



# What We've Done for You Lately With PDSE

Cecilia Carranza Lewis, IBM  
z/OS DFSMS Architecture, Design and Development  
carranc@us.ibm.com

August 8, 2011  
Session 9939

# Agenda

- History
- Quality Activities
- Enhancements
- Quality Trends
- References

# PDSE History



- **PDSE's (Partitioned Data Sets Extended) were first introduced in 1989 (DFP 3.2) to replace PDS (Partitioned Data Sets).**
  - Result of a mid-80's SHARE white paper describing PDS pain points, such as:
    - No reuse of space...requiring need for compress
    - Non-expandable directory size
    - Limited to 65,535 tracks
    - Limited to 16 extents
    - No ability to stop directory from being overwritten
  - **PDSE original function included:**
    - **\*\*Disk space reclaimed without compressing\*\***
    - Effectively unlimited directory space
    - Unlimited number of tracks
    - Up to 123 extents
    - Improved member integrity
    - PDSE's had to be SMS managed



## PDSE History cont.



- **After the introduction, additional functional enhancements were added:**
  - Extended sharing (DFSMS 1.1)
  - Storing executable code in PDSE's (DFSMS 1.1)
  - Non-SMS managed PDSE's (DFSMS 1.4)

# Quality Activities – Historical Issues of the mid-90’s

- **Issues**

- Outages due to excessive use of Extended Common Storage Area (ECSA)
- System hangs due to failure and cancel situations
- No tools to display/diagnose problem situations

- **Solutions**

- PDSE Address Space (SMSPDSE)
  - SPE delivered in 2002 On OS/390 R10 thru z/OS V1R5
    - *Over 98% of PDSE structures moved from Common into address space*
- Improved Diagnosis tooling
  - PDSE Monitor tool
  - Analysis commands
  - Freelatch command
- Enhanced internal test
  - Developed internal test tool, Valkyrie
  - Rebuilt and fortified PDSE test workloads
    - *Including, addition of robust cancel scenarios*

- **Result**

- Eliminated outages due to excessive ECSA usage (increased scalability)
- Reduced IPLs due to system hangs
- Provided storage administrators tools for monitoring and diagnosing
- Increased scalability

## Quality Activities continued

- Continued focus on high availability and outage eliminations
- **Solution**
  - PDSE Restarable Address Space (z/OS V1R6)
    - Non-disruptive, transparent restart
    - New address space (SMSPDSE1)
  - Enable via Parmlib member IGDSMSxx
    - PDSESHARING(EXTENDED) and PDSE\_RESTARTABLE\_AS(YES)
  - Restart
    - Vary SMS,PDSE1,Restart
- **Result**
  - Further reduced outages and need to re-IPL

## Enhancements: z/OS V1R8

- **64-bit Virtual Storage**
  - Directory buffers and control blocks for member connections moved to 64-bit virtual, allowing for greater number of concurrently opened members as well as scalability for caching directory buffers.
  - Some tuning may be done using IGDSMSxx in PARMLIB, or issuing a SETSMS command specifying 64-bit virtual storage to manage how long the unreferenced directory pages stays in the directory cache.

## 64-bit virtual cont.

- IGDSMSxx to change PDSE 64-bit virtual storage size used to cache PDSE directory buffers.
  - PDSE\_DIRECTORY\_STORAGE (nnn | 2G)
  - PDSE1\_DIRECTORY\_STORAGE(nnn | 2G)
- Operator command to change virtual size used to cache PDSE directory buffers.
  - SETSMS PDSE1\_DIRECTORY\_STORAGE(nnn)



## Enhancements: z/OS V1R8

- **Buffer Beyond Close**
  - Provides an option to retain buffers (directory and member data) after the last close of a PDSE dataset.
  - Improves performance for PDSEs that are opened/closed repeatedly.

## Buffer Beyond Close: Usage

- To enable buffer beyond close, modifications must be made to IGDSMSxx in PARMLIB or issuing a SETSMS command specifying buffer beyond close.
  - IGDSMSxx to invoke PDSE buffer beyond close
    - PDSE\_BUFFER\_BEYOND\_CLOSE(YES | NO)
    - PDSE1\_BUFFER\_BEYOND\_CLOSE(YES | NO)
  - Operator command to change the buffer beyond close option
    - SETSMS PDSE1\_BUFFER\_BEYOND\_CLOSE(YES | NO)

# Buffer Beyond Close: Performance

- Retaining PDSE buffers longer could raise a virtual storage concern.
- PDSE Cache Least Recently Used (LRU) Tuning Knobs.
  - SMSPDSE Address Space
    - PDSE\_LRUTIME
    - PDSE\_LRUCYCLES
  - SMSPDSE1 Address Space
    - PDSE1\_LRUTIME
    - PDSE1\_LRUCYCLES
- Recommend taking defaults (LRUTIME=60, LRUCYCLES=15)

