

SHARE in Atlanta, March 2012



zEnterprise eXposed Part 1: The zEnterprise Intersection of WLM, RMF, and Unified Resource Manager Performance Management

Session 10906

Glenn Anderson, IBM Technical Training



© 2011 IBM Corporation

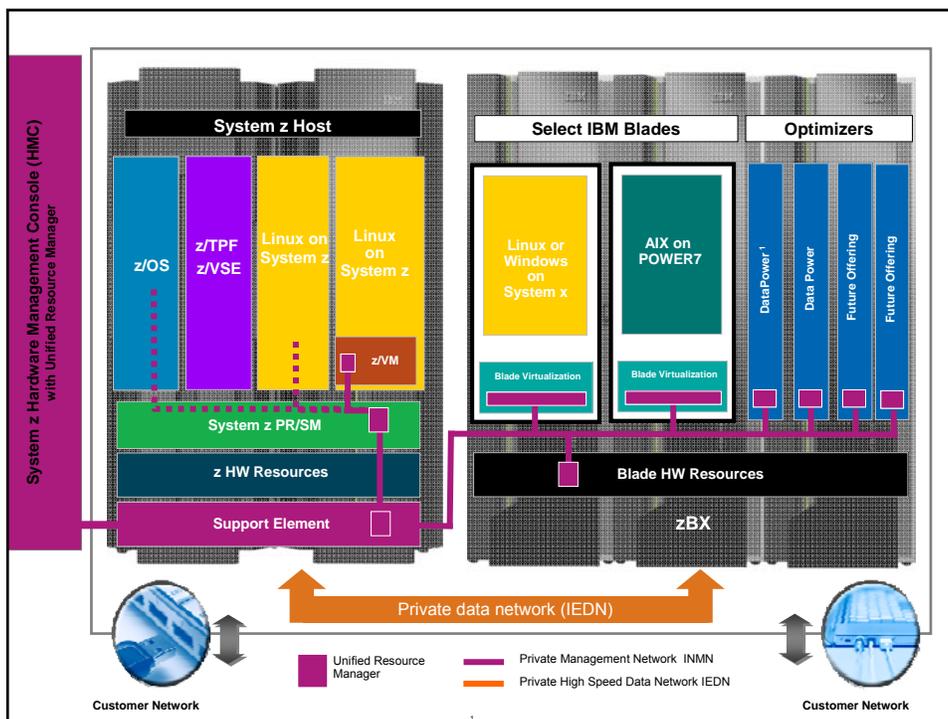


Agenda

- zEnterprise Workload Management
 - z/OS Virtual Servers
 - WLM and IRD
 - z/VM Guests and zBX Blade Virtual Servers
 - PPM

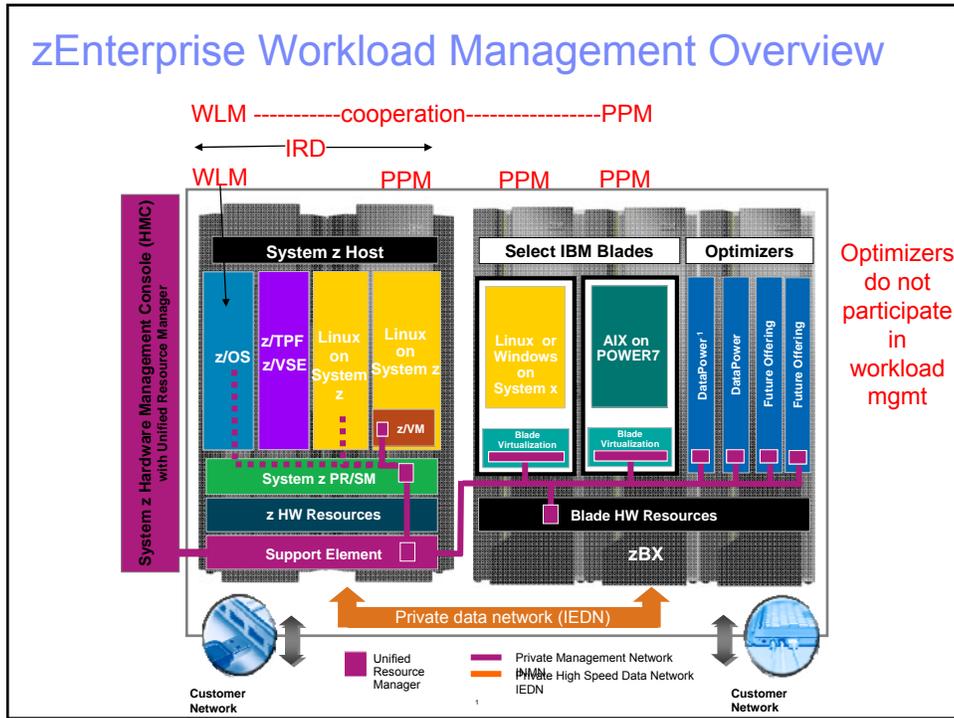
- zEnterprise Resource Monitoring
 - z/OS Virtual Servers
 - RMF
 - z/VM Guests and zBX Blade Virtual Servers
 - PPM
 - RMF XP

© Copyright IBM Corporation 2012



zEnterprise Workload Management

© Copyright IBM Corporation 2012



Platform Performance Manager

zEnterprise Platform Performance Manager

- **Platform management component responsible for goal-oriented resource monitoring, management, and reporting across the zEnterprise Ensemble**
 - Core component responsible for definition and implementation of goal-oriented management policy
 - Extend goal oriented approach of z/OS WLM to platform managed resources
 - Common approach to monitoring / management of platform resources across zEnterprise
 - Orchestration of autonomic management of resources across virtual servers
 - Provide Intelligent Resource Director like function across the zEnterprise
 - Pushes management directives to the SE, Hypervisors, and OS agents as required across the zEnterprise
- **Integration of HMC console support**
 - Integrated UI for monitoring, display of workload topology relationships, status alerts, etc
 - Definition of Performance Management Goals and Policy Administration
- **Functionality integrated into the Unified Resource Manager**
 - Code structured and packaged as System Z firmware
 - Inter-Component communication over trusted internal platform management network

© Copyright IBM Corporation 2012

PPM Components

- **HMC**
 - HMC is management server and console
 - Provides ensemble wide aggregation of performance data
 - UI for defining workloads, performance policy and reporting data
 - Pushes management directives to all the nodes of ensemble
- **Support Element (SE)**
 - Provides node (or CPC) level aggregation of performance data
 - Pushes management directives to all the hypervisors in the node.

The diagram illustrates the PPM components. On the left, a vertical purple bar represents the HMC. To its right, a yellow box represents the SE (Support Element), which contains PR/SM, Z CPU, Memory and IO, and SE. Further right, a grey box represents the z Blade Extension, which contains two main sections: Power7 and System x. The Power7 section includes zOS, z/VM, and z/VM Mgmt Guest. The System x section includes AIX, VIOS, Linux, Windows, Power VM, xHyp, and Optimizer. An AMM (Application Management Module) is shown at the bottom of the z Blade Extension.

