

# Using New CP Features in z/VM 5.4

Session 9111 – SHARE 111 – San Jose, CA  
August, 2008

Romney White – System z Software Strategy

[romneyw@us.ibm.com](mailto:romneyw@us.ibm.com)

# Trademarks

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

DB2*	FICON*	Parallel Sysplex*	VisualAge*
DFSMS/VM*	FlashCopy*	Performance Toolkit for VM	VM/ESA*
Distributed Relational Database Architecture	GDDM*	QMF	z9
DirMaint	HiperSockets	RACF*	z/Architecture*
DS4000	IBM*	REXX	z/OS*
DS6000	IBM eServer	System Storage	z/VM*
DS8000	IBM logo*	System z	z/VSE
Enterprise Storage Server*	Language Environment*	System z9*	zSeries*
Enterprise Systems Architecture/390*	MQSeries*	System z10	zSeries Entry License Charge
ESCON*	OMEGAMON*	Tivoli*	

\* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

InfiniBand is a registered trademark of the InfiniBand Trade Association.

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation in the United States and other countries.

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

## Notes:

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

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

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

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

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

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

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

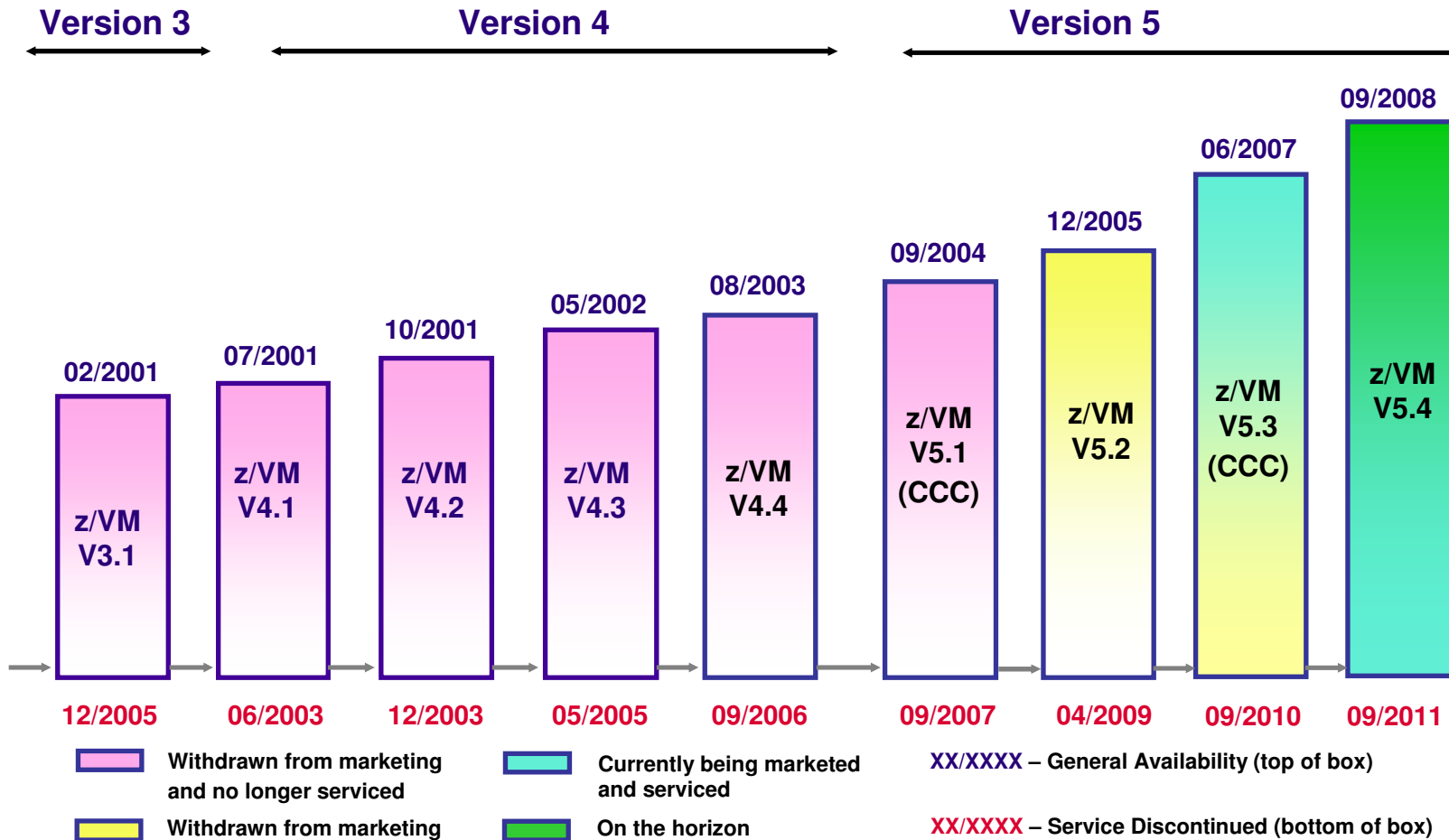
# Agenda

- **z/VM® Evolution**
- **z/VM Version 5 Release 3 (V5.3) Service**
- **z/VM Evaluation Edition**
- **z/VM Version 5 Release 4 (V5.4) Enhancements**
  - Enhanced scalability and constraint relief
  - Virtualization technology and Linux enablement
  - Network virtualization
  - Miscellaneous
- **Statement of Direction**



