

ibm.com



Parallel Sysplex Availability Checklist

Frank Kyne - ITSO Poughkeepsie



Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2004

Objective

All mainframe shops have (at least!) two challenges:

- Provide better availability
- Do it with fewer staff!

To help, this presentation provides *SOME* recommended ways of improving the availability of your applications in a sysplex environment. The information is provided in a checklist format, to make it easy to identify features that you are not currently using.



© 2004 IBM Corporation

ibm.com/redbooks

Parallel Sysplex availability

There are a number of facets to availability

- You can have *planned or unplanned* outages
 - Vast majority of outages are planned
 - As machines and systems get larger, planned outages are getting harder and harder to schedule
 - Planned outage avoidance consists of exploiting capabilities in the software and hardware to make dynamic changes, plus good planning
 - Unplanned outage avoidance consists of good configuration planning, keeping up-to-date on service, careful system monitoring, and so on.



© 2004 IBM Corporation

ibm.com/redbooks

Parallel Sysplex availability

Facets to availability.....

- You want to avoid outages, but for unavoidable ones, you want to minimize (mitigate) the impact of the outage
 - In the case of an outage, you want your users to still be able to use their applications.
 - The only way to achieve this is to exploit Parallel Sysplex data sharing and dynamic workload balancing
 - You also want to get back up and running as quickly as possible - this requires education, automation, planning, and testing



© 2004 IBM Corporation

ibm.com/redbooks

Parallel Sysplex availability

Facets to availability.....

- Performance

- There comes a point when performance is *so* bad that the application must be viewed as being unavailable
- Do you have agreed thresholds? Do you have monitoring to:
 - ▶ Warn you when the threshold is reached
 - ▶ Tell you *what part* of the application is causing the problem? There is no point re-IPLing z/OS if the cause of the problem is a faulty network router.....
- Do you have agreed actions to take immediately when those thresholds are reached? Or do you extend the "outage" by arguing over what action should be taken?



© 2004 IBM Corporation

ibm.com/redbooks

Parallel Sysplex availability

Facets to availability.....

- Build from the bottom up

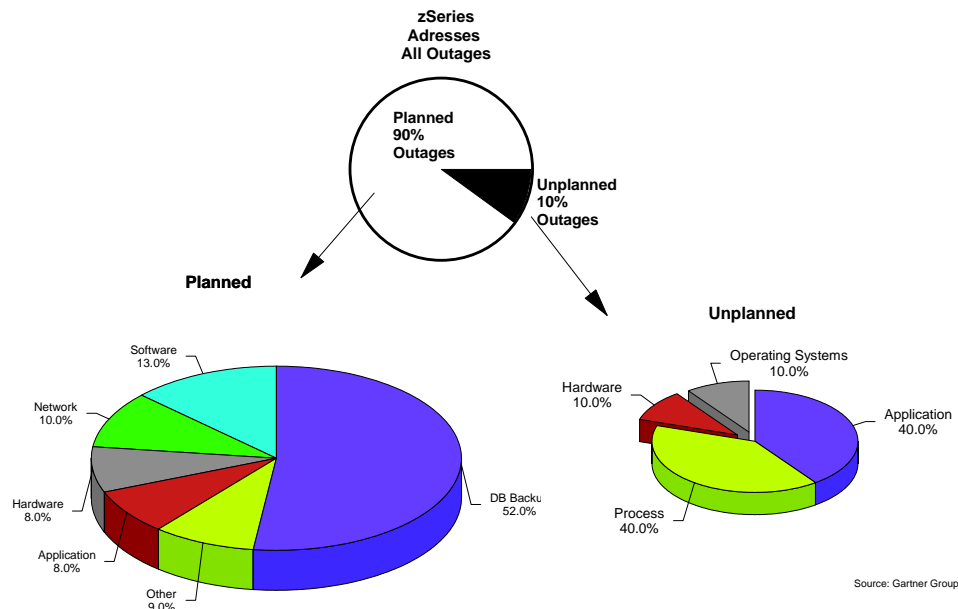
- While sysplex is designed to help you survive an outage of a single component, to have a highly available sysplex, you need highly available components
- So, a lot of the material in this presentation is aimed at improving the availability of a single system - this provides a sound basis for building a highly-available sysplex



© 2004 IBM Corporation

ibm.com/redbooks

Continuous Application Availability



© 2004 IBM Corporation

ibm.com/redbooks

Parallel Sysplex Availability

To get the availability benefits enabled by a Parallel Sysplex, you *must* make the effort to exploit the capabilities it provides

