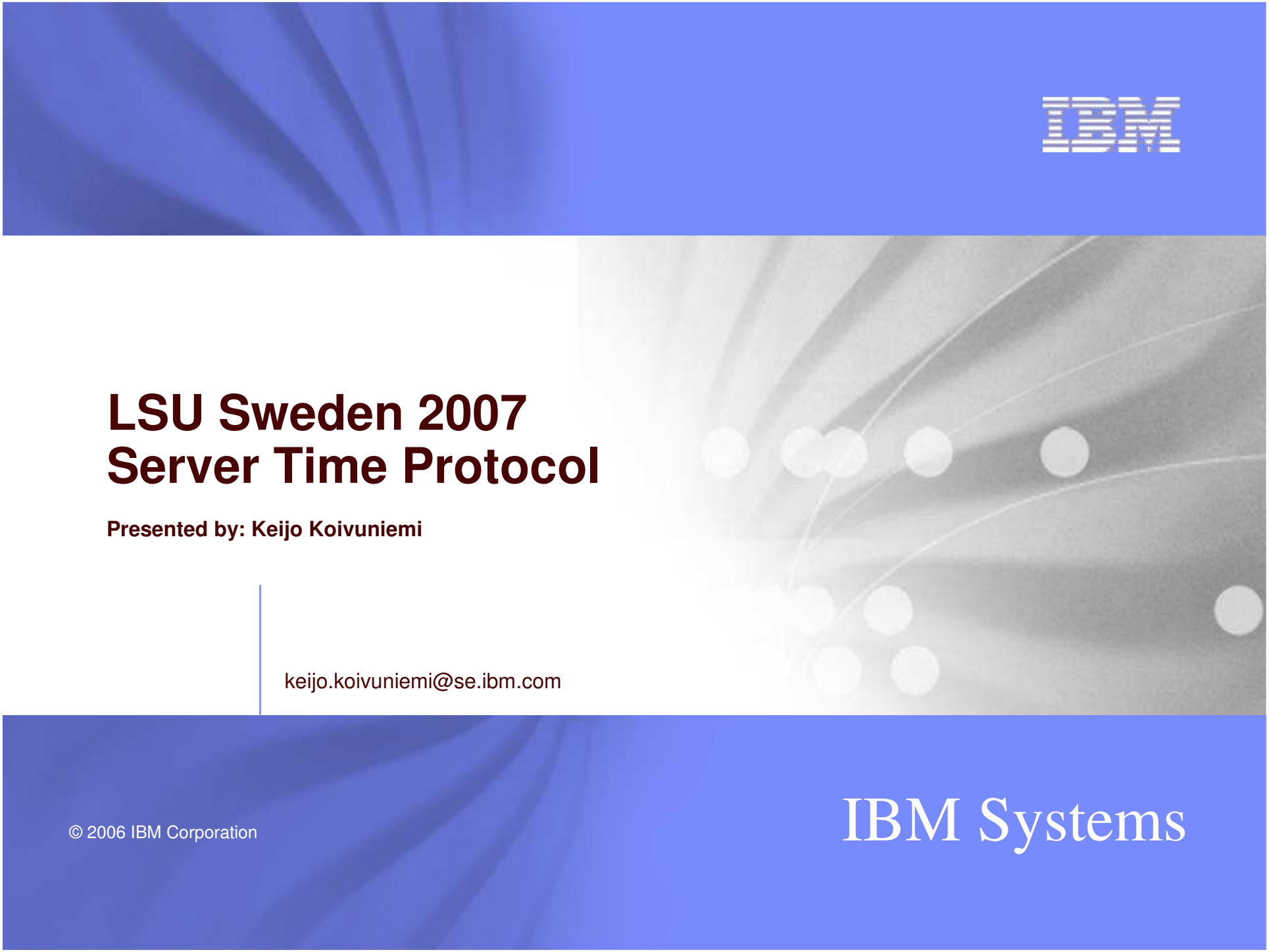


LSU Sweden 2007
Server Time Protocol

Presented by: Keijo Koivuniemi

keijo.koivuniemi@se.ibm.com

IBM Systems

Server Time Protocol Overview

Time Synchronization – The Next Generation

IBM Systems

Trademarks

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

GDPS*
IBM*
IBM eServer
IBM logo*
MVS
Parallel Sysplex*
PR/SM
Redbook
Resource Link
Sysplex Timer*
System z9
z/OS
zSeries

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Intel is a trademark of Intel Corporation in the United States, other countries, or both.

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.

SAP and Netweaver are registered trademarks of SAP AG in Germany and in other countries.

* 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.

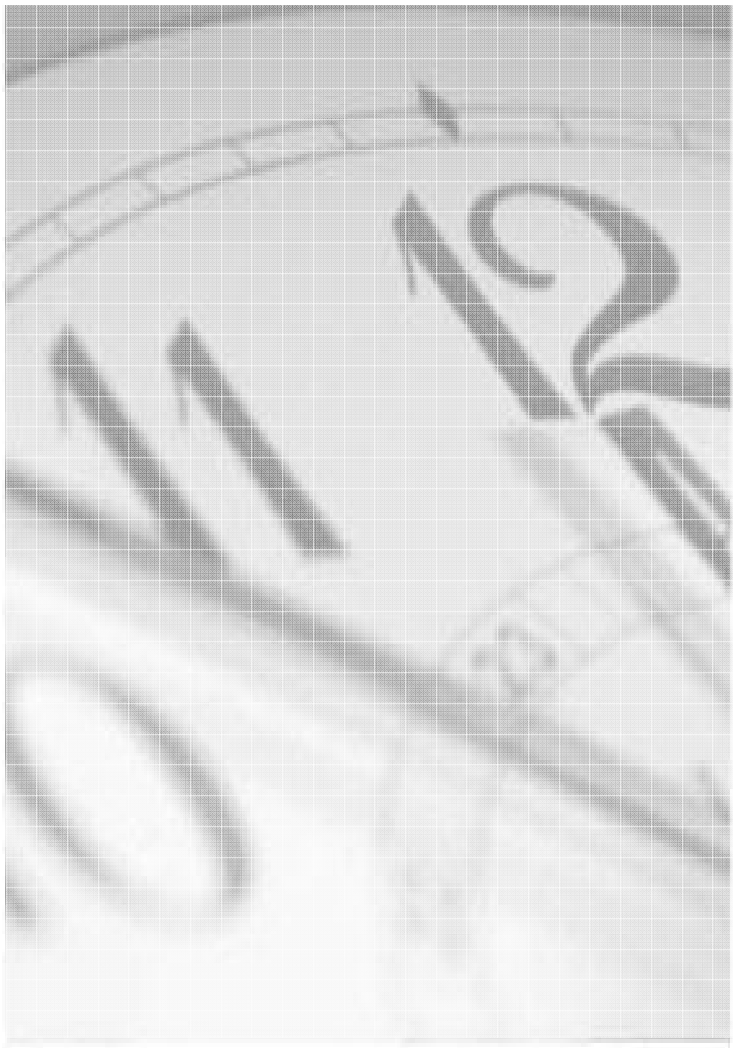
Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Server Time Protocol Technical Overview - Agenda

- Description
- Key Attributes Summary
- Value
- Terminology
- Mixed Coordinated Timing Network
- STP-only Coordinated Timing Network
- Prerequisites
- Summary
- Additional Information

What is Server Time Protocol (STP)?

