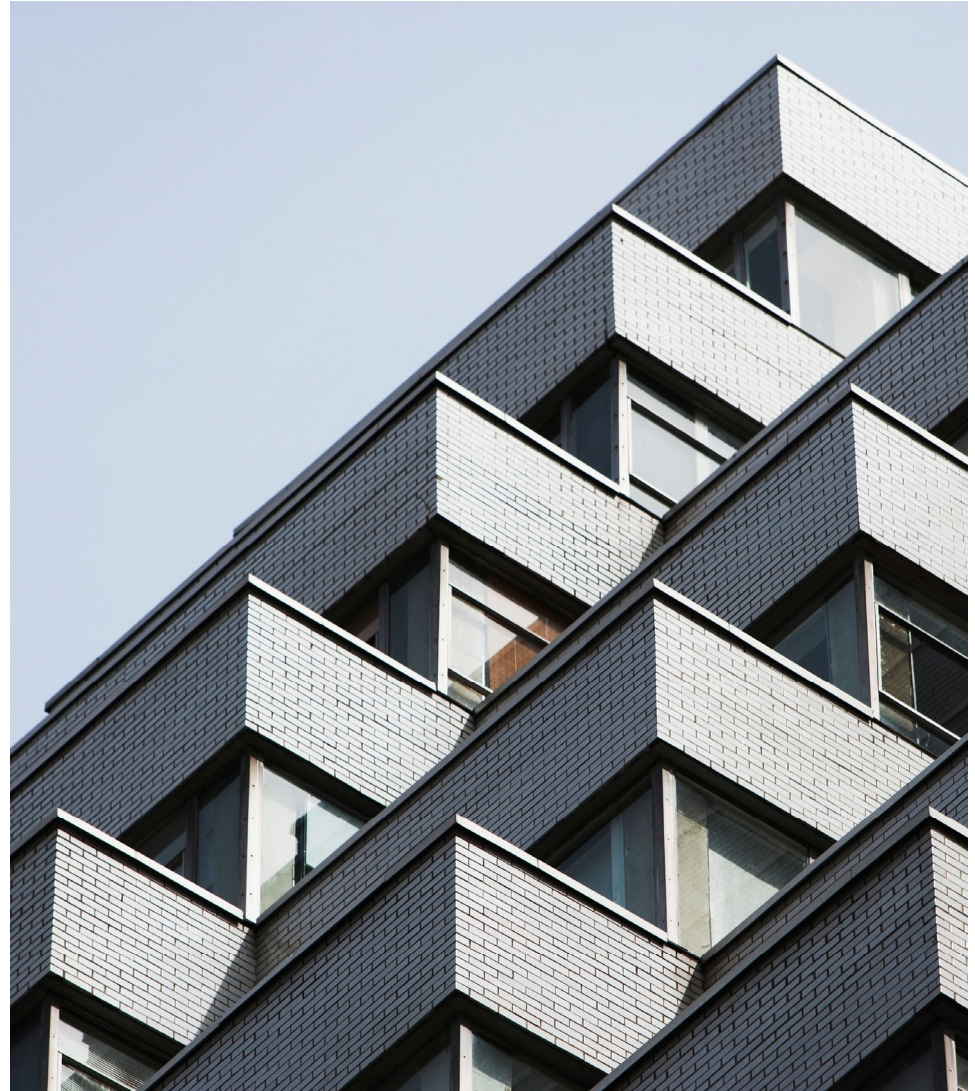


# Common z/VM Hurdles

And How to Overcome Them

Dan Shekhtman  
z/VM CP Development & Level 2 Support  
MVMUA, January 2024

[daniel.shekhtman@ibm.com](mailto:daniel.shekhtman@ibm.com)





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

Db2*	FlashCopy*	IBM eserver	OMEGAMON*	XIV*	z10 BC	zSecure
DirMaint	FlashSystem	IBM (logo)*	PR/SM	z13*	z10EC	zSeries*
DS8000*	GDPS*	IBM Z*	RACF*	z13s	z/Architecture*	z/VM*
ECKD	ibm.com	LinuxONE*	System z10*	z14	zEnterprise*	z Systems*
FICON*	IBM Cloud*	LinuxONE Emperor	System 390*	z15	zPDT	
		LinuxONE Rockhopper	WebSphere*		z/OS*	