## Buffer Beyond Close: Validation

- Once buffer beyond close has been invoked either through PARMLIB or operator command, issue D SMS,OPTIONS command to ensure the line item is in effect.
- Refer to the SMF cache statistics:
  - SMF Type 14/15 with extended information segment type 6
  - SMF Type 42 subtype 1

## Enhancements: z/OS V1R8 cont

- **Monitor/Analysis Command Improvements**
  - On Monitor, new DISPLAY and DUMPNEXT options
    - DISPLAY turns monitor on and causes any error messages displayed on console
    - DUMPNEXT turns monitor on and causes dump to be taken on next error
  - Analysis enhanced to provide option to take a dump if any problem found
- **Error Redrive**
  - On an ABEND0F4 due to invalid directory cache, enhance PDSE error recovery to percolate back to the entry of PDSE, purge the cache, and redrive the PDSE request....thus eliminating the 0F4.
- **PDSE LINKLIST Disconnect**
  - By removing existing member connections, enables a PDSE library to be deleted after it has been removed from the LNKLIST

## Enhancements: z/OS V1R9

- **Data Set Level Caching Statistics**
  - Provided data set caching statistics in SMF Type 14 and 15 records
    - A new optional section called *PDSE Data Set Caching Statistics (type6)* added

# PDSE Cache Statistics in SMF 14/15

## Extended Information Section Descriptor Word

Offsets		Name	Length	Format	Description
0	0	SMF14ES L	2	binary	Size of this extended information section (size of variable length fields including this 4 byte section descriptor word).
2	2		1	binary	Reserved
3	3	SMF14ST Y	1		Section type code.  Type Meaning When Set 1 Compressed format data set section 2 SMS class information section 3 Step Information section 4 ISO/ANSI Version 4 CCSID (coded character set ID) information section 5 Additional data set characteristics section <b>6 PDSE data set caching statistics</b> 7 Key label data for tape encryption.

# PDSE Cache Statistics in SMF 14/15

## PDSE Data Set Caching Statistics (type6)

This section describes the caching statistics for a PDSE data set.

Offsets	Name	Length	Format	Description
4 4	SMF14DRD	4	binary	Directory read request count
8 8	SMF14DRDH	4	binary	Directory read hit count
12 C	SMF14MRD	4	binary	Member read request count
16 10	SMF14MRDH	4	binary	Member read hit count
20 14	SMF14MCE	4	binary	Member Cache Eligible
24 18	SMF14MST	4	binary	Member Cache Stolen
28 1C	SMF14MNC	4	binary	Member Cache Eligible but not cached
32 20	SMF14MCF	4	binary	Member Cache Eligible but cache full



## Enhancements: z/OS V1R9 cont.

- **Reclaim Physical Space in PDSE**
  - Release space previous used rather than release to the highest used (formatted) track

# Enhancements: z/OS V1R10

- **PDSE DISPLAY SMS Enhancement**
  - Enhanced DISPLAY SMS command with new options to display point-in-time PSDE cache information in real time and overall effectiveness of PDSE caching
    - Such as use of member caching in hyperspace, including the size of the hyperspace, current LRUTIME value, current LRUCYCLE value, datasets eligible for caching, and datasets which are in cache.
  - Caching statistics will be displayed at the data set level
  - Allows the user to determine what, if any, changes are required to the original cache settings in order to improve system performance

## Enhancements: z/OS V1R12

- **EAV Support**
  - PDSE as EAS eligible
- **Identifying corrupted PDSE's in LNKLIST**
  - Validates PDSEs in LNKLIST using internal validation tooling
  - A new message (IGW037I) identifying any corrupted PDSE data sets that are found in the LNKLIST at IPL time. The message identifies the PDSE before a possible wait state begins, allowing time to restore the data set and re-IPL the system.
- **First phase of PDSE fairness**
  - Internal support to enhance locking

# Enhancements: z/OS V1R13

- IEBPDSE Utility
- New PDSE commands
- PDSE Fairness/Sharing Policy

# Enhancements: z/OS V1R13 IEBPDSE

- **IEBPDSE Utility**
  - PDSE Validation tool designed to verify the structure of the PDSE directory.
    - First phase of the validation tool was integrated into z/OS 1.12 to identify corrupt PDSE's in LNKLST and at NIP time.
  - Invoked via JCL: EXEC PGM=IEBPDSE
    - Optional PARM keyword: PARM [DUMP|NODUMP]
      - *DUMP option indicates that the PDSE validation utility is to issue an ABEND in the PDSE address space when an error has been found in the analysis of the PDSE.*

