

**SHARE**

Technology • Connections • Results

# What Do I Not Know About Modern Disk and Tape and Why do I Really Care

John Iczkovits [iczkovit@us.ibm.com](mailto:iczkovit@us.ibm.com)  
IBM

March 2, 2011  
Session #8386



# Disclaimer

- © Copyright IBM Corporation 2010. 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](http://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

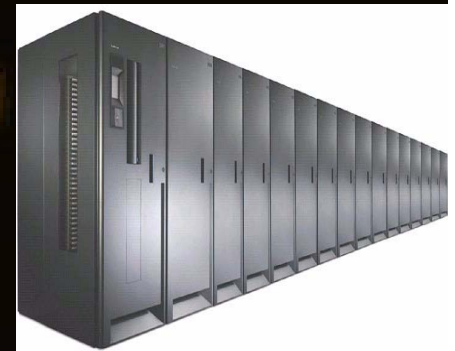
<http://www.redbooks.ibm.com/abstracts/sg247823.html?Open>



**Want to understand  
how DB2 works with tape?**

**Learn about DB2/tape  
best practices:**

**<http://www.ibm.com/support/techdocs/atmastr.nsf/WebIndex/PRS2614>**



## 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

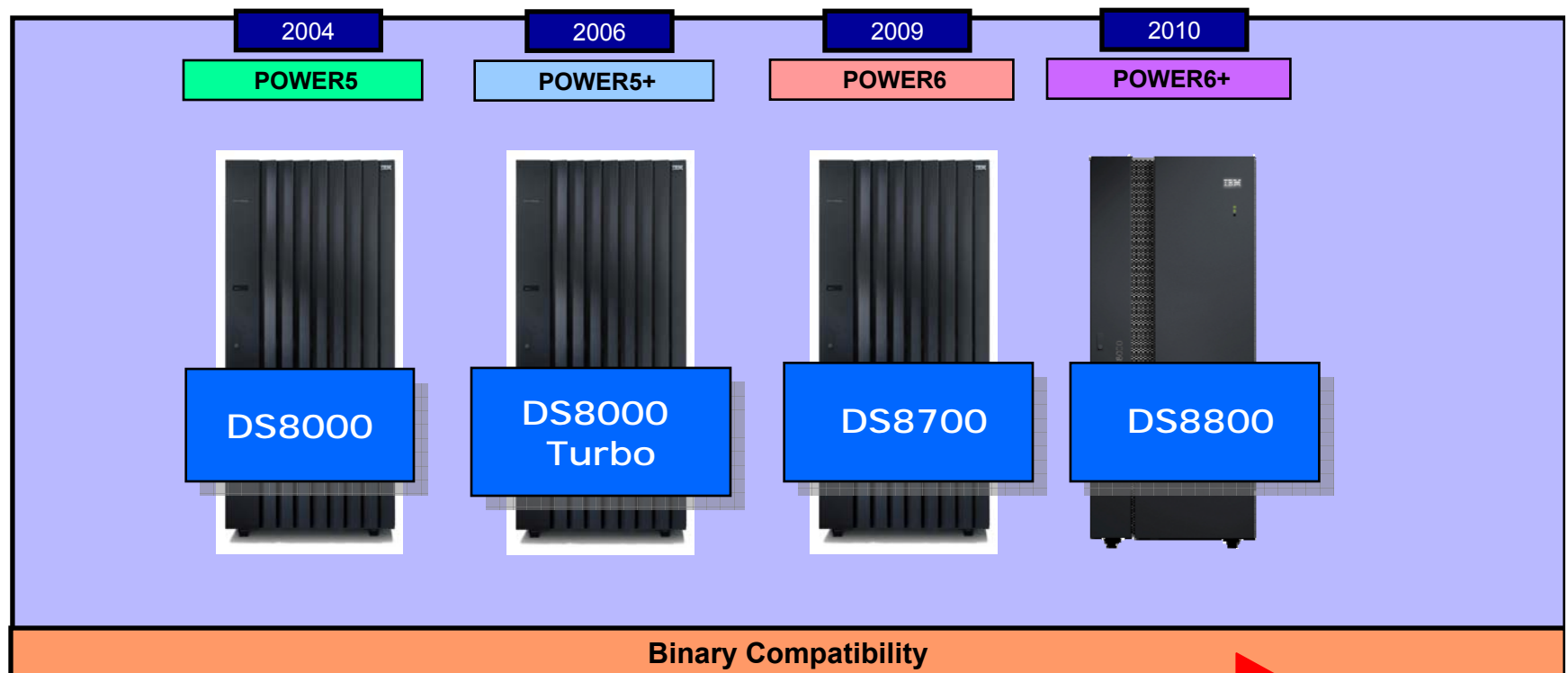


# 4<sup>th</sup>-generation DS8000 enterprise disk system



The IBM POWER processor has been behind the success of IBM enterprise storage beginning with the Enterprise Storage Server in 1999

