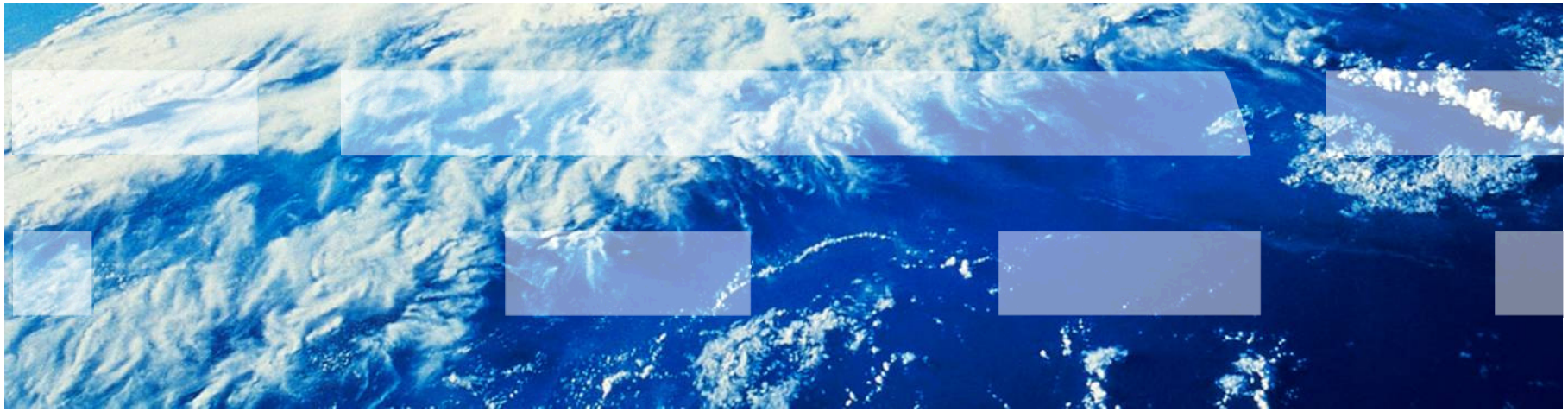
z/VM Platform Update



Brian W. Hugenbruch, CISSP
z/VM Security Architect
bwhugen@us.ibm.com

Version 7

Trademarks

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

IBM*	System z10*	System z196
IBM Logo*	Tivoli*	System z114
DB2*	z10 BC	
Dynamic Infrastructure*	z9*	
GDPS*	z/OS*	
HiperSockets	z/VM*	
Parallel Sysplex*	z/VSE	
RACF*	zEnterprise*	
System z*		

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

OpenSolaris, Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. 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.
INFINIBAND, InfiniBand Trade Association and the INFINIBAND design marks are trademarks and/or service marks of the INFINIBAND Trade Association.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

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.

Acknowledgments – Platform Update Team

- Alan Altmark
- Bill Bitner
- Miguel Delapaz
- Glenda Ford
- John Franciscovich
- Les Geer
- Susan Greenlee
- Dan Griffith
- Brian Hugenbruch
- Romney White