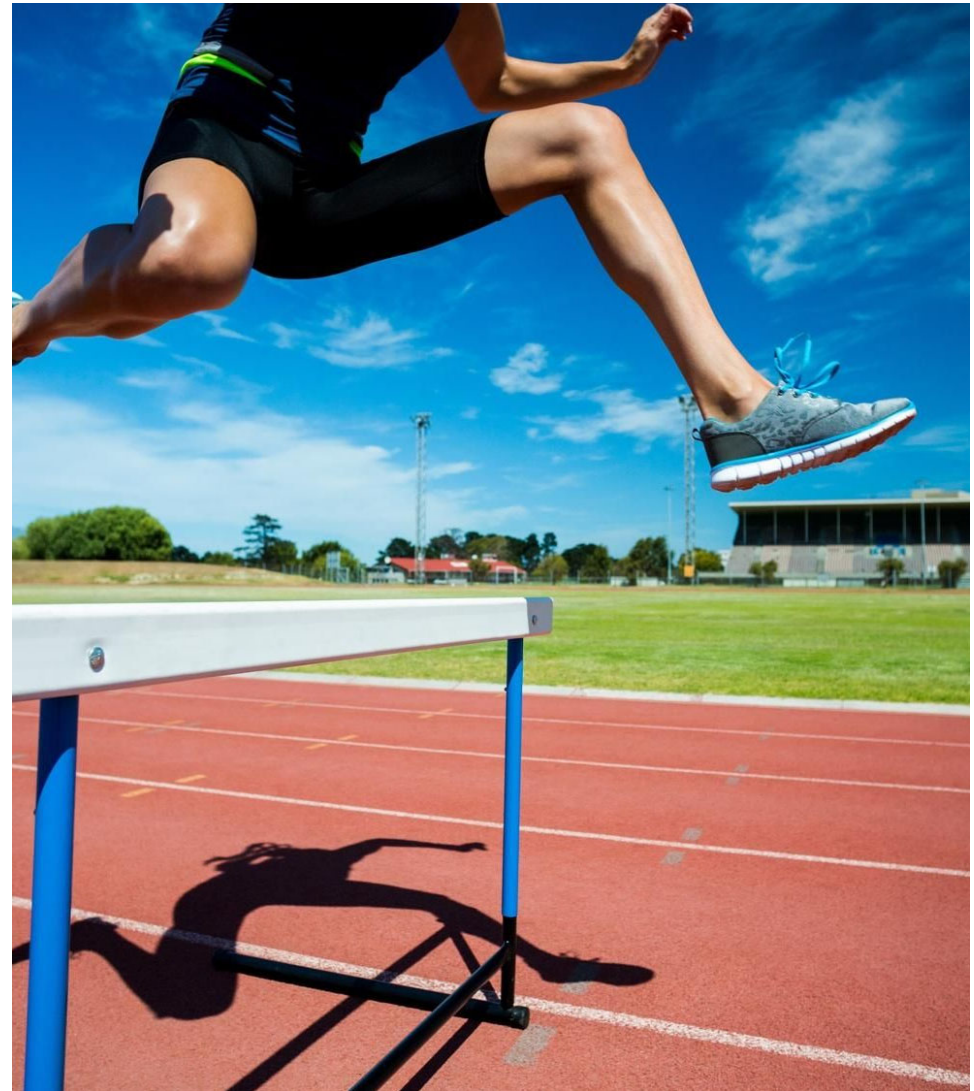
\* Registered trademarks of IBM Corporation

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.  
IT Infrastructure Library is a Registered Trade Mark of AXELOS Limited.  
ITIL is a Registered Trade Mark of AXELOS Limited.  
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 other countries.  
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.  
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.  
Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.  
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.  
UNIX is a registered trademark of The Open Group in the United States and other countries.  
VMware, the VMware logo, VMware Cloud Foundation, VMware Cloud Foundation Service, VMware vCenter Server, and VMware vSphere are registered trademarks or trademarks of VMware, Inc. or its subsidiaries in the United States and/or other jurisdictions.  
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.

# Agenda

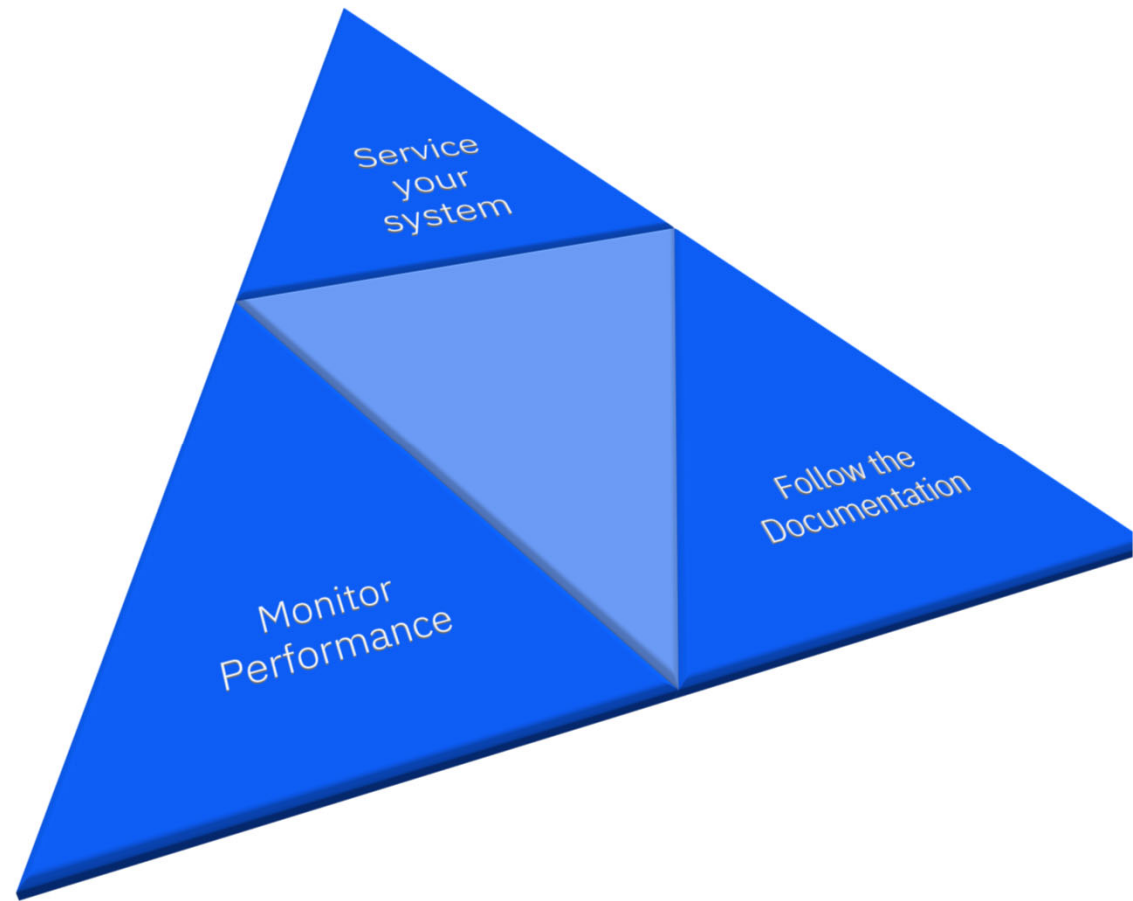
1. Orientation
  - What defines a “hurdle”
  - When are you most likely to hit a z/VM “hurdle”
  - What areas to watch out for
2. General Mitigation and Avoidance Tactics
3. Regularly Hit Hurdles
4. Tips and Tricks
5. Parting Thoughts