**SHARE**  
Technology • Connections • Results



**DS8800 builds on a market-proven, reliable code base!**

**SHARE**  
in Anaheim  
2011

# DS8000 family models



*Two base models with scalable controllers and capacity*



## DS8700

- POWER6 controllers (2-way and 4-way)
- 4 Gb/s and 2 Gb/s host and device adapters
- 3.5" Enterprise Fibre Channel drives



## DS8800

- POWER6+ controllers (2-way and 4-way)
- 8 Gb/s host and device adapters
- 2.5" Enterprise SAS-2 drives

## Extents – what are they?

- Lots of discussions and opinions regarding extents.
  - Do they still matter?
  - How many extents are too many extents?
  - Based on how many extents should objects be REORGed?
- Extents are more logical than physical
  - Mostly people view an extent as a long stripe of data under one read/write head
  - In older technology, data was viewed from a vertical point of view. Tracks are placed in the same area on different read/write heads on different platters. Tracks refer to the read/write head. Depending on the technology, cylinder 1 track 1 would be on platter 1 using read/write head 1, cylinder 1 track 2 would be on platter 2 using read/write head 2, etc. Data is aligned vertically on different platters.
  - Newer technology does not use this approach. Tracks are spread out using a different approach. Tracks are no longer aligned with read/write heads on different platters.



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

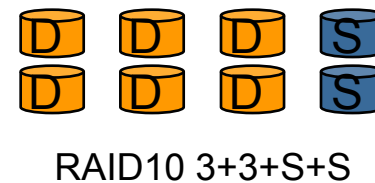
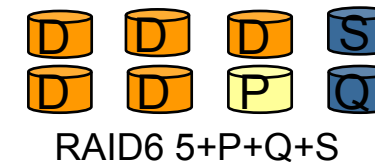
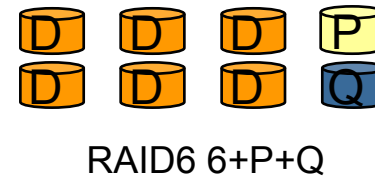
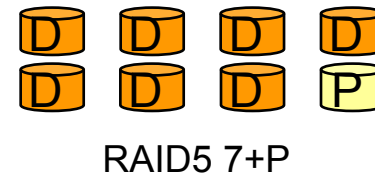
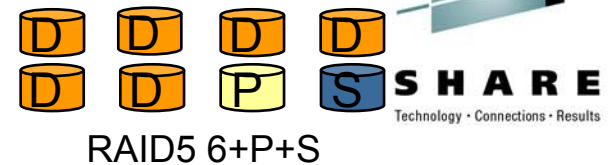
- RAID 0 - data striping without parity
- RAID 1 - dual copy
- RAID 2 - synchronized access with separate error correction disks
- RAID 3 - synchronized access with fixed parity disk
- RAID 4 - independent access with fixed parity disk
- **RAID 5 – data striping, independent access with floating parity (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**

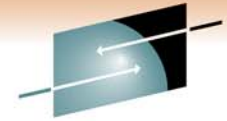
Parity is additional data, “internal” to the RAID subsystem, that enables a RAID device to regenerate complete data when a portion of the data is missing. Parity

- 9 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.

