Leveraging the Newest Capability in z/VM 6.4

John Franciscovich
IBM - z/VM Design and Development
francisj@us.ibm.com

March 2019

Trademarks

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

BladeCenter*	GDPS*	IBM z13*	PR/SM	System z9*	zSecure
DB2*	HiperSockets	IBM z14	RACF*	System z10*	z/VM*
DS6000*	HyperSwap	OMEGAMON*	Storwize*	Tivoli*	z Systems*
DS8000*	IBM LinuxONE Emperor	Performance Toolkit for VM	System Storage*	zEnterprise*	
ECKD	IBM LinuxONE Rockhopper	Power*	System x*	z/OS*	
FICON*	IBM Z*	PowerVM	System z*		

* Registered trademarks of IBM Corporation

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.
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.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.
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.
Java and all Java based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.
Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and
Linux is a registered trademark of Linus Torvalds 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.
OpenStack is a trademark of OpenStack LLC. The OpenStack trademark policy is available on the OpenStack website.
TEALEAF is a registered trademark of Tealeaf, an IBM Company.
Windows Server and the Windows logo are trademarks of the Microsoft group of countries.
Worklight is a trademark or registered trademark of Worklight, an IBM Company.
UNIX is a registered trademark of The Open Group in the United States and other countries.

* Other product and service names might be trademarks of IBM or other 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.

This information provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs) ("SEs"). IBM authorizes customers to use IBM SE only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the “Authorized Use Table for IBM Machines” provided at www.ibm.com/systems/support/machine_warranties/machine_code/aut.html (“AUT”). No other workload processing is authorized for execution on an SE. IBM offers SE at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

Notice Regarding Specialty Engines (e.g., zIIPs, zAAPs and IFLs):

Any information contained in this document regarding Specialty Engines ("SEs") and SE eligible workloads provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs). IBM authorizes customers to use IBM SE only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the “Authorized Use Table for IBM Machines” provided at www.ibm.com/systems/support/machine_warranties/machine_code/aut.html (“AUT”).

No other workload processing is authorized for execution on an SE.

IBM offers SEs at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

Agenda

- z/VM 6.4 release information
  - Installing z/VM on a z14
- Continuous delivery enhancements to z/VM 6.4
- Overview of base z/VM 6.4 enhancements

Release Status and Information

z/VM Release Status Summary

z/VM Level	GA	End of Service	End of Marketing	Minimum Processor Level	Maximum Processor Level	Security Level
7.1	3Q2018			zEC12 & zBC12		
6.4	11/2016			z196 & z114®	-	Common Criteria Complete! FIPS 140-2 In Progress

z/VM 6.4

- A release born of customer feedback
  - z Systems Business Leaders Council (zBLC)
  - SHARE dialogues
  - IBM internal T3s (Teach the Teacher)

- Prioritizations set by customers and adjusted by IBM resources and skills

- Two major areas:
  - Technical enhancements that continue to improve TCO and bring direct value
  - Improved quality of life for z/VM system programmers

- New Architecture Level Set (ALS)
  - z196 and z114 or newer
  - Drops z10 EC and BC support

z/VM 6.4 Supported Hardware

- Following z Systems servers:
  - z14
  - z13
  - z13s
  - LinuxONE Emperor and Emperor II
  - LinuxONE Rockhopper
  - IBM zEnterprise EC12
  - IBM zEnterprise BC12
  - IBM zEnterprise 196
  - IBM zEnterprise 114

- Electronic and DVD install
  - No tapes

RACF Considerations

- **Validate the database before up-leveling RACF database template**
  - RACUT200 utility checks database integrity
  - Always run RACUT200 before issuing RACFCONV

- **Database best practices**
  - Have a procedure for database backups
  - Integrity-check your back-up databases
  - Automate around RACF initialization
- Whitepaper - **Validating and Repairing RACF Database Integrity on z/VM**
  - https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=ZSW03366USEN&
    - *Brian W. Hugenbruch, CISSP - IBM Z Virtualization Security*
  - More information available at http://www.vm.ibm.com/security/

Fresh Install Considerations for z/VM 6.4

- Supports 3390 mod-27 DASD (32760 cylinders)

- Default location for components is **now SFS** instead of minidisks
  - Minimizes future disruption for increasing minidisks
  - Can select to use minidisks instead
  - Different component names (e.g. dirmsfs instead of dirm)

- Install must be done to **full pack** minidisks (***cannot*** define as n-1 cylinders)
  - Minimum install sizes:
    - 3390 mod-3 - 3339 cylinders (changed from 3338)
    - 3390 mod-9 - 10017 cylinders (changed from 10016)
    - 3390 mod-27 - 32760 cylinders
  - Refreshed install image allows n-1 cylinders for 3390 mod-9 and mod-27
    - available August 25, 2017

Upgrade-in-Place

- Enables a smoother upgrade of z/VM 6.2 and z/VM 6.3 systems to z/VM 6.4
  - Especially helpful in a Single-System-Image (SSI) environment
  - Avoids a full and fresh install
- Includes processes to:
  - Apply vendor and customer modifications
  - Back out upgrade changes
- Requires appropriate service on the old z/VM level
- See the Install Guide for the complete list of prerequisites
- Unlike z/VM 6.3, z/VM 6.4 requires TCP/IP machine to be shut down at one point, so will need alternate method to get to z/VM (OSA ICC is an excellent choice)
- See Live Virtual Class for session on Upgrade-in-Place May 31st, 2017
  - http://www.vm.ibm.com/education/lvc/

DFSMS/VM Considerations for z/VM 6.4

- There are no changes in DFSMS/VM from z/VM 6.3 to 6.4
  - You do **not** need to re-install if you already have it on z/VM 6.3
  - If not on z/VM 6.3, a fresh install is required

- Upgrade-in-Place **does not** install DFSMS/VM
  - If you order DFSMS/VM for z/VM 6.4 you will receive a deliverable file
    - Ignore if you already have DFSMS/VM on z/VM 6.3

TCP/IP IPWIZARD Utility

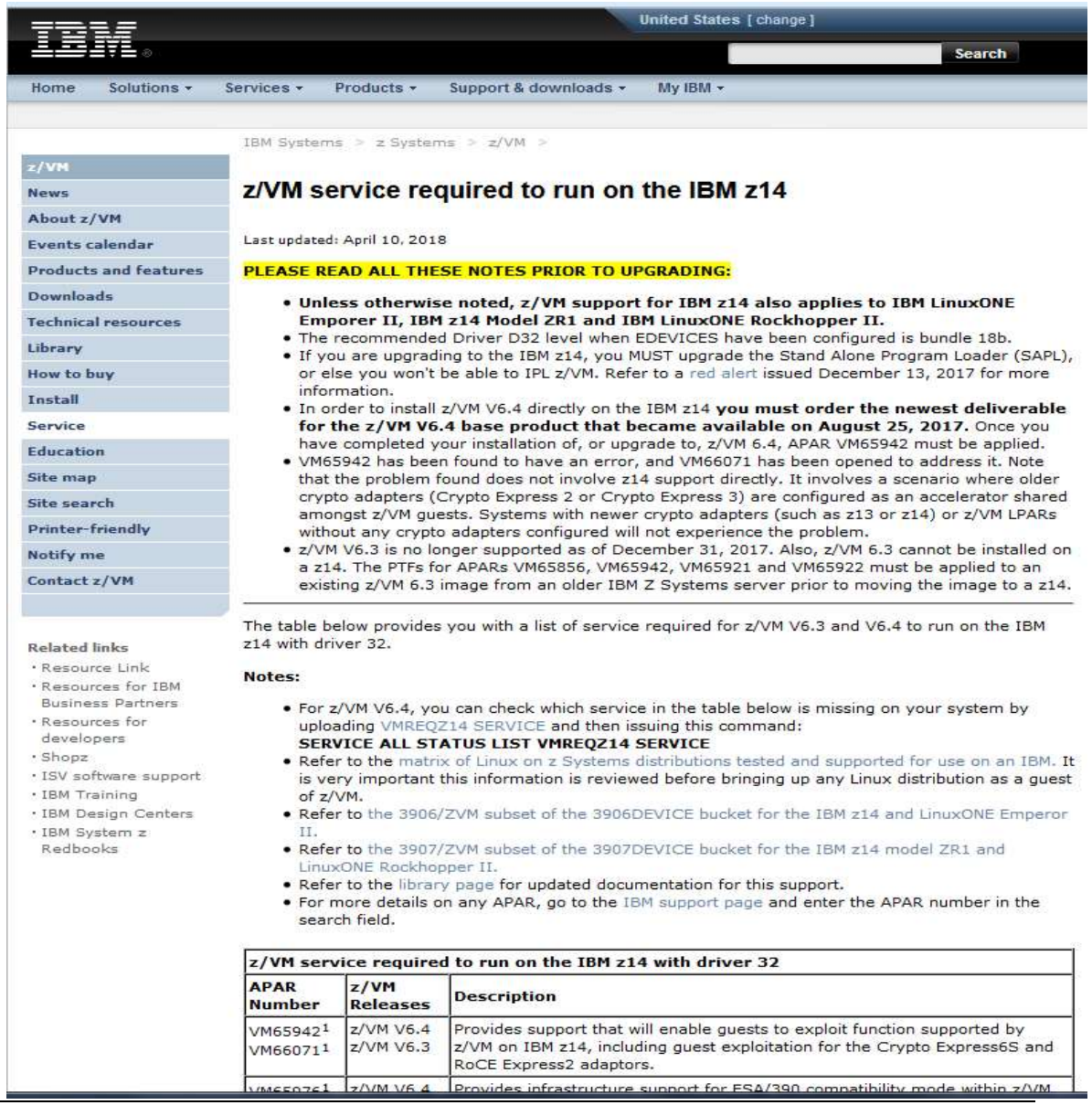
- Need to circumvent a problem before running IPWIZARD on a newly-installed z/VM 6.4 system
  - TCPIP DATA file needs to be created
- See http://www.vm.ibm.com/related/tcpip/tcpipwiz.html for details
- APAR PI70089 corrects this

Expanded Storage

- z/VM 6.4 fulfills Statement of Direction to drop support for all use of expanded storage
- Convert any expanded storage to central storage (real memory) before bringing up z/VM 6.4
- The memory management changes made in z/VM 6.3 made expanded storage obsolete

Installing z/VM on a z14

- Required service information at http://www.vm.ibm.com/service/vmreqz14.html



IBM

United States [change]

Search

Home | Solutions | Services | Products | Support & downloads | My IBM

IBM Systems > z Systems > z/VM >

z/VM
News
About z/VM
Events calendar
Products and features
Downloads
Technical resources
Library
How to buy
Install
Service
Education
Site map
Site search
Printer-friendly
Notify me
Contact z/VM

Related links
- Resource Link
- Resources for IBM Business Partners
- Resources for developers
- Shopz
- ISV software support
- IBM Training
- IBM Design Centers
- IBM System z Redbooks

z/VM service required to run on the IBM z14

Last updated: April 10, 2018

PLEASE READ ALL THESE NOTES PRIOR TO UPGRADING:

- **Unless otherwise noted, z/VM support for IBM z14 also applies to IBM LinuxONE Emporer II, IBM z14 Model ZR1 and IBM LinuxONE Rockhopper II.**
- The recommended Driver D32 level when EDEVICES have been configured is bundle 18b.
- If you are upgrading to the IBM z14, you MUST upgrade the Stand Alone Program Loader (SAPL), or else you won't be able to IPL z/VM. Refer to a red alert issued December 13, 2017 for more information.
- In order to install z/VM V6.4 directly on the IBM z14 **you must order the newest deliverable for the z/VM V6.4 base product that became available on August 25, 2017.** Once you have completed your installation of, or upgrade to, z/VM 6.4, APAR VM65942 must be applied.
- VM65942 has been found to have an error, and VM66071 has been opened to address it. Note that the problem found does not involve z14 support directly. It involves a scenario where older crypto adapters (Crypto Express 2 or Crypto Express 3) are configured as an accelerator shared amongst z/VM guests. Systems with newer crypto adapters (such as z13 or z14) or z/VM LPARs without any crypto adapters configured will not experience the problem.
- z/VM V6.3 is no longer supported as of December 31, 2017. Also, z/VM 6.3 cannot be installed on a z14. The PTFs for APARs VM65856, VM65942, VM65921 and VM65922 must be applied to an existing z/VM 6.3 image from an older IBM Z Systems server prior to moving the image to a z14.

The table below provides you with a list of service required for z/VM V6.3 and V6.4 to run on the IBM z14 with driver 32.

Notes:

- For z/VM V6.4, you can check which service in the table below is missing on your system by uploading VMREQZ14 SERVICE and then issuing this command:
  SERVICE ALL STATUS LIST VMREQZ14 SERVICE
- Refer to the matrix of Linux on z Systems distributions tested and supported for use on an IBM. It is very important this information is reviewed before bringing up any Linux distribution as a guest of z/VM.
- Refer to the 3906/ZVM subset of the 3906DEVICE bucket for the IBM z14 and LinuxONE Emperor II.
- Refer to the 3907/ZVM subset of the 3907DEVICE bucket for the IBM z14 model ZR1 and LinuxONE Rockhopper II.
- Refer to the library page for updated documentation for this support.
- For more details on any APAR, go to the IBM support page and enter the APAR number in the search field.

z/VM service required to run on the IBM z14 with driver 32		
APAR Number	**z/VM Releases**	**Description**
VM65942[1] VM66071[1]	z/VM V6.4 z/VM V6.3	Provides support that will enable guests to exploit function supported by z/VM on IBM z14, including guest exploitation for the Crypto Express6S and RoCE Express2 adaptors.
VM65976[1]	z/VM V6.4	Provides infrastructure support for ESA/390 compatibility mode within z/VM

Installing z/VM 6.4 on a z14

1. *Fresh install requires new install media*
   - Available August 25, 2017
   - Look for “-01” suffix, (e.g. LCD7-7040-01 for 3390)

2. *Apply VM65942 immediately after installing the above*
   - PE VM66071; limited to shared Crypto Express 2 & 3

3. *If running SSI:*
   - Apply VM65976 to all members ***before*** IPLing any member on a z14

- All of the above APARs are on RSU 1702

Checking for Valid Service (z/VM 6.4)

- File with list of service required on
  http://www.vm.ibm.com/service/vmreqz14.html

- For z14 model ZR1 and LinuxONE Rockhopper II
  - Get file VMREQZR1 SERVICE
  - Issue: **`SERVICE ALL STATUS LIST VMREQZR1 SERVICE`**

- For z14 and LinuxONE Emperor II
  - Get file VMREQZ14 SERVICE
  - Issue: **`SERVICE ALL STATUS LIST VMREQZ14 SERVICE`**

Migrating z/VM 6.3 to a z14

- ***NOTE:*** z/VM 6.3 is out of service as of December 31, 2017
- *z/VM 6.3 cannot be installed directly on a z14*
  - Must be migrated from a prior server after applying required service

1. *Apply VM65942, VM65921, and VM65922 and IPL on pre-z14 server*

2. *Run **SALIPL** utility (updated with VM65856) to create a new **SAPL** (Stand-Alone Loader) program on the IPL disk*

3. *Run the SDINST utility to create a new stand-alone dump program*
   - Detailed instructions for running SDINST are provided in Chapter 11 of *CP Planning and Administration*

4. *If running SSI:*
   - Apply VM65976 to all members before IPLing any member on a z14
   - If a mixed-release cluster (z/VM 6.3 and z/VM 6.4), apply VM65867 to all z/VM 6.3 members

      • These can be applied at the same time as the APARs in Step 1.

Stay Informed about Future New Function

- New web page to subscribe to:
  - http://www.vm.ibm.com/newfunction/

- Lists enhancements IBM is pursuing and gives:
  - Tentative dates for planning purposes
  - A high level view of impact and compatibility
  - Interaction with ISV products, Linux, and hardware

- Allows clients to
  - Express interest in being a sponsor user for the item
  - Plan for upcoming new support
  - Avoid surprises

Stay Informed about New-Function PTFs

- Off z/VM service page http://www.vm.ibm.com/service/ is new page for new-function APARs
  - http://www.vm.ibm.com/service/vmnfapar.html

- Applies to z/VM operating system and related products:
  - Operations Manager for z/VM
  - Backup and Restore Manager for z/VM
  - OMEGAMON XE on z/VM and Linux
  - Etc.

