GSE Frühjahrestagung 2013
Z/VSE und z/VM mit Linux auf System z,
U30 - zTalents
Leipzig, 22.-24. April 2013

VM06 z/VM Topics

Dr. Manfred Gnirss
IBM Client Center, IBM Germany Lab
gnirss@de.ibm.com

© 2013 IBM Corporation

z/VM 6.2 Kurs

z/VM 6.2 Update Kurs ist im Kundenkurskatalog buchbar und wird auch problemlos gefunden, wenn man z.B. den Suchbegriff z/VM Update eingibt

2. - 6. September 2013
IBM Mainz, Hechtsheimer Str. 2, Gebäude 20

Enrollment:
http://www-304.ibm.com/jct03001c/services/learning/ites.wss/de/de?pageType=page&c=a0003393
Price: 3.100 EUR

z/VM 6.2.0-Update Course – Course Code ZOVME3DE
This course introduces you to install multiple z/VM Systems interconnected within a cluster which share some system resources. The class focuses on the required steps of implementing DirMaint defining a Linux server and installing service in a single z/VM System. You will participate in hands-on labs where you will do a live relocation of a Linux server in a Single-System Image Cluster.

© 2013 IBM Corporation

Agenda

- Release Status and Information
- z/VM Version 6 Release 2
- z/VM Version 6 Release 3
- Value of z/VM
- Futures and Statements of Direction

© 2013 IBM Corporation

Release Status and Information

© 2013 IBM Corporation

z/VM Release Status Summary

z/VM Level		GA	End of Service	End of Marketing	Minimum Processor Level	Security Level
Ver 6	Rel 3	3Q / 2013	TBA		z10	TBA
	Rel 2	12 / 2011	4 / 2015	3Q / 2013	z10	-
	Rel 1	10 / 2009	4 / 2013	12 / 2011	z10	EAL 4+[1] OSPP-LS
	Rel 4	9 / 2008	12 / 2014[2]	3 / 2012	z800, z900	-
Ver 5	Rel 3	6 / 2007	9 / 2010	9 / 2010	z800, z900	EAL 4+ CAPP/LSPP

[1] Available since Feb. 2013
[2] Or later (Announced August 7, 2012)

Marketed & Serviced

Serviced, but not Marketed

End of Service & Marketing

Extended support contracts are available.

© 2013 IBM Corporation

z/VM Version 5 Release 4

- The last release of z/VM to support IBM System z9 and older processors
  - **No longer available as of March 12, 2012**
  - Also supports the zEC12

- End of Service has been extended to **December 31, 2014** or end of IBM service for System z9, whichever is ***later***
  - Was September 30, 2013
  - Later, but not *too* much later!
  - Be on the lookout for Delayed Onset Panic Syndrome

© 2013 IBM Corporation

z/VM Version 6
Security Certification Plans

- Common Criteria (ISO/IEC 15408)
  - Statement of Direction issued July 22, 2010
  - **Evaluation is complete (BSI-DSZ-CC-0752)**
  - Security Target: Operating System Protection Profile (OSPP) at EAL 4+
    - Virtualization extension
    - Labeled Security extension

- Federal Information Protection Standard (FIPS) 140-2
  - z/VM 6.1 System SSL is FIPS 140-2 Validated(TM)
  - Enablement requirements for certificate database and servers

    - http://csrc.nist.gov/groups/STM/cmvp/documents/140-1/1401val2012.htm#1735

- z/VM 6.2 is <u>designed to conform</u> to both Common Criteria and FIPS 140-2 evaluation requirements

TM *A Certification Mark of NIST, which does not imply product endorsement by NIST, the U.S. or Canadian Governments.*

© 2013 IBM Corporation

z/VM Version 6 Release 2

© 2013 IBM Corporation

z/VM Version 6 Release 2

- Generally available December 2, 2011
- Will be withdrawn from marketing 3Q 2013
  - Concurrent with z/VM V6.3 GA
- Major changes include:
  - Single System Image
  - Live Guest Relocation
  - Turnkey support for Unified Resource Manager
- Replaced z/VM V6.1
  - No longer available
  - Ends service Real Soon!