# z/VM Evolution

*Expand your opportunities with z/VM*



# z/VM Version 5 Terms and Conditions

- **International Program License Agreement (IPLA)**

- Program Use License
  - One-time charge (OTC) for standard or IFL processors
    - Engine-based Value Units
  - Can be transferred within an enterprise
  - Service by mail, fax, and e-mail only under basic warranty
- Subscription and Support (S&S)
  - Comparable service to traditional ICA products
  - Not required but highly recommended
  - Must decline when ordering if not desired
  - Annual renewable charge per processor
  - Adds telephone support
  - No additional charges for updates, new versions and releases

- **SoftwareXcel available for an additional charge**

- **IPLA applies to z/VM base code and the optional features**

- DirMaint, RSCS, RACF Security Server for z/VM, and the Performance Toolkit for VM

- **No-charge upgrade to the Performance Toolkit for VM for:**

- Customers who purchased z/VM V4 S&S for the RTM, PRF, or Performance Toolkit for VM features
- Customers who purchased the FCON/ESA program (5788-LGA)

## z/VM V5.3 Service

- **System z10 Processor Compatibility (VM64180, UM32277/32278)**
  - Guest support for Execute-Extensions Facility
- **System z10 IOP Subchannel Recovery (VM64242, UM32285/32286)**
- **Memory Management Improvement (VM64349, UM32250/32251)**
- **Encryption Re-Key Support – z/VM 5.3 only (VM64260, UM32221)**
  - SET TAPE REKEY
- **Virtual Switch Port Isolation (VM64281, UM32269/32270)**
  - Isolate guests on VLAN-unaware virtual switch (DROP or FORWARD inter-guest traffic)
- **System z10 Performance Toolkit OMEGAMON Enhancements (VM64337, UM32233/32234)**
- **System z10 Performance Toolkit Enhancements (VM64369, UM32257/32258)**
- **System z10 IOCP Support (VM64302, UM32279/32280)**
- **Missing adapter interruption detection (VM64398, UM32371/UM32372)**

## z/VM Evaluation Edition (GA: July, 2008)

- **DVD-RAM containing z/VM 5.3 (also available via Web download)**
  - Boot loader
  - z/VM 5.3 Nucleus
  - RAMdisk – one-pack system
    - CP
    - CMS
    - DIRMAINT
    - Performance Toolkit
- **Insert in HMC DVD drive and IPL (“Load from CD-ROM or server”)**
  - Requires 3GB or larger logical partition on System z10
  - Automatic invocation of customization script
- **Save customizations on DVD**
  - New guest definitions
  - Additional DASD allocations
  - System configuration changes

[www.vm.ibm.com/eval](http://www.vm.ibm.com/eval)

## z/VM V5.4 CP Enhancements (GA: Sept 12, 2008)

- **Enhanced scalability and constraint relief**
  - System z10 processor instruction exploitation
  - Dynamic memory upgrade
  - DAT Table performance enhancements
- **Virtualization enhancements for Linux and other guests**
  - Virtual CPU SHARE redistribution
  - Mixed-engine partition enhancements
  - Dynamic memory upgrade virtualization
  - Guest FCP dump
  - DCSS above 2GB
- **Network virtualization**
  - Multi-port OSA
  - Virtual switch networking management
- **Miscellaneous**
  - SHUTDOWN verification
  - OSA, FCP, and Hipersockets MIH support
  - Logical volume expansion compatibility
  - Enable PerfKit Diagnose 4 independence
  - Virtual Switch port isolation



# Enhanced Scalability and Constraint Relief

- System z10 processor instruction exploitation
- Dynamic memory upgrade
- DAT Table performance enhancements

