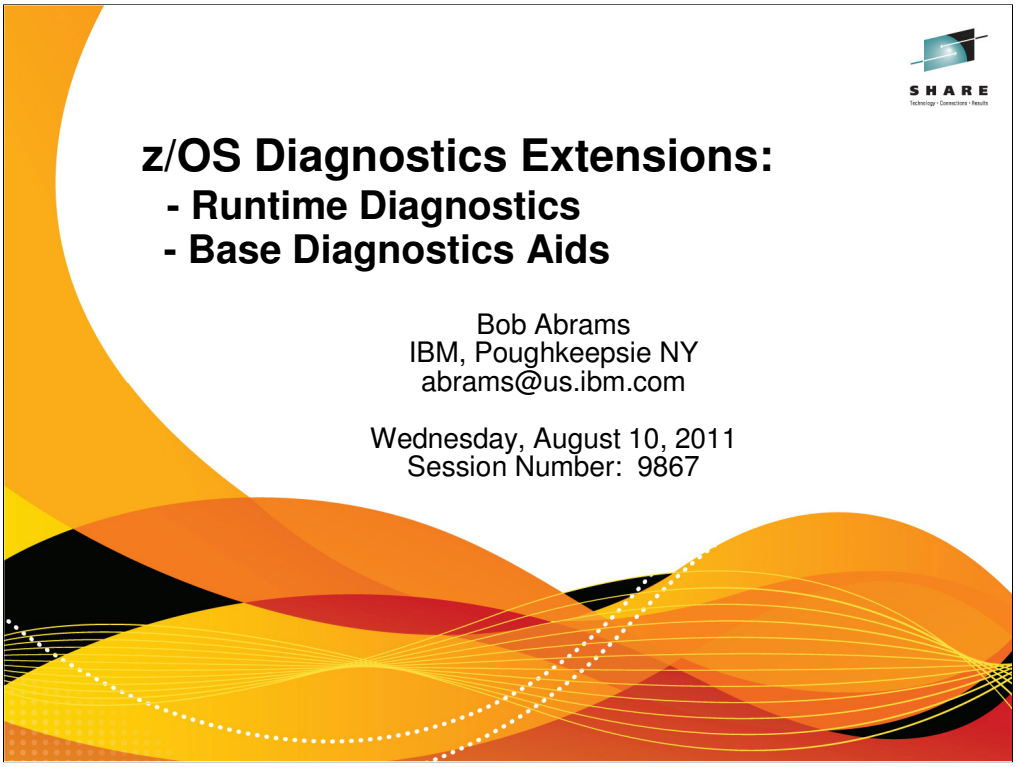

SHARE
Technology · Connections · Results
z/OS Diagnostics Extensions:
- Runtime Diagnostics
- Base Diagnostics Aids
Bob Abrams
IBM, Poughkeepsie NY
abrams@us.ibm.com
Wednesday, August 10, 2011
Session Number: 9867

Problem Determination Update - Agenda

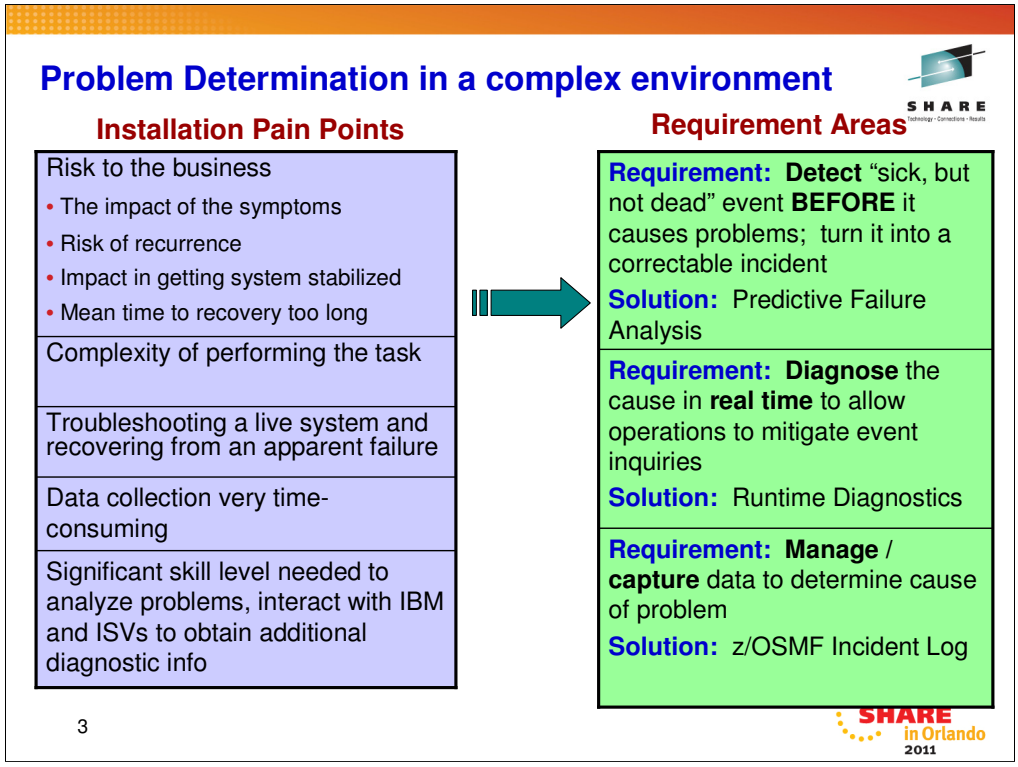
- Requirements for Problem Determination Improvements
  - Distinguish z/OS PD solutions
- Problem identification: Runtime Diagnostics
  - Help analyze, diagnose a soft failure-related problem
- z/OS Service Aids enhancements (R10-R13)
  - Large System Effects
  - Availability
  - Simplification

2

Goal of this presentation is to show areas where IBM z/OS is investing in simplifying problem determination tasks in

Problem identification: Runtime Diagnostics

Service Aids enhancements



Problem Determination pain points, requirement areas, and z/OS functions in these areas

There have been a lot of requirements analysis in the past, focused on gaining an understanding on what the requirement areas are related to Problem Determination.

What is the need that we are trying to address: troubleshooting ... System programmers are spending too much time on actives that do not help them , that are not helping their business, it's complicated to collect the right information and documentation, it is time consuming, and in addition to the problem analysis, it also includes how do you manage the problem data – the diagnostic data that is needed to even begin debugging. How do you reduce the time and skill required to do that.

When a problem is encountered on a z/OS system today, the system programmer has to take many manual and time consuming steps to collect diagnostic data like dumps and appropriate excerpts from logs and then send those materials elsewhere, such as to IBM or ISV support, via FTP. Getting a consolidated list of the abend related problems across a sysplex, that the system programmer can work with, is also a challenging task right now.

