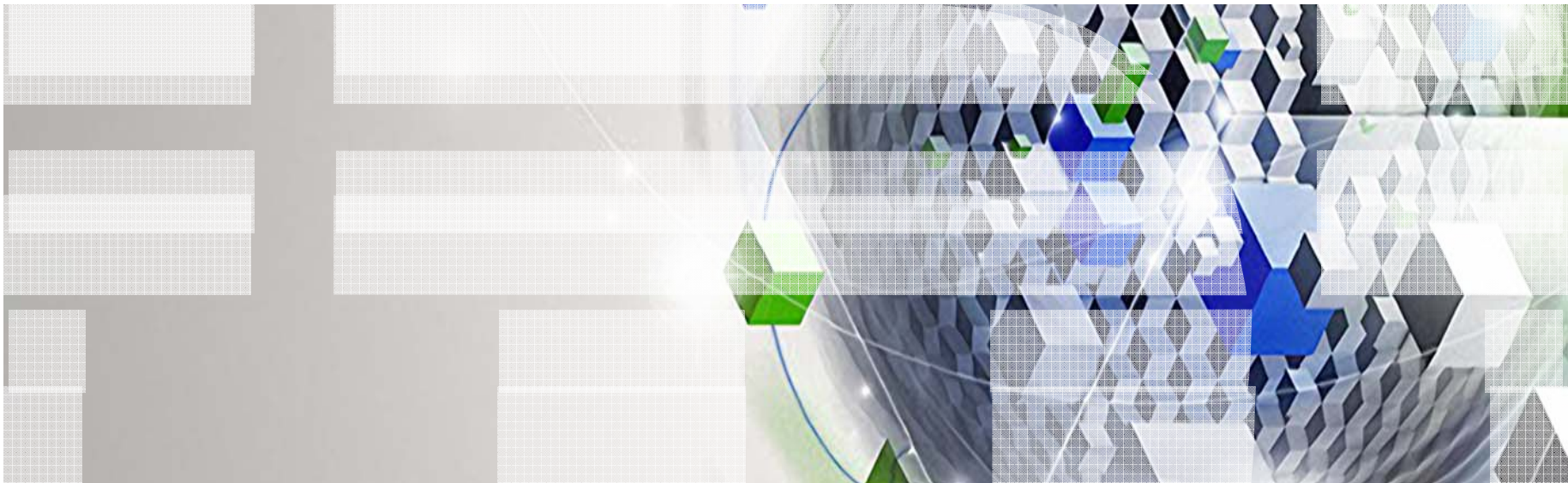


z/VM Update

Romney White
System z Architecture and Technology
romneyw@us.ibm.com



Agenda

- **z/VM 6.3 Themes**

- **z/VM 6.3 Content**
 - **Processor**
 - **Memory**
 - **I/O**
 - **Network**
 - **Install and Service**
 - **Systems Management**
 - **Performance**
 - **Miscellaneous**

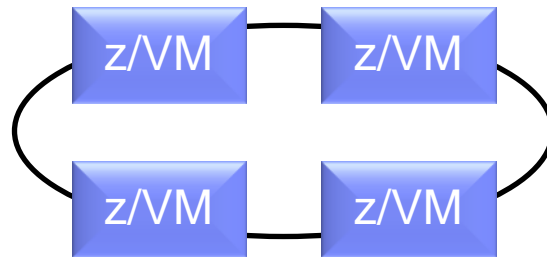
- **Q&A**

z/VM 6.3 – Making Room to Grow Your Business

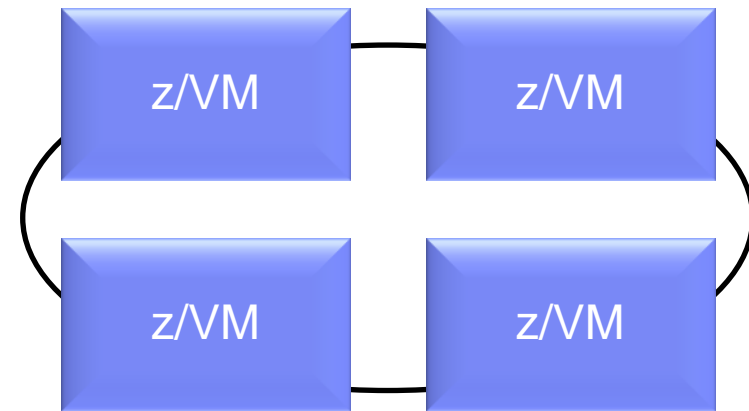
- **Preview Announcement introducing z/VM 6.3 on February 5, 2013**
- **Full Announcement on July 23, 2013**
- **General Availability on July 26, 2013**



z/VM 6.1



z/VM 6.2



z/VM 6.3

z/VM 6.3 Themes

- **Reduce the number of z/VM systems you need to manage**
 - Expand z/VM systems constrained by memory up to four times
 - Increase the number of Linux virtual servers in a single z/VM system
 - Exploit HiperDispatch to improve processor efficiency
 - Allow more work to be done per IFL
 - Support more virtual servers per IFL
 - Expand real memory available in a Single System Image Cluster to 4 TB

- **Improved memory management flexibility and efficiency**
 - Benefits for z/VM systems of all memory sizes
 - More effective prioritization of virtual server use of real memory
 - Improved management of memory on systems with diverse virtual server processor and memory use patterns

z/VM 6.3 Content

- **Processor**
 - HiperDispatch

- **Memory**
 - Memory Scalability
 - Large Memory Dump support
 - IPL Changes for NSS in Linux Dump

- **I/O**
 - MSS and DS8K Synergy (stage 1)
 - FCP Data Router support

z/VM 6.3 Content ...