- Provides 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)
- Server-wide facility implemented in IBM System z9 EC, z9 BC, IBM eServer™ zSeries® 990 and 890 (z990, z890) Licensed Internal Code (LIC)
  - Single view of "time" to PR/SM™
  - PR/SM can virtualize this view of time to the individual partitions (LPARs)
- Message-based time synchronization protocol
  - **Similar to Network Time Protocol (NTP) industry standard**
  - Timekeeping information transmitted over Coupling Links
  - NOT standard NTP but…

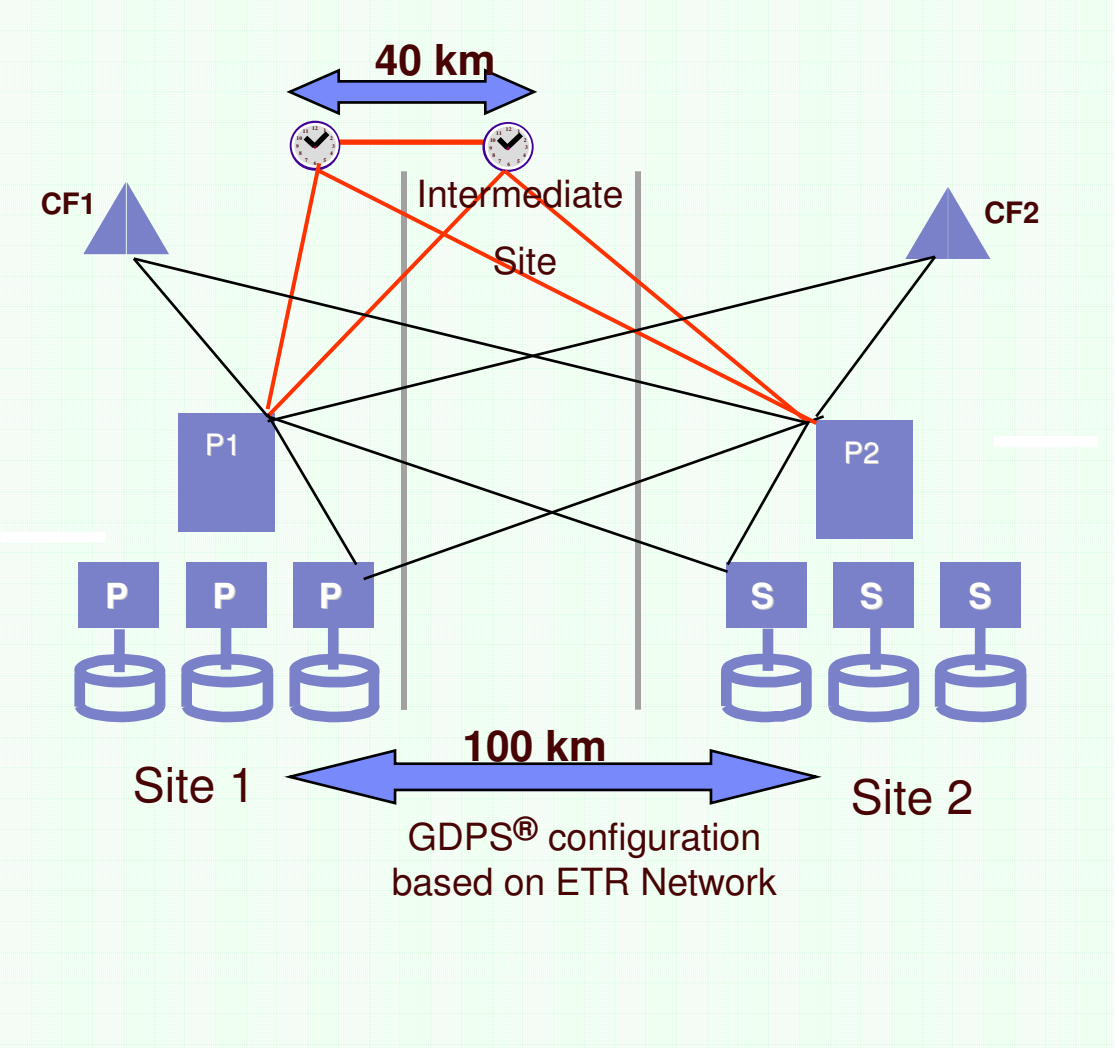


Key Attributes

- Designed to provide improved time synchronization, compared to Sysplex Timer®, for servers in a Sysplex or non-Sysplex configuration
- Supports a multi-site timing network of up to 100 km over fiber optic cabling
  - Allows a Parallel Sysplex® cluster to span **up to** 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

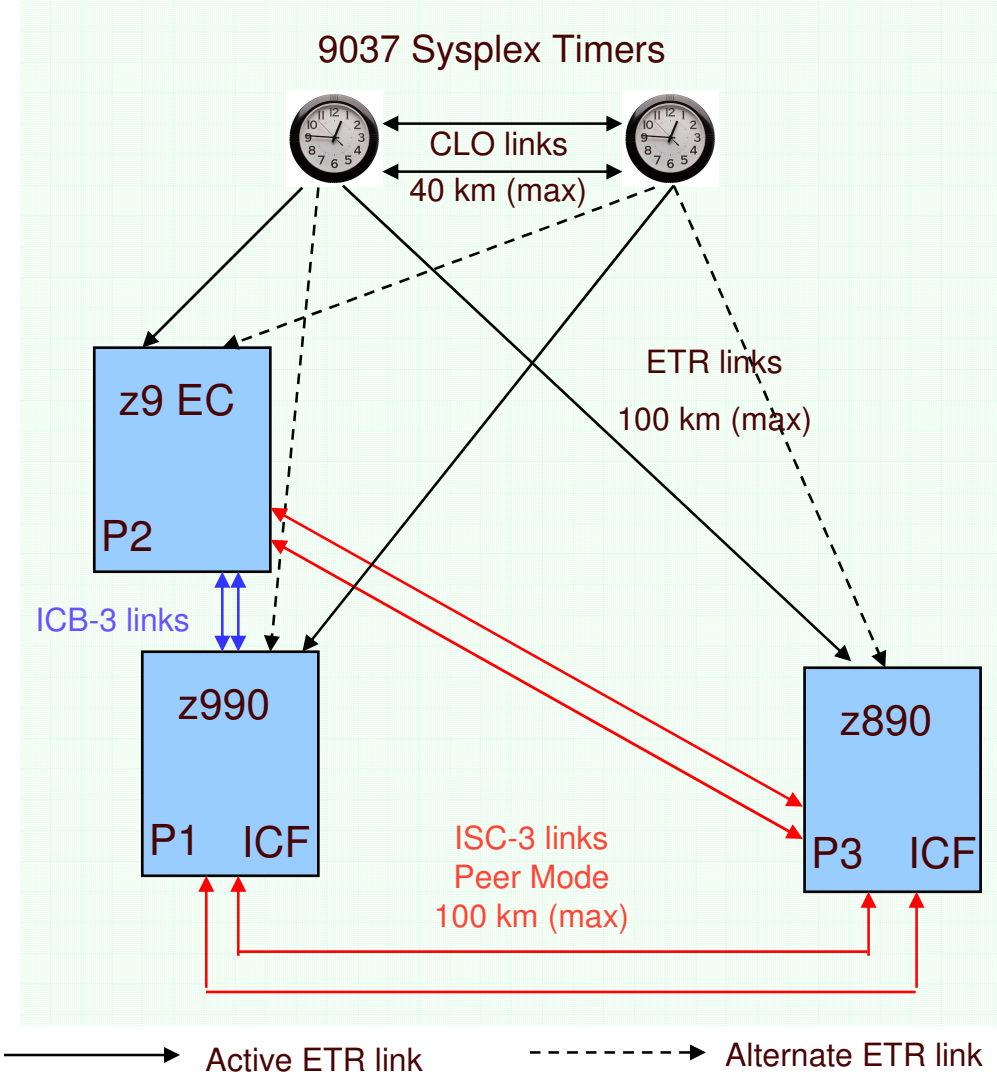
STP Enhancements over ETR Network

- **STP supports a multi-site timing network of up to 100 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



