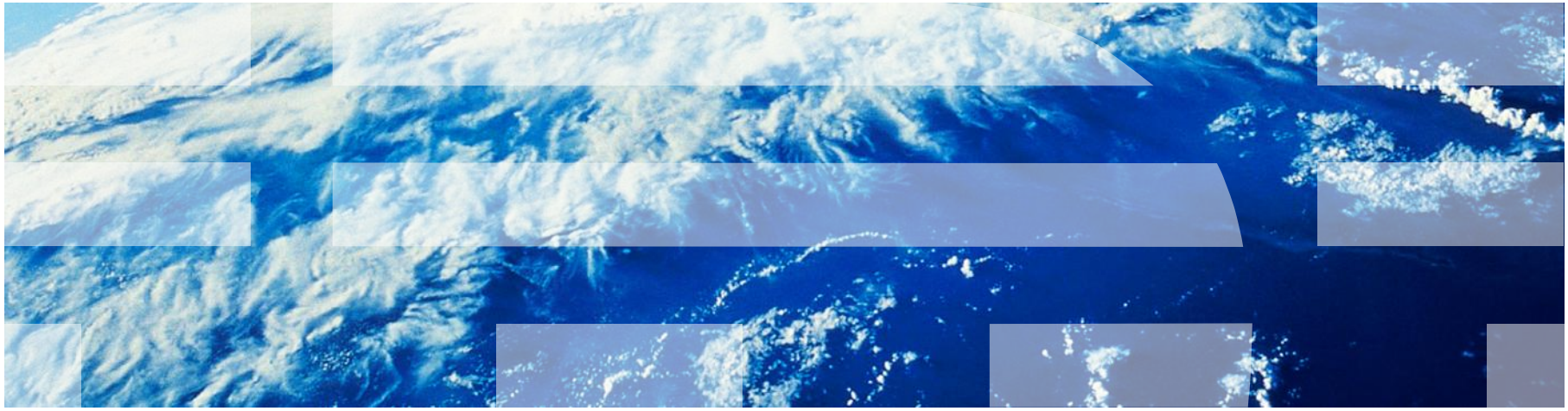
Nicole M. Fagen – Parallel Sysplex Support Team Lead
nfagen@us.ibm.com
February 28, 2011

Session: 8921
Coupling Facility Non-disruptive Serialized Dump



© 2011 IBM Corporation

Trademarks

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

IBM*	FICON*	System x*
IBM (logo)*	IMS	System z*
ibm.com*	Parallel Sysplex®	System z9®
AIX*	POWER7	System z10
BladeCenter*	ProtecTIER*	Tivoli*
DataPower*	RACF*	WebSphere*
CICS*	Rational*	XIV*
DB2*	Redbooks®	zEnterprise
DS4000*	Sysplex Timer®	z/OS*
ESCON®	System Storage	z/VM*
* Registered trademarks of IBM Corporation		z9®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license there from.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
InfiniBand is a trademark and service mark of the InfiniBand Trade Association.
Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* 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.

© 2011 IBM Corporation

Agenda

- Coupling Facilities
  - Structures
  - Processing Requests
    - Simplex
    - Duplex
- Problem Scenarios
  - Performance
  - Loss of Connectivity
  - CF Terminates
  - Breakduplex
- Street Light Analogy
- Testing Recommendation
- z196 “Non-disruptive” Serialized Dumps
- Additional Testing Recommendations

© 2011 IBM Corporation

Coupling Facility

- Piece of computer hardware which allows multiple processors to access the same data
- Firmware running Coupling Facility Control Code (CFCC)
- Memory
- Accessed by z/OS using CF links
- Required for parallel sysplex
- No I/O devices
- All data resides in memory
- No applications run on the coupling facility

- Any IBM mainframe can serve as a coupling facility
  - Internal CF – resides on a CEC with a z/OS system using the CF
  - External CF – resides on a CEC that does not have z/OS systems on it, or at least not z/OS systems from the sysplex using the CF