▪ Network

- Live Guest Relocation support for Port Based Virtual Switches
- Virtual Switch VEPA support
- HiperSockets Completion Queue Guest Exploitation
- Virtual Switch Recovery Stall Prevention
- Manual Virtual Switch Recovery
- IPv6 support for SSL
- Uplevel MPRROUTE to z/OS 1.13
- A220/CLAW/DHCP/LPD Removal
- Support TLS V1.2

▪ Install and Service

- Installation Upgrade in Place
- Highest Level Part Handling for Multiple Releases

z/VM 6.3 Content ...

- **Systems Management**
 - **SMAPI enhancements**
 - **Disable CSE**
 - **OVERRIDE utility removal**
 - **OpenStack enablement**
 - **xCAT appliance integration**

- **Performance Toolkit**
 - **Memory Scalability**
 - **HiperDispatch**
 - **Guest Fibre-Channel eXtensions (FCX)**
 - **HiperSockets Bridge**
 - **Live Guest Relocation**

z/VM 6.3 Content ...

- **Miscellaneous**
 - Uplevel DFSMS Binder to z/OS 1.13
 - RACF CMS 27 and LE enablement

Processor

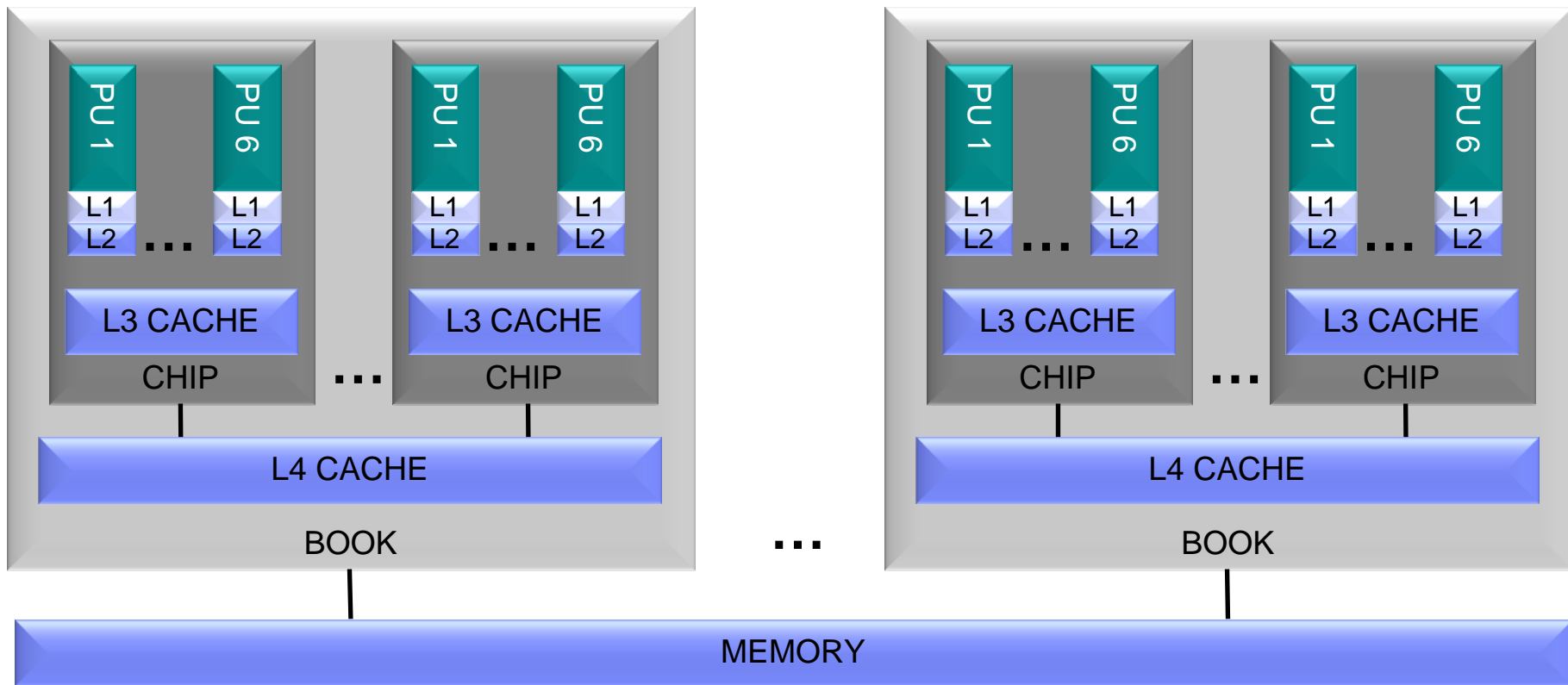
HiperDispatch

- **Improves processor efficiency**
 - **Better n-way curves**
 - **Supported processor limit remains at 32**
 - **Better use of processor cache to take advantage of cache-rich system design**

- **Two components**
 - **Dispatching Affinity: dispatching cognizant of processor cache topology**
 - **Vertical CPU Management: cooperation with PR/SM to distribute physical processor resources to logical processors more efficiently for some configurations**

