# SMS is Now Mandatory for DB2 V10!
# SMS Best Practices for DB2

John Iczkovits
IBM

March 3, 2011
Session #8385

# Disclaimer

- *© Copyright IBM Corporation 2011. All rights reserved. U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.*
- *THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. WHILE EFFORTS WERE MADE TO VERIFY THE COMPLETENESS AND ACCURACY OF THE INFORMATION CONTAINED IN THIS PRESENTATION, IT IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. IN ADDITION, THIS INFORMATION IS BASED ON IBM'S CURRENT PRODUCT PLANS AND STRATEGY, WHICH ARE SUBJECT TO CHANGE BY IBM WITHOUT NOTICE. IBM SHALL NOT BE RESPONSIBLE FOR ANY DAMAGES ARISING OUT OF THE USE OF, OR OTHERWISE RELATED TO, THIS PRESENTATION OR ANY OTHER DOCUMENTATION. NOTHING CONTAINED IN THIS PRESENTATION IS INTENDED TO, NOR SHALL HAVE THE EFFECT OF, CREATING ANY WARRANTIES OR REPRESENTATIONS FROM IBM (OR ITS SUPPLIERS OR LICENSORS), OR ALTERING THE TERMS AND CONDITIONS OF ANY AGREEMENT OR LICENSE GOVERNING THE USE OF IBM PRODUCTS AND/OR SOFTWARE.*
- **IBM, the IBM logo, ibm.com, DB2 are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml**

For more detailed information review

Redpaper - Disk storage access with DB2 for z/OS

http://www.redbooks.ibm.com/redpieces/pdfs/redp4187.pdf

DB2 9 for z/OS and Storage Management SG24-7823

http://www.redbooks.ibm.com/abstracts/sg247823.html?Open

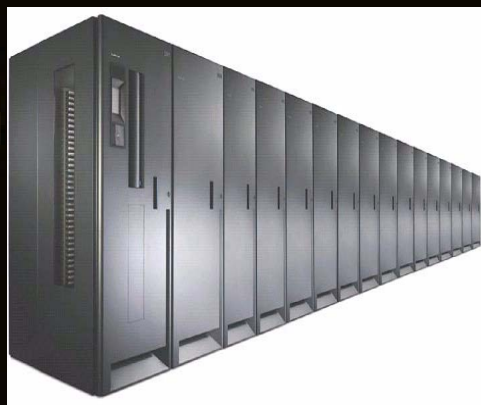
Want to understand

how DB2 works with tape?

Learn about DB2/tape

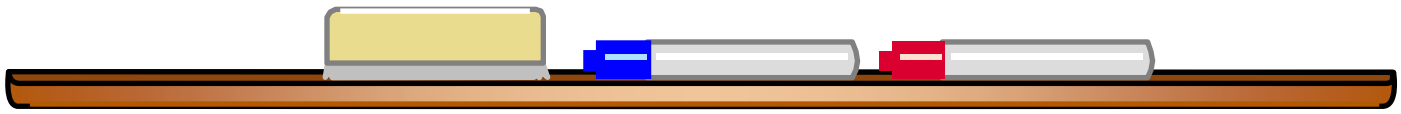
best practices:

http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS2614



# Agenda

- DB2 V10 SMS Required -
- First, What is DASD Physically? -
- DFSMS Basics for DB2 Professionals –
-z/OS 1.11 and 1.12 enhancements useful for DB2 -
- Space problems -





SHARE in Anaheim 2011

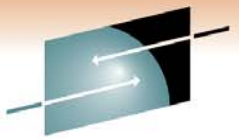
# DB2 V10
# SMS

# Required!

# DB2 10 requires that the DB2 catalog and directory be SMS managed

- New Catalog and Directory data sets created during CM require SMS, with EF (Extended Format) and EA (Extended Addressability) enabled.
- During ENFM, some of the DB2 catalog and directory objects are converted to PBG with a DSSIZE 64G.
- VSAM Objects > 4GB are required to be SMS managed.
  - Data Class in the ISMF panel - EF (Extended Format) and EA (Extended Addressability) must be turned on.
  - Data Class ACS routine must set the DB2 catalog and directory objects to this Data Class with EF/EA enabled. Setting the value is easily done as we know that the 3rd qualifier is DSNDB01 or DSNDB06.
- DSNTIJSS (in SDSNSAMP) provides SMS classes for customers without SMS in use.
  - The environment created by DSNTIJSS is ONLY for DB2 Catalog and Directory data sets.
    - Other DB2 data sets such as logs and BSDS not covered.
    - Parts of DSNTIJSS use NaviQuest procedures

# Installation process changes

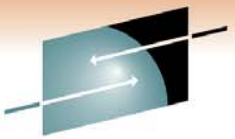


**Introducing job DSNTIJSS: Creates a sample SMS environment**

- Background:
  - In V10, data sets being defined for the catalog and directory are managed by DB2 – not the user
  - These data sets also must be associated with an SMS data class for allocating them in extended format and using extended addressability.
  - You are not required to convert existing DB2 catalog and directory data sets to the SMS environment before migrating. These can remain non-SMS managed indefinitely, but will be converted to SMS management the next time the related table space is reorganized
- DSNTIJSS shows how to create a sample SMS environment for DB2 catalog and directory data sets that consists of:
  - An SMS Source Control Data Set (SCDS) for storing the base configuration
  - An SMS storage class
  - An SMS storage group with 3 enabled volumes
  - An SMS data class for allocating data sets in extended format and to use extended addressability
  - SMS Automatic Class Selection (ACS) routines for managing the DB2 catalog and directory data sets in the storage group, storage class, and data class
- The sample environment is to help you get started – it is not the ideal or recommended SMS environment for your system
- After running the job, use the following commands to make the SCDS active
  SETSMS SCDS(*scds-name*)

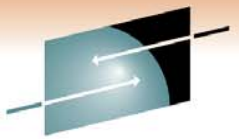
- **Serviceability:**
  - **Common problems:**
    - Symptom: After installing and starting DB2 V10, job DSNTIJTC fails with abend S04E, reason code 00C200EF
    - Cause: The catalog and directory data sets are not defined in a valid SMS environment
    - Response: Provide an SMS environment with an SMS data class that allocates the data sets in extended format and for extended addressability.
  - **Diagnostics: Obtain and review joblog**

# DB2 catalog and directory are now DB2 managed



- Not only does the DB2 catalog an directory require SMS, the data sets will also be DB2 managed instead of user managed. No more IDCAMS DEFINE requirement after ENFM for DB2 Catalog and Directory objects.
- During ENFM, some catalog and directory data sets are restructured requiring EF/EA, and will be converted to DB2 managed.
  - During ENFM most catalog and directory data sets do not require EF/EA or to be SMS managed.
  - For customers that do not SMS manage the catalog and directory data sets, moving your data sets to SMS managed volumes is not required.
    - New catalog and directory objects must be SMS managed with EF/EA enabled
    - Catalog and directory objects that are REORGed will automatically become DB2 managed, and SMS managed with EF/EA required for the newly allocated data sets.
    - During REORG of catalog and directory objects, an internal STOGROUP that cannot be viewed is used. The STOGROUP was created with VOLUMES("*").
    - Data sets that are not REORGed can remain on the non-SMS volumes until REORG.
    - This requirement is for the DB2 catalog and directory objects only and do require the BSDS or logs to be SMS managed.
    - User defined data sets for the Catalog and Directory remain the same and do not require any changes.
    - An alternative option is to use a process, such as dss to move the catalog and directory data sets to SMS managed volumes with the ACS routines assigning the Data Class with EF/EA enabled.

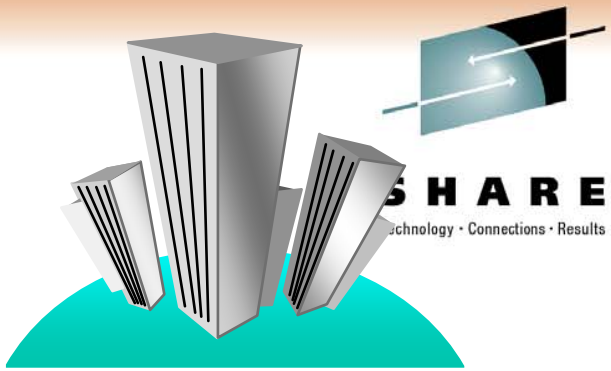
# DB2 managed data sets - Benefits



- Minimize user's effort to maintain data sets.
  - No need to allocate data sets and extended pieces (A002, A003 etc…)
  - No need to allocate data sets as part of the migration (next next release)
    - We still have DSNTIJIN in this release because there are new catalog indexes to be created during migration and we cannot make them DB2 managed until leaving CM.
  - No need to allocate data sets for the shadow for online REORG.
  - No need to estimate the size for the data sets (had to provide space allocations in the DEFINEs)
    - DB2 will use a default for the primary and a sliding scale for secondary.
    - Minimize the chance of running out of extends before reaching the maximum data set size.
  - SMS will determine which dataset goes to which volume.
  - Minimize outage due to improper data set allocations.

- New fields in installation/migration panels (CLIST)
  - SMS information (data class, storage class, management class) stored in ZPARM.

# DB2 Catalog Evolution




***The DB2 catalog continues to grow with every DB2 release.***

| DB2 Version | Table Spaces | Tables | Indexes | Columns | Table Check Constraints |
|---|---|---|---|---|---|
| V1 | 11 | 25 | 27 | 269 | N/A |
| V3 | 11 | 43 | 44 | 584 | N/A |
| V5 | 12 | 54 | 62 | 731 | 46 |
| V6 | 15 | 65 | 93 | 987 | 59 |
| V7 | 20 | 84 | 117 | 1212 | 105 |
| V8 | 21 | 87 | 128 | 1286 | 105 |
| V9 | 28 | 106 | 151 | 1668 | 119 |
| **DB2 10** | **81 (88-7)** | **124** | **195** | **1701** | **119** |

**Does not include objects for XML Schema Repository!**



SHARE in Anaheim 2011

# Catalog Restructure enhancement summary

- DB2 (SMS) managed catalog and directory data sets
- New CLOB/BLOB columns to the catalog
  - Merge records that store SQL statements' text
- Reduce catalog contention
  - Removal of links
  - Change to row-level locking
- Convert some catalog and directory table spaces to partition-by-growth (PBG)
  - With MAXPARTS = 1
- Combine SYSUTIL and SYSUTILX into a single table SYSUTILX

# Eliminate 64Gb limit on Catalog & Directory – e.g. SPT01

- **V10 will relieve SPT01 space problems**
- SPT01 can grow beyond 64G, when system in DB2 10 NFM.
- Hence Application Developer creates packages using:
  - Creates Packages using BIND PACKAGE.
  - Creates Packages using CREATE TRIGGER.
  - Creates Packages using SQL procedures.
  - Binds a package, using plan stability.
- DB2 extends SPT01 beyond 64G when needed.