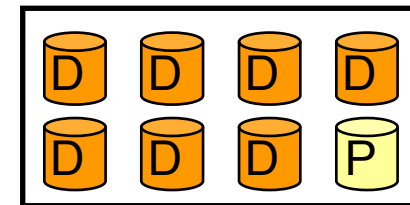
# Array

- 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





DS8000 CKD rank **SHARE**  
Technology • Connections • Results



RAID5 7+P

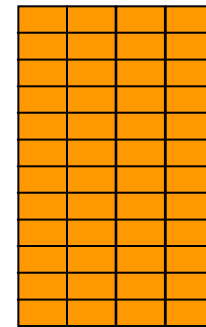
# Rank

- 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

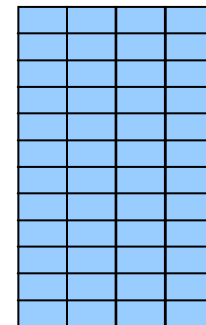
## Rank (continued)

- 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



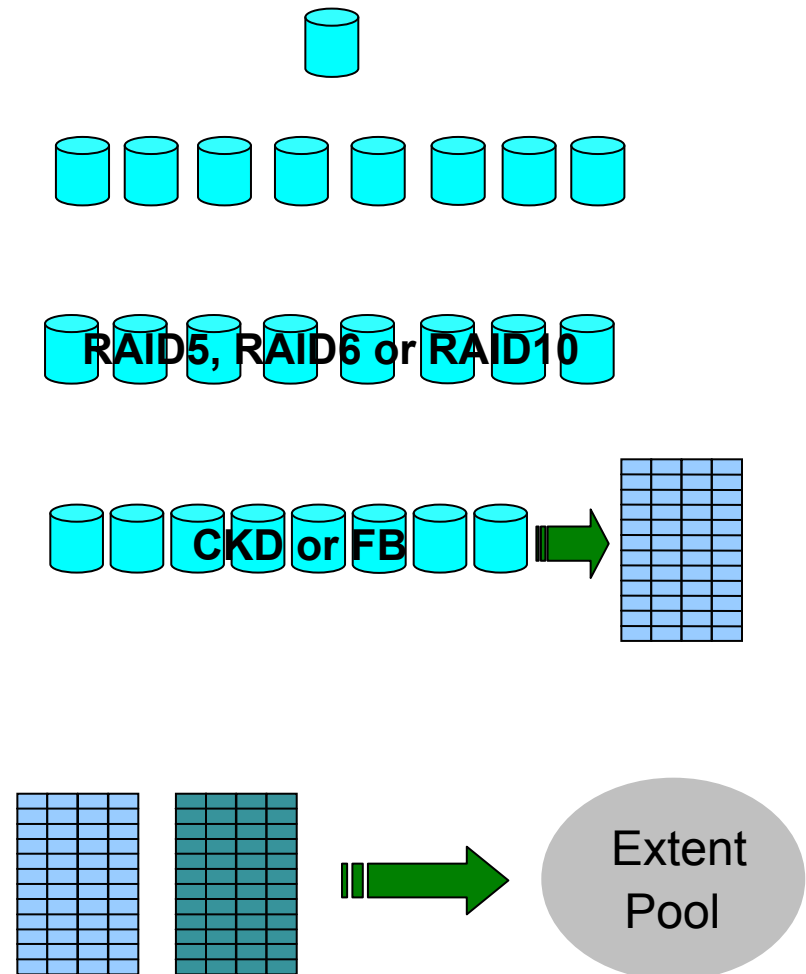
FB Rank



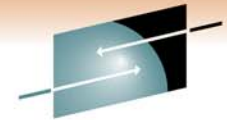
# Storage Resource Summary



- 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



**SHARE**

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 <b>EAV</b>	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.

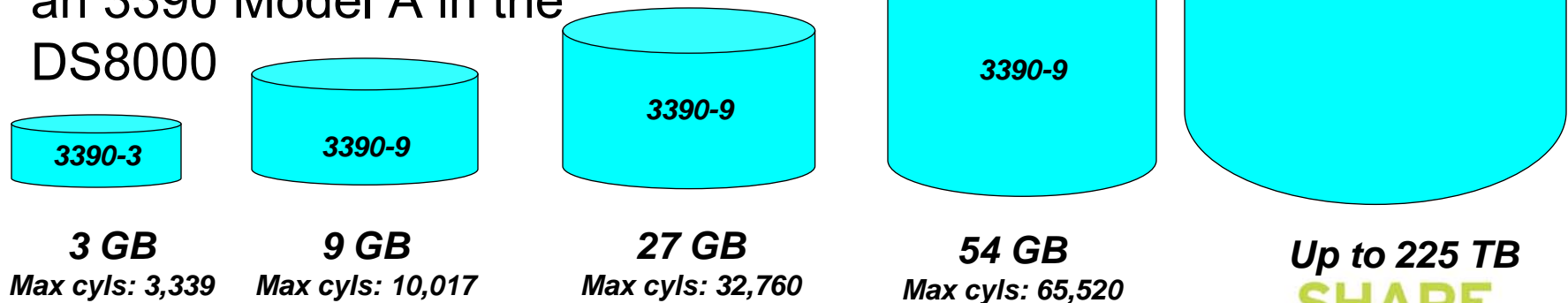
\*\* Bytes per track refers to the device capacity only. The actual allocation for a DB2 LDS data set is 48 KB, not the total of 56 KB. This allows for 12 - 4K DB2 pages to be allocated per track.