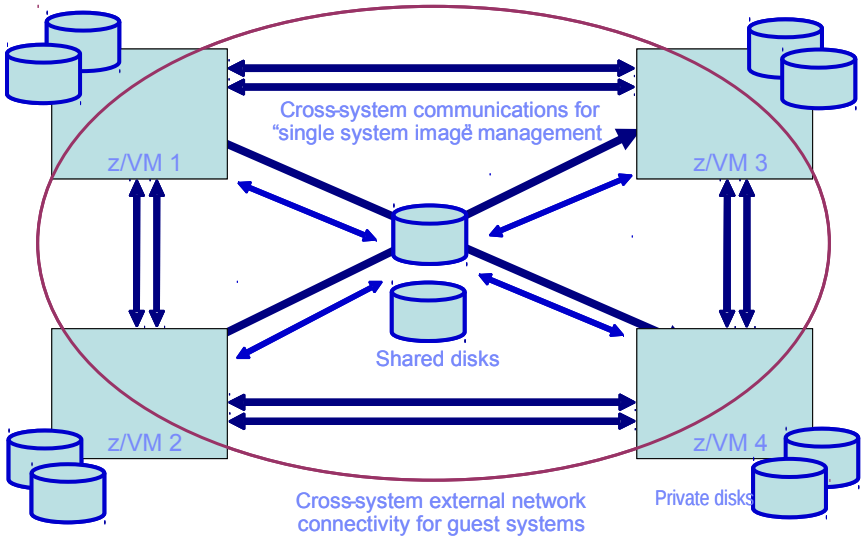
z/VM V6.2
End of Service
April 30, 2015

z/VM V6.1
End of Service
April 30, 2013

© 2013 IBM Corporation

Single System Image (SSI) Feature
Clustered Hypervisor with Live Guest Relocation

- An optional priced feature
- Connect up to four z/VM systems as members of a Single System Image cluster
- Provides a set of shared resources for member systems and their hosted virtual machines
- Cluster members can be run on the same or different System z servers
- Simplifies management of a multi-z/VM environment
  - Single user directory
  - Cluster management from any member
    - Apply maintenance to all members in the cluster from one location
    - Issue commands from one member to operate on another
  - Built-in cross-member capabilities
  - Resource coordination and protection of network and disks



© 2013 IBM Corporation

SSI Cluster Management – Features for Greater Reliability

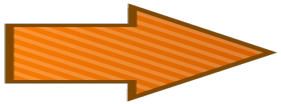
- Cross-checking of configuration details as members join cluster and as resources are used:
  - SSI membership definition and identity
  - Consistent definition of shared spool volumes
  - Compatible virtual network configurations (MAC address ranges, VSwitch definitions)
- Cluster-wide policing of resource access:
  - Volume ownership marking to prevent dual use
  - Coordinated minidisk link checking
  - Autonomic minidisk cache management
  - Single logon enforcement
- Communications failure “locks down” future resource allocations until resolved
- Comprehensive checking for resource and machine feature compatibility during relocation:
  - Adjustment of “virtual architecture level” to support customer relocation policy

© 2013 IBM Corporation

Single System Image Feature
Clustered Hypervisor with Live Guest Relocation

- Dynamically move Linux guests from one member to another with Live Guest Relocation
  - Reduce planned outages
  - Enhance workload management
  - Non-disruptively move work to available system resources and non-disruptively move system resources to work

- When combined with Capacity Upgrade on Demand, Capacity Backup on Demand, and Dynamic Memory Upgrade, you will get the best of both worlds



Bring additional resources to the workload!

Move the workload to the resources!

© 2013 IBM Corporation

Single System Image Feature
Imbedded Relocation Wizard

- Eligibility checks done multiple times throughout the relocation process.
- Check more than just eligibility to move the virtual machine, but also check if it is “safe” to move.
  - Overrides are available
- Checks for:
  - Does virtual machine really have access to all the same resources and functions?
  - Will moving the virtual machine over commit resources to the point of jeopardizing other workload on the destination system?
- Pacing logic to minimize impact to other work in more memory constrained environments

© 2013 IBM Corporation

Single System Image Feature

- xCAT 2.8 intends to support SSI and LGR
  - Already supports z/VM today
- Unified Resource Manager (zManager) *does not* support SSI and LGR
- IBM Systems Director *does not* support SSI and LGR
- Suggested best practice is to not combine SSI and LGR with zManager or Systems Director
  - Work with your IBM Sales Team, IBM Lab Services, or z/VM Development Lab to determine which technologies are most critical to your environment and business.

© 2013 IBM Corporation

z/VM Storage Support

- z/VM 6.2 supports
  - DS8000 Series (FCP or FICON)
  - DS6000 Series (FICON)
  - XIV (FCP)
  - IBM San Volume Controller (FCP)
  - IBM Storwize V7000 (FCP)
    - See http://www.ibm.com/support/docview.wss?uid=ssg1S1003703#_zvm
  - As well as many of the older storage devices
- Note: The System Storage Interoperation Center (SSIC) support page has some omissions of the above support
  - http://www.ibm.com/systems/support/storage/ssic/interoperability.wss
- The z/VM 6.2 General Information Manual has additional information, but had not been updated for Storwize, see URL above for requirements.