Installing a Coupling Facility does not automagically eliminate all outages

There is NO magic silver bullet to provide continuous availability

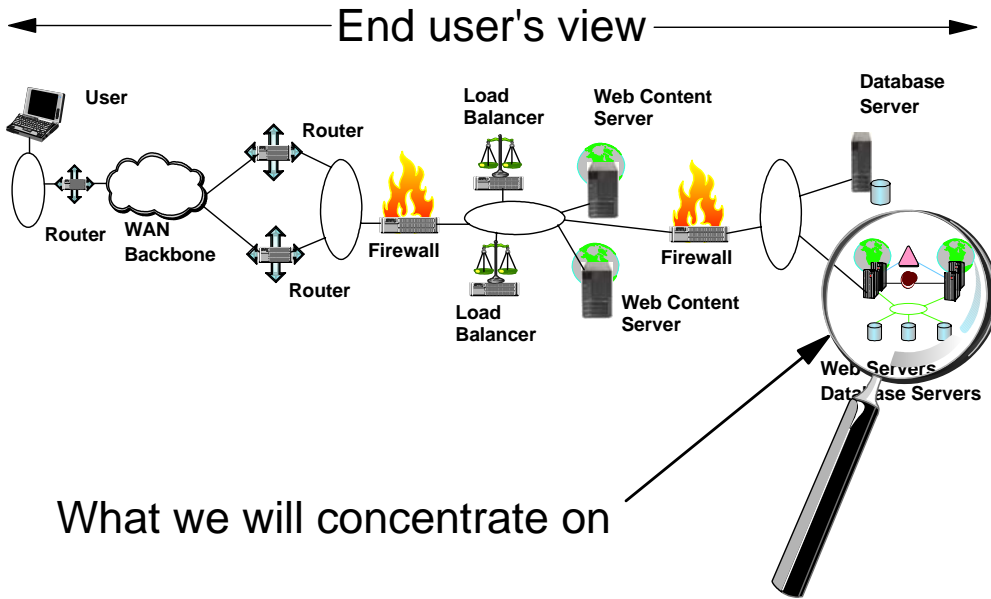
BUT... there are many small changes you can make - when added together, these can make a significant improvement to availability



© 2004 IBM Corporation

ibm.com/redbooks

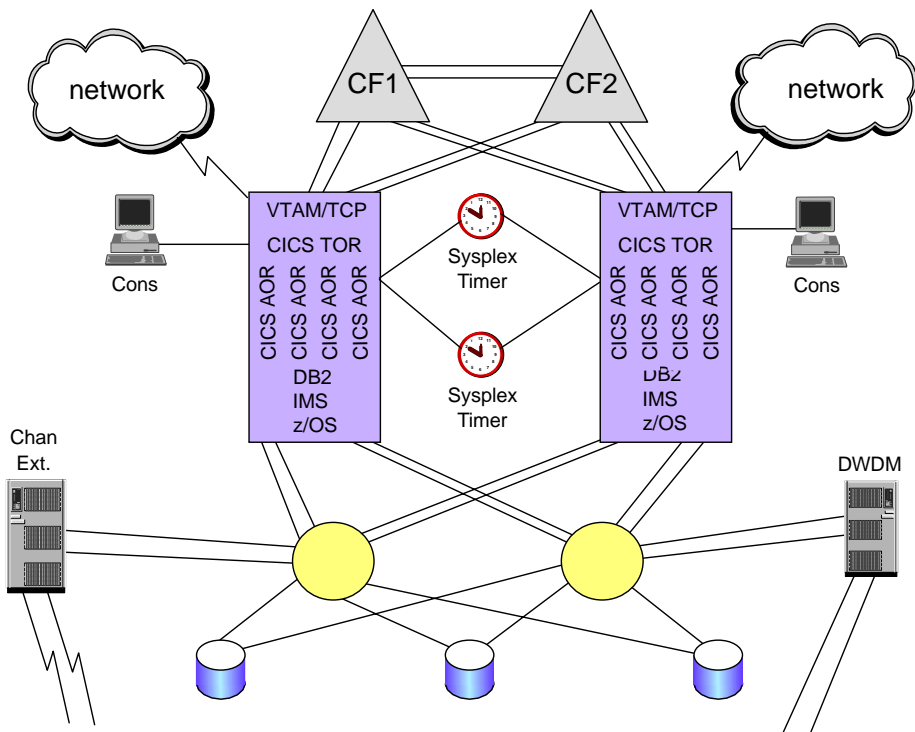
Parallel Sysplex Availability



© 2004 IBM Corporation

ibm.com/redbooks

Application service delivery components

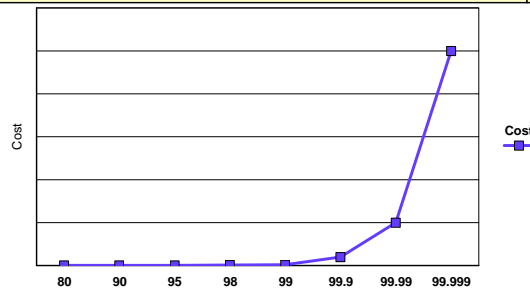


© 2004 IBM Corporation

ibm.com/redbooks

Management considerations

Consideration	Addressed?
Which are your really critical applications - the ones where all the phones ring when they are down? Does your IT spending reflect the criticality of the applications?	
When a new software release goes in, is education about the new features provided for ops and appl developers?	
Is there an executive that has responsibility for availability AND has the budget to improve availability?	
Does the group that is responsible for availability (typically Ops) have the authority to stop an application going into production if that application does not support high availability?	
Education - vital for both System Programmers and Operators	
Does everyone that plays a role in availability have availability as one of their objectives? Ops, sysprogs, and appl developers MUST all work together, but often these groups have different objectives	