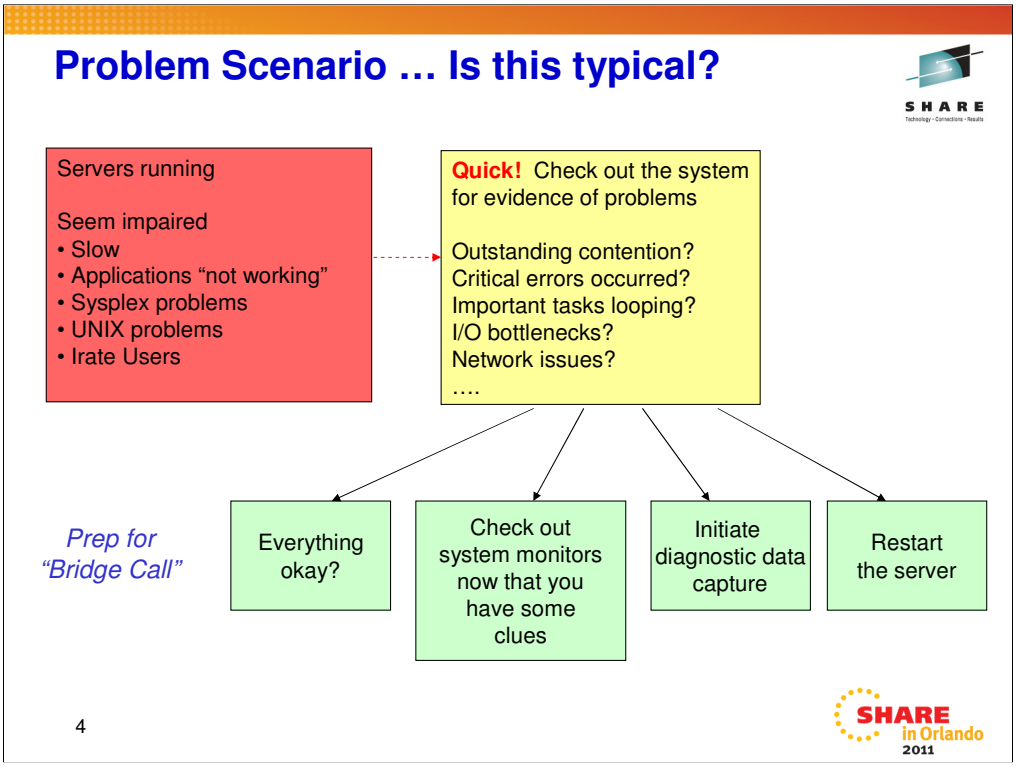
Historically, z/OS has had a rich set of diagnostic service aids, such as different forms of dumps and traces, commands to initiate dumps and the ability to set traps in suspicious running code. This is great for those who are skilled in the art of z/OS debugging, but additional help is required for the installation.

"Sick but not dead" problems have been identified as among the most critical area needing help, because z/OS is engineered to recover, and yet the recovery may be identifying areas where the system programming staff should focus. Predictive Failure Analysis uses data analysis technology to identify storage consumption and "damaged system" issues based on arrival rates of different event triggers.

Diagnosing the system in "real time" is another critical area, to identify symptoms and culprits causing underlying problems that in aggregate may be the cause of sick but not dead issues. This is the focus of Runtime Diagnostics.

The need to simplify how diagnostic materials are identified and handled is addressed by the z/OSMF Incident Log. "Incidents" are created whenever an SVC dump is written to a data set, and common diagnostic data is captured for those problems. Log Snapshots created at that time also collect additional data from the system and tie them all together with each Incident. We collect 30 minutes of Operlog, 1 hour of Logrec detail, and 4 hours of Logrec summary. Not only can you use this information to review all the incidents on your sysplex, you can drill down on those to see what they are what data is associated with them and also FTP the doc to IBM, ISV or elsewhere for further debugging. In addition, interaction with DAE to indicate that the "next dump" be taken for an incident is simplified.

This presentation focuses on Runtime Diagnostics and enhancements made over the last several releases related to capture and processing of Diagnostic service data.

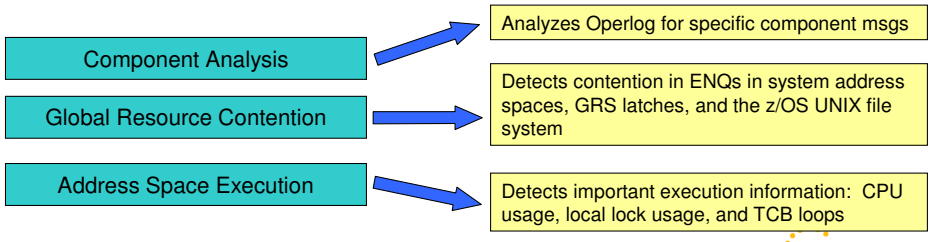


Problem occurs … how often do you have people running around trying to figure out what's going on, and trying to determine the next steps?

Attempts to isolate the problem depending on symptoms.

Runtime Diagnostics

- Analyzes a “sick, but not dead” system in a timely manner
- Performs analysis similar to a very experienced system programmer
  - But faster – goal of 60 seconds or less
  - More comprehensive
  - Looks for specific evidence of “soft failures”
  - Provides suggested next steps
- Runtime Diagnostics
  - Is not automation or a monitor
  - Takes no corrective action, but recommends next steps
  - Has no background processing and minimal dependencies on system services



5

SHARE in Orlando 2011

Looking at the operating system only!

Diagnose sick system by identifying symptoms that could lead to identifying the culprit, and offering next steps to take.

3 areas:

•Component analysis (messages)

•Global resources (ENQs)

•Local address space characteristics

Runtime Diagnostics Benefits

- Reduces the skill level needed by a system programmer for investigating soft failures
  - Provides timely, comprehensive analysis at a critical time period
  - *Also great productivity aid for experienced system programmers!*
- Allows you to *quickly discover next actions* to take such as
  - which jobs to cancel
  - what to investigate further
    - Such as classes of resources or a single address space using a monitor like RMF or Tivoli Omegamon
- Use Runtime Diagnostics …
  - when the help desk or operations reports a problem on the system
  - to get ready for the "bridge call"
  - when PFA detects abnormal behavior

6

Use it when getting ready for a bridge call.

Discreet symptoms

Runtime Diagnostics Invocation

- z/OS 1.12 – Started task – "Run" the analysis via a START command
  - **START HZR,SUB=MSTR**
  - Invokes HZR PROC
  - Will only run on R12 system, but other systems in the Sysplex do not need to be R12
  - Can override HZROUT to specify a data set, for example:
    - //HZROUT DD DISP=SHR,DSN=MY.DATA,RECFM=FB,BLKSIZ=0,LRECL=121
- z/OS 1.13 – Address space – started with the START command above
  - Address space needs to be available for PFA integration
    - Recommend to start address space at IPL (e.g., COMMNDxx)
    - Allocate HZROUT with DISP=SHR … allows you to view the external file without stopping HZR
  - "Run" the analysis via a MODIFY command
    - **f hzr,analyze**
  - Migration Action: If you used Runtime Diagnostics in z/OS 1.12, ensure you update the hzrproc to use PGM=HZRINIT instead of PGM=HZRMAIN.

7

Invocation … Runtime Diagnostics is based on the start task, HZR. START it when needed

ENQ and operlog when run against a different system.