HiperDispatch: Dispatching Affinity

- Processor cache structures increasingly complex and critical to performance
- Goal is to re-dispatch work close (in terms of topology) to where it last ran



HiperDispatch: Dispatching Affinity

- **Dispatcher changed to be aware of cache topology and dispatch work accordingly**
 - **Attempt to dispatch virtual CPU near cache where its data may be, based on where it was last dispatched**

- **Potentially increases cache efficiency, lowering processor costs by reducing CPI (Cycles Per Instruction)**

- **Previously, z/VM used soft affinity to processor when dispatching virtual CPUs**
 - **No awareness of chip or book**

HiperDispatch: Vertical CPU Management

- **Attempts to align system workload processor resource requirements to a minimal number of logical processors to improve efficiencies of system management and cache use**

- **With current horizontal management, weight of particular processor type is evenly distributed across the logical processors of the z/VM LPAR**

- **Example**
 - **10 Physical IFLs**
 - **LPAR A: 8 Logical IFLs, weight of 400 out of 1000**
 - **LPAR A's entitlement is 4 IFLs, so 8 is more than necessary if LPAR A only receives its entitlement**
 - **z/VM & LPAR can cooperate**
 - **z/VM concentrates the workload on a smaller number of logical processors**
 - **LPAR redistributes the partition weight to give a greater portion to each of the (smaller number of) logical processors**

Memory

Large Memory Support

- **Increase the real memory limit from 256GB to 1TB**
 - Proportionately increases total virtual memory based on tolerable over commitment levels and workload dependencies
 - Individual virtual machines limit unchanged at 1TB

- **Paging DASD utilization and requirements change**
 - Proactive writing of pages to DASD increases need to have properly configured paging subsystem
 - Removed the need to provide double the paging space on DASD
 - Some additional space will continue to be recommended to avoid problems.

- **Expanded Storage continues to be supported with current limit of 128GB**

Large Memory Support

- **Page selection algorithms rewritten**
 - Reorder processing removed

- **Improved effectiveness of the CP SET RESERVE command**
 - Reserved pages protected better than in previous releases
 - Support for reserving pages of NSS or DCSS space
 - E.g., Monitor Segment (MONDCSS)
 - Ability to limit the total number of reserved pages in the system

Large Memory Dump

- **Ensure that customers can obtain dumps of z/VM systems that exploit larger real memory**
 - **Extended z/VM to create dumps of real memory configurations up to 1 TB**
- **New Stand-Alone Dump utility can dump to either ECKD or SCSI devices**
- **Enhancements made to speed up hardabend dump processing**
 - **Reduce the effect of using the SNAPDUMP command on production systems**
- **No enhancements to existing VM Stand-Alone Dump Utility, VMDUMP, and DUMPLOAD**
 - **VM Stand-Alone Dump Utility supports storage sizes up to 512 GB**
 - **VMDUMP dumps virtual machines and is not recommended for dumping large amounts of memory**
 - **DUMPLD2 (DUMPLOAD replacement) can load dump to multiple files/devices**

IPL Changes to Retain NSS

- **Linux users can share the kernel via NSS**
 - Requires a specific setup in Linux
 - All sharing Linux systems at same kernel level

- **NSS released when guest IPLs**

- **Linux has its own tools to create a stand-alone dump**
 - Linux standalone dump tool IPLed in virtual machine
 - NSS not included in dump because it is released by IPL
 - NSS contains kernel code that is often needed for debug purposes

- **Enhancement allows stand-alone dump tools to include NSS contents in the dump**
 - CP IPL command enhanced with new NSSDATA option
 - Changed specifically to support Linux standalone dump

I/O

MSS and DS8K Synergy (Stage 1)

- **Support new disk subsystem capabilities**
 - Preemptive HyperSwap notifications to improve Hyperswap reliability
 - Summary event notifications to improve PPRC suspend scalability

- **Allow PPRC secondary volumes to reside in an alternate subchannel set**
 - Complement z/OS alternate subchannel set support
 - Simplify I/O subsystem configuration

- **Provide related infrastructure for exploitation by GDPS 3.8**

FCP Data Router Support

- **Initial FCP adapter design used store-and-forward approach**
 - Intermediate buffer in adapter
 - Limited throughput

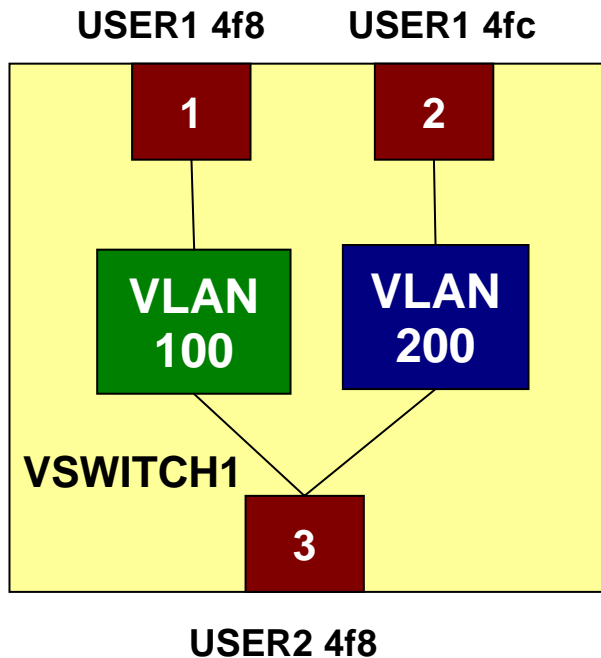
- **zEnterprise GA2 introduced DMA capability in FCP adapter**
 - Optional feature enabled by guest
 - Improves throughput by eliminating intermediate buffer use
 - Allows errors to be reported after initial acceptance of request
 - Enables multiple I/Os to be acknowledged in one response

