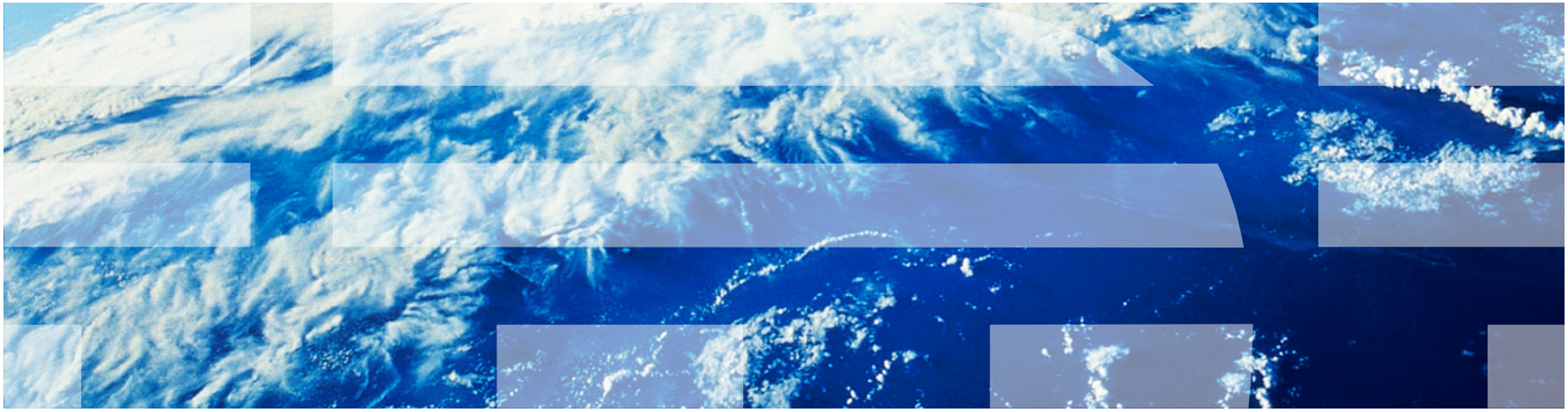
z/VM I/O Enhancements

Steve Wilkins
wilkinss@us.ibm.com
10/20/2011



Trademarks

The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

z/VM® z10™ z/Architecture® zEnterprise™

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

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

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

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

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

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

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Disclaimer

The information contained in this document has not been submitted to any formal IBM test and is distributed on an "AS IS" basis without any warranty either express or implied. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

In this document, any references made to an IBM licensed program are not intended to state or imply that only IBM's licensed program may be used; any functionally equivalent program may be used instead.

Any performance data contained in this document was determined in a controlled environment and, therefore, the results which may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environments.

All statements regarding IBM's plans, directions, and intent are subject to change or withdrawal without notice, and represent goals and objectives only. This is not a commitment to deliver the functions described herein

What's New in z/VM I/O Support?

- This presentation provides insight into those “significant” updates that describe new I/O features and facilities supported by z/VM.

- The contents of this presentation is expected to change over time, as new function is made available via the service stream or regular releases.

Topics

- Single System Image (SSI) enhancements
- Live Guest Relocation
- Extended Remote Copy (XRC)
- Hyperswap Improvements
- Extended Address Volumes (EAV)
- FlashCopy / Space-Efficient

Single System Image (SSI)

SSI – Channel-to-Channel Adapters (CTCAs)

- Required for ISFC communication between members of the SSI Cluster
  - Every member in the cluster needs to have at least one
  - Installation allows you to define up to two connections between members; more can be added later if needed/desired
- Combine multiple CTCAs into a single logical link between members
  - 1-16 CTCAs can be combined in this fashion
- Use two logical links, on distinct CHPIDs, for the purposes of redundancy in the event of a failure
- Four CTCAs per FICON CHPID seems to be the maximum configuration that can be supported before saturating the channel during a relocation

IOCP – CTCA Example

```
CHPID PCHID=143,PATH=(CSS(3),F2),TYPE=FC,SWITCH=42,PART=(MCPX4)

CNTLUNIT CUNUMBR=1340,PATH=(CSS(3),F2),LINK=(19),UNIT=FCTC,
              UNITADD=((00,16))

IODEVICE ADDRESS=(1340,16),CUNUMBR=1340,UNIT=FCTC,UNITADD=00
```

Config – CTCA Example

```
/*********************************************************/
/* Activate ISLINK statements                            */
/*********************************************************/

/*********************************************************/
/* ISFC links for Member 1                               */
/*********************************************************/
VMSYS01: Activate ISLINK 0A2B 1A2B Node VMSYS02
VMSYS01: Activate ISLINK 0A2C 1A2C Node VMSYS03
VMSYS01: Activate ISLINK 0A2D 1A2D Node VMSYS04

/*********************************************************/
/* ISFC links for Member 2                               */
/*********************************************************/
VMSYS02: Activate ISLINK 0B2A 1B2A Node VMSYS01
VMSYS02: Activate ISLINK 0B2C 1B2C Node VMSYS03
VMSYS02: Activate ISLINK 0B2D 1B2D Node VMSYS04
```

SSI – Device Equivalency

- The Single System Image cluster requires a mechanism for identifying devices on different LPARs that are either identical or equivalent
  - Preferably in a fashion that does not rely on items that may be different between CECs and LPARs, such as device numbers
  - z/VM will automatically generate this identifier wherever possible, but will provide a mechanism for a human to define an identifier for the scenarios where they are needed, but cannot be created
- These identifiers will be required in preparation of a guest relocation, to ensure that a guest will have access to the same resources before and after the relocation
  - Any problems that occur with these identifiers will prevent a relocation from occurring, and a helpful error message will be displayed

SSI – Device Equivalency (cont.)