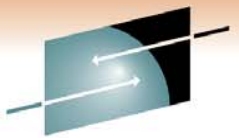
**Binds and rebinds never fail for lack of SPT01 space.**



SHARE in Anaheim 2011

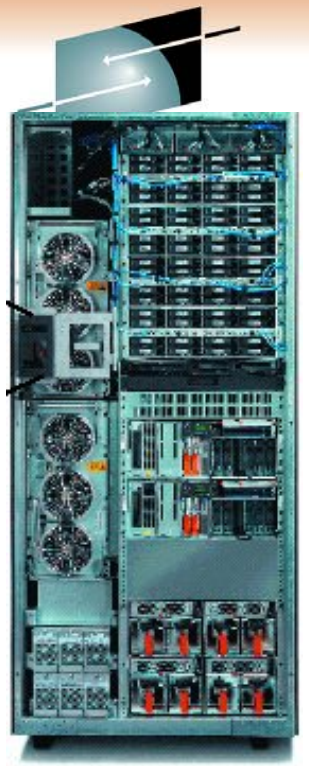
# Work files with V10

- DB2 9 changed the work file allocations from user managed to DB2 managed, however did not require that the data sets be DB2 managed.
- DB2 managed work file data sets are not required for V10. Customers can continue to use user defined work file table spaces in V10.
- With V10 NFM work files can be allocated as PBG data sets allowing keywords MAXPARTITIONS and DSSIZE to be used. This is very useful when having problems with large DGTTs.
  - Using keywords MAXPARTITIONS and DSSIZE can limit the size of the data sets and thereby limit runaway transactions
    - For example, to limit the space used to 3 GB, you could set MAXPARTITIONS 3 DSSIZE 1G. With DB2 managed classic segmented table spaces, this was was not possible. In this case you could only limit the growth at 2 GB or less (via PRIQTY nK SECQTY 0)
  - With PBGs, a DGTT can span multiple 64GB data sets, however it is limited to one table
  - PBGs must be DB2 managed data sets
  - Limiting the size can be done today using IDCAMS for user managed work files, but in this scenario PBGs are not used
- For work file data sets that exceed 4GB, Data Class EF/EA must be enabled
- If ZPARM WFDBSEP=NO (the default), DB2 will try to use work file PBG table spaces for DGTT only, however if there is no other table space available DB2 will also use it for workfiles (e.g. sorts, CGTTs). With WFDBSEP=YES, DB2 will only use work file PBG table spaces for DGTT and if there is no other table space available, a work file application (e.g. sorts, CGTTs) will get a 'resource unavailable' message.




# First,
# What is DASD Physically?

# IBM TotalStorage DS8000 Series - 2004



- Different than RVA and ESS, but still using disk array concept.
- Capacity scales linearly up to 1,024 TB
- FlashCopy
- Supports PAV and MA
- With the implementation of the POWER5 (POWER6+ with DS8800) Server Technology in the DS8000 it is possible to create storage system logical partitions (LPAR)s, that can be used for completely separate production, test, or other unique storage environments
- New caching algorithms (all variations still used):
  - 2004 - ARC (Adaptive Record Cache)
    - Dynamically partitions the read cache between random and sequential portions
  - 2007 – AMP (Adaptive Multi-Stream Pre-Fetch – R3.0)
    - Manages the sequential read cache and decides what, when and how much to prefetch
  - 2009 – IWC (Intelligent Write Cache – R4.2)
    - Manages the write cache and decides what order and rate to destage

# IBM System Storage DS8000 Turbo – Powerful Innovation (information below are not turbo specific)

## Innovation that extends DS8000 world class performance

- **Storage Pool Striping** –new volume configuration option to maximize performance without special tuning
- **AMP-**breakthrough caching technology can dramatically improve sequential read performance to reduce backup times, processing for BI/DW, streaming media, batch
- **z/OS Global Mirror Multiple Reader**- IBM unique Innovation to improve throughput for z/OS remote mirroring

## Innovation to simplify and increase efficiency

- **IBM FlashCopy SE** (space efficient snapshot capability) can lower costs by reducing capacity needed for copies.
- **Dynamic Volume Expansion** - Easier, online, volume expansion to support growth
- **Expansion frame warranty intermix** - Increased upgrade flexibility and investment protection through base and expansion frame machine type (warranty) Intermix
- **SSL**-New secure connection protocol option for call home support and additional audit logging
- **IBM System Storage Productivity Center** -Enhanced user interface with single pane control and management

***Enabled through new release of microcode***
***New and existing customers can leverage the new capabilities!!***

# DS8000 OEL Family Releases

**2004**
- 255 LCUs Supported
- RAID5/RAID10
- RMC/zGM/PTC/PAV
- 64K Logical Volumes
- 2GB FCP/FICON
- 73/146/300GB DDs

**10-2006**
- HyperPAV
- HMC CIM Agent
- 3rd & 4th expansion frame

**2-2008**
- SSPC Support (upgr)
- DS8000 M/T intermix

**10-2008 (?)**
- zHPF
- zGM Incremental Recovery

**7-2009**
- Thin Provisioning
- Quick init
- zHPF Multi-track support

**4-2010**
- Easy Tier
- Thin Provisioning
- Quick Init
- 600GB 15K
- 2TB SATA
- Multi-GM
- zHPF multi-trk support.

**2006**
- Turbo Models
- 500GB FATA
- 4GB FCP/FICON
- 242x Machine Types
- Synergy Items

**2007**
- SSPC Support (new)
- Storage Pool Striping
- FC Space Efficient
- Dynamic Volume Expansion

**5-2008**
- Extended Address Volumes
- Variable LPAR
- IP v6

**2-2009**
- Solid State Drives
- 1 TB SATA
- Intelligent Write Cache
- Full Disk Encrypt
- Remote pair FC/IC Writes (4.2.5).

**10-2009**
- DS8700

**10-2010**
- DS8800

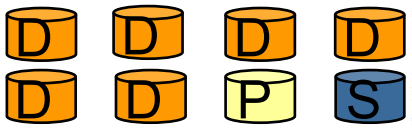
# RAID (Redundant Arrays of Inexpensive Disks) Technology- Disk Type Dependent

- RAID 0 - data striping without parity
- RAID 1 - dual copy
- RAID 2 - synchronized access with separate error correction disks
- RAID 3 - synchronized access with fixed parity disk
- RAID 4 - independent access with fixed parity disk
- **RAID 5 – data striping, independent access with floating parity (tolerates loss of 1 physical volume (DDM))**
- **RAID 6 – data striping, dual redundancy with floating parity (tolerates loss of 2 physical volumes (DDMs))**
- **RAID 10 (DS8000 and some ESS) - RAID 0 + RAID 1, mirrored, data striping but with no parity**

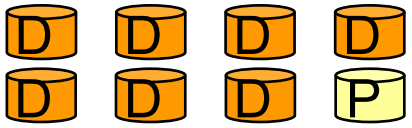
Parity is additional data, “internal” to the RAID subsystem, that enables a RAID device to regenerate complete data when a portion of the data is missing. Parity works on the principle that you can sum the individual bits that make up a data block or byte across separate disk drives to arrive at an odd or even sum.

# Array

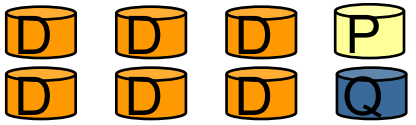
- DS8000 – 8 DDM arrays
  - 1 array site
  - RAID5
    - 6+P
    - 7+P
    - Parity is striped across all disks in array but consumes capacity equivalent to one disk
  - RAID 6
    - 6+P+Q
    - 5+P+Q+Spare
  - RAID10
    - 3+3
    - 4+4



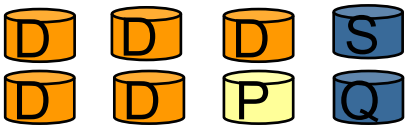

RAID5 6+P+S



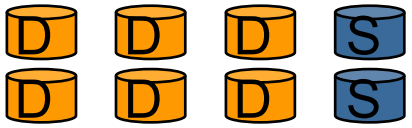

RAID5 7+P



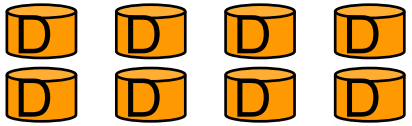

RAID6 6+P+Q




RAID6 5+P+Q+S




RAID10 3+3+S+S

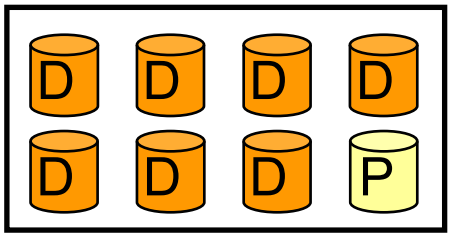



RAID10 4+4

# Rank

- RAID array with storage type defined
  - CKD or FB
- One-to-one relationship between an array and a rank
  - One RAID array becomes one rank
  - DS8000 – 8 DDMs

**DS8000 CKD rank**



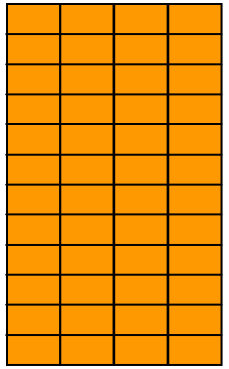

RAID5 7+P

# Rank (continued)

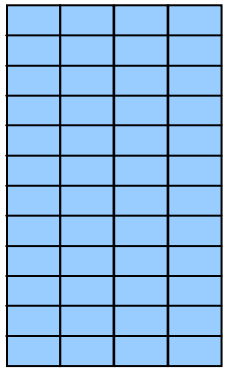
- Ranks have no pre-determined or fixed relation to:
  - Server0 or Server1
  - Logical Subsystems (LSSs)

- Ranks are divided into ‘extents’
  - Units of space for volume creation
  - CKD rank
    - Extents equivalent to a 3390M1
    - 1113 cylinders or .94GB
  - FB rank
    - 1GB extents

CKD Rank

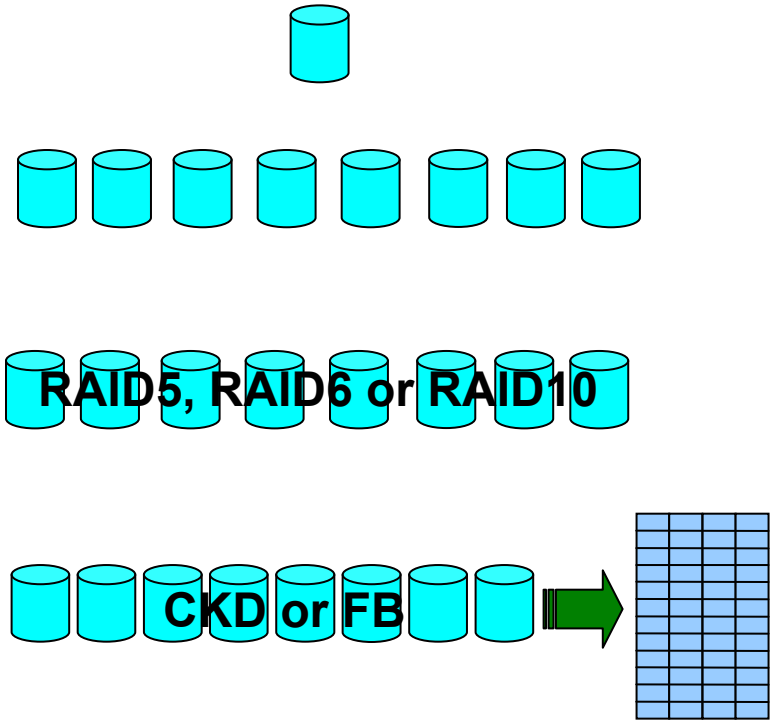


FB Rank



# Storage Resource Summary

- Disk
  - Individual DDMs
- Array Sites
  - Pre-determined grouping of DDMs of same speed and capacity (8 DDMs for DS8000; 4 DDMs for DS6000)
- Arrays
  - One 8-DDM Array Site used to construct one RAID array (DS8000)
- Ranks
  - One Array forms one CKD or FB Rank (8 DDMs for DS8000; 4 or 8 DDMs for DS6000)
  - No fixed, pre-determined relation to LSS
- Extent Pools
  - 1 or more ranks of a single storage type (CKD or FB)
  - Assigned to Server0 or Server1



# Disk Space Numbers - 3390 Emulated Volumes

| Model | Cylinders | Tracks | Bytes/Volume | Bytes/Track ** |
|---|---|---|---|---|
| 3390-1 | 1113 | 16695 | 946 MB | 56664 |
| 3390-2 | 2226 | 33390 | 1.89 GB | 56664 |
| 3390-3 | 3339 | 50085 | 2.83 GB | 56664 |
| 3390-9 * | 10017 | 150255 | 8.51 GB | 56664 |
| 3390-27 * | 32760 | 491400 | 27.84 GB | 56664 |
| 3390-54 * | 65520 | 982800 | 55.68 GB | 56664 |
| 3390-A EAV | 262,668 | * * * | 223 GB | 56664 |

* Storage Administrators refer to 3390 mod 9, 27, and 54 as large-volume support 10017 cylinders, 32760 cylinders, and 65520 cylinders respectively. These large volumes are all considered mod 9s with varying number of cylinders.

