

# z/VM 6.4: Preparation & Use

June 25, 2017 Version 13

Bill Bitner  
z/VM Dev Lab Client Focus & Care  
[bitnerb@us.ibm.com](mailto:bitnerb@us.ibm.com)



# Trademarks

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

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

\* 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](http://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](http://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.

## Abstract

z/VM 6.4 was generally available on November 11, 2016 and brought a lot of new capability to the z/VM environment. This presentation reviews some of the key changes in the release. First we'll look at your existing system and discuss what will be different when you go to z/VM 6.4. Next we'll look at what you should change prior to IPLing z/VM 6.4, things like expanded storage or changes for HyperPAV paging. We'll wrap things up with a discussion of how you exploit some of the new features after you are running z/VM 6.4.

# Agenda

- z/VM 6.4 Release Notes
  
- Things you need to look at or do while on older system for a smooth transition and to avoid problems
  
- Things to do as you bring up z/VM 6.4
  
- Things you need to look at to get the most out of z/VM 6.4 after it is up and running

# 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
6.4	11/2016			IBM System z196 & z114®	-	Statement of Direction
6.3	7/2013	12/2017 <sup>[1]</sup>	11/2016	IBM System z10®	-	EAL 4+ OSPP-LS
6.2	12/2011	6/2017 <sup>[2]</sup>	7/2013	IBM System z10®	z13	-
5.4	9/2008	12/2017 <sup>[3]</sup>	3/2012	IBM eServer zSeries 800& 900	zEC12	-

<sup>[1]</sup> Announced February 3, 2015

<sup>[2]</sup> Announced February 2, 2016

<sup>[3]</sup> Announced August 2, 2016

## z/VM 6.4

- General Availability November 11, 2016
- A release born from customer feedback
- Key components:
  - Enhanced technology for improved scaling and total cost of ownership
  - Increased system programmer and management capabilities
- New Architecture Level Set (ALS) of z196 and higher





## z/VM 6.4 Supported Hardware

- Following z Systems servers:
  - z13
  - z13s
  - LinuxONE Emperor
  - LinuxONE Rockhopper
  - IBM zEnterprise EC12
  - IBM zEnterprise BC12
  - IBM zEnterprise 196
  - IBM zEnterprise 114
  
- Electronic and DVD install
  - No tapes

## Upgrade In Place

- Enables a smoother upgrade of existing z/VM 6.2 and z/VM 6.3 systems to z/VM 6.4, especially in a Single System Image (SSI) environment, and 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 pre-requisites
- Unlike z/VM 6.3, z/VM 6.4 requires TCP/IP machine to be shutdown at one point, so will need alternate method to get to z/VM
- See Live Virtual Class for session on Upgrade in Place May 31<sup>st</sup>, 2017
  - <http://www.vm.ibm.com/education/lvc/>