© Copyright IBM Corporation 2012

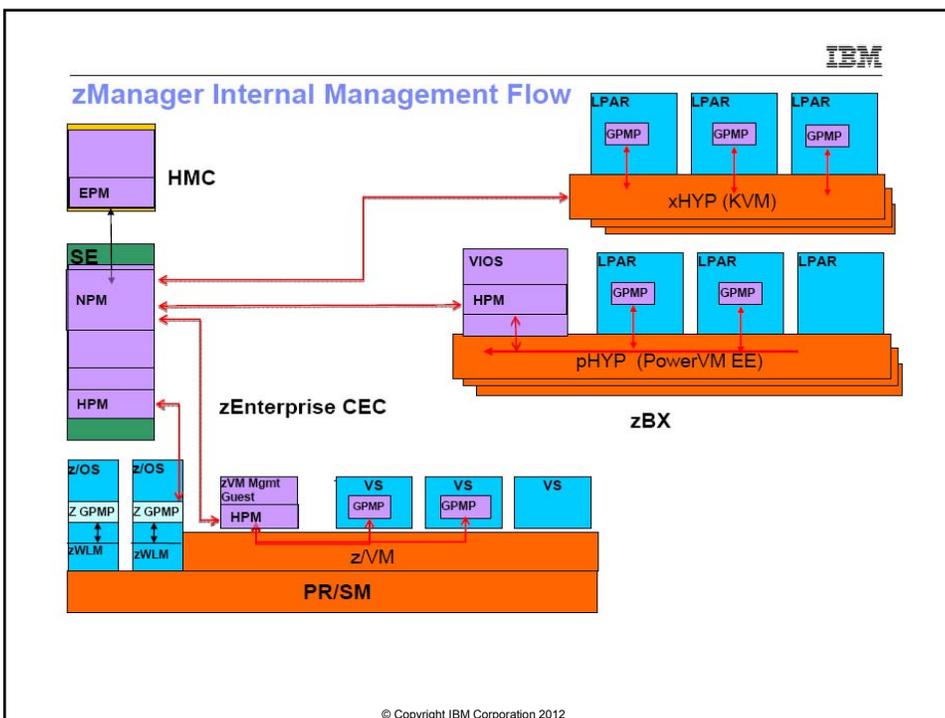
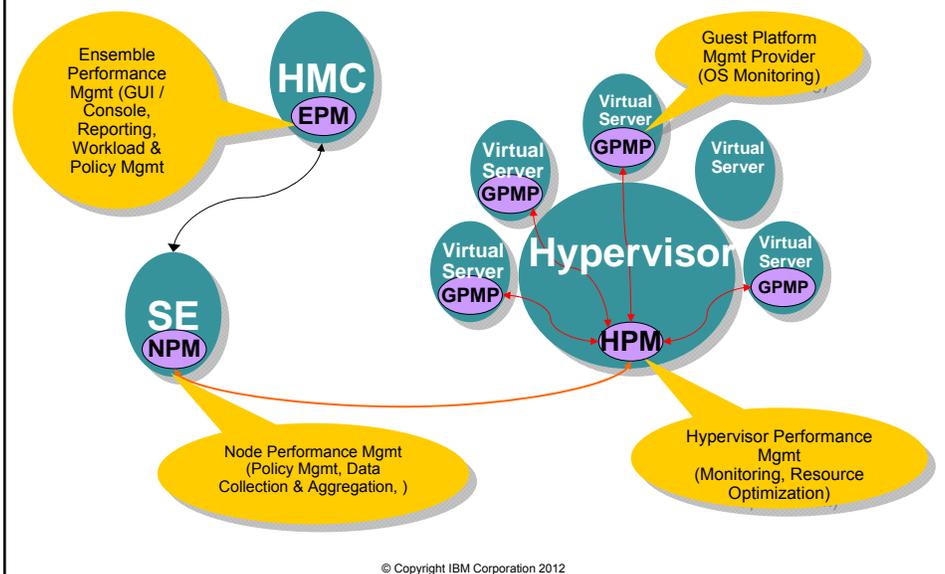
PPM Components

- **Hypervisors**
 - Monitors goal defined in performance policy and performs dynamic resource mgmt (z/VM and Power VM) to achieve performance goal where applicable
 - Collects virtual server statistics from hypervisor and guest platform management providers. Pushes aggregated metrics to SE
- **Virtual Servers**
 - **Optional** Guest Platform Management Provider software deployed in Virtual Server
 - Collects monitoring data from Operating system and ARM instrumented applications and pushes to hypervisors.

The diagram illustrates the PPM components, identical to the one above. It shows the HMC, SE, and z Blade Extension with its Power7 and System x sections, including various operating systems and management tools.

© Copyright IBM Corporation 2012

Platform Performance Management Structure



IBM

zManager CPU Mgmt Functions

- z/VM and PowerVM Hypervisors
 - Virtual Server CPU Management provides the ability to manage CPU resources across virtual servers based on a goal-oriented performance policy.
- System x xHyp (KVM based) Hypervisor
 - Does not currently participate in CPU Management
- PR/SM Hypervisor
 - Does not make resource management adjustments based on PPM Policy. Only IRD dynamically influences the PR/SM hypervisor

© Copyright IBM Corporation 2012

IBM

Platform Workload

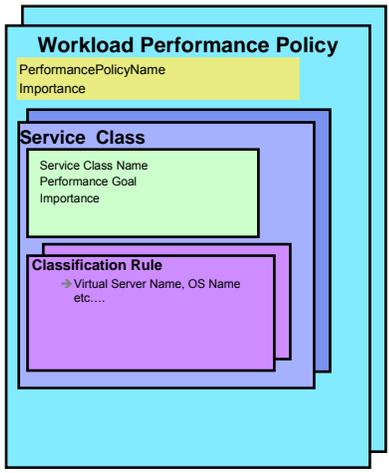
- A Platform Workload is a grouping mechanism and “management view” of virtual servers supporting a business application
- Provides the context within which associated platform resources are presented, monitored, reported, and managed
- Management policies are associated to Platform Workload
 - Currently supports Performance Policy

The diagram illustrates the Platform Workload architecture. It shows two Central Electronic Complexes (CEC 1 and CEC 2). CEC 1 contains a Web Server for Payroll app on Linux and a Web Server for HR app on Linux. CEC 2 contains a WebSphere for Payroll app on Linux and a WebSphere for HR app on Linux. Both CECs share a DB2 database on z/OS. Two Performance Policy boxes are shown, one pointing to the Payroll workload and one pointing to the HR workload. The Payroll workload is enclosed in a blue dashed box labeled 'Workload = Payroll', and the HR workload is enclosed in a pink dashed box labeled 'Workload = HR'.

© Copyright IBM Corporation 2012



Workload Performance Policy...