- **Enabled for guest exploitation**

Network

Port-Based Virtual Switch LGR Support

- Logical ports on a VSWITCH are reserved for a guest (one or more)
- Each port associated with a VLAN
- 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 USER1
set vswitch vswitch1 portnumber 2 userid USER1
set vswitch vswitch1 portnumber 3 userid USER2 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]
```

Virtual NIC Relocation Eligibility Requirements

- **Improved eligibility checks for Virtual Switch**
 - **Additional attributes now checked to make sure source and target match**
 - **CP Authorizations from source will be set on target if missing or not the same**
 - **Authorization to access Virtual Switch**
 - **VLAN attributes on target**
 - **OSDSIM authorization**
 - **MACPROTECT setting for virtual NIC**
 - **PROMISCUOUS setting**

Virtual NIC Relocation Eligibility Requirements

- **Requirements checked**
 - **Same Virtual Switch name**
 - **Same Virtual Switch type (QDIO, IEDN, INMN)**
 - **Guests coupled to IEDN or INMN VSWITCH not eligible for relocation**
 - **Uplink ports must be operational and have same EQID**
 - **Same Virtual Switch protocol (ETHERNET or IP)**
 - **Same user-based or port-based attribute**
 - **Same VLAN attributes**
 - **VLAN aware or unaware**
 - **Native VLAN ID**
 - **Default VLAN ID**
 - **Same isolation setting (ON or OFF)**
 - **For port-based Virtual Switch, same USERID and PORT number definitions**
 - **Same OSDSIM setting**
 - **Same VLAN port type (ACCESS or TRUNK)**

Virtual Switch Recovery Stall Prevention

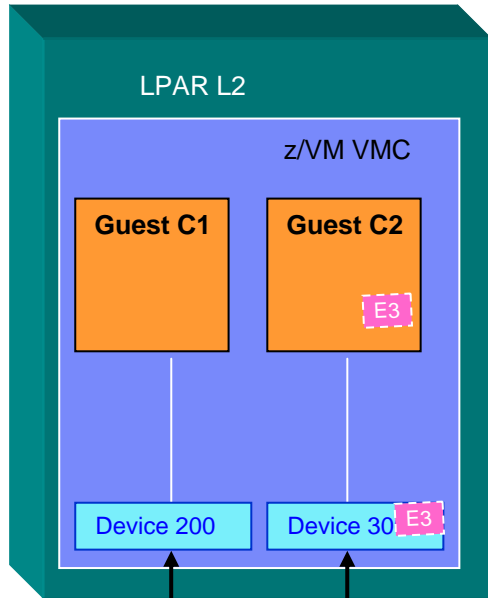
- **Virtual Switch controller can become hung trying to clear failure condition and failover to another uplink device**
 - **CP enhanced to try to clear conditions causing a hang so that normal failover processing can continue**
 - **Exploits z/VM Missing Interrupt Handler selectively based on CCW command code**
 - **CP takes measures to force failover to an alternate controller in extreme cases**
 - **Primary controller will remain hung**
 - **Alternate controller will accomplish failover to a backup uplink device**

Virtual Switch VEPA Support

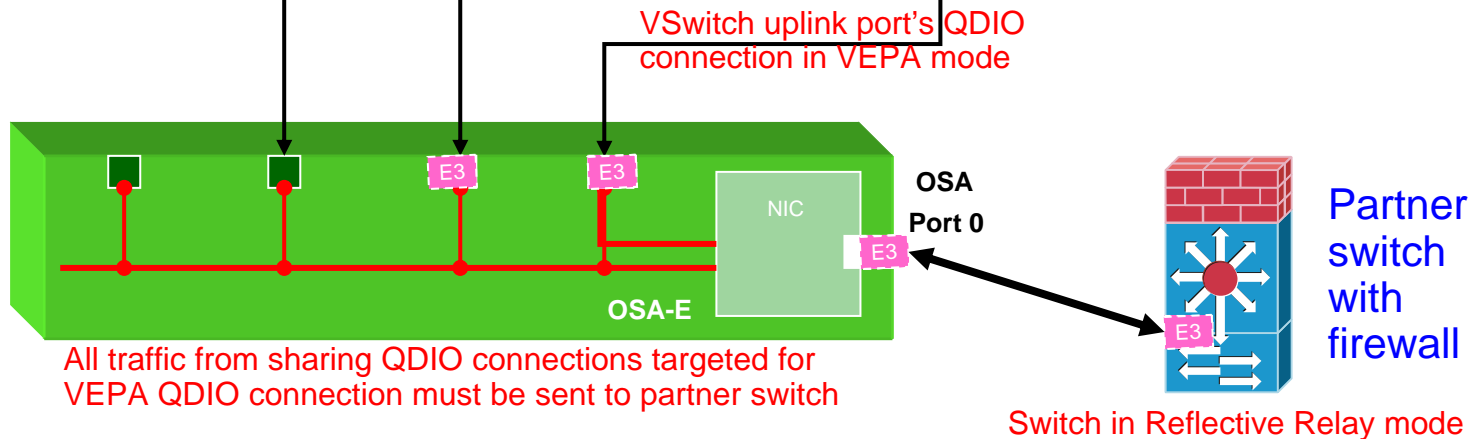
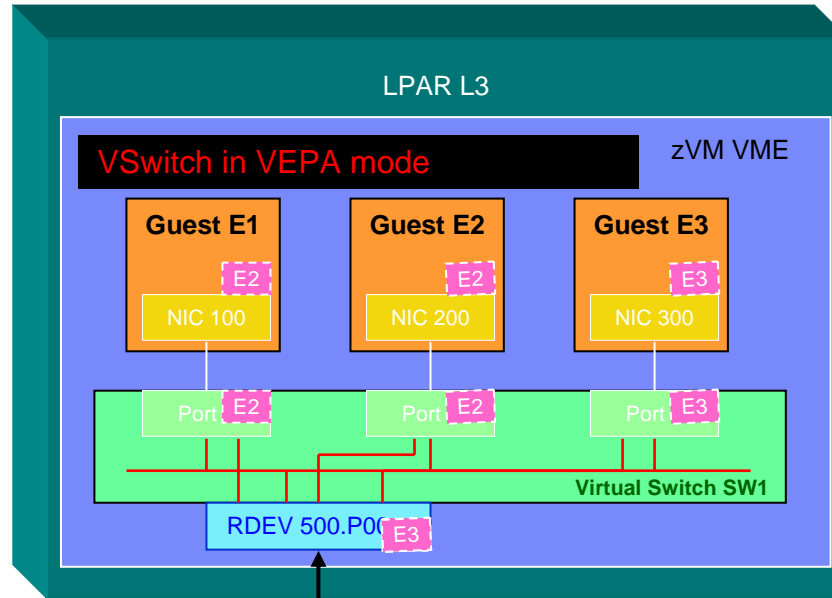
- **Implemented in the z/VM Virtual Switch and the OSA feature**
 - **Based on existing OSA QDIO Connection Isolation**
- **Supports both IEDN and customer external (QDIO) Ethernet (Layer 2) networks**
- **VEPA mode instantiated through the SET VSWITCH command (VEPA option)**
 - **Initiates a VEPA mode QDIO connection between the z/VM VSWITCH and its OSA uplink port**
 - **Initiates a VEPA mode connection between the partner switch and the OSA**