Why larger disk sizes?

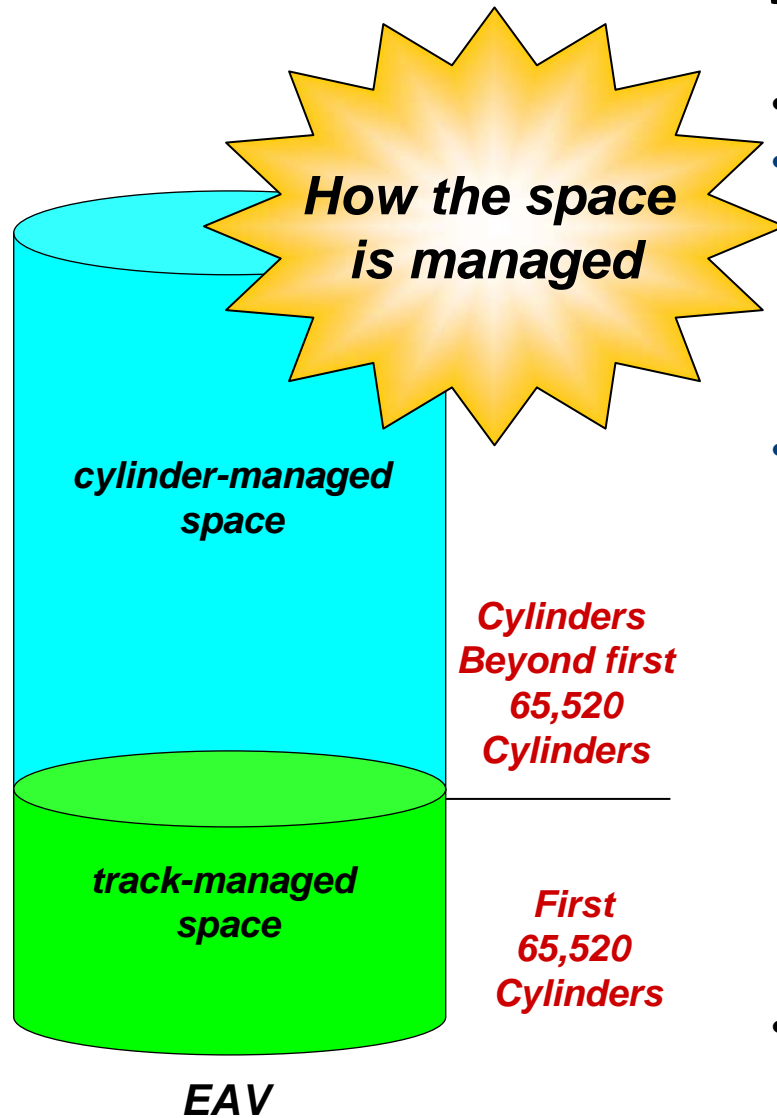
- reduces issues with MVS addresses (maximum number of devices met)
- simpler management of storage

# 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

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

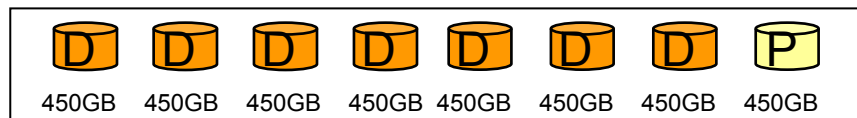
- Three different Fibre Channel (FC) DDM types (available in both non-encrypted and encrypted versions):
  - 146 GB, 15K RPM drive
  - 300 GB, 15K RPM drive
  - 450 GB, 15K RPM drive
  - See next page regarding 2.5 inch SAS drives for the DS8800, which includes 600 GB, 10K RPM drives
- One Serial Advanced 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.

