Workload Manager Technical Overview
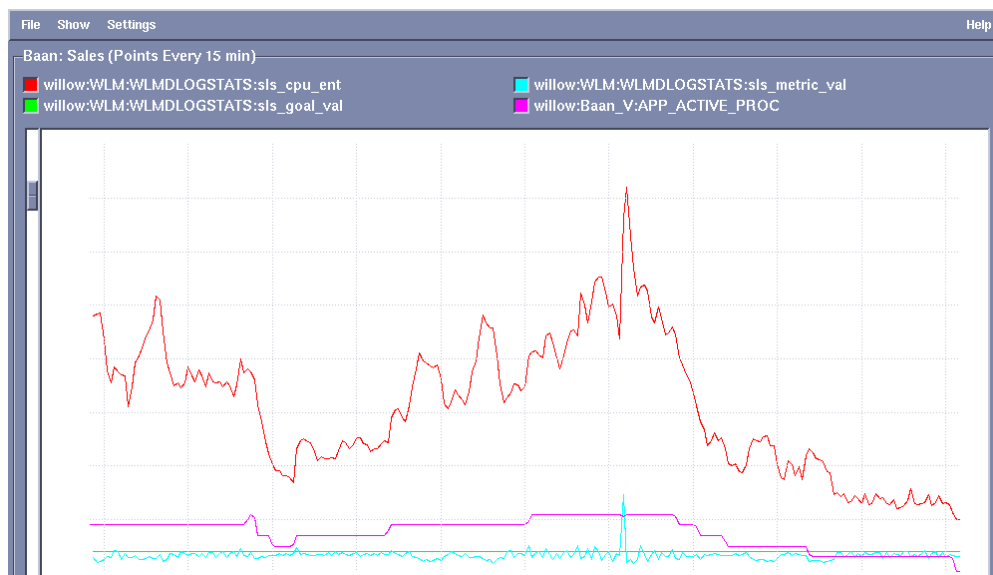
Siegfried Becker
Presales Consultant

Target Problem

- Handling Peaks in Load on Mission Critical Applications
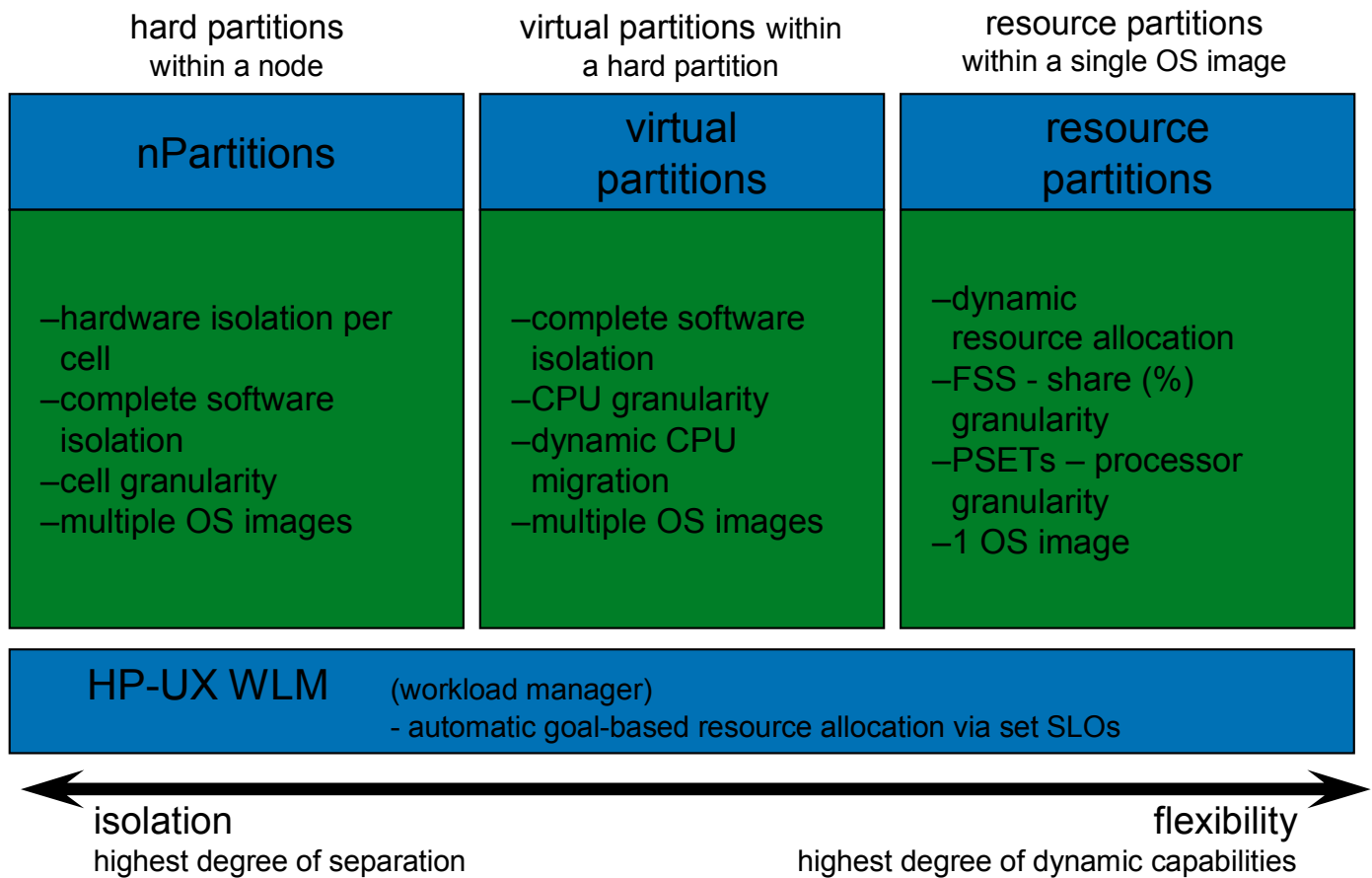
Traditional Approach