- **VEPA support requires a network switch that supports Reflective Relay (VEPA) and is exclusive to the IBM zEC12 processor and beyond**

Virtual Switch VEPA Support

OSA LPAR to LPAR communications.



VSwitch guest port communications



HiperSockets Completion Queue Guest Exploitation

- **Transfer HiperSockets messages asynchronously**
- **Used when traditional synchronous queues are full**
- **Automatic enablement; no z/VM configuration required**
- **Helpful when traffic is “bursty”**
- **z/VM 6.2 support provided for z/VM Virtual Switch**
 - **Support for guest exploitation added**

Manual Virtual Switch Recovery

- **New SET VSWITCH UPLINK SWITCHOVER CP command**
 - Forces fast failover to the backup uplink device
 - Use when primary uplink device needs to be made available for hardware maintenance
 - Link aggregation (GROUP) uplink ports not supported
 - SET PORT GROUP LEAVE/JOIN is preferred method for Link Aggregation
 - Virtual Switch must have active and connected uplink port
 - NIC uplink port is not supported
 - Virtual Switch with NOUPLINK is not supported
 - At least one backup uplink device must be configured
 - Bridge Ports not supported
 - Use SET VSWITCH BRIDGE DISCON to accomplish failover

TCP/IP

IPv6 Support for SSL

- **Since January 2011, US Federal Government has required IT companies doing business with Government agencies to support IPv6-based communication**
- **In support of this requirement, an IPv6 roadmap was developed for z/VM to outline a schedule for its TCP/IP clients and servers to become IPv6-capable**
- **IPv6 support for the z/VM SSL Server is delivered in z/VM 6.3 in accordance with this roadmap**
 - **Parts of the TCP/IP stack that communicate with the SSL Server updated to support IPv6 addressing**
 - **Servers that currently support secure IPv4 connections (FTP, Telnet, and SMTP) modified to also support secure IPv6 connections**

Uplevel MPROUTE to z/OS 1.13

- **Improve serviceability (comply with z/OS DoU)**

A220/CLAW/DHCP/LPD Removal

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

Support TLS V1.2

- **Enhance SSL server to support TLS 1.2**
 - **Allow clients to adhere to requirements of NIST Special Publications 800-131a and 800-131b**
 - **Prepare for potential requirements of the Common Criteria Operating System Protection Profile (OSPP), the basis of z/VM's current Target of Evaluation**

Installation and Service

Release Migration Procedures

- Migration to z/VM 6.3 can use one of three procedures
 - **Migration** – Documented in the migration section of *z/VM Installation Guide*
 - E.g.,
 - 5.4 or 6.1 → 6.3 non-SSI
 - 5.4 or 6.1 → 6.3 SSI
 - **Use Cases** – Documented in Part 5 of *z/VM CP Planning and Administration*