Runtime Diagnostics Invocation (continued)

- The output of Runtime Diagnostics is a multi-line WTO
  - Can also be directed to a sequential dataset using HZROUT DD
- SYSNAME option targets system other than HOME
  - Operlog and ENQ analysis are done for specified system
    - Operlog is suggested to allow message analysis
  - Example: OPTIONS=(SYSNAME=SYS2)
  - z/OS 1.12 – SYSNAME option on *START* command
  - z/OS 1.13 – SYSNAME option on *MODIFY* command
- DEBUG option for use under IBM Service guidance
  - Takes a dump to help debug analysis
  - Options specific to type of analysis and when found or not found
  - Example: OPTIONS=(DEBUG=(LOOP,NOENQ))
  - z/OS 1.12 – DEBUG option on *START* command
  - z/OS 1.13 – DEBUG option on *MODIFY* command

8

Invocation … Runtime Diagnostics is based on the start task, HZR. START it when needed

ENQ and operlog when run against a different system.

Runtime Diagnostics Output

- Success

```
f hzr,analyze
HZR0200I RUNTIME DIAGNOSTICS RESULT 974
SUMMARY: SUCCESS
 REQ: 001 TARGET SYSTEM: SY1       HOME: SY1       2010/12/21 - 11:30:57
 INTERVAL:  60 MINUTES
 EVENTS:
  FOUND: 05 - PRIORITIES: HIGH:05  MED:00  LOW:00
  TYPES: CF:04
  TYPES: HIGHCPU:01
```

- Success, although Operlog not active (text modified in R13)

```
f hzr,analyze
HZR0200I RUNTIME DIAGNOSTICS RESULT 239
SUMMARY: SUCCESS
 REQ: 001 TARGET SYSTEM: SY1      HOME: SY1      2011/08/05 - 09:15:19
 INTERVAL:  60 MINUTES
 EVENTS:
  FOUND: 01 - PRIORITIES: HIGH:01  MED:00  LOW:00
  TYPES: HIGHCPU:01
  PROCESSING BYPASSED:
   OPERLOG....OPERLOG IS NOT ACTIVE.
 ------------------------------------------------------------------------
EVENT 01: HIGH - HIGHCPU      - SYSTEM: SY1      2011/08/05 - 09:15:20
ASID CPU RATE:99%     ASID:002E   JOBNAME:IBMUSERX
STEPNAME:STEP1    PROCSTEP:         JOBID:JOB00051 USERID:IBMUSER
JOBSTART:2011/08/05 - 09:12:05
  ERROR: ADDRESS SPACE USING EXCESSIVE CPU TIME. IT MIGHT BE LOOPING.
 ACTION: USE YOUR SOFTWARE MONITORS TO INVESTIGATE THE ASID.
```

9

SHARE in Orlando 2011

In R12, if Operlog is not active, this is represented as a cryptic IXGCONN Connect Error. It was modified in R13 to indicate that Operlog is not active. Other behavior remains the same.

Runtime Diagnostics: Critical Message Analysis

- Component-specific, critical messages in OPERLOG
  - "Needles in a haystack"
  - Looks one hour back, if available
  - For some messages, additional analysis done
    - Groups related messages into a single event
    - Weeds out shortage and relieved critical messages
    - In some cases, will only show last message if a critical message for the same resource name is repeated, say every 10 minutes
  - Message summary found listed in Runtime Diagnostics output

```
EVENT 02: HIGH - CF           - SYSTEM: SY1      2011/02/15 - 14:47:03
IXC585E STRUCTURE LIST01 IN COUPLING FACILITY TESTCFN,
PHYSICAL STRUCTURE VERSION C7565A8D E48F6410,
IS AT OR ABOVE STRUCTURE FULL MONITORING THRESHOLD OF  80%.
ENTRIES:  IN-USE:         491 TOTAL:         583,  84% FULL
ELEMENTS: IN-USE:         508 TOTAL:        1167,  43% FULL
  ERROR: INDICATED STRUCTURE IS APPROACHING FULL MONITORING THRESHOLD.
 ACTION: D XCF,STR,STRNAME=strname TO GET STRUCTURE INFORMATION.
 ACTION: INCREASE STRUCTURE SIZE OR TAKE ACTION AGAINST APPLICATION.
```


These are the areas analyzed by Runtime Diagnostics in z/OS R12

Runtime Diagnostics: ENQ Contention Checking

- Looks for a system address space that is an ENQ “waiter” for over 5 seconds
- Lists both waiter and blocker
- Equivalent to D GRS,AN,WAITER

```
f hzr,analyze
HZR02001 RUNTIME DIAGNOSTICS RESULT 581
SUMMARY: SUCCESS
 REQ: 004 TARGET SYSTEM: SY1      HOME: SY1      2010/12/21 - 13:51:32
  INTERVAL:   60 MINUTES
  EVENTS:
   FOUND: 04 - PRIORITIES: HIGH:04  MED:00  LOW:00
   TYPES: HIGHCPU:01
   TYPES: LOOP:01 ENQ:01 LOCK:01
_____________________________________________________________________
EVENT 01: HIGH - ENQ            - SYSTEM: SY1      2010/12/21 - 13:51:32
ENQ WAITER  - ASID:0038 - JOBNAME:IBMUSER2 - SYSTEM:SY1
ENQ BLOCKER - ASID:002F - JOBNAME:IBMUSER1 - SYSTEM:SY1
QNAME: TESTENQ
RNAME: TESTOFAVERYVERYVERYVERYLOOOOOOOOOOOOOOOOOOOOONGRNAME1234567...
  ERROR: ADDRESS SPACES MIGHT BE IN ENQ CONTENTION.
 ACTION: USE YOUR SOFTWARE MONITORS TO INVESTIGATE BLOCKING JOBS AND
 ACTION: ASIDS.
```


These are the areas analyzed by Runtime Diagnostics in z/OS R12

Runtime Diagnostics: CPU Analysis

- Takes two quick samples over 1 second interval
- Any task using > 95% of a single CPU is considered a potential problem
- The usage reported might be > 100% if an address space has multiple TCBs and several are using a high percentage of the capacity of a CPU

```
f hzr,analyze
HZR0200I RUNTIME DIAGNOSTICS RESULT 581
SUMMARY: SUCCESS
 REQ: 004 TARGET SYSTEM: SY1      HOME: SY1      2010/12/21 - 13:51:32
 INTERVAL:  60 MINUTES
 EVENTS:
  FOUND: 04 - PRIORITIES: HIGH:04  MED:00  LOW:00
  TYPES: HIGHCPU:01
  TYPES: LOOP:01 ENQ:01 LOCK:01
-------------------------------------------------------------------------
EVENT 02: HIGH - HIGHCPU      - SYSTEM: SY1      2010/12/21 - 13:51:33
ASID CPU RATE:99%     ASID:002E    JOBNAME:IBMUSERX
STEPNAME:STEP1     PROCSTEP:         JOBID:JOB00045 USERID:IBMUSER
JOBSTART:2010/12/21 - 11:22:51
  ERROR: ADDRESS SPACE USING EXCESSIVE CPU TIME. IT MIGHT BE LOOPING.
 ACTION: USE YOUR SOFTWARE MONITORS TO INVESTIGATE THE ASID.
```


Runtime Diagnostics provides analysis for **Local lock suspension and CPU analysis**:

•**CPU analysis**: a point in time check of any address space that is using more than 95% of the capacity of a single CPU, which might indicate the address space is in a loop. The analysis is a one second sample interval based on the capacity of a single CPU within the LPAR. It is possible for the usage to be reported greater than 100% if the address space has multiple TCBs and several of the TCBs are individually using a high percentage of the capacity of a CPU.