- ## Overprovisioning
  - Lots of dedicated Unix servers
  - Excess capacity on each
  - Gartner states that the average IT organization utilizes their infrastructure at approximately 35% of capacity

- ## Drawbacks
  - Cost of underutilized capacity
  - Difficult to manage many systems

New Solutions
The Adaptive Infrastructure

- ## Dynamically reconfigurable partitions
  - nPars with iCOD
  - Virtual Partitions
  - Resource Partitions

- ## Capacity on Demand
  - iCOD
  - iCOD Temporary Capacity (TiCOD)
  - Pay Per Use (PPU)

- ## Application Consolidation
  - Run multiple workloads on a single Unix system

- ## Spare Capacity Consolidation
  - Provide spare capacity for multiple apps on the same

HP Partitioning Continuum

hard partitions within a node	virtual partitions within a hard partition	resource partitions within a single OS image
nPartitions	**virtual partitions**	**resource partitions**
–hardware isolation per cell –complete software isolation –cell granularity –multiple OS images	–complete software isolation –CPU granularity –dynamic CPU migration –multiple OS images	–dynamic resource allocation –FSS - share (%) granularity –PSETs – processor granularity –1 OS image

HP-UX WLM   (workload manager)
- automatic goal-based resource allocation via set SLOs

← **isolation**
highest degree of separation

flexibility →
highest degree of dynamic capabilities

HP Workload Manager

- ## HP WLM is a state-of-the-art dynamic workload manager for HP-UX servers
  - It automatically adapts the partition configuration based on the loads on the applications running in those partitions and your business priorities
  - Supports:
    - Resource partitions and vPars
    - Automatic activation/deactivation of iCOD and pay-per-use CPUs
    - Resource partition memory reallocation when workloads are activated/deactivated due to failover or batch job activation
- ## WLM helps you comfortably increase utilization while still ensuring that your mission critical applications maintain their performance requirements

WLM Features

- Goal based allocation of CPU between partitions
- Automatic reallocation of resources on ServiceGuard failover
- Automatic activation and deactivation of iCOD and PPU processors
- Automatic movement of CPUs between vPars to meet goals
- Out-of-the-box integration toolkits for Oracle, GlancePlus, Job Duration/SAS, Apache/Java and PPU meter