© 2011 IBM Corporation

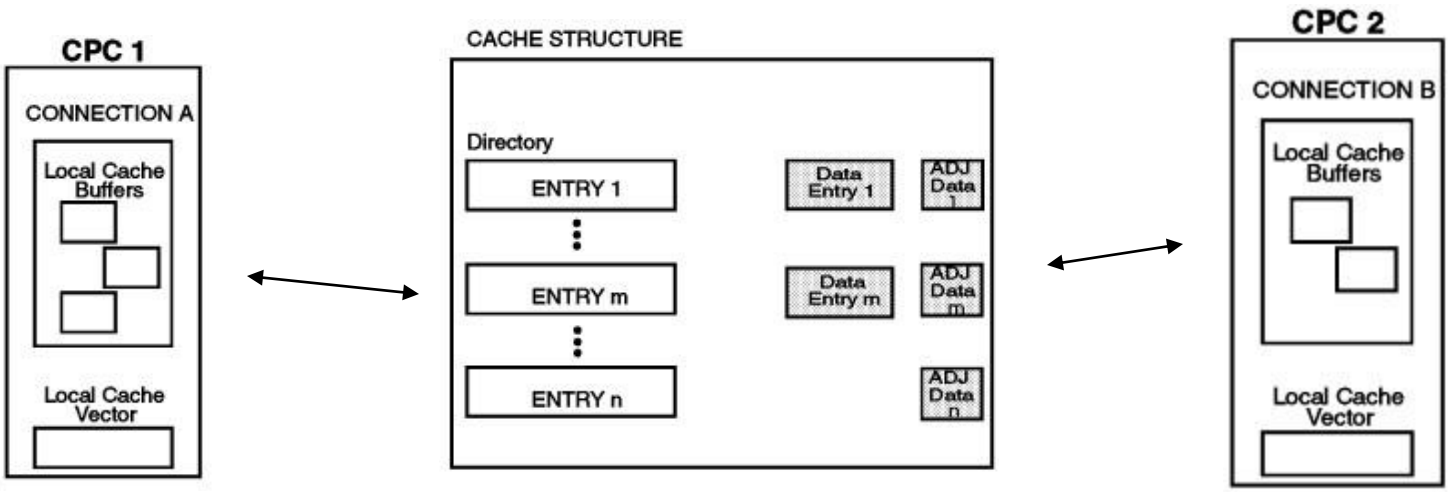
Coupling Facility Structures

- Defined in the CFRM policy
- Types of structures
  - List
  - Lock
  - Cache
  - Serialized List
- z/OS application requests to connect to the structure
- XES initiates request to the CF to allocate the structure
- CF allocates the structures and notifies XES the structure is available
- XES notifies the z/OS application (connector) that it has access to the structure
- All connectors to a particular structure can directly access the same data while maintaining data integrity and data consistency throughout the sysplex while using the functions provided by the CF to maintain data integrity and data consistency.
- Connectors may make various uses of the data
  - Read
  - Write
  - Delete
  - Move
  - Lock
  - Unlock
  - Etc
- z/OS may dump the user contents of the structure via STRLIST option for SLIP or DUMP

© 2011 IBM Corporation

Structure Example - Cache Structure

- Provides connections with data consistency and high-speed access to data
- A mechanism called "cross-invalidate" informs connections of changes to data.
- Connections test the local buffers for validity using IXLVECTR macro
- Accessing data stored in the local cache buffer is the quickest way for a user to access the shared data.
- Use the IXLCACHE macro to access the data.



© 2011 IBM Corporation

z/OS | CF Interaction - Simplex Structures

- Application must connect to a structure
  - IXLCONN
- Application issues a XES macro to place data in the structure, read data from the structure, delete data from the structure
  - IXLLIST
  - IXLLOCK
  - IXLCACHE
  - Etc.
- XES
  - MCB – message command block
- MCB flows across the CF link to the CF
- CF receives the MCB
- CF process the MCB (black box)
- CF responds to XES with answer
  - MRB – message response block
- XES notifies the connector of the response

© 2011 IBM Corporation

z/OS | CF Interaction – System Managed Duplex Structures

- Application issues a XES macro place data in the structure, read data from the structure, delete data from the structure
- Request A goes to CF 1, Request B goes to CF2
- CF1 and CF2 receive the requests
- CF1 tells CF2, “I’m ready to execute!”
- CF2 tells CF1, “I’m ready to execute!”
- CF1 and CF2 execute the request
- IF DUPLEXCF16 is DISABLED
  - CF1 tells CF2, “I’m complete”
  - CF2 tells CF1, “I’m complete”
- If DUPLEXCF16 is ENABLED
  - Exchange completion signals is asynchronous
  - Control returns to z/OS faster than DUPLEXCF16
- CF1 and CF2 send responses to XES
- XES processes the responses
- XES notifies the connectors of the response

© 2011 IBM Corporation

A Peek Inside the Coupling Facility

- At the core, the CF is similar in nature to z/OS
  - Storage foot print
  - Control blocks
  - Tasks to process requests from z/OS
  - Latches to serialize work
  - Queues with elements representing work to do