- The term Equivalency ID (EQID) is used for this identifier, but comes in two forms
  - System-generated identifiers
    - Generated at VARY ON, or device initialization during IPL
    - Visible as EBCDIC, 50 characters in length
    - Cannot be modified or overwritten
  - User-generated identifiers
    - Generated at RDEVICE level, and can be set using its command or configuration file statement
    - Limited to eight EBCDIC characters

SSI – Device Equivalency (cont.)

- Certain devices, such as DASD, need to be identified as identical, in order to access the same data both before and after relocation
  - Identifiers need to be unique per device
- Certain devices, such as OSA, need to be equivalent, in order to access the same resources (e.g., network) both before and after relocation
  - Identifiers need to be unique per equivalent resource
  - For example, subchannels on a given CHPID, could be equivalent. But depending on the device it could instead be true that they are different themselves, as each subchannel could be accessing unique data (think NPIV-enabled FCP subchannels)

SSI – Device Equivalency Query

```
q eqid for 1290-1295
1290: EQID = MYOSA
1293: EQID = 002107900IBM750000000000004191000400000000000FFFF
1295: EQID = POOL1
Ready;
```

SSI – Device Equivalency Commands

- ATTACH command can be used to dedicate a resource to a guest (or z/VM SYSTEM) based on this identifier, instead of a device number or volume serial
  - In the event of multiple devices with the same identifier (such as an OSA), the first matching RDEV that is FREE will be used
- DETACH can also use this information, but the command will fail if multiple matching devices are attached to a resource
- Directory entries (DEDICATE, MDISK, etc.) were NOT expanded to do the same thing

SSI – Data Integrity

- The SSI Cluster necessitates the specification of ownership information for CPOWNED volumes, to ensure that z/VM in a particular LPAR doesn't allocate warm start, checkpoint, page, spool, temp disk, or directory data on a volume owned by another member
  - Ownership information will be the combination of both the z/VM system (LPAR) and SSI Cluster names
  - Can be applied to non-SSI system, if desired

- Established and permanently recorded by disk utilities like CPFMTXA and ICKDSF, and is propagated with DASD copy utilities
  - OWN and SHARE parameters on CPOWNED statements/commands are ignored when IPLing in an SSI cluster, deferring to the information on this disk and the configuration in which it is being used

SSI – QUERY CPOWNED

```
SLOT VOL-ID RDEV TYPE  STATUS                    SSIOWNER SYSOWNER
1    M01RES C4A0 OWN   ONLINE AND ATTACHED      CLUSTERA VMSYS01
2    ------ ---- ----- RESERVED                  -------- --------
3    ------ ---- ----- RESERVED                  -------- --------
4    ------ ---- ----- RESERVED                  -------- --------
5    VMCOM1 C4FF OWN   ONLINE AND ATTACHED      CLUSTERA --------
6    ------ ---- ----- RESERVED                  -------- --------
7    ------ ---- ----- RESERVED                  -------- --------
8    ------ ---- ----- RESERVED                  -------- --------
9    ------ ---- ----- RESERVED                  -------- --------
10   M01S01 C4A8 OWN   ONLINE AND ATTACHED      CLUSTERA VMSYS01
11   M02S01 C4B8 SHARE ONLINE AND ATTACHED      CLUSTERA VMSYS02
12   M01S02 C4A9 OWN   ONLINE AND ATTACHED      CLUSTERA VMSYS01
13   M02S02 C4B9 SHARE ONLINE AND ATTACHED      CLUSTERA VMSYS02
14   M01S03 C4AA DUMP  ONLINE AND ATTACHED      CLUSTERA VMSYS01
15   M02S03 C4BA DUMP  ONLINE AND ATTACHED      CLUSTERA VMSYS02
16   ------ ---- ----- RESERVED                  -------- --------
...
253  ------ ---- ----- RESERVED                  -------- --------
254  M01P02 C4C1 OWN ONLINE AND ATTACHED        CLUSTERA VMSYS01
255  M01P01 C4C0 OWN ONLINE AND ATTACHED        CLUSTERA VMSYS01
```

Data Integrity (What extents are brought online in an SSI?)

SSI Name on Disk	System Name on Disk	PERM or PARM (shareable)	SPOOL (owner shared)	PAGE, TDISK, DRCT, WARM, CKPT (private)
NULL	NULL	NO	NO	NO
NULL	This System	YES	YES (Owner, single-member cluster only)	YES
NULL	Other System	NO	NO	NO
This SSI	NULL	YES	NO	NO
This SSI	This System	YES	YES (Owner)	YES
This SSI	Other System (within this SSI)	YES	YES (Shared)	NO
This SSI	Other System (outside of SSI)	NO	NO	NO
Other SSI	Anything	NO	NO	NO

Data Integrity (What extents are brought online in a non-SSI?)

SSI Name on Disk	System Name on Disk	PERM or PARM (shareable)	SPOOL (owner shared)	PAGE, TDISK, DRCT, WARM, CKPT (private)
NULL	NULL	YES	YES	YES
NULL	This System	YES	YES (Owner)	YES
NULL	Other System	NO	YES (Shared)	NO
Any SSI	Anything	NO	NO	NO

SSI – Minidisk Handling

- Enhancements to the z/VM LINK code will perform additional checking when in an SSI (both for LINK command and LINK directory statement)
  - This means communication between SSI members!
  - Link conflict analysis now applies to all members of the SSI cluster, to ensure data integrity when a new minidisk is created
  - Ensures Minidisk Cache does not expose data integrity problems for minidisks links on multiple systems (not cross-system cache)
- Minidisks can be shared with systems outside the SSI (SHARED=YES)
  - But, z/VM provides no management/security of this behavior, and is ineligible for minidisk cache
  - BE CAREFUL!!!

SSI – Minidisk Handling (cont.)