WLM 2.1- Major New Features

- Itanium Support on HP-UX 11i version 2
- Automatic PSET CPU Migration
- BEA Weblogic toolkit to collect load metrics from Weblogic
- Monitoring GUI – graphing of WLM allocation of resources and actual utilization by workloads
- Auditing (billing) utilities – utilities that accumulate the actual usage of resources by each workload over time, csv formatted for upload to your favorite billing package
- Advisory mode - to allow customers to monitor their workloads without turning on WLM controls
- Transient group support – Resource partitions are created when an application starts (eg. on failover, or batch job

New Features in WLM 2.2

- Support for goal-based CPU allocation across nPars using available iCOD processors
- Support for Temporary iCOD activation/deactivation
- Remote Monitoring GUI
- Remote Configuration GUI
- Support for PSET based transient groups
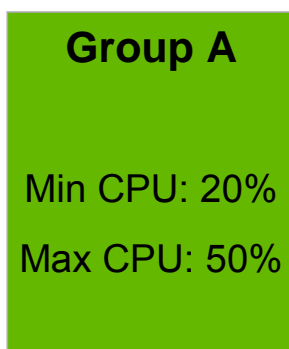- Configuration Wizard enhancements

WLM Service Level Objectives

SLO's use goals, constraints, and conditions.

An SLO consists of:

- **A workload (partition)**
- **Constraints (min, max cpu)**
- **A goal**
- **Priority**
- **Conditions (time of day, event, etc)**

Group A

Min CPU: 20%

Max CPU: 50%

Group A receives 3 shares for each additional user.

Policy applies 9am to 5pm AND

WLM goal types

- Any of the following can be used to allocate resources to a workload:
  - resource utilization
    - CPU entitlement based on utilization of current entitlement
  - direct measurement of the performance of the workload
    - response time
    - throughput
  - measurement of load on application
    - number of users/processes
    - queue length
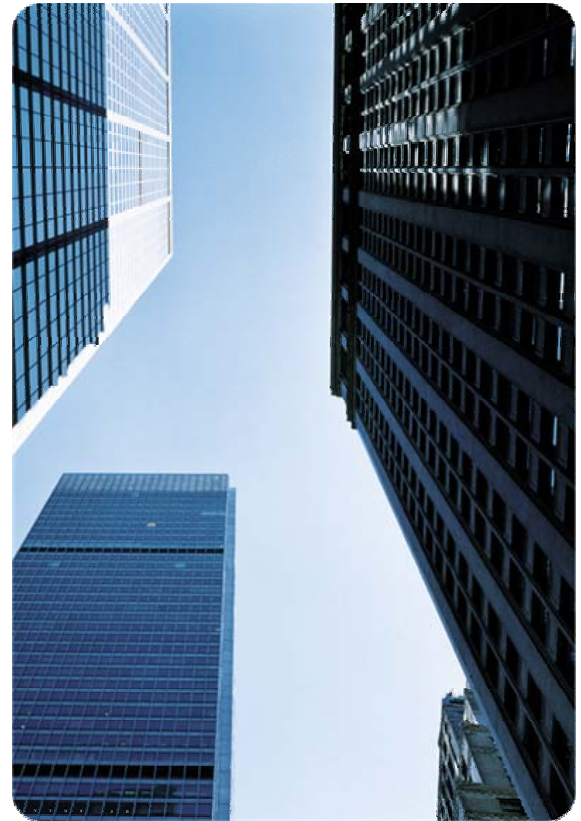
Multi-Partition WLM Architecture