- As with any technology …

Sometimes there are problems …

© 2011 IBM Corporation

Performance

- Performance
  - Sync service times
  - Async service times
  - CF % busy
  - Volume of requests converted to async
  - Subchannel busy
  - Path busy

© 2011 IBM Corporation

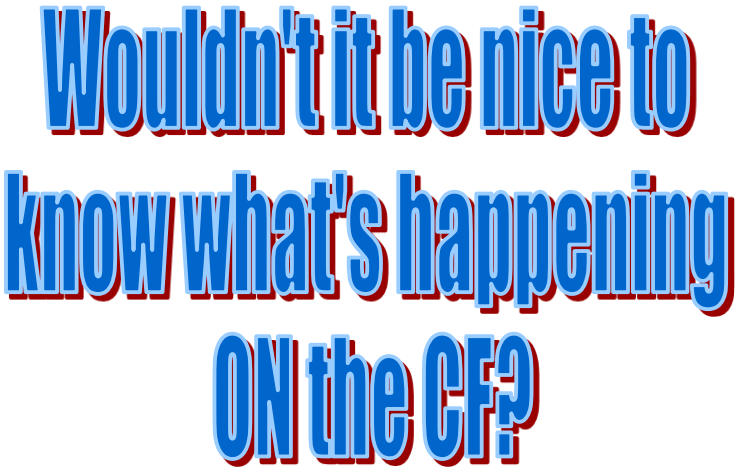
Potential Performance Impact(s)

- Impact varies
  - No noticeable impact
  - End users delayed
  - End user timeouts on transactions
  - Overall drag on the sysplex

© 2011 IBM Corporation

Performance Problem Determination

- Investigate workload changes
- Review CF Activity Report
  - Initial focus on structure related to impact
  - Identify top structures consuming resources
  - Compare utilization to “good” time
  - Consider
    - Increase in user requests?
    - Change in # CPs on CF?
    - Observe service times
- Review Partition Data Reports
  - Identify any capacity issues
- Review operlog
  - Any activity on the CF
    - Connects
    - Disconnects
    - Rebuilds
    - Alter processing
  - Application error messages
- z/OS Dump of XCFAS and connectors address space
  - SYSXES ctrace active
  - SDATA option XESDATA



© 2011 IBM Corporation

Loss of Connectivity

- What you'll see …

One for each link to the CF

IXL158I PATH *chpid* IS NOW NOT-OPERATIONAL TO CUID: *cuid* COUPLING FACILITY 002817.IBM.yy.0000000xxxxxx PARTITION: *partition side* CPCID: *cpcid

IXC518I SYSTEM IT2T NOT USING

COUPLING FACILITY 002817.IBM.yy.0000000xxxxxx

PARTITION: pp CPCID: 00

System (XCF) indicates it is no longer using the CF

NAMED cf_name

REASON: CONNECTIVITY LOST.

REASON FLAG: 13300005.

Possibly zOS issued messages indicating IFCCs

IXL044I COUPLING FACILITY cf_name HAS EXPERIENCED

INTERFACE CONTROL CHECKS ON CHPID zz DURING THE LAST 125 SECONDS

© 2011 IBM Corporation

Loss of Connectivity

- zOS will drop a link to a coupling facility if it has not received a response for any of the outstanding requests to the CF on that link in 2.5 seconds.
- Losing one link is not necessarily a problem
- Breadth of loss of connectivity
  - One system could lose connectivity
  - X of Y systems could lose connectivity
  - All systems on one CEC could lose connectivity
  - All system in the sysplex could lose connectivity
- Losing all the links at the same time from one z/OS system results in the z/OS system ceasing to use the coupling facility and initiation of sysplex recovery actions to rebuild or fail over to structures in another CF image

© 2011 IBM Corporation

Potential Loss of Connectivity Impact(s)

- Breakduplex
  - Structure transitions from duplex mode to simplex mode
  - Instance of the structure in the CF which did not lose connectivity is maintained
- For simplex structures, connectors will be notified of the loss of connectivity
- Structures may start to rebuild to another coupling facility with connectivity
- End user applications may experience delays
- If connectivity is re-established immediately the impact may be minimal
  - If the CF did not terminate the prior instances of the structures will still be available
  - Applications can reconnect to prior instances
- z/OS resources consumed to process all of the recovery by connectors
  - Recovery Time Varies
    - Quantity of rebuilds
    - Size of structures
    - Connector rebuild protocols
    - CFRM MSGBASED