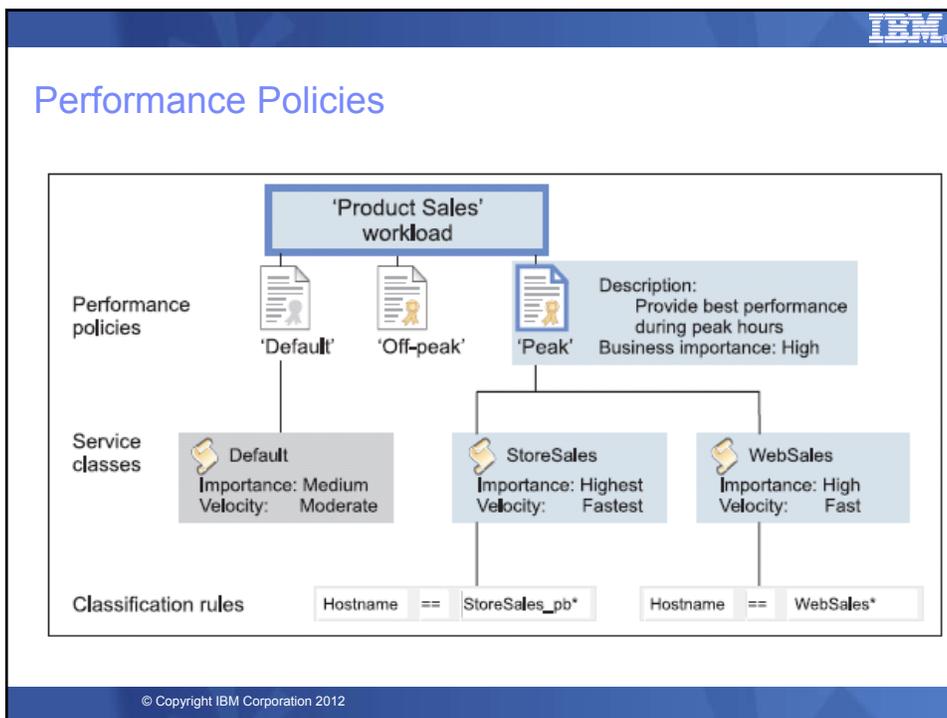
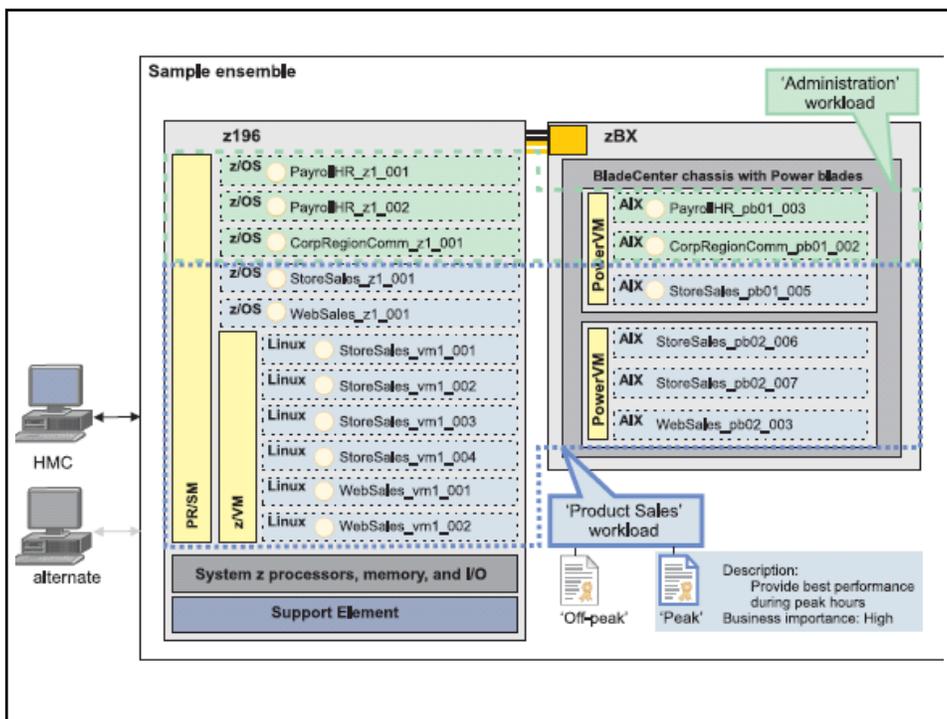
- Policy structure:
 - Policy contains a set of service classes
 - Classification rules map each virtual server within the workload to a service class
 - A service class assigns a performance goal and importance
- HMC as console for policy creation and editing
 - Wizard for policy creation
 - Repository for policies under development and saved policies
 - Links to Workload based performance reporting

© Copyright IBM Corporation 2012

Elements of a Service Class

- **Performance Goal (managed at the virtual server level)**
 - Velocity: Fastest, Fast, Moderate, Slow, Slowest
 - Discretionary: No performance goal
- **Business Importance:** Highest, High, Medium, Low, Lowest
- **Classification Rule**
 - Use Virtual Server Name as qualifier to assign Service Class
 - Virtual Servers under the PR/SM and System x hypervisors should be classified into a Service Class for resource monitoring purposes

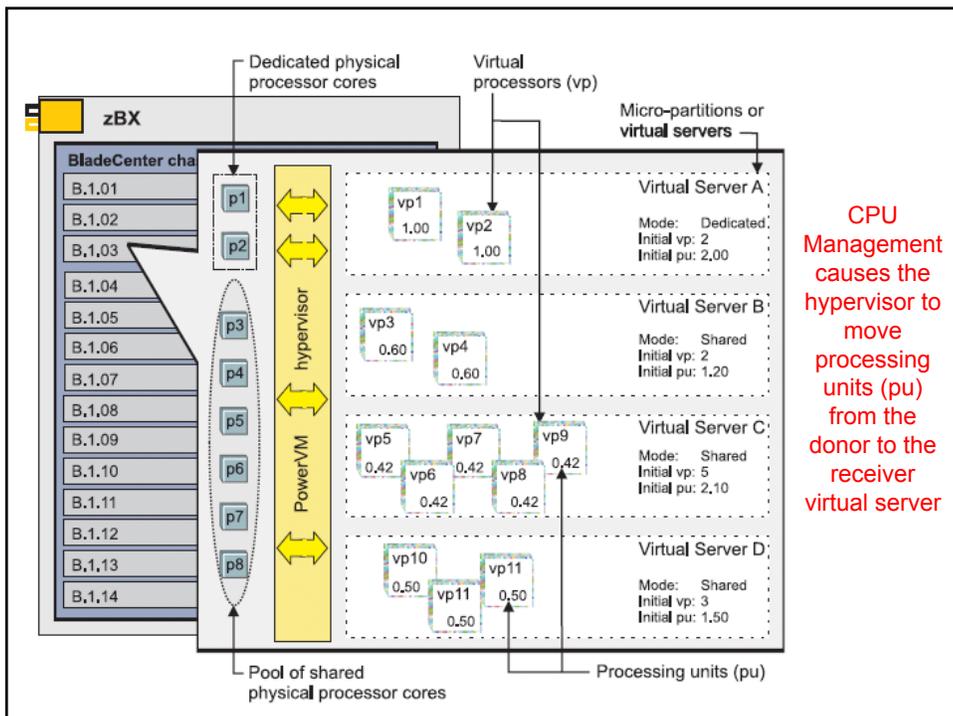
© Copyright IBM Corporation 2012



Managing Resources across Virtual Servers on P7 blade

- Manage resources across virtual servers to achieve workload goals
 - Detect that a virtual server is part of Workload not achieving goals
 - Determine that the virtual server performance can be improved with additional resources
 - Project impact on all effected Workloads of moving resources to virtual server
 - If good trade-off based on policy, redistribute resources
 - Initially support CPU management