WLM Data Collector Interfaces
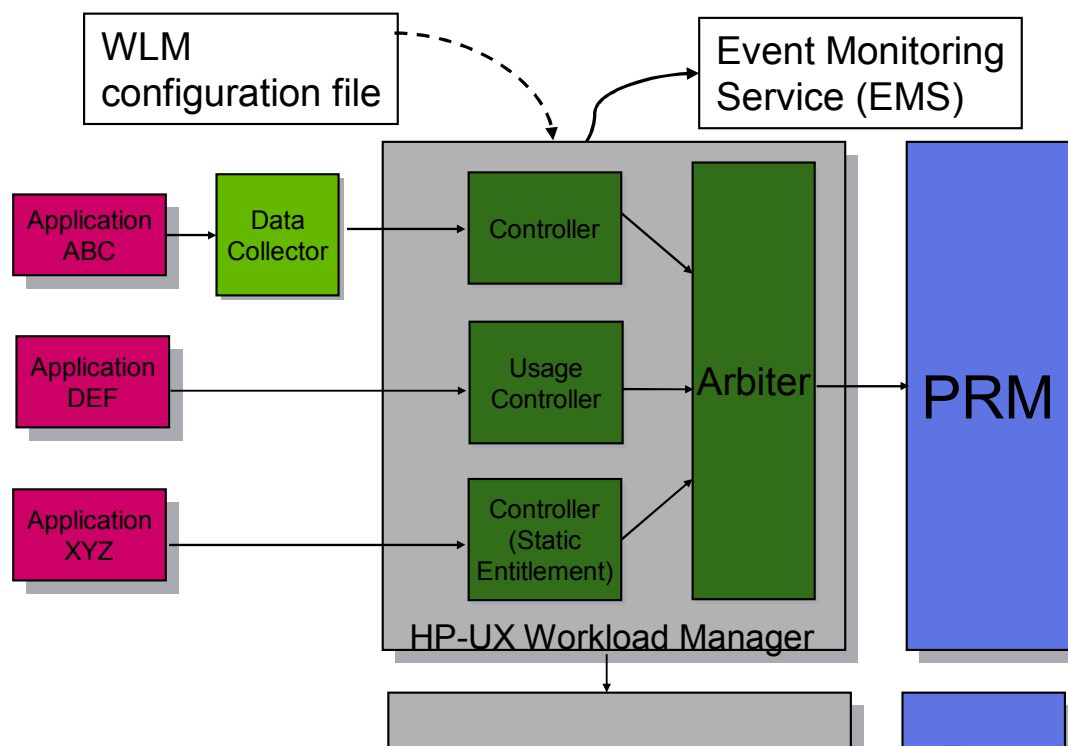
- Usage Controller (requires no data)
- ISV Toolkits
  - Oracle, WebLogic, SAS, Apache/Java, SNMP
- Scripting Interface
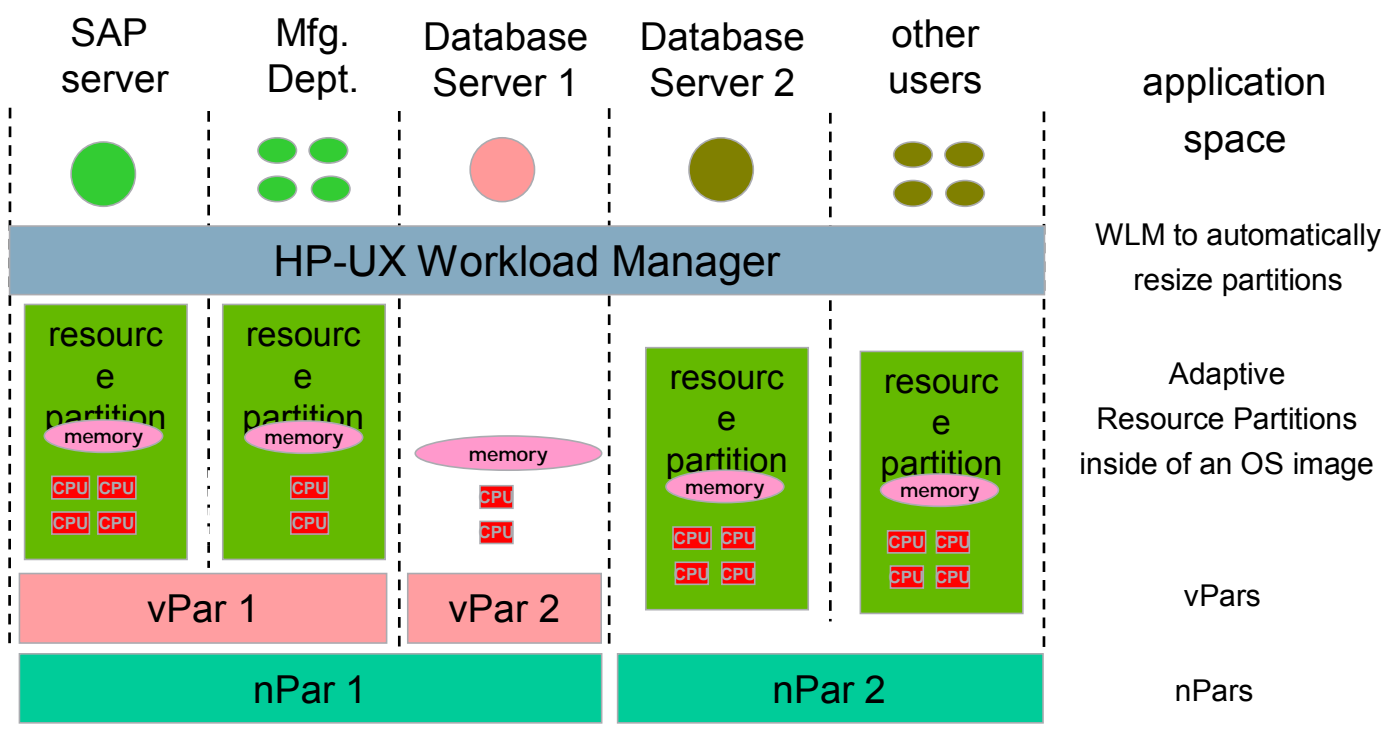- GlancePlus Toolkit
- ARM Toolkit
- WLM API

