yourdotcom

International Technical Support Organization and Authoring Services

IBM Unified Resource Manager

www.ibm.com/redbooks

Part 0 – General information

Redbooks Workshop

How are we doing?

Introduction	09:00 - 09:15
Architectural overview	09:15 - 10:45
Break	*10:45 - 11:00*
Monitoring and Managing	11:00 - 12:30
Lunch	*12:30 - 13:30*
Implementing the ensemble	13:30 - 14:30
Walk-through - Part I	14:30 - 15:00
Break	*15:00 - 15:15*
Walk-through - Part II	15:15 - 16:15
What's new	16:15 - 16:45
Discussion	16:45 - 17:00

09:15 - 09:30	Intro
09:30 - 10:35	Architectural Overview
10:35 - 10:50	Break
10:50 - 11:15	Architectural Overview
	Monitoring
12:45 - 13:45	Lunch
13:45 - 15:30	Implementation I
15:30 - 15:50	Break
15:50 - 16:50	Implementation II
16:50 - 17:10	What's new

Contact

Thomas Schulze

Senior IT Specialist
IBM STG Lab Services – R&D Böblingen

Wilhelm-Fay-Str. 30 – 34
D-65936 Frankfurt

Mobil: 0049-172-6354066
Mail: tschulze@de.ibm.com

Agenda for today

I Architectural Overview

II Monitoring and Managing

III Implementing the ensemble

IV Ensemble Managing and Problem Determination

V What's new

General information for today

- Schedule for the day

Introduction	**09:00 - 09:15**
Architectural overview	**09:15 - 11:15**
Break	*11:15 - 11:30*
Monitoring and Managing	**11:30 - 12:30**
Lunch	*12:30 - 13:30*
Implementing the ensemble	**13:30 - 14:30**
Walk-through - Part I	**14:30 - 15:00**
Break	*15:00 - 15:15*
Walk-through - Part II	**15:15 - 16:15**
What's new	**16:15 - 16:45**
Discussion	**16:45 - 17:00**

- Questions can be asked at any time
- Please make sure that you are back on time after the breaks

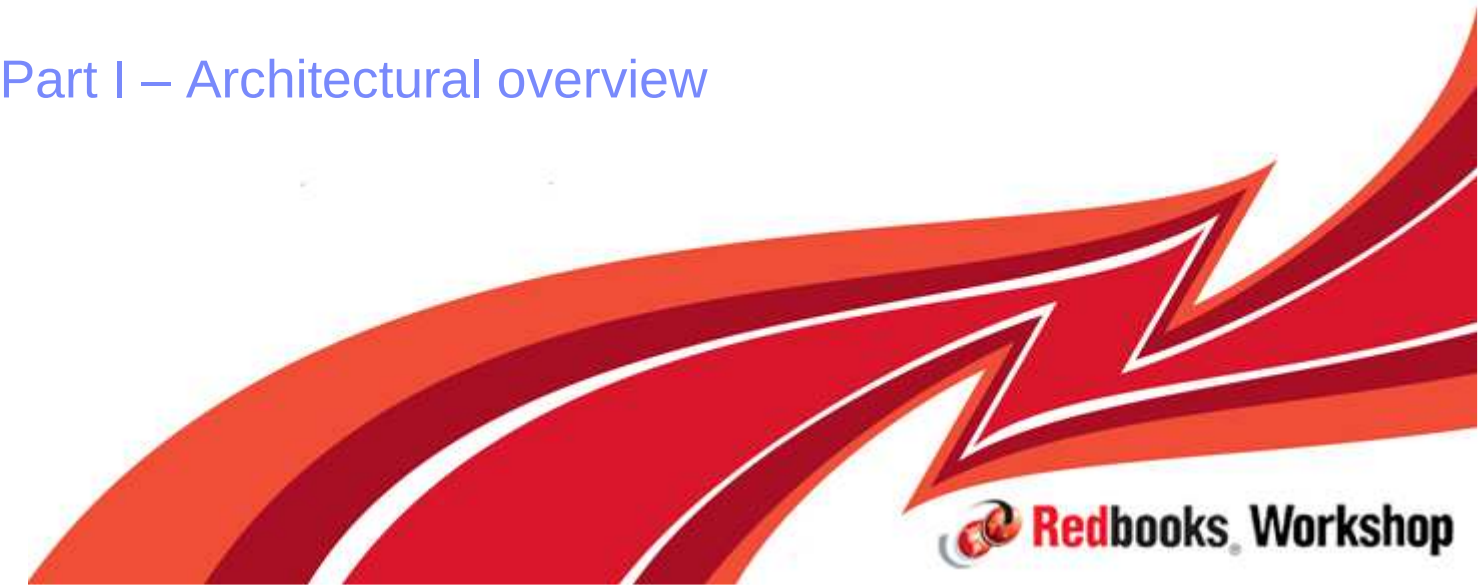
yourdotcom

International Technical Support Organization and Authoring Services

IBM Unified Resource Manager

www.ibm.com/redbooks

Part I – Architectural overview



Agenda Part I – Architectural overview

1. System zEnterprise Ensemble overview
2. Hardware building blocks
3. Hypervisors
4. Storage
5. Networks
6. Virtual Servers

Agenda Part I – Architectural overview

1 **System zEnterprise Ensemble overview**

- ***Today's IT environment***
- ***zEnterprise Architecture***
- ***Ensemble Overview***

2 Hardware building blocks

3 Hypervisors

4 Storage

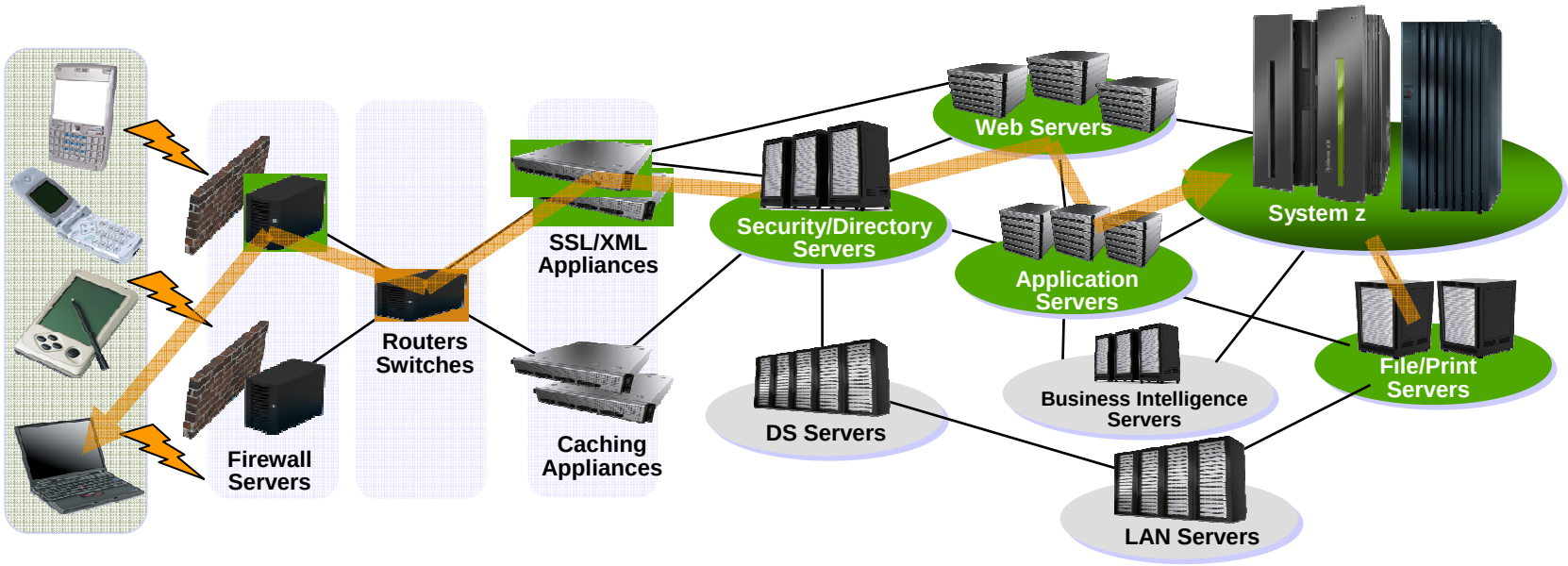
5 Networks

6 Virtual Servers

Today's IT infrastructure

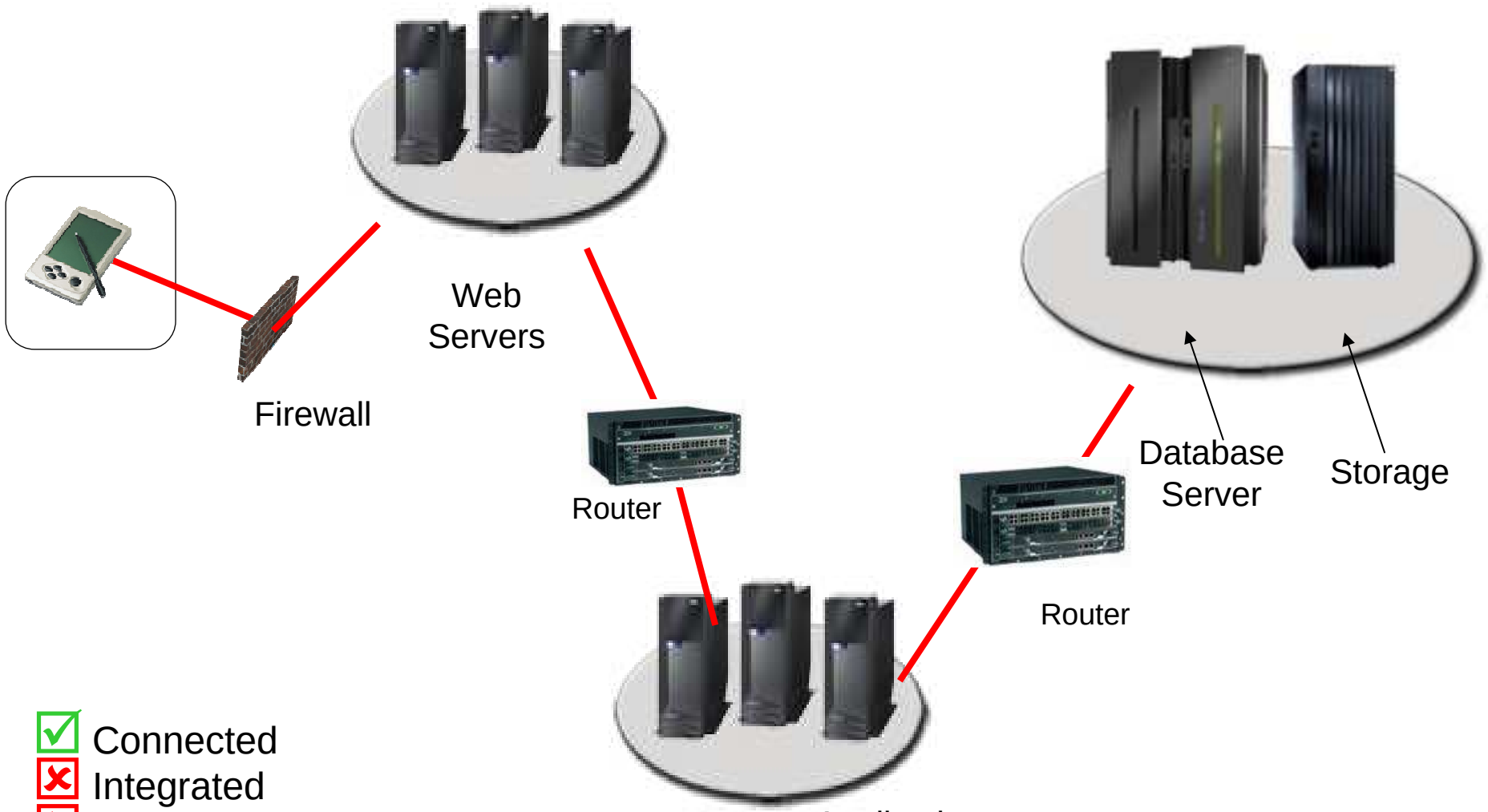
A business function delivered through a collection of IT resources

Information Technology today is limited by it's architecture and configurations available



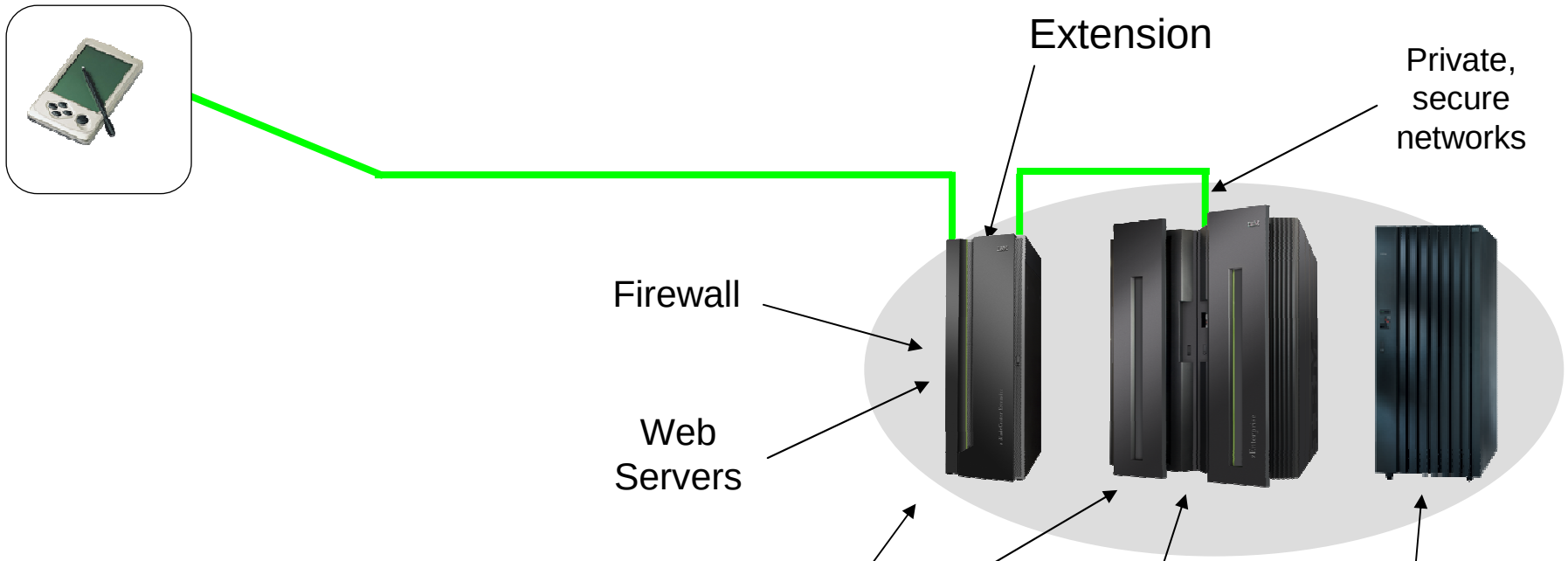
- ☑ Connected
- ☒ Integrated
- ☒ Flexible, Dynamic, and Responsive

Subset of resources representing a specific workload



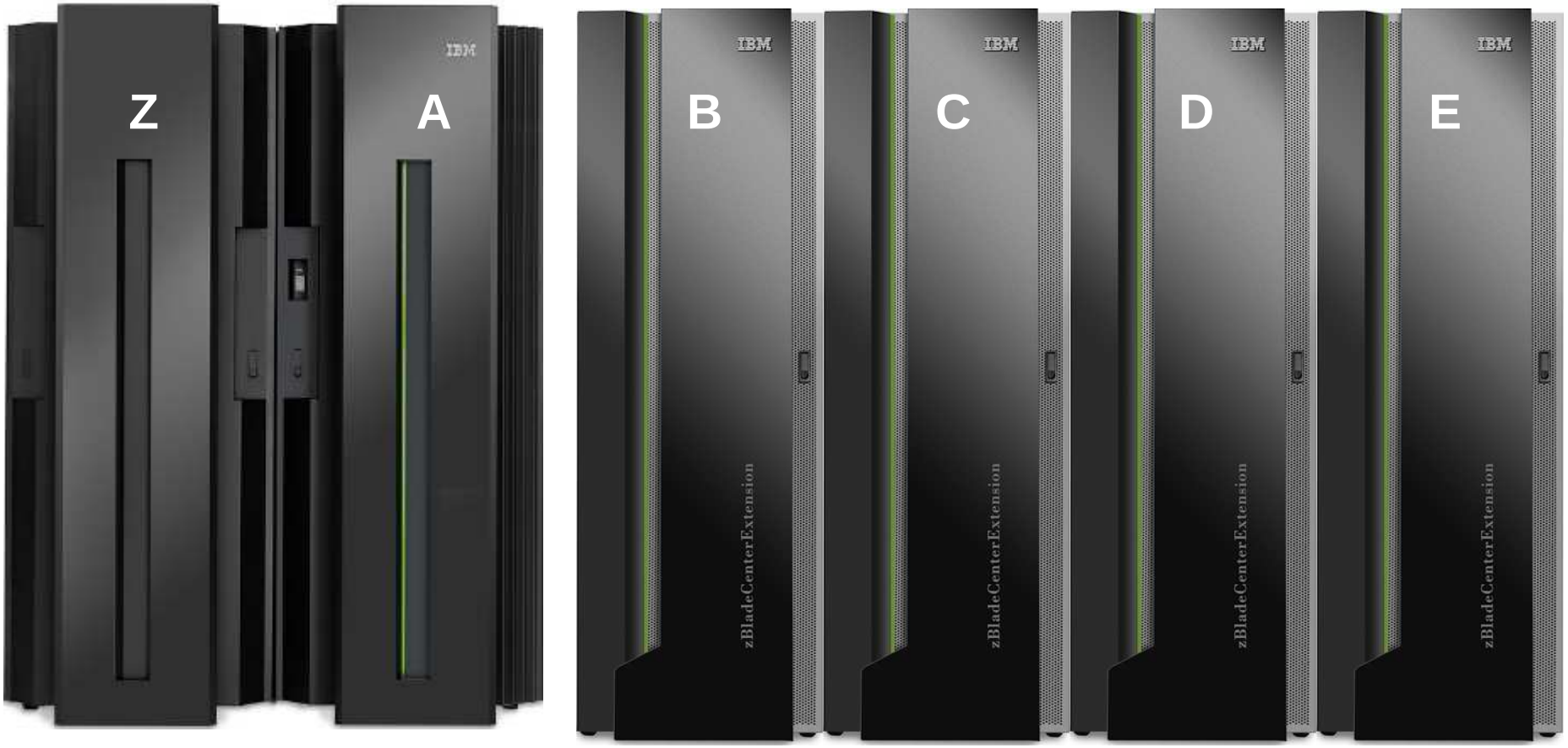
- ☑ Connected
- ☒ Integrated
- ☒ Flexible, Dynamic, and Responsive

IBM zEnterprise™ - technologies for a smarter system



- ☑ Connected
- ☑ Integrated
- ☑ Flexible, Dynamic, and Responsive

zEnterprise System introduction

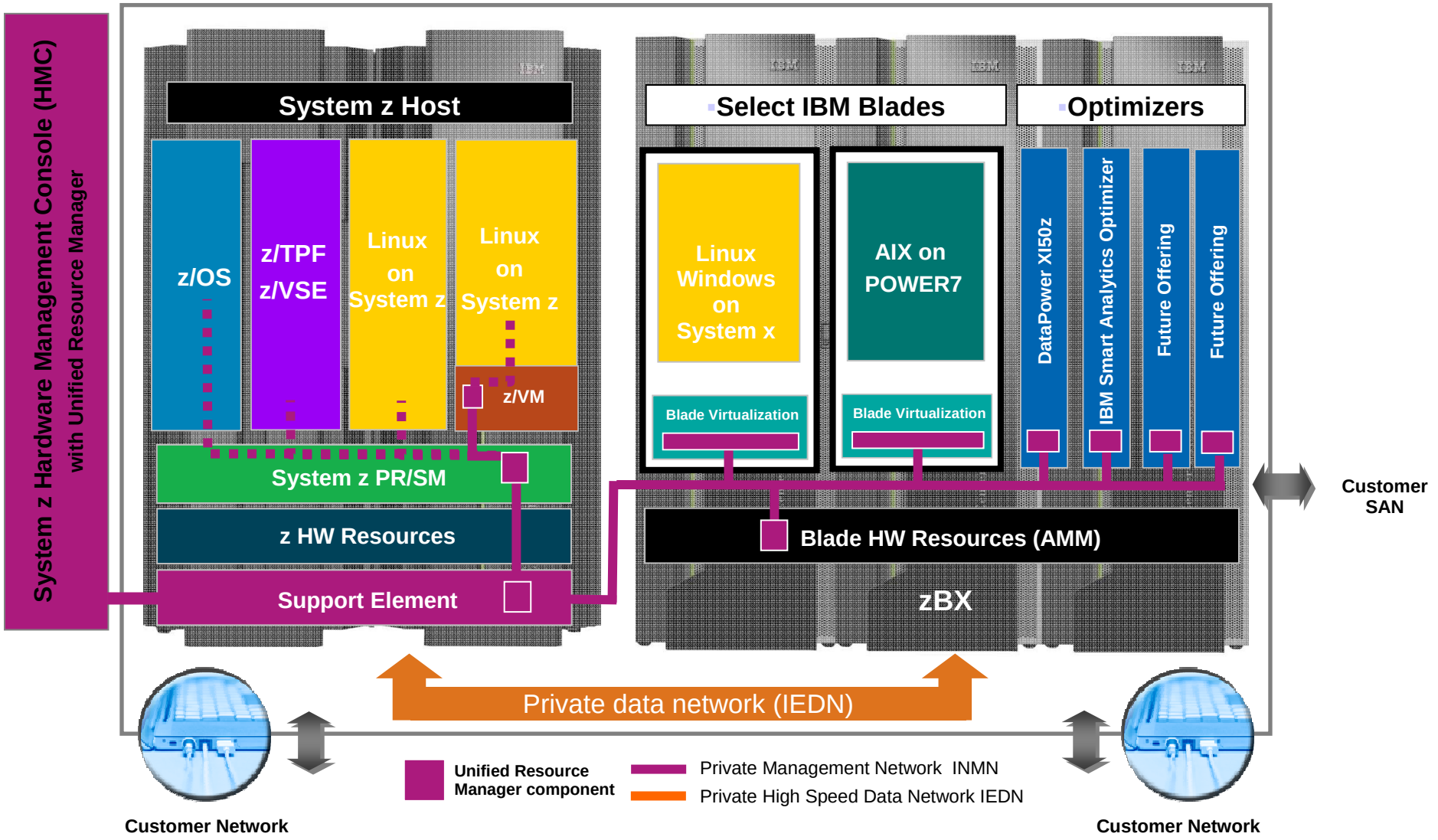


IBM zEnterprise z196/ z114

IBM zEnterprise BladeCenter Extension (zBX)

IBM zEnterprise Unified Resource Manager

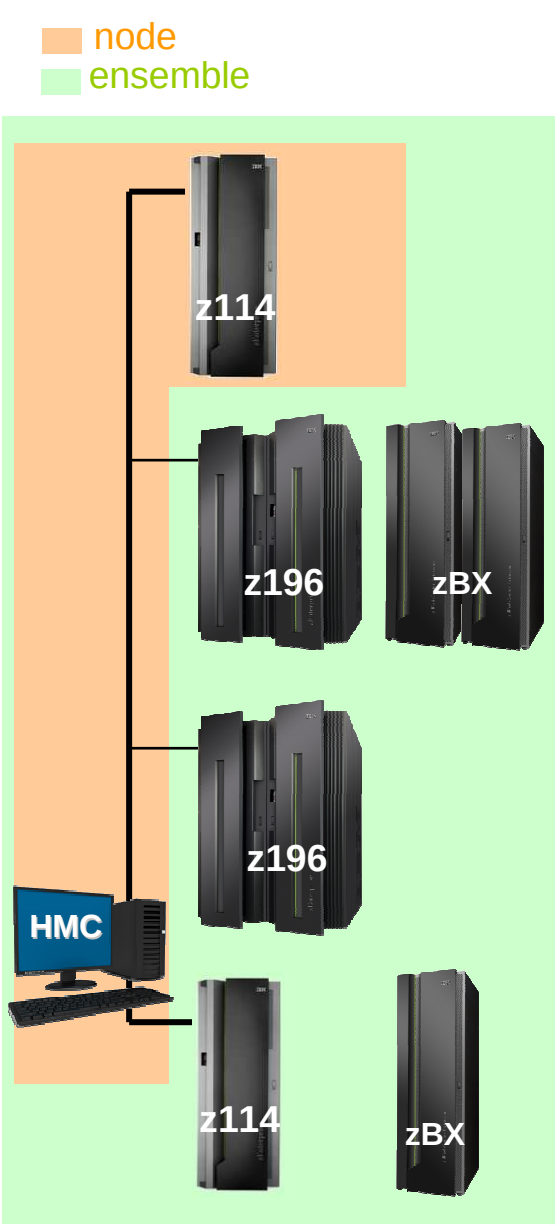
zEnterprise System – Architectural overview



Ensemble Overview

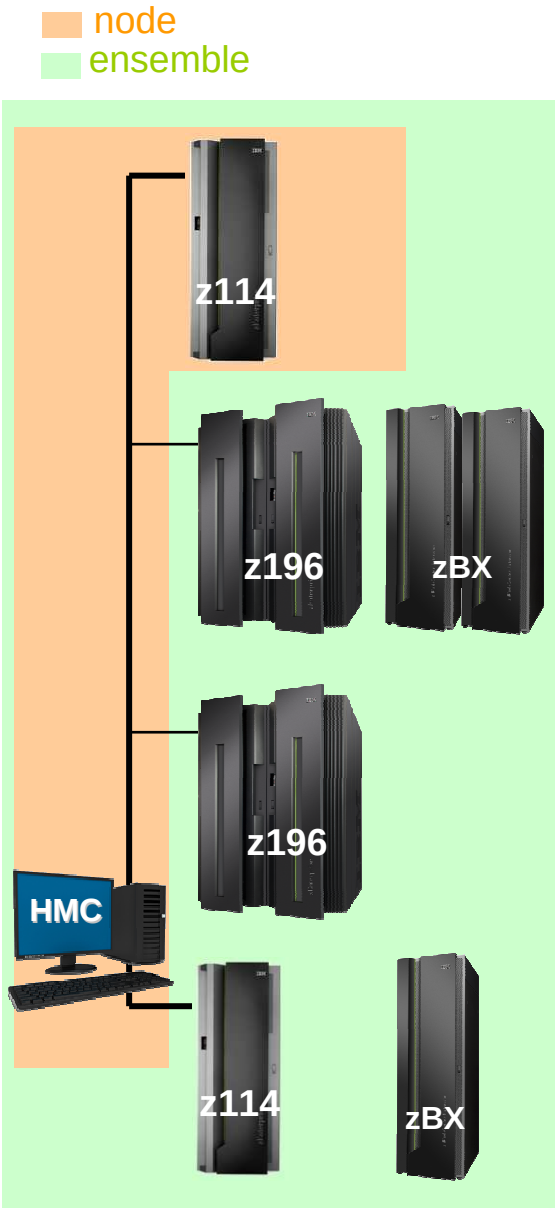
- zEnterprise Ensemble
  - An ensemble is a collection of one or more zEnterprise NODES managed as a single virtualized pool of server resources
  - A zEnterprise Node is either a z196 or a z196 with attached z BladeCenter Extension (zBX)
  - A zEnterprise ensemble is managed by a primary hardware management console (HMC)
  - The Primary HMC requires an Alternate HMC for redundancy purposes

- Platform management
  - All platform management actions for a zEnterprise node are managed through the primary HMC
  - The HMC interfaces with the components of the Ensemble node through the Support Element (SE)
  - Managing resources of the ensemble is the responsibility of Unified Resource Manager



Important IBM zEnterprise™ terms

- Node
  - A single z196 and any optionally attached zBX. A node can be a member of only one ensemble
- Ensemble
  - A collection of one or more zEnterprise nodes (including any optionally attached zBX) that are managed as a single logical virtualized system by the Unified Resource Manager using a Hardware Management Console (HMC)
- Ensemble member
  - A zEnterprise node that has been added to an ensemble using the Hardware Management Console
- Hypervisor
  - A program that allows multiple instances of operating systems or virtual servers to run simultaneously on the same hardware device
  - A hypervisor can run directly on the hardware, can run within an operating system, or can be imbedded in platform firmware. Examples of hypervisors include PR/SM™ z/VM®, and PowerVM™
- Virtual server
  - A logical construct that appears to comprise processor, memory, and I/O resources conforming to a particular architecture. A virtual server can support an operating system, associated middleware, and applications. A hypervisor creates and manages virtual servers



Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

- ***zEnterprise CPC (z196 and z114)***
- ***zEnterprise BladeCenter Extension (zBX)***
- ***Blades***

3 Hypervisors

4 Storage

5 Networks

6 Virtual Servers

zEnterprise System – z196

IBM zEnterprise 196 (2817)

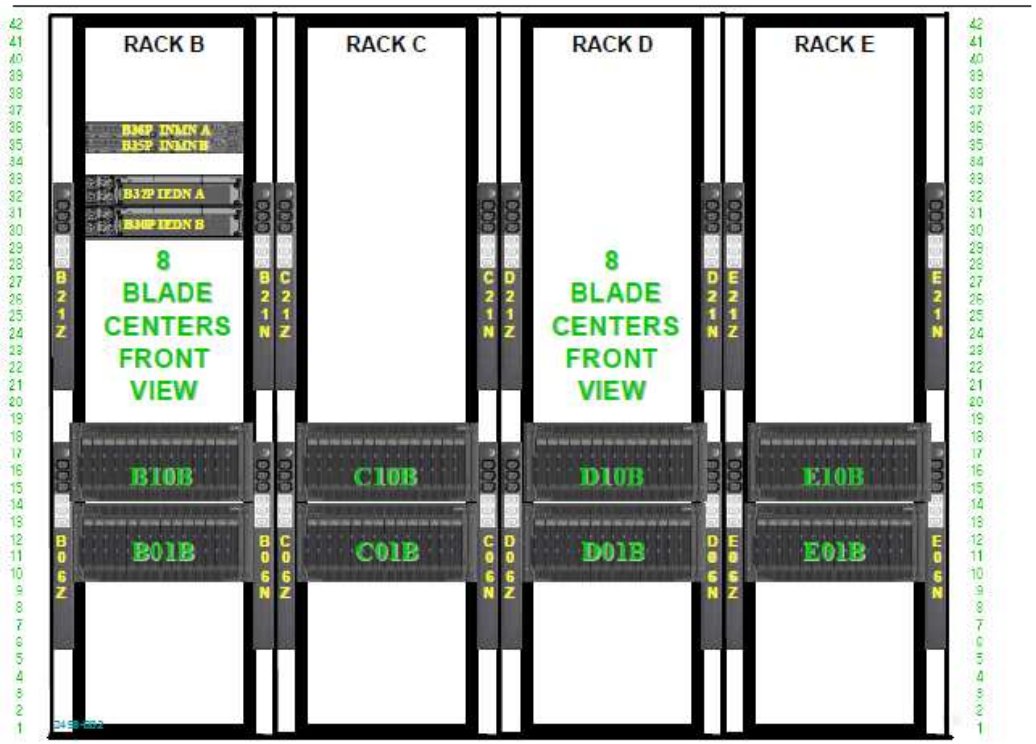
- **Announced 7/10 – Server w/ up to 96 PU cores**
- **5 models – Up to 80-way**
- **Granular Offerings for up to 15 CPs**
- **PU (Engine) Characterization**
  - CP, SAP, IFL, ICF, zAAP, zIIP
- **On Demand Capabilities**
  - CoD, CIU, CBU, On/Off CoD, CPE
- **Memory – up to 3 TB for Server and up to 1 TB per LPAR**
  - 16 GB Fixed HSA
- **Channels**
  - Four LCSSs
  - 3 Subchannel Sets
  - MIDAW facility
  - Up to 240 ESCON channels
  - Up to 288 FICON channels
  - FICON Express8 and 8S
  - zHPF
  - OSA 10 GbE, GbE, 1000BASE-T
  - InfiniBand Coupling Links
- **Configurable Crypto Express3**
- **Parallel Sysplex clustering**
- **HiperSockets – up to 32**
- **Up to 60 logical partitions**
- **Enhanced Availability**
- **Unified Resource Manager**
- **Operating Systems**
  - z/OS, z/VM, z/VSE, z/TPF, Linux on System z

zEnterprise System – z114

- **Machine Type**
  - 2818
- **2 Models**
  - M05 and M10
  - Single frame, air cooled
  - Non-raised floor option available
  - Overhead Cabling and DC Power Options
- **Processor Units (PUs)**
  - 7 PU cores per processor drawer (One for M05 and two for M10)
  - Up to 2 SAPs per system, standard
  - 2 spares designated for Model M10
  - Dependant on the H/W model - up to 5 or 10 PU cores available for characterization
    - Central Processors (CPs), Integrated Facility for Linux (IFLs), Internal Coupling Facility (ICFs), System z Application Assist Processors (zAAPs), System z Integrated Information Processor (zIIP), optional - additional System Assist Processors (SAPs)
  - 130 capacity settings
- **Memory**
  - Up to 256 GB for System including HSA
    - System minimum = 8 GB (Model M05), 16 GB (Model M10)
    - 8 GB HSA separately managed
    - RAIM standard
    - Maximum for customer use 248 GB (Model M10)
    - Increments of 8 or 32 GB
- **I/O**
  - Support for non-PCIe Channel Cards
  - Introduction of PCIe channel subsystem
    - Up to 64 PCIe Channel Cards
  - Up to 2 Logical Channel Subsystems (LCSSs)
- **STP - optional** (No ETR)

zEnterprise Blade Center Extension - zBX

- BladeCenter frame to host multiple resource types
  - Select IBM blades – POWER7 and IBM x86
  - Optimizers –
    - IBM Smart Analytics Optimizer
    - WebSphere DataPower appliance
- More System z-like service than BladeCenter ……
  - Fully integrated system with tested and supported components
  - Redundant switches, power, blowers, network to improve availability
  - A simplified service method to provide tested code updates
  - 24x7 warranty support
  - Support for z “call home”, diagnostics and maintenance by IBM zSSR
  - No MIPS/MSU rating for zBX

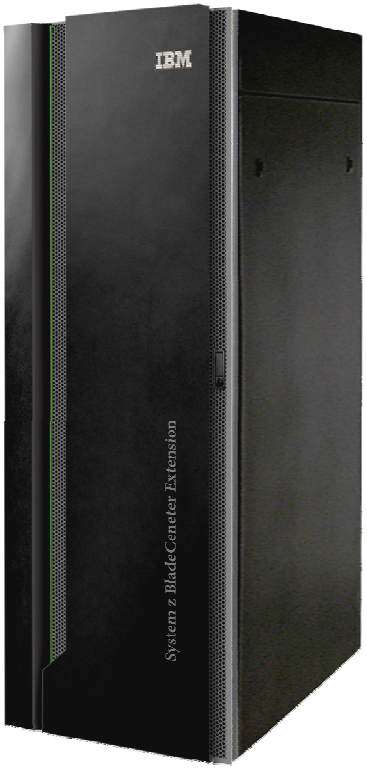


- Connected to z196/ z114 with dual 10Gbit/s secure link
- Additional connection between zBX and z196/ z114 for systems management

For details on Physical Planning refer to: *zEnterprise BladeCenter Extension Installation Manual for Physical Planning 2458-002*

zEnterprise Blade Center Extension – zBX – available doors

- Multiple doors available at customers choice



Front – zBX Standard

Front –acoustic door (optional)

Back –Rear door heat exchanger (optional)

Available POWER BladeCenter configuration

- Technical information
  - IBM BladeCenter PS701 Express
  - 8 core POWER7 processor 64-bit 3.0 GHz
  - 256KB L2 and 4MB L3 cache per core
  - QLogic 2-port 10Gb Converged Network Adapter (CFFh)
  - QLogic 8Gb Fibre Channel Expansion Card (CIOv)

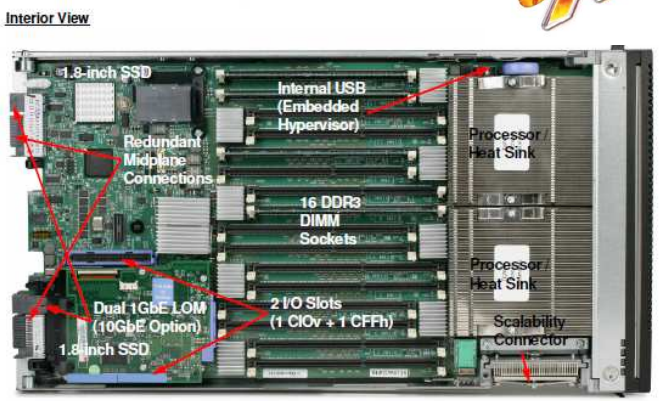
- General information
  - Must be **separately purchased and configured**
  - **Internal disk used only for License Internal Code**
  - PowerVM Enterprise Edition license needed

Config variations	Config 01	Config 02	Config 03
Processor	8-core 3.0GHz	8-core 3.0GHz	8-core 3.0GHz
DIMMS	32GB (4 x 8GB)	64GB (8 x 8GB)	128GB (8 x 16GB)
Internal Disk	1 x 300 GB HDD	1 x 300 GB HDD	1 x 300 GB HDD
CFFh I/O expansion	10 GbE	10 GbE	10 GbE
CIOv I/O expansion	8 Gbps FC	8 Gbps FC	8 Gbps FC

Before ordering the BladeServer check document *ZSY03019* for supported configuration
http://www-03.ibm.com/systems/z/hardware/zenterprise/zbx.html

Available x86 BladeCenter configuration

- Technical information
  - 16 cores available
  - Different memory configurations up to 256GB
  - Broadcom 10 GbE Ethernet (expansion card)
  - QLogic 8Gb Fibre Channel (expansion card)



- General information
  - Must be **separately purchased and configured**
  - **Internal disk used only for License Internal Code**
  - No additional hypervisor license needed
  - Supports only 64Bit operating systems

Config variations	Config 01 (7873-A4x)	Config 02 (7873-A5X)
Processor	2x8 Core Intel 2.13GHz	2x8 Core Intel 2.13GHz
DIMMS	64GB (8x8GB)	128GB (16x8GB)
Internal Disk	2x50GB MCL SSD	2x50GB MCL SSD
CFFh I/O expansion	10 GbE	10 GbE
CIOv I/O expansion	8 Gbps FC	8 Gbps FC

Before ordering the BladeServer check document *ZSL03128* for supported configuration
http://www-03.ibm.com/systems/z/hardware/zenterprise/zbx.html

Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 **Hypervisors**

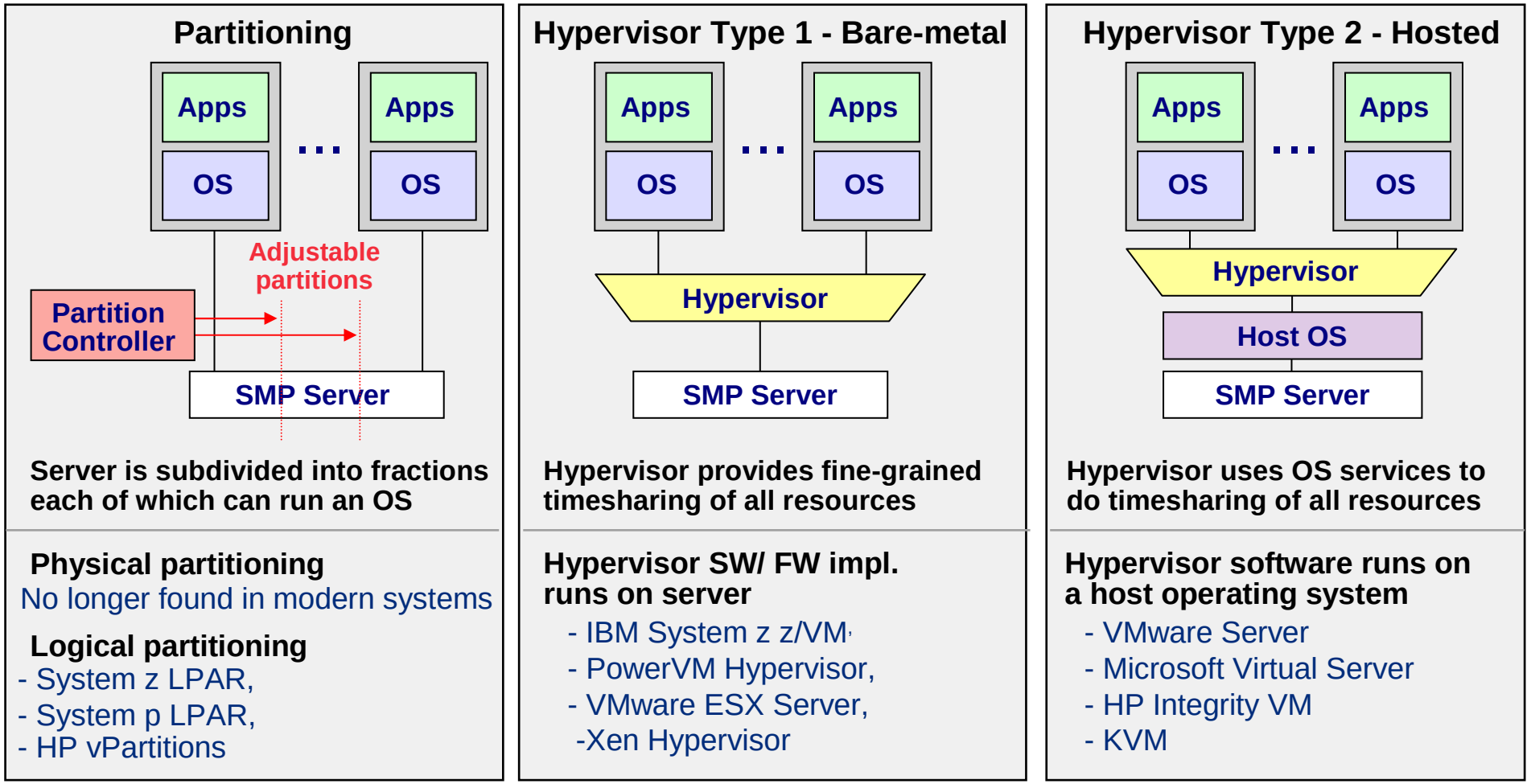
- ***Introduction to Hypervisors***
- *POWER VM*
- *PR/SM*
- *z/VM*
- *KVM*

4 Storage

5 Networks

6 Virtual servers

The Basics - Server virtualization approaches



- **Characteristics**
  - Bare-metal (firmware) hypervisors offer high efficiency and availability
  - Hosted hypervisors are useful for clients where host OS integration is important

Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 **Hypervisors**

- *Introduction to Hypervisors*
- ***POWER VM***
- *PR/SM*
- *z/VM*
- *KVM*

3 Storage

4 Networks

5 Virtual servers

PowerVM hypervisor concepts

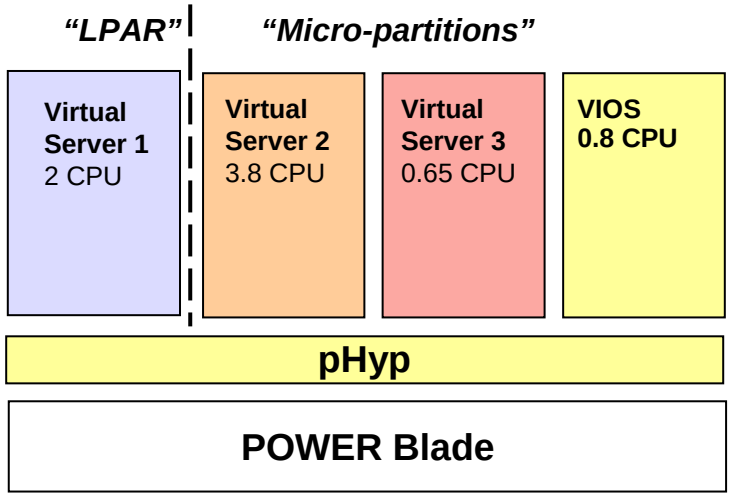
- PowerVM is built of multiple components of which **pHyp** is the base layer
  - The pHyp layer does not require HW capacities
- Different Editions are available in a general environment → A license is required on any System p hardware
  - PowerVM Express Edition
  - PowerVM Standard Edition
  - **PowerVM Enterprise Edition** *(required for Unified Resource Manager environments)*
- Concepts very similar to system z
  - LPAR
  - Dedicated/ shared processors
  - Capped/ **uncapped** processing mode *(only uncapped mode used in Unified Resource Manager)*
  - Logical processors
  - Internal, virtual communication
  - Shared I/O

Note: Though concepts are similar, looking at them in detail reveals some differences between PowerVM and the System z world.

PowerVM hypervisor concepts - LPAR

- **Approach I:** The "*normal*" LPAR
  - Whole processors only
  - Uses dedicated resources

- **Approach II:** Micro-Partition
  - Fractions of a processor → processing units
  - Uses shared resources
  - Requires a "System guest/ server" that performs the I/O → VIOS (Virtual I/O Server)



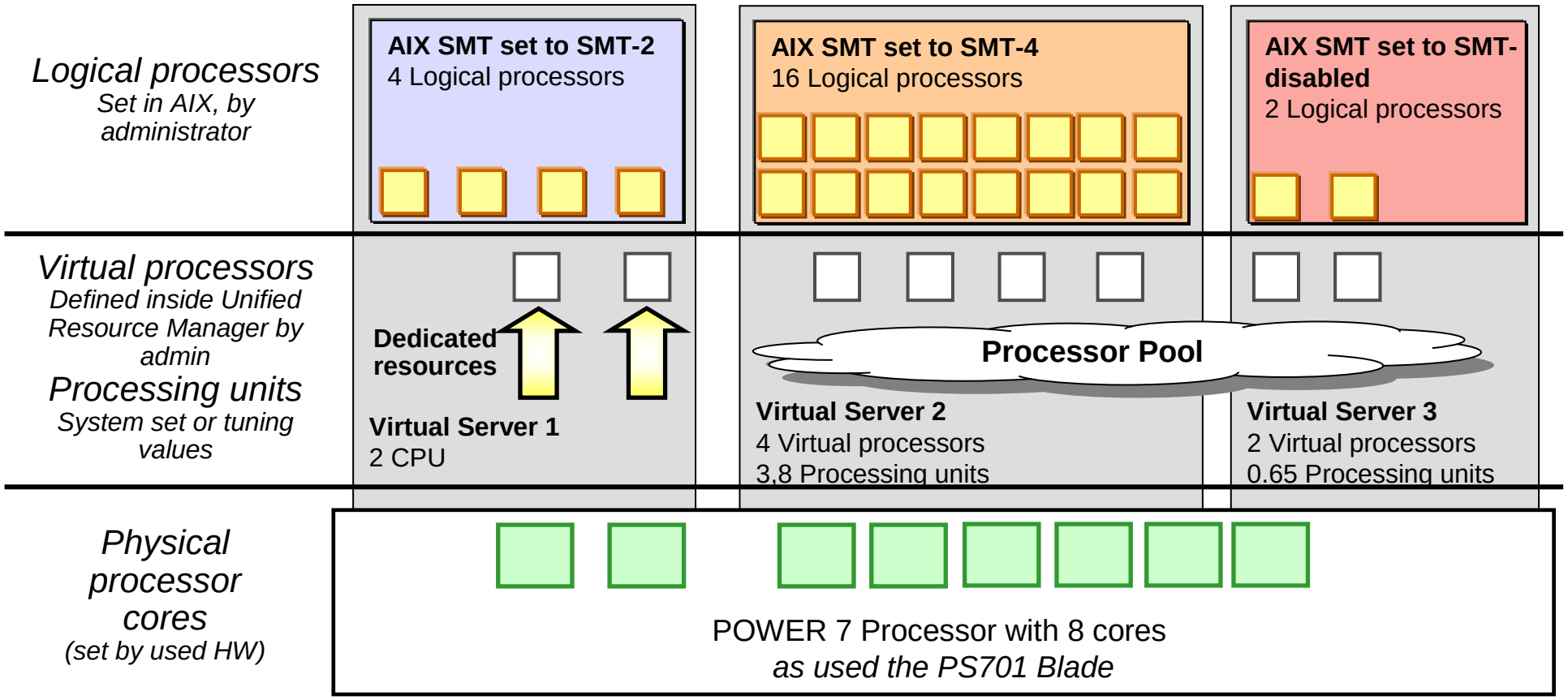
- PowerVM can mix virtual servers of both approaches on one hardware

- Assigned resources can be adjusted on-the-fly using manual actions

Note: There is nothing like "Intelligent Resource Manager" available in a PowerVM environment to automatically adjust resources
→ but with Unified Resource Manager

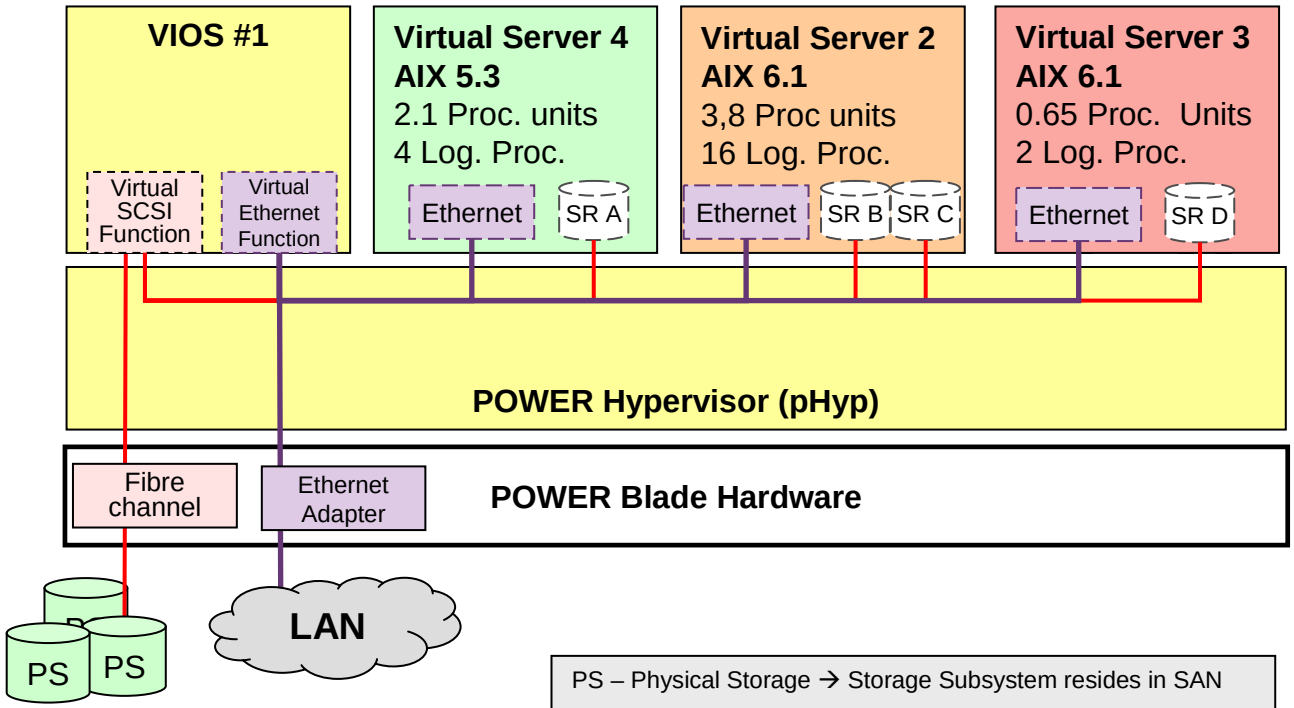
PowerVM hypervisor concepts – The Big picture on processors

- Multiple definitions can be made at different levels of the overall architecture
- Virtual servers with shared processor resources can be adjusted on the fly



PowerVM hypervisor concepts – Connectivity

- Storage resources for a VIOS
  - Virtual SCSI
  - Fibre Channel
  - DVD
- Virtual Ethernet
  - VLAN and link aggregation support
  - LPAR to LPAR using VLANs
  - Shared Ethernet Adapter Failover



Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 **Hypervisors**

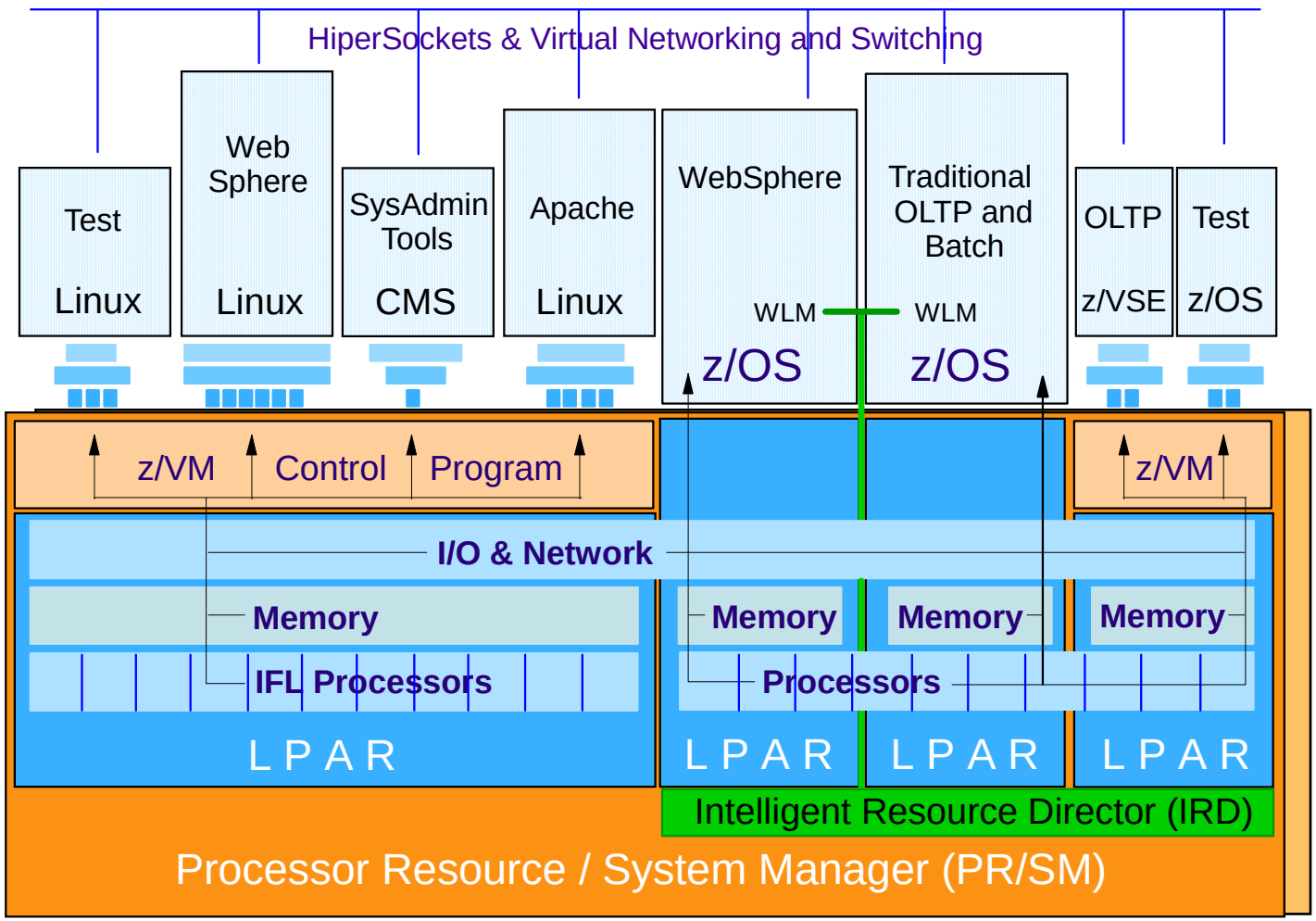
- *Introduction to Hypervisors*
- *POWER VM*
- ***PR/SM***
- *z/VM*
- *KVM*

4 Storage

5 Networks

6 Virtual servers

System z PR/SM Virtualization Architecture



- System z provides logical (LPAR) and software (z/VM) partitioning
- PR/SM enables highly scalable virtual server hosting for LPAR *and* z/VM virtual machine environments
- IRD coordinates allocation of CPU and I/O resources among z/OS and non-z/OS LPARs*