 - E.g.,
 - Converting a z/VM System to a Single-Member z/VM SSI Cluster
 - Combining Two Non-SSI z/VM Systems to Create a z/VM SSI Cluster
 - **Upgrade in Place** – Described in *z/VM Installation Guide*

Installation Upgrade In Place

- **New INSTUPGR command**
 - Designed to facilitate new shared environment supported by SSI
 - Also supports non-SSI systems

- **Stage 1**
 - Does not affect workload running on the system being upgraded
 - Does not affect the way other members in an SSI cluster are running
 - Does not change the current function of other members in a cluster
 - Does not require an IPL or taking down a virtual machine
 - Can be backed out

- **Stage 2**
 - May affect the workload running on the system being upgraded
 - Can be backed out only by restoring a DDR backup
 - May affect another member in the SSI cluster

Highest Level Part Handling

- **When SSI cluster members run different release or service levels, certain parts are required to be at the highest release level on all members in order to manage shared resources**
 - **User Directory: DIRECTXA MODULE, DISKMAP EXEC, DIRMAP MODULE, RPIDIRECT EXEC**
 - **System-owned DASD: CPFMTXA EXEC, CPFMTXA MODULE**
 - **SSI Persistent Data Record: FORMSSI MODULE**

- **In z/VM 6.2, these parts were moved to a common disk**

- **In z/VM 6.3, new common test/production build disk set used to ship and maintain these parts**
 - **Disks located on common (cluster-wide) volume**
 - **Installation and service programs and files recognize these parts and process them to ensure they are copied to the common disks to prevent back leveling**

Systems Management

SMAPI Enhancements

- **Restructure SMAPI Directory Exit (DMSSIXDM)**
 - Directory exit is 11,000 lines of code and provides more than 70 APIs
 - Makes service updates and new development very difficult
 - Split into several common routines and a separate EXEC for each API

- **New capabilities for existing API functions**
 - **System_Disk_Query**
 - **System_WWPN_Query**
 - **Virtual_Network_Vswitch_Create_Extended**
 - **Virtual_Network_Vswitch_Query_Extended**
 - **Virtual_Network_Vswitch_Set_Extended**

SMAPI Enhancements ...

- **New API functions**
 - **Image_Lock_Query_DM**
 - **Profile_Lock_Query_DM**
 - **Image_MDISK_Link_Query**
 - **SMAPI_Status_Capture**
 - **SMSTATUS** (stand-alone command, not callable as an API)
 - **System_Disk_I/O_Query**
 - **System_EQID_Query**
 - **System_Information_Query**
 - **System_Page_Utilization_Query**
 - **System_Performance_Info_Query**
 - **System_Service_Status_Query**
 - **System_Shutdown**
 - **System_Spool_Utilization_Query**
 - **Network_IP_Interface_Create**
 - **Network_IP_Interface_Modify**
