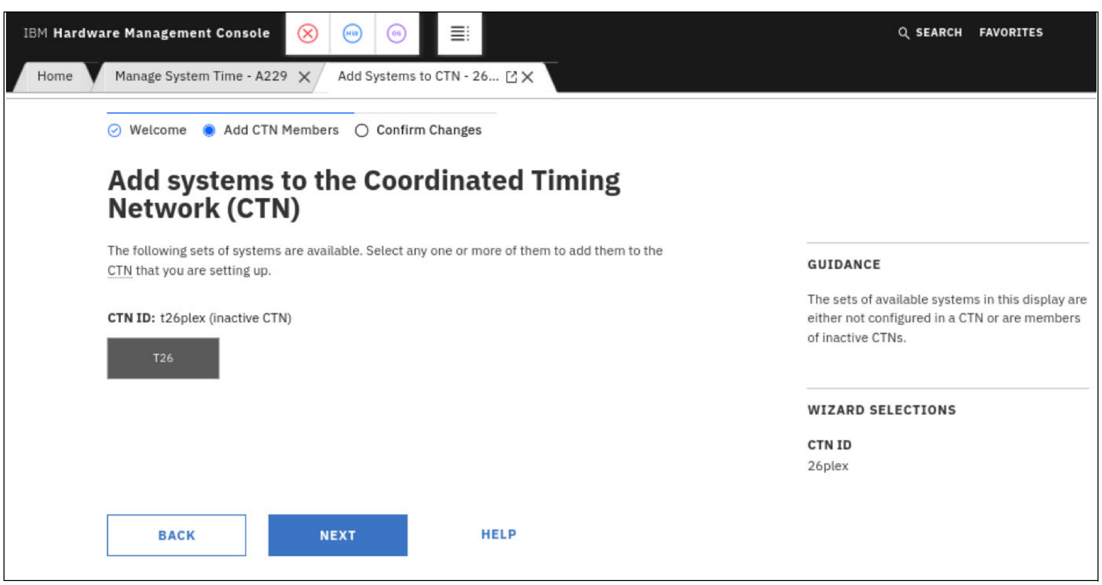


Figure 4-16 Adding a server to an existing CTN: Selecting the server

2. A window opens that shows the target configuration. Accept the configuration by clicking **APPLY**. The next window provides an overview of the systems that will join.

   The message that is shown in Figure 4-17 informs us that the CTN ID for the server that we selected to join the existing CTN will change to match the target CTN ID.

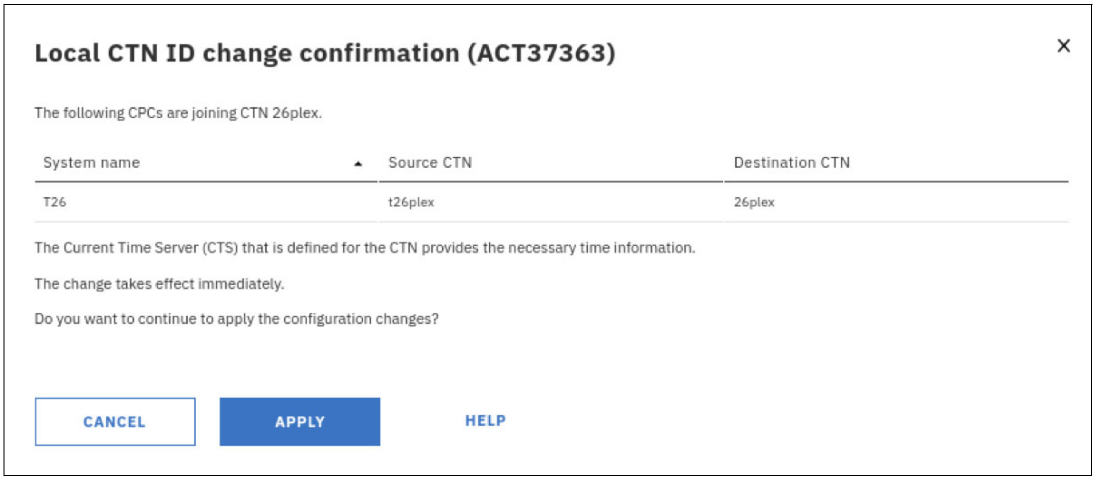


Figure 4-17 CTN ID change confirmation

3. Click **APPLY** and the changes are performed. The new configuration is active immediately. A window opens and confirms the activation with the message ACT39275I Systems added to CTN successfully.

4.3.2 Splitting a CTN

To split a CTN into two distinct CTNs, use the **Split to new CTN task** under the **Advanced Actions** menu under the STP ACTIONS pane. The systems to be split off the existing CTN must *not* have a role that is assigned. All these roles must be removed, as described in 4.2.4, “Changing the server roles” on page 125.

In this example, we split T26 off the CTN with the ID 26plex into the newly created CTN ID t26plex. T26 also becomes the PTS and CTS of t26plex. A229 remains the PTS and the CTS of the existing CTN 26plex.

Important: Before splitting a CTN, make sure that members of the new CTNs are properly interconnected through coupling links to other members of their respective CTN.

The current configuration of 26plex is shown in Figure 4-18.