* Excluding non-shared resources like Integrated Facility for Linux processors

Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 **Hypervisors**

- *Introduction to Hypervisors*
- *POWER VM*
- *PR/SM*
- ***z/VM***
- *KVM*

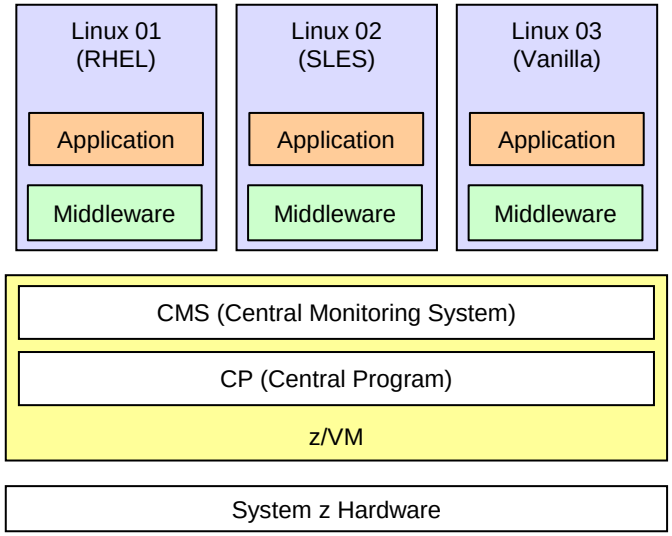
4 Storage

5 Networks

6 Virtual servers

z/VM overview

- General
  - Current Version 6.1
  - Available only on System z hardware
  - Supports HW functions (crypto, HiperSockets, OOCoD, …)
  - Runs with CP‘s or IFL processors (not mixed)

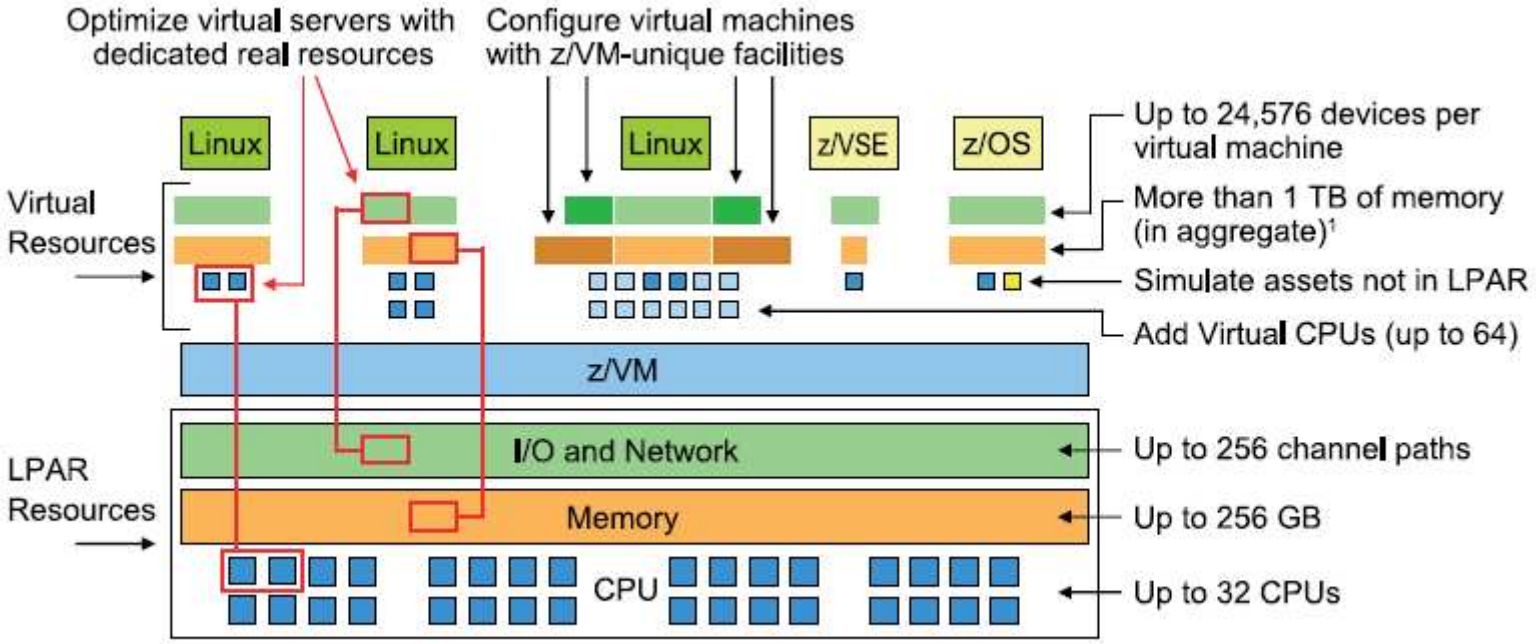


- Supported guest operating systems
  - Linux (all major distributions are supported)
  - z/OS - z/TPF
  - z/VM - z/VSE

- Features
  - Virtual networks conforming IEEE standard
  - VSwitches on OSI Layer 2
  - Cooperative Memory Management for dynamic adjustment of Linux memory
  - Shared Memory Segments, „read-only“ memory in main storage
  - Dynamic activation/ deactivation of physical resources → benefit for all guests

Note: When managing z/VM with Unified Resource Manager make sure that there are no manual changes to the z/VM directories.
This might lead to an undefined system state.

z/VM the big picture

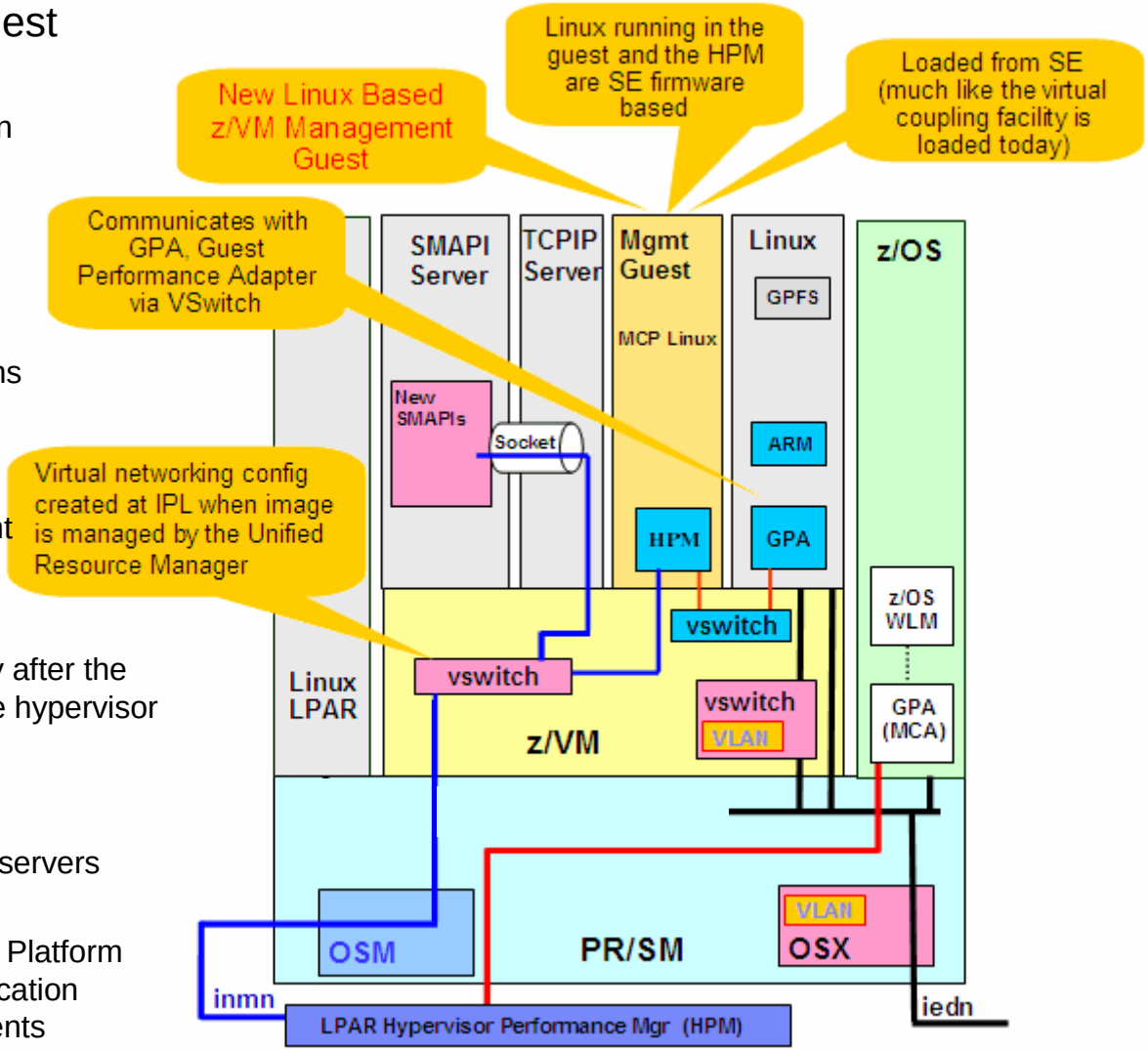


[1] *The actual amount of usable real and virtual memory is dependent on the amount of real memory in the z/VM logical partition, the hardware server model, firmware level and configuration, and the number of guests and their workload characteristics.*

- z/VM supports “mini-automation”
- Resources delivered as required, automatically, based on business-oriented goals
- New OS images can be started without affecting ongoing work

z/VM interaction with Unified Resource Manager

- New Linux based management guest
  - Shipped with the firmware, not with z/VM
  - Loaded from SE during ensemble creation
- Required z/VM guests
  - TCP/IP server
  - System Management API (SMAPI) server to update z/VM guest configurations
- Communication
  - Management guest uses the management network (INMN) to communicate
  - Configuring virtual switches
  - Components are configured automatically after the z/VM LPAR is enabled to be an ensemble hypervisor
- Tasks
  - Manage (start/ stop/ modify) other virtual servers running in z/VM
  - Receive performance metrics from Guest Platform Management Provider (GPMP) and Application Response Measurement (ARM) components



Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 Hypervisors

- *Introduction to Hypervisors*
- *POWER VM*
- *PR/SM*
- *z/VM*
- ***KVM***

4 Storage

5 Networks

6 Virtual servers

Hypervisor on x86 Blades

- x86 Blades in the zBX use always a hypervisor to virtualize physical resources
  - Exact one hypervisor per Blade
- Kernel-based Virtual Machine (KVM) used in a Linux guest
- Supported operating systems versions/editions* on zBX
  - Linux (RHEL5.5 and SLES 11 SP1)
  - Windows Server 2008 R2 Datacenter Edition
- One Blade can run Windows and Linux parallel

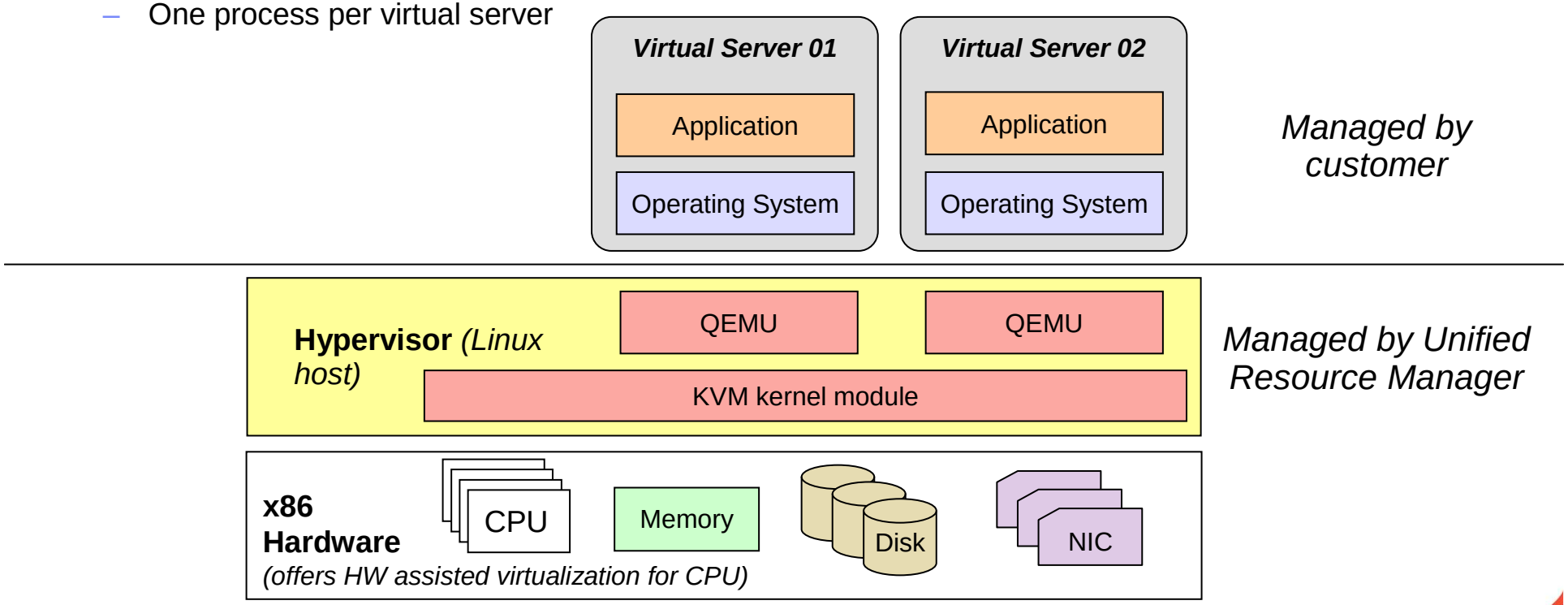
Important

Hypervisor is treated as part of the zEnterprise License Internal Code.

No customer modifications are possible. Fully managed by the Unified Resource Manager.

Virtualizing the different components of a virtual server

- CPU
  - Exploits HW assisted virtualization implemented in the CPU (Intel VT *or AMD-V* [1])
  - Each virtual processor appears as a process in the hypervisor Linux
- Memory
  - Virtualized by KVM for all virtual servers
- I/O
  - Modified QEMU (Quick Emulator) used
  - One process per virtual server



[1] Not supported in the zBX

Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 Hypervisors

4 **Storage**

- ***Storage resource overview***
- *Fibre Channel Protocol basics*
- *Blade attached storage*

5 Networks

6 Virtual Servers

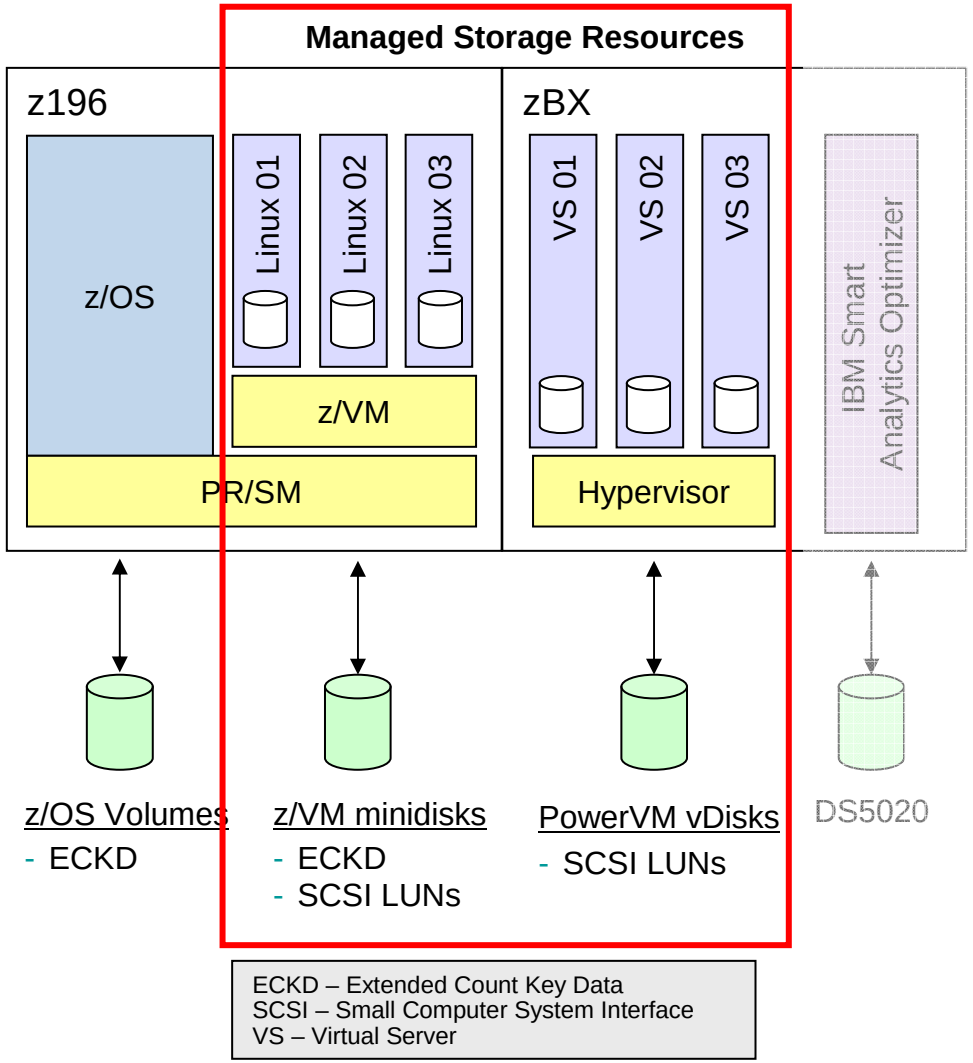
Unified Resource Manager - Storage Resources overview

- Types of storage devices
  - HMC attached
    - Virtual DVD drive
    - HMC USB drive
  - **Storage resources on SAN**
  - No direct attached tapes supported

- Storage categories
  - **Managed Storage** → Provided to servers through hypervisors
  - **Non-managed Storage** → Storage attached to LPARs or servers in the ensemble

Storage management per Hypervisor

- z/VM devices are only managed if created through Unified Resource Manager, CP created devices are ignored
- PowerVM storage can only be managed through Unified Resource Manager



For details on supported storage subsystems refer to *ZSP03437* available at:
http://www-03.ibm.com/systems/z/hardware/zenterprise/zbx.html

Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 Hypervisors

4 Storage

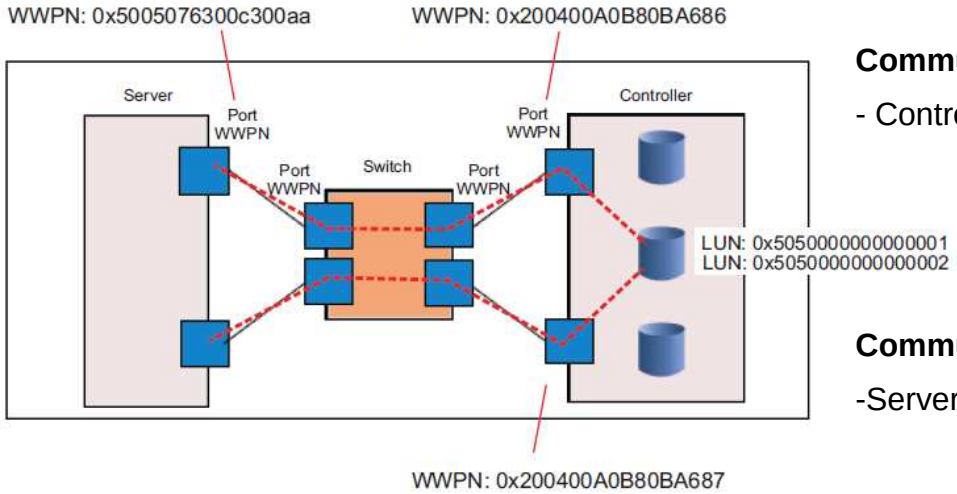
- *Storage resource overview*
- ***Fibre Channel Protocol basics***
- *Blade attached storage*

5 Networks

6 Virtual Servers

The Basics – Fibre Channel protocol (FCP)

- Terminology
  - Node
    - Server accessing fibre channel based storage
  - Ports
    - Connection point for Fibre-Channel cable
    - Identified by unique World-Wide Port name (WWPN)
  - Logical Unit
    - Virtual SCSI disk drive on Storage Controller
    - Addressed by WWPN/Logical Unit Number (LUN)
- Access Control
  - Zoning
    - Logical grouping of host and storage controller ports in switches
    - Only members within a group can see/talk to each other
    - WWPN based
  - LUN Masking
    - Enforced by storage controller
    - Host port WWPN based



Communication server → Controller:

- Controller WWPN and LUN needed

Communication Controller → server

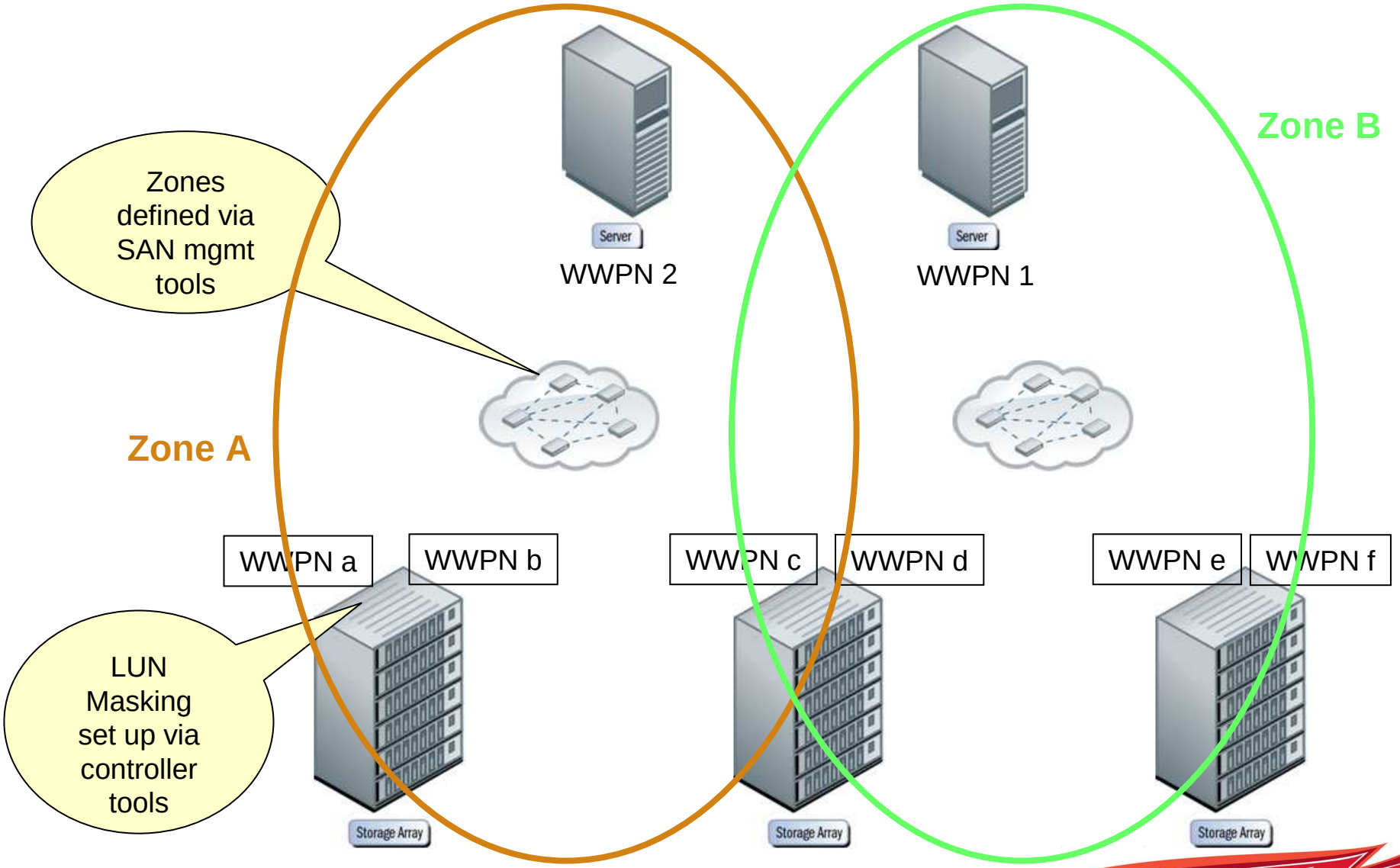
-Server WWPN needed

Zoning and **LUN Masking** are covered on the next slides

The Basics – SAN zoning

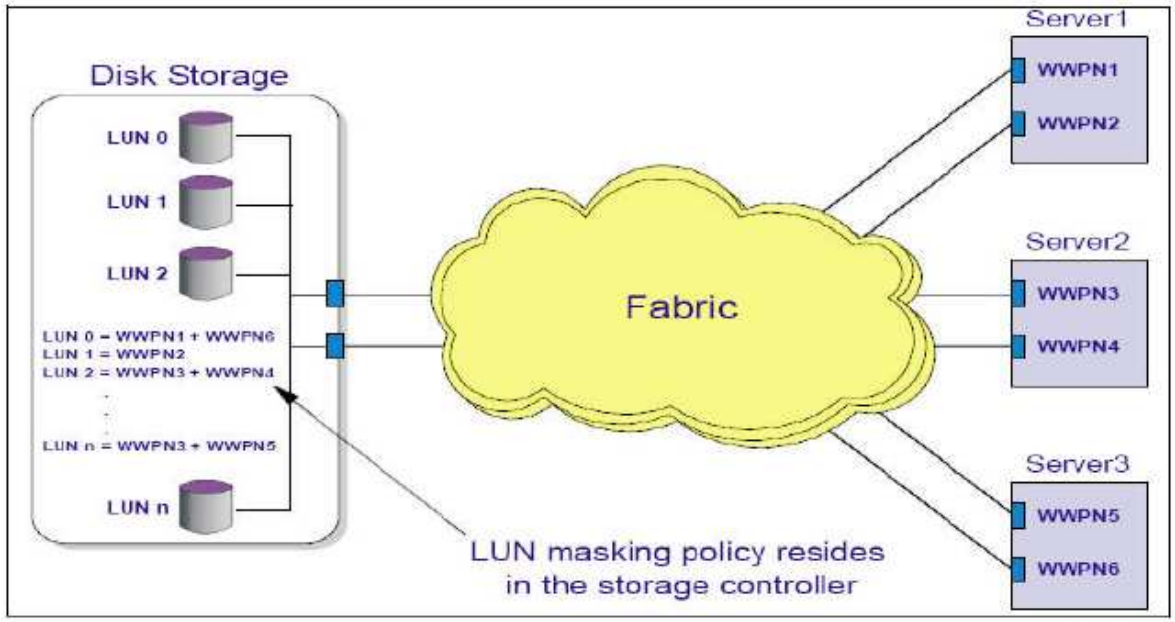
- Logical grouping of Fibre Channel devices
- Configured in the Fibre Channel Fabric
- Utilized to implement compartmentalization of data for security
- Devices may be placed into multiple zones

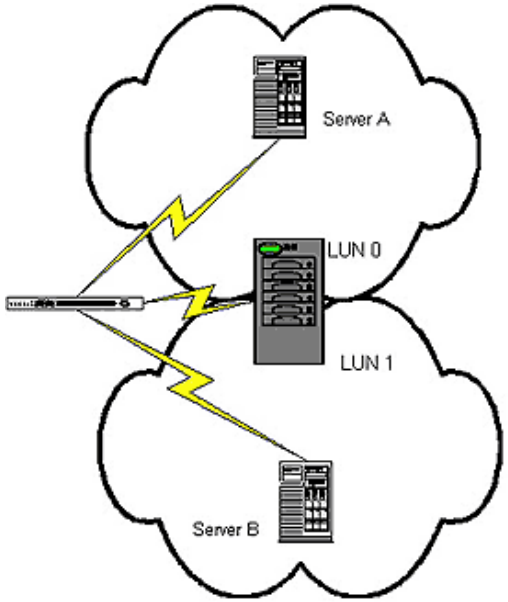
The Basics – SAN zoning – an example



The Basics – LUN masking

- Authorization process that hides/ shows LUNs to requesting hosts
- Some storage controllers also support LUN Masking
- LUN Masking is important because Windows based servers attempt to write volume labels to all available LUNs. This can render the LUNs unusable by other operating systems and can result in data loss





Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 Hypervisors

4 Storage

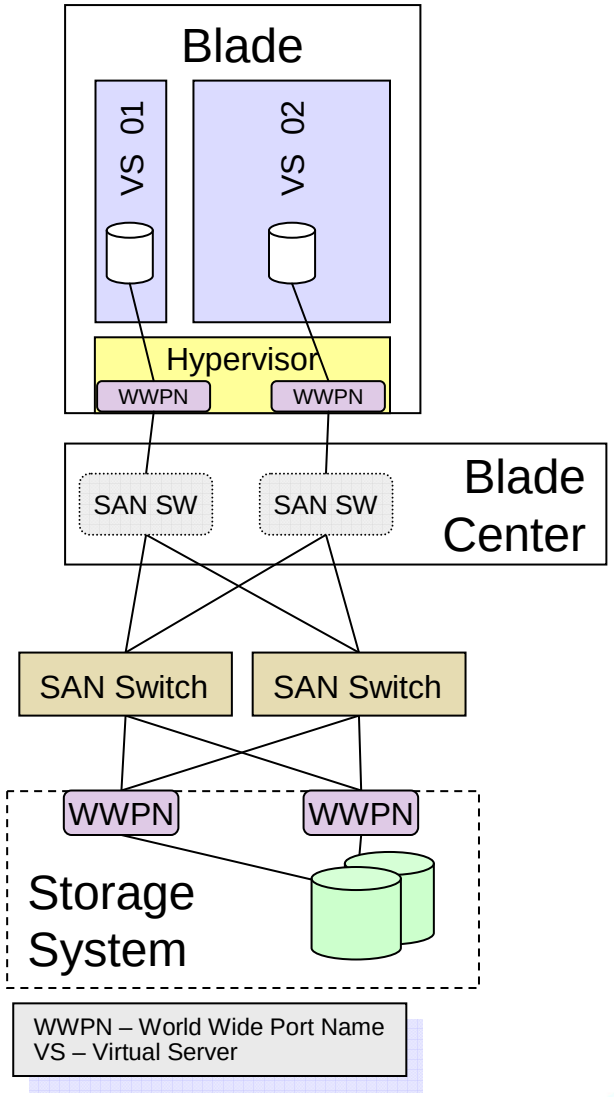
- *Storage resource overview*
- *Fibre Channel Protocol basics*
- ***Blade attached storage***

5 Networks

6 Virtual Servers

Storage Resource mapping – Blade attached Storage

- Blade
  - Supports only SAN attached storage
  - All virtual servers on a Blade share the same physical WWPN
  - Hypervisors
    - Uses Blades physical WWPNs
    - Fills the gap between shared WWPN and virtual server, mapping data packages to the target
- BladeCenter SAN Switch
  - Accumulate 14 internal ports to 6 external
  - Work in pass-through mode → don't act as switch but as N_ports
- External SAN Switch
  - Needs to support NPIV to support the addressing of 14 Blades using 1 physical adapter in the external SAN switch



Multipathing

Hypervisor performs multipathing under the cover through redundant physical links. The virtual server's operating system only sees one path to the LUN.

Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 Hypervisors

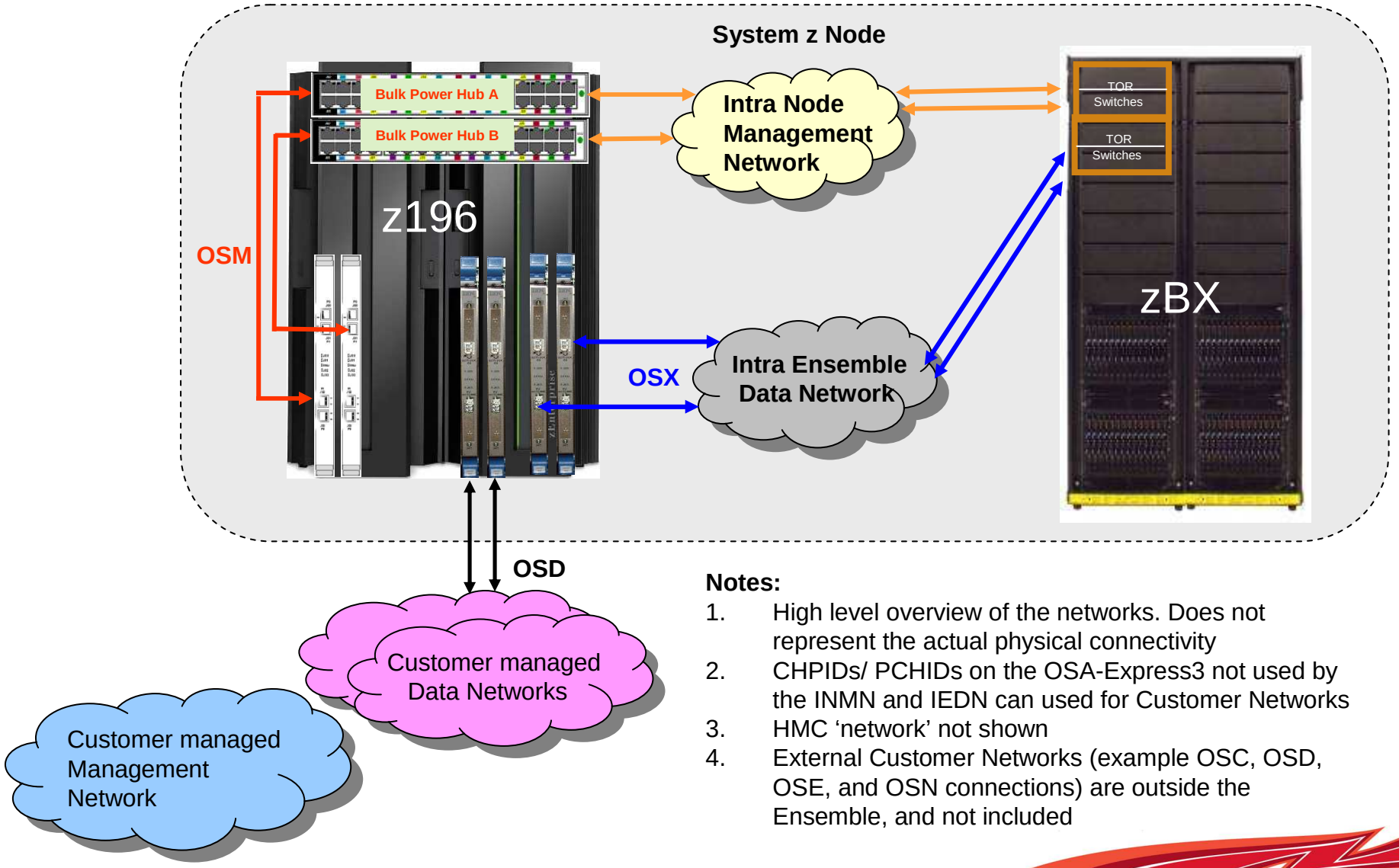
4 Storage

5 **Networks**

- ***Ensemble network overview***
- *Network virtualization*

6 Virtual Servers

Overview of physical networks in an ensemble



Notes:

1. High level overview of the networks. Does not represent the actual physical connectivity
2. CHPIDs/ PCHIDs on the OSA-Express3 not used by the INMN and IEDN can used for Customer Networks
3. HMC 'network' not shown
4. External Customer Networks (example OSC, OSD, OSE, and OSN connections) are outside the Ensemble, and not included

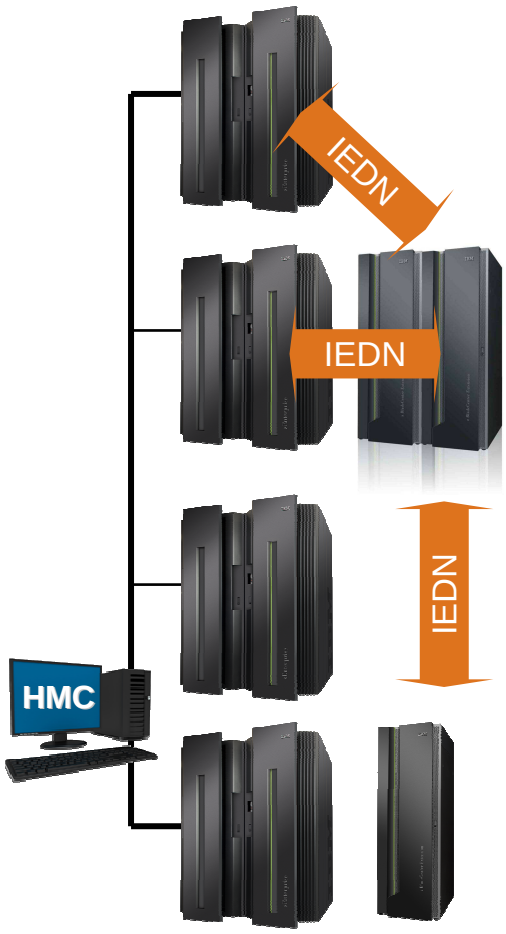
Management Network

- Intra Node Management Network (INMN)
  - Network scope is a single node in an Ensemble
- Physical 1GbE links
  - 1 link pair (2 cables) from z196 to zBX
- Used for management purposes
  - Management/ configuration of physical resources (TOR switches, Blades, …)
  - Management/ configuration of virtual resources (Virtual servers, VLANs, …)
- Implemented as
  - Layer 2 flat network
  - IPv6 network
- System z LPARs need CHPID OSM definied in the IOCDS

NOTE: This network is not customer managed. As part of the Unified Resource Manager "black box" it is transparent to other external networks/ means of communication.

Data Network

- Intra Ensemble Data Network (IEDN)
  - Spans the overall Ensemble
- Physically available in pairs of 10GbE links
  - Max. of 8 different z196 can be connected
  - 1-8 link pairs (2-16 cables) from zBX to „any“ z196
- Implemented as
  - Layer 2 flat network
  - Supports IPv4 and IPv6
- Used for customer data
  - Application data
  - Customer administration network (for operating system, middleware or applications)
  - Installation of middleware
- Isolation of multiple LAN‘s using VLAN technology and VLAN tagging



NOTE: The cables used for the IEDN are customer provided and needs to be ordered using the customers normal procurement process.

Agenda Part I – Architectural overview

1 System zEnterprise Ensemble overview

2 Hardware building blocks

3 Hypervisors

4 Storage

5 **Networks**

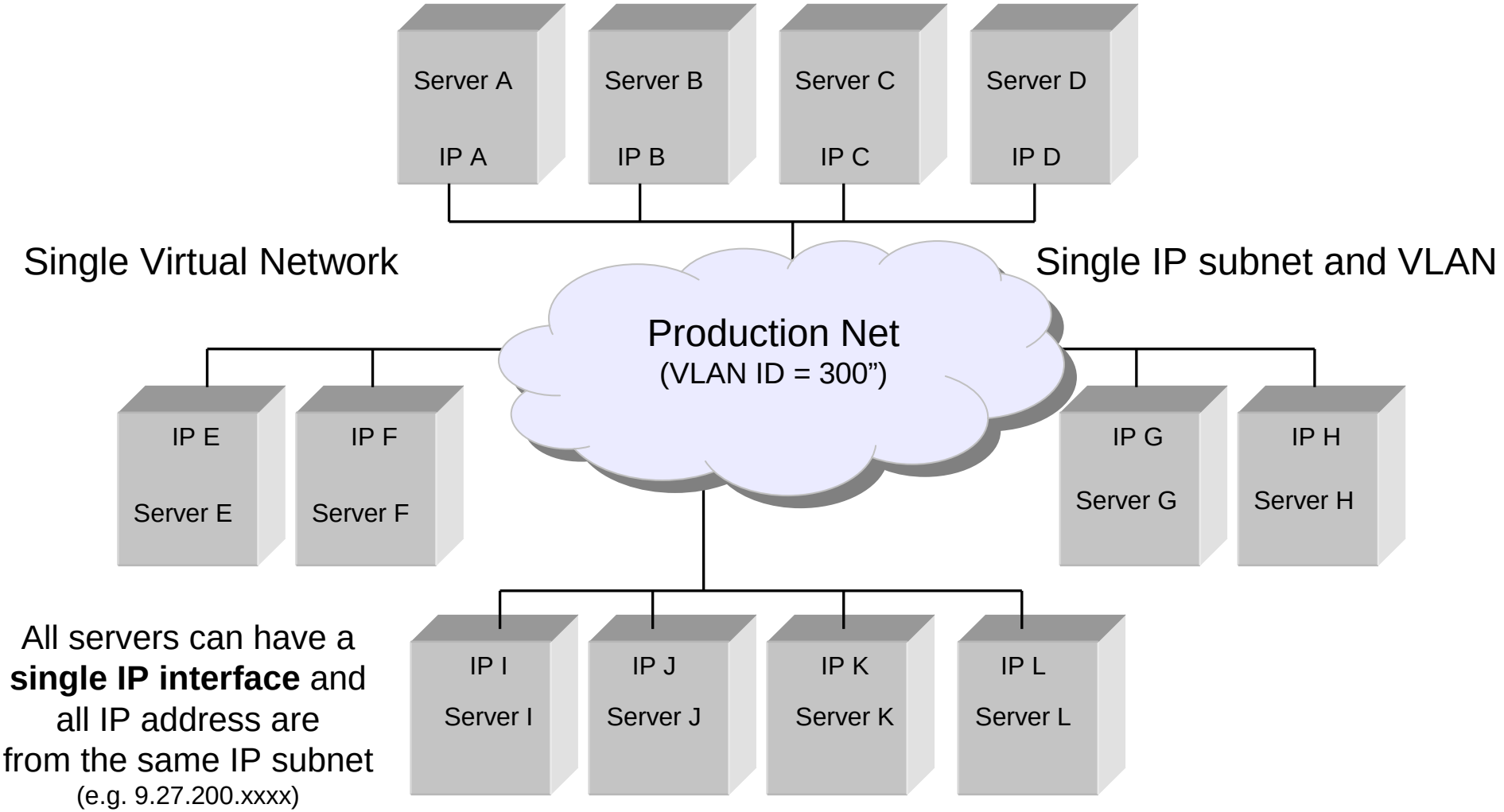
- *Ensemble network overview*
- ***Network virtualization***

6 Virtual Servers

The basics - Network virtualization

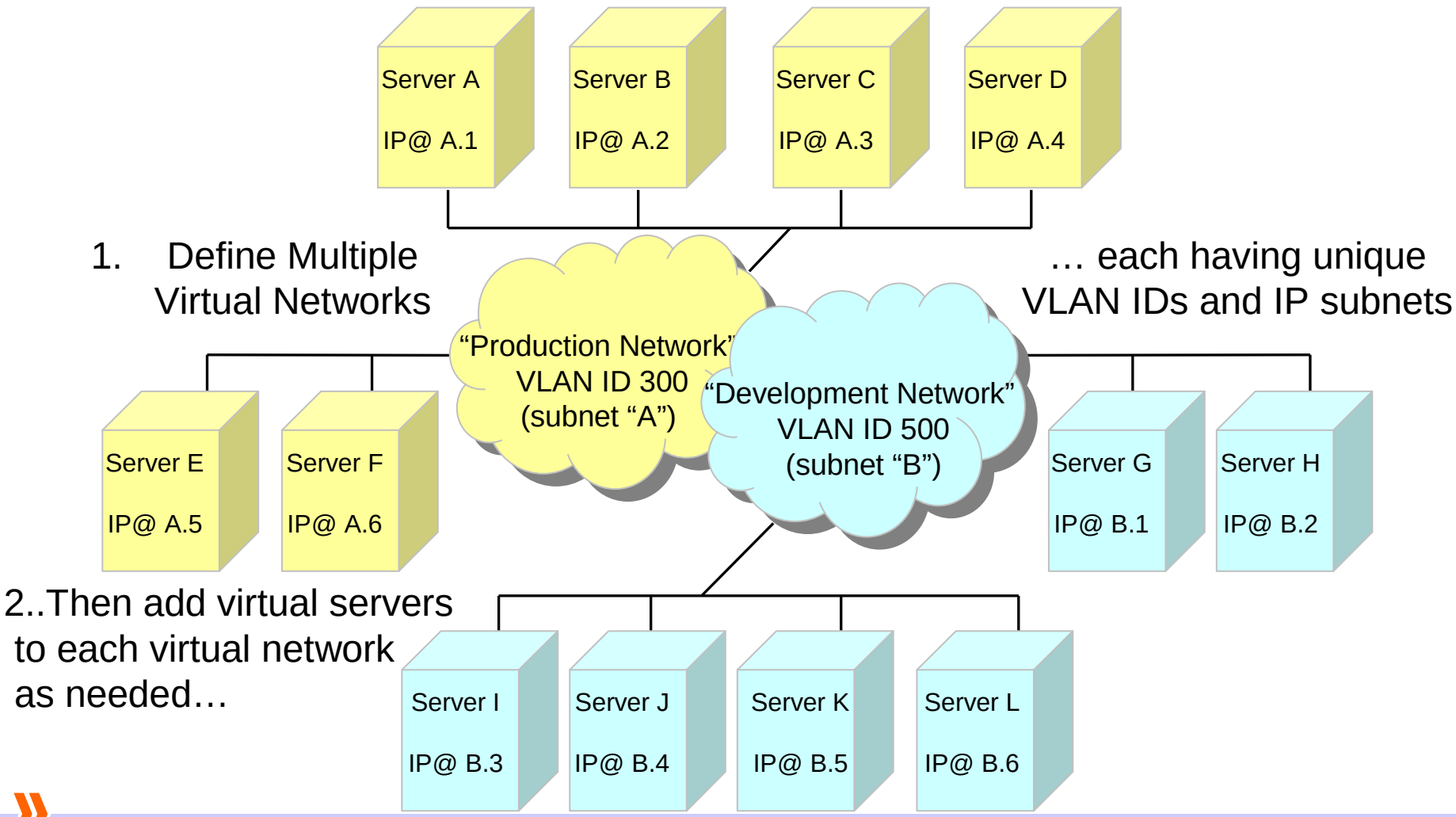
- Virtual local area network (VLAN) purposes
  - Group multiple hosts to a single network, regardless of their physical location
  - Segmentation of a single (physical) LAN into multiple (logical) ones for reasons of
    - Scalability/ performance
    - Security
    - Management flexibility
- VLANs are an industry standard commonly used
- VLAN implementation
  - Using OSI Layer 2 → MAC addresses for communication
  - Uses switches to connect multiple physical networks into one single logical
  - “VLAN Tags” used to assign frames to a specific VLAN

Exploitation of VLANs – Single network



All servers can have a **single IP interface** and all IP address are from the same IP subnet
(e.g. 9.27.200.xxxx)
Multiple Interfaces are created for redundancy!

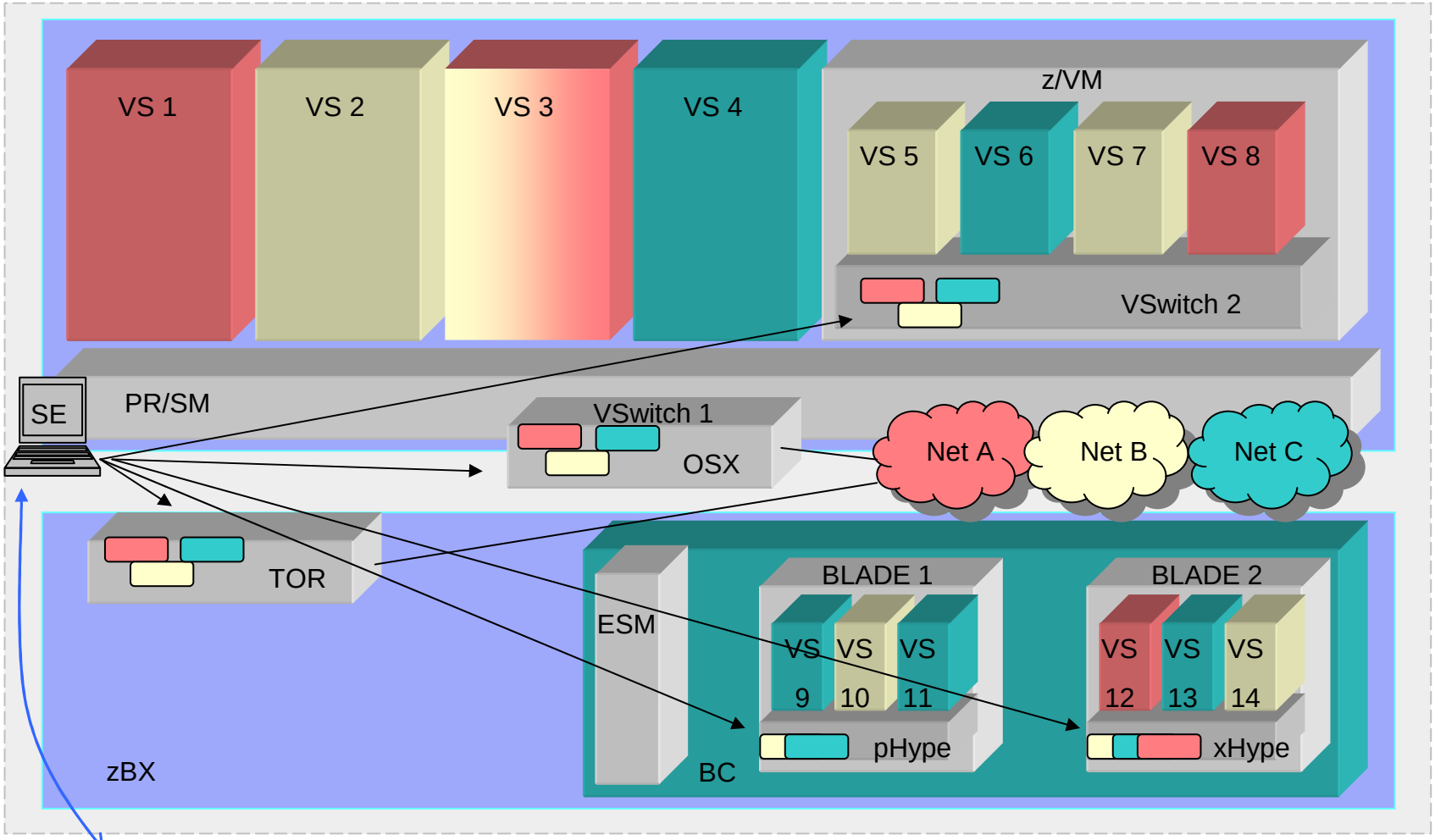
Exploitation of VLANs - Isolation



Note: IBM Unified Resource Manager can be configured to use a single or multiple VLANs. In addition the connecting infrastructure, the switches is also virtualized.

Fully virtualized network infrastructure in an ensemble

- The VLANs on the IEDN are managed using virtual switches (vswitch)



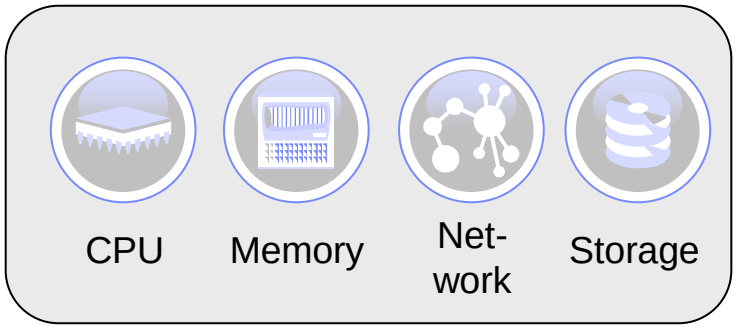
Unified Resource Manager pushes virtual network access control information to the node and the SE propagates to control points (OSX and Hypervisors)

Agenda Part I – Architectural overview

1. System zEnterprise Ensemble overview
2. Hardware building blocks
3. Hypervisors
4. Storage
5. Networks
6. **Virtual Servers**

Virtual server - overview

- Container for customer workload
  - Processor capacity
  - Memory
  - Network
  - Storage
  - Workload information [optional]



Basic characteristics for a virtual server

- Needs an operating system to be installed by customer
- Definitions can be changed dynamically
  - As manual action
  - Based on workload policies [optional]
- Connected to a hypervisor
  - Can be migrated to an other hypervisor of the same architecture
  - Multiple virtual servers can co-exist in the same hypervisor
  - A virtual server is limited to the physical boundary of a Blade

yourdotcom

International Technical Support Organization and Authoring Services

IBM Unified Resource Manager

www.ibm.com/redbooks

Part II – Monitoring and Management

Agenda Part II – Monitoring and Managing

1. **Monitoring**
2. Performance Management
3. Energy Management

Agenda Part II – Monitoring and Managing

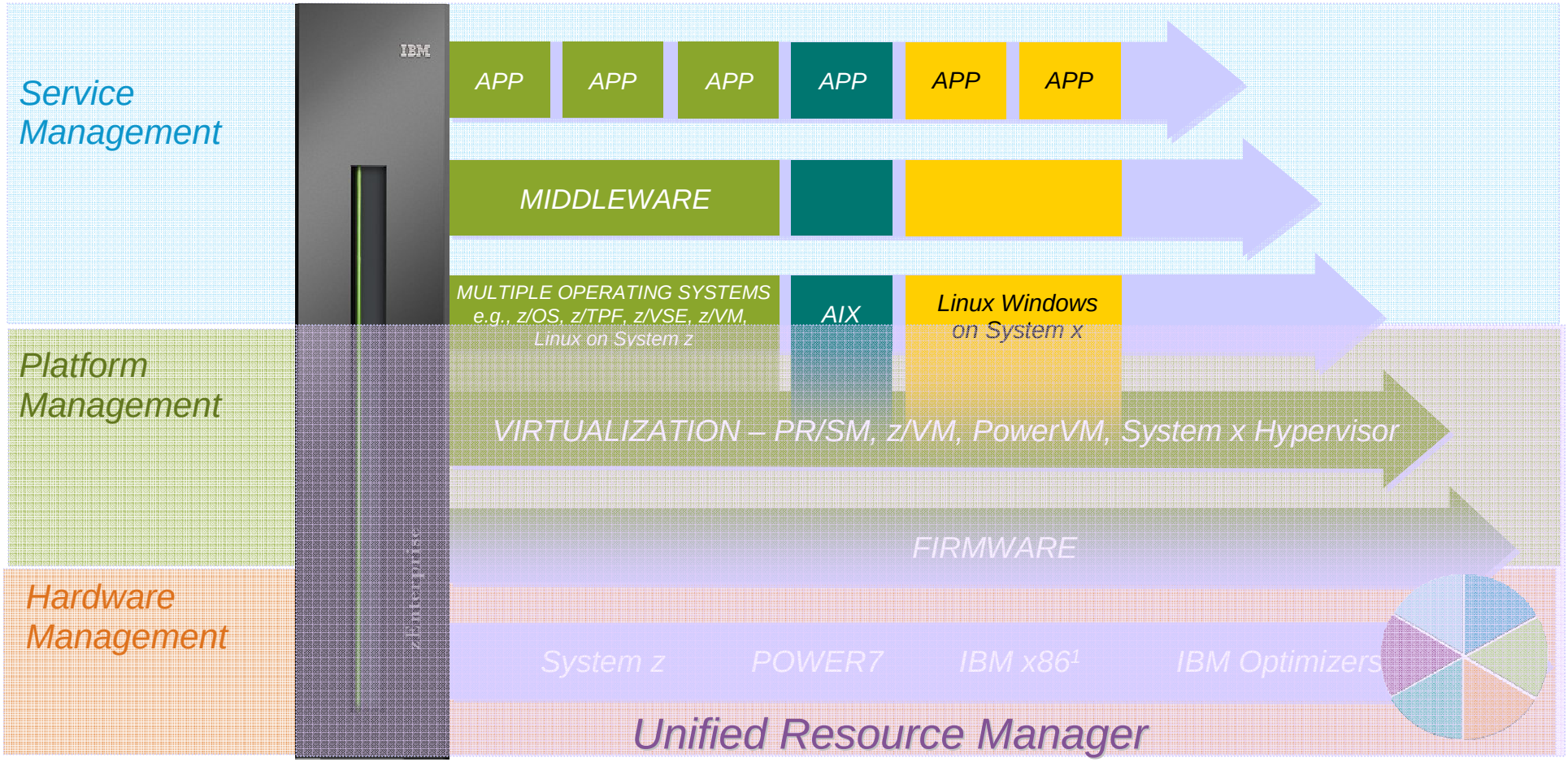
1 **Monitoring**

- *Monitoring levels*
- *Application/ middleware monitoring*
- *Operating system monitoring*
- *Hardware monitoring*
- *Network monitoring – IEDN metrics*

2 Performance Management

3 Energy Management

Monitoring levels in an Unified Resource Manager environment



Focused, collaborative innovation

A "complete systems" approach

Application/ middleware monitoring

- *"Watch the behavior of a specific (typically self-written) applications."*

- Different cases
  - Case 1: Provided by **special monitoring software** that is installed on top of an operating system
    - This monitoring software/ option is certified for specific operating system versions/ levels
    - Agents or scripts that provide information can be installed as usual
  - Case 2: "Embedded applications" like applications in an application server are **monitored by this middleware**
    - This monitoring software/ option is certified for specific operating system versions/ levels
    - Agents or scripts that provide information can be installed as usual

- Unified Resource Manager options
  - The scope of Unified Resource Manager is not applications or middleware therefore no monitoring options are offered for application/ middleware Monitoring

Response time information monitoring through Unified Resource Manager

- If the right infrastructure is available Unified Resource Manager allows customers to monitor response times of single transactions.
- This option is described later on in this presentation

Operating system monitoring

- *"Watch behavior of a specific operating system as a total."*

- Different cases
  - Case 1: Monitoring provided by **agents**
    - Script/ little peace of software than runs inside the operating system
    - Typically certified for specific operating system versions/ levels
    - Tools can be installed as usual
  - Case 2: Monitoring provided by **virtualization layer** (hypervisor)
    - Information is accessed through default options offered by the used hypervisor or through an agent that is installed into the hypervisor by the customers
    - Since the used hypervisors on the Blades is considered License Internal Code that is not accessible, this monitoring option will not work on a zBX

- Unified Resource Manager options
  - Several reports are available in the HMC that give detailed information about resource usage
  - With the use of the Systems API these information can be accessed from external systems (RYO or ISV software)
  - They are described later on in this presentation

Note: Some ISV certify their software only for non-virtualized environments. Therefore check whether or not a virtualized environment is supported. This however is no zBX limitation.