 - **Network_IP_Interface_Query**
 - **Network_IP_Interface_Remove**

Disable CSE

- **CSE (Cross-System Extensions) is pre-SSI VM clustering technology**
 - Replaced by and incompatible with SSI
 - In z/VM 6.2, can choose either CSE or SSI for clustering

- **In z/VM 6.3, all CSE functions are disabled except XLINK**
 - XLINK used to share minidisks across z/VM systems
 - Supports more than four systems
 - Still supported in non-SSI environments

- **External interfaces required for CSE are disabled**
 - Publication references removed

OVERRIDE Utility Removal

- **User Class Restructure (UCR) first introduced in VM/SP Release 6 to allow changes to the privilege classes associated with CP commands and DIAGNOSE codes**
- **OVERRIDE utility is “compiler” used to create special UCR files in the spooling subsystem**
- **Function replaced by MODIFY COMMAND capability in VM/ESA**
 - **Use the CP MODIFY COMMAND command or SYSTEM CONFIG statement**

OpenStack Enablement

- **OpenStack**

- “[...] global collaboration of developers and cloud computing technologists producing the ubiquitous open source cloud computing platform for public and private clouds” (openstack.org)
- **Building z/VM adapter to present OpenStack APIs leveraging z/VM SMAPI**
- **Basis for SmartCloud offerings for System z**

- **Schedule**

- **2012-12-10: Initial code drop to OpenStack community**
- **2013-04-18: Final code drop to OpenStack community for z/VM 6.3 support**

xCAT Appliance Integration

- **eXtreme Cloud Administration Tool**
 - Open source tool to manage, provision, and monitor physical and virtual machines
 - Grew up in HPC environments

- **Available for z/VM today as GTS service offering**

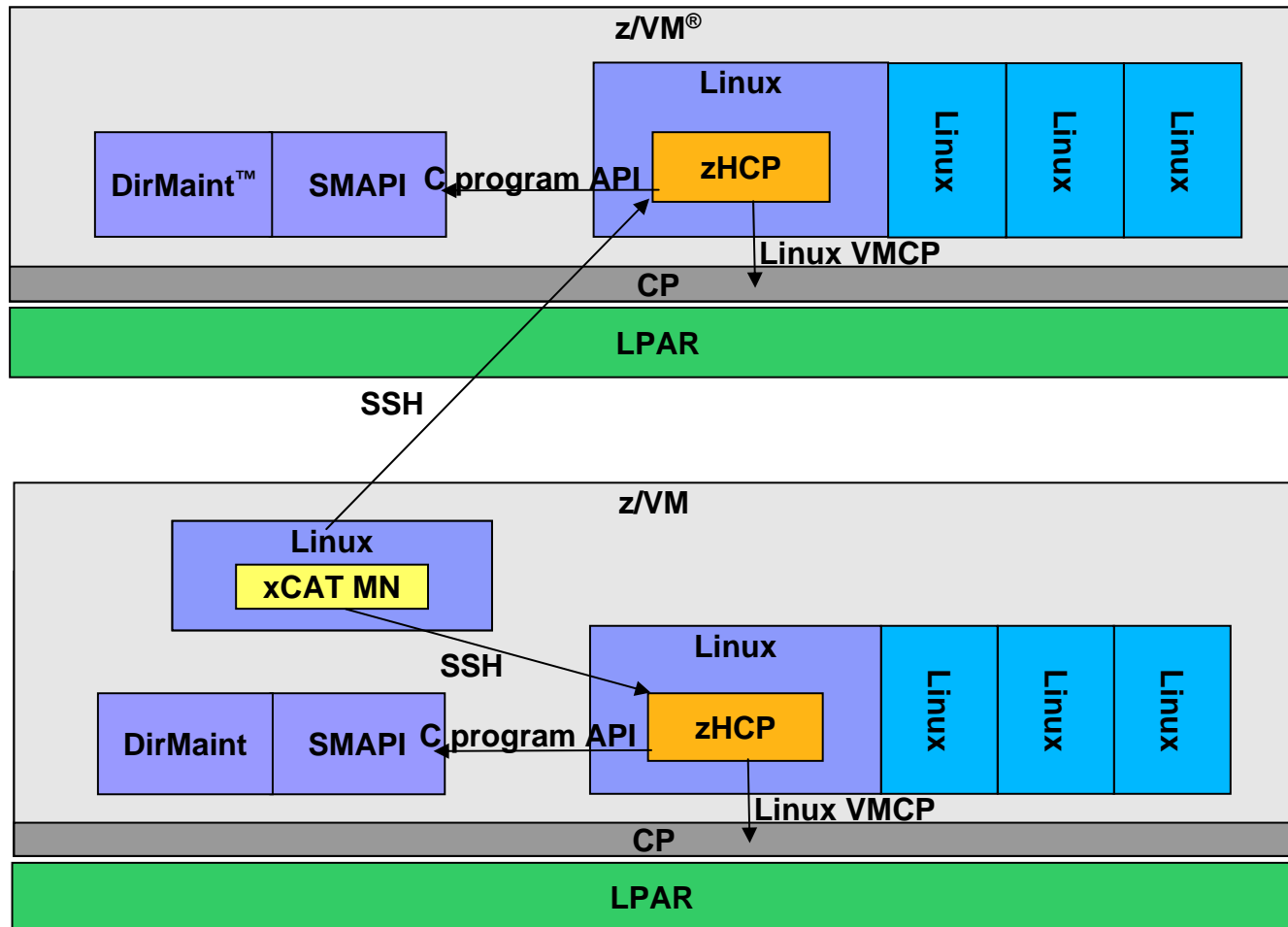
- **Integrated into z/VM 6.3**
 - Virtual machine definitions
 - Management node and zHCP executables
 - Serviceable via full-part replacement

xCAT Appliance Integration ...

- **Functionality**
 - Virtual machine lifecycle management
 - Inventory
 - Image management
 - Network management
 - Storage management
 - Operating system management
 - Monitoring

- **xCAT management node (MN) runs on preconfigured Linux virtual server**
 - Manages each z/VM partition using a System z hardware control point (zHCP) running on second preconfigured Linux virtual server
 - zHCP interfaces with z/VM systems management API (SMAPI), directory manager (DirMaint), and Control Program (CP) to manage the z/VM partition
 - C socket interface to communicate with the SMAPI layer
 - VMCP Linux module to communicate with the CP layer

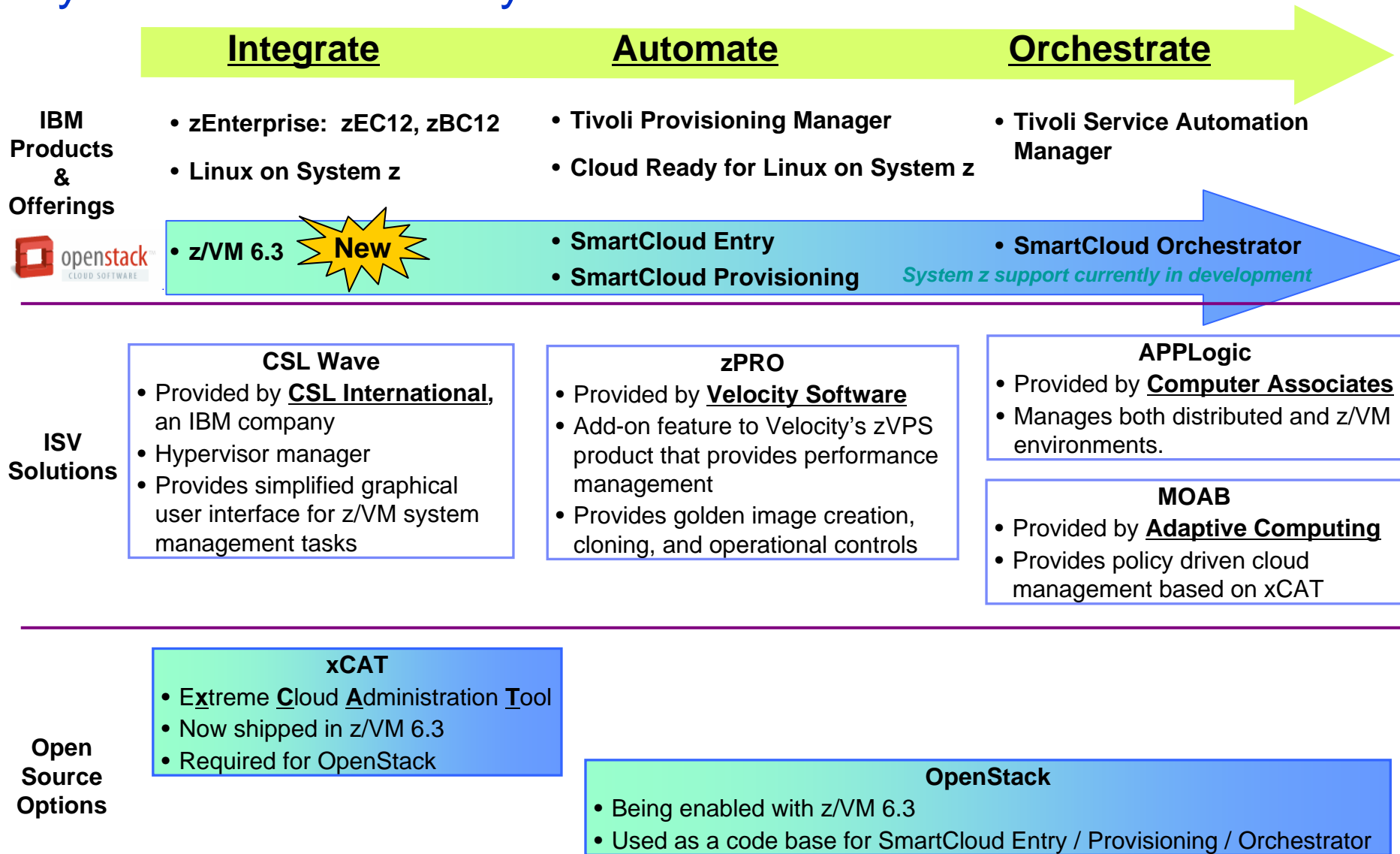
xCAT Architecture



zManager Support Removal

- **Unlike zEnterprise blade hypervisors, z/VM is not encapsulated**