## Enhancements: z/OS V1R13 IEBPDSE cont.

- **IEBPDSE Utility Return Codes**

- 00 (X'00')            Successful completion
- 04 (X'04')            Slightly damaged PDSE
  - This return code describes data sets that can be opened normally but have some form of corruption. Currently, the only instance of a “slightly” damaged PDSE is when the free space list marks free pages as used, wasting space that could normally be reclaimed. This does not prevent the user from opening the PDSE but the user should copy the PDSE to a new data set.
- 08 (X'08')            Corrupted PDSE
- 12 (X'0C')            PDSE could not be opened
- 16 (X'10')            Input data set not a PDSE

## Enhancements: z/OS V1R13 IEBPDSE cont.

- JCL Examples

```
//STEPCHK EXEC PGM=IEBPDSE  
//SYSPRINT DD SYSOUT=A  
//SYSLIB DD DSN=IBMUSER.SIMPLE.V2.PDSE,DISP=OLD
```

```
//STEPCHK2 EXEC PGM=IEBPDSE  
//SYSLIB DD DSN=IBMUSER.SIMPLE.V2.PDSE,DISP=OLD  
// DD DSN=IBMUSER.SIMPLE.V3.PDSE,DISP=OLD  
// DD DSN=SYS1.TCPIP.SEZALOAD,DISP=SHR
```

```
//STEPLINK EXEC PGM=IEBPDSE,PARM='DUMP'  
//SYSLIB DD DSN=SYS1.TCPIP.SEZALOAD,DISP=SHR
```

# Enhancements: z/OS V1R13 Commands

- **Requirement**
  - Customers have expressed the need to discard cached directory pages in order to recovery from possible incorrect pages in the directory cache.
- **Solution**
  - Provide a command to discard all buffered data for a PDSE
    - V SMS,PDSE(1),REFRESH
  - Messages
    - IGW052I - REFRESH is successful
    - IGW053I - REFRESH is unsuccessful



# Enhancements: z/OS V1R13 Commands

- **Requirement**
  - Customers have expressed the need to determine the users of any given PDSE, and determine which users are blocking data set access.
- **Solution**
  - Provide a command to display the connections for a PDSE. Used to help determine scope of a problem with a particular PDSE. Allows Level 2 and customer to decide whether to cancel jobs or restart PDSE address space.
    - Display  
SMS,PDSE(1),CONNECTIONS,DSNAME(dsn),VOLSER(volume)
  - Message
    - IGW051I PDSE CONNECTIONS REPORT

# Enhancements: z/OS V1R13 Fairness

- **PDSE Fairness Policy**
  - PDSE processing is enhanced to reduce delays that can occur when multiple systems are concurrently accessing a PDSE while it is being updated.
    - The PDSE locking structure is changed to prevent one system from dominating access to a PDSE that is being accessed across systems.
    - This project is also designed to simplify the code associated the global PDSE lock structure.
  - Only impact PDSE's used under extended sharing rules. If a PDSE is being used with normal sharing rules, Fairness Policy changes will have no effect.

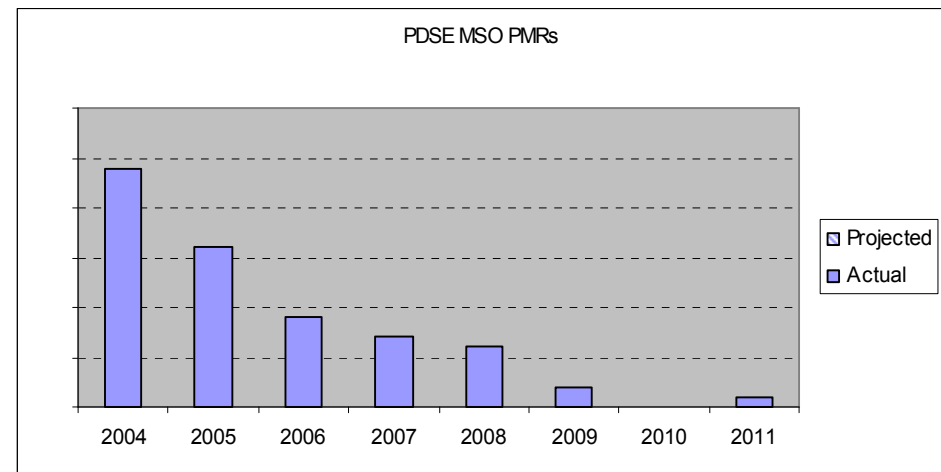
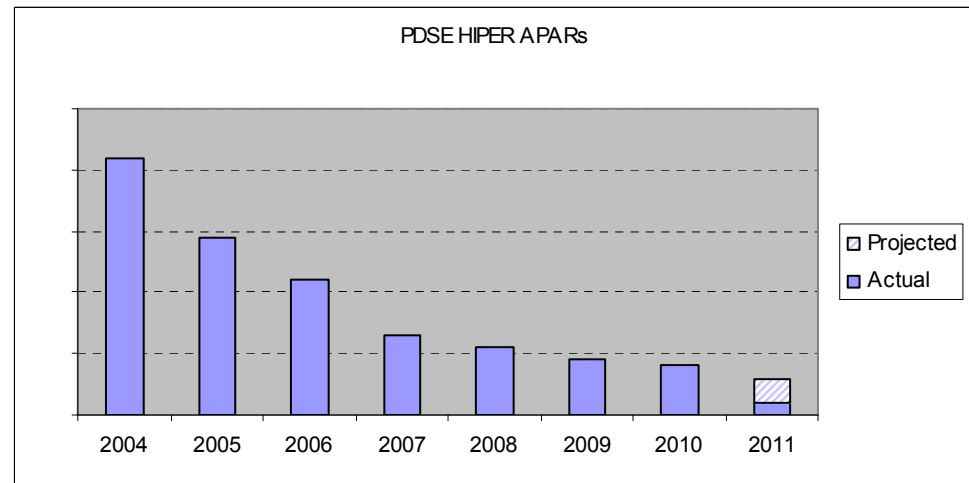
## Enhancements: z/OS V1R13 Fairness cont.