** Bytes per track refers to the device capacity only. The actual allocation for a DB2 LDS data set is 48 KB, not the total of 56 KB. This allows for 12 - 4K DB2 pages to be allocated per track.

Why larger disk sizes?

- reduces issues with MVS addresses (maximum number of devices met)
- simpler management of storage

# What is a DDM (Disk Drive Modules) physical drive?

- Three different Fibre Channel (FC) DDM types (available in both non-encrypted and encrypted versions):
  - 146 GB, 15K RPM drive
  - 300 GB, 15K RPM drive
  - 450 GB, 15K RPM drive
  - See next page regarding 2.5 inch SAS drives for the DS8800, which includes 600 GB, 10K RPM drives
- One Serial Advanced Technology Attachment (SATA) DDM drive type:
  - 1 TB, 7.2K RPM drive
  - 2 TB, 7.2K RPM drive
- Two different Solid State Drive (SSD) types:
  - 73 GB
  - 146 GB
  - See next page regarding for the DS8800, which includes the new 300 GB drives
- Disks are 2.5 inch SAS (Serial Attached SCSI) for DS8800, otherwise 3.5 inches FC
- Each DDM (physical disk) can hold many logical volumes. For example, a mod 3 device is 2.83 GB. Many mod 3s can reside on one DDM. When you lose one DDM, how many logical volumes are you really losing? Think about storage pool striping in the next few slides as well.
- Three types of drives:
  - HDD (Hard Disk Drive) Fibre Channel. Almost all mainframe disk is HDD.
  - SATA. Too slow for most mainframe uses. Some companies will use SATA drives for Data Warehousing if slower performance is acceptable, or for HSM Level 2, or backups, such as image copies. Much slower performance must be acceptable when putting data on SATA drives. Generally SATA drives are not used for DB2 environments.
  - SSD. VERY expensive and not generally used for DB2 environments. Very useful when access is random. Random access is better on SDD vs. HDD or SATA.

# DS8800 Hardware Changes

- Drives
  - Seagate Hornet 15K 146 GB non-FDE (Full Disk Encryption)
  - Seagate Firestorm 10K RPM 450 GB and 600 GB both FDE and non-FDE.
  - STEC Hikari SSD 300 GB non-FDE
    - SSD drive sets are not supported in RAID-6 or RAID-10 configurations
- All vendors including IBM depend on other vendors for their physical disk. For example, for IBM devices, Seagate provides all vendors disk for HDD and SATA, while STEC provides all vendors for SSD. What is very different is how each vendor chooses to implement the software that drives the hardware.

# SSD (Solid State Disk) Candidates

- SSD has no moving parts. This also means that traditional latency times are reduced to primarily data transfer, since disk seek, scan, etc. are no longer performed for SSD.
- Today SSD is only practical if there is high access density (IO/sec/GB)
- Data sets with good cache hit ratios or low I/O rates should remain on spinning disks
- Sequential data sets should remain on spinning disks
- Data sets with high I/O rates and poor cache hit ratios are good candidates for SSD, but they are likely to be large data sets

# Easy Tier - Overview

- Easy Tier is a DS8700 feature that supports online dynamic relocation of data at the sub-volume/LUN level
  - Workload learning algorithms collect I/O statistics and generate a heat map of data to be optimized on SSD
  - Data can be relocated to/from SSD and FC/SATA
    - Automatic Mode
    - Manual Mode
  - Storage Tier Advisor Tool (STAT) for I/O analysis and projected benefit
- DS8700 Hardware feature
  - Easy Tier is a new LIC feature available with Release 5.1
  - Supported by all server platforms with no additional software

# Storage Pool Striping

- Example of using storage pool striping with rotating extents across 3 ranks with RAID 5 (7+P) for a mod 3 device (2.83GB - 3,339 cylinders) – 450GB DDMs

**Rank 1 – 1,113 cylinders**
**Logical volume - .94GB**

**Rank 2 – 1,113 cylinders**
**Logical volume - .94GB**

**Rank 3 – 1,113 cylinders**
**Logical volume - .94GB**

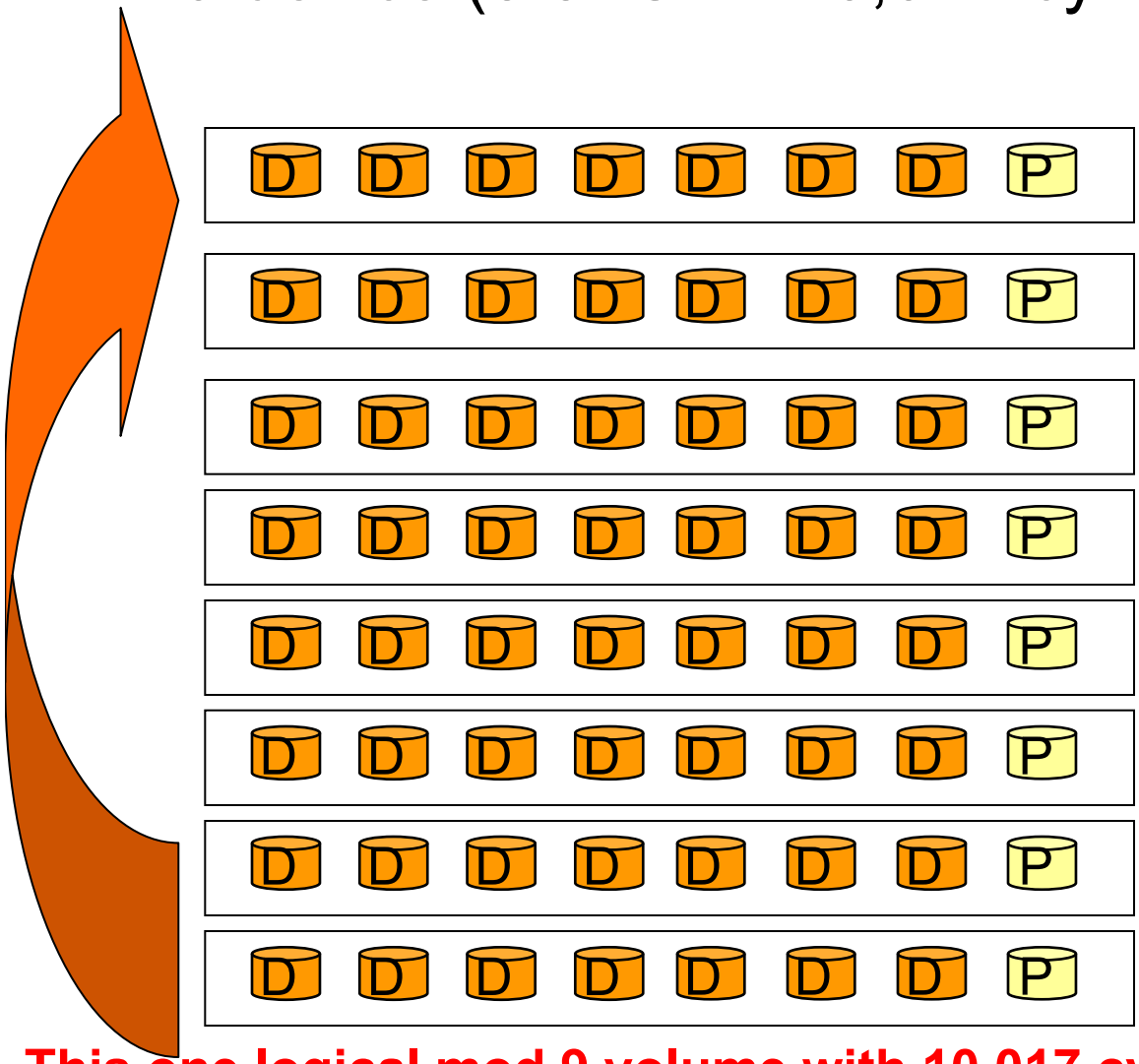
**This one logical mod 3 volume with 3,339 cylinders really resides on 24 physical devices (DDMs)**

**In this scenario when you backup or Flash Copy one logical volume you are really accessing 24 physical volumes.**

- Parity is striped across all disks in array but consumes capacity equivalent to one disk

# Storage Pool Striping – same example, but 8 ranks and a mod 9 device

- Example of using storage pool striping with rotating extents across 8 ranks with RAID 5 (7+P) for a mod 9 device (8.51GB - 10,017 cylinders) - 450GB DDMs




**Rank 1 – 1,113 cylinders**

**Logical volume - 1.88GB**

**Rank 2 – 1,113 cylinders**

**Logical volume - .94GB**

**Rank 3 – 1,113 cylinders**

**Logical volume - .94GB**

**Rank 4 – 1,113 cylinders**

**Rank 5 – 1,113 cylinders**

**Rank 6 – 1,113 cylinders**

**Rank 7 – 1,113 cylinders**

**Rank 8 – 1,113 cylinders**

**Logical volume - .94GB**

**This one logical mod 9 volume with 10,017 cylinders really resides on 64 physical devices (DDMs). Mod 9 devices require the equivalent of 9 ranks. Since only 8 ranks exist, the last 1,113 cylinders will be added back to rank 1. When adding to ranks, full ranks are skipped.**

# Storage Pool Striping - Advantages

- Technical Advantages
  - Method to distribute I/O load across multiple Ranks
  - DS8000 optimized for performance
  - Far less special tuning required for high performance data placement.
    - Means less work for the storage administrator
  - Reduces storage administrator work needed to optimize performance
  - Less potential for hot spots

# How many DDMs does my single volume data set reside on?

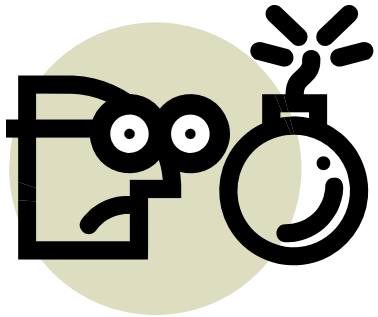
- From a VTOC perspective an allocation on a single volume will display one (logical) volume. In reality the data set resides on several different physical volumes.
- For our examples we are using a RAID 5 7+P configuration with at least 3 ranks
- Data set is on a single logical volume with 2,500 cylinders allocated. Can be 1 or multiple extents, in this case it does not matter. In this scenario the data set is allocated on 24 DDMs (3 ranks * 8 DDMs).
- Data set is on a single logical volume with 25 cylinders allocated in 1 extent. Allocation can be on 8 DDMs (1 rank), or 16 DDMs (2 ranks). Since the VTOC is not aware of ranks, it is possible there is a chunk of space that spans ranks that they are both used. For something as small as 25 cylinders, 3 ranks would not be used as they would not exceed 1 extent on 2 ranks.
- Data set has 2 extents on a single logical volume, each extent is 1 track. It depends. For example, extent 1 was allocated on 1 array, extent 2 is not allocated until a month later. Extent 2 could be on any rank with sufficient space. It is possible that extent 1 resides on an array in one rank, and that extent 2 resides on a different array in a different rank.