© 2004 IBM Corporation

ibm.com/redbooks

Systems Management considerations

Consideration	Addressed?
Do you know your current availability? If you don't know what you are achieving today, how will you know if changes make things better or worse?	
Where is your pain? What is the root cause of your past outages? The pattern varies from installation to installation, so you need to be able to identify where your weaker areas are. If a particular component is causing a large percent of outages, can it be removed or replaced? You also need this information to justify expense of improving availability	
A recent (non-IBM) survey, found that 80% of unplanned outages are as a result of poor Systems Management practices	
Communication (Problem and Change Management) must cover <i>all</i> the people that touch any part of an application, including application developers, database admin, network, distributed servers, ops and the MVS System Programmers	
Software updates - large percent of outages are caused by problems where the fix has been available for more than 6 months. Consider using DR tests as a way of testing the new software levels.	
Subscribe to Red Alerts and apply/address identified fixes as soon as possible	
Security - Do only the people that really should have access actually have access to critical system resources?	
Documentation - are procedures for failure scenarios documented, understood, and <i>regularly tested</i> ?	
Problem determination - how often are you told that you must recreate a problem to get the information required to debug the problem? If this happens a lot, perhaps your procedures should be reviewed? You learn NOTHING from an unexplained outage	
Have you identified all Single Points of Failure? Do you repeat this process regularly? Do you have procedures to handle the recovery should one fail?	



© 2004 IBM Corporation

ibm.com/redbooks

CPU considerations

Feature	Using?
Use latest technology you can afford - each new processor generation has availability/reliability and dynamic upgrade/downgrade improvements over previous generations	
Exploit concurrent upgrade capabilities like CUoD (all types of PUs), On/Off Capacity Demand (all PU types), CBU (CPs only), Concurrent PU conversion, I/O upgrade plan ahead, and Concurrent Memory Upgrade - See GM13-0257-01 for summary of capabilities.	
Ensure every multi-path critical device is configured with <i>at least</i> two channels, through different switches/directors. Use CHPID Mapping Tool to obtain minimum single points of failure, both within the CPC and within the channel subsystem	
Ensure Driver levels are kept up to date, and install Hiper MCLs in a timely manner (nearly all patches are non-disruptive) - establish procedures and communication with hardware service representative	
Implement <i>at least</i> two HMCs and at least two HMC Phone Servers, so that CPCs can report a problem even if one HMC is in the middle of downloading patches	
Consider placing two LPARs from each sysplex on each CPC - ensures that MIPS on that CPC will still be available to the sysplex even if 1 LPAR is shutdown. Use IRD to optimize capacity allocation across LPARs	
Pre-define spare LPARs on z990 (partition name of *) - can then be renamed and used without a POR	
Every LPAR should have some Reserved CPs in LPAR Profile to allow more CPs to be added non-disruptively. On a z990, INITIAL+RESERVED can = 24 (16 if < z/OS 1.6), even on an A08...	
Enable Automatic I/O Reset Facility in CPC Reset profile	
On z900 or earlier, configure sufficient HSA to allow for Dynamic I/O Reconfigurations. On z890/z990, set the MAXDEV value in HCD high enough to provide I/O configuration definition growth	
Don't configure an LPAR with extremely low weights (<5% of CPC) in a production sysplex	
LPARs should either be successfully IPLed and running, or they should be DEACTIVATED. Don't have LPARs that are ACTIVATED but not IPLed, or ones that have failed IPLs that are not addressed (typically sysprog test LPARs)	
Capacity planning - if consolidating onto a smaller number of larger CPUs (z990 for example), do you have sufficient capacity to failover in case of a planned or unplanned CPU outage?	



© 2004 IBM Corporation

ibm.com/redbooks

Switches, directors, and timer considerations

Feature	Exploiting?
Always spread paths to critical devices across more than one director/switch	
Define switch Control Unit Port in HCD, to ensure errors can be reported back to the operating system	
If appropriate, order High Availability configuration when purchasing FICON switch	
Use Model 2 Sysplex Timers - support on Model 1s ran out at the end of 2003. See Redpaper on migration options at: http://www.redbooks.ibm.com/redpapers/pdfs/redp3666.pdf	
Implement the Expanded Availability Configuration with Sysplex Timer Model 2s	
Especially when using recent technology, ensure that microcode levels are kept up to date.	
Ensure that fibre quality and DB loss are within IBM-recommended limits	
z/OS 1.5 adds ability to re-enable Index VTOCs without taking the volume offline to all sharing systems	
z/OS 1.5 enhances REFORMAT command to support dynamically extending VTOC and Index VTOCs	
If you are a GDPS/PPRC user, we <i>strongly</i> recommend the use of HyperSwap to improve the availability of the DASD subsystems	



© 2004 IBM Corporation

ibm.com/redbooks

CF considerations

Feature	Exploiting?
CFLevel 11 is only supported CF Level on 9672 G5/G6. Ensure you are running this level and remember that upgrading to this level is disruptive	
CFLevel 12 (and 13 soon) is only supported CF Level on 2064 and 2066. Ensure you are running this level and remember that upgrading to this level is disruptive. Also, remember that migrating to CFLevel 12 can significantly increase your structure's storage requirements.	
Most CFLevel and CFCC service level upgrades should no longer require a POR following z990 GA3 (or z890 GA1)	
CFLevel 13 is not expected to significantly impact structure sizes compared to CFLevel 12	
If you are a large DB2 data sharer, especially if you have both data and indexes in the same GBP, install CF Level 13	
MUST have at least 2 CFs, even for Resource Sharing	
CFs must have sufficient capacity (CF Links, storage, MIPS) to take over workload from a failed CF. Should not run higher than 50% busy (max)	
If doing data sharing and not using SM Duplexing, at least one CF must be failure-isolated and all 'category 3' structures should be in that CF. See WSC FLASH 98029	
Regularly make sure that all structures are in the CF that you expect them to be in, especially if one CF is an ICF - msys for Operations helps with this; also SETXCF START,REALLOCATE in APAR OA03481 (z/OS 1.4 and later).	