# Storage Pool Striping

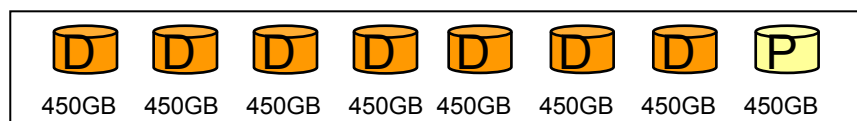
- Example of using storage pool striping with rotating extents across 3 ranks with RAID 5 (7+P) for a mod 3 device (2.83GB - 3,339 cylinders) – 450GB DDMs



**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)**

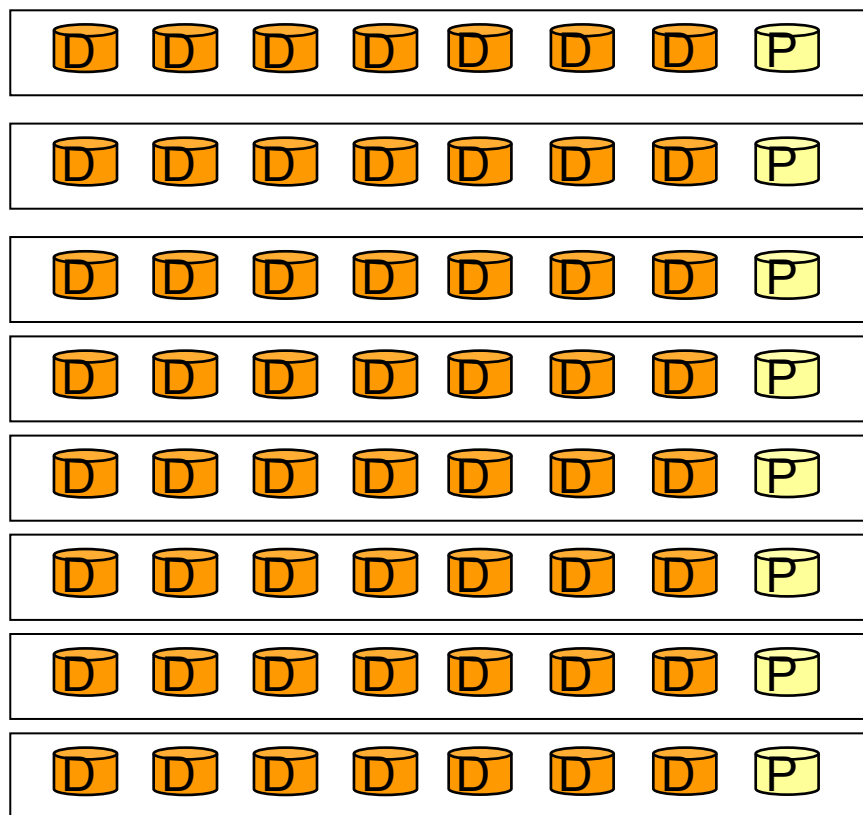
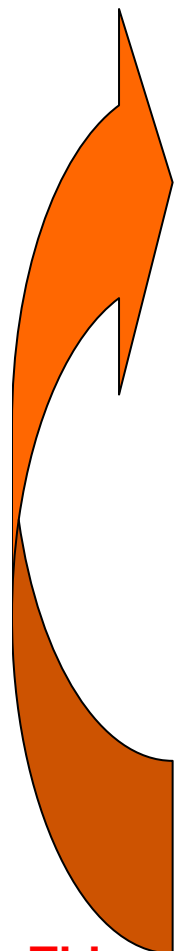
**In this scenario when you backup or Flash Copy one logical volume you are really accessing 24 physical volumes.**

- 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 skinned.

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



E  
results

- 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.

## Do extents really matter?

- On what physical volumes do my extent reside? No idea!
- 1 extent does not mean the data is necessarily contiguous, nor ever on the same array and rank
- From my very rudimentary tests:
  - 1 data set with 1 extent allocated with 100-200 cylinders of data
  - Have another data set with 100 extents allocated with 100-200 cylinders of data. Data inserted is the same across both data sets
  - SELECT \* for both data sets took about the same time.
  - Open time for the 100 extent data set was roughly double the one for the 1 extent data set. MVS open is one of the most expensive operations.
    - VSAM data sets can now be 7,257 extents when overriding the Data Class attribute for your LDS. If 100 extents took 2x open over the 1 extent data set, how long would it take to open a 3,000 extent data set?
  - Tests were based on read (SELECT) and not insert, which would take additional time to create the extents, as well as a very small additional time for sliding secondary if on