SHARE in Anaheim 2011

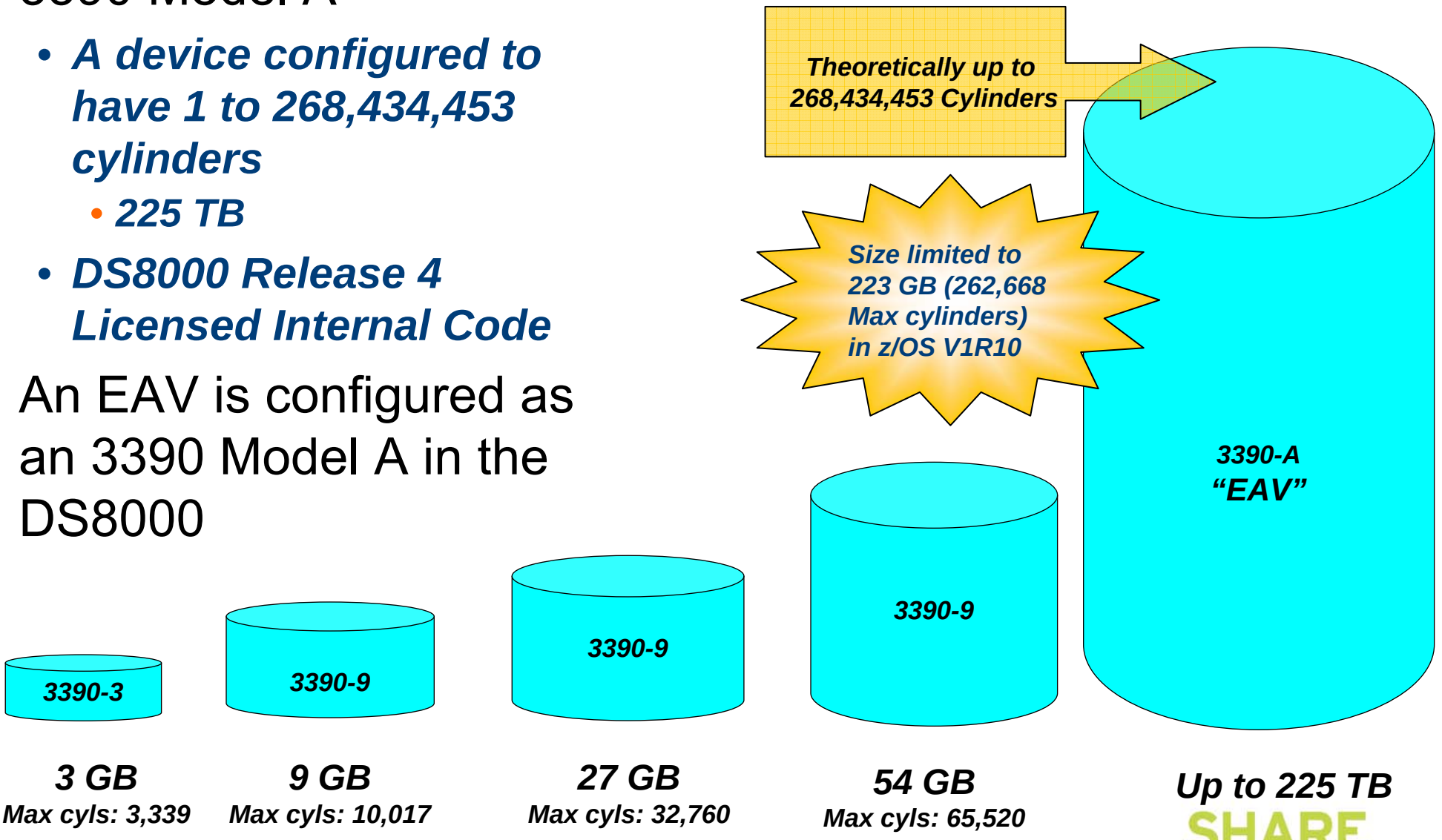
# MTBF (Mean Time Between Failure)

- Depending on the manufacturer, the MTBF for disk is generally 1,000,000 (1M) + hours.
- MTBF also depends on type of disk, where SDD ranges between 1.5M and 2M hours.
- Even at the lower number of 1M, that is 114 years as mean time between failures!
- MTBF refers to the physical disk, it does not include problems such as:
  - Outside of the physical disk itself, such as a bad card in the controller
  - VTOC being ZAPed incorrectly causing the volume to become unusable

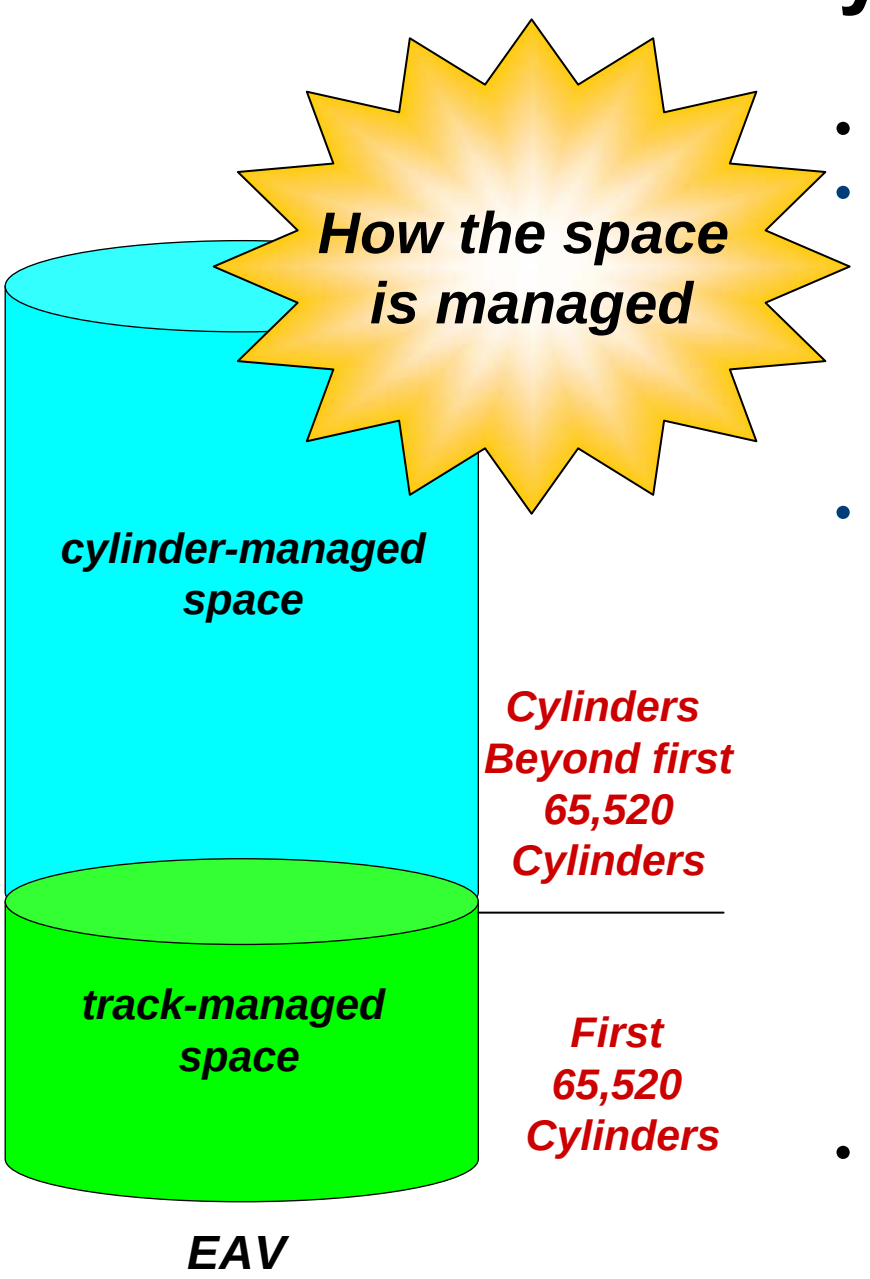


# Overview - DS8000 Support

- 3390 Model A
  - ***A device configured to have 1 to 268,434,453 cylinders***
    - ***225 TB***
  - ***DS8000 Release 4 Licensed Internal Code***
- An EAV is configured as an 3390 Model A in the DS8000



# Overview – EAV Key Design Points




- Maintains 3390 track format
- **Track-managed space:** the area on an EAV located within the first 65,520 cylinders
  - Space is allocated in track or cylinder increments
  - Storage for "small" data sets
- **Cylinder-managed space:** the area on an EAV located above the first 65,520 cylinders
  - Space is allocated in **multicylinder units**
    - A fixed unit of disk space that is larger than a cylinder. Currently on an EAV it is 21 cylinders
    - System may round space requests up
  - Storage for "large" data sets
- Track-managed space comparable to same space on non-EAVs

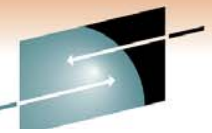
***DB2 data sets may reside in EAS, by z/OS release with appropriate maintenance applied***

**** eligible starting in DB2 V8***

***** eligible starting in DB2 V9***

| | z/OS 1.10 | z/OS 1.11 | z/OS 1.12 |
|---|---|---|---|
| Tables and Indexes * | Yes | Yes | Yes |
| BSDS * | Yes | Yes | Yes |
| Active Logs * | Yes | Yes | Yes |
| Archive Logs ** | No | Yes, if EF sequential | Yes |
| Utilities sequential input and output data sets * | No | Yes, if EF sequential | Yes |
| Utilities partitioned data sets and PDSEs | No | No | Yes |
| Sort Work data sets | No | No | Yes, if using DFSORT for DB2 utilities |
| DB2 Installation data sets (clists, panels, samples, macros, etc) | No | No | Yes |
| SDSNLINK, SDSNLOAD | No | No | Yes |
| HFS | No | No | No |

# VSAM Data Striping - ESS, DS8000, and some RVA

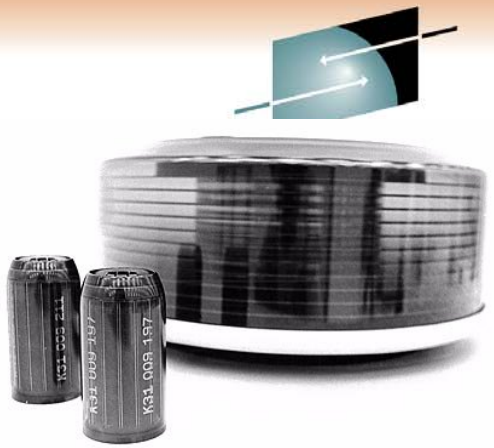


- Mainframes have RAID 5, 6, and/or 10 devices. All of these RAID types single stripe data. VSAM and sequential striping are used for multi striping data.
  - LISTC only displays stripe related information when multi-striped
- Spreads the data set among multiple stripes on multiple control units (this is the difference from hardware striping which is within the same disk array)
- An equal amount of space is allocated for each stripe
- Striped VSAM data sets are in extended format (EF) and internally organized so that control intervals (CIs) are distributed across a group of disk volumes or stripes.
  - DB2 V8 now allows striping for all page types, while V7 only allows striping for 4K pages. See APAR PQ53571 For issues with objects greater than 4K pages prior to DB2 V8.
- Greater rate for sequential I/O
- Recommended for DB2 active log data sets
- I/O striping is beneficial for those cases where parallel partition I/O is not possible. For example: segmented tablespace, work file, non partitioning indexes (NPIs, NPSIs), DPSIs, LOBs, etc.