# The 3 Best Ways to Avoid z/VM Problems

1. Stay current on z/VM Service
2. Have a strategy for monitoring performance
3. Follow the Documentation

**Bonus Tip:** Automate where you can, document when you can't



O 8/ 4 :\$ :/54

How do we  
define a  
Hurdle?

A Hurdle is anything  
that causes you to:



Open A  
Service Ticket



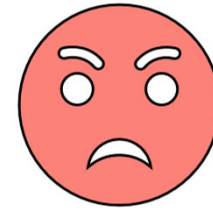
Ask Someone  
For Help



Be  
Inconvenienced  
by an Otherwise  
Avoidable  
Problem

# How does a hurdle manifest?

1. You get stalled during service/install/upgrade procedure
2. The z/VM Hypervisor abends
3. The z/VM Hypervisor hangs
4. You or your clients experience responsiveness and/or performance issues



– You



– Your Boss



– Your Clients



– z/VM Support



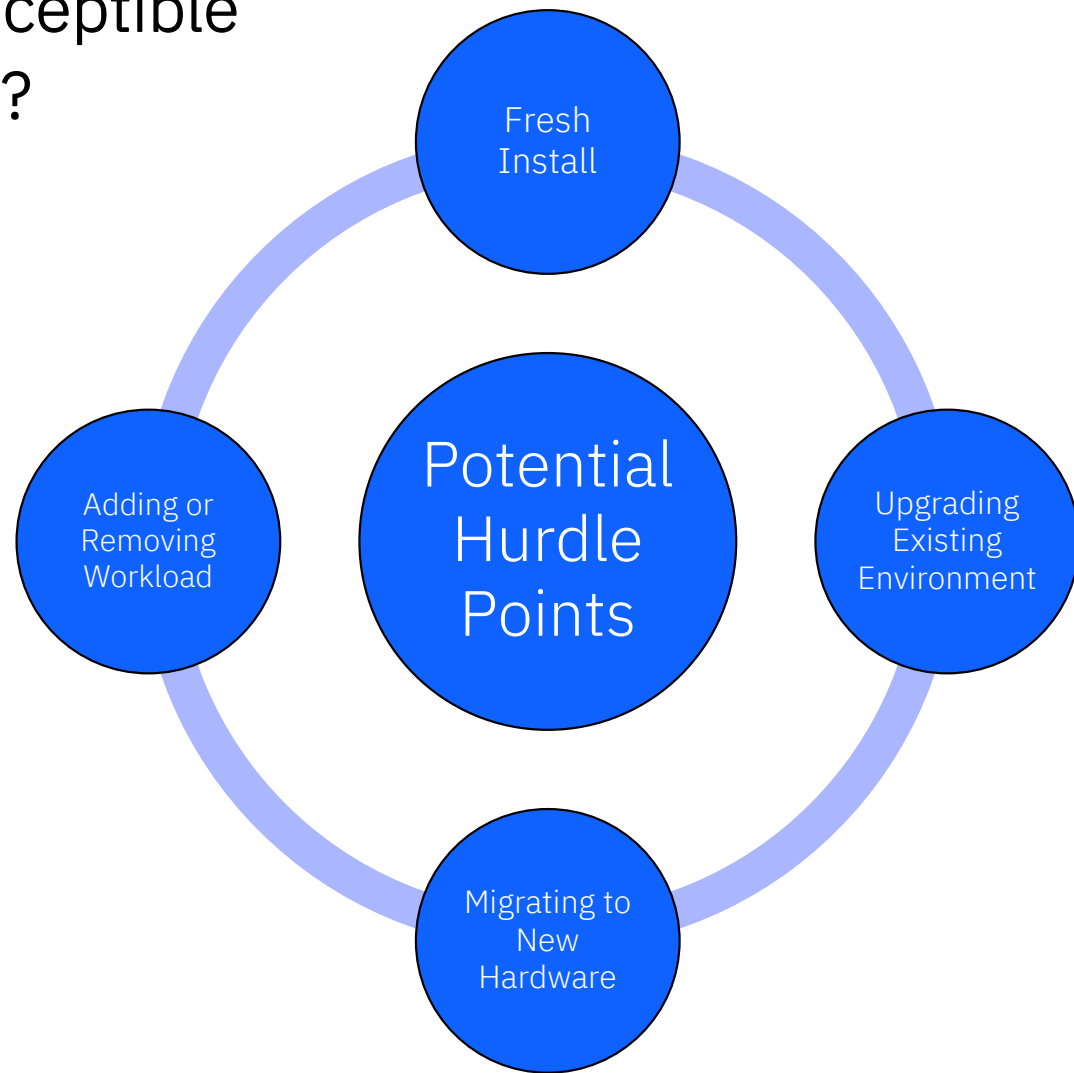
# Why do we hit these stumbling blocks?