© Copyright IBM Corporation 2012



Managing Resources across z/VM Virtual Machines

- Manage resources across z/VM virtual machines
 - Detect that a virtual machine that is part of Workload is not achieving goals
 - Determine that the virtual machine performance can be improved with additional resources
 - Project impact on all effected Workloads of moving resources to virtual machine
 - If good trade-off based on policy, redistribute resources
 - Initially support CPU management

© Copyright IBM Corporation 2012

Ensemble WorkLoad Balancing

© Copyright IBM Corporation 2012

Platform Performance Manager

PPM Load Balancing Function

- **Objective: Influence workload balancing decisions across a System z Ensemble**
 - Use SASP protocol to make recommendations for workload balancers (e.g. IP switches / routers that load balance)
 - HMC hosts SASP function
 - Scope of recommendations is non-z/OS virtual servers within the ensemble
 - z/OS Load Balancing Advisor (LBA) provides SASP recommendations for z/OS
 - Same SASP client code can interact with both LBA and HMC SASP implementations to provide complete coverage of z environment
 - HMC recommendations based on the platform performance manager's understanding of the current performance of the members of a load balancing group
 - Recommendation based on overall utilization and delays experienced by virtual servers
 - If IP address and port used to register members of a load balancing group, port is used to determine application availability on each member of load balancing group. Weight of 0 given to members where port is not open

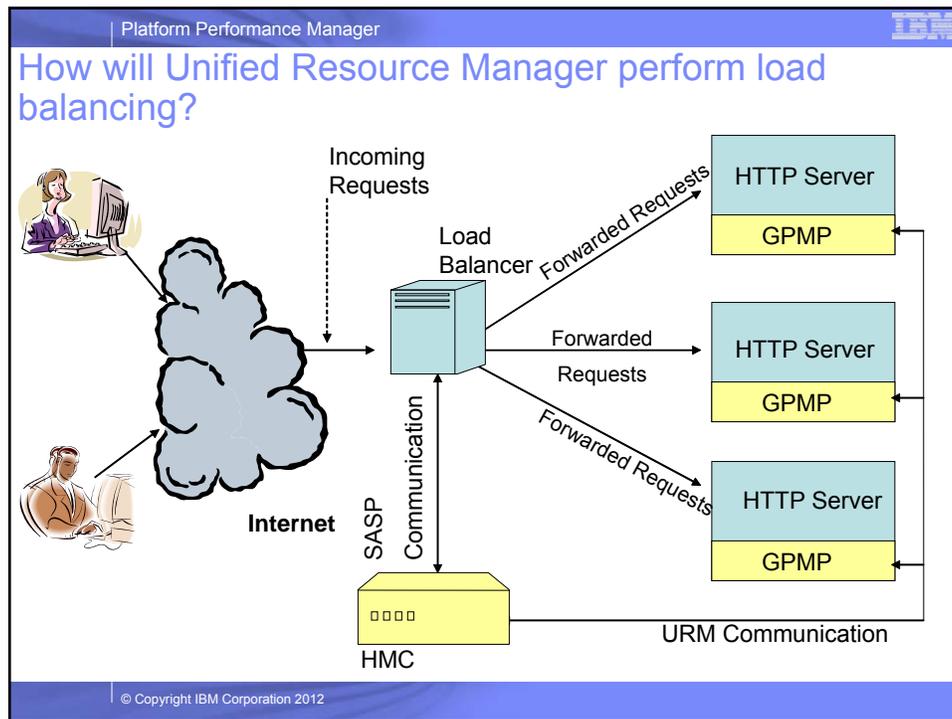
© Copyright IBM Corporation 2012

Platform Performance Manager

Unified Resource Manager View

- **GPMP will sample system statistics**
- **GPMP will know whether application is currently running or not**

© Copyright IBM Corporation 2012

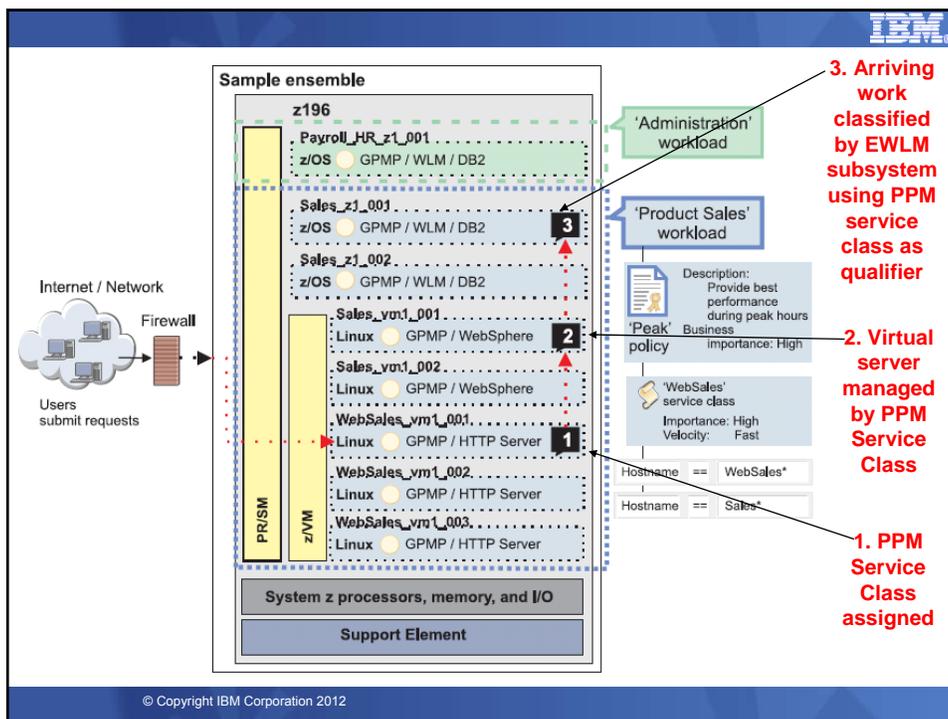


IBM

Co-operative management with z/OS WLM

- z/OS provides differentiated service to PPM classified work
- Transaction coming to z/OS needs to be ARM instrumented via Guest Platform Management Provider (GPMP) implementation
- WLM service definition needs to map PPM service classes to z/OS WLM service classes via EWLM classification rules
- PPM service class associated with transaction is used by WLM to classify work unit to a different WLM service class.
- WLM manages the resources based on the goal assigned to this specific service class.