© 2004 IBM Corporation

ibm.com/redbooks

CF considerations

Feature	Exploiting?
Each CF should have enough storage to hold <i>all</i> structures, especially if you only have 2 CFs - but remember to allow for duplexed structures. Remember Control Storage if CF still CFLevel 11 - check available control space before attempting to empty one CF into the other	
Monitor CF and structure storage utilization as you add structures and upgrade CF Levels	
CF with 1 CP should not run at >50% CPU utilization. If there is more than 1 CP, higher utilization is possible without impacting performance	
Must consider impact of high CF CPU utilization on CF response times, especially during CF or system recovery	
Should be <i>at least</i> 2 sender or peer links from every CPC to every connected CF	
If possible, avoid the use of shared engines for production CFs - See ITSO Hint and Tip entitled "Use of Shared Engines for Coupling Facilities"	
Use highest speed links possible, especially for structures with large data transfers - don't permit path busy conditions caused by shared CF links to exceed 20%.	
If considering using System Managed Duplexing, refer to White Paper http://www.ibm.com/servers/eserver/zseries/library/techpapers/gm130103.html	



© 2004 IBM Corporation

ibm.com/redbooks

z/OS Health Checker

Feature	Exploiting?
Use z/OS HealthChecker either in SA/390, msys for Ops, or standalone and ensure you have the latest level - download from http://www14.software.ibm.com/webapp/download/search.jsp?go=y&rs=hchk	
Once you have the HealthChecker running cleanly, schedule it to run every day on every system . You should only ever have to take manual action if there is a non-zero RC	
If you have automation that addresses some of the things HealthChecker monitors for, modify it to work with the WTOs that HealthChecker can put out	
HealthChecker future plans....	



© 2004 IBM Corporation

ibm.com/redbooks

z/OS considerations

Feature	Exploiting?
Paging volumes - should be high performance, support DASD fast write, and use <i>dynamic</i> PAV if available (z/OS 1.3 and later only)	
z/OS 1.3 and later protects page data sets from use by multiple systems by placing system name and other information in page data set and checking this at IPL	
Stripe standalone dump data sets across multiple volumes (4 seems to be optimum number) using AMDSADDD Rexx exec (delivered in SAMPLIB) to minimize SAD duration	
Extended format sequential data sets larger than 64K tracks will be useable for Stand Alone dumps - supported in z/OS 1.6 - rolled back by APAR OA04140	
Check that RNLs for RACF resources conform to guidance in GRS Planning manual	
GRS Star - higher performance, superior availability and recoverability compared to GRS Ring. Considered a pre-req for large sysplexes	
Enable GRS SYNCHRES option in GRSCNF member (or dynamically via SETGRS command)	
Exploit IBM-provided automation in SA/390 and msys for Ops	
EXITs - eliminate where possible (is the exit still needed?), for those that remain, use Dynamic Exit support where possible.	
OPERLOG - used for sysplex-wide problem determination - can be used together with syslog	
LOGREC - sysplex-wide problem recording medium that never fills	
z/OS 1.5 Security Server adds support for dynamic update of RACF Templates (avoids an IPL) and prevents accidentally installing downlevel templates	
z/OS 1.6 Security Server adds support for dynamic updates of RACF Class Descriptor Tables (avoid an IPL)	
z/OS 1.6 Security Server removes the need for a router table entry (ICHRFR0X) for most user-defined classes (avoids an IPL)	



© 2004 IBM Corporation

ibm.com/redbooks

z/OS considerations

Feature	Exploiting?
Place JES2 Checkpoint in CF - equivalent-to-better performance, especially with mixed CPC sizes, and elimination of Reserves	
JES2 Health Monitor started automatically in z/OS 1.4 and later - monitors health of JES2 from separate address space. Use \$J commands to communicate with it.	
z/OS 1.4 adds checking to JES2 to avoid corruption if the wrong checkpoint data set is accidentally used (HASP863)	
Sysplex HFS sharing - improved flexibility and availability for end-users and system programmers	
New REMOUNT option lets you unmount and remount a sysplex-shared HFS without having to unmount all filesystems under the one being remounted - APAR OA02584	
Use of system symbols in Parmlib, Proclib, VTAM, TCP, NetView, OPC, AOC, Clists to minimize multi-system maintenance overhead - with z/OS 1.4, you can have up to 800 symbols	
Exploit Console Restructure feature in z/OS 1.5 and available as a feature on top of z/OS 1.4 - designed to minimize impact of WTO buffer shortages	
z/OS 1.6 will add the ability to have multiple GTF traces running concurrently	
Prior to z/OS 1.6, DPS array initialization during MSI was a serial process. Changes to IOS in z/OS 1.6 parallelize this process, resulting in faster IPLs, especially with large I/O configurations - available back to 1.4 via APAR OA07335	
z/OS 1.6 moves EXCP control blocks above the line - will help relieve constraint below the 16MB line	
In z/OS 1.6, size of buffers for SMF data can be increased to 1GB from the prior maximum of 128MB	