1. "With great power comes great responsibility"  
- Uncle Ben
2. Care and feeding of your VM system was not in the original job description
3. When it's new, it can be confusing
4. Expectations != Reality

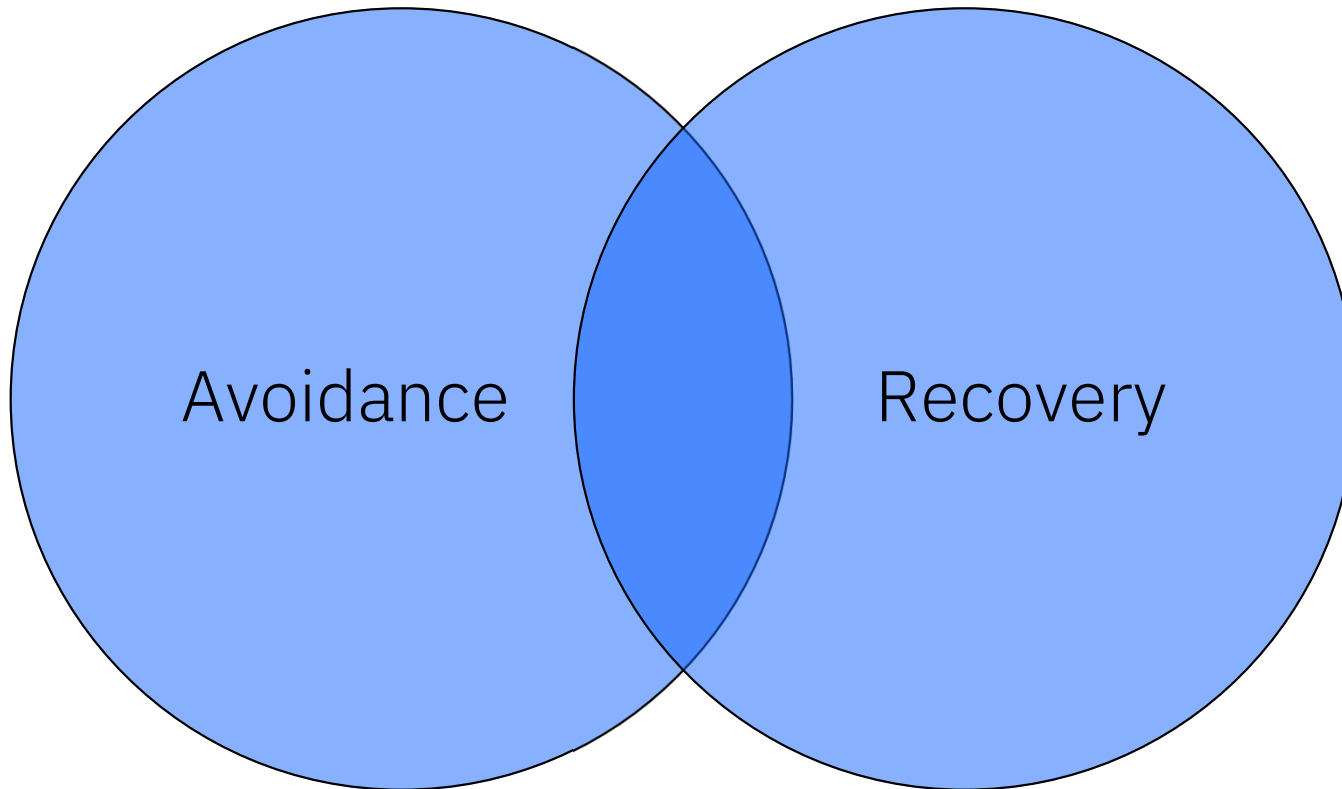




When are you most susceptible to hitting a z/VM hurdle?



Why can we do?



A<5 / ( / 4 + : . )

P 85 & 2) 3

# Avoiding The Problem: Staying Current on Service

**Scenario:** Your system has been impacted by an issue and the IBM Support personnel identify the root cause as service that is available but not applied to your system.

## Hints & Tips:

- Have a plan to stay current on the RSU level:  
<https://www.vm.ibm.com/service/rsu/rsuplan.HTML>
- Don't forget about the PSP bucket!
- If managing eight or fewer z/VM partitions, consider SSI to reduce service impact.
- If managing more than eight<sup>1</sup> z/VM partitions, you could consider Centralized Service Management (CSM).  
(<https://www.vm.ibm.com/service/csm serv.html>, Chapters 8-11 in z/VM: Service Guide)

<sup>1</sup> z/VM CSM supports up to 54 systems.