# After I REORG my data set in many extents, I always see performance improvements even on reads



- Does the problem relate to extents, or rather disorganized data?
- Many times the problem is not extents, rather some common problems that reorganizing the data resolves as opposed to the number of extents:
  - Pseudo deleted RIDs
  - Cluster ratio < 94% (fixed in DB2 V10)
  - High page splits

## LDS limit of 7,257 extents

- LDS (Linear Data Sets – almost all of your DB2 VSAM data sets) can now reach 7,257 extents. dfp rules state that an LDS can reside on:
  - $59 \text{ volumes} * 123 \text{ extents per volume} = 7,257 \text{ extents}$
  - 7,257 extents is the architectural limit. You have 10 volumes in the Storage Group,  $10 \text{ volumes} * 123 \text{ extents per volume} = 1,230$  extents. Another example, SG = 10 volumes, with no more than 40 extents per volume,  $10 \text{ volumes} * 40 \text{ extents per volume} = 400$  extents.
  - By default the maximum is still 255 extents per LDS. The Storage Administrator must override the 255 extent rule in the Data Class.
    - From my experience, most customers do not have Data Class assigned for the DB2 LDSes. This means that 255 extents is used as it is the default.
  - To exceed the 255 extent limit, data set must be SMS managed



# Why is it hard to test performance on LDSes with more than 255 extents?

**59 volumes \* 123 extents per volume = 7,257 extents (architectural limit)**



- Only SMS managed data sets can exceed 255 extents
- SMS uses adjacent extent removal
- Having an SMS managed data set reach 123 extents on 1 volume is very difficult – the volume would need to be extremely fragmented. Having 59 volumes that are extremely fragmented allowing for 123 extents per volume for an LDS is even more difficult.
- To test 7,257 extents, because of adjacent extent removal:
  - Turn off sliding secondary in ZPARMS (MGEXTSZ=NO). Do not run this test with other work for DB2 depending on sliding secondary.
  - Allocate a table space and insert an amount of data on 1 volume
  - Allocate a 1 track sequential data set residing right behind your LDS allocation so that the table space takes a new extent on next insert to avoid adjacent extent removal.
  - More inserts to fill up the next extent
  - Allocate a 1 track sequential data set residing right behind your next extent
  - Do this 123 times per volume on 59 volumes!
  - Want to try a smaller test, just 3,000 extents?
    - 123 extents per volume \* 25 volumes
      - *You may only get 119 extent per volume, to test 3,000 extents with 119 extents per volume, you will need 26 volumes*

## z/OS 1.12 Faster Data Set Open

- See APARs:
  - PM00068
  - PM17542
  - PM18557
- I have not yet tested the effects of these APARs in regards to DB2 and number of extents

## Disk Hardware

- 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.



## IBM System Storage DS8800 Architecture and Implementation SG24-8886

“Storage Pool Striping can enhance performance significantly, but when you lose one rank (in the unlikely event that a whole RAID array failed due to a scenario with multiple failures at the same time), not only is the data of this rank lost, but also all data in this Extent Pool because data is striped across all ranks. To avoid data loss, mirror your data to a remote DS8000.”

# Why not have just one large SMS Storage Group?

- Most Storage Administrators allocate the least amount of Storage Groups possible. This helps with overall workload for the Storage Administrator and management of the devices. Then when should multiple Storage Groups be used?
  - Availability - When using Flash Copy or any other full volume dump/restore mechanism, Storage Groups should contain data sets from only one DB2 Data Sharing group, or subsystem. If multiple ones exist on the same volumes, flashing back or recovering a volume will effect not just the intended member/subsystem, but rather all others as well.
    - Specific data sets, such as the BSDS and active log data sets that are duplexed should be on allocated at a minimal on different extent pools
  - Performance – although volumes are logical and many logical volumes can be placed on the same physical DDMs there are still times when applications do not coexist well from an I/O perspective and cause hot spots.
  - Space requirements – often architecting a large and small, or small, medium, and large pools based on space will avoid many space related problems. This is commonly a solution for times as when DB2 allocates a shadow data set for REORGs.