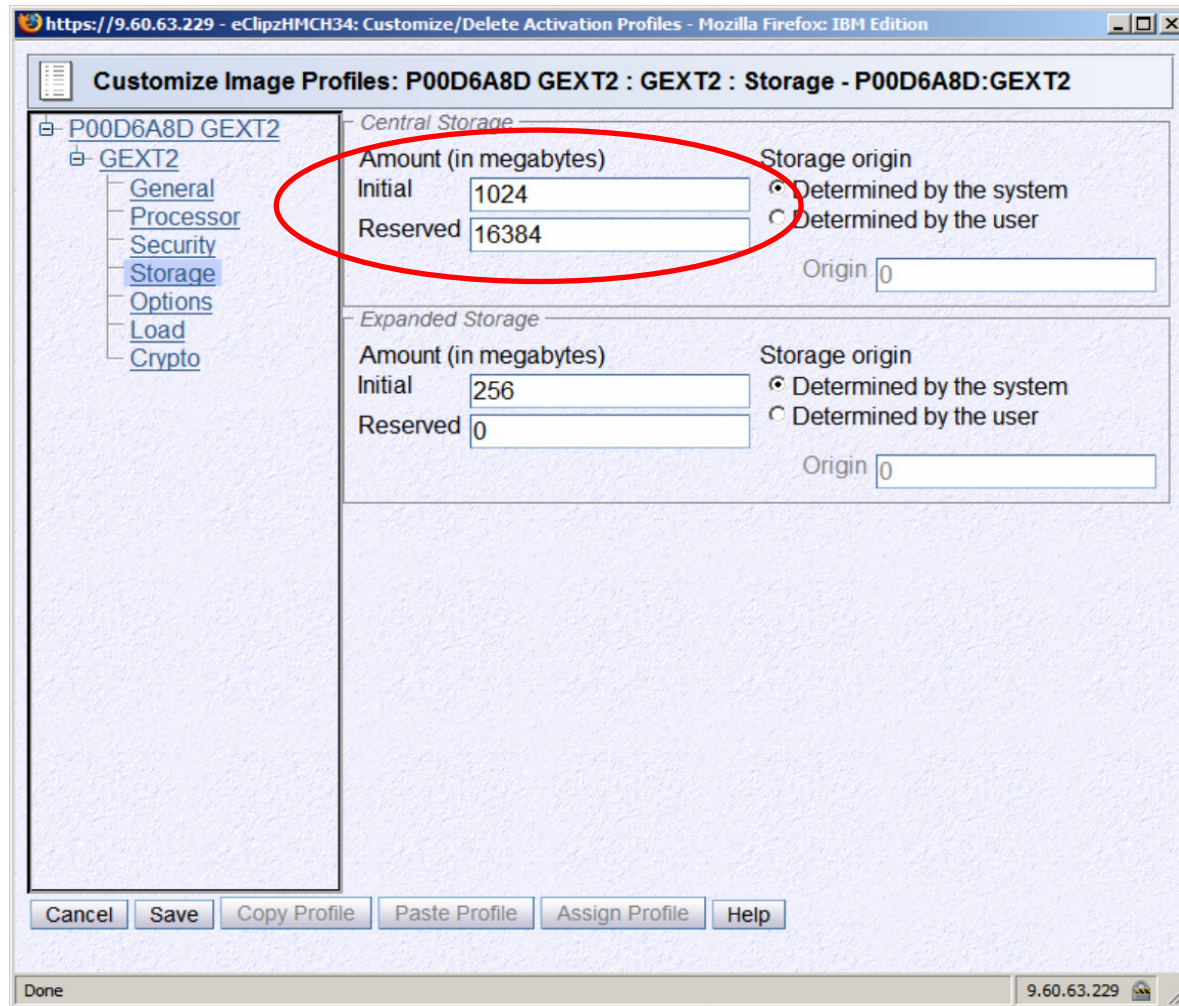
# System z10 New Instruction Support

- **Execute-Extensions facility**
  - Execute Relative long (EXRL)
- **General-Instruction-Extension Facility**
  - 71 new instructions, 4 new instruction formats, 7-character mnemonics
- **Parsing-Enhancement Facility**
  - Translate and Test Extended, Translate and Test Reverse Extended

# Dynamic Memory Upgrade

- **Allow z/VM to dynamically increase amount of real memory**
  - Define partition with Initial and Reserved amounts
  - Enable portion of Reserved amount to be brought online
  - Appropriate planning can enable memory growth to meet demand
- **Does not allow storage to be deconfigured dynamically**
  - Difficult technical challenge
  - May not meet some customer needs
- **Complements existing ability to add processors and I/O resources non-disruptively**

# Dynamic Memory Upgrade ...

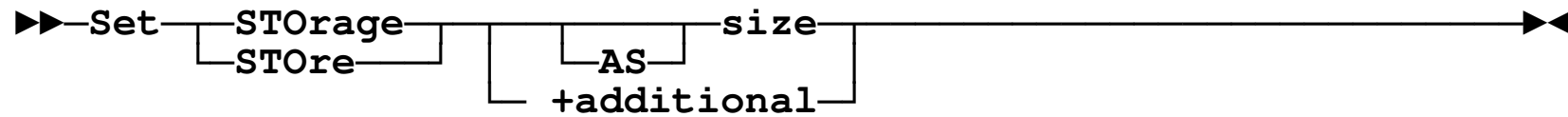


## Dynamic Memory Upgrade ...

- **Storage defined in increments**
- **Increment size determines multiples in which storage can be added**
- **Increment size**
  - Is a function of potential storage size (Initial + Reserved)
  - Is expressed in megabytes
  - Must be a power of 2
- **Cannot be more than 512 increments**