© 2013 IBM Corporation

z/VM System Management

- **Operations Manager for z/VM V1.4**
  - Facilitates automated operations
  - Monitor, view, and interact with consoles without logging on to service machines or Linux guests
  - Take actions based on service machine console messages and other system events
  - Schedule events for immediate execution or on a regular schedule
- **OMEGAMON XE on z/VM and Linux V4.2**
  - Performance monitoring of z/VM and Linux guests
  - Part of the Tivoli Management Services, including Tivoli Enterprise Portal
  - Uses IBM Performance Toolkit for VM as its data source
- **Backup and Restore Manager for z/VM V1.2**
  - Backup and restore file level data for CMS minidisks and Shared File System
  - Backup and restore images of Linux guests and/or z/VM volumes
    - Use Tivoli Storage Manager for file level backup and restore of Linux data
- **Tape Manager for z/VM V1.3**
  - Manage tapes: retention, access control, data security erase
  - Manage devices: share with other z/VM and non-z/VM systems
  - Manage mount requests for ATL, VTS, and manual mount devices
- **Archive Manager for z/VM V1.1**
  - Users and administrators manage disk space more efficiently and effectively
  - Archive infrequently used or large files to tape or other disk
- **zSecure Manager for RACF z/VM V1.11.1**
  - Automate complex, time consuming z/VM security management tasks
  - Quickly identify and prevent problems in RACF
  - Create comprehensive audit trails

© 2013 IBM Corporation

z/VM 6.2 SSI and GDPS Support

- All supported GDPS releases (3.7, 3.8, & 3.9) supported with non-SSI environment
  - See GDPS PSP buckets for required service (z/OS, Linux, & z/VM)
  - If GDPS environment shared with older z/VM releases, z/VM service is required on them before adding z/VM 6.2

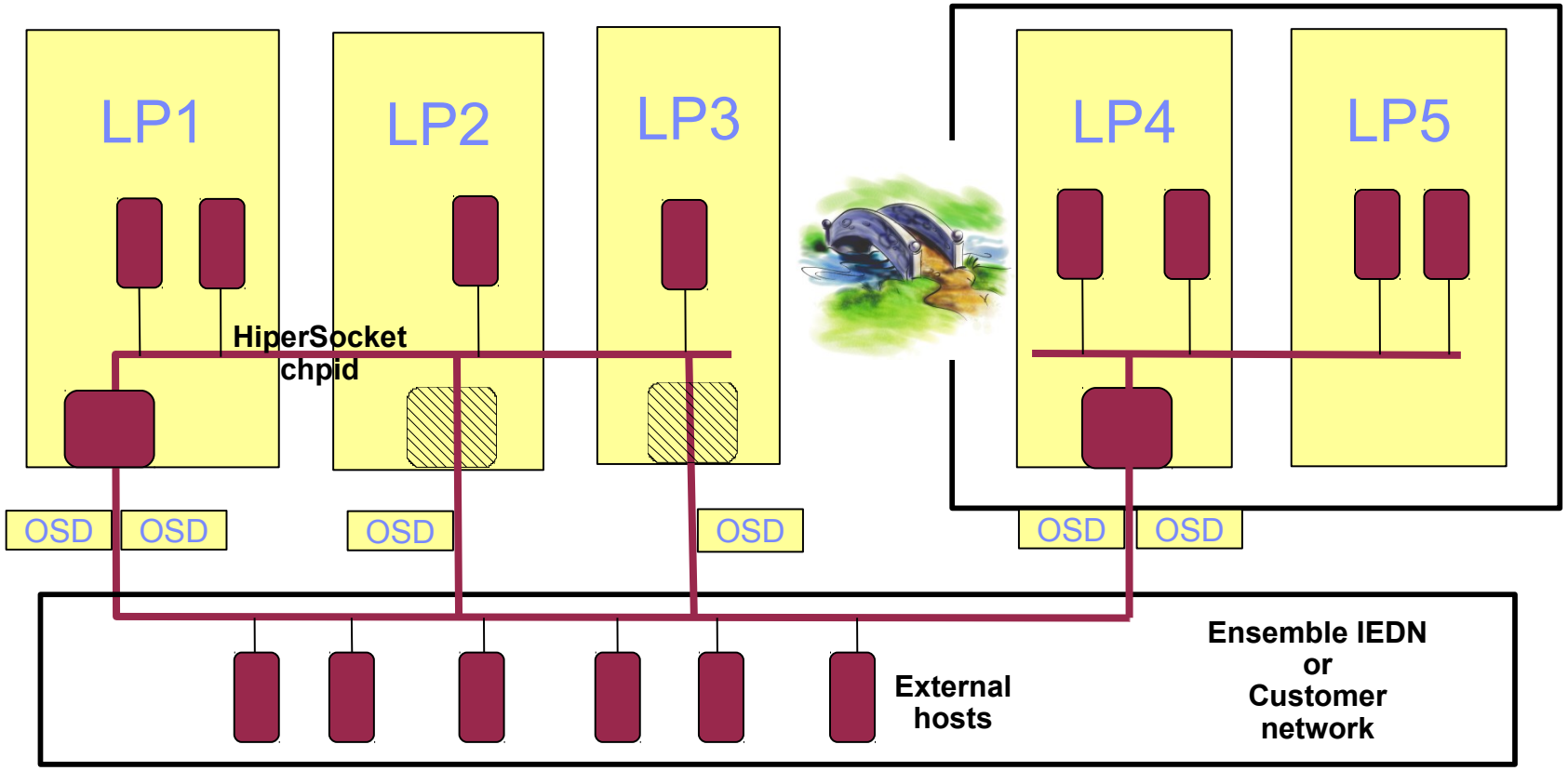
- GDPS/PPRC xDR 3.9 supports SSI configuration
  - APAR PM64211
  - Requires z/VM APAR VM65176