# What 3390 type should I emulate for my DB2 objects?

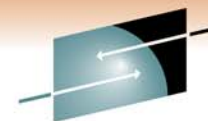
- Since volumes are logical, mod 1, 2, 3, 9, 27, 54, and EAV are recommended types. Some customers find a better fit by allocating for example mod 18 and 45 devices for their DB2 data sets, even though they are not device types commonly used. So long as space is used efficiently and within proper bounds, volumes can be allocated to fit your DB2 needs.
- With the DS8000, volumes can be dynamically expanded. This means that so long as the VTOC, VTOC index, and VVDS are large enough, a mod 3 for example can be dynamically expanded to a larger size emulated disk.
- How many data sets are or will be used for DB2?
- How large are the data sets? Most customers have a natural break between large and small data sets. The value between large and small is arguable. Most customers find the line between 200 and 400 cylinders. Typically 10% of the data sets are above 200-400 cylinders, 90% are below.
- Many customers still use 3390 mod 3 emulated disk. Allocating just 1 DSSIZE 64GB data set would require at least 32 mod 3 devices. If this same customer had 10 – 64GB data sets, at least 320 devices are required.
  - An EAV device can house not just 1 - 64GB data set, rather about 2-3. 2 mod 54 devices can house a 64GB data set.
  - Management of devices can be greatly reduced by allocating data sets to device types meant to hold a specific amount of data

## Tapes – what options are available?

- Virtualized tape – TS7700 family, VTS, TS7680. Use of disk with logical tape volumes that may create data sets on physical tape volumes depending on product installed (see next page).
  - Virtual tape can be used together. For example, a TS7720 and TS7740 can be used together almost as an hsm ML1 and ML2 approach. Data sets can be allocated to the TS7720 tapeless unit and then later migrated to the TS7740.
  - Interesting fact – the repository that manages tapes in IBM virtual tape systems is DB2 for LUW!
- ATL – Automated Tape Libraries. Not front ended by disk or using disk. Real physical tapes are used with robotics instead of manual intervention.
- Manually mounted tapes. No disk used.
- VTFM – virtual tape, but allocated on native disk (DS8000, etc.)
- TMM (Tape Mount Management) – tape data set allocations are converted to disk allocations

Blue can be used in the mainframe environment

# IBM Backup & Archive Portfolio



**SHARE**  
Technology • Connections • Results



## Tape Drives

## Tape Libraries

## Virtualization

- **LTO5 tape drive**
  - Encryption capable
  - 1.5TB Native capacity cartridge
  - Up to 140 MB/sec throughput
  - LTFS support
  - Dual ported drive
- **TS1130 tape drive/controller**
  - Third generation tape drive
  - Controller supports FICON & ESCON
  - Tape drive data encryption
  - 128, 640 and 1TB cartridge capacity
  - 32 to 160 MB/sec throughput
  - Auto Virtual Backhitch

- **TS3100 tape library (up to 19.2TB)**
- **TS3200 tape library (up to 38.4TB)**
- **TS3310 tape library (up to 316.8TB)**
  - Stackable modular design
  - LTO Tape drives
- **TS3400 tape library (up to 18TB)**
  - TS1100 tape drives
- **TS3500 tape library (up to 30PB with LTO5 or up to 15PB with TS1130)**
  - Linear, scalable, balanced design
  - High Availability
  - High Density
  - Fastest robotics in industry
  - LTO and TS1100 tape drive

- Mainframe Virtual Tape**
- **TS7720 (Virtual Tape)**
    - Tapeless
    - Up to 800 MB/s throughput
    - Up to 440 TB native cache
    - Standalone or GRID (PtP)
    - Up to 6-way Grid
    - Hybrid Grid support
  - **TS7740 (Virtual Tape)**
    - Up to 600 MB/s throughput
    - Up to 28 TB native cache
    - Standalone or GRID (PtP)
    - “Touchless” with Export options
    - Up to 6-way Grid
    - Hybrid Grid support
  - **TS7680G (Dedup)**
    - 600Mb/s INLINE
    - Up to 1PB Repository
    - 100% Data Integrity
    - Data / Disk Agnostic
    - Native Replication
    - High Availability

