

IBM

IBM TotalStorage®

## SMS Volume Selection

Or: Why did my data set go there???



z/Series Expo  
Session Z30  
September, 2005  
Ruth Ferziger  
[ruthf@us.ibm.com](mailto:ruthf@us.ibm.com)

© 2003 IBM Corporation

IBM

IBM TotalStorage®

## Trade Marks

DFSMSdftp	DFSMSdss	DFSMShsm
DFSMS/MVS	DFSORT	IPCS
RAMAC	SnapShot	FlashCopy
z/OS	ESS	RVA
ESCON	PAV	

© 2003 IBM Corporation

IBM TotalStorage®

IBM

## Abbreviations

DC/Dataclas	Data class
DFW	DASD fast write
EOV	End of volume
IART	Initial access response time
MSR	Millisecond response
PCU	Physical control unit
PPRC	Peer to peer remote copy
SC/Storclas	Storage class
SDR	Sustained data rate
SG/Storgrp	Storage group

© 2003 IBM Corporation

IBM TotalStorage®

IBM

## Volume Classification

The diagram illustrates the classification of volumes. At the top, a light green box labeled "Selected Volume" is positioned above a large, downward-pointing arrow. This arrow is composed of four segments, each representing a different volume type: "Primary volume" (light blue), "Secondary volume" (medium blue), "Tertiary volume" (dark blue), and "Rejected volume" (very dark blue). The segments are arranged horizontally and point downwards towards a common point at the bottom of the arrow.

© 2003 IBM Corporation

IBM TotalStorage®

IBM

## Data Set Separation

- Allows you to designate groups of data set which are to be physically separated
- SMS attempts to allocate the data sets behind different control units
- A data set separation profile must be provided
- The name of the data set containing the profile must be specified in the SMS base configuration
- Cannot be used with non-SMS-managed data sets or with full volume copy utilities such as PPRC

© 2003 IBM Corporation

IBM TotalStorage®

IBM

## Recommended Use of Separation

- Use only when absolutely required to ensure no single point of failure
- Use only for a small set of mission-critical data sets
  - ▶ Volume rejection because of separation may drastically reduce the number of eligible volumes
  - ▶ Data set separation can affect system performance
  - ▶ Take care when using separation with striping
  - ▶ May require constant updating if used with GDGs

© 2003 IBM Corporation

## Specifying Separation

```
SEPARATIONGROUP | SEP
    FAILLEVEL | FAIL( { PCU | NONE } )
    DSNLIST | DSNS | DSN (data set name{, data set name,...});
```


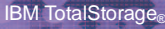
Example:

```
/* SMS CONTROL Data Sets */
SEPARATIONGROUP FAILLEVEL(NONE) -
    DSNLIST (SMS.SCDS, SMS.ACDS, SMS.COMMDS);

/* JES CHECKPOINT Data Sets */
SEP FAIL(PCU) DSNS(SYS1.JESCKPT1, - ! primary
                  SYS1.JESCKPT2, - ! secondary
                  SYS1.JESCKPT3); ! tertiary
```

## Multiple Separation Profiles


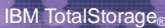
- You can create multiple separation profiles in different data sets or PDS members
- You can only specify one separation profile in the configuration base
- If you have multiple configurations, they can all share the same profile
- Or they can all have separate profiles
- Profile is read when SMS initializes or restarts and whenever a new configuration is activated



## When Separation Does Not Work

- The allocation is not SMS-managed.
- The allocation is performed on a down-level system.
- The separation profile cannot be accessed.
- The separation profile is invalid.
- The allocation uses a temporary data set name.
- Two data sets are allocated on different systems.
- A volume is varied online during allocation.
- An IODF change occurs during allocation.

© 2003 IBM Corporation



## When Separation Does Not Work

- A data set name not in the profile is specified during HSM recover.
- The profile was modified after configuration activation.
- SMS does not perform separation during:
  - ▶ Rename.
  - ▶ HSM migration to level 1 or 2.
  - ▶ Full volume image copy.