© 2011 IBM Corporation

Loss of Connectivity - Problem Determination

- HW review CF logs
- Review CF Activity Report
  - If the problem situation was “building” the CF Activity Report may provide some insight
  - If the problem happened all within one recording interval, much of the data will be lost
  - Identify top structures consuming resources
  - Compare utilization to “good” time
  - Consider
    - Increase in user requests?
    - Change in # CPs on CF?
    - Investigate workload changes
- Review operlog
  - Locate IXL158I and IXC518I
  - Any activity on the CF
  - Connects, disconnects, rebuilds, alter processing
- Review any z/OS dumps taken around the time of the loss of connectivity
- Review logrec
  - IFCCs

© 2011 IBM Corporation

CF Terminates (Abort)

- What you'll see …

- From z/OS, the CF terminating looks like a loss of connectivity on all systems at the same time

Incredibly Rare Situation
Mentioned Here for Completeness

© 2011 IBM Corporation

Potential CF Termination Impact

- Loss of connectivity to CF
- All instances of structures in the failed CF lost
  - Breakduplex
  - Rebuild structures
- CF terminates and stays down
  - There must be enough space in the other CFs to allow all deleted structures to rebuild
- CF may “bounce” right back
  - Rebuilds
    - Into another CF
    - Into the CF which bounced (if it was fast enough)
- z/OS resources consumed to process all of the recovery by connectors
  - Recovery Time Varies
    - Quantity of rebuilds
    - Size of structures
    - Connector rebuild protocols
    - CFRM MSGBASED

© 2011 IBM Corporation

CF Terminates (Abort) Problem Determination

- CF disruptive dump
  - Analogous to z/OS standalone dump
  - Automatically taken by HW
- Extremely high degree of success at identifying root cause

Clear picture of what was happening ON the CF is collected.

© 2011 IBM Corporation

Breakduplex

- What you'll see …

IXC522I SYSTEM-MANAGED DUPLEXING REBUILD FOR STRUCTURE

Structure_name IS BEING STOPPED

TO SWITCH TO THE NEW STRUCTURE DUE TO

FAILURE OF A DUPLEXED REQUEST

SYSTEM CODE: 09260310

© 2011 IBM Corporation

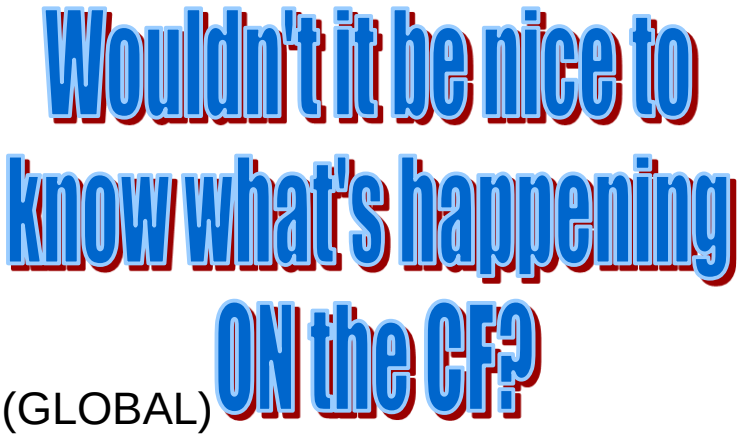
Potential Breakduplex Impact

- Structure transitions to simplex mode
- Application may experience a delay while duplexing is being disabled
- Duplexing may be automatically restarted depending on configuration if DUPLEX(ENABLED)
- Application may experience a delay while duplexing is being established

© 2011 IBM Corporation

Breakduplex Problem Determination

- Review CF Activity Report
  - Observe the structures with the greatest workload
    - # Requests
    - % CF Utilization



- Recreate needed?
  - Turn on SYSXES ctrace on all systems
    - GLOBAL

      TRACE CT,16M,COMP=SYSXES,SUB=(GLOBAL)

      R id,OPTIONS=(REQUEST,HWLAYER,LOCKMGR),END
    - Connectors

      TRACE CT,16M,COMP=SYSXES,SUB=(STRUCTURENAME)

      Note: Add single quotes around the structure name if there is a special character

      R id,OPTIONS=(REQUEST,HWLAYER,LOCKMGR),END

  - Turn on IXLBREAKDUPLEX Trap
    - TRAPS NAME(IXLBREAKDUPLEX)
    - SET DIAG=xx

© 2011 IBM Corporation

Street Light

- CF views prior to z196
  - Un-serialized dumps could be initiated from the SE
    - Un-serialized picture does not allow for accurate problem determination
  - Disruptive Serialized dump of a CF
    - CF terminates
    - CF dumps
    - CF must be reactivated
    - The impact of documentation collection is too high
      - For break duplex, disruptive serialized dumps on both CFs at the same time are needed
      - Cannot collect such documentation in the field