- No support for Live Guest Relocation of monitored Linux guests or z/VM services

© 2013 IBM Corporation

HiperSocket VSWITCH Bridge

Available: April 13, 2012



- Built-in failover and failback
- CHPARM=x2 to bridge IEDN
- CHPARM=x4 to bridge customer network
- Same or different LPAR
- One active bridge per CEC
- PMTU simulation

© 2013 IBM Corporation

HiperSocket VSWITCH Bridge

Available: April 13, 2012

- Virtual Switch bridge between Ethernet LAN and HiperSockets
  - zEnterprise IEDN (OSX) or customer network (OSD) connections
  - Guests can use simulated OSA or dedicated HiperSockets
  - VLAN aware
  - One HiperSocket chpid only
- Full redundancy
  - Up to 5 bridges per CEC
  - One bridge per LPAR
  - Automatic takeover
  - Optionally designate one “primary”
    - Primary will perform “takeback” when it comes up
  - Each bridge can have more than one OSA uplink
- CP: VM65042 / UM33691, TCP/IP: PM46988 / UK77220

© 2013 IBM Corporation

High Performance FICON

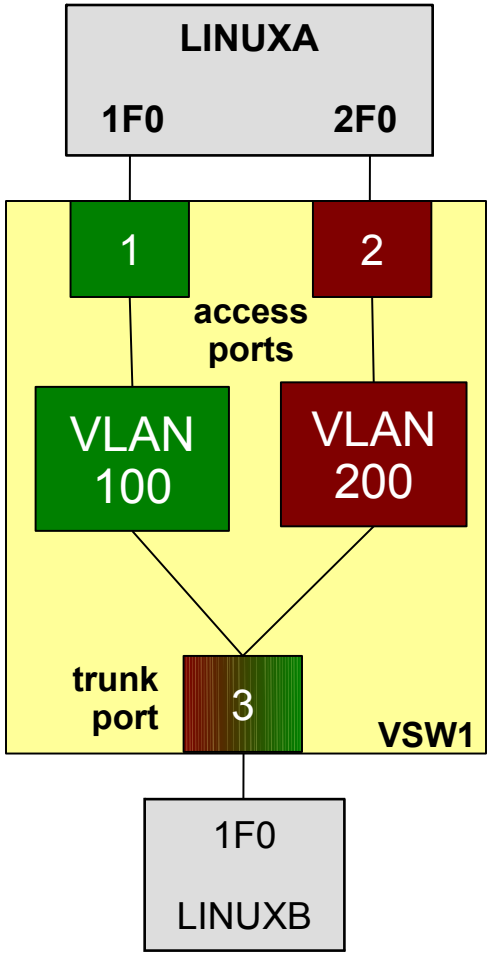
Available: April 13, 2012

- Enable guests to use High Performance FICON for System z (zHPF)
  - Different I/O model
  - Single and multiple track I/O
  - CP APAR VM65041 PTF UM33646
  - DVF APAR VM65144 PTF UM33647
- Requires host and control unit compatibility
  - Consult a storage specialist for details
- z/OS and Linux provide exploitation
- Performance results available at:
  - http://www.vm.ibm.com/perf/reports/zvm/html/620jb.html

© 2013 IBM Corporation

VSWITCH: Multiple virtual access ports per guest

- New PORTBASED attribute for the VSWITCH
- VLAN and port management similar to physical switches
- One or more virtual ports on a the VSWITCH are reserved for a guest
- Each port is associated with a single VLAN
  - Each VLAN on the vswitch contains a list of associated ports
  - VLAN change takes effect immediately
- GRANT is not required
- ESM authorization requirement unchanged
  - Must detach and re-COUPLE the virtual NIC



© 2013 IBM Corporation

Scalability and Performance Enhancements

- Reduction in the amount of memory and CPU required in order to manage larger memory sizes
- Integration of prior performance-related PTFs
  - SET REORDER
  - Reduce overhead of guest page release
  - Coalesce adjacent page frames
  - More accurace LIMITHARD
  - Reduce incidence of DIAGNOSE 0x9c and 0x44
  - Eliminate unexpected pauses in busy systems with large CPU over-commitment

© 2013 IBM Corporation

TCP/IP Enhancements

- Stack
  - RFC 4191: Router selection preferences
  - RFC 5175: IPv6 router advertisement flags extension
- FTP
  - IPv6
  - Passwords suppressed in server traces
  - Wildcards supported for BFS files
- SMTP
  - IPv6
  - Includes IPv6 support in CMS NOTE and SENDFILE