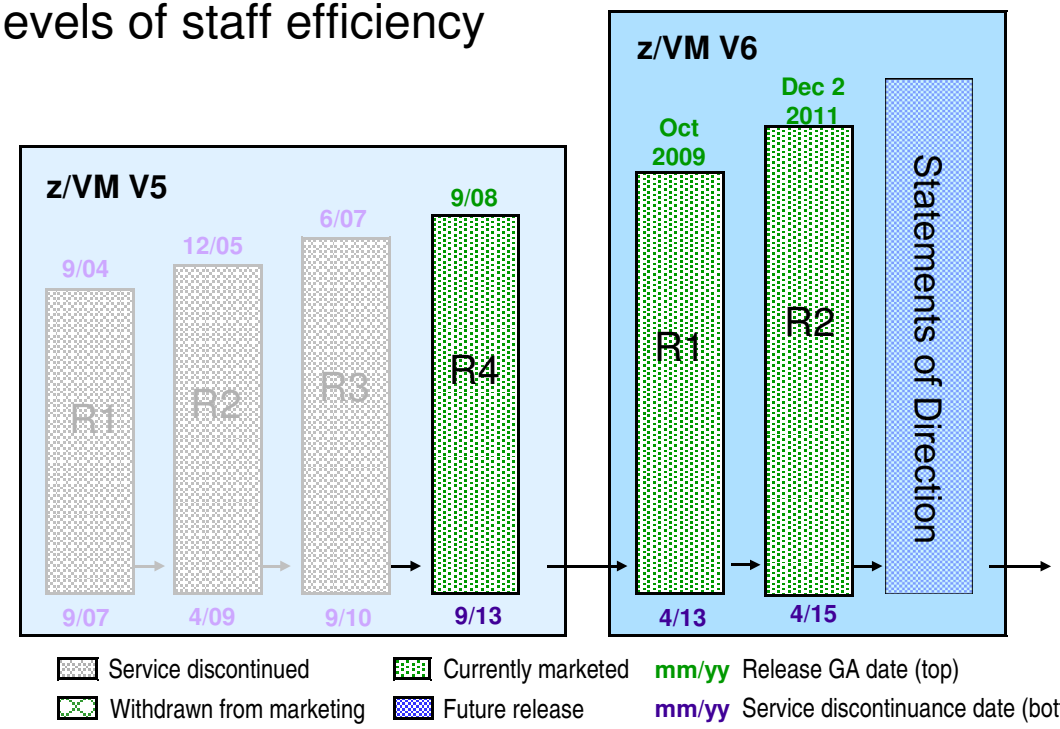
Agenda

- z/VM Timeline
- A word about z/VM V5.4...
- Introducing z/VM V6.2
- The Future: IBM Statements of Direction

z/VM Release Status

z/VM: helping clients "do more with less"

Higher core-to-core consolidation ratios
Higher levels of resource sharing and utilization
Higher levels of staff efficiency



Future enhancements:

- **Hardware and I/O**
- **Ensemble management**
- **z/OS and Linux synergy**
- **Single System Image**
- **Virtualization**
- **Networking**
- **Security**

IBM received EAL 4+ certification of z/VM V5.3 from the German Federal Office of Information Security (Bundesamt für Sicherheit in der Informationstechnik) for conformance to the Controlled Access and Labeled Security protection profiles (CAPP and LSPP) of the Common Criteria standard for IT security, ISO/IEC 15408. z/VM V6.1 is currently undergoing evaluation against OSPP with the labeled security extension at EAL 4+.

z/VM Version 5 Release 4
System z9 and older

- End of Service for z/VM V5.4 is September 30, 2013
- z/VM V5.4 and z/VM V6 are available concurrently
- Clients with System z9 or prior generations should acquire z/VM V5.4 now
  - Excellent time to also look at moving to newer processor technology with recent z114 availability.
    - z114 Servers are fast enough to provide the equivalent of six z9 EC IFLs for most workloads
    - Need to validate processor and memory requirements for these migrations.

z/VM Version 6 Release 1
Security Certification Plans

- IBM intends to evaluate z/VM 6.1 under Common Criteria (ISO/IEC 15408)
  - Statement of Direction issued 22 July 2010
  - **Evaluation in progress (BSI-DSZ-CC-0752)**
  - Security Target: Operating System Protection Profile (OSPP) at EAL 4+
    - Virtualization extension
    - Labeled Security extension
  - RACFVM and z/VM SSL must be configured
  - *z/VM Secure Configuration Guide* will be updated