Storage Size (s)	Increment Size
$s \leq 512\text{M}$	1M
$512\text{M} \leq s \leq 1\text{G}$	2M
$1\text{G} \leq s \leq 2\text{G}$	4M
$2\text{G} \leq s \leq 4\text{G}$	8M
$4\text{G} \leq s \leq 8\text{G}$	16M
$8\text{G} \leq s \leq 16\text{G}$	32M
$16\text{G} \leq s \leq 32\text{G}$	64M
$32\text{G} \leq s \leq 64\text{G}$	128M
$64\text{G} \leq s \leq 128\text{G}$	256M
$128\text{G} \leq s \leq 256\text{G}$	512M

## Dynamic Memory Upgrade ...



## Dynamic Memory Upgrade ...

- **QUERY STORAGE reports additional information**

`STORAGE = d CONFIGURED = c INC = i STANDBY = s RESERVED = r`

- **Initialization of dynamically added storage above 2G is lazy**

- Immediate if demand exists (may be partial)
- When idle otherwise

- **Monitor records**

- MRMTRMCC (Event) created when partition memory configuration changes
- MRMTRMEM (Sample) includes standby and reserved storage sizes
- MRSTOADD (Event) created when memory added via SET STORAGE

# DAT Table Performance Enhancements

- **Support 64-bit host Control Registers**
  - CR 1 (Primary Space)
  - CR 7 (Secondary Space)
  - CR 13 (Home Space)
- **Enables Region and Segment tables that describe guest storage to reside above 2G**
- **Mitigates “needle in haystack” problem**
  - Upper-level DAT tables may require multiple contiguous frames
  - Can be difficult to find in constrained below-2G storage

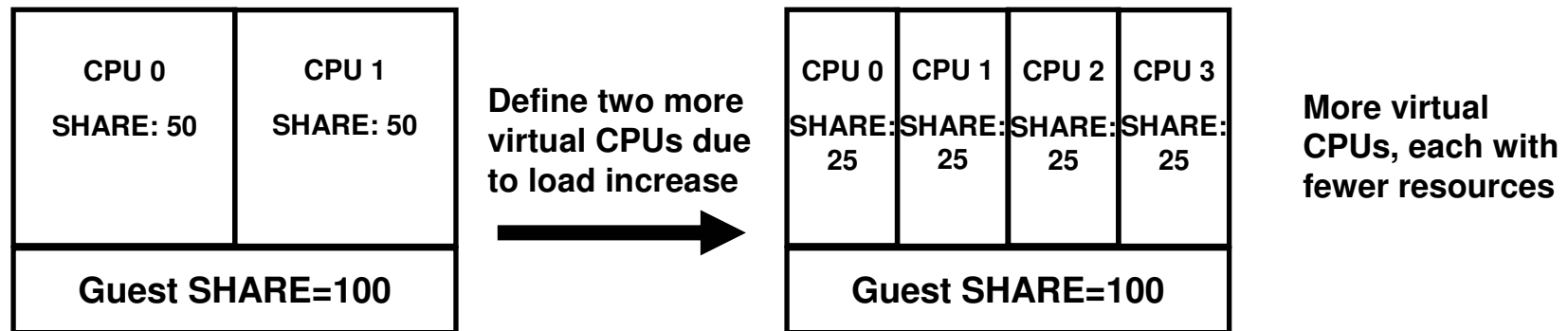


# Virtualization Technology and Linux Enablement

- **Virtual CPU SHARE redistribution**
- **Mixed-engine partition enhancements**
- **Dynamic memory upgrade virtualization**
- **Guest FCP dump**
- **DCSS above 2GB**

# Virtual CPU Share Redistribution

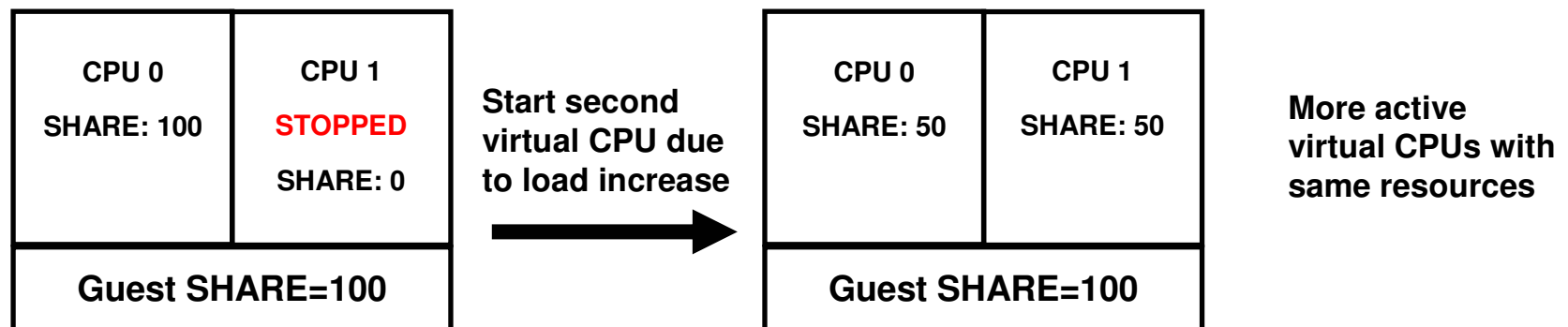
- **SHARE distributed evenly among guest virtual CPUs**
- **Prevents dynamic accommodation of changes in multiprogramming need**