# Avoiding The Problem:

## Have a strategy for monitoring performance

**Scenario:** You found out there was a performance problem because an SLA was missed, or a client complained. In hindsight there were warning signs of a sick system.

### Hints & Tips:

- Plan to monitor data continuously, not just when a problem occurs
- Plan to monitor data before and after any system changes
- Monitor ALL parts of the stack that matter (Linux data, DB Perf, Monwrite, etc.)
- Collect all pertinent logs
- Have a well-defined measure of “success”
- Take advantage of automation solutions (eg, Infrastructure Suite for z/VM and Linux)
- Have a test partition for validating changes
- If you have questions....ask
- <https://www.vm.ibm.com/perf/tips/>



# Avoiding The Problem: Follow the Documentation

**Scenario:** You are performing a z/VM service, upgrade, or maintenance task and run into an issue (usually an error message)

## Hints & Tips:

- Follow the documentation corresponding to the release level you're working with
- Complete all worksheets
- Don't skip steps
- If you aren't sure about something.....ask



A C522) ' :/54 5\*  
C53 3 54 I99; ) 9



# Hurdle: PGT004 Abends (Allocating enough Paging space)

**Scenario A:** Your SSI-enabled z/VM system has been running as usual, but you decide to LGR more workload onto it in anticipation of a scheduled outage on another member of the cluster.. A few hours later during some regular processing, your system takes a hard PGT004 abend.

**Scenario B:** Your newly configured z/VM system is running well, and all seems dandy. When your workload peaks, you experience a hard PGT004 abend.

## How to Overcome:

- Periodically issue QUERY ALLOC PAGE to monitor paging space utilization and consider adding more paging space once your usage approaches 75% of the available space.
- Consider the logon and maximum allowed storage sizes for each virtual machine.
- <https://www.vm.ibm.com/perf/tips/prgpage.html>
- Consider System Growth:
  - Any changes to the CP Directory to define more users, to increase virtual storage sizes for existing users, and to increase the number of allowable VDISKS must be monitored to ensure your existing paging capacity can handle the growth.
  - When operating within an SSI cluster, you must take into consideration the possibility for guests to relocate from other members of the cluster. Consideration should be given to the size of the guests to be relocated, including the size of VDISKS allocated to those guests.

# Hurdle: LGR Architecture incompatibility (SSI Only)

**Scenario:** You upgraded to a newer CPU and/or version of z/VM. You were planning on taking advantage of newer architectural features. After bringing up your guests, you realize these features are still disabled.

## Background:

- The default relocation domain for a single configuration virtual machine (USER) is the cluster wide relocation domain SSI.
- Each domain has an Architecture Description Block (ARDBK), which describes the supported architecture level of that domain.
- The Virtual Architecture Level (VAL) is calculated to be the lowest common denominator in a given domain, in terms of both hardware and software level. This enforces LGR compatibility among all guests in a domain.
- In a z/Linux VM guest, `/proc/cpuinfo` will report information on the virtual CPU, while `/proc/sysinfo` will report information on the real CPU. If there is a VAL discrepancy, it will show here.

## How to overcome:

- Double check your SSI relocation domains. (QUERY VMRELOCATE and QUERY RELODOMAIN)
- If you do not intend to ever relocate a particular guest, place it into the singleton domain. (VMRELOCATE OFF in its directory entry)
- When decommissioning a member of an SSI cluster, don't forget to make the slot available with the SET SSI command. This clears the ARDBK for that domain and forces a recalculation of the VAL.
- Use RELODOM utility from VM Download Packages. It will help you (or us) figure out what Virtual Architecture Level (VAL) CP is setting for your relocation domain(s).
- Upgrade the back-level members of your relocation domain.

# Hurdle: Right-Sizing of z/VM Partitions or Guests

**Scenario:** You've moved workload to z/VM, but the guest or partition is either under-sized or over-sized in its processor allocation

## How to Overcome:

- If FCX304 PRCLOG is well below capacity, you can shed logical cores.
- If FCX304 PRCLOG is bumping up against all CPUs 100% busy, you need to add logical cores.
- If FCX112 USER %CPU is well below 100 times virtual processors, you can shed virtual processors.
- If FCX112 USER %CPU is bumping into 100 times virtual processors, you need to add virtual processors.

# Hurdle: Right-Sizing of z/VM Partitions

**Scenario:** If FCX304 PRCLOG is well below capacity, you can shed logical cores.

**Hints & Tips:**

- Choose data that matches your heaviest load.
- Look for peaks. Core utilization will vary over time.

From Perfkit FCX304 PRCLOG  
(fabricated example to illustrate concept)

Interval						Pct		Total
END Time	CPU	TYPE	PPD	Ent.	DVID	Park Time	%Susp	
08:25:41	00	IFL	VH	100	0000	0	.1	99.2
08:25:41	01	IFL	VH	100	0001	0	.1	85.4
08:25:41	02	IFL	VH	100	0002	0	.1	53.8
08:25:41	03	IFL	VH	100	0003	0	.1	29.1
08:25:41	04	IFL	VH	100	0004	0	.1	11.2
08:25:41	05	IFL	VH	100	0005	0	.1	.6
08:25:41	06	IFL	VH	100	0006	0	.1	.0
08:25:41	07	IFL	VH	100	0007	0	.1	.0
08:25:41	08	IFL	VH	100	0008	0	.1	.5
08:25:41	09	IFL	VH	100	0009	0	.1	.9
08:25:41	0A	IFL	VH	100	000A	0	.1	.0

# Hurdle: Right-Sizing of z/VM Partitions

**Scenario:** If FCX304 PRCLOG is bumping up against all CPUs 100% busy, you need to add logical cores.

**Hints & Tips:**

- Choose data that matches your heaviest load.
- Look for peaks. Core utilization will vary over time.