Hardware monitoring

- *"Watch behavior of a specific piece of hardware."*

- Different cases
  - Case 1: Monitoring of "virtualized hardware"
    - Information is accessed through default options offered by the used hypervisor or through an agent that is **installed into the hypervisor** by the customers
    - Since the used hypervisors on the Blades is considered License Internal Code that is not accessible, in general tools can't be used without modification
  - Case 2: Monitoring of non-virtualized hardware
    - Script/ little peace of software than runs **inside the operating system** on a physical hardware
    - Scripts and software can be installed, but data collected does not reflect the physical hardware since the Unified Resource Manager is a fully virtualized environment.

- Unified Resource Manager options
  - Several reports are available in the HMC that give detailed information about resource usage
  - With the use of the Systems API these information can be accessed from external systems (RYO or ISV software)
  - They are described later on in this presentation

Network monitoring – IEDN metrics

- *"Watch physical usage of ports, utilization of bandwidth and vSwitch activity."*

- Different cases
  - Traditional monitoring products and tools
  - Case 1: Monitoring tools (IBM Tivoli Network Monitoring (ITNM), OMEGAMON XE for Mainframe)
    - Traditional monitoring products and tools can still be used to monitor network at Layer 3 (IP)
  - Case 2: Operating system information
    - Operating systems running SNMP can provide standard TCP/IP MIB-based metrics to SNMP-based management tools

- Unified Resource Manager options
  - Monitor IEDN (Layer 2) network resources, collect metrics
  - Display information using the graphical user interface or the system management API
  - Network monitor dashboard
  - Metrics shown
    - Performance statistics (transfer rates over intervals)
    - Cumulative and interval metrics
    - Raw metrics as collected
    - Resource state/status (where available)
    - Resource utilization relationships (e.g., Servers using an OSA)

Agenda Part II – Monitoring and Managing

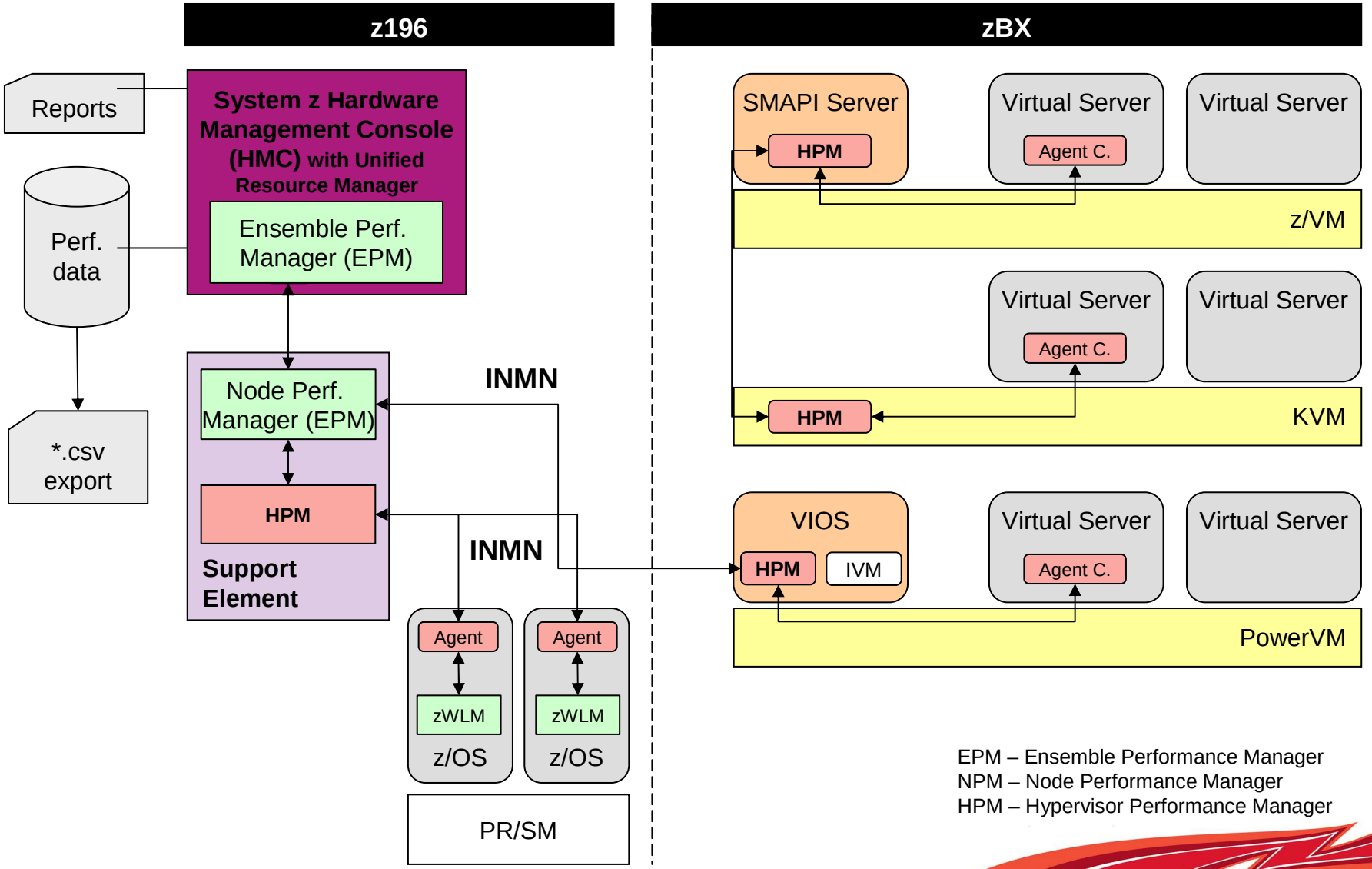
1 Monitoring

2 Performance Management

- ***Overview***
- *Monitoring/ reporting*
- *Implementation*

3 Energy Management

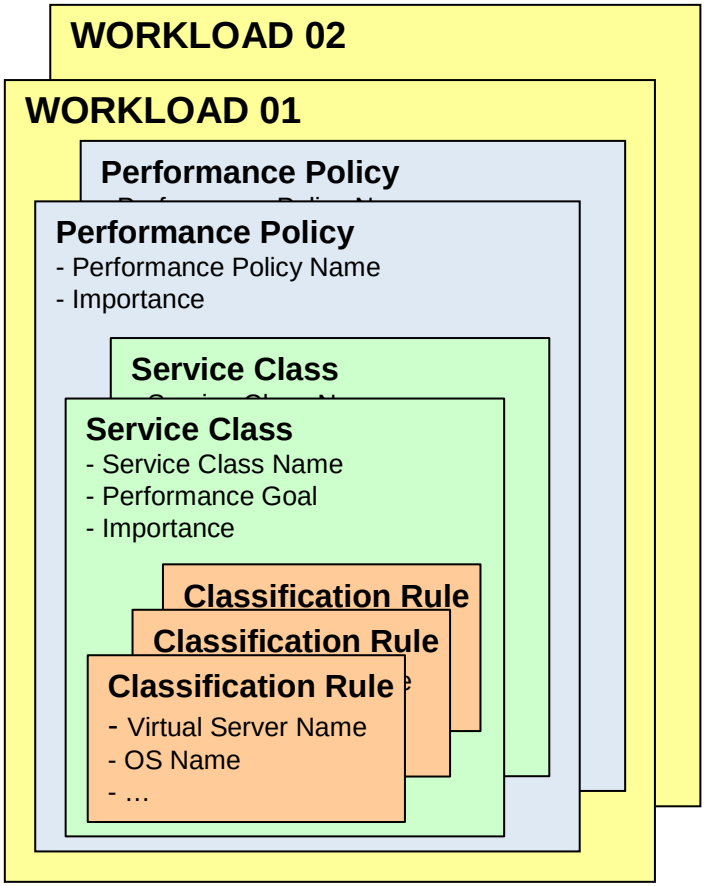
Workload Management – The big picture



EPM – Ensemble Performance Manager
NPM – Node Performance Manager
HPM – Hypervisor Performance Manager

Workload Management – The basic concepts

- Workload
  - Holistic description of performance goals for resource collection
  - *"What kind of production is managed?"*
- Workload Performance Policy
  - Different performance descriptions for a workload, for different times
  - *"When is my production managed?"*
- Service Class
  - Actual definition of performance goals
  - Velocity goal/ discretionary
- Classification Rule
  - Rule set for classifying incoming work to a service class
- Performance index
  - Indication on how good goals are met



Workload Management – New concepts

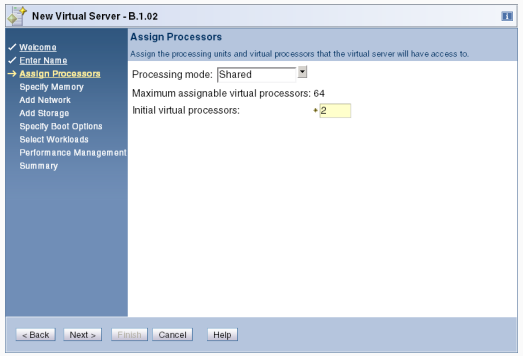
- Business importance (Policy & Service class)
  - Defined for workloads and service classes
  - Sets the importance of a workload/ service class compared to other workloads/ service classes
  - Used in a resource contention situation to check whether a server will donate/ receive resources
  - Possible (relative) values set by customer
    - Highest
    - High
    - Medium <*default*>
    - Low
    - Lowest

- Virtual server management instead of application management
  - Granularity of the Unified Resource Manager‘s workload management is a single virtual server
  - Depending on the Performance Index (PI) the whole virtual server can become eligible to receive or donates resources in a contention situation
    - PI =< 1 → server can donate resources
    - PI > 1 → server needs more resources

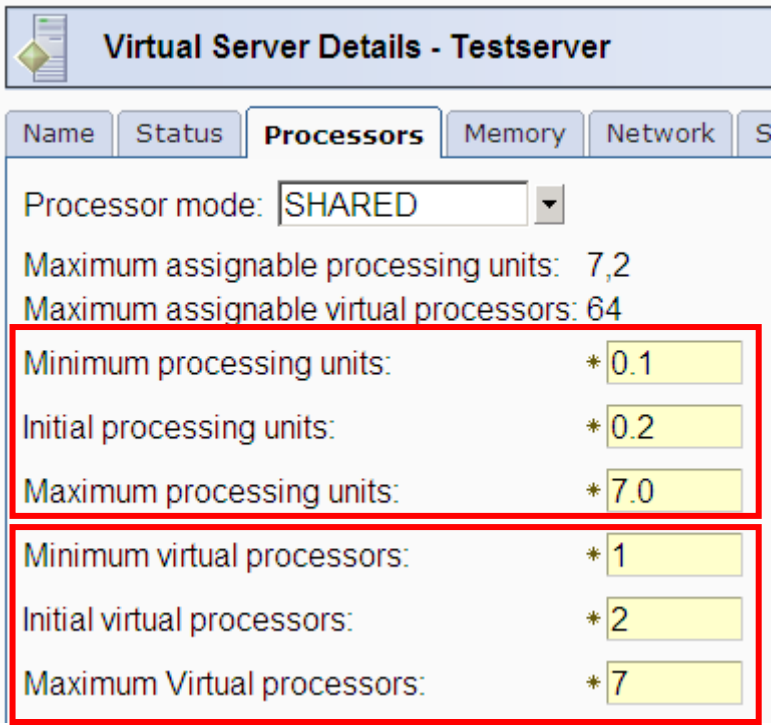
Note: Workload management is not available for optimizers in an ensemble. Only Application Serving Blades will benefit from the management capabilities.

The truth on virtual servers CPU information

- During virtual server creation only one value is used to specify the processor power



- When looking at the server details after server creation there are multiple values
  - Processing units specify the pure processing power in terms of CPU time slices
  - Virtual processors describe the parallelism for processing a virtual server
  - Minimum/ Maximum set the lower and the upper limit for workload management
  - Initial value is the starting value when a server is activated



Resource adjustment process

- General statements
  - Interacting servers can be assigned to different workloads
  - Virtual servers need to be defined in the same hypervisor (the same hardware) to interact with each other
  - A server can never shrink below his amount of „minimum“ processing resources defined
- Non-contention situation
  - *„A virtual server as a PI>1 and needs resources. Free resources are available in the processor pool.“*
  - Assign resources from the free pool to the virtual server
- Contention situation
  - *„A virtual server as a PI>1 and needs resources. No free resources are available in the processor pool.“*
  - Scenario 1: Other virtual servers with PI<1 available
    - ? Receiver has less then its maximum processing units defined after adjustment?
    - ? Donor will still have PI =< 1 after adjustment?
    - ? Donor will not shrink below its minimum processing units defined?
    - Re-arrange processing units from other virtual servers without any server missing their goals
  - Scenario 2: No virtual servers with PI<1 available or no resources left to donate
    - ? Receiver has less then its maximum processing units defined after adjustment?
    - ? Donor will not shrink below its minimum processing units defined?
    - Searches for „discretionary“ defined servers
    - Compares importance for business importance of service class and workload

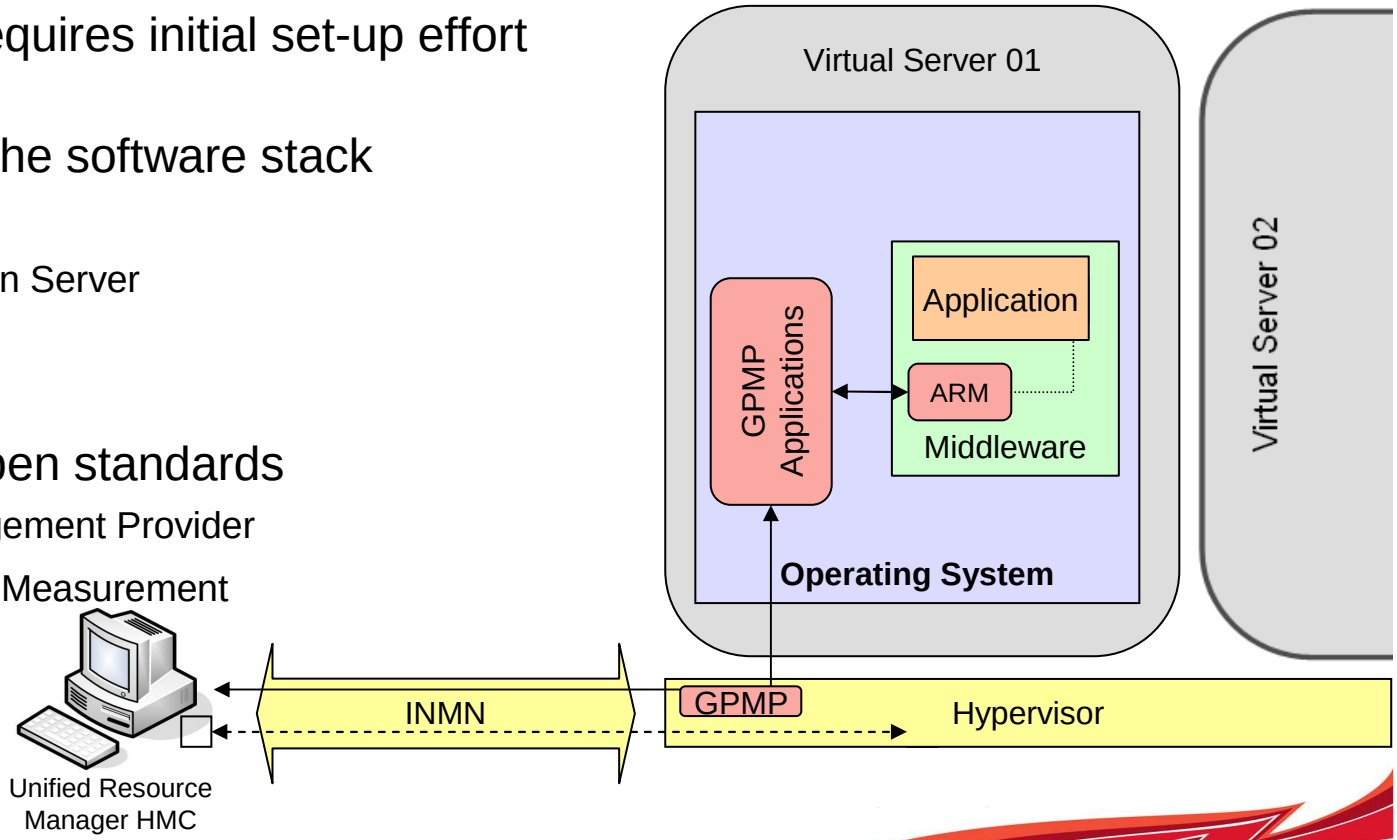
Comparing importance for virtual servers

- Internally Unified Resource Manager maps the importance to a numeric value
  - Highest 1
  - High 2
  - Medium 3
  - Low 4
  - Lowest 5
- Values are then aggregated for the performance policy (HLQ) and the service class (LLQ)
- **Lower aggregated value means higher importance for the donation process**

Server name	*Description*	*Policy name and business importance*		*Service class name and business importance*		*Aggregated importance value*
PRODONLIN01	Server in daytrader production environment used for background processing.	DAYSHIFT	HIGHEST	ONLINE	HIGH	HIGHEST + HIGH 1 + 2
PRODBATCH01	Server in daytrader production environment used for background processing.	DAYSHIFT	HIGHEST	BATCH	LOW	HIGHEST + LOW 1 + 4
TESTSERV01	JATS Environment SW quality for daytrader development	TESTALL	LOW	TESTTRADE	HIGHEST	LOW + HIGHEST 4 + 1

Workload monitoring using GPMP and ARM

- Enhanced workload management capabilities can be used with Unified Resource Manager
  - Monitoring of transactions
  - Finer grained classification rules for workload management
  - Additional reports for performance monitoring
- No charge option, requires initial set-up effort
- Usage depends on the software stack
  - IBM http server
  - WebSphere Application Server
  - DB2
- Used technology/ open standards
  - Guest Platform Management Provider
  - Application Response Measurement



Workload management depends on the available suite

- Repetition: Product bundles available for Unified Resource Manager
  - Manage suite
    - FC0019
    - Offers all basic capabilities needed to built and operate an ensemble
  - Automate suite
    - FC0020
    - Adds management option for energy related tasks
    - Adds workload management capabilities

- Workload management based functions available depending on the suite

Function	*Manage suit*	*Automate suite*
Service class default available	Yes, only service class available	Yes
Define own service classes	No	Yes
Use Classification rules	No	Yes
Reporting options based on	<*default*> service class only	Individual service classes
Automatic resource adjustments for virtual server	No	Yes
Manual adjustments to CPU resources	Yes	Yes

Note: The Automate suite (FC0020) in general offers more capabilities for an ensemble. Future functions will also exploit the automate suite.

Agenda Part II – Monitoring and Managing

1 Monitoring

2 Performance Management

- *Overview*
- ***Monitoring/ reporting***
- *Implementation*

3 Energy Management

Workload management – Available reports

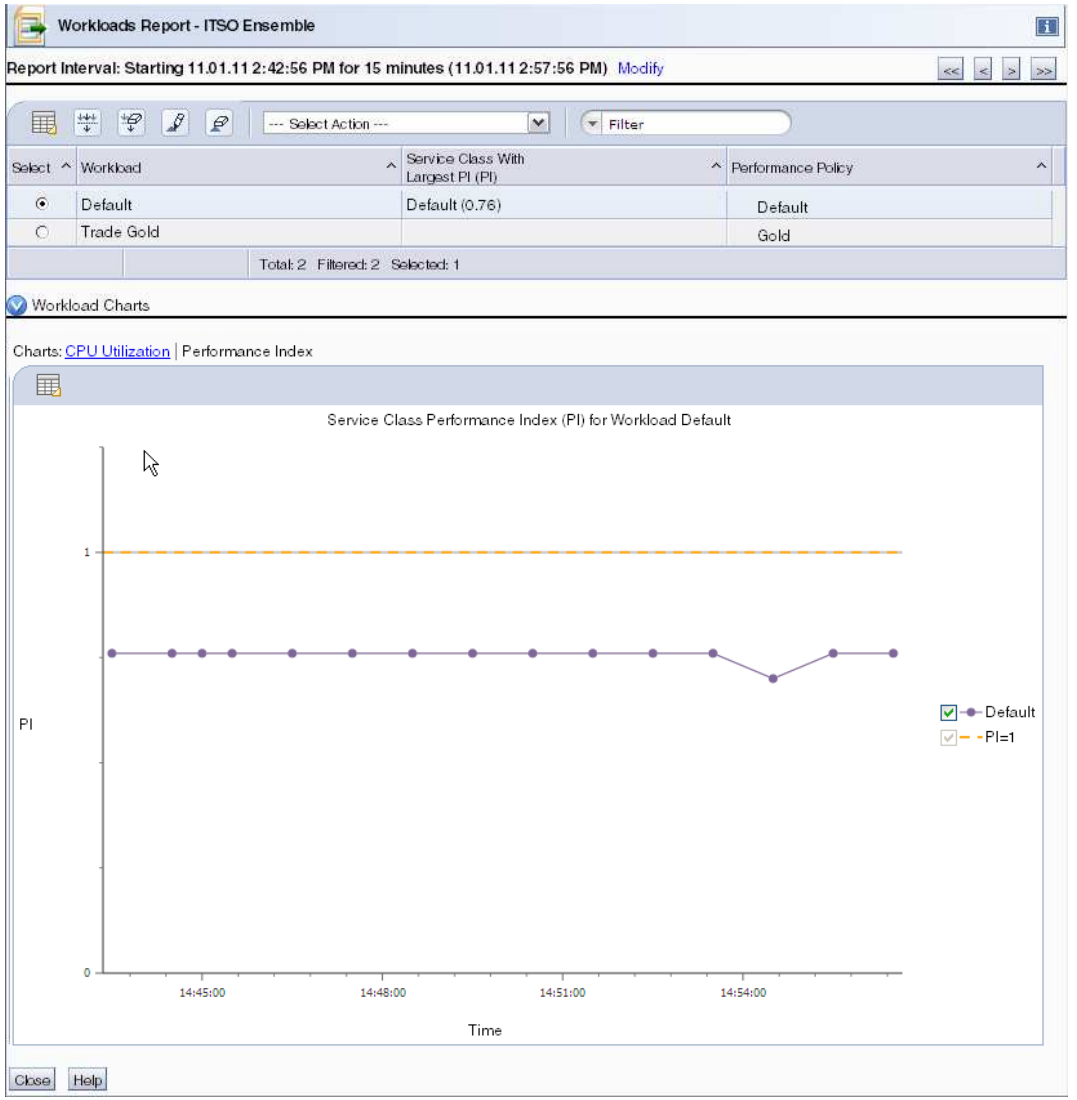
- Information provided in graphical way, but can be exported as "*.csv" to local workstation
- Data can be collected through the Unified Resource Manager API offered through the HMC
- Reports show per default last 15 minutes of activity
  - Intervals can be changed to display different intervals
- General reports <*explained in detail*>
  - Workloads report
  - Service class report
  - Virtual server report
  - Hypervisor report
  - Resource adjustment report
- New reports through GPMP
  - Topology report
  - Hops report → showing which server takes what response time

Implementation hint

The task role "Performance Management Operator" is required to access the full range of available reports.

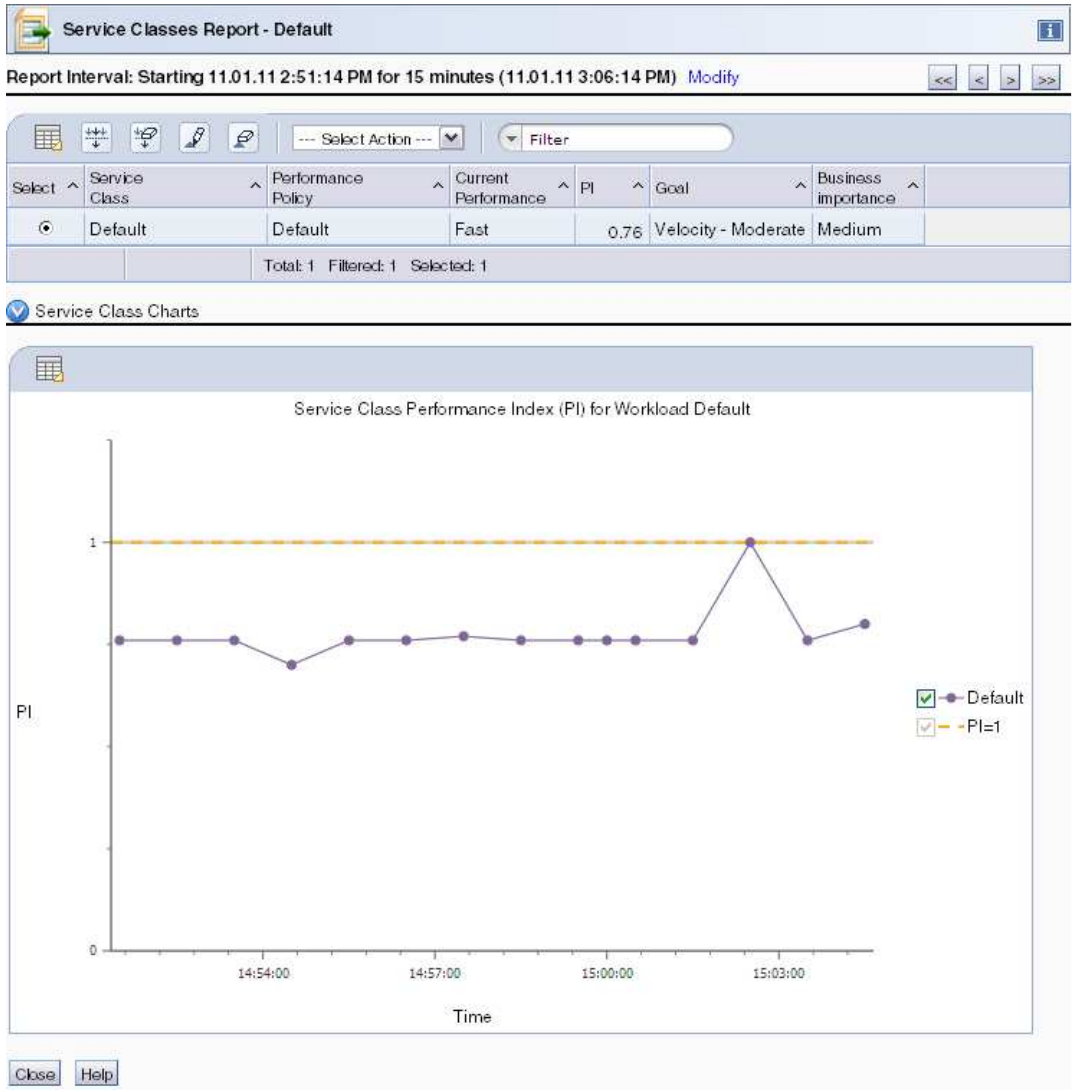
Workload management – Workload Report

- *"Indicating the worst service class of a workload!"*
- High level system view
- Further information in charts
  - Performance index for service classes
  - CPU utilization of virtual servers in the workload
- Links to other reports
  - Monitor System Events
  - Workload Resource Adjustments Report
  - Virtual Servers Report
  - Service Classes Report



Workload Management – Service class report

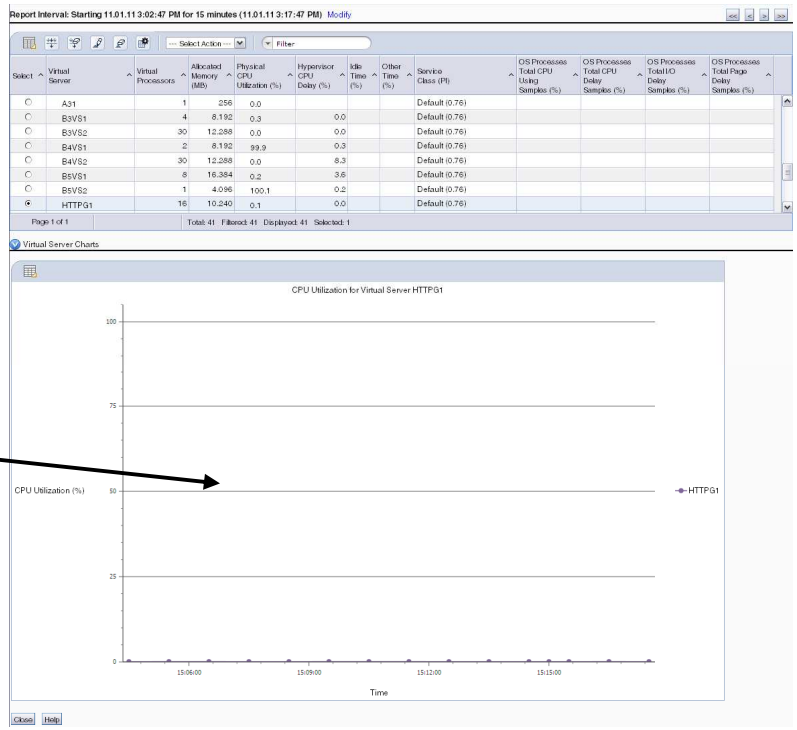
- *"Does my service class achieves it's goal?"*
- For each service class a detailed performance index chart can be displayed
- Information displayed
  - Current achieving
  - History information (up to 36 hours)



Workload management – Virtual server report

- *„Everything I need to know on a specific server for the last interval and a short CPU usage history“*

- Information displayed
  - Virtual Server
  - Hypervisor
  - Hypervisor Type
  - OS Name
  - OS Type
  - OS Level
  - Hostname
  - Virtual Processors
  - Allocated Memory (MB)
  - **Physical CPU Utilization (%) [Hypervisor view]**
  - Hypervisor CPU Delay (%) [Hypervisor view]
  - Idle Time (%) [Hypervisor view]
  - Other Time (%)
  - Service Class (PI)
  - **OS Processes Total CPU Using Samples (%)*
  - ** OS Processes Total CPU Delay Samples (%)*
  - ** OS Processes Total I/O Delay Samples (%)*
  - ** OS Processes Total Page Delay Samples (%)*



- For each virtual server a detailed CPU utilization chart can be displayed

* Information available only when using GPMP

Workload management – Hypervisor report

- *„What virtual servers are running in a specific hypervisor and how are their current CPU values?“*

- Information displayed
  - Virtual server name
  - # processors
  - Processor usage
  - Allocated memory
  - Dedicated/ shared processors
  - CPU weight
  - Min/ max CPU

- Exact type of information depends on the hypervisor type (PR/SM, PowerVM, KVM, z/VM) but is always the whole ensemble information of the virtual server

Hypervisor Report - A01

Report Interval: Starting 11.01.11 3:39:24 PM for 15 minutes (11.01.11 3:54:24 PM) Modify

Hypervisor Details:

Hypervisor:	SCZP301	Processor count:	38	Total memory allocated for virtual servers:	164,864 MB
Hypervisor type:	PR/SM	Total shared CP consumption:	3.0%	Total memory:	540,672 MB

Virtual Servers:

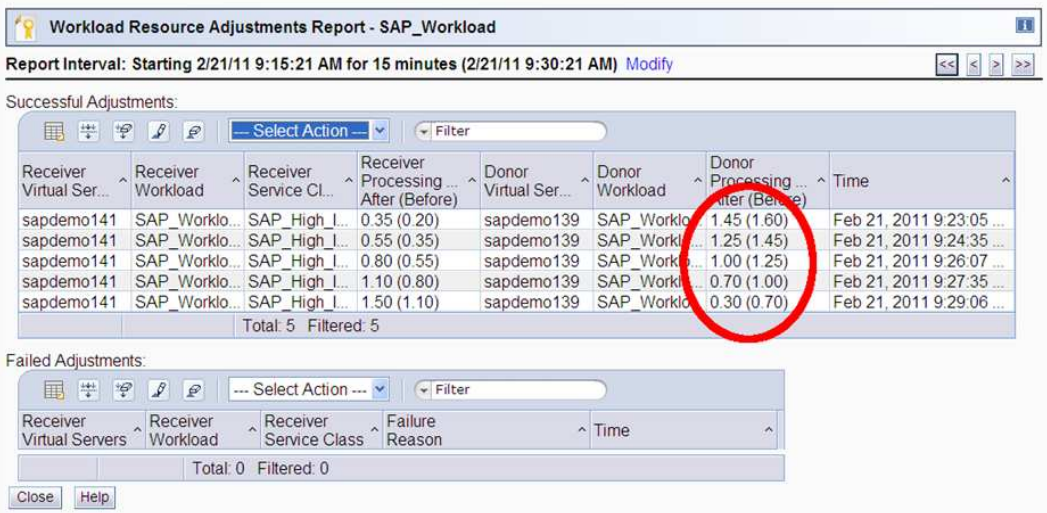
--- Select Action --- Filter

Virtual Server	Logical Processors	Consumed Processors	Allocated Memory (MB)	Dedicated	CPU Weight	Min CPU Weight	Max CPU Weight
A01	2	0.00	8,192	–	10	0	0
A02	3	0.00	6,144	–	10	0	0
A03	2	0.00	8,192	–	10	0	0
A04	2	0.00	8,192	–	10	0	0
A05	2	0.00	8,192	–	10	0	0
A06	2	0.00	8,192	–	10	0	0
A07	2	0.00	6,144	–	10	0	0
A0B	1	0.00	256	–	10	0	0
A0D	1	0.00	1,024	–	10	0	0
A0E	1	0.00	1,024	–	10	0	0
A0F	1	0.00	1,024	–	10	0	0
A11	4	0.00	4,096	–	10	0	0
A12	4	0.00	6,144	–	10	0	0
A13	4	0.00	4,096	–	10	0	0
A16	4	0.00	4,096	–	10	0	0

Page 1 of 1 Total: 32 Filtered: 32 Displayed: 32

Workload management – Resource adjustment report

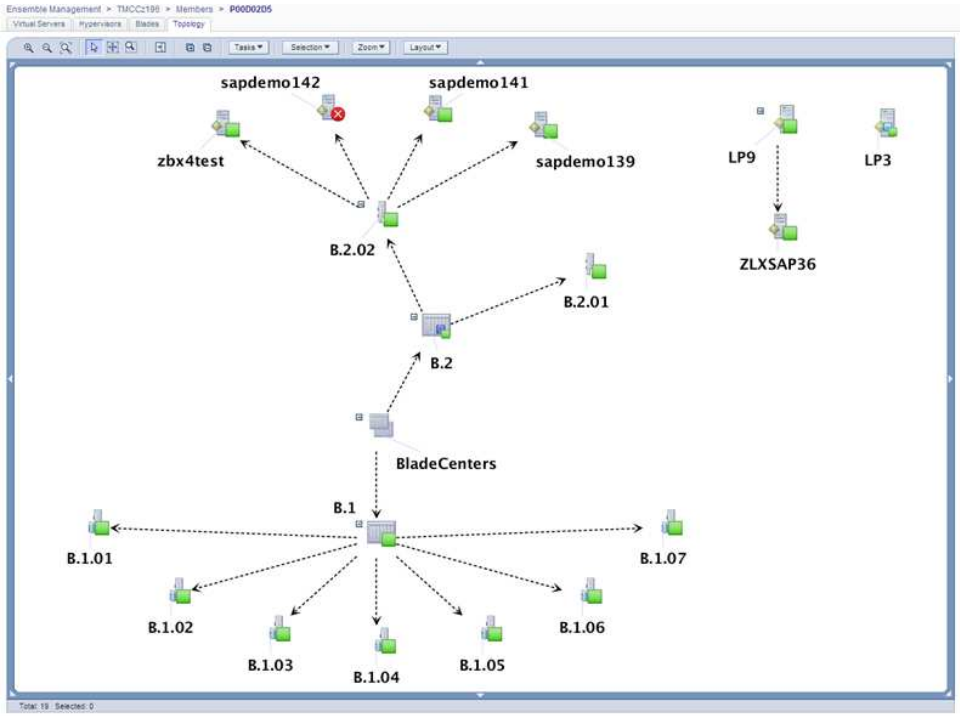
- *„Do my workload settings work and should I adjust the CPU min/max values for a server?“*
- Lists the resource adjustments depending on the result
  - Successful
  - Failed
- Includes capacity information before and after adjustment
- For failed adjustments the failure reason is displayed for better tuning
  - Unknown
  - Not enough capacity
  - No potential donors
  - Entitled capacity not achievable
  - Processor is not fully utilized
  - More shares than max requested
  - Not enough virtual CPUs



Note: As long as there are free resources in the pool, you can‘t see donating adjustments in this report. Only in an resource contention situation this report shows donations.

Workload management – Topology report

- *"What does my ensemble infrastructure look like?"*
- Only available when GPMP and ARM is used
- Only accessible from the Service Class Report
- Shows relationships between virtual servers that are running the workload and providing the resources to complete the work.



Implementation hint: Project view

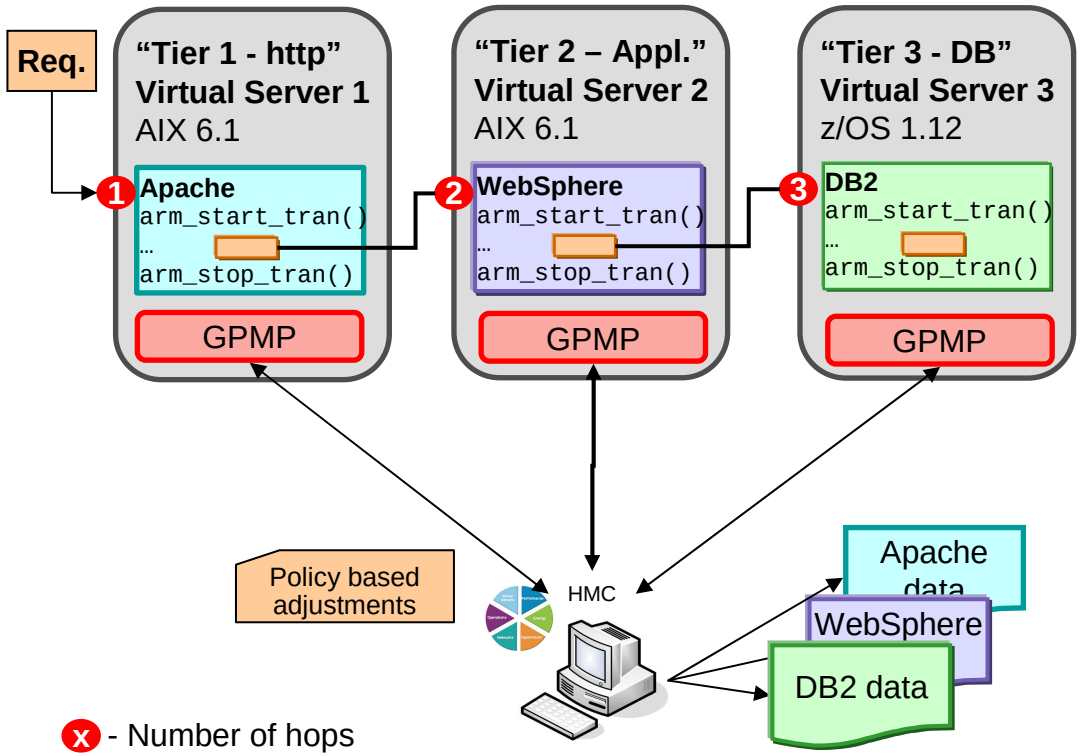
Depending on the size of an infrastructure the picture becomes quite complex → Implement a user ID that only sees parts of the overall infrastructure

Workload management – Performance reporting using ARM

- Usage of ARM enabled middleware with GPMP allows detailed performance monitoring

- Response times for any step (or hop) in an infrastructure

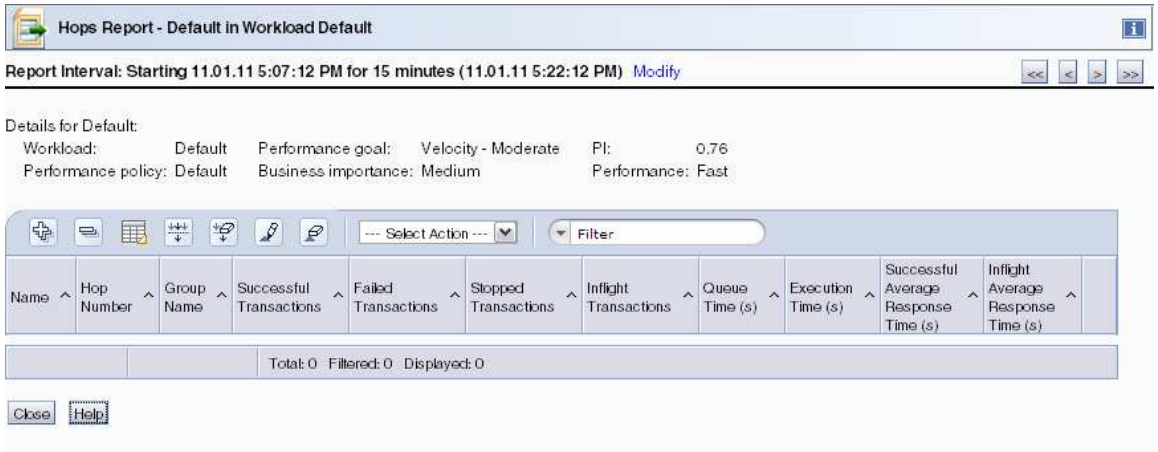
1. End-user request enters system
2. **HTTP server is 1. hop**
3. Apache processes request and starts a new transaction, calling WebSphere
4. **WebSphere (WAS) is 2. hop**
5. WebSphere processes transaction and calls DB2
6. **DB2 is 3. hop**
7. DB processes transaction and flow goes backwards, to “parent” WAS, and from there to HTTP and then to end-user



“Parents” do not count as additional hop, since they have been already used.

Workload management – Hops report

- *"Where is my request and how does it perform?"*
- Only available when GPMP and ARM is used
- Shows the different steps (hops) that a request made between different virtual servers
- Information listed
  - Information about the service class the report focuses on
  - Name
  - Hop Number
  - Group Name
  - Successful Transactions
  - Failed Transactions
  - Stopped Transactions
  - In-flight Transactions
  - **Queue Time(s)**
  - Execution Time(s)
  - **Successful Average Response Time(s)**
  - **In-flight Average Response Time(s)**



Agenda Part II – Monitoring and Managing

1 Monitoring

2 Performance Management

- *Overview*
- *Monitoring/ reporting*
- ***Implementation***

3 Energy Management

Tasks to be performed for workload management

Gather requirements	Planning	Implementing	Testing & Tuning
1. Identifying the requirements for ▪ Performance ▪ Monitoring ▪ Reporting	2. Prerequisites 3. Server sizing 4. Number of workloads 5. Different policies 6. Number and definition of service classes	7. Create Workload 8. Assign virtual servers 9. Create Performance Policy 10. Create Service Class 11. Create Classification Rule 12. Activate policy 13. Plan for scheduled operations	14. Test the environment using the usual test methods 15. Tune/ adjust environment 16. Repeat tests until they match the requirements

The planning phase is described in more detail on the following slides, implementation is covered during the walk-through

Workload management – Software requirements

- Supported operating systems
  - z/OS Version 1.12, 1.11, or 1.10 with PTF‘s
  - z/VM 6.1 with PTF's
  - Red Hat Enterprise Linux (RHEL) 6, 5.4 or 5.3
  - Novell SUSE Linux Enterprise Server (SLES) 11 or 10
  - AIX 6.1 TL 5 or V5.3 TL12

- Supported middleware
  - WebSphere Application Server 6.1 or 7.0 (z/OS and Multiplatform)
  - Web server plug-ins that the WebSphere Application Server version provides or supports
    - IBM HTTP Server and Apache HTTP Server
    - Microsoft Internet Information Services (IIS)
    - Lotus Domino®
  - DB2
    - DB2 Database for Linux, UNIX, and Windows Version 9
    - DB2 for z/OS Versions 9 and 10

Server sizing still needed as prerequisite

- Workload management allows shifting CPU resources within a specified range
  - It doesn't do magic
  - The boundaries need to be set by administrator

- **Two scenarios**
  - Migrating an existing environment
    - Reuse the performance information available
  - Greenfield approach
    - Contact your IBM representative to assist you in performing a system planning/sizing

- For further workload management planning you need to define the minimum and maximum boundaries for the consumable CPU resources on a server base
  - These values may be a starting point and have to be verified(adjusted in a test & tune phase

Planning – Workload assessment and planning

- Assess the environment
  - Applications and the platforms present
  - Workload types
  - Service Level Agreements (SLA)
  - Needed performance due to workload behavior

- Plan for workloads
  - Determine the number of workloads
  - Map the virtual servers to each workload
  - Check what performance management functions are needed for each workload or server
    - Automatic CPU management
    - Data collection using GPMP

Planning – Performance policies

- Determine the needed performance policies for each defined workload
- Compare the importance of all defined policies for all workloads and sort them into the available categories
  - Highest
  - High
  - Medium <default>
  - Low
  - Lowest

Planning – Service classes

- Plan for the service classes within each policy
  - Name
  - Goal type (velocity/ discretionary)
  - Importance
  - What servers should be assigned to them

- Plan for the classification rules
  - Unified Resource Manager needs a characteristic to decide in which service class a request will be put
  - For each service class add
    - Filter type
    - Filter value

Workload management related user roles and tasks

- HMC Task roles
  - Workload administrators
    - Create and manage workloads and
    - Add or remove virtual servers to workloads
  - Performance management administrators
    - Create or modify performance policies
    - Create or modify service classes for each policy
  - Performance management operators
    - Activate a policy
    - Schedule policy activation
    - Define thresholds for event notification

- HMC Resource roles
  - IBM Blade Virtual Server Objects
  - Workload Objects
  - z/VM Virtual Machine Objects

Agenda Part II – Monitoring and Managing

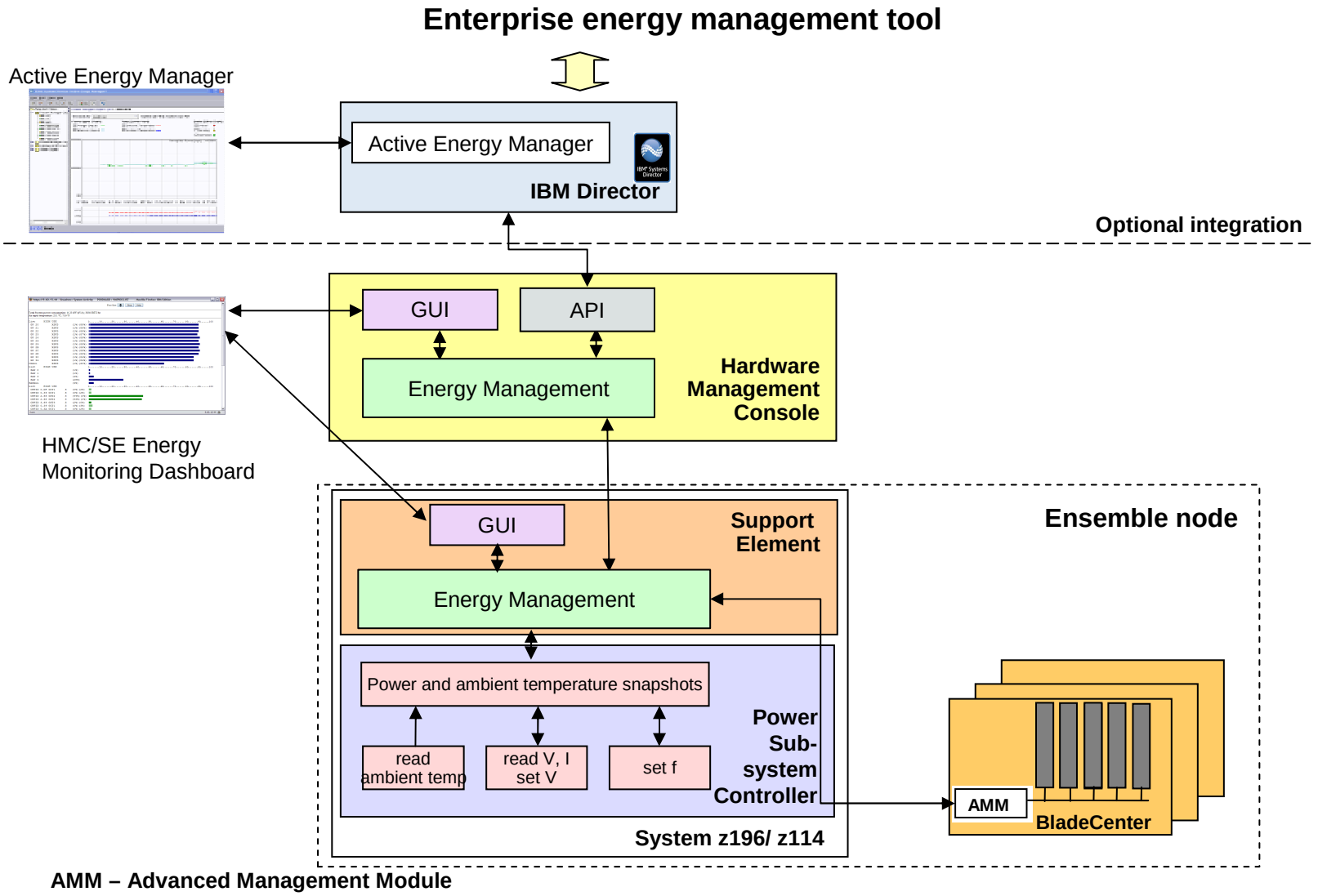
1 Monitoring

2 Performance Management

3 Energy Management

- ***Overview***
- *Monitoring capabilities*
- *Management capabilities*

Energy Management – The big picture



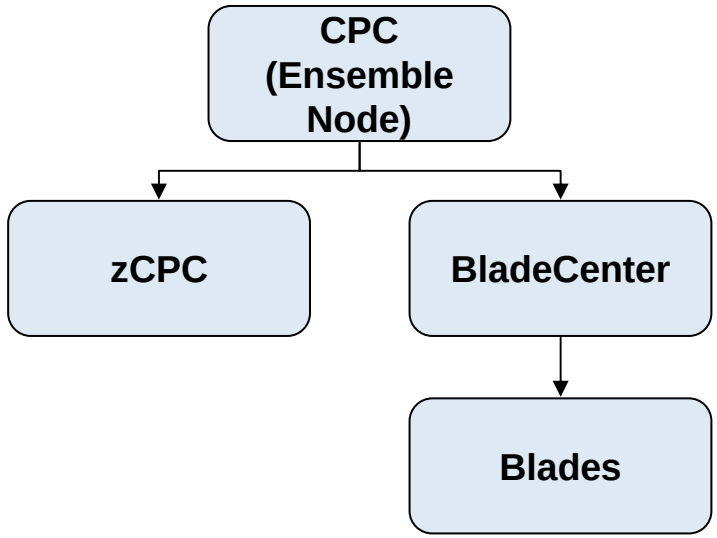
AMM – Advanced Management Module

Energy Management – Group concept

- Group definition

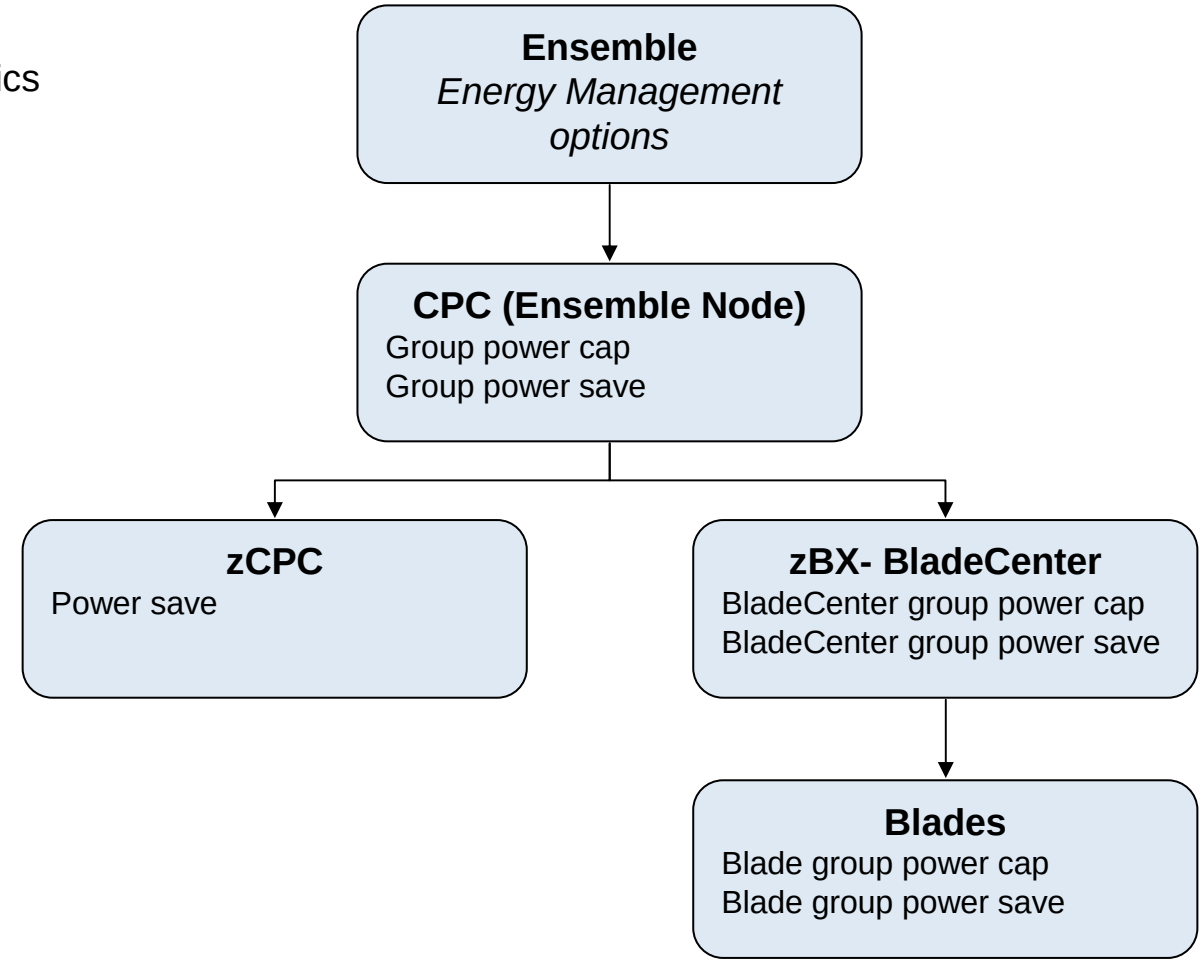
> *A group is a set of similar objects that are included in a higher-level object.*
>
> Definition

- Characteristics
  - Group settings replace individual settings
  - Group settings are done implicitly by changing the higher-level object
  - To disable group settings
    - Change one member of the group
    - Change high-level member



Energy Management – Capabilities

- Energy monitoring
  - *Manage suite (FC0019)*
  - Monitors dashboard
  - Environmental Efficiency Statistics
  - Details view
- Energy management
  - *Automate suite (FC0020)*
  - Power Capping
  - Power Save
- Management capabilities can be combined with the „Scheduled Operations“ task in the HMC