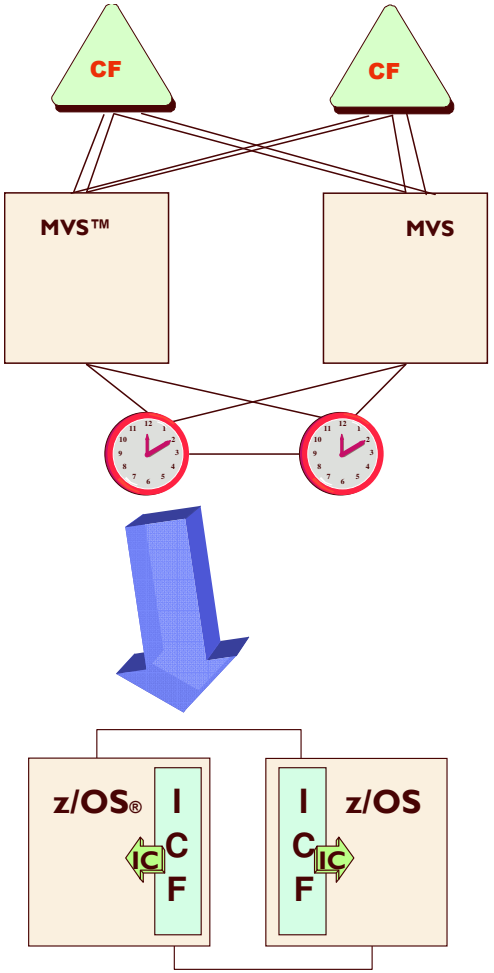
STP Enhancements over ETR Network *(cont)* …

- **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 (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



Terminology

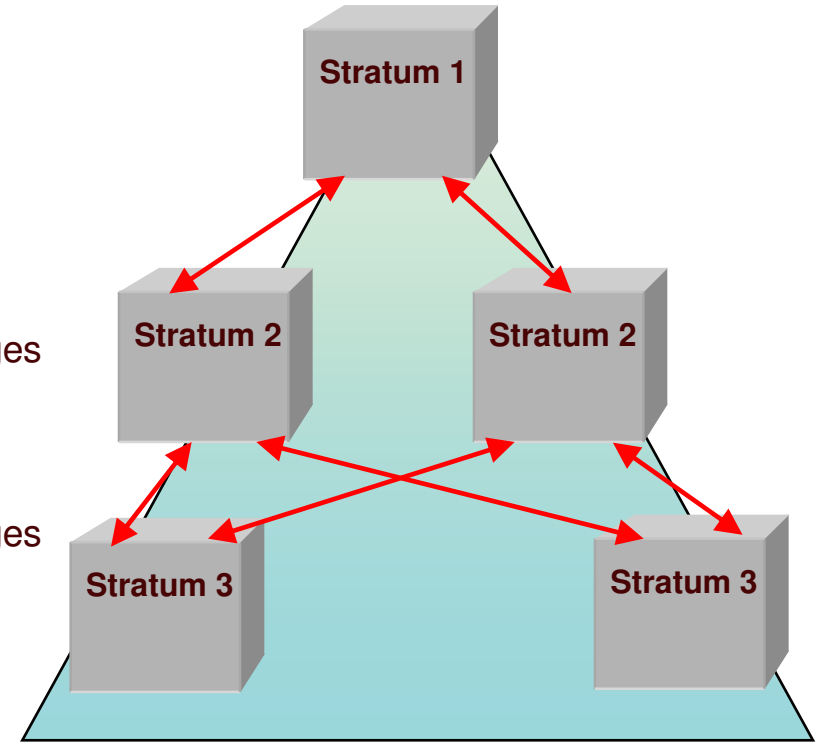
- STP-capable server/CF
  - z9 EC, z9 BC, z990, z890 server/CF with STP LIC installed
- STP-enabled server/CF
  - STP-capable server/CF with STP FC 1021 installed
    - STP panels at the HMC/SE can now be used
- STP-configured server/CF
  - STP-enabled server/CF with a CTN ID assigned
    - STP message 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 *(cont)*

- 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
  - 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)

Terminology (cont) …

- 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 S3



Time message will find a new path if needed

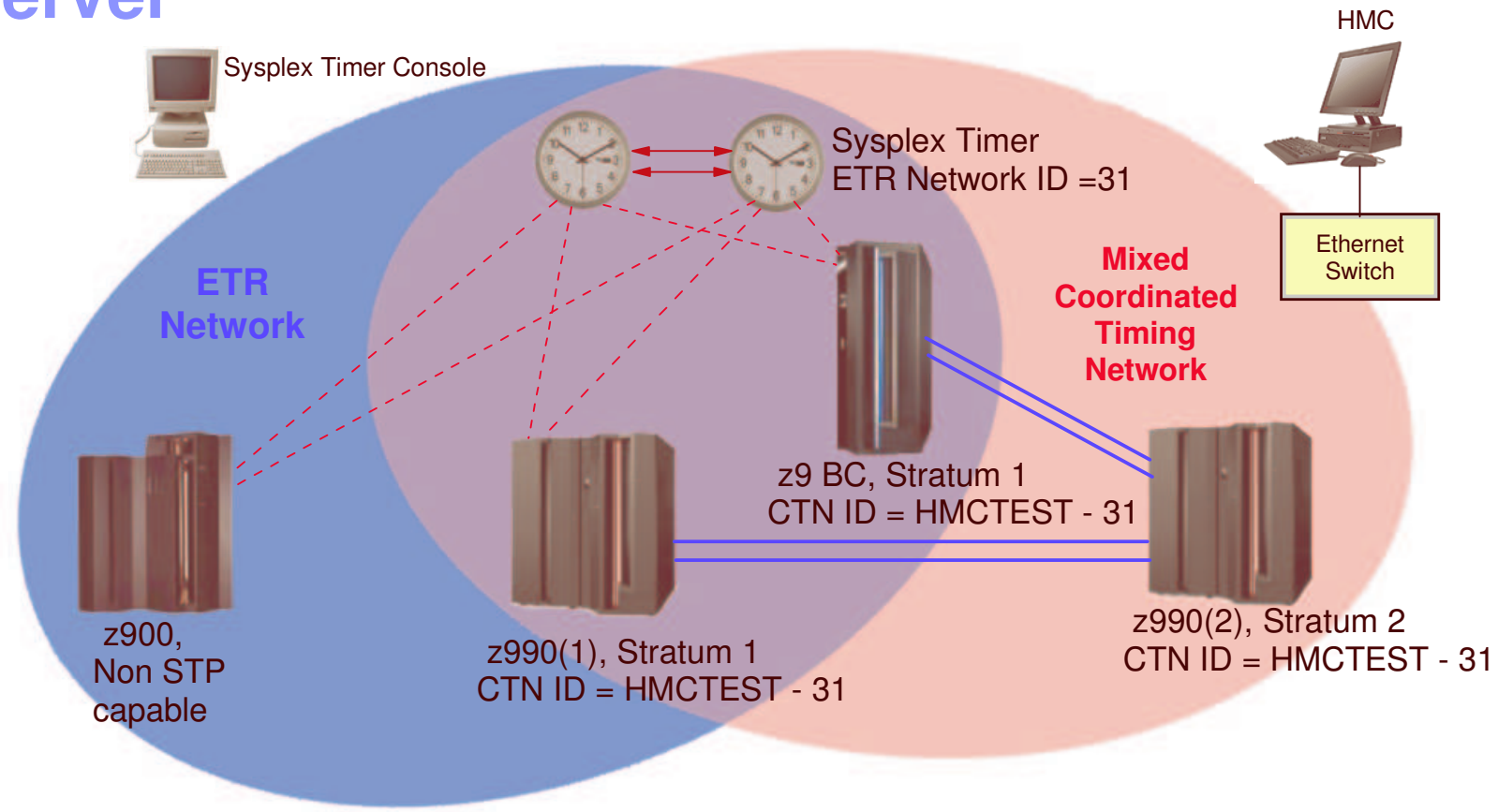
Mixed Coordinated Timing Network (CTN)