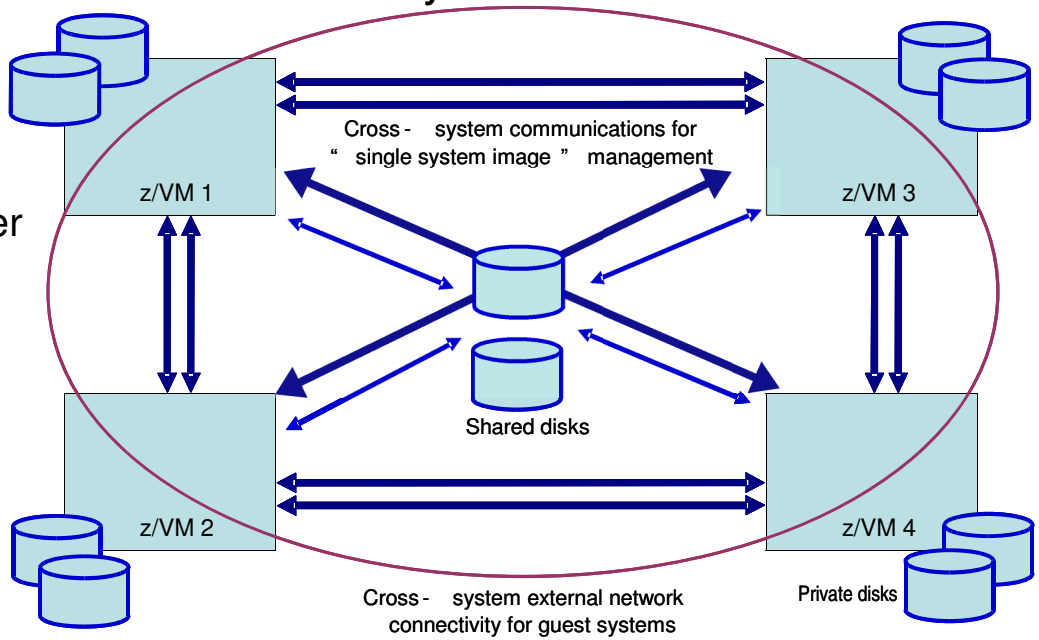
- Federal Information Protection Standard (FIPS) 140-2
  - z/VM 6.1 SSL is designed to support FIPS mode
  - Enablement for both server and certificate database
  - Validation of AES ciphers complete:
    - http://csrc.nist.gov/groups/STM/cavp/documents/aes/aesval.html#1712

z/VM Version 6 Release 2

- Announced **October 12, 2011**
- z/VM V6.2 may be ordered on **November 29, 2011**
  - z/VM V6.1 will be withdrawn when V6.2 becomes orderable
  - If order is placed prior to this date, z/VM V6.1 will be shipped
- Generally available **December 2, 2011**
- End of service **April 30, 2015**
- Major changes include:
  - Single System Image
  - Live Guest Relocation
  - Turnkey support for Unified Resource Manager

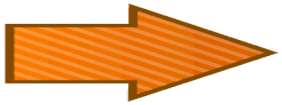
Single System Image Feature
Clustered Hypervisor with Live Guest Relocation

- Provided as an optional priced feature.
- Connect up to four z/VM systems as members of a Single System Image (SSI) 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 systems 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



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
- Complements existing HA solutions
  - Not designed to replace them
- 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!

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

Single System Image Feature
Clustered Hypervisor with Live Guest Relocation

- Unified Resource Manager does not support SSI and LGR
- IBM Director does not support SSI and LGR
- Suggested best practice is to not combine SSI and LGR with the above offerings
  - 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.

z/VM Single System Image and Live Guest Relocation Implementation Services

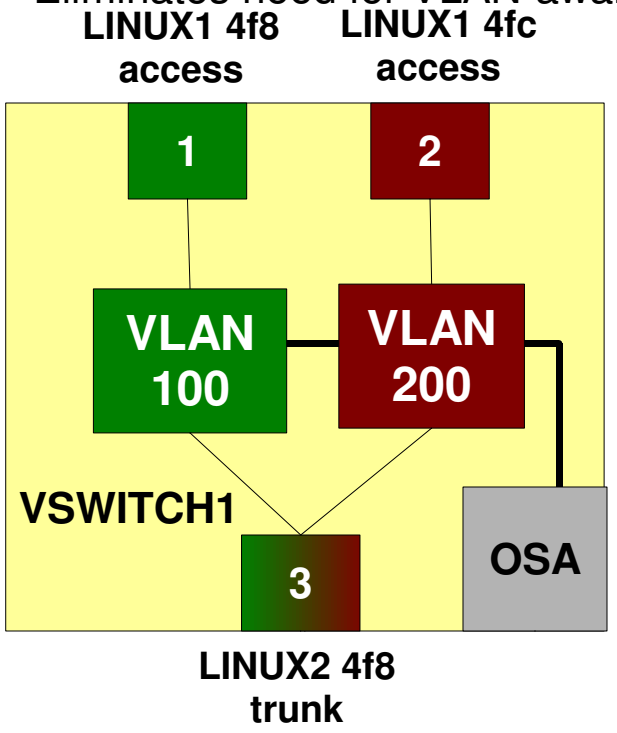
IBM System z Lab Services Offering:

- In-depth education on the functions of VMSSI
- Cluster planning and deployment assistance
- Operational guidance and recommendations
- Migration assistance for users of CSE
- Demonstrate the technology in your own environment.
- Help you create system configuration files
- Analyze how SSI and LGR will affect your system initialization, recovery, and automation procedures
- Early identification of any inhibitors to use
- Identification of any required z/VM or Linux operating system patches

For more information, contact **systemz@us.ibm.com**

VSWITCH: Multiple access ports per guest

- One or more virtual ports on a VSWITCH are reserved for a guest
- Ports are associated with a VLAN – implicit authorization (exc. RACF)
- Authorization changes take effect immediately
- Eliminates need for VLAN-aware guests




```
define vswitch vswitch1 portbased vlan aware native none
set vswitch vswitch1 portnumber 1 userid LINUX1
set vswitch vswitch1 portnumber 2 userid LINUX1
set vswitch vswitch1 portnumber 3 userid LINUX2 porttype trunk
set vswitch vswitch1 vlanid 100 add 1 3
set vswitch vswitch1 vlanid 200 add 2 3

USER1:
Couple 4f8 to system vswitch1 portnum 1
Couple 4fc to system vswitch1 portnum 2

USER2:
Couple 4f8 to system vswitch1 [portnum 3]
```

Switch port number not available on NICDEF. Use `COMMAND COUPLE` in the directory.

Scalability and Performance Enhancements
Available by PTF to prior releases where shown

- Reduction of memory and CPU resources required to manage larger memory sizes
- Control of the guest page re-ordering process, improving the performance characteristics of guests with large memory footprints (VM64774)
- Reduced system overhead of guest page release function, thereby helping to increase guest throughput (VM64715)
- Improved contiguous frame coalescing algorithms help to increase system throughput (VM64795)

Scalability and Performance Enhancements
Available by PTF to prior releases where shown

- More accurate scheduling algorithm for guests that have LIMITHARD shares (VM64721)
- Reduce LPAR suspend time by reducing the number of DIAGNOSE 0x9C and 0x44 instructions issued when obtaining system locks (VM64927 for z/VM 6.1 only)
- Improve workload dispatch algorithm to eliminate erratic virtual machine pause in busy systems with more than 14:1 total virtual to logical CPU over-commitment (VM64887)

Advances in Processor Performance

- The CPU Measurement Facility is a System z hardware facility that characterizes the performance of the CPU and nest:
  - Instructions, cycles, cache misses, and other processor related information
  - Available on z10 EC/BC, z196, z114
- IBM will be using data from this facility to influence future processor design and benchmark validation of those designs.
- Will also increase accuracy of future processor capacity sizing tools
- To assist, by providing sample Monwrite data containing the counters, please contact Richard Lewis (rflewis@us.ibm.com)