- **Clients tell us they cannot perform all z/VM management tasks via zManager**
 - Hypervisor configuration and management
 - Backup
 - Accounting
 - Clustering and guest mobility
 - Performance analysis and tuning
 - Dedicated device management
 - ...
- **Need to continue to use existing management interfaces conflicts with zManager**
- **Unrealistic to expect zManager to address gaps and restrictions, especially as z/VM evolves**
- **z/VM 6.2 will be last release supported by zManager and only on zEC12 / zBC12 and earlier**
 - z/VM will continue to be supported in a logical partition but zManager will not provide management of z/VM guests
 - Alternatives will be provided for establishing IEDN connectivity

System z Cloud Ecosystem



Performance

Performance Toolkit Feature Enhancements

- **New memory displays/reports**
 - **User paging menu**
 - **User Page Activity**
 - **User Page Activity (benchmarked)**
 - **User Page Utilization Data**
 - **User Page Utilization Data (benchmarked)**
 - **Available List Data Below 2G, by Time**
 - **Available List Data Above 2G, by Time**
 - **Steal Statistics, by Time**

Performance Toolkit Feature Enhancements ...

- **New HiperDispatch displays/reports**
 - **System topology machine organization**
 - **Logical PU Organization Log**
 - **Logical PU Configuration Log**
 - **Dispatch Vector Configuration Log**
 - **Dispatch Vector Activity**
 - **Real CPU Utilization Log**
 - **DSVBK Steals per Processor Log Screen**
 - **Processor Log Screen**
 - **Logical Partition Activity Menu**
 - **Logical Partition Share**
 - **Logical Partition Logs Menu**

Performance Toolkit Feature Enhancements ...

- **Guest Fibre-Channel eXtensions (FCX)**
 - Additions to existing I/O displays/reports

- **HiperSockets Bridge**
 - Additions to existing network displays/reports

- **New Live Guest Relocation displays/reports**
 - Live Guest Relocation Event Log
 - Live Guest Relocation Data

Other

Uplevel DFSMS Binder to z/OS 1.13

- **Support PRIV parameter on STRIPSEC function**
 - **Allows Private Code sections to be removed**
- **Support RLD conditional sequential resolution**
- **Improve serviceability (comply with z/OS DoU)**

RACF LE and CMS 27 Enablement

- **RACF/VM supplies its own CMS system**

- **Upgrade CMS level one provided with z/VM 6.3**
 - **Move to a supported release of CMS (was CMS 14!)**
 - **Remove CMS-imposed restrictions on running Language Environment**

- **Enable Language Environment (LE) and LE-dependent programs, specifically System SSL, to run in the RACF/VM virtual machine**
 - **System SSL requires LE C run-time support**
 - **Enable future support for RACDCERT (RACF digital certificates)**

Thank You