- Subscribe to receive notifications automatically when new-function APARs become available

- Obtain lists of previously shipped new-function APARs

Continuous Delivery Enhancements: Enhancements for CPU Resource Management

How to Get Enhancements for CPU Resource Management

- Infrastructure changes to facilitate possible management tools
- Available April 11, 2018 for z/VM 6.4

Component	APAR	PTF
CP	VM66105	UM35303
DIRMAINT	VM66109	UV99330

- Improves methods for collecting information and managing CPU resources
  - Extensions to the STHYI instruction
  - CPU Pools renamed to Resource Pools
  - Ability to collect subset of monitor records to capture changes via commands
  - New class 0 events for *VMEVENT system service

Effect of Running with PTF

- By default most changes do not affect processing

- Monitor command information enabled by:

  `MONITOR EVENT ENABLE COMMAND`

  – Causes Event records to be generated for VARY ON, LOGOFF, SET SHARE, etc.

- **`DEFINE CPUPOOL`** and **`QUERY CPUPOOL`** are superseded by **`DEFINE RESPOOL`** and **`QUERY RESPOOL`**

  – The output of QUERY CPUPOOL has changed slightly. If you have EXECs that parse it, please check.

- Now a limit of 1000 virtual machines per pool

QUERY Examples

Before PTF

```
Pool name       CPU     Type   Members
LINUXP2     8.00 Cores  IFL          0
LINUXP1       NoLimit   IFL          6
CPPOOL10      12 %       CP          8
LINUXP3       30 %      IFL         20
```

After PTF

```
Pool name       CPU     Type   Storage Trim Members
LINUXP2     8.00 Cores  IFL   NoLimit ----       0
LINUXP1       NoLimit   IFL   NoLimit ----       6
CPPOOL10      12 %       CP   NoLimit ----       8
LINUXP3       30 %      IFL   NoLimit ----      20
```

Continuous Delivery Enhancements: Virtual Networking Enhancements

How to Get VSwitch Load Balance Enhancements

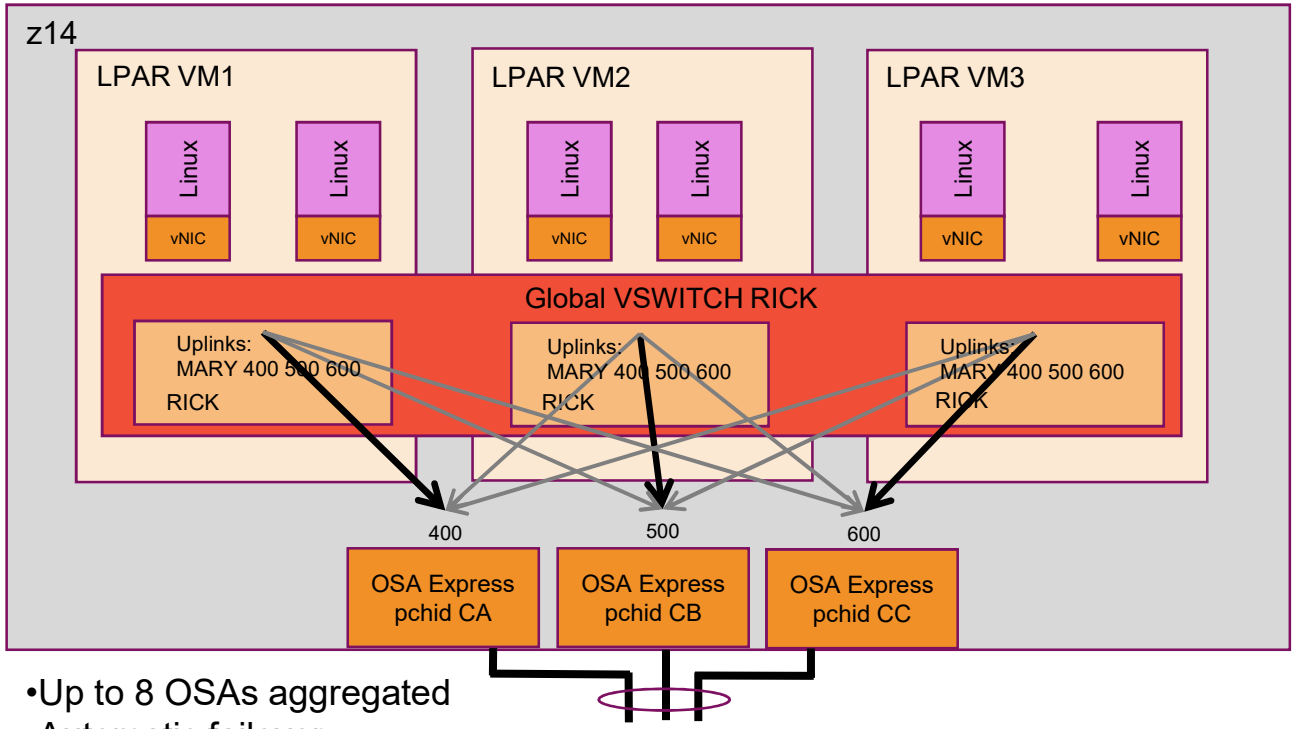
- Available December 15, 2017 for z/VM 6.4

Component	APAR	PTF
CP	VM65918	UM35253

- Apply PTF and reIPL z/VM

- Improves rebalancing algorithms
  - Especially for Multi-VSwitch Link Aggregation configurations.
  - Especially for configurations with more than 2 OSA.

Global z/VM Virtual Switch



- Up to 8 OSAs aggregated
- Automatic failover
- Automatic balancing

NICDEF Security Controls

- We call this function “Directory Network Authorization” (DNA)
  - Define virtual network configuration on **NICDEF** directory statement
    - Consolidates virtual NIC and its network properties in single, secure location
  - Eliminates need for additional statements and commands to authorize and connect user to virtual network
    - **MODIFY VSWITCH** statement / **SET VSWITCH GRANT** command
    - **COUPLE** command
  - Eliminates operational differences between **PORTBASED** and **USERBASED** VSwitches
    - Live Guest Relocation still requires **PORTBASED** or **USERBASED** designation for a VSwitch to match on the source and target members.

How to Get NICDEF Security Controls

- Available August 4, 2017 for z/VM 6.4
  - Included on RSU 1702

Component	APAR	PTF
CP	VM65925	UM35189
RACF (RPIDIRCT)	VM65931	UV61339
DIRMAINT	VM65926	UV61338

Virtual NIC - User Directory

- The Virtual NIC is now fully configured by statements in the CP directory entry:

```
NICDEF vdev TYPE QDIO
               [LAN SYSTEM switch]
               [DEVICES nn]
               [MACID hhhhhh]
```

Combined with VMLAN USERPREFIX to create virtual MAC

New! ->

```
               [PORTNUMBER n]
               [PORTTYPE ACCESS|TRUNK]
               [VLAN vidset]
               [PROMISCUOUS|NOPROMISCUOUS]
```

```
Example:

NICDEF 1100 TYPE QDIO LAN SYSTEM SWITCH1
NICDEF 1100 MACID B10006
NICDEF 1100 VLAN 57
```

Define and Connect to VSwitch

```
DEFINE VSWITCH VSW1 ETHERNET
                    PORTBASED

                    RDEV E00 F00

                    VLAN AWARE
                    NATIVE NONE
```

```
NICDEF E00 TYPE QDIO LAN SYSTEM VSW1 MACID B10006 VLAN 57
```

Best Practice: "VLAN AWARE NATIVE NONE"

Best Practice: Use PORTBASED

VSwitch Access List

- All authorization and configuration in **NICDEF**
  - **NICDEF** overrides any prior **SET VSWITCH GRANT**

- **SET VSWITCH** can be used to change authorizations dynamically

```
SET VSWITCH vswitch_name GRANT userid VLAN vid
```

  - Immediate effect for PORTTYPE, VLAN, PROMISCUOUS

- Can revert to old behavior with… (but why would you?)

```
VMLAN   DNA DISABLE
SET VMLAN DNA DISABLE
```

  - Results in HCP3224I (NICDEF network configuration ignored)

PORTNUMBER n

- **PORTNUMBER** is now optional for **PORTBASED** VSwitches
  - (as it has been for **USERBASED** VSwitches)

- If you select a port, it must be 1-2048
  - **COUPLE** will fail if there is a conflict

- If you don't select a port, CP will choose one 2176-4095

  - Cannot **VMRELOCATE** to a pre-DNA member
    - Ports above 2048 are not supported

  - If in an SSI cluster and do not use **PORTNUMBER**, apply the PTF to all members

Best Practices for VLAN-Aware VSwitch

- Use **NICDEF** to assign VLANs and port numbers (NEW)

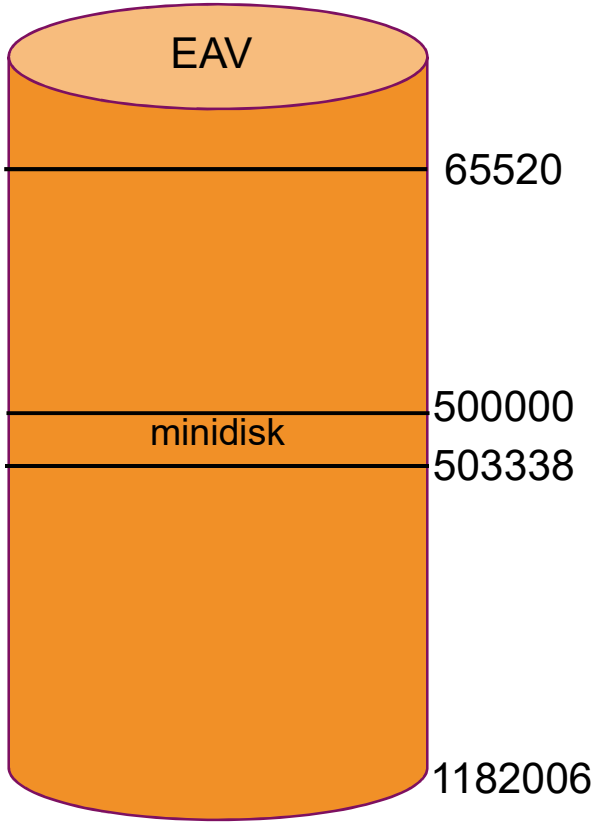
- Define VSwitch with "**VLAN AWARE NATIVE NONE**"
  - Guest that has not been given access will get errors
  - No chance of untagged frames escaping from z/VM

- Use ESM and groups to manage VLAN assignments
  - Simplifies VLAN changes
  - Overrides VLAN specification on **NICDEF**
  - CP will use **NICDEF** if ESM defers

Continuous Delivery Enhancements:
EAV Minidisks

EAV Minidisks: Overview

- Enhanced minidisk support for Extended Address Volumes (EAV)
  - 3390 Model A with more than 65520 cylinders
- Minidisks can be defined anywhere on an EAV
  - Can be up to 1,182,006 cylinders (1 TB)
  - Previously, only fullpack minidisks could use cylinders 65520 and higher



How to Get Enhanced EAV Support

- Available August 25, 2017 for z/VM 6.4

Component	APAR	PTF
CP	VM65943 VM66140	UM35187 UM35296
CMS	VM65945	UM35204
ICKDSF R17 (IF REQ of CP APAR)	PI85943	UI49579

EAV Minidisk Usage

- Disks that can be located anywhere on an EAV
  - Minidisks
  - Temporary disks
  - PARM disks
- Temporary disks and non-CMS minidisks can be up to 1,182,006 cylinders
- **DDR** and **CP FLASHCOPY** can be used for virtual DASD located anywhere on an EAV
- Most common diagnose codes and I/O interfaces have been enhanced
  - Some are limited to maximum disk sizes of 65520 cylinders

EAV Support Limitations

- Disk usage that must remain below cylinder 65520
  - **DRCT, PAGE, SPOL** extents
  - **Checkpoint and warmstart** areas
    - Can start no higher than cylinder 65511 (can be up to 9 cylinders)
  - **MAPMDISK** is restricted to minidisks that are entirely below real cylinder 65520
  - **XLINK** supports only the first 65520 cylinders (0-65519) on an EAV

- Maximum disk size of 65520 cylinders
  - **CMS minidisks**
  - **PARM** disks
  - Minidisks manipulated using **DFSMS/VM**

- **Minidisk cache** is not supported for any minidisk defined on an EAV
  - MDC is supported for real 3390-A volumes that are 65520 cylinders or less
    - Requires APAR VM65741 (PTF UM34922)

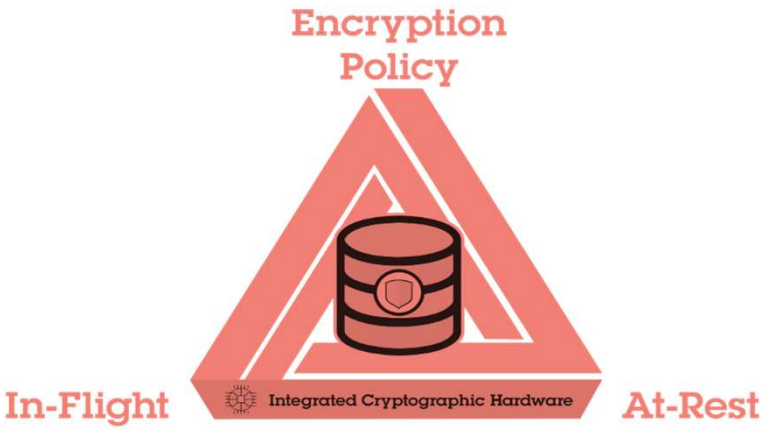
EAV Support – Additional Hints

- Live Guest Relocation (LGR)
  - To relocate guests with non-fullpack minidisks above cylinder 65519
    - Both source and target members must have the Enhanced EAV PTF installed
- SFS file pools that use VM Data Spaces
  - User storage group minidisks located on an EAV must end below cylinder 65520
- If you have GDPS
  - Consult GDPS PSP buckets for possible required service
- Contact vendors for appropriate updates if you have ISV products for
  - Directory management
  - Backup and restore

Continuous Delivery Enhancements:
z/VM Encrypted Paging

IBM Z Pervasive Encryption

- Pervasive Encryption is an important IBM Z strategic theme
- z/VM Encrypted Paging supports this theme
- *Transparent and consumable approach to enable extensive encryption of data in-flight and at-rest to substantially simplify & reduce the costs associated with protecting data & achieving compliance mandates*



z/VM Encrypted Paging: Overview

- Encrypts guest and VDISK pages when writing to paging disks
  - Exploits **z14** and **LinuxONE Emperor II** and **LinuxONE Rockhopper II** hardware capability

- Ciphering occurs as data moves from memory to CP paging volumes
  - ECKD, SCSI, or native FBA

- Makes customer data defensible from an attack or breach of volumes
  - Including cases where a system administrator has unintended access