Runtime Diagnostics: Local Lock Suspension

- Lists any address space where its local lock suspension time is over 50%
  - Lock, Cross Memory Local (CML) Lock; CMS locks

```
f hzr,analyze
HZR0200I RUNTIME DIAGNOSTICS RESULT 581
SUMMARY: SUCCESS
 REQ: 004 TARGET SYSTEM: SY1      HOME: SY1      2010/12/21 - 13:51:32
 INTERVAL:  60 MINUTES
 EVENTS:
  FOUND: 04 - PRIORITIES: HIGH:04  MED:00  LOW:00
  TYPES: HIGHCPU:01
  TYPES: LOOP:01 ENQ:01 LOCK:01

EVENT 04: HIGH - LOCK          - SYSTEM: SY1      2010/12/21 - 13:51:33
HIGH LOCAL LOCK SUSPENSION RATE - ASID:000A JOBNAME:WLM
STEPNAME:WLM      PROCSTEP:IEFPROC  JOBID:******** USERID:********
JOBSTART:2010/12/21 - 11:15:08
  ERROR: ADDRESS SPACE HAS HIGH LOCAL LOCK SUSPENSION RATE.
 ACTION: USE YOUR SOFTWARE MONITORS TO INVESTIGATE THE ASID.
```

13

SHARE in Orlando 2011

•**Local lock suspension:** a point in time check of local lock suspension for any address space. For the local lock suspension, Runtime Diagnostics calculates the a mount of time an address space is suspended waiting for the local lock. If an address is suspended more than 50% of the time waiting for a local lock, Runtime Diagnostics reports it as an event.

Runtime Diagnostics: Loop Detection

- Investigates all tasks in all address spaces looking for TCB loops
  - Takes a snapshot of the system trace
  - Looks for consistent, repetitive activity that typically indicates a loop
- When both HIGHCPU and LOOP events occur for the same job, there is a high probability that the task in the job is in a loop.
- Normal, corrective action is to cancel the job.

```
f hzr,analyze
HZR0200I RUNTIME DIAGNOSTICS RESULT 581
SUMMARY: SUCCESS
 REQ: 004 TARGET SYSTEM: SY1      HOME: SY1      2010/12/21 - 13:51:32
 INTERVAL:  60 MINUTES
 EVENTS:
  FOUND: 04  PRIORITIES: HIGH:04  MED:00  LOW:00
  TYPES: HIGHCPU:01
  TYPES: LOOP:01 ENQ:01 LOCK:01
-------------------------------------------------------------------------
EVENT 02: HIGH - HIGHCPU        - SYSTEM: SY1      2010/12/21 - 13:51:33
ASID CPU RATE:99%     ASID:002E    JOBNAME:IBMUSERX
STEPNAME:STEP1    PROCSTEP:         JOBID:JOB00045 USERID:IBMUSER
JOBSTART:2010/12/21 - 11:22:51
 ERROR: ADDRESS SPACE USING EXCESSIVE CPU TIME. IT MIGHT BE LOOPING.
 ACTION: USE YOUR SOFTWARE MONITORS TO INVESTIGATE THE ASID.
-------------------------------------------------------------------------
EVENT 03: HIGH - LOOP           - SYSTEM: SY1      2010/12/21 - 13:51:14
ASID:002E   JOBNAME:IBMUSERX   TCB:004FF1C0
STEPNAME:STEP1    PROCSTEP:         JOBID:JOB00045 USERID:IBMUSER
JOBSTART:2010/12/21 - 11:22:51
 ERROR: ADDRESS SPACE MIGHT BE IN A LOOP.
 ACTION: USE YOUR SOFTWARE MONITORS TO INVESTIGATE THE ASID.
```

14

SHARE in Orlando 2011

Loop detection: Runtime Diagnostics looks through all tasks in all address spaces to determine if a task appears to be looping. Runtime Diagnostics does this by examining various system information for indicators of consistent repetitive activity that typically appears when a task is in a loop. When both a HIGHCPU event and a LOOP event (shown in the example) list the job name, there is a high probability that a task in the job is in a loop. The normal corrective action is to cancel the job name listed.

Runtime Diagnostics: z/OS UNIX File System Latch Contention

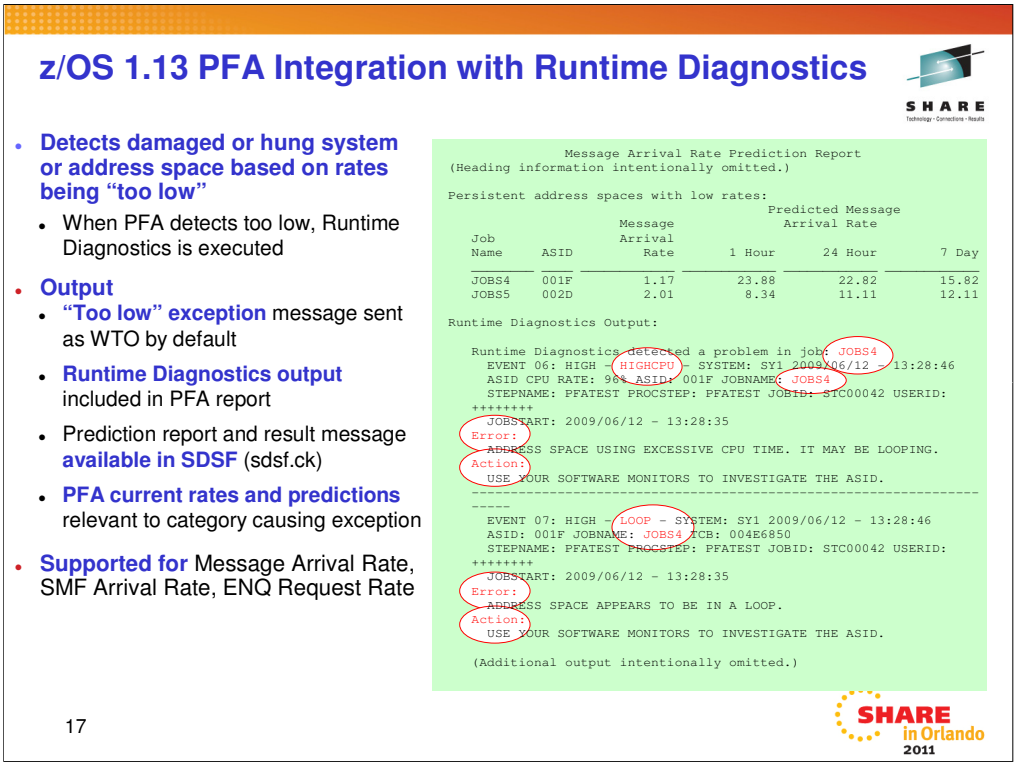
- New in z/OS 1.13
- If z/OS UNIX file system latch contention or waiting threads exist for > 5 minutes in z/OS UNIX, a Runtime Diagnostics OMVS event is created.
- Normal action is to issue D OMVS,W,A to get the ASID and job names of the waiters

```
F HZR,ANALYZE
HZR0200I RUNTIME DIAGNOSTICS RESULT 692
SUMMARY: SUCCESS
 REQ: 009 TARGET SYSTEM: SY1      HOME: SY1      2010/12/21 - 14:24:29
 INTERVAL:  60 MINUTES
 EVENTS:
  FOUND: 02 - PRIORITIES: HIGH:02  MED:00  LOW:00
  TYPES: OMVS:01
  TYPES: LOCK:01
 ------------------------------------------------------------------------
EVENT 01: HIGH - OMVS          - SYSTEM: SY1      2010/12/21 - 14:24:29
ASID:000E - JOBNAME:OMVS
MOUNT LATCH WAITERS: 1
FILE SYSTEM LATCH WAITERS: 0
XSYS AND OTHER THREADS WAITING FOR Z/OS UNIX: 1
  ERROR: z/OS UNIX MIGHT HAVE FILE SYSTEM LATCH CONTENTION.
  ACTION: D OMVS,W,A TO INVESTIGATE z/OS UNIX FILE SYSTEM LATCH
  ACTION: CONTENTION, ACTIVITY AND WAITING THREADS. USE YOUR SOFTWARE
  ACTION: MONITORS TO INVESTIGATE BLOCKING JOBS AND ASIDS.
```

