Agenda

- zEnterprise Workload Management
  - z196 / z114
    - z/OS Virtual Servers - WLM
    - IRD Clusters
  - 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
System z Hardware Management Console (HMC) with Unified Resource Manager

- Select IBM Blades
- Optimizers

- z/OS
- z/VSE
- Linux on System z
- Linux on System z

- System z PR/SM
- z HW Resources
- Support Element

- Blade HW Resources
- zBX

Private data network (IEDN)

Customer Network

Unified Resource Manager

Private Management Network INMN

Private High Speed Data Network IEDN

zEnterprise Workload Management
z/OS Workload Manager (WLM)

- A contract between the installation and the z/OS operating system

- Installation
  - Classifies work running on z/OS in distinct Service Classes
  - Defines goals that express the expectation of how work should perform

- WLM
  - Uses goal definitions to manage work across all systems of a sysplex through distribution of resources
z/OS WLM Workload Mgmt in an LPAR

Intelligent Resource Director (IRD)

- What is IRD?
  - Set of functions that distribute CPC resources based on business importance
  - Problem areas being addressed:
    - Workloads may change over the course of a day, week, month, year...
    - Business priorities may change over course of a day, week, month, year...
      - E.g., online vs. batch, production vs. test, workload fluctuations, or periodic work
    - Single static configuration may be sub-optimal to handle different workload mixes
      - Distribute resources based on workload and service level agreements (WLM goals)
    - Reliability problems, e.g. caused by single points of failure

- Consists of
  - LPAR CPU Management
    - LPAR Weight Management
    - LPAR Vary CPU Management
  - Dynamic Channel Path Management (DCM)
  - Channel Subsystem Priority Queuing (CSSPQ)

(c) IBM Corporation 2011
LPAR Clusters

- Scope of IRD management is the LPAR Cluster (LPC)
- LPC = "Set of LPARs on same CPC, which are part of same Sysplex"
  - For CPU management of zLinux systems, specify the sysplex name as CP management cluster name on Customize Image Profiles panel
  - For DCM, IOCLUSTER keyword in the IODF must match Sysplex name.
- LPARs with dedicated CPs can still join a cluster
  - But will not be enabled for WLM LPAR Weight and Vary CPU Management.
- Multi-image/Sysplex LPCs require a CF structure (except for CSSPQ)

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
  - Workload monitoring and reporting based on management policy
  - Common approach to monitoring / management of platform resources across zEnterprise
  - Extend goal oriented approach of z/OS WLM to platform managed resources
  - Orchestration of autonomic management of resources across virtual servers
    - Provide Intelligent Resource Director like function across the zEnterprise
    - Management function will evolve over time
    - 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
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.

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.
zManager Workload Mgmt Functions

- **z/VM and PowerVM Hipervisors**
  - Virtual Server CPU Management provides the ability to manage CPU resources across virtual servers based on a goal-oriented performance policy.

- **System x (KVM based) Hipervisor**
  - Does currently participate in dynamic resource management

- **PR/SM Hipervisor**
  - Does not make resource management adjustments based on PPM Policy. Only IRD dynamically influences the PR/SM hipervisor

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
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

Elements of a Service Class

- **Performance Goal**
  - 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
Performance Policies

- **'Product Sales' workload**
  - **Performance policies**
    - 'Default'
    - 'Off-peak'
    - 'Peak'
  - **Service classes**
    - Default: Importance: Medium, Velocity: Moderate
    - StoreSales: Importance: Highest, Velocity: Fastest
    - WebSales: Importance: High, Velocity: Fast
  - **Classification rules**
    - Hostname == StoreSales
    - Hostname == WebSales

- **Description**
  - Provide best performance during peak hours
  - Business importance: High

(c) IBM Corporation
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 affected Workloads of moving resources to virtual server
  - If good trade-off based on policy, redistribute resources
  - Initially support CPU management

CPU Management causes the hypervisor to move processing units (pu) from the donor to the receiver virtual server.
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

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.
1. PPM Service Class assigned
2. Virtual server managed by PPM Service Class
3. Arriving work classified by EWLM subsystem using PPM service class as qualifier

Setup for co-operative mgmt with z/OS WLM
zEnterprise eXposed Part 3: zManager and z/OS Workload Manager (Brad Snyder)

Tuesday, 3:00pm
zEnterprise Resource Monitoring Overview

- **System z Host**: z/OS, z/TPF, Linux on System z
- **System z PR/SM**: z/VM
- **z HW Resources**: zBX, Blade HW Resources
- **Optimizers**: IBM Smart Analytics Optimizer, z HW Resources, z/OS, Linux on System z, z/TPF, z/VM
- **Future Offering**: Blade Virtualization

Optimizers do not participate in resource monitoring.

Private data network (IEDN)

Customer Network
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
Application Programming Interfaces (APIs) to Unified Resource Manager (July 12, 2011 announcements)

IBM intends to offer Application Program Interfaces (APIs) to Unified Resource Manager. These APIs provide access to the same underlying functions that support the Unified Resource Manager user interface and can be exploited to enable Discovery, Monitoring, and Provisioning use cases.

IBM intends to extend the Tivoli® Integrated Service Management for System portfolio to take advantage of the zEnterprise ensemble monitoring and management capabilities provided by the Unified Resource Manager APIs.

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
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
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
PPM Processor Mgmt results (SC Report)
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.
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.

Topology report with GPMP active
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
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)

RMF XP – Component Overview
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
//STEP2 EXEC PGM=BPXBATCH,TIME=NOLIMIT,REGION=0M,
//       PARM='PGM /usr/lpp/gpm/bin/gpm4cim cfg=/etc/gpm/gpm4cim.cfg'
//STEP2 DD PATH=('/tmp/gpm4cim.log',PATHOPTS=(OWRONLY,OCREAT,OTRUNC),
//       PATHMODE=(SIRUSR,SIWUSR,SIRGRP)),
//STEP2 DD PATH=('/tmp/gpm4cim.err',PATHOPTS=(OWRONLY,OCREAT,OTRUNC),
//       PATHMODE=(SIRUSR,SIWUSR,SIRGRP)),
//S GPM4CIM.GPM4A,OS=A
//S GPM4CIM.GPM4X,OS=X
//S GPM4CIM.GPM4Z,OS=Z
```
**z/OS Tuning Courses from IBM Training**

- **Basic z/OS Tuning Using the Workload Manager (WLM)**
  - ES545
  - 4.5 Days, Hands-on Lab Exercises

- **Advanced z/OS Performance: WLM, Sysplex, Unix Services, Web**
  - ES851
  - 4.5 Days

- [ibm.com/training](http://ibm.com/training)