- SET MDC will communicate with the SSI cluster, and turn off/on minidisk cache appropriately
- No cross-system link checking performed for private minidisks, temp disks, virtual disks, or minidisks created with DEFINE command
  - Also not performed when the LINK can be fully resolved based on existing links within the system
- Non-FP Minidisks defined to allow virtual reserve/release (“V-suffix” in directory), or with working allegiance (“WRKALLEG”) can only be linked to one member of the SSI cluster at a time
  - Also, they will cause the linking user to be ineligible for relocation

SSI – Minidisk Query

- Query LINKS output will display all userids that hold a link to a minidisk, regardless of what system they are on in the SSI cluster
- Query MDISK DETAILS output has been expanded to display whether a minidisk is restricted to a particular system (e.g., VDISK) or is available across the SSI cluster
  - Local = This minidisk is restricted to a particular member in an SSI Cluster
  - Global = This minidisk is not restricted to any member in an SSI
  - Local – DEFINE MDISK = Created via DEFINE MDISK command
- Query MDISK has a new RESERVED operand that will display the reservation state of a minidisk
- Query MDC can inform the user if MDC is suppressed for a VDEV because of a write link on a local or remote system

SSI - Commands

```
query links 191
SYSTEM1 : userid1  0191 R/W, userid2  0191 R/0, userid3  0191 R/0
SYSTEM2 : userid1  0191 R/0, userid4  0191 R/0
Ready;
```

```
query mdisk 190 details
TargetID Tdev OwnerID  Odev Minidisk  DEVNO Duplex  Scope
REESELIS 0191 REESELIS 0191 Regular   No    No      Local
REESELIS 0495 REESELIS 0495 Regular   No    No      Global
REESELIS 0690 REESELIS 0690 Regular   No    No      Local - DEFINE MDISK
Ready;
```

Live Guest Relocation

Live Guest Relocation

- Guests that are to be relocated need to meet certain criteria, in order to be made “eligible” for relocation
  - Short version: Standard Linux configurations
  - Long version: There's a long list of pros/cons

- One of these requirements is for the guest to enable (via directory entry) both CHPID and Path Group ID Virtualization
  - Previously, guests would see the real information as presented by LPAR
  - As Path Groups would differ between LPARs, and CHPIDs may not match across LPARs/CECs, z/VM needs to mask this information from guests in order to prevent changes from being exposed across a relocation
  - Both are enabled by Option CHPIDVirtualization ONE or GLOBALOpts CHPIDVirtualization ONE statements in the directory
    - Default is CHPIDVirtualization OFF

Extended Remote Copy (XRC

Extended Remote Copy (XRC)

- Premier disaster recover solution for System z storage systems
  - Applies only to ECKD disks; SCSI volumes not covered
  - Also known as z/OS Global Mirror
- Requires a combination of software and hardware capabilities
  - Asynchronous copy between primary and secondary sites
  - Consistent data set at secondary site
  - Timestamp is placed on all write I/Os, in order to reassemble non-committed changes in the event of a disaster

XRC (cont.)

- z/VM announced (July 22, 2010) support for XRC
- Support is available for z/VM 5.4.0 and 6.1.0, with PTFs for APARs VM64814 and VM64816
  - VM64814 contains bulk of support; has no impact if not enabled
  - VM64816 contains additional fixes and feature enhancements (e.g., Monitor)
- Together, the APARs provide baseline Server Time Protocol support
  - Synchronizes z/VM TOD clock with STP server at IPL
  - Maintains a delta value of TOD changes over lifetime of z/VM IPL
  - Supports STP timezone management
  - CPC must be either:
    - A member of an STP-Only CTN
    - Stratum 2 or higher member of a mixed-CTN
- Inserts timestamp on write I/O, for use by XRC-enabled hardware
  - Occurs when XRC LIC is installed on DASD and z/VM LPAR is part of STP

XRC Config Options

- New FEATURES statements for SYSTEM CONFIG to enable support:
  - STP_Timestamping
    - Timestamps will be added to write channel programs issued to all DASD devices that have the XRC LIC installed.
  - STP_TIMEZone / STP_TZ
    - System time zone will be derived from the STP server.
  - XRC_OPTional
    - System behaves differently when STP is suspended – Specifically, we will stop timestamping whenever STP sync is lost but continue issuing I/O. Without this, all I/O that is to be timestamped will be deferred until STP sync is restored
  - XRC_TEST
    - Only allowed 2nd-level. This will enable STP_Timestamping without STP availability. Manually-specified TOD value is used for timestamping. Intended for vendor test support.

Extended Remote Copy (XRC) – Externals

- Query STP
  - Server Time Protocol facility not enabled
    - The STP LIC has not been enabled on your CPC. Cannot use XRC Timestamping functionality until this is resolved.
  - Server Time Protocol available but not enabled for CP use
    - The STP LIC has been enabled, but no STP features have been enabled in the System Configuration file.
  - Server Time Protocol synchronization activated for timestamping
    - The TOD clock was synchronized with STP at IPL and timestamping is active for XRC-capable DASD.
  - Server Time Protocol synchronization suspended. I/O to XRC-capable DASD will be delayed until synchronization completes.
    - Connectivity to the STP server has been severed. I/Os to XRC-capable DASD will queue until connectivity is restored.
  - Server Time Protocol synchronization suspended. I/O to XRC-capable DASD will be issued without timestamps until synchronization completes.
    - Connectivity to the STP server has been severed. I/O to XRC-capable DASD will continue without timestamps.