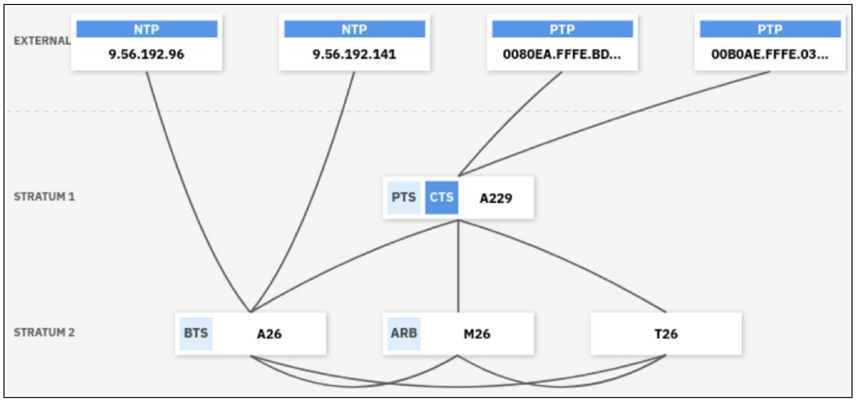


Figure 4-18 Starting point of the "Split CTN" scenario

> **Important:** Because it is not recommended to split a running sysplex, a warning message appears, as shown in Figure 4-19. Make sure that you properly reconfigured your sysplex workloads to avoid splitting sysplexes.

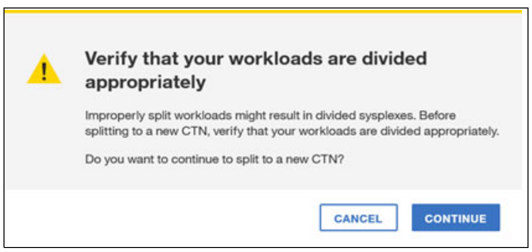


Figure 4-19 Warning message

Complete the following steps:

1. If everything in Figure 4-18 has been double-checked, click **CONTINUE**. On the next window, specify the name of the (new) CTN ID to create. In our example, the new CTN ID is named t26plex, as shown in Figure 4-20.

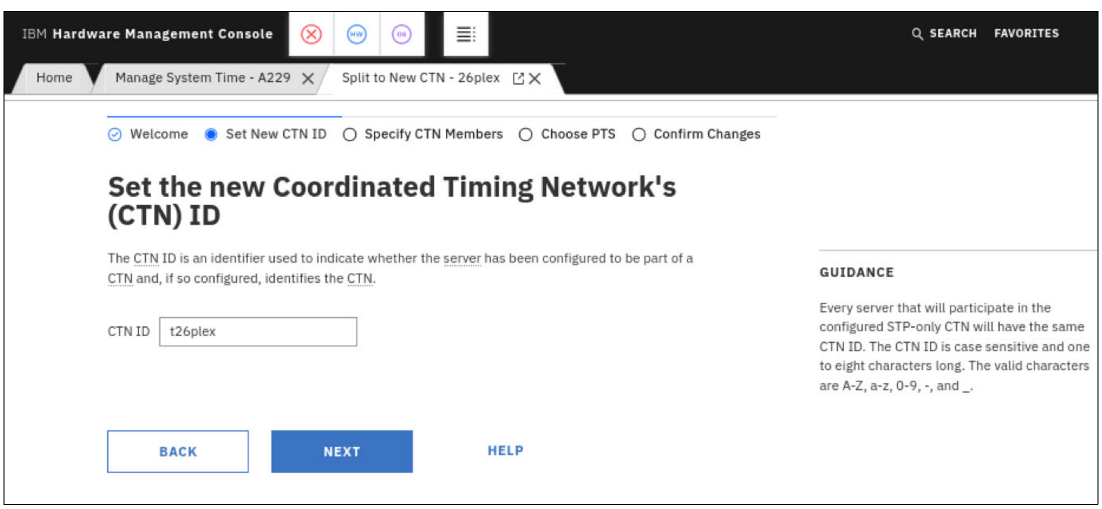


Figure 4-20 New CTN ID

2. 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 that is assigned. Remove the role from the server and start over. In our example, we choose to split T26 from 26plex, as shown in Figure 4-21.

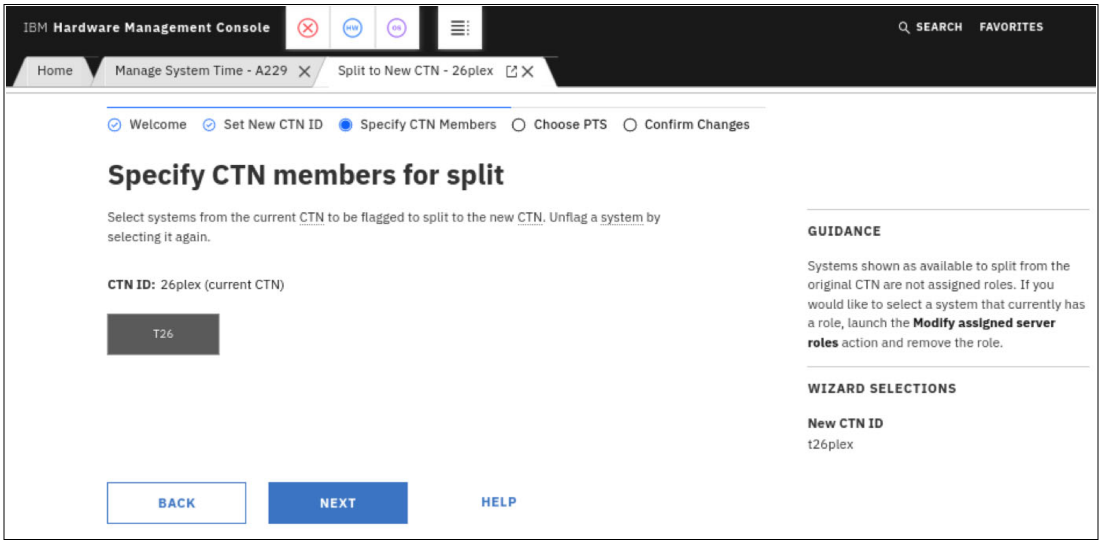


Figure 4-21 Servers to be split into the new CTN

Attention:

- If there are systems detected that are running in a sysplex with the same name that would be separated by the split, a warning message appears (see Figure 4-22 on page 132). Make sure that you really want to split before continuing.
- Also, if you are splitting multiple CPCs from their original CTN, make sure that they are actively interconnected through coupling channels.

