IBM

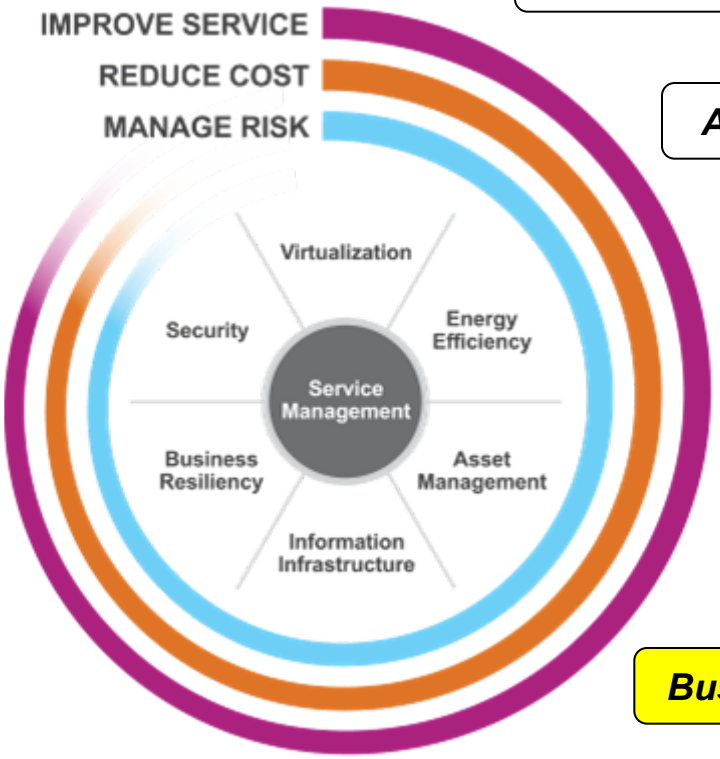
Server Time Protocol (STP) Overview
and Enhancements

Time Synchronization for the Next Generation

STP Overview

- Learning Objectives:
  - At the conclusion of this activity, participants should be able to describe:
    - ✓ Key attributes
    - ✓ Terminology
    - ✓ Supported configurations
    - ✓ External Time Source options
    - ✓ Hardware and software prerequisites
  - You will also be able to understand
    - ✓ Enhancements announced October 2008
    - ✓ Enhancements announced April 2009
    - ✓ Enhancements of September 2011
  - ❑ Understanding STP concepts is the first step to plan for:
    - ✓ Concurrently migrating from an ETR network or
    - ✓ Installing a new STP-only CTN

Building a dynamic infrastructure.



Service Management

Provide visibility, control and automation across all the business and IT assets to deliver higher value services.

Asset Management

Maximizing the value of critical business and IT assets over their lifecycle with industry tailored asset management solutions.

Virtualization

Leadership virtualization and consolidation solutions that reduce cost, improve asset utilization, and speed provisioning of new services.

Energy Efficiency

Address energy, environment, and sustainability challenges and opportunities across your business and IT infrastructure.

Business Resiliency

Maintaining continuous business and IT operations while rapidly adapting and responding to risks and opportunities.

Security

End to end industry customized governance, risk management and compliance solutions.

Information Infrastructure

Helping businesses achieve information compliance, availability, retention, and security objectives.

Agenda

- **Introduction**
  - Description
  - Key Attributes
  - Terminology
- Configurations
  - Mixed Coordinated Timing Network
  - STP-only Coordinated Timing Network
- External Time Source Options
  - NTP server
  - NTP server with PPS
- Timing-only Links
- Enhancements (Oct 08/Jan 09)
- Enhancements (April 28, 2009 announce)
- Enhancements (September 2011)
- Hardware and Software Prerequisites
- Summary

Statements of Direction (fulfilled)

- **The IBM System z10™ will be the last server to support connections to the Sysplex Timer® (9037).**
  - Servers that require time synchronization, such as to support a base or Parallel Sysplex®, will require Server Time Protocol (STP). STP has been available since January 2007 and is offered on the IBM z13, z13s, z10EC, z10BC, IBM System z9®, and IBM eServer™ zSeries® 990 and 890 (z990, z890) servers.
- **The zEnterprise servers will be the last to support dialing out to an external time source via modem.**

What is Server Time Protocol (STP)?

- Designed to provide the capability for multiple servers to maintain time synchronization with each other and form a Coordinated Timing Network (CTN)
  - CTN: a collection of servers that are time synchronized to a time value called Coordinated Server Time (CST)
- IBM Server-wide facility implemented in z14, z13, z13s, zEC12, zBC12, z196, z114, z10, z9, z990, z890 Licensed Internal Code (LIC)
  - Single view of “time” to PR/SM™
  - PR/SM can virtualize this view of time to the individual logical partitions (LPARs)
- Message based time synchronization protocol
  - Similar to Network Time Protocol (NTP) – an industry standard
  - Timekeeping information transmitted over Coupling Links
    - ISC-3 links (Peer mode), ICB-3, ICB-4, PSIFB and ICA SR
  - NOT standard NTP

Notes:
The STP Feature code can not be ordered if the server has been withdrawn from marketing.
STP follows the same Sysplex n-2 rule.

Key Attributes

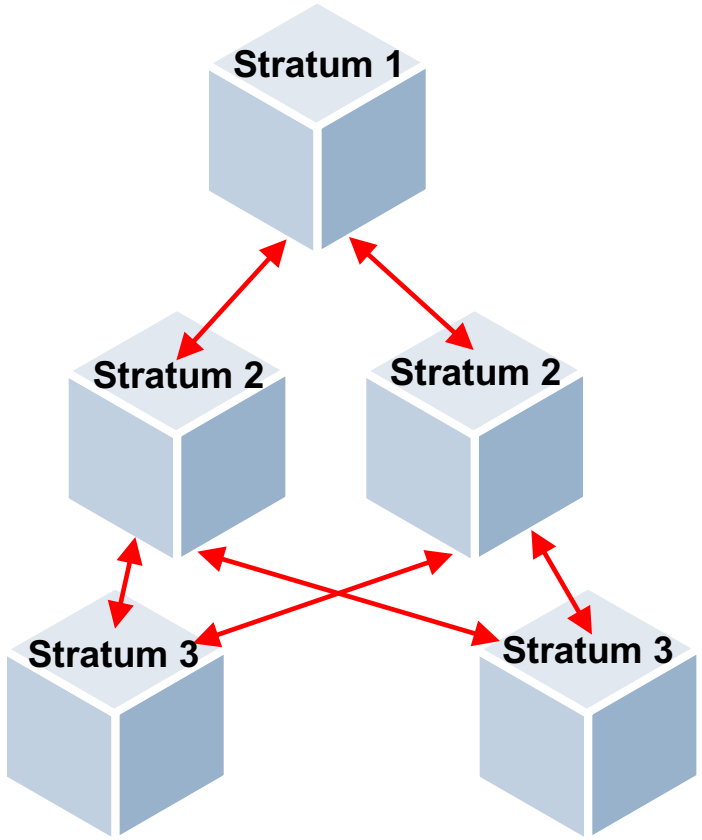
- Designed to provide improved time synchronization, compared to Sysplex Timer, for servers in a Sysplex or non-Sysplex configuration
- Can scale with distance
  - Generally, servers exchanging data over fast short links require more stringent synchronization than servers exchanging data over long distances
- Supports a multi-site timing network of up to 200 km over fiber optic cabling
  - Parallel Sysplex cluster can span up to 200 km
  - RPQ 8P2263 required for distances greater than 100 km
- Potentially reduces the cross-site connectivity required for a multi-site Parallel Sysplex cluster
  - Dedicated links not required to transmit timekeeping information
- With proper planning, allows concurrent migration from an existing External Time Reference (ETR) network
- Allows coexistence with ETR network
- Allows automatic updates of DST offset based on time zone algorithm
- Allows adjustment of CST up to +/- 60 seconds

Terminology

- STP-capable server/CF
  - IBM z14, z13, z13s, zEC12, zBC12, z196, z114, z10, z9, z990, z890 server/CF with STP LIC installed
- STP-enabled server/CF
  - STP-capable server/CF with STP FC 1021 installed
  - STP panels at the HMC/SE can now be used
- STP-configured server/CF
  - STP-enabled server/CF with a CTN ID assigned
  - STP message exchanges can take place
- CTN
  - Collection of servers that are time synchronized to a time value called Coordinated Server Time (CST)
- CTN ID
  - Servers / Coupling Facilities (CFs) that make up a CTN are all configured with a common identifier CTN ID

Terminology *(continued) …*

- Sysplex Timer transmits timekeeping information to attached servers in a star pattern
- STP transmits timekeeping information in layers or Stratums
- Stratum 1 (S1)
  - Highest level in the hierarchy of timing network that uses STP to synchronize to CST
- Stratum 2 (S2)
  - Server/Coupling Facility (CF) that uses STP messages to synchronize to Stratum 1
- Stratum 3 (S3)
  - Server/Coupling Facility (CF) that uses STP messages to synchronize to Stratum 2
- STP supports configurations up to Stratum 3 (S3)
  - Statum 4 supported on an HMC at Driver 32 and higher.



Time message will find a new path if needed

Agenda

- Introduction
  - Description
  - Key Attributes
  - Terminology
- **Configurations**
  - Mixed Coordinated Timing Network
  - STP-only Coordinated Timing Network
- External Time Source Options
  - NTP server
  - NTP server with PPS
- Timing-only Links
- Enhancements (Oct 08/Jan 09)
- Enhancements (April 28, 2009 announce)
- Enhancements (September 2011)
- Hardware and Software Prerequisites
- Summary

Configurations

- **Two types of CTN configurations possible:**
  - **Mixed CTN**
    - Allows servers/CFs that can only be synchronized to a Sysplex Timer (ETR network) to coexist with servers/CFs that can be synchronized with CST in the “same” timing network
    - Sysplex Timer provides timekeeping information
    - CTN ID format
      - STP network ID concatenated with ETR network ID

Note: Focus on n-2 Sysplex rules. A Mixed CTN can exist only if a z10 or older is in the network.

  - **STP-only CTN**
    - All servers/CFs synchronized with CST
    - Sysplex Timer is NOT required
    - CTN ID format
      - STP network ID only (ETR network ID field has to be null)

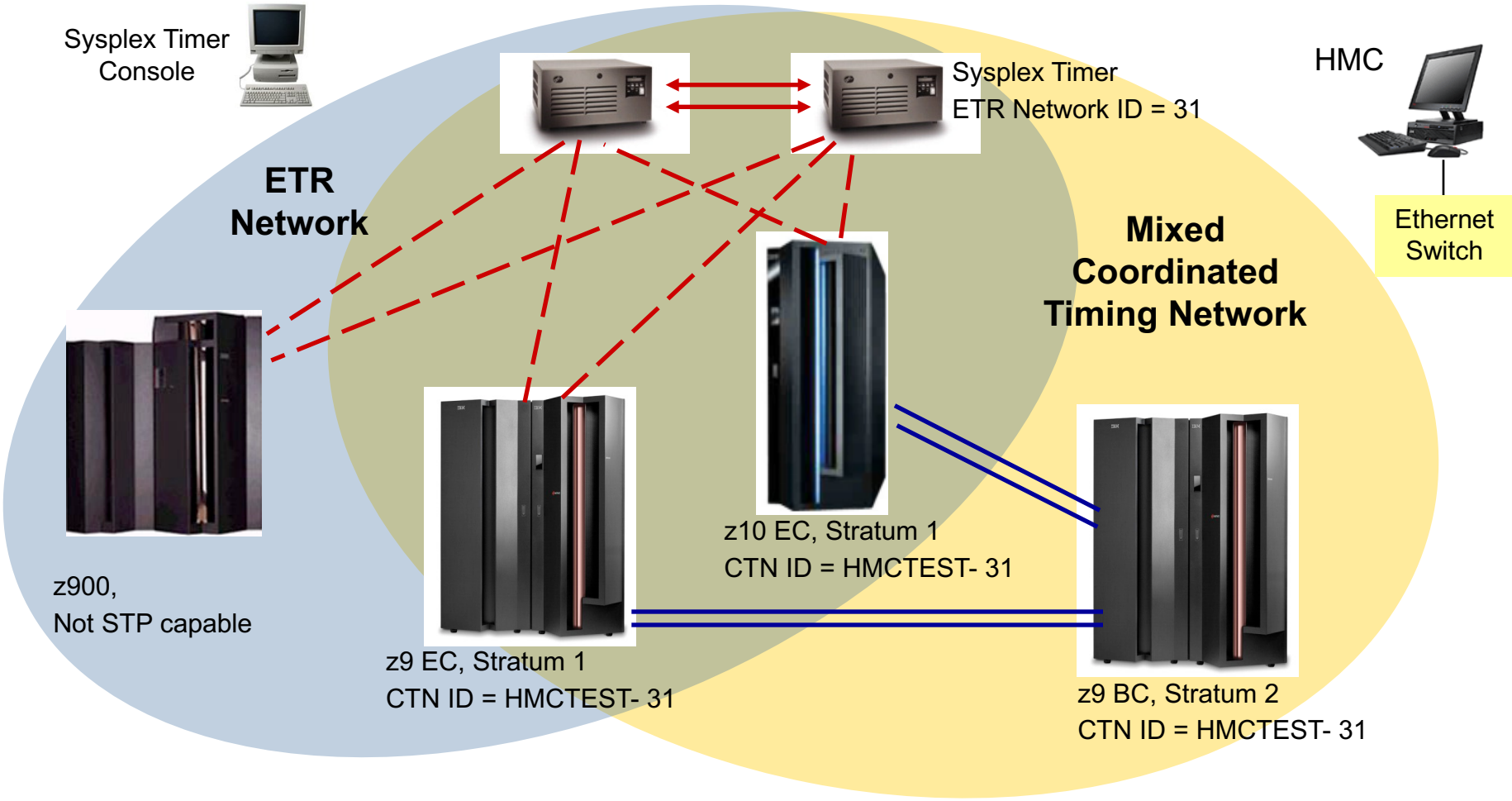
Mixed Coordinated Timing Network (CTN)

- Need at least one STP-enabled server to configure Mixed CTN
  - Selected STP-enabled server MUST also be synchronized to the Sysplex Timer
    - Automatically becomes a Stratum 1 server for the Mixed CTN
    - Can have multiple Stratum 1 servers in Mixed CTN
  - Stratum 2 server/CF uses Stratum 1 as clock source
  - Stratum 3 server/CF uses Stratum 2 as clock source
- Sysplex Timer provides timekeeping information for Mixed CTN
- Sysplex Timer console continues to be used for all timing related functions of the Mixed CTN
  - Initialize time; Set Time Zone, Daylight Saving Time (DST), Leap seconds offsets
  - Schedule DST and Leap seconds offset changes
  - Adjust time up to +/- 4.999 seconds
- Hardware Management Console (HMC) must be used for Mixed CTN ID initialization and modification

Mixed CTN (continued)

- Concurrent migration from ETR network (with proper planning)
  - Concurrently migrate from existing ETR network to Mixed CTN
  - Concurrently migrate from Mixed CTN to ETR Network
    - Allows testing Mixed CTN during change window and go back concurrently to
    - ETR network at start of production
- Coexistence in the “same” timing network of:
  - Servers/CFs that can only be synchronized to a Sysplex Timer (ETR network) and
  - Servers/CFs that can be synchronized with CST
  - In a Parallel Sysplex configuration, the only non STP-capable server and CF that can coexist are the IBM eServer zSeries 900, 800 (z900, z800) server and CF
    - Non STP-capable Server/CF MUST support Message Time Ordering Facility (MTOF)
    - Non STP-capable server/CF MUST be attached to Sysplex Timer