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

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
```

Access controls for dedicated or attached devices

- The CP ATTACH and GIVE commands, as well as the DEDICATE statements in the directory will now engage ESM access controls
- Integrated ASCII console on the HMC is also managed
- Full discretionary and mandatory access controls
- RACF support included

Mandatory access controls for virtual consoles

- SET SECUSER and SET OBSERVER are now available when mandatory access controls (security labels) are active.
- Virtual security zones (“color coding” of users and resources) can now co-exist with system automation functions.
- Also applies to the user ID specified on CONSOLE directory statement.
- Users in different zones cannot see or manage each others' virtual console
  - Console cannot be given
  - Console cannot be taken
  - System administrators and automation solutions can use label SYSNONE to allow them access to all consoles

RACF Security Server

- Single System Image
  - Automatic propagation of most RACF commands
- Protected Users
  - User without a password or password phrase will not be revoked due to too many invalid password attempts or inactivity
- Real device protection
  - ATTACH, GIVE, DEDICATE
  - New VMDEV class
  - Profiles: RDEV.*device.system_id*
    - Qualified by system ID in order to accommodate shared database across CECs
    - Device “SYSASCII” used for HMC integrated ASCII console
- Support for Diagnose 0xA0 Subcode 0x48
  - Obtain information about installed ESM (not just RACF!) in architected format

RACF Security Server

- High Level Assembler no longer required for most common customizations
- ALTER (MW) access for VMMDISK no longer conveys the ability to change the access list for the minidisk

- RPIDIRCT updates:
  - Create VMLAN profiles from NICDEF statements
  - Create VMDEV profiles from DEDICATE statements
  - Recognize IDENTITY and SUBCONFIG definitions
  - Passwords AUTOONLY, LBYONLY, and NOPASS cause user to be Protected
  - Password NOLOG causes user to be revoked unless required for POSIX
    - POSIX users will be Protected

LDAP Server Upgrade

- z/OS R12 level
- Management and change logging of general resources
- Password management policy support to improve LDAP authentication from open systems such as Linux
  - Expiry warnings
  - Interactive password change when password has expired
  - Password rule validation

Additional z/OS R12 Equivalency Upgrades

- Language Environment (LE) runtime libraries
- MPROUTE
- Program Management Binder
  - COMPAT supports ZOSV1R10, ZOSV1R11, ZOSV1R12
  - New suboptions on RMODE
  - Compiler parameters can be read from IEWPARMS DDNAME
  - New C/C++ API
- Support for **IBM XL C/C++ Compiler for z/VM, V1.3** (5654-A22)
  - Details can be found in US announcement letter 211-369
- System SSL

z/CMS

- Previously shipped with z/VM as a sample program, now supported as an optional CMS
  - IPL ZCMS
- Enables CMS programs to use z/Architecture instructions and 64-bit registers
- Existing ESA/390 architecture programs continue to run unchanged
  - CMS not exploit memory above 2 GB
  - CMS does not provide 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 architectural support for XC mode
  - VM Data Spaces not available

Installation Improvements

- Significant changes to system layout to support Single System Image
  - "This isn't your grandfather's z/VM!"
- 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
  - Installation initiated with a single command