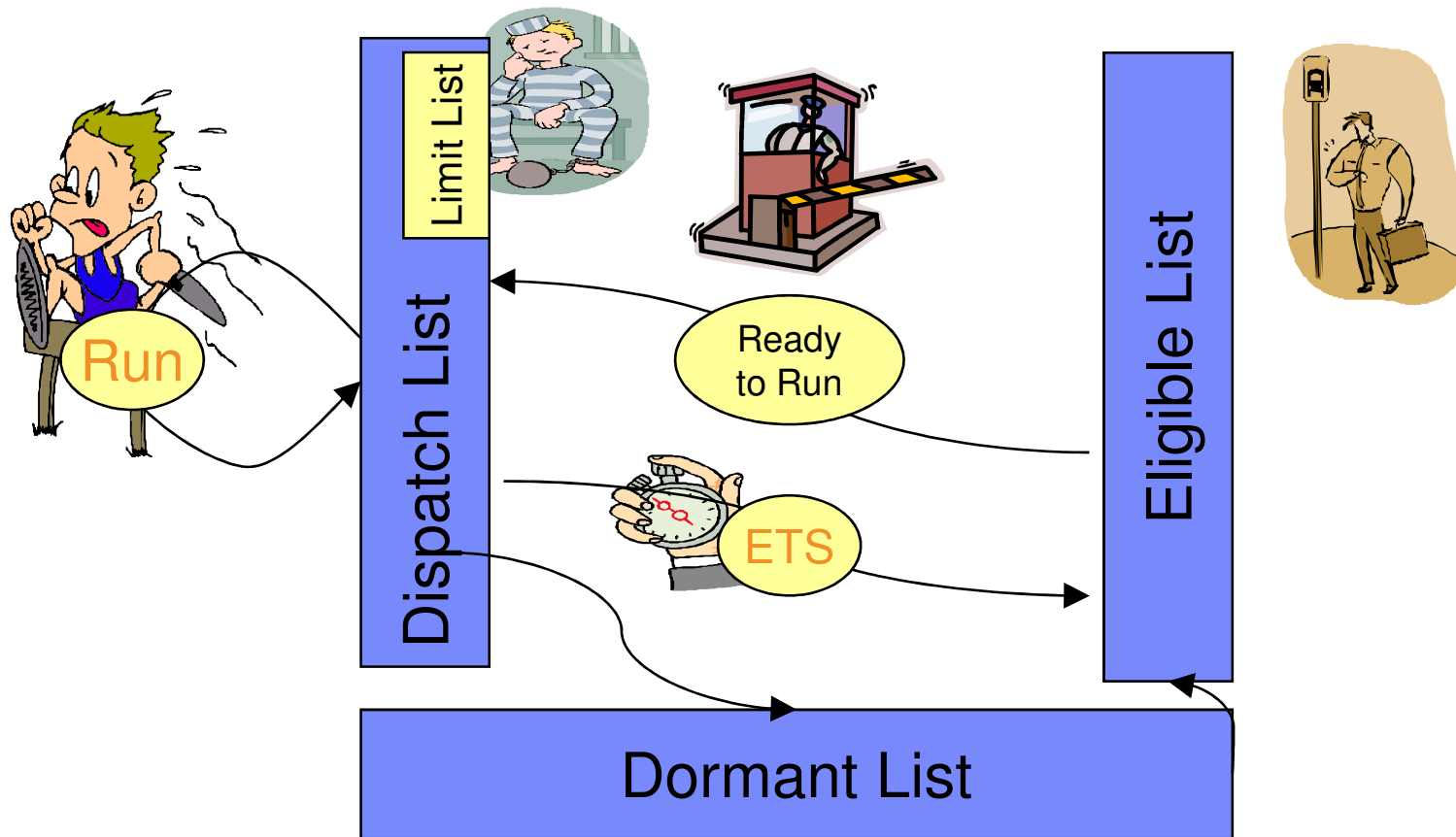
## Fresh Install Considerations

- 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)
- Installing to almost full pack minidisks (n-1 cylinders) not possible
  - 3390 mod-3 minimum install size changed to 3339 cylinders from 3338
  - 3390 mod-9 minimum install size changed to 10017 cylinders from 10016
    - IBM may be able to change mod-9 back to 10016 at a later time.

## 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) when bringing up z/VM 6.4 or 6.3
- The memory management changes made in z/VM 6.3 made expanded storage obsolete

# 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 factored into the normalization, not just virtual machines in the dispatch and eligible lists.

# Surplus Share Distribution: Background

- Shares are relative to other users 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	z/VM 6.4	z/VM Prior
			Correct Distribution	Problem Scenario
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%)