Agenda Part II – Monitoring and Managing

1 Monitoring

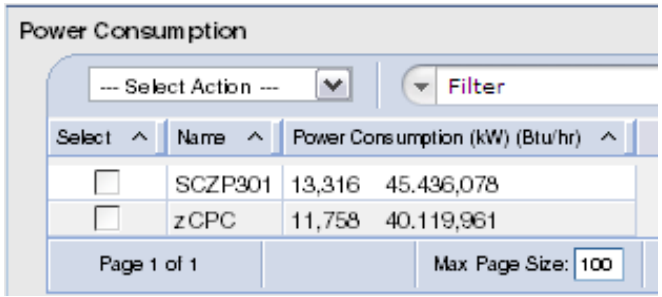
2 Performance Management

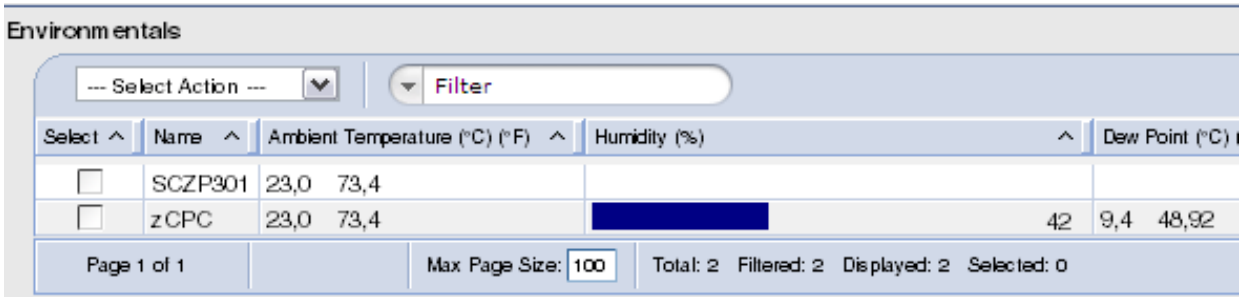
3 Energy Management

- *Overview*
- ***Monitoring capabilities***
- *Management capabilities*

Energy monitoring – Monitors dashboard

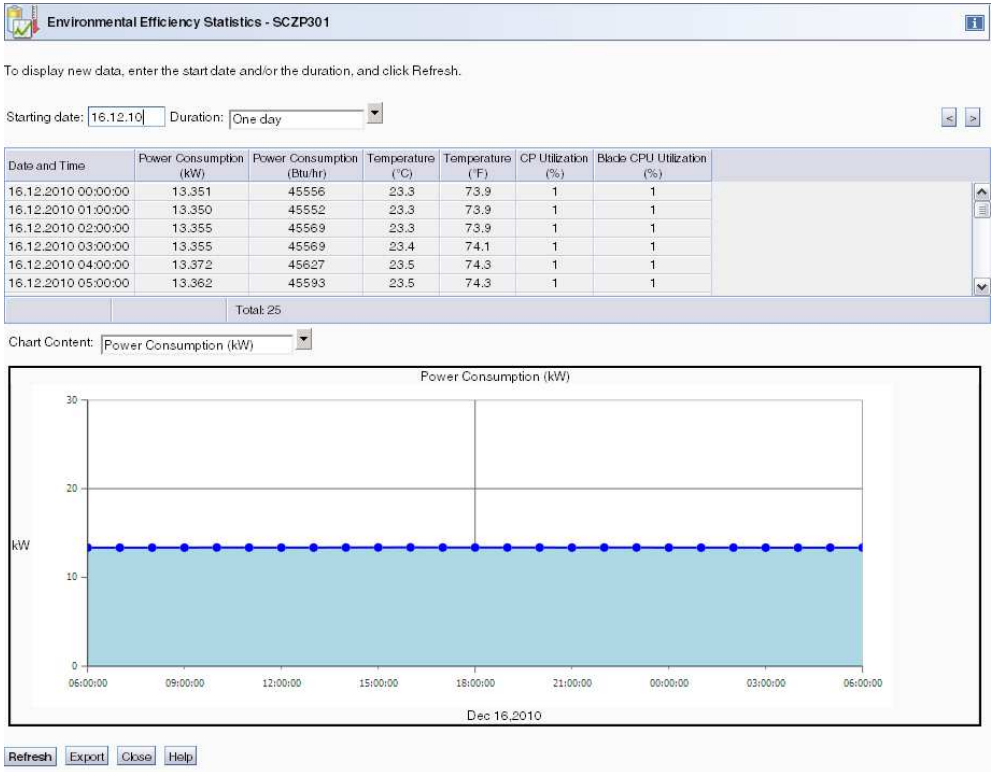
- Accessed using the HMC
  - User needs “System Programmer Tasks” task role
- Real-time data for
  - Power consumption
    - Ensemble
    - zCPC
    - zBX (BladeCenters Blades)
  - Air Temperature
- History monitoring
  - Can be exported as *.csv file
  - Covers last 36 hours





Energy monitoring – Environmental efficiency statistics

- Accessed using the HMC
  - User needs “System Programmer Tasks” task role
- Displays table & graph for up to one week
- Real-time data for
  - Power consumption (kW and Btu)
  - Temperature (Celsius and Fahrenheit)
  - CP utilization percentage
  - Blade CPU utilization percentage
- History monitoring
  - Can be exported as *.csv file



Energy monitoring – Detail view

- Shows real-time data for power and heat
- Available at different levels
  - Ensemble node (Total/ z196/ z114)
  - BladeCenter (not zBX)
  - Blades
- User ID needs permission to view the HW blocks to see the details

Ensemble details view

CPC

Power rating:	36844 W
Power consumption:	13258 W
Power saving:	Not supported
Power capping:	Custom
Cap range:	35400 W - 115050 W
Current cap:	115050 W

zCPC

Power rating:	27400 W
Power consumption:	11726 W
Ambient temperature:	22.8°C (73.0°F)
Exhaust temperature:	33.0°C (91.4°F)
Humidity:	41 %
Dew point:	9.1°C (48.4°F)
Heat load:	40039 BTU/hr.
Heat load (forced-air):	40039 BTU/hr.
Heat load (water):	0 BTU/hr.
Maximum potential power:	11366 W
Maximum potential heat load:	38808 BTU/hr.
Power saving:	High performance
Power capping:	Disabled
Cap range:	11366 W - 27400 W
Current cap:	27400 W

BladeCenter details view

Instance Information | Acceptable Status | Product Information | **Energy Management Information**

Power rating:	9444 W
Power consumption:	1570 W
Ambient temperature:	20.0°C (68.0°F)
Exhaust temperature:	26.5°C (79.7°F)
Power saving:	Not supported
Power capping:	Custom
Cap range:	17954 W - 9444 W
Current cap:	9444 W

Blade details view

Instance Information | Acceptable Status | Product Information | **Energy Management Information**

Power rating:	2834 W
Power consumption:	155 W
Power saving:	Not supported
Power capping:	Enabled
Cap range:	2835 W - 2834 W
Current cap:	300 W

Agenda Part II – Monitoring and Managing

1 Monitoring

2 Performance Management

3 Energy Management

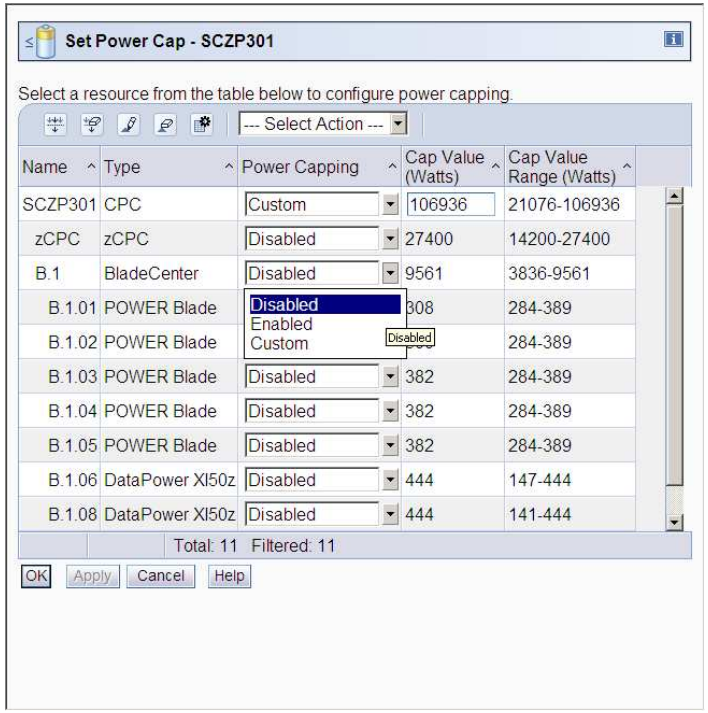
Overview

Monitoring capabilities

Management capabilities

Energy Management – Power cap

- Limit the maximum amount of energy used
- Intended for dynamic workload environments
- Systems hardware will adjust processing speed to reduce the amount of energy used
- Can be set on different levels
  - Ensemble node
  - BladeCenter
  - zCPC → Query Max Potential Power



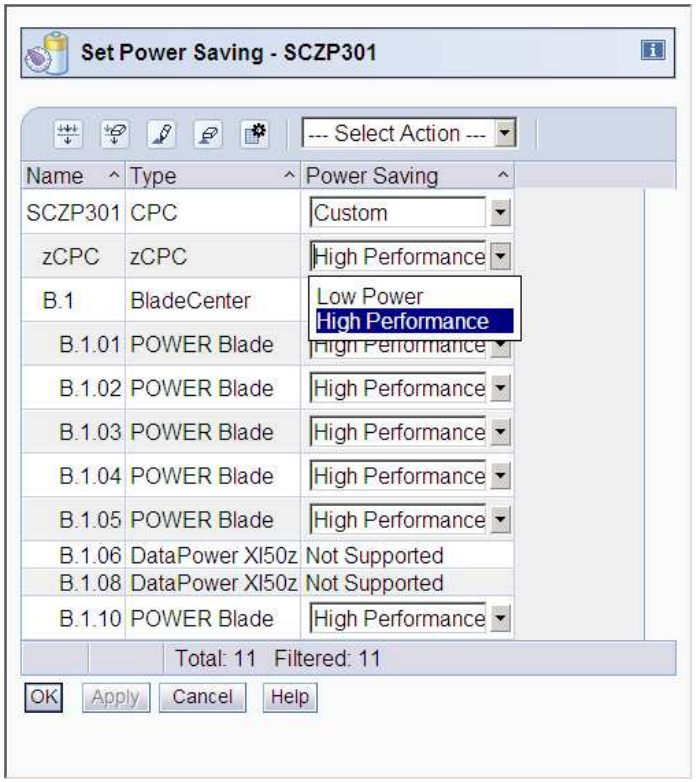
Displayed min/max values

Capping values listed as min/max are higher than the ones in the technical specs.

This is the result of a hot-room test performed in a datacenter that is heated up artificially to represent a worst-case-scenario.

Energy Management – Power saving mode

- Place systems in a power saving mode, reducing the CPU frequency
- Intended to reduce power in a static way over a longer period of time → *during night time*
- Possible settings
  - High performance
  - Low performance



Attention

- The power saving mode can not be used on a z114.
- If a z114 is part of an ensemble the power save mode the option can't be set for the zCPC in this node, but still for the other components of the ensemble

yourdotcom

International Technical Support Organization and Authoring Services

zEnterprise
Unified Resource Manager

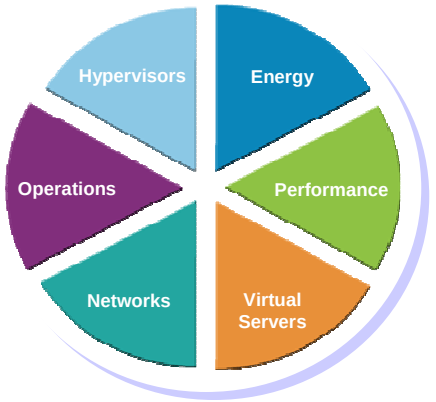
www.ibm.com/redbooks

Part III – Implementing the Ensemble

Redbooks Workshop

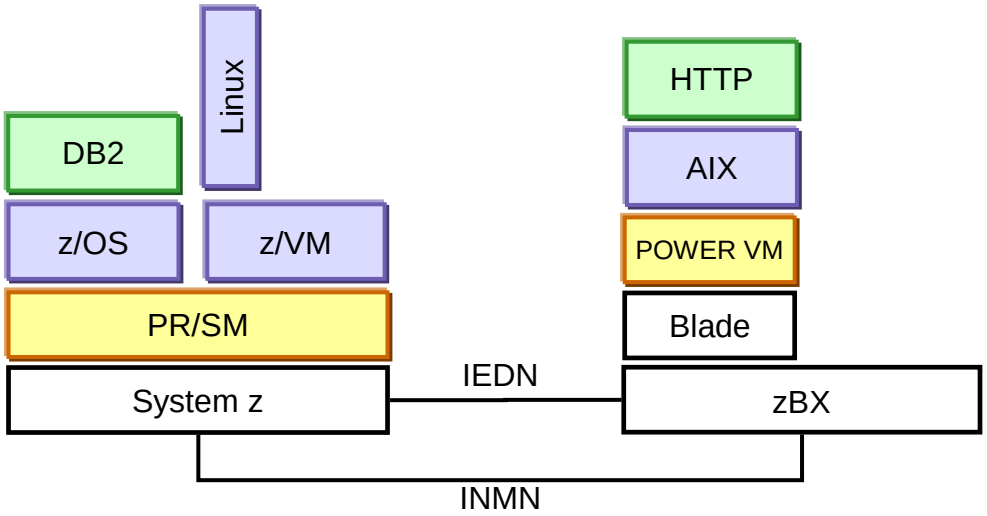
Our Case study scenario

- Workload components
  - z/OS DB2
  - Linux on z with WAS on z/VM
  - AIX on Power Web (HTTP) servers
  - Daytrader project workload



- Existing infrastructure
  - System z CPC
  - z/OS LPAR with DB2
  - z/VM LPAR with Linux
  - System p running HTTP

- New infrastructure -> Target environment
  - zBX Model 2
  - POWER7 Blades
  - INMN and IEDN
  - Unified Resource Manager with Automate suite
    - Workload management
    - Monitoring



Part III – Implementing the Ensemble

1. Designing and planning the Ensemble
2. Unified Resource Manager prerequisites
3. Implementing the Ensemble

Part III – Implementing the Ensemble

1 **Designing and planning the Ensemble**

- ***Design & planning process***
- *Gathering requirements*
- *Environment inventory*
- *Reconciling & Optimizing*
- *Designing and planning the solution*

2 Unified Resource Manager prerequisites

3 Implementing the Ensemble

Our design and planning process

Gather requirements

1. Mapping requirements to
   - Features
   - Functions
   - Topology
   - Technology
2. Create your high level design of the ensemble

Physical and logical inventory

3. Collect and review current environment documentation
4. Understand the current infrastructure physical resources
5. Identify the components and functions of the current workload
6. Identify the protocols and data flows between the components

Reconcile and optimize

7. Group components for placement and consolidation
8. Use available tools to determine the utilization and performance of resources

Design the solution

9. Create a high-level design that incorporates the necessary components, functions, and requirements for redeploying the existing workload.

Part III – Implementing the Ensemble

1 **Designing and planning the Ensemble**

- *Design & planning process*
- ***Gathering requirements***
- *Environment inventory*
- *Reconciling & Optimizing*
- *Designing and planning the solution*

2 Unified Resource Manager prerequisites

3 Implementing the Ensemble

Gathering requirements

- Business requirements
  - Availability
  - Performance
  - Reporting capabilities
  - Security requirements
  - Flexibility, efficiency, responsiveness, business performance
- Operational requirements
  - Monitoring
  - Management capabilities
  - Polices and Automation
  - Security requirements
- Organizational requirements
  - Which departments are involved
  - Who is responsible for what action
  - What processes are used

Part III – Implementing the Ensemble

1 **Designing and planning the Ensemble**

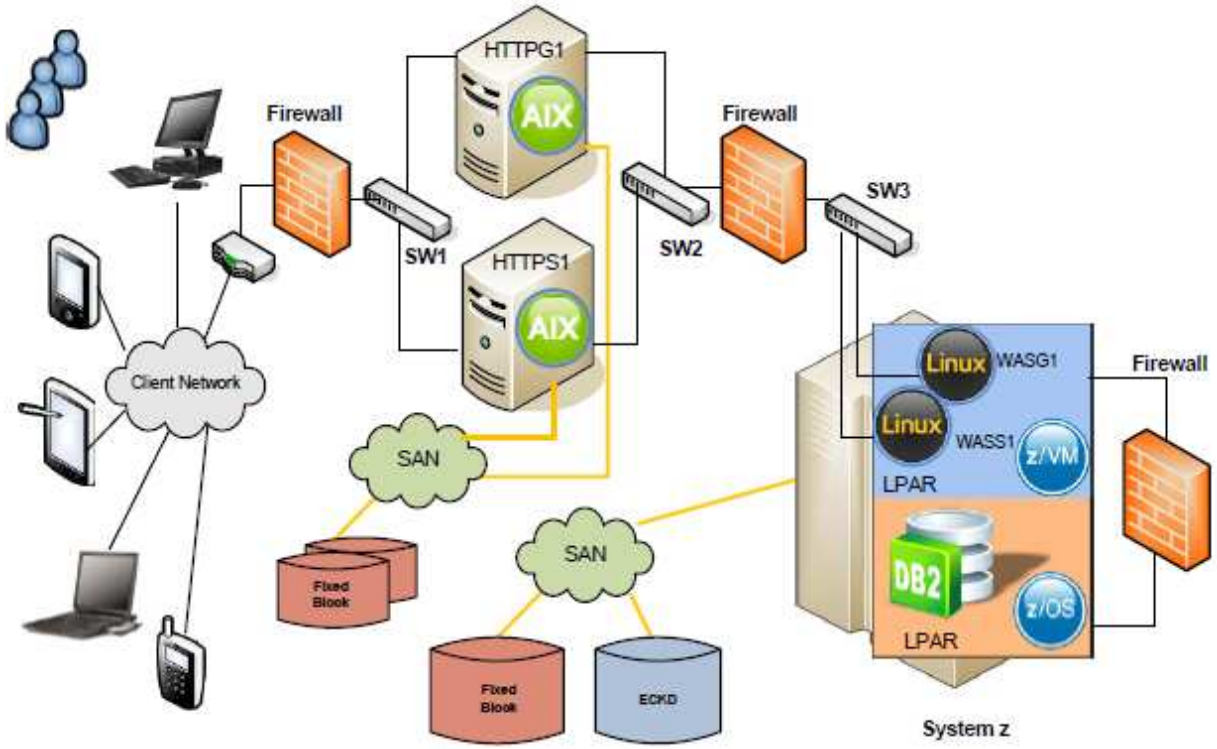
- *Design & planning process*
- *Gathering requirements*
- ***Environment inventory***
- *Reconciling & Optimizing*
- *Designing and planning the solution*

2 Unified Resource Manager prerequisites

3 Implementing the Ensemble

Environment inventory

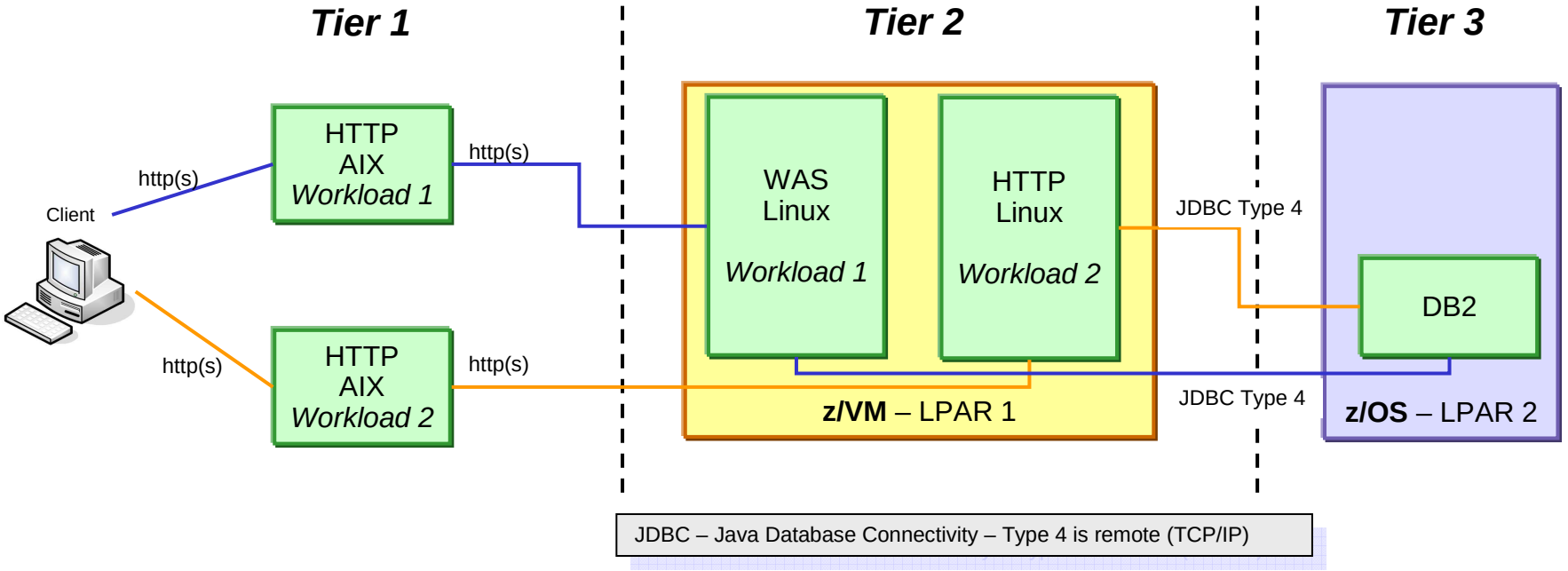
- Physical resources
  - Identify and verify the installed equipment
    - Visual inspection
    - Review available documentation / connectivity diagrams
- Logical resources
  - Identify the components and functions being used
    - Review data flows between components
    - Review protocols used by the existing workload



Note: For security reasons, data flows between components are isolated using firewalls.

Environment inventory – Existing workload

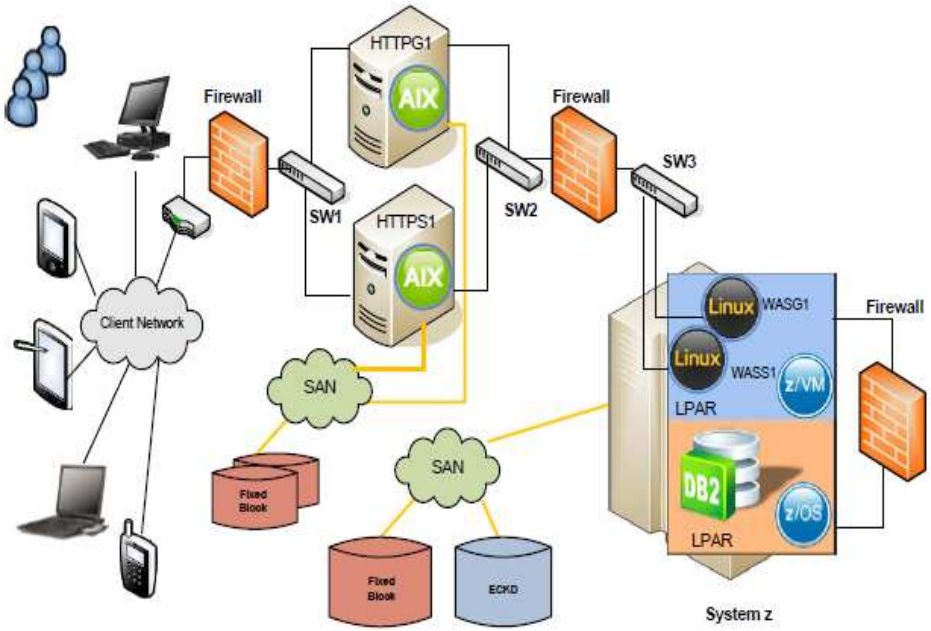
- Understand the data flows and the functions of the application
- Plan and design the target network and virtual server infrastructure



- *The existing application environment is deployed using a 3-tier architecture*
  - *Tier 1- presentation layer: HTTP servers on AIX*
  - *Tier 2 - business logic layer: WAS servers on Linux on System z in the z/VM LPAR*
  - *Tier 3 - data serving layer: DB2 server on z/OS*

Environment inventory – Existing Servers

- Characteristics
  - Types of servers and Operating Systems
  - Servers configuration and utilization
    - Memory
    - Processors
    - I/O and etc.

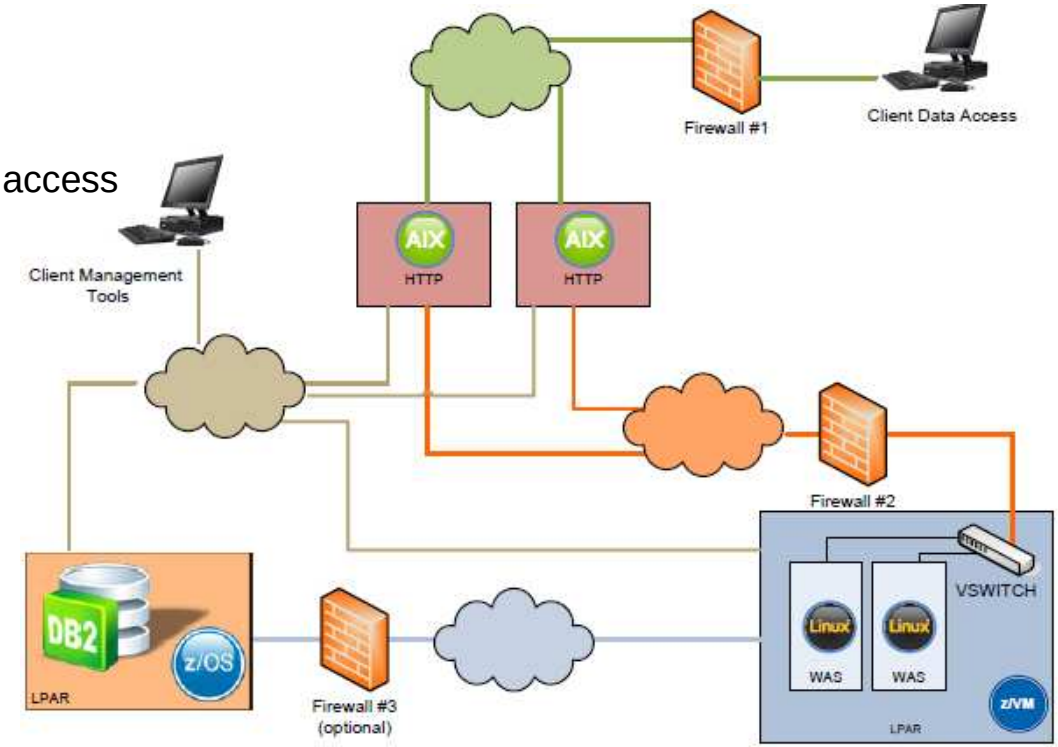


- Performance
  - Use existing server's reports to clearly evaluate
    - Utilization
    - Performance (achieved x required x projected)
  - Evaluate and make adjustments to support the target environment

- *Existing environment configuration consists of*
  - *Distributed environment servers*
  - *System z server*

Environment inventory – Existing Network

- Current network infrastructure
  - Switches, routers and firewalls
- Identify and evaluate
  - Network security implementation
  - Traffic segmentation and application layer access
    - Physical and logical
  - Network management
    - Monitoring, backup, Hw control, etc.

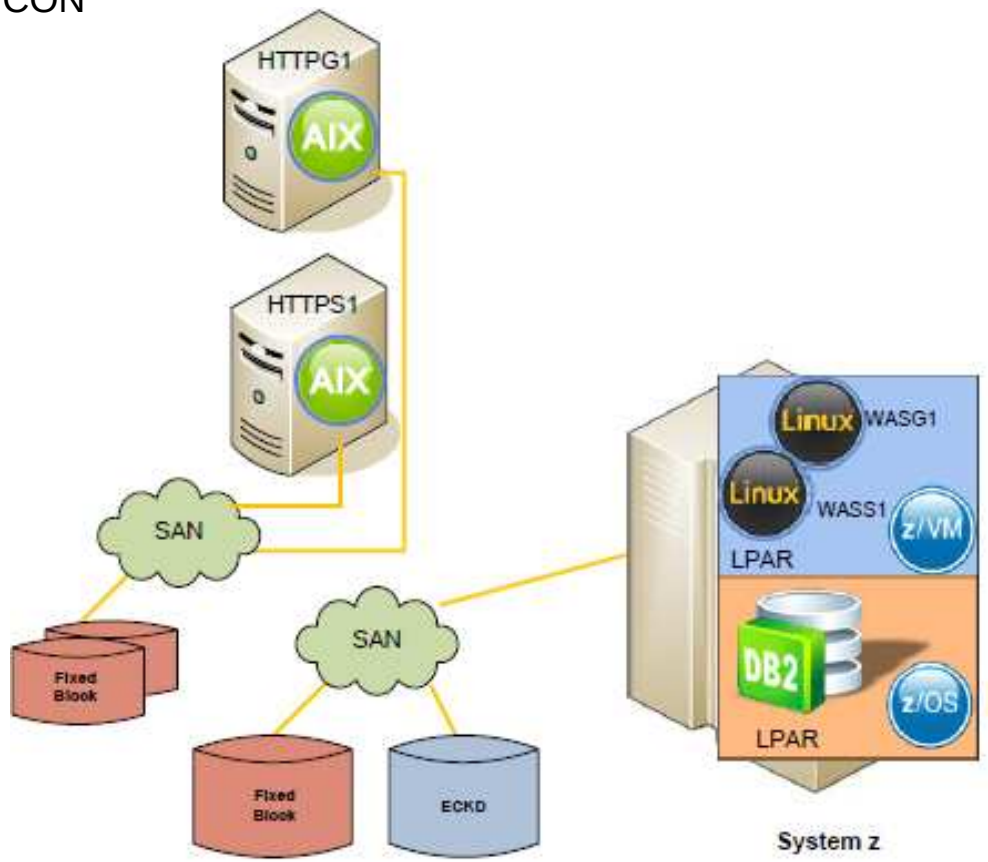


- *Existing network segments*
  - *Client management tools*
  - *Client network < > HTTP servers*
  - *HTTP server < > WAS server*
  - *WAS server < > DB2 server*
    - *Applications and network traffic are segregated by tiers*
    - *Firewalls control communication between servers and clients*

Important: Knowing how the data flows between the components and the protocols being used, helps us to understand the placement of firewalls and if they can be removed or replaced by VLAN isolation.

Environment inventory – Existing storage

- SAN Specifics
  - Type of storage and protocol
    - Fixed-block, ECKD, Fibre Channel and FICON
    - Storage size for each individual server, OS and application
    - SAN zoning, LUN masking, etc
- Information collected includes
  - Volume names for ECKD devices
  - WWPNs and LUN IDs for the FCP and SCSI devices
  - Storage controllers configuration (ECKD / SCSI)
  - Operating systems using storage and mode / protocol
- *Example Consisted of two separated Storage Area Networks (SAN)*
  - *System z and AIX storage*



Part III – Implementing the Ensemble

1 **Designing and planning the Ensemble**

- *Design & planning process*
- *Gathering requirements*
- *Environment inventory*
- ***Reconciling & Optimizing***
- *Designing and planning the solution*

2 Unified Resource Manager prerequisites

3 Implementing the Ensemble

Reconciling and optimizing

- Physical and logical inventory is used to Identify components and functions that can be grouped, consolidated, placed or incorporated in the high level design
  - Servers, OS, applications and parts of the workload
  - Storage resources (SAN)
  - Networking, access (security) zones and segments
- Review operational aspects
  - Network topology
  - Hardware nodes, locations, etc. “What runs where” (placement)
  - Service level characteristics (performance and availability)
  - Management and operations
- Identify deployment units
  - Grouped components
  - Nodes
  - Operating platforms
  - Execution, data and installation aspects may be separately placed
- “Clean up the environment”
  - Consolidate hardware
  - Optimize resources
  - Use (or extend the use of) virtualization

Part III – Implementing the Ensemble

1 **Designing and planning the Ensemble**

- *Design & planning process*
- *Gathering requirements*
- *Environment inventory*
- *Reconciling & Optimizing*
- ***Designing and planning the solution***

2 Unified Resource Manager prerequisites

3 Implementing the Ensemble

Designing the solution

- Transposing the existing environment to the target environment

Existing environment	*Target environment*
One z10 running z/OS and Linux under z/VM	One zEnterprise ensemble consisting of:
	✓One z196 running z/OS and Linux under z/VM ✓One zBX rack with installed IBM Power blades
Two IBM POWER6 servers running AIX	
SAN - Storage resources	✓SAN – Storage resources
Network segments and firewalls	✓Network segments and VLAN
	✓Unified Resource Manager (Unified Resource Manager)

- The output from the reconcile and optimize task provides a solid foundation for the ensemble design

- Essentially, it defines the target environment along with it's required functions (detail configuration is part of planning)
  - Server images + storage
  - Network structure

Planning the solution

- With the output of the design process a detail planning for the various components can be performed

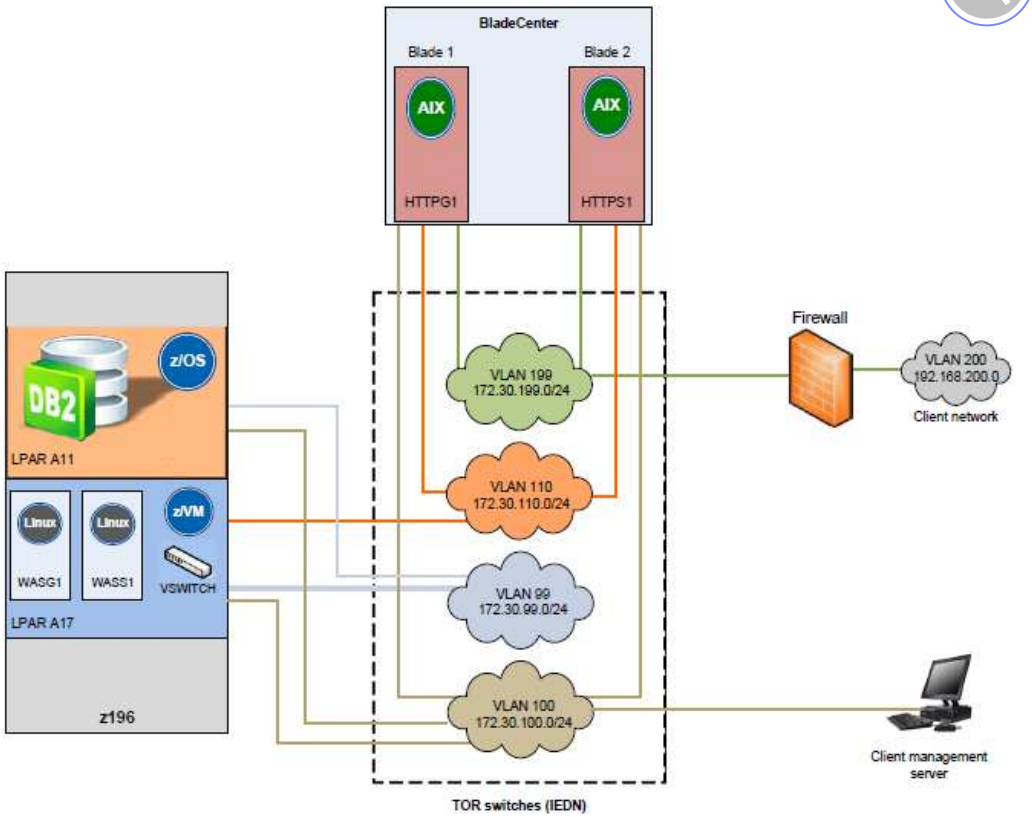
- Proposed order for planning
  - Virtual networks
  - Storage resources
  - Virtual servers

- Planning workloads
  - Initaly group servers deping on their importance for the environment
    - Production/ test
    - Hours of operation (batch/ online)
  - If dynamic workload management is needed make sure to add a test & tuning phase to your project

Planning target environment – Virtual networks

- The worksheet shown below lists the network information needed to configure our virtual networks

- In this worksheet we have the names of our virtual networks with
  - VLAN IDs
  - IP subnets and masks
  - Connected virtual servers



Virtual network name	*VLAN ID*	*(IP subnet/mask)*	*Platform (virtual server)*
Client access - HTTP server VLAN	199	172.30.199.0/24	**AIX** (HTTPG1/HTTPS1)
System management and admin VLAN	100	172.30.100.0/24	**AIX** (HTTPG1/HTTPS1) **z/VM LPAR A17** (WASG1/WASS1) **Z/OS LPAR A11** (DB2 Server)
HTTP server - WAS VLAN	110	172.30.110.0/24	**AIX** (HTTPG1/HTTPS1) **z/VM LPAR A17** (WASG1/WASS1)
WAS - DB2 VLAN	99	172.30.99.0/24	**z/VM LPAR A17** (WASG1/WASS1) **Z/OS LPAR A11** (DB2 Server)

Suggested Network implementation tasks

1. Define VLAN IDs for each required VLAN
2. Create an IP addressing scheme for each VLAN
3. Determine the hosts and networking components
4. Determine the IEDN TOR switch ports that will connect to each VLAN
5. Create diagrams and collect information
   a. Network names
   b. VLAN IDs
   c. IP subnets and masks
   d. TOR ports numbers and types (internal or external)
   e. VLAN modes (access or trunk)
   f. Hosts names for the servers
6. Fill out worksheets with all necessary information before you start

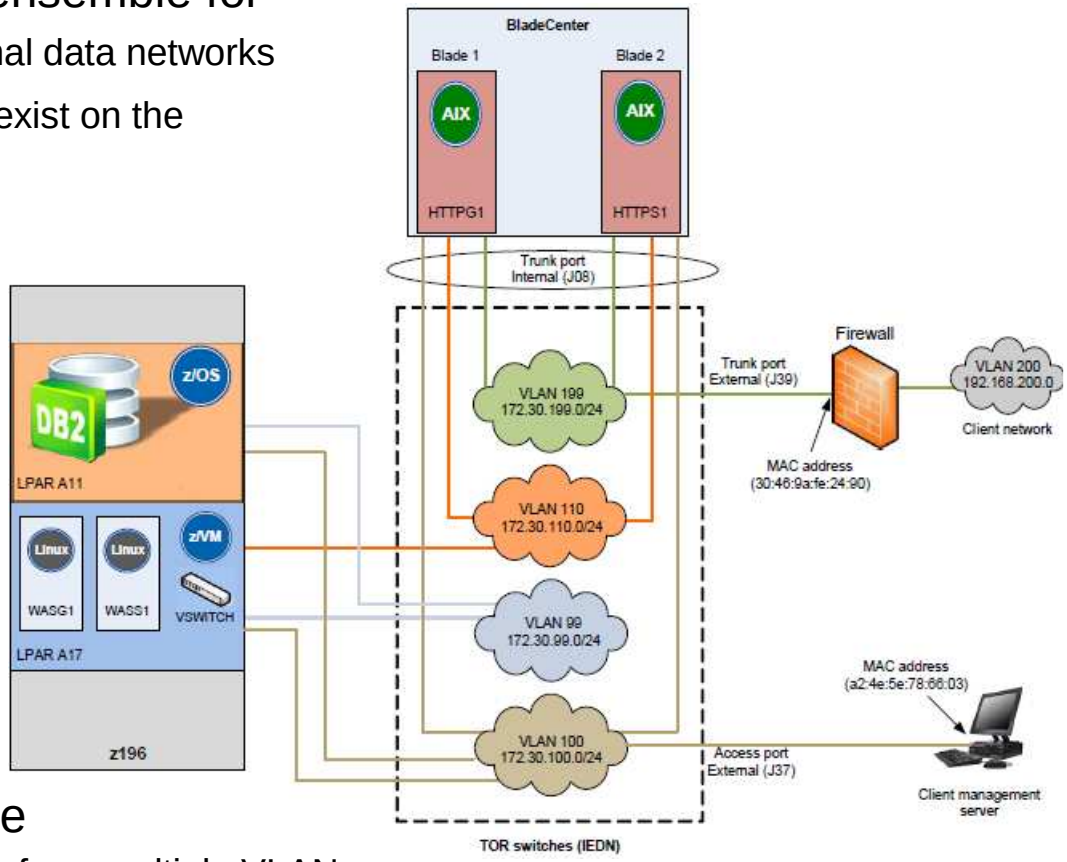
Attention

- The use of worksheets shown in the next charts is strongly suggested!

Planning considerations for the IEDN TOR switches

- Core of data communication in the ensemble for
  - Virtual servers, Virtual switches and external data networks
  - Mix of trusted and untrusted traffic may coexist on the IEDN TOR switch

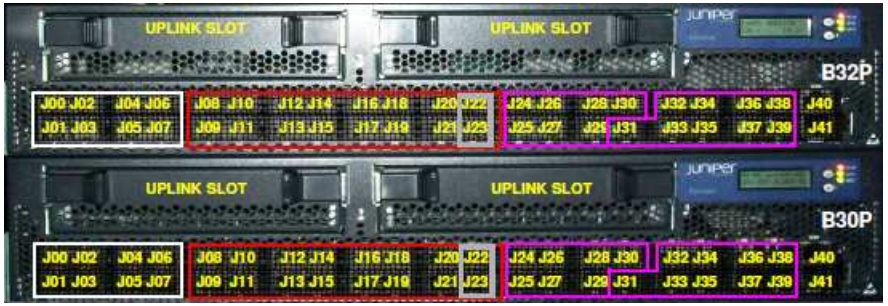
- Use VLAN IDs checking
  - Enforces security
  - *Trunk* and *Access* modes
    - In *Trunk* mode, the IEDN TOR expects all inbound frames to have VLAN tags
    - In *Access* mode, the IEDN TOR will apply VLAN tags to each inbound frame



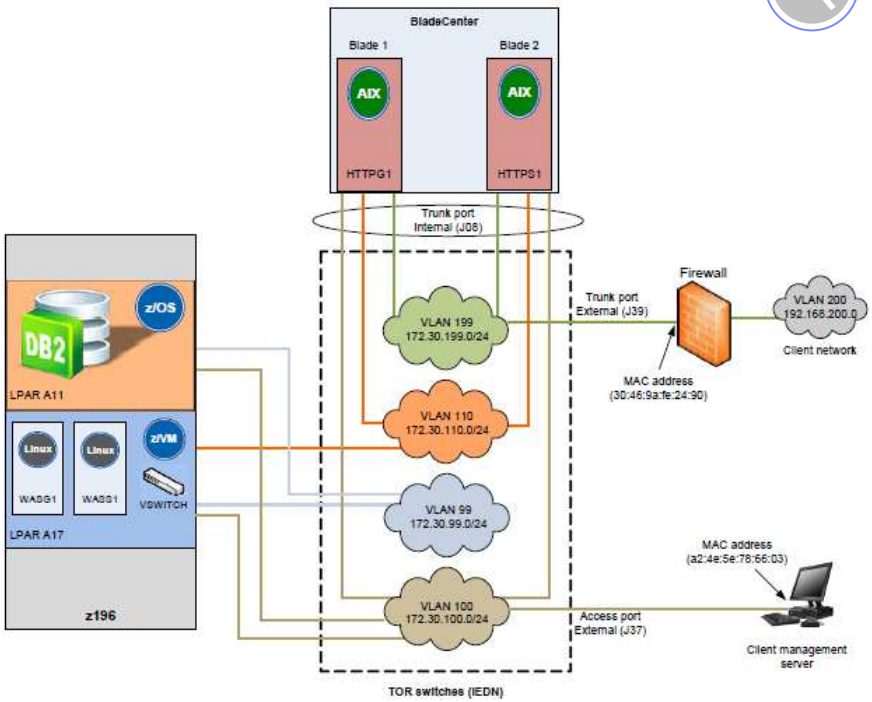
- Number of VLAN IDs in use will determine the switch operating mode
  - Trunk mode when the port receives frames from multiple VLANs

- Use MAC Filtering
  - Can be implemented for the IEDN TOR external ports
  - Ensures that only authorized devices connect to the IEDN TOR switch

Planning target environment – TOR switches



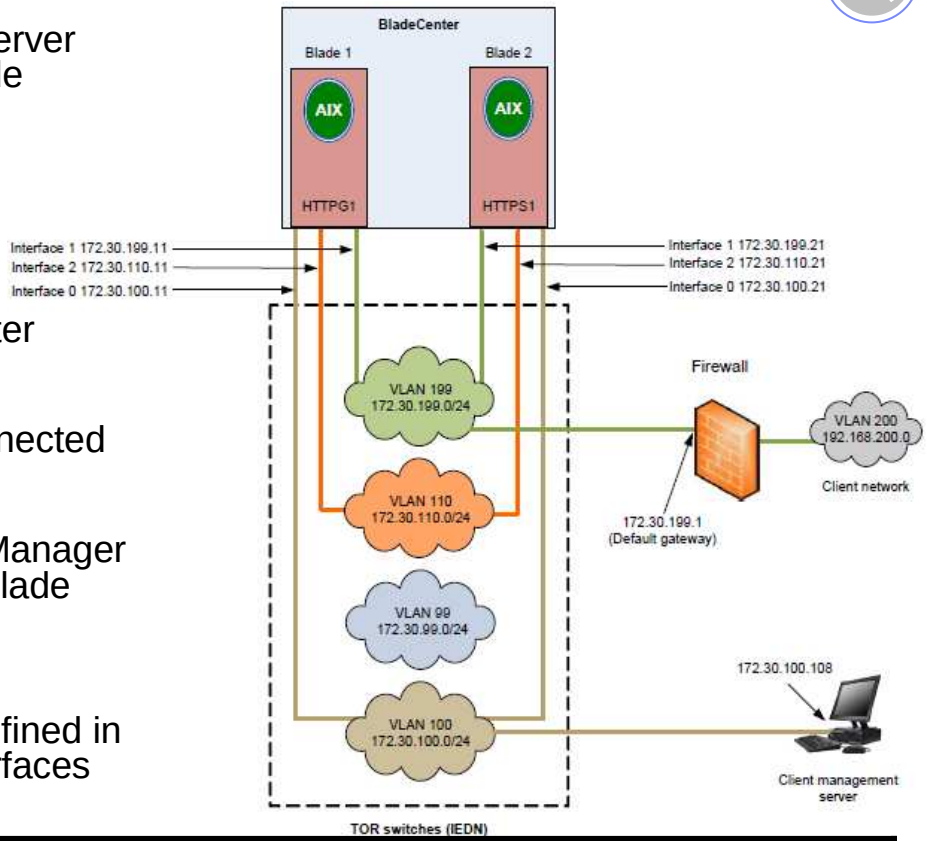
zBX – IEDN TOR switch



Virtual network name Description	*VLAN ID*	*Port # type*	*Port mode*	*Platform Host name*	*MAC Address*
Client access - HTTP server VLAN	199	J39 External	Trunk	**Firewall/router** (Client Network) **POWER7** HTTPG1/HTTPS1	30:46:9a:fe:24:90
System management and administration VLAN	100	J37 External	Access	**Management server** **POWER7** HTTPG1/HTTPS1 **z/VM LPAR A17** WASG1/WASS1 **z/OS LPAR A11** DB2 server	a2:4e:5e:78:66:03
HTTP server - WAS VLAN	110	J08 Internal	Trunk	POWER7 HTTPG1/HTTPS1 **z/VM LPAR A17** WASG1/WASS1	N/A
WAS - DB2 VLAN 99	99	J08 Internal	Trunk	**z/VM LPAR A17** WASG1/WASS1 **z/OS LPAR A11** DB2 server	N/A

Planning target environment – Blade server network

- The virtual network adapters defined for each virtual server connect into the Power Hypervisor switch in each Blade
- The Virtual I/O server provides the Ethernet bridge between the Ethernet switch defined in the Power Hypervisor and the physical dual port 10GbE adapter installed in each Blade
- The 10GbE adapters are connected via the BladeCenter backplane into the 10GbE switches (CFFh)
- The 10GbE switches in the BladeCenter are each connected into the TOR IEDN switches (J08)
- The VLAN configuration defined in Unified Resource Manager is propagated from to the TOR IEDN switches to the Blade Center 10GbE switches (CFFh) down to the Power Hypervisor defined switches
- For the current implementation, the virtual adapters defined in Unified Resource Manager are standard Ethernet interfaces (not VLAN aware (802.3q)



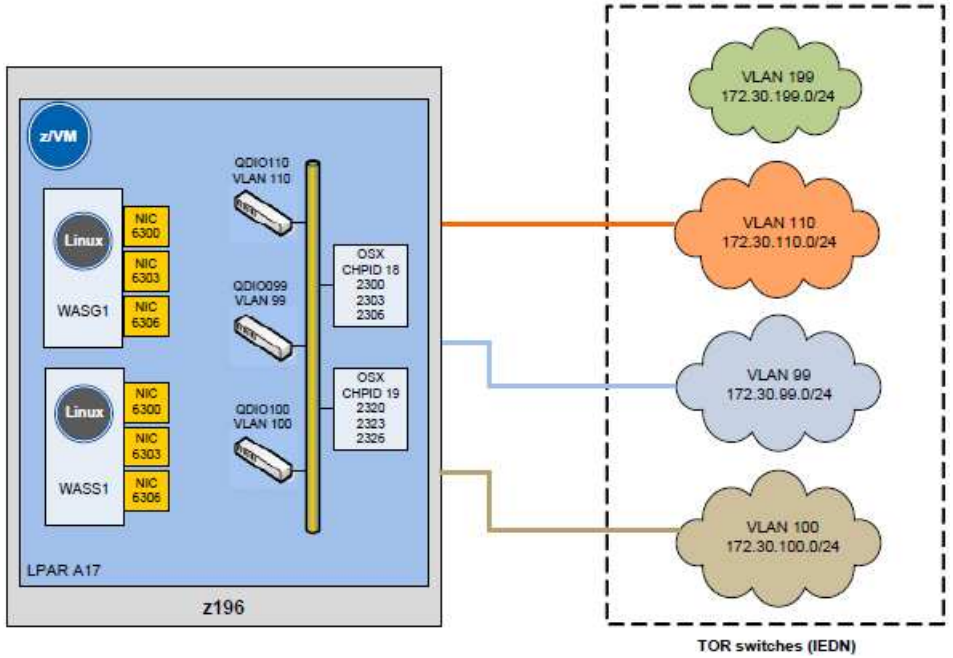
Virtual Server	*Virtual network name*	*Virtual Interface*	*VLAN ID*	*IP subnet add / mask*
AIX HTTPG1	System management and administration VLAN	en0[a]	100	172.30.100.11/24
	Client access - HTTP server VLAN[b]	en1	199	172.30.199.11/24
	HTTP server - WAS VLAN	en2	110	172.20.110.11/24
AIX HTTPG2	System management and administration VLAN	en0[a]	100	172.30.100.21/24
	Client access - HTTP server VLAN[b]	en1	199	172.30.199.21/24
	HTTP server - WAS VLAN	en2	110	172.30.110.21/24

a. If an AIX NIM server is used as a boot source (virtual sever boot mode is System Management Services (SMS) and boot source is Network Adapter (ID 0)), then the virtual interface must be en0

b. Requires a default gateway (172.30.199.1)

Planning target environment – Network and VSWITCHes for z/VM servers

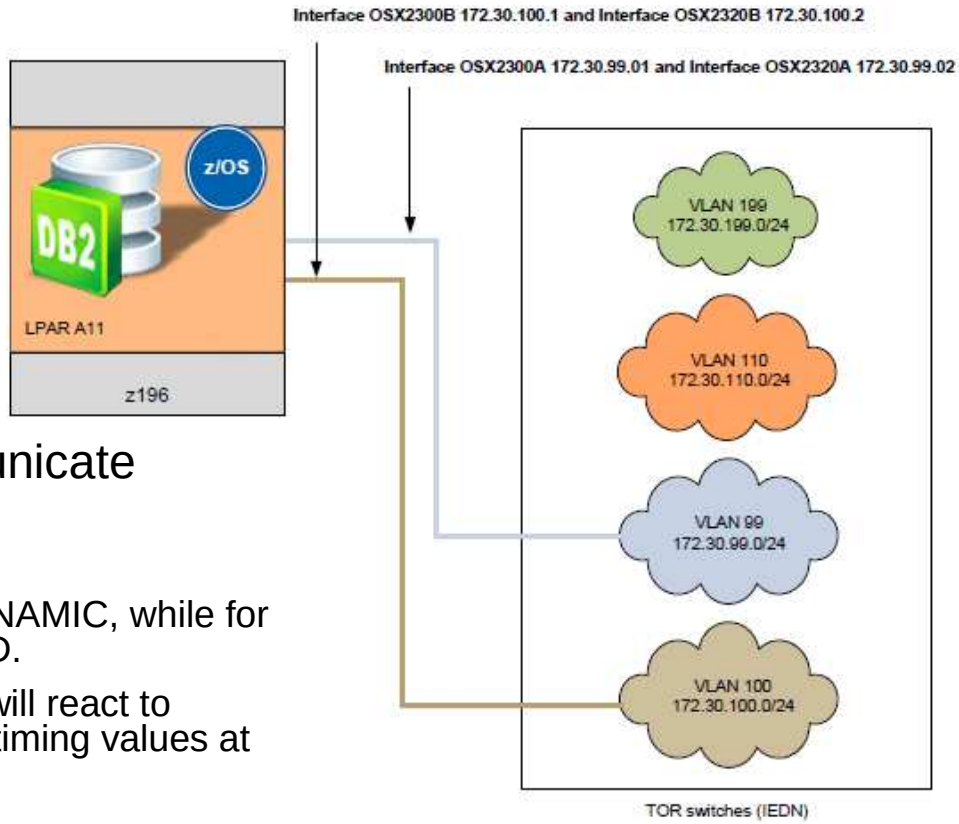
- Guest systems can attach to a OSX CHPID directly if the operating system has OSX support. The z/VM guest operating systems that use OSX CHPIDs should have a VLAN ID assigned to each network interface (NIC)
- z/VM guest operating systems that do not support OSX CHPIDs, require an IEDN VSWITCH to simulate an OSD connection to the guest operating system. The IEDN VSWITCH enforces the activation of that connection using the appropriate VLAN ID. The IEDN VSWITCH points to three real addresses of the OSX CHPID
- A QDIO type virtual NIC is needed to connect the guest operating system to the IEDN VSWITCH. The user directory for that z/VM guest is created and contains a NICDEF statement with a virtual NIC, which links to a virtual address
- The commands required by a z/VM security manager are not run automatically. The security manager command to grant a z/VM guest access to a VSWITCH must still be issued manually



Virtual switch name	*Virtual switch type*	*Virtual network name*	*VLAN ID*	*Uplink (CHPID – device)*
QDIO100	IEDN	**System Management and administration** VLAN	100	CHPID 18 – 2303 CHPID 19 – 2323
QDIO110	IEDN	**HTTP server < > WAS** VLAN	110	CHPID 18 – 2306 CHPID 19 – 2326
QDIO099	IEDN	**WAS < > DB2** VLAN	99	CHPID 18 – 2300 CHPID 19 – 2320

Planning target environment – Network for the z/OS LPAR

- z/OS accesses the IEDN through 10 GbE OSX (OSX CHPID type)
- OSX must be configured as Layer 3 IP interfaces to z/OS with an IP address and subnet mask
- Definitions must be consistent with the other virtual servers VLANs to which the z/OS TCP/IP stack will communicate
- TCP/IP INTERFACE statement changes
  - For OSX interfaces, INBPERF will default to DYNAMIC, while for OSD interfaces, INBPERF default to BALANCED.
  - The DYNAMIC setting for INBPERF parameter will react to changes in traffic patterns, and set the interrupt-timing values at the point where throughput is maximized



Virtual Server	*Virtual network name*	*Interface*	*VLAN ID*	*IP subnet add / mask*
z/OS LPAR A11 (DB2 Server	System management and administration VLAN	OSX2300B OSX2320B	100	172.30.100.1/24 172.30.100.2/24
	WAS < > DB2 VLAN	OSX2300A OSX2320A	99	172.30.99.1/24 172.30.10.2/24

Planning target environment – Storage resources worksheet

- Table below shows the data exported by Unified Resource Manager export WWPN function
- The shaded area contains the default export WWPN fields for each storage resource type
- In the example shown, we exported WWPN for Blade B.1.01 (BladeCenter 1, Blade 01)

FCP_DEF:	name	size	desc	location	Host WWPN	WWPN	LUN		
#ECKD_DEF:	name	size	desc	location	devno	VOLSER	LCU/UA	model	
#ZVM_FCP_DEF:	Name	size	desc	location	Host WWPN	WWPN	LUN	devno	VOLSER
FCP				SCZP301:B.1.01	21000024ff2a428c				
FCP				SCZP301:B.1.01	21000024ff2a428d				

- *We used the information provided by the Unified Resource Manager export WWPN function as a starter for the Storage Worksheet, as shown below*

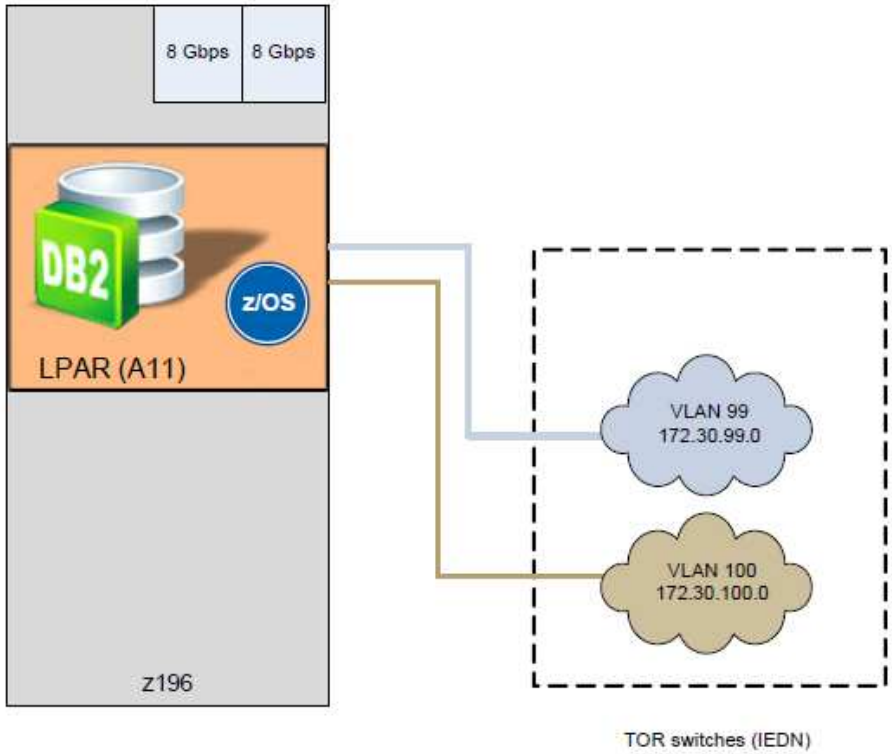
Proposed Storage Worksheet (PowerVM)					
		Information provided by the Unified Resource Manager WWPN export function		*Information added by the Storage Administrator*	
Name	***Size***	***Location***	***Host WWPN***	***Target WWPN***	***LUN***
freeB101	*32212254720*	*SCZP301:B.1.01*	*21000024ff2a428c*	*202600a0b847d6d0*	*0002000000000000*
freeB101	*32212254720*	*SCZP301:B.1.01*	*21000024ff2a428d*	*202600a0b847d6d0*	*0002000000000000*
freeB102	*32212254720*	*SCZP301:B.1.01*	*21000024ff2a428c*	*202600a0b847d6d0*	*0004000000000000*
freeB102	*32212254720*	*SCZP301:B.1.01*	*21000024ff2a428d*	*202600a0b847d6d0*	*0004000000000000*
freeB103	*32212254720*	*SCZP301:B.1.01*	*21000024ff2a428c*	*202600a0b847d6d0*	*0006000000000000*
freeB103	*32212254720*	*SCZP301:B.1.01*	*21000024ff2a428d*	*202600a0b847d6d0*	*0006000000000000*