Mixed CTN example



- **z10 EC, z9 EC, z900 synchronized to Sysplex Timer**
- **z10 EC, z9 EC are Stratum 1 servers**
  - Two Stratum 1 servers recommended to help avoid single point of failure
- **z9 BC synchronized to either z10 EC or z9 EC via STP is a Stratum 2 server**
  - z9 BC does not need ETR link connections
  - z9 BC can be located up to 200 km away from z10 EC, z9 EC

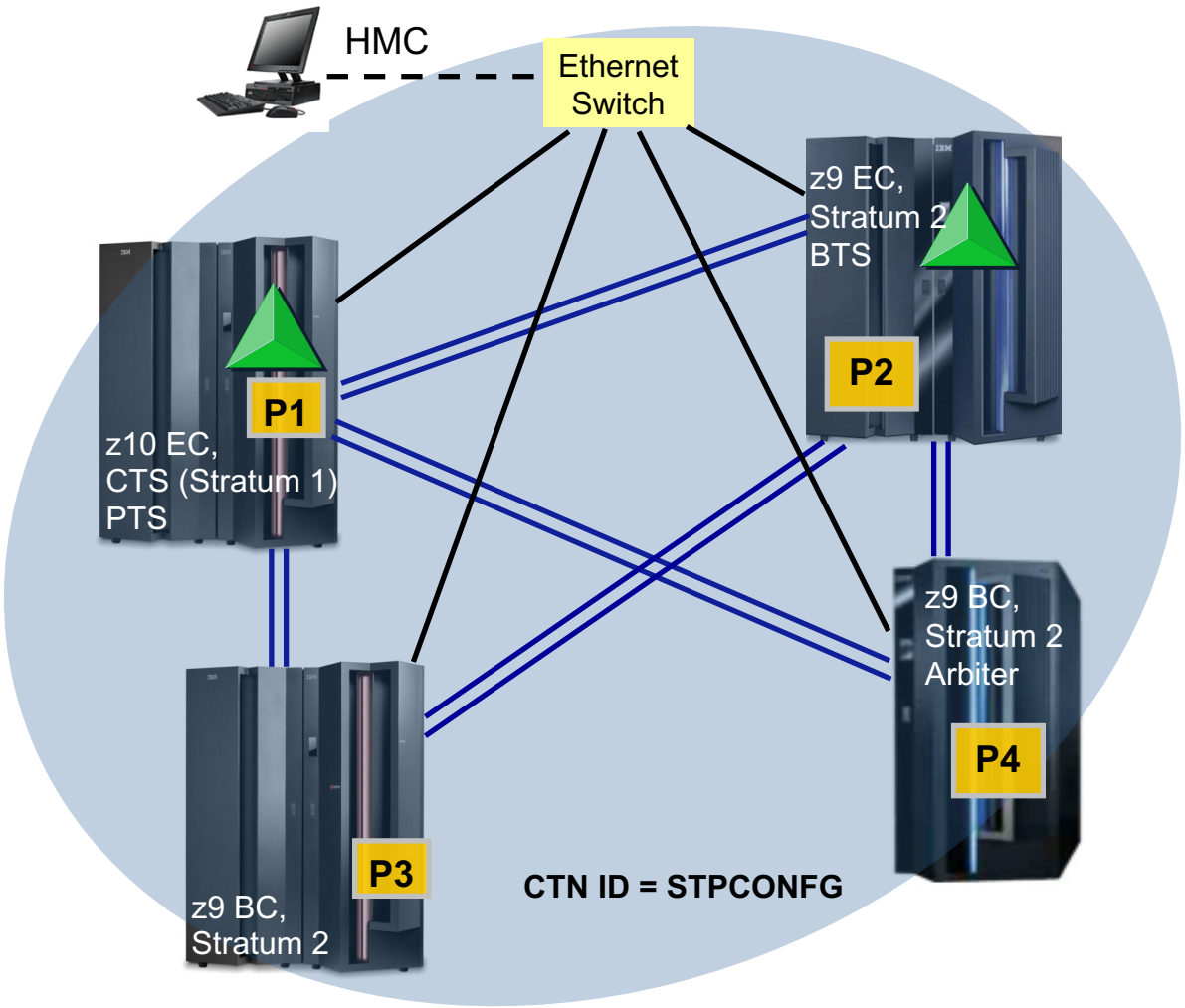
STP-only CTN

- All servers in STP-only CTN have to be STP capable
  - 9037s no longer required
- Server roles
  - Preferred Time Server (PTS)
    - Server that is preferred to be the “active” Stratum 1 server
  - Backup Time Server (BTS)
    - Role is to take over as the Stratum 1 under planned or unplanned outages, without disrupting synchronization capability of STP-only CTN
  - Current Time Server (CTS)
    - “Active” Stratum 1 server
      - Only one “active” S1 allowed
      - Only the PTS or BTS can be assigned as the CTS
      - Normally the PTS is assigned the role of CTS (recommended)
  - Arbiter
    - Provides additional means to determine if BTS should take over as the CTS under unplanned outages

STP-only CTN (continued)

- HMC must be used to provide the following functions:
  - Initialize Coordinated Server Time (CST) manually
  - Initialize CST to an External Time Source (ETS)
  - Configure an External Time Source (ETS) to set CST accurately
  - Set Time Zone Offset, Daylight Saving Time Offset, Leap seconds Offset
  - Schedule and change Offsets (Daylight Saving, Leap seconds)
    - Automatic scheduling of Daylight Savings Time based on algorithm
  - Adjust time by up to +/- 60 seconds (9037 allows 4.999 seconds)
  - Define, modify, view the STP-only CTN ID
- Concurrent migration (with proper planning)
  - Concurrently migrate from Mixed CTN to STP-only CTN or
  - Concurrently migrate from existing ETR network to STP-only CTN
  - Concurrently migrate from STP-only CTN to Mixed CTN
    - Allows testing STP-only CTN during change window and go back concurrently to Mixed CTN or ETR network at start of production.

STP-only CTN example



STP can use existing Coupling links

- Configuration has to be defined
  - Must assign PTS and CTS
  - Optionally assign BTS
    - Strongly recommended to allow near-continuous availability
  - Optionally assign Arbiter
    - Recommended for configurations of 3 or more servers/CFs
    - Can improve recovery

P1, P2, P3, P4 – systems in a Parallel Sysplex

Agenda

- Introduction
  - Description
  - Key Attributes
  - Terminology
- Configurations
  - Mixed Coordinated Timing Network
  - STP-only Coordinated Timing Network
- **External Time Source Options**
  - NTP server
  - NTP server with PPS
- Timing-only Links
- Enhancements (Oct 08/Jan 09)
- Enhancements (April 28, 2009 announce)
- Enhancements (September 2011)
- Hardware and Software Prerequisites
- Summary

External Time Source Options

- Dial-out from HMC to time services (available on z10, z9, z990, z890)
  - NIST Automated Computer Time Service (ACTS)
  - NRC Canadian Time Service (CTS)
  - IEN Telephone Date Code (CTD)
  - Scheduling of periodic dial-outs to time services so that CST can be gradually steered to time provided by dial-out time service
  - Time accuracy +/- 100 ms of time provided by dial-out time service
- NTP server (available on z9 and higher)
  - Addresses the requirements of customers who want time accuracy across heterogeneous platforms (System z and non-System z)
  - Time accuracy +/- 100 ms of time provided by NTP server
- NTP server with Pulse per second (PPS) output (available on z9 and higher)
  - Pulse per second (PPS) provides enhanced accuracy
    - 10 microseconds vs 100 milliseconds
- Continuous Availability of ETS
  - ETS configured for Backup Time Server (BTS) can be used for time adjustments to maintain ETS accuracy, if ETS configured for Preferred Time Server (PTS) not accessible
  - Applies to NTP server support with or without PPS
- Encryption and authentication
  - STP supports external time sources that support encryption. Announced with the zEC12.

NTP server functions not available on z990, z890

Network Time Protocol (NTP) client support

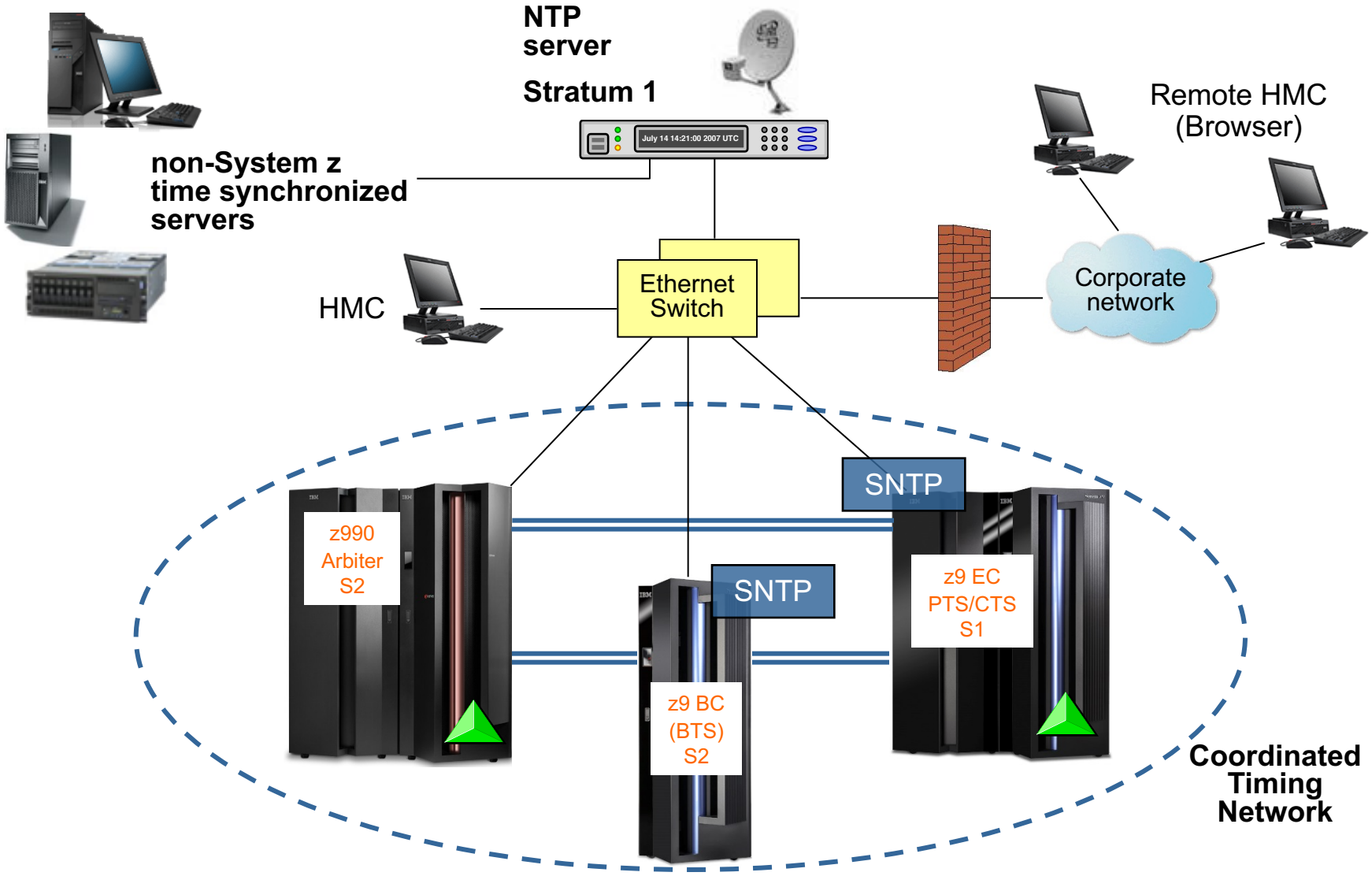
- Purpose of this function is to allow the same accurate time across an enterprise comprised of heterogeneous platforms
- Simple Network Time Protocol (SNTP) client support added to Support Element (SE) code of z9 servers/CFs and higher to enable interfacing with Network Time Protocol (NTP) servers
  - SNTP client support not available on z990 and z890 servers
  - z990 or z890 server can still be configured in an STP-only CTN
    - Cannot be assigned role of CTS
- NTP server becomes the single time source (External Time Source (ETS)) for STP, as well as other non-System z servers (UNIX®, Windows NT®, other) that have NTP clients
- Time reference for NTP server can be GPS, dial-out, DCF-77 etc.
- Accuracy is within +/- 100 ms of the time provided by an NTP Server
  - Actual time accuracy, relative to UTC, is dependent on how accurate the NTP server time is with respect to UTC.
- Only the Current Time Server (CTS) steers the time based on:
  - Timing information sent by Support Element (SE) code

NTP Server Planning

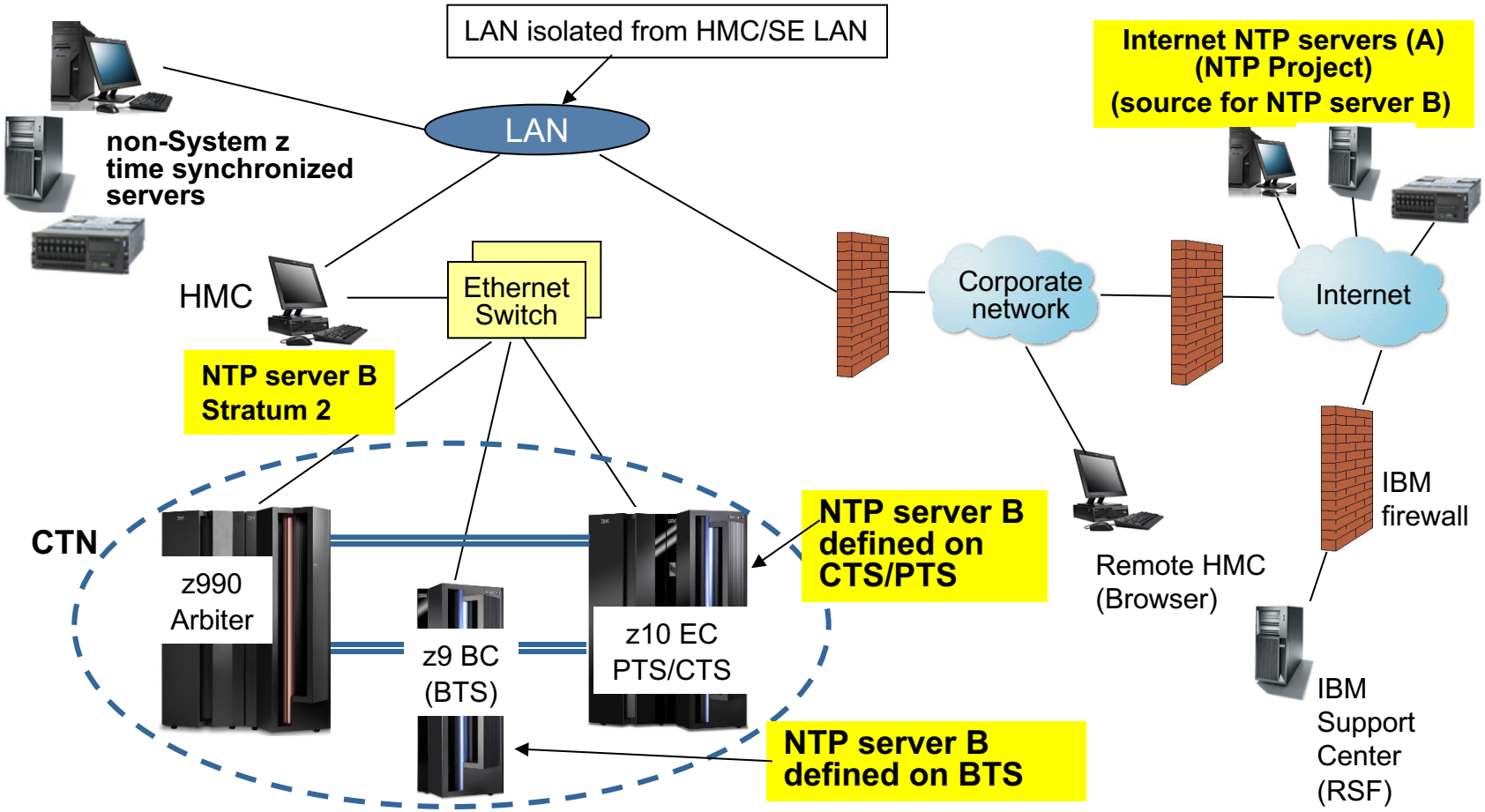
- NTP/SNTP server can be :
  - An external device available from several vendors
  - **NTP server on the HMC**
  - A local workstation (example, UNIX, Linux®, Windows) running the NTP/SNTP server code
    - HMC or workstation may use an internet NTP server for its time source
- NTP server should be directly connected to the SE LAN
  - SE LAN considered to be a private dedicated LAN
    - Isolated from the corporate and public networks
- NTP server on HMC addresses security concerns
  - Provides LAN isolation for SE NTP Client
  - The HMC has two LAN ports, physically isolated:
    - HMC/SE LAN used by NTP client code
    - Second LAN used by HMC to access an NTP time server to set its time