# Sequential Data Striping

- Recommendations –
  - Stripe all possible objects that are input or output for utilities when the associated table space is striped. For example (when the associated table space is striped):
    - Stripe the image copy data sets
    - Stripe input files for LOADs, etc.
- Only disk data sets can be striped.
- You can now stripe your disk archive log data sets beginning in DB2 9.
  - You can now also compress archive log data sets providing a considerable savings in disk space.

# CCHH (Cylinder Head) for newer disk



Messages will still show CCHH information for newer disk (RVA, ESS, DS8000). For example:

LISTCAT output

```
EXTENTS:
LOW-CCHH-----X'079D0000'
HIGH-CCHH----X'07AC000E'
```

IEHLIST utility with FORMAT option

```
EXTENTS   NO   LOW(C-H)    HIGH(C-H)     NO   LOW(C-H)    HIGH(C-H)
            0      0  1          0  1       1      0 12          0 14
```

DSNU538I RECOVER ERROR RANGE OF DSN=dataset name ON VOLUME=volser FROM CCHH=X'ccchhhh' TO CCHH=X'ccchhhh' CONTAINS PHYSICAL ERROR

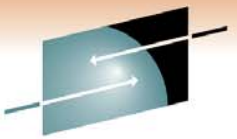
- The disk controller itself will track allocations between the VTOC and where data really resides. Data sets no longer really reside at the CCHH reported.



SHARE in Anaheim 2011

# dfp Enhancements and DB2

- As with DB2, dfp introduces enhancements on a periodic basis. These enhancements may effect data sets used by DB2, however, keep in mind that many times a DB2 APAR is required before DB2 can take advantage of the dfp enhancement.
  - For example, the dfp enhancement to increase the VSAM data set maximum to 7,257 extents required some DB2 APARs in order to utilize the enhancement. Another example is when dfp introduced large sequential data sets in z/OS 1.7, DB2 could not take advantage of utilizing this enhancement for such things as archive log data sets and image copies until the introduction of DB2 9.
  - This is true for other dfp enhancements as well. For example, extended format sequential in z/OS 1.11 is not available for use for utilities with the current versions of DB2 (8 and 9).
- Verify if any DB2 maintenance is required to take advantage of dfp enhancements.
- DB2 generally releases a new version every 3 years. MVS (z) on the other hand provides a release about once a year. There is 1 or several disk related updates once a year. Make sure you understand how these other newer releases relate to DB2.




Simplified data allocation

Improved allocation control

Improved performance management

# DFSMS Basics for DB2 Professionals

Automated disk space management

Improved data availability management

Simplified data movement

# Managing Data With SMS

No SG info!

```
ISPF 3.4 Data Set listing:
General Data
 Management class . . : MCDB2
 Storage class  . . . : SCDB2
  Volume serial . . . : DB2810
  Device type . . . . : 3390
 Data class . . . . . : DCDB2
```

ACS (Automatic Class Selection) Routines

- What does it look like?
- What are its requirements?
- Where is it placed?

- Data Class
- Storage Class
- Management Class
- Storage Group

ACS routines invoked in order (DC,SC,MC,SG)

# Data Set Separation - Part of the SMS Base Configuration

- Allows you to designate groups of data sets which are to be physically separated
- Volume separation is now available with z/OS 1.11
  - This new function may be useful when you have problems with hotspots. For example, you do not want partition 1 and 2 allocated on the same volume. More important is to make sure dual software copied data sets such as the BSDS and active logs are allocated on different extent pools.
  - SMS attempts to separate data sets listed in a separation group onto different extent pools (in the storage subsystem) and volumes.
    - Separation Profile can be coded so that 2 data sets must never reside on the same logical volume, however it cannot guarantee they will reside in different extent pools.
      - You can verify the extent pool a logical volume belongs to using the RMF Cache Subsystem Activity Report.
- For DB2, you may want to consider using this technique for the BSDS and active log data sets. This is an availability benefit for DB2.
  - My personal preference for the BSDS and active log data sets is to use the Storage Class Guaranteed Space attribute and hand place these data sets. The Separation Profile can be used along with Guaranteed Space to ensure data sets are allocated correctly.
- SMS attempts to allocate the data sets behind different physical control units and/or volumes. NOTE – control unit separation is different than the new volume separation profile.
- A separation profile data set must be provided

# Interesting Data Class Information for DB2 Professionals

- **A collection of allocation and space attributes - for new data sets only, e.g. - LRECL, BLKSIZE, space, etc.**
  - **Space specified by user can be overridden with Data Class space attributes.**
- **Can be used for SMS or non-SMS managed data sets. Some attributes such as EF or bypassing the 255 extent rule cannot be associated with non-SMS managed data sets. Attributes such as DCB and space can be associated with non-SMS managed data sets.**
- **For VSAM and sequential data sets**
- **Formats can be: EXTENDED Format (EF), Extended Addressable (EA), HFS, LIBRARY (PDSE), PDS, and LARGE format sequential**
  - **For DB2 partitions, LOBs, and XML greater than 4GB require EF/EA enabled.**
  - **When VSAM or sequential striping, EF must be enabled + Storage Class value for SDR**
- **Provides space constraint relief for VSAM and sequential data sets – avoids many X37 abends**
  - SMS can be set up to allocate a percentage of the requested quantity if not enough space is available. This option is not available for data sets using Guaranteed Space.
  - It is possible to set up SMS whereby the 5 extent rule is bypassed for initial allocation, as well as when extending a data set to a new volume.
  - VSAM and non-VSAM multistriped data sets do not support space constraint relief. However, single-striped VSAM and non-VSAM data sets can use space constraint relief.
- **Allows for data sets to be multi volumed - e.g. image copy data sets**
- **Allows certain VSAM data sets to exceed the 255 extent limit, see 'Extent Constraint Removal'**
- **Compress (Compact) objects – archive log**
- **CA (Control Area) reclaim - indicates whether the DASD space for empty CAs will be reused on z/OS 1.12 or later systems**

# Data Class – VOLUME COUNT and DYNAMIC VC

- Volume count and dynamic volume count are similar. They both add volumes at EOV. Volume count will add candidate volumes, dynamic volume count will not.
  - VC shows candidate volume, DVC does not
- Recommendation:
  - For DB2 managed data sets, keep the default of 1 for volume count and do not have a value for dynamic volume count. This has to do with the way DB2 adds volumes at EOV.
  - For user managed data sets, such as the DB2 catalog and directory, keep the default of 1 for volume count and allow for a reasonable value for dynamic volume count. You may want to set VC=1, DVC=2 or 3. This would allow a DB2 catalog and directory data set to expand to the number of volumes set for DVC.
    - Using the VOLUMES parameter on DEFINE overrides VC and DVC
    - DEFINE using no Guaranteed Space, but with VOLUMES parameter specifying several volumes results in the allocation of the number of volumes requested, but only the first volume is allocated, the remaining volumes are candidate volumes with "*" for the VOLSER. For example, GS=N, specify 6 volsers for VOLUMES, VC=3, DVC=5, allocation will result on 1 physical volume for the primary allocation, and 5 candidate volumes with "*" as the volser.
    - DEFINE using Guaranteed Space with no VOLUMES parameter results in the allocation of volumes for VC and DVC. Volumes are allocated with no candidate volumes. For example, GS=Y, no VOLUMES, VC=3, DVC=5, allocation will result on 3 physical volumes, with the primary allocation value on all 3 volumes because of GS.
  - NOTE – when DB2 Catalog and Directory data sets are converted to DB2 managed in DB2 V10, VC and DVC are ignored even when specified. DB2 managed data sets use the normal EOV process as opposed to VC/DVC specifications.
    - Most of the time SMS rules apply, however there are certain circumstances such as above where DB2 code overrides SMS code.

# EF/EA Recommendation

- EF (Extended Format) and EA (Extended Addressable) are primarily used in DB2 environments to be able to exceed the 4GB limit of LOBs, XML files, and partitions.
- Enabling EF/EA in the SMS Data Class is recommended for DB2 LDS with newer disk controllers and channels. This would provide DBAs the flexibility to assign objects requiring a large DSSIZE without the need to ask the Storage Administrators for a special change.
- If Storage Administrators do not want to assign EF/EA to all DB2 LDS, one Data Class can be assigned EF/EA and in DB2 9 DBAs can CREATE or ALTER a DB2 STOGROUP and specify the assigned Data Class. Objects requiring a larger DSSIZE can then be created using this special Data Class.
- NOTE - An EF data set requires a 32 byte suffix for every physical block. The number of physical blocks per CI depends on the CI size. If the CI size is such that the best use of DASD space is to create two physical blocks per CI (32K object), the 32 byte suffix is for every block. This does NOT add space for the data set – the additional bytes are generally hidden for DB2 objects.

# Large Sequential Data Sets – see DB2 notes

- Removes the size limit of 65535 tracks (4369 cylinders) per volume for sequential data sets
  - BSAM, QSAM, and EXCP
    - Data sets do not have to be in SMS managed and in extended format
  - 16 extents is still the limit
  - Architectural limit is 16,777,215 tracks
  - JES2/JES3 spool can now be larger than 64K tracks
    - Must still be a single extent
- Changed APIs supports all sequential and partitioned data sets
- DFSMShsm support of migration/recall, backup/restore and ABACKUP/ARECOVER of large format data sets
- Addresses limits on capacity in customers' systems
  - Direction is to support large capacity devices
  - Systems support up to 64,512 (63K) devices

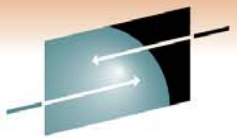
# Specifying Large Sequential Data Sets

- DSNTYPE DD statement, dynamic allocation text unit, TSO ALLOCATE keyword support 4 new values:
  - LARGE
    - If the data set is sequential or DSORG is omitted, then large format data set
  - EXTREQ
    - If the data set is VSAM or sequential or DSORG is omitted, then extended format data set is required
  - EXTPREF
    - If the data set is VSAM or sequential or DSORG is omitted, then extended format data set is preferred
      - *If not possible, then neither extended format, nor large format*
  - BASIC
    - If the data set is VSAM or sequential or DSORG is omitted, then neither extended format nor large format
- Data class can provide the new DSNTYPE
  - New &DSNTYPE variable for ACS routines with values of either LARGE or BASIC



SHARE in Anaheim 2011

# DB2 9 and Large Sequential Data Sets



- Can be non SMS managed.
- In DB2 9 NFM supported for:
  - Utility input data sets.
  - Utility output data sets when coded in the SMS Data Class or DD includes DSNTYPE=LARGE. NOTE – at this time you cannot use DSNTYPE in your TEMPLATE statement, nor can you use a model DSCB that uses LARGE. You can however point to the TEMPLATE to a DATA CLASS with LARGE enabled.
  - Archive log data sets can be allocated on one disk as 4 GB -1 byte. Archive data sets can be created on disk only in NFM or ENFM modes, however, DB2 tolerates reading them in CM.
    - LARGE and EF are mutually exclusive. If you want to compress the archive log data sets you must specify EF which will exceed the 4,369 cylinder limit as well.
    - ZPARM values for the archive log Data Class do not exist. The Data Class ACS routine must associate the archive log with the correct Data Class.