© 2004 IBM Corporation

ibm.com/redbooks

XCF considerations

Feature	Exploiting?
There should be 1 XCF structure per Transport Class per CF. Every structure should be defined as both PATHIN and PATHOUT. If you already have CTCs defined, use those in addition to CF structures, especially if there is a large distance between systems	
Ensure every transport class has at least two paths defined	
Place CDSs on dedicated, high performance volumes with no RESERVEs	
Use a backup method that does not cause RESERVE problems - consider Disaster Recovery requirements	
Place primary, alternate, and spare CDSs on <i>separate devices and separate physical control units</i>	
Place the primary sysplex, CFRM, and LOGR CDSs on different volumes	
Ensure there is automation in place to add a spare CDS if primary or alternate is lost	
APAR OW56355 provides benefits for users of VTAM/XCF experiencing NO BUFFER conditions	
z/OS 1.6 increases the maximum number of members in an XCF group from 1023 to 2047 - OA 04034 rolls support back to z/OS 1.4	



© 2004 IBM Corporation

ibm.com/redbooks

XCF considerations (cont)

Feature	Exploiting?
If APAR OA03481 is applied, always use the SETXCF START,REALLOCATE command repopulate a CF - if not, use the POPULATECF command	
Implement automation to check that each structure is in the first CF in its preference list - REALLOCATE will do this for you as well.	
Implement automation to act on potential hang situations (messages IXL040E and IXL041E).	
If XCF message rates above 1000/sec, see Setting Up a Sysplex for tuning recommendations	
Use enhanced D XCF,PI,STRNM=nnnn command to get response time information for XCF paths - this information is also available in the RMF SMF records, but not in the RMF PP reports.	



© 2004 IBM Corporation

ibm.com/redbooks

CFRM considerations

Feature	Exploiting?
Ensure that System-Managed Rebuild is enabled - use D XCF,C,TYPE=CFRM command (should say SMREBLD(1)).	
Enable Structure Full Monitoring for appropriate structures and set up automation to monitor for resulting messages (to alert appropriate tech support staff)	
AutoAlter enabled for <i>recommended</i> structures. Plus monitor for messages so that CFRM policy can be adjusted accordingly	
Always update the CFRM policy any time a structure size is changed using SETXCF ALTER or Auto Alter	
All structures should have at least 2 CFs in preference list	
Ensure a CFRM policy naming convention is used so that policy changes can be easily backed out	
Use appropriate structure sizes - validate with CFSizer (for rough figure) or check RMF reports (for more accurate indication if existing structure is large enough)	
Never use REBUILDPERCENT > 1 - recommend to allow REBUILDPERCENT to default	
APAR OW48624 provides performance enhancements when partitioning a system out of the sysplex	
z/OS 1.4 delivered further enhancements to reduce the elapsed time for structure rebuilds	



© 2004 IBM Corporation

ibm.com/redbooks

Logger considerations

Feature	Exploiting?
Ensure LOGR CDS is at HBB7705 format level (D XCF,C,TYPE=LOGR) - get this by defining LOGR CDS with SMDUPLEX option	
If all systems > z/OS 1.2, and at least one system is at z/OS 1.3, most log stream attributes can now be changed while the log stream is connected	
Ensure that APAR OW51854 is applied, and set up automation to monitor for IXG310I/IXG311I/IXG312E messages - all relate to offload problems (APAR integrated into z/OS 1.4 and rolled back to OS/390 2.10)	
z/OS 1.5 adds Logger Task monitoring and messages IXG271I, IXG272E. See the section entitled "Offload and Service Task Monitoring" in Setting up a Sysplex for guidance on handling these messages	
In z/OS 1.5, System Logger address space will be automatically restarted unless it is terminated using FORCE ARM (rolled back to 2.10 - 1.4 via OW53349)	
Set up automation to monitor for IXG267I message - issued at data set allocation time and indicates that offload or staging data sets do not have the correct share options (3 3)	
z/OS 1.4 increased number of DASDONLY connections on a single system from 512 to 1024. (APAR OW51437)	
z/OS 1.2 added support for use of System Managed Duplexing with CF Log streams	
Never delete offload or staging data sets manually without first checking to ensure Logger has no knowledge of those data sets - always do a DELETE LOGSTREAM to remove log stream definitions and all associated data sets.	



© 2004 IBM Corporation

ibm.com/redbooks

Logger considerations

Feature	Exploiting?
Put logstreams with similar average buffer size and request rates in the same structure. Do <i>not</i> mix idle and busy logstreams in same structure	
Even if you don't intend to use staging data sets, specify STG_DUPLEX(YES), DUPLEXMODE(COND) (or apply the PTF for OA03001) and at least STG_SIZE and STG_DATACLAS for every log stream	
Use data class with 24K CI Size for all offload data sets - staging data sets must still have 4K CI Sizes	
When using CF Log streams and duplexing to DASD, ensure that staging data sets are sufficiently large that they are not hitting the High Threshold before the log streams in the CF	
Have at least 2 different systems connected to each LOGR structure - allows peer recovery in case of a system failure	
For CICS DFHLOG, IGWLOG, and RRS Active log streams, monitor offload data sets - if sequence number is large, adjust HIGHOFFLOAD and LOWOFFLOAD or structure size	
Refer to recent ITSO Redbook on System Logger (z/OS Systems Programmers Guide to: System Logger (SG24-6898)) for detailed information on Logger usage and tuning	



