

Z31

Server Time Protocol

Noshir Dhondy (dhondy@us.ibm.com)



September 19 - 23, 2005

San Francisco, CA

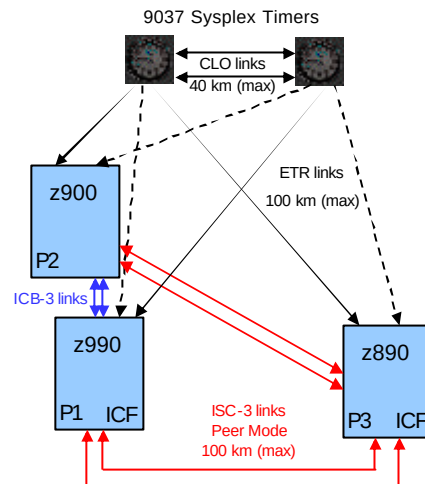
Agenda

- ETR Network Limitations
- Key future time synchronization requirements
- STP description
- Mixed Timing Network
 - Message Time Ordering Facility (MTOF) requirements with STP
- STP-only Timing Network
- Prerequisites
- Timing Modes
 - Important timing considerations for STP
- Implementation Assistance program (IAP)
- Summary

ETR Network Limitations

- Fiber distance between 9037 Sysplex Timers cannot exceed 40 km
 - Requires intermediate site for second timer if data centers more than 40 km apart
- “Best case” messaging times over ICB links in Parallel Sysplex (8 us approx) approaching “Worst case” TOD synchronization between CECs stepping to 9037s 40 km apart (4 us approx)
- Statement of Direction July 26, 2005 that IBM intends to withdraw 9037 Model 2 from marketing in 2006

— Active ETR link
 - - - Alternate ETR link



Key Future Time Synchronization Requirements

- Solution must
 - Provide improved time synchronization (compared to ETR) for zSeries servers in a Sysplex or non-Sysplex configuration
 - Scale with distance
 - Servers exchanging messages over fast short links require more stringent synchronization than servers exchanging messages over long distances
 - Scale with Server and Link technologies
 - For example: solution reused with appropriate changes when coupling link technologies change in future
 - Support a multi-site Sysplex of at least 100 km
 - Should not preclude going to longer distances in future
 - Provide at least all the functions provided by 9037 console functions
 - Allow concurrent migration from ETR network
 - Allow coexistence with ETR network
 - Be generally available before 9037 withdrawal from marketing