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

- In z/VM 6.4 the sums include *all* users, not just those in the dispatch and eligible lists
  - Watch for systems where:  
$$\sum \text{relative\_share} \gg \sum \text{relative\_share\_disp\_list}$$
- In all releases, 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
```

## 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

## 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
  - **ADD**: adds FCP subchannel and WWPN 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 now 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



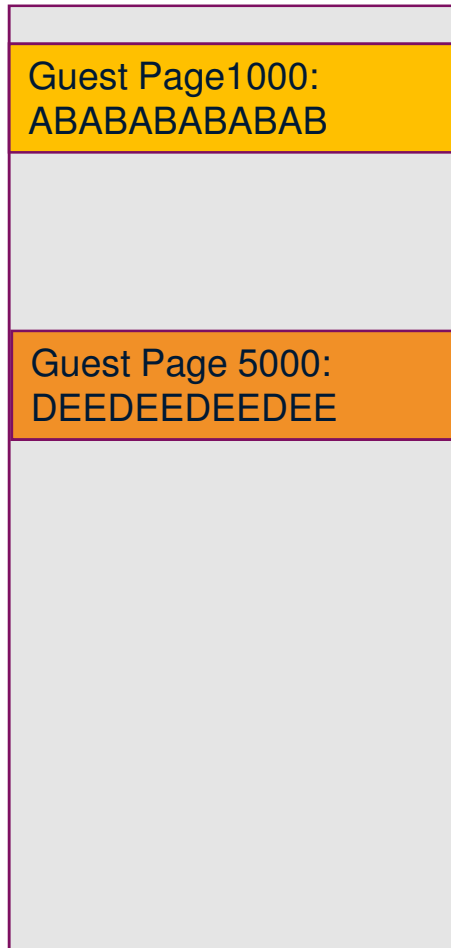