© 2011 IBM Corporation

Street Light

- Look at z/OS documentation to identify the pattern of requests being initiated
- Look at the z/OS documentation to see the responses from the CF when z/OS processes them
- Look at z/OS to see the performance of the CF
- Recreate road is very long
  - Bring workload “in-house” or try to simulate pattern
  - TMCF to collect two hard dumps
- Really want to look at what is happening on the CF at the time of the issue!!
- Move the “street light” to where the error is

© 2011 IBM Corporation

z196 Non-disruptive Serialized CF Dump

- New “non-disruptive” serialized CF dumps can be initiated by firmware or by z/OS

- Benefits
  - Dump contains a snapshot/capture of critical CF storage areas into a "capture" area set aside for that purpose
  - Enables CFCC development to see the requests being processed at the time of the command on the CF timed out

- Situations where non-disruptive serialized dump automatically initiated by firmware
  - Locally initiated by the CFCC in response to a set of triggering events
    - CF command timeout
    - Breakduplex
  - Remotely initiated by coupling link firmware in response to a set of triggering events
    - CF link protocol issues
      - Coordinated with CF link diagnostic logouts

© 2011 IBM Corporation

Breakduplex - CF Non-disruptive Serialized Dump

- What you'll see …

IXC522I SYSTEM-MANAGED DUPLEXING REBUILD FOR STRUCTURE

Structure_name IS BEING STOPPED

TO SWITCH TO THE NEW STRUCTURE DUE TO

FAILURE OF A DUPLEXED REQUEST

SYSTEM CODE: 09260310

***IXL010E NOTIFICATION RECEIVED FROM**
COUPLING FACILITY 002817.IBM.02.00000000xxxx
PARTITION: 1F CPCID: 00
NAMED CF02
A NON-DISRUPTIVE DUMP WAS TAKEN BY THE CF.

© 2011 IBM Corporation

More Triggers for z196 Non-disruptive Serialized CF Dump

- New Function APAR, **OA31387**, enables z/OS to initiate non-disruptive serialized CF dumps on z196 (z/OS 1.10 & z/OS 1.11)
- Support provided in base of z/OS 1.12

- With the PTFs applied, z/OS will automatically initiate non-disruptive serialized dumps in the following situations
  - By default (automatically)
    - Inconsistent lock data detected, aka, missing record data entry
  - DUPLEXCFDIAG function ENABLED
    - Link timings exceed 200 milli seconds
  - IXLDUPOUTOFSYNCH DIAG set
    - Duplex out of sync condition detected

© 2011 IBM Corporation

z/OS Initiated “Non-Disruptive” Serialized CF Dumps

- IXL051E issued on z/OS
  - Missing record data
  - Duplex out of Sync

IXL051E NON-DISRUPTIVE DUMP OF COUPLING FACILITY CF4

002187.IBM.02.0000000xxxxx PARTITION: 14 CPCID: 00

INITIATED BY THE SYSTEM FOR STRUCTURE MQGPCSQ_ADMIN SID 0077

DIAG DATA: DUMPCF invoked CF dump by STRNME

- IXL010E issued after the dump has been taken

© 2011 IBM Corporation

z/OS Documentation to Accompany CF Non-disruptive Serialized Dump

- ABEND026 RSN 07070001 taken on all systems with connections to the CF which took the non-disruptive serialized CF dump for the following situations
  - Break duplex
  - Link diagnostic time out
  - CF command timeout
  - Note: CF dumping can only trigger z/OS to take the dumps when the z/OS images resides on z196. z/OS initiates the software dumps when a new CRW on the z196 is received. The CRW is not provided to z/OS images residing on z10.
- If z/OS initiates the CF dump, z/OS will also ABEND026 on all systems with connections to the structure
  - RSN 09260002, 09260003, or 09260009 for duplex out of synch
  - RSN 0C090020, 0C2A0020, or 0C5B0020 for record data entry not found

© 2011 IBM Corporation

"Non-disruptive" Serialized CF Dump Experience

- Timings ..
  - Approximately 40-60 msec to actually take the dump
    - CF activity quiesced
  - Approximately 10 seconds to write the dump to SE hard drive
    - CF activity not quiesced
  - IXL010E issued
  - IXL010E must be manually deleted if AMRF is active
    - K C,CE,id
- Impact …
  - CF is quiesced while the dump is being captured
    - On the order of a 40-60 msec delay to the CF
    - End users do not generally notice 40-60 msecs
  - Note: If the dumps are taken for breakduplex then at least one operation exceeded 300 msec. The 40-60 microseconds to take the dump did not cause the breakduplex.
- Time between dumps …
  - Only one non-disruptive serialized dump per CF within 5 minute interval
  - Additional requests rejected