ISV Toolkits

- We are developing toolkits for databases and major applications
- These will allow customers to quickly implement policies most appropriate to each app right out of the box
- Current Toolkits:
  - Oracle Database
  - BEA WebLogic
  - Job Duration – SAS
  - Apache/Java/Oracle Apps
  - SNMP/Pay Per Use
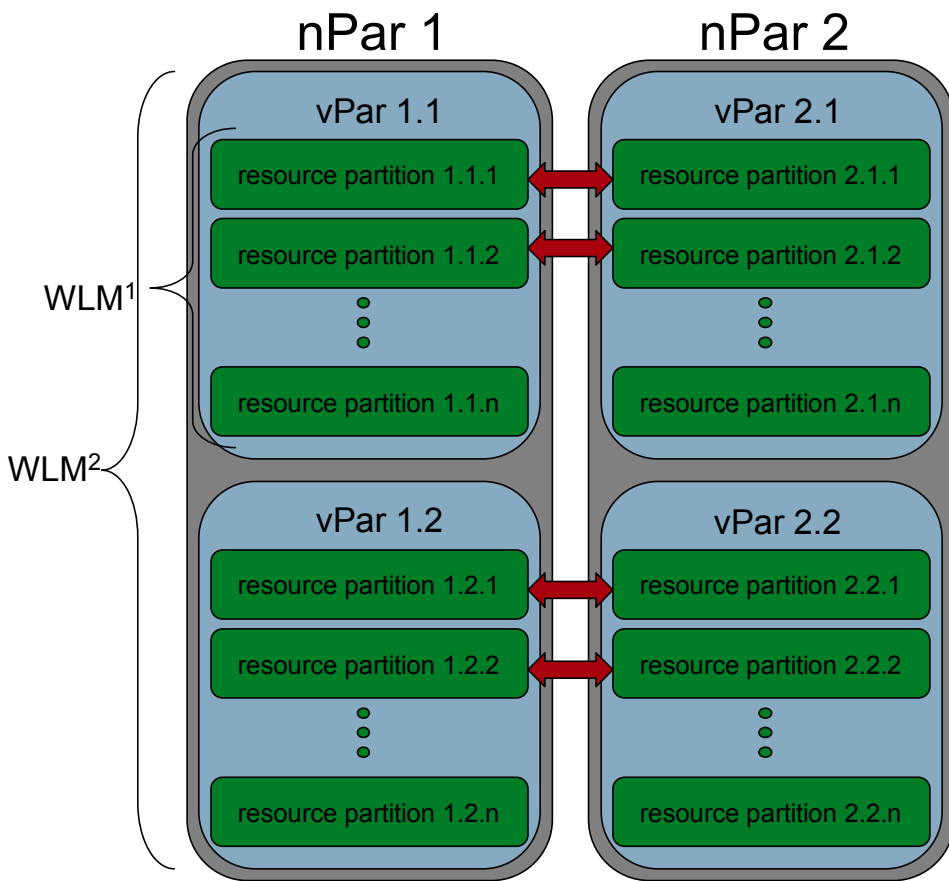
Self-Tuning Utility Computing