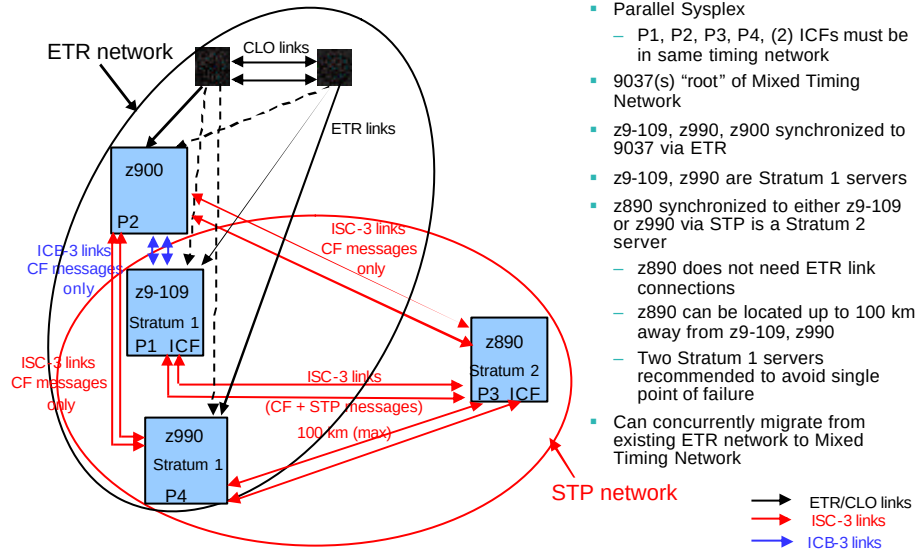
STP description

- Provides capability for multiple servers to maintain time synchronization with each other and form a Timing Network
- Server-wide facility implemented in IBM System z9™, z990, z890 Licensed Internal Code (LIC)
 - Single view of “time” to LPAR
 - LPAR can virtualize this view of time to the individual partitions (for example z/OS)
 - STP not available on z900, z800 or 9672 Gx servers
- Message based time synchronization protocol
 - Based on Network Time Protocol (NTP) – an industry standard
 - Time message information transmitted over all defined Coupling Facility Links
 - ISC-3 links (Peer mode), ICB-3 and ICB-4 links
 - **NOT standard NTP**
- Supports a multi-site Sysplex of at least 100 km (62 miles)
 - Does not preclude going to longer distances in the future
- Allows use of dial-out time services to set time to international time standard (UTC) to within +/- 100 ms as well as adjust to UTC on a periodic basis

Mixed Timing Network

- Allows co-existence of ETR and STP networks
- 9037s are the “root” for both the ETR and the STP networks
- Need at least one server in Mixed Timing Network capable of:
 - Time coordination using ETR (attach to 9037) and time coordination using STP message-based protocol
- Any STP capable server (z9, z990, z890) attached to 9037 can be a Primary Time Server (Stratum 1) for the STP network
 - Secondary Time Servers can be Stratum 2 or 3
 - Stratum 2 uses Stratum 1 as clock source
 - Stratum 3 uses Stratum 2 as clock source
- 9037 console continues to be used for all timing related functions of the Mixed Timing Network
- Hardware Management Console (HMC) must be used for Mixed Timing Network ID initialization and modification

Mixed Timing Network

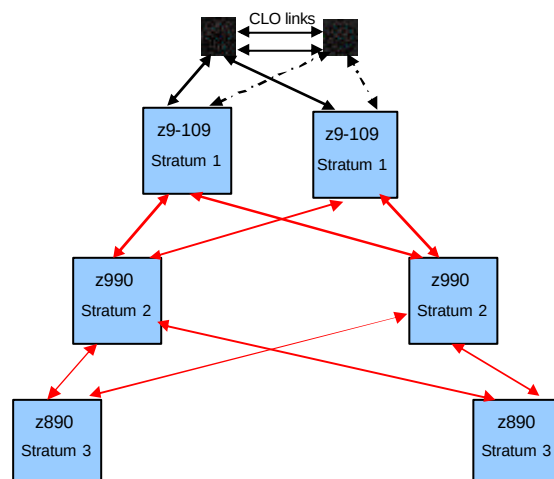


- Parallel Sysplex
 - P1, P2, P3, P4, (2) ICFs must be in same timing network
- 9037(s) "root" of Mixed Timing Network
- z9-109, z990, z900 synchronized to 9037 via ETR
- z9-109, z990 are Stratum 1 servers
- z890 synchronized to either z9-109 or z990 via STP is a Stratum 2 server
 - z890 does not need ETR link connections
 - z890 can be located up to 100 km away from z9-109, z990
 - Two Stratum 1 servers recommended to avoid single point of failure
- Can concurrently migrate from existing ETR network to Mixed Timing Network

7

© 2005 IBM Corporation

Mixed Timing Network example (Stratum 1-3)



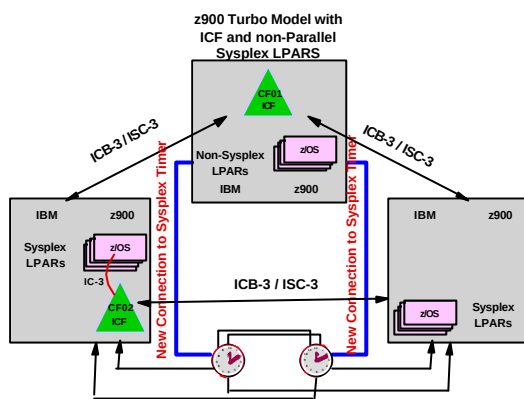
Maximum Stratum level for STP is 3

NOTE: Example may not be a typical Parallel Sysplex configuration

8

© 2005 IBM Corporation

Message Time Ordering Facility (MTOF) (also known as CF Request Time Ordering) announced Apr 30th 2002)