© 2011 IBM Corporation

CF Storage

- CFCC uses more CF storage on z196 then z10 due to new function in z196
  - Non disruptive serialized CF dump
  - Added code
  - Increased the numbers the link buffers(64 chpids-->128)
  - Increased number of SIDS(1024-->2048).
  - Increased number of retry buffers and had support for >7sch's.
- Likely not an issue in a production environment
- May need further consideration in a test environment

© 2011 IBM Corporation

What to do if you see a Non-disruptive Serialized CF Dump

- Report a HW Problem.
  - Today, the z196 does not phone home when a “non-disruptive” CF dump is taken.

- HMC
  - Choose the Report a Problem ICON which is under Service Tasks
  - Click HMC problem
  - Add a brief description

- Support Element
  - Choose the Report a Problem ICON which is under Service Tasks
  - Choose the area of the issue being reported
  - Add a brief description

- Report any follow on software issues to the software support center

© 2011 IBM Corporation

Test Environment Testing Recommendations

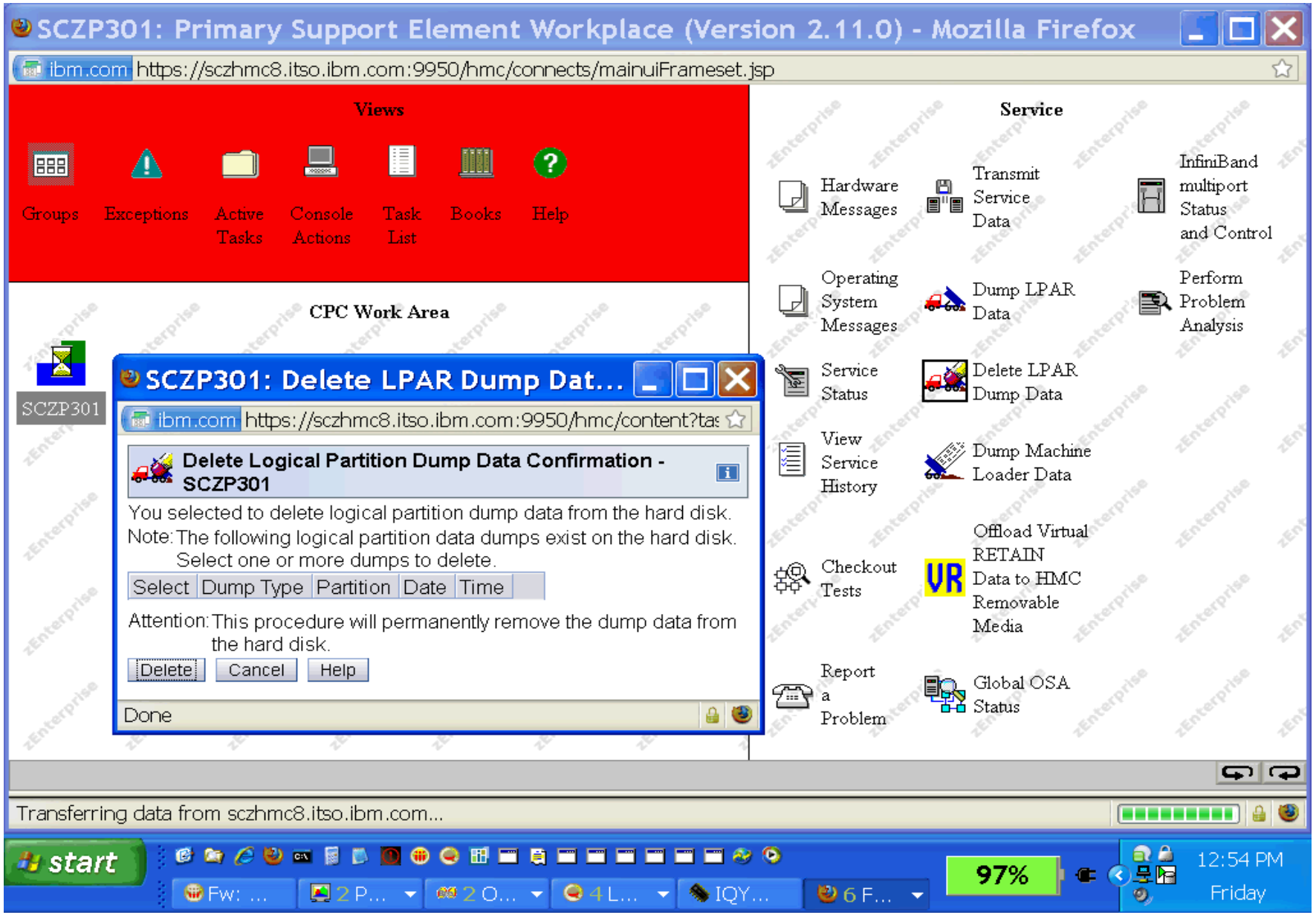
- Test z196 non-disruptive serialized CF Dump
- Observe impact to applications (none)
- Gain experience (get “comfortable”) with non-disruptive serialized CF Dumps