- 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
  - 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

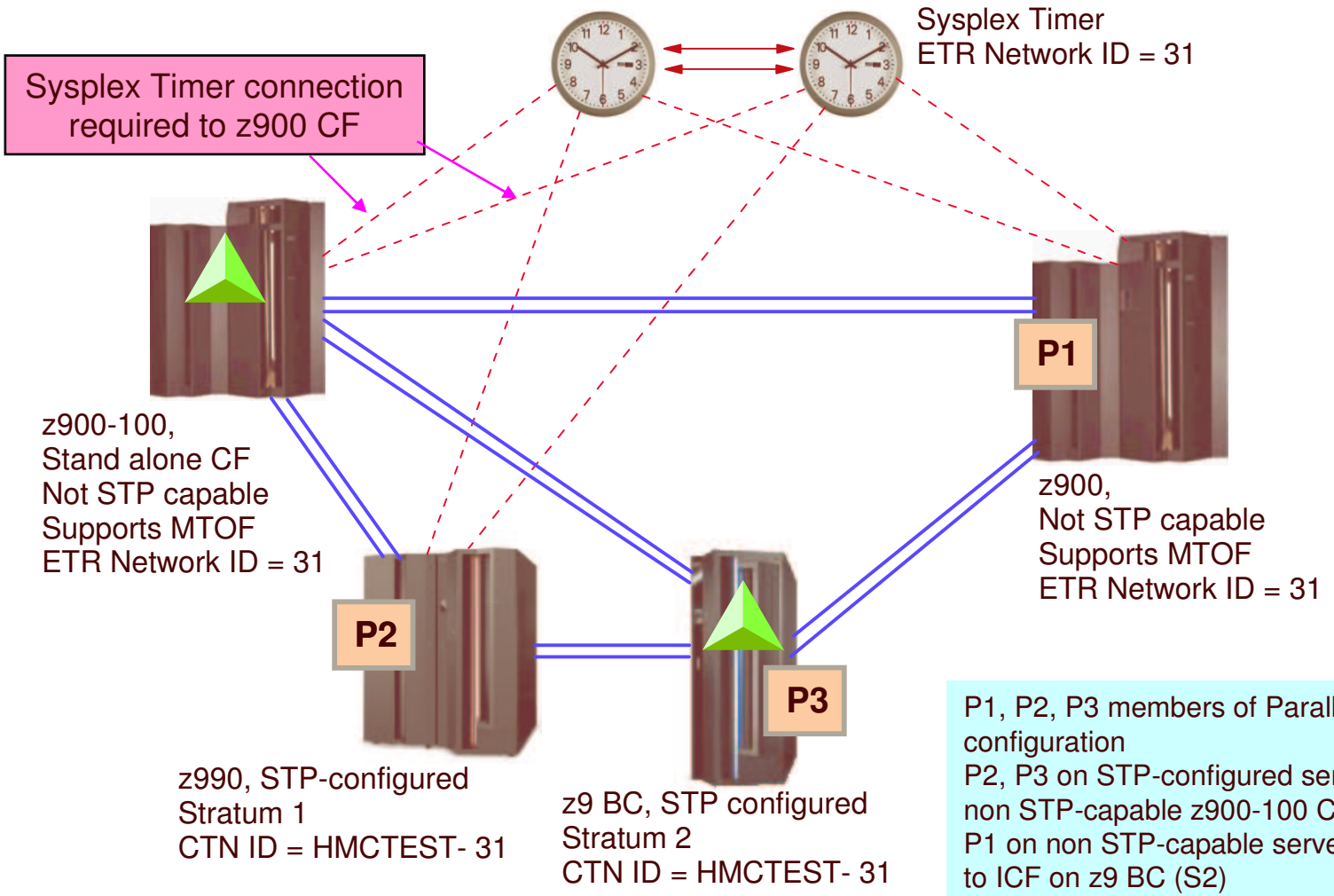
- **Sysplex Timer console continues to be used for all timing related functions of the Mixed CTN**

Example: Coexistence with non STP-capable server



- z9 BC, z990(1), z900 synchronized to Sysplex Timer
- z9 BC, z990(1) are Stratum 1 servers
  - Two Stratum 1 servers recommended to avoid single point of failure
- z990(2) synchronized to either z9 BC or z990(1) via STP is a Stratum 2 server
  - z990(2) does not need ETR link connections
  - z990(2) can be located up to 100 km away from z9 BC, z990(1)