Extended Remote Copy (XRC) – Externals

- Query TIMEZONE
  - Represents provided by STP
  - The boundary defines the time and date at which the next scheduled timezone change will occur
  - The inactive timezone will become the active timezone at the boundary time/date

```
Zone  Direction   Offset   Status          Boundary
GMT     ----     00.00.00  Inactive
UTC     ----     00.00.00  Inactive
EDT     West     04.00.00  Active-(STP)
EST     West     05.00.00  Inactive-(STP) 10/31/09 02:00:00
```

Extended Remote Copy (XRC)

- For more information, consult the redbook “GDPS Family – An Introduction to Concepts and Capabilities”, order number SG24-6374

Hyperswap Improvements

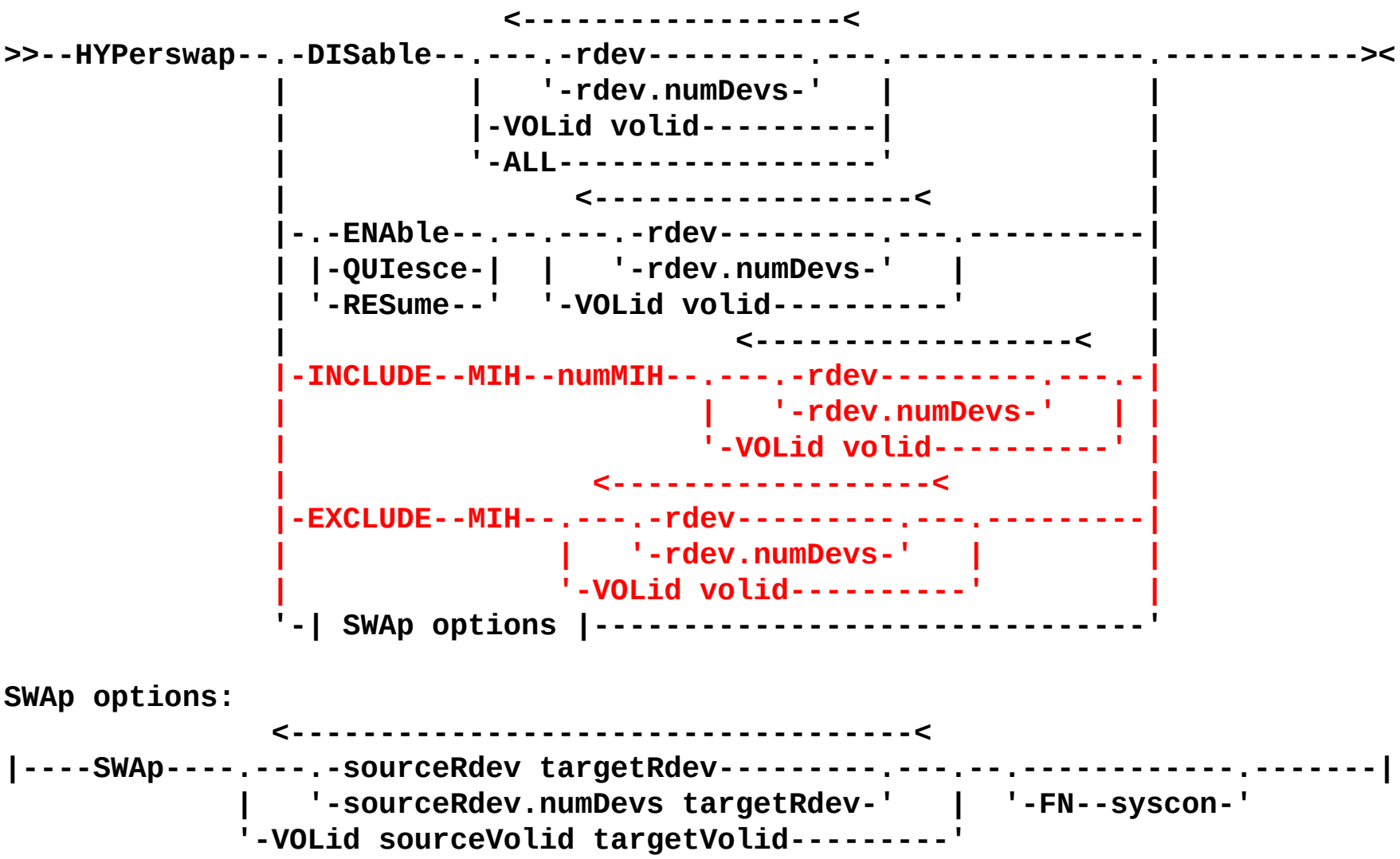
Hyperswap Improvements

- Hyperswap is a technology for switching between mirrored DASD in the event of a Disaster Recovery event
- z/VM has a Class B command, HYPERSWAP, that performs this function
  - Used in conjunction with GDPS/PPRC & GDPS/MzGM solutions

Hyperswap Improvements

- In 2010 and 2011, many APARs have been released that fix defects and improve the functionality of z/VM's Hyperswap support
- VM64815 is the main new function APAR for this support
  - Adds controls to include/exclude devices from an automatic quiesce in the event of an MIH condition (or a certain number of MIH conditions)
  - Management of PAV/HyperPAV is improved
  - Wait state (HCP9060W) in the event of an abend occurs while a Hyperswap is in progress; used in conjunction with non-mirrored dump devices on Site2

Hyperswap Improvements



Extended Address Volumes (EAV)

Extended Address Volumes (EAV)

- Mechanism for creating and using volumes larger than 65,520 cylinders (the so-called “mod-54” variant of a 3390-9), or about 48.3 GB
  - Identified as 3390-A devices, initially ~193 GB maximum

EAV Architecture

- Large cylinder addressing is done by “stealing” 12 bits from head value of CCHH identifier for use as high-order bits to cylinder
  - No DASD has ever had more than 15 heads (x0000-x000E)
  - Often documented as ccccCCCh, where the complete cylinder is rearranged as CCCccccc

Cylinders (Dec)	Cylinders (Hex)	Highest ccccCCCh
3339	x0D0B	x0D0A000E
65520	xFFF0	xFFEF000E
65536	x10000	xFFFF000E
65537	x10001	x0000001E
65667	x10083	x0082001E

