

DFSMS Advanced: PDSE Diagnostics and Recovery

Speaker: Thomas Reed

IBM Corporation

Session: 16958





SHARE is an independent volunteer-run information technology association that provides education, professional networking and industry influence.







Agenda

- How PDSE externalizes problems.
- PDSE's had a problem, now what?
- Identifying and solving the problem:
 - Tools
 - -0F4's
 - Hangs
 - Corruption





What is a PDSE?

- PDSE: <u>Partitioned DataSet Extended</u>
- A PDSE is a collection of directory and data pages
- At V2R1 there are 2 dataset formats V1 and V2 PDSEs
- PDSE server consists of one or two address spaces (SMSPDSE and SMSPDSE1)
- The SMSPDSE(1) address spaces serve client access requests for PDSE datasets
- Under the hood SMSPDSE(1) also manages PDSE serialization and buffering





How PDSE Externalizes Problems

- ABENDS
 - 0F4
 - -0Cx
- IGW messages
 - IGW038A/IGW031I
 - IGW702I
 - IGW007E/IGW077E
- CSV031I





0F4 Abends

- Indicate a PDSE "Logic Error"
 - In other words something unexpected happened
- RC (Reg 15) is not very useful
- RSN (Reg 0) tells us where the error occurred and what the error was
 - Errors can percolate up the calling chain
 - Many error codes may be issued from the same place in code





Reading the 0F4 Dump Title

- CompID
 - DF115 = PDSE
 - DF104 = FAMS
- CSECT + offset
- PTF number
- ABEND code
- Return code
- Reason code

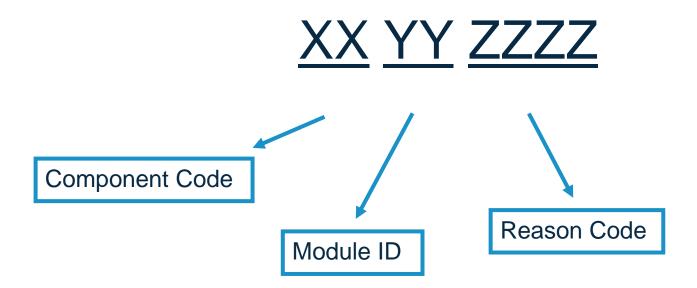
COMPID=DF115, CSECT=IGWDAV00+10C4, MAINT ID=UA69488, ABND=0F4, RC=00000024, RSN=01188022





Reason Code Translation

Reason Code Format:







Reason Code Translation Example

RSN: 150BC008

This error generally the result PDSE corruption due to being shared outside the SYSPLEX

How it's broken down:

15 = SubComp = PDSE BMF IO Control

0B = Module ID = IGWBIEX1

C008 = Reason = RSNS_IO_ERROR_DIRECTORY





IGW Messages and Meanings

- IGW038A and IGW031I
 - Indicates a PDSE latch or lock has been held for longer than the PDSE Monitor's set duration
- IGW702I
 - IEBPDSE (The PDSE Validation Tool) completion message
 - Indicates that IEBPDSE did not complete RC 0
 - IGW700I is the successful completion message





IGW Messages and Meanings

VOLSER:XPR10F

RC:8 RS:01188011 R14:849B0CD0

RPN: 27904 VPTVFN: 1





IGW Messages and Meanings

IGW700I PDSE Directory Validation Successful

DSN:TREED.PDSE.GOOD

ADPages:462 IXRecords:23606

ADPagesInCore: 3 ADPagesRead: 459

ADTreeLevels:3

NDPages:104 IXRecords:9438

NDPagesInCore:1 NDPagesRead:103

NDTreeLevels:2

AD ND Tree Nodes:9438

ADPercentFree:41 NDPercentFree:36

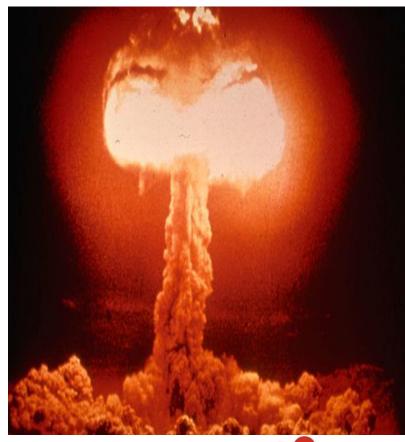
Version:1





IGW Messages and Meanings Cont'd

- IGW007E and IGW077E
 - SMSPDSE and SMSPDSE1 address space failures respectively
 - IGW007E will require an IPL to recover
 - IGW077E may be recoverable via the PDSE1,ACTIVATE command







IGW Messages and Meanings Cont'd

CSV031I

LIBRARY {SEARCH | ACCESS} FAILED FOR MODULE *mod*, RETURN CODE *xx*, REASON CODE *reason-code*, DDNAME *ddname*

- Not a PDSE message
- RC and RSN code are percolated from PDSE
- A "canary" for PDSE corruption







OCx ABENDs

- 0C4 ABENDs
 - Generally due to a bad pointer
 - Variety of causes
 - Expect an APAR
- 0C1 ABENDs
 - Very rare
 - Some 0C1's are diagnostic branches to a 0h





So, PDSE has just had a problem...