Example: Non STP-capable server and CF in a z/OS Parallel Sysplex configuration

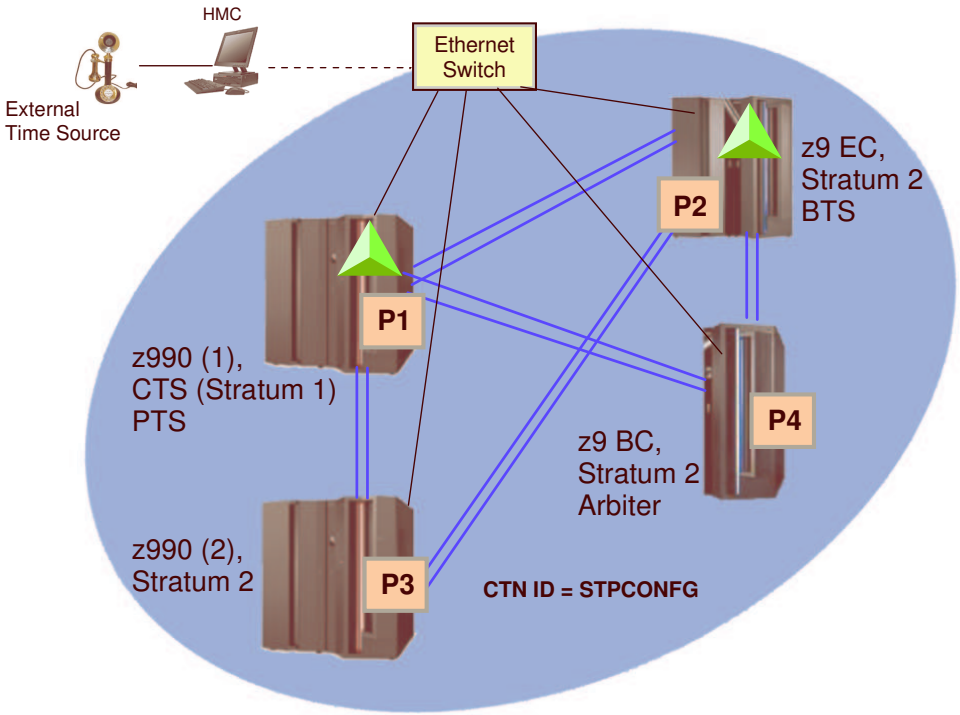


P1, P2, P3 members of Parallel Sysplex configuration
P2, P3 on STP-configured servers can connect to non STP-capable z900-100 CF
P1 on non STP-capable server (z900) can connect to ICF on z9 BC (S2)

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
  - **Arbiter**
    - Provides additional means to determine if BTS should take over as the CTS under unplanned outages

Example of STP-only CTN

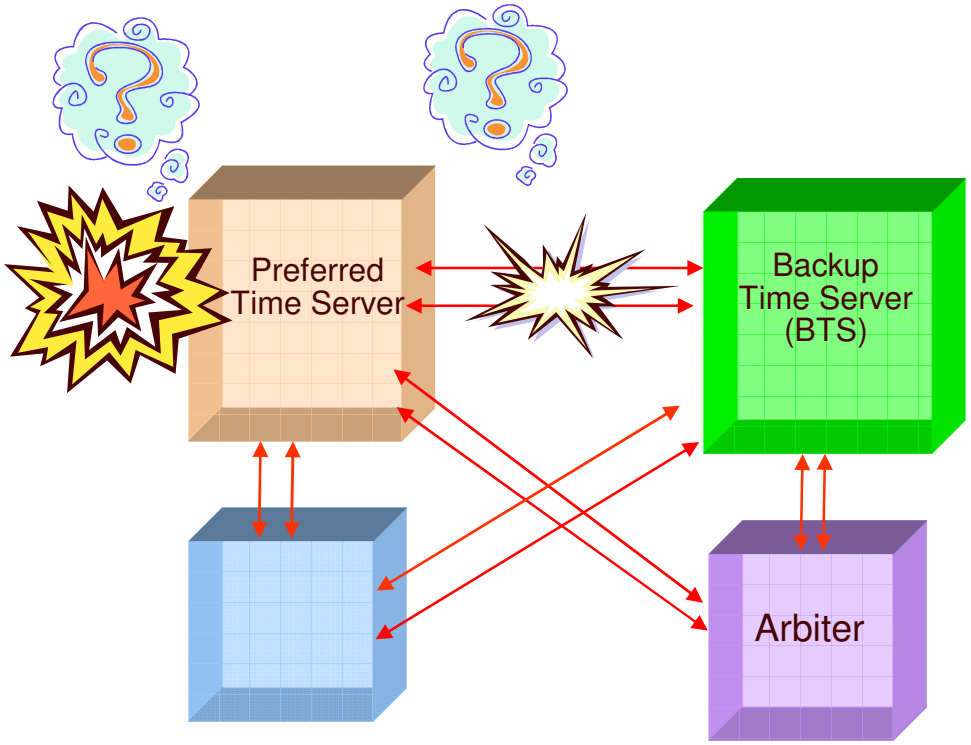


- Configuration has to be defined
  - Must assign PTS and CTS
  - Optionally assign BTS
    - Strongly recommended to allow near-continuous availability
  - Optionally assign Arbiter
    - Recommended for configurations of 3 or more servers/CFs
    - Can improve recovery

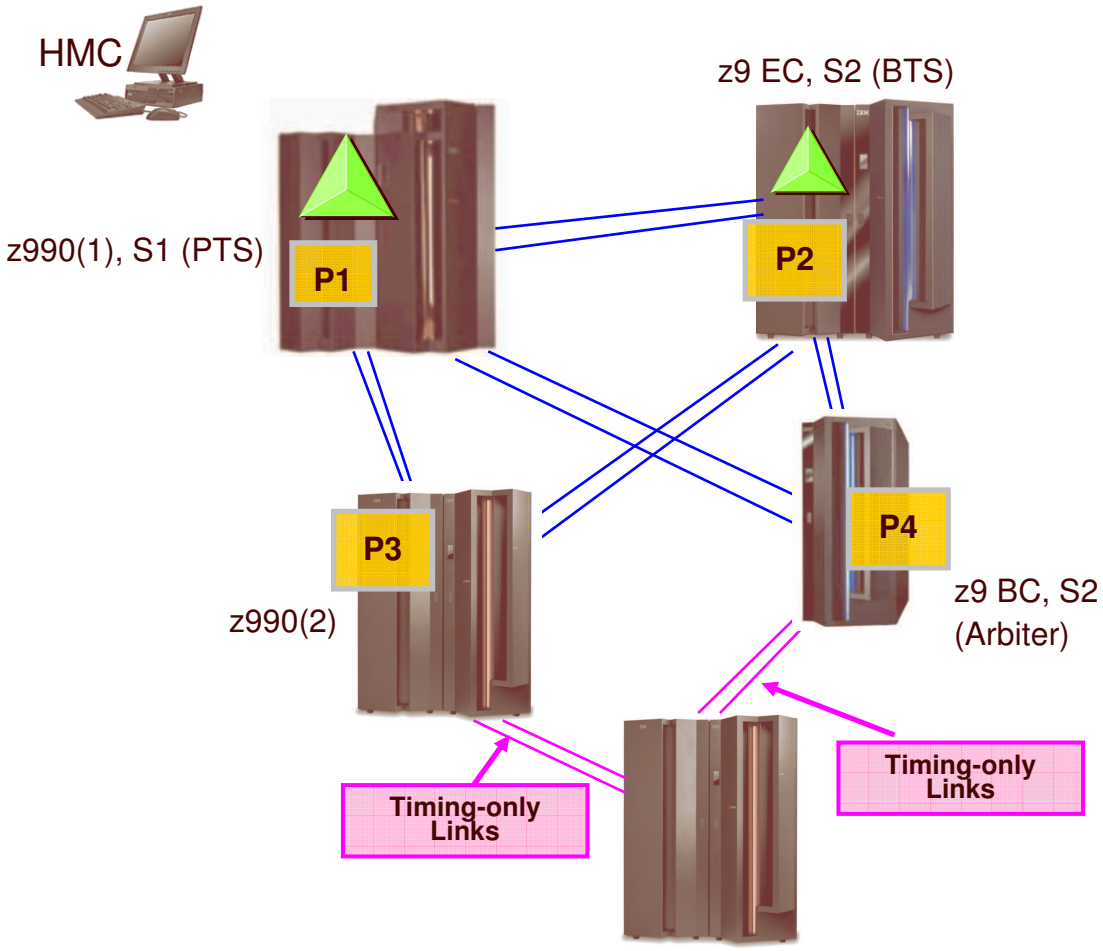
STP can use existing Coupling links

Example of Backup Time Server / Arbiter

- If the Backup Time Server loses communication with PTS on all established paths, what does it do?
  - It asks the Arbiter for its status
  - If Arbiter still getting signals (just links are lost), BTS becomes S3
  - If Arbiter also lost communication, (assume PTS failed), BTS takes over as CTS (Stratum 1)