Z

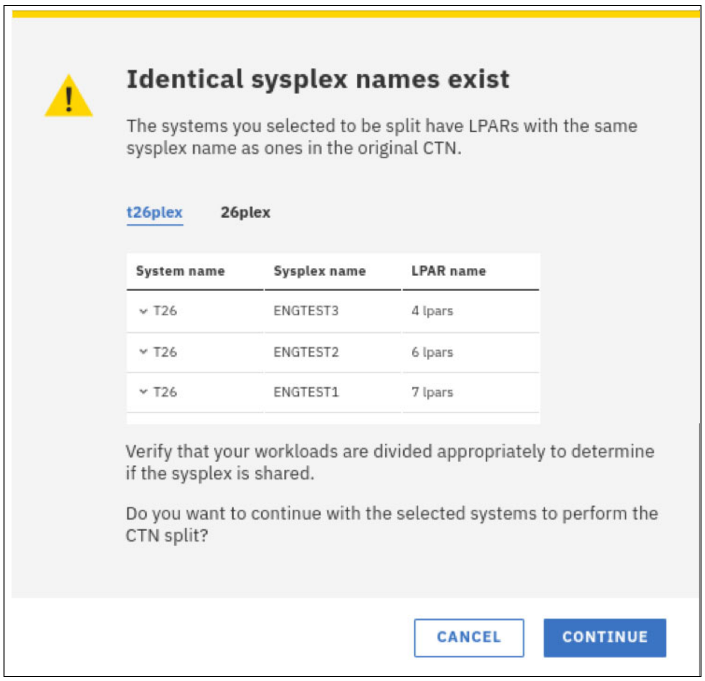


Figure 4-22 Final warning message before splitting the CTN

3. If everything has been checked, click **CONTINUE**.
4. On the next window, you can choose which system should be the PTS. In our example, we select **T26**, as shown in Figure 4-23 on page 133.

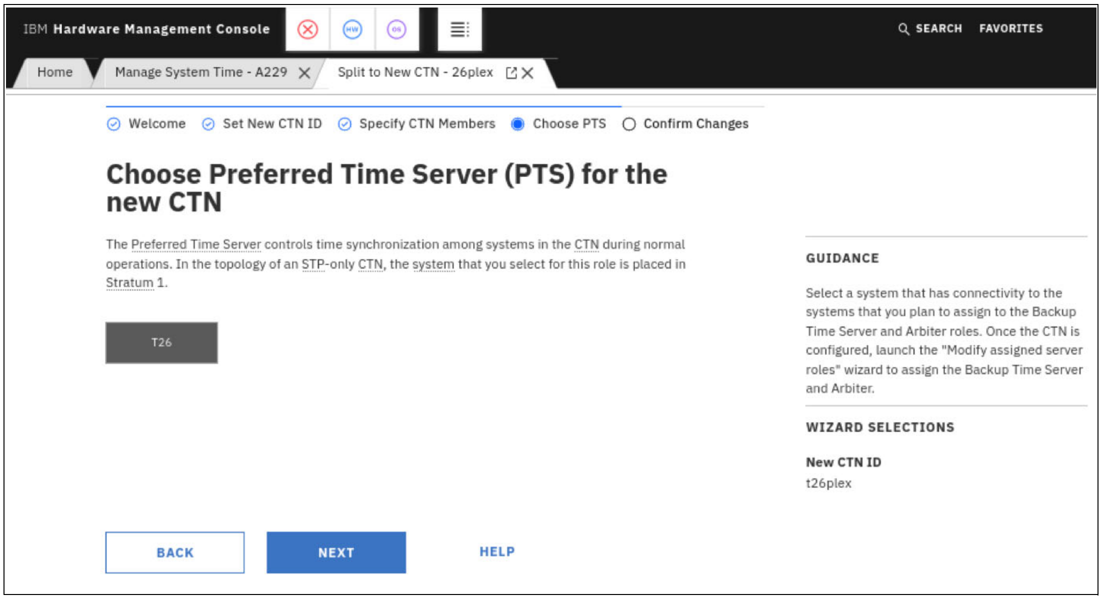


Figure 4-23 Server roles for the new CTN

5. On the next window, we verify both CTNs. If both CTNs match our expectations, we click **APPLY**. The changes are performed immediately. After a short period, a successful split is confirmed by the message ACT39329I “The CTN split task completed successfully”. You now have two valid CTNs.

4.3.3 Merging two CTNs

To merge two CTNs into one CTN, use the **Join existing CTN** task under **Advanced Actions** in the STP ACTIONS pane. Make sure that you first select the CTN ID to be joined into an existing CTN. The CTN that is selected first to be joined will no longer exist after the successful execution of the merge CTN operation.

Attention:

- If either of the two CTNs that are being joined or merged is membership-restricted, the join operation fails and might lead to an outage. Therefore, to have a successful join or merge operation, neither of the two CTNs may be restricted.
- With a restricted or bounded CTN, you cannot add or remove a CPC from the CTN.
- A restricted CTN cannot be split.

Important: As a best practice, make sure that the Coordinated Server Time (CST) of the CTNs are close before you start a *merge* operation so that you can reduce the amount of time that is required to synchronize and merge the two CTNs. Otherwise, this procedure can take a considerable amount of time.

Attention: If the difference in the CST of the two CTNs being merged is *greater* than 1 minute, you cannot do a merge or join.

In our example, we merge t26plex into 26plex. The systems that are defined in t26plex will become part of the CTN with the ID 26plex, and *t26plex* will no longer exist.

The initial configuration is as follows:

- 26plex is a 3-server CTN with the members A229 (PTS and CTS), A26 (BTS), and M26.
- t26plex is a single-server CTN with the member T26.

Complete the following steps:

1. First, select t26plex as the current CTN ID. Then, start the merge CTN operation by clicking **Join existing CTN**. A new tab opens, where you select the CTN to join, as shown in Figure 4-24.