How to Get Encrypted Paging

- Available December 11, 2017 for z/VM 6.4

Component	APAR	PTF
CP	VM65993	UM35257

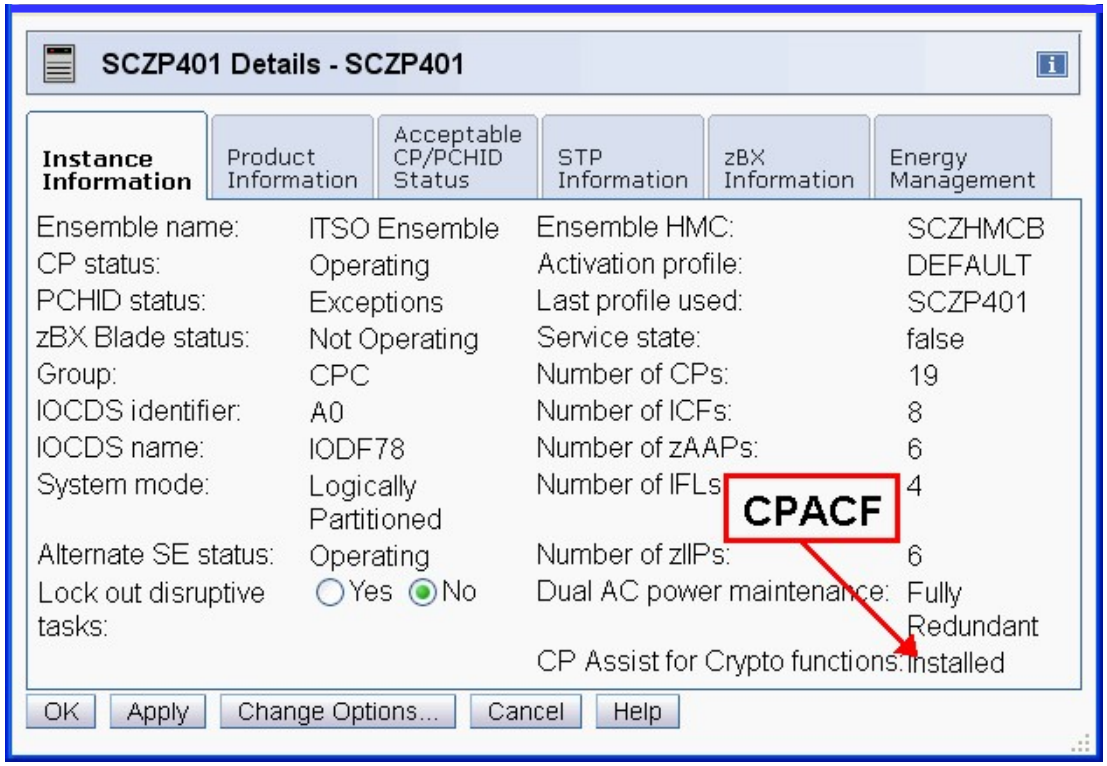
Getting Started with Encrypted Paging

1. Starting point: z/VM partition on a **z14** or **LinuxONE Emperor II** or **LinuxONE Rockhopper II**
   – **CPACF** hardware feature must be enabled
2. Enable encryption in the system configuration file
   – Can be changed dynamically
3. Generates an ephemeral *n*-bit AES encryption key during IPL
4. If ENCRYPT PAGING is **ON**, then data is encrypted/decrypted as it moves to/from paging volumes
5. Use monitor records to determine the performance impact for workloads

Relevant User Roles: Data Owner, Security Admin, Auditor

CP Assist for Cryptographic Functions (CPACF)

- No-charge feature but needs to be ordered with **z14** or **LinuxONE Emperor II** or **LinuxONE Rockhopper II**
- Must be enabled on the Support Element



Using Encrypted Paging for z/VM

- Enable and specify encryption algorithm
  - **ENCRYPT PAGING** statement in system configuration file
    - **OFF | ON | REQUIRED** (**OFF** is the default)
    - **AES256** is the default cipher algorithm
  - **SET ENCRYPT** command
    - Same operands as **ENCRYPT** statement
  - If encryption is enabled with **REQUIRED**, settings are locked until the next IPL
  - The cipher algorithm can be set at only the first enabling of encryption
    - IPL is required to change it
- Use **QUERY ENCRYPT** command to view settings

```
Encrypt Paging settings:
     Currently: Required AES256
     At IPL: Off

Ready;
```

Using Encrypted Paging for z/VM (cont.)

- Only way to ensure 100% compliance is to IPL your z/VM system with
  - **`ENCRYPT PAGING ON ALGORITHM AES256`**
- If encryption is changed from **ON** to **OFF**, pages will still be decrypted when read into memory
- Auditing with monitor records
  - D1 R4 – system configuration and current status thereof
  - D1 R34 (new) – change record for status (**SET ENCRYPT**), with userid
  - D3 R2 – pages encrypted/decrypted, CPU consumed for encryption/decryption
- Auditing with SMF records
  - Auditing in RACF automatically covers new CP commands, per above
  - Just enable tracking in your VMXEVENT profile

Encrypted Paging: Notes on 'REQUIRED' Operand

- Please note that **REQUIRED** means **REQUIRED.**
  - Cannot be changed, cannot be broken
  - Meant to assure 100% compliance for the customers who need it

- If you have configured **REQUIRED** on a system which does not support the feature, ***your system will not IPL***
  - Double-check system labels in an SSI cluster – exclude back-level systems
  - CPACF not enabled on new CEC – turn on CPACF
  - z13 and earlier hardware – not supported
  - May be a problem for DR sites

- IBM recommends:
  1. Test Encrypted Paging with **ON** before switching to **REQUIRED**
  2. Consider either:
     a) Switching from **ON** to **REQUIRED** in AUTOLOG1 (during system IPL)
     b) Putting **SET ENCRYPT PAGING REQUIRED** on a COMMAND statement for OPERATOR
  3. Have a backup system configuration file (with setting **ON**) for emergency purposes
  4. Double-check DR plans for hardware availability of z/VM systems

Encrypted Paging: SSI and LGR Implications

- It is OK to enable encrypted paging on only some of the members of an SSI cluster

- Ephemeral keys are not shared; there is one ephemeral key per member
  - When relocating a guest
    - Its pages are decrypted before they are relocated to the target member
    - Target member re-encrypts the guest's pages using its own ephemeral key

- Relocation domains can be defined based on guests' security requirements, such as
  - Access to hardware facilities such as z14 CPACF
  - Encrypted paging (requires z14 partitions)

Encrypted Paging: Sample Performance Data

- High paging workload with scaling of logical processors & memory
- Non-SMT with default algorithm AES256
- Workloads paged between 150K to 163K pages / sec (considered high)
- Stats computed from data in D3 R2 monitor records

Logical Processors/ Memory (GB)	En+Decrypted paging Rate (D3R2)	En+Decrypted CPU Bsy (D3R2) (% of one IFL CPU)
08 / 512	159997.55	19.57
16 / 1024	163467.08	19.57
24 / 1536	150619.01	17.96
32 / 2048	155924.89	18.88

Encrypted Paging: Sample Performance *Results*

- Used between **18% to 20%** of one logical processor
  - Increased as paging rate increased
- The percent CPU used to do encryption was greater than to do decryption
  - This is a function of the CPACF facility
- Total CPU/tx increase did not exceed **5%** when encryption was enabled
  - z14 with encrypted paging performed better than z13 without encryption

Logical Processors/ Memory (GB)	En+Decrypted paging Rate (D3R2)	En+Decrypted CPU Bsy (D3R2) (% of one IFL CPU)
08 / 512	159997.55	19.57
16 / 1024	163467.08	19.57
24 / 1536	150619.01	17.96
32 / 2048	155924.89	18.88

Encrypted Paging: D3R2EC Tool

- Tool from z/VM Performance team to track encrypted paging monitor values
- Available on z/VM download page
- Possible Performance Toolkit updates at a later date

D3R2 Encrypted paging report for file: A05Y9152 MONDATA

Interval			<------- Rate of Pages -------->			<------ Percent CPU busy ------>		
__Ended_	Type	LPU_	_Enc+Dec__	Encrypted_	Decrypted_	_Enc+Dec__	__Encrypt_	__Decrypt_
>>Mean>>	IFL	0	19451.25	11662.78	7788.47	2.45044	1.71205	0.73840
>>Mean>>	IFL	1	19036.57	9766.84	9269.73	2.31351	1.43584	0.87766
>>Mean>>	IFL	2	19153.36	9761.35	9392.01	2.32062	1.43352	0.88710
>>Mean>>	IFL	3	19010.73	9657.54	9353.18	2.32729	1.43122	0.89607
>>Mean>>	IFL	4	19131.78	9685.10	9446.68	2.33772	1.43319	0.90453
>>Mean>>	IFL	5	21139.60	9656.43	11483.17	2.50907	1.42566	1.08341
>>Mean>>	IFL	6	21351.01	9744.53	11606.48	2.53488	1.44154	1.09333
>>Mean>>	IFL	7	21167.82	9827.81	11340.01	2.52316	1.45072	1.07244
>>Total>	....	8	159442.12	79762.38	79679.73	19.31669	11.76374	7.55294
15:27:27	IFL	0	14500.07	9057.13	5442.94	1.83363	1.33507	0.49856
15:27:27	IFL	1	15452.78	8950.06	6502.72	1.91393	1.31984	0.59409
15:27:27	IFL	2	15215.59	8310.86	6904.73	1.85513	1.22522	0.62991
15:27:27	IFL	3	14394.43	7823.19	6571.24	1.78056	1.17005	0.61051
15:27:27	IFL	4	14700.28	8225.17	6475.11	1.82524	1.22422	0.60102
15:27:27	IFL	5	18332.57	8317.30	10015.27	2.14883	1.23835	0.91048
15:27:27	IFL	6	18304.86	8439.71	9865.15	2.15040	1.25402	0.89638
15:27:27	IFL	7	18117.23	8296.26	9820.97	2.12680	1.23287	0.89393
>>Total>	....	8	129017.81	67419.68	61598.13	15.63452	9.99964	5.63488
15:27:57	IFL	0	20984.71	11808.29	9176.42	2.58744	1.71926	0.86818
15:27:57	IFL	1	20038.51	8859.42	11179.09	2.34774	1.29137	1.05637
15:27:57	IFL	2	20170.38	9001.16	11169.22	2.36140	1.30838	1.05302
15:27:57	IFL	3	19741.21	8430.19	11311.02	2.31781	1.23350	1.08431
15:27:57	IFL	4	19681.81	8459.56	11222.25	2.30965	1.23409	1.07556
15:27:57	IFL	5	22587.21	8467.49	14119.72	2.56253	1.23307	1.32946
15:27:57	IFL	6	22904.38	8472.96	14431.42	2.59338	1.23633	1.35705
15:27:57	IFL	7	23478.97	9439.22	14039.75	2.70212	1.37671	1.32541
>>Total>	....	8	169587.18	72938.29	96648.89	19.78207	10.63271	9.14936

Best Practices for z/VM Encrypted Paging

- System configuration: Use **ON**, not **REQUIRED**
  - Safer for DR scenarios
  - Prevents accidental lockout
  - Switch to **REQUIRED** in AUTOLOG1 (before RACF is IPL'd)

- Test your workloads vs. ephemeral key size
  - Find the encryption strength which works best for you
  - Guidance from IBM on z/VM Performance website
    http://www.vm.ibm.com/perf/reports/zvm/html/640ep.html
  - Consider your security needs when enabling encryption at one level vs. another

- Audit your encryption
  - Monitor records – watch for updates to Performance Toolkit etc.
  - SMF records – mind your security at all times

Continuous Delivery Enhancements:
High PR/SM LPAR Management Time Relief

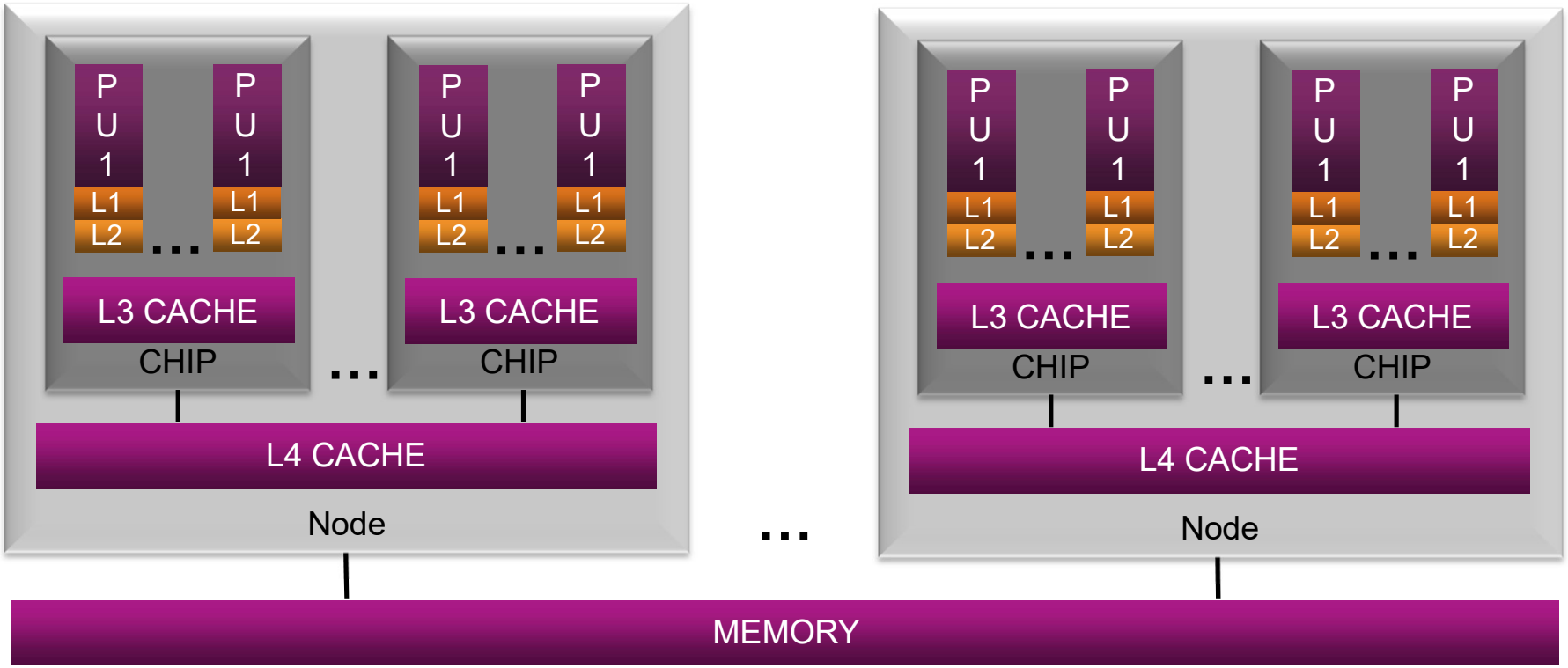
High PR/SM LPAR Management Time Relief

- New logical processor unparking schemes help prevent unnecessary use of logical processors, especially vertical-low logical processors, thereby decreasing dispatch contention inside PR/SM.
- Available October 23, 2017
- More information at http://www.vm.ibm.com/perf/tips/unpark.html

Component	APAR	PTF
CP	VM66063	UM35232

HiperDispatch – Dispatching Affinity

- Processor cache structures become increasingly complex and critical to performance
- Goal is to re-dispatch work close (in terms of topology) to where it last ran



High PR/SM LPAR Management Time Relief

- Consider a CPC like this:
  - 96 physical IFL cores
  - Four all-IFL LPARs, each with 32 logical IFL cores
  - LPAR weights set pretty close to equally
  - This yields, per LPAR:
    - 22 vertical-high logical cores
    - 2 vertical-medium logical cores
    - 8 vertical-low logical cores

- When the CPC is only moderately busy, all four z/VMs will do this:
  - They'll all sense there is plenty of spare CPC power, so…
  - They'll all try to run guests on all of their vertical-lows, which means…
  - There will be PR/SM dispatch contention on the (96-88) = 8 physical IFL cores available to run the (4 x (2+8)) = 40 vertical-mediums and vertical-lows