Runtime Diagnostics: GRS Latch Contention

- New in z/OS 1.13
- Obtains latch contention information from GRS
- Omits z/OS UNIX file system latch contention
- Returns the longest waiter for each latch set

```
F HZR,ANALYZE
HZR0200I RUNTIME DIAGNOSTICS RESULT 692
SUMMARY: SUCCESS
 REQ: 002 TARGET SYSTEM: SY1      HOME: SY1      2010/12/21 - 14:32:01
 INTERVAL:  60 MINUTES
 EVENTS:
  FOUND: 02 - PRIORITIES: HIGH:02  MED:00  LOW:00
  TYPES: LATCH:02
_____________________________________________________________________
EVENT 01: HIGH - LATCH        - SYSTEM: SY1      2010/12/21 - 14:32:01
LATCH SET NAME: SYSTEST.LATCH_TESTSET
LATCH NUMBER:3          CASID:0039  CJOBNAME:TSTLATCH
TOP WAITER - ASID:0039 - JOBNAME:TSTLATCH - TCB/WEB:004E2A70
TOP BLOCKER- ASID:0039 - JOBNAME:TSTLATCH - TCB/WEB:004FF028
  ERROR: ADDRESS SPACES MIGHT BE IN LATCH CONTENTION.
 ACTION: D GRS,AN,LATCH,DEP,CASID=0039,LAT=(SYSTEST.L*,3),DET
 ACTION: TO ANALYZE THE LATCH DEPENDENCIES. USE YOUR SOFTWARE
 ACTION: MONITORS TO INVESTIGATE BLOCKING JOBS AND ASIDS.
```



When an exception for an abnormally low condition is found, a health check exception will be issued explaining the problem. The PFA report will include the current rates and predicted rates for the category that was failing. In addition it will include the Runtime Diagnostics output received when PFA called Runtime Diagnostics to verify the problem.

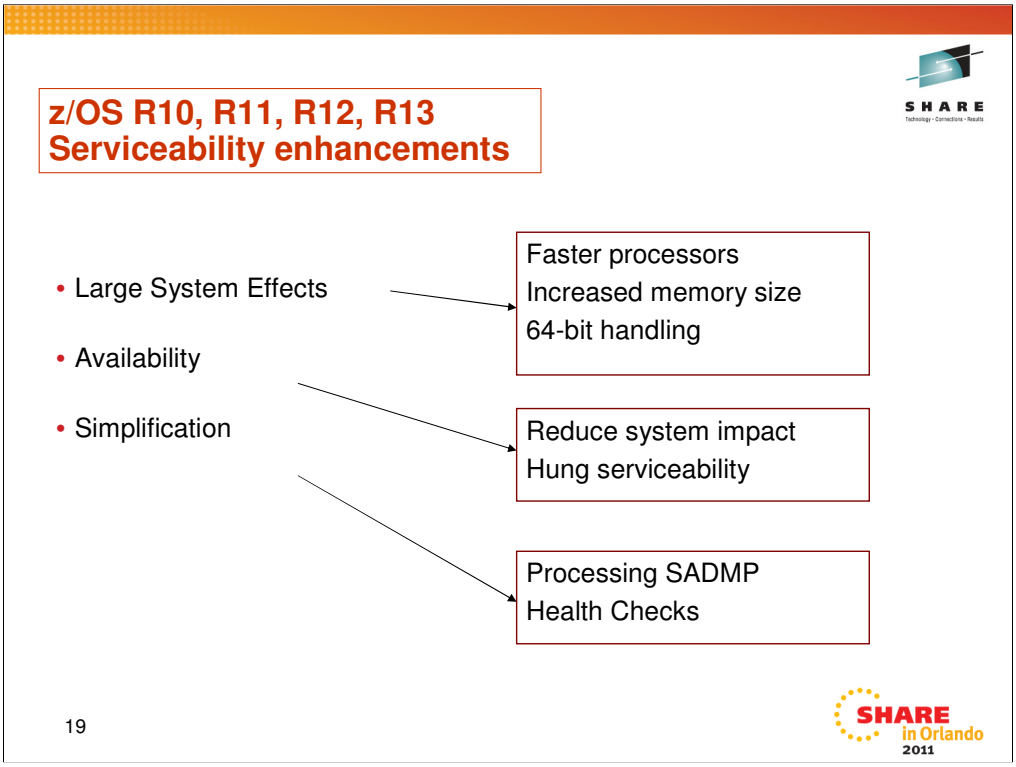
Note that in this example, PFA indicated that jobs JOBS4 and JOBS5 had a Message Arrival Rate that was too low when compared to their expected rates for any of the time ranges. Runtime Diagnostics verified that there could be a problem by detecting both a HIGHCPU and a LOOP event for JOBS4. Therefore, the abnormally low message arrival rate coupled with the results of Runtime Diagnostics show that JOBS4 is very likely looping. The Runtime Diagnostics output for JOBS5 were similar, but were purposely omitted from this display due to lack of space.

Just like the other PFA prediction reports, the PFA prediction reports for abnormally low conditions are available in SDSF.

The combination of PFA & RTD is catching a lot of potential problems in internal testing!!