© 2011 IBM Corporation

Clear out Old Coupling Facility Dumps

- Drag-n-drop CPC icon over to Service task called **Delete LPAR Dump Data.**
- Select /check boxes for ALL dumps(both CFCC and LPAR).
  - The CFCC dumps are stored on the Support Element in **/console/ffdc/dumps** directory as either iqzqccf*.* for hard dumps or iqzqcfn*.* for non-disruptive dumps.
- Close the Delete LPAR Dump Data iconcc

© 2011 IBM Corporation

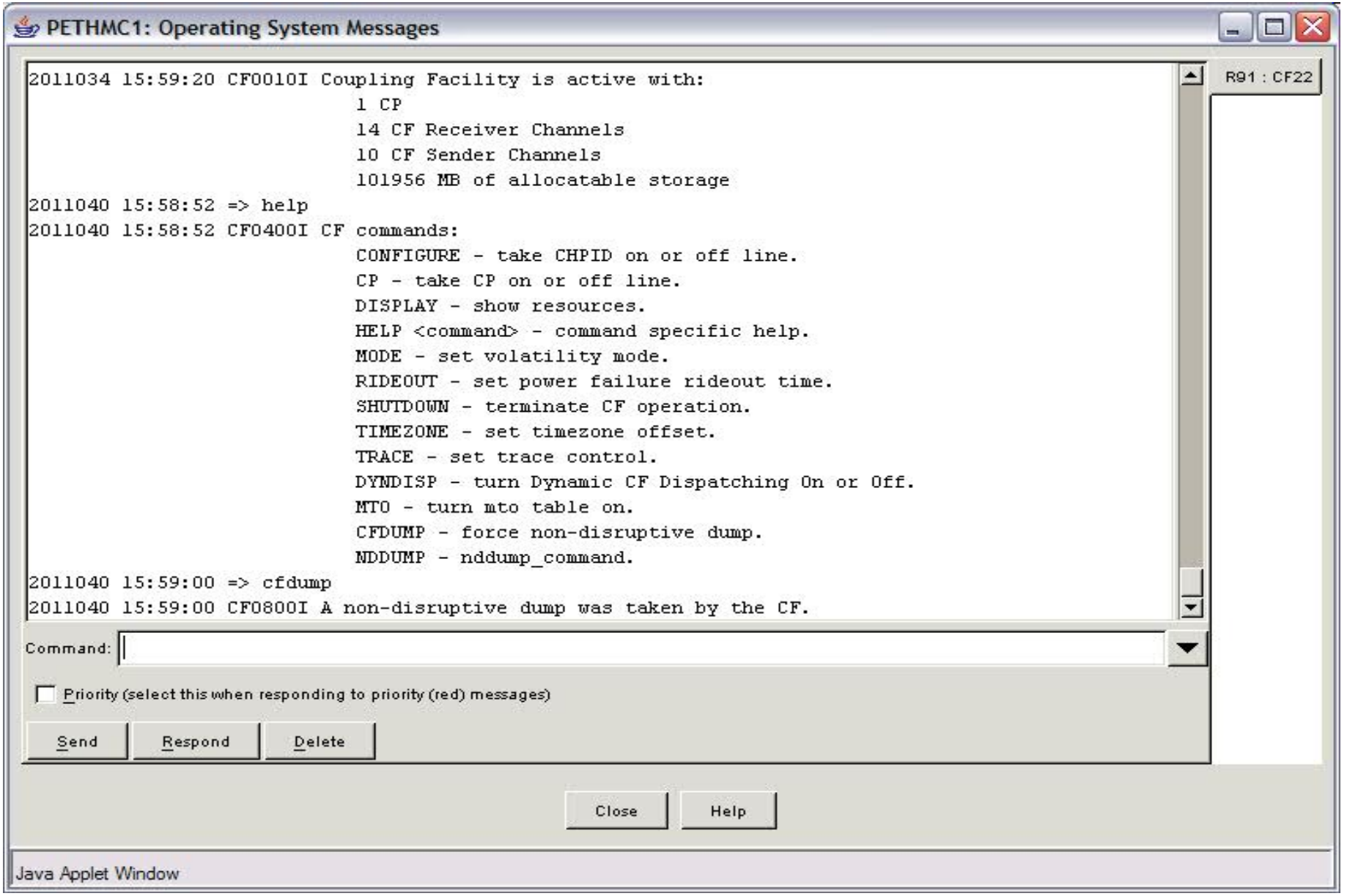


© 2011 IBM Corporation

Command to Initiate z196 “Non-disruptive” CF Dump

- Open the Opermsg console for the CFCC image
- Use the **CFDUMP** command to initiate the dump
- Message CF0800I will appear indicating the dump has occurred.

© 2011 IBM Corporation



© 2011 IBM Corporation

Document CFDUMP Testing Observations

- Any impact noticed when the dump was taken?
- Using the interface

- SHARE
  - CFDUMP command and observations with co-workers

© 2011 IBM Corporation

IBM's Commitment to Serviceability

- Under Development – OA35342
- z/OS operator command to initiate non-disruptive CF dumps
- Enables support center to leverage automation based on z/OS events