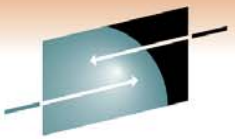
# DB2 9 and Large Sequential Data Sets

- Large sequential data sets are eligible for EAV devices. Do not confuse large sequential data sets with Extended format sequential data sets which are new types of data sets on the EAV introduced by z/OS 1.11.

# Large Block Interface (LBI)

- Large Block Interface allows block size > 32,760 bytes
  - Up to 40% reduction elapsed time for COPY and RECOVER RESTORE phase
- With LBI the maximum block size for most tapes increases from 32,760 to 256KB.
- LBI is supported on all tapes, DASD, dummy data sets, and z/OS UNIX files.
  - Although LBI works with DASD, you cannot exceed a block size of 32,760
  - With LBI, IBM tape devices such as the VTS, TS7700, ATL, 3590, and 3592 can use a block size of 256KB.
  - Maximum block size limits are documented in the *DFSMS Using Data Sets* manual under "Optimum and Maximum Block Size Supported"
  - DB2 archive log data sets cannot exceed a block size of 28K, even of tape
- No minimum level of z or tape microcode is required. The z/OS software asks the drive its maximum and optimum block size
- Block size can be set:
  1. **BLKSZLIM value in the DD statement or dynamic allocation.**
  2. **ISMF Data Class panel for option "Block Size Limit"**
     - **The Data Class ACS routine can automatically set allocation for a Data Class with LBI enabled, or can be specified by using the DATACLAS parameter on a DD statement if allowed by the Storage Administrator.**
     - **LBI enabled data sets can be SMS or non-SMS managed.**
  3. **TAPEBLKSZLIM value in the DEVSUP*xx* member of SYS1.PARMLIB. A system programmer sets this value, which is in the Data Facilities area (DFA) (see *z/OS DFSMSdfp Advanced Services*)**
     - Only one option is required to use LBI. Any combination or all 3 options can be specified. JCL overrides the Data Class, and the Data Class overrides PARMLIB.
- When migrating to DB2 V10:
  - migrate from V8 -> LBI only allowed in V10 NFM, disallowed in CM8, ENFM8.
  - migrate from V9 -> LBI allowed in V9 NFM, CM9, ENFM9, and V10 NFM.

- **Separates data set performance objectives and availability from physical storage – it does not represent any physical storage, but rather provides the criteria that SMS uses in determining an appropriate location to place a data set.**
- **SMS determines at this time if a data set is SMS managed or not. Ones that are not do not continue down the ACS routines and stop here.**
- **Most Storage Class definitions are no longer generally required, such as MSR rate, accessibility, etc. Storage Class is used in DB2 environments for three general purposes:**
  - Guaranteed Space
  - SDR rate for VSAM and sequential striping
  - Enable multi-tiered Storage Groups
  - Direct allocations to SSD devices MSR=1, or HDD devices MSR=10 if both exist
- **Defining the Guaranteed Space Attribute (honor volser and space request)**
  - Not recommended for DB2 user data sets, especially when using PAV and MA (avoids **MOST** hot spots). DB2 BSDS and active log data sets can still be hand placed using guaranteed space for availability as opposed to performance, even when using PAV and MA. Also review the SMS data set separation profile with your Storage Administrator. Guaranteed Space and the Separation Profile can be used together.
  - When you allocate a data set with Guaranteed space and a data set becomes multi volumes, the extend to a new volume for the secondary extent will actually be what the primary for the initial extent, not the secondary. This is true for DB2 data sets that are user managed, however for DB2 managed data sets the extend to a new volume will be created as the secondary amount. Review APARS PK83108/PK50113 as well.




- **Other Specs**
  - Allowance for multi-tiered SG (Storage Group) - see Storage Group section for additional information
  - Determine if PAV volumes should be part of the Storage Group selection - see Storage Group section for additional information
- **Recommendations:**

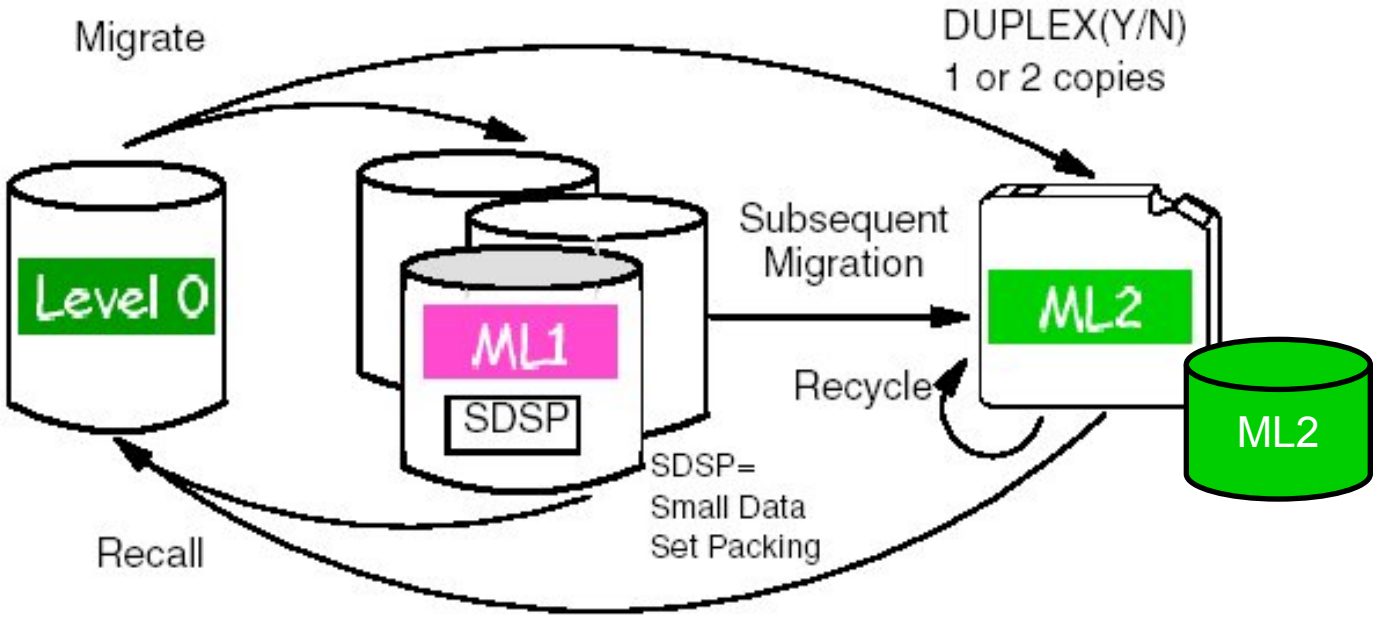
- If your installation has a mix of all types, models, and configurable disk:
  - Determine if specific targeted response time rates are required, as well as such things as RAID, and guaranteed space.
  - For example: If your installation is all ESS and DS8000, you may not need to worry about these attributes for performance when using PAVs and MA. Place your BSDS and active log data sets on the fastest devices possible (DS8000), leaving your DB2 user data on the ESS.
  - Sit down with your Storage Administrator and determine the best use of data and technology. Do not assume that your Storage Administrator understands DB2 as well as you do.

# Interesting Management Class Information for DB2 Professionals - page 1 of 3

- **A management class is a list of data set migration, backup and retention attribute values, as well as expiration criteria which uses DFSMShsm and DFSMSdss. Some issues to consider are:**
  - Requirements for releasing over-allocated space
  - Migration requirements
  - Retention criteria
  - Treatment of expired data sets
  - Frequency of backup
  - Number of backup versions
  - Retention of backup versions
  - Number versions
  - Retain only version
  - Retain only version unit
  - Retain extra versions
  - Retain extra versions unit
  - Copy serialization
  - Generation data group (GDG) information

- **Recommendations:**
  - BEWARE! If your Storage Administrator has turned on the 'Partial Release Attribute', DFSMShsm may compress any created VSAM data sets in extended format not using the guaranteed space attribute.
    - In a non-production environment releasing unused space in the DB2 LDSes can be of a great benefit.
  - Discuss with your Storage Administrator the 'GDG Management Attributes' if you are creating GDGs for such things as image copies. Determine how many GDGs should be retained on disk.
  - Keep in mind DR situations! ABARS may be used to backup your application's ML1 and ML2 data sets if required.
  - Objects related to DB2 commonly migrated:
    - Archive log data sets
    - Image Copy data sets
    - Some customers migrate table spaces and indexes. Review ZPARM values for RECALL and RECALLD. Not recommended, especially in production.

# Interesting Management Class Information for DB2 Professionals - page 3 of 3




## Space Management Activities - DFSMShsm

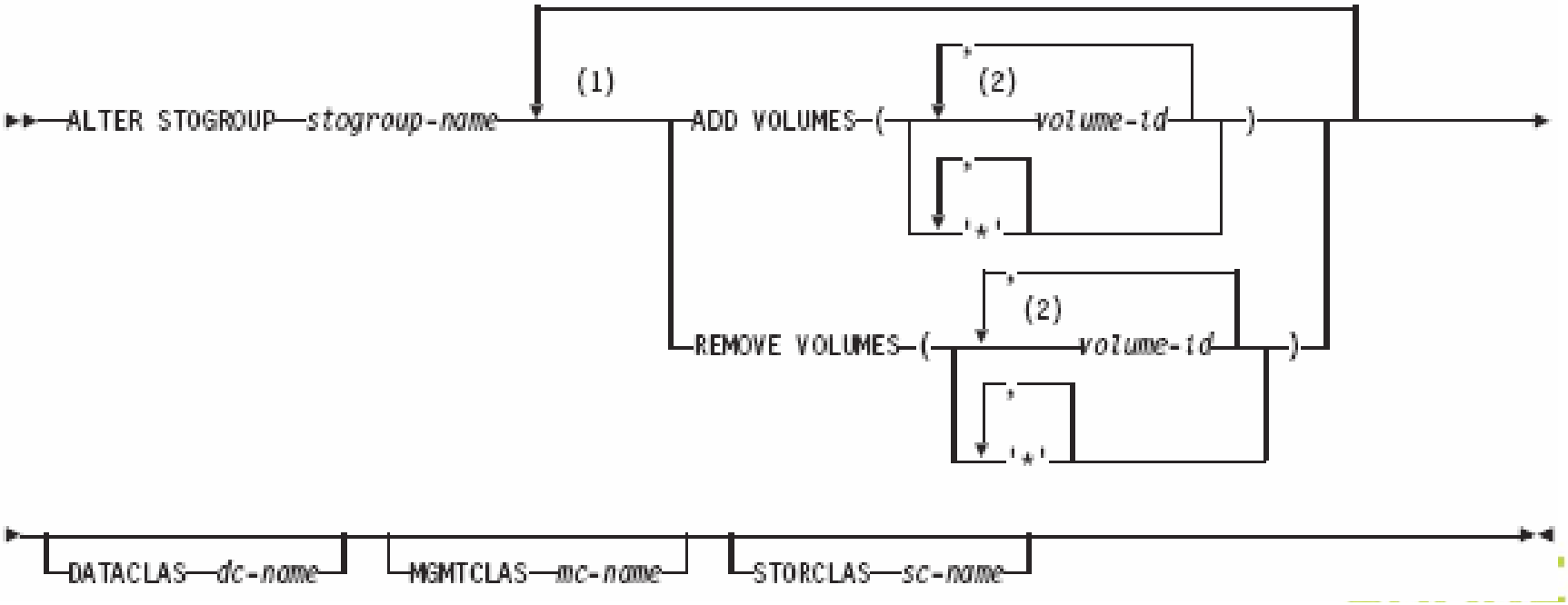
Level 0 - user volume where data is migrated from

ML1 (Migration Level 1) - DFSMShsm owned disk where data can be migrated. Data can then remain on this volume or further migrate to ML2.