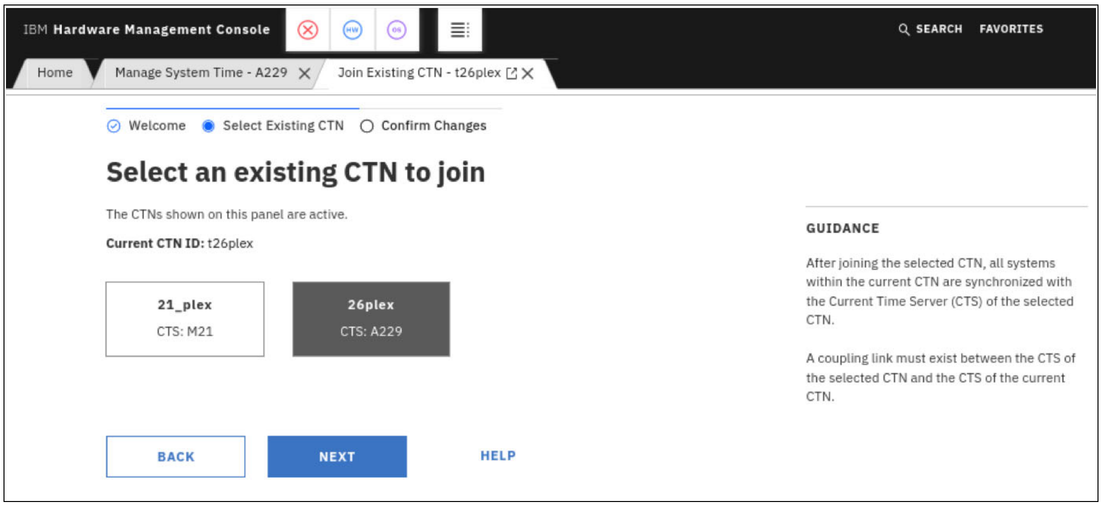


Figure 4-24 Initiating the merge operation

2. Select **26plex** and click **NEXT**. The Current CTN ID is t26plex. On the next window, verify the result. If you are satisfied with the result, click **APPLY**. The join process starts, and a progress window opens (see Figure 4-24).
3. When the join process finishes, another window opens and confirms that the merge was successful (see Figure 4-25). 26plex is now the current CTN ID, and t26plex no longer exists.

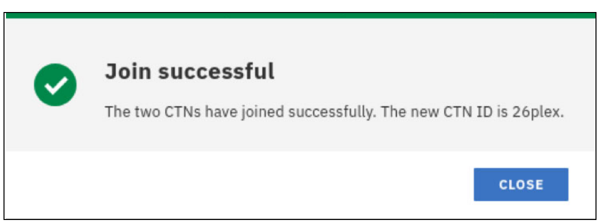


Figure 4-25 Successful join message

4.3.4 Removing a system from the CTN

To remove a system from a CTN, use the **Remove systems from CTN** task under the STP ACTIONS pane. The system to be removed must not have a role that is assigned. All roles must be removed by running the **Changing the server roles** task, which is described in 4.2.4, “Changing the server roles” on page 125.

In this example, we remove A21 from the CTN with the ID 21_plex.

Important: Before removing a system, make sure that the remaining members of the CTN remain interconnected through the coupling links.

Note: Removing a system from an active CTN results in the system becoming stratum 0 (unsynchronized) in an inactive CTN. This situation might result in an unintended service call. If the system remains in an active CTN, follow the steps that are described in 4.3.2, “Splitting a CTN” on page 129.

The current configuration of ID 21_plex is shown in Figure 4-26.

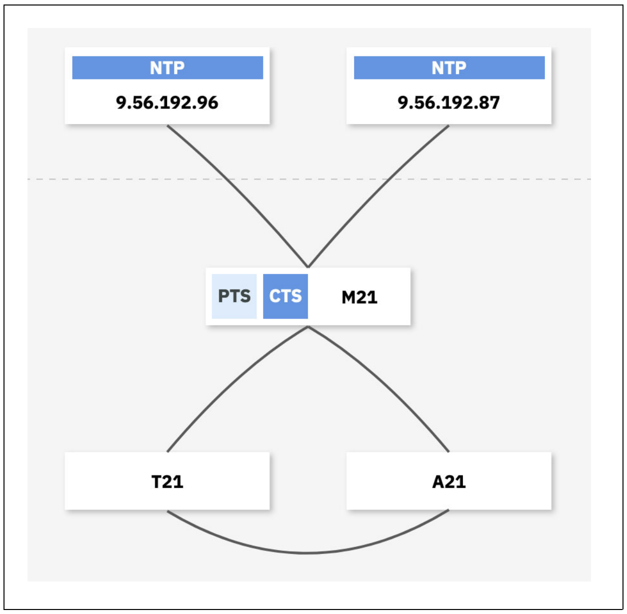


Figure 4-26 Initial ID 21_plex server configuration

The next window lists all members that are eligible to be removed from the CTN. If a system does not appear, it might still have a role that is assigned. Remove the role from the server and start over. We choose **A21**, as shown in Figure 4-27.