- **Cannot reduce number of virtual CPUs with DETACH because of associated virtual system reset**
- **However, guest can start and stop virtual CPUs to control its multiprogramming level**

## Virtual CPU Share Redistribution ...

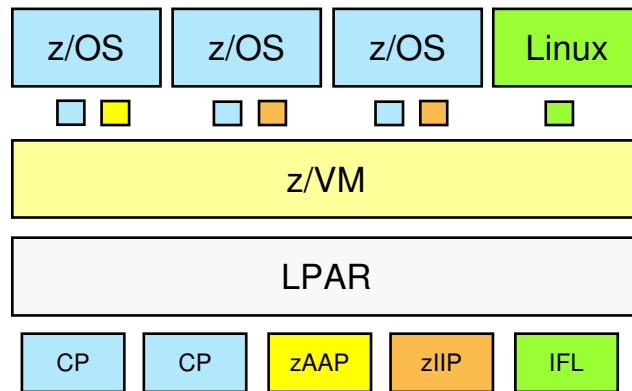
- Virtual CPU SHARE redistribution eliminates need for DETACH by not distributing SHARE to stopped virtual CPUs



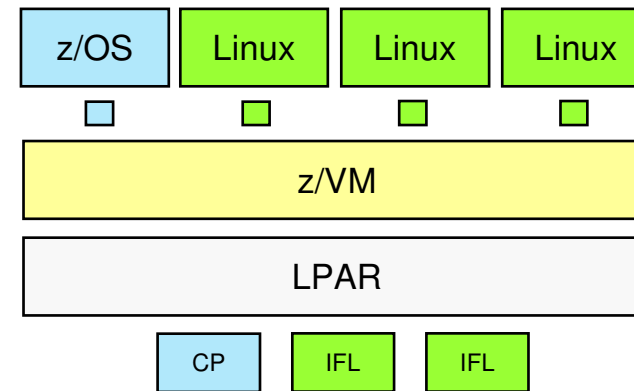
- Linux CPU hotplug daemon starts/stops CPUs based on Load Average
- Monitor records
  - MRMTRUSR, MRUSEACT, MRUSEINT – Added number of virtual processor starts/stops

# Mixed-Engine Partition Enhancements

- System z10 introduced “z/VM-mode” partitions
- Allow CPs, IFLs, zIIPs, zAAPs, and ICFs in a single partition
- z/VM manages guest dispatch appropriately



Add Linux on IFLs



Add z/OS on CPs

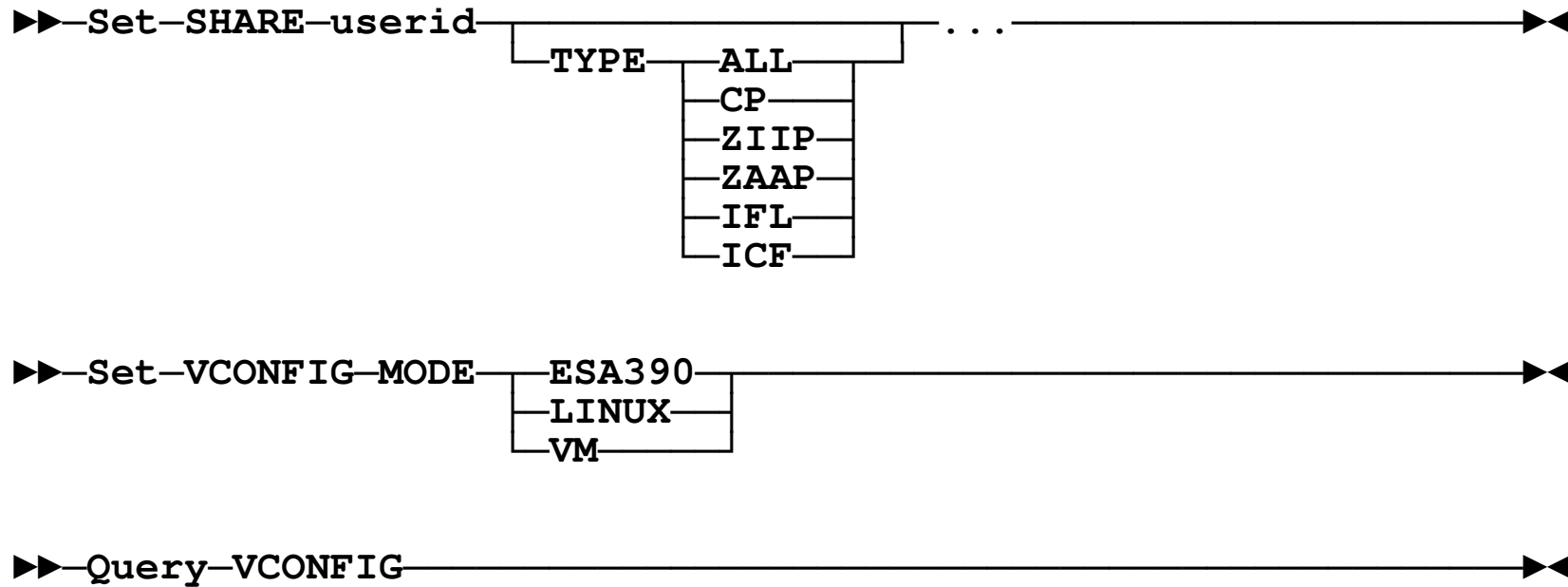
## Mixed-Engine Partition Enhancements ...