- Time stamps included in messages transferred between z/OS and CF
- Time stamps compared to local TOD
 - If local TOD is behind time stamp received, message processing is delayed, until local TOD has caught up to time stamp value
- CF configured as ICF on any z900 (2C1 thru 216 models) or z990 requires connectivity to 9037
 - If ICF on same server as Sysplex member no additional connectivity
 - If ICF on server with non-Sysplex LPARs, additional connectivity required

MTOF requirements in a Parallel Sysplex **without STP**

CF → z/OS ↓	G5 or G6	z800 or z900 (except 2C1-2C16)	z900 (2C1-2C16)	z9-109 or z990 or z890
G5 OR G6	MTOF Not-required Not-capable	MTOF Not-required Not-capable	MTOF Not-required Not-capable	MTOF Not-required Not-capable
z800 or z900 (all models)	MTOF Not-required Not-capable	MTOF Not-required Capable	MTOF Required Capable	MTOF Required Capable
z9-109 or z990 or z890	MTOF Not-required Not-capable	MTOF Not-required Capable	MTOF Required Capable	MTOF Required Capable

IBM Systems and Technology Group

IBM

Mixed Timing Network – Supported “Legacy” servers for Parallel Sysplex

- Prerequisite for STP in Parallel Sysplex environment is Message Time Ordering Facility (MTOF)
- STP capable servers are MTOF capable
- In a Mixed Timing Network, “Legacy” server or standalone CF MUST be MTOF capable and attached to 9037s
 - z900 or z800
- For example:
 - P1, P2, P3, P4 members of sysplex
 - P1, P2, P3 can only connect to CFs on “legacy” servers with MTOF capability
 - z900 Model 100 CF is allowed as long as it has ETR connections
 - Similarly, ICF on z890 can only connect to P4 on “legacy” server with MTOF capability
 - P4 on z900 allowed

→ ETR/CLO links

→ ISC-3 links

→ ICB-3 links

11

© 2005 IBM Corporation

IBM Systems and Technology Group

IBM

MTOF requirements in a Parallel Sysplex with STP

CF z/OS → ↓	G5 or G6 (Server NOT STP capable)	z800 or z900 (Server NOT STP capable)	z900 (2C1-2C16) (Server NOT STP capable)	z9-109, z990, z890 (STP capable)
G5 OR G6 (Server NOT STP capable)	MTOF Not-required Not-capable	MTOF Not-required Not-capable	MTOF Not-required Not-capable	MTOF Required Not-capable NOT ALLOWED
z800 or z900 (Server NOT STP capable)	MTOF Not-required Not-capable	MTOF Not-required Capable	MTOF Not-required Capable	MTOF Required Capable
z9-109, z990, z890 (STP capable)	MTOF Required Not-capable NOT ALLOWED	MTOF Required Capable	MTOF Required Capable	MTOF Required Capable