NTP server on HMC requires HMC code to be at Driver 76D (z10 HMC, v2.10.1 and higher)

NTP Server as ETS (example 1 – vendor box)



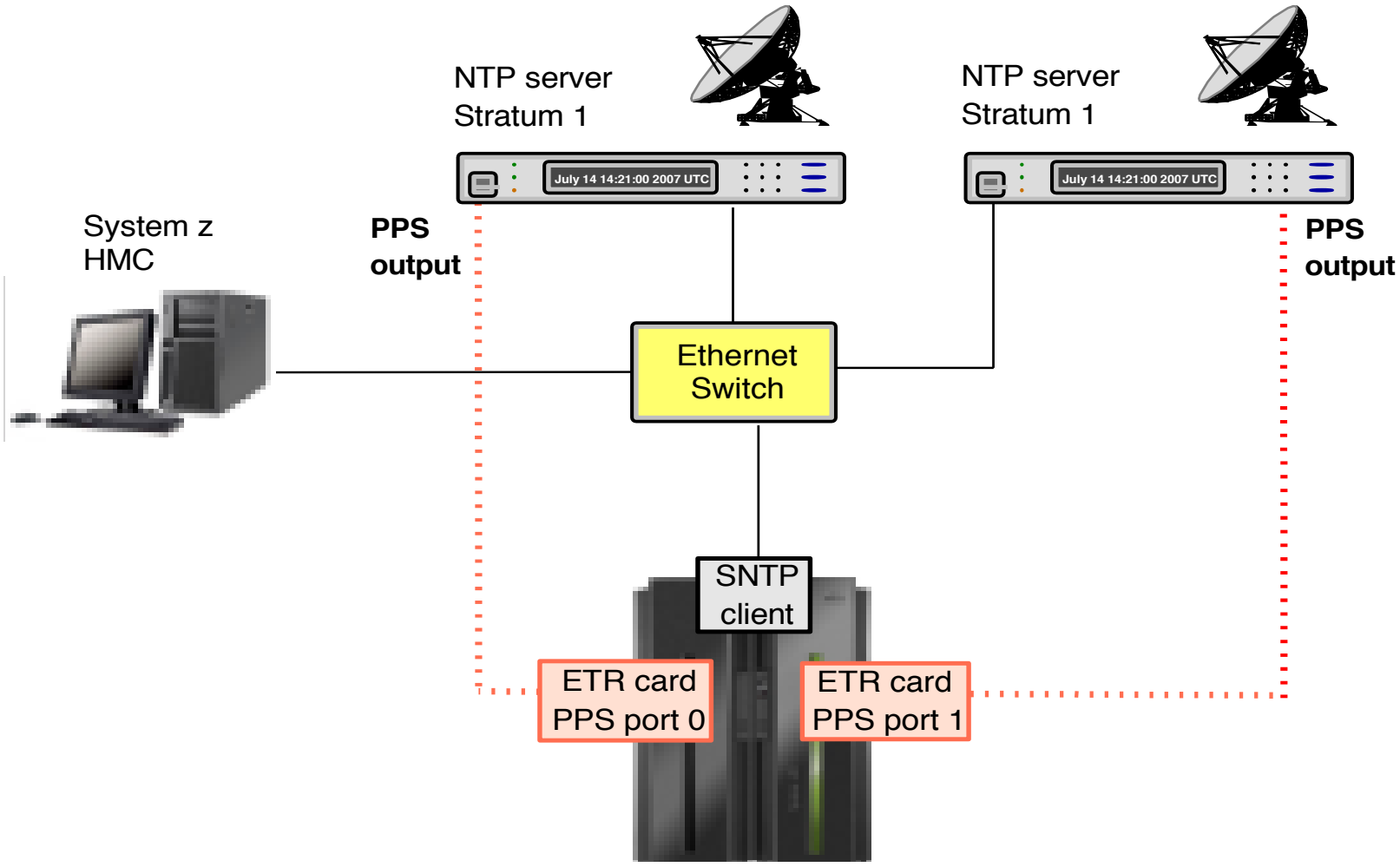
NTP Server as ETS (example 2 – HMC)



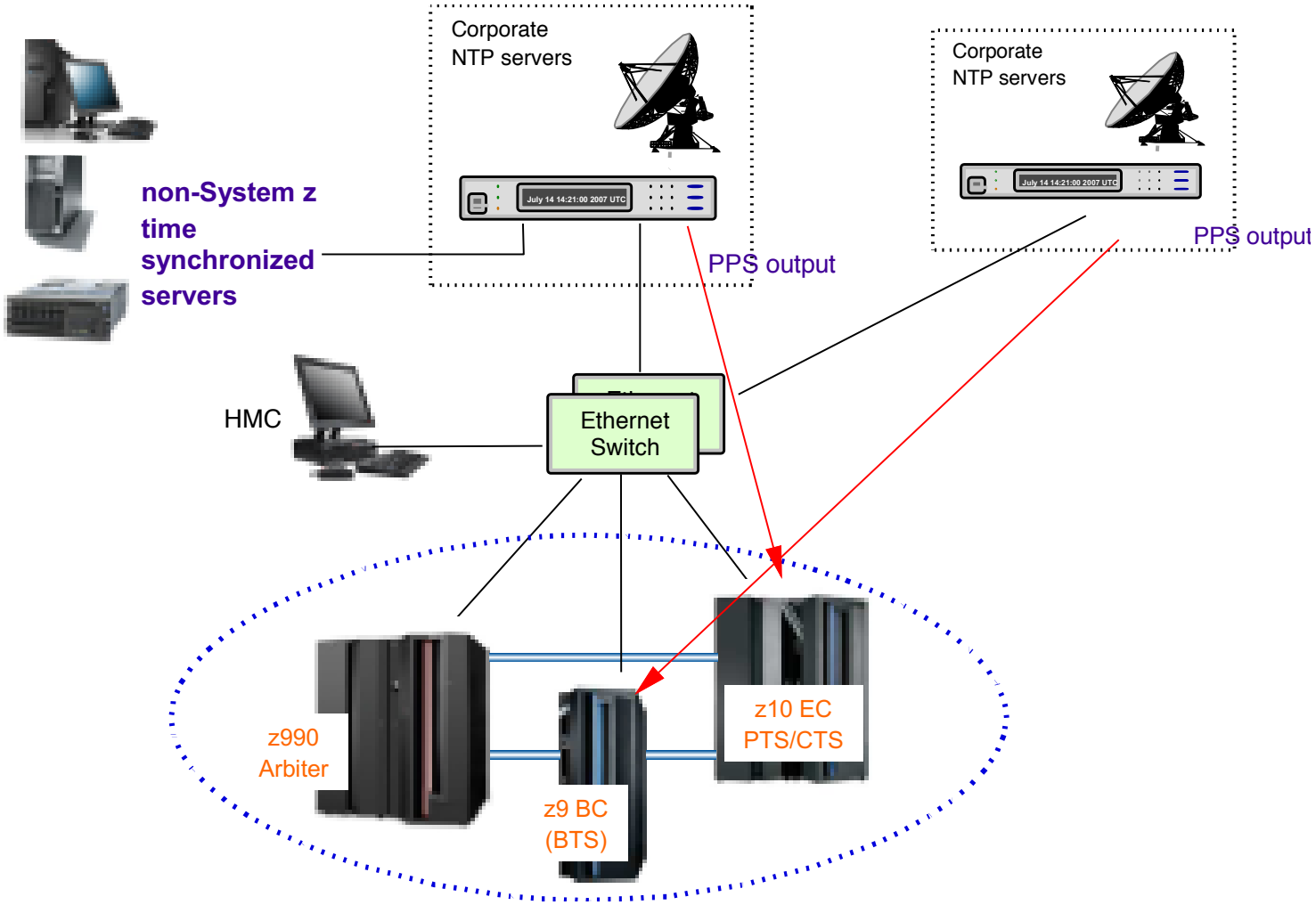
Enhanced Accuracy to an External Time Source (ETS)

- Simple Network Time Protocol (SNTP) client support added to Support Element (SE) code of z9 and above servers to interface to Network Time Protocol (NTP) servers
  - NTP Client support GA October 2007
- Time reference for NTP server can be GPS, dial-out, DCF-77 etc
- NTP servers provide multiple NTP outputs
  - Time information
- Some NTP servers also provide a highly stable, accurate “Pulse per second” (PPS) output
  - ETR card of z10 EC, z10 BC, z9 EC and z9 BC has additional PPS input
  - ECF or OSC card on zEnterprise
  - Each System z has 2 ETR, ECF or OSC cards providing capability of attaching to two NTP Servers with PPS output for redundancy
  - **There is no PPS connection available on the HMC.**
- STP utilizes the PPS signal to provide additional accuracy
  - Accuracy to PPS – 10 microseconds
- Only the Current Time Server (CTS) steers the time based on:
  - Timing information sent by Support Element (SE) code
  - PPS signal received by PPS port on ETR card

Enhanced Accuracy to an External Time Source (ETS) - ETS redundancy on same server (PTS/CTS) example

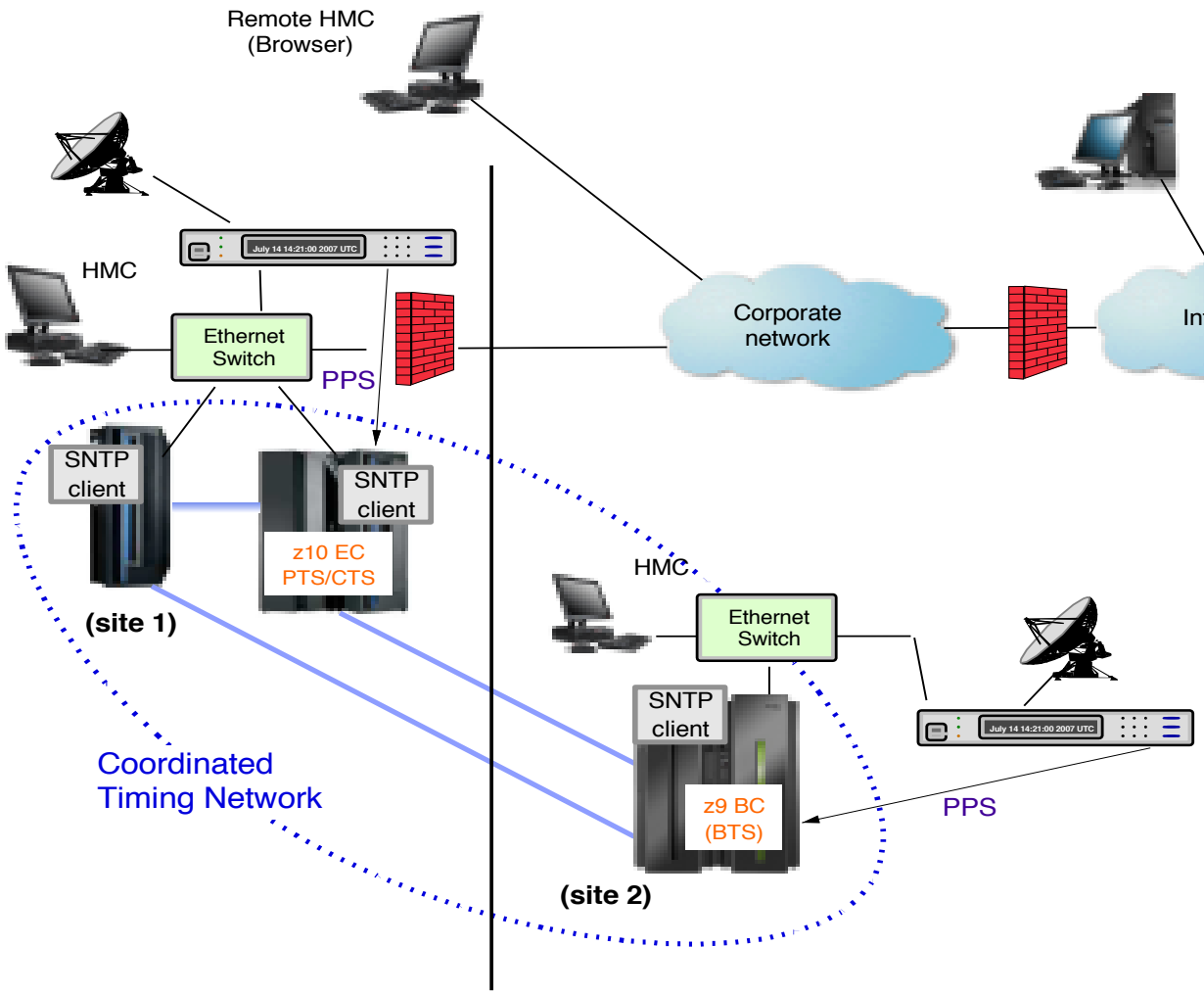


Enhanced Accuracy to an External Time Source (ETS) - ETS redundancy on PTS and BTS example