© 2013 IBM Corporation

TCP/IP Enhancements OSA Diagnostics

- The NETSTAT command has been updated to provide details taken from the OSA Address Table (OAT) via new OSAINFO option.
- OSA/SF no longer required to obtain device details
- OSA-Express3 and later

```
VM TCP/IP Netstat Level 620       TCP/IP Server Name: TCPIP

Device K4L3VSW6640DEV: data as of 09/23/11 01:05:21
  OSA Generation:                          OSA-Express3
  OSA Firmware Level:                      00000766
  Port Speed/Mode:                         1000 Mbs / Full Duplex
  Port Media Type:                         Multi Mode (SR/SX)
  PCHID:                                   0291
  CHPID:                                   0053
  Manufacturer MAC Address:                00-14-5E-78-17-F2
  Configured MAC Address:                  00-00-00-00-00-00
  Data Device Sub-Channel Address:         6640
  CULA:                                    00
  Unit Address:                            40
  Physical Port Number:                    0
  Number Of Output Queues:                 1
  Number Of Input Queues:                  1
  Number Of Active Input Queues:           0
  QDIO CHPID Type:                         OSD
  QDIO Connection:                         Not Isolated
  IPv4 L3 VMAC:                            00-00-00-00-00-00
  IPv4 VMAC Router Mode:                   No
  IPv4 L3 VMAC Active:                     No
  IPv4 L3 VMAC Source:                     n/a
  IPv4 L3 Global VLAN ID Active:           No
  IPv4 Global VLAN ID:                     0
  IPv4 Assists Enabled:                    00001C71
  IPv4 Outbound Checksum:                  00000000
  IPv4 Inbound Checksum:                   00000000

  IPv4 Address:                             IPA Flags:
-----------------------                   -----------------
9.60.29.53                                00000002

IPv4 Multicast Address:                   MAC Address:
-----------------------                   -----------------
224.0.0.1                                 01-00-5E-00-00-01
```


© 2013 IBM Corporation

RACF Security Server

- Single System Image Support
  - Automatic propagation of most RACF commands
  - Also works with multiple RACF servers on same z/VM system
- Protected Users
  - User without a password or password phrase will not be revoked due to too many invalid password attempts or inactivity
- High Level Assembler no longer required for most common customizations
- Real device protection
  - ATTACH, GIVE, DEDICATE
  - New VMDEV class
  - Profiles: RDEV.*device.system_id*
- Support for Diagnose 0xA0 Subcode 0x48
  - Obtain information about any ESM in architected format

© 2013 IBM Corporation

z/CMS

- Previously shipped with z/VM as a sample program, now supported
  - IPL ZCMS or MAINT 990
- Enables CMS programs to use z/Architecture instructions and 64-bit registers
- Existing ESA/390 architecture programs continue to run unchanged
  - CMS does not exploit memory above 2 GB
  - CMS *does* provide basic memory management API for memory above 2 GB
- Programs that examine or change architecture-sensitive memory locations (NUCON) must be updated in order to use z/CMS
- No z/Architecture extensions for DAT-off virtual machines to use access registers
  - VM Data Spaces not available

© 2013 IBM Corporation

Installation Improvements

- Significant changes to system layout to support Single System Image
- Choose a non-SSI system or a complete 1- to 4-member SSI cluster
  - First or second level
- All installation information is gathered at one time
- All DASD volumes can be labeled at installation time, including the system residence volume
- Turnkey support for zEnterprise ensembles enables clients new to z/VM to easily get started with Unified Resource Manager (zManager)
  - If you have purchased a directory and/or security manager, decline this option during installation; manual enablement is required!

© 2013 IBM Corporation

XEDIT – Default changed to mixed case

- For those coming to z/VM from an open system background, the folding of mixed case to upper case is surprising

- Many comments along the lines of “it hurts when you do that”
  - Linux can read CMS files
  - Often case-sensitive
- Default for other file types changed to CASE MIXED RESPECT
  - No folding
  - To get old behavior, update PROFILE XEDIT to
    SET CASE UPPER RESPECT
  - Or you may wish to consider
    SET CASE MIXED IGNORE

© 2013 IBM Corporation

Removed Functions

- Kerberos authentication system
  - IBM Software Announcement 208-249
- CMS-based Domain Name Server (NAMESRV)
  - IBM Software Announcement 209-207
- RESOURCE option of VMSES/E VMFINS command
  - IBM Software Announcement 210-234
- z/VM Manageability Access Point (zMAP) agent and Platform agent for IBM Systems Director for Linux on System z
  - Both previously shipped with z/VM V6.1

© 2013 IBM Corporation

z/VM Version 6 Release 3 Preview

Previews provide insight into IBM plans and direction. Availability, prices, ordering information, and terms and conditions will be provided when the product is announced.