Planning target environment – Virtual servers in general

- Review prerequisites for the planned servers
  - Hardware → amount of CPU or memory resources
  - Software requirements → SW versions and special products/ tools
- Prepare and complete the proposed hypervisor's worksheets
- Determine virtual server resource requirements and configuration
  - Number of processors
  - Memory assignment
  - Other server resources
- Review the Virtual Servers Operating Systems installation and boot requirements
- Collect requirements from the previous prepared worksheets
  - Storage needed
  - Network resource access

Planning target environment – z/OS virtual servers

- z/OS LPARs become part of the ensemble when it is created
- Unified Resource Manager can‘t be used to create a new z/OS LPAR



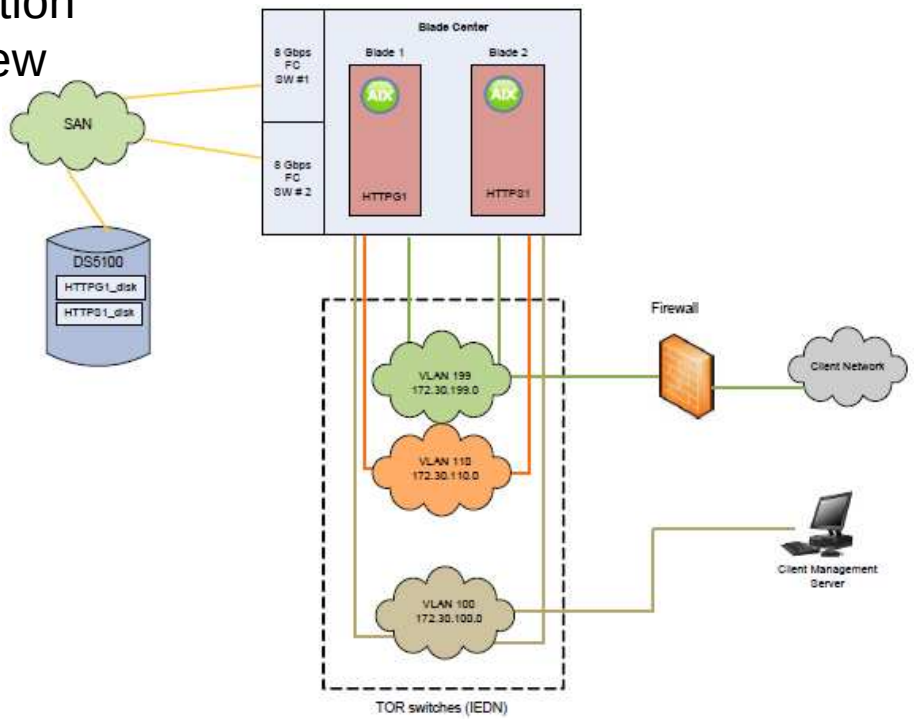
Virtual Server	*LPAR name*	*Description*	*VLANs*
DB2 Server	A11	z/OS LPAR	WAS – DB2 VLAN (99) System management and administration VLAN (100)

Planning target environment – AIX virtual servers

- Prepare initial values for the AIX servers
  - These values are needed for the „create virtual server task“

- Values can be changed after implementation for tuning purposes using the „Details“ view

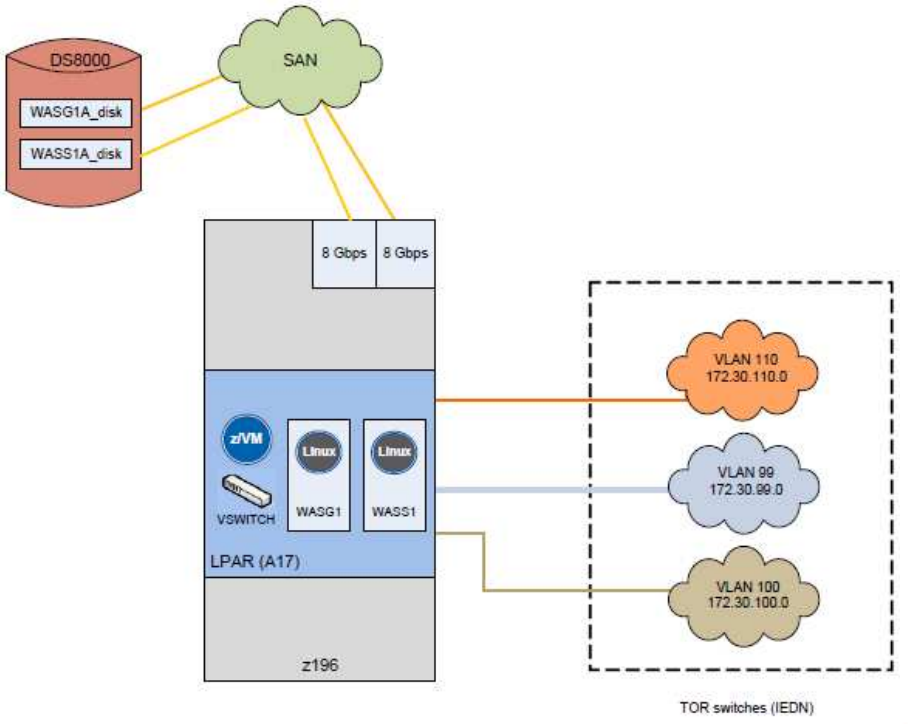
Virtual server parameter	*Virtual server 1*	*Virtual server 2*
Name (64 chars)	HTTPG1	HTTPS1
Description	HTTP virtual server 1	HTTP virtual server 2
IP address		
Target hypervisor	PowerVM	PowerVM
Processor mode (shared/ dedicated)	shared	Shared
Initial virtual processors	4	3
Initial dedicated memory	8 GB	8 GB
Networks (VLANs)	110,199,100	110,199,100
Storage resource name (persistens)		
Storage drives	30 GB	30 GB



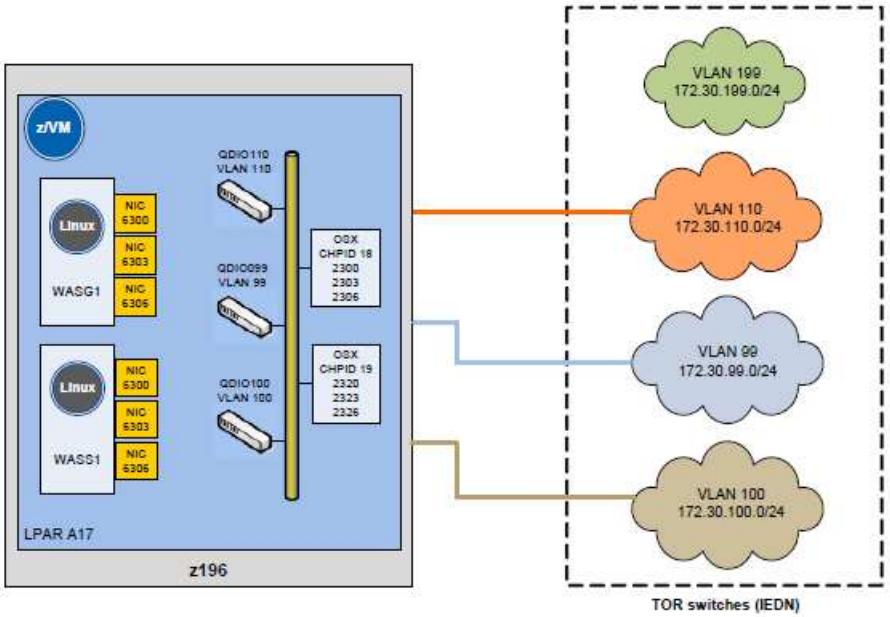
Planning target environment – z/VM Linux guest servers

Virtual server parameter	Virtual server 1	Virtual server 2
Name (8 chars)	WASG1	WASS1
Description (256 chars)	WAS server #1 on Linux on System z	WAS server #2 on Linux on System z
IP address		
Target hypervisor	zVM	zVM
Processor type – (CP/IFL)	CP	CP
Initial virtual processors	2	2
Initial memory	2 GB	2 GB
Networks (VLANs)	099,110,100	099,110,100
Storage drives	fullpack 30GB	fullpack 30GB
Planned amount of storage (GB)		
IPL Boot device		
IPL Load parameters		
z/VM Options	IPL CMS, G class	IPL CMS, G class
Storage resource name		
Min. processing units		
Max. virtual processors	4	4
Shared mode – relative/absolute		
Share limit		
Initial relative shares		
Maximum share		
Initial memory		
Maximum memory		
Privilege class		

- Existing servers can be included in the ensemble
- Decide whether servers are managed by Unified Resource Manager or manually changing the user directory
  - Do not use both approaches in parallel



Planning target environment – z/VM virtual switches*



Virtual switch name	*Virtual switch type*	*Virtual network name*	*VLAN ID*	*Uplink (CHPID - device)*
QDIO100	IEDN	System management and administration VLAN	100	CHPID 18 – 2303 CHPID 19 – 2323
QDIO110	IEDN	HTTP server – WAS VLAN	110	CHPID 18 – 2306 CHPID 19 – 2326
QDIO099	IEDN	WAS – DB2 VLAN	99	CHPID 18 – 2300 CHPID 19 – 2320

* **Note:** The definition of the VSWITCHes are required when the Linux guests don't have the required OSX CHPID support.

Part III – Implementing the Ensemble

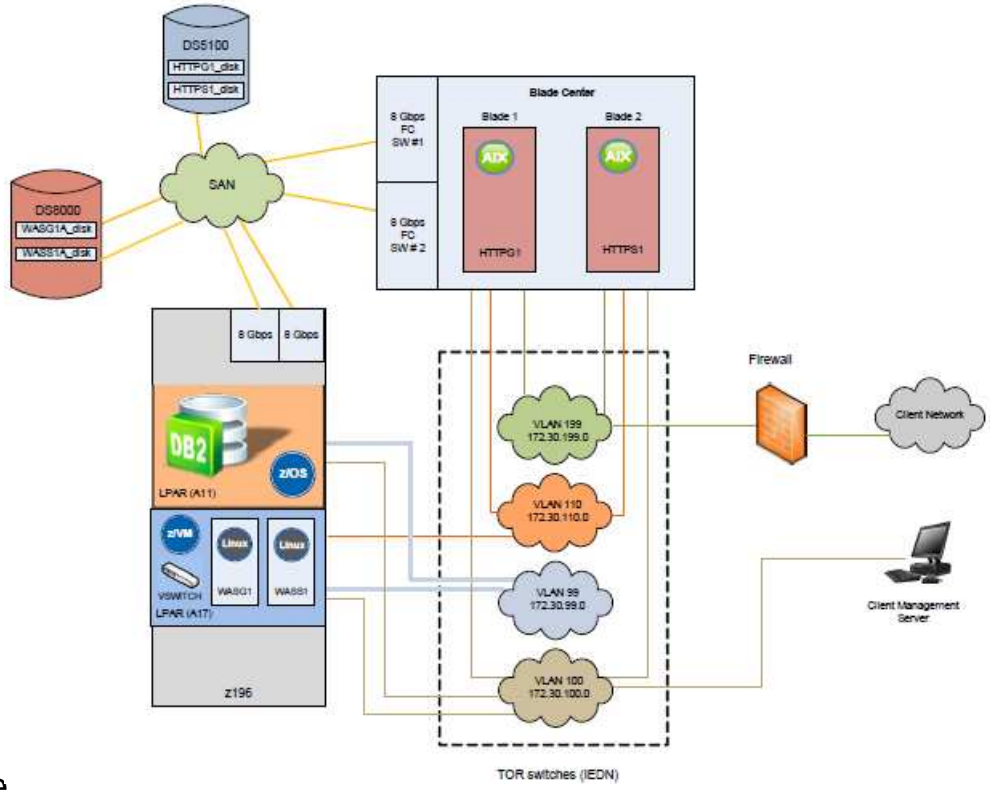
1 Designing and planning the Ensemble

2 Unified Resource Manager prerequisites

3 Implementing the Ensemble

Operating system prerequisites

- System z
  - z/OS Version 1 Release 10 (V1R10) + PTFs
  - z/OS Version 1 Release 11 (V1R11) + PTFs
  - z/OS Version 1 Release 12 (V1R12) + PTFs
  - z/OS Version 1 Release 13 (V1R13) + PTFs
  - z/VM Release 6.1 + PTFs
- Blades
  - AIX 5.3 (Technology Level 12) and later (deployed on an IBM Blade in the zBX)
  - AIX 6.1 (Technology Level 5) and later (deployed on an IBM Blade in the zBX)
  - Red Hat Enterprise Linux *(Version 5.5 or higher)*
  - SUSE Linux Enterprise Server *(Version 10 or higher)*
  - Microsoft Windows Server 2008 R2 Enterprise (x86-64)



Unified Resource Manager pre-requisites

- HMC
  - Unified Resource Manager requires a pair of FC 0091 HMCs
- zEnterprise Feature Codes
  - FC 0019 – Manage suite
  - FC 0020 – Automate suite
- “Entitlements” allow customers to activate and use Blades or Optimizers

Feature Code	Quantity needed	Description	Charged per
0019	0-1	Manage Firmware Suite	No charge for z/OS and IFL‘s
0020	0-1	Automate Firmware Suite	No charge for z/OS
0025	0-1	Ensemble membership	No charge
0039	0-56	Manage FW ISAOPT	Per Blade → Bundle
0040	0-28	Manage FW DP	Entitlement - per Blade
0041	0-112*	Manage FW Pwr Blade	Entitlement - per Blade
0043	0-56	Automate FW ISAOPT	
0044	0-28	Automate FW DP	
0045	0-112	Automate FW Pwr Blade	Entitlement - per Blade
0052	# of IFL on z196/z114	Automate FW IFL	Entitlement - per Blade
0091	2	HMC w/Dual EN	

* Minus the number of purchased DataPower Entitlements

Part III – Implementing the Ensemble

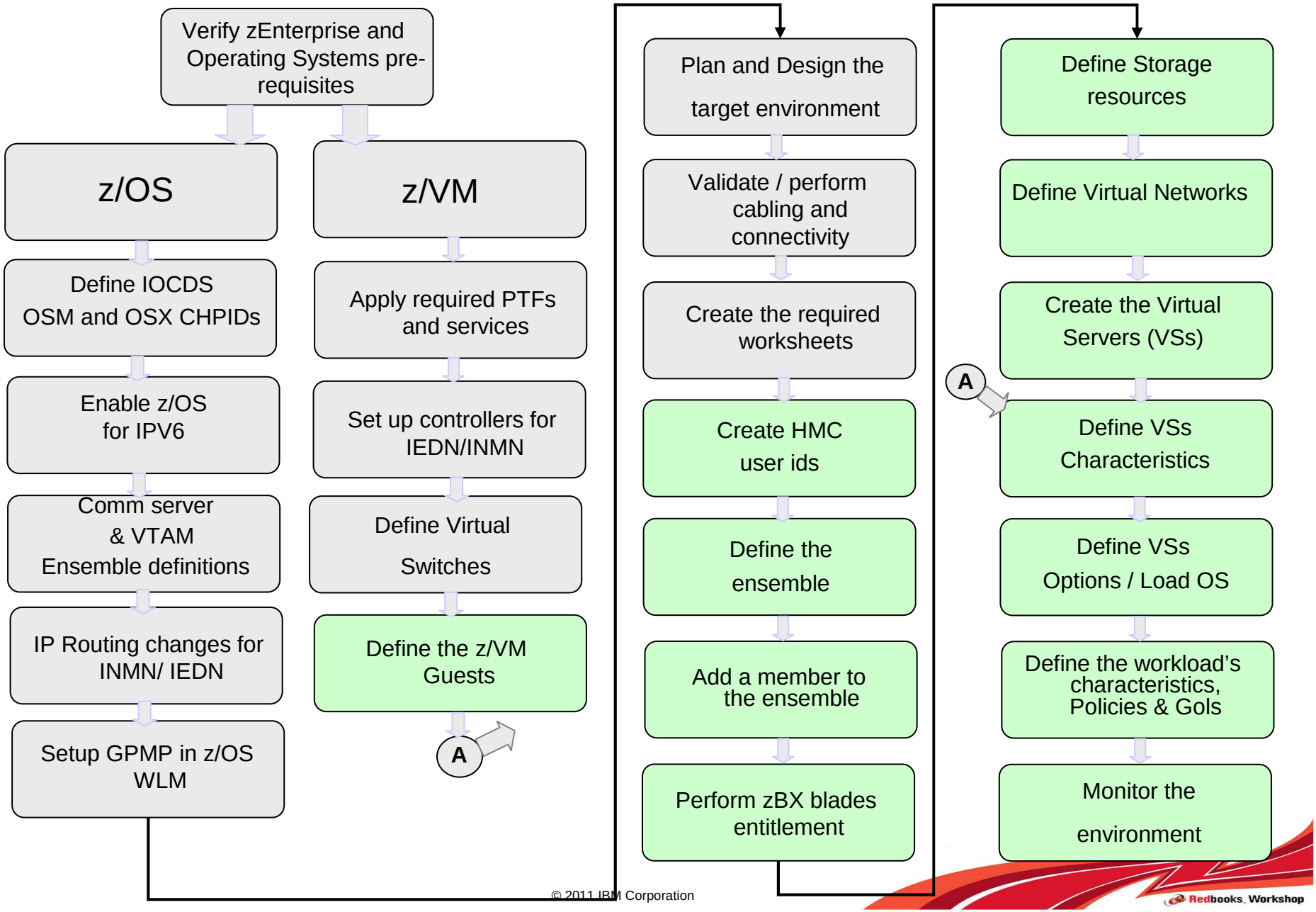
1 Designing and planning the Ensemble

2 Unified Resource Manager prerequisites

3 Implementing the Ensemble

- ***Implementation process overview***
- *HMC tasks*
- *Implementation walk-through*

Building the new infrastructure (target environment) flow



Reference material

Form Number	Document Title
SG24-7921	IBM zEnterprise Unified Resource Manager RedBook
GC27-2606	zEnterprise System HMC Operations Guide for Ensembles
GC27-2607	zEnterprise System Ensemble Performance Management
GC27-2608	IBM zEnterprise System Ensemble Planning and Configuring Guide
GC27-2609	zEnterprise System Introduction to Ensembles
SB10-7037-08	Input/Output Configuration Program User's Guide for ICP IOCP
SC31-8778	z/OS Communications Server V1R12 SNA Network Definition Reference
SC31-8775	z/OS Communications Server V1R12 IP Configuration Guide
SC31-8776	z/OS Communications Server V1R12 IP Configuration Reference
SC31-8885	z/OS Communications Server V1R12 IPv6 Network and Application Design Guide
SC24-6234-01	z/VM System Management Application Programming
SC24-6166-01	z/VM 6.1 CMS Commands and Utilities Reference
SC24-6175-01	z/VM 6.1 CP Commands and Utilities Reference
SC24-6174-01	zVM 6.1 Connectivity
SC24-6178-01	z/VM CP Planning and Administration Guide

Part III – Implementing the Ensemble

1 Designing and planning the Ensemble

2 Unified Resource Manager prerequisites

3 Implementing the Ensemble

Implementation process overview

HMC tasks

Implementation walk-through

HMC related tasks

- Plan physical HMC placement
  - Exact 2 HMC's are available
  - Both HMC's need to be in the same subnet but can be on remote sites
- Prepare a HMC to become an "Alternate HMC"
  - Remove System z objects defined to it (`"Remove object definition"`)
  - Disable the RSF option of the alternate HMC (`Disabel "Call home"`)
  - HMC is not allowed to receive LIC update (`LIC Management "disabled"`)
  - Perform "`Manage Alternate HMC`" with user ENSADMIN to make HMC eligible as alternate HMC
- HMC user related tasks
  - Plan and implement user roles → define a set of task and resources
  - Implement remote access of users who have a "need to connect"
  - Implement user authentication (local or against a remote LDAP)

Tip: We highly suggest not to use a generic "superuser" (UID0) like ENSADMIN for day-to-day tasks. Instead implement functional users, exploiting task and resource roles.

HMC changes – New default users on the HMC

- Traditional HMCs userids
  - ACSADMIN, SYSPROG, OPERATOR, ADV OPERATOR, SERVICE
- Ensemble HMCs new userids and roles
  - ENSADMIN – ensemble administrator
  - ENSOPERATOR – ensemble operator

Ensemble Administrator (ENSADMIN)	
Objects	Tasks
• *All zCPC Managed objects* • *BladeCenter objects* • *Defined zCPC Managed objects* • *Ensemble object* • *IBM Blade object* • *IBM Blade Virtual Server object* • *ISAOPT Blade object* • *Storage Resource objects* • *Virtual Network objects* • *Workload objects* • *zVM Virtual Machine objects*	•*Energy Mgt Administrator* •*Ensemble Administrator* •*Performance Mgt Administrator* •*Storage Resource Administrator* •*System Programmer* •*Virtual Network Administrator* •*Virtual Server Administrator* •*Workload Administrator* •*zVM Virtual Machine tasks*

Ensemble Operator (ENSOPERATOR)	
Objects	Tasks
•*All zCPC Managed objects* •*Blade Center objects* •*Defined zCPC Managed objects* •*Ensemble object* •*IBM Blade object* •*IBM Blade Virtual Server object* •*ISAOPT Blade object* •*Storage Resource objects* •*Virtual Network objects* •*Workload objects* •*zVM Virtual Machine objects*	•*Advanced Operator* •*Performance Mgt Operator* •*Virtual Server Operator*

HMC changes – New Ensemble task roles details

Role	Description
Ensemble Administrator	Responsible for creating and managing the zEnterprise ensemble. *Create Ensemble, Add Member…*
Virtual Network Administrator	Responsible for Managing Virtual Networks, Hosts, and MAC Prefixes. *Manage Virtual Networks, Add Hosts to Virtual Networks, Create VLAN IDs…*
Virtual Server Administrator	Responsible for managing virtual servers. *New /Modify Virtual Server, Add Virtual Disk, Migrate…*
Virtual Server Operator	Responsible for performing and scheduling virtual server activation /deactivation, mounting virtual media *Activate, Deactivate, Mount Virtual Media, Console session…*
Storage Resource Administrator	Responsible for managing storage resources – Storage Access Lists(SAL), WWPNs, z/VM Storage Groups *Export WWPN, Import SAL, Add Storage Resources…*
Workload Administrator	Responsible for managing workloads. *New /Modify workload, Add / Remove Virtual Servers..*
Performance Management Administrator	Responsible for managing performance policies *New /Modify performance policy, Import policy*
Performance Management Operator	Responsible for performing and scheduling policy activations and creating threshold notifications *Activate, Export Policy, Monitor System Events*
Energy Management Administrator	Responsible for managing power settings including power capping and power savings modes. *Set Power Cap, Set Power Savings Mode, set zBX Power Policy*

Preparing for Ensemble Implementation

Checklist – before creating the ensemble:

- Have all the hardware **MCL changes** for Ensemble participation been implemented?
- Have all the **PSP buckets** for z/OS and z/VM been researched and the appropriate fixes applied?
- Do you have the appropriate **diagrams, worksheets, and manuals** to complete the IOCDS task? *(Refer to the IOCP and Ensemble reference documents)*
- Did you verify that the required **CHPIDs, CNTLUNITs, and DEVICEs are online** to the appropriate LPARs and **accessible**?
- Do you have the appropriate diagrams, worksheets, and manuals to complete definitions and the **inter-frame cabling** task? *(Refer to the z/BX documentation)*
- Did you implement the **Ensemble user IDs**, roles, and tasks in the HMC?
- Did you complete the **operating system changes for z/OS** to participate in the Ensemble? *(Refer to the z/OS Communications Server manuals)*
- If the **Guest Management Platform Provider** is to be used in your Ensemble, have their implementation tasks task been planned and assigned?
- Did you complete the **z/VM required changes** for it's participation in the Ensemble?

Part III – Implementing the Ensemble

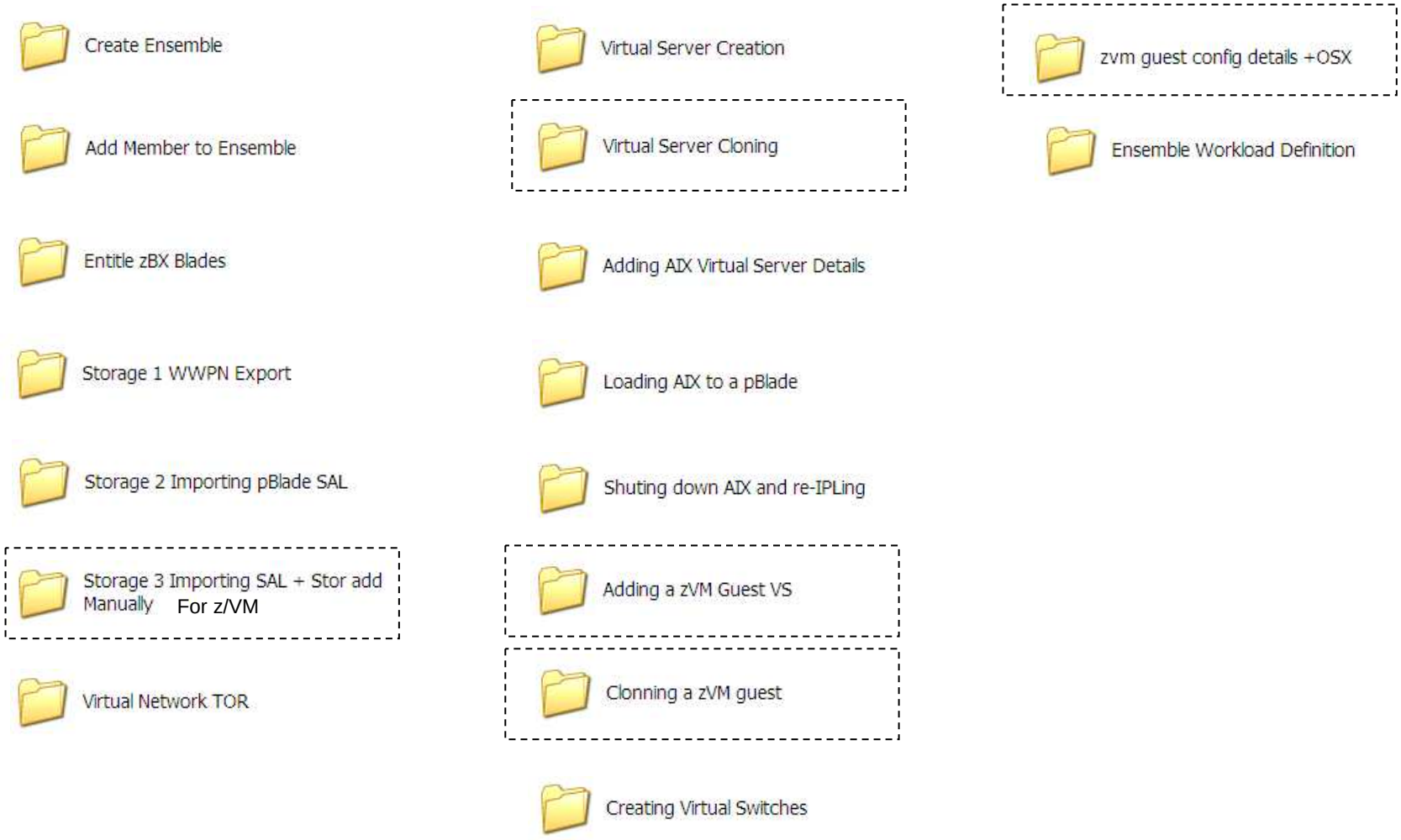
1 Designing and planning the Ensemble

2 Unified Resource Manager prerequisites

3 Implementing the Ensemble

4 Implementation Walk-through

Ensemble Demos



yourdotcom

International Technical Support Organization and Authoring Services

zEnterprise
Unified Resource Manager

www.ibm.com/redbooks

Part IV – Ensemble Managing and Problem Determination

Redbooks Workshop

Part IV – Ensemble Managing and Problem Determination

1. Managing the Environment
2. z/VM in an Ensemble
3. z/OS in an Ensemble
4. AIX
5. zBX
6. Glossary

Note: This chapter is intended to be a reference when implementing an ensemble.

Part IV – Ensemble Managing and Problem Determination

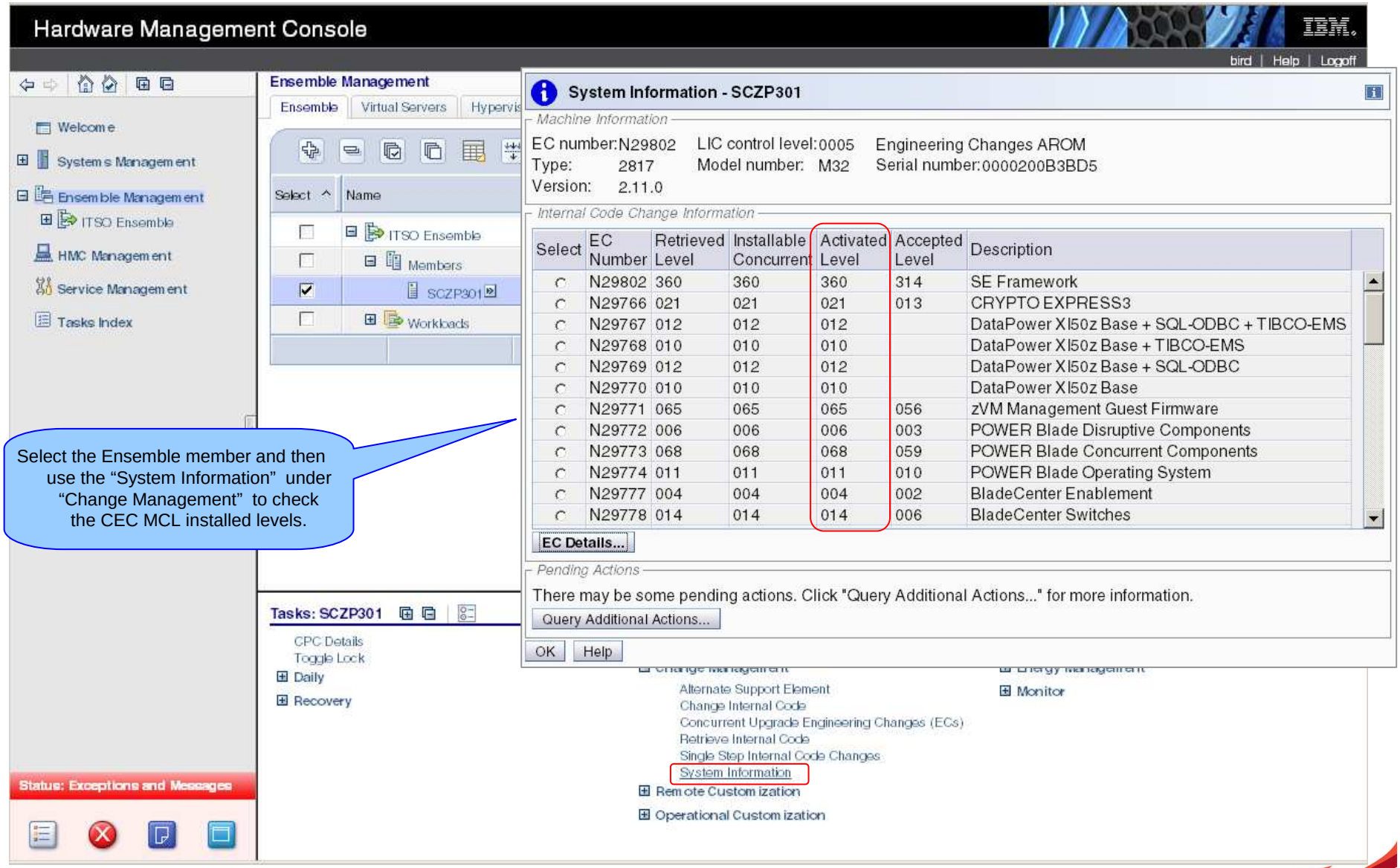
1 Managing the Environment

- *zBX firmware*
- *Installed firmware (MCLs)*
- *Management levels*
- *Performance Management*
- *Ensemble / HMC info*
- *zBX power*
- *Checking blades entitlement*
- *HMC Monitoring functions*

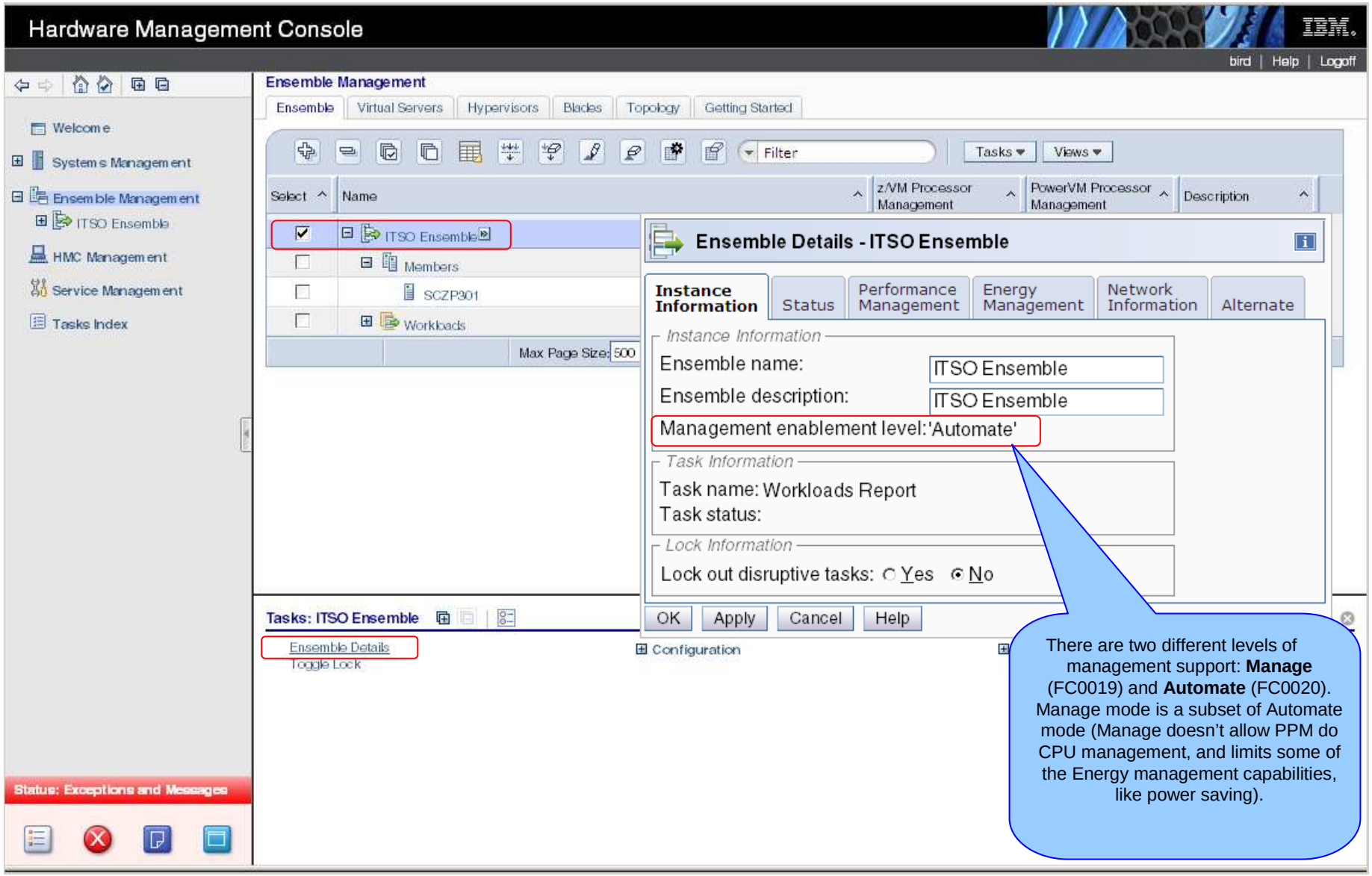
Unified Resource Manager related Firmware

- zBX Rack
  - Top-Of-Rack Switch
  - INMN TOR SW firmware
  - IEDN TOR SW firmware
  - BladeCenter
  - AMM, External Switch Modules (10GbE SW, 8Gb FC SW)
- ISAOPT (IBM Smart Analytics OPTimizer) Blade
  - HS22 component firmwares : uEFI(BIOS), iMM, I/O Adapter, Diagnostics
  - ISAOPT Operating System (SLES) and Application
  - ISAOPT Agent Code (SE Agent Code)
- Power Blade
  - Lower Layer FW : PHYP, Partition FW, FSP, I/O Adapter
  - Image FW : VIOS (Virtual I/O Server - AIX, Virtualization, IVM)
  - Component FW : SE Agent, HPM, FFDC, Auth, RAS, Tools Surveillance Daemon
  - OSes which are running in Virtual Servers considered Software (not firmware)
- DataPower XI50z
  - Lower Layer FW : uEFI, iMM, I/O Adapter, Diagnostics
  - Image FW : MCP
  - Application : AMP Server, DP Application
- System x Blade
  - Lower Layer FW : uEFI, iMM, I/O Adapter, Diagnostics
  - Image FW : RHEV-H (XHYP & Redhat - Virtualization on Redhat linux)
  - Component FW : SE Agent, HPM, FFDC, Auth, RAS, Libs, Tools Surveillance Daemon
  - OSes which are running in Virtual Servers considered Software (not firmware)
- zVM
  - zVMMG - zVM Management Guest

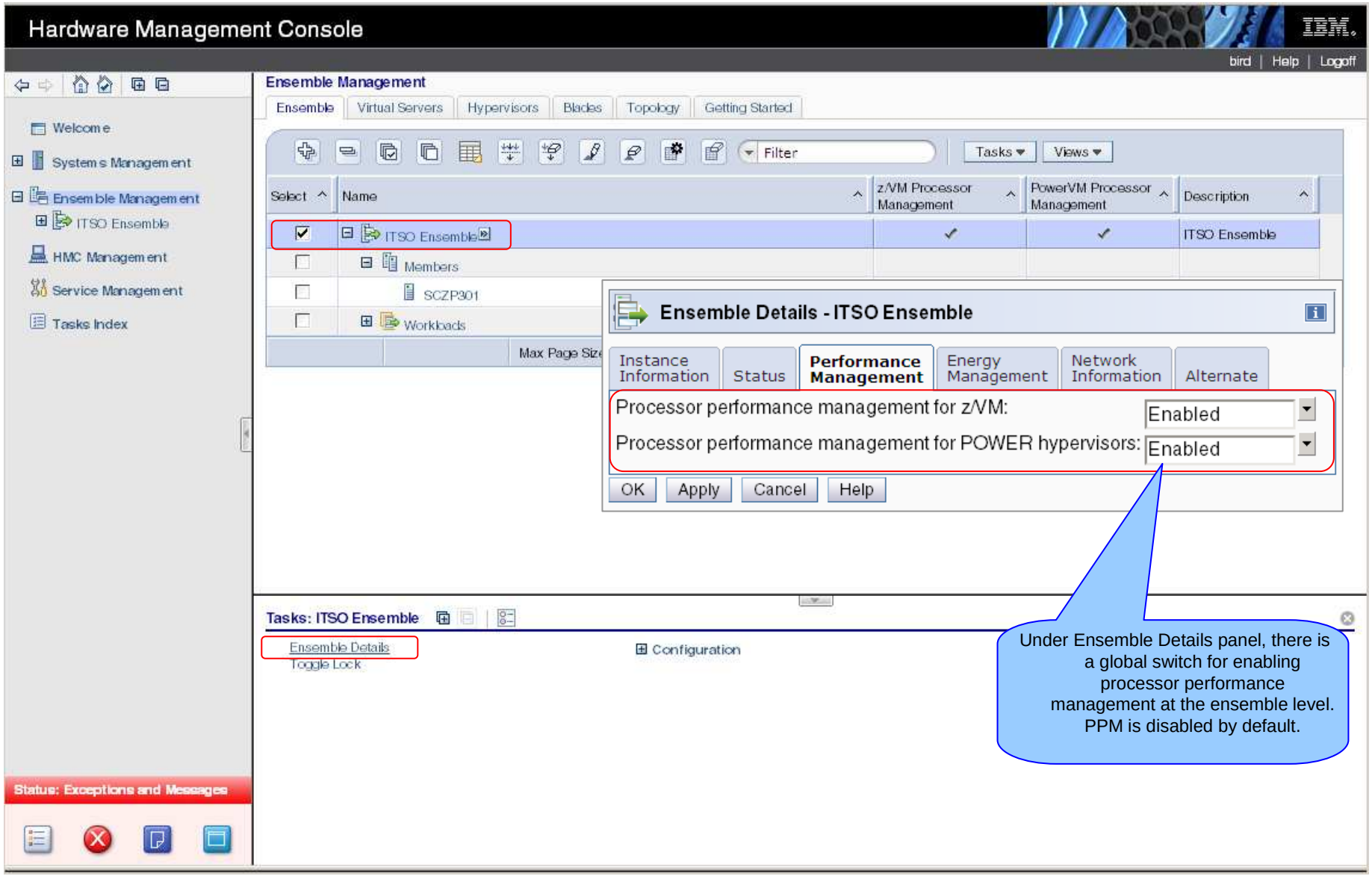
Verifying Ensemble Firmware EC (microcode) Levels



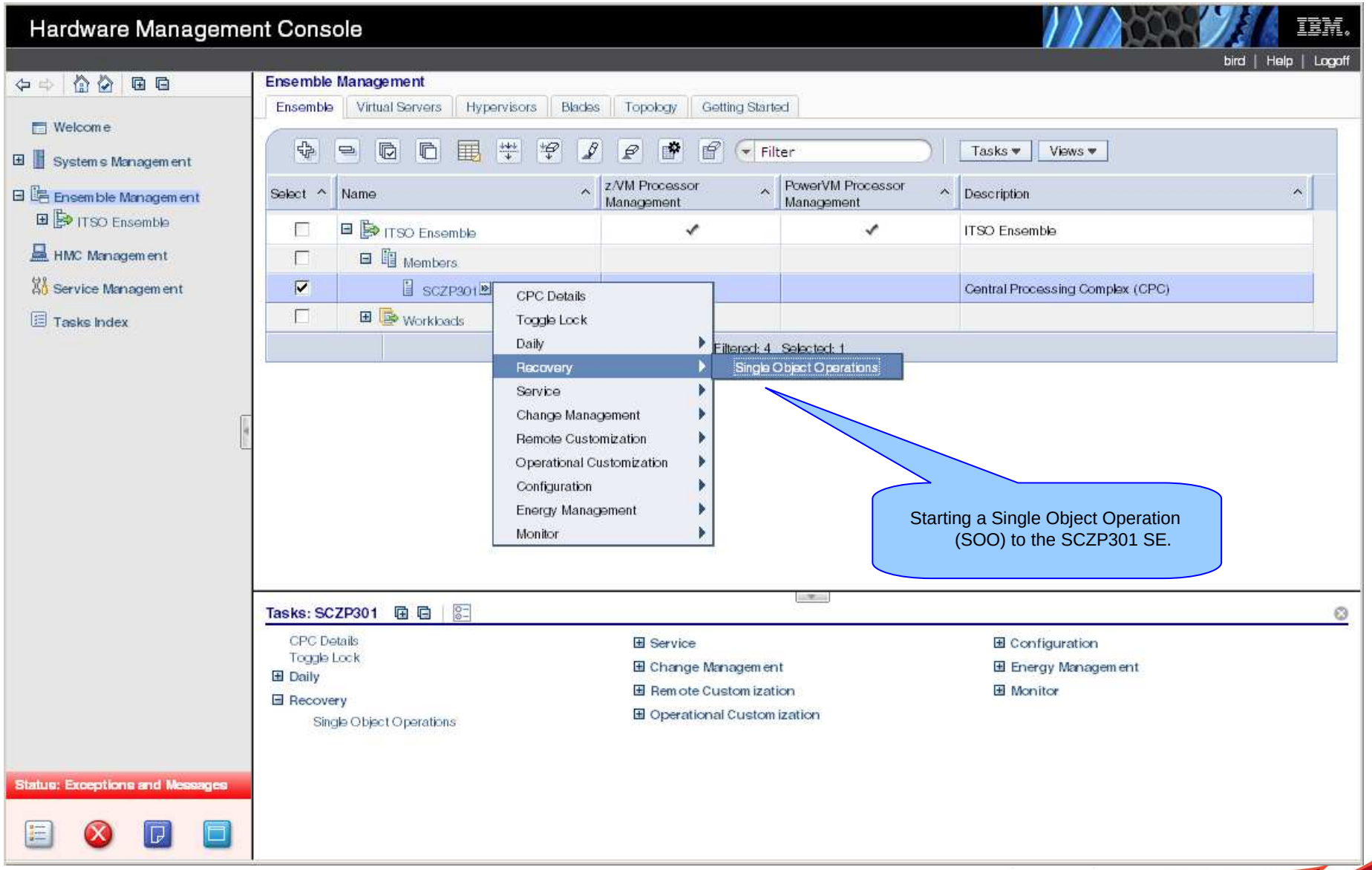
Checking the Ensemble management levels



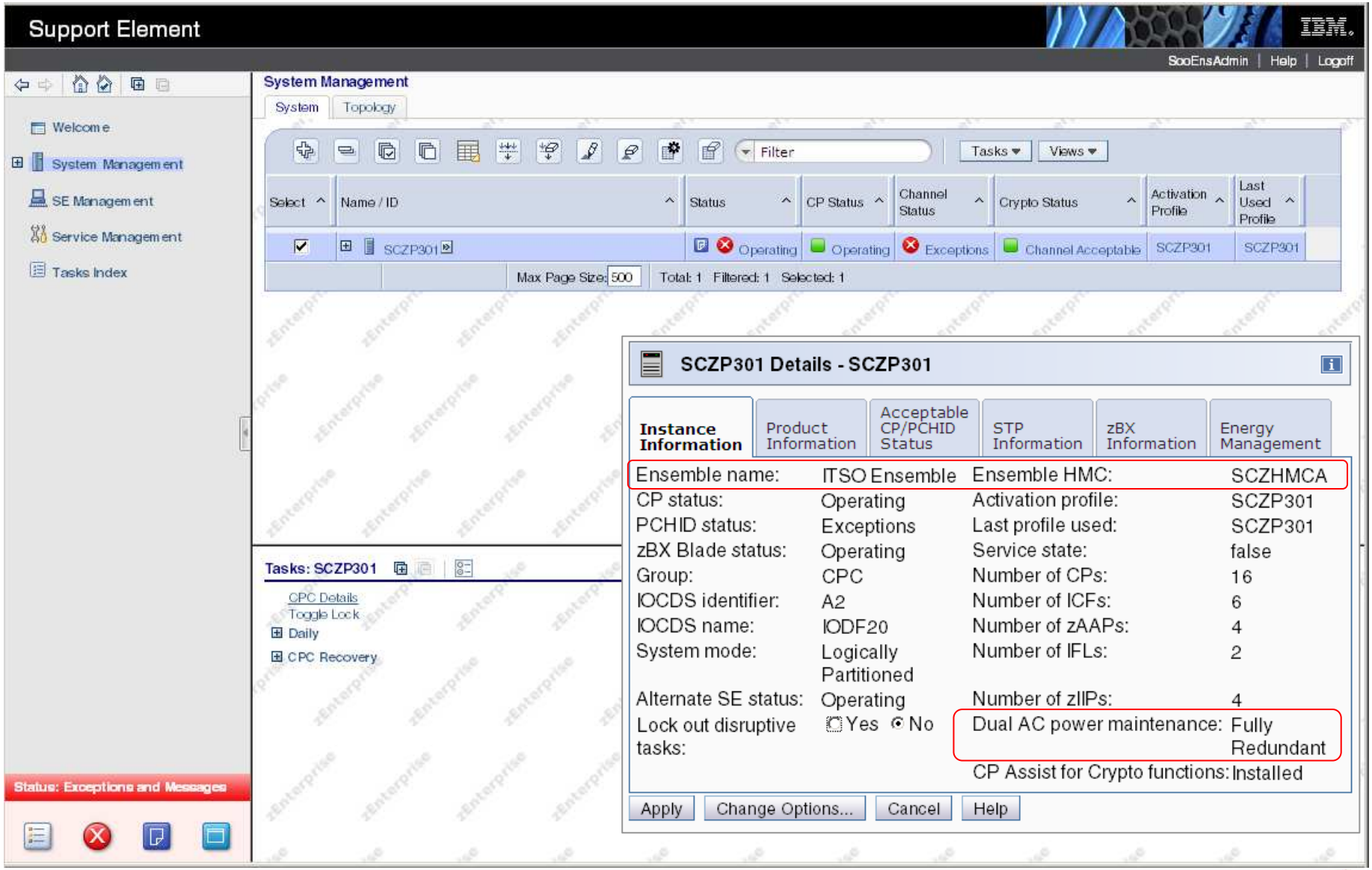
Enabling the Ensemble Performance Management



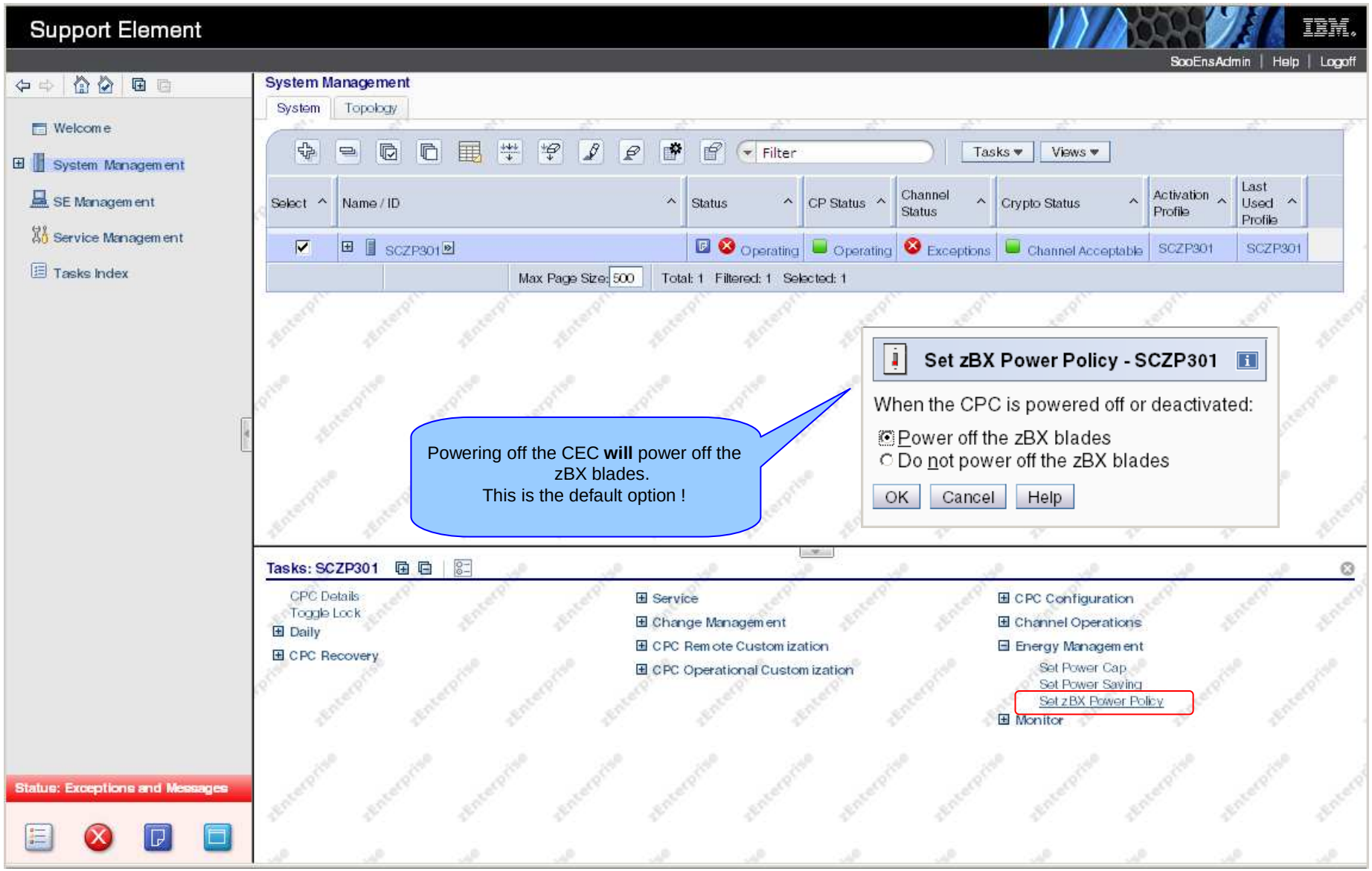
SOO to the Ensemble member SE



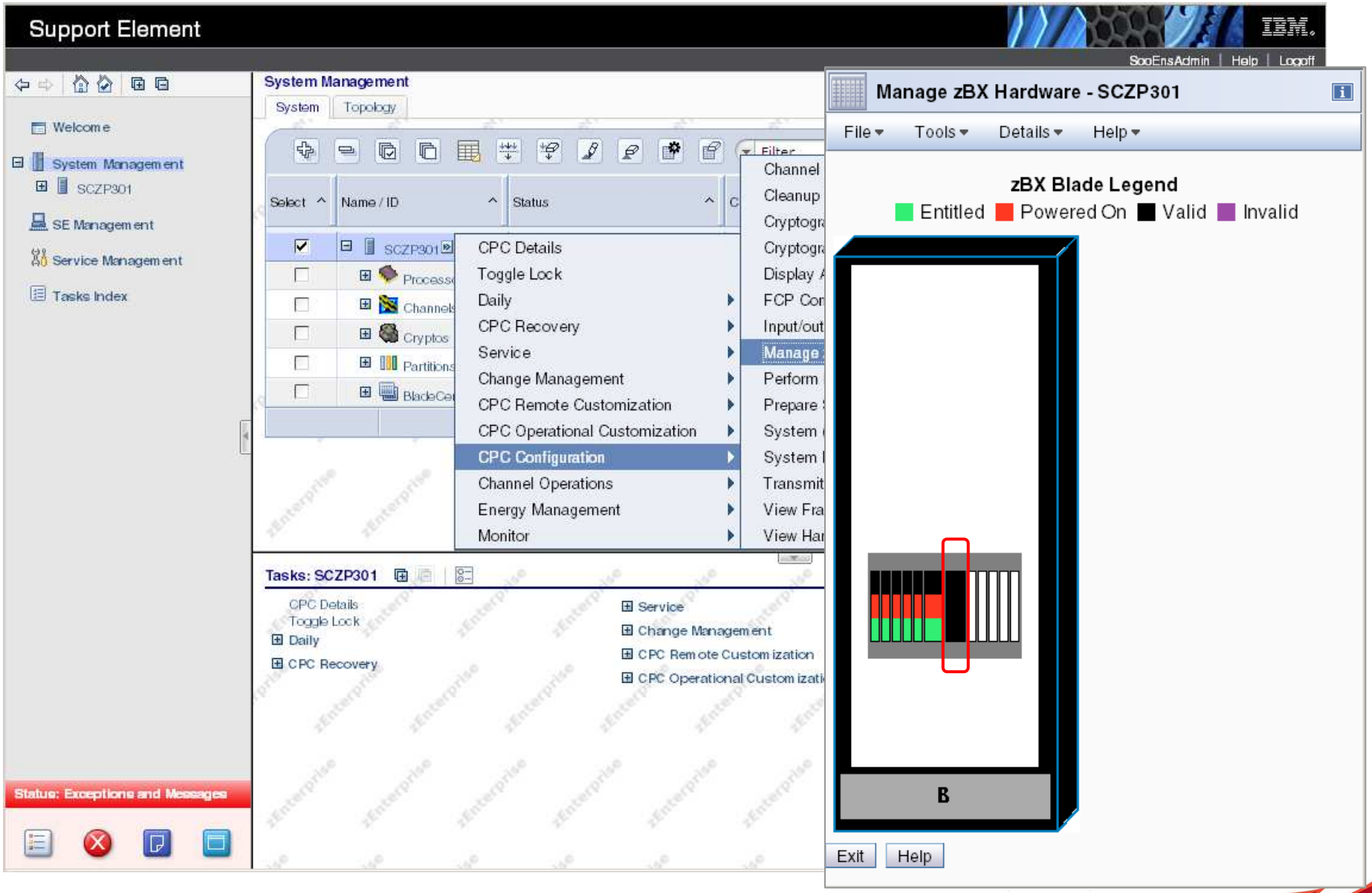
Checking SE Ensemble membership



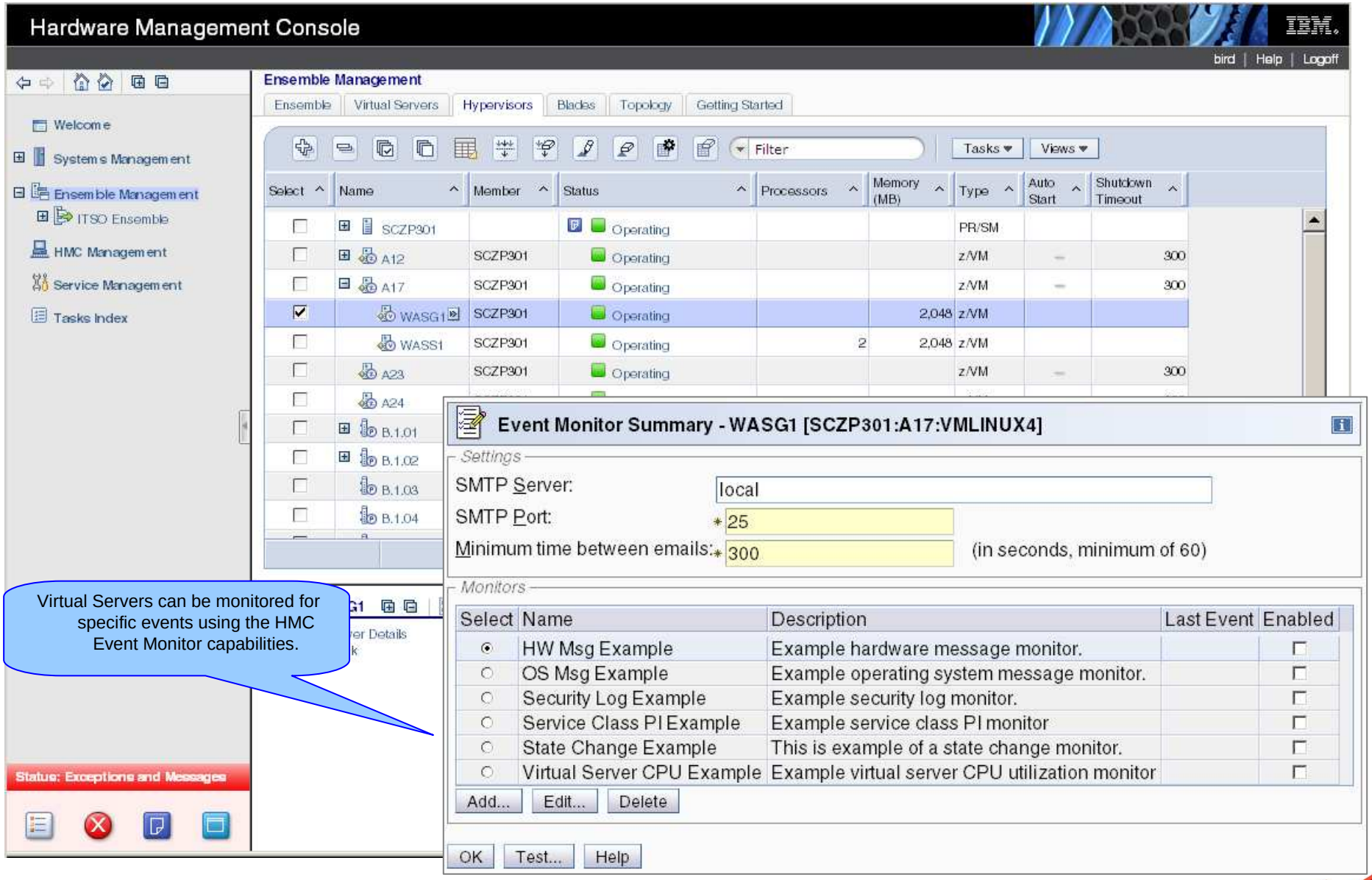
zBX and z196 Power On / Power Off option



zBX – Blade Center – Checking blades entitlement



Monitoring System events



Part IV – Ensemble Managing and Problem Determination

2 z/VM in an Ensemble

- *Implementation*
- *IPL*
- *z/VM SMAPI Servers*
- *System Management APIs (SMAPI)*
- *Enabling z/VM*
- *Implementation tips*
- *Checking the environment*
- *Management Guest*
- *Network Management*
- *Virtual Switches*
- *Shared File System (SFS)*

z/VM Support for zEnterprise 196 and Unified Resource Manager

- z/VM 5.4
  - Compatability → support for hosts/guest on z196 at the z10 functional level with limited exploitation of new functions
    - Support available as PTFs
    - Includes PTFs for EREP, IOCP, HCD, HCM, and Performance Toolkit
    - See http://www.vm.ibm.com/service/vmreqze.html
- z/VM 6.1
  - Exploitation → support allowing the Unified Resource Manager to provide hypervisor and virtual server management for z/VM
    - available with current RSU and PTFs: VM64822 + VM64904, VM64917, VM64956, VM64957
    - Requires minimum Bundle 41Z or higher on the z196
    - Check pre-reqs here: http://www.vm.ibm.com/service/vmrequrm.html
- Software
  - Supported SLES and RHEL distributions
    - Optional – GPMP: Guest Management Platform Provider
    - IEDN / INMN (OSX/OSM) NIC support
    - Legacy NIC connection to IEDN or INMN via Virtual Switch (OSDSIM support)

z/VM Implementation – IPL System

- When CP is loaded:
  - Checks to see if CEC is ensemble managed
  - Keeps OSM and OSX devices offline
  - Does not define any INMN and IEDN vswitches