Runtime Diagnostics Summary

- Helps you analyze a soft failure, diagnose the problem, and take corrective action in a timely manner
- References
  - z/OS Problem Management: G325-2564
  - z/OS Hot Topics Newsletter:
    http://www.ibm.com/systems/z/os/zos/bkserv/hot_topics.html
    - #23 (GA22-7501-19) – *Runtime to the Rescue! Using Runtime Diagnostics to find out your problems fast*
      by Bob Abrams, Don Durand, and Dave Zingaretti

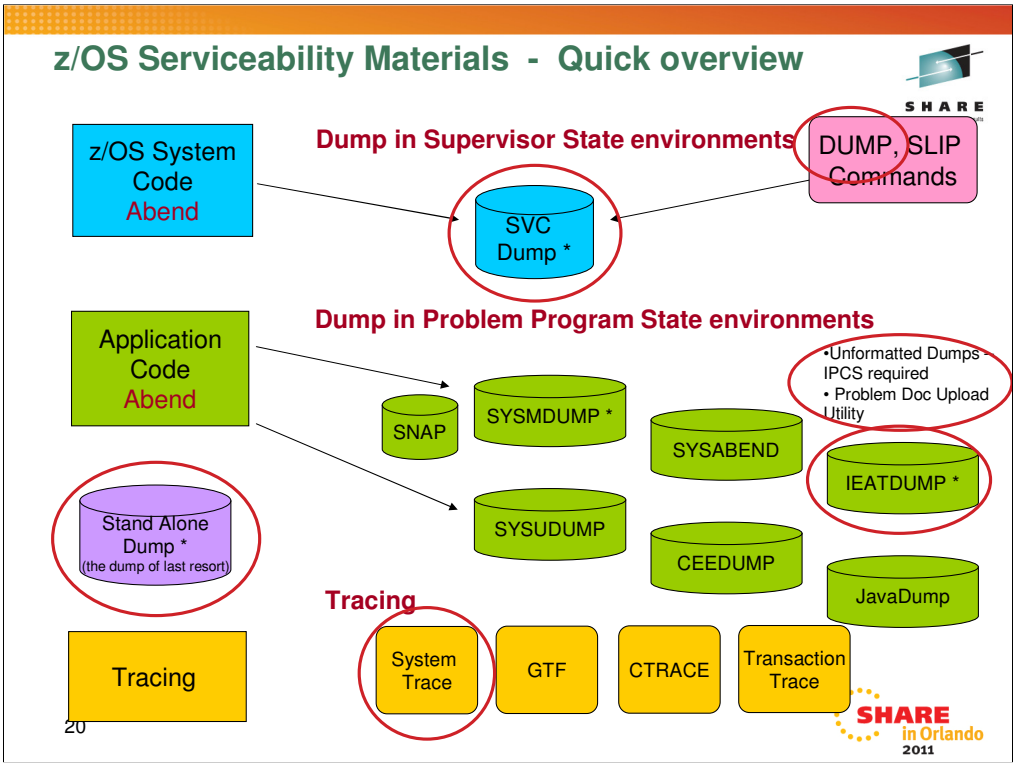


Categories:

•Large system effects

•Availability enhancements

•Simplification



All of Diagnostics Aids at a glance. The circled areas are discussed in the subsequent charts.

Large System Effects

Release	Problem	Solution area
R10 SysTrace	System Trace growth	System Trace buffers move above the Bar (2G) Default/Min trace buffer per processor: 256K → 1M
R12 IPCS SYSTRCE	Make it easier to locate info with more/ faster CPUs	IPCS System Trace formatting - more filtering options on IPCS SYSTRCE cmd SORTCPU: Number of trace entries before & after specified date/time for each CPU Date, Time, N (default=10)
R10 TDUMP	Unable to capture large IEATDUMP for address spaces exploiting 64-bit virtual (e.g., 64-bit Java) … Dynamic data set alloc preferred usage model	2G limit on TDUMP size now removed … dump will span over multiple data sets (dynamically allocated) … &DS in data set names to sequence/correlate
R10 SDUMP	Run out of room dumping storage in ascending order before dumping critical data → WAS 64-bit users with large heaps can end up with dumps where the system and LE stacks needed to debug them are truncated	SDUMP support for 64-bit storage - Priority established for high virtual memory objects - SDUMP honors that priority - Below the bar storage dumped as before - SDATA SQA/CSA/LSQA/RGN include high virtual

21

SHARE in Orlando 2011

System Trace enhancement

Problem: z/OS LPARs and sysplexes are growing in complexity and size; So we needed to increase system trace capacity. System trace buffers are virtualized in LSQA in TRACE address space, limiting the actual space available

Solution: System Trace buffers are moved above the bar (2G)

- The size of each trace buffer remains at 4K. The combined trace structure size of all processors will not exceed half the amount of online real storage
- The default trace buffer size per processor has been increased from 256K to 1M. The minimum trace buffer size per processor is 1M.
- Changes to the trace structure size will be allowed in Megabyte or Gigabyte units only.
- System trace buffers will use 'Large Pages' when the System z10 hardware is available and the 'LFAREA' parameter in 'IEASYSxx' is set.

>2G TDUMPs

Problem: Exploitation of large real and large virtual storage is resulting in larger dumps. Users of TDUMP prefer dynamic dump dataset allocation instead of the task of pre-allocating datasets. TDUMP can only capture up to 2G of data per dump if dynamic dataset allocation is enabled. For address spaces exploiting 64-bit virtual storage, the 2G limit is frequently too restrictive. In 64-bit Java environment, large heaps can result in truncated Tdumps and requirement for problem recreates
Solution: The 2G limit is removed, allowing TDUMPs to scale to whatever size necessary while using dynamic dump dataset allocation, over multiple data sets. The installation must specify their naming patterns using a specific token that allows TDUMP to extend to multiple datasets as necessary. Data captured for dumps >2G will be written immediately to the dump datasets; no caching of dump records in storage will occur as it does (and will continue) with 2G limited TDUMPs.

SDUMP for 64-bit storage

SDUMP dump sizes are on the rise, and with the increase in system capacity and utilization of high virtual storage, will continue to increase. IBM applications such as Websphere or DB2 obtain large ranges of high virtual storage, some of which have contents that are not critical, nor sometimes even necessary, for problem determination efforts. Temporary data structures, data caches are examples of data areas that may not need to be routinely dumped, and/or can safely be sacrificed in lieu of more important areas of storage.

Large System Effects – Smarter SVC Dump processing

- Objectives
  - Never cause an outage taking a dump
  - Capture diagnostic data before it is overwritten (capture it fast enough)
  - Cause minimal performance disruption
    - Due to their memory intensive nature, dumps cannot be processed transparently, but their impact should be mitigated to be just what is essential

- Prior to z/OS V1R11:
  - MAXSPACE defines the maximum amount of virtual storage for DUMPSRV to use
    - CD SET,SDUMP,MAXSPACE=xxxxxxxxM
  - Default is 500M
  - SysProg must estimate the impact upon auxiliary (paging) storage
  - SVCDUMP processing truncates the dump when MAXSPACE is reached or SRM detects that 85% of paging space is used up

- z/OS V1R12:
  - CHNGDUMP AUXMGMT=ON/OFF
  - MAXSNDSP=sss (SDUMP max system non-dispatchability time)
  - AUX utilization continually monitored during SDUMP
    - Detects AUX storage utilization changes more rapidly
    - Improves management of virtual storage when SVC Dump taken
    - Allows a dump to complete if sufficient AUX storage is provided

Large System Effects – Smarter SVC Dump processing …

- AUXMGMT=ON (default)
  - AUX storage utilization 50%: No new dumps are allowed
  - AUX storage utilization 68%: Current dump data capture stops
  - Once the limit is exceeded, new dumps will not be processed until the AUX storage utilization drops below 35%
  - Always honor MAXSPACE when it is more restrictive than AUXMGMT. (i.e. When MAXSPACE=35Meg, stop SVC dumps when MAXSPACE is exceeded even if AUX utilization is only 3%)
- AUXMGMT=OFF
  - SDUMP virtual storage management reverts to control via MAXSPACE
  - Dump in progress is stopped, made Partial, when critical AUX storage shortage (85%) detected or MAXSPACE exceeded
  - After critical AUX storage shortage, AUX storage utilization must be 35% or less before dump capture will resume
  - Must turn AUXMGMT off to get previous behavior
- Maximum system non-dispatchability
  - MAXSNDSP set to 15 seconds by default; can be modified via CHNGDUMP
- SmartCopy
  - If source data on AUX, move it directly to SDUMP buffer
  - Ensure data remains "unreferenced" after the capture
  - SDUMP capture space made to "look old" so it's a top candidate to be paged out
    - Avoid paging out your important data