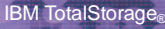

© 2003 IBM Corporation

## Conventional Volume Selection

- Used for all non-striped data set
- Used for all data sets with zero or blank SDR
- Uses a preference sequence to sort volumes in the candidate storage groups into:
  - ▶ Primary
  - ▶ Secondary
  - ▶ Tertiary
  - ▶ Rejected

## Volume Selection Evaluation Process

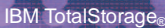

Criteria	Preferences	Pref. Value
VIO	A VIO storage group is made eligible by the storage group ACS routine and is eligible only when the data set is VIO capable	8192
Data Set Separation	Volume not on same PCU as data set from which it is separated.	4096
Volume Count	Volume is in a storage group that can satisfy the volume count.	2048
High Threshold	Volume has sufficient space without exceeding high threshold.	1024

## Volume Selection Evaluation Process

Criteria	Preferences	Pref. Value
SMS Status	Volume and its storage group are both enabled.	512
EOV Extend	For EOV extend, volume does not reside in the extend storage group.	256
Non-Overflow	Volume resides in a non-overflow storage group.	128

© 2003 IBM Corporation

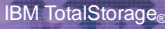




## Volume Selection Evaluation Process...

Criteria	Preferences	Pref. Value
IART	Volume is mountable and IART specified is non-zero.	64
Snapshot	Volume is on same snapshot capable controller as the data set.	32
Accessibility	Controller for volume supports accessibility & value is PREF. Or controller does not support it and value is STANDARD	16
Availability	Controller for volume supports availability & value is PREF. Or controller does not support it and value is STANDARD	8



© 2003 IBM Corporation



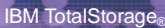




## Volume Selection Evaluation Process...

Criteria	Preferences	Pref. Value
Extended Format	Volume is on a control unit that supports extended format and IF EXT is PREF.	4
Millisecond Response (MSR)	Volume provides the requested response time specified in direct or sequential MSR.	2
	Volume provides a faster response time that requested in the direct or sequential MSR	1

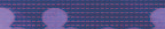




© 2003 IBM Corporation

## Example

<b>Volume A</b>	
Does not have sufficient space	0
Volume & storage group are enabled	512
Volume is in a non-overflow storage group	128
Volume is on a control unit that supports accessibility & value is PREF.	16
<b>Total preference value for Volume A:</b>	<b>656</b>
<b>Volume B</b>	
Has sufficient space	1024
Volume & storage group are enabled	512
Volume is not in a non-overflow storage group	0
Volume is not on a control unit that supports accessibility & value is PREF.	0
<b>Total preference value for Volume B:</b>	<b>1536</b>

© 2003 IBM Corporation



IBM TotalStorage<sup>®</sup>

IBM

## The Primary List

- Meet data set separation requirement
- SMS storage group and volume statuses are enabled
- MVS status is online
- IART requirement is met
- Number of volumes in storage group  $\geq$  volume count
- Accessibility requested can be met
- Availability requested can be met
- Meets the guaranteed space requirement
- Can perform the allocation & stay below high threshold
- For MSR=999, volume is non-cached
- Data class extended format request can be met

© 2003 IBM Corporation

IBM TotalStorage<sup>®</sup>

IBM

## The Secondary List

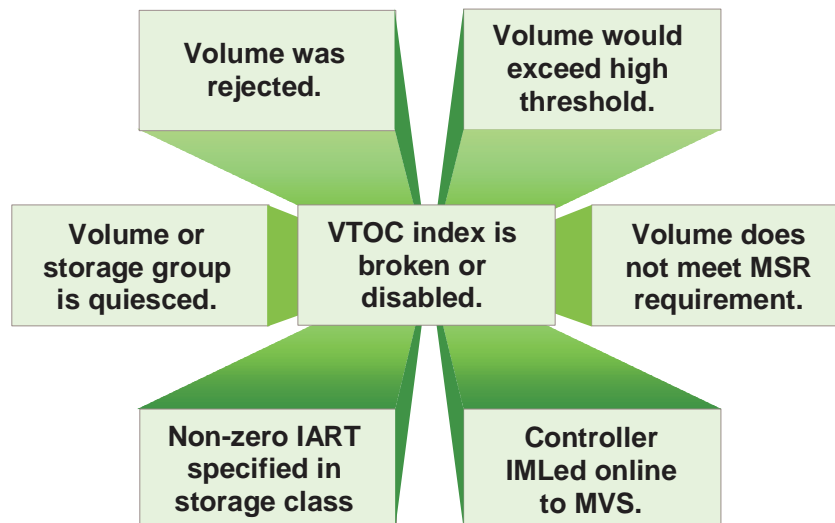
- ABEND X37 prevention - the most available space
- Meet data set separation requirement
- Meet volume count requirement
- Can perform the allocation without going more than 20% over high threshold
- SMS storage group and volume status
- Honors tiering of storage groups
- Spill/Overflow volumes
- Volume characteristics
  - ▶ Availability
  - ▶ Accessibility
  - ▶ Extended format
  - ▶ Guaranteed space
- Mount time performance

© 2003 IBM Corporation

## The Tertiary List

- Only used for:
  - ▶ Non-guaranteed space requests
  - ▶ Non-VSAM data sets
- Consists of volumes in storage groups that do meet the volume count requested

## Why Isn't My Volume Primary?



IBM TotalStorage® IBM

## Why Isn't My Volume Primary?

Storage class accessibility value.