Can participate in a Mixed Timing Network

12

© 2005 IBM Corporation

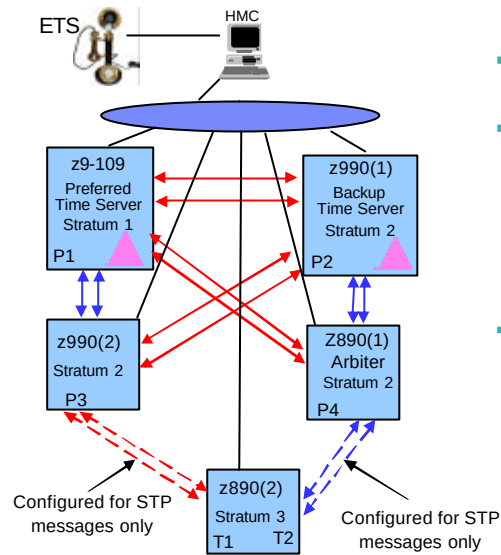
STP-only Timing Network

- All servers in STP-only Timing Network have to be STP capable
- 9037s no longer the “root” of the timing network as in the Mixed Timing Network
- Configuration has to be defined
 - Preferred Time Server (Stratum 1)
 - Either the Preferred or Backup Time Servers can be the Stratum 1 of the STP-only Timing Network
 - Backup Time Server can take over as the Stratum 1 under unplanned or planned exceptions, without disrupting synchronization capability of STP-only Timing Network
 - Arbiter (optional)
 - Provides additional means to determine if Backup Time Server should take over as the Stratum 1

STP-only Timing Network (continued)

- HMC must be used to provide the following functions:
 - Set time manually
 - Includes Time Zone Offset, Daylight Savings Time Offset, Leap Second Offset
 - Set and maintain time to an international time standard (UTC)
 - External Time Source (ETS) capability
 - Dial-out from HMC to set time to international time standard (UTC) to within +/- 100 ms as well as adjust to UTC on a periodic basis
 - Schedule and change Offsets (Daylight Savings, Leap seconds)
 - Automatic scheduling of Daylight Savings Time based on algorithm
 - Adjust time by up to +/- 60 seconds (currently 9037 allows 4.999 seconds)
 - Define, modify, view the STP-only Timing Network configuration

STP-only Timing Network example



- Parallel Sysplex
 - P1, P2, P3, P4, (2) ICFs
- If Servers configured with ICFs selected as Preferred and Backup Time Servers, connectivity in place for CF messaging can be used for STP message communication
 - Any STP capable server can be selected as Preferred or Backup
- If Timing Network scope is greater than Sysplex scope, additional ISC or ICB links may need to be configured for STP
 - T1, T2 on z890(2) not in P1-P4 Parallel Sysplex but in same Timing Network as P1-P4

→ ISC-3 links
→ ICB-3 links

15

© 2005 IBM Corporation

Prerequisites

- Hardware
 - z9-109, z990 or z890 server
 - Concurrent MCLs on Driver 55 (z990, z890) to install STP Licensed Internal Code (LIC)
 - Concurrent MCL to enable STP
 - z9-109 HMC
 - New HMC
 - Upgrade of existing HMC
- Software
 - z/OS 1.7 or higher
 - STP code in z/OS 1.7 (default=disabled)
 - PTFs for STP support
 - PTF to enable STP
 - Toleration PTFs for z/OS 1.4, 1.5, 1.6 in Mixed Timing Network

16

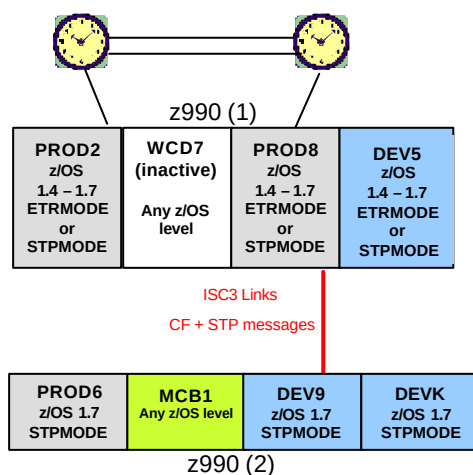
© 2005 IBM Corporation

CLOCKxx statements

- OPERATOR PROMPT|NOPROMPT
- TIMEZONE W|E hh.mm.ss
- ETRMODE YES|NO
- ETRZONE YES|NO
- SIMETRID nn
- 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
 - Where:
 - nn = 0 - 31
 - ss = 0 - 99 seconds