  - Performance
  - Loss of connectivity
  - Dropping of links
  - z/OS reporting higher than expected volume of IFCCs

© 2011 IBM Corporation

IBM's Commitment to Serviceability (Cont.)

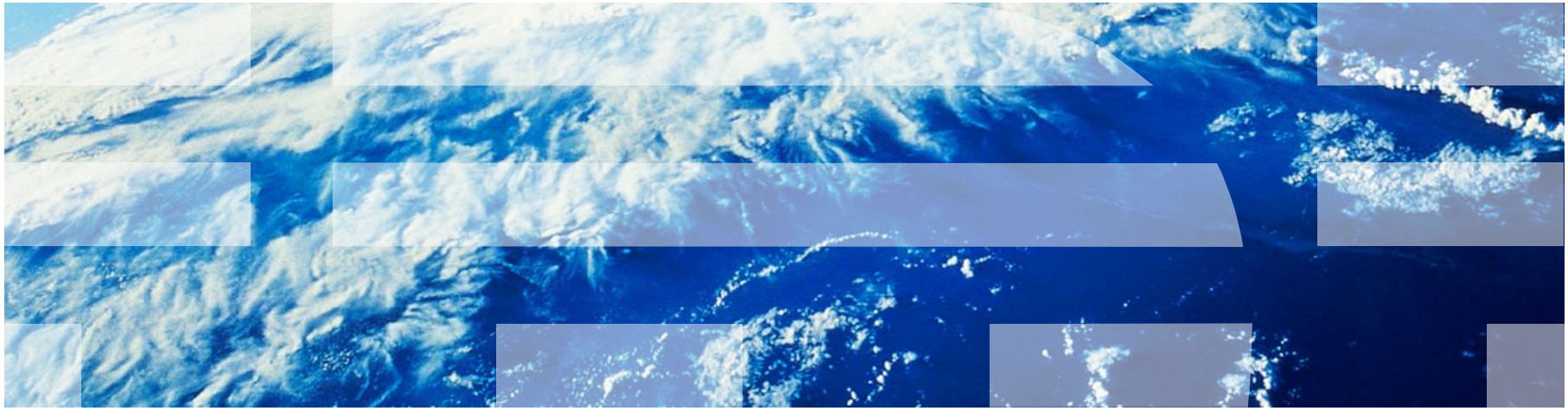
- **Non-disruptive serialized dump support is being rolled back to z10!**
  - **CFCC 16 Service Level 4.01**

- **OA33723 provides z/OS support to initiate non-disruptive serialized dumps on z10**

Reliability, Availability, Serviceability

© 2011 IBM Corporation

Questions?



© 2011 IBM Corporation

Additional **Test** Environment Testing Recommendations

- While you are testing …
- If you have time & interest …
- Consider stress testing application recovery related to CF outages

- Scenario 1: Test CF Terminates and stays down
  - From HMC deactivate CF image
  - Validate application recovery

- Scenario 2: Test CF Terminates and comes right back up
  - From HMC reactivate CF image
  - Validate application recovery

- Verify resiliency of enterprise
  - Additional resiliency information check out SHARE Session 9028: Parallel Sysplex Resiliency

© 2011 IBM Corporation