Storage class availability value.

Data class extended format value.

Insufficient volumes in storage group.

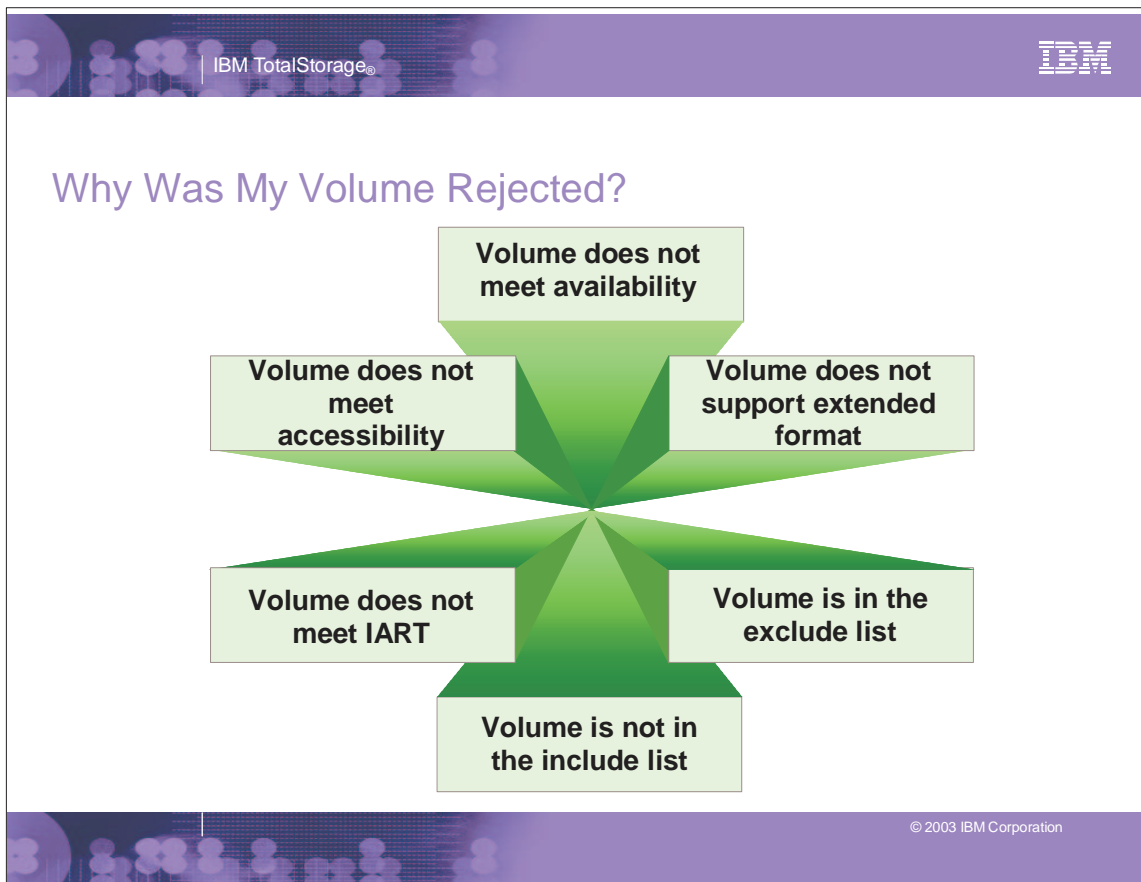
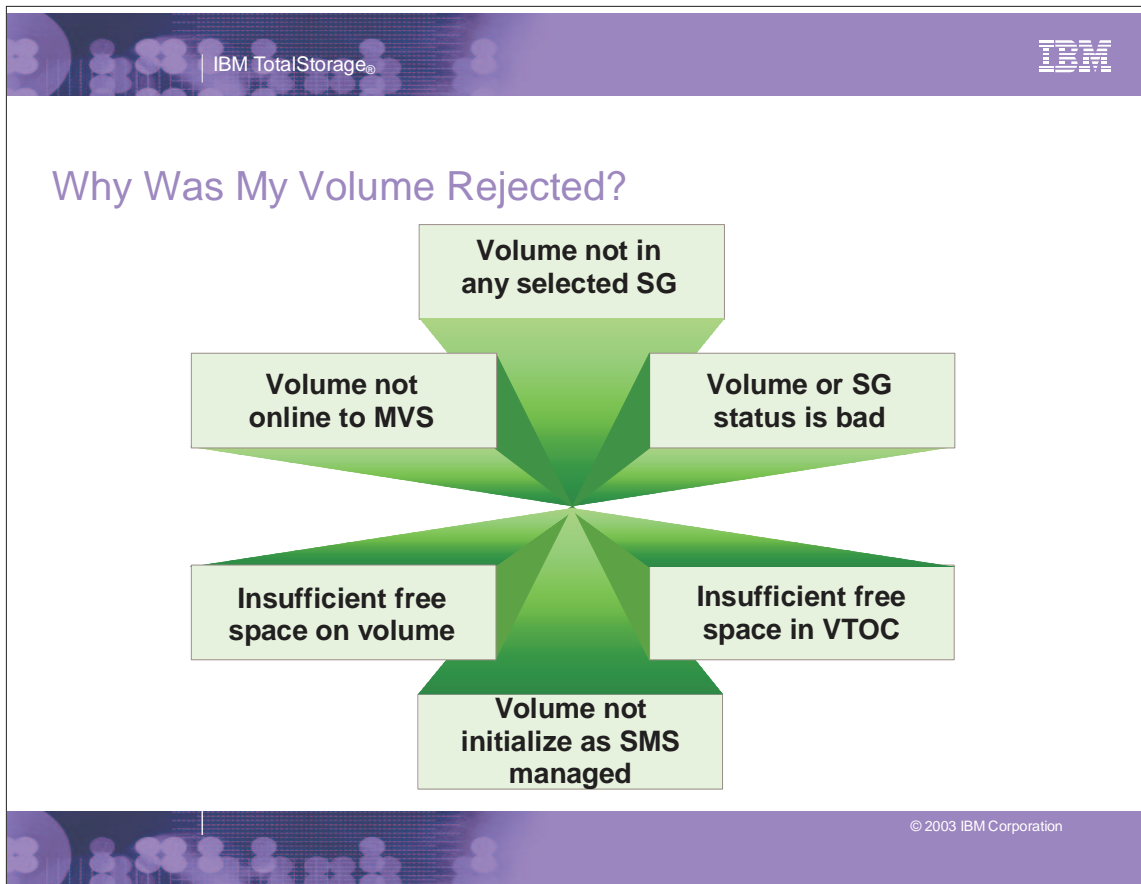
© 2003 IBM Corporation