© Copyright IBM Corporation 2012



Setup for Co-operative Mgmt with z/OS WLM

```

Session B - [24 x 80]
File Edit View Communication Actions Window Help
-----
Subsystem-Type Xref Notes Options Help
-----
Modify Rules for the Subsystem Type      Row 3 to 10 of 16
Command ==>                             Scroll ==> CSR_

Subsystem Type . : EWLM      Fold qualifier names?  Y (Y or N)
Description . . : Rules for testing PPM/GPMP  RJD

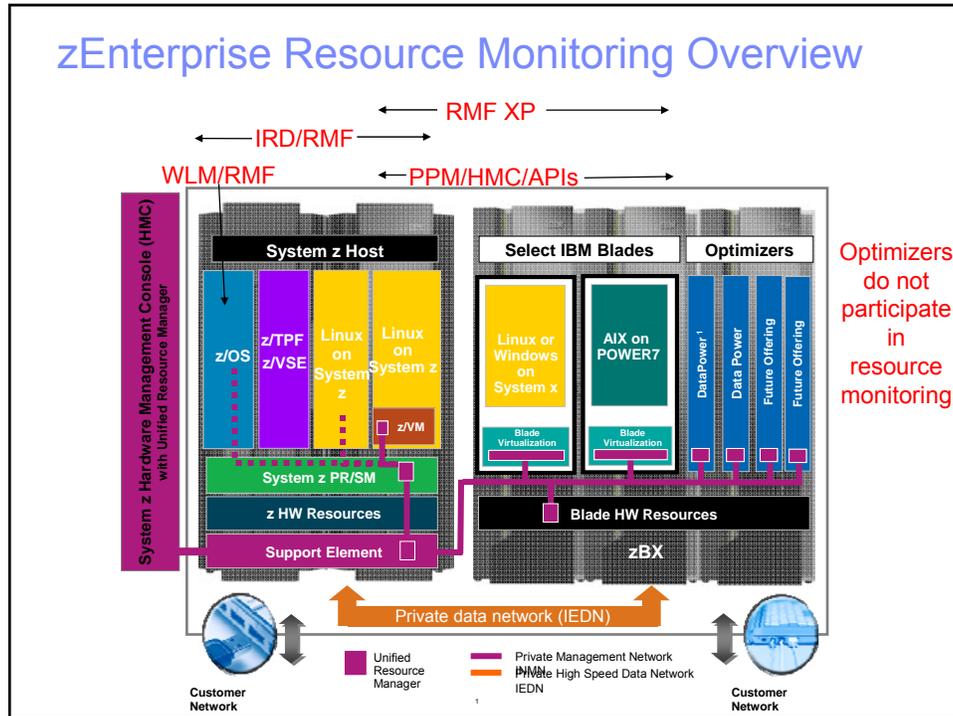
Action codes:  A=After      C=Copy      M=Move      I=Insert rule
                B=Before    D=Delete row R=Repeat  IS=Insert Sub-rule
                More ==>

-----Qualifier-----
Action  Type  Name      Start  Service  Report
-----
1  ESC  SrvClsFo 1  -      -        -
2  ESC  rFastest 9  -      -        -
3  ESC  Highest  17    GPFSTHST
1  ESC  SrvClsFo 1  -      -        -
2  ESC  rFastHig 9  -      -        -
3  ESC  h        17    GPFASHIG
1  ESC  SrvClsFo 1  -      -        -
2  ESC  rModerat 9  -      -        -

-----Class-----
DEFAULTS: EWLMDEFA

M  b                                     21 / 040
-----
Connected to remote server/host 9.12.41.91 using lu/pool TCP00114 and port 23
    
```

© Copyright IBM Corporation 2012



PPM Workload Based Monitoring and Reporting

- Provide reporting capability that shows usage of platform resources in a Workload context within a zEnterprise Ensemble scope
 - Across virtual servers / partitions supporting the Workload
- Workload goal vs actual reporting
- Drill down from overall Workload “performance health” view to contributions of individual virtual server
- Graphical views
 - Topology, trending graphs, etc
- Links to system activity displays to show hardware utilization views
- Reporting is limited to platform level resources, not trying to replicate tools that report on intra-OS resources and performance



Workload Monitoring Overview

- Provide monitoring on the HMC based on a Workload context
- Display of current data and fairly recent history
 - Current stake in the ground is 36 hours of history
 - Interval of data displayed is user selectable
 - Granularity of data kept in repository changes over time
 - 1 minute granularity kept for most recent hour
 - 15 minute interval data kept after first hour

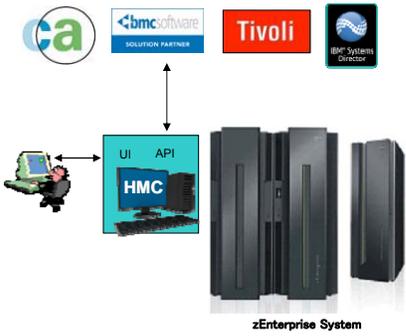
© Copyright IBM Corporation 2012



Unified Resource Manager APIs

Enabling External Management Tools

- New API support allows programmatic access to the same underlying functions exploited by the HMC user interface (UI)
 - ▶ Same resource types, instances and policies
 - ▶ API functions corresponding to views and tasks in the UI
 - Listing resource instances
 - Creating, changing, deleting resource instances
 - Operational control of resource instances
- Access to functions will enable management of Unified Resource Manager from external (to HMC) tools
- Initially the priority scenarios will be the discovery, monitoring, and provisioning use cases