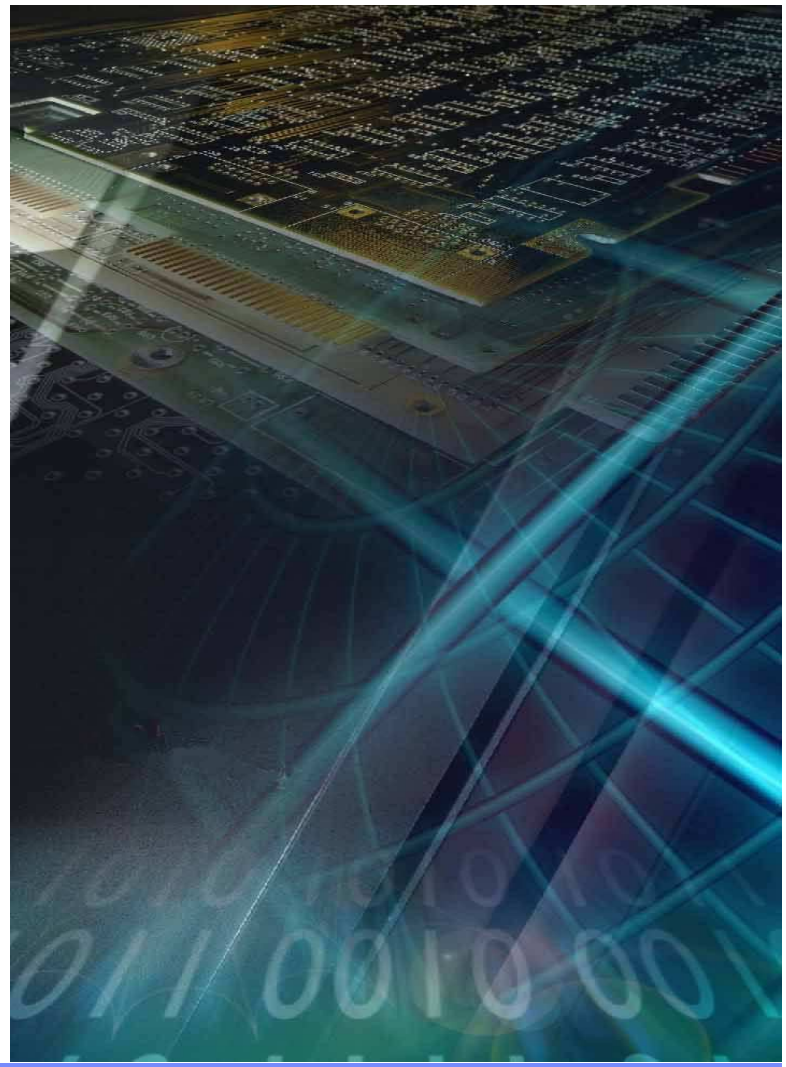
Timing-only Links



- Coupling links that allow two servers to be synchronized when a CF does not exist at either end of link
  - Typically required when synchronization needed not in a Parallel Sysplex configuration (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

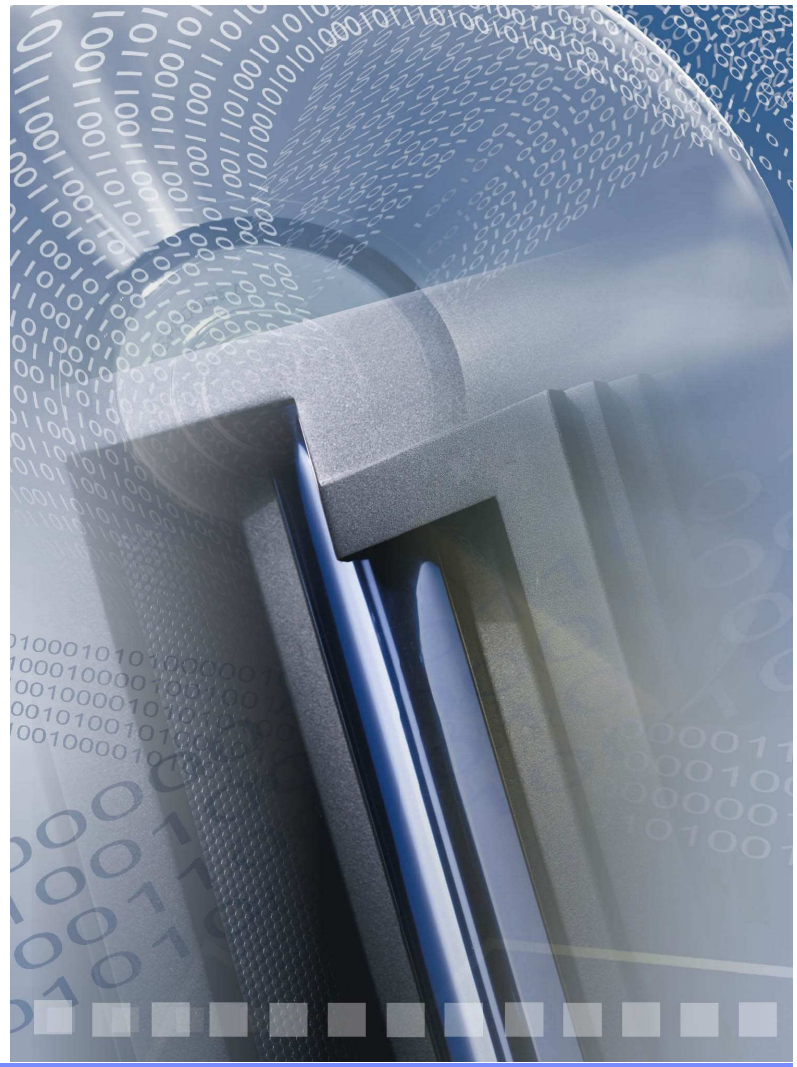
Prerequisites

- Hardware
  - 9037-002 concurrent LIC upgrade (if migrating from ETR network)
    - 9037 code changes to support STP Mixed CTN
  - System z9 server must be at EC Driver level 63J
    - Concurrently install all of the latest MCLs for this driver
    - STP prerequisite MCLs (LIC) will be installed
  - z990 and z890 must be at EC Driver level 55K
    - Concurrently install all of the latest MCLs for this driver
    - STP prerequisite MCLs (LIC) will be installed
  - HMC v2.9.1 (EC Driver level 64) or higher
    - Can upgrade z890/z990 HMC to new HMC code level
  - Concurrently install STP Enablement MCL (FC 1021)