EAV APARs

- Two APARs for z/VM 5.4.0 and 6.1.0
  - VM64709 (CP)
  - VM64711 (CMS)
- PTFs available December 2009

EAV CPOWNED Volumes

- EAV volumes can be attached to the SYSTEM, generally for the purposes of minidisks
- Non-PERM extents on CPOWNED EAV volumes are restricted to the first 65,520 cylinders
  - Extents that cross this line will be truncated, with message HCP138E
  - Extents that exist entirely above this line will be ignored, with message HCP139E
- CPFMTXA will enforce this boundary requirement; ICKDSF does not

EAV User Volumes

- z/VM support for dedicated devices, and fullpack minidisks
  - “Fullpack” is defined as 0-END, or DEVNO
  - The ending cylinder must be “END”, even if the number would equal the size of the volume
- Partial pack minidisks on 3390-A volumes are supported, provided they exist completely below cylinder 65,520

EAV Diagnose xA8

- Diagnose xA8 will operate on any area of disk, provided the application is aware of EAV addressing constructs
  - New SGIOPTS byte, immediately after SGILPM, is added, with flag bit SGIEAV (x80) defined to signal this awareness
  - Any application attempting to reference above cylinder 65,520 without enabling this bit will fail with CC=1 R15=2 (unsupported device)

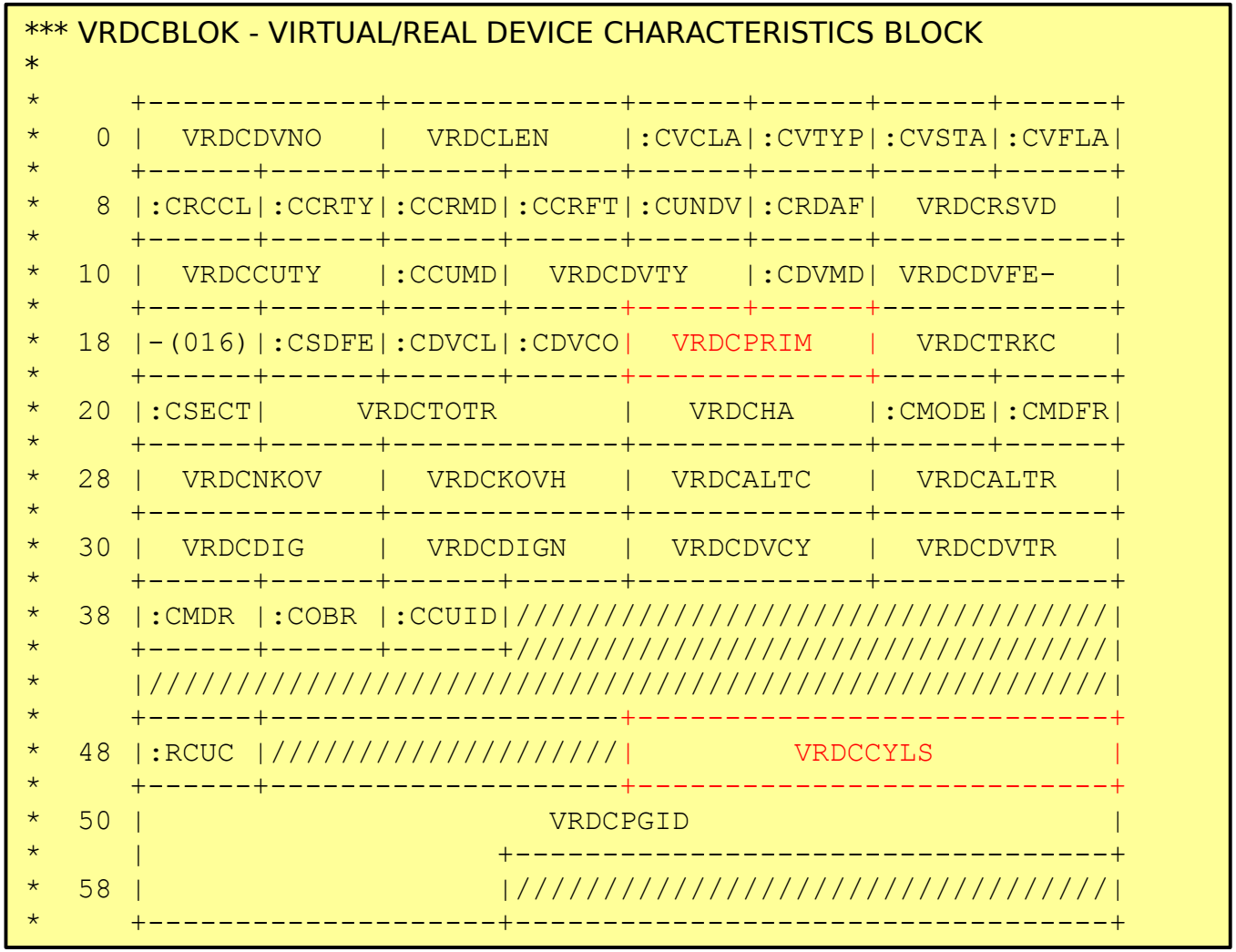
EAV SGIOP

```
*** SGIOP - Synchronous General I/O Parameters
*
*      +-------------+------+------+--------------------------+
*    0 |  SGIDEVNO   |SGIKEY|SGIFLG|         SGIRESV1         |
*      +-------------+------+------+--------------------------+
*    8 |          SGICPA           |         SGIRESV2         |
*      +---------------------------+------+------+------------+
*   10 |         SGICCWA           |:DEVST|:SCHST|  SGIRESCT  |
*      +------+------+-------------+------+------+------------+
*   18 |SGILPM|:OPTS |  SGIRESV3   |  SGIRESV4   |  SGISNSCT  |
*      +------+------+-------------+-------------+------------+
*   20 |         SGIRESV5          |         SGIRESV6         |
*      +---------------------------+--------------------------+
*   28 |         SGIRESV7          |         SGIRESV8         |
*      +---------------------------+--------------------------+
*   30 |         SGIRESV9          |         SGIRESVA         |
*      +---------------------------+--------------------------+
*   38 |                                                      |
*      =                      SGISDATA                        =
*      |                                                      |
*      +------------------------------------------------------+
*   58
*
*** SGIOP - Synchronous General I/O Parameters
```