Continuous Availability of ETS - two site example

Announced Oct 08

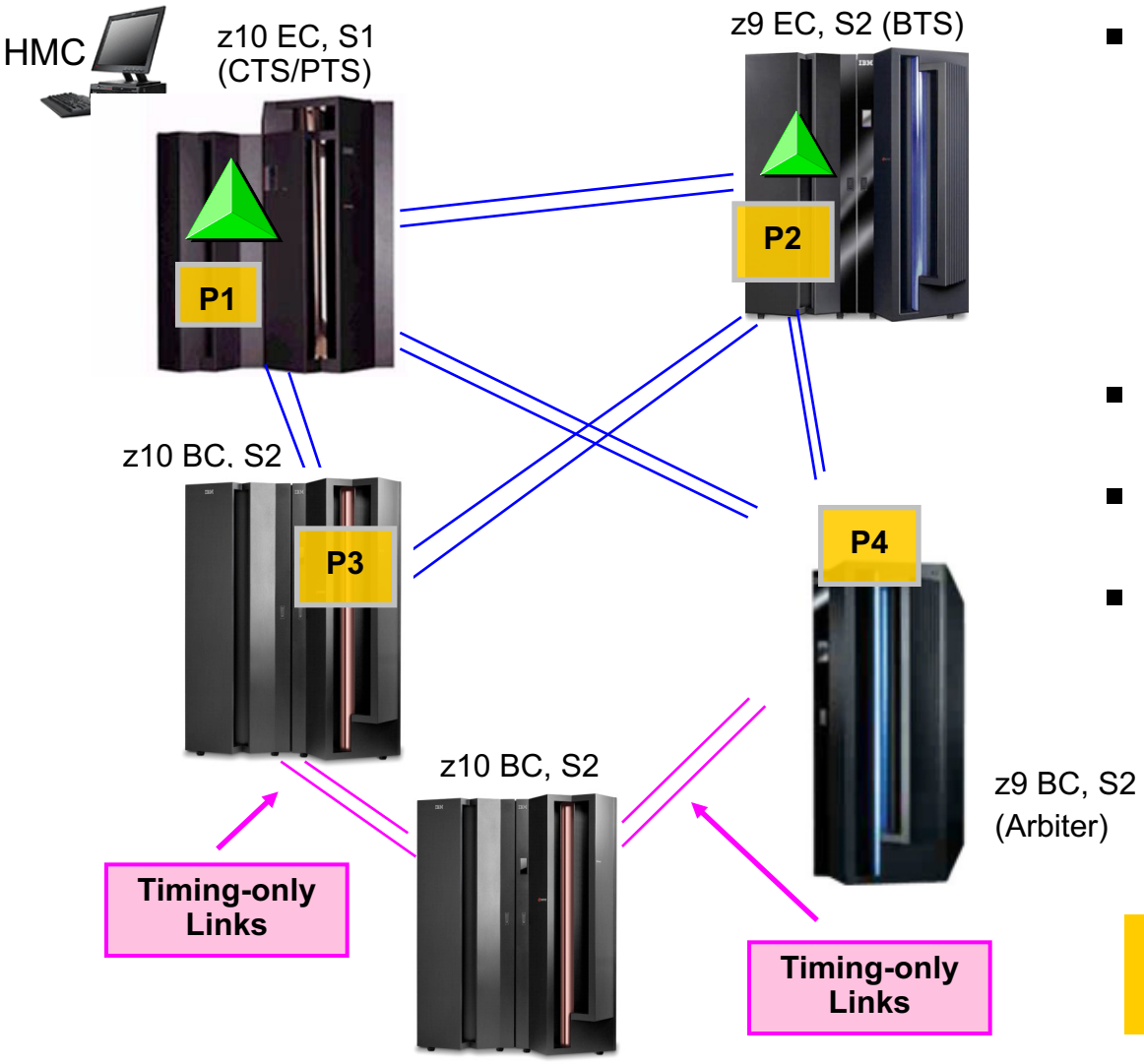


- BTS transmits adjustment information to PTS/CTS based on its NTP/PPS data
- When PTS/CTS detects failures associated with its NTP/PPS data
  - PTS/CTS switches to using NTP/PPS data from BTS
  - CTS role DOES NOT switch to BTS

Agenda

- Introduction
  - Description
  - Key Attributes
  - Terminology
- Configurations
  - Mixed Coordinated Timing Network
  - STP-only Coordinated Timing Network
- External Time Source Options
  - NTP server
  - NTP server with PPS
- **Timing-only Links**
- Enhancements (Oct 08/Jan 09)
- Enhancements (April 28, 2009 announce)
- Enhancements (September 2011)
- Hardware and Software Prerequisites
- Summary

Timing-only Links



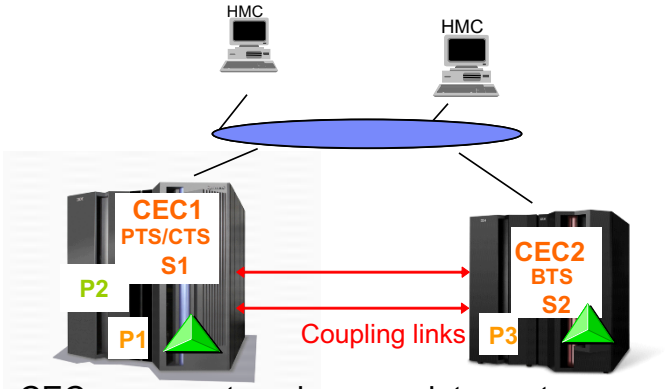
- Coupling links that allow two servers to be synchronized when a CF does not exist at either end of link
  - Typically required when servers requiring synchronization not in a Parallel Sysplex (for example XRC)
- HCD enhanced to define Timing-only links
- Can be defined in either Mixed CTN or STP-only CTN
- Timing-only links used to transmit STP messages only

P1, P2, P3, P4 – systems in a Parallel Sysplex

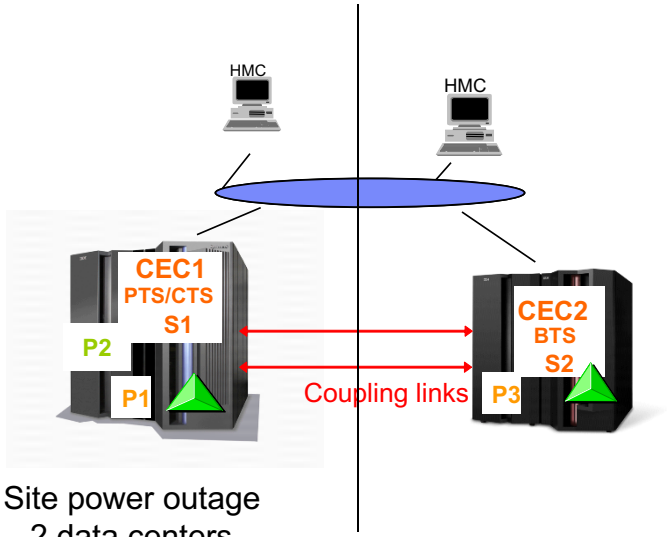
Agenda

- Introduction
  - Description
  - Key Attributes
  - Terminology
- Configurations
  - Mixed Coordinated Timing Network
  - STP-only Coordinated Timing Network
- External Time Source Options
  - NTP server
  - NTP server with PPS
- Timing-only Links
- **Enhancements (Oct 08/Jan 09)**
- Enhancements (April 28, 2009 announce)
- Enhancements (September 2011)
- Hardware and Software Prerequisites
- Summary
- Additional Information

Power Outage PTS/CTS with Internal Battery Feature



CEC power outage in same data center



Site power outage
– 2 data centers

- With Internal Battery Feature (IBF) on CEC1
  - CEC1 power outage, enters IBF state
  - CEC1 notifies CEC2 it is running on IBF
  - CEC2 waits for 30 seconds to take action
    - Could be a power glitch
    - If notified within 30 seconds that CEC1 back to "normal power", no further action
  - If CEC1 in IBF state > 30 seconds,
    - CEC2 takes over as the CTS
    - CEC1 becomes S2 until IBF no longer functional and power drops
  - CEC1 power resumes
    - Automatic re-takeover as PTS/CTS

For IBF recommendations please refer to:
STP Planning Guide SG24-7280

Not available on z990, z890

Automation to reconfigure STP-only CTN

- Targeted for Geographically Dispersed Parallel Sysplex™ GDPS® and non-GDPS customers
  - Fulfilled customer requirement to remove manual reconfiguration required without API.
- Provides automation to reconfigure STP-only CTN for the following:
  - Planned action when disruptive action needs to be performed on CTS
  - Unplanned outage when BTS takes over as CTS
  - Planned or Unplanned outage of BTS or Arbiter

Not available on z990, z890

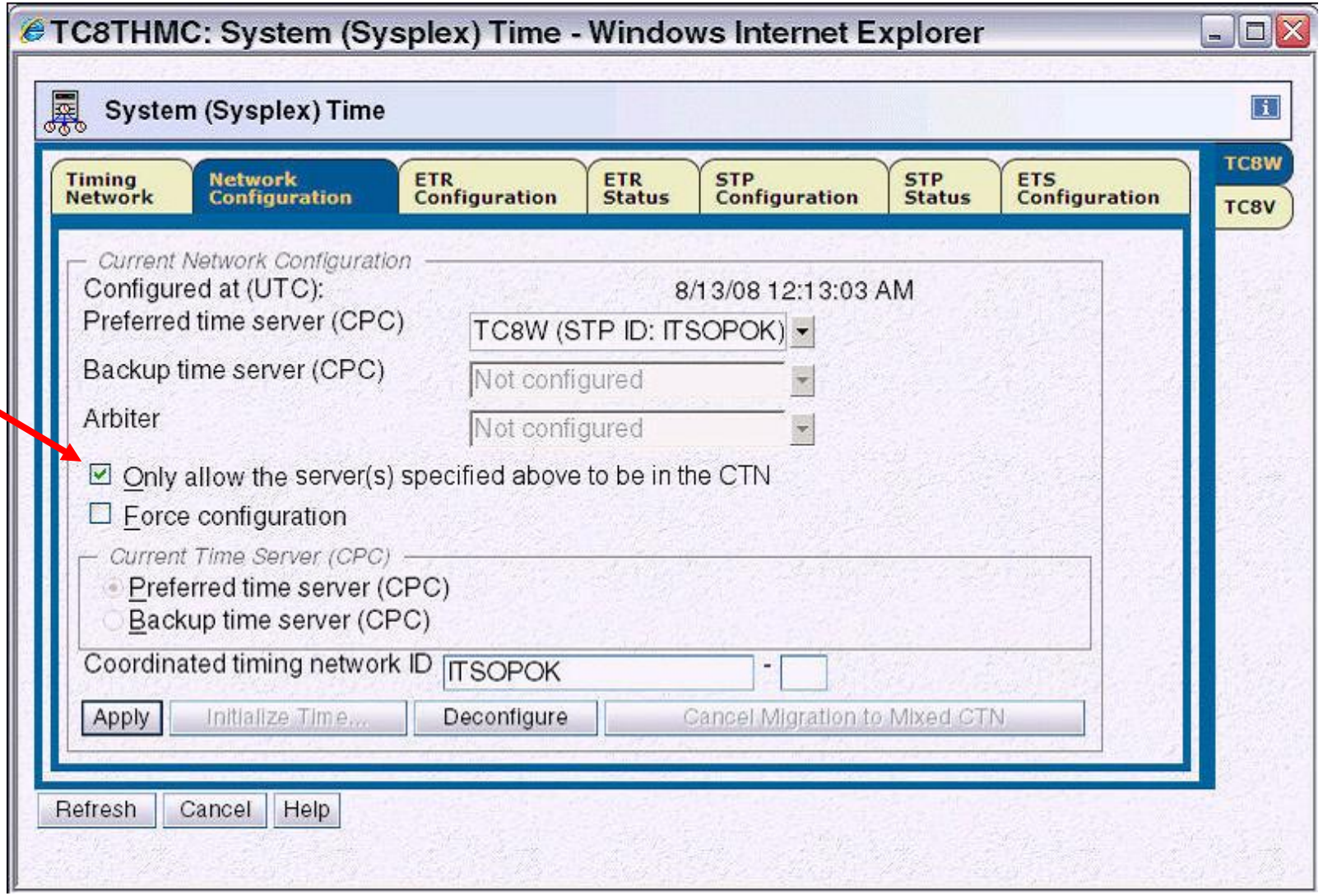
STP configuration and time information saved across Power on Resets or power outages

- **Single server CTN**
  - Previously
    - For an already operational single server CTN, when IML/POR done, user is required to:
      - ✓ Initialize the Time
      - ✓ Redefine that server as the stratum 1
  - System Management enhancement
    - Customer selects the option to not allow any other servers to join the single server CTN when configuring the CTN
    - STP CTN configuration (PTS/CTS) will be reinstated after IML/POR
      - ✓ Initialize Time not required

Also available on z990, z890

See http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105103

STP configuration and time information saved across Power on Resets or power outages (continued)



Agenda

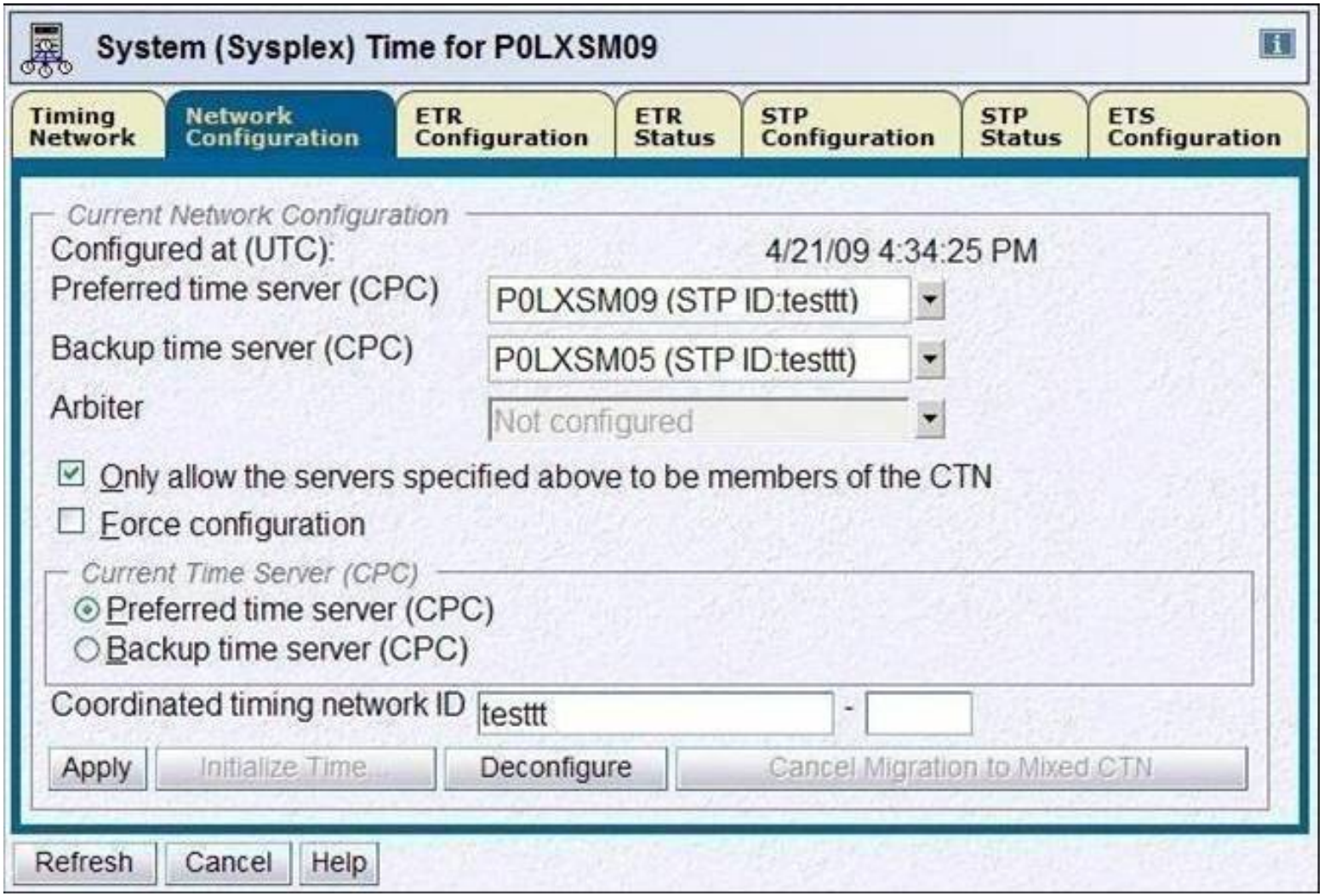
- Introduction
  - Description
  - Key Attributes
  - Terminology
- Configurations
  - Mixed Coordinated Timing Network
  - STP-only Coordinated Timing Network
- External Time Source Options
  - NTP server
  - NTP server with PPS
- Timing-only Links
- Enhancements (Oct 08/Jan 09)
- **Enhancements (April 28, 2009 announce)**
- Enhancements (September 2011)
- Hardware and Software Prerequisites
- Summary