Prerequisites

- Software
  - z/OS 1.7 or higher
  - Additional software maintenance required for z/OS 1.7, 1.8
    - Includes STP enablement APAR
    - Maintenance can be applied using “rolling IPL” process
  - Check Preventive Service Planning (PSP) buckets
    - Listed in the 2084DEVICE, 2086DEVICE, 2094DEVICE and 2096DEVICE PSP buckets for the z990, z890, z9 EC and z9 BC respectively
  - 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
  - Mixed CTN can include pre-1.7 systems
    - PTFs required for toleration code

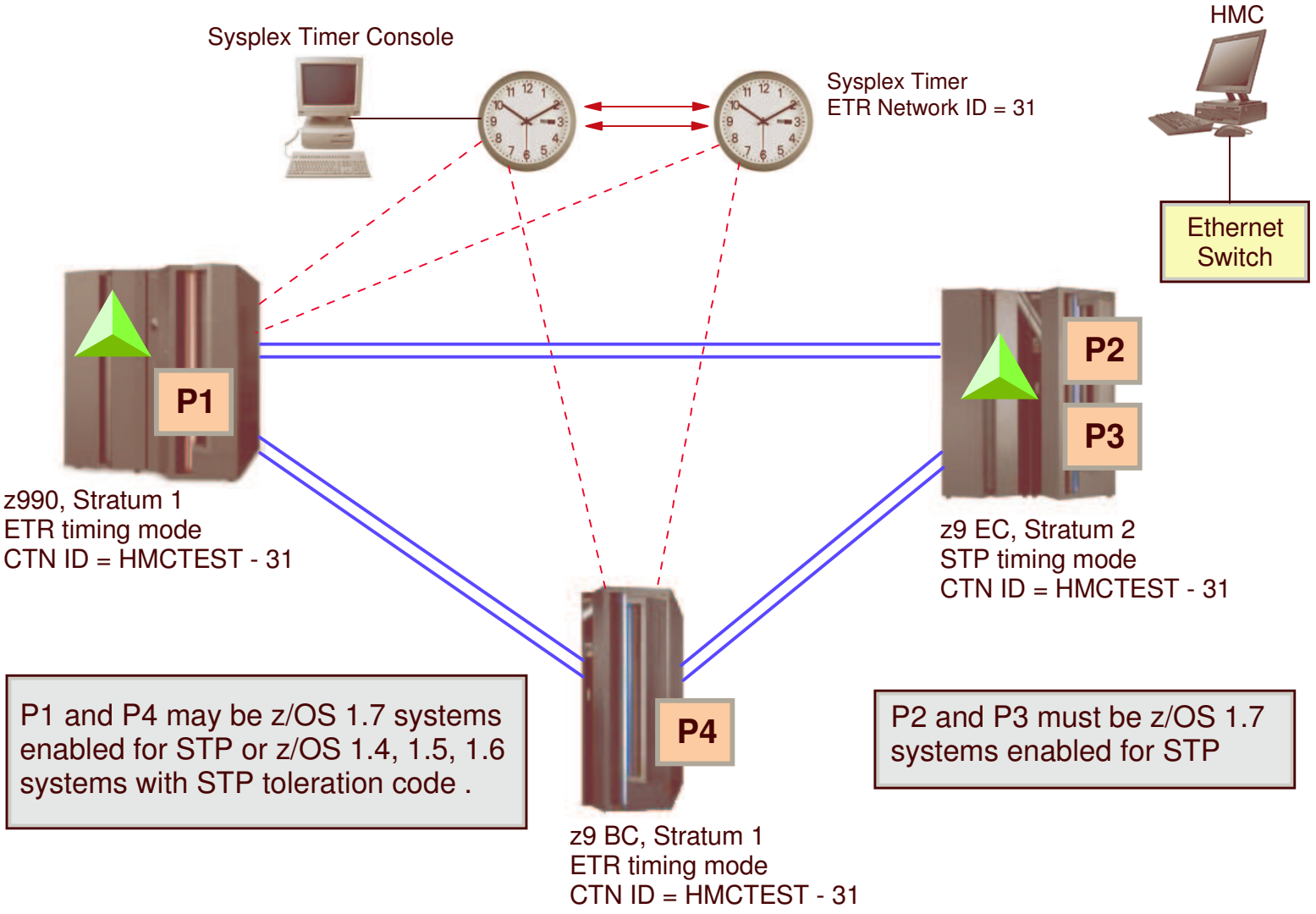


CLOCKxx statements

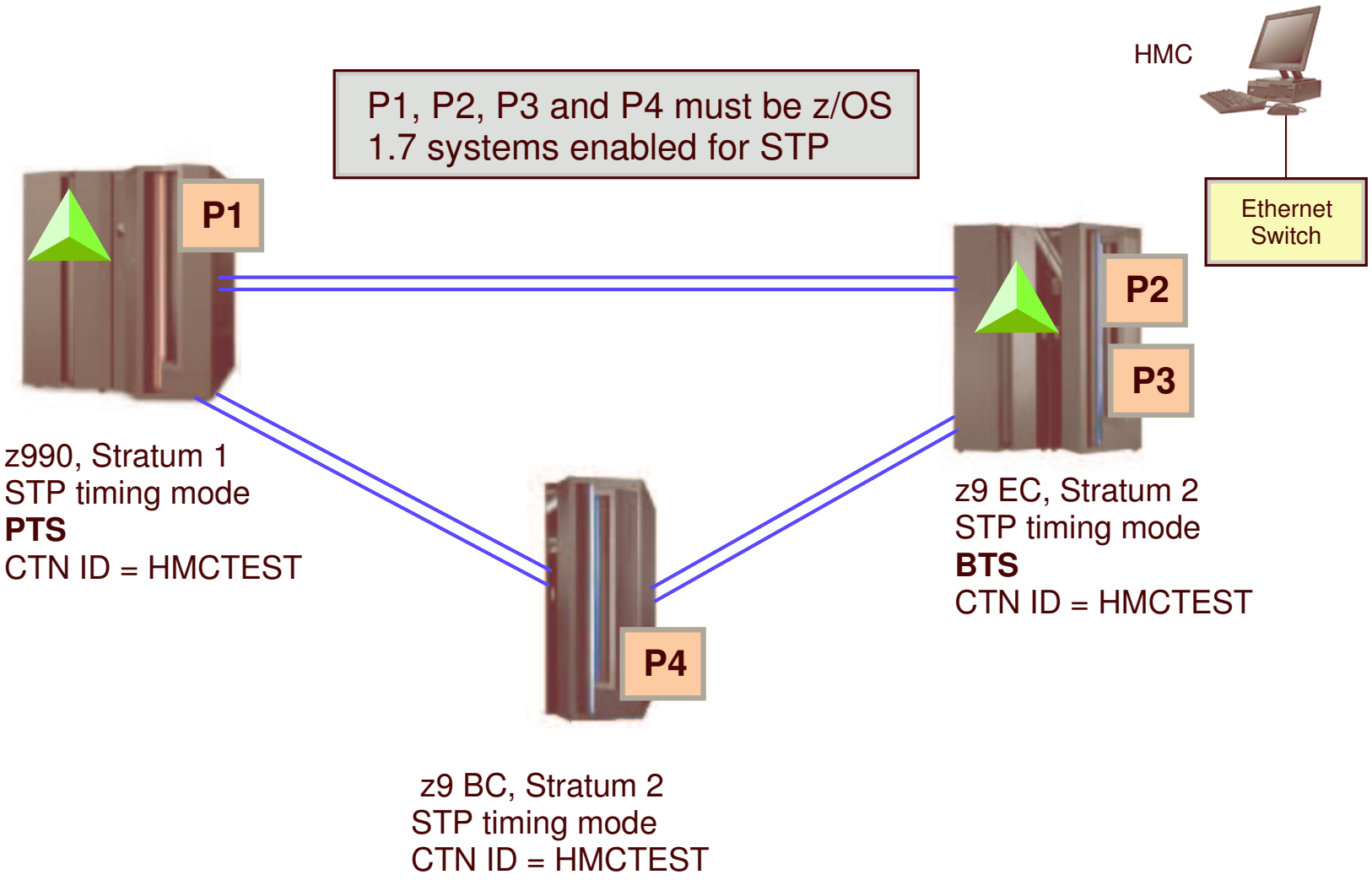
- OPERATOR PROMPT|NOPROMPT
- TIMEZONE W|E hh.mm.ss
- ETRMODE YES|NO
- ETRZONE YES|NO
- SIMETRID nn
  - nn = 0 – 31
- STPMODE* YES|NO
  - Specifies whether z/OS is using STP timing mode
  - STPMODE YES default
- STPZONE* YES|NO
  - Specifies whether the system is to get the time zone constant from STP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