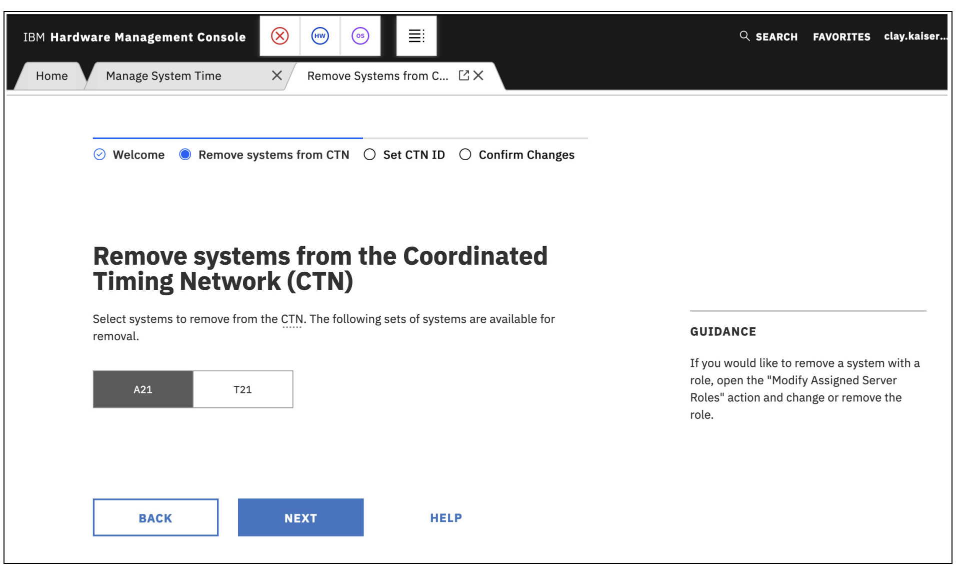


Figure 4-27 Selecting a system to be removed from the CTN

Complete the following steps:

1. In the window that is shown in Figure 4-28, specify the name of the new (inactive) CTN ID. In our example, the new CTN ID is named 26plex, as shown in Figure 4-28. Click **NEXT**.

> **Note:** The CTN of the systems that are selected in the **Remove system from a CTN** task is an inactive CTN. Any systems in this new CTN become stratum 0 (unsynchronized). If the removed systems should remain in an active CTN, follow the steps that are described in 4.3.2, “Splitting a CTN” on page 129.

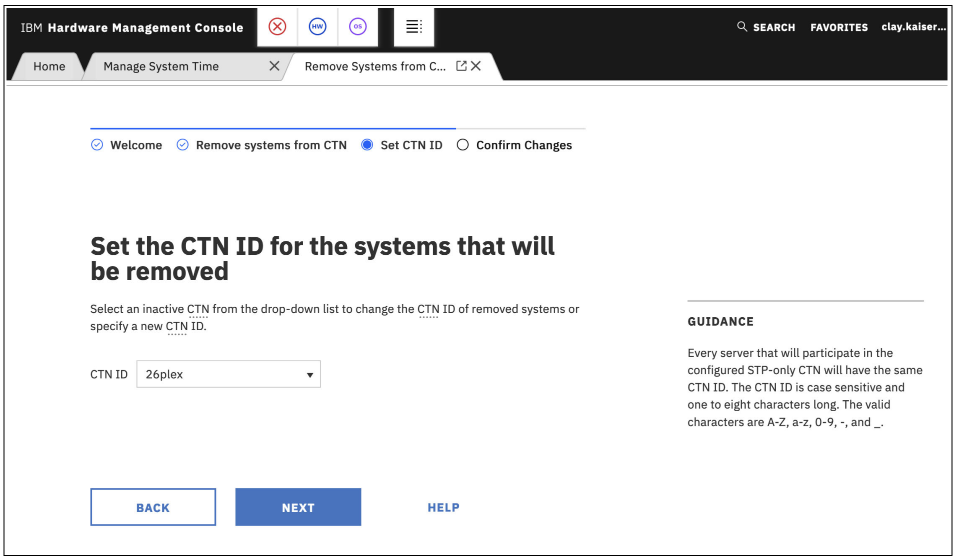


Figure 4-28 A new inactive CTM_id that is named 26plex is created

2. Verify the remaining CTN. If the CTN matches expectations, click **APPLY**. The changes are performed immediately. After a short period, a message appears and confirms that the removal was successful.

A

IEEE 1588 Precision Time Protocol introduction and overview

This appendix provides an introduction and overview of the Precision Time Protocol (PTP). The information that is provided introduces the PTP technology and the relevant standards; summarizes the regulatory requirements in the industry that led to IBM moving forward with PTP; offers a brief comparison with Network Time Protocol (NTP); and describes what clients can do to prepare their data center to use PTP.

The following topics are covered in this appendix:

- Introduction
- Background on PTP: History
- Background on PTP: Technology overview
- Regulatory requirements
- A comparison of NTP and PTP
- Preparing for PTP
- Conclusion

Introduction

As announced in the September 2019 IBM z15 product announcement's statement of general direction, IBM introduced IEEE 1588 PTP as an External Time Source (ETS) for IBM Z Server Time Protocol (STP) running in an IBM Z Coordinated Timing Network (CTN). The initial implementation provided PTP connectivity through a Support Element (SE) in the CPC frame. However, with IBM z16, PTP connectivity is established directly through a CPC drawer. Configuration and management of the ETS is performed by using the IBM Z Hardware Management Console (HMC).

For the implementation that was announced on 14 April 2020, there is no change to the usage of STP CTNs for time coordination, other than the potential to use a PTP-based ETS. Future implementation[1] plans to reintroduce of the concept of a Mixed CTN, with support for traditional STP (time synchronization across IBM Z CPCs through coupling or timing links) and native PTP implementations to enhance the role of IBM Z in a hybrid cloud environment.

Background on PTP: History

The NTP was introduced in 1985. During the 1990s, public discussion started on standardizing the technology and techniques for synchronizing clocks in devices that were not associated with NTP, such as devices that were typically used in measurement and control applications. At the time, these discussions were taking place between engineers developing systems that were 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 the formation of a committee that was 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 on18 July 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 that were built at Agilent Technologies under the leadership of John Eidson between 1990 and 1998. As conceived, PTP is capable of near-nanosecond precision while leveraging a network infrastructure that is similar to what NTP uses.

The IEEE 1588-2002 standard with more 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, that is, it is not possible to have a mix of Version 1 and Version 2 devices in the same timing network.