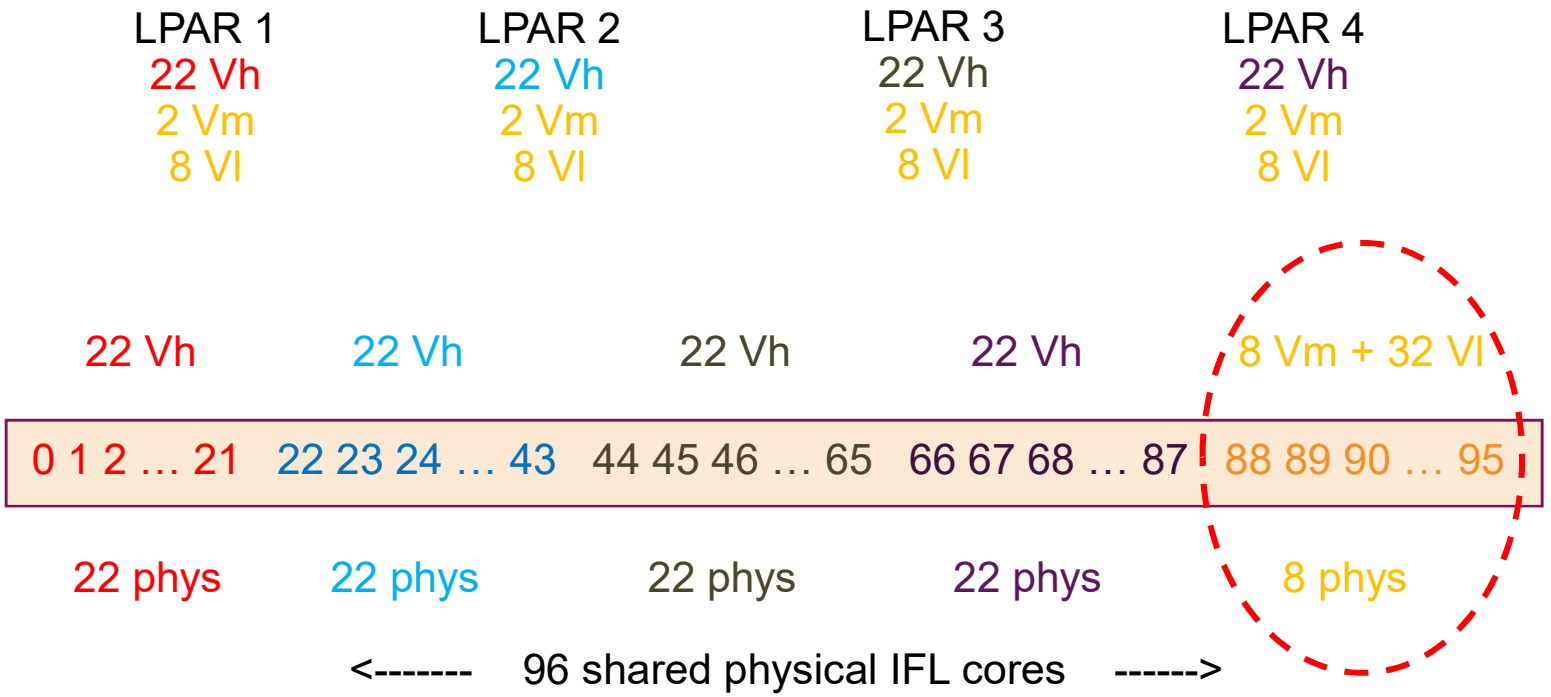
High PR/SM LPAR Management Time Relief

- Consider a CPC like this:
  - 96 physical IFL cores
  - Four all-IFL LPARs, each with 32 logical IFL cores
  - LPAR weights set pretty close to equally
  - This yields, per LPAR:
    - 22 vertical-high logical cores
    - 2 vertical-medium logical cores
    - 8 vertical-low logical cores

- When the CPC is only moderately busy, all four z/VMs will do this:
  - They'll all sense there is plenty of spare CPC power, so…
  - They'll all try to run guests on all of their vertical-lows, which means…
  - There will be PR/SM dispatch contention on the (96-88) = 8 physical IFL cores available to run the (4 x (2+8)) = 40 vertical-mediums and vertical-lows

High PR/SM LPAR Management Time Relief

Here's what PR/SM does. You can now see why this configuration is a problem if the operating systems misuse their vertical-lows.



The overcommit is not straight 128/96 as we might think.
It's more like four 1:1 and one 40:8.

Effect of Misuse of Vertical-Lows

```
FCX304  Run 2018/04/09 14:47:47          PRCLOG
                                         Processor Activity, by Time
From 2018/04/04 00:01:00
To   2018/04/04 07:59:00
For  28680 Secs 07:58:00                 Result of PMxxxxxx Run
```

Interval End Time	CPU	Type	PPD	Ent.	DVID	Pct Park Time	%Susp	<--- Percent Busy ----> Total	User	Syst	Emul
>>Mean>>	0	IFL	Vh	100	0000	0	7.9	60.0	48.0	12.1	44.8
>>Mean>>	1	IFL	Vh	100	0000	0	8.0	58.8	47.4	11.4	44.3
>>Mean>>	2	IFL	Vh	100	0001	0	7.8	59.4	47.5	11.9	44.3
>>Mean>>	3	IFL	Vh	100	0001	0	7.9	58.4	47.1	11.3	44.0
>>Mean>>	4	IFL	Vh	100	0002	0	6.9	60.5	47.2	13.3	44.2
>>Mean>>	5	IFL	Vh	100	0002	0	7.0	59.6	48.7	10.8	45.9
>>Mean>>	6	IFL	Vh	100	0003	0	7.0	60.6	48.9	11.8	45.9
>>Mean>>	7	IFL	Vh	100	0003	0	7.1	59.3	48.4	10.9	45.5
>>Mean>>	8	IFL	Vh	100	0004	0	6.8	60.1	48.8	11.3	45.8
>>Mean>>	9	IFL	Vh	100	0004	0	6.8	58.8	48.3	10.5	45.5
>>Mean>>	10	IFL	Vh	100	0005	0	6.8	59.7	48.4	11.3	45.5
>>Mean>>	11	IFL	Vh	100	0005	0	6.8	58.4	47.9	10.5	45.1
>>Mean>>	12	IFL	Vh	100	0006	0	6.8	58.7	47.5	11.2	44.7
>>Mean>>	13	IFL	Vh	100	0006	0	6.9	57.4	47.0	10.3	44.2
>>Mean>>	14	IFL	Vm	52	0007	0	49.0	31.2	24.9	6.3	23.5
>>Mean>>	15	IFL	Vm	52	0007	0	49.2	29.9	24.2	5.7	22.9
>>Mean>>	16	IFL	Vl	0	0008	0	48.8	30.5	24.5	6.0	23.2
>>Mean>>	17	IFL	Vl	0	0008	0	48.9	29.3	23.9	5.4	22.7
>>Mean>>	18	IFL	Vl	0	0009	0	48.2	29.7	23.9	5.8	22.6
>>Mean>>	19	IFL	Vl	0	0009	0	48.4	28.7	23.5	5.3	22.2
>>Total>	20	IFL	MIX	1504	MIX	1	393.1	1009	816.0	193.0	766.8

High PR/SM LPAR Management Time Relief

- Solution: unpark only what is needed **and** powered.

- Now there are three unparking models:
  - LARGE: unparks like today - aggressive use of vertical-lows
  - MEDIUM: unparks all vertical-highs and vertical-mediums, and only the vertical-lows it appears are *needed **and** powered*
  - SMALL: unparks only the logical processors it appears are needed **and** powered
    - (this will park vertical-highs and vertical-mediums)

- Parking unneeded logical processors can help reduce PR/SM dispatch contention

- **CP SET SRM UNPARKING {LARGE | MEDIUM | SMALL}**

- System configuration file statement: **SRM UNPARKING** …

Continuous Delivery Enhancements: z-Thin Provisioning Enhancements

z-Thin Provisioning

- Users of DS8K Extent-Space-Efficient (ESE) devices will now be able to fully monitor and manage ESE disk pool space from z/VM and use those devices for any use case.
- Available March 28, 2018

Component	APAR	PTF
CP	VM66098 PE: VM66153	UM35296 UM35317
DIRMAINT	VM66108	UV99329

z-Thin Provisioning - Prior Support

- Minor support for Track-Space-Efficient (TSE) devices
- Extent-Space-Efficient (ESE) devices could be used for guest disks
- Space-Efficient devices were not allowed for CP-owned space
  - enforced at IPL (wait state) and **ATTACH to SYSTEM**
- No recognition or reporting of pool space events.
- No method of managing/returning disk space no longer in use.
- **QUERY DASD SPACE-EFFicient** and **QUERY CU SELC** provided TSE pool space statistics.
  - Did not recognize ESE devices as a valid type of Space-Efficient device

z-Thin Provisioning – New Support

- Guests can now use Space-Efficient management CCWs for:
  - Full pack and 1-END minidisks
  - Dedicated DASD

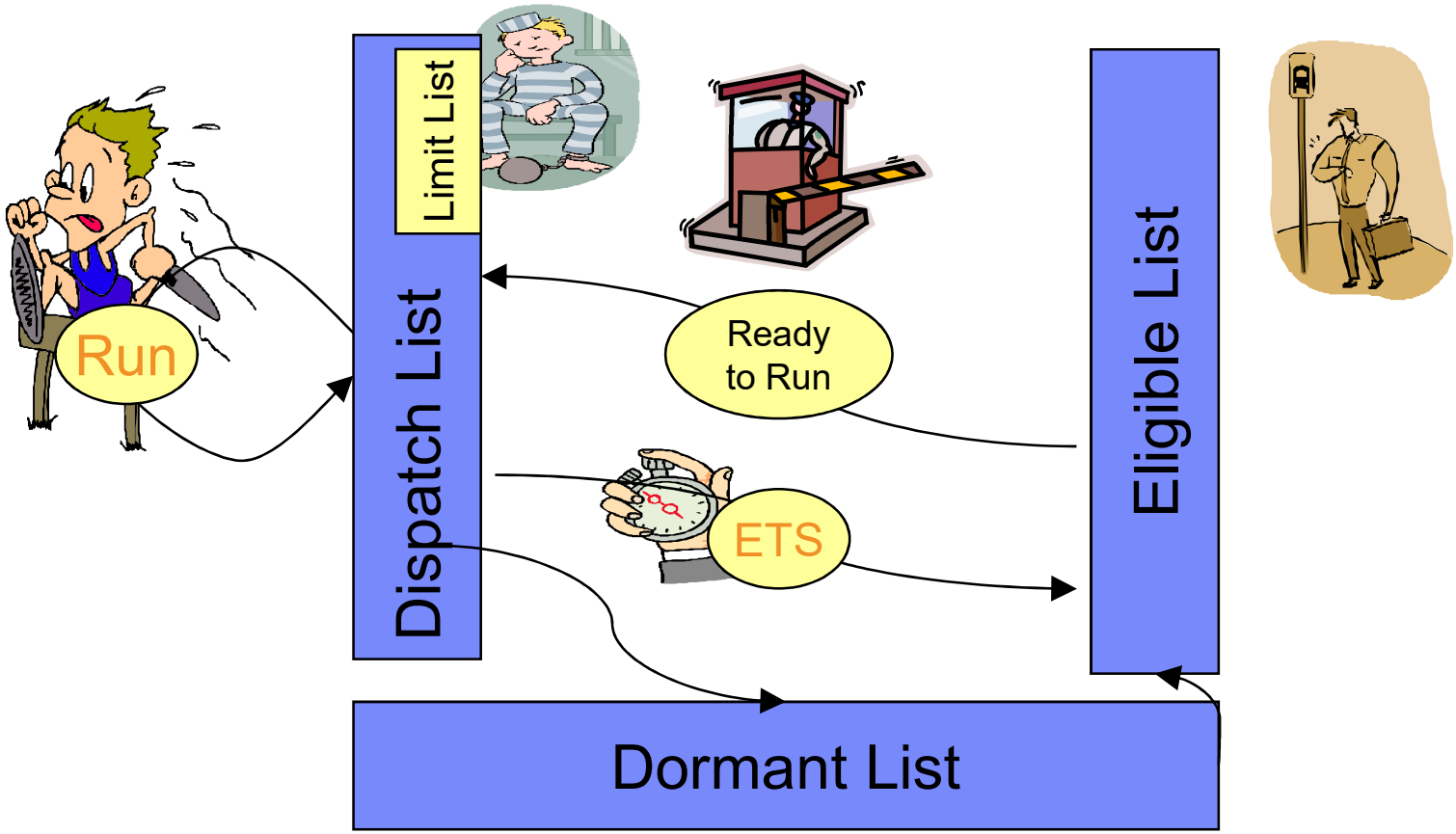
- Host recognition and exploitation
  - Dynamic reporting of pool space events to OPERATOR
    - Usage warning percentage reached
    - Out of space
    - Space constraint has been relieved
  - ESE pool usage statistics via **Q DASD SPACE-EFFicient** and **Q CU SELC**
  - New **RELSPACE** command returns extents to their associated ESE pool
    - Contents are erased when extents are returned
    - Exploited by DirMaint

z/VM 6.4 Base Enhancements

Summary of Base z/VM 6.4 Enhancements

- Scheduler improvements
- Security enhancements
- No longer require SVC to use FlashSystems for system volumes
- New query and display information for disk devices
- Ability to verify FCP definitions
- 2 TB real memory support
- Enhanced DAT – lets guest operating systems use 1 MB pages
- Ability to free up paging disk space (KEEPSLOT = NO)
- Use of HyperPAV and High Performance FICON (zHPF) for paging
- Dynamic SMT
- CP Environment Variables
- Query CPSERVICE
- SHUTDOWN enhancements
- VMREVIEW tool

Scheduler Lists



Eligible List

- z/VM 6.4 no longer places virtual machines into the eligible list. The eligible list is still defined and is displayed in various commands.
  - In the past, the wrong virtual machines went into the eligible list for too long
  - No longer need to worry about SET SRM STORBUF and LDUBUF settings
  - Need to ensure that you have sufficient system resources to avoid thrashing scenarios

- Check to see if you have had eligible lists forming in a case where they were needed.
  - Performance Toolkit SCHEDLOG report can show this
  - If you have had these scenarios, contact IBM to discuss options

- The QUICKDSP option on a virtual machine was used in past to ensure critical virtual machines always bypassed the eligible list.
  - Current recommendation is to not remove this option from machines where it is currently set.

Scheduler Changes

- z/VM 6.4 improves the accuracy in the distribution of processor power
  - Existing problem where surplus 'share' is not distributed appropriately has been addressed

- The algorithms were changed to help accommodate this fix resulting in share values being normalized differently
  - All virtual machines are factored into the normalization, not just virtual machines in the dispatch and eligible lists.

Surplus Share Distribution: Background

- Shares are relative to other virtual machines that want to run (in dispatch and eligible lists)

- Example:
  - Four compute-bound virtual machines on a real 1-way:
    - LINUX01 Relative 100 = 17%
    - LINUX02 Relative 100 = 17%
    - LINUX03 Relative 200 = 33%
    - LINUX04 Relative 200 = 33%
  - Total Shares = 600
  - What happens if LINUX04 wants to use only **3%**?