* New statements for STP

Timing Modes in a Mixed CTN (example)



Timing Mode in an STP-only CTN (example)



Statements of Direction October 2006

- IBM intends to enhance the accuracy of initializing and maintaining Coordinated Server Time to an international time standard such as Coordinated Universal Time (UTC). The then current server is planned to have the capability of attaching to an external time source, such as a Global Positioning System (GPS) receiver.
- Network Time Protocol (NTP) client support: IBM intends to enhance the STP design to provide Network Time Protocol (NTP) client capability, so that Coordinated Server Time may be initialized and maintained to time provided by an NTP server. The purpose of this function is to allow the same time across an enterprise comprised of heterogeneous platforms.

* All statements regarding IBM's plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these Statements of General Direction is at the relying party's sole risk and will not create liability or obligation for IBM.

October 2007 Announcement:

NTP Client Support

- IBM is announcing the general availability of the NTP Client function for Server Time Protocol (STP). The STP design has been enhanced to include code for a Simple Network Time Protocol (SNTP) client on the Support Element (SE) of the IBM System z9 Enterprise Class (z9 EC) and IBM System z9 Business Class (z9 BC) servers. This support will allow customers to initialize the time of an STP-only Coordinated Timing Network to the time provided by a Network Time Protocol (NTP) server, and maintain time accuracy. This allows an enterprise comprised of heterogeneous platforms to track to the same time source.

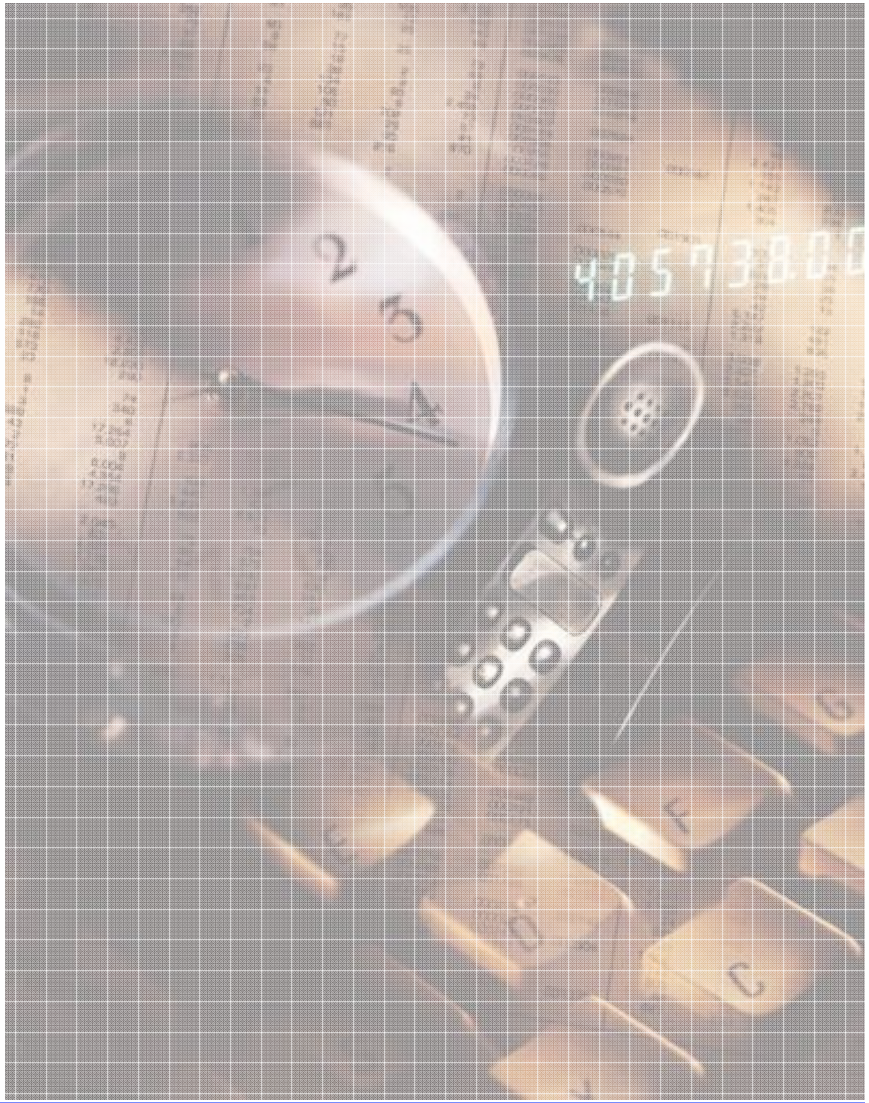
- NTP Client support will be generally available effective October 15, 2007 on System z9 EC or System z9 BC servers that are at **Driver 67L** and have the STP Feature 1021 installed. When the latest MCLs are installed, NTP client support is available on System z9 servers.

Sample implementation scenario

- Conduct Systems Assurance review for STP
- Create an implementation plan with dates and actions (sample Checklist will be provided by IBM)
- Order FC 1021
- Review Link and Sysplex Timer configuration
- Configure Mixed Mode CTN
- Run
- Configure STP Only CTN with external Time Source if desired (Dial-out or NTP Client)
- Run

Summary

- Server Time Protocol:
  - Allows Parallel Sysplex distances to extend beyond the current 40 km limit
    - Limits set by coupling protocol and links
  - Can help meet more stringent time synchronization requirements
  - **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**
  - Allows coexistence with ETR network



Additional Information

- Redbook™
  - Server Time Protocol Planning Guide SG24-7280
  - Server Time Protocol Implementation Guide SG24-7281
  - **Server Time Protocol NTP Client Support (REDP-4329)**
- 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) and to complete the Systems Assurance Confirmation Form via Resource Link

Questions ?