- WLM will first attempt to allocate resources between resource partitions to meet SLOs

- If needed, WLM can migrate CPUs between vPars to meet SLOs

- Finally, if critical SLOs are still not being met, WLM can activate Utility CPUs

- Resources are released as soon as the load decreases
  - ensuring that utility costs are minimized
  - ensuring resources are not reserved for high priority workloads that are not using them

ALL AUTOMATICALLY
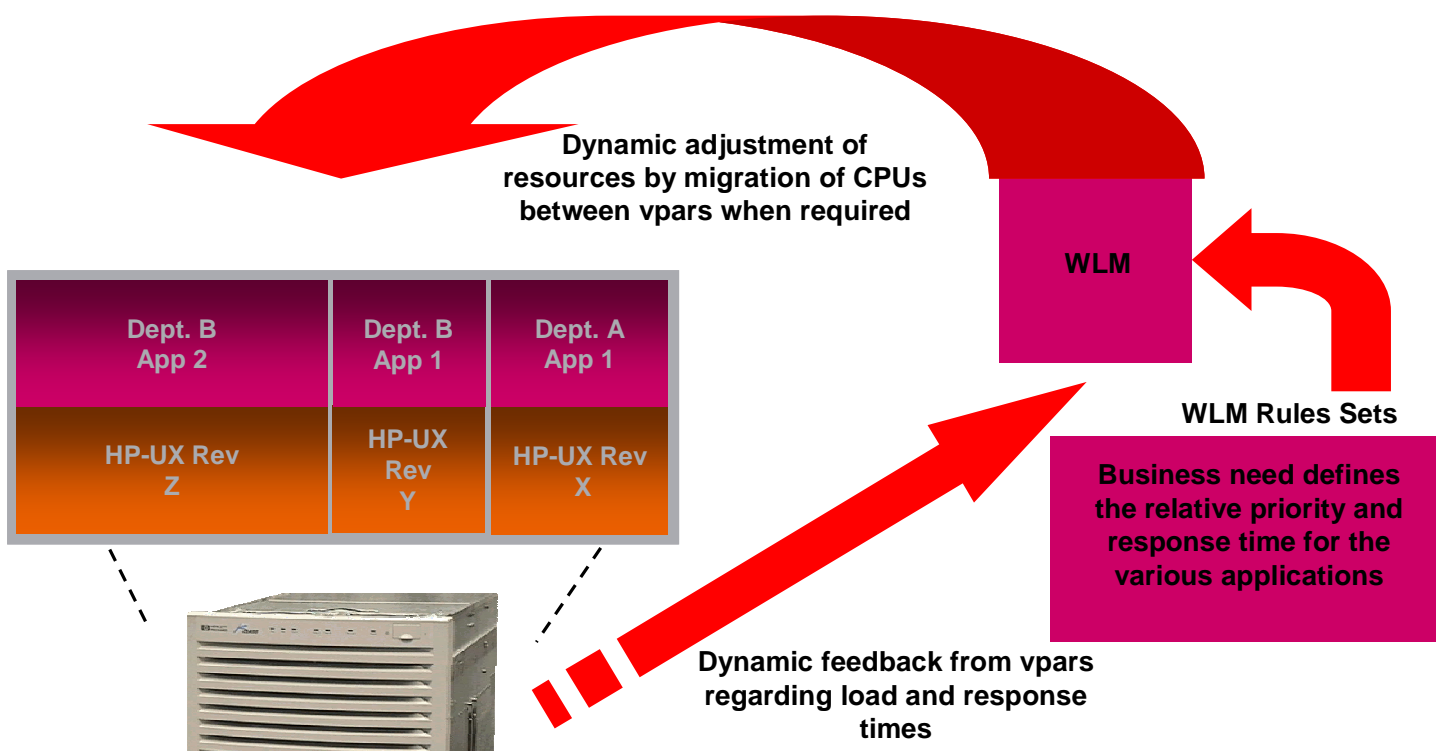
HP-UX = Broadest Partitioning Portfolio

| SAP server | Mfg. Dept. | Database Server 1 | Database Server 2 | other users | application space |