ML2 (Migration Level 2) - DFSMShsm owned tape or disk (traditionally tape) where data can be migrated to either from ML1 or directly from level 0. Discuss with your Storage Administrator which approach works best for you.

# DB2 9 and SMS Classes

- Starting in DB2 9, you can use SQL with CREATE STOGROUP or ALTER STOGROUP to specify a DATACLAS, STORCLAS, and/or MGMTCLAS. You cannot specify an SMS Storage Group.
- VOLUME clause is now optional: it can be omitted if any of the DFSMS classes are specified.
  - **NOTE! If you do not add VOLUMES, SMS is in control – by default the DC has VC=1, DVC=NULL. No multi volume data sets!**



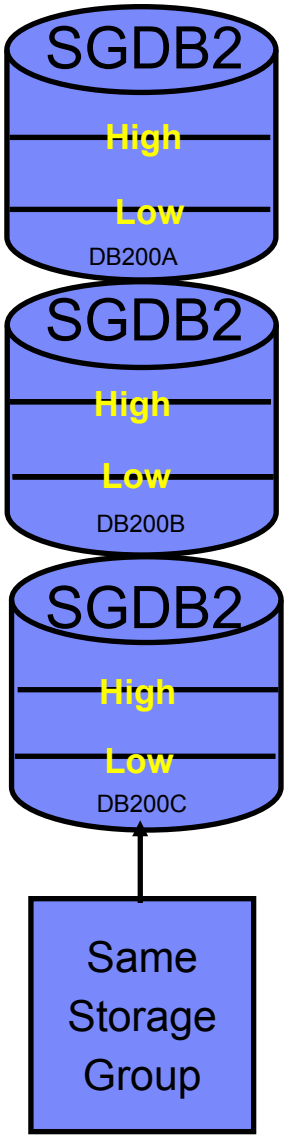

Note – not all Storage Administrators will allow you to use the SMS classes. Verify this before using SMS Classes with STOGROUP.

# What is an example of using STOGROUP with SMS Classes?

- For example, my DB2 LDSes go through a Data Class that does not have EF/EA enabled for partitions, LOBS, or XML objects greater than 4GB.
- The Storage Administrator created a special Data Class with EF/EA enabled.
- I create a DB2 Stogroup pointing to the special newly created Data Class.
- I create new partitions, LOBS, or XML objects that require a DSSIZE > 4GB. My CREATE statement includes USING STOGROUP created specially for data sets > 4GB in the previous step.

- **SMS uses the Storage Class attributes, volume and Storage Group SMS status, MVS volume status, and available free space to determine the volume selected for the allocation.**
- **DFSMShsm functions to migrate data sets, backup (data sets incrementally), and dump (volume level) are decided at the Storage Group level.**
- Recommendations:
  - For production DB2 environments, do not allow DFSMShsm to migrate or backup (data set level) table spaces or index spaces. However, depending on your installation's backup and recovery scenarios, you may want DFSMShsm to dump (full volume) your volumes, even while DB2 is up. It is much better though if DB2 is down if at all possible.
  - My recommendation for non production environments is the same as for production environments. However, some installations will allow data to be migrated. If you decide to allow for migration, keep in mind ZPARM values for RECALL and RECALLD. Some considerations are:
    - How many objects need to be recalled at the same time?
    - How many objects reside on the same DFSMShsm ML2 tape?
    - Will serial recalls complete in a timely manner?

SHARE
Technology · Connections · Results




# Migration and Allocation Thresholds - high and low values

## High and low values are set for all volumes in a Storage Group

**HIGH** - Used for disk volume allocation threshold

- If set to 75% for example, it will level all allocations to avoid volumes greater than 75% full
- If set to 75%, then it allows for 25% growth in extents
- Recommendation, start at 75% and then review
- Also used to determine if volume migration will be executed

**LOW** - Used for disk volume migration threshold

- Controls the low threshold by percentage of volume that eligible data sets can be migrated, thereby reducing stress on DFSMShsm
- e.g., set low value to 50%, migrate after 100 days and now 90% of volume is eligible, only migrate down to 50%
- Recommendation, For image copy volumes that need to totally empty all data sets to allow for space for the next image copy run, set the low value 0. Otherwise, set the value higher to reduce the stress on DFSMShsm for migration

# FAQ – What should HIGH be set to?

- It is not a one size fits all value.
  - For SMS Storage Groups housing active log and BSDS data sets – 99%. Active log data sets do not have secondary allocations and the BSDS has small secondaries. Data sets are not migrate, take LOW default.
  - For SMS Storage Groups housing the DB2 catalog and directory – 99%. Typically the DB2 catalog and directory sits on one or two volumes. You may want to consider lowering this value if they reside on many more volumes.
  - For SMS Storage Groups housing DB2 user data sets – IT DEPENDS! Between 75 and 85% may be a very good place to start as it would provide data sets the room to grow for secondaries.
    - If you have multiple Storage Groups based on size, you may want a higher value for your LARGE pool. You may need a much higher ceiling for large data sets.
    - Be careful when SGs have different flavors of emulated disk. 75% of a mod 3 is very different than 75% of a mod 54.
  - For SMS Storage Groups housing archive log data sets – 99%. If you have specified the correct ZPARM values for the primary allocation there should be no secondary allocations. Some customers do not have uniform sizes for the active log data sets, in which case you may need to reduce the value.
    - If you migrate the archive log data sets, the HIGH must be set lower to trigger the archive. Determine what size LOW is required based on the amount of data migrated.
  - For SMS Storage Groups housing image copy data sets – 99%. Typically allocations should fit into the primary application. Consider lowering the value If your utility or tool breaks down the allocation between primary and secondary.
    - If you migrate the archive log data sets, the HIGH must be set lower to trigger the archive. Determine what size LOW is required based on the amount of data migrated.

# Above HIGH Value Message

- Storage group high allocation threshold exceeded message Messages IGD17380I and IGD17381I directed to the console so they will be recorded in SYSLOG and may be detected by automation, are issued. These messages alert storage administrators when a storage group's space usage has gone above its high threshold.
- One of the places you will see this message besides SYSLOG is on the DBM1 job log. As an example:
- **IGD17380I STORAGE GROUP (*name*) IS ESTIMATED AT 86% OF CAPACITY,**

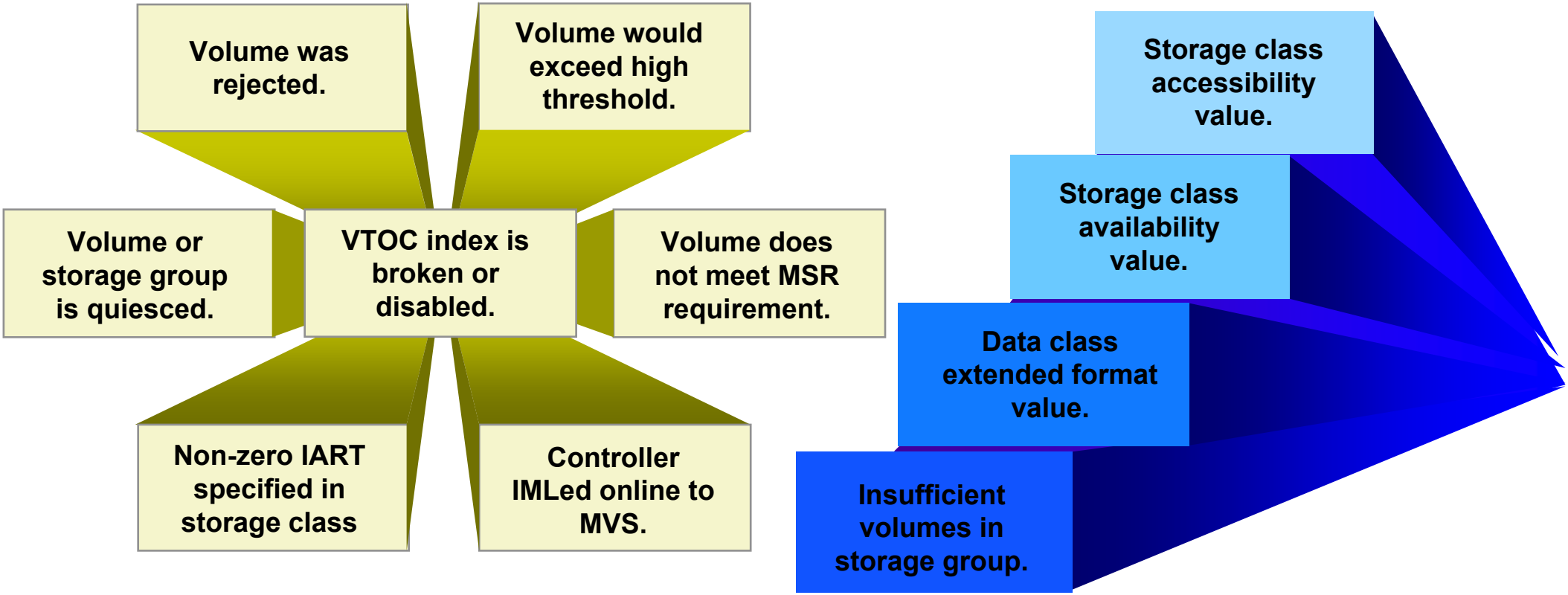
  **WHICH EXCEEDS ITS HIGH ALLOCATION THRESHOLD OF 85%**
- **Setting HIGH to 99% could defeat the purpose of receiving this message and taking action.**

- **Extend SG Name:**
  - Data sets can be extended from one Storage Group to another, if there is insufficient amount of storage in the primary group.
  - An extend SG name can also be an overflow Storage Group.
- **Overflow Storage Group:**
  - An overflow Storage Group is used when non overflow Storage Groups are above their thresholds.
  - An overflow storage group may also be specified as an extend storage group.
- Discuss with your Storage Administrator if using extend and/or overflow Storage Groups will help you.
- **Be aware of where your DB2 data sets are allocated. Review periodically for space related problems and inclusion for SnapShot and FlashCopy operations!**
- **COPY POOL BACKUP STORAGE GROUP - used starting in DB2 V8 to allow new BACKUP function.**

# Multi-Tiered Storage Groups

- **Specify Multi-Tiered SG Y in the Storage Class**
- **Example:**
  - SET &STORGRP = 'SG1', 'SG2', 'SG3'
- **Result:**
  - SMS selects volumes from SG1 before SG2 or SG3.
  - If all enabled volumes in SG1 are over threshold, then SMS selects from SG2.
  - If all enabled volumes in SG2 are over threshold, then SMS selects from SG3.
  - If all volumes are over threshold, then SMS selects from the quiesced volumes in the same order.