From Perfkit FCX304 PRCLOG  
(fabricated example to illustrate concept)

Interval							Pct Park		
END Time	CPU	TYPE	PPD	Ent.	DVID	Time	%Susp	Total	
08:25:41	00	IFL	VH	100	0000	0	.1	99.2	
08:25:41	01	IFL	VH	100	0001	0	.1	100.0	
08:25:41	02	IFL	VH	100	0002	0	.1	98.0	
08:25:41	03	IFL	VH	100	0003	0	.1	98.4	
08:25:41	04	IFL	VH	100	0004	0	.1	99.2	
08:25:41	05	IFL	VH	100	0005	0	.1	99.6	
08:25:41	06	IFL	VH	100	0006	0	.1	100.0	
08:25:41	07	IFL	VH	100	0007	0	.1	100.0	
08:25:41	08	IFL	VH	100	0008	0	.1	97.5	
08:25:41	09	IFL	VH	100	0009	0	.1	98.9	
08:25:41	0A	IFL	VH	100	000A	0	.1	99.0	

# Hurdle: Right-Sizing of z/VM Guests

Scenario 1: If FCX112 USER %CPU is well below 100 times virtual processors, you can shed virtual processors.  
Scenario 2: If FCX112 USER %CPU is bumping into 100 times virtual processors, you need to add virtual processors.

From Perfkit FCX112 USER  
(fabricated example to illustrate concept)

<----- CPU Load ----->				
<-Seconds-> T/V				
Userid	%CPU	TCPU	VCPU	Ratio
. . .				

User Data:				
PAPABEAR	4.21	2.560	1.520	1.68
MAMABEAR	2.88	1.730	1.124	1.54
BABYBEAR	6.18	6.180	4.760	1.30

## vCPU Allocation

PAPABEAR : 10 vCPUs ← too many  
MAMABEAR : 3 vCPUs ← too few  
BABYBEAR : 8 vCPUs ← just right

## Hints & Tips:

- Choose data that matches your heaviest load.
- Look for peaks. vCPU utilization will vary over time.

\*CAUTION: If you have RELATIVE SHARES setup, they may need to be changed when you add or remove vCPUs to/from a guest.

# Hurdle: Failure to adjust SHARE Settings

**Scenario:** You add vCPUs to a z/VM guest, but performance of that guest degrades instead of improves..

- Check SHARE Settings
- Default SHARE settings for all virtual machines is “Relative 100”
- Recommendation for a starting point is to set SHARE RELATIVE (100 \* number of vCPUS defined). Creates a level playing field
- Changing vCPU allocation will introduce parity in priority under heavy load

GUEST	vCPUs	SHARE
AARON	6	6000
BOB	8	8000
CHARLIE	10	10000
DIANE	<del>6</del> 8	6000



# Hurdle: Improper LPAR Weights create unusable entitlement

**Scenario:** The LPAR weights on a z/VM partition are set too high, creating entitlement that cannot be used. This deprives other LPARs of entitled power.

## How to Overcome:

- decrease weights of LPARs with unusable entitlement
- Make sum of weights =  $10 * (\text{\# of shared physical cores})$
- This makes entitlement =  $\text{weight}/10$
- Handy Entitlement Calculator:  
<https://www.vm.ibm.com/perf/tips/calcent.cgi>

From Perfkit FCX306 LSHARACT

Core counts:	CP	ZAAP	IFL	ICF	ZIIP
Dedicated	0	0	0	0	0
Shared physical	1	0	108	0	0
Shared logical	1	0	322	0	0

(edited to show IFL cores only)

Core Type	Partition Name	Core Count	Load Max	LPAR Weight	Entlment	Cap	AbsCap	GrpCapNm	GrpCap	<CoreTotal,%>	Core
										Busy Excess	Conf
IFL	LPAR01	64	6400	10	133.3	No	...	...	...	.2	.0 o
<b>IFL</b>	<b>LPAR02</b>	<b>1</b>	<b>100</b>	<b>10</b>	<b>133.3</b>	<b>No</b>	<b>...</b>	<b>...</b>	<b>...</b>	<b>.1</b>	<b>.0 u &lt;--</b>
IFL	LPAR03	30	3000	60	800.0	No	...	...	...	81.8	.0 o
IFL	LPAR04	20	2000	60	800.0	No	...	...	...	57.5	.0 o
IFL	LPAR05	20	2000	60	800.0	No	...	...	...	135.3	.0 o
IFL	LPAR06	20	2000	60	800.0	No	...	...	...	82.2	.0 o
IFL	LPAR07	20	2000	60	800.0	No	...	...	...	58.9	.0 o
IFL	LPAR08	20	2000	60	800.0	No	...	...	...	199.6	.0 o
IFL	LPAR09	12	1200	60	800.0	No	...	...	...	1.4	.0 o
IFL	LPAR10	30	3000	60	800.0	No	...	...	...	1.2	.0 o
IFL	LPAR11	30	3000	60	800.0	No	...	...	...	1.4	.0 o
IFL	LPAR12	4	400	10	133.3	No	...	...	...	25.0	.0 o
<b>IFL</b>	<b>LPAR13</b>	<b>22</b>	<b>2200</b>	<b>200</b>	<b>2666.7</b>	<b>No</b>	<b>...</b>	<b>...</b>	<b>...</b>	<b>602.1</b>	<b>.0 u &lt;--</b>
IFL	LPAR14	6	600	10	133.3	No	...	...	...	2.9	.0 o
IFL	LPAR15	8	800	10	133.3	No	...	...	...	176.6	43.3 o
IFL	LPAR16	7	700	10	133.3	No	...	...	...	7.4	.0 o
IFL	LPAR17	8	800	10	133.3	No	...	...	...	1.7	.0 o