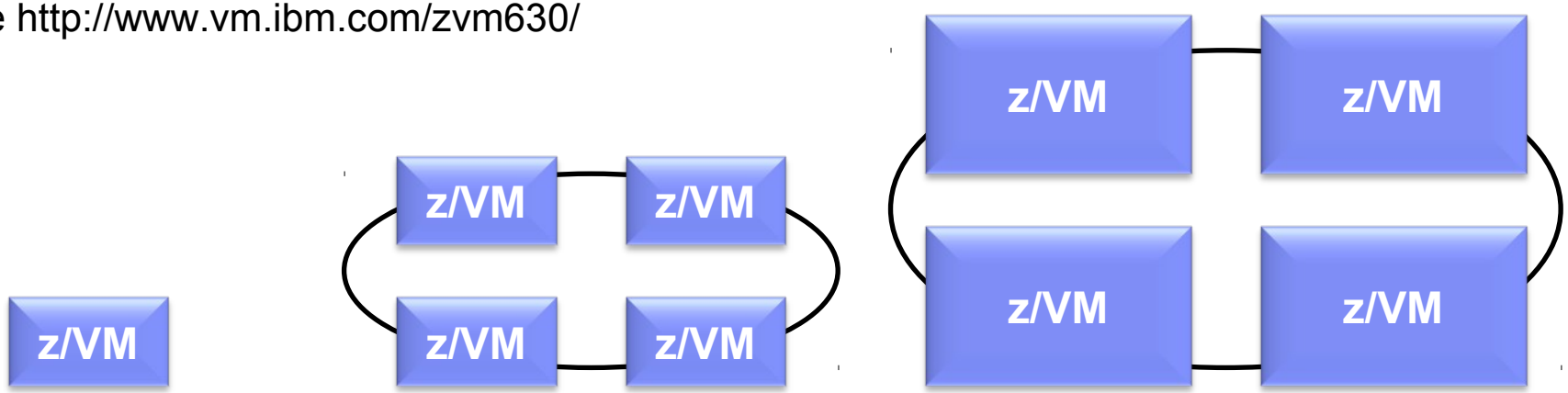
© 2013 IBM Corporation

z/VM Version 6 Release 3 Preview
Making Room to Grow Your Business

- Preview announcement on February 5, 2013
- Planned Availability: 3rd Quarter 2013

- Major Enhancements for Scalability and Performance
  - Support for larger amounts of real memory
  - Increased processor efficiency

- See http://www.vm.ibm.com/zvm630/



© 2013 IBM Corporation

z/VM Version 6 Release 3 Preview
Large Memory Support

- Real memory limit raised from 256GB to **1TB**
  - Proportionately increases total virtual memory based on tolerable over- commitment levels and workload dependencies
- Virtual machine memory limit remains unchanged at **1TB**
- Paging DASD utilization and requirements change
  - Removed the need to double the paging space on DASD
  - Paging algorithm changes increase the need to have a properly configured paging subsystem
- Expanded Storage continues to be supported with a limit of **128GB**

© 2013 IBM Corporation

z/VM Version 6 Release 3 Preview
Large Memory Support

- Reorder processing removed

- Improved effectiveness of the CP SET RESERVE command
  - Stronger “glue” to hold reserved pages in memory
  - Support for reserving pages of NSS or DCSS
    - Example: Use with the Monitor Segment (MONDCSS)
  - Ability to limit the overall number of reserved pages for the system