- **Problem**
  - Today, a writer to a PDSE can be blocked by high read activity against the PDSE on other systems.
- **Solution**
  - Fairness is designed to improve the efficiency of sharing a PDSE when it is being shared across systems and updated repeatedly. Algorithm for PDSE locking is changed to request ENQs unconditionally instead of conditionally.
- **Benefit**
  - Performance of PDSE is more consistent
  - There may be less GRS activity associated with PDSE processing

# Quality Trends

# PDSE High Impact Problems

- Results of PDSE improvements
  - Significant reduction (>80%) in HIPER APARS
  - Nearly eliminated Multi-System Outage PMRs
  
- In addition,
  - Field APAR rate has been reduced by nearly **60%** since 2004
  - PMR rate has been reduced by nearly **two-thirds** since 2004



# References

- Partitioned Data Set Extended Usage Guide, RedBooks
  - SG24-6106-01 or [Ibm.com/redbooks](http://Ibm.com/redbooks)
- z/OS DFSMS: Using Data Sets
  - SC26-7410-10
- z/OS DFSMS Macro Instructions for Data Sets
  - SC26-7408-09
- z/OS DFSMSdfp Utilities
  - SC26-7414-07
- z/OS MVS System Management Facilities (SMF)
  - SA22-7630-22
- Techdoc PRS4280
  - Why You Want to Use Partitioned Data Set Extended (PDSE)



***THANK YOU***



# Trademarks and Disclaimers



The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml):

<b>CICS*</b>	FICON*	Lotus*	System Storage
<b>DB2*</b>	FlashCopy*	MQSeries*	Tivoli*
<b>DFSMS</b>	GDDM*	Multiprise*	TotalStorage*
<b>DFSMSdfp</b>	GDPS*	OMEGAMON*	Virtualization Engine
<b>DFSMSdss</b>	geoManager*	OS/390*	VisualAge*
<b>DFSMShsm</b>	HiperSockets	Parallel Sysplex*	VM/ESA*
<b>DFSMSrmm</b>	HyperSwap	PR/SM	VSE/ESA
<b>DFSORT</b>	IBM*	QMF	VTAM*
<b>DFSMS</b>	IBM logo*	RACF*	WebSphere*
<b>DS4000</b>	ImagePlus*	Rational*	z/Architecture*
<b>DS6000</b>	IMS	RMF	z/OS*
<b>DS8000</b>	Intelligent Miner	System i	z/VM*
<b>Enterprise Storage Server*</b>	Language Environment*	System z	z/VSE
<b>ESCON*</b>		System z9	zEnterprise
			zSeries*
			zSeries Entry License Charge

The following are trademarks or registered trademarks of other companies:

Java and all Java based trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries or both

Microsoft, Windows, Windows NT and the Windows logo are registered trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries or both.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

## NOTES:

Any performance data contained in this document was determined in a controlled environment. Actual results may vary significantly and are dependent on many factors including system hardware configuration and software design and configuration. Some measurements quoted in this document may have been made on development-level systems. There is no guarantee these measurements will be the same on generally-available systems. Users of this document should verify the applicable data for their specific environment.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

Information is provided "AS IS" without warranty of any kind.





# Trademarks and Disclaimers (continued)



## NOTES:

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices are suggested US list prices and are subject to change without notice. Starting price may not include a hard drive, operating system or other features. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM makes no representation or warranty regarding third-party products or services including those designated as ServerProven, ClusterProven or BladeCenter Interoperability Program products. Support for these third-party (non-IBM) products is provided by non-IBM Manufacturers.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. Send license inquires, in writing, to IBM Director of Licensing, IBM Corporation, New Castle Drive, Armonk, NY 10504-1785 USA.