STP configuration and time information saved across power outages – extended to two server CTN

- Two server CTN
  - Previously
    - If **both** servers experienced an unplanned power outage, user is required to:
      - ✓ Initialize the Time; Redefine the PTS, BTS, CTS roles
  - System Management enhancement
    - Customer selects the option to not allow any other servers to join the dual server CTN when configuring the CTN
    - STP CTN configuration (PTS/CTS, BTS) will be reinstated after IML/POR
      - ✓ Initialize Time not required
    - Configuration restoration rules (after power restored)
      - ✓ If servers can communicate via coupling links
        - PTS->S1 (CTS); BTS-> S2
      - ✓ If servers cannot communicate via coupling links, can communicate via LAN
        - LAN path used to determine state of servers and which server becomes S1
      - ✓ If state cannot be determined or cannot communicate on any path, both servers stay unsynchronized (S0)

Not available on z990, z890

STP configuration and time information saved across power outages – extended to two server CTN



Improved System Management with new z/OS messaging

- Previously
  - If ETS failure condition is detected
    - Hardware message generated
    - No z/OS® message generated, possibly delaying problem determination and correction

- System management enhancement
  - A z/OS message will be issued when an STP alert is raised
    - Message indicates event that caused the alert
  - z/OS V1.11 (rolled back to z/OS V1.10 and z/OS V1.9)

Not available on z990, z890

Agenda

- Introduction
  - Description
  - Key Attributes
  - Terminology
- Configurations
  - Mixed Coordinated Timing Network
  - STP-only Coordinated Timing Network
- External Time Source Options
  - NTP server
  - NTP server with PPS
- Timing-only Links
- Enhancements (Oct 08/Jan 09)
- Enhancements (April 28, 2009 announce)
- **Enhancements (September 2011)**
- Hardware and Software Prerequisites
- Summary

Why design enhancements?

- Current STP recovery design
  - Handles recovery of single failures in an STP-only CTN
  - Provides safeguards against planned disruptive actions affecting time source or last link providing timing information for CTN

- Enhancements required to handle planned and unplanned actions that could impact all servers with roles in an STP-only CTN with three or more servers
  - Note : Enhancements will not handle Double failures of the PTS and BTS - will still result in a sysplex outage

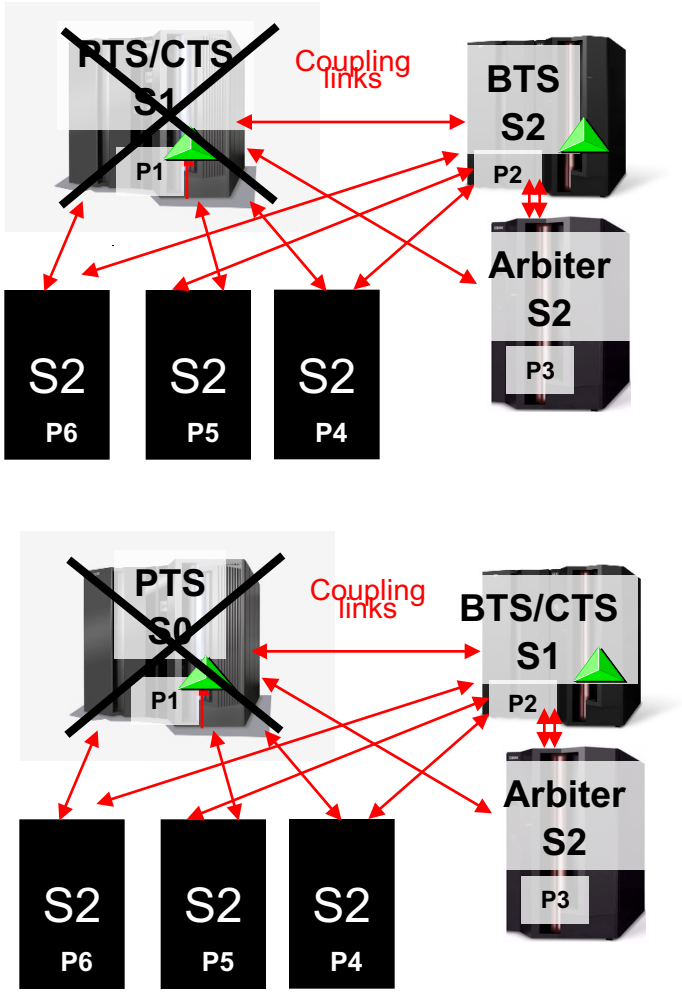
Recovery rules for 3 or more servers in CTN

- In order to keep the complex discussion that follows as understandable as possible, it will be assumed that
  - PTS has been assigned as the CTS.
  - BTS is a Stratum 2 server capable of becoming the CTS.
- Note that you can have only one CTS in the CTN, and it can be either the PTS or BTS.

- **CANNOT have two Stratum 1 servers in timing network to ensure data integrity**
- Backup Time Server (BTS) can take over as Current Time Server (CTS), only if either:
  - Preferred Time Server (PTS) can indicate it has “failed”
  - BTS can unambiguously determine the PTS has “failed”

- **Recovery rules based on “voting” (2 out of 3)**
- **If BTS and Arbiter agree they cannot communicate with PTS, then safe for BTS to takeover as S1**
- **If PTS cannot communicate with BTS and Arbiter, it can no longer remain a S1**
  - **Even if it is operational and providing a time source to rest of CTN**

STP-only CTN (Preferred, Backup and Arbiter assigned)
CTS failure or power outage



- **BTS loses communication with CTS on all established paths**
- **BTS and Arbiter communicate to establish if Arbiter also cannot communicate with CTS**
- **If both BTS and Arbiter cannot communicate with CTS**
  - BTS takes over as CTS (S1)
- **Since only 1 CTS (S1) can exist,**
  - PTS surrenders role of CTS

P1 – P6 in Parallel Sysplex

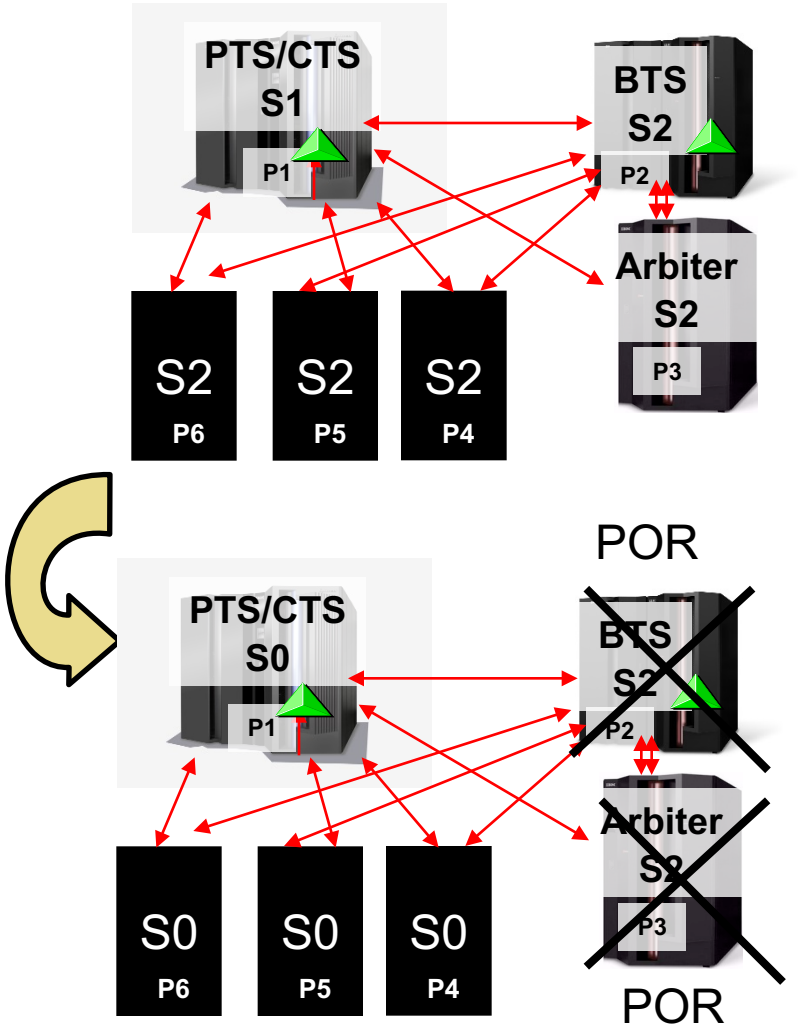
PTS >S0 server <<< After Recovery>>> BTS >S1 server

Original design – disruptive planned operations

- If planned disruptive action attempted on PTS (assuming it is the CTS),
  - **STP SE code blocks action until role reassigned to another server in CTN or all Chpids configured offline**

- If planned disruptive action attempted on BTS or Arbiter,
  - **STP code DOES NOT block the action**
  - **Potential Sysplex outage** if disruptive actions on BTS and Arbiter performed as part of same task or sequentially **without reassigning or removing roles**

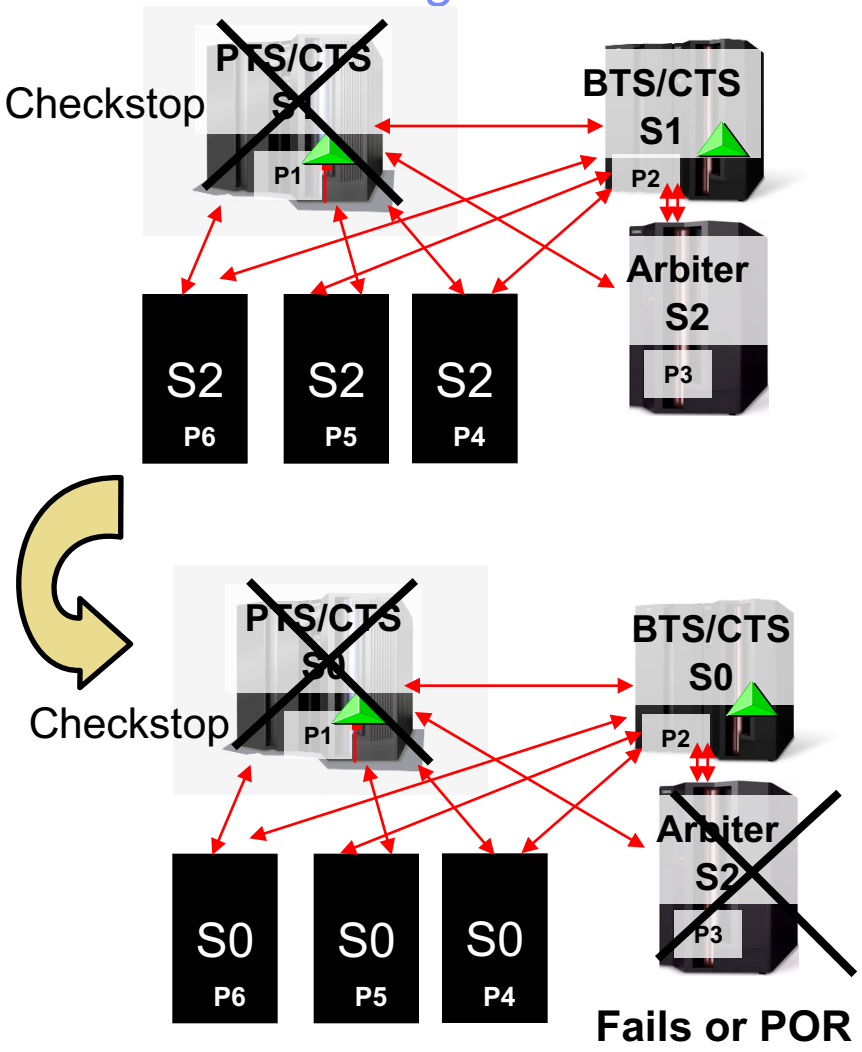
Potential Hazard case – roles not reassigned prior to planned disruptive actions on special role servers



- Planned disruptive action initiated from HMC to POR BTS and Arbiter
  - Could be same task or
  - Could be sequential
- **Roles not reassigned as recommended**
- PTS loses communication to both BTS and Arbiter
  - Surrenders S1 role
  - Since no other clock source available, PTS becomes unsynchronized (S0)
  - CECs with P4, P5, P6 also become unsynchronized
- **Sysplex wide outage**

Coupling links

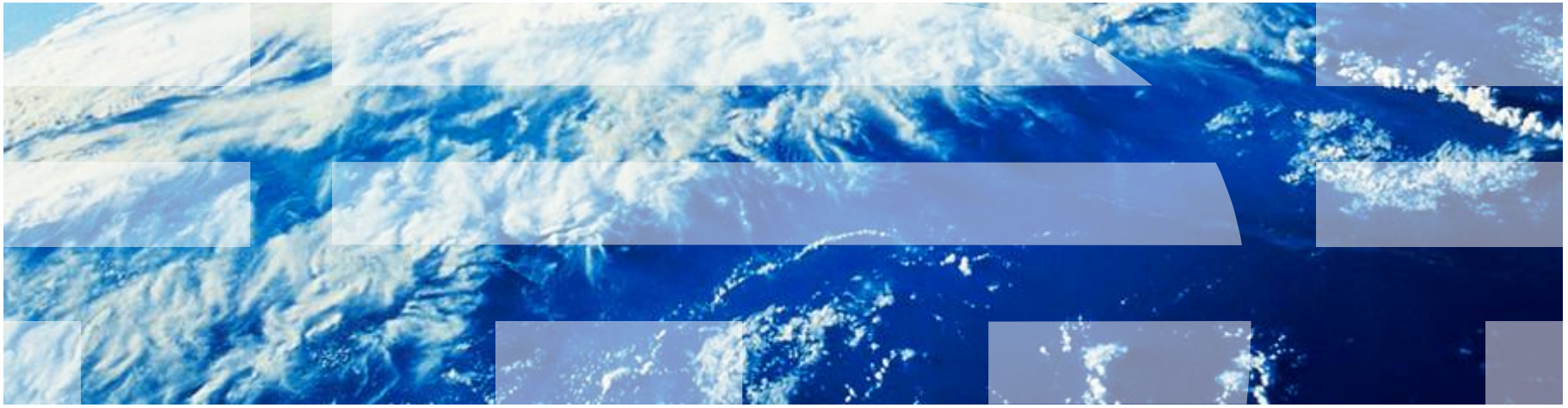
Potential Hazard Case – Roles not reassigned after BTS takeover as S1