Excess Share Distribution Problem

User ID	Share	Normalize	Correct Distribution (z/VM 6.4)	Problem Scenario (z/VM Prior)
LINUX01	100	17%	**24.5%**	**17%**
LINUX02	100	17%	**24.5%**	**17%**
LINUX03	200	33%	**48%**	**63%**
LINUX04	200	33%	**3%**	**3%**

Normalization Change

- z/VM 6.3 and earlier normalization
  - z/VM kept accumulated share values for virtual machines in the dispatch and eligible lists; one for absolute shares and one for relative shares
  - For absolute share:
    - If sum of absolute shares of virtual machines > 99%, prorate to 99%
    - Else absolute share → normalized share
  - For relative share:
    - Determine what is left over from absolute shares (always at least 1%)

$$normalized\ share \quad = (100 \ - \ \sum absolute\_share\_disp\_list\ ) \times \frac{relative\_share}{\sum relative\_share\_disp\_list}$$

- In z/VM 6.4 the sums include *all* users, not just those in the dispatch and eligible lists
  - Watch for systems where:

$$\sum relative\_share \gg \sum relative\_share\_disp\_list$$

- The sum of absolute and relative shares is provided in the SCHEDLOG Performance Toolkit report
- This is done for each processor type in all releases

Security Changes

- z/VM SSL Server
  - Default in z/VM 6.4 is TLS 1.2, with TLS 1.0 disabled
- System Config file
  - Passwords_on_cmds feature now defaults to “No”
- Logon error message
  - If an incorrect password is given for a valid userid, the error message no longer indicates that the userid was valid
  - If invalid userid is entered without password, we still prompt for password
  - Meant to prevent phishing

```
HCPLGA050E LOGON unsuccessful - incorrect userid and/or password
```

Using FlashSystems for z/VM system volumes

- Prior to z/VM 6.4, you needed a San Volume Controller (SVC) to use FlashSystems for z/VM volumes
  - Could be connected to Linux guests without the SVC

- New device attribute (driver) for **EDEVICE** statement or **SET EDEVICE** command

- System configuration file:

EDEVICE *edev* TYPE FBA ATTRIBUTES FLASH FCP_DEVICE *rdev* WWPN *wwpn* LUN *lun*

New Query Info for Disk Devices

- Extended Information on QUERY commands
  - **Query EDEV nnnn details** – added LUN serial number
  - **Query DASD nnnn details** – added serial number

```
Query edev 1111 details

EDEV 1111 TYPE FBA ATTRIBUTES 2105
  VENDOR: IBM       PRODUCT: 2105F20          REVISION: .293
  BLOCKSIZE:         512  NUMBER OF BLOCKS: 390656
  PATHS:
    FCP_DEV: B908  WWPN: 5005076300CD04DA  LUN: 5144000000000000
      CONNECTION TYPE: POINT_TO_POINT STATUS: ONLINE
  EQID: ABCDEFGH
  SERIAL NUMBER: 2146561344562
```

New Query Info for Disk Devices

- Extended Information on **QUERY EDEVICE**
  - New **inquiry** option to provide data from the device: Standard Inquiry Info and Vital Product Data

```
q edev 111 inquiry

- Begin - EDEV 0111 - Standard Inquiry Page -
00000532 9F101002 49424D20 20202020 32313037 39303020 20202020 20202020
2E323034 37353034 31393131 34303020 20202020 20202020 00600DA0 0A000300
03200000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00323037 35003236 34303400 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00080000
- End - EDEV 0111 - Standard Inquiry Page -

q edev 111 inquiry page 83

- Begin - EDEV 0111 - Vital Product Data Page 83 -
00830024 01030010 60050763 03FFC09C 00000000 00001400 01140004 00000032
01150004 00000000
- End - EDEV 0111 - Vital Product Data Page 83 -
```

Additional Information on DASD

- For ECKD disks get Read Device Characteristics (RDC) and Read Configuration Data (RCD)
  - **QUERY DASD** with **CHARACTERISTICS** option

```
q dasd char 521d
- Begin - RDEV 521D - Read Configuration Data -
DC010100 F0F0F2F1 F0F7F9F0 F0C9C2D4 F7F5F0F0 F0F0F0F0 F0E8F5F8 F1F1071D
D4020000 F0F0F2F1 F0F7F9F3 F2C9C2D4 F7F5F0F0 F0F0F0F0 F0E8F5F8 F1F10700
D0000000 F0F0F2F1 F0F7F9F3 F2C9C2D4 F7F5F0F0 F0F0F0F0 F0E8F5F8 F1F00700
F0000001 F0F0F2F1 F0F7F9F0 F0C9C2D4 F7F5F0F0 F0F0F0F0 F0E8F5F8 F1F10700
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
80000310 2D001E00 05070013 62131325 000CC01D 3905FADB 03100000 0000F200
- End - RDEV 521D - Read Configuration Data -
- Begin - RDEV 521D - Read Device Characteristics -
2107E833 900A5F8C 5FF72024 01F4000F E000E5A2 05940222 13090674 00000000
00000000 00000000 24241F02 DFEE0001 0677080F 007F4A00 003C0000 000001F4
- End - RDEV 521D - Read Device Characteristics -
```

IOEXPLOR Exec – FCP Example

- Used to format new output.

IOEXPLOR 7FFF CHAR

-Begin: Characteristics Data for device 7FFF
Serial Number _____ 052a62e2052a-0000002a-00002c
Standard Inquiry Data
Peripheral Qualifier/Peripheral Device Type _____ 000b/00h
Vendor Identification _____ IBM
Product Identification _____ FlashSystem-9840
Product Revision Level _____ 1217
Version Descriptor _____ SAM-3 (no version)
Version Descriptor _____ FC-PH-3 (no version)
Version Descriptor _____ FC-AL-2 (no version)
Version Descriptor _____ FCP-3 (no version)
Version Descriptor _____ SPC-3 (no version)
Version Descriptor _____ SBC-2 (no version) (cont'd….)

IOEXPLOR Exec – FCP Example

```
                Device Identification
Cluster Identification _____________________________052a62e2052a
  IO Group _________________________________________ 0000
  Vdisk Number _____________________________________ 002a
  LUN Identification _______________________________ 00002c
  IEEE Company Identification ______________________ 005076
  Cluster Alias ____________________________________ 12A62E2052A
  Slot Number ______________________________________ 0C
  Channel Number ___________________________________ 01
                Device Characteristics
  Device class code ________________________________ 21
  Unit type ________________________________________ 11
  Bytes per track __________________________________ 56832
  Bytes per cylinder _______________________________ 397824
  Bytes per block __________________________________ 512
  Device size ______________________________________ 2097152 blocks
 -End: Characteristics Data for device 7FFF
 Ready; T=0.01/0.01 21:20:09
```

IOEXPLOR Exec – FICON Example

IOEXPLOR 19E CHAR

```
-Begin: Characteristics Data for device 19E
                        I/O Device Information
Device type-model _____________________________________2107-900
Device manufacturer ___________________________________ IBM
Serial number (plant-seq#) ____________________________ 75-Y5811
Logical Volume Number _________________________________1040
            Control Unit Information
Device type-model _____________________________________2107-932
Serial number (plant-seq#) ____________________________75-Y5811
Logical Subsystem Number ______________________________10
            Additional Device Information
Device manufacturer ___________________________________ IBM
Device type-model ____________________________________ 2107-932
Serial number (plant-seq#) ____________________________ 75-Y5810
Logical Subsystem Number ______________________________10
            Additional Device Information
Device manufacturer ___________________________________ IBM
Device type-model ____________________________________ 2107-900
Serial number (plant-seq#) ____________________________75-Y5811
Logical Subsystem Number ______________________________10
```

(cont'd...)

IOEXPLOR Exec – FICON Example

General NEQ

Interface id	0230
Missing Interrupt Timer Interval	30 seconds
Secondary Missing Interrupt Timer Interval	0 seconds
Controller System Adapter ID (SAID)	0230
Logical paths supported	61952

Device

Host CU type-model	2107-E8
Device type-model	3390-0A

Storage Directory Facilities

VM non-full pack minidisk	Yes
MIDAW Capability supported	No
Parallel Access Vol. state	HyperPAV Enabled
XRC Functions	Enabled
Peer-to-Peer Remote Copy	Not Enabled
Striping and Compaction	Supported
Locate Record Erase	Supported
Cache Fast Write	Supported
Multi-Path Lock	Supported
Track Cache	Supported

(cont'd...)

IOEXPLOR Exec – FICON Example

DASD Fast Write _____ Supported
24 Byte Compatibility sense _____ Yes
Device class code _____ 20
Device type code _____ 24
Primary cylinders _____ 500
Tracks per cylinder _____ 15
Number of Sectors _____ 224
Track length _____ 58786
HA + R0 length _____ 1428
Capacity formula _____ 2
Capacity factors F1-F6 _____ 34 19 9 6 116 6
MDR Record ID _____ 24
OBR Record ID _____ 24
Storage director Type _____ 1F
Read Trackset length _____ 2
Max Record zero length _____ 57326 (cont'd...)

IOEXPLOR Exec – FICON Example

```
                Storage Class
Data Encrypted device _____________________________ No
Solid State drive _____________________________ No
Enterprise Disk _____________________________ No
SATA Disk _____________________________ No
Flash Storage _____________________________ No
Tiered Storage Pool _____________________________ No
Track Set Size _____________________________1
Concurrent Copy Lower _____________________________ 0F
Concurrent Copy Upper _____________________________ 7F
                Generic Device / CU functions
Mirrored Device _____________________________ No
RAID Device _____________________________Yes
Transparent subsystem cache _____________________ No
Split CE/DE _____________________________Yes
Device capable of Mirroring _____________________ No
XRC Device Management enabled _____________ Yes
RVA Snapshot supported _____________________ No
Real Control Unit code _____________________ 00
Real Device Code _____________________________3C
-End: Characteristics Data for device 19E
```

FCP Problem Determination

- New CP Command **EXPLORE FCP** allows for testing of FCP subchannels and WWPN ports
  - **ADD**: adds FCP subchannel and WWPN port to list of devices to be tested
    - (can also **REMOVE**)
  - **START**: activates FCP subchannels and opens WWPN ports in list of SCSI devices to be tested
    - (can also **STOP**)
  - **QUERY**: displays the FCP subchannels and WWPN ports in the list of SCSI devices to be tested and their current activation status

Performance Toolkit

- Performance Toolkit for z/VM 6.4 runs in a **z/CMS** virtual machine
  - Allows exploitation of
    - more memory for processing large amounts of data
    - z/Architecture instructions for performance benefits

- Ensure virtual machines that utilize Performance Toolkit can run in **z/CMS**
  - z/CMS and XC mode virtual machines are incompatible
  - No exploitation of z/VM data spaces
    - SFS dircontrol file directories

2 TB Real Memory Support

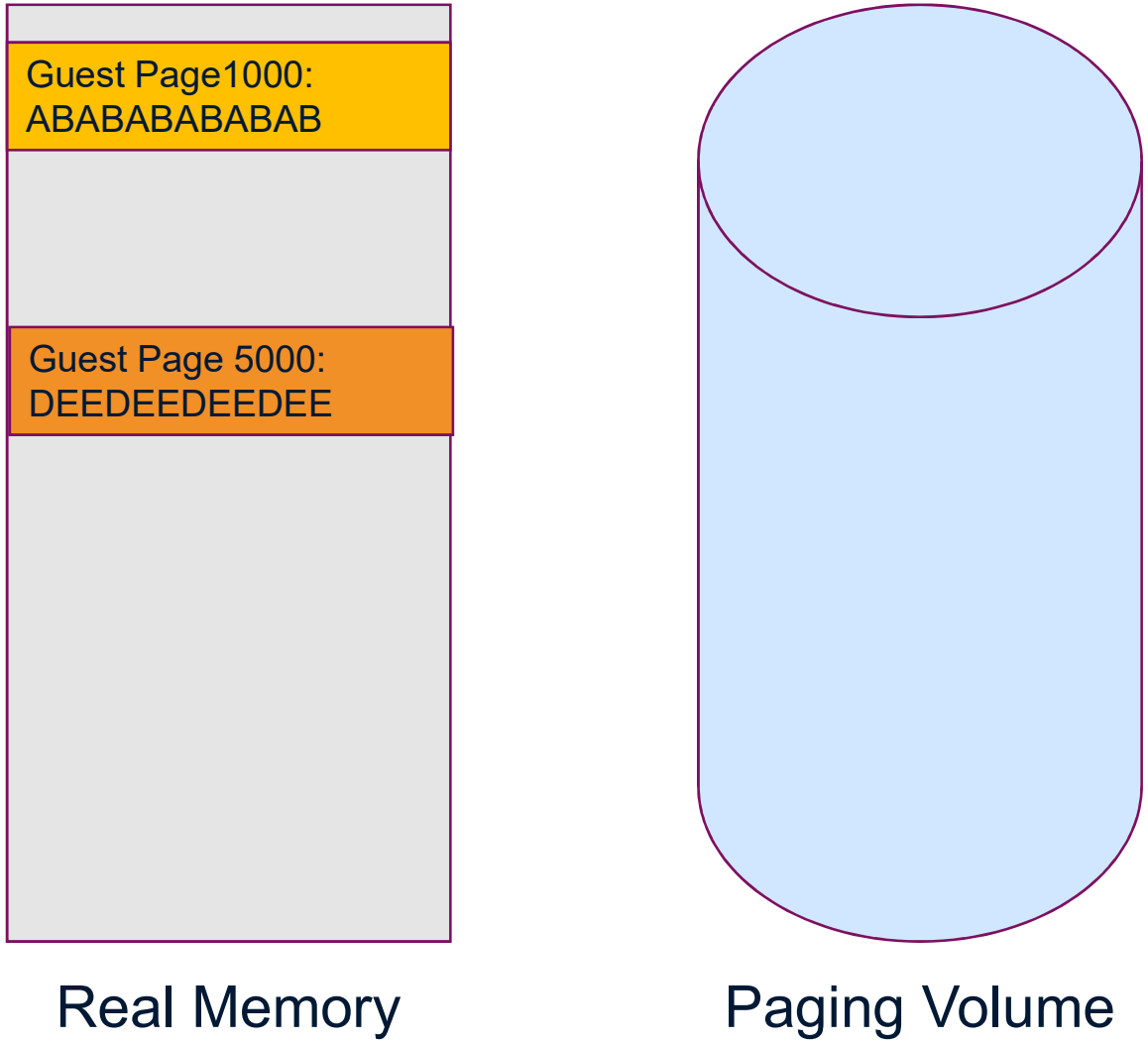
- z/VM 6.4 increases supported real memory from 1 TB to 2 TB
- Virtual machine limit remains at 1TB
- If exploiting, ensure
  - Sufficient dump space
  - Sufficient paging space
- Even if not increasing memory used, a good time to double check space guidelines

Guest Large Page

- z/VM 6.4 adds guest support for Enhanced DAT, providing 1 MB pages for guest.
  - Continue to be managed as 4 KB pages at the z/VM host level
  - Reduces memory requirements for guest

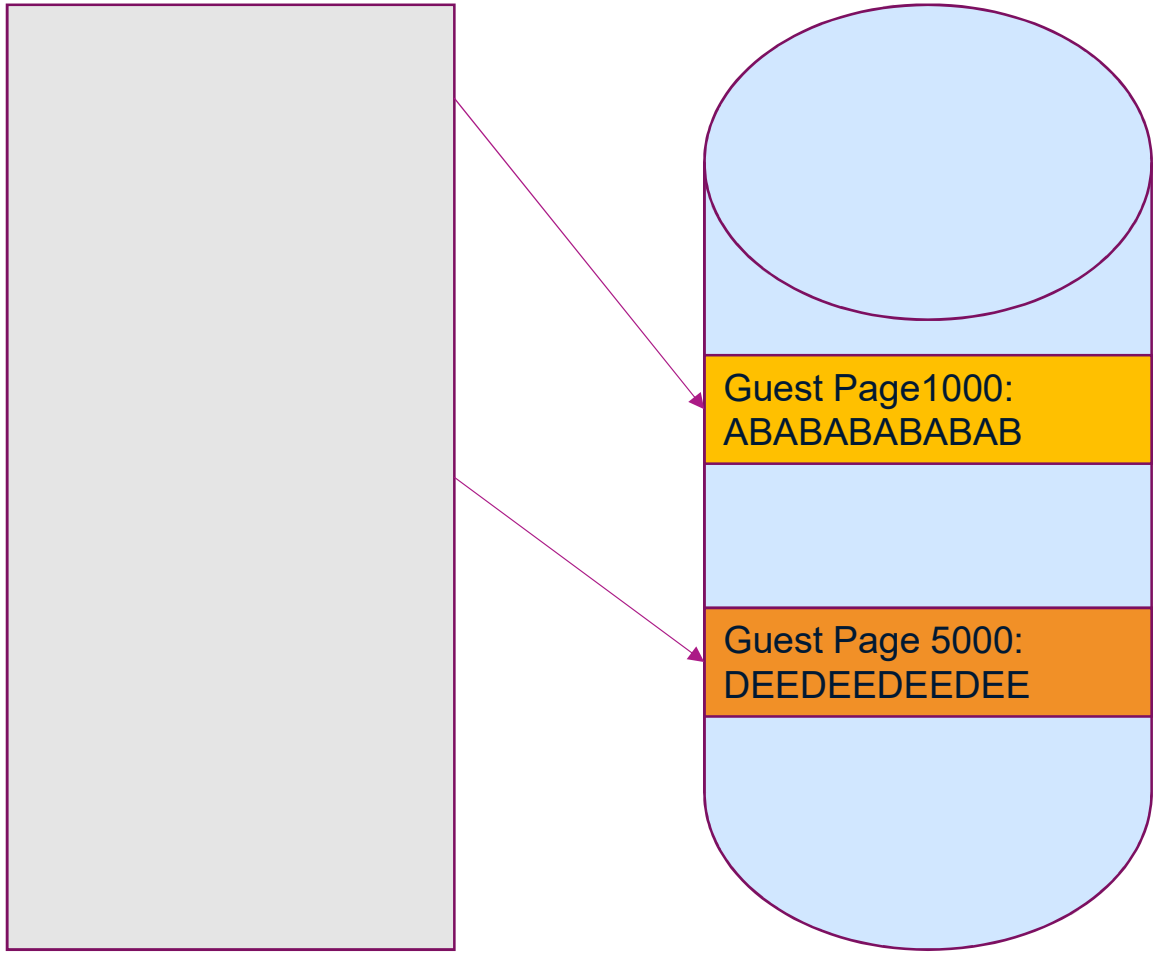
- To use this from Linux:
  - Build a kernel containing large page exploitation (this is the default build)
  - Add hugepages=<n> kernel parameter (number of large pages to be allocated at boot time)
  - If desired, set sysctl variable to enable allocating large pages from moveable memory