DON'T PANIC



Step 1: Identify the Problem

- Start narrow by defining the symptoms
- What are we experiencing?
 - A 0F4 dump?
 - Many 0F4 dumps?
 - IGW messages?
 - Processing hangs?
 - Dataset corruption?
 - CPU/Storage consumption?





Step 2: Gather Doc!

- Initial PDSE doc gathering comes down to 4 items
 - 1. An SVCDump including the PDSE[1] address space
 - 2. SYSLOG/OPERLOG
 - 3. LOGREC/RAW EREP
 - 4. DSS physical dump of the affected dataset





Step 2 Cont'd

- For dumps, gather:
 - The SVC Dump (for multiple, start with the first)
 - SYSLOG and LOGREC covering the incident
- For hangs, gather:
 - A console dump of the PDSE[1] address space
 - SYSLOG and LOGREC starting prior to the incident
- For corruption, gather:
 - The SVC Dump (for multiple, start with the first)
 - SYSLOG and LOGREC starting prior to the incident
 - A DSS Physical dump of the dataset

*See APPENDIX for JCL and Parameters





Step 2 Cont'd

- DO NOT SUPPRESS 0F4 Dumps
- DO NOT SUPPRESS IGW038A messages







Step 3: Call L2

Shameless L2 plug!







Solving the Problem





Tools of the Trade

- IPCS
 - IP ST (IP STATUS)
 - SMSXDATA
 - Q
 - AS
 - MSGS
 - POOLS
- IEBPDSE PDSE Validation Tool





IP STATUS

- Gives a quick overview of the error
- For PDSE 0F4s the most important are the registers
 - Return Code in REG 15
 - Reason Code in REG 0





VERBX SMSXDATA 'f(func) jobname(job) comp(comp)'

- 'comp' stands for component within PDSE
 - Many components with reports
 - CLM contains the most important
- 'f' stands for function
 - Q = Quick view of error TCB
 - AS = Address spaces summary





VERBX SMSXDATA 'f(Q) jobname(job) comp(CLM)'

```
FSMS verbexit processing
Title:COMPID=DF115.CSECT=IGWLHA10+244E.DATE=02/20/12.MAINTID=UA64274 .ABND=0F4.RC=00000024.RSN=13F2FE1E
SN::
unction=Q
ump compatible with IGWFPMAN version:HDZ1C10
 Primary Asid:0009 Secondary Asid:0009 PSATOLD:008FAE88
 Looking for SSF thread by Register 13:7B9261E0 FVSA:7B923040 Asid:0009 SMSPDSE1
 FVSP:7B926190 IGWFCPER <<<<iinput FVSP
 FVSP:7B924EE0 IGWLHA10
 FVSP:7B924438 IGWLHA00
 Displaying Recovery FVSA associated sequences
 FVSA:73C26040 Asid:0009 SMSPDSE1
 FVSP:73C284B8 IGWFDDMP
  FVSP:73C27A80 IGWFCLRR
 FVSP:73C27828 IGWFERTR
 FVSP:73C27438 IGWFEFRR
 Looking for SSF thread by Control Reg15:7F6E3388 Linkage Stack
 Tcbs Description---
        Tcb: 0
```





VERBX SMSXDATA 'f(AS) jobname(SMSPDSE[1]) comp(CLM)'

	Latch Name	Holder	Waiter(s)
	IGWASRB(7FF93380)+18	(81)0032:008BD4E0	(76)0032:0089D030 (84)0032:008BDBE8
			(82)0032:008BD7E8
			(77)0032:0089D1C8
			(79)0032:008AE380
	Commercial products personally products.		(78)0032:008AE0F0
	IGWHTAB(7FF9F000)+38	(41)8009:008F0900	(16)0009:008F9E00
			(13)0009:008FF230
2	TOURTOD/ZEEOEOOO\.40	(41)0000,00050000	(595)0200:008FF260
•	IGWHTAB(7FF9F000)+48	(41)0009:008F0900	(14)0009:008FA1D0 (81)0032:008BD4E0
			(34)0009:008F6CF0
			(33)0009:008F6E88
			(32)0009:008F7190
			(61)0009:008EA3B0
			(26)0009:008F8360
	IGWHTAB(7FF9F000)+58	(41)0009:008F0900	(1530)02BE:008CA7B
			(1242)0257:008FF26
	IGWHTAB(7FF9F000)+60	(41)0009:008F0900	(1295)0270:008FF26
	IGWHTAB(7FF9F000)+70	(41)0009:008F0900	(1243)0259:008CA48