© 2004 IBM Corporation

ibm.com/redbooks

SFM considerations

Feature	Exploiting?
USE SFM. Single change that would have eliminated the largest number of multi-system outages	
If the policy is defined and started, SFM can: <ul style="list-style-type: none"> ↳ Partition a system out of the plex if status update is missing ↳ Partition a system out of the plex if XCF signalling to the system is lost ↳ Assist with V XCF,xxx,OFFLINE processing ↳ Control structure rebuild processing based on REBUILDPERCENT 	
Don't set INTERVAL in COUPLExx - let system determine it dynamically	
You must have a CF to be able to use SFM's system isolation function	
Set CLEANUP interval to 15 seconds - should be sufficient for planned shutdowns	
Do <i>not</i> specify PROMPT in the SFM - use ISOLATETIME(0) instead to get sick system immediately removed from sysplex	
Don't specify OPNOTIFY- let it default to the calculated INTERVAL plus 3 seconds	
Ensure that "Enable I/O Interface Reset" option is enabled in CPC Reset profile	



© 2004 IBM Corporation

ibm.com/redbooks

DFSMS features

Feature	Exploiting?
Use DFSMSHsm Multiple Address Space HSM - lets you break HSM tasks across multiple address spaces, running higher priority tasks (like recall) in one address space, and lower priority tasks in another (OS/390 2.10)	
Enable Fast Subsequent Migration, but be aware that this will result in larger MCDS	
Use DFSMSHsm Common Recall Queue - balances HSM recall processing across hosts in a sysplex and provides persistency of recall requests should the HSM address space fail or be stopped - see 2.3.1 in HSM Stor Admin Guide	
z/OS 1.6 adds support for multiple HSM hosts to concurrently run all phases of Secondary Space Management, except for L1 DASD to L2 DASD migration	
OS/390 2.10 increased number of concurrent DFSMSHsm backup tasks from one per system to up to 64 per system. You can also backup data sets directly to tape. See SETSYS DSBACKUP keyword.	
Recommend a max of 6 tasks per function per host. To increase throughput, increase the number of hosts running each function, not the number of tasks.	
Use DUMPIO(3 4) not (4 4)	
Place CDSs on SMS-managed volume to prepare for Extended Attribute.	
If there are a very large number of HSM-managed data sets or volumes, use VSAM/RLS for CDSs to provide improved performance for CDS-intensive activities	
z/OS 1.5 adds the ability to use CDSQ or CDSR serialization with VSAM Extended Attribute HSM CDSs (removes requirement for VSAM/RLS)	
For more tuning suggestions, refer to Chapter 2.5 of DFSMSHsm Implementation and Customization	
If you still have a DFSMSHsm performance bottleneck, IBM offers a fee service to assist with DFSMSHsm Tuning - ask for Stan Kissinger	



© 2004 IBM Corporation

ibm.com/redbooks

DFSMS features

Feature	Exploiting?
If you have >1 HSMplexes in the same sysplex, implement DFSMSHsm Single GRSplex support to remove false contention on the CDSs	
Enable DFSMSHsm Secondary Host Promotion (OS/390 2.8)	
Specify RESTART='(a,b)' on HSM startup JCL to automatically restart HSM following a failure	
Exploit DFSMSHsm support for creating duplex tapes	
DFSMSHsm supports Concurrent Copy for all data sets, both SMS-managed and non-SMS managed	
DFSMSHsm in z/OS 1.5 adds copy pool support that is used by DB2 V8 to take point in time copies (using FlashCopy) of an entire DB2 subsystem	



© 2004 IBM Corporation

ibm.com/redbooks

DFSMS features

Feature	Exploiting?
Enhanced Catalog Sharing - performance equivalent to non-shared DASD environment	
Warning messages (IEC361I) issued if catalog space utilization exceeds a user-specifiable threshold - set using F CATALOG,NOTIFYEXTENT (z/OS 1.3)	
Ensure that SYSIGGV2 and SYSZVDS ENQs are handled "correctly" in GRS RNLs - see GRS Planning manual for details	
Enhanced Catalog Sharing structure supports User Managed Rebuild, permitting recovery of the structure following a CF failure - OS/390 2.10	
Data set separation support feature (z/OS 1.3) allows you to give SMS a list of SMS-managed data sets that should not reside on the same physical control unit	
OAM sysplex support - ability to access OAM data from any system in the sysplex, and remove OAM affinity to just one system	
In z/OS 1.5, GDS reclaim processing default can be changed without an IPL	



© 2004 IBM Corporation

ibm.com/redbooks

DFSMS features

Feature	Exploiting?
"Space Constraint Relief" (OS/390 2.4) decreases chances of a data set allocate or extend failing due to lack of DASD space - see Data Class definitions. Also, see "Reduce Space By %" parameter also in Data Class definition	
Specify Overflow and Extend Storage Groups for production storage groups to reduce out-of-space abends (z/OS 1.3)	
z/OS 1.5 adds ability to automatically kick off space management when a storage group reaches its high threshold. Also, SMS issues WTOs that can be used to kick off automation actions	
Also in 1.5, SMS will place data sets in storage groups in the sequence they are specified in the ACS routines. The first specified storage group will be used till it fills, then the next one till it fills, and so on	
Ability to rename duplicate data sets - support added in OS/390 2.10. Protected by RACF FACILITY profile called STGADMIN.DPDSRN.*	
z/OS 1.6 adds ability to take existence of PAV into account when allocating SMS-managed data sets (new PAV keyword in Storage Class)	
Exploit Enhanced Integrity for sequential data set support in z/OS 1.5 - see member IFGPSEDI in SYS1.PARMLIB	
OS/390 2.10 delivered improvements in trace facility for VSAM data sets	
Peer to Peer VTS support provides ability to keep a duplex copy of volumes in a Virtual Tape Server - delivered in OS/390 2.10	
DFSMSdss support for data set extent reduction while DEFRAGING a volume - added in OS/390 2.10 (CONSOLIDATE keyword on DEFRAG)	
Data set level FlashCopy (requires FlashCopy V2) available with DFSMSdss in z/OS 1.3 and rolled back to OS/390 2.10.	