EAV Diagnose x210

- Diagnose x210 returns the size of a volume in field VRDCPRIM, but this field cannot hold anything larger than 16-bits (65,535)
  - Field practical maximum is 65,520 (xFFF0)
- If VRDCPRIM contains xFFFE, indicates the field has overflowed (large, EAV volume)
- New 32-bit field VRDCCYLS at offset x4C
  - Always contains cylinder size

EAV VRDCBLOK



EAV Other Diagnoses

- Other Diagnoses (x18, x20, xA4, x250, and *BLOCKIO) cannot operate on devices that exist wholly or partially above cylinder 65,520
  - Attempts to do so will be rejected, with the proper “unsupported device” return/condition code combination for the Diagnose

EAV Other Functions

- New DDR fully supports EAV volumes for backup purposes
- Updates to FlashCopy ensure it fully supports EAV volumes
- New fullword fields in MRSEKSEK (Domain 7, Record 1) monitor record less likely to wrap for larger volumes. Equivalent to:
  - CALCURCY, CALSKCYL, IORPOSSM, CALECYL

EAV CMS

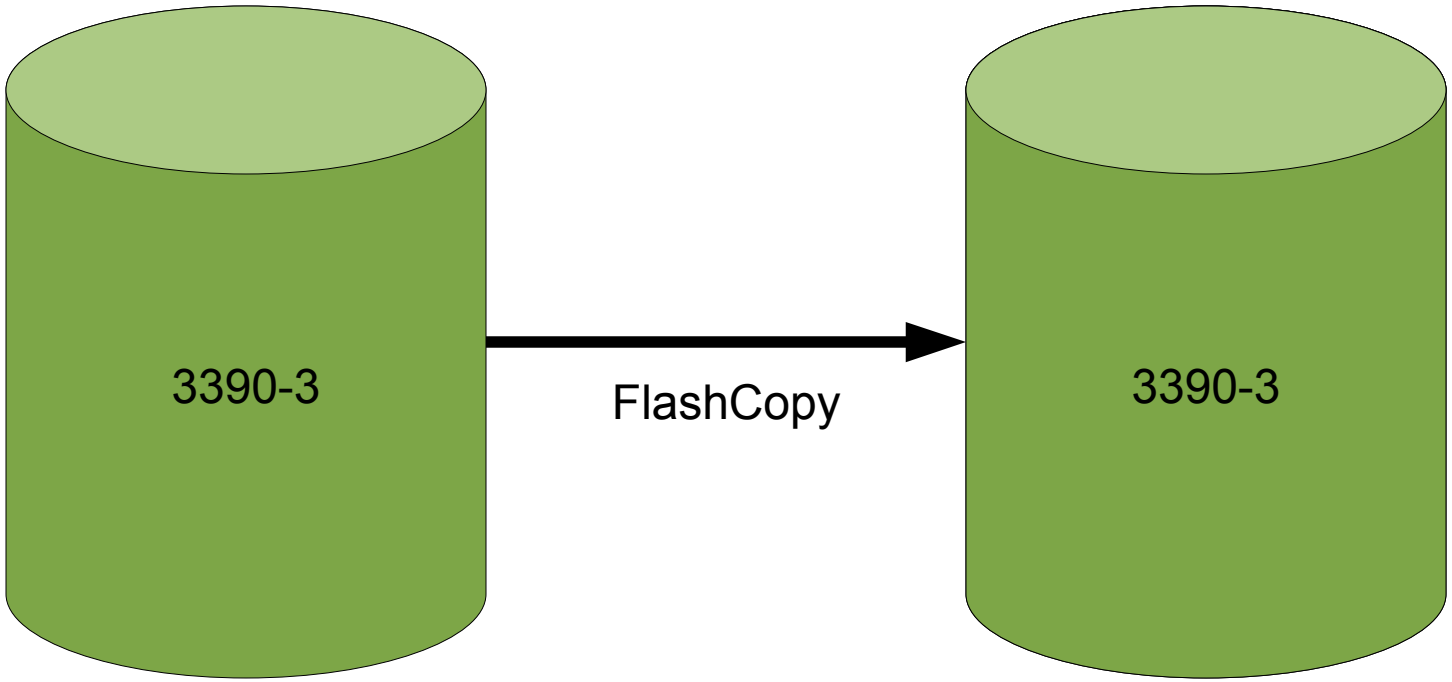
- CMS had supported volumes up to 32,767 cylinders in size
- With aforementioned APAR, CMS is updated to support, via FORMAT and ACCESS, volumes and minidisks up to 65,520 cylinders in size
  - Since this does not provide support for EAV-sized volumes, can be applied separately from CP APAR
- Does not eliminate the requirement of file status and control information that is stored below the 16MB line of the CMS virtual machine
  - If volumes larger than 32,767 cylinders in size are used, care should be taken to avoid large numbers of small files on the disk. Otherwise, the CMS file system may encounter problems accessing and using the device.