* New statements for STP

Timing Modes in a Mixed Timing Network (example)



- Prod2, Prod6, Prod8 (Prod Sysplex)
- Dev5, Dev9, DevK (Dev Sysplex)
- z990 (1) server in ETR timing mode
 - Synchronized to 9037(s)
- z990 (2) server in STP timing mode
 - Synchronized to z990(1) using STP
- Prod2, Prod 8, Dev5 can be at z/OS 1.4, 1.5, 1.6 or 1.7
 - If z/OS pre-1.7
 - Timing Mode can be ETRMODE only
 - Need toleration PTF
 - If z/OS 1.7
 - Timing Mode can be STPMODE or ETRMODE
- Prod6, Dev9, DevK MUST be at z/OS 1.7 (w/STP enablement PTF)
 - MUST have STPMODE YES defined

Timing mode in a STP-only Timing Network (example)

z990 (1) – Preferred Time Server (Stratum 1)

PROD2 z/OS 1.7 STPMODE	WCD7 (inactive) Any z/OS level	PROD8 z/OS 1.7 STPMODE	DEV5 z/OS 1.7 STPMODE
-------------------------------------	--	-------------------------------------	------------------------------------

ISC3 Links
CF + STP messages

PROD6 z/OS 1.7 STPMODE	MCB1 Any z/OS level	DEV9 z/OS 1.7 STPMODE	DEVK z/OS 1.7 STPMODE
-------------------------------------	-------------------------------	------------------------------------	------------------------------------

z990 (2) – Backup Time Server (Stratum 2)

- Prod2, Prod6, Prod8 (Prod Sysplex)
- Dev5, Dev9, DevK (Dev Sysplex)
- z990 (1 & 2) servers in STP timing mode
 - z990(2) synchronized to z990(1) using STP
- Prod2,6,8 and Dev5,9,K MUST be at z/OS 1.7 (w/STP enablement PTF)
 - MUST have STPMODE YES defined

Important Timing Considerations

- Within a Mixed Timing Network:
 - “Server timing mode” may be different than the “Image timing mode” of a z/OS image running in a Parallel Sysplex on the same server.
 - For example,
 - “Server timing mode” ETRMODE
 - “Image timing mode” STPMODE
 - Servers with different “Server timing modes” may exist.
 - For example,
 - Servers synchronized to 9037 in ETRMODE
 - Servers synchronized using STP in STPMODE
 - z/OS images within a Sysplex may have different “Image timing modes”.
- Within a STP-only Timing Network:
 - “Server Timing Mode” of all servers is STPMODE
 - “Server timing mode” and the “Image timing mode” of z/OS images running in a Parallel Sysplex on the same server are the same
 - STPMODE.

Implementation Assistance Program (IAP)

- Program to accelerate customer adoption of STP with IBM's assistance.
- You will receive:
 - technical planning assistance
 - review of migration plans to an STP network
 - technical support during the implementation of STP
- Potential candidates for IAP
 - Require a multi-site sysplex across more than 40 km
NOTE: using STP at a distance greater than 40km is not a requirement for IAP.
 - New Parallel Sysplex implementation – single or multi site
Note: a SOD was included in the z9-109 announcement that IBM plans to withdraw the 9037-002 from marketing in 2006
 - Latest Parallel Sysplex technology
- Target date to start IAP on z990 and z890 is 10/05, on z9-109 is 1Q06
- Contact Angelo Corridori (angeloc@us.ibm.com) if interested or have questions

Summary

- Server Time Protocol will:
 - Meet future more stringent time synchronization requirements
 - Scale with technology as processors and messaging technology improve
 - Not require dedicated Timer links
 - Uses same hardware and protocols as data
 - Allow GDPS distances to extend beyond the current 40 Km limit
 - Limits set by coupling protocol and links
 - Allow concurrent migration from an ETR network
 - Allow coexistence with ETR network
 - Allow time to be set to a dialup service to within 100 ms
 - STP plans to improve this capability in 2006