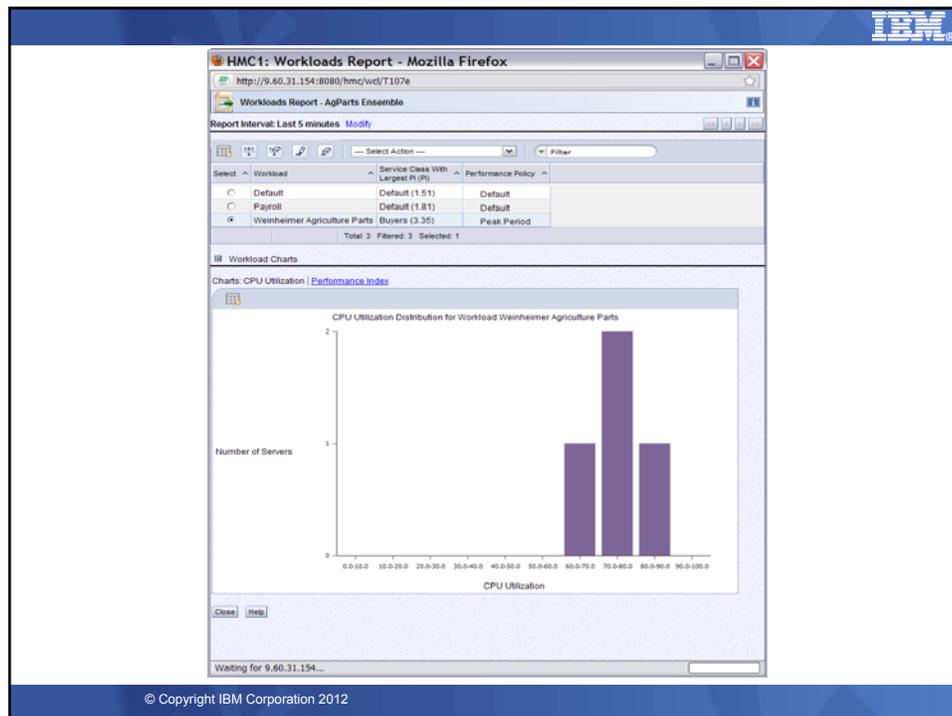
zEnterprise System

© Copyright IBM Corporation 2012 IBM Confidential
© 2012 IBM Corporation

Workload Monitoring Overview...

- Workload Report
 - Display high level view of “performance health” of each Workload
 - Indication if a Workload contains service class missing goals
 - Worst performing service class / performance index
 - Details of specific Workloads
 - Graph of PI of worst performing service class
 - Option to graph other service classes
 - Bar graph of virtual server utilization distribution
 - Visualize view of workload overall load
 - Drill down to Workload’s service class report

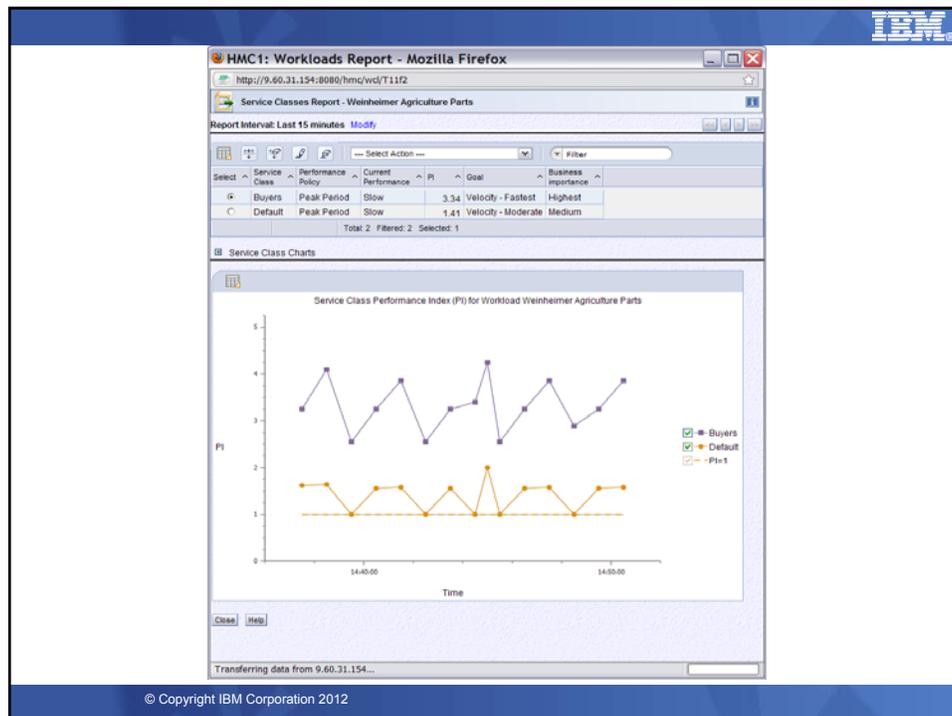
© Copyright IBM Corporation 2012



Workload Monitoring Overview...

- Service Class Report
 - High-level view of each service class in Workload's performance policy
 - Goal and importance
 - Actual performance
 - Indication if monitoring event is established for service class and event is triggered
 - Service class details
 - Graph of service class performance index
 - Drill down to virtual server report for Workload
- Event Monitoring
 - Initial support:
 - Leverage HMC event monitoring
 - Send e-mail when selected metrics reach threshold
 - Service Class PI threshold
 - Virtual Server CPU Utilization threshold

© Copyright IBM Corporation 2012



© Copyright IBM Corporation 2012

IBM

Workload Monitoring Overview...

- **Workload virtual server report**
 - List of virtual servers in a service class
 - Virtual server velocity
 - Resource usage
 - Physical CPU utilization
 - OS view of CPU utilization
 - Physical memory used
 - Hypervisor delay percentage
- **Resource adjustment report**
 - Resource adjustment actions taken over report interval

© Copyright IBM Corporation 2012

IBM

HMC1: Workloads Report - Mozilla Firefox
<http://9.60.31.154:8080/hmc/wcd/T1325>
 Virtual Servers Report - Weisheimer Agriculture Parts
 Report Interval: Last 15 minutes [Modify](#)

Select	Virtual Server	Virtual Processors	Allocated Memory (MB)	Physical CPU Utilization (%)	Hypervisor CPU Delay (%)	Idle Time (%)	Other Time (%)	Service Class With Largest P (P)	OS Processes Total CPU Using Samples (%)	OS Processes Total CPU Delay Samples (%)	OS Processes Total IO Delay Samples (%)	OS Processes Total Page Delay Samples (%)
<input checked="" type="checkbox"/>	Buyer 1	1	1,024	88.7	94.0			Buyers (1.64)				
<input type="checkbox"/>	Buyer 2	1	1,024	83.0	66.0			Buyers (1.64)				
<input type="checkbox"/>	Vendor 1	1	1,024	83.0	66.0			Default (0.76)				
<input type="checkbox"/>	Vendor 2	1	1,024	66.6	66.1			Default (0.76)				
Total: 4, Filtered: 4, Selected: 1												

Virtual Server Charts

CPU Utilization for Virtual Server Buyer 1

The chart shows CPU Utilization (%) on the y-axis (ranging from 25 to 100) over time. The utilization for Buyer 1 fluctuates between approximately 80% and 95%.