- VSMGUARD[1] started by AUTOLOG1
  - Checks SMAPI configuration
  - Checks to see if CEC ensemble managed, and CP capable of being ensemble managed
  - Set Ensemble
    - Vary on OSM and OSX chpids, setup VSWITCHes, VSWITCH controllers
    - Get ensemble ID, set MACPREFIX, tell CP it is now in managed state
  - Autolog all SMAPI servers and Management Guest

- Normal customer guests and networking connections started as usual

- INMN and IEDN connections are delayed until SMAPI sets up ensemble environment

[1] VSMGUARD replaces VSMWORK1 as preferred machine to autolog for initiating SMAPI

z/VM SMAPI – System Management APIs

- As part of the support for the IBM zEnterprise Unified Resource Manager, new SMAPI servers were also created:
  - AF_MGMT request server – Used to communicate between the SE and SMAPI
  - INET6 request severs – Use IPv6 to connect with clients
  - VSMGUARD worker server – Guard server to provide resiliency and error recovery
  - Management Guest (ZVMLXAPP) – Automatic instantiation by the Unified Resource Manager
  - New Systems Management APIs added

- SMAPI Servers
  - VSMGUARD – The VSMGUARD sever is a new worker server that provides better resiliency and error recovery. You start this server and it automatically start the remaining SMAPI servers and management guest. Unlike the worker servers, VSMGUARD does not process any request.
  - VSMREQIM – The VSMREQIM is a AF_MGMT request server. The AF_MGMT request server is used to communicate between the support element and the z/VM SMAPI server environment, only when z/VM is managed by the Unified Resource Manager. There can be one and only one AF_MGMT request server.
  - VSMREQI6 – The VSMREQI6 is the AF_INET6 request server. This server handles requests over the IPV6 sockets • VSMREQIN VSMREQIN is the AF_INET request server. This server handles request over the IPV4 sockets.
  - VSMPROXY - The VSMPROXY is the AF_SCLP request server. This server is used for communication between the support element and the z/VM SMAPI server environment. There can be one and only one AF_SCLP server.

  more…

z/VM SMAPI – System Management APIs

- SMAPI Servers (cont.)
  - VSMREQIU – The VSMREQIU is the AF_IUCV request server. There can be one or more AF_IUCV request servers.
  - VSMWORK1 – The VSMWORK1 is the short call request server. It is one of the three default worker servers. There must always be at least one short call worker server. The default SFS directories are owned by the “short call” request server VSMWORK1.
  - VSMWORK2 – The VSMWORK2 is a long call request server. This is the one of two long call request servers. If all the request servers are busy, the request will be queued until on becomes available.
  - VSMWORK3 – The VSMWORK3 is a long call request server. This is the one of two long call request servers. If all the request servers are busy, the request will be queued until on becomes available.
  - ZVMLXAPP – The ZVMLXAPP is a the new Management Guest. The Management Guest is automatically instantiated by the Unified Resource Manager. It is also part of the INMN network communication past to Linux guest.

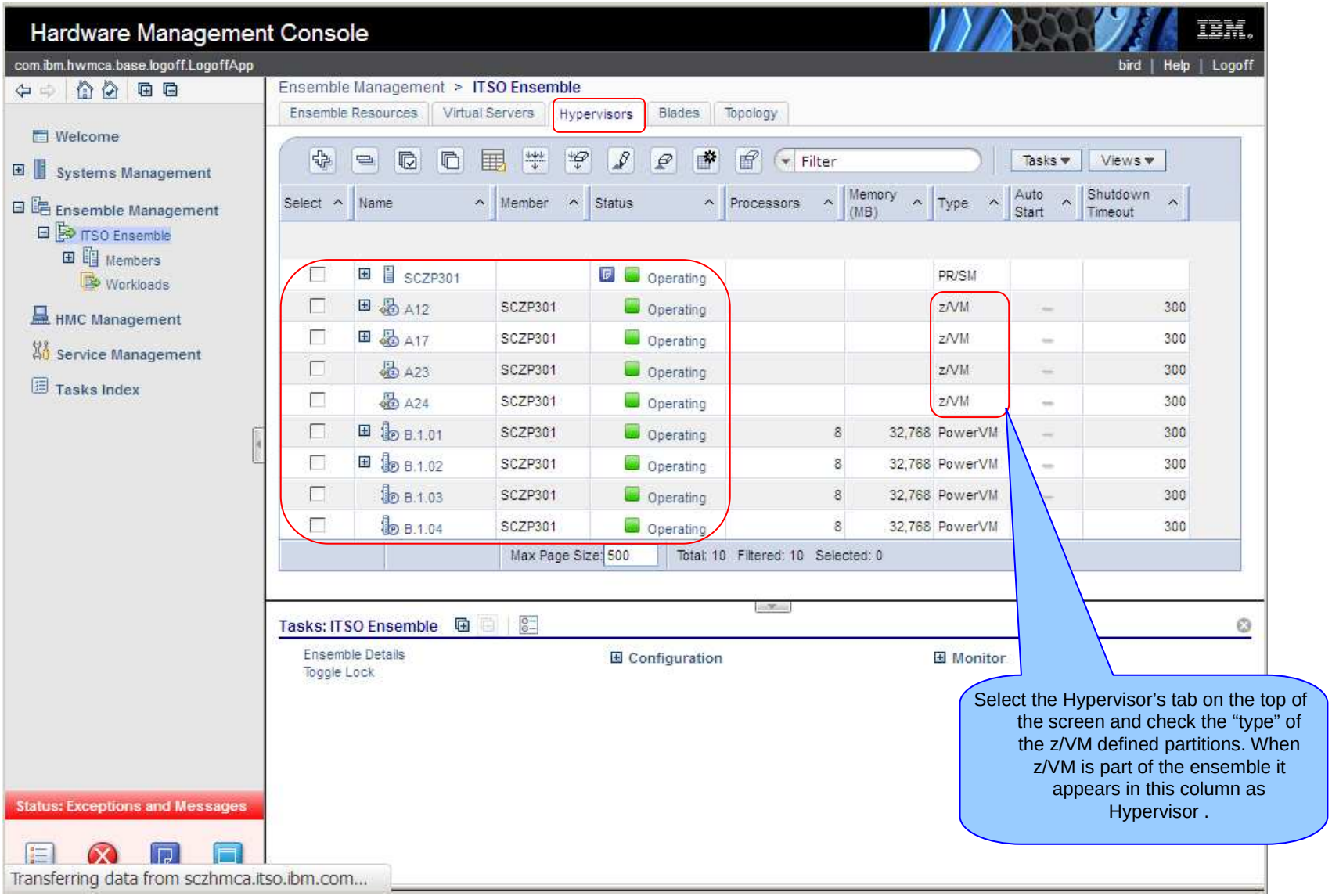
Enabling z/VM to be managed by Unified Resource Manager

- (Must have) References
  - CP Planning and Admin Guide (SC24-6178-01)
    - Chapter 16 for all the detailed installation steps
  - z/VM System Management Application Programming (SC24-6234-01)
  - z/VM CP Commands and Utilities Reference SC24-6175-01
  - z/VM Directory Maintenance Facility Commands Reference SC24-6188-01
  - zEnterprise Ensemble Performance Management Guide GC27-2607-01
  - zEnterprise Ensemble Planning and Configuration Guide GC27-2608-01
  - IBM zEnterprise Unified Resource Manager Redbook SG24-7921
    - Chapter 8 for all the detailed installation steps

z/VM implementation tips

- Existing SMAPI servers need their existing PROFILE EXECs updated
- You can NOT manually define via CP commands an IEDN VSWITCH and attach to a guest. This must be done using the Unified Resource Manager on the HMC
- z/VM Ensemble configuration and logging in is done in VMSYS file system.
  - Back it up with the rest of your system.
- Resources must work without Unified Resource Manager if they are going to work with Unified resource Manager.
  - FCP devices must be able to access LUNs without Unified Resource Manager if they are going to be accessed with Unified Resource Manager
- Console output from VSMGUARD, VSMWORK1, and VSMREQIU can be used as a source for Problem Determination
- If you vary all devices offline in the SYSTEM CONFIG and then vary on only the ones you know about, Unified Resource Manager defined FCP devices could be a problem. You may want to have a predefined range of devices for this
- Unified Resource Manager is not a RACF security administration application
- DIRMAINT actions driven by Unified Resource Manager can still invoke the DIRMAINT RACF interface
- If ZVMLXAPP does not start, the other SMAPI service machines will not be started

z/VM and Linux on System z Virtual Servers



Checking z/VM environment using “maint” id

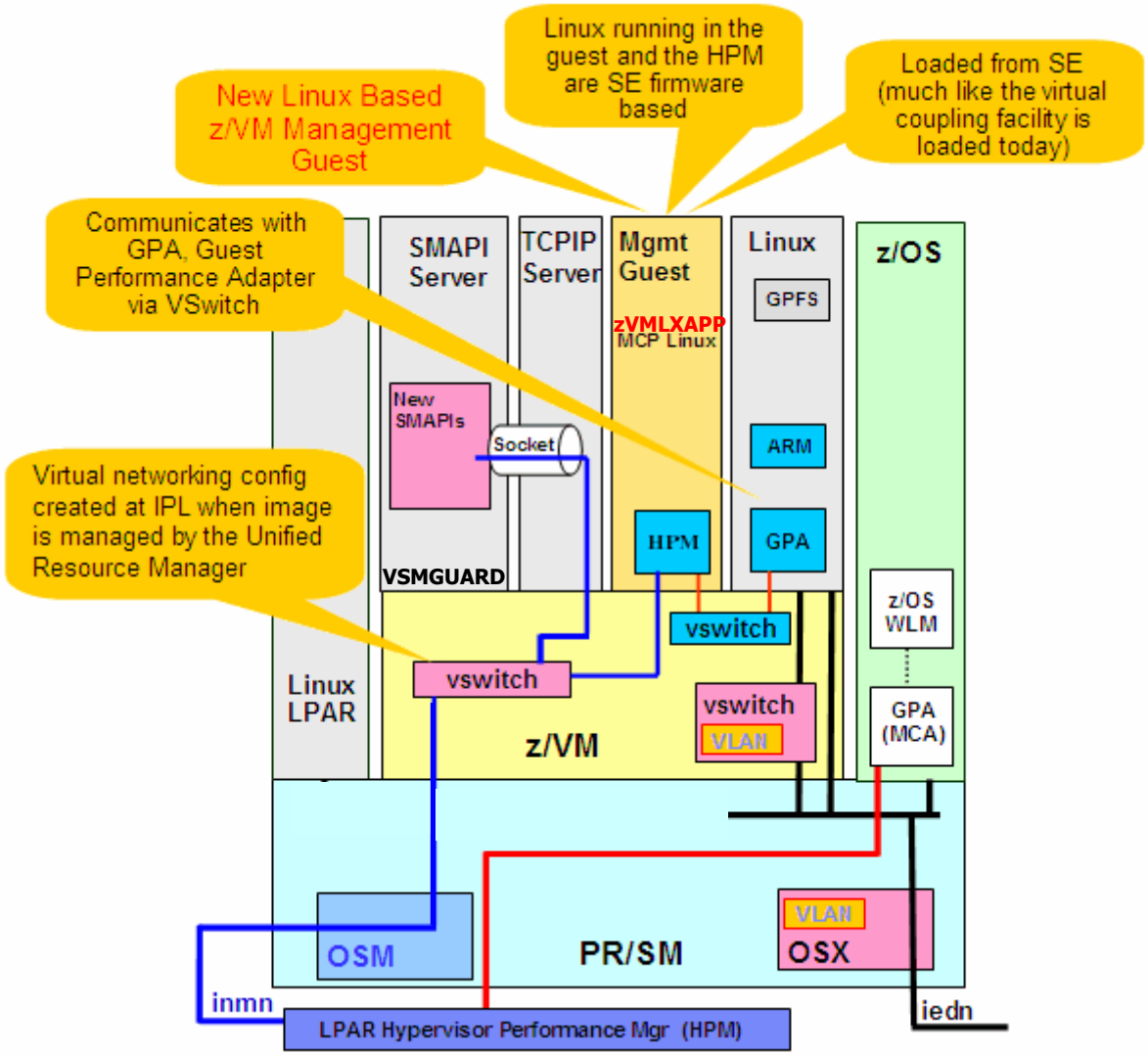
- z/VM SMAPI Servers
  - Various service machines are used for SMAPI and mgmt guest support
    - VSMREQIU -- Default AF_IUCV server
    - VSMPROXY -- Default AF_SCLP server
    - VSMREQIM -- Default Management Network Server (communicates over OSM)
    - VSMREQIN -- SMAPI request Server
    - VSMGUARD -- This server starts ZVMLXAPP and z/VM ensemble bring up
    - VSMWORK1 -- Default short call server
    - VSMWORK2 -- Default long call server
    - VSMWORK3 -- Default long call server
    - DTCENS1 -- VSWITCH controller for red network
    - DTCENS2 -- Backup VSWITCH controller for red network
    - ZVMLXAPP -- z/VM Management Guest
    - DIRMAINT -- Directory maintenance server

```
q n
WASG1    - DSC , FTPSERVE - DSC , TCPIP    - DSC , VSMPROXY - DSC
OP1      - DSC , HAIMO    - DSC , VSMREQIU - DSC , VSMREQI6 - DSC
VSMREQIN - DSC , VSMWORK3 - DSC , VSMWORK2 - DSC , VSMWORK1 - DSC
ZVMLXAPP - DSC , VSMREQIM - DSC , DTCENS2  - DSC , DTCENS1  - DSC
VSMGUARD - DSC , RSCS     - DSC , PVM      - DSC , DATAMOVE - DSC
DIRMAINT - DSC , DTCVSW2  - DSC , DTCVSW1  - DSC , VMSERVR  - DSC
VMSERVU  - DSC , VMSERVS  - DSC , GCS      - DSC , OPERSYMP - DSC
OPERATOR - DSC , WASS1    - DSC , MAINT    -L0004
VSM      - TCPIP
Ready; T=0.01/0.01 11:22:15
```

- NOTE: When z/VM is participating in an ensemble, SMAPI server will automatically start DirMaint (if it isn’t started already) and use it to do all directory updates.

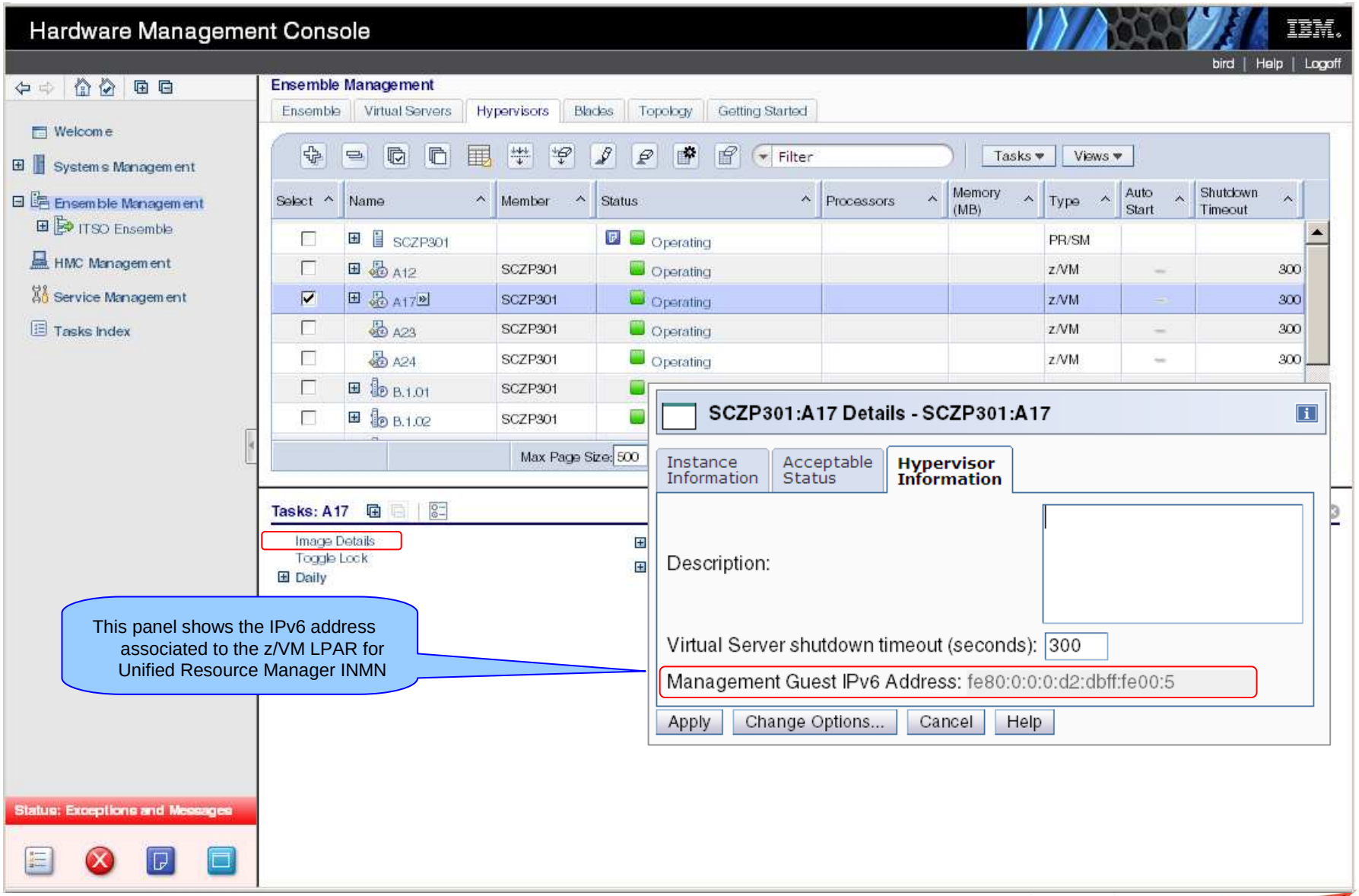
z/VM Management Guest

- ZVMLXAPP is the z/VM management guest
  - Runs the PPM Hypervisor Performance Manager (HPM) for z/VM
  - It is started by service machine VSMGUARD
  - It runs MCP Linux automatically loaded from the SE
  - The command Q NAMES will show if ZVMLXAPP is running

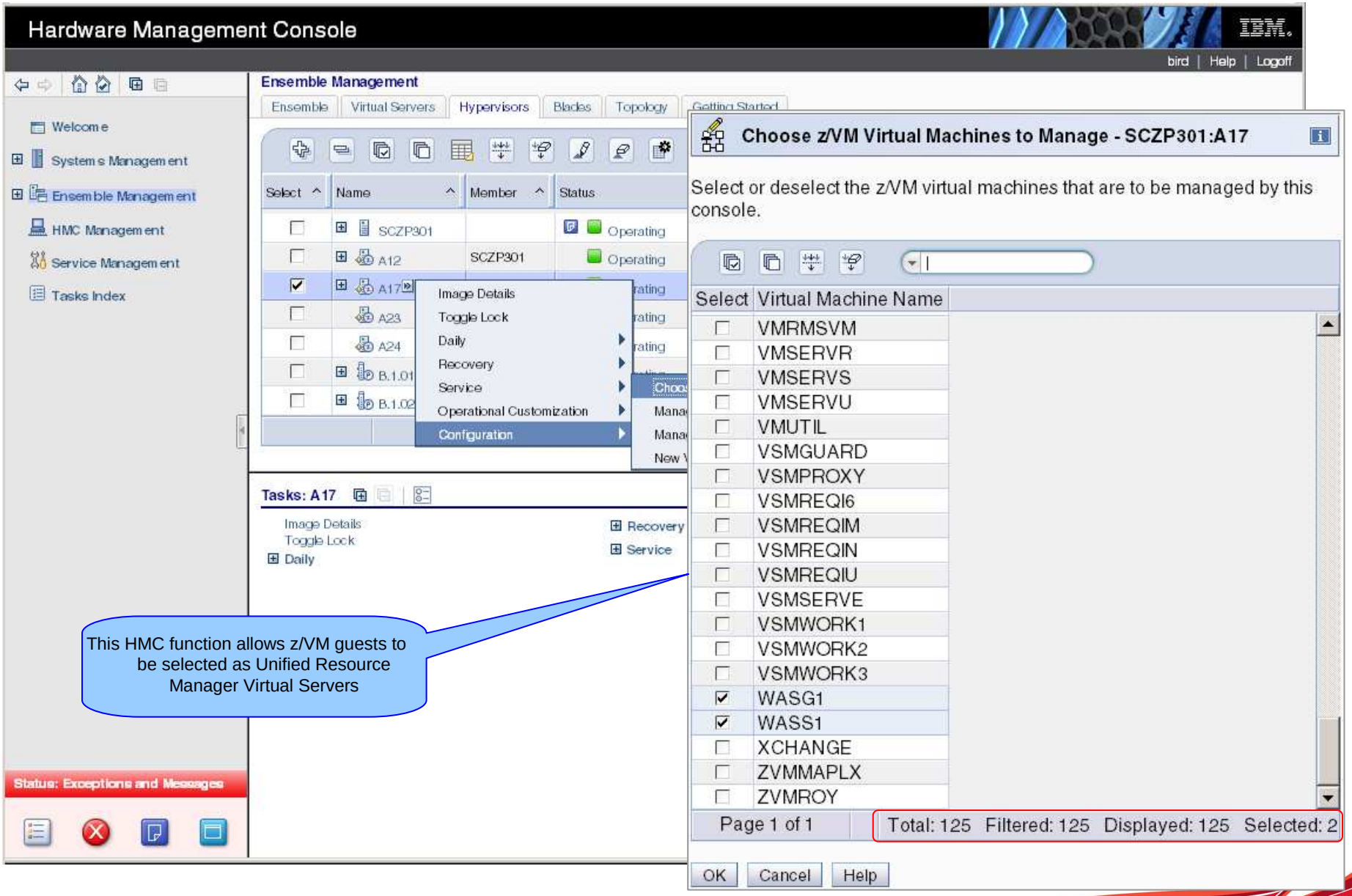



```
q n
WASG1     - DSC , FTPSERVE - DSC , TCPIP     - DSC , VSMPROXY - DSC
OP1       - DSC , HAIMO    - DSC , VSMREQIU  - DSC , VSMREQI6 - DSC
VSMREQIN  - DSC , VSMWORK3 - DSC , VSMWORK2  - DSC , VSMWORK1 - DSC
ZVMLXAPP  - DSC , VSMREQIM - DSC , DTCENS2   - DSC , DTCENS1  - DSC
VSMGUARD  - DSC , RSCS     - DSC , PVM       - DSC , DATAMOVE - DSC
DIRMAINT  - DSC , DTCVSW2  - DSC , DTCVSW1   - DSC , VMSERVR  - DSC
VMSERVU   - DSC , VMSERVS  - DSC , GCS       - DSC , OPERSYMP - DSC
OPERATOR  - DSC , WASS1    - DSC , MAINT     -L0004
VSM       - TCPIP
Ready; T=0.01/0.01 11:22:15
```

z/VM and management network



Choosing z/VM Guests to be managed



z/VM and Linux on System z Virtual Servers

- To the right is the output of a ***QUERY VMLAN*** command entered from the maint id.
  - Note the Unified Resource Manager status shown. (it shows the expected command output)
- Below is the output of a ***QUERY OSA TYPE ENSEMBLE*** command entered from the maint id.
  - Note the OSA devices attached to the DTCENS1 and DTCENS2 switch controllers. (the expected command output is shown)

```
q vmlan
VMLAN maintenance level:
  Latest Service: VM64780
VMLAN MAC address assignment:
  System MAC Protection: OFF
  MACADDR Prefix: 020000 USER Prefix: 020000
  MACIDRANGE SYSTEM: 000001-FFFFFF
             USER:   000000-000000
VMLAN Unified Resource Manager status:
  Hypervisor Access: YES       Status: MANAGED
  ID: 52BD737254BF11E0B85A0010184CB262
  MAC Prefix: 02D2DB
VMLAN default accounting status:
  SYSTEM Accounting: OFF       USER Accounting: OFF
VMLAN general activity:
  PERSISTENT Limit: INFINITE   Current: 5
  TRANSIENT  Limit: INFINITE   Current: 0
Ready; T=0.01/0.01 17:39:08
```

```
q osa type ens
OSA  2300 ATTACHED TO DTCENS2  2300 DEVTYPE IEDN          CHPID 18 OSX
OSA  2301 ATTACHED TO DTCENS2  2301 DEVTYPE IEDN          CHPID 18 OSX
OSA  2302 ATTACHED TO DTCENS2  2302 DEVTYPE IEDN          CHPID 18 OSX
OSA  2303 ATTACHED TO DTCENS2  2303 DEVTYPE IEDN          CHPID 18 OSX
OSA  2304 ATTACHED TO DTCENS2  2304 DEVTYPE IEDN          CHPID 18 OSX
OSA  2305 ATTACHED TO DTCENS2  2305 DEVTYPE IEDN          CHPID 18 OSX
OSA  2306 ATTACHED TO DTCENS2  2306 DEVTYPE IEDN          CHPID 18 OSX
OSA  2307 ATTACHED TO DTCENS2  2307 DEVTYPE IEDN          CHPID 18 OSX
OSA  2308 ATTACHED TO DTCENS2  2308 DEVTYPE IEDN          CHPID 18 OSX
OSA  2309 FREE                      DEVTYPE IEDN          CHPID 18 OSX
OSA  230A FREE                      DEVTYPE IEDN          CHPID 18 OSX
OSA  230B FREE                      DEVTYPE IEDN          CHPID 18 OSX
OSA  231D FREE                      DEVTYPE IEDN          CHPID 18 OSX
OSA  231E FREE                      DEVTYPE IEDN          CHPID 18 OSX
OSA  231F FREE                      DEVTYPE IEDN          CHPID 18 OSX
OSA  2320 ATTACHED TO DTCENS1  2320 DEVTYPE IEDN          CHPID 19 OSX
OSA  2321 ATTACHED TO DTCENS1  2321 DEVTYPE IEDN          CHPID 19 OSX
OSA  2322 ATTACHED TO DTCENS1  2322 DEVTYPE IEDN          CHPID 19 OSX
OSA  2323 ATTACHED TO DTCENS1  2323 DEVTYPE IEDN          CHPID 19 OSX
OSA  2324 ATTACHED TO DTCENS1  2324 DEVTYPE IEDN          CHPID 19 OSX
OSA  2325 ATTACHED TO DTCENS1  2325 DEVTYPE IEDN          CHPID 19 OSX
OSA  2326 ATTACHED TO DTCENS1  2326 DEVTYPE IEDN          CHPID 19 OSX
OSA  2327 ATTACHED TO DTCENS1  2327 DEVTYPE IEDN          CHPID 19 OSX
```

Validating the configuration – VSWITCHes controllers

```
q controller
Controller DTCVSW2   Available: YES   VDEV Range: *         Level 610
  Capability: IP ETHERNET VLAN_ARP GVRP    LINKAGG    ISOLATION
              NO_ENSEMBLE NO_INMN
Controller DTCVSW1   Available: YES   VDEV Range: *         Level 610
  Capability: IP ETHERNET VLAN_ARP GVRP    LINKAGG    ISOLATION
              NO_ENSEMBLE NO_INMN
Controller DTCENS1   Available: YES   VDEV Range: *         Level 610
  Capability: IP ETHERNET VLAN_ARP GVRP    LINKAGG    ISOLATION
              ENSEMBLE    INMN
   SYSTEM DTCINMN    Primary          Controller: DTCENS1   VDEV: 236D
   SYSTEM DTCINMN    Backup           Controller: DTCENS1   VDEV: 234D
   SYSTEM QDIO099    Backup           Controller: <list>    VDEV: 2320
   SYSTEM QDIO100    Backup           Controller: *         VDEV: 2323
   SYSTEM QDIO110    Backup           Controller: *         VDEV: 2326
   SYSTEM QDIO310    Backup           Controller: *         VDEV: 2329
Controller DTCENS2   Available: YES   VDEV Range: *         Level 610
  Capability: IP ETHERNET VLAN_ARP GVRP    LINKAGG    ISOLATION
              ENSEMBLE    NO_INMN
   SYSTEM QDIO099    Primary          Controller: <list>    VDEV: 2300
   SYSTEM QDIO100    Primary          Controller: *         VDEV: 2303
   SYSTEM QDIO110    Primary          Controller: *         VDEV: 2306
   SYSTEM QDIO310    Primary          Controller: *         VDEV: 2309
Ready; T=0.01/0.01 17:41:30
```

Validating the configuration – query VSWITCHes

```
q vswitch dtcinmn
VSWITCH SYSTEM DTCINMN  Type: INMN    Connected: 2    Maxconn: INFINITE
  PERSISTENT  RESTRICTED    ETHERNET                  Accounting: OFF
  VLAN Unaware
  MAC address: 02-3C-90-00-00-01    MAC Protection: Unspecified
  State: Ready
  IPTimeout: 5         QueueStorage: 8
  Isolation Status: ON
Uplink Port:
  RDEV: 236D.P00 VDEV: 236D Controller: DTCENS1
  RDEV: 234D.P00 VDEV: 234D Controller: DTCENS1  BACKUP
```

```
q vswitch sw2
VSWITCH SYSTEM SW2      Type: INMN    Connected: 1    Maxconn: INFINITE
  PERSISTENT  RESTRICTED    ETHERNET                  Accounting: OFF
  VLAN Unaware
  MAC address: 02-3C-90-00-00-03    MAC Protection: Unspecified
  State: Ready
  IPTimeout: 5         QueueStorage: 8
  Isolation Status: ON
Uplink Port:
  NIC: ZVMLXAPP  VDEV: 0200
```

Shared File System (SFS)

- Shared File System (i.e., VMSYS:) is required by the SMAPI service machines
- z/VM Ensemble management will not work if the SFS is not started
  - Check if SFS service machine, **VMSERVS**, is started using *q names* command, from *maint* id

```
q n
WASG1    - DSC , FTPSERVE - DSC , TCPIP    - DSC , VSMPROXY - DSC
OP1      - DSC , HAIMO    - DSC , VSMREQIU - DSC , VSMREQI6 - DSC
VSMREQIN - DSC , VSMWORK3 - DSC , VSMWORK2 - DSC , VSMWORK1 - DSC
ZVMLXAPP - DSC , VSMREQIM - DSC , DTCENS2  - DSC , DTCENS1  - DSC
VSMGUARD - DSC , RSCS     - DSC , PVM      - DSC , DATAMOVE - DSC
DIRMAINT - DSC , DTCVSW2  - DSC , DTCVSW1  - DSC , VMSERVR  - DSC
VMSERVU  - DSC , VMSERVS  - DSC , GCS      - DSC , OPERSYMP - DSC
OPERATOR - DSC , WASS1    - DSC , MAINT    -L0004
VSM      - TCPIP
Ready; T=0.01/0.01 11:22:15
```

- If SFS fails, you'll see this message:
  - **VSMWORK1 DMSJCA1152S File pool VMSYS is unavailable; accessed directories for this file pool are released**
- *QUERY FILEPOOL CONNECT* command will show the SMAPI servers connected to the SFS

```
query filepool connect
Userid     Connected
VSMGUARD   Yes
VSMREQIM   Yes
VSMWORK1   Yes
VSMREQIN   Yes
VSMPROXY   Yes
VSMREQI6   Yes
VSMWORK1   Yes
VSMWORK3   Yes
VSMWORK2   Yes
VSMREQIU   Yes
MAINT      Yes
Ready; T=0.01/0.01 12:45:00
```

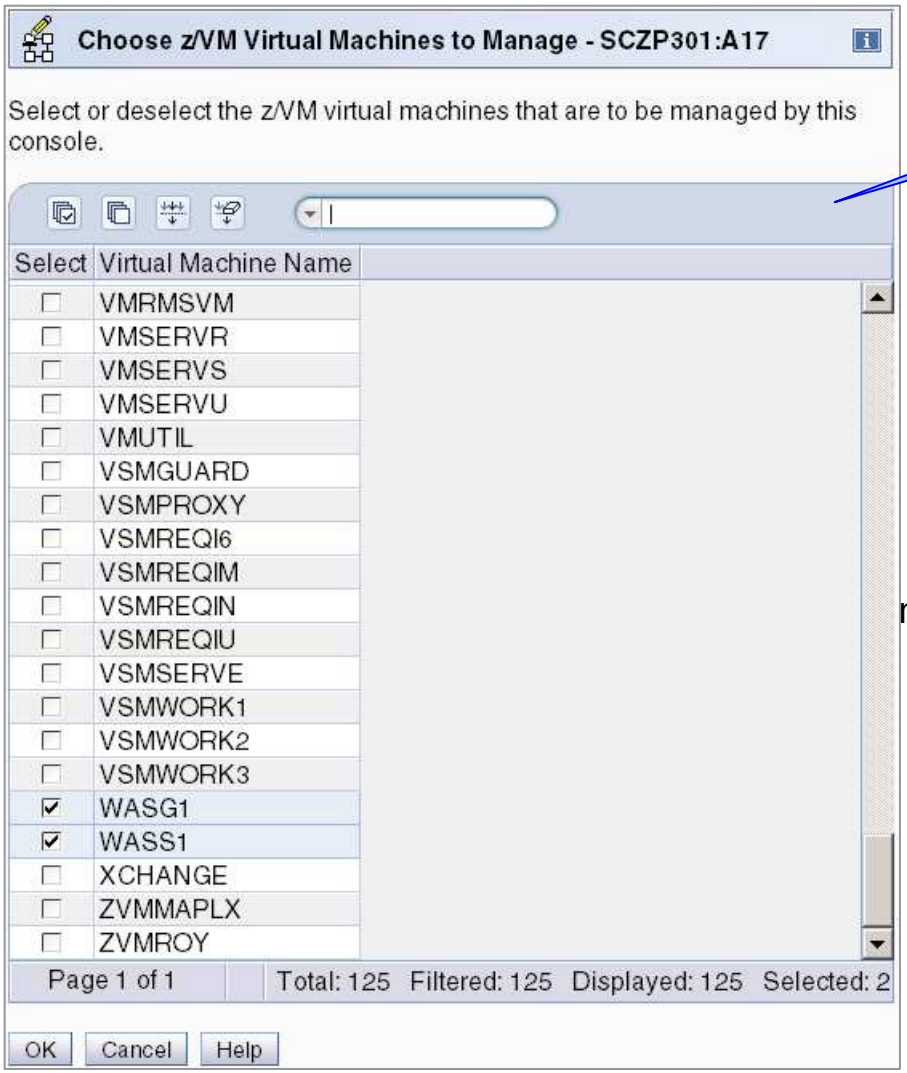
Checking the SMAPI service machines access to SFS

```
q auth vmsys:vsmwork1.
Directory = VMSYS:VSMWORK1.
Grantee  R  W  NR NW
MAINT    X  X  X  X
VSMWORK1 X  X  X  X
VSMGUARD X  X  X  X
VSMPROXY X  -  X  -
VSMREQIM X  X  X  X
VSMREQIN X  -  X  -
VSMREQIU X  -  X  -
VSMREQI6 X  -  X  -
VSMWORK2 X  -  X  -
VSMWORK3 X  -  X  -
Ready; T=0.01/0.01 17:30:02
```

```
q auth vmsys:vsmwork1.data.
Directory = VMSYS:VSMWORK1.DATA
Grantee  R  W  NR NW
MAINT    X  X  X  X
VSMWORK1 X  X  X  X
FTPSERVE X  -  -  -
VSMGUARD X  X  X  X
VSMPROXY X  X  X  X
VSMREQIM X  X  X  X
VSMREQIN X  X  X  X
VSMREQIU X  X  X  X
VSMREQI6 X  X  X  X
VSMWORK2 X  X  X  X
VSMWORK3 X  X  X  X
Ready; T=0.01/0.01 17:37:26
```

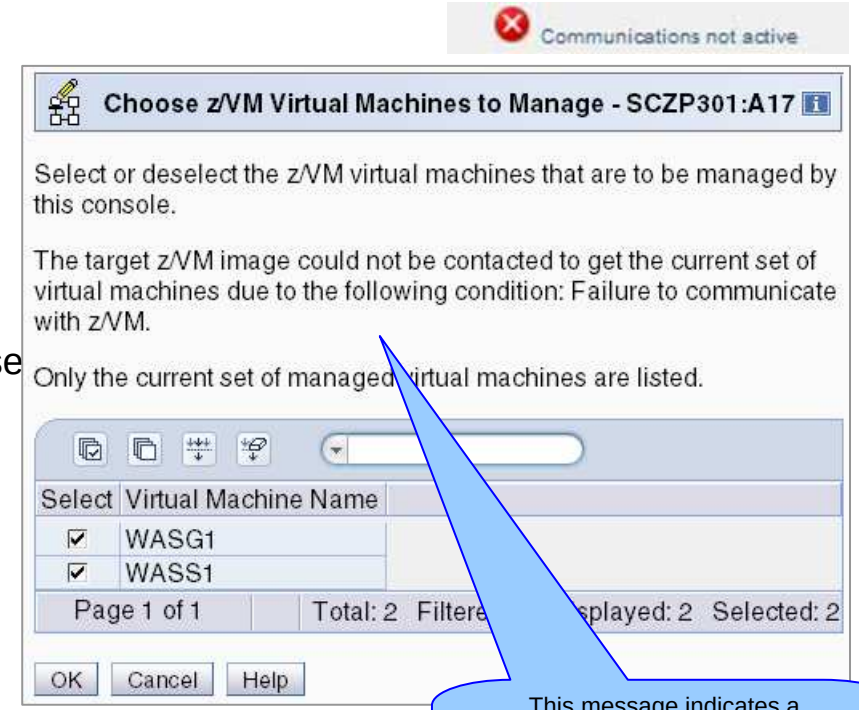
z/VM Virtual Server and Unified Resource Manager communication problem example

- Adding z/VM guests (Virtual Servers) as managed



Example of adding z/VM Virtual machines to be managed by Unified Resource Manager.

nse



This message indicates a communication problem between the Virtual Servers and Unified Resource Manager.

Shared File System (SFS)

- *QUERY FILEPOOL REPORT* command from *maint* id will provide SFS information below when it is up and running
  - The *QUERY FILEPOOL REPORT* command has an extensive output – use *enter / clear* keys to read it

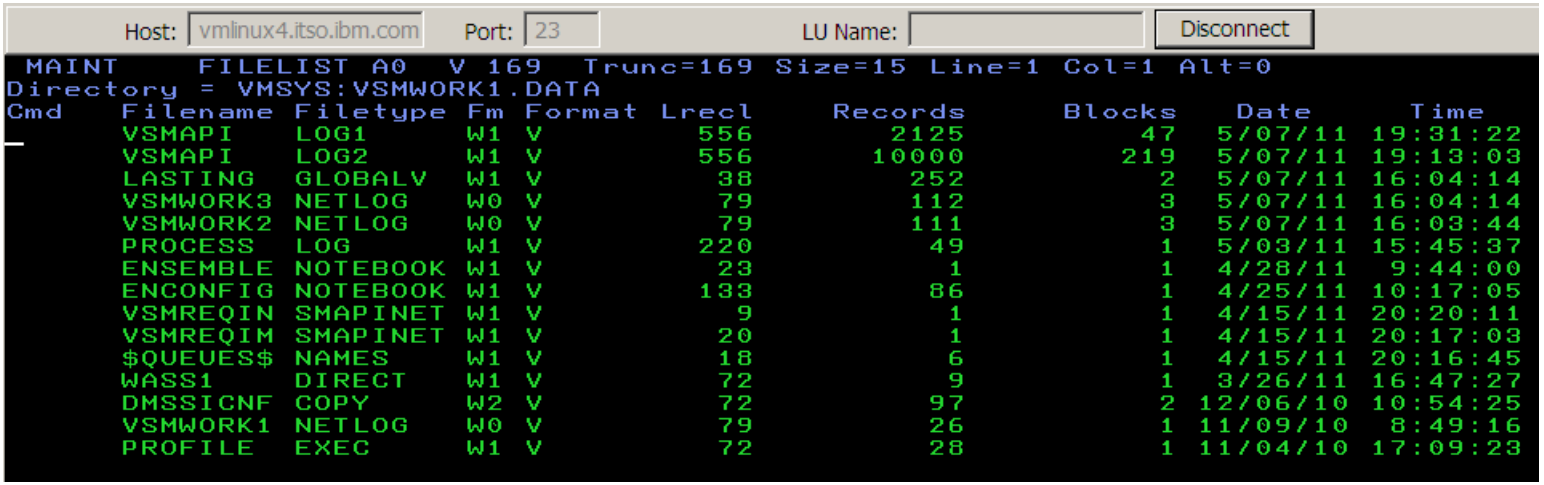
The QUERY FILEPOOL REPORT command lists:

- File Pool counter information
- CRR counter information (only if CRR Recovery Server)
- File pool compare information
- File pool information
- Currently defined minidisk information
- Agent information
- Repository file pool log information
- Catalog space information (optional)

```
query filepool report
                    VMSYS    File Pool Report

Start-up Date 04/15/11                                  Query Date 05/08/11
Start-up Time 20:16:19                                  Query Time 05:33:48
================================================================================
FILE POOL OVERVIEW INFORMATION

   67108864  Virtual Storage Size in Bytes (     65536 in KB)
      62262  Virtual Storage Highest Value in KB
          0  Virtual Storage Requests Denied
         19  Virtual Storage Reclaim Value

       2000  Maximum Number of Connections
         18  Connections Highest Value

         16  Total Number of Agents
          7  Active Agents Highest Value

        500  Maximum Number of Storage Groups
          2  Storage Groups in Use

        500  Maximum Number of Minidisks
          2  Minidisks in Use

   56074240  Potential Addressable 4K Blocks in File Pool
      39005  Defined Addressable 4K Blocks in File Pool
   56035235  Undefined Addressable 4K Blocks in File Pool
================================================================================
STORAGE GROUP MINIDISK INFORMATION

 Storage       Minidisk          4K Blocks         4K Blocks          Virtual
 Group No.     Number              In-Use             Free             Address
     1             1             145 -    2%         8838              0304
     2             2            1618 -    5%        28404              0305
================================================================================
STORAGE GROUP MINIDISK TOTALS

 Storage          4K Blocks         4K Blocks
 Group No.          In-Use            Free
     1            145 -    2%         8838
```

More about SFS

- z/VM guest management problems can be caused by the SFS not being available
  - See if you can access SFS files (VMSYS:VSMWORK1 files) from MAINT
    - Issue: *acc vmsys:vsmwork1.data w*
    - Below is the output of the vsmwork1.data w file list



  - Ensure VSMPROXY is not in CP READ state
    - If so, IPL CMS and make sure it's in RUNNING state before disconnecting
  - Check if TCP/IP communication to the HMC and SE is still active
    - Try access the Ensemble HMC from the Web
    - Check for any other networking issues

z/VM Virtual Switches

- **INMN**
  - Shown below is the output of the *QUERY VSWITCH SW2 ACCESS* command from maint

```
q vswitch sw2 acc
VSWITCH SYSTEM SW2       Type: INMN     Connected: 1     Maxconn: INFINITE
  PERSISTENT  RESTRICTED     ETHERNET                    Accounting: OFF
  VLAN Unaware
  MAC address: 02-D2-DB-00-00-03     MAC Protection: Unspecified
  State: Ready
  IPTimeout: 5           QueueStorage: 8
  Isolation Status: ON
    Authorized userids:
      SYSTEM    ZVMLXAPP
 Uplink Port:
  NIC: ZVMLXAPP  VDEV: 0200
Ready; T=0.01/0.01 12:07:20
```

- **IEDN**
  - Virtual IEDN switches need to be defined when z/VM guests don't support OSX CHPIDs
  - Shown below is the output of the *QUERY VSWITCH QDIO100 ACCESS* command from maint
  - QDIO1000 switch has been added to provide IEDN OSX connectivity to Virtual Servers (z/VM Guests)

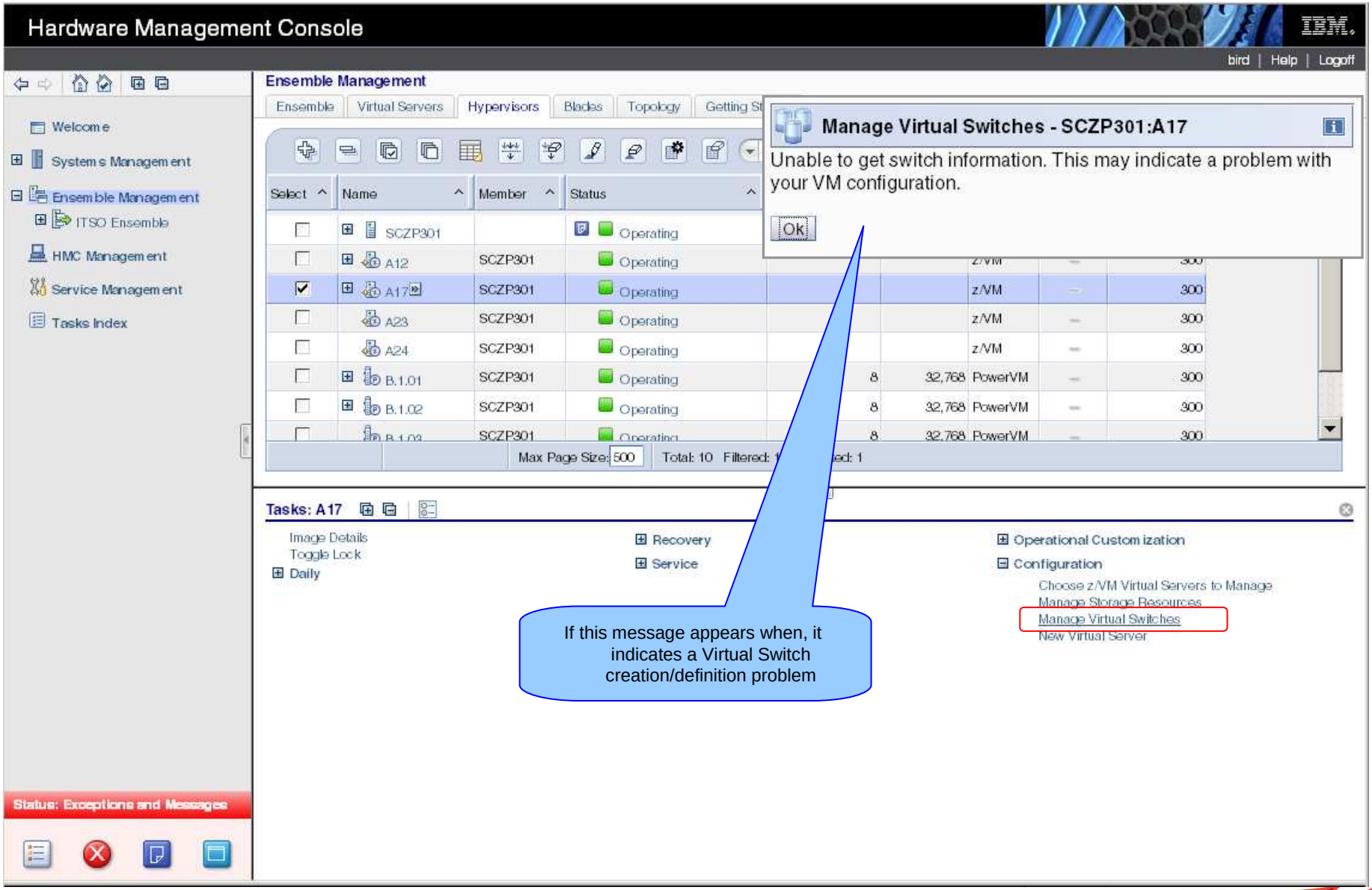
```
VSWITCH SYSTEM QDIO100   Type: IEDN     Connected: 2     Maxconn: INFINITE
  PERSISTENT  RESTRICTED     ETHERNET                    Accounting: OFF
  VLAN Aware  Default VLAN: NONE     Default Porttype: Access  GVRP: Enabled
              Native  VLAN: NONE     VLAN Counters: OFF
  MAC address: 02-D2-DB-00-00-07     MAC Protection: OFF
  State: Ready
  IPTimeout: 5           QueueStorage: 8
  Isolation Status: OFF
    Authorized userids:
      SYSTEM    Porttype: Access VLAN: NONE
      WASG1     Porttype: Access VLAN: 0100
      WASS1     Porttype: Access VLAN: 0100
 Uplink Port:
  RDEV: 2303.P00 VDEV: 2303 Controller: DTCENS2
  RDEV: 2323.P00 VDEV: 2323 Controller: DTCENS1  BACKUP
Ready; T=0.01/0.01 13:56:28
```