© 2004 IBM Corporation

ibm.com/redbooks

DFSMS features

Feature	Exploiting?
z/OS 1.4 creates a new address space, SMSPDSE, to contain all PDSE-related control blocks, removing them from ESQA and ECSA	
Add ability to restart one of the PDSE address spaces (SMSPDSE1), potentially avoiding an IPL if a restart is required - planned for z/OS 1.6. Must specify PDS_RESTARTABLE_AS(YES) in IGDSMSxx member	
The SMS SCDS data set can be recreated from the ACDS using the SETSMS SAVESCDS command - new in z/OS 1.5 (See dfp Stor Admin Ref for details)	
Extended distance remote copy without XRC support added by PPRC XD in z/OS 1.3. PPRC Global Mirror (Asynch PPRC) builds on this capability	
Coupled XRC supports lets you XRC more volumes with time consistency across the full set. Also, XRC parameters can now be placed in parmlib members rather than having to issue XSET (z/OS 1.3)	



© 2004 IBM Corporation

ibm.com/redbooks

Network/sysplex considerations

Feature	Exploiting?
See Alfred Christensen's presentations Z15 and ZW06	
VTAM Generic Resources for: CICS DB2 IMS TSO NetView Access Services	
VTAM Multi Node Persistent Sessions for: CICS (CICS TS 2.2 supports signon persistence) IMS Rapid Network Recovery	
TCP/IP can use Dynamic XCF for intra-sysplex IP communication	
TCP/IP use XCF signalling to provide auto notification when TCP/IP stacks join or leave the sysplex	
TCP/IP Virtual IP Addressing	
TCP/IP Dynamic VIPA movement between systems in the sysplex	
TCP/IP uses the Coupling Facility to coordinate assignment of client port numbers when a Sysplex-wide source VIPA address is in use	
TCP/IP Sysplex Distributor (See SG24-5235-02)	
TCP/IP uses the Coupling Facility to coordinate Sysplex-wide IPSEC security associations	
TCP/IP WLM/DNS support - start replacing with Sysplex Distributor if you are using this feature	



© 2004 IBM Corporation

ibm.com/redbooks

CICS considerations

Feature	Exploiting?
Make sure all application programs (COBOL, PL/I) are running above the line	
Make sure all application programs (COBOL, PL/I) are running above the line (worth saying it twice!)	
Enable CICS storage protection to protect CICS code and control blocks from errant user programs	
Make sure CICS transaction isolation is used to protect CICS transactions from each other	
If possible, use the TCP/IP sockets support provided with CICS rather than the CICS sockets support provided with Communications Server	
Define CICS TORs as VTAM Generic Resources	
Define two cloned TORs for each CICSplex on each MVS image	
Use a separate Sockets Owning Region (SOR) on each image	
If using CICS Java support, define a separate Java Owning Region (JOR)	
Make sure there is enough capacity (MIPS, below the line storage, MAXTASKS, DB2 and DBCTL threads, etc) in the CICS regions to handle failover should an MVS image fail	
CICS supports DB2 group attach in CICS TS 2.2 - be sure to specify RESYNCMEMBER(YES) if you restart CICS and DB2 on another MVS image in case of failures	



© 2004 IBM Corporation

ibm.com/redbooks

CICS considerations

Feature	Exploiting?
Use CICS Shared Temporary Storage in CF to eliminate non-recoverable temporary storage as an affinity - also, CICS TS 2.2 supports System Managed Rebuild and Duplexing for this structure	
Use CICS Named Counter Server in CF to provide unique counter values across multiple CICS regions - also, CICS TS 2.2 supports System Managed Rebuild and Duplexing for this structure	
Use CICS Global ENQ/DEQ to propagate application ENQs for CICS resources across multiple CICS regions in a CICSplex, removing the affinity to a particular region	
Use CICS Data Tables in CF to share information currently held in a CICS data table across multiple CICS regions - also, CICS TS 2.2 supports System Managed Rebuild and Duplexing for this structure	
Use CICS VSAM Record Level Sharing to remove CICS FORs as a single point of failure	
Use CICSplex System Manager (part of CICS TS since CICS TS 1.1) for CICSplex management and dynamic transaction routing	
Use CICS Transaction Affinities Utility or CICS Interdependency Analyzer to identify, remove, or manage transaction affinities	