In June 2013, an IEEE Standards Association Project Authorization Request (PAR) was approved to revise IEEE 1588-2008. From 2008 - 2013, more industries with use cases beyond the ones that were associated with the early PTP implementations emerged. The stated purpose of the revision was to reflect the common needs of the various industries that used the standard to meet the requirements of more security in a precision clock synchronization protocol that is also suitable for higher speed applications in the sub-nanosecond range. It also was decided that PTP needed to support both IPV4 and IPV6.

[1] 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 that is described for IBM products remain at the sole discretion of IBM.

The revision was approved by the IEEE Standards Association in November 2019, and was officially published in June 2020 as IEEE V2.1, and as IEEE 1588-2019.

Background on PTP: Technology overview

The IEEE 1588 standard provides a scope for the standard. This scope summarizes the standard, its goals, and the technology. The latest scope (from IEEE V2.1) is summarized in the following section. This scope is the overall scope, so 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 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 through 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) and 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 that is external to the system, if time that is traceable to international standards or other sources 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. There are options 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, and the behavior of devices regarding 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 defines conformance requirements.

Optional specifications are provided for protocol security. The PTP standard documents specify conditions under which Version 2.1 is compatible with IEEE 1588-2008.

The PTP underlying technology

This section introduces the basic terminology and components of PTP, including the various types of clocks, the synchronization methodology, and the basics of the delay mechanisms that are 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 several protocol messages and classifies each as either event messages or general messages. Management and signaling messages fall into the general messages category.

The first stage of the PTP process is the usage of an algorithm, which is known as the Best TimeTransmitter Clock Algorithm (BTCA), to organize all clocks in a 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 BTCA self-organizing hierarchy. The second stage of the 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 standards. PTP defines a hierarchy that is based on criteria that describe a combination of a device's timekeeping capability and the traceability of its time source. The time transmitter serves as a reference for one or more ordinary devices. The process of selecting the top TimeTransmitter clock (the PTP Grandmaster) from the list of participating devices in the hierarchy is defined in the PTP BTCA. The BTCA is a key part of the “simple, administration-free” aspects of PTP objectives as a protocol. PTP devices transmit “Announce” messages at configurable intervals. The devices in the PTP domain process these announce messages according to the BTCA and select the Grandmaster. Attributes that are contained in the “Announce” message that play a role in the BTC selection include the following ones:

- The device time source (atomic clock, GPS, or free-running oscillator)
- The device clock ID
- The device's priority (determined and configurable by administrators)

The IEEE 1588 PTP standard defines 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 or domain (making it the ultimate source of time for all other devices in the network or PTP domain). An ordinary clock synchronizes to another clock that can serve 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 that is driven by the upper-level TimeTransmitter, and the remainder are TimeTransmitter ports). Therefore, boundary clocks can be said to terminate and then start the PTP time distribution. The boundary clock function is 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 that 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 that are traversing the device. These clocks are especially suited for networks with redundant paths. In such networks, the PTP message always contains the actual delay that it experiences on the network, regardless of the path that is taken.
- The final component is a *PTP management node*, which is a network-connected device that is used to configure and monitor PTP devices.

The concept of PTP profiles

One of the goals of the IEEE 1588 PTP standard is to make it possible to set up and run a PTP network with minimal device settings and administrative impact. 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 meet the performance requirements of applications and eliminate or minimize device configuration settings. By using 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 or industry use cases while 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 International Telecommunication Union (ITU).

PTP profiles change some aspects of the standard and further extend the standard by mandating some of the optional features in the standard:

1. A PTP profile may define its own BTC algorithm, configuration, monitoring, and management mechanism.
2. A PTP profile may also define the usage of a specific transport mechanism, usage of multi-cast or unicast, and whether end-to-end or peer-to-peer path delay mechanisms are used.
3. PTP profiles also define any IEEE 1588 optional components that are required, permitted, or prohibited.
4. PTP profiles may also define new transport mechanisms, such as optical transport networks (OTNs) and data types.

The IEEE 1588 PTP standard defines 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 meet the needs of almost any application or use case. This flexibility has proven highly useful to the energy and telecommunications industries. As more profiles are developed, this flexibility will likely prove invaluable to the financial and banking industry too.

Regulatory requirements

This section provides a comprehensive background on time synchronization regulations. This section is important because understanding this topic, especially the most recent regulatory changes, makes it clear why IBM decided PTP was an important technology for their IBM Z clients.

The world's financial markets, even though they often depend on IBM Z, effectively operate as high-performance distributed systems where the timestamp of any trade can have a huge influence on the financial fortunes of millions of investors both large and small. The integrity and trust that are 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 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 making 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 that was required to run transactions and the “spread size” decreased. High Frequency Trading (HFT) became a widespread practice. Today, according to Financial Review, half of all transactions in today's stock markets are the result of HFT.[2]

The rapid expansion of computer-based trading with its highly sophisticated algorithms increased the need for synchronization of trading systems and traceability to a common reference time scale to help prevent irregularities and aid forensics. The practice of HFT makes it essential for all stock exchanges, trading platforms, and associated IT hardware to document that their timestamps are accurate. Institutions that interact with financial exchanges require a high accuracy time for several reasons:

- Accurate time is necessary for institutions to control their own trading traffic.
- Accurate time is also required to settle disagreements and prevent fraud. The timestamps that are collected by each institution must be examined when there are disagreements or errors in processing trades. The most famous example 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[3].” Timestamps that re created by different networks or systems can be meaningfully compared only if they are based on the same reference. For time and computer systems, the global reference is UTC.

[2] Source: https://www.nasdaq.com/glossary/h/high-frequency-trading.
[3] Source: https://support.sas.com/resources/papers/proceedings17/SAS0729-2017.pdf

UTC is the time scale that underpins GPS and broadcasts time signals and precise time services. The Bureau International des 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 or 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 produce a physical signal. 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 that is 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 5 days. Circular T allows each of the participating laboratories to know the recent difference of its local UTC time scale regarding UTC, which establishes 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)). The SI second is defined as the duration of a specified number of cycles of the microwave radiation that correspond to a transition within Cesium 133 atoms.) The US 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 10-year period (2006 - 2015), the average UTC-UTC(NIST) was 4.2 nanoseconds.