- All DASD volumes can be labeled at installation time, including the system residence volume
- Turnkey support for zEnterprise ensembles
  - Enable clients new to z/VM to get started with Unified Resource Manager
  - Those who purchase DIRMAINT or another directory manager, or who require an external security manager, need to perform manual enablement
    - Decline this option during installation

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 the platform agent for IBM Systems Director for Linux on System z, previously shipped with z/VM V6.1

Previously shipped Functional Enhancements Included in z/VM V6.2

- XRC timestamps
- Hyperswap improvements
- SSL Server Reliability and Scalability
- zEnterprise Unified Resource Manager
- CPU Measurement Counter Facility Host support
- zEnterprise Unified Resource Manager

APAR numbers shown apply to z/VM 6.1 and z/VM 5.4 unless otherwise stated

XRC Timestamps

VM64814 and VM64816

- CP will sync with STP at IPL and, optionally, obtain time zone and leap seconds from STP
  - No need to deactivate/activate LPAR
- Correct time will be placed in all host and guest I/O
  - CP will monitor STP time signals
- Enabled via SYSTEM CONFIG with option to skip timestamp or delay I/O if CP is unable to sync with STP
- No virtualization of STP
  - Option for 2nd level systems to stamp I/O without use of STP

Hyperswap Improvements

VM64815 and VM64816

- CP HYPERSWAP command now has additional controls for missing interrupt handling
  - Do not trigger automatic quiesce (default)
    - GDPS will not be notified
  - Trigger automatic quiesce after specified number of MI detection intervals
    - GDPS will be notified

- Better management of PAV and HyperPAV devices

- Avoid unnecessary hyperswaps due to normal maintenance activities
  - Concurrent storage controller upgrade

- New wait state 9060 if abend occurs when Hyperswap is in progress
  - no checkpoint taken, no automatic dump
  - restart dump if dedicated dump volume, else standalone dump

SSL Server Reliability and Scalability

PK97437, PK97438, PK75662

- Major rewrite
- Multiple SSL servers with ‘resume’ cache manager and shared database
  - Can balance total number of sessions against number of sessions per server
- Significant performance improvements
  - Interactive workloads such as telnet
  - Session establishment costs, particularly during mass ‘reconnect’
- Updates to TCP/IP stack, as well
- Migration required
  - http://www.vm.ibm.com/related/tcpip/tcsslspe.html

CPU Measurement Facility Counters – Host Support
VM64961

- Sets of counters for each logical processor that count events such as cycle, instruction, and cache directory-write counts
  - Same COUNTER information as z/OS partitions
- Accumulation is a relatively low-overhead activity and is performed automatically by the machine when the counters are authorized, enabled, and activated
- Authorization controlled by a logical partition's Security settings in its activation profile
- Enablement, activation, and data collection controlled by z/VM MONITOR command

zEnterprise Unified Resource Manager

VM64822, VM64904, VM64917, VM64956, VM64957

- z/VM V6 only
- Turnkey installation option to enable virtual server management via zEnterprise Unified Resource Manager (z/VM V6.2 only)
- Enables Unified Resource Manager to perform system and virtual server management tasks
  - Virtual server configuration
  - Disk storage management
  - Virtual network management
  - Performance monitoring
- CP, CMS, LE, TCP/IP, DIRMAINT, Performance Toolkit, HCD