# Hurdle: Improper LPAR Weights creates excess logical cores

**Scenario:** The LPAR weights on a z/VM partition are set too low, creating too many unentitled cores on which to dispatch work. This can create PR/SM overhead and dispatch delays.

## How to Overcome:

- increase weights of LPARs with excess logical cores
- Make sum of weights = 10 \* (# of shared physical cores)
- This makes entitlement = weight/10
- Handy Entitlement Calculator:  
<https://www.vm.ibm.com/perf/tips/calcent.cgi>

From Perfkit FCX306 LSHARACT

LPAR Data, Collected in Partition LPAR03

Core counts:	CP	ZAAP	IFL	ICF	ZIIP
Dedicated	0	0	0	0	0
Shared physical	0	0	29	0	0
Shared logical	0	0	42	0	0

LPAR03 logical cores:

VH: 3                  VM: 2                  VL: 6

Core Type	Partition Name	Core Count	Load Max	LPAR Weight	Entlment	Cap	AbsCap	GrpCapNm	GrpCap	<CoreTotal,%> Busy	Excess	Core Conf
IFL	LPAR01	2	200	9	26.1	No	...	...	...	1.2	.0	o
IFL	LPAR02	29	2900	850	2467.5	No	...	...	...	1507.6	.0	o
IFL	LPAR03	11	1100	140	406.4	No	...	...	...	717.7	311.3	o

I4 9 :\$ 2\$ :/54 <S) 8</' )  
H ; 8(2) 9

# Hurdle: RSUs and Red Alerts

- <https://www.vm.ibm.com/service/redalert/>
- This page documents critical service information, examples:
  - APARs that are important though the abstract might not be obvious as to why you would need it.
  - Problems related to different factors (e.g. millicode change and unrelated z/VM feature)
  - Potential high-impact problems where APAR or PTF may not be available yet.
- Low number of red alerts, ~ 1 to 3 per year
- Highly recommend you subscribe for notification of changes to this page.

- <https://www.vm.ibm.com/service/rsu/index.html>
- This page documents RSU Content and key information:
  - Plans for next RSU (recommend you subscribe)
  - Specific RSU Installation information
- You also need to look at PSP bucket information
  - NOTE: previous PSP bucket search application is being sunset. Use [IBM Support](#) search instead.

# Hurdle: Reworking of Local Modifications

**Scenario:** You reworked a local modification (eg. RACF buildlist) and things worked fine. Some time later (usually when trying to use an IBM function related to the update)...an error occurs.

**How to overcome:**

Take extra Precaution when following the Service Guide for Local Modifications

- z/VM Service Guide Chapter 4: Reworking a Local Modification
  - 4a) run LOCALMOD
  - 4b) Reply to any Prompts
  - 4c) Make your changes to the displayed file
  - 4d) File your Changes

D / 8 M \$ / 4 : H ; 8 ( 2 ) 9

# Hurdle: DirMaint is misconfigured

**Scenario:** You're hitting errors with DirMaint, but you can't pinpoint what exactly is wrong.

**How to overcome:**

- Use the DIRM IVP command to validate the configuration (available in z/VM 7.2 and newer)





RACF H ; 8(2) 9

# Hurdle: Help! I'm locked out!

**Hurdle:** Something has gone wrong with your RACF server, and you need to recover your system.

## Best Practices for system recovery:

- Always ensure you have a SPECIAL user that is not revoked
- Always have a non-RACF enabled CPLOAD MODULE available
- Always have the current USER DIRECT stored off system
- Have OPERATOR logged on through the HMC if possible



# Staying Safe with RACF

Check out "The Junior Woodchuck's Guide to Repairing your RACFVM Database"

<https://www.vm.ibm.com/devpages/hugenbru/RACDBREP.PDF>

The Highlights:

- Make regular backups of your primary and secondary databases
- Perform health checks of your backups
- Check your Database Level (RACFCONV)
- RACUT200 (database verification utility) is your friend.....use it regularly