The concept of traceability for measurements of the time of day refers to a continuous chain of comparisons extending from a time comparison, timestamp, or clock synchronization that is 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 performed with some inherent level of inaccuracy (uncertainty of the measurement). The uncertainties of each link in the overall chain can 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 demonstrating 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 more requirements for demonstrating time traceability to a common reference:

- The timing equipment must be calibrated so that its unknown internal delays do not bias its time output.
- The equipment must be monitored continuously so that any fault or anomaly can be detected, and the time output is not used until the equipment is working correctly again. The calibration evidence and monitoring results should be archived so that the status of the timing equipment at any point in time can be verified later.

Overview of regulations: Past, present, and future

Before the advent of HFT and distributed electronic trading platforms, many of the clocks that were maintained for financial transactions were mechanical devices that physically stamped the time (ink) onto paper documents that were used to record the transactions. These clocks were seldom synchronized to the granularity that is 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 US and in Europe. Because there are some substantial differences, let us examine the US and European Union (EU) requirements in the separate following sub-sections.

United States: National Association of Securities Dealers, Order Audit Trail System, Financial Industry Regulatory Authority, and the consolidated audit trail

In August 1996, the US 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 Order Audit Trail System (OATS) *Rule 6953*, which was 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 timestamps are accurate. It also required the synchronization to be within 3 seconds of the NIST atomic clock, which while not 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, which are 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 US stock market transactions. Rule 7340 made the synchronization requirement more stringent by a factor of three, now requiring synchronization within 1 second of NIST time. Rule 7340 also required that this 1-second synchronization be maintained always when markets are open.

Shortly, 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. Then, SEC issued its Regulatory Notice 16-23 in July 2016. Under the rule change, effective 15 August 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 the following parts:

- The difference between the NIST standard and a time provider's clock
- Transmission delay from the source
- The amount of drift in the member's clock

Firms were given 6 months from the effective date until 20 February 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 19 February 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 that is specified in Rule 4590. This log should be maintained for the period and accessibility that are specified in SEC Rule 17a-4(b), and the log should be maintained and preserved for the required period in paper format or in a format that is permitted under SEC Rule 17a-4.

This documentation and log requirements stemmed in part from the 2012 SEC *Rule 613*, which required FINRA and US-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 US markets. The rules also required the sending of complete documentation about orders to a central repository by 8 AM Eastern Time the day following a trade. CAT requires timestamps 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 or canceled
5. The time when the order ran

Europe: MiFID II

The European Securities and Markets Authority (ESMA) is the closest equivalent in Europe to the SEC in the US. The ESMA goals are ensuring fair and equitable financial markets and protecting investors. ESMA is empowered by the Market in Financial Instruments Directive (MiFID) to draft the Regulatory Technical Standard (RTS) for European financial markets and firms. MiFID has been applicable across the EU since November 2007. ESMA began the process of updating MiFID with a new Directive, MiFID II, in 2011. MiFID II was formally adopted by the EU Parliament in June 2014, with the new requirements applicable 3 January 2018. The MiFID II clock synchronization requirements are more stringent than the latest US requirements. The specific section of MiFID II that details these new clock synchronization requirements is Commission Delegated Regulation (EU) 2017/574, which is better known as 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 by using either a link to one of the laboratories that maintain a UTC(k) realization of UTC or the time signals that are disseminated by GPS or another satellite system. If you use 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, that is, the maximum divergence from UTC that should be allowed by the operators of financial trading venues while accounting for the gateway to gateway latency of their trading systems. This topic 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 applies to the business clocks of members or participants of financial trading venues. This topic is 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, functions, 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 tighter synchronization than typically is achievable with NTP. PTP can synchronize devices within nanoseconds of each other over a common networking infrastructure. PTP is capable of picosecond-level synchronization through implementation of the high accuracy optional component of the latest IEEE 1588 standard along with the ITU-T Synchronous Ethernet technology.

For many use cases and industries, NTP is “good enough”. If you need the tighter synchronization that is required in the financial and banking industry, PTP is likely a better option for you. As these regulations become increasingly strict, you likely will see increased adoption of PTP in the industries that are subjected to these regulations.

The key reason why PTP is capable of tighter and 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 (OS). Also, unlike NTP, PTP devices account for device latency by accounting for the amount of time that synchronization messages spend in each device.

The PTP BMC algorithm allows for PTP to more easily adapt to changing conditions and failure scenarios than NTP. The PTP BMC algorithm ensures that 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 before it is used, and its performance suffers when messages must traverse network elements such as switches.

Preparing for PTP

How do you prepare your environment for implementing PTP? The process is similar to what you would do before introducing any other new technology into your data center: evaluate what you have and need; evaluate your budget; and take care of the gaps.

In terms of evaluating what you have, look at several things:

- Look at the time servers that are connected to the outside world (such as an antenna receiving GPS time) and are providing your IBM Z Sysplex CTN with its STP External Time Reference (ETR). Most likely, your ETR is NTP. Hopefully, your time server can also function as a PTP Grandmaster. It likely can, but you want to verify, and ask the vendor what if any code upgrades are needed for PTP support.
- The networking equipment in your time synchronization network must support PTP. At a minimum, your network must function as a PTP transparent clock. The good news is that PTP has been a standard since 2002, so it is likely that the switches and routers in your data center network, especially in the mainframe environment, are “PTP-compliant.” You might need to enable PTP only in the management software, or worst case, do a code or OS update. Make sure UDP ports 319 and 320 are open on any intermediate firewalls.
- You might want to use PTP for other server environments in the data center, especially for servers that play a role for applications that are impacted by regulations. In this case, look at the server hardware, such as NICs, to see whether they support PTP. Does the server OS support PTP, and if so, do you have the driver levels that are required? What about the applications themselves?
- Make certain you meet all the IBM Z requirements.