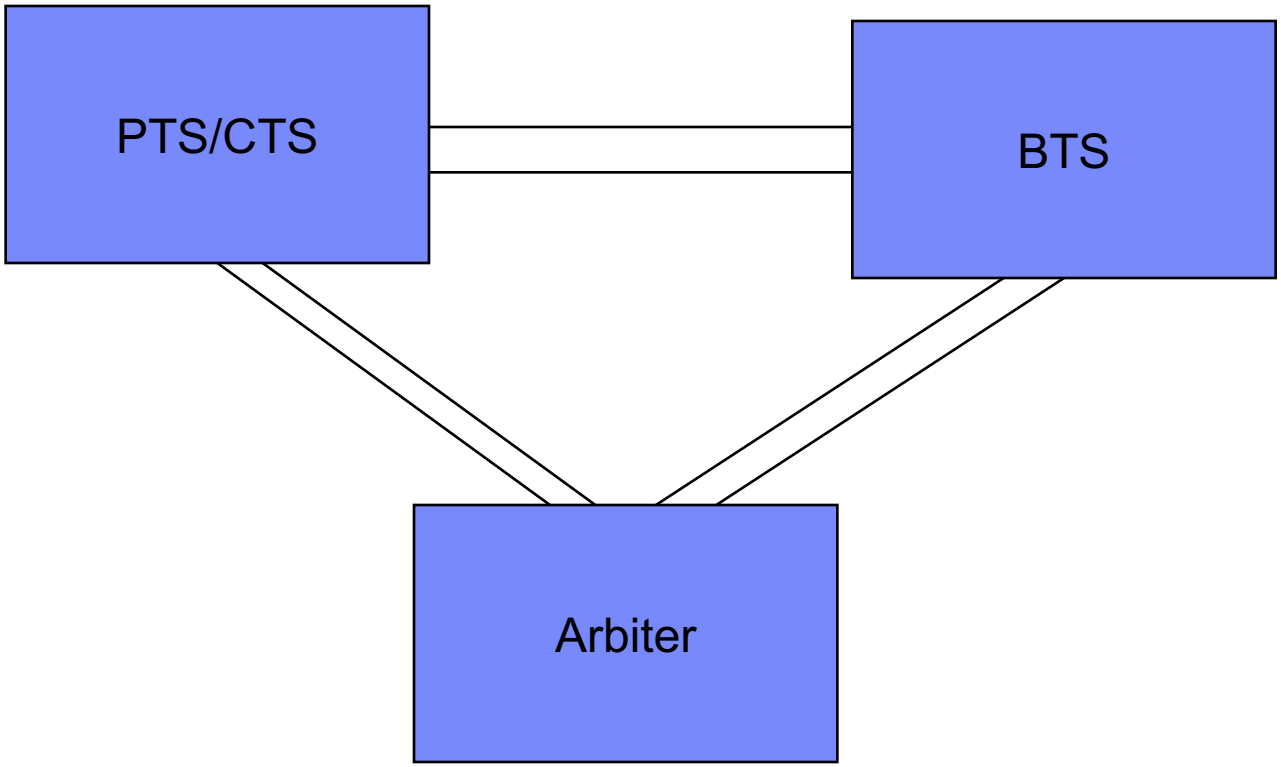
- PTS checkstop
  - BTS take over as S1
  - **Customer does not reassign roles as recommended**
- If subsequently BTS fails, no other server available to take over as S1 or
- If subsequently Arbiter fails or is PORed (planned action)
  - BTS loses communication to both PTS and Arbiter
    - Surrenders S1 role
    - Since no other clock source available, BTS becomes unsynchronized (S0)
    - CECs with P4, P5, P6 also become unsynchronized
- **Sysplex wide outage**

Coupling links

Triad Recovery changes

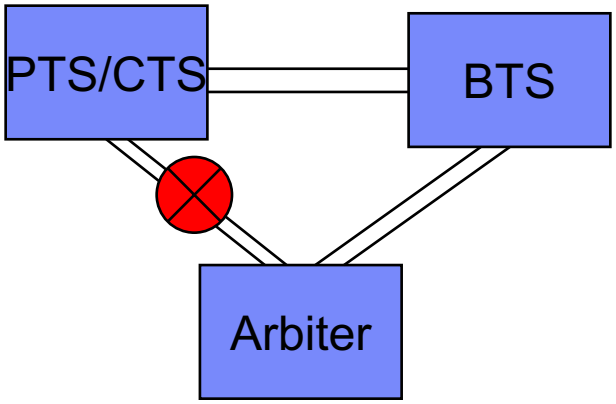


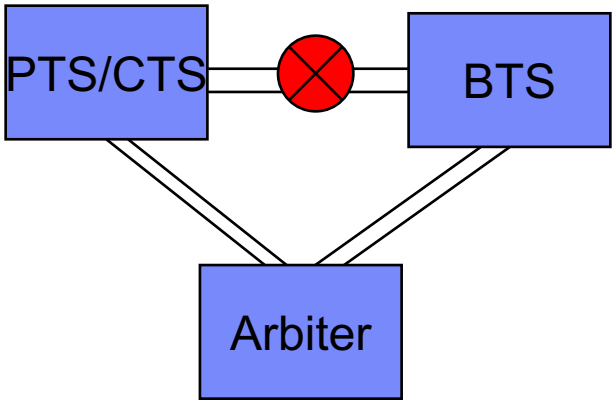
Full triad attachment state

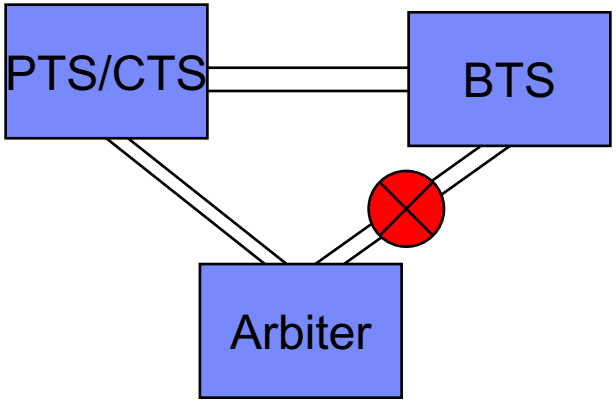


All special role servers attached to each other to form a triangle

Partial triad attachment state







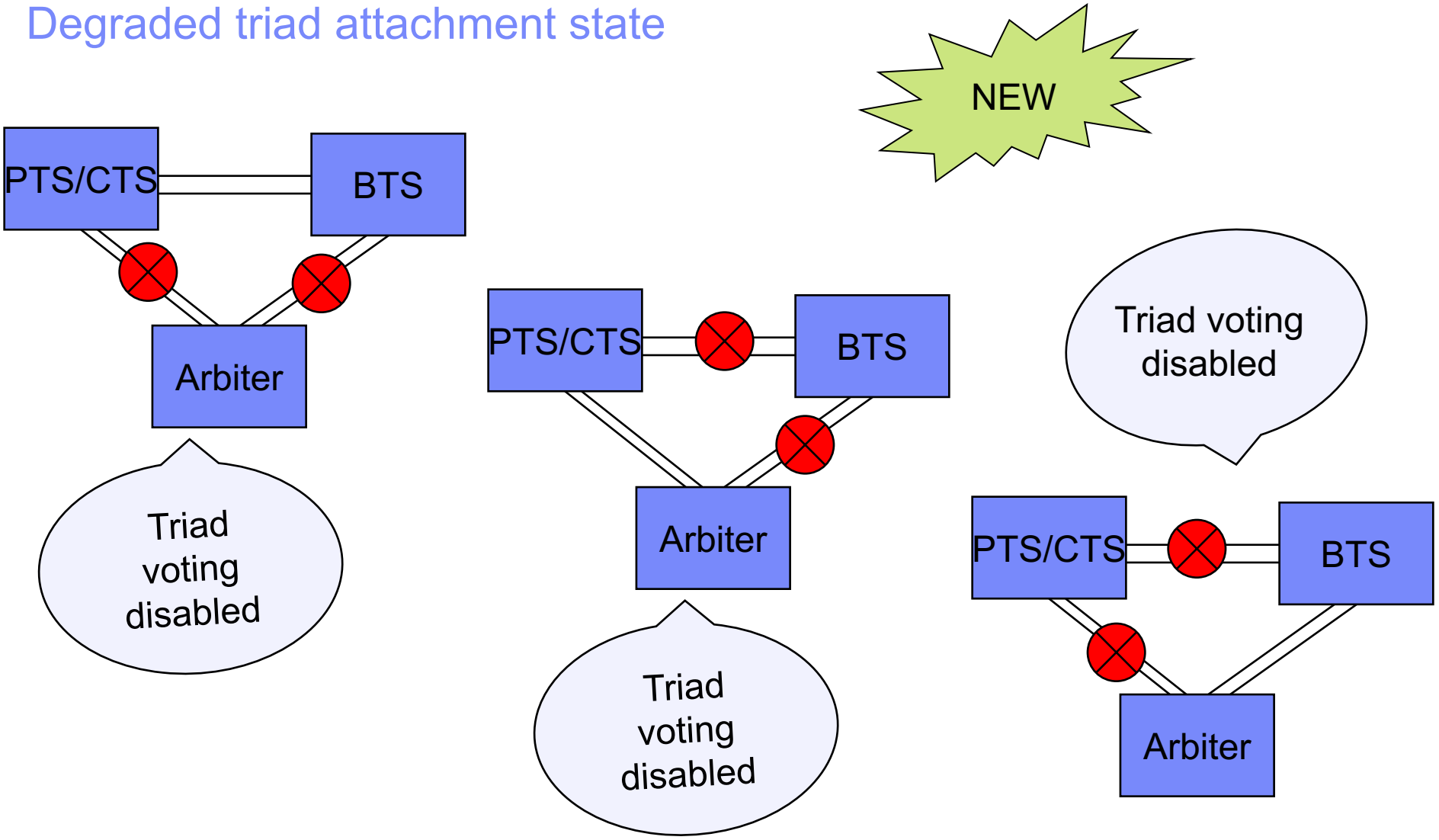
TRIAD VOTING

BTS can take over as CTS
using Arbiter Assisted recovery.

CTS will surrender its role when it loses
attachment to the remaining special role server!

One pair of special role servers not attached

Degraded triad attachment state



When any two of the special role servers cannot communicate with the third role

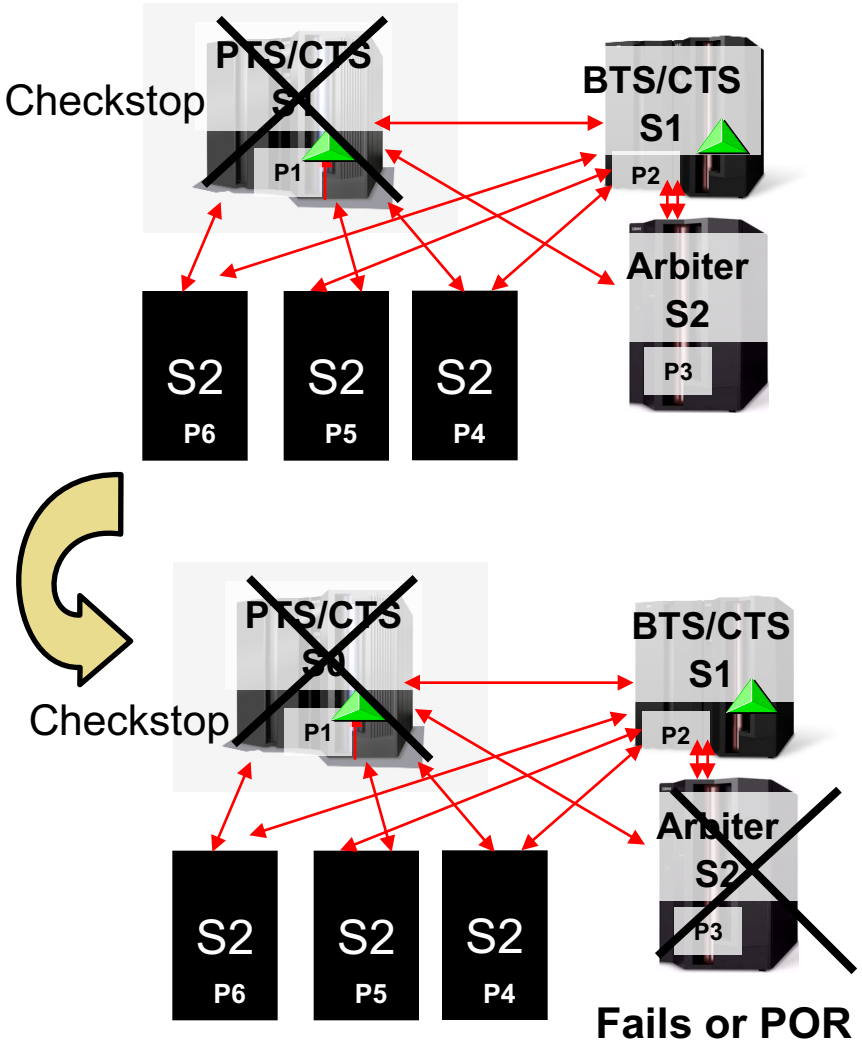
STP design changes – degraded triad state

- When triad state transitions to degraded triad attachment state, either from full attachment state or partial attachment state, triad voting is disabled by the CTS

- When triad voting has been disabled
  - BTS cannot take over as CTS using Arbiter Assisted recovery
  - CTS will not surrender its role when it loses attachment to the remaining special role server
  - BTS can still take over as CTS using either Console Assisted Recovery (CAR) or the STP Going Away Signal (GOSIG)

- Triad voting is re-enabled when triad state transitions back to full attachment state