- Open Systems Virtual Tape**
- **TS7610 App Express (Dedup)**
    - 80Mb/s INLINE
    - 4TB & 5.4TB Useable capacity
    - 100% Data Integrity
    - Data Agnostic
    - Native replication
    - Many to one replication support
  - **TS7650 Appliance (Dedup)**
    - 500Mb/sec INLINE
    - 7TB to 36TB Useable capacity
    - 100% Data Integrity
    - Data Agnostic
    - Many to one replication support
    - High Availability (36TB option)
  - **TS7650G (Dedup)**
    - 1GB/s (Cluster) INLINE
    - Up to 1PB Useable capacity
    - 100% Data Integrity
    - Data / Disk Agnostic
    - Native replication
    - Many to one replication support
    - High Availability





# Which tape should I use for my DB2 environment?

**S H A R E**

Technology • Connections • Results

- First, what types of tapes are available for DB2?
- Keep in mind, when data sets are allocated on tape, even virtual tape, MVS returns a tape device to DB2. This means that even if the data set actually resides on disk, requests are serialized because MVS understands the device to be a tape.
- How many tapes are required? How much data do they need to hold? What are the requirements for write time as well as recall time?
- If recalling using hsm, are the SMS Storage Groups large enough to hold the amount of data required for recalls as well as for daily processing?
- If you are using manual tapes and migrating DB2 LDSes, make sure ZPARM values for RECALL and RECALLD are set correctly to avoid resource unavailable messages.
- Make sure that dual copied tapes do not wind up on the same physical devices. If there is the possibility, make sure another copy is available, such as using a grid system, or for hsm manages tapes using a duplexed ML2 tape, or copying and sending off one of the copies.
- When using virtual tape, make sure your tapes are allocated onto a large enough virtual tape via the SMS Data Class to avoid multi volume tapes when not necessary.
- When allocating archive log data sets to the TS7740 or VTS, with the understanding that the data sets will be retained on disk for at least 24 hours, make sure there is enough disk available to contain the data sets.

## Where should I place my archive log data sets? E jts

- What is available? Disk, tape, if tape, what type of tape?
- Keep 24-48 hours of combined active log and archive log on disk for fast recovery.
  - For fast and concurrent recovery, make sure archive log and other required data sets reside MVS acknowledged real disk.
- Is hsm being used for recall back to the SMS Storage Group for faster recovery?
- Review the previous page

# What happens if I lose a physical disk or tape?

- Disk
  - It depends on the RAID solution. For example, RAID5 and you lost 1 DDM in an array – volume should be automatically rebuilt utilizing the data + floating parity. RAID 5, but you lost 2 DDMs in an array – unrecoverable, you have lost all of the logical volumes associated with the failed DDMs. With RAID5 and a double DDM outage, but Hyperswap is available, disk would be switched to the mirrored set without an outage.
- Tape
  - It depends on the type of failure. Most outages can be resolved if a grid system is in place. Keep in mind the large capacity of tapes. What happens if a tape snaps, and there is no grid in place or copy of the data sets on the tape? In this case your data is lost.

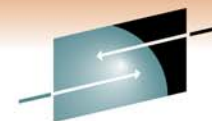
## What relationship is there between SMS and virtual disk and tape?

- Very little
- From a disk perspective SMS sees volumes as we do – from a logical volumes perspective
- SMS still sees disk and tape from older storage perspective, not virtualized
- VTOCs and other structures used for disk for example do not know what DDMs, arrays, ranks, etc. are. VTOCs do not recognize that virtualization exists
- Understanding virtualization is very important to understand how your technology works with DB2

# Now that I understand SMS to physical mapping, should I still permit the use of Guaranteed Space?



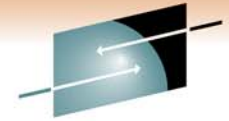
- The SMS Storage Class Guaranteed Space attribute does what it sounds like:
  - The amount of space you requested must be available. With GS, space cannot be reduced by such features as the Data Class space constraint relief.
  - The volser requested is the only eligible volser in the Storage Group
- GS is not enforced (it is ignored) for DB2 managed data sets, only enforced for user managed data sets
  - With GS, primary allocations are propagated for the first allocation of each volser for multi volume data sets (not for DB2 managed data sets)
- Use of GS can result in very unpredictable consequences. Do not specify GS unless required. From a DB2 perspective, consider the use of GS in combination with the Separation Profile for the BSDS and active log data sets **only**.



**SHARE**

Technology • Connections • Results

# Questions?



**SHARE**  
Technology • Connections • Results  
**Gracias**  
Spanish

**Dank u**  
Dutch

**Merci**  
French

**Спасибо**  
Russian

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

**SHARE**  
in Anaheim  
2011