## 2 TB Real Memory Support

- 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

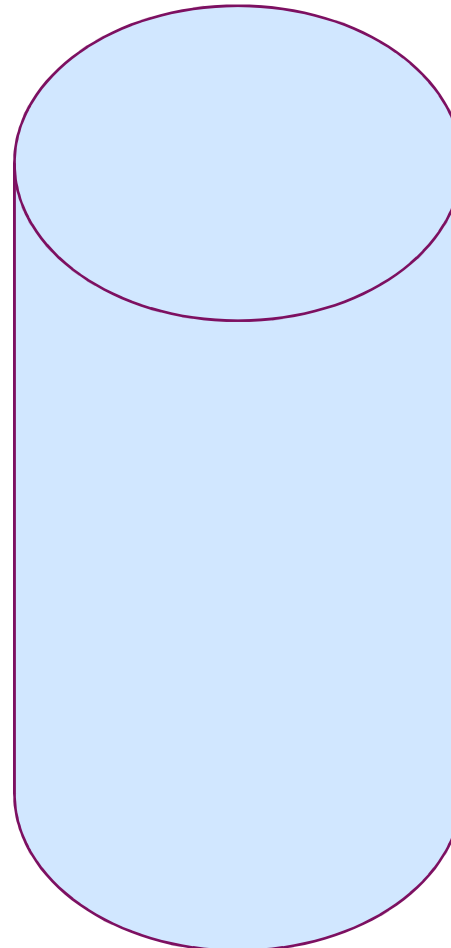
## Guest Large Page

- 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



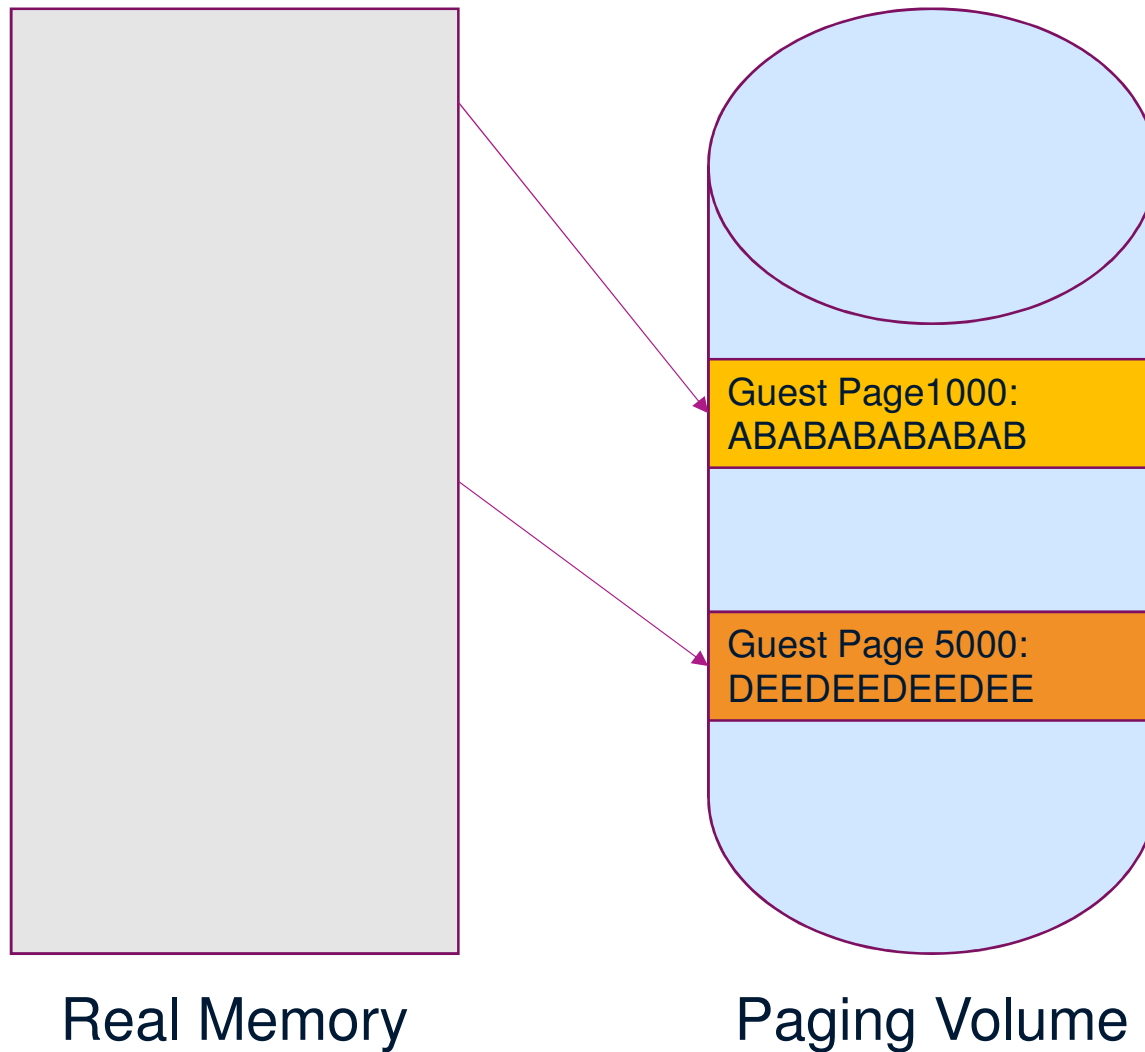
Real Memory



Paging Volume

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

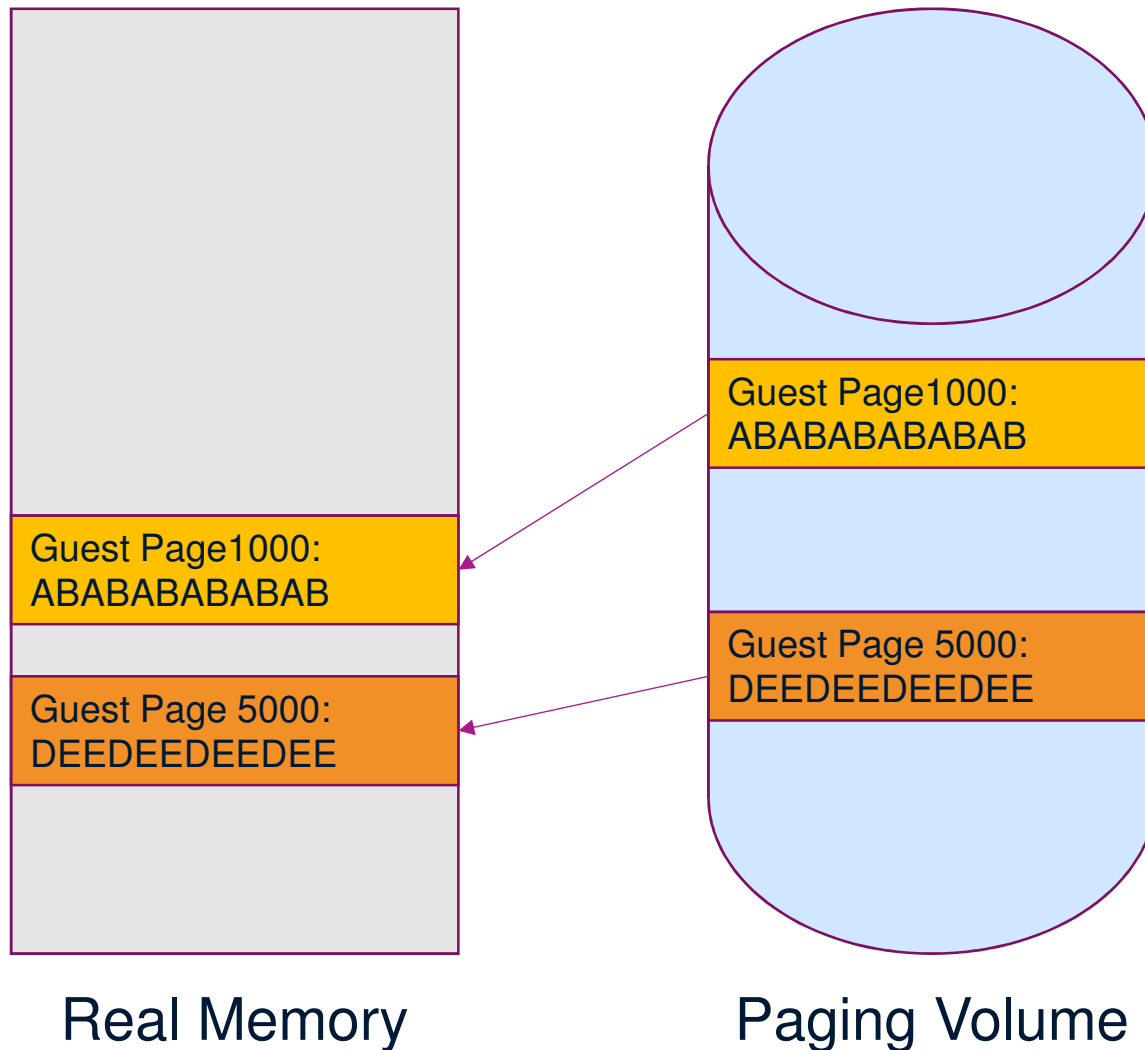
## KEEPSLOT - Background



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).

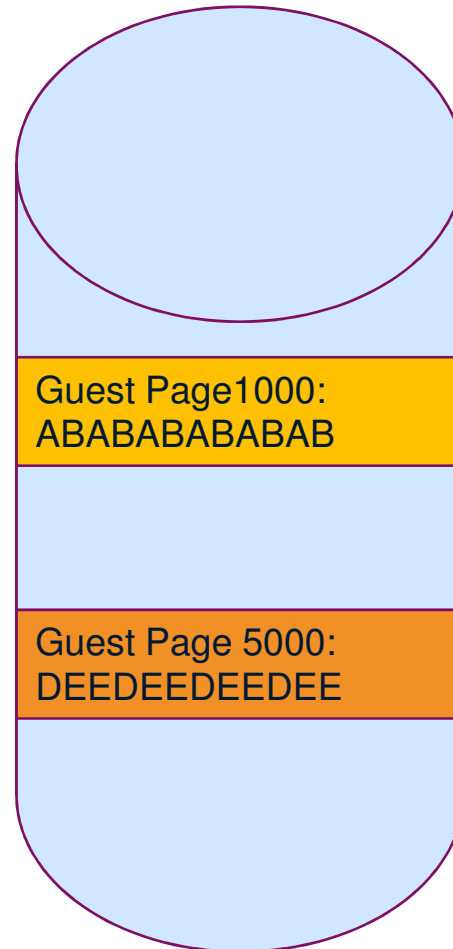
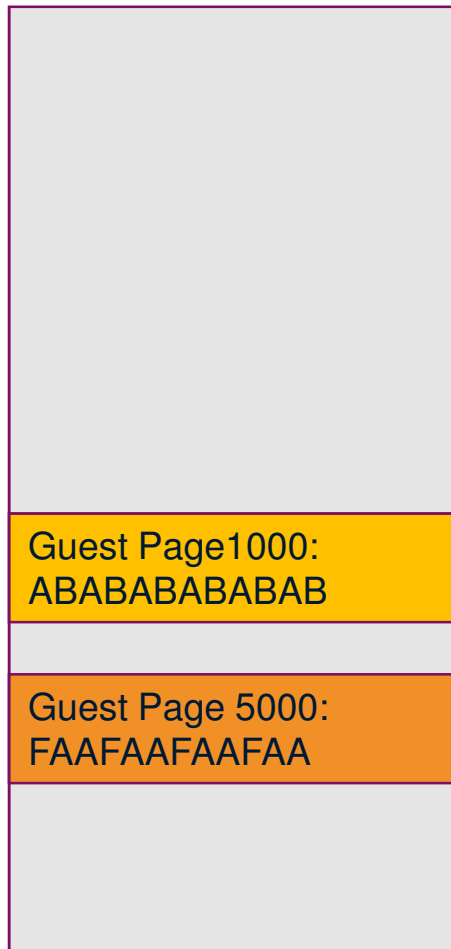
## KEEPSLOT - Background



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.

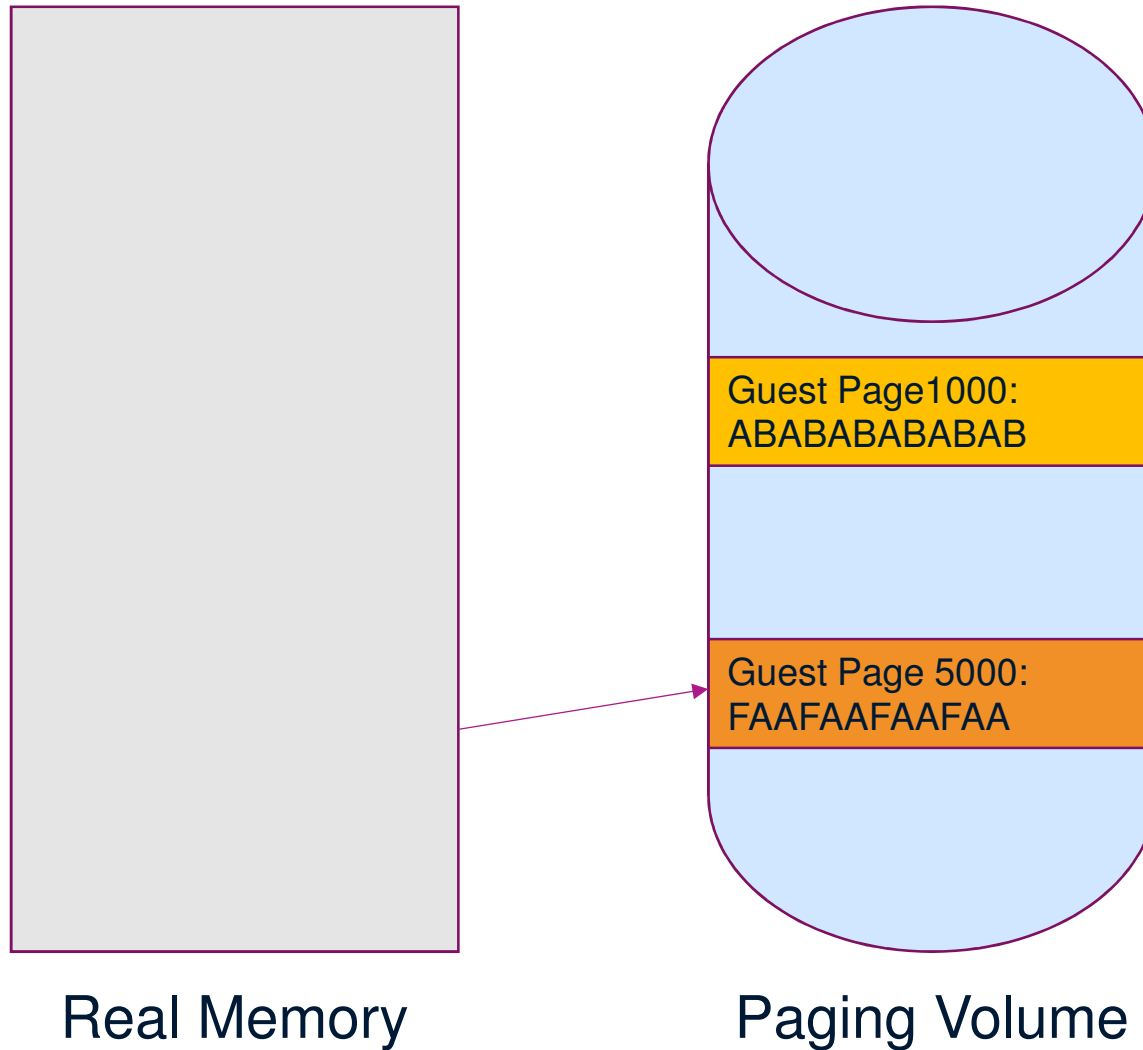
# KEEPSLOT - Background



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 change.

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 *nnnnn* PAGING\_SHARE *nnnnn***
  - Command:  
**SET CU ALIAS MDISK\_SHARE *nnnnn* PAGING\_SHARE *nnnnn* *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\_HP**

# 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 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 2<sup>nd</sup> 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

- Live Guest Relocation (LGR) supports relocation domains
  - Which allows for relocation across SSI cluster members which do not have identical configurations/capabilities.
  
- z/VM 6.4 introduces two additional scenarios where architectures, from the guest perspective, may appear incompatible when the cluster includes z/VM 6.4 and 6.3 systems:
  - Enhanced DAT (large page)
  - Transactional Execution Facility
  
- If guests are in relocation domains that span members with z/VM 6.4 and older z/VM releases, realize the guests will not see the new capabilities.
  - **SET VMRELOCATE USER** *userid* **DOMAIN** *ssi\_member\_name*

# 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
```



```
iplvar=PRODUCTION
```

```
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**

<b>APAR</b>	<b>PTF</b>
<b>VM65371</b>	<b>UM34046</b>

# 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