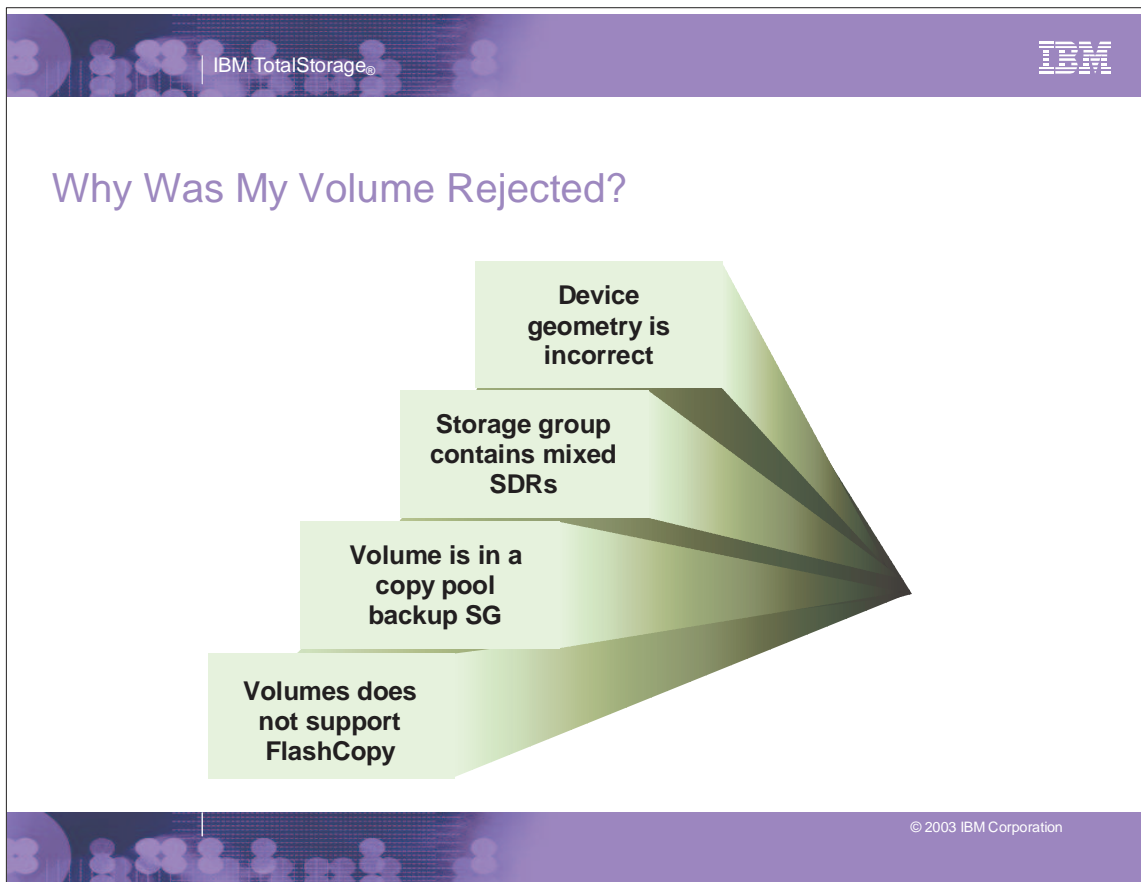
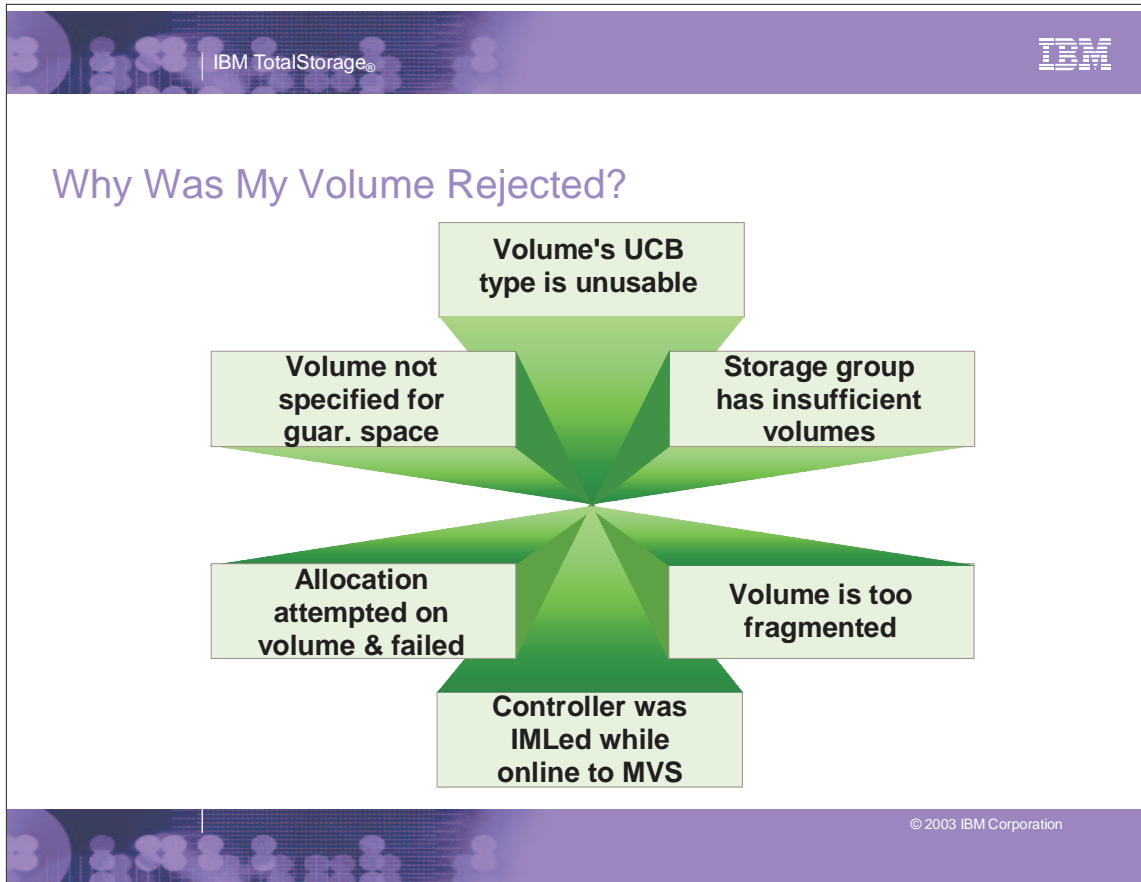
IBM TotalStorage® IBM