When you understand what you have, 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, make sure that you buy equipment that supports PTP, and update software and your OS as needed.

Conclusion

This appendix explained the basics of PTP, including its history, an overview of the technology, the regulations that led to IBM to support PTP on IBM Z, and some advice about how to prepare for implementing PTP. IBM Z started to investigate PTP in 2012 - 2013. These investigative and 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 SDOs. Then, IBM developed support for PTP with IBM Z CTNs, and announced a statement of direction for IBM z15 and follow-on IBM Z platforms.

Abbreviations and acronyms

AID	adapter ID
BIPM	Bureau International des Poids et Mesures
BMCA	Best Master Clock Algorithm
BPA	Bulk Power Assembly
BTCA	Best TimeTransmitter Clock Algorithm
BTS	Backup Time Server
CAR	Console Assisted Recovery
CAT	consolidated audit trail
CE LR	Coupling Express Long Reach
CF	Coupling Facility
CFCC	Coupling Facility Control Code
CMAS	CICSPlex System Manager address space
CMG	Computer Measurement Group
CP	Control Program
CST	Coordinated Server Time
CTN	Coordinated Timing Network
CTS	Current Time Server
DR	disaster recovery
DST	Daylight Saving Time
EDA	Enhanced Drawer Availability
EPSPT	Enhanced Preventive Service Planning Tool
ESMA	European Securities and Markets Authority
ETR	External Time Reference
ETS	External Time Source
EU	European Union
FINRA	Financial Industry Regulatory Authority
FSP	flexible support processor
GAS	Going Away Signal
GDPS	Geographically Dispersed Parallel Sysplex
GMT	Greenwich Mean Time
HA	high availability
HADR	high availability and disaster recovery
HCA	Host Channel Adapter
HCD	Hardware Configuration Dialog
HCM	Hardware Configuration Manager
HFT	High Frequency Trading
HMC	Hardware Management Console
IBF	Internal Battery Feature
IBM	International Business Machines Corporation
IBM SSR	IBM System Services Representative
ICA SR	Integrated Coupling Adapter Short Reach
IC	Internal Coupling
ICB	Integrated Cluster Bus
IERS	International Earth Rotation and Reference Systems Service
iPDU	intelligent Power Distribution Unit
ISC	Inter-System Channel
ITU	International Telecommunication Union
LAN	local area network
LIC	Licensed Internal Code
LPAR	logical partition
MCL	maintenance change level
MIF	Multiple Image Facility
MiFID	Market in Financial Instruments Directive
MIS	Management Information Systems
MRO	Multi-region Operation
MTOF	Message Time Ordering Facility
NASD	National Association of Securities Dealers
NAT	network address translation
NIST	National Institute of Standards and Technology
NTP	Network Time Protocol
OADM	optical add-drop module
OATS	Order Audit Trail System
OS	operating system
OSC	oscillator card
OTN	optical transport network
PAR	Project Authorization Request
PCHID	physical channel identifier
PDU	Power Distribution Unit

POR	power-on reset
PPS	Pulse Per Second
PR/SM	Processor Resource/Systems Manager
PSP	Preventive Service Planning
PTP	Precision Time Protocol
PTS	Preferred Time Server or Primary Time Server
RII	Redundant I/O Interconnect
RTS	Regulatory Technical Standard
SDO	standards development organization
SE	Support Element
SEC	Securities and Exchange Commission
SFP	small form-factor pluggable
SI	System of Units
SNTP	Simple Network Time Protocol
SPE	small programming enhancement
STCK	STORE CLOCK
STCKE	STORE CLOCK EXTENDED
STCKF	STORE CLOCK FAST
STP	Server Time Protocol
TAI	Temps Atomique International also known as International Atomic Time
TOD	time-of-day
UDP	User Datagram Protocol
UPS	uninterruptible power supply
UT1	Universal Time 1
UTC	Coordinated Universal Time also known as Universal Time Coordinated also known as temps universel coordonné
V2	Version 2
VCHID	Virtual Channel Path Identifier
WDM	Wavelength Division Multiplexing
XRC	Extended Remote Copy
z/TPF PUT	z/Transaction Processing Facility Product Update

Related publications

The publications that are listed in this section are considered suitable for a more detailed description of the topics that are covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide more information about the topics in this document. Some publications that are referenced in this list might be available in softcopy only.

- *IBM z14 (3906) Technical Guide*, SG24-8451
- *IBM z14 Technical Introduction*, SG24-8450
- *IBM z14 Model ZR1 Technical Introduction*, SG24-8550
- *IBM z14 ZR1 Technical Guide*, SG24-8651
- *IBM z15 (8561) Technical Guide*, SG24-8851
- *IBM z15 (8562) Technical Guide*, SG24-8852
- *IBM z15 Technical Introduction*, SG24-8850
- *IBM Z Connectivity Handbook*, SG24-5444

You can search for, view, download, or order these documents and other Redbooks, Redpapers, web docs, drafts, and additional materials, at the following website:

ibm.com/redbooks

Other publications

The following publication is also relevant as a further information source:

Hardware Management Console (HMC) Operations Guide Version 2.15.0:

https://www.ibm.com/support/pages/hardware-management-console-operations-guide-version-2150 (requires IBMid authentication).

Online resources

The IBM Resource Link for documentation and tools website is also relevant as another information source:

http://www.ibm.com/servers/resourcelink (requires IBMid authentication)

Help from IBM

IBM Support and downloads:

ibm.com/support

IBM Global Services:

ibm.com/services

Redbooks

IBM Z Server Time Protocol Guide

(0.2"spine)
0.17"<->0.473"
90<->249 pages

SG24-8480-01

ISBN 0738461377

Printed in U.S.A.

Get connected

ibm.com/redbooks