KEEPSLOT - Background



z/VM determines it needs to page out Guest Pages 1000 and 5000

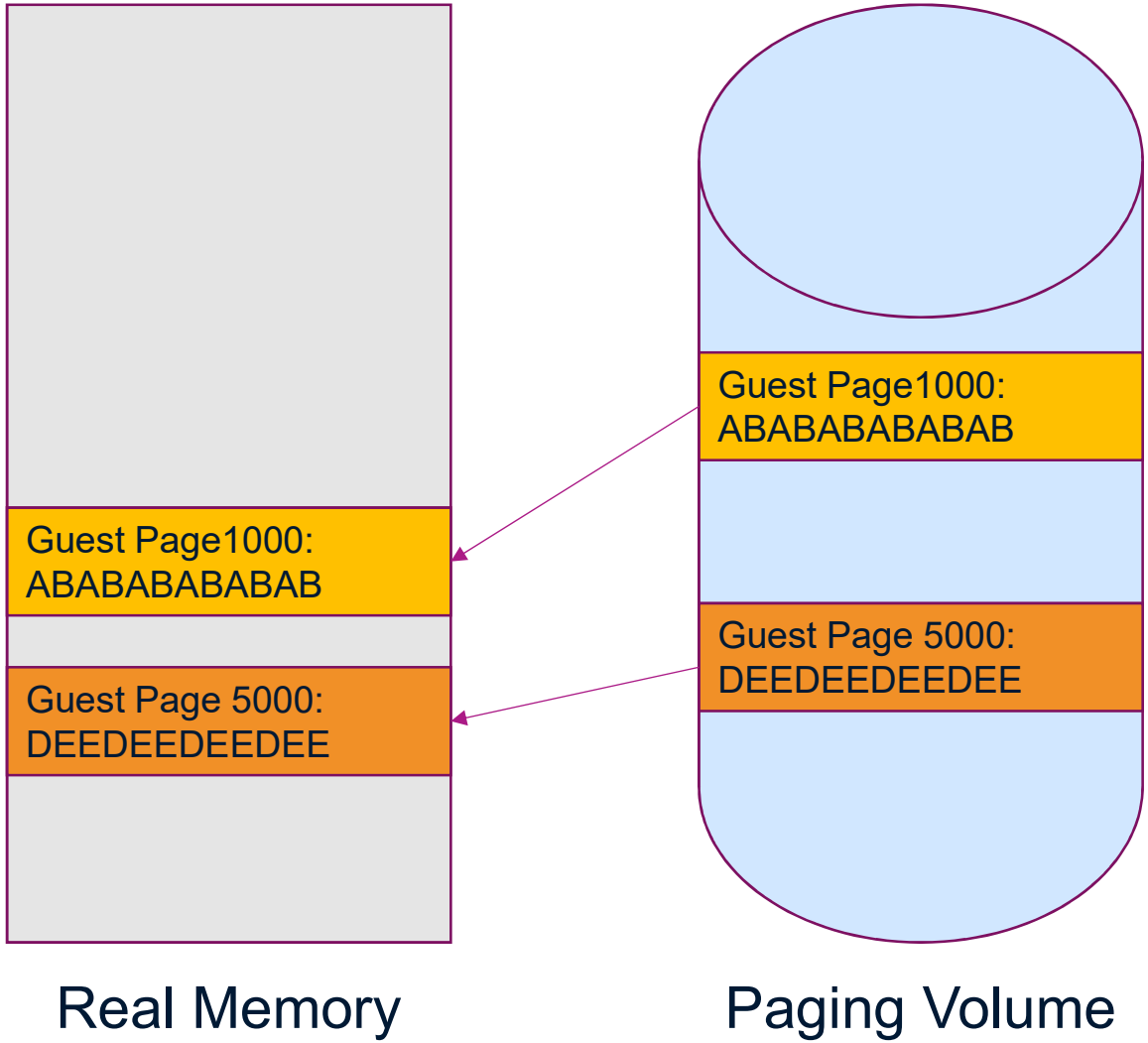
KEEPSLOT - Background



z/VM will select slots on a paging volume and write out the page.

(Actually it writes out a “set” of pages with this I/O).

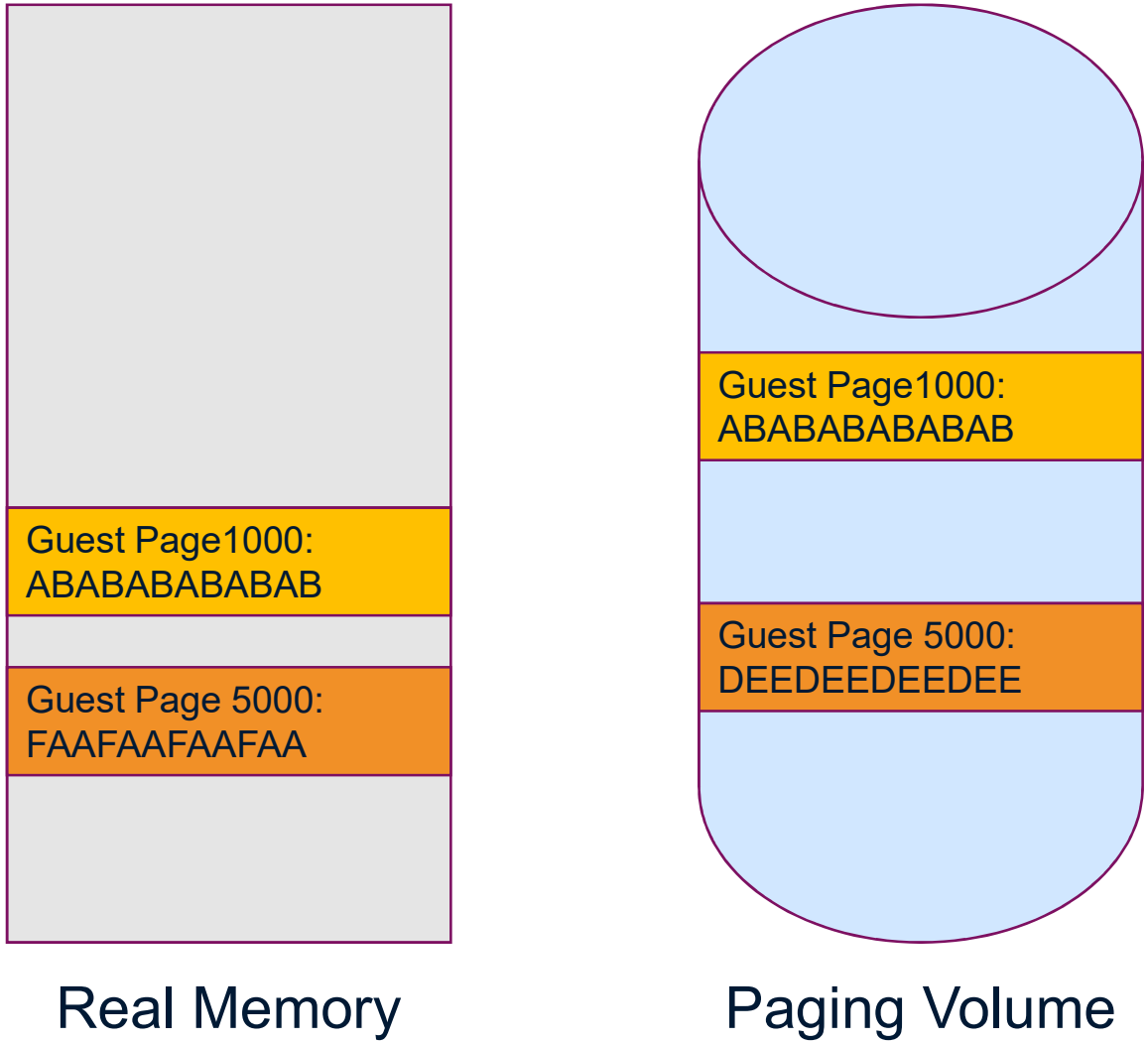
KEEPSLOT - Background



At some future time, the guest may reference the page that was paged out and z/VM page it back into real memory. But we leave the page in the disk slot as well.

This means we actually have two copies of the guest pages at this time.

KEEPSLOT - Background



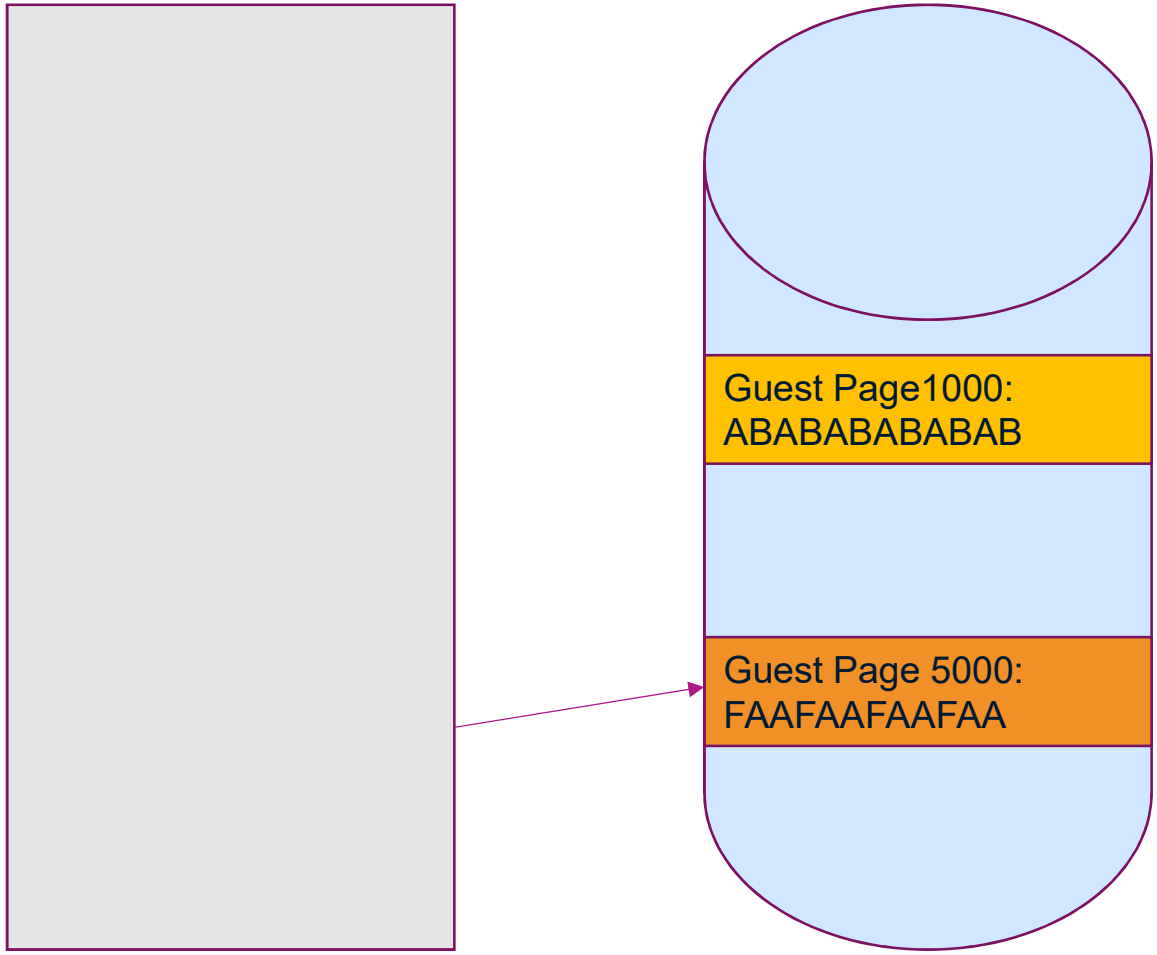
Real Memory

Paging Volume

Over time, lets assume that Page 5000 is changed.

Now the copy on disk doesn't match what's in memory.

KEEPSLOT - Background



When we have to steal frames again, we do not need to write out page 1000 because that has not changed.

Page 5000 will be re-written because it changed since it was paged in.

KEEPSLOT

- z/VM does not remove guest pages from disk when they are paged in *("keeps the slots")*
  - Avoids the need to re-write pages that have not changed

- Downside - this can result in larger paging space requirements
  - Especially after z/VM 6.3, where early writes were introduced

- z/VM 6.4 introduces a new **AGELIST** option to disable this
  - For environments where the overcommit level is low and large amounts of real memory are being used, you will want to consider disabling early writes and keeping disk slots

    - Command

      SET AGELIST EARLYWRITES NO KEEPSLOT NO

    - System configuration file:

      STORAGE AGELIST EARLYWRITES NO KEEPSLOT NO

Paging Use of HyperPAV

- Applies to paging I/O to ECKD volumes on storage servers that support HyperPAV
  - Allows a pool of alias volumes to be associated with base volumes, allowing z/VM to start more than one I/O at a time.

- On existing systems check for queuing on z/VM paging volumes
  - Performance Toolkit FCX109 DEVICE CPOWN report
  - Page queues not reflected on the FCX108 DEVICE report

- On existing systems check for impact to virtual machines of queuing
  - Performance Toolkit FCX114 User State Sampling report shows page wait in %PGW and %PGA columns

- Set up HyperPAV paging
  - Recommend enabling via command and if no surprises, update system configuration file
    - Command: **SET PAGING ALIAS ON**
    - Configuration file: **FEATURES ENABLE PAGING_ALIAS**
  - Can also be controlled at control unit level

Paging Use of HyperPAV

- Recommend using a single logical control unit (LCU) for paging and other z/VM system volumes
- If you mix user volumes and paging volumes that exploit HyperPAV in the same LCU there can be contention
- Controls added to help influence bias for alias use between minidisk and paging usage
  - Configuration file:
    CU HYPERPAV *ssid* ALIAS MDISK_SHARE *nnnn* PAGING_SHARE *nnnn*
  - Command:
    SET CU ALIAS MDISK_SHARE *nnnn* PAGING_SHARE *nnnn ssid*
- Exploitation of HyperPAV makes use of larger paging volumes more feasible
- Still recommend having at least as many paging volumes as you have logical processors for the z/VM system

Paging Use of High Performance FICON (zHPF)

- z/VM 6.4 introduced use of zHPF, transport mode, for z/VM system I/O (paging) for ECKD devices on storage servers that support zHPF

- Set up paging with zHPF
  - Recommend enabling via command and if no surprises, update system configuration file
    - Command: **SET PAGING HPF ON**
    - Configuration file: **FEATURES ENABLE PAGING_HPF**

Dynamic SMT

- z/VM 6.4 allows one to dynamically change the number of active threads per core when SMT has been enabled in the system configuration file.
- Requires z14, z13, z13s, LinuxONE Emperor or LinuxONE Rockhopper
- Decide if more than 32 cores are required, if so cannot use SMT even with one active thread per core
- System configuration file statement enables SMT-1 (1 thread per core)

  MULTITHREADING ENABLE TYPE ALL 1

- Once z/VM has started, toggle between 1 and 2 threads via CP command:

  SET MT TYPE ALL 2

  – May take a few seconds to transition.