© Copyright IBM Corporation 2012

HMC1: Workloads Report - Mozilla Firefox
 http://9.60.31.154:8080/hmc/vcd/71400

Virtual Server Resource Adjustments Report - Buyer 1

Report Interval: Last 15 minutes [Modify](#)

Successful Adjustments

Receiver Virtual Servers	Receiver Workload	Receiver Service Class	Receiver Processing Units After (Before)	Donor Virtual Servers	Donor Workload	Donor Processing Units After (Before)	Time
Buyer 1	Weinheimer Agriculture Parts	Buyers	0.52 (0.50)	Payroll App	Payroll	0.49 (0.50)	Jul 11, 2010 4:13:18 PM
Buyer 1	Weinheimer Agriculture Parts	Buyers	0.52 (0.50)	Vendor 1	Weinheimer Agriculture Parts	0.49 (0.50)	Jul 11, 2010 4:13:18 PM
Total: 2 Filtered: 2							

Failed Adjustments

Receiver Virtual Servers	Receiver Workload	Receiver Service Class	Failure Reason	Time
Total: 0 Filtered: 0				

Done

© Copyright IBM Corporation 2012

R90HMC1: Virtual Servers Report - Mozilla Firefox
 https://9.12.16.241/hmc/vcd/743de48a6e70p_jcb66c66

Processor count: 8 Total CPU consumption: 4.8%
 Total memory allocated for LPARs: 32,768 MB Total memory: 65,536 MB
 Total processor entitlement: 6.84

PPM Hypervisor Report

Virtual Servers

Virtual Server	Processor Management Status	Processor Management Reason	Virtual Processor Count	Consumed Processors	Hypervisor Processing Unit Delay (%)	Allocated Memory (MB)	LPAR Capped	Uncapped Weight	Current Entitled Capacity	Defined Entitled Capacity	Min Entitled Capacity
r901b207v2	Active	None	2	0.01	0.0	4,096	-	128	180	180	10
r901b207v3	Active	None	2	0.02	0.0	4,096	-	128	20	20	20
r901b207v4	Active	None	2	0.01	0.0	4,096	-	128	139	139	10
r901b207v5	Active	None	2	0.01	0.0	4,096	-	128	140	140	10
Total: 8 Filtered: 8											

Successful Adjustments

Receiver Virtual Servers	Receiver Workload	Receiver Service Class	Receiver Processing Units After (Before)	Donor Virtual Servers	Donor Workload	Donor Processing Units After (Before)	Time
r901b207v7	WkldForModerateMedium	SvcClsForModerateMedium	0.45 (0.25)	r901b207v2	Default	1.72 (1.80)	Sep 29, 2010 11:33:19 AM
r901b207v7	WkldForModerateMedium	SvcClsForModerateMedium	0.45 (0.25)	r901b207v4	Default	1.35 (1.39)	Sep 29, 2010 11:33:19 AM
r901b207v7	WkldForModerateMedium	SvcClsForModerateMedium	0.45 (0.25)	r901b207v5	Default	1.36 (1.40)	Sep 29, 2010 11:33:19 AM
r901b207v7	WkldForModerateMedium	SvcClsForModerateMedium	0.45 (0.25)	r901b207v6	Default	1.36 (1.40)	Sep 29, 2010 11:33:19 AM
Total: 4 Filtered: 4							

Failed Adjustments

Receiver Virtual Servers	Receiver Workload	Receiver Service Class	Failure Reason	Time
Total: 0 Filtered: 0				

Done

© Copyright IBM Corporation 2012

IBM

Benefits of GPMP

- Guest Platform Management Provider (GPMP) is a lightweight component of PPM that provides additional monitoring data
- Allows cooperative management with z/OS WLM
- Allows virtual server to be classified using additional attributes such as HostName, SystemName, OS Level etc.
- With instrumented middleware support, GPMP provides metrics that allows detailed transaction topology as transaction hops through heterogeneous platforms in zEnterprise

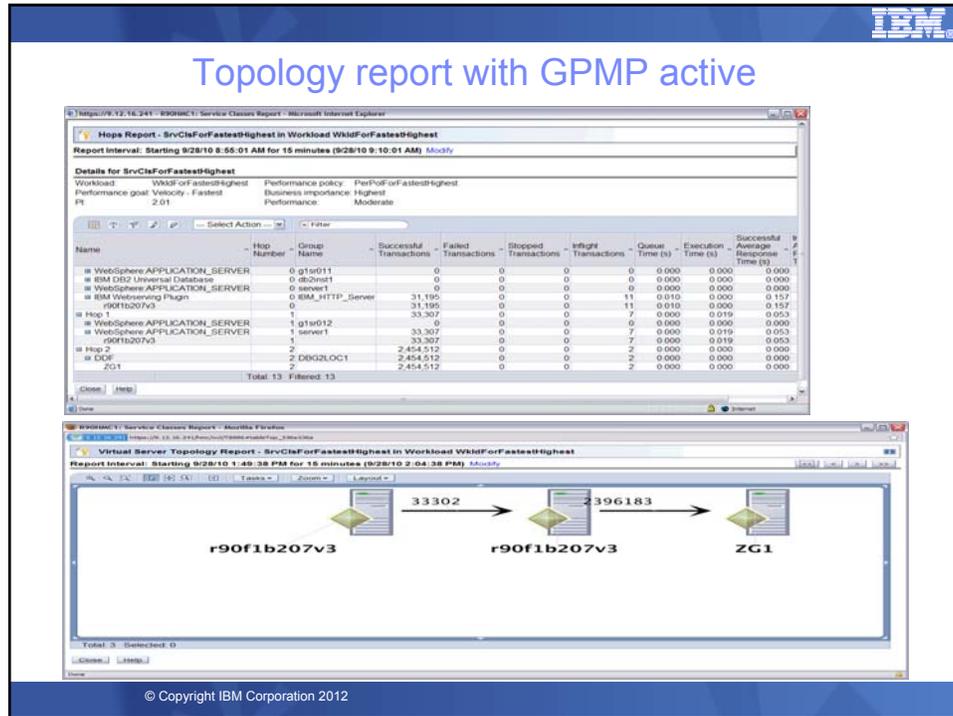
© Copyright IBM Corporation 2012

IBM

Benefits of Middleware instrumentation

- Transaction response time reporting
- Multi-tiered work request flow across environments
- Relationship to server resources being consumed
- Same reasoning lead to instrumentation of z/OS subsystems (CICS, IMS, DB2, etc) for z/OS WLM
- OpenGroup Application Response Measurement (ARM) standards based instrumentation.