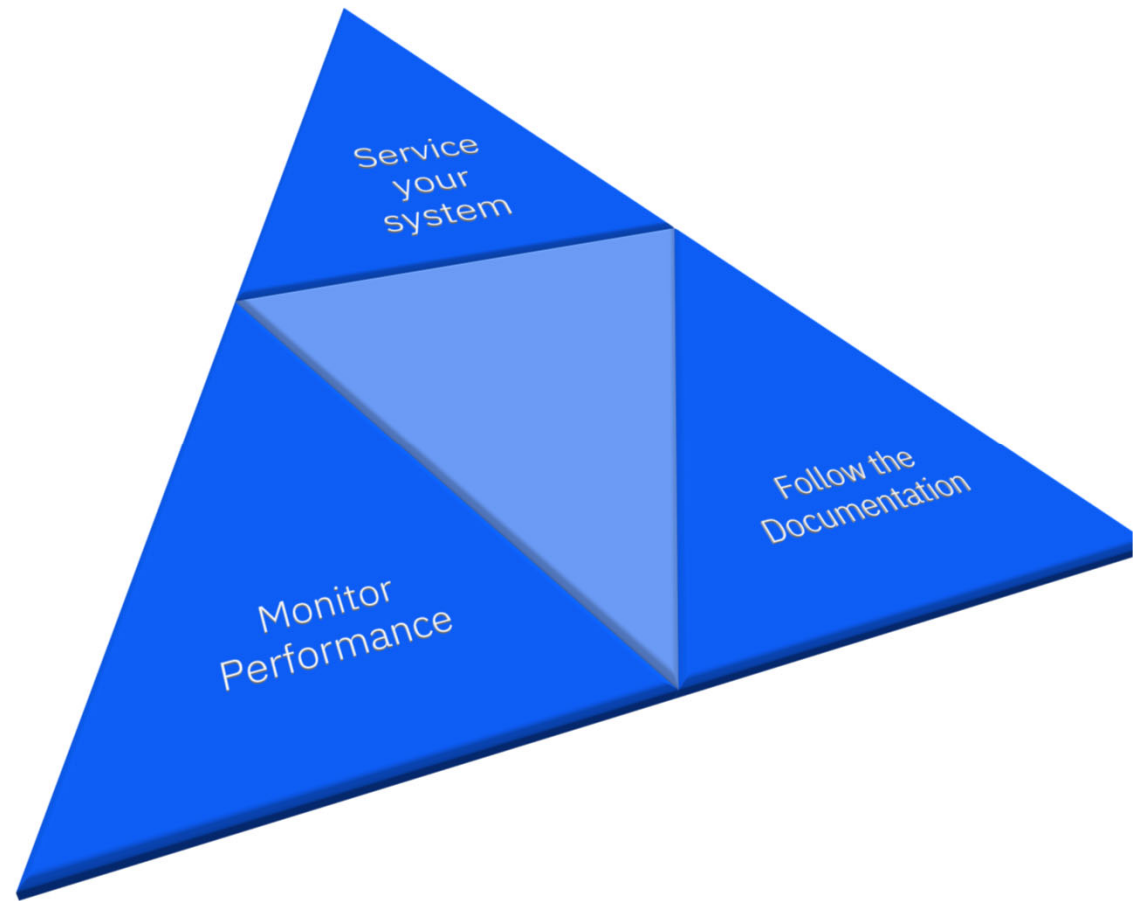
© 1954, Walt Disney Corporation

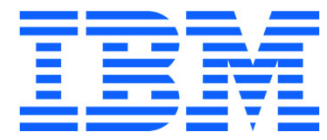
P\$8:/4+ T.5;+. :9

# The 3 Best Ways to Avoid z/VM Problems

1. Stay current on z/VM Service
2. Have a strategy for monitoring performance
3. Follow the Documentation

**Bonus Tip:** Automate where you can, document when you can't





B\$'1;6 C.\$8:9



# Hurdle: DirMaint is failing during Upgrade

**Scenario:** DirMaint errors received during z/VM upgrade procedure.

## How to overcome:

If error messages indicate problem with DASD:

- Make sure to add the release volumes to the EXTENT CONTROL file prior to upgrade.
- Make sure the updated EXTENT CONTROL file is placed on the DIRMAINT server
- Make sure the new EXTENT CONTROL file has been placed online (via DIRM RLDEXTN command)

Other things to check:

- Ensure MIGMAINT authorization is specified correctly. Must be authorized for **ONLY**

```
ALL MIGMAINT * 140A ADGHMOPS  
ALL MIGMAINT * 150A ADGHMOPS
```

- Ensure all required config files exist. The minimum files needed are:

- CONFIGxx DATADVH
- AUTHFOR CONTROL
- EXTENT CONTROL
- WHERE TO DATADVH

# Hurdle: DirMaint is failing during Upgrade in an SSI

**Scenario:** You are running into errors with DirMaint while upgrading members of an SSI cluster.

**How to overcome:**

- Be sure to bring up the DirMaint server on the member being upgraded.
  - Running the server on a different member is valid but can lead to timeout conditions that aren't always detected properly.

# Hurdle: DIRM PURGE of a user (or DIRM DMDISK) is delayed

**Scenario:** You have issued a DIRM PURGE, or DIRM DMDISK with the CLEAN option, and the operation has delayed (or hung)

## How to overcome:

- Prior to purging a USER or MDISK, ensure there are no active links to any MDISKS.
- If removing an MDISK, ensure the volume for the MDISK is attached to system.

# Hurdle: Where are my MDISKs?

**Scenario:** New volumes added but MDISK allocation are not available.

**How to overcome:**

- Make sure to add the new volumes to the EXTENT CONTROL file prior to upgrade.
- Make sure the updated EXTENT CONTROL file is placed on the DIRMAINT server
- Make sure the new EXTENT CONTROL file has been placed online (via DIRM RLDEXTN command)

Hurdle: You messed with an  
IBM supplied ID