Dynamic SMT

- With SMT-1, the real processor addresses will all be even, skipping the 2nd processor that would be shown with SMT-2
  - SMT-1

```
Query processor
PROCESSOR 00 MASTER IFL
PROCESSOR 02 ALTERNATE IFL
PROCESSOR 04 ALTERNATE IFL
```

  - SMT-2

```
Query processor
PROCESSOR 00 MASTER IFL
PROCESSOR 01 ALTERNATE IFL
PROCESSOR 02 ALTERNATE IFL
PROCESSOR 03 ALTERNATE IFL
PROCESSOR 04 ALTERNATE IFL
PROCESSOR 05 ALTERNATE IFL
```

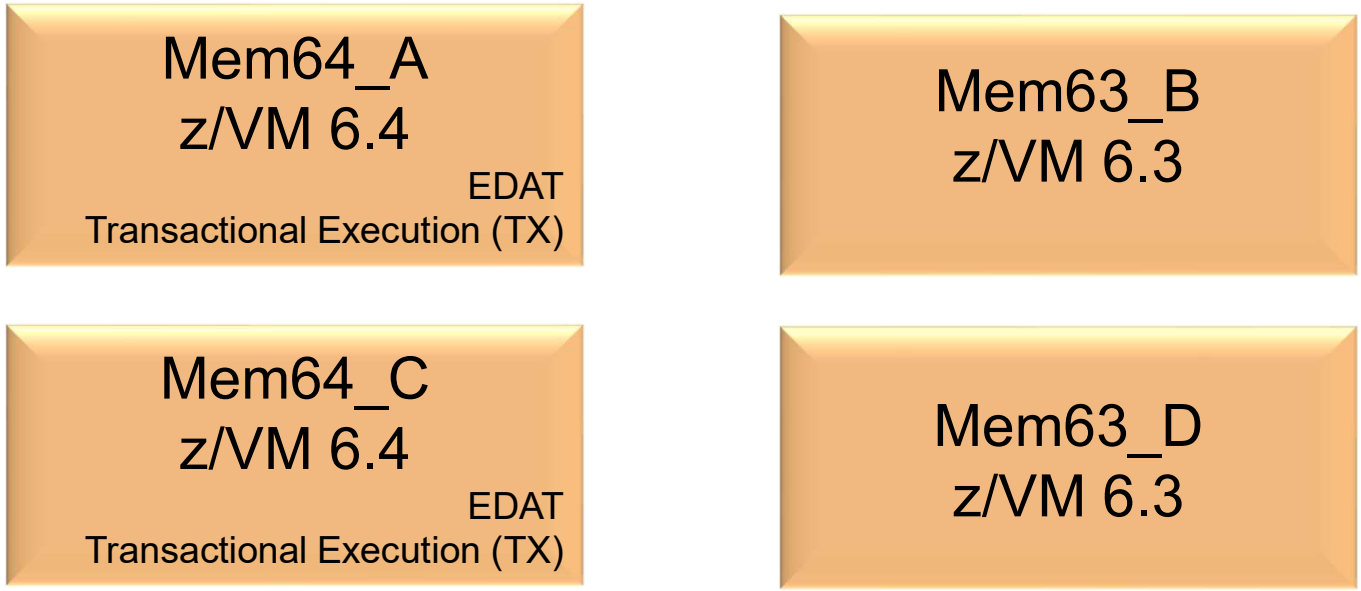
Live Guest Relocation Considerations

- Live Guest Relocation (LGR) supports relocation domains
  - Allows relocation among SSI cluster members which do not have identical configurations/capabilities.

- A relocation domain provides a common architectural level among all members regardless of differences in the facilities of its individual members
  - This is the maximal common subset of all of the members' facilities
    - Only the facilities in this subset are available to guests assigned to the domain

Relocation Domains

- z/VM 6.4 introduces guest support for two architectures that are not available to guests on z/VM 6.3:
  - Enhanced DAT (large page)
  - Transactional Execution Facility (TX)
- Guests that are assigned to relocation domains which include members running both z/VM 6.4 and z/VM 6.3 will not see these new facilities

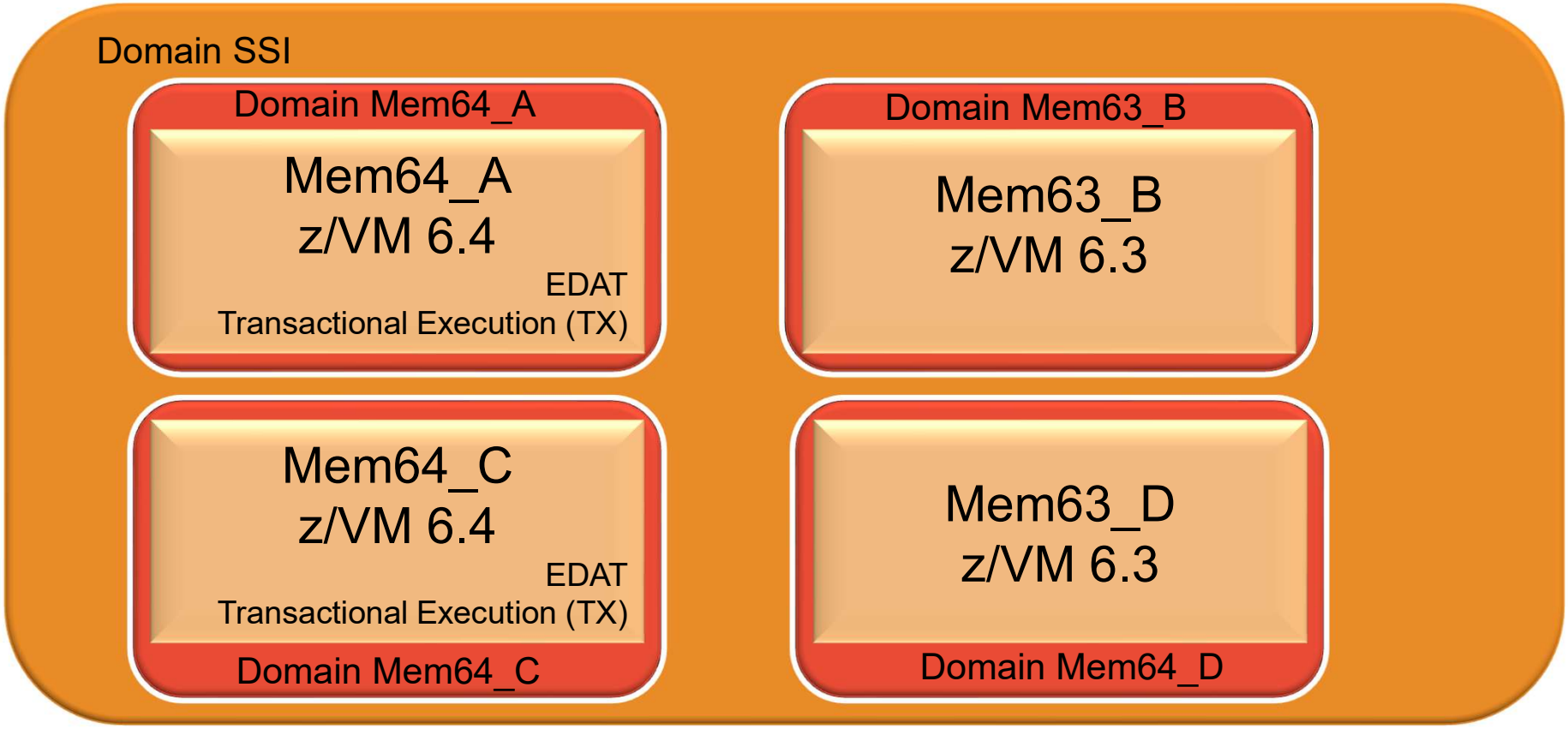


Relocation Domains – SSI Domain

- The SSI domain is a default relocation domain that includes all members of an SSI Cluster
  - Guests are assigned to the SSI domain by default when they are logged on.
  - In this cluster, guests that are assigned to the SSI domain will not have access to the EDAT and TX facilities since they are available on z/VM 6.4 but not z/VM 6.3.

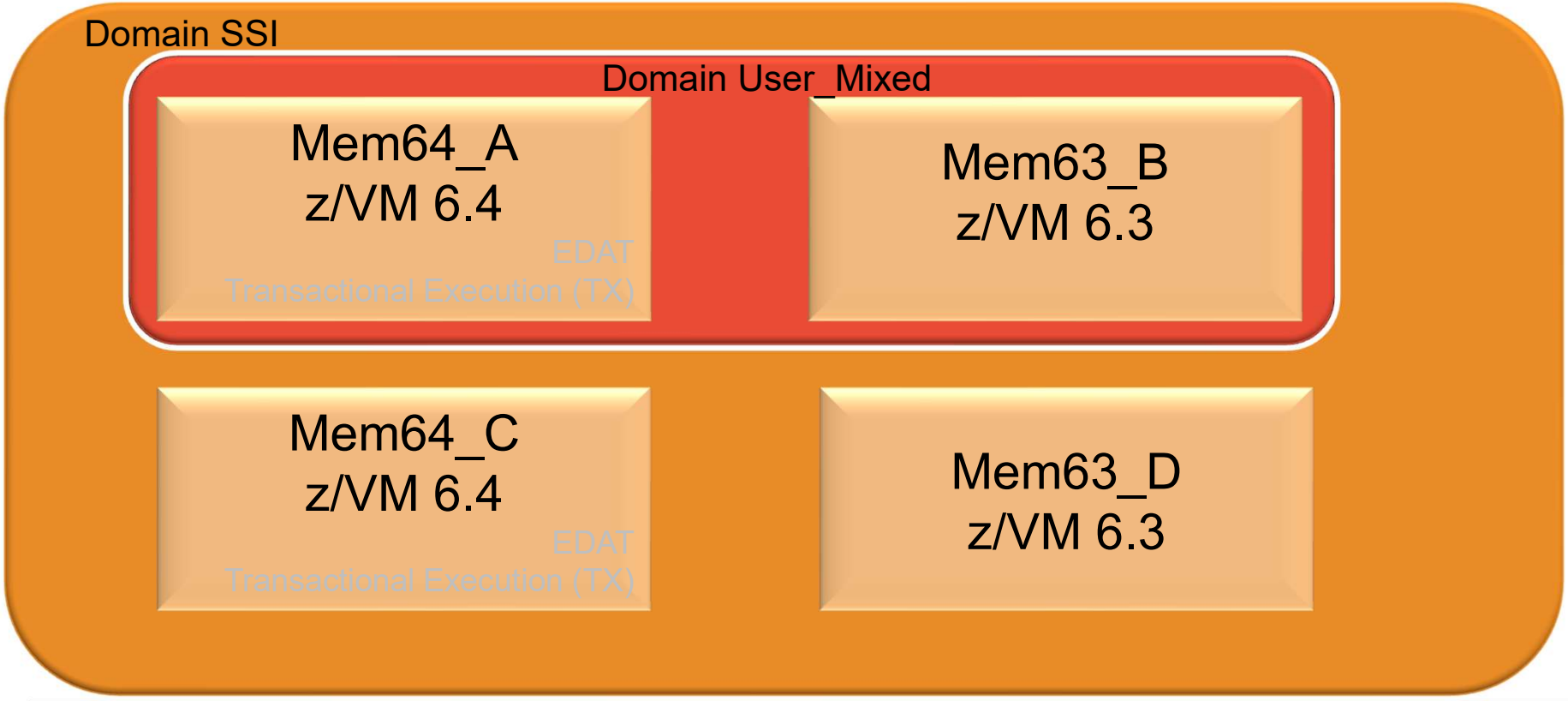
Relocation Domains – Member Domains

- Each member of a SSI cluster has its own default "singleton" relocation domain.
  - Guests that are assigned to the singleton domains have access to all of the facilities on that member.
    - Mem64_A and Mem64_C domains will have access to the EDAT and TX facilities



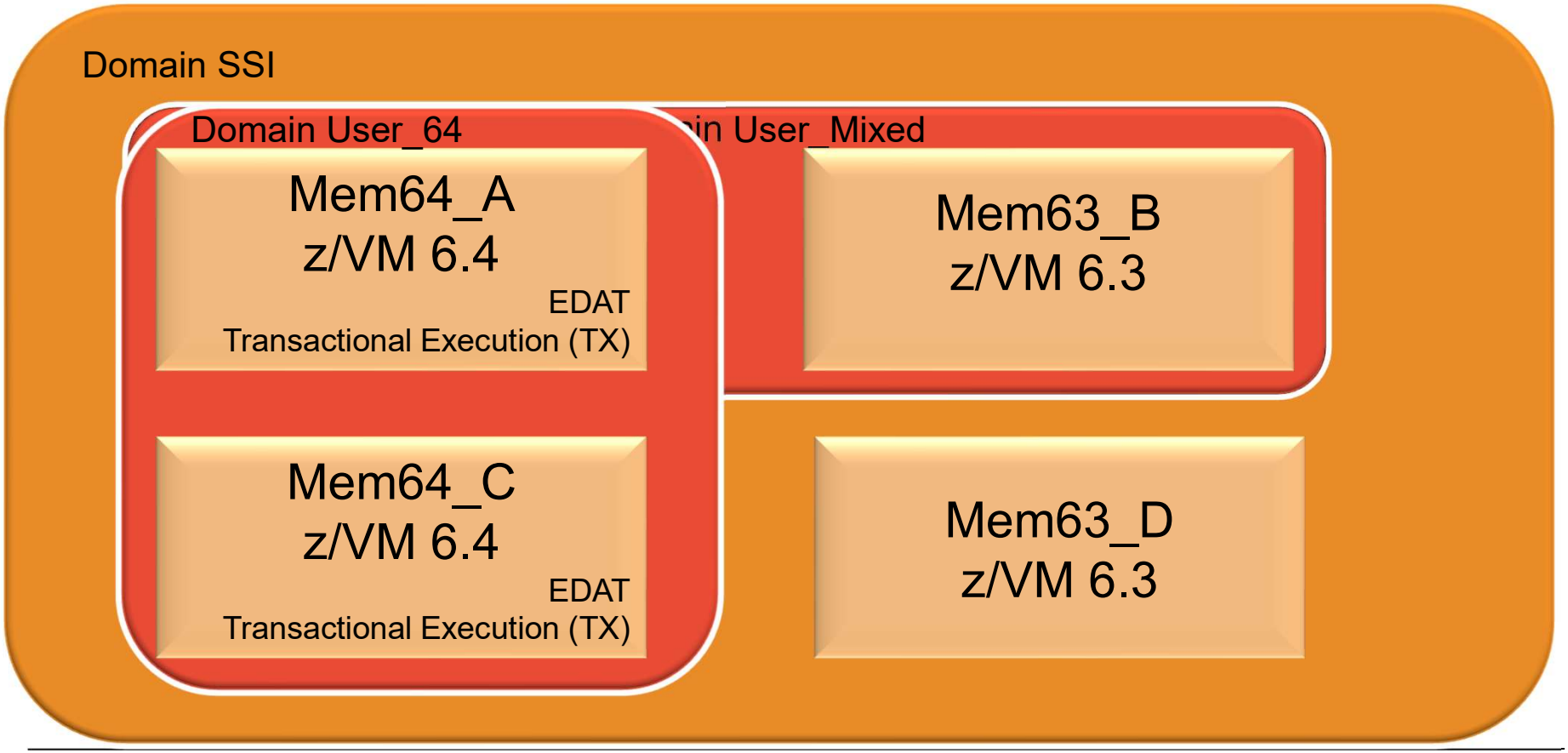
Relocation Domains – User-Defined Domains

- You can define additional relocation domains which can include any combination of members in the SSI cluster
  - Like the SSI domain, guests assigned to the User_Mixed domain will not have access to the EDAT and TX facilities since they are available on z/VM 6.4 but not z/VM 6.3.



Relocation Domains – User-Defined Domains

- Guests assigned to the User_64 domain will have access to the EDAT and TX facilities since both members in this domain are z/VM 6.4 and include these facilities.