More about z/VM Virtual Switches

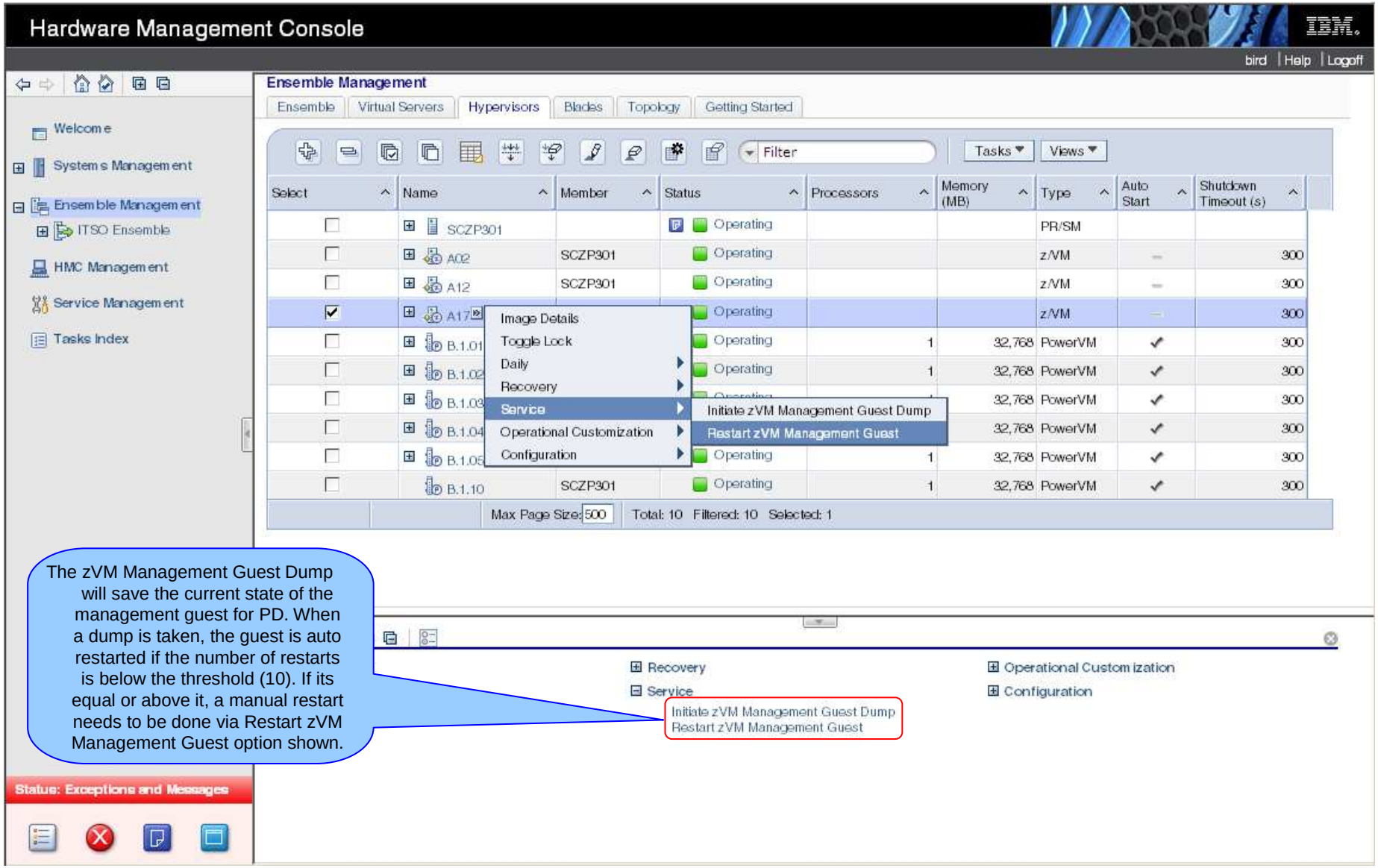
- VSWITCH Definitions
  - Persistent data for QDIO (OSD) VSWITCH created by the Unified Resource Manager is saved in **RESTART NOTEBOOK** file on the VMSYS:VSMWORK1.DATA SFS
  - IEDN VSWITCH, (no OSX Guest support) persistent data is stored in **ENCONFIG NOTEBOOK** file on the VMSYS:VSMWORK1.DATA SFS.
    - Use acc vmsys:vsmwork1.data w to check these files

```
ENCONFIG NOTEBOOK W1   V 133  Trunc=133 Size=86 Line=0 Col=1 Alt=0
00000 * * * Top of File * * *
00001 DEFINE VSWITCH QDIO099 TYPE IEDN RDEV NONE CONNECT QUEUESTORAGE 8 CONTRO
00002 SET VSWITCH QDIO099 GRANT SYSTEM
00003 SET VSWITCH QDIO099 MACPROTECT ON
00004 SET VSWITCH QDIO099 QUEUESTORAGE 8
00005 SET VSWITCH QDIO099 IPTIMEOUT 5
00006 SET VSWITCH QDIO099 CONTROLLER *
00007 SET VSWITCH QDIO099 RDEV 2300
00008 SET VSWITCH QDIO099 RDEV 2300.P00 2320
00009 DEFINE VSWITCH QDIO100 TYPE IEDN RDEV NONE CONNECT QUEUESTORAGE 8 CONTRO
00010 SET VSWITCH QDIO100 GRANT SYSTEM
00011 SET VSWITCH QDIO100 MACPROTECT ON
00012 SET VSWITCH QDIO100 QUEUESTORAGE 8
00013 SET VSWITCH QDIO100 IPTIMEOUT 5
00014 SET VSWITCH QDIO100 CONTROLLER *
00015 SET VSWITCH QDIO100 RDEV 2303
00016 SET VSWITCH QDIO100 RDEV 2303.P00 2323
00017 DEFINE VSWITCH QDIO110 TYPE IEDN RDEV NONE CONNECT QUEUESTORAGE 8 CONTRO
00018 SET VSWITCH QDIO110 GRANT SYSTEM
00019 SET VSWITCH QDIO110 MACPROTECT ON
00020 SET VSWITCH QDIO110 QUEUESTORAGE 8
00021 SET VSWITCH QDIO110 IPTIMEOUT 5
00022 SET VSWITCH QDIO110 CONTROLLER *
00023 SET VSWITCH QDIO110 RDEV 2306
00024 SET VSWITCH QDIO110 RDEV 2306.P00 2326
00025 SET VSWITCH QDIO099 GRANT WASG1 PORTTYPE TRUNK
00026 SET VSWITCH QDIO099 GRANT WASG1 OSDSIM ON
```

Manage Virtual Switches from Unified Resource Manager



z/VM Management Guest



Part IV – Ensemble Managing and Problem Determination

3 z/OS in an Ensemble

- *z/OS Ensemble Participation*
- *OSM & OSX*
- *IOCP definitions*
- *z/OS OSM PD*
- *z/OS OSX PD*
- *z/OS GPMP*
- *z/OS CommServer definitions*
- *IPL'ing z/OS*

z/OS Ensemble participation

- z/OS participation in an Ensemble requires:
  - Enabling z/OS for IPv6
    - Needed to participate in the INMN
    - Enable IPv6 for TCPIP in BPXPRMxx member
  - VTAM definitions
    - Ensemble = YES in VTAM member ATSCTR00
  - IEDN definitions
  - RACF definitions
    - Check URM Redbook SG24-7921 for above definitions

Example BPXPRMxx entry for enabling IPv6

```
FILESYSTYPE TYPE(INET)ENTRYPOINT(EZBPFINI)
NETWORK DOMAINNAME(AF_INET)
DOMAINNNUMBER(2)
MAXSOCKETS(2000)
TYPE(INET)
NETWORK DOMAINNAME(AF_INET6)
DOMMAINNUMBER(19)
MAXSOCKETS(3000)
TYPE(INET)
```

- Check if z/OS is a “member” of the Ensemble, issuing:

```
D NET,VTAMOPTS,FORMAT=COMPLETE
 IST097I DISPLAY ACCEPTED
 IST1188I VTAM CSV1R12 STARTED AT 21:22:59 ON 07/19/10
 IST1349I COMPONENT ID IS 5695-11701-1C0
 IST1348I VTAM STARTED AS NETWORK NODE
 IST1309I START OPTION  CURRENT VALUE     ORIGINAL VALUE    ORIGIN

 IST1310I ENSEMBLE      YES               YES               ATCSTREN
```

This is an indication that z/OS is participating in an Ensemble environment

```
D NET,VTAMOPTS,FORMAT=COMPLETE
 IST097I DISPLAY ACCEPTED
 IST1188I VTAM CSV1R12 STARTED AT 21:22:59 ON 07/19/10
 IST1349I COMPONENT ID IS 5695-11701-1C0
 IST1348I VTAM STARTED AS NETWORK NODE
 IST1309I START OPTION  CURRENT VALUE     ORIGINAL VALUE    ORIGIN

 IST1310I ENSEMBLE      YES               ***NA***          DEFAULT
```

This line indicates that z/OS is not enable to participate in the Ensemble

OSM CHPID Definitions for INMN

IOCP Definitions

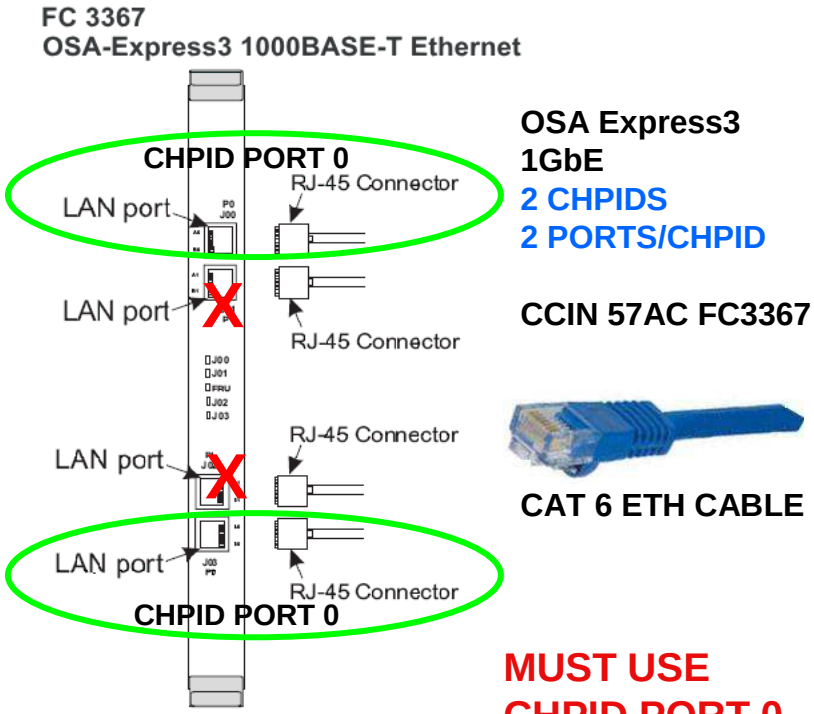
CHPID PATH=(CSS(1),0A),SHARED,
PARTITION=((A11,A13,A16,A17,A18),(=))),
CHPARM=02,PCHID=531,**TYPE=OSM**

CNTLUNIT CUNUMBR=2340,
PATH=((CSS(1),0A),**UNIT=OSM**

IODEVICE
ADDRESS=(2340,016),MODEL=M,UNITADD=00,
CUNUMBR=(2340),
UNIT=OSA

z/OS Definitions

OSM Connections are dynamically created when an Ensemble is created



Sample IOCP Definitions for OSM (INMN)

```
ID    MSG1='ITSO',MSG2='SYS6.IODF57 - 2011-08-01 10:47',              *
            SYSTEM=(2817,1),LSYSTEM=SCZP301,                          *
            TOK=('SCZP301',00800006991E2094104707650111213F00000000,*
            00000000,'11-08-01','10:47:07','SYS6','IODF57')
      RESOURCE PARTITION=((CSS(0),(A0A,A),(A0B,B),(A0C,C),(A0D,D),(A*
            0E,E),(A0F,F),(A01,1),(A02,2),(A03,3),(A04,4),(A05,5),(A*
            06,6),(A07,7),(A08,8),(A09,9)),(CSS(1),(A1B,B),(A1E,E),(*
            A1F,F),(A11,1),(A12,2),(A13,3),(A14,4),(A15,5),(A16,6),(*
            A17,7),(A18,8),(*,9),(*,A),(*,C),(*,D)),(CSS(2),(A2E,E),*
            (A2F,F),(A21,1),(A22,2),(A23,3),(A24,4),(A25,5),(A28,8),*
            (*,6),(*,7),(*,9),(*,A),(*,B),(*,C),(*,D)),(CSS(3),(A3E,*
            E),(A3F,F),(A31,1),(A34,4),(A35,5),(Y32,2),(*,3),(*,6),(*
            *,7),(*,8),(*,9),(*,A),(*,B),(*,C),(*,D)))
      CHPID PATH=(CSS(0,1,2),0A),SHARED,                              *
            PARTITION=((CSS(0),(A02,A07),(=)),(CSS(2),(A25),(=))),    *
            NOTPART=((CSS(1),(A14,A15,A1B,A1E,A1F),(=))),CHPARM=02,   *
            PCHID=531,TYPE=OSM
      CHPID PATH=(CSS(0,1,2),0B),SHARED,                              *
            PARTITION=((CSS(0),(A02,A07),(=)),(CSS(2),(A25),(=))),    *
            NOTPART=((CSS(1),(A14,A15,A1B,A1E,A1F),(=))),CHPARM=02,   *
            PCHID=101,TYPE=OSM
      CNTLUNIT CUNUMBR=2340,                                          *
            PATH=((CSS(0),0A),(CSS(1),0A),(CSS(2),0A)),UNIT=OSM
      IODEVICE ADDRESS=(2340,016),MODEL=M,UNITADD=00,CUNUMBR=(2340),*
            UNIT=OSA
      CNTLUNIT CUNUMBR=2360,                                          *
            PATH=((CSS(0),0B),(CSS(1),0B),(CSS(2),0B)),UNIT=OSM
      IODEVICE ADDRESS=(2360,016),MODEL=M,UNITADD=00,CUNUMBR=(2360),*
            UNIT=OSA
```

OSX CHPID Definitions for IEDN

IOCP Definitions

CHPID PATH=(CSS(,1),18),SHARED,
PARTITION=((A11,A13,A16,A17,A18),(=)),
PCHID=590,**TYPE=OSX**

CNTLUNIT CUNUMBR=2300,
PATH=((CSS(1),18)),**UNIT=OSX**

IODEVICE ADDRESS=(2300,016),MODEL=X,UNITADD=00,
CUNUMBR=(2300), **UNIT=OSA**

z/OS Definitions

VTAM Definitions

Dynamic or Manually defined TRLEs

TCP/IP Definitions

INTERFACE IPAQENET

INTERFACE IPAQENET6

OSX (IEDN)

OSA Express3
10 GbE
2 CHPIDS
1 PORT/CHPID

CCIN 57A3 FC3370 (LR)
Single Mode 9 micron LC duplex

CCIN 57AD FC3371 (SR)
Multi Mode 50/62.5 micron LC duplex

Supports IOCP CHPID types:
OSD and OSX (ONLY 10 GbE).
PCHID = xxx0 & xxx1

Sample IOCP Definitions for OSX (IEDN)

```
ID    MSG1='ITSO',MSG2='SYS6.IODF57 - 2011-08-01 10:47',              *
            SYSTEM=(2817,1),LSYSTEM=SCZP301,                          *
            TOK=('SCZP301',00800006991E2094104707650111213F00000000,*
            00000000,'11-08-01','10:47:07','SYS6','IODF57')
      RESOURCE PARTITION=((CSS(0),(A0A,A),(A0B,B),(A0C,C),(A0D,D),(A*
            0E,E),(A0F,F),(A01,1),(A02,2),(A03,3),(A04,4),(A05,5),(A*
            06,6),(A07,7),(A08,8),(A09,9)),(CSS(1),(A1B,B),(A1E,E),(*
            A1F,F),(A11,1),(A12,2),(A13,3),(A14,4),(A15,5),(A16,6),(*
            A17,7),(A18,8),(*,9),(*,A),(*,C),(*,D)),(CSS(2),(A2E,E),*
            (A2F,F),(A21,1),(A22,2),(A23,3),(A24,4),(A25,5),(A28,8),*
            (*,6),(*,7),(*,9),(*,A),(*,B),(*,C),(*,D)),(CSS(3),(A3E,*
            E),(A3F,F),(A31,1),(A34,4),(A35,5),(Y32,2),(*,3),(*,6),(*
            *,7),(*,8),(*,9),(*,A),(*,B),(*,C),(*,D)))
      CHPID PATH=(CSS(0,1,2),18),SHARED,                              *
            PARTITION=((CSS(0),(A02,A06,A07),(=)),(CSS(2),(A25),(=))*
            ),NOTPART=((CSS(1),(A14,A15,A1B,A1E,A1F),(=))),           *
            PCHID=590,TYPE=OSX
      CHPID PATH=(CSS(0,1,2),19),SHARED,                              *
            PARTITION=((CSS(0),(A02,A06,A07),(=)),(CSS(2),(A25),(=))*
            ),NOTPART=((CSS(1),(A14,A15,A1B,A1E,A1F),(=))),           *
            PCHID=510,TYPE=OSX
      CNTLUNIT CUNUMBR=2300,                                          *
            PATH=((CSS(0),18),(CSS(1),18),(CSS(2),18)),UNIT=OSX
      IODEVICE ADDRESS=(2300,032),MODEL=X,UNITADD=00,CUNUMBR=(2300),*
            UNIT=OSA
      CNTLUNIT CUNUMBR=2320,                                          *
            PATH=((CSS(0),19),(CSS(1),19),(CSS(2),19)),UNIT=OSX
      IODEVICE ADDRESS=(2320,032),MODEL=X,UNITADD=00,CUNUMBR=(2320),*
            UNIT=OSA
```

z/OS Ensemble OSA CHPIDs display outputs

- INMN
  - D M=CHP(0A)

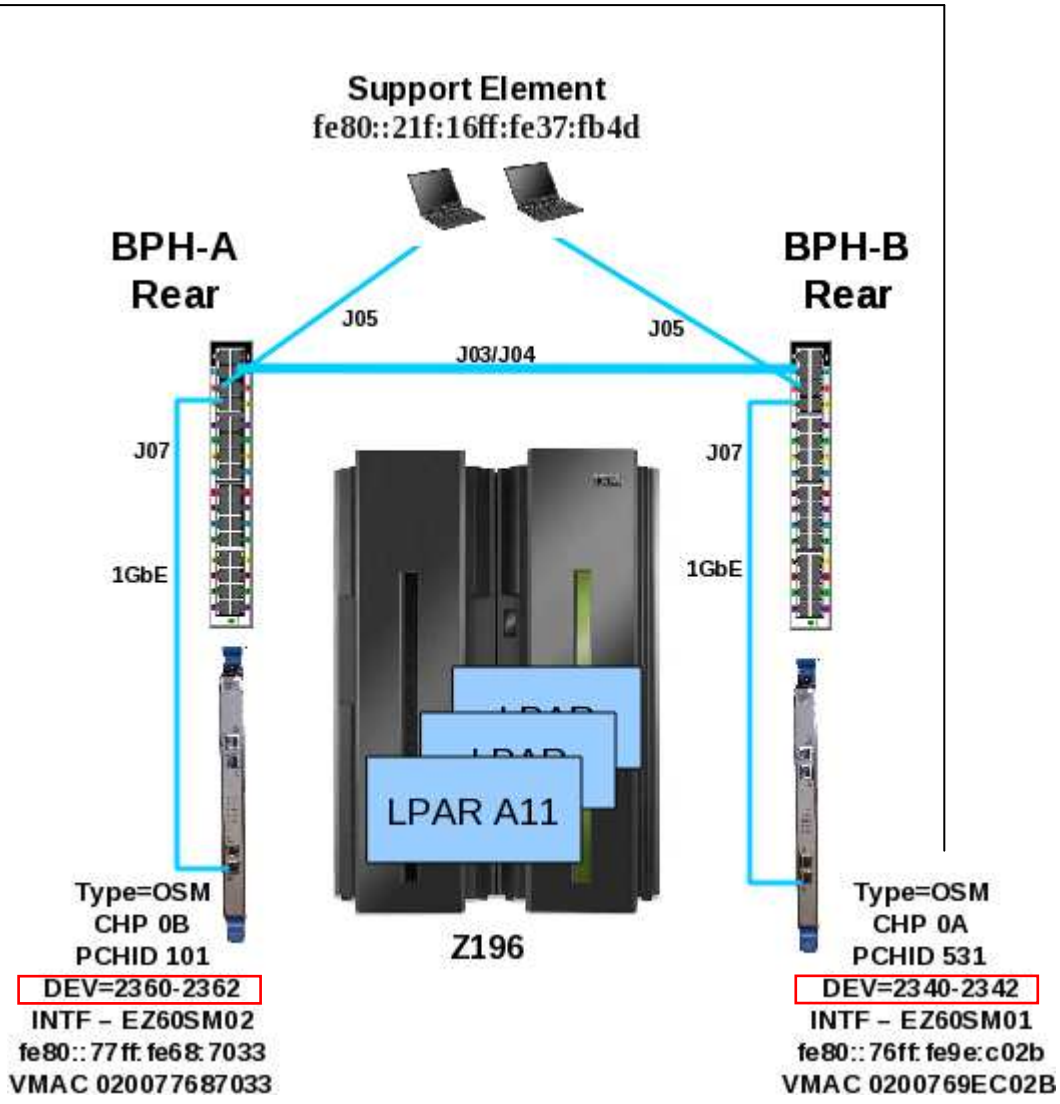
```
D M=CHP(0A)
IEE174I 16.53.26 DISPLAY M 702
CHPID 0A:  TYPE=31, DESC=OSA ZBX MANAGEMENT, ONLINE
DEVICE STATUS FOR CHANNEL PATH 0A
     0  1  2  3  4  5  6  7  8  9  A  B  C  D  E  F
0234 +  +  +  +  +  +  +  +  +  +  +  +  +  +  +  +
SWITCH DEVICE NUMBER = NONE
PHYSICAL CHANNEL ID = 0531
************************ SYMBOL EXPLANATIONS ************************
+ ONLINE    @ PATH NOT VALIDATED   - OFFLINE    . DOES NOT EXIST
* PHYSICALLY ONLINE   $ PATH NOT OPERATIONAL
```

- IEDN
  - D M=CHP(18)

```
D M=CHP(18)
IEE174I 17.06.28 DISPLAY M 709
CHPID 18:  TYPE=30, DESC=OSA ZBX DATA, ONLINE
DEVICE STATUS FOR CHANNEL PATH 18
     0  1  2  3  4  5  6  7  8  9  A  B  C  D  E  F
0230 +  +  +  +  +  +  +  +  +  +  +  +  +  +  +  +
0231 +  +  +  +  +  +  +  +  +  +  +  +  +  +  +  +
SWITCH DEVICE NUMBER = NONE
PHYSICAL CHANNEL ID = 0590
************************ SYMBOL EXPLANATIONS ************************
+ ONLINE    @ PATH NOT VALIDATED   - OFFLINE    . DOES NOT EXIST
* PHYSICALLY ONLINE   $ PATH NOT OPERATIONAL
```

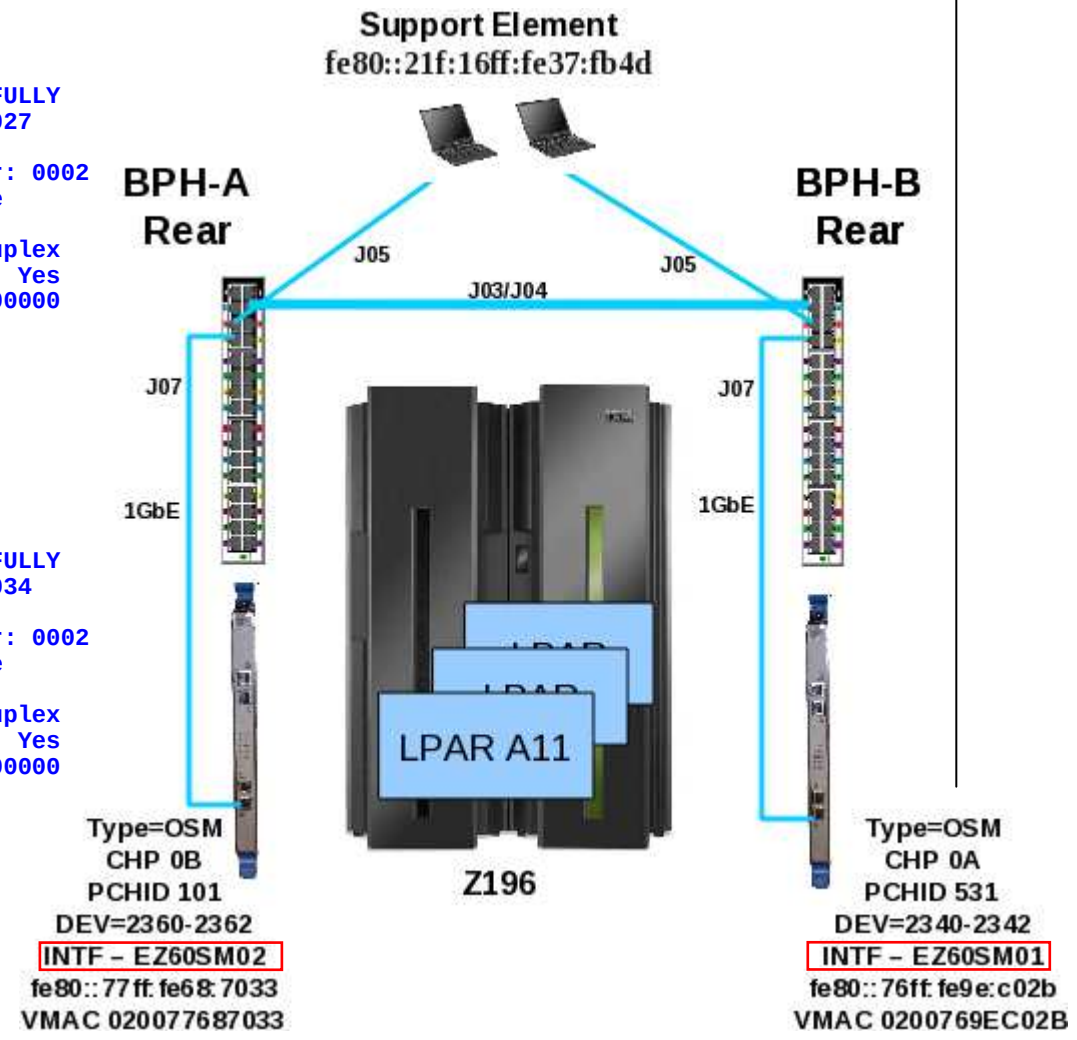
z/OS **OSM** PD Commands

```
D M=DEV(2340)
IEE174I 09.43.58 DISPLAY M 000
DEVICE 2340   STATUS=ONLINE
CHP                  0A
ENTRY LINK ADDRESS   ..
DEST LINK ADDRESS    0D
PATH ONLINE          Y
CHP PHYSICALLY ONLINE Y
PATH OPERATIONAL     Y
MANAGED              N
CU NUMBER            2340
MAXIMUM MANAGED CHPID(S) ALLOWED:  0
DESTINATION CU LOGICAL ADDRESS = 00
SCP CU ND         = 001730.005.IBM.02.2817000B3BD5.0A00
SCP TOKEN NED     = 001730.005.IBM.02.2817000B3BD5.0A00
SCP DEVICE NED    = 001732.001.IBM.02.2817000B3BD5.0A00

D M=DEV(2360)
IEE174I 10.02.46 DISPLAY M 014
DEVICE 2360   STATUS=ONLINE
CHP                  0B
ENTRY LINK ADDRESS   ..
DEST LINK ADDRESS    0D
PATH ONLINE          Y
CHP PHYSICALLY ONLINE Y
PATH OPERATIONAL     Y
MANAGED              N
CU NUMBER            2360
MAXIMUM MANAGED CHPID(S) ALLOWED:  0
DESTINATION CU LOGICAL ADDRESS = 00
SCP CU ND         = 001730.005.IBM.02.2817000B3BD5.0B00
SCP TOKEN NED     = 001730.005.IBM.02.2817000B3BD5.0B00
SCP DEVICE NED    = 001732.001.IBM.02.2817000B3BD5.0B00
```



z/OS **OSM** TCPIP PD Commands

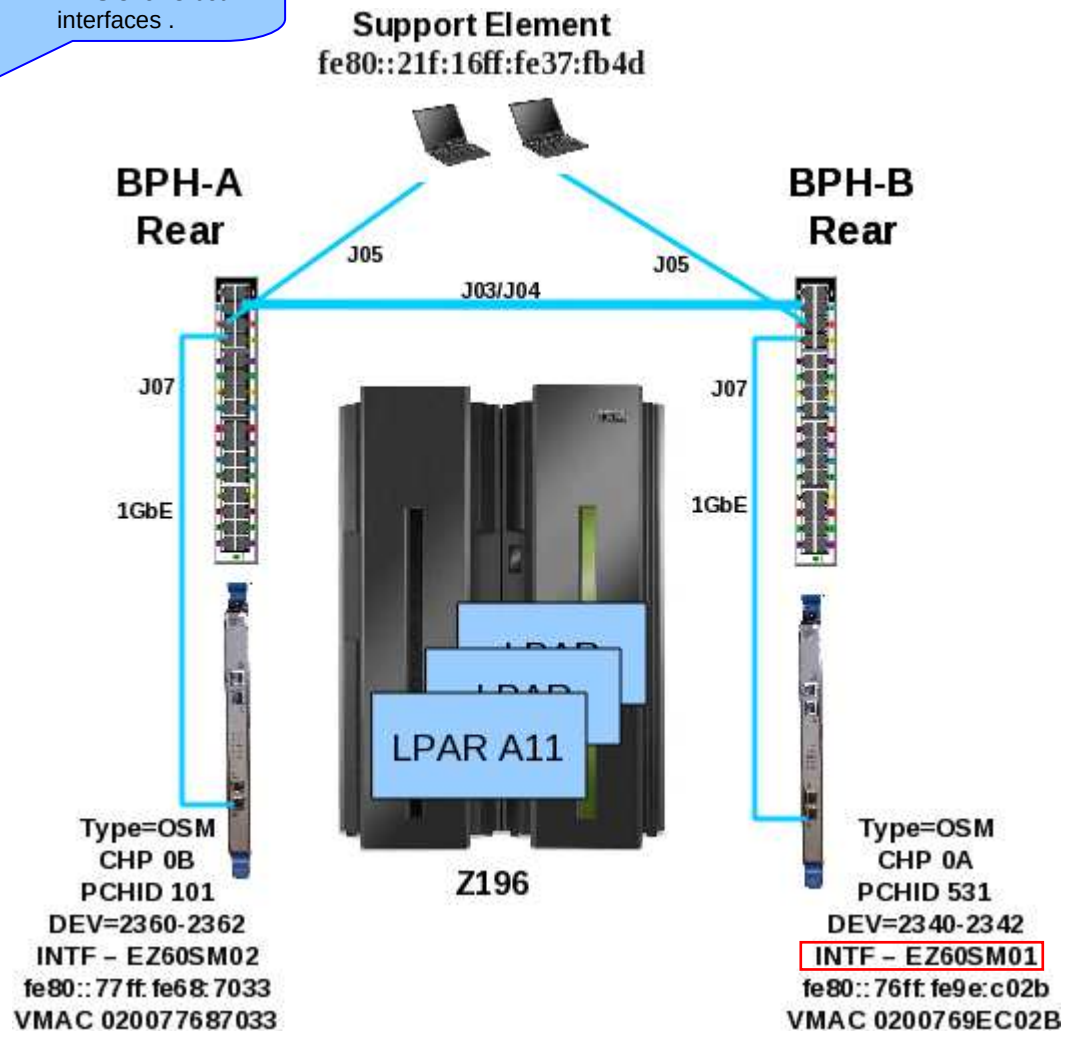
```
D TCPIP,TCPIP,OSAINFO,INTFNAME=EZ6OSM01
EZZ0053I COMMAND DISPLAY TCPIP,,OSAINFO COMPLETED SUCCESSFULLY
EZD0031I TCP/IP CS V1R13  TCPIP Name: TCPIP     10:18:08 027
Display OSAINFO results for IntfName: EZ6OSM01
PortName: IUTMP00A  PortNum: 00  Datapath: 2342   RealAddr: 0002
PCHID: 0531         CHPID: 0A    CHPID Type: OSM  OSA code
level: 0909
Gen: OSA-E3         Active speed/mode: 1000 mb/sec full duplex
Media: Copper                  Jumbo frames: Yes  Isolate: Yes
PhysicalMACAddr: 00145E7769EC  LocallyCfgMACAddr: 000000000000
Queues defined Out: 1  In: 1   Ancillary queues in use: 0
Connection Mode: Layer 2
SAPSup: 0009F603               SAPEna: 0008A603
Layer 2 attributes:
  VLAN ID: N/A             VMAC Active: Yes
  VMAC Addr: 0200769EC02B  VMAC Origin: OSA

D TCPIP,TCPIP,OSAINFO,INTFNAME=EZ6OSM02
EZZ0053I COMMAND DISPLAY TCPIP,,OSAINFO COMPLETED SUCCESSFULLY
EZD0031I TCP/IP CS V1R13  TCPIP Name: TCPIP     10:21:31 034
Display OSAINFO results for IntfName: EZ6OSM02
PortName: IUTMP00B  PortNum: 00  Datapath: 2362   RealAddr: 0002
PCHID: 0101         CHPID: 0B    CHPID Type: OSM  OSA code
level: 0909
Gen: OSA-E3         Active speed/mode: 1000 mb/sec full duplex
Media: Copper                  Jumbo frames: Yes  Isolate: Yes
PhysicalMACAddr: 00145E776878  LocallyCfgMACAddr: 000000000000
Queues defined Out: 1  In: 1   Ancillary queues in use: 0
Connection Mode: Layer 2
SAPSup: 0009F603               SAPEna: 0008A603
Layer 2 attributes:
  VLAN ID: N/A             VMAC Active: Yes
  VMAC Addr: 020077687033  VMAC Origin: OSA
```



z/OS **OSM** TCPIP PD Commands

```
D TCPIP,TCPIP,NETSTAT,DEVLINKS
EZD0101I NETSTAT CS V1R13 TCPIP 003
DEVNAME: LOOPBACK DEVTYPE: LOOPBACK
DEVSTATUS: READY
LNKNAME: LOOPBACK LNKTYPE: LOOPBACK LNKSTATUS: READY
ACTMTU: 65535
ROUTING PARAMETERS:
.
.
.
INTFNAME: EZ6OSM01 INTFTYPE: IPAQENET6 INTFSTATUS: READY
PORTNAME: IUTMP00A DATAPATH: 2342 DATAPATHSTATUS: READY
CHPIDTYPE: OSM
QUESIZE: 0 SPEED: 0000001000
VMACADDR: 0200769EC02B VMACORIGIN: OSA VMACROUTER: ALL
DUPADDRDET: 1
CFGMTU: NONE ACTMTU: 1500
VLANID: NONE VLANPRIORITY: DISABLED
READSTORAGE: GLOBAL (4096K)
INBPERF: DYNAMIC
WORKLOADQUEUEING: NO
CHECKSUMOFFLOAD: NO SEGMENTATIONOFFLOAD: NO
SECCLASS: 255 MONSYSPLEX: NO
ISOLATE: YES OPTLATENCYMODE: NO
TEMPPREFIX: NONE
INPUT ===>
MULTICAST SPECIFIC:
MULTICAST CAPABILITY: YES
GROUP: FF02::1:FF9E:C02B
REFCNT: 0000000001 SRCFLTMD: EXCLUDE
SRCADDR: NONE
GROUP: FF01::1
REFCNT: 0000000001 SRCFLTMD: EXCLUDE
SRCADDR: NONE
GROUP: FF02::1
REFCNT: 0000000001 SRCFLTMD: EXCLUDE
SRCADDR: NONE
INTERFACE STATISTICS:
BYTESIN = 5994
INBOUND PACKETS = 10
INBOUND PACKETS IN ERROR = 43
INBOUND PACKETS DISCARDED = 0
INBOUND PACKETS WITH NO PROTOCOL = 0
BYTESOUT = 534
OUTBOUND PACKETS = 5
OUTBOUND PACKETS IN ERROR = 0
OUTBOUND PACKETS DISCARDED = 0
```

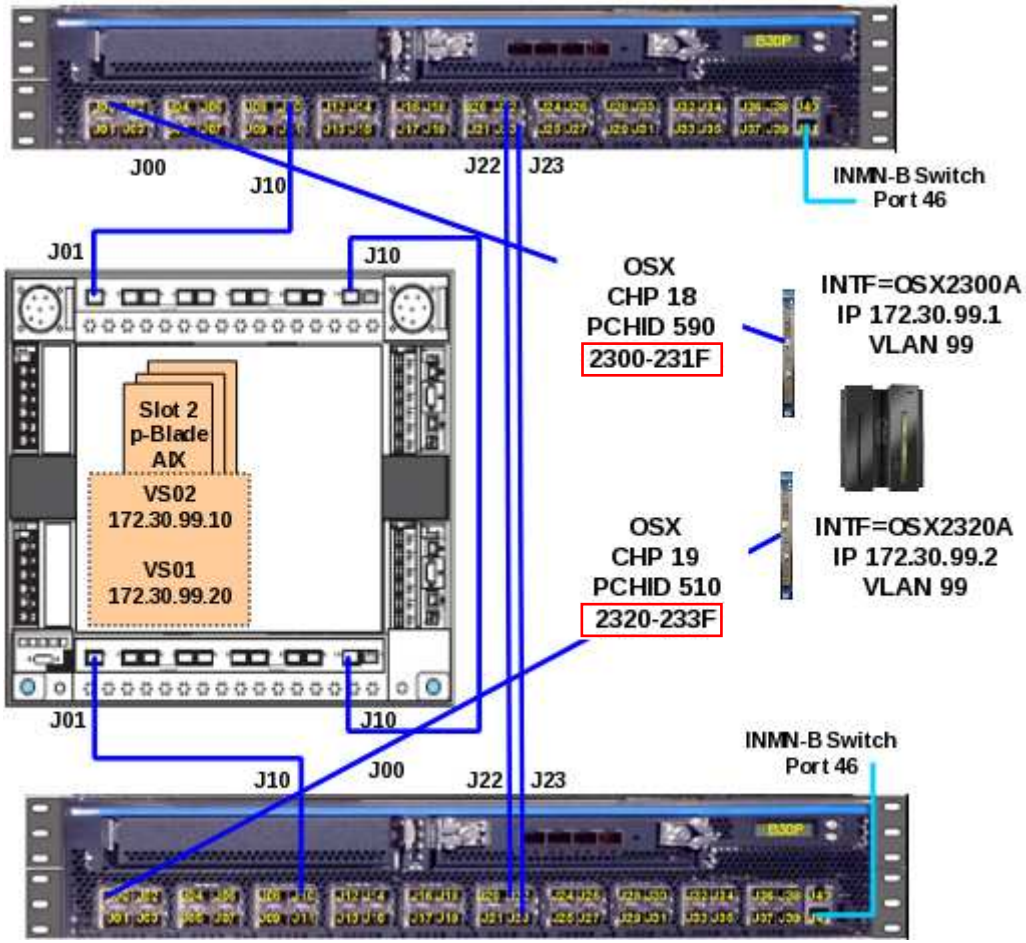
Only the output for EZ60SM01 is shown here. The DEVLINKS shows both interfaces .



z/OS **OSX** PD Commands

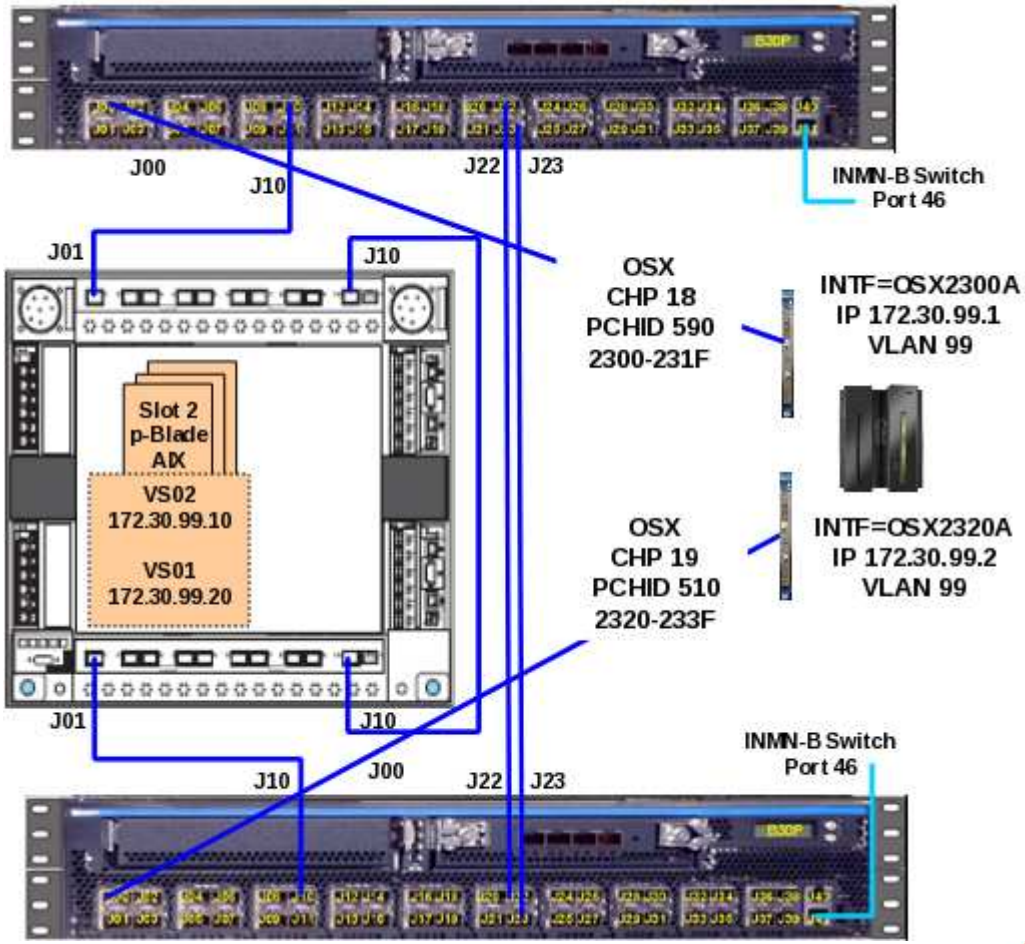
```
D M=DEV(2300)
IEE174I 20.31.58 DISPLAY M 317
DEVICE 2300   STATUS=ONLINE
CHP                  18
ENTRY LINK ADDRESS   ..
DEST LINK ADDRESS    0D
PATH ONLINE          Y
CHP PHYSICALLY ONLINE Y
PATH OPERATIONAL     Y
MANAGED              N
CU NUMBER            2300
MAXIMUM MANAGED CHPID(S) ALLOWED:  0
DESTINATION CU LOGICAL ADDRESS = 00
SCP CU ND         = 001730.005.IBM.02.2817000B3BD5.1800
SCP TOKEN NED     = 001730.005.IBM.02.2817000B3BD5.1800
SCP DEVICE NED    = 001732.001.IBM.02.2817000B3BD5.1800
```

```
D M=DEV(2320)
IEE174I 20.32.27 DISPLAY M 319
DEVICE 2320   STATUS=ONLINE
CHP                  19
ENTRY LINK ADDRESS   ..
DEST LINK ADDRESS    0D
PATH ONLINE          Y
CHP PHYSICALLY ONLINE Y
PATH OPERATIONAL     Y
MANAGED              N
CU NUMBER            2320
MAXIMUM MANAGED CHPID(S) ALLOWED:  0
DESTINATION CU LOGICAL ADDRESS = 00
SCP CU ND         = 001730.005.IBM.02.2817000B3BD5.1900
SCP TOKEN NED     = 001730.005.IBM.02.2817000B3BD5.1900
SCP DEVICE NED    = 001732.001.IBM.02.2817000B3BD5.1900
```



z/OS **OSX** TCPIP PD Commands

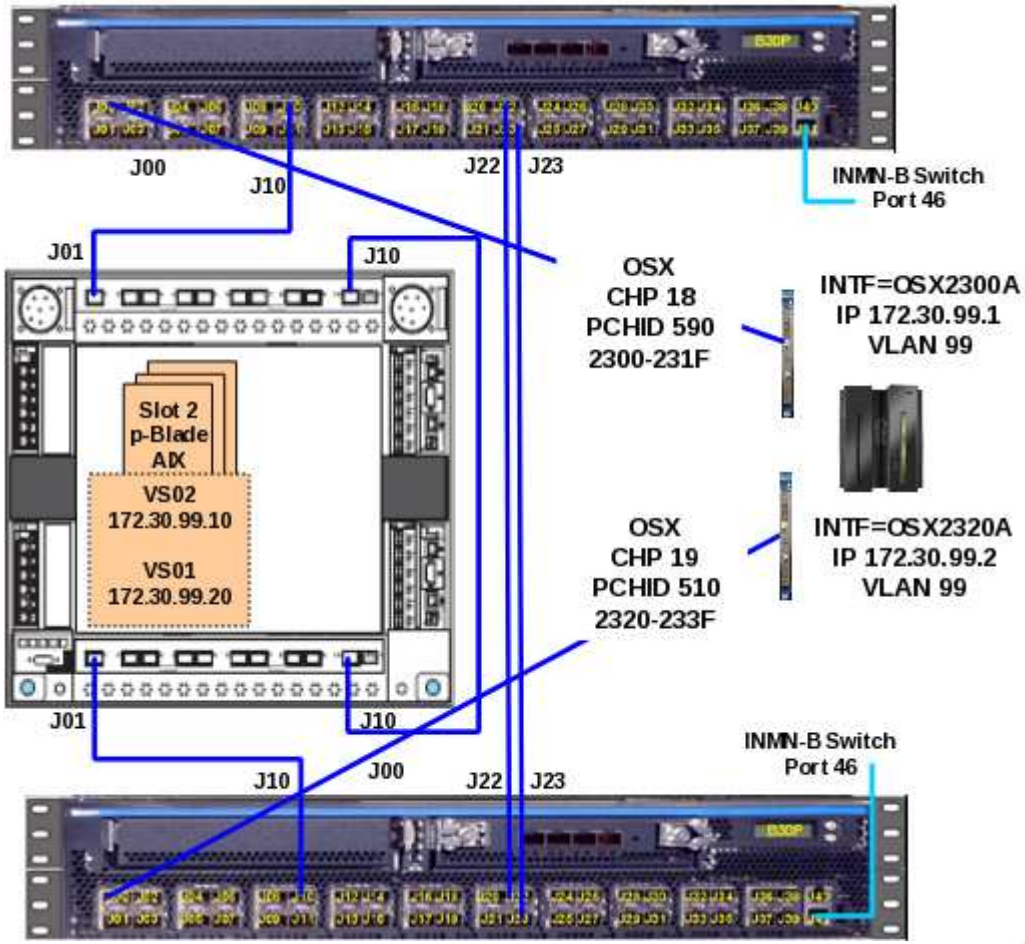
```
D TCPIP,TCPIP,OSAINFO,INTFNAME=OSX2300A
EZZ0053I COMMAND DISPLAY TCPIP,,OSAINFO COMPLETED SUCCESSFULLY
EZD0031I TCP/IP CS V1R13  TCPIP Name: TCPIP     20:45:53 326
Display OSAINFO results for IntfName: OSX2300A
PortName: IUTXP018  PortNum: 00  Datapath: 2302   RealAddr: 0002
PCHID: 0590         CHPID: 18    CHPID Type: OSX  OSA code level:
0D0C
Gen: OSA-E3         Active speed/mode: 10 gigabit full duplex
Media: Multimode Fiber          Jumbo frames: Yes  Isolate: No
PhysicalMACAddr: 001A643B2135  LocallyCfgMACAddr: 000000000000
Queues defined Out: 4  In: 1   Ancillary queues in use: 0
Connection Mode: Layer 3       IPv4: Yes  IPv6: No
SAPSup: 000FF603               SAPEna: 0008A603
IPv4 attributes:
  VLAN ID:   99            VMAC Active: Yes
  VMAC Addr: 0207E300001B  VMAC Origin: OSA       VMAC Router: All
  AsstParmsEna: 00200C57   OutCkSumEna: 0000001A  InCkSumEna:
0000001A
Registered Addresses:
  IPv4 Unicast Addresses:
    ARP: Yes  Addr: 172.30.99.1
    Total number of IPv4 addresses:      1
  IPv4 Multicast Addresses:
    MAC: 01005E000001  Addr: 224.0.0.1
    Total number of IPv4 addresses:      1

D TCPIP,TCPIP,OSAINFO,INTFNAME=OSX2320A
EZZ0053I COMMAND DISPLAY TCPIP,,OSAINFO COMPLETED SUCCESSFULLY
EZD0031I TCP/IP CS V1R13  TCPIP Name: TCPIP     21:24:05 352
Display OSAINFO results for IntfName: OSX2320A
PortName: IUTXP019  PortNum: 00  Datapath: 2322   RealAddr: 0002
PCHID: 0510         CHPID: 19    CHPID Type: OSX  OSA code level:
0D0C
Gen: OSA-E3         Active speed/mode: 10 gigabit full duplex
Media: Multimode Fiber          Jumbo frames: Yes  Isolate: No
PhysicalMACAddr: 001A643B2020  LocallyCfgMACAddr: 000000000000
Queues defined Out: 1  In: 1   Ancillary queues in use: 0
Connection Mode: Layer 3       IPv4: Yes  IPv6: No
SAPSup: 000FF603               SAPEna: 0008A603
IPv4 attributes:
  VLAN ID:   99            VMAC Active: Yes
  VMAC Addr: 021502000017  VMAC Origin: OSA       VMAC Router: All
  AsstParmsEna: 00200C57   OutCkSumEna: 0000001A  InCkSumEna:
0000001A
Registered Addresses:
  IPv4 Unicast Addresses:
    ARP: Yes  Addr: 172.30.99.2
    Total number of IPv4 addresses:      1
  IPv4 Multicast Addresses:
    MAC: 01005E000001  Addr: 224.0.0.1
    Total number of IPv4 addresses:      1
```



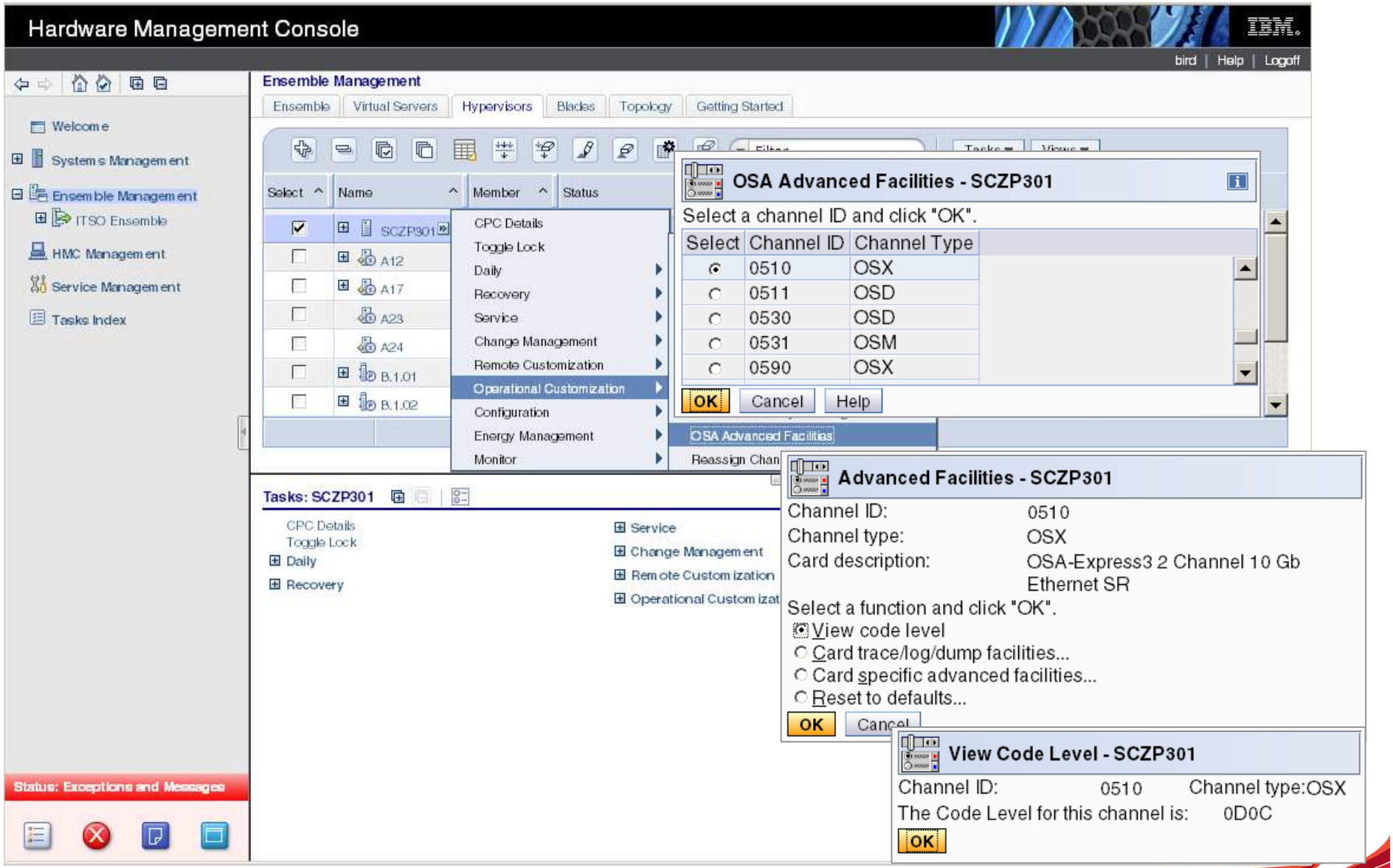
z/OS **OSX** TCPIP PD Commands

```
D TCPIP,TCPIP,NETSTAT,DEVLINKS
EZD0101I NETSTAT CS V1R13 TCPIP 003
DEVNAME: LOOPBACK DEVTYPE: LOOPBACK
DEVSTATUS: READY
LNKNAME: LOOPBACK LNKTYPE: LOOPBACK LNKSTATUS: READY
ACTMTU: 65535
ROUTING PARAMETERS:
.
.
.
INTFNAME: OSX2300A INTFTYPE: IPAQENET INTFSTATUS: READY
PORTNAME: IUTXP018 DATAPATH: 2302 DATAPATHSTATUS: READY
CHPIDTYPE: OSX CHPID: 18
SPEED: 0000010000
IPBROADCASTCAPABILITY: NO
VMACADDR: 0207E300001B VMACORIGIN: OSA VMACROUTER: ALL
ARPOFFLOAD: YES ARPOFFLOADINFO: YES
CFGMTU: 8992 ACTMTU: 8992
IPADDR: 172.30.99.1/24
VLANID: 99 VLANPRIORITY: DISABLED
DYNVLANREGCFG: NO DYNVLANREGCAP: YES
READSTORAGE: GLOBAL (4096K)
INBPERF: DYNAMIC
WORKLOADQUEUEING: NO
CHECKSUMOFFLOAD: YES SEGMENTATIONOFFLOAD: NO
SECCLASS: 255 MONSYSPLEX: NO
ISOLATE: NO OPTLATENCYMODE: NO
MULTICAST SPECIFIC:
MULTICAST CAPABILITY: YES
GROUP REFCNT SRCFLTMD
224.0.0.1
0000000001 EXCLUDE
SRCADDR: NONE
INTERFACE STATISTICS:
BYTESIN = 0
INBOUND PACKETS = 0
INBOUND PACKETS IN ERROR = 0
INBOUND PACKETS DISCARDED = 0
INBOUND PACKETS WITH NO PROTOCOL = 0
BYTESOUT = 84
OUTBOUND PACKETS = 1
OUTBOUND PACKETS IN ERROR = 0
OUTBOUND PACKETS DISCARDED = 0
.
.
INTFNAME: OSX2320A INTFTYPE: IPAQENET INTFSTATUS: READY
PORTNAME: IUTXP019 DATAPATH: 2322 DATAPATHSTATUS: READY
CHPIDTYPE: OSX CHPID: 19
```

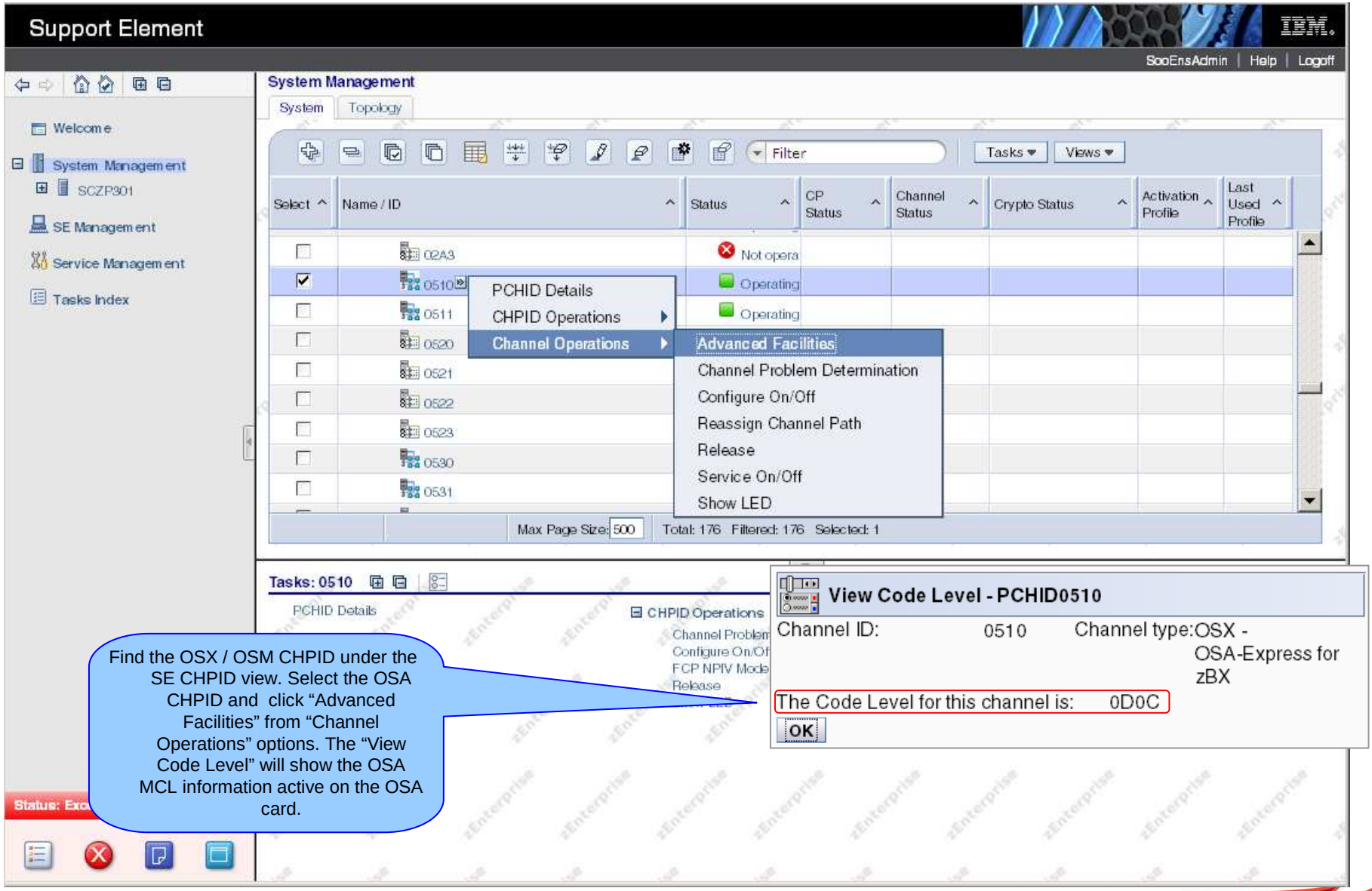
Only the output for OSX2300A is shown here. The DEVLINKS shows both interfaces .