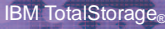
## Why Isn't My Volume Primary?

**RAID devices  
are eliminated if  
AVAILABILITY  
is STANDARD**

© 2003 IBM Corporation




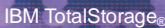




## Wrong Volume Selected?

- Check construct assignments
- Check channel path utilization
- Check storage group/volume utilization
- No volumes in primary volume list
- Expected volume was on tertiary list
- Expected volume was rejected
- Products which hook into system code (such as SRM) can create unexpected results

© 2003 IBM Corporation



## Effect of MSR and Bias

- Uses only the stored MSR; cached if cache is active, native otherwise
- Devices close to the requested MSR are placed on the primary list
- Devices not close to the requested MSR are placed on the secondary list

© 2003 IBM Corporation

## Effect of MSR and SDR

- MSR is 999
  - ▶ Devices with a native value of 25 are placed on the primary list
- SDR is greater than 0
  - ▶ Causes striping volume selection to be used
  - ▶ May cause MSR, availability, accessibility, and free space criteria to be ignored
  - ▶ Considers controllers over volume attributes



© 2003 IBM Corporation

## Effect of Availability

Availability	Effect
Continuous	Dual Copy and Array DASD only
Preferred	Array DASD preferred
Standard	Simplex volumes preferred
No Preference	All volumes except dual copy considered equal



© 2003 IBM Corporation








## Effect of Accessibility

First Choice	Second Choice	Third Choice	Specify:
Versioning device	None	None	Access=C Version=Y Backup=N
Method 1 backup device	None	None	Access=C Version=N Backup=Y
Versioning or backup device	None	None	Access=C Version=blank Backup=blank
Versioning or backup device	Non-accessibility device	None	Access=P Version=blank Backup=blank






© 2003 IBM Corporation

## Effect of Accessibility....

First Choice	Second Choice	Third Choice	Specify:
Versioning or method 2 backup device	Method 1 backup device	Non-accessibility device	Access=P Version=Y Backup=Y
Versioning device	Non-versioning device	None	Access=P Version=Y Backup=N
Non-accessibility device	Versioning or backup device	None	Access=S Version=blank Backup=blank
Any device	None	None	Access=N Version=blank Backup=blank

© 2003 IBM Corporation

IBM TotalStorage<sup>®</sup>

IBM

## 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

© 2003 IBM Corporation

IBM TotalStorage<sup>®</sup>

IBM

## Parallel Access Volumes

- Feature of the Enterprise Storage Server (ESS)
- Available only when the PAV option is enabled
- Use the Parallel Access Volume Storage Class attribute:
  - ▶ Required: Only volumes with the PAV feature are selected
  - ▶ Preferred: Only volumes with the PAV feature are primary
  - ▶ Standard: Only volumes without the PAV feature are primary
  - ▶ Nopreference: All volumes, PAV and non-PAV are treated equally

© 2003 IBM Corporation

IBM TotalStorage®

IBM

## Secondary/Tertiary List Preferencing

1. Data set separation
2. Storage group volume count
3. Can perform the allocation without going more than 20% over high threshold
4. High threshold
5. SMS status of volume/storage group
6. Honors tiering of storage groups
7. IART greater than zero
8. Preferred volume for fast replication requirement
9. Accessibility value
10. Data class extended format value
11. Availability value
12. MSR value
13. Never cache (MSR=999)

© 2003 IBM Corporation

IBM TotalStorage®

IBM

## Spreading Allocations across Volumes

- Storage Class Attributes
  - ▶ MSR = Blank
  - ▶ Bias = Blank
  - ▶ Accessibility = NOPREF
  - ▶ Availability = NOPREF
  - ▶ Guaranteed Space = N
- Data Class Attribute
  - ▶ If Ext = Blank

© 2003 IBM Corporation

## Striping Volume Selection

- Used only for:
  - ▶ Initial allocation of extended format preferred or required data sets with SDR > 0
  - ▶ Recall/Recover of multistripe data sets
- Not used for Recall/Recover of single-striped multivolume data sets
- Selects storage groups based on unique controllers
- Attempts to meet requested MSR

© 2003 IBM Corporation

## Striping Volume Selection....

1. No SGs with mixed device types
2. Volumes must remain < threshold
3. Enabled/quiesced treated the same
4. Number of volumes computed from SDR
5. Temporary data set with volume count > 1 treated as non-striped
6. Volume must be able to satisfy primary space requested