Setting Up Relocation Domains

- *Relocation domains can be defined:*
  - In system configuration file:

```
RELOCATION_DOMAIN User_Mixed MEMBERS Mem64_A Mem63_B
RELOCATION_DOMAIN User_64    MEMBERS Mem64_A Mem64_C
```

  - Dynamically (by command):

```
DEFINE RELODOMAIN User_Mixed MEMBERS Mem64_A Mem63_B
DEFINE RELODOMAIN User_64    MEMBERS Mem64_A Mem64_C
```

Setting Up Relocation Domains …

- *Guests can be assigned to relocation domains:*
  - In their user directory entry (one per user):

    VMRELOCATE ON DOMAIN Mem64_A

    VMRELOCATE ON DOMAIN Mem63_B

    VMRELOCATE ON DOMAIN User_Mixed

    VMRELOCATE ON DOMAIN User_64

  - Dynamically (by command):

    SET VMRELOCATE *userid* DOMAIN Mem64_A

    SET VMRELOCATE *userid* DOMAIN Mem63_B

    SET VMRELOCATE *userid* DOMAIN User_Mixed

    SET VMRELOCATE *userid* DOMAIN User_64

CP Environment Variables

- z/VM 6.4 introduces a framework to handle meta data
  - Limit of 1000 variables
  - Variables starting with ‘CP.’ are reserved for IBM use

- System programmers with class B privilege can set variables
  - Additionally, one can be passed in via IPLPARMS on the SAPL screen
    - IPLVAR=variable on SAPL screen
    - CP.IPLPARMS.IPLVAR is the environment variable
  - Command or system configuration file statement:

    SET VARIABLE SYSTEM *name string*

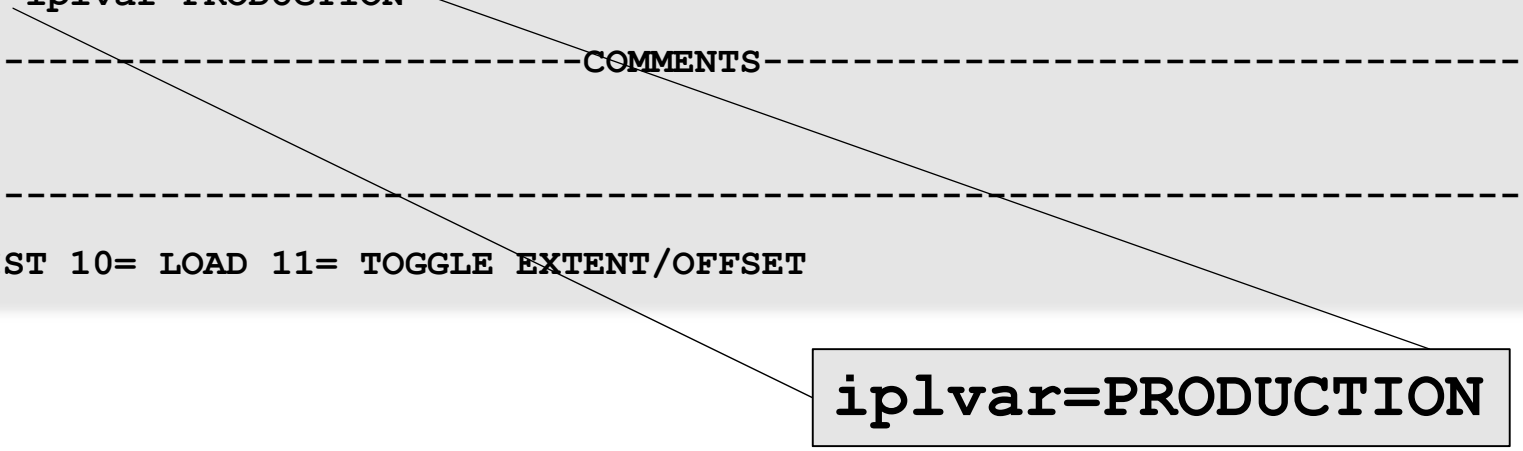
- Read the fields via query command from any class G virtual machine:

  QUERY VARIABLE ALL
  QUERY VARIABLE NAME *variable_name*

Setting the IPLVAR Environment Variable

```
STAND ALONE PROGRAM LOADER: z/VM VERSION 6 RELEASE 4.0
DEVICE NUMBER: 018B MINIDISK OFFSET: 35 EXTENT: -
MODULE NAME: CPLOAD LOAD ORIGIN: 2000
--------------------------------IPL PARAMETERS-------------------------------
cons=0080 iplvar=PRODUCTION
-----------------------------------COMMENTS-----------------------------------


------------------------------------------------------------------------------
9= FILELIST 10= LOAD 11= TOGGLE EXTENT/OFFSET
```




```
cp.iplparms.iplvar ← PRODUCTION
```

Query CP Service

- Very simple **QUERY CPSERVICE**
  - Options to limit output to local mods, PTFs, APARs
  - Option to ask for a particular update
  - Wildcards with '*'

- Shows service for the CPLOAD module that is currently running.
  - Does not show service for standalone utilities and other CP parts

- May want to use with CMS Pipelines if you use the default "ALL" option
  - PIPE CP QUERY SERVICE | > cpservice output a

- Checking for a specific APAR

```
QUERY CPSERVICE APAR VM65371

APAR      PTF
VM65371   UM34046
```

Orderly Shut Down of Guests

- Guests can be enabled to receive a signal to shut down
  - For Linux guests, put the following in the *etc/inittab* file:

```
# z/VM or LPAR is shutting down
ca:12345:ctrlaltdel:/sbin/shutdown -h now
```

(make sure you issue **–h** instead of **–r**)

- Specify time interval allowed for guests that receive the signal to shut themselves down
  - In your system configuration file:

```
Set ,
 Signal ShutdownTime 500,
 ShutdownTime 30 /* amount of time reserved for z/VM shutdown
```

  - Can also be set or changed with **SET SIGNAL** and **SET SHUTDOWNTIME** commands

- z/VM does not shut down until either:
  - All signaled guests indicate that they have shut down
  - The specified time interval expires

Shutdown Enhancements

- New **QUERY SHUTDOWN** command
  - Provides information about shutdown time and status of a pending shutdown
  - Class G guests and service virtual machines can obtain shutdown status information
  - Can help automate an orderly shutdown of the z/VM system and guests

```
query shutdown
System shutdown time: 30 seconds; previous shutdown duration: 9 seconds
SHUTDOWN initiated at 2017-02-27 14:58:33 by MAINT
Signaled users have 490 seconds left to shut down
```

- **SHUTDOWN** sends a message to the operator console when shutdown is started or cancelled

```
HCPSHU2116I SHUTDOWN issued at 2017-02-27 14:43:54 by MAINT
```

- **SIGNAL SHUTDOWN ALL** or **SIGNAL SHUTDOWN <userid>** sends a message to the operator console

```
HCPSIG2118I SIGNAL SHUTDOWN ALL issued at 2017-02-27 14:51:50 by MAINT
```

- **FORCE** sends a message to the operator console when the forced-off user is enabled for signals

```
HCP2118I Shutdown signal sent to USER1 because a FORCE was issued at
2017-02-27 15:05:40 by MAINT
```

IBM Tape Manager for z/VM

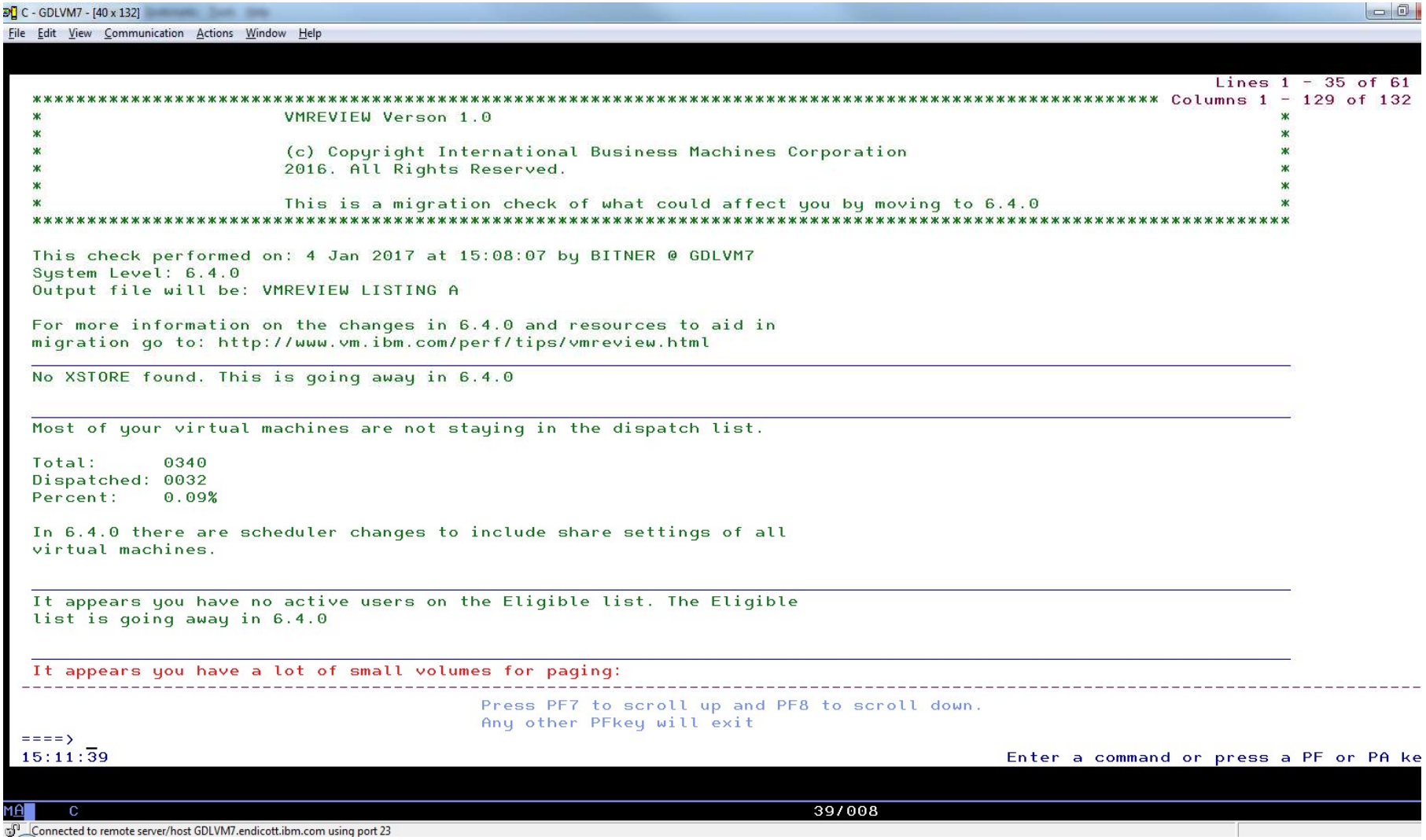
- Tape Manager for z/VM V1.3 supports z/VM 6.4

- In shared catalog environment that mixes z/VM 6.3 and z/VM 6.4
  - Communications error possible
  - Apply PTF UI45318 for Tape Manager V1.3
    - APAR PI77465 fixes the problem with Pipelines stage conflicts

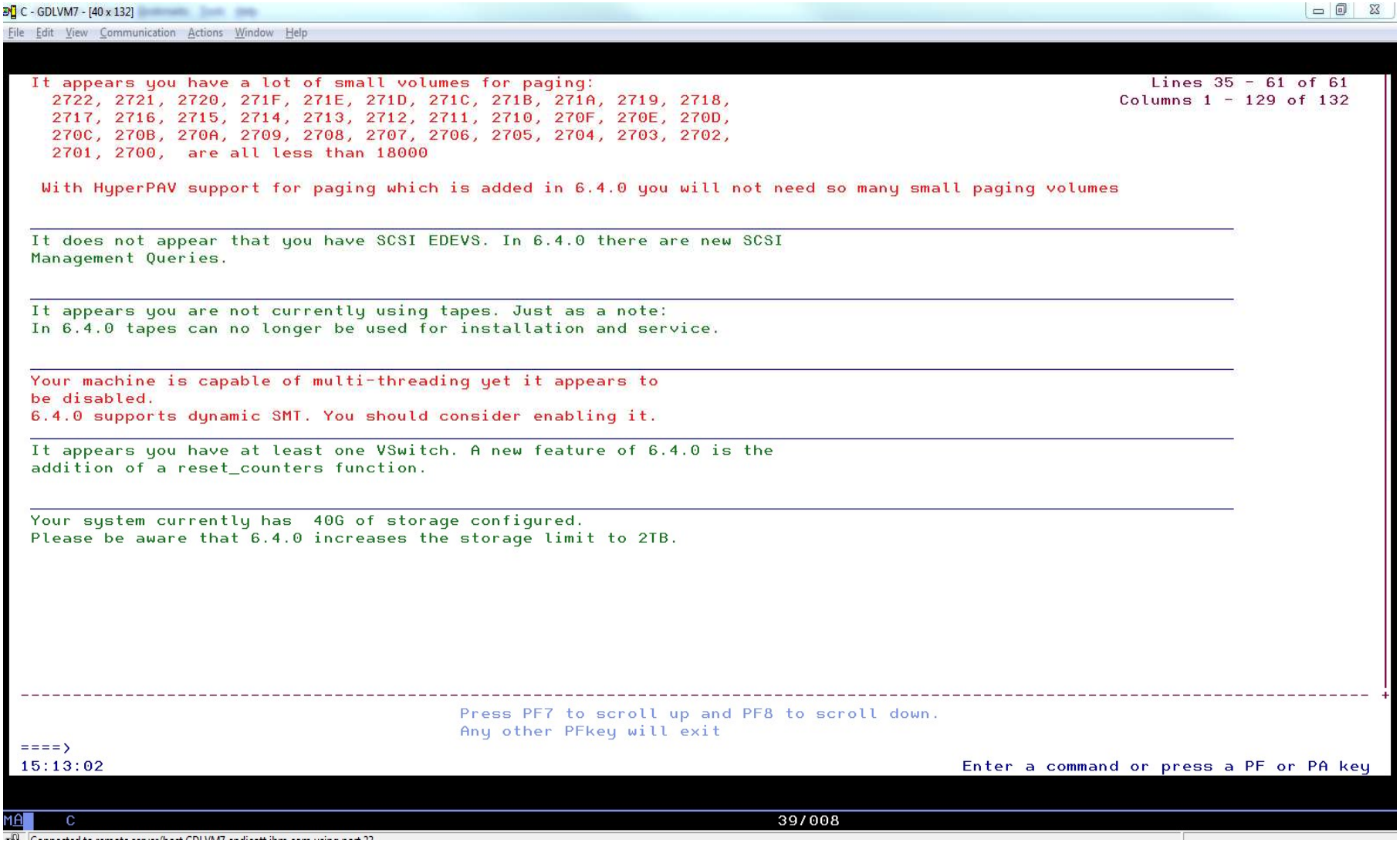
How do you know what to expect in z/VM 6.4?

- New VMREVIEW utility on z/VM download page
  - Run on existing z/VM 5.4, 6.1, 6.2, or 6.3 systems
  - Will highlight:
    - Things that should be changed prior to going to z/VM 6.4
    - Value that could be gained by going to z/VM 6.4
    - Other interesting things in regard to this environment being on z/VM 6.4
  - Envision this being a work in progress
    - Interested in feedback for other things it should do
- Started as an extra project by some of the newer members of the z/VM team
- http://www.vm.ibm.com/perf/tips/vmreview.html

VMREVIEW Output



VMREVIEW highlights considerations



Summary

Summary - Checklist

- Before you go to z/VM 6.4
  - ❑ Check service for z/VM Upgrade-in-Place if you plan to use it
  - ❑ Check for formation of eligible list
  - ❑ If planning to use additional memory, plan for additional dump and paging space
  - ❑ Acquire a z196, z114 or newer machine
  - ❑ Check for queues on paging devices
  - ❑ Download and run VMREVIEW utility
  - ❑ Validate RACF DB before and after uplevel
  - ❑ Collect Monwrite performance data and application performance data
- When you bring up z/VM 6.4
  - ❑ Configure expanded storage as central storage
  - ❑ To prepare for Dynamic SMT, enable multithreading with 1 thread per core
  - ❑ Check relocation domain considerations
  - ❑ Collect Monwrite performance data and application performance data
- To exploit capabilities with z/VM 6.4
  - ❑ Ensure guest configured to use large page as appropriate
  - ❑ If memory-rich, consider using KEEPSLOT
  - ❑ Enable HyperPAV for paging if appropriate
  - ❑ Enable zHPF for paging
  - ❑ Investigate uses for environment variables
  - ❑ Collect Monwrite performance data and application performance data