# Why Isn't My Volume Primary? - One example of why my data was allocated to a volume I did not expect




SMS can be set up to allow for volumes to be selected as primary, secondary, tertiary, or rejected. Discuss further with the Storage Administrator.

# Volume selection, is primary always what you want?

- SMS allows the System Resources Manager (SRM) to select a DASD volume from the primary candidate list. SRM has a very short history span, typically seconds, generally not more than 2 minutes.
- When SRM is used, all volumes are in the primary list first and SMS uses an elaborate weighing algorithm, with one of the factors being the most amount of free space. This technique does not pick volumes at random.
  - With SRM, when many data sets are created in a short period, SMS will at times place the data sets on the same volume and not randomly allocate data sets on other volumes.
- If random volume section is desired over SRM, set the Storage Class value for IART>0. Volumes are picked from the secondary candidate list and randomized.
  - May not be a good idea with data sets that have Guaranteed Space
- TEST, TEST, and TEST more! Make sure random volume selection works properly for you, especially when using overflow or spill volumes.

# CREATE STOGROUP VOLUMES

- For allocations with non-Guaranteed Space, specify VOLUMES("*") with one asterisk only. Using a real volser without GS results in unpredictable results under certain conditions.
- Using a non-SMS volser for VOLUMES does not make the volume non-SMS managed.

# Thinking about SMS managing all of your DB2 volumes (good idea!) and for STOGROUP specifying VOLUMES("*") (another good idea!)?

- Creating a STOGROUP and specifying VOLUMES("*") allows SMS to choose a volume for allocation.
- Here is the issue at hand – at times a Storage Administrator sets up the SMS ACS routines to check for specific volume types for allocation. Much of the time an * is not provided for the test. Without the ACS routine having an * as part of a valid test, your allocation will fail.
- If you will be creating STOGROUPs with VOLUMES("*") make sure your Storage Administrator adds an * to their test of device allocations in the ACS routines.



SHARE Technology · Connections · Results
SHARE in Anaheim 2011

# Differences Between DB2 STOGROUP and SMS Storage Group

| DB2 STOGROUP | SMS Storage Group |
|---|---|
| Different STOGROUPs can share the same disk volume(s) | One disk volume can only belong to one SMS Storage Group |
| VOLSERs are specific | Recommendation - code VOLUMES("*") to allow for SMS management. Avoid Guaranteed Space and specific VOLSERs where possible since this defeats the purpose of SMS. |
| SYSIBM.SYSVOLUMES has a row for each volume in column VOLID | When created with VOLUMES("*") SYSIBM.SYSVOLUMES has an * for column VOLID |
| Limited to management of 133 volumes | No volume limit |
| Volume selection based on free space | Volume selection based on SMS algorithms |

# z/OS 1.11 and 1.12 Enhancements useful for DB2

# z/OS 1.11 enhancements useful for DB2

- IDCAMS DELETE command is enhanced to include a new function called DELETE MASK. It allows users to specify the data set name selection criteria desired with a mask-entry-name and a keyword "MASK".
  - A mask-entry-name (also called as filter key) can have two consecutive asterisks (**) or one or more percentage signs (%)
  - MASK keyword is the keyword to turn on the new feature. For example,
    - *DELETE A.B.** MASK*
    - *DELETE A.BC.M%%K MASK*
  - AMS deletes only the first 100 data sets that matched the selection criteria
- SMS expands the data set separation function to the volume level to reduce I/O contention and risk of single points of failure
  - SMS provides new syntax for data set separation function
  - SMS attempts to separate data sets listed in volume separation groups onto different extent pools and volumes
- SMS striping restrictions lifted
- EAV can be used for more data types.

# z/OS 1.11 enhancements useful for DB2

- HSM
  - Provide the capability to specify a retention period on the data set backup command, which allows users to keep an individual backup copy for either a shorter or longer than normal period of time
  - Enable individual backup copies to be retained for an indefinite amount of time
  - Maintain more than 100 backup copies
  - Larger data sets (>58 K tracks) will now be DFSMShsm managed large format sequential data sets when migrated or backed up.
  - OVERFLOW volumes can now be used for migration and backup.
  - An installation can adjust the values that direct data sets to OVERFLOW versus NOOVERFLOW volumes and the threshold of the OVERFLOW volume pool.

- FlashCopy
  - Fast Replication data set recovery will use saved catalog information to recover deleted and moved data sets
  - The ALLOWPPRCP keywords from the FRBACKUP and FRRECOV commands can now be set in the SMS copy pool definition rather then specifying them on the command line
  - User catalogs will automatically be unallocated prior to a FRRECOV of a copy pool version
  - The ARC1803E message will include DFSMSdss messages and return codes for each failing volume
  - Fast replication copy pool backup storage group volumes will be automatically reinitialized when DFSMShsm has determined that the VTOC may become corrupt

# z/OS 1.12 enhancements useful for DB2

- Ability to open more data sets concurrently
- Enhance VSAM and VSAM RLS to reclaim empty space (control areas) in KSDSs. Applies to SMS and non-SMS data sets.
  - Will resolve the main reasons for KSDS reorganizations:
  - Reclaim space
  - Improve both sequential and direct performance
  - Can be used to reclaim space in an ICF catalog if running out of space or extents without down time
- EAV supports almost all types of data sets
- ICF catalogs can be allocated greater than 4GB when Extended Addressability (EA) is enabled. ICF Catalogs can grow beyond the 4 GB size to a maximum size of 4 GB times 32K depending on CI size.
- IDCAMS command DELETE GDG FORCE will no longer require that migrated GDSs be recalled before being deleted. GDSs having a volser of “MIGRAT” invoke HDELETE instead and will not recall the data sets first.

# z/OS 1.12 enhancements useful for DB2

- IDCAMS command DELETE is enhanced to allow users to specify a single asterisk (*) as the member name of a PDS or PDSE. This will delete all the members in that PDS or PDSE.
  - For example, DELETE A.B(*) will delete all the members residing in partitioned data set A.B
- MVS will call the PDSE validation tool to check the integrity of a PDSE in LNKLST when the system is coming up and when LNKLST is being changed. With this change, users will learn of PDSE structure errors earlier.
- Partial release can be used for Extended format (EF) data sets which are over allocated and span volumes (e.g. best fit data sets). Space on the last volume containing data and all unused empty volumes will be scratched and marked as candidates.

# z/OS 1.12 enhancements useful for DB2

- hsm
  - Using Fast Reverse Restore, you can:
    - Recover a copy pool without waiting for physical background copy to complete. This is a significant benefit as the background copy may take a very long time to complete.
    - Recover a copy pool from a disk version in either COPY or NOCOPY environment
    - Recover a copy pool from space efficient volumes in NOCOPY environment. Space efficient volumes are beneficial for certain DB2 related functions, such as cloning.
  - Restore from dump tape:
    - Multitasking volume recovery from dump tape can concurrently recover up to 64 volumes
    - Recovery of a copy pool version from dump can be done with single command. Before z/OS 1,12, A FRR (Fast Reverse Recovery) volume recovery command is needed for each volume when recovering a copy pool version from tape
    - Copy pools can be recovered from partially dumped versions
    - The Fast Replication Recover function will resume an interrupted recovery of a copy pool version unless RESUME(NO) is specified
    - Recover processing will reduce mounts and demounts by automatically scanning the Recover queue for additional recovery requests that need the currently mounted tape before demounting a tape.
  - hsm space management:
    - Perform Primary Space Management (PSM), Interval Migration (IM) and Command Volume Migration in multiple overlapping phases.

# Space problems

## Space related information not make sense? How did this happen?
## Primary 25 cylinders, secondary 1 cyl, total 70 cyl, extents=3

- Some reasons this can happen:
  - Part of the data set is allocated on an EAV in the cylinder managed area.
  - SMS adjacent extent reduction
  - DB2 sliding secondary
  - DB2 ALTER executed increasing the space without running REORG or LOAD REPLACE to recreate the data set
  - DEFRAG was run with the CONSOLIDATE keyword
  - dss was run with the CONSOLIDATE function
- What really caused this? There is no easy way to tell. It can be a combination of factors above.
  - 32K objects will throw off numbers as well, but will not cause the above issue.

# Want to see what the secondary really was?

- Using OMPE:

```
DB2PM GLOBAL(TIMEZONE(+7:00)
        INCLUDE (IFCID(258)))
        RECTRACE TRACE LEVEL(LONG)
       EXEC
```

```
START TRACE (PERFM)
 CLASS(30)
 IFCID(258)

IFCID 258 is already on with
statistics class 3, if class 3 is not
on you can limit the tracing by
specifying AUTHID, PLANID, etc.
```

PRIQTY=720 (1 cylinder) 180*4. SECQTY=1440 (2 cylinders) 360*4

High allocated before – 1 cylinder (180*4=720), after - 3 cylinders (4*(180+360)=2160)

2GB data set. 524288*4096=2147483648

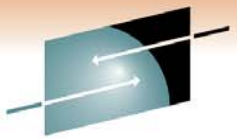
```
|-------------------------------------------------------------------------------------------------------
|DATA SET NAME      : DB9AU.DSNDBD.DSNDB04.JOHNEXT.I0001.A001                TIMESTAMP    : 06/30/10 19:21:10.736508
|DATABASE NAME      : DSNDB04                  DBID             :          4
|TABLESPACE NAME    : JOHNEXT                  PSID             :        562
|PRIMARY QUANTITY   :        180               SEC. QUANTITY    :        360
|HIGH ALLOC BEFORE  :        180               HIGH ALLOC AFTER :        540     MAX DS SIZE  :     524288
|EXTENTS BEFORE     :          1               EXTENTS AFTER    :          1     MAX EXTENTS  :        251
|VOLUMES BEFORE     :          1               VOLUMES AFTER    :          1     MAX VOLUMES  :         59
|-------------------------------------------------------------------------------------------------------
```

SMS managed data set – adjacent extent reduction

DFP max, not current data set

QUANTITY values must be multiplied my 4

Sliding secondary data set with adjacent extent reduction. Started out with a 1 cylinder data set with 1 extent, took an extent (equivalent to extent 2 – 2 cylinders) and the total became 3 cylinders (logical extent 2).




Dank u
Dutch

Merci
French

Спасибо
Russian

Gracias
Spanish

شكراً
Arabic

감사합니다
Korean

Tack så mycket
Swedish

धन्यवाद
Hindi

תודה רבה
Hebrew

Obrigado
Brazilian Portuguese

谢谢
Chinese

Dankon
Esperanto

# Thank You

ありがとうございます
Japanese

Trugarez
Breton

Danke
German

Tak
Danish

Grazie
Italian

நன்றி
Tamil

děkuji
Czech

ขอบคุณ
Thai

go raibh maith agat
Gaelic