BTS takeover as CTS example – after design change



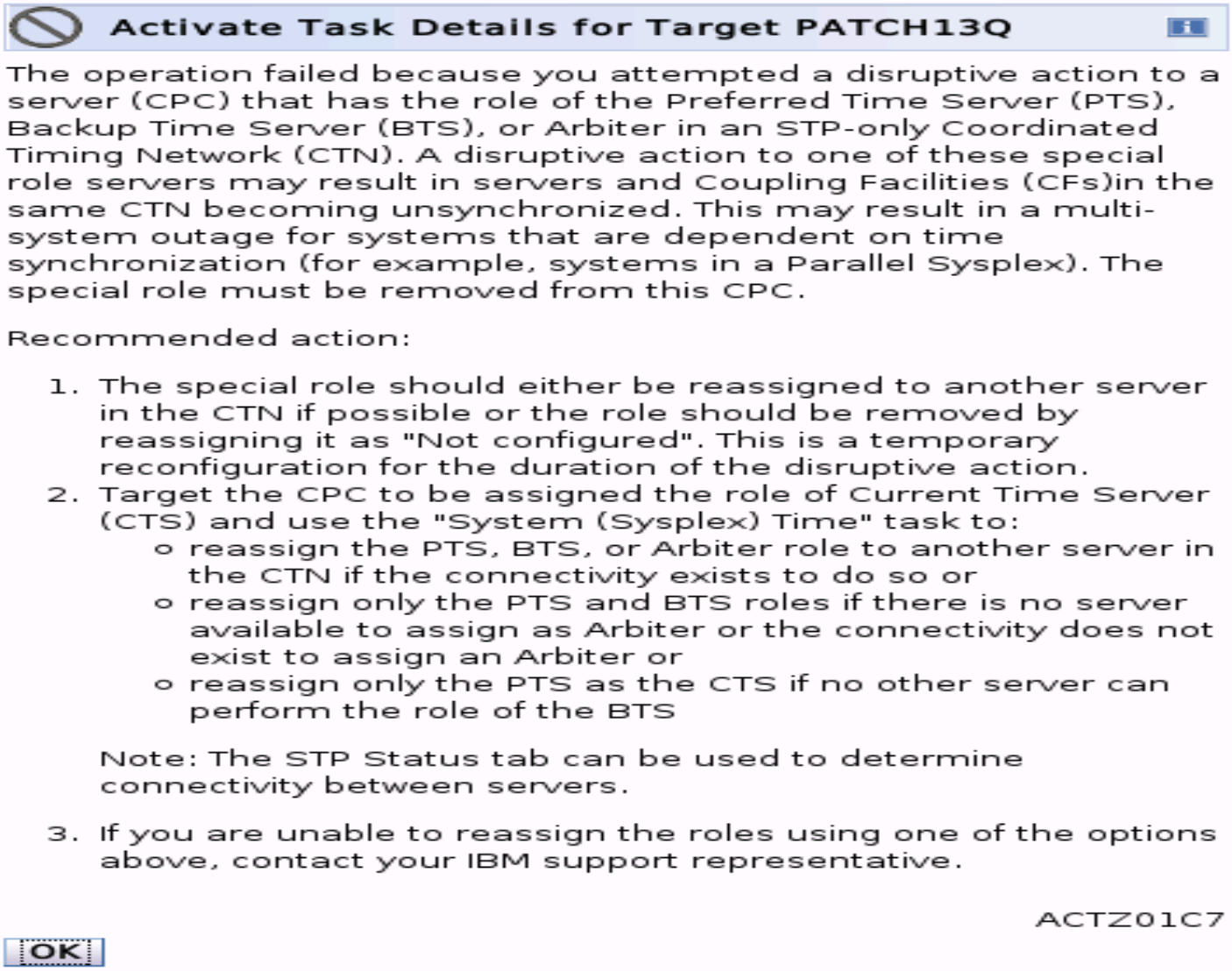
- PTS checkstop
  - BTS take over as S1
  - **Triad voting disabled**
- **Assume roles not reassigned**
- If subsequently BTS fails, no other server available to take over as S1
  - **Customer should still reassign PTS, BTS, Arbiter to maintain robust configuration**
- If subsequently Arbiter fails or is PORed (planned action)
  - **BTS/CTS does not surrender S1 role**
    - CECs with P4, P5, P6 stay synchronized as S2 servers
- **Minimal impact to Sysplex**

Coupling links

Block disruptive planned actions

- Block planned disruptive action from proceeding on the BTS and Arbiter using similar design used to block the PTS/CTS
  - Pros:
    - Safeguard for disruptive planned actions performed as part of same HMC task on both the BTS and Arbiter or done sequentially
      - ✓ Forces role to be removed or reassigned prior to proceeding with disruptive action
    - Consistency – same procedure for all 3 special role servers
  - Cons:
    - Does not safeguard against unplanned double failures
    - Customers will have to have special operational procedures for 3 servers in the CTN, instead of just the PTS/CTS
      - ✓ Remove or reassign role before planned disruptive action
      - ✓ Remember to restore assignment of roles after disruptive action
- See http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102019

HMC message



Activate Task Details for Target PATCH13Q

The operation failed because you attempted a disruptive action to a server (CPC) that has the role of the Preferred Time Server (PTS), Backup Time Server (BTS), or Arbiter in an STP-only Coordinated Timing Network (CTN). A disruptive action to one of these special role servers may result in servers and Coupling Facilities (CFs)in the same CTN becoming unsynchronized. This may result in a multi-system outage for systems that are dependent on time synchronization (for example, systems in a Parallel Sysplex). The special role must be removed from this CPC.

Recommended action:

1. The special role should either be reassigned to another server in the CTN if possible or the role should be removed by reassigning it as "Not configured". This is a temporary reconfiguration for the duration of the disruptive action.
2. Target the CPC to be assigned the role of Current Time Server (CTS) and use the "System (Sysplex) Time" task to:
   - reassign the PTS, BTS, or Arbiter role to another server in the CTN if the connectivity exists to do so or
   - reassign only the PTS and BTS roles if there is no server available to assign as Arbiter or the connectivity does not exist to assign an Arbiter or
   - reassign only the PTS as the CTS if no other server can perform the role of the BTS

   Note: The STP Status tab can be used to determine connectivity between servers.
3. If you are unable to reassign the roles using one of the options above, contact your IBM support representative.

ACTZ01C7

OK

Hardware Messages

The STP preferred time server and backup time server have both lost connectivity to the arbiter.
STP recovery is in a degraded state. Please re-establish connectivity or reassign STP server roles.

The STP preferred time server and arbiter have both lost connectivity to the backup time server.
STP recovery is in a degraded state. Please re-establish connectivity or reassign STP server roles.

The STP backup time server and arbiter have both lost connectivity to the preferred time server.
STP recovery is in a degraded state. Please re-establish connectivity or reassign STP server roles.

There is no connectivity between the STP preferred time server and arbiter.
Please re-establish connectivity or reassign STP server roles.

There is no connectivity between the STP preferred time server and the backup time server.
Please re-establish connectivity or reassign STP server roles.

There is no connectivity between the STP backup time server and the arbiter.
Please re-establish connectivity or reassign STP server roles.

There exists connectivity between each of the STP preferred time server, the backup time server and the arbiter. This is normal operations and all is well in the world.

Agenda

- Introduction
  - Description
  - Key Attributes
  - Terminology
- Configurations
  - Mixed Coordinated Timing Network
  - STP-only Coordinated Timing Network
- External Time Source Options
  - NTP server
  - NTP server with PPS
- Timing-only Links
- Enhancements (Oct 08/Jan 09)
- Enhancements (April 28, 2009 announce)
- Enhancements (September 2011)
- **Hardware and Software Prerequisites**
- Summary

Hardware Prerequisites

- 9037-002 concurrent LIC upgrade (if migrating from ETR network)
  - 9037 code changes to support STP Mixed CTN
- z14
- z13
- zEnterprise
- System z10
- System z9
- z990 and z890
- Concurrently install STP Enablement MCL (Feature Code 1021)

Software Prerequisites

- z/OS 1.7 current releases
- Additional PTFs may be required
  - Maintenance can be applied using “rolling IPL” process
- Check Preventive Service Planning (PSP) buckets
  - Listed in the PSP buckets for the servers and coupling facilities
- To simplify identification of PTFs for STP, functional PSP bucket created
  - Use the Enhanced Preventive Service Planning Tool (EPSPT)
  - www14.software.**ibm.com**/webapp/set2/psp/srchBroker
- Coexistence with z/OS 1.4 through z/OS 1.6
  - Mixed CTN can include pre-1.7 systems
  - PTFs required for toleration code
  - z/OS 1.4 through 1.6 – service has been discontinued

Update CLOCKxx statements

- OPERATOR PROMPT|NOPROMPT
- TIMEZONE W|E hh.mm.ss
- ETRMODE YES|NO
- ETRZONE YES|NO
- SIMETRID nn
  - nn = 0 – 31
- STPMODE* YES|NO
  - Specifies whether z/OS is using STP timing mode
  - STPMODE YES default
- STPZONE* YES|NO
  - Specifies whether the system is to get the time zone constant from STP
- ETRDELTA ss | TIMEDELTA* ss
  - ss = 0 – 99 seconds

* New statements for STP

Agenda

- Introduction
  - Description
  - Key Attributes
  - Terminology
- Configurations
  - Mixed Coordinated Timing Network
  - STP-only Coordinated Timing Network
- External Time Source Options
  - NTP server
  - NTP server with PPS
- Timing-only Links
- Enhancements (Oct 08/Jan 09)
- Enhancements (April 28, 2009 announce)
- Hardware and Software Prerequisites
- **Summary**

Summary

- Allows multi site Parallel Sysplex distances to extend to 200 km (with RPQ)
  - Limits set by coupling protocol and links
- Can help meet more stringent time synchronization requirements and expected to scale with technology as processors and messaging technology improve
- Does not require dedicated Timer links
  - Uses same hardware and protocols as data
- Allows concurrent migration from an ETR network with proper planning
- Three external Time Source options
  - Dial-out (100ms accuracy)
  - NTP server (100ms accuracy)
    - Time coordination across platforms
  - NTP server with PPS (10 us accuracy)
- Continuous availability of ETS without switching server roles

Additional Information

- Redbooks®
  - Server Time Protocol Planning Guide SG24-7280
  - Server Time Protocol Implementation Guide SG24-7281
  - Server Time Protocol Recovery Guide SG24-7380
- Education
  - Introduction to Server Time Protocol (STP)
    - Available on Resource Link™
    - www.**ibm.com**/servers/resourcelink/hom03010.nsf?OpenDatabase
- STP Web site
  - www.**ibm.com**/systems/z/pso/stp.html
- Systems Assurance
  - The IBM team is required to complete a Systems Assurance Review (SAPR Guide SA06-012).

IBM Implementation Services for System z – Server Time Protocol (6948-J56)

Offering Description	• **This offering is designed to assist clients to quickly and safely implement Server Time Protocol within their existing environments. STP provides clients with the capability to efficiently manage time synchronization within their multi-server infrastructure. Following best practices and using detailed planning services, IBM helps clients identify various implementation models and engage in the appropriate configuration required to effectively support STP for driving a more responsive business and IT infrastructure.**
Program, Play, Industry Alignment	• **Infrastructure Improvement; Energy Efficiency; Better performance and lower operational cost**
Client Value ***(enables customers to...)***	• **Swift and secure implementation of STP for improved availability, integrity and performance** • **Improves multi-server time synchronization without interrupting operations** • **Enables integration with next generation of System z infrastructure**
Target Audience	• **Primarily core, Large Enterprise customers.** • **Existing z midrange clients**
Key Competitors	• **In house staff**
Competitive Differentiation	• **Leverages best practices with secure implementation** • **Short implementation time – lower risk** • **Provides support and facilitates knowledge sharing through IBM's mainframe expertise**
Proof Points & Claims for Client Value / Differentiation	• **Need to safely implement a reliable replacement for Sysplex Timer® while maintaining continuous operations** • **Cost of providing and maintaining hardware, floor space and solution support for additional Sysplex Timer intermediate site** • **Lack of in-house expertise, skills and resources for implementing Server Time Protocol**
Engagement Portfolio	• **http://spimweb1.boulder.ibm.com/services/sosf/dyno.wss?oid=50423&loc=All&langcd=en-US#1**
Offering Manager	• **Anna Lee/Southbury/IBM, 512-590-8914, T/L: 268-9318**

IBM Announces –
IBM Implementation Services for System z – Server Time Protocol

Implementation of STP for improved availability and performance

Offering

Assist clients to quickly and safely implement Server Time Protocol within their existing environments. IBM helps clients identify various implementation models and engage in the appropriate configuration required to effectively support STP for driving a more responsive business and IT infrastructure

Customer Value:

- Improves multi-server time synchronization without interrupting operations
- Enables integration with next generation of System z infrastructure
- Swift and secure implementation of STP for improved availability, integrity, and performance
- Reduces hardware maintenance and power costs while eliminating intermediate site requirements for Sysplex Timer

Leverages IBM's knowledge and best practices to help implementation of Server Time Protocol

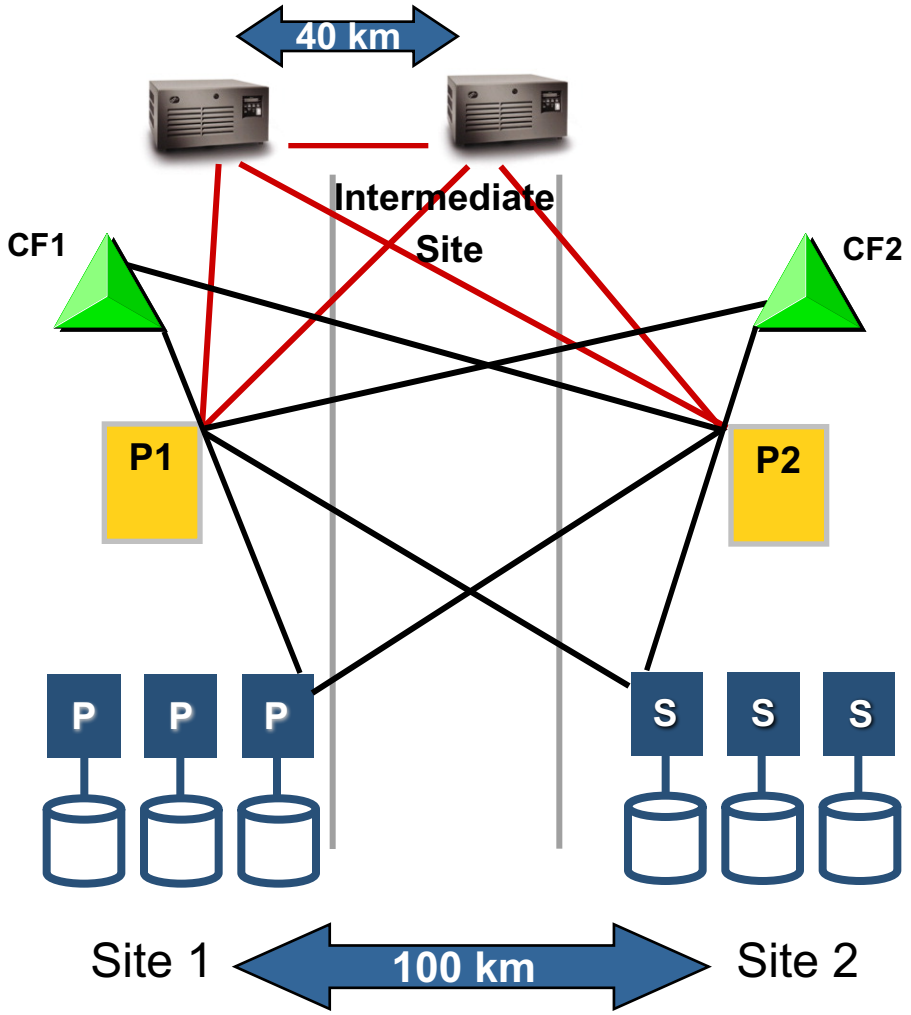
Reference Material - Terminology

APAR	Authorized Program Analysis Report
ARB	Arbiter
BTS	Backup Time Server
CF	Coupling Facility
CTS	Current Time Server
CTN	Coordinated Timing Network
DWDM	Dense Wave Division Multiplexer
ETR	External Time Reference
ETS	External Time Source
FC	Feature Code
HMC	Hardware Management Console
HCA	Host Channel Adapter
ICB	Integrated Cluster Bus
IPL	Initial Program Load
ISC	InterSystem Coupling Channel
LAN	Local Area Network
LIC	Licensed Internal Code
LPAR	Logically Partition
NTP	Network Time Protocol
PR/SM	Processor Resource / Systems Manager
PSIFB Infiniband	Parallel Sysplex
PTF	Temporary Program Fix
PTS	Preferred Time Server
SW and	Software (programs operating systems)
SE	Support Element
TPF	Operating System
UTC	Coordinated Universal Time
zVM	Operating System
zVSE	Operating System
z/OS	Operating System
z/VM	Operating System