Comp



VERBX SMSXDATA 'f(POOLS) jobname(SMSPDSE[1])'

Max to the bottom for the important parts

	Extents	Total_	Size_of_the_Pool	_TotCells	InUseCells	Fixed	FProt	Key	Description
***	*****	k******	******	******	*******	****	****	***	*************
	45		47,185,920	245745	186334	NO	YES	50	IGWHL3B STORAGE POOL
	4		4.194.304	20164	4412	NO	YES	50	IGWAJB STORAGE POOL
	44		46.137.344	262108	197644	NO.	YES		IGWLRE STORAGE POOL
	400		4 104 204	400	112	NO	VEC	E0	TABLE IN AN
	33		2,214,592,512	540672	521	NO	YES	50	DEP BMF PAGE BUFFER POOL
								30	
	24		25,165,824	524280	521	NO	YES	20	DFP BMF BCB POOL
	41		42,001,010	200013	101040	HU.	150	30	CONTROL FOR TON TON
	59		61,865,984	38645	198	NO	NO	50	JCDM DSC POOL FOR FIB
	326		341.835.776	213530	186129	NO	NO	50	JCDM DSC POOL FOR FIB DREFD

- TotCells = Allocated storage in cells
- InUseCells = In use storage in cells





VERBX SMSXDATA 'f(MSGS) jobname(SMSPDSE[1]) comp(SSF)'

- Formats the PDSE address space's message buffer
- IGW031I messages are recorded in the buffer regardless of whether the ANALYSIS command is issued
- Newest messages at the bottom





IEBPDSE

- The PDSE validation tool
 - 1.13 and above
 - Use the highest available version if possible
 - 1.13 and 2.1 versions are NOT identical
 Do not expect identical results
 - Works on both V1 and V2 datasets





0F4 ABENDs



Step 1: Getting the reason code

IP ST [REGS]

```
CPU STATUS:
PSW=07540000 80000000 00000000 06ACAAFA
    (Running in PRIMARY, key 5, AMODE 31, DAT ON, SUPERVISOR STATE)
   Disabled for PER
  ASID(X'0009') 06ACAAFA. IGWDLACO+BAFA IN EXTENDED PLPA
 ASCB68 at FC0780, JOB(GUDB2035), for the home ASID
 ASXB68 at BFD860 and TCB68E at B99E88 for the home ASID
 HOME ASID: 0044 PRIMARY ASID: 0009 SECONDARY ASID: 0044
 General purpose egister values
                             000000000_040F4000
          000000000_160D5476
          000000000160D5476
                              00000000_000000020
          000000000...000000000
                              000000000 00002DFC
          000000000_000000000
                              000000000_000000000
          000000000_00FE0CF0
                              000000000_06ACC
          00000048_00FE0C80
                              00000000 06ACB26F
    12-13 000000000_06ACA270
                              000000000_7EB510B8
    14-15 000000000_86ACAAF0
                              000000000_000000020
```

In this case: JCDM_NO_LSSM_RECORD





Step 2: Get the calling sequence

```
Primary Asid:0009 Secondary Asid:0044 PSATOLD:00B99E88
ooking for SSF thread by Register 13:7EB510B8
FVSA:7F180040 Asid:0009 SMSPDSE1
                       <<<<iirbut FVSP
    FDES:7F185BD8 DUB:7EA107F0
         FileToken:01-S28000-000104-00000000003-00000000AA6E-0000
FVSP:7F185260
                                                           Requested function
  CDM Destroy File:
                  DirConnToken: 000000007EA107F0 Defer: 1 DestroyAll: 0 KEEP_PRIMARY_NAME: 0
                     CreatePendingDelete:0
                   D:01-S28000-000104-000000000003-000000000AA6E
                                 Httb:7FF6AB30 DSN: Sys1.samp.pdse
      Dub: 7EA107F0 Dib: 7F384720
                                                                The dataset we're
                                                                connecting to
                                  Vsqt:01-S28000-000104
                  Intent: OUTPUT
           of IGWDBIP1 running
                                                               Connecting for output
Displaying Recovery FVSA associated sequences
```



Step 3: Now What!?