- **Consolidate partitions**
  - Reduce systems management effort
  - Improve processor, memory and I/O utilization
  - Simplify connectivity (Guest LAN, Virtual Switch)
- **Run IPLA products on IFLs**
- **Run MLC products on CPs**
- **Exploit spare ICF capacity**
- **Also fully virtualized**

## Mixed-Engine Partition Enhancements ...

- **QUERY PROCESSORS EXPANDED** reports partition mode
  - ESA/390, LINUX-ONLY, or z/VM
- **SET SHARE** allows different share settings for each processor type
  - Only applicable if CPUAFFINITY ON

# Mixed-Engine Partition Enhancements ...



# Mixed-Engine Partition Enhancements ...

## ▪ Monitor records

- MRUSEINT – Virtual configuration mode
- MRMTRSYS – Real configuration mode
- MRMTRUSR, MRSCCLADL, MRSCCLAEI, MRSCCLDDL, MRSCCLSHR, MRUSELOF, MRUSELON – CPU SHARE settings
- MRMTRUSR, MRSCCLSHR, MRUSEACT, MRUSEATE, MRUSELOF, MRUSELON – Per-CPU type SHARE settings
- MRMTRUSR – Number of virtual ICFs
- MRMTRPRP, MRPRCPRP, MRPRCVON, MRSTORSP, MRSTOSXP, MRSYTCOM, MRSYTTPRP, MRSYTRSP, MRSYTSCP, MRSYTSXP, MRSYTSYP, MRSYTXSP, MRUSEDFC, MRUSEDTC, MRUSEINT, MRUSEITE, MRUSELOF, MRUSERDC, MRUSETRE – New engine type code for ICF
- MRSYTUWT – Dispatch list counts for ICFs



# Dynamic Memory Upgrade Virtualization

- **Virtualized dynamic storage reconfiguration architecture**
  - Enables guests to dynamically increase or decrease memory
  - Limited by maximum storage size in User Directory
- **Supported by z/OS and z/VM guests**
- **Linux guest exploitation anticipated**

# Dynamic Memory Upgrade Virtualization ...



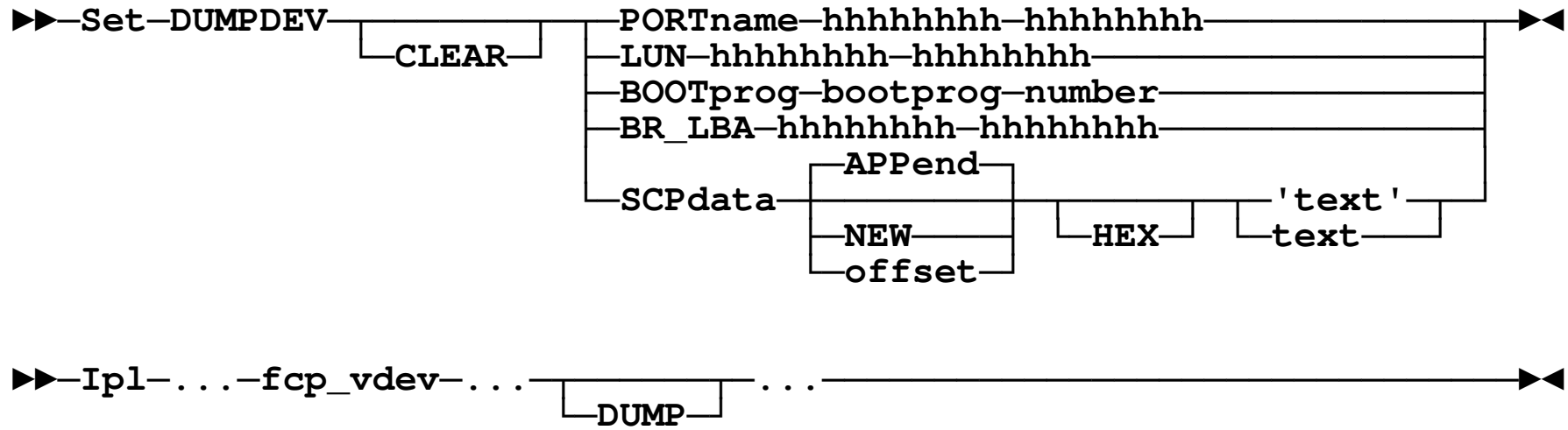
## Dynamic Memory Upgrade Virtualization ...