© Copyright IBM Corporation 2012



Cross Platform Performance Monitoring with RMF XP

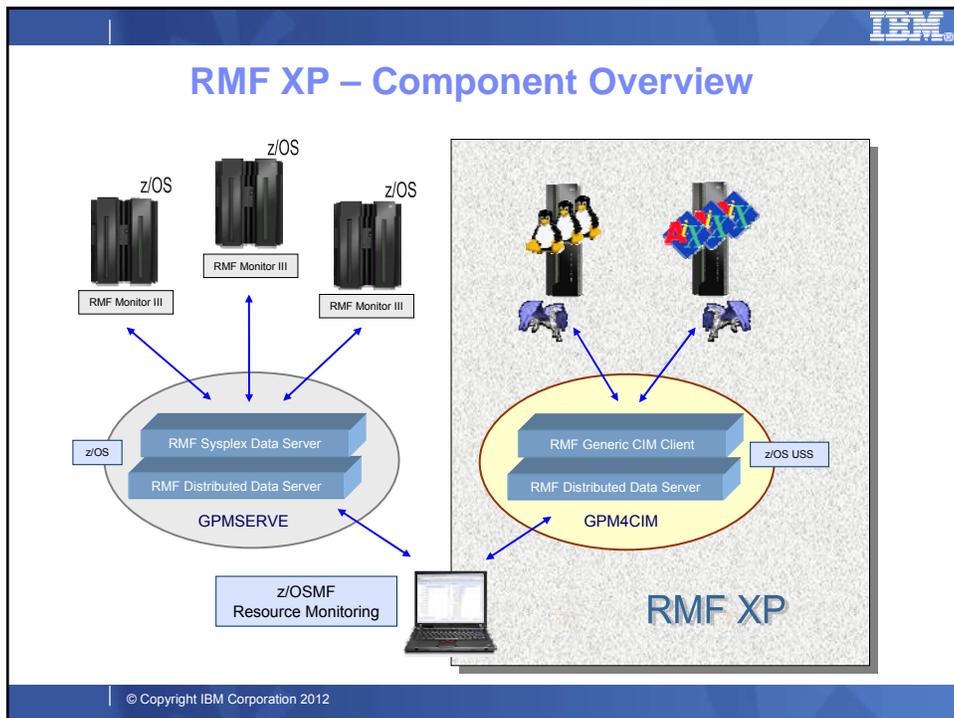
- The Common Information Model (aka CIM) instrumentation is available for almost all operating systems of this planet
- RMF has the infrastructure already in place to
 - combine performance data from multiple systems to a Sysplex wide view
 - display performance data by means of state-of-the-art graphical frontends
- RMF XP brings these two well-proven things together
- RMF XP supports the following operating systems:
 - AIX on System p
 - Linux on System x
 - Linux on System z

IBM

RMF XP

- Seamless performance monitoring solution for z/OS and distributed platforms
- z/OS as management platform for distributed environments
- Easy to setup, almost no customization needed
- Two graphical frontends
 - ▶ Instant access via web browser
 - ▶ z/OSMF with advanced capabilities
- zIIP exploitation helps to reduce costs
- Available with z/OS V1R13 RMF and z/OS V1R12 RMF (APAR OA36030)

© Copyright IBM Corporation 2012



IBM

Invocation

- Started Task: SYS1.PROCLIB(GPM4CIM)
- Runs in USS Environment via BPXBATCH
- Multiple instances can run in parallel: one STC per platform
 - ▶ S GPM4CIM.GPM4A,OS=A
 - ▶ S GPM4CIM.GPM4X,OS=X
 - ▶ S GPM4CIM.GPM4Z,OS=Z

```

//GPM4CIM PROC OS=X
//STEP1 EXEC PGM=BPXBATCH, TIME=NOLIMIT, REGION=OM,
// PARM=' PGM /usr/lpp/gpm/bin/gpm4cim cfg=/etc/gpm/gpm40S..cfg'
//STDENV DD PATH=' /etc/gpm/gpm4cim.env'
//STDOUT DD PATH=' /var/gpm/ogs/gpm4cim&OS..out' ,
// PATHOPTS=(OWRONLY, OCREAT, OTRUNC) ,
// PATHMODE=(SI RUSR, SI WUSR, SI RGRP)
//STDERR DD PATH=' /var/gpm/ogs/gpm4cim&OS..trc' ,
// PATHOPTS=(OWRONLY, OCREAT, OTRUNC) ,
// PATHMODE=(SI RUSR, SI WUSR, SI RGRP)
//SYSPRINT DD SYSOUT=*
//SYSOUT DD SYSOUT=*
// PEND
            
```

© Copyright IBM Corporation 2012

IBM

Resource Models

AIX_SYSTEM_COMPLEX

- ACTIVE_MEMORY_EXPANSION
- ACTIVE_MEMORY_SHARING
- ALL_DISKS
- DISK
- ALL_LOGICAL_PROCESSORS
- LOGICAL_PROCESSOR
- ALL_NETWORK_PORTS
- NETWORK_PORT
- ALL_LOCAL_FILE_SYSTEMS
- LOCAL_FILE_SYSTEM
- ALL_PROCESSES
- PROCESS
- ALL_SHARED_ETHERNET_ADAPTERS
- SHARED_ETHERNET_ADAPTER
- ALL_VIRTUAL_TARGET_DEVICES
- VIRTUAL_TARGET_DEVICE
- PARTITION
- MEMORY

XLINUX_SYSTEM_COMPLEX

- XLINUX_IMAGE
- ALL_LOCAL_FILE_SYSTEMS
- LOCAL_FILE_SYSTEM
- ALL_IP_PROTOCOL_ENDPOINTS
- IP_PROTOCOL_ENDPOINT
- ALL_LOGICAL_PROCESSORS
- LOGICAL_PROCESSOR
- ALL_NETWORK_PORTS
- NETWORK_PORT
- ALL_PROCESSES
- PROCESS
- ALL_KVM_GUESTS
- KVM_GUEST
- ALL_XEN_GUESTS
- XEN_GUEST

ZLINUX_SYSTEM_COMPLEX

- CEC
- LPAR
- ZLINUX_IMAGE
- ALL_LOCAL_FILE_SYSTEMS
- LOCAL_FILE_SYSTEM
- ALL_IP_PROTOCOL_ENDPOINTS
- IP_PROTOCOL_ENDPOINT
- ALL_LOGICAL_PROCESSORS
- LOGICAL_PROCESSOR
- ALL_NETWORK_PORTS
- NETWORK_PORT
- ALL_PROCESSES
- PROCESS
- ALL_CHANNELS
- CHANNEL
- ALL_VOLUMES
- VOLUME

© Copyright IBM Corporation 2012

RMF XP – Resource Tree

The screenshots illustrate the RMF Performance Data Portal interface. The top-left screenshot shows the main portal with a 'Welcome, you are connected to: WEBPLEX_AIX_SYST' message and a table of resources under the path 'Children of: WEBPLEX_AIX_SYSTEM_COMPLEX'. The top-right screenshot provides a detailed view of this resource tree, listing resources like 'IMMO-123-131.AIX_IMAGE' and their attributes. The bottom screenshot shows the 'Children of: tmcc-123-141, ALL_NETWORK_PORTS' resource tree, listing network-related resources such as 'IMMO-123-141.640.NETWORK_PORT'.

© Copyright IBM Corporation 2012

RMF XP – z/OSMF Integration

The screenshot displays the IBM z/OSMF Management Facility interface. It features a 'Resource Monitoring' section with a 'WEBPLEX Health Check (Running)' status. The interface includes several performance charts: 'Overall Processor Utilization' showing CPU usage for various nodes; 'Memory Usage' with bar charts for memory consumption; 'Process CPU Consumption' showing CPU usage for individual processes; and 'File System Space' monitoring. The bottom of the screen shows a timeline for the health check results.

© Copy