- Open a PMR, a 0F4 means something's gone awry and L2 should look at it
- Repeated 0F4's on access of the same dataset can indicate corruption





Hangs

PDSE Resource Serialization Issues





Step 1: IGW038A Messages

 IGW038A messages indicate that there is a PDSE resource that has been held longer than the PDSE monitor's set duration time

IGW038A Possible PDSE Problem(s) (SMSPDSE)
Recommend issuing V SMS, PDSE, ANALYSIS

 Issuing the suggested ANALYSIS command produces an IGW031I report





Step 1: IGW031I Messages

```
IGW031I PDSE ANALYSIS Start of Report(SMSPDSE )
++ Unable to latch ASRBULCH:00000007F04DE00
  Latch:00000007F04DE18 Holder(0008:009F6CF0)
  Holding Started Task: SMSPDSE
++ Unable to latch ASRBULCH:00000007F050560
  Latch:00000007F050578 Holder(0008:009F75D0)
  Holding Started Task: SMSPDSE
++ Unable to latch ASRBULCH:00000007F050640
  Latch:00000007F050658 Holder(0008:009F6AD0)
  Holding Started Task:SMSPDSE
++ Unable to latch ASRBULCH:00000007F0503A0
   Latch:00000007F0503B8 Holder(00E5:009AEE88)
  Holding Started Task:TREED12
-----data set name-----
                                           01-SYSQ01-000109
SYS2.SOMETHING.LOADLIB
++ Unable to latch HL1b:00000007FF62E40
  Latch:00000007FF62E50 Holder(0008:009F6CF0) IGWLHPRG
  Holding Started Task: SMSPDSE
PDSE ANALYSIS End of Report(SMSPDSE)
```

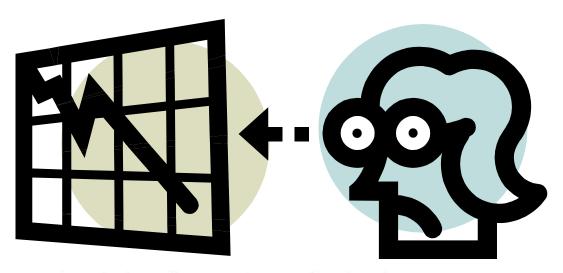
- ANALYSIS examines both LATCHes and LOCKs
- Global LOCKs serialize between systems
- Local LOCKs serialize within a system
 - LATCHes serialize internal structures within a system





Step 1: Reading the Trends

- The PDSE monitor runs every 90 seconds by default
 - A single IGW038A message is OK
 - Can be tripped in a busy system
 - Repeating messages every ~2min generally indicates a problem







Step 2: Get a Dump

- Using the V SMS,PDSE[1],MONITOR,DUMPNEXT makes getting a good dump easy
- Alternately taking a console dump of both PDSE address spaces works well
 - Make sure to get the involved PDSE ASID
- If there are unresponded messages to a system
 - Take a console dump from that system as well
 - Again, make sure to get the involved PDSE ASID





Step 3: Finding the Root of the Hang

- Do we have a real hang?
 - Is it a single holder?
 - Are holders changing?
 - Are the IGW038A message ongoing?
- Changing holders can indicate processing delays but not necessarily a real hang
 - Check CPU utilization
 - CPU starvation can cause delays with jobs releasing PDSE resources







- Finding the key latch
- IGW031I report is not what we need....
- AS report to the rescue!

```
Latch Caller(s)
     IGWASRB(7F04DE00)+18
                               (35)0008:009F6CF0
Tcb Latch Summaru
                                                        AWaited Latch(es)
                              Held Latch(es)
                                                         IGWASRB(7F04DE00)+18
34.
      0008:009F6E88
35.
      0008:009F6CF0
                              IGWASRB(7F04DE00)+18
      0008:009F6CF0
                              IGWASRB(7F04DE00)+18
                              IGWHL1B(7FF62E40)+10
      00E5:009AEE88
                                                         IGWHL1B(7FF62E40)+10
```





Step 5: Reading the AS Latch Report

- Track down the originating latch
- Note the TCB and ASID

```
atch Summaru
                                                      Waiter(s)
                                                                               Latch Caller(s)
                             (35)0008:009F6CF0
    IGWASRB(7F04DE00)+18
Tcb Latch Summaru
*********
                            Held Latch(es)
                                                      AWaited Latch(es)
     0008:009F6E88
                                                      IGWASRB(7F04DE00)+18
     0008:009F6CF0
                             IGWASRB(7F04DE00)+18
     0008:009F6CF0
                             IGWASRB(7F04DE00)+18
                             IGWHL1B(7FF62E40)+10
     00E5:009AEE88
                                                      IGWHL1B(7FF62E40)+10
```