- **QUERY VIRTUAL STORAGE response enhanced**
  - `STORAGE = d MAX = m INC = i STANDBY = s RESERVED = r`
- **Increment rules same as for real storage**
- **Initial reserved storage size limited to maximum-defined**

# Guest FCP Dump

- **Enable guest to dump to SCSI device**
  - Capacity
  - Performance
- **Supported by Linux and z/VSE**

# Guest FCP Dump ...



## DCSS Above 2G

- **Problem exploiting DCSS support (e.g., with XIP)**
  - DCSS below 2G limited machine size and complicated configuration
  - DCSS  $\leq$  2047M limited size of shared memory
- **Enabling 2047M DCSS above 2G addresses both problems**
  - Some added complexity for exploiters (Linux)
  - New 64-bit interfaces to Diagnose 64

## DCSSS Above 2G ...

- **DEFSEG allows page ranges up to X'7FFFFFFF' (512 GB)**
  - SPACE, SECURE, and SC ranges not allowed
- **Number of page ranges for DEFSYS and DEFSEG reduced to 64 from 128**
- **QUERY NSS MAP response changed for segments above 2G**

```
FILE FILENAME FILETYPE BEGPAG          ENDPAG          TYPE CL #USERS
spid fn          DCSSG          nnnnnnnnnnnnnnnn nnnnnnnnnnnnnnnn type c  nnnnn
```

- **New Diagnose X'64' subcodes**
  - X'20' (LOADSHR), X'24' (LOADNSHR), X'2C' (FINDSEG), X'38' (SEGEXT)
  - Correspond to existing (31-bit) subcodes X'00', X'04', X'0C', and X'18'
  - Subcodes X'20', X'24', and X'2C' produce 64-bit addresses
  - New output area formats for subcode X'38' accommodate 64-bit addresses

# Network Virtualization

- **Multi-port OSA**
- **Virtual switch networking management**



## Multi-Port OSA

- **System z10 OSA-Express3 support**
  - Multiple ports per adapter on one CHPID
- **Allow port numbers to be specified for Virtual Switch real devices**
- **Virtualization enables any single port number to be used**
- **Report port number in QUERY VSWITCH, QUERY PORT, QUERY LAN, QUERY NIC**



## Multi-Port OSA ...

- **Report port number in QUERY VSWITCH, QUERY PORT, QUERY LAN, QUERY NIC**
- **Monitor records**
  - Add port number to MRIODVSW, MRIODVSF, and MRIODVSR
- **Diagnose X'26C' Subcode X'20'**
  - Add real port number to output

# Virtual Switch Management Improvements

- **Record Virtual Switch activity counters at either individual VLAN or entire network level**
- **Improve Virtual Switch/Guest LAN Query selectivity**
- **Recognize failures of Virtual Switch backup OSA**
  - Detect and correct problems when they happen
  - Avoid problems when failover occurs

# Virtual Switch Management Improvements ...

▶▶ Set-VSWITCH-switchname-...-VLAN\_counters ON  
OFF ... ▶▶

▶▶ MODify-VSWITCH-switchname-...-VLAN\_counters ON  
OFF ... ▶▶

▶▶ Query-LAN-laname-... 

VLAN ANY
VLAN-vlanid
VLAN-NONE
USERid-userid

 ... ▶▶

▶▶ Query-VSWITCH-switchname-... 