© 2004 IBM Corporation

ibm.com/redbooks

DB2 considerations

Feature	Exploiting?
Use DB2 data sharing	
Always duplex DB2 GBPs. If the DB2 Lock or SCA structure shares a single point of failure with a connected DB2, use System Managed Duplexing for those structures	
If restarting DB2 on another image after a system failure, use the DB2 RESTART LIGHT option added in DB2 V7 - Note that RESTART LIGHT will not connect to other resource managers to resolve IN-DOUBT work. THIS IS CHANGED IN DB2 V8	
Especially for batch jobs, do NOT let them do many many updates without issuing a COMMIT	
Partition databases so the whole database does not need to be offline for maintenance. DB2 V6 allows you to change partition boundaries and only impact the affected partitions, NOT the whole tablespace (SQL also supports query parallelization across partitions - the more partitions, the more potential parallelization)	
Prior to CFLevel 13, use different buffer pools for data and index portions of databases. CFLevel 13 addresses high CF CPU utilization that can occur during large online reorg.	
Online reorg allows you to reorg tablespaces or partitions transparently to the user (slight pause while DB2 switches to the new copy)	
In every new DB2 release, the performance and function of DB2 utilities is enhanced to provided better data availability - do your Application developers and DBAs exploit these new capabilities?	
As DB2 reliability and availability improve, an increasing percent of outages are caused by application problems. Do you have tested and consistent plans for restoring databases to a point in time?	
RRS considerations - If RRS did NOT fail, DB2 must restart using the same RRS instance. If RRS DID fail, DB2 can restart on another image, and all involved resource managers must start on the same system UNLESS you are z/OS 1.2 or later (removes restriction on grouping). z/OS 1.6 will let DB2 and other RMs start anywhere, even if the original RRS is still up.	
Change IRLM timeout interval from default of 300 seconds to 10-30 seconds with F IRLM,SET,TIMEOUT=xx,DBMS. Also, new IRLM supports 64-bit and ONLY PC=YES	



© 2004 IBM Corporation

ibm.com/redbooks

IMS considerations

Feature	Exploiting?
Have multiple cloned IMSs - cloning critical so apps can run in any of the IMSs	
Share all databases, and use Fast Database Recovery (FDBR) region to free up locks in case of a failure	
Use shared queues to protect from failure of a control region	
Use Extended Terminal Option (ETO) to add IMS terminals dynamically	
Use VTAM Generic Resources to mask control region unavailability from logging-on users	
If using VTAM GR, set up IMS so that it releases the terminal back to VTAM if IMS dies. The method to do this varies with IMS version - look up GRAFFIN to find how it applies to your version of IMS.	
For TCP/IP connected devices, use Sysplex Distributor and/or Websphere Edge Server	
Use IMSGROUP to allow spreading of BMPs across data sharing members	
If IMS V8 or later, implement IMS Common Service Layer with Sysplex Terminal Management	
Exploit High Availability Large databases (HAL DBs) to let you manage parts of large databases without impacting the whole database	
Exploit concurrent backup and reorg capabilities	



© 2004 IBM Corporation

ibm.com/redbooks

Other MVS-related exploiters

Feature	Exploiting?
Utilize WLM-Managed Batch Initiators where appropriate - can adjust to provide an "appropriate" number of batch initiators if work gets re-routed around a system outage - improved in z/OS 1.4	
Use WLM Resource Affinity Scheduling together with automation to control where jobs can run	
Set up an OPC Hot Standby Controller to take over in case of planned or unplanned outage of the OPC Controller	
Exploit the OPC Interface to WLM to alter WLM Service Class for selected jobs/applications	
Sysplex enhancements in System Automation for OS/390 (also available to non-SA/390 customers as "msys for Operations")	
GRS Enhanced Contention Analysis (enhanced D GRS command)	
SNA Console support - full function MVS console (except NIP support) from any SNA or IP screen	
Use DFSMS Data Set Separation feature to keep software-duplexed data sets on separate control units	



© 2004 IBM Corporation

ibm.com/redbooks

Testing (that horrible word)

Feature	Exploiting?
Do you have a test environment? Is it used for destructive testing or just to test things that you know will work?	
How closely does the test environment reflect the production one?	
What percentage of bugs (software or hardware) are found in the test system? What percentage in production?	
When you hit a bug in production, do you analyze why it wasn't found in the test environment? Do you adjust the test environment to rectify this?	
Do you have a Training environment for operators and systems programmers or do they try out their procedures on the production system?	
Do the operators <i>understand</i> their procedures? Do you provide training so they understand what is happening (IBM has an offering called Parallel Sysplex Training Environment that provides this capability)	
Do you test all changes you will make to the production system? Do you test the backout procedures?	



© 2004 IBM Corporation

ibm.com/redbooks

Parallel Sysplex Availability

Other sources of information:

- [SG24-6061 Sysplex Availability Guide \(currently in development\)](#)
- [SG24-2086 Continuous Availability S/390 Technology Guide](#)
- [SG24-4502 Parallel Sysplex Continuous Availability Presentation Guide](#)
- [SG24-4503 Parallel Sysplex Continuous Availability SE Guide](#)
- [SG24-4593 Planning for CICS Continuous Availability in an MVS/ESA Environment](#)
- [SG24-5680 Implementing ESS Copy Services on S/390](#)
- [Parallel Sysplex Training Environment, available from IBM Learning Services](#)
- [Parallel Sysplex Availability Checklist, available at:](#)
 - http://www.ibm.com/servers/eserver/zseries/library/whitepapers/pdf/availchk_parsys.pdf
- [Other Parallel Sysplex availability documents available at:](#)
 - <http://www.ibm.com/servers/eserver/zseries/psa/>



© 2004 IBM Corporation

ibm.com/redbooks