Large System Effects …

Release	Problem	Solution area
R10 SADMP	Reduce time to initialize SADMP Elongated SADMP transmit time	Export dump directory created by COPYDUMP IPCS COPYDDIR IPCS JOBLIST & EASYCOPY keywords … create copy with base system address spaces
R12 SADMP	Address spaces not in summary list tend to be higher ASIDs and not captured (out of space, stopped the dump)	Stand Alone Dump ASID prioritization - Add addr space names of value to the "summary list" (ADDSUMM: ASIDs, job names) - ANTMAIN, CONSOLE, XCFAS, IOSAS, SMXC, WLM, CATALOG, GRS, SMF , ALLOCAS , ANTAS000, DEVMAN, DUMPSRV, GRS, IEFSCHAS, IXGLOGR, JESXCF, JES2, JES3, OMVS, PCAUTH, RASP, SMSPDSE, SMSPDSE1, SMSVSAM, TRACE
R12 DUMP command	Significant amount of time in addr space non-dispatchability while global dump exits running - on DUMP command, getting a consistent view of Global & addr space storage may not be as critical - *more important to reduce the dump's impact to the system*	Reduce impact of DUMP command dump to system - *Defer setting tasks non-dispatchable* on DUMP command - DEFERTND option … on CHNGDUMP & DUMP command - Delay setting addr space non-dispatchable until after global capture completed

Stand Alone Dump ASID prioritization

It's possible to have an address space that is not part of the default list and has higher numeric value but is more important to customer system

Some important address spaces may not be dumped or are truncated because the installation runs out of space during dump capture which will cause difficulty in problem diagnosis

Solution:

Provide an ability for customers to add address space(s) they value more to an enhanced summary address space list and

Enhanced the existing summary address space list by using the COPYDUMP address space list

Allows important address spaces with high ASIDs to be dumped earlier and decrease the chance to not be dumped or truncated

Defer setting tasks non-dispatchable on DUMP command

A new option DEFERTND=(Yes/No) was provided on CHNGDUMP and DUMP command REPLY

CD SET,SDUMP,DEFERTND=YES

Id IEE094D SPECIFY OPERAND(S) FOR DUMP COMMAND

R id,DEFERTND=Yes/NO,…

The new option is also supported through parmlib member IEADMCxx

Allows installation to specify whether SDUMP processing should delay the setting of the tasks non-dispatchable in the address spaces being dumped until after the global capture completed

IPCS COPYDUMP enhancement

Problem: The growth rate of the largest z/OS dumps is outpacing the rate at which bandwidth is allowing such dumps to be transmitted. As a result, it is taking longer to initialize dumps, delaying the start of dump analysis and resolution of the problems.

Solution: IPCS COPYDUMP enhancement to create subset dump (JOBLIST & EASYCOPY keywords). COPYDUMP has been changed to perform dump initialization while it is transcribing the dump.

COPYDDIR supports: EXPORT dump directory records pertaining to one source; IMPORT dump directory records generated by the EXPORT function into the dump directory for the current session.

IPCS DOCPU command – Obtain SADMP data for multiple CPUs (R13)

- Obtain CPU-related data from SADMP with 1 command rather than repeat command for each CPU
  - Works only for Stand Alone Dumps

- Command Syntax:

```
{ DOCPU }
        [ ( CPU ( cpu-address-range-list ) ) |
          CPUTYPE ((ZAAP|ZA)  | (ZIIP|ZI)  | (STANDARD | CP | S) )  |
          CPUMASK ( cpumask ) ]
          EXEC ((ipcs-subcommand))
```

- Flexible CPU range specification:
  - CPU(0)
  - CPU(5:10)
  - CPU(0 5:10)
  - CPU(0,3,5:10)
  - CPU(X'A')
  - CPUMASK(FFF)
  - CPUMASK(F0F0)
  - CPUMASK(80)

Display 4 bytes of storage at 414 for CPU 0 and CPU 1:

```
DOCPU CPU(0,1) EXEC((L 414 LEN(4)))
  CPU(X'00'):
LIST 0414.CPU(X'00') ASID(X'0001') LENGTH(X'04') AREA
00000414.04454000
  CPU(X'01'):
   LIST 0414.CPU(X'01') ASID(X'0001') LENGTH(X'04') AREA
00000414.027EF000
```

25

Use the DOCPU subcommand to gather stand-alone dump data for tasks that need to be repeated for each of the specified processors.

For example, to display contents of a processor-related control block for a group of processors.

With this command, you can obtain processor-related diagnostic data from a stand-alone dump with one command rather than repeating the command for each processor.

IPCS SYSTRACE: CPUTYPE & CPUMASK (R13)

- CPUTYPE and CPUMASK are also supported on IPCS SYSTRACE command
- Limits formatting of System Trace info only to trace entries produced on the specified processors
  - CPUTYPE ((ZAAP|ZA) | (ZIIP|ZI) | (STANDARD | CP | S))
  - CPUMASK (cpumask)

1. To show all data for processors from 0 to 11
   - SYSTRACE ALL CPUMASK(FFF)
2. To show all data for processors from 0 to 3 and from 8 to 11:
   - SYSTRACE ALL CPUMASK(F0F0)
3. To show data for processors from 0 to 127:
   - SYSTRACE CPUMASK(FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF)
4. To show all data for processors 2, 3, 5, 8, 9, 10, 11:
   - SYSTRACE ALL CPUMASK(34F)
5. To show all data for ZAAP and ZIIP processors, for processor 0, 2, 5, 7, and for processor from 8 to 11:
   - SYSTRACE ALL CPUTYPE(ZA ZIIP) CPU(0,2,5,7) CPUMASK(00F)

SDUMP Availability

Release	Problem	Solution area
R10	Only 1 SDUMP capture can be in progress at a time - SDUMP could not be taken because system indicates a dump is in progress, but there's no dump in progress. *- Current dump collection may not complete, but future dumps are put on hold indefinitely* - Canceling DUMPSRV results in lost captured dumps - You don't know SDUMP is locked until a dump is needed and cannot be taken	Hung SDUMP Detection - System detects DUMPSRV hang situation - Uses IEATDUMP to dump DUMPSRV and all its data spaces - Issues message, warning installation that DUMPSRV can be recycled

Hung SDUMP Detection

Solution:

System detects the hung situation and takes a TDUMP from DUMPSRV address space

TDUMP will include system data SDATA=(NUC,CSA,LPA,RGN,TRT,GRSQ,SQA) and DUMPSRV data spaces

Message "IEA044E Dumping Services Function is Unavailable" is issued

Benefit:

Improved communication of dumping services availability so it can be recycled by the customer

TDUMP of DUMPSRV can help diagnose the situation that led to the failed dump or improper lock condition