Backup slides

STP Enhancements over ETR Network

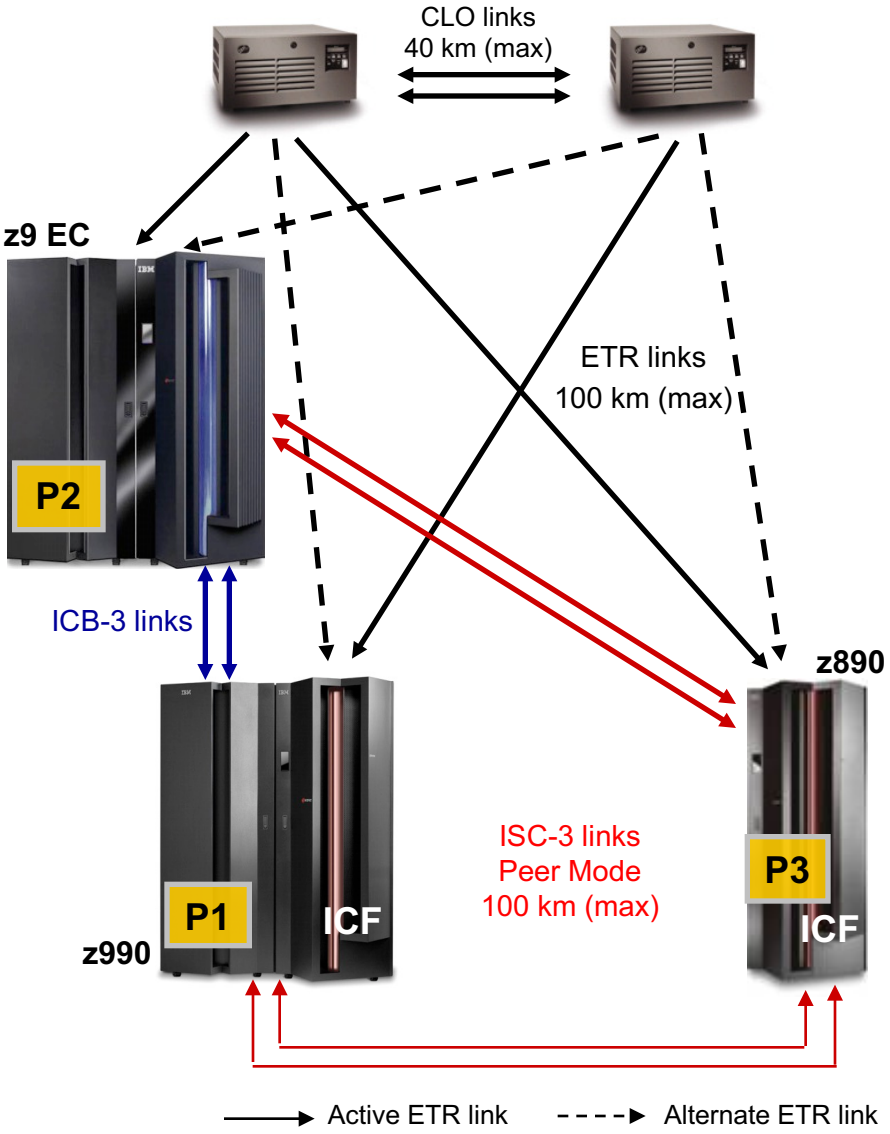
- STP supports a multi-site timing network of up to 200 km without requiring an intermediate site
- Fiber distance between Sysplex Timers cannot exceed 40 km
  - Intermediate site to locate second timer recommended to avoid a single point of failure, if data centers more than 40 km apart



GDPS configuration based on ETR Network

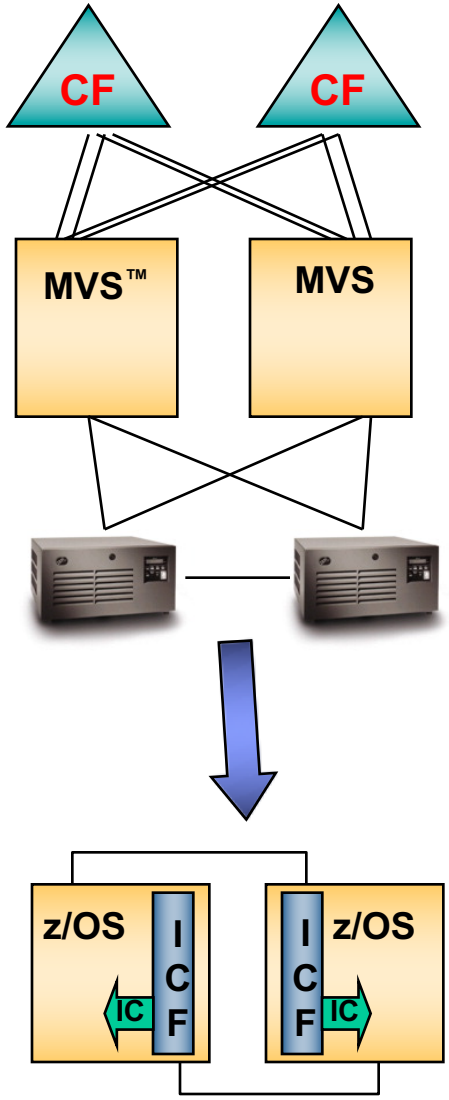
STP Enhancements over ETR Network *(cont)* …

- STP design can allow more stringent synchronization
  - Time of Day (TOD) clocks of servers must be synchronized within the fastest messaging time between servers
    - In this diagram, between IBM System z9 Enterprise Class (z9 EC) and z990 (shortest messaging time in example) compared to between z990 and z890 (longer messaging time in example)
  - "Best case" messaging times over ICB links in Parallel Sysplex configuration (8 us approximately) approaching "Worst case" TOD synchronization between CECs stepping to 9037s 40 km apart (4 us approx.)

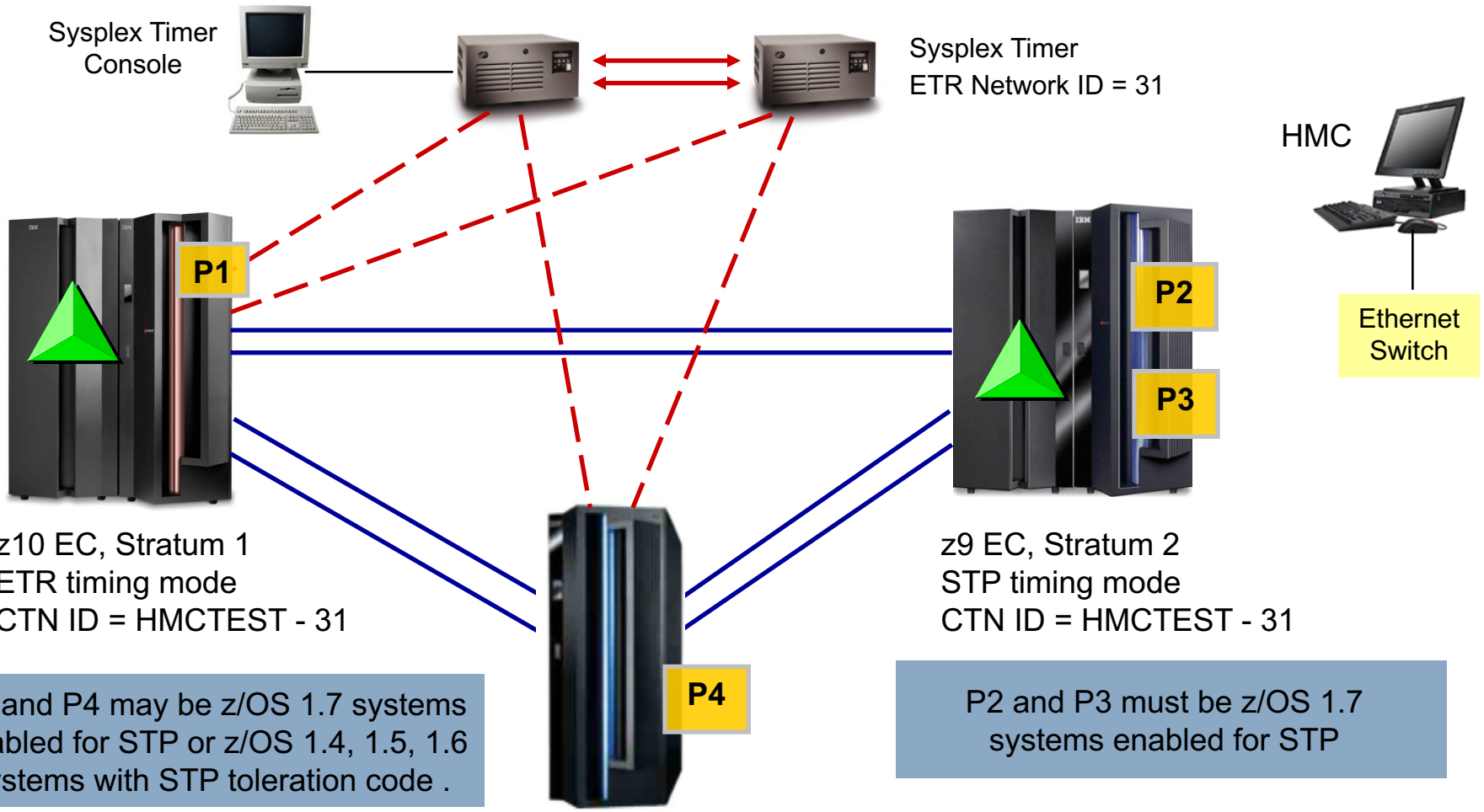


Server Time Protocol Potential Value

- Helps Eliminate
  - Infrastructure requirements (space, power, etc.) to support Sysplex Timers
  - Sysplex Timer maintenance costs.
  - Dark fiber between sites for ETR and CLO links
- Helps Reduce
  - Fiber optic infrastructure requirements for DWDM ports, patch/trunk cables
- Helps improve System Management
  - Allows automatic adjustment of Daylight Saving Time offset based on time zone algorithm
    - With ETR network, you need to schedule DST offsets at least twice a year manually at the Sysplex Timer console
  - Allows gradual time adjustment of up to +/- 60 seconds
    - Sysplex Timer allows time adjustments of up to +/- 4.999 seconds



Timing Modes in a Mixed CTN (example)



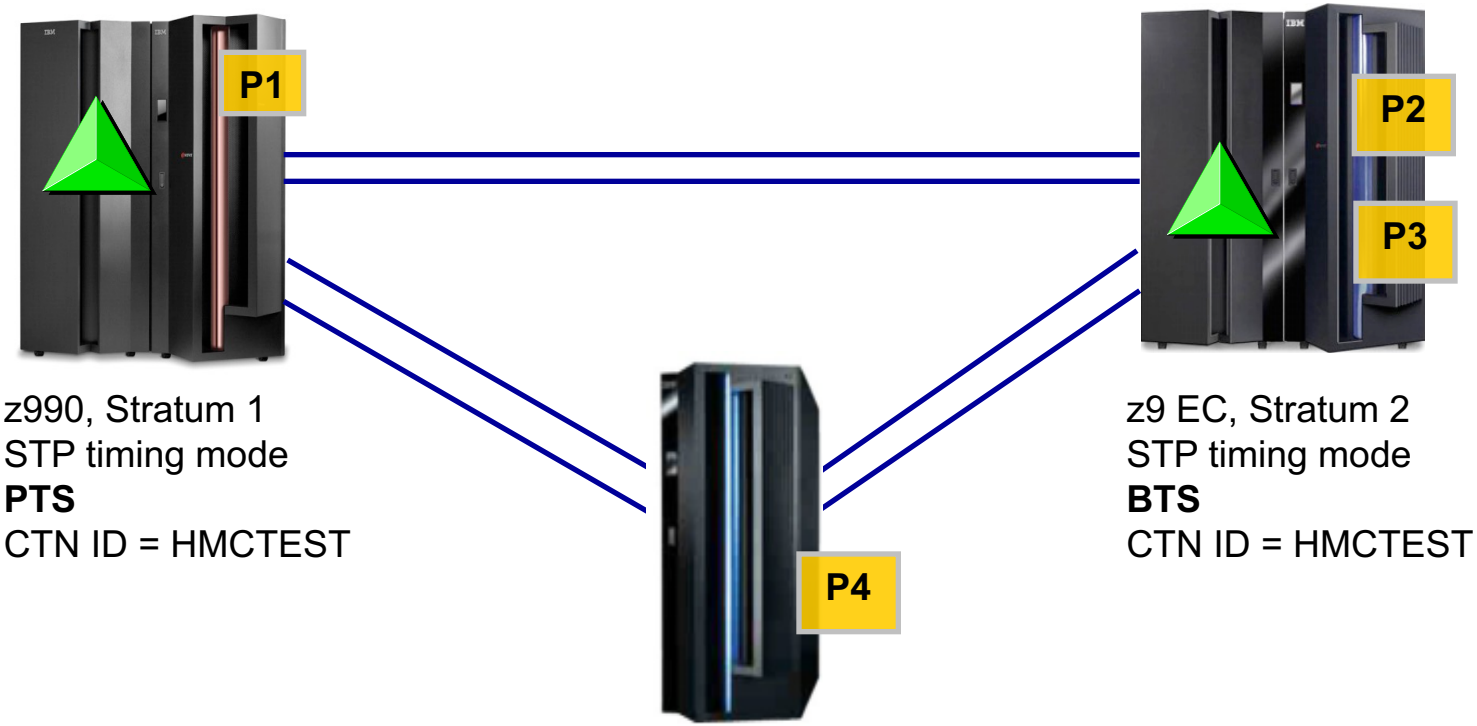
P1 and P4 may be z/OS 1.7 systems enabled for STP or z/OS 1.4, 1.5, 1.6 systems with STP toleration code .

P2 and P3 must be z/OS 1.7 systems enabled for STP

P1, P2, P3, P4 – systems in a Parallel Sysplex

Timing Mode in an STP-only CTN (example)

P1, P2, P3 and P4 must be z/OS 1.7 systems enabled for STP



z990, Stratum 1
STP timing mode
PTS
CTN ID = HMCTEST

z9 EC, Stratum 2
STP timing mode
BTS
CTN ID = HMCTEST

z9 BC, Stratum 2
STP timing mode
CTN ID = HMCTEST

P1, P2, P3, P4 – systems in a Parallel Sysplex

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

GDPS*	Parallel Sysplex*	System z*	z10
Geographically Dispersed Parallel Sysplex	PR/SM	System z9*	z10 BC
IBM	Redbooks*	System z10	z10 EC
IBM eServer	Resource Link	System z10 Business Class	z/OS*
IBM (logo)	Sysplex Timer*	z9*	zSeries*

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

INFINIBAND, InfiniBand Trade Association and the INFINIBAND design marks are trademarks and/or service marks of the INFINIBAND Trade Association

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

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

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

Red Hat, the Red Hat "Shadow Man" logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

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

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the 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.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

ZSP03176-USEN-00