- Standalone dump to disk

© 2013 IBM Corporation

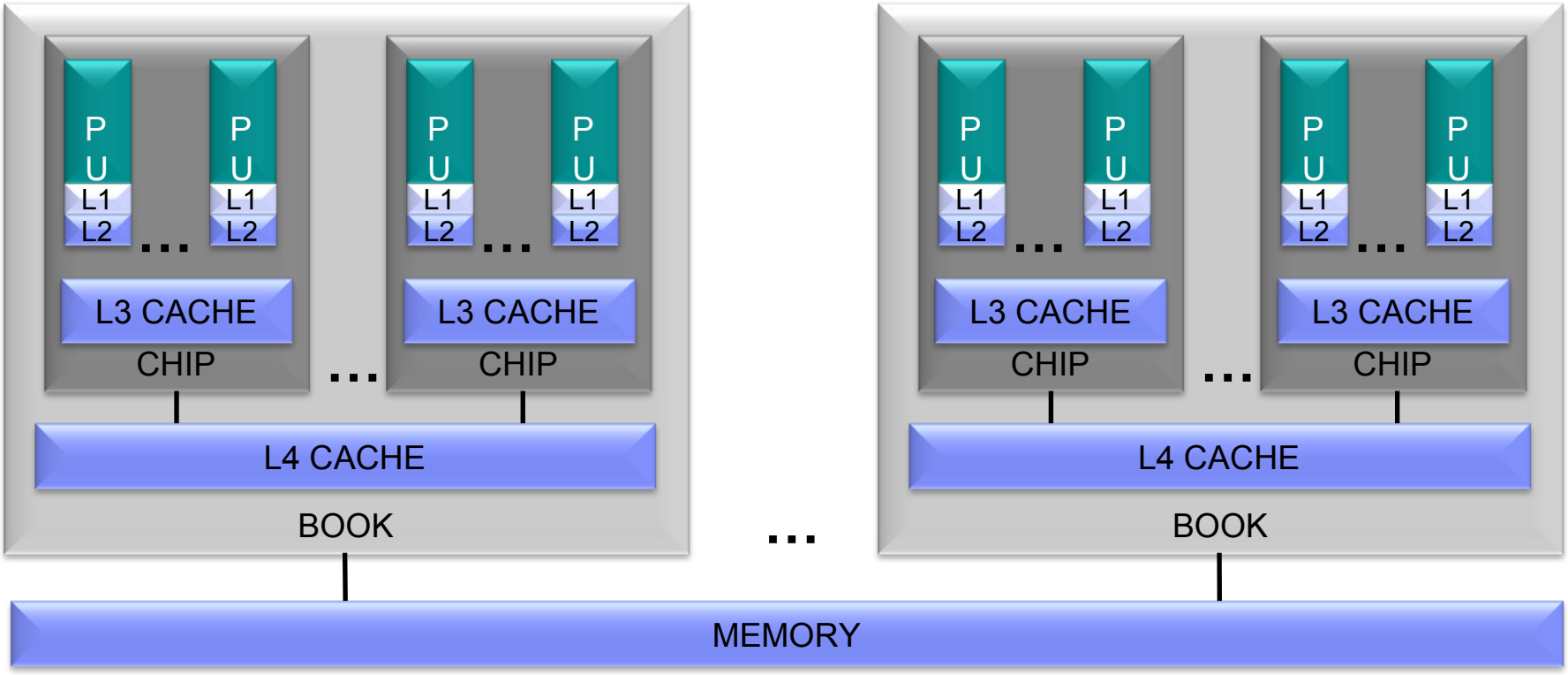
z/VM Version 6 Release 3 Preview
HiperDispatch

- Improved processor efficiency
  - Better n-way curves
    - Supported processor limit of 32 remains unchanged
  - Better use of processor cache to take advantage of cache-rich system design of more recent machines
- Two components:
  - Dispatching Affinity
  - Vertical CPU Management

© 2013 IBM Corporation

z/VM Version 6 Release 3 Preview
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



© 2013 IBM Corporation

z/VM Version 6 Release 3 Preview
HiperDispatch: Dispatching Affinity

- Dispatcher is aware of the cache and memory topology
  - Dispatch virtual CPU near where its data may be in cache based on where the virtual CPU was last dispatched
  - Keep virtual CPUs of the same virtual machine near one another
- Better use of cache can reduce the execution time of a set of related instructions
- z/VM V6.2 and earlier uses “soft” affinity to dispatch virtual CPUs
  - No awareness of chip or book

© 2013 IBM Corporation

z/VM Version 6 Release 3 Preview
HiperDispatch: Vertical CPU Management

- Today's “horizontal” management distributes the LPAR weight evenly distributed across the logical processors of the z/VM LPAR
- “Vertical” management attempts to minimize the number of logical processors, allowing LPAR to similarly manage logical CPUs

Example:

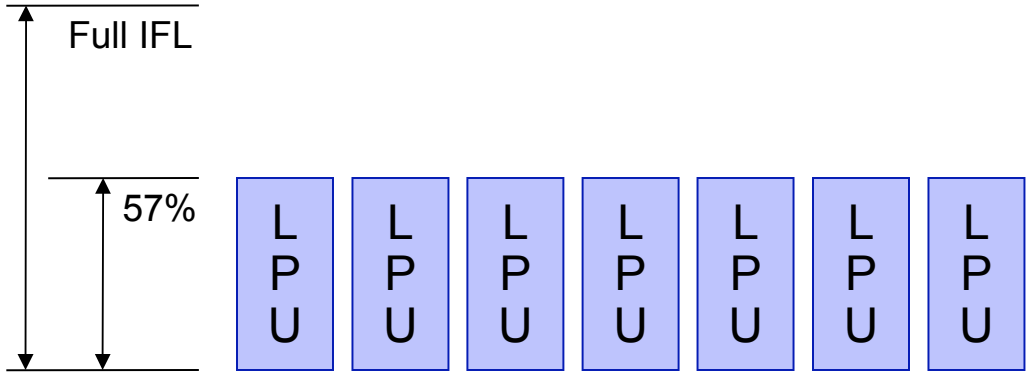
- 10 Physical IFLs, 7 logical IFLs, weight of 400 out of 1000
  - Each logical IFL (LPU) entitled to 57% of an IFL
- When CEC is constrained, the LPAR’s entitlement is reduced to 4 IFLs, so 7 is more than required
- z/VM & LPAR will cooperate
  - z/VM will concentrate the workload on a smaller number of logical processors
  - LPAR will redistribute the partition weight to give a greater portion to this smaller number of logical processors (~100% of 4 CPUs)