Data that has been captured in DUMPSRV data spaces will not be lost

Simplification: Configuration / Usability

Release	Problem	Solution area
R10	Confusion over what needs to be extracted from Stand Alone Dump	Simplified running COPYDUMP - EasyCopy - Defaults set, can be overridden on panel
R11	Configuration gotchas	Health Checks - Determine if AUTOIPL used in GDPS environment - Determine if AUTOIPL is configured - Validate devices specified for SADMP & MVS AUTOIPL - DAE data set configuration
R12	Run IPCS, *save results to pass to other analysts*, reduce overall time spent on problem	Extract & store important info in PDS - Stand Alone Dump or SDUMP - Allocate PDS to IPCSPDS DD - SETDEF PDS … SETDEF NOPDS - Output of each IPCS subcommand (or REXX exec) written to separate member - Supports IPCS subcommand and REXX execs (not CLISTs)

Problem Documentation Upload Utility – AMAPDUPL (R13)

- Facilitates sending large amounts of documentation in an efficient manner
  - Drives multiple simultaneous FTP sessions, allowing higher utilization of your networking infrastructure
  - Shorter transmission time for very large data sets
  - No need to use AMATERSE or TRSMAIN to compress the input data set
- Formerly a Service download … delivered in R13
  - MTFTPS alias maintained for compatibility

- Encapsulates entire process – single job step
- Hardware compression and encryption
  - Always compresses
  - Industry standard encryption (192 bit 3-DES) optional
    - Make sure you update PMR with ENCRKEY=xxx or CIPHERKEY=xxx
    - Uses a built-in encryption instruction, available on all processors starting with the z990 and z890
- Maintains dataset characteristics
- Enforces PMR file naming convention
- Performance improvements
  - Parallel FTP sessions – 20 max; 5 recommended

Problem Documentation Upload Utility (PDUU) …

- AMAPDUPL resides in SYS1.MIGLIB … STEPLIB DDNAME not needed
  - Was necessary with MTFTPS tool
- Data set types supported:
  - *Members* of partitioned data sets (PDS) and partitioned data sets extended (PDSE)
  - Large format (DSNTYPE=LARGE) and traditional sequential data sets
  - Extended format sequential data sets
  - Fixed and variable, blocked and unblocked, unspanned record formats (RECFM = F,FB,FBS,V,VB)
  - Data sets with records containing ISO/ANSI or machine code control characters
  - Data sets in cylinder-managed space
- PDUU does **not** support following types of input data sets:
  - Large block interface (LBI) (no BLKSIZE value)
  - VSAM and direct (DSORG=DA) data sets
  - Data sets with keys (KEYLEN)
  - Spanned record formats (VBS)
  - Partitioned data sets (PDS) and partitioned data sets extended (PDSE)
  - z/OS UNIX files
  - Any data set with an undefined-length record format (RECFM=U)

Summary: z/OS Service Aids update

- z/OS Serviceability keeping up with new environments
- Significant set of enhancements, R10-R12
  - Large System Effects
  - Availability
  - Simplification
  - SVC Dump, Stand Alone Dump, System Trace, PDUU, IPCS, …
- References
  - z/OS IPCS Commands, SA22-7594
  - z/OS Diagnostics: Tools & Service Aids, GA22-7589
  - z/OS Authorized Assembler Services Guide, SA22-7608
- Runtime Diagnostics (R12, R13)
  - Helps you analyze a soft failure, diagnose the problem, and take corrective action in a timely manner

Backup

PDUU invocation – DD statements

SYSPRINT

Specifies the job output data set

The data set can be either SYSOUT or a sequential data set

Must be RECFM=FB,LRECL=134

SYSUT1

Specifies the sequential input data set to transfer to IBM

SYSIN

Specifies the sequential data set that uses following control statements. (The data set must be RECFM=FB,LRECL=80)

TARGET_SYS – The name of the TCP/IP system to transfer the files to using FTP

USERID – The user ID on the target system that is used to send the files

PASSWORD – The password for the USERID on the target system

ACCOUNT – The account data that is sent when an FTP session is started.

WORK_DSN – The prefix for the data set names of work files on the sending system

WORK_DSN_SIZE – The size of the work files in megabytes

KEEP_WORK – The parameter to save the work data sets that are dynamically allocated for each FTP session.

DATACLAS – The data class to use when allocating the work files on the sending system

STORCLAS – The storage class to use when allocating the work files on the sending system

CC_FTP – The number of parallel FTP sessions to use when transmitting the files

DIRECTORY – The directory on the target system where the files will be sent with FTP

PMR – The PMR number that this file is to be associated with

CIPHER_KEY – The encryption key to use for 192-bit triple DES encription

FTPCMDS

An optional DD statement that provides additional flexibility for traversing firewall or proxy servers

PDUU invocation – JCL

```
//FTP EXEC PGM=AMAPDUPL
//SYSPRINT DD SYSOUT=*
//SYSUT1 DD DISP=SHR,DSN=IPCS.PROBLEM.DUMP
//SYSIN DD*
USERID=anonymous
PASSWORD=anonymous
TARGET_SYS=testcase.boulder.ibm.com
TARGET_DSN=wessamp.bigfile
WORK_DSN=wes.ftpout
CC_FTP=03
WORK_DSN_SIZE=50
DIRECTORY=/toibm/mvs/
PMR=12345.123.123
//
```

Large System Effects (continued)

See “Smarter SVC Dump processing” (next chart)

Release	Problem	Solution area
R11 SDUMP	SDUMP system non-dispatchability running too long (enough to partition system from sysplex!!)	Limit SDUMP non-dispatchability with huge dumps - MAXSNDSP (default = 15 seconds) on CHNGDUMP command
R11 SDUMP	Running out of AUX space when taking SVC Dumps … MAXSPACE not sufficient - prevent WAIT 03C (paging space exhausted)	Allow system to determine if sufficient AUX available! SDUMP continually monitors AUX usage during dump. - AUXMGMT=ON (on CHNGDUMP cmd) - alternative to MAXSPACE (honored when specified) - Makes Availability higher priority than FFDC
R12 SDUMP/ RSM	Long SDUMP AUX capture times	RSM “smartcopy”, moving AUX frames directly to DUMPSRV; **Dramatic** SDUMP performance improvement when source data on AUX!

35

SHARE in Orlando 2011

Limiting SDUMP non-dispatchability

System may stay non-dispatchable long enough to result in the system being partitioned from the sysplex

Taking down or inhibiting the functions of customer’s system in order to take an svc Dump is certainly not desirable

MAXSNDSP (Maximum System Non-Dispatchability) is supported on CHNGDUMP

AUXMGMT for SVC Dumps

Prior to z/OS V1R11: MAXSPACE defines the maximum amount of virtual storage for DUMPSRV to use [CD SET,SDUMP,MAXSPACE=xxxxxxxxM], where the default is 500M.

You must estimate the impact upon auxiliary (paging) storage. Meanwhile, SDUMP processing truncates the dump when SRM detects that 85% of paging space is used up

This causes a large exposure that dumping could cause a WAIT state 03C RSN01 (paging space exhausted).

Solution:

SDUMP continually monitors AUX utilization during the dumping process. AUXMGMT=ON.

Aux Monitoring is improved to detect AUX storage utilization changes more rapidly.

Improve the management of virtual storage when an SVC DUMP is taken.

Allows a dump to complete if the customer has provided sufficient AUX storage.