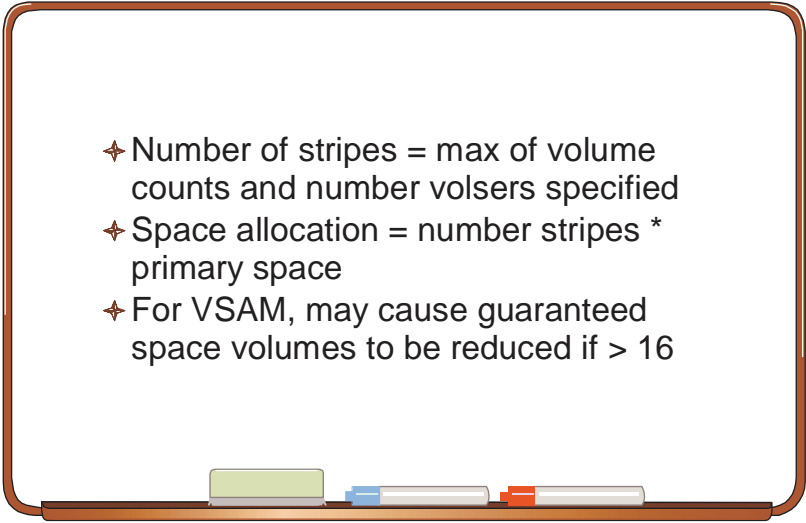
© 2003 IBM Corporation

IBM TotalStorage<sup>®</sup>

IBM

## Striping and Guaranteed Space

- ✦ Number of stripes = max of volume counts and number volsers specified
- ✦ Space allocation = number stripes \* primary space
- ✦ For VSAM, may cause guaranteed space volumes to be reduced if > 16



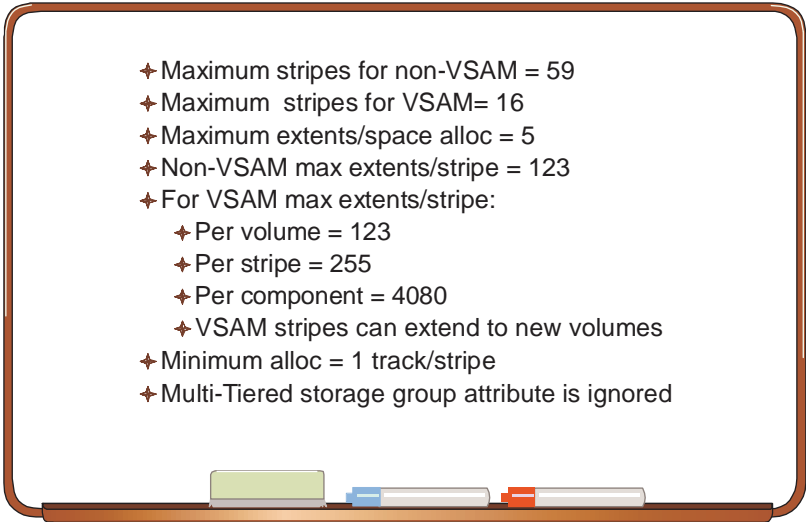
© 2003 IBM Corporation

IBM TotalStorage<sup>®</sup>

IBM

## Rules for Striped Data Sets

- ✦ Maximum stripes for non-VSAM = 59
- ✦ Maximum stripes for VSAM = 16
- ✦ Maximum extents/space alloc = 5
- ✦ Non-VSAM max extents/stripe = 123
- ✦ For VSAM max extents/stripe:
  - ✦ Per volume = 123
  - ✦ Per stripe = 255
  - ✦ Per component = 4080
  - ✦ VSAM stripes can extend to new volumes
- ✦ Minimum alloc = 1 track/stripe
- ✦ Multi-Tiered storage group attribute is ignored



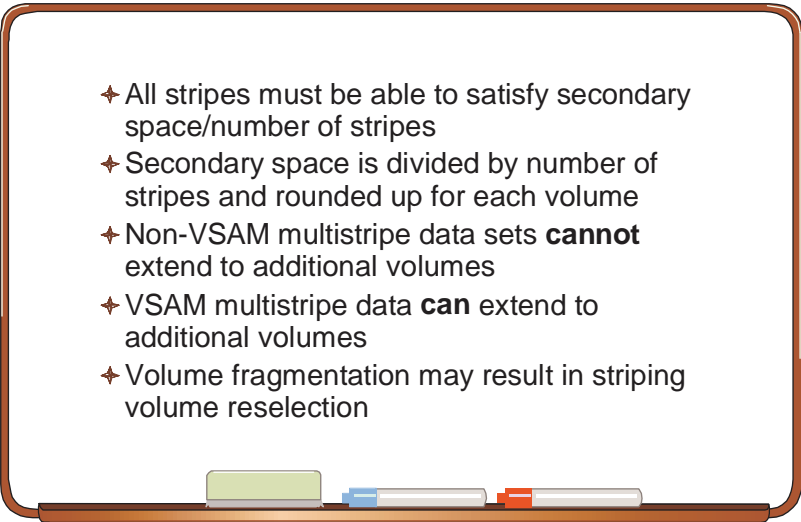
© 2003 IBM Corporation

IBM TotalStorage®

IBM

## Extending Striped Data Sets

- ✦ All stripes must be able to satisfy secondary space/number of stripes
- ✦ Secondary space is divided by number of stripes and rounded up for each volume
- ✦ Non-VSAM multistripe data sets **cannot** extend to additional volumes
- ✦ VSAM multistripe data **can** extend to additional volumes
- ✦ Volume fragmentation may result in striping volume reselection