FlashCopy / Space Efficient

FlashCopy non-Persistent

- Point-in-time copy technology, good for duplicating volumes where differences presented over time are acceptable
- Target volume is an exact replica of the source upon completion of the copy
- Both source and target volume can be used for applications when FlashCopy is finished

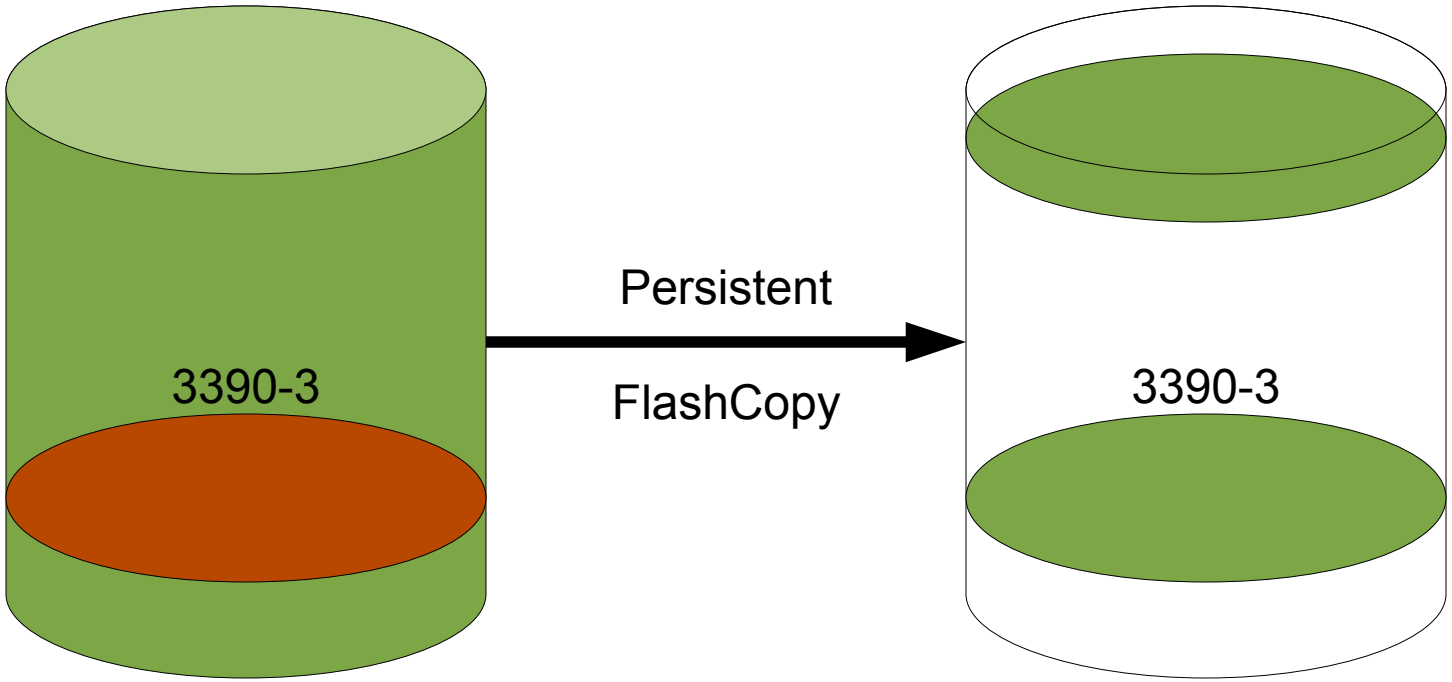
FlashCopy non-Persistent Example



FlashCopy Persistent

- Persistent FlashCopy relationships are used to maintain pointers between source and target volumes, and changes made to either one
- Still a point-in-time copy; does not suggest a mirroring technology in any fashion
- The no-background-copy option (typically NOCOPY) is used to prevent tracks being written to target device until they are actually needed by storage subsystem.

Persistent FlashCopy with NOCOPY Example



FlashCopy / Space-Efficient (SE)

- Space-Efficient volumes are those created solely for the purpose of being targets of persistent FlashCopy relationships
  - A special LIC feature of DS8000
- Tracks for the volume are only allocated as needed to maintain the point-in-time image as source/volume change
  - Still have the look-and-feel of a fully populated device to the user

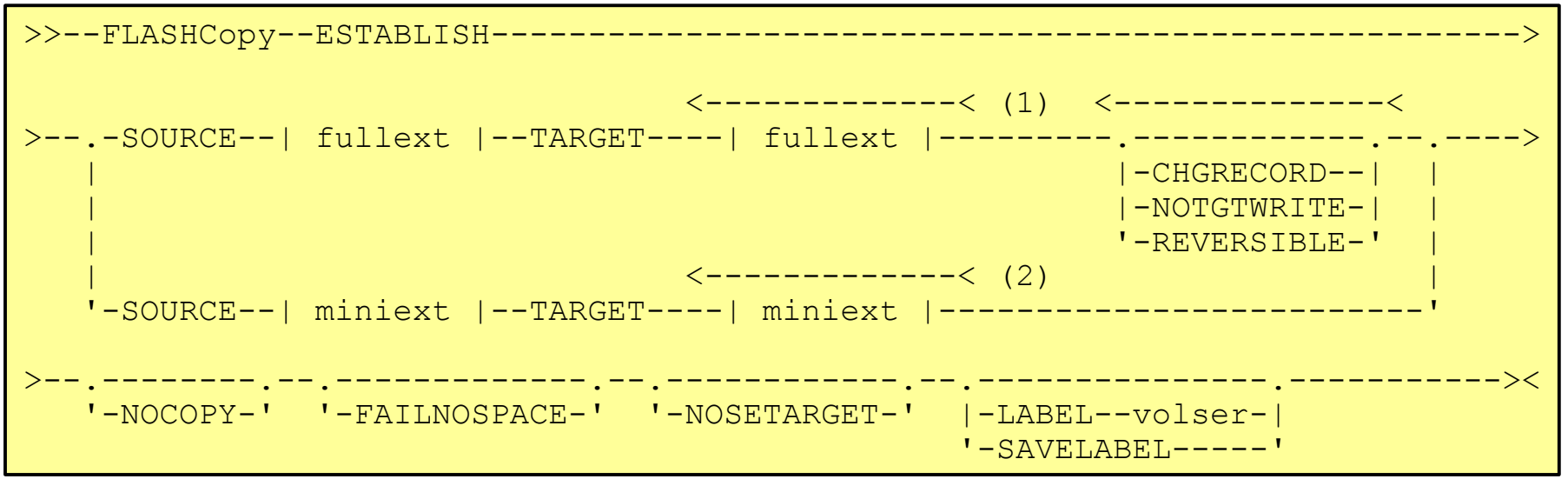
FlashCopy / SE (cont.)