HL1B latches are ALWAYS suspicious





Step 6: Learn about the holder

- Is it still alive?
 - Orphaned latches have a holder that's gone
 - Live holders are generally waiting
 - Check for IGWLHSUS (SSF Suspend)
- Is it in the PDSE address space?
 - This can indicate more serious issues
 - Can lead to a queue of holders obtaining and hanging on the latch
 - Call L2





Step 7: Make sure the holder's dead

For holders NOT in the PDSE address space

- If the holder is dead: Good!
- If the holder is alive:
 - Get it out of the system
 - Cancel is the preferred option
 - Force if necessary
 - Recovery MAY resolve the latch on it's own





Step 8: FREELATCH

V SMS,PDSE|PDSE1,FREELATCH(< latch address>, asid,tcb)

- Don't forget to add the latch offset!!
- WARNING: The holder MUST be out of the system





The Aftermath

- Clearing resources may take a few minutes, wait for the next IGW038A message
- No Change:
 - Holder changed but still hung
 - PDSE1 restart (if applicable)
- Partial Clearing:
 - Get a new dump, repeat the process
 - Tasks may take 0F4's if a latch is freed out from under them
- Fully Cleared
 - PARTY!





Corruption

In-core and on DASD corruption





In-core or on DASD?

- In-Core Corruption
 - If the PDSE is accessable from other systems in the PLEX the corruption is in-core
 - Issue is in the PDSE index cache and the cache is local to each PDSE ASID
- Corrupt on DASD
 - Will fail on all sharing systems
 - IEBPDSE will fail to run RC 0





In-Core Corruption

- For either PDSE or PDSE1:
 - V SMS,PDSE|PDSE1,REFRESH,DSN(dsname)[,VOL(volser)]
 - Throws out in-core index and re-loads from DASD
 - 1.13 and above
- PDSE1 only
 V SMS,PDSE1,RESTART
 [,QUIESCE(duration | 5)[,COMMONPOOLS(NEW|REUSE)]
 - PDSE1 restart will re-load ALL index pages from DASD and reconnect





Corrupt on DASD

- Gather a PHY Dump of the dataset
 - See appendix for JCL
 - L2 has formatting tools to help determine how the dataset broke
- An IEBCOPY may correct minor index issues
- Recovering from the last backup is often the only way to recover





Questions? Comments?





Please Fill Out the Survey!







Appendix

Parameters, Commands and JCL





SHARE is an independent volunteer-run information technology association that provides education, professional networking and industry influence.

Insert Custom Session QR if Desired.





Appendix: Parameters, Commands and JCL

PDSE Console Dump Parameters

```
COMM=(PDSE PROBLEM)

JOBNAME=(*MASTER*, SMSPDSE*),

SDATA=(PSA, CSA, SQA, GRSQ, LPA, LSQA, RGN, SUM, SWA, TRT, COUPLE, XESDATA), END
```

- IGDSMSxx Parameters:
 - SMSPDSE1 restartable address space:

```
PDSE_RESTARTABLE_AS(NO | YES)
```

– PDSE Sharing Modes:

PDSESHARING(EXTENDED|NORMAL)





Appendix: Parameters, Commands and JCL

IEBPDSE JCL (1.13 and above only)

```
//VALIDATE EXEC PGM=IEBPDSE
//SYSPRINT DD SYSOUT=*
//SYSIN DD DUMMY
//SYSLIB DD DISP=SHR, DSN=INPUT.PDSE.BAD
```



Appendix:

SHARE,

Parameters, Commands and JCL

DSS PHYSICAL dump JCL

```
//DUMP
           EXEC PGM=ADRDSSU
//SYSPRINT DD SYSOUT=*
//OUT
           DD UNIT=3390,
               VOL=SER=XXXXXX,
               DISP=(NEW, KEEP),
               SPACE=(CYL, (100, 100)),
               DSN=hilev.DSSDUMP,
               DCB=BLKSIZE=32760
/SYSIN
          DD
      PIDY(vvvvv) -
DUMP
    OUTDD(OUT)
     DATASET(INCLUDE(pdse.dataset.name)) -
     ALLDATA(
*
```





Checking the Error Information

- Check Reg 0 for the reason Code
- Check Reg 15 for the return code

Time of Error Information

PSW: 07541000 80000000 00000000 08251698

Instruction length: 02 Interrupt code: 000D

Failing instruction text: 58F05048 0A0DB219 0000A788

Breaking event address: 00000000 00000000

Registers 0-7

Registers 8-15