© 2003 IBM Corporation


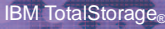
IBM TotalStorage®

IBM

## Requirements for Striping

- Volumes behind one of the following controllers:
  - ▶ ESCON-attached controller that supports T0 copy
  - ▶ 3990-6 controllers
  - ▶ 3990-3 controllers with Extended Platform that are ESCON-attached
  - ▶ 3990-3 controllers with RAMAC support-level microcode
  - ▶ 9394 controllers
  - ▶ 9343 controllers with cache
  - ▶ IBM RAMAC Virtual Array
  - ▶ IBM Enterprise Storage Server
- Volumes must be ENABLED or QUIESCED and varied ONLINE

© 2003 IBM Corporation





## If All Else Fails....

Data class contains two values which can be used to influence volume selection:

- Space Constraint Relief
- Reduce Space Up To (%)
- If you specify the second, the first must be Y

© 2003 IBM Corporation



## If You Use Space Constraint Relief...

- Very large allocations may succeed with large enough volume count.
- Existing data sets may end up with less space than requested on extents.
- New data sets may be smaller than requested.
- Fewer extents may be available when the data set extends.
- May result in more than 5 primary extents
- X37 abends should occur less frequently.

© 2003 IBM Corporation



IBM TotalStorage<sup>®</sup>

IBM

## How Many Extents Are Allowed?

- Non-VSAM, non-extended format: Up to 16 on the volume
- Non-VSAM, extended format: Up to 123
- PDSE and HFS: Up to 123 on the volume
- VSAM: Up to 255 per component, but only 123 per volume per component
- Striped VSAM: Up to 4080 per data component

© 2003 IBM Corporation

IBM TotalStorage<sup>®</sup>

IBM

## The Retry Process...

- If the volume count is 1:
  - ▶ SMS retries the allocation after reducing the requested space as indicated
  - ▶ SMS removes the 5 extent limit
- If the volume count is greater than 1:
  - ▶ First, SMS uses best-fit volume selection
  - ▶ If this fails, SMS reduces the space quantity and removes the 5 extent limit

© 2003 IBM Corporation

IBM TotalStorage<sup>®</sup> IBM

## Requesting Assistance

- Turn SMS tracing on:
  - ▶ SETSMS TRACE(ON),TYPE(ALL),SIZE(1M),DESELECT(ALL),  
SELECT(MSG,VTOCC,VTOCA,MODULE),JOBNAME(jobname)
- Run the job
- Turn SMS tracing off
  - ▶ SETSMS TRACE(OFF)
- Make note of the dump data set name
- Take a dump of the SMS address space
  - ▶ DUMP COMM=(any dump title you desire)
  - ▶ R #,JOBNAME=SMS,CONT
  - ▶ R #,SDATA=(LPA,CSA,ALLNUC,GRSQ,LSQA,SWA,PSA,SQA,TRT,RGN,SUM)


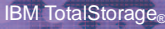
© 2003 IBM Corporation

IBM TotalStorage<sup>®</sup> IBM

## Requesting Assistance....

- Activate IPCS from a TSO session.
- Set the defaults (dump data set name) using option 0
- Go to the IPCS COMMAND option (IPCS option 6)
- Issue: VERBX SMSDATA 'TRACE'
- If possible, use IPCS PRINT to create a hard copy of the trace


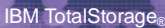
© 2003 IBM Corporation



## References: APARs

- II07464 - reasons for volume selection failure
- II08004 - reasons why wrong volume selected
- II08442 - volume selection and DCME settings
- II08618 - striping volume selection information
- II08987 - continuation of II08004
- OW04270 - extended format data set volume selection
- OW08472 - volumes and MSR capabilities
- OW08630 - storage class AVAILABILITY options

© 2003 IBM Corporation



## References: Publications

- z/OS DFSMSdfp Storage Administration Reference (SC26-7402)
- z/OS DFSMS: Implementing System-Managed Storage (SC26-7407)
- MVS/ESA SML: Managing Storage Groups (SC26-3125)
- z/OS DFSMSHsm Storage Administration Reference (SC35-0422)
- z/OS DFSMSHsm Storage Administration Guide (SC35-0421)
- z/OS DFSMSdss Storage Administration Reference (SC35-0424)
- z/OS DFSMSdss Storage Administration Guide (SC35-0423)
- z/OS DFSMS Advanced Copy Services (SC35-0428)

© 2003 IBM Corporation