VLAN ANY
VLAN-vlanid
VLAN-NONE
USERid-userid
RDEV-nnnn

 ... ▶▶

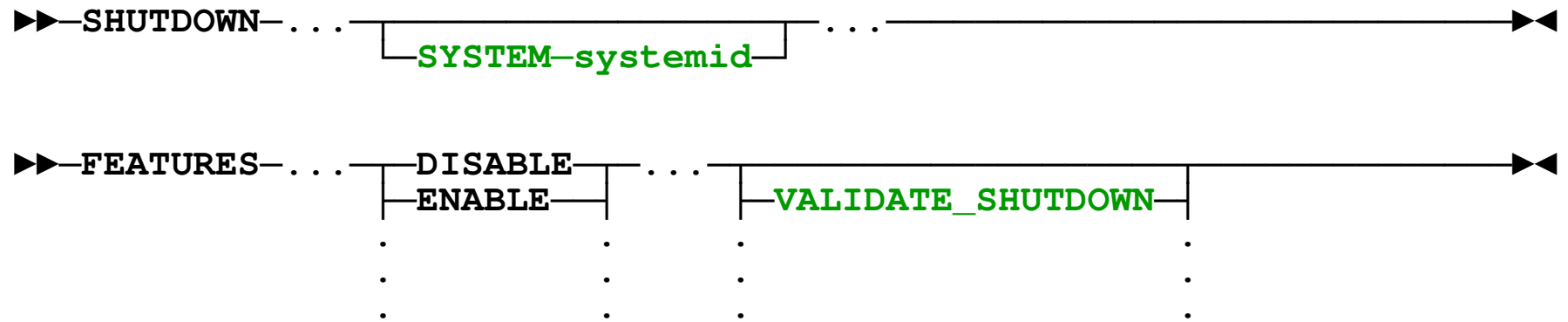
## Miscellaneous

- **SHUTDOWN verification**
- **Adapter MIH support**
- **Logical volume expansion compatibility**
- **Enable Performance Toolkit Diagnose 4 independence**
- **Virtual Switch port isolation**

## SHUTDOWN Verification

- **System configuration option to require system name on SHUTDOWN command**
- **May prevent inadvertent SHUTDOWN of first-level system**
  - Better to not give second-level guest Class A privileges

# SHUTDOWN Verification ...





## Adapter MIH Support

- **Detect and attempt to correct missing interruptions on OSA, FCP, and Hipersockets adapters**
- **Allow reasonable period of time (five seconds) for expected interruptions to be presented**
- **Attempt to recover up to ten times**
- **APAR VM64398 for z/VM 5.2 and 5.3**

## Adapter MIH Support ...

- **New Operator message issued**

HCPVAP2154E Missing adapter interruption

- **Monitor data**

- New fields in MRSYTSYG

- Number of missing adapter interruptions detected
- Number of unrecoverable missing adapter interruptions

# Logical Volume Expansion Compatibility

- **DS8000 added support to allow volumes to be expanded dynamically**
  - ECKD: 3390-3 and 3390-9, SCSI
  - Up to 65,520 cylinders
- **z/VM compatible with this feature**
  - Recognizes when expansion has occurred
    - Dedicated: Reflects interrupt to guest
    - Non-dedicated: Prevents use of expanded area until device varied off/on
  - Device model number may change
    - Expand 3390-3 beyond 3339 cylinders => 3390-9
- **APAR VM64305 (PE: VM64354)**

## Logical Volume Expansion Compatibility ...

- **QUERY DASD DETAILS** reports pending expansion

PENDING EXPANSION DETAILS: DEVTYPE = 3390-0C, CYLS = 65535

- **QUERY EDEVICE DETAILS** reports pending expansion

PENDING EXPANSION NUMBER OF BLOCKS: 147483640

## Enable PerfKit Diagnose 4 Independence

- **Performance Toolkit used Diagnose 4 to obtain fields not in Monitor Data**
- **Monitor data enhanced to eliminate that dependency**
  - MRPRCPRP – CPU operating state
  - MRSTOXSG – Additional page migration counters
  - MRSYTPRP – Additional fast-path function counters
  - MRSYTXSG – Additional Minidisk Cache counters
  - MRUSEACT – Number of Minidisk Cache inserts

## Virtual Switch Port Isolation

- **Prevent CP from switching packets between guests on same Virtual Switch**
  - Allows sharing of OSA(s) without enabling inter-guest communication
  - Can drop packets or forward through OSA for disposition (e.g., by firewall)

## Virtual Switch Port Isolation ...

▶▶ `Set-VSWITCH-switchname-...-ISOLation`

OFF
DROP
FORWARD

 ... ▶▶

▶▶ `MODIFY-VSWITCH-switchname-...-ISOLation`

OFF
DROP
FORWARD


 ... ▶▶

## Statement of Direction

- **MMC (Mainframe to Micro Channel) card**  
IBM intends to withdraw support in a future z/VM release for the MMC card, which enabled communication between the PS/2 (PWSCS) and VM (PWSCF or ISFC).
- **z/VM Flashcopy enhancements**  
IBM intends to enhance z/VM FlashCopy capabilities to support the FlashCopy SE function of the IBM DS8000 with the PTF for APAR VM64449 in fourth quarter 2008. FlashCopy SE offers a space-efficient snapshot capability that reduces the storage capacity needed for point-in-time copies. This function is especially useful for short-lived testing or backups, such as flash to intermediate volume for backup to tape.



**Questions?**

The image shows a close-up, low-angle view of a server rack. The perspective is from the bottom, looking up at the server units. The server units are dark grey or black, with a prominent diagonal line running across the frame. In the center, the IBM logo is displayed in a white, stylized font. Below the logo, the text '@server' is visible in a smaller, red font. The background is dark, and the lighting creates a sense of depth and texture on the server surfaces.

IBM

@server