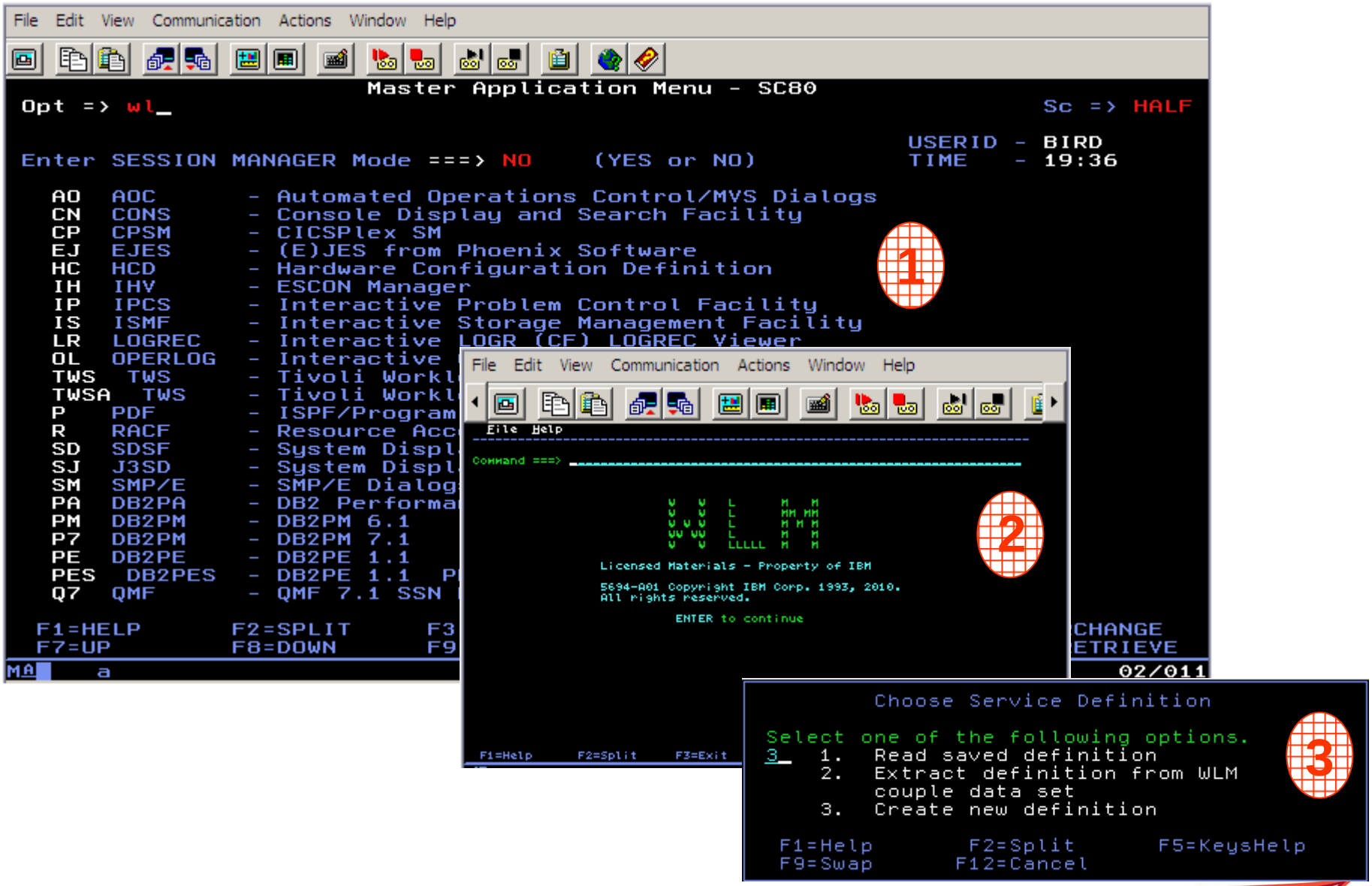
Checking OSX and OSM code Levels using the HMC



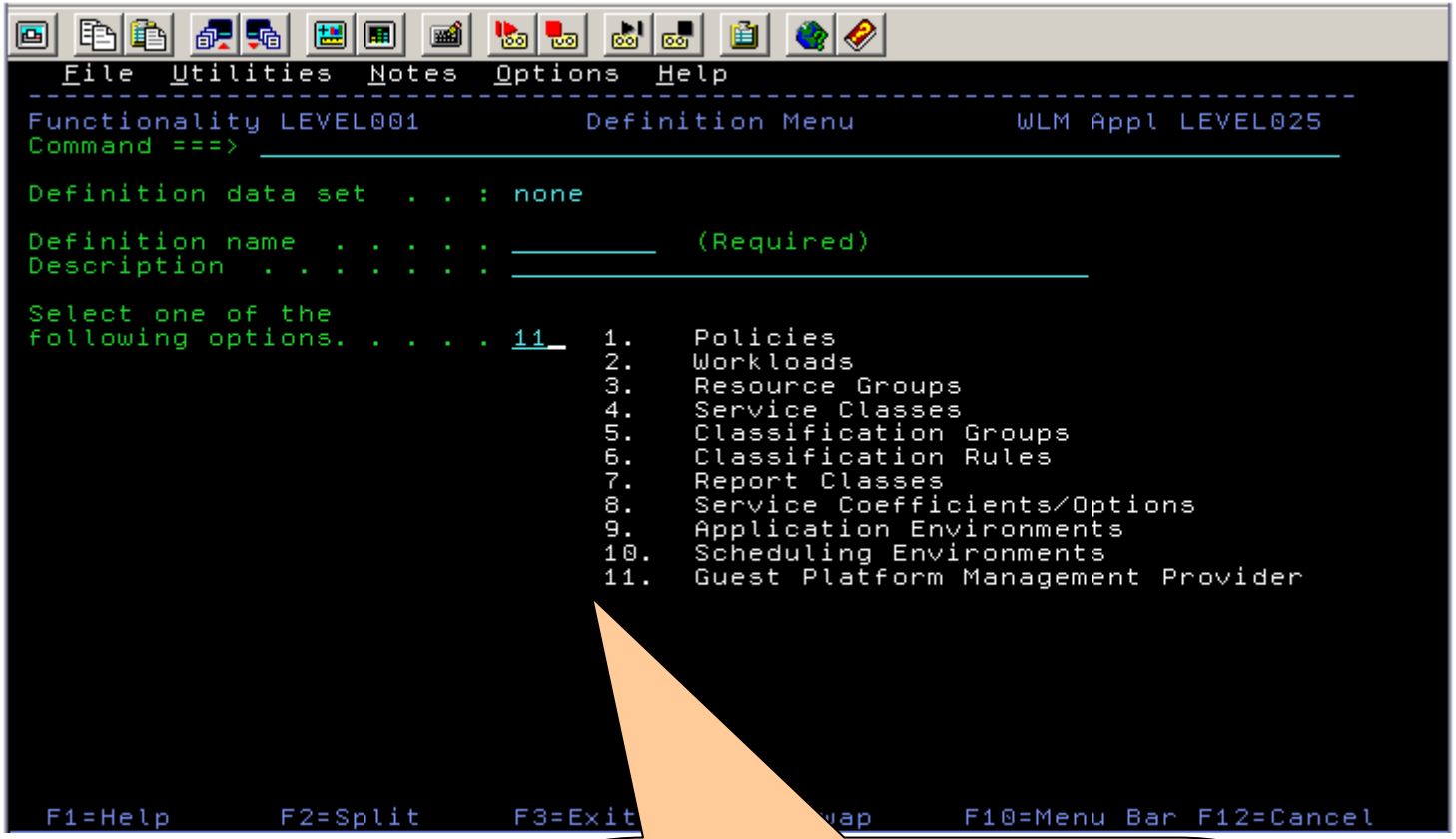
Checking OSX and OSM code Levels using the SE



z/OS GPMP – Guest Platform Management Provider



z/OS GPMP – Guest Platform Management Provider



The guest platform management provider (GPMP) is the interface between the intra-node management network (INMN) and the z/OS Workload Manager (WLM).

It provides policy information to WLM about the platform wide performance goals of the workloads in which the z/OS system is participating.

Enable GPMP on z/OS – Start GPMP Task

- Make sure HVEMCA procedure is in SYS1.PROCLIB
- Make certain that ARM is enabled
  - F WLM,AM=ENABLE
- Start GPMP
  - F WLM,GPMP,START

```
F WLM,GPMP,START
$HASP100 HVEMCA   ON STCINRDR
IEF695I START HVEMCA   WITH JOBNAME HVEMCA   IS ASSIGNED TO USER
HVEMCA1 , GROUP HVEMCA
$HASP373 HVEMCA   STARTED
IRR812I PROFILE ** (G) IN THE STARTED CLASS WAS USED 602
      TO START BPXAS WITH JOBNAME BPXAS.
$HASP100 BPXAS    ON STCINRDR
$HASP373 BPXAS    STARTED
BPXP024I BPXAS INITIATOR STARTED ON BEHALF OF JOB HVEMCA RUNNING IN ASID 0010
FEW0600I GPMP STARTED
```

Displaying the status of z/OS GPMP

D WLM,SYSTEMS,GPMP

```
IWM025I  15.37.07  WLM DISPLAY 597
  ACTIVE WORKLOAD MANAGEMENT SERVICE POLICY NAME: POL01
  ACTIVATED: 2010/11/15  AT: 12:57:47  BY: BARI      FROM: SC30
  DESCRIPTION: Initial Policy set up by Bob
  RELATED SERVICE DEFINITION NAME: WLMDEF01
  INSTALLED: 2010/11/15  AT: 12:57:19  BY: BARI      FROM: SC30
  WLM VERSION LEVEL:       LEVEL025
  WLM FUNCTIONALITY LEVEL: LEVEL003
  WLM CDS FORMAT LEVEL:    FORMAT 3
  STRUCTURE SYSZWLM_WORKUNIT STATUS: CONNECTED
  STRUCTURE SYSZWLM_3BD52817 STATUS: CONNECTED
  STATE OF GUEST PLATFORM MANAGEMENT PROVIDER (GPMP): ACTIVE
  *SYSNAME*  *MODE*  *POLICY*  *WORKLOAD MANAGEMENT STATUS*
  SC30       GOAL    POL01     ACTIVE
  SC31       GOAL    POL01     ACTIVE
  SC32       GOAL    POL01     ACTIVE
  SC33       GOAL    POL01     ACTIVE
  *SYSNAME*  *GPMP STATUS*
  SC30       INACTIVE
  SC31       INACTIVE
  SC32       INACTIVE
  SC33       INACTIVE
```

D WLM,AM

```
IWM075I  16.47.41  WLM DISPLAY 696
 ARM SERVICES ARE ENABLED
 GUEST PLATFORM MANAGEMENT PROVIDER JOBNAME=HVEMCA ASID=0010
 GPMP POLICY IS ACTIVE
 NUMBER OF REGISTERED PROCESSES=1, APPLICATIONS=1
```

Communications Server Definitions

- Sample ATCSTRxx

```
APPNCOS=NONE,
ENSEMBLE=YES,
CONFIG=00,
CONNTYPE=APPN,
CPCP=YES,
CSALIMIT=0,
CDRDYN=YES,
DYNLU=YES,
DYNADJCP=YES,
ISTCOSDF=INDLU,
NETID=USIBMWZ,
SSCPID=&SUBAREA.,
SSCPNAME=S&SUBAREA.CDRM,
SUPP=NOSUP,
CRPLBUF=(33,,4,,1,8),
IOBUF=(100,508,8,,14,15),
LFBUF=(2,,0,,1,1),
LPBUF=(36,,0,,1,1),
NODETYPE=EN,
SFBUF=(14,,0,,1,1),
XNETALS=YES
```

- Sample TCPPARMS

```
;OSX ---- TSYS CHPID 18 ----------
INTERFACE OSX2300A
      DEFINE IPAQENET
      CHPIDTYPE OSX
      IPADDR 172.30.99.1/24
      CHPID 18
      MTU 8992
      VLANID 99
      VMAC ROUTALL

;OSX ---- TSYS CHPID 19 ----------
INTERFACE OSX2320A
      DEFINE IPAQENET
      CHPIDTYPE OSX
      IPADDR 172.30.99.2/24.11/24
      CHPID 19
      MTU 8992
      VLANID 99
      VMAC ROUTALL
```

z/OS IPL Messages

- The z/OS IPL expected messages are shown below

```
EZZ4340I INITIALIZATION COMPLETE FOR INTERFACE OSX2300A
EZZ4340I INITIALIZATION COMPLETE FOR INTERFACE OSX2320A
IEF196I IEF285I   SYS1.LINKLIB                    KEPT
IEF196I IEF285I   VOL SER NOS= ZBLRS1
EZZ4340I INITIALIZATION COMPLETE FOR INTERFACE EZ6OSM02
EZZ4340I INITIALIZATION COMPLETE FOR INTERFACE EZ6OSM01
EZD1176I TCPIP HAS SUCCESSFULLY JOINED THE TCP/IP SYSPLEX GROUP EZBTCPCS
```

Part IV – Ensemble Managing and Problem Determination

4 AIX

AIX GPMP Status

AIX Virtual Server Details

Installing the GPMP package in AIX

- Make a mount point and mount the virtual media

```
mkdir /cdrom
mount –V cdrfs –o ro /dev/cd0  /cdrom
```

- Change to the mount point directory and install GPMP package appropriate for AIX version

```
cd /cdrom
rpm –ivh <gpmp_Package>
```

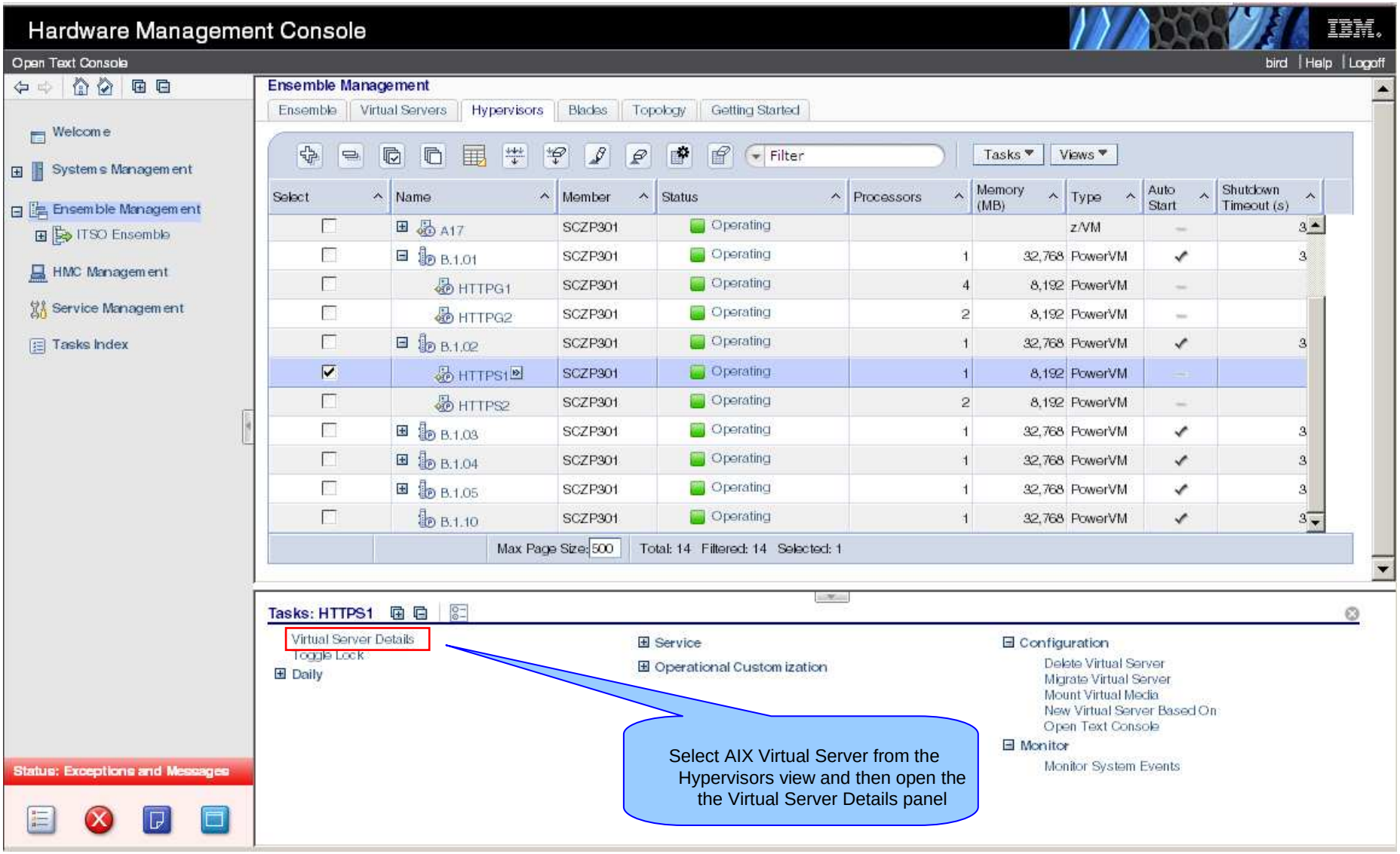
- Optionally execute the post-install script to add user-id for running GPMP

```
/opt/ibm/gpmp/post-install-config
```

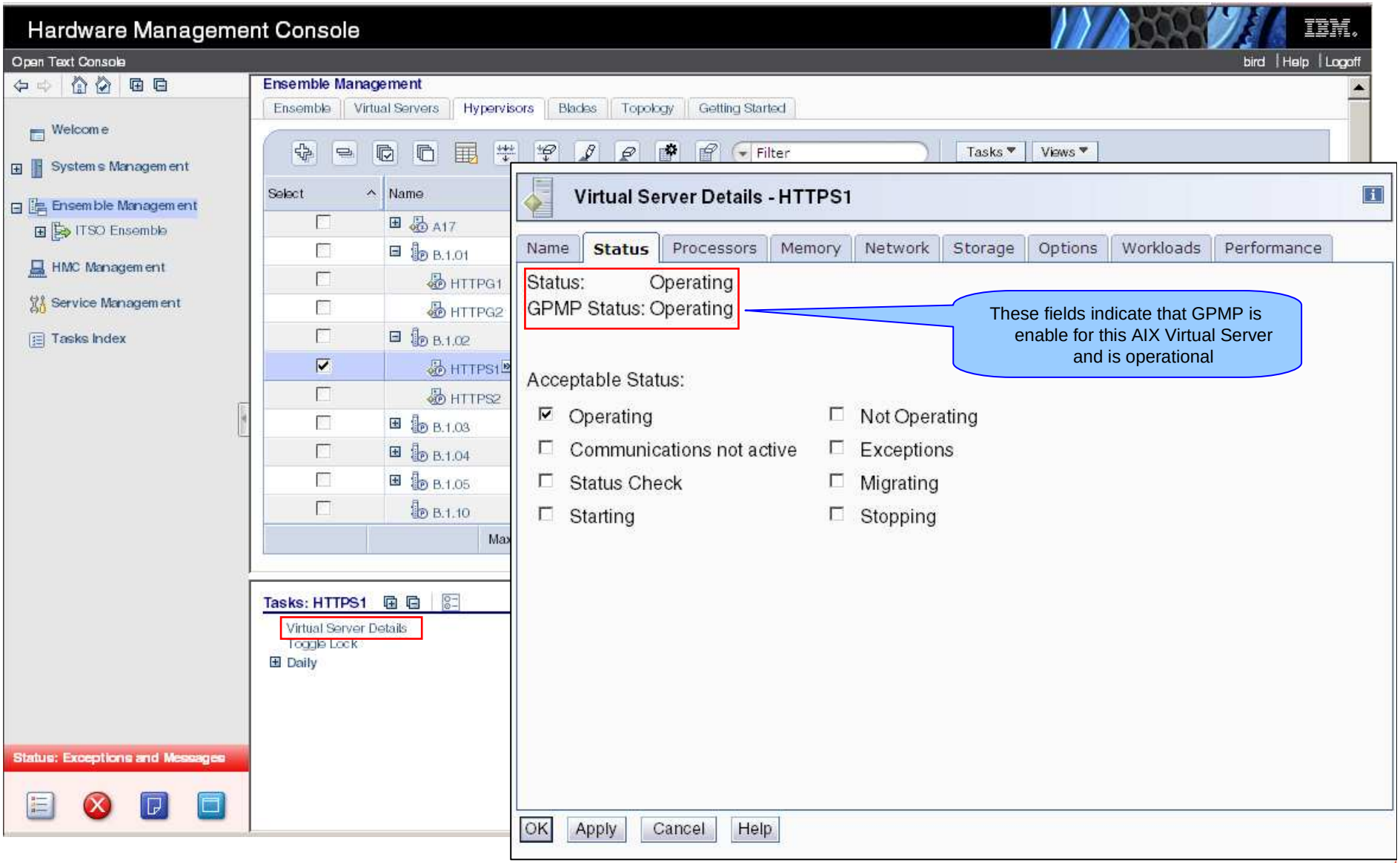
- Start the GPMP

```
/opt/ibm/gpmp/gpmp start
```

Checking GPMP status in AIX



AIX Virtual Server GPMP status

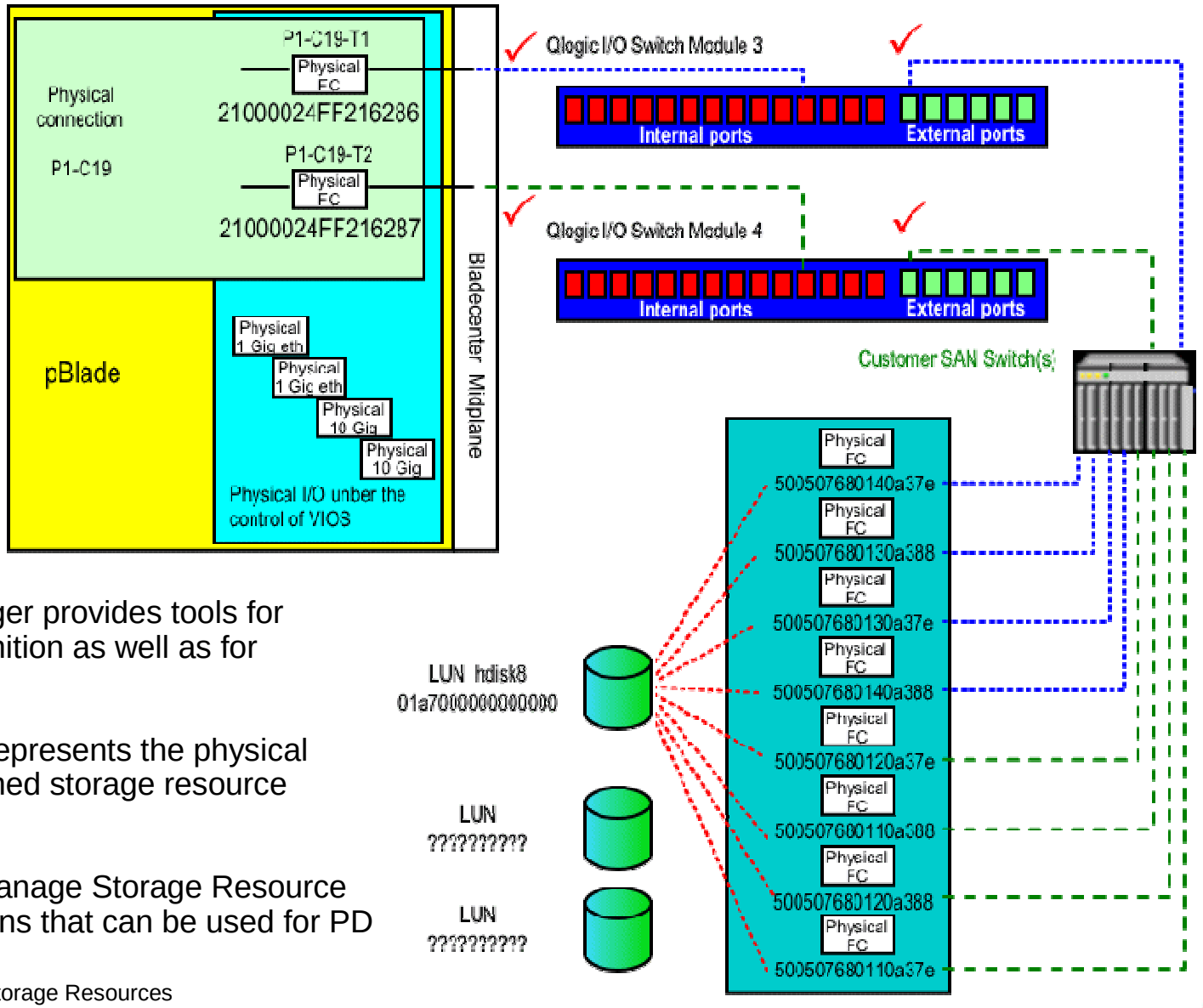


Part IV – Ensemble Managing and Problem Determination

5 zBX

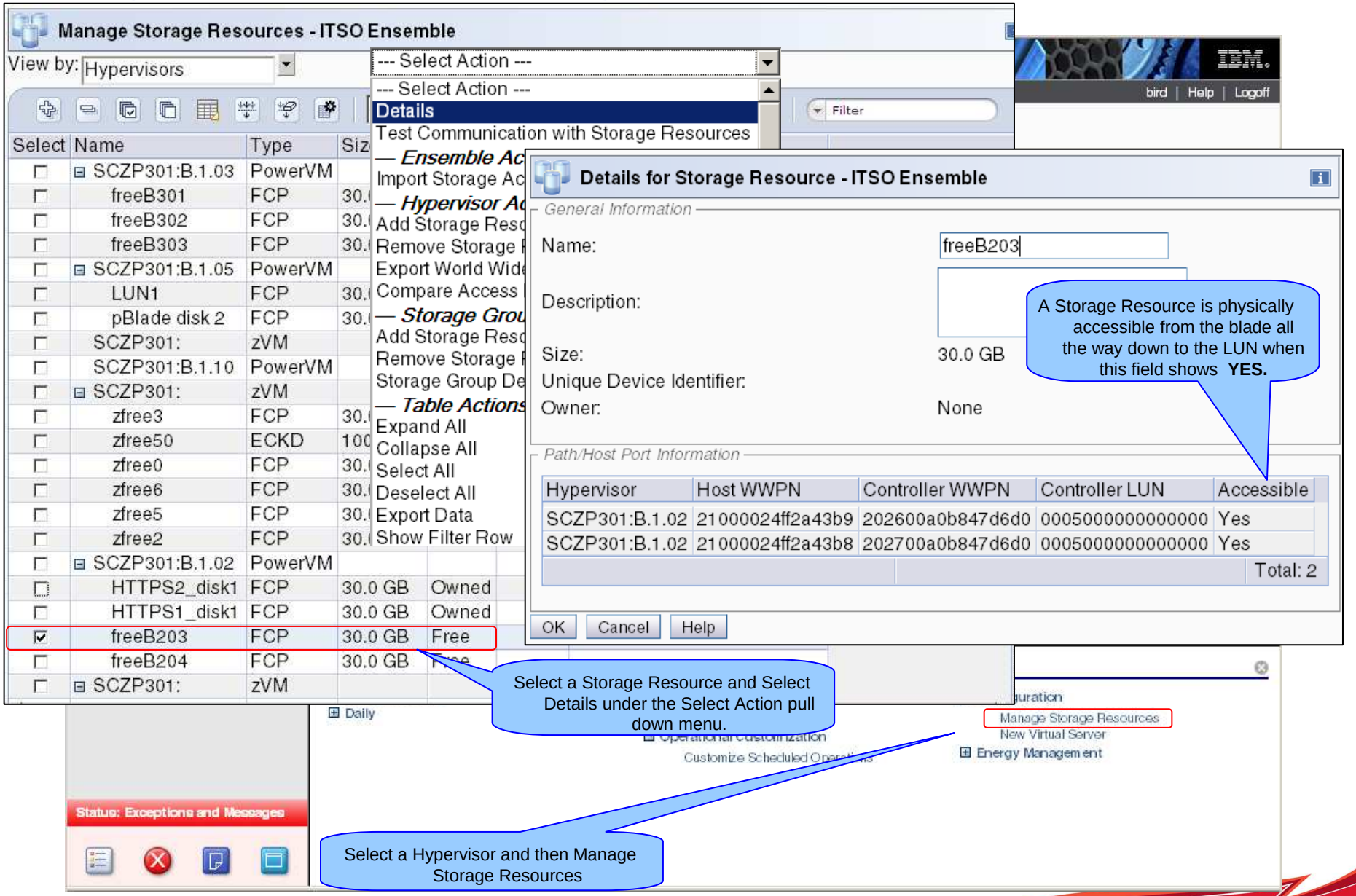
Some blade Storage PD

zBX Storage PD



- Unified Resource Manager provides tools for Storage Resources definition as well as for Problem Determination
- The picture to the right represents the physical connectivity path to defined storage resource hdisk8
- Under hypervisor and Manage Storage Resource options there are 2 options that can be used for PD
  - Display storage Details
  - Test Communications with Storage Resources

Storage resource connectivity PD



Part IV – Ensemble Managing and Problem Determination

6 Glossary

Glossary

Blade	Hardware that provides application-specific services and components. The consistent size and shape (or form factor) of each blade allows it to fit in a BladeCenter chassis.
BladeCenter chassis	A modular chassis that can contain multiple blades, allowing the individual blades to share resources such as the management, switch, power, and blower modules
IBM BladeCenter PS701 Express	The supported POWER7 blade that can be installed in a zBX Model 002
IBM Smart Analytics Optimizer	IBM Smart Analytics Optimizer
AMM	Advance Management Module
Node	A single z196 and any optionally attached zBX. A node can be a member of only one ensemble
Ensemble	A collection of one or more zEnterprise nodes (including any optionally attached zBX) that are managed as a single logical virtualized system by the Unified Resource Manager using a Hardware Management Console (HMC)
BPH	Bulk Power Hub (redundant hw switch inside z196 and z114 CECs)
BCN	BladeCenter H
HMC	Hardware Management Console (primary or alternate)
GPFS	General Parallel File System
IEDN	intraensemble data network
INMN	intranode management network
OSM	OSA Channel path identifier (CHPID) for INMN
OSX	OSA Channel path identifier (CHPID) for IEDN
PDN	Private Data Network
Unified Resource Manager	IBM zEnterprise Unified Resource Manager – The Unified Resource Manager provides energy monitoring and management, goal-oriented policy management, increased security, virtual networking, and data management for the physical and logical resources of a given ensemble
PSN	Private Support Network
PSCN	Power System Control Network.
SOD	Statement of Direction
zBX	IBM zEnterprise BladeCenter Extension

Glossary

EM	
PPM	
HVM	
NVM	Network Virtualization Manager – component on HMC responsible for MAC address and VLAN assignments for all servers running in the ensemble
VSM	
SVM	
API	Application Programming Interface
SMAPI	System Management APIs
GPMP	

yourdotcom

International Technical Support Organization and Authoring Services

IBM Unified Resource Manager

www.ibm.com/redbooks

Part V – What's new

Part V – What's new

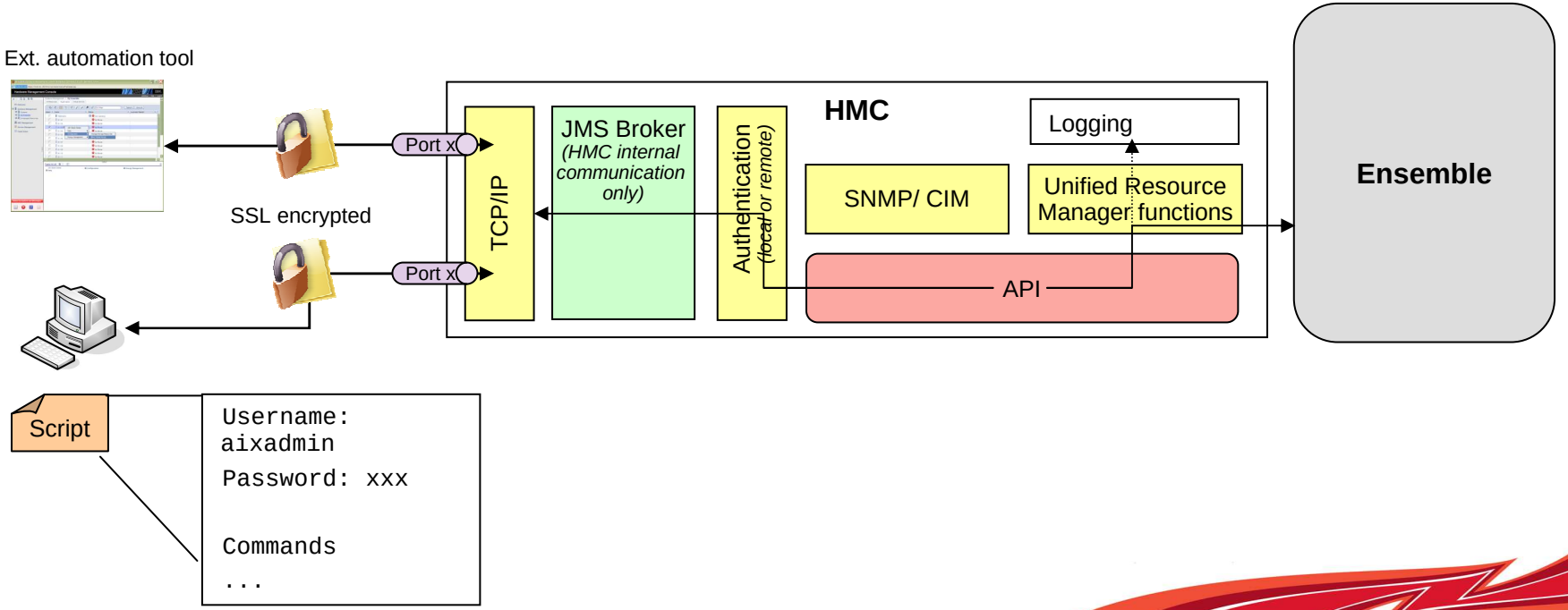
1. System management API's
2. Externalize IEDN Metrics & Network Monitors Dashboard
3. Storage Virtualization Management Enhancements

Part V – What's new

1. **System management API's**
2. Externalize IEDN Metrics & Network Monitors Dashboard
3. Storage Virtualization Management Enhancements

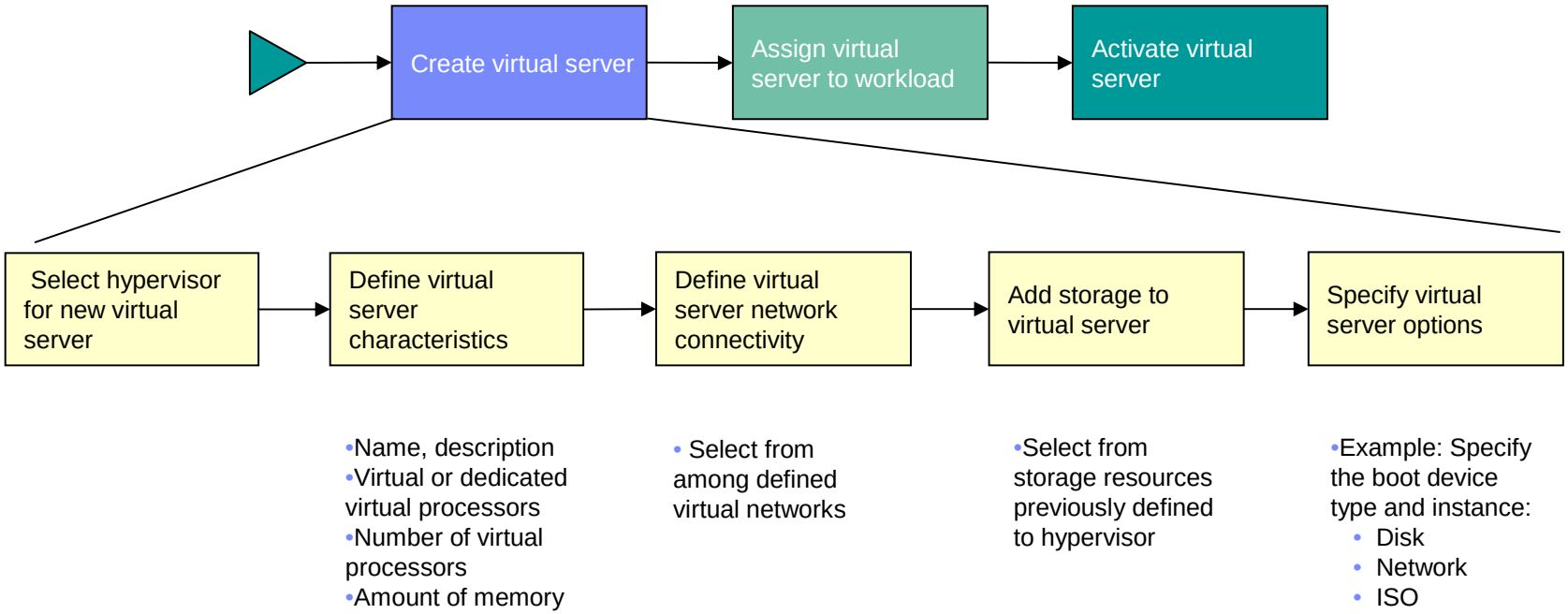
Management using external API

- Management API will be a new API implementation in the HMC
  - HMC UI remains in place, supported and will continue to be extended
- Architecture / design based on current industry design practices
- Security
  - SSL for connection security to the HMC
  - Requests always performed under an HMC user context → need the right user profile settings
- HTTP web services orientation facilities access from scripting environments
  - Modern scripting languages (Perl, Python, etc.) have HTTP support libraries



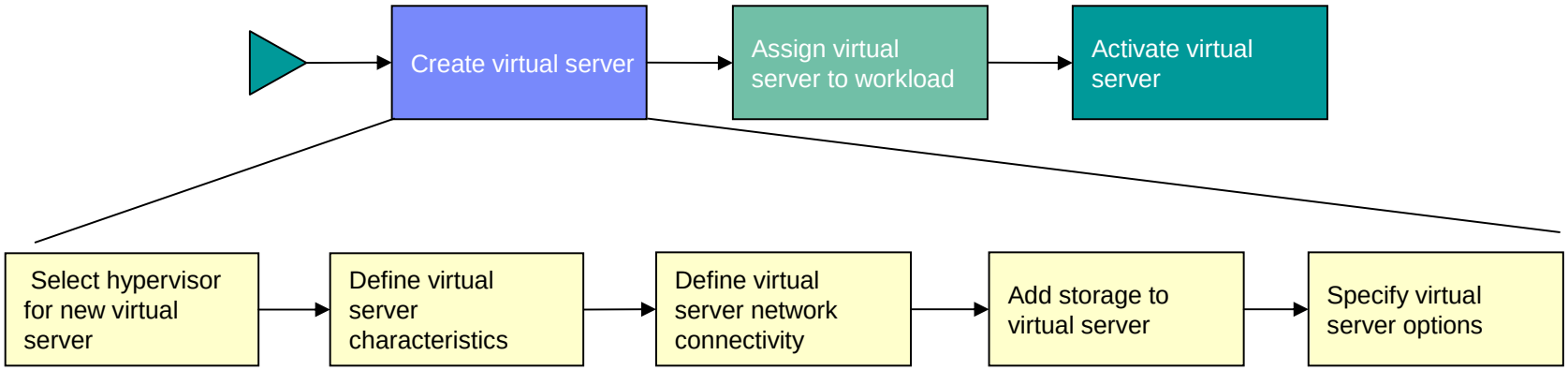
HMC API and UI Provide Same Level of Function

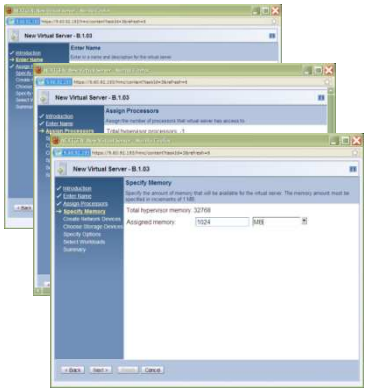
- Example: Creating a Virtual Server on an IBM Blade
- Regardless of the interface used, this is accomplished through a series of steps:

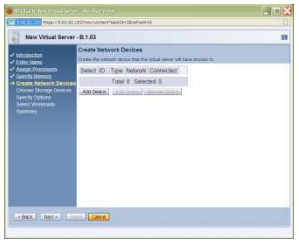


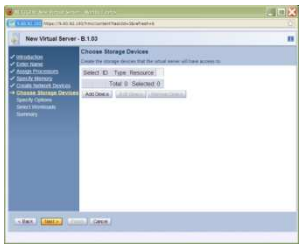
HMC API and UI Provide Same Level of Function (con't)

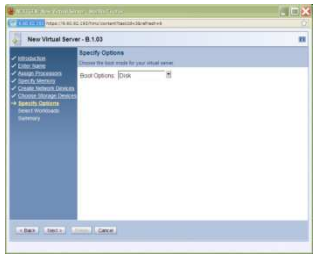
- HMC UI: Steps are accomplished using panels in a wizard-style task





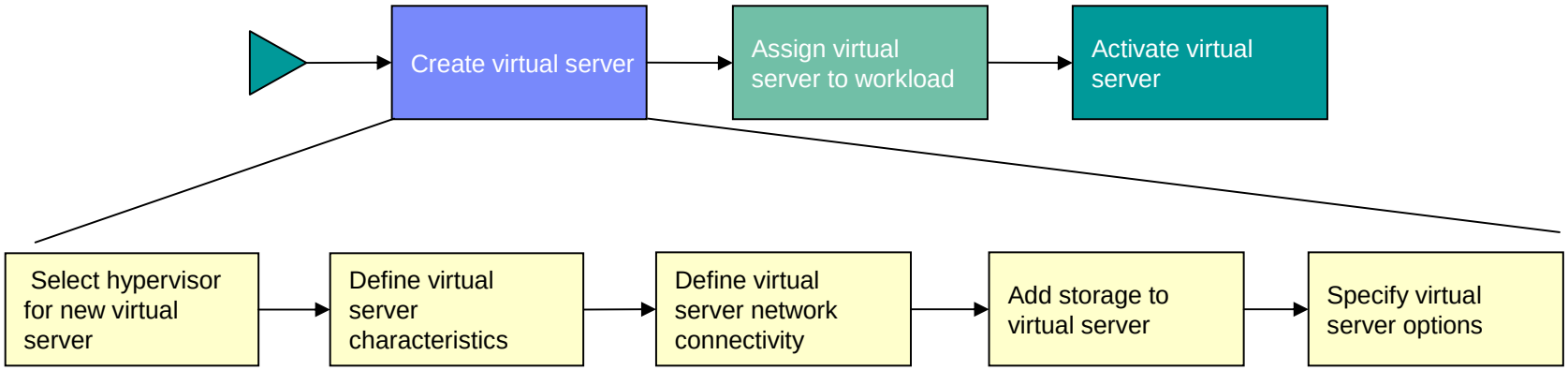






HMC API and UI Provide Same Level of Function (con't)

- Unified Resource Manager API: Steps are accomplished by calling management primitives of the API
- Note: Function names listed below are conceptual, not the actual API syntax



- Call List-Hypervisors function to obtain a list of hypervisors
- <Invoking application selects desired hypervisor>

- Call Create-VS function specifying selected hypervisor as target and basic VS parameters to get base VS created

- Call List-VNetworks function to obtain current virtual networks
- <Select network>
- Call Add-VNIC function specifying new VS as target and virtual network parameters

- Call List-Stg-Resourcces function to obtain list of available volumes
- <Select volume>

Call Add-VDisk function specifying new VS as target and selected storage resource

- <Select boot device>
- Call Update-VS function to set boot device

Summary of API Functional Scope

- Core (Traditional) Entities
  - List, Get/Set properties
  - Start/Stop/Restart/etc.
  - For CPC, Image, LPAR, CF, Groups, Capacity records, Console, etc.
- Ensemble
  - List, Get/Set properties
  - Add/Remove CPC members
- Workloads
  - List, Create, Delete
  - Get/Set properties
  - Add/Remove virtual-server members
  - Create/Delete performance policies
  - Get performance report data
- Virtual Networks
  - List, Create, Delete
  - Get/Set Properties
  - Recovery actions
- Hypervisors (KVM, POWER VM, z/VM)
  - List, Get/Set properties
  - Start/Stop
- Virtual servers (in KVM, POWER VM, z/VM)
  - List, Create, Delete
  - Get/Set properties
  - Start/Stop
- Storage
  - Define, List
  - Assign to hypervisor
  - Assign to storage groups
- zBX infrastructure (BladeCenters, Blades)
  - List, Get/Set properties
  - Get/Set energy management modes
- Service Oriented Functions
  - Metrics retrieval
  - Inventory

Sample of a script

- Generic
  - Scripted in Jython (modern scripting language)
  - Can be triggered from any remote workstation that has TCP/IP access to the HMC
- Requirements
  - HMC enabled for API communication
    - HMC API activated
    - Open port on the HMC
  - User
    - Use an ID that is defined on the HMC
    - ID and pwd are sent as arguments with the execution command (KEEP IN MIND SSL IS USED)
- Uses http commands/ mechanism
  - URI information
  - Return codes

```
# ##############################################################
# Get Ensemble
# ##############################################################

def getEnsemble():
    params = ""
    response = request("GET","/api/ensembles",params, headers)
    #print response.status
    if response.status == 400:
        errorResponseDisplay( "exit due to bad response from GE
        sys.exit()

    objs = json.loads(response.read())
    ens=[]
    if 'ensembles' in objs:
        for o in objs['ensembles']:
            if 'object-uri' in o:

                ens.append(o['object-uri'])


        return ens
    else:
        sys.exit("no ensemble found")


# ##############################################################
# Get Virtual Serve of Ensemble
# ##############################################################
def getVserver (ensemble):
    ensemble += '/virtual-servers'
    params=""
    response = request("GET",ensemble,params,headers)
    #print response.status
    if response.status == 404:
        errorResponseDisplay( "exit due to bad response from GE
        sys.exit()
```

Part V – What's new

1 System management API's

2 Externalize IEDN Metrics & Network Monitors Dashboard

3 Storage Virtualization Management Enhancements

Network Monitors Dashboard

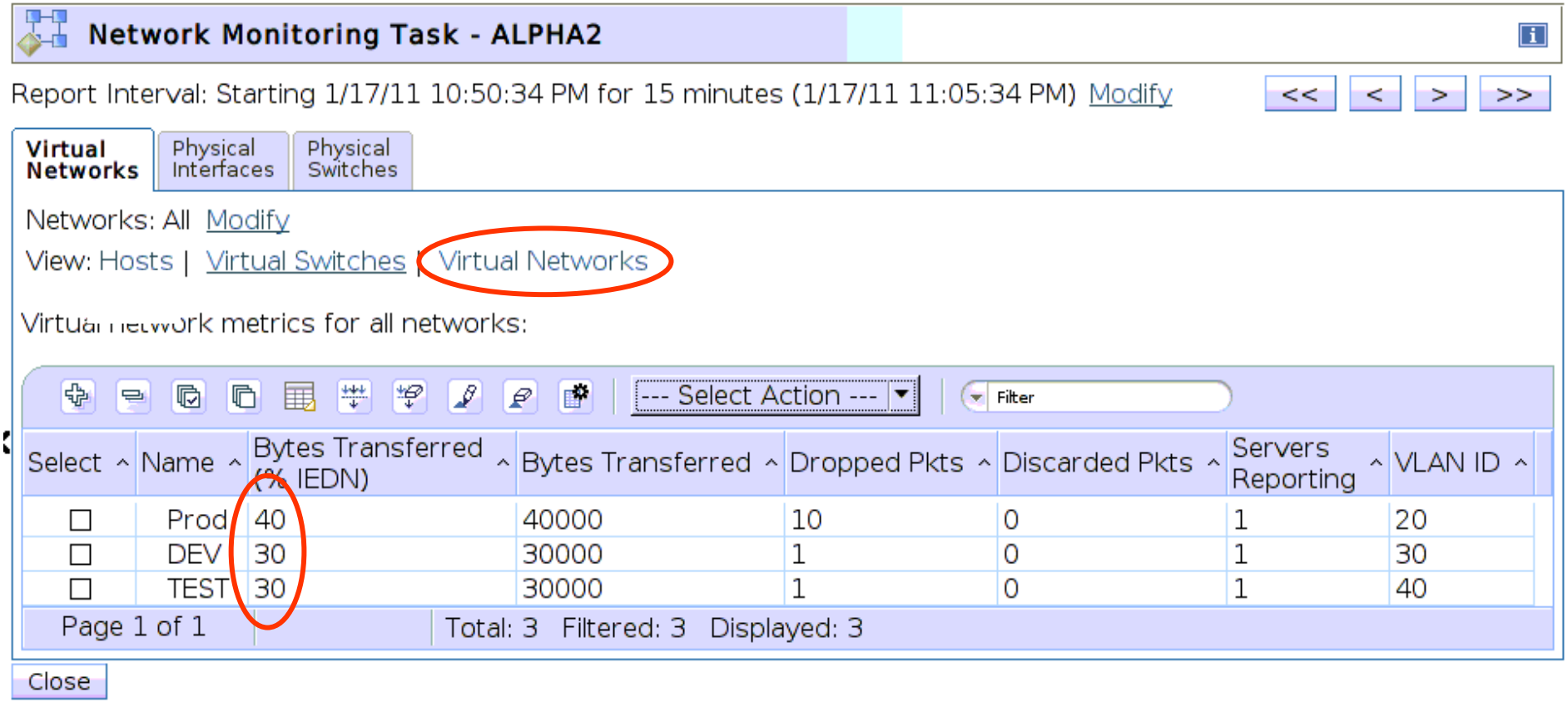
- New User Interface display that shows the following major views
  - Virtual Networks across IEDN, allowing drilling down to resources that comprise the virtual network: virtual switches, virtual servers, virtual network interfaces
    - These views show the performance between the virtual switch and the guest
    - Provide data “by” virtual network (i.e., by VLAN) and across all virtual networks
  - Physical Interfaces to the IEDN
    - Virtual Switch Uplink interfaces to the physical interfaces that connect to the IEDN
    - External TOR ports connected to a router
  - Physical Switches - TOR and ESM (Enterprise Switching Module – IEDN switch in BladeCenters)

- Metrics shown
  - Performance statistics (transfer rates over intervals)
  - Cumulative and interval metrics
  - Raw metrics as collected
  - Resource state/status (where available)
  - Resource utilization relationships (e.g., Servers using an OSA)

- Network Monitors Dashboard provides both current and historical displays

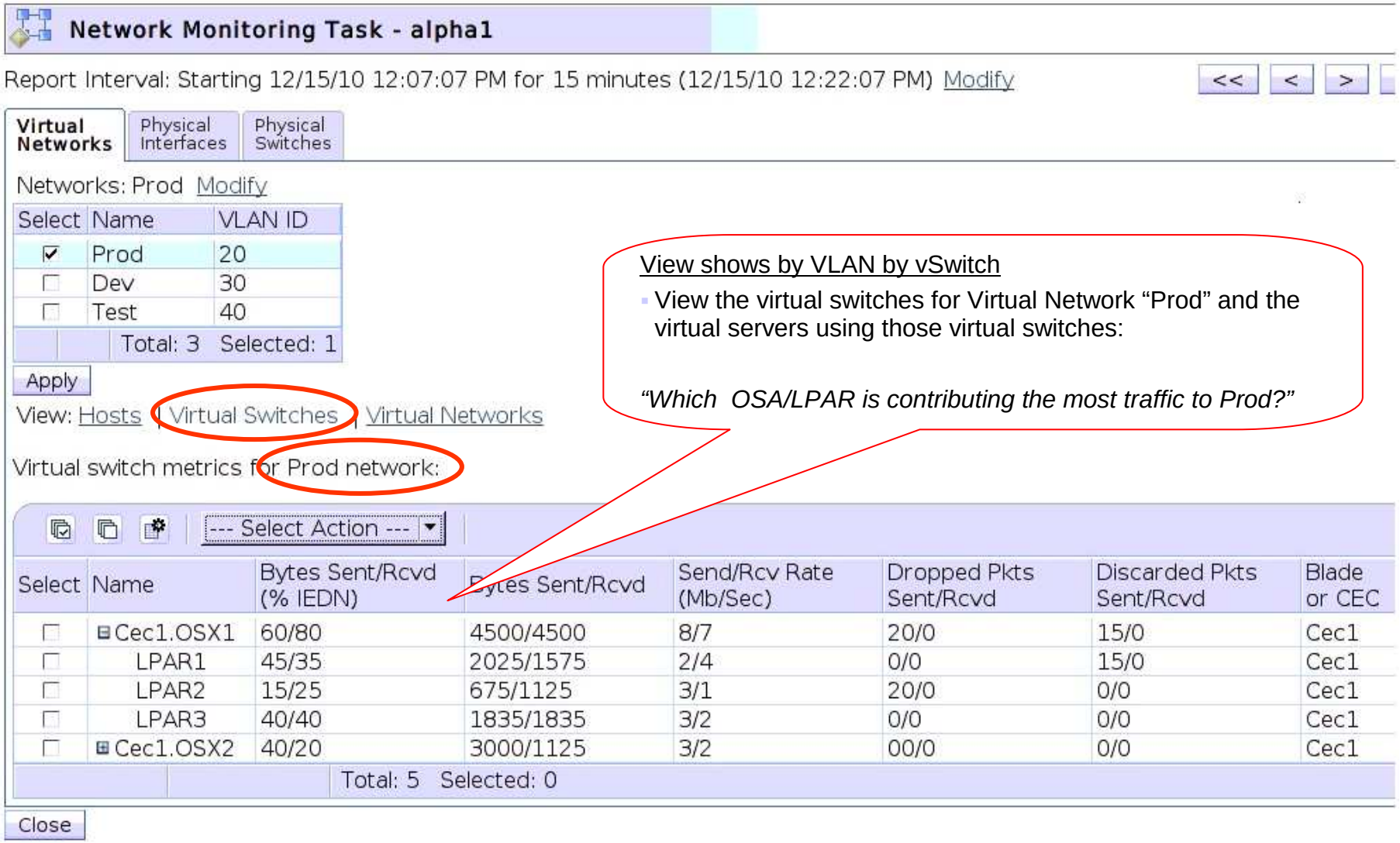
Sample Dashboard View: By Virtual Networks

- Resource usage by VLAN as % of IEDN bandwidth



Sample Dashboard View: By VLAN By vSwitch

- "Which OSA/LPAR is contributing the most traffic to Prod?"



View shows by VLAN by vSwitch

- View the virtual switches for Virtual Network "Prod" and the virtual servers using those virtual switches:

"Which OSA/LPAR is contributing the most traffic to Prod?"

Part V – What's new

1 System management API's

2 Externalize IEDN Metrics & Network Monitors Dashboard

3 Storage Virtualization Management Enhancements

Storage Virtualization Management Enhancements

- Current support
  - FCP-based storage resources have to be added manually by entering addressing information (WWPN/LUN pairs)
- Enhanced support
  - Allow the server admin to trigger dynamic discovery of additional storage resources through the User Interface.
  - The U/I will provide a list of newly detected storage resources that are not part of the current configuration
  - Server admin has to assign/add
    - unique name
    - VOLSER (z/VM only)
    - device number (z/VM only)
  - Server admin can optionally provide
    - description
    - size information

Thomas Schulze

IBM STG Lab Services
R&D Böblingen, Germany
Mail: tschulze@de.ibm.com

Special thanks to the following:

- Carl Mayer, IBM Germany
- Falk Steinbrueck, IBM Germany
- Friedemann Baitinger, IBM Germany
- Gary Tressler, IBM US
- Ira Siegel, IBM US
- Jerry Stevens, IBM US
- Karlo Petri, IBM Germany
- Leslie Geer III, IBM US
- Linda Flanders, IBM US
- Romney White, IBM US
- Steve Fellenz, IBM US
- Utz Bacher, IBM Germany

Dank u
Dutch

Merci
French

Спасибо
Russian

شكراً
Arabic

감사합니다
Korean

Tack så mycket
Swedish

धन्यवाद
Hindi

תודה רבה
Hebrew

Gracias
Spanish

Obrigado
Brazilian Portuguese

Dankon
Esperanto

Thank You

谢谢
Chinese

ありがとうございます
Japanese

Trugarez
Breton

Danke
German

Tak
Danish

Grazie
Italian

நன்றி
Tamil

děkuji
Czech

ขอบคุณ
Thai

go raibh maith agat
Gaelic