- http://www.vm.ibm.com/service/vmrequrm.html

zEnterprise Unified Resource Manager
Ensemble Membership

- If configured to participate in an ensemble, z/VM will automatically join the ensemble at IPL

- Configuration tasks
  - Set up OSM and OSX channel paths
  - Set up controllers for IEDN and INMN networks
    - Pre-defined controllers DTCENS1 and DTCENS2 for exclusive use by ensemble networks
    - DTCENS1 automatically creates a VSWITCH to provide SMAPI connectivity to INMN network
  - Configure directory manager (REQUIRED)
  - Configure SMAPI servers

- See chapter "Configuring z/VM for an Ensemble" in CP Planning and Administration manual

Statements of Direction

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

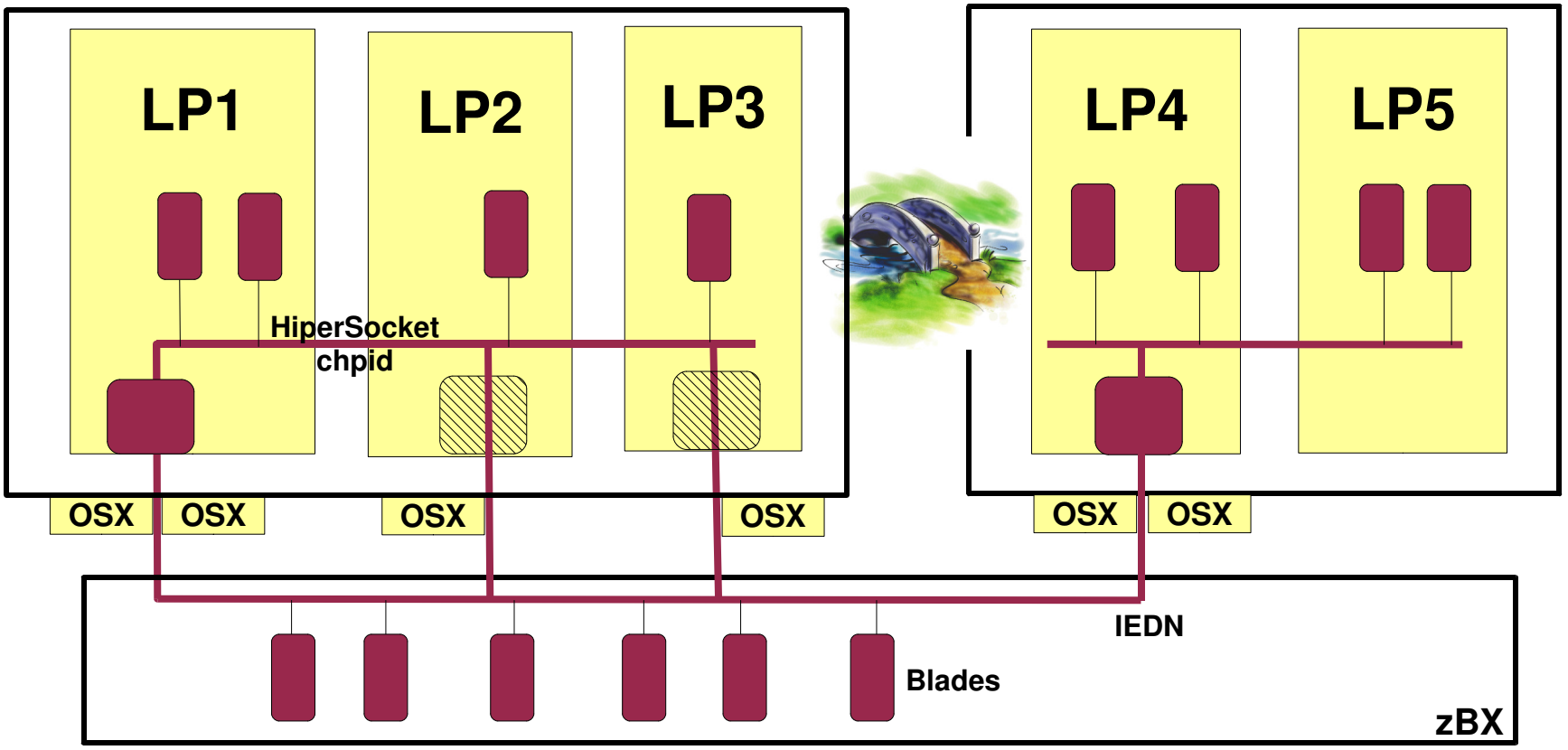
Note for withdrawals: Unless otherwise stated, it is IBM's intent that z/VM V6.2 will be the last release of z/VM to support the indicated function.

HiperSocket VSWITCH Integration with zEnterprise IEDN
z/VM Statement of Direction: New function

- Virtual Switch bridge between Ethernet LAN and HiperSockets
  - zEnterprise IEDN (OSX) 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

HiperSocket VSWITCH Integration with zEnterprise IEDN
z/VM Statement of Direction: New function



- z/VM guest only
- Built-in failover and failback
- Special IOCP definition will be required

- Same or different LPAR
- One active bridge per CEC
- PMTU simulation

HiperSockets Completion Queues
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

High Performance FICON
z/VM Statement of Direction: New function

- Enable guests to use High Performance FICON for System z (zHPF)
  - Different I/O model
  - Single and multiple track I/O
- Requires host and control unit compatibility
  - Consult a storage specialist for details
- z/OS and Linux provide exploitation

IBM Performance Toolkit for VM: 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

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, z196 and z114

- z/VM virtual server management will continue to be supported using the zEnterprise Unified Resource Manager on the z196 and later

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

- A220 HYPERchannel devices
- CLAW devices
- DHCP daemon
- LPSERVE (LPD)
  - RSCS LPD will continue to be provided at no charge
  - Does not affect LPR

User Class Restructure and OVERRIDE utility
z/VM Statement of Direction: Withdrawal

- User Class Restructure (UCR) was first introduced in VM/SP Release 6 to allow changes to the privilege classes associated with CP commands and DIAGNOSE subcodes.

- OVERRIDE utility was a “compiler” used to create special UCR-type files in the spool

- Function was replaced by MODIFY COMMAND capability in VM/ESA
  - Use the CP MODIFY COMMAND command or SYSTEM CONFIG statement

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
  - Cross-system LINK (XLINK)
  - Shared spool
  - Shared source directory

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

Support for GDPS/PPRC 3.8
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 to place PPRC secondary devices

References

- z/VM Home Page: http://www.ibm.com/vm/
- z/VM Version 6 Release 2 Resources: http://www.ibm.com/vm/zvm620/
  - Includes links to announce letters
- z/VM Single System Image Overview: http://www.ibm.com/vm/ssi/
- Information on service required for z/VM when used managed by Unified Resource Manager: http://www.ibm.com/vm/service/vmrequrm.html

Thanks!

Contact Information:

Brian W. Hugenbruch, CISSP
z/VM Security Design and Development
bwhugen@us.ibm.com
+1 607.429.3660