```

C - GDLMV7 - [40 x 132]
File Edit View Communication Actions Window Help
Lines 1 - 35 of 61
Columns 1 - 129 of 132
*****
*                VMREVIEW Verson 1.0                *
*                                                    *
*                (c) Copyright International Business Machines Corporation *
*                2016. All Rights Reserved.          *
*                                                    *
*                This is a migration check of what could affect you by moving to 6.4.0 *
*****

This check performed on: 4 Jan 2017 at 15:08:07 by BITNER @ GDLMV7
System Level: 6.4.0
Output file will be: VMREVIEW LISTING A

For more information on the changes in 6.4.0 and resources to aid in
migration go to: http://www.vm.ibm.com/perf/tips/vmreview.html

-----
No XSTORE found. This is going away in 6.4.0

-----
Most of your virtual machines are not staying in the dispatch list.

Total:          0340
Dispatched:    0032
Percent:       0.09%

In 6.4.0 there are scheduler changes to include share settings of all
virtual machines.

-----
It appears you have no active users on the Eligible list. The Eligible
list is going away in 6.4.0

-----
It appears you have a lot of small volumes for paging:

-----
Press PF7 to scroll up and PF8 to scroll down.
Any other PFkey will exit

====>
15:11:39                                     Enter a command or press a PF or PA ke
MÁ C 39/008
Connected to remote server/host GDLMV7.endicott.ibm.com using port 23

```

# VMREVIEW highlights considerations

```
C - GDLMV7 - [40 x 132]
File Edit View Communication Actions Window Help

It appears you have a lot of small volumes for paging:
2722, 2721, 2720, 271F, 271E, 271D, 271C, 271B, 271A, 2719, 2718,
2717, 2716, 2715, 2714, 2713, 2712, 2711, 2710, 270F, 270E, 270D,
270C, 270B, 270A, 2709, 2708, 2707, 2706, 2705, 2704, 2703, 2702,
2701, 2700, are all less than 18000
Lines 35 - 61 of 61
Columns 1 - 129 of 132

With HyperPAV support for paging which is added in 6.4.0 you will not need so many small paging volumes

It does not appear that you have SCSI EDEVS. In 6.4.0 there are new SCSI
Management Queries.

It appears you are not currently using tapes. Just as a note:
In 6.4.0 tapes can no longer be used for installation and service.

Your machine is capable of multi-threading yet it appears to
be disabled.
6.4.0 supports dynamic SMT. You should consider enabling it.

It appears you have at least one VSwitch. A new feature of 6.4.0 is the
addition of a reset_counters function.

Your system currently has 40G of storage configured.
Please be aware that 6.4.0 increases the storage limit to 2TB.

-----
Press PF7 to scroll up and PF8 to scroll down.
Any other PFkey will exit

====>
15:13:02 Enter a command or press a PF or PA key

MÁ C 39/008
```

# 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
  - Collect Monwrite 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
- 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