- z/VM APAR VM64449 (December 2008) was released to add persistent FlashCopy, and by extension, FlashCopy/SE, support to the z/VM FLASHCOPY command
  - Several APARs released in 2009; check for related!
- Several new commands are added to create and manage FlashCopy relationships

FlashCopy / SE Establish

- FLASHCOPY ESTABLISH
  - Create Persistent Relationships
  - Needed for SE volumes, but can be used with traditional volumes, if desired
  - LABEL and SAVELABEL options also exist on traditional (non-persistent) FLASHCOPY command

FlashCopy / SE Syntax



See FLASHCOPY Command description for full explanation

FlashCopy / SE Withdraw

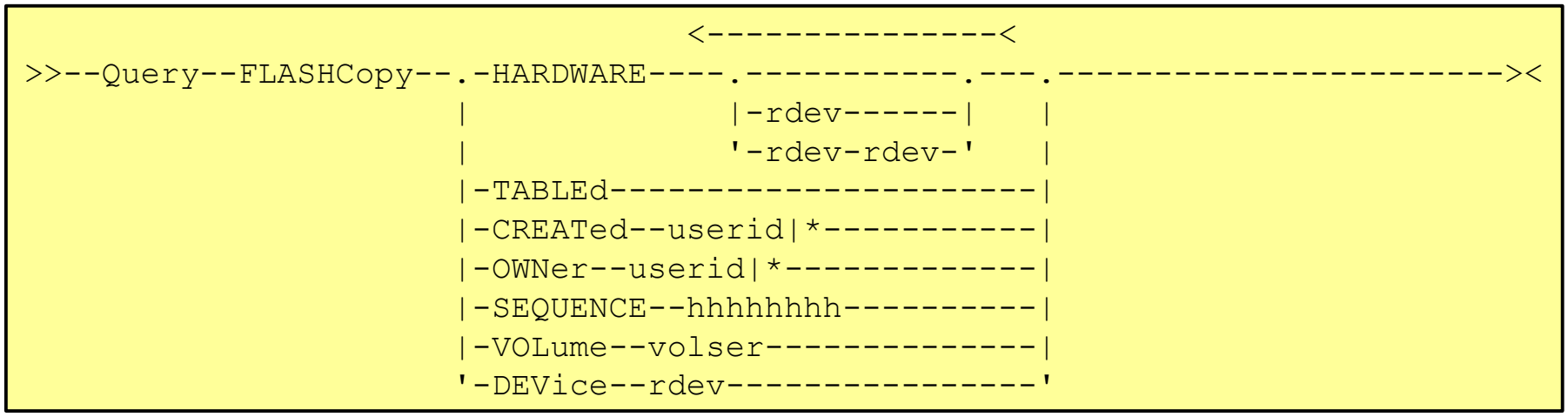
- FLASHCOPY WITHDRAW
  - Removes Persistent FlashCopy relationship between paired volumes
  - FORCE option required when background copy has not completed, as a result of NOCOPY option on Establish
    - FLASHCOPY BACKGNDCOPY can be used to initiate background copy without withdrawing relationship, but cannot be used on space-efficient targets

FlashCopy / SE More Options

- FLASHCOPY RESYNC
  - For persistent relationships, will establish a new point-in-time copy between source and target volume(s)
- FLASHCOPY TGTWRITE
  - Disables the write-inhibit option on target volume, if specified on original Establish

FlashCopy / SE Query

- QUERY FLASHCOPY
  - Privileged command to interrogate status of FlashCopy relationship(s) on a real device
  - State on the hardware, and options for retrieving information about a relationship created within a z/VM session during current IPL



See QUERY FLASHCOPY Command description for full explanation

FlashCopy / SE Query Output Examples

```
q flashcopy hardware
             --------SOURCE-------- --------TARGET--------
SEQUENCE FLGS RDEV VOLSER CC...CC/HH RDEV VOLSER CC...CC/HH REMAINING/TOTAL
4B145C34 0800 5100 PACK01     100/00 5101 PACK02     100/00 0/1500
4B14700A 0800 5100 PACK01     200/00 5101 PACK02     200/00 0/1500
4B145C34 8800 5100 PACK01     100/00 5101 PACK02     100/00 0/1500
4B14700A 8800 5100 PACK01     200/00 5101 PACK02     200/00 0/1500

q flashcopy tabled
SEQUENCE ---DATE--- --TIME--  RDEV  VOLSER CREATOR  OWNER    VDEV
4B145C34 2008-04-30 15:19:55  5100> PACK01 RWS      RWS      0200
4B145C34 2008-04-30 15:19:55 >5101  PACK02 RWS      RWS      0300
4B14700A 2008-04-30 15:20:00  5100> PACK01 RWS      RWS      0210
4B14700A 2008-04-30 15:20:00 >5101  PACK02 RWS      RWS      0310
Ready; T=0.01/0.01 15:20:40
```

Thanks!

Contact Information:

Stephen (Steve) G. Wilkins
IBM
z/VM Development
Endicott, NY

wilkinss@us.ibm.com

A PDF of this presentation will be available at:
http://www.vm.ibm.com/devpages/wilkinss