© 2013 IBM Corporation

Horizontal vs. Vertical CPU Management

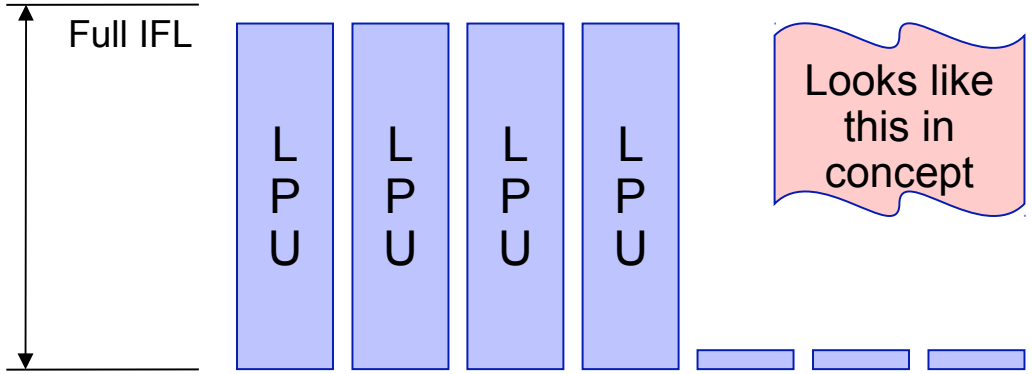
Horizontal:

- The logical processors are all created/treated equally.
- z/VM dispatches work evenly across the 7 logical processors



Vertical:

- The logical processors are skewed to where some get greater share of the weight.
- z/VM dispatches work accordingly to the heavier weighted workload.



© 2013 IBM Corporation

z/VM Version 6 Release 3 Preview
Items of Interest

- Virtual Switch updates

- Installation enhancements

© 2013 IBM Corporation

Statements of Direction

Subject to change or withdrawal without notice,
representing IBM goals and objectives only.

Note for withdrawals: Unless otherwise stated, it is IBM's intent that z/VM V6.2 will be the last release of z/VM to support the indicated function.

© 2013 IBM Corporation

HiperSockets Completion Queues

Available April 13, 2012

z/VM Statement of Direction: New function

- Transfer HiperSockets messages asynchronously
- Used whenever traditional synchronous queues are full
- Automatic enablement; no z/VM configuration required
- Helpful when traffic is “bursty”
- Exploitation by CP VSWITCH only; no guest simulation

© 2013 IBM Corporation

z/VM Performance Toolkit: RMFPMS agent
z/VM Statement of Direction: Stabilize existing function

- Performance Toolkit processing of the output from Linux rmfpms agent, part of the z/OS RMF PM offering, will no longer be updated

- Performance Toolkit may give incorrect results as the underlying rmfpms agent evolves

- Support for the Linux rmfpms agent has already been withdrawn, but continues to be available on an as-is basis

© 2013 IBM Corporation

HMC non-ensemble z/VM System Management
z/VM Statement of Direction: Withdrawal of existing function

- z/VM V6.2 is the last release of z/VM that will be supported by the non-ensemble z/VM System Management functions of the System z10 and later

- IBM intends that z/VM V6 virtual server management will continue to be supported using the zEnterprise Unified Resource Manager on the System z196 and later.

© 2013 IBM Corporation

TCP/IP Devices and Daemons
z/VM Statement of Direction: Withdrawal

- A220 HYPERchannel devices
- CLAW devices
- DHCP daemon
- LPSERVE (LPD)
  - RSCS LPD is provided at no charge
  - Does not affect LPR (client)

© 2013 IBM Corporation

Cross System Extensions (CSE)
z/VM Statement of Direction: Withdrawal

- The z/VM Single System Image (VMSSI) feature replaces the functions provided by CSE:
  - Logon once in the cluster, with exceptions
  - Cross-system MESSAGE and QUERY commands
  - Shared spool
  - Shared source directory

- VMSSI has additional value such as autonomic minidisk cache management and a single point of maintenance

- XLINK shared disk support is not affected

© 2013 IBM Corporation

Support for GDPS/PPRC
z/VM Statement of Direction: New function

- Disk subsystem preemptive HyperSwap
  - Storage controllers will notify host when failure is predicted
  - HyperSwap before I/O errors are generated

- HyperSwap scalability
  - Summary “PPRC Suspend” event notification by storage controller
  - Avoid separate notification for each disk

- Future z/VM release support for an alternate subchannel set in which to place PPRC secondary devices

© 2013 IBM Corporation

Summary

© 2013 IBM Corporation

Leadership

z/VM continues to provide additional value to the platform as the strategic virtualization solution for System z

Innovation

z/VM 6.2 introduced horizontal scalability and guest mobility through Single System Image clustering and Live Guest Relocation with RAS in the forefront of the design

Growth

z/VM 6.3 increases the vertical scalability and efficiency to complement the horizontal scaling introduced in z/VM 6.2, because we know our customers' systems continue to grow

© 2013 IBM Corporation

Thanks!!

Contact Information:

Dr. Manfred Gnirss
gnirss@de.ibm.com

© 2013 IBM Corporation