HP-UX Workload Manager — WLM to automatically resize partitions

resource partition memory CPU CPU CPU CPU — Adaptive Resource Partitions inside of an OS image

resource partition memory CPU CPU

memory CPU CPU

resource partition memory CPU CPU CPU CPU

resource partition memory CPU CPU CPU CPU

vPar 1   vPar 2 — vPars

nPar 1   nPar 2 — nPars

adaptive infrastructure

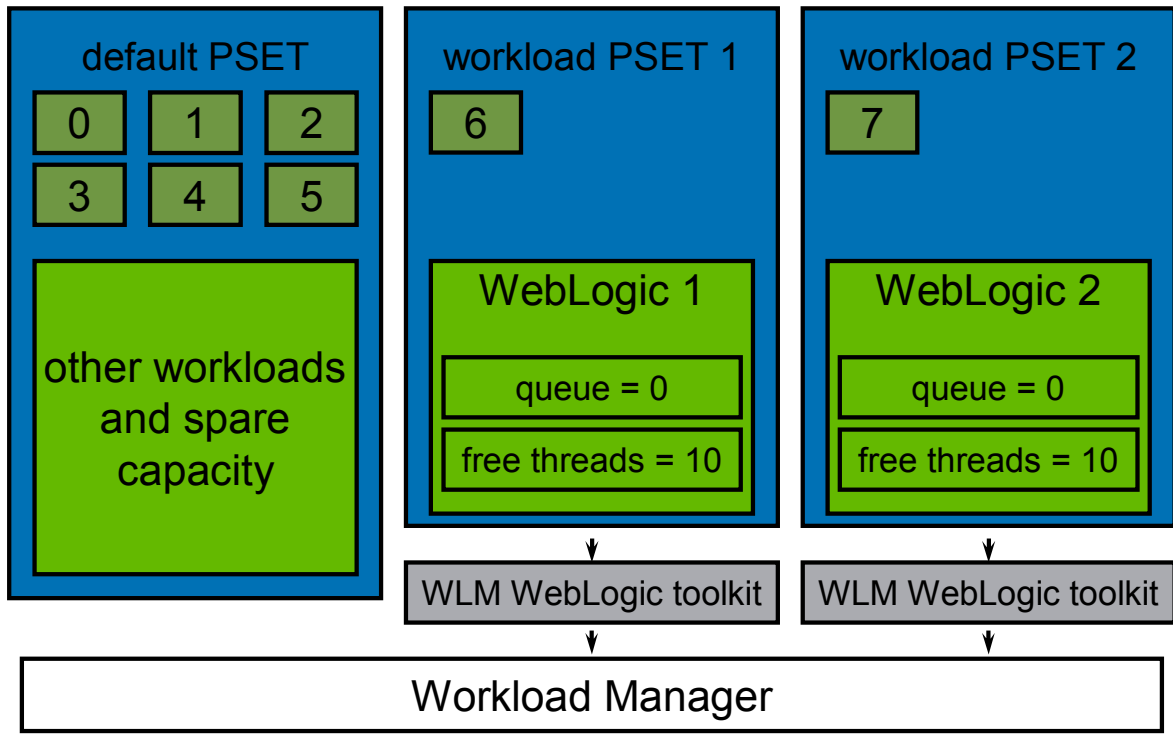| nPar 1	nPar 2
vPar 1.1	**vPar 2.1**
resource partition 1.1.1	resource partition 2.1.1
resource partition 1.1.2	resource partition 2.1.2
resource partition 1.1.n	resource partition 2.1.n
vPar 1.2	**vPar 2.2**
resource partition 1.2.1	resource partition 2.2.1
resource partition 1.2.2	resource partition 2.2.2
resource partition 1.2.n	resource partition 2.2.n

WLM^1
WLM^2

- 2 nPars provides
  - hardware fault isolation
- 2 vPars within each nPar provides
  - software fault isolation
  - OS version isolation
- Any number of resource partitions (one for each major application, or group of same priority minor applications) in each vPar provides:
  - resource isolation
- WLM^1 automatically allocates CPU resources as needed to resource partitions
- WLM^2 automatically allocates CPUs as needed to vPars
- Failover across nPar boundary (indicated by ⟷) provides
  - HA for both hardware and software faults
  - WLM will reallocate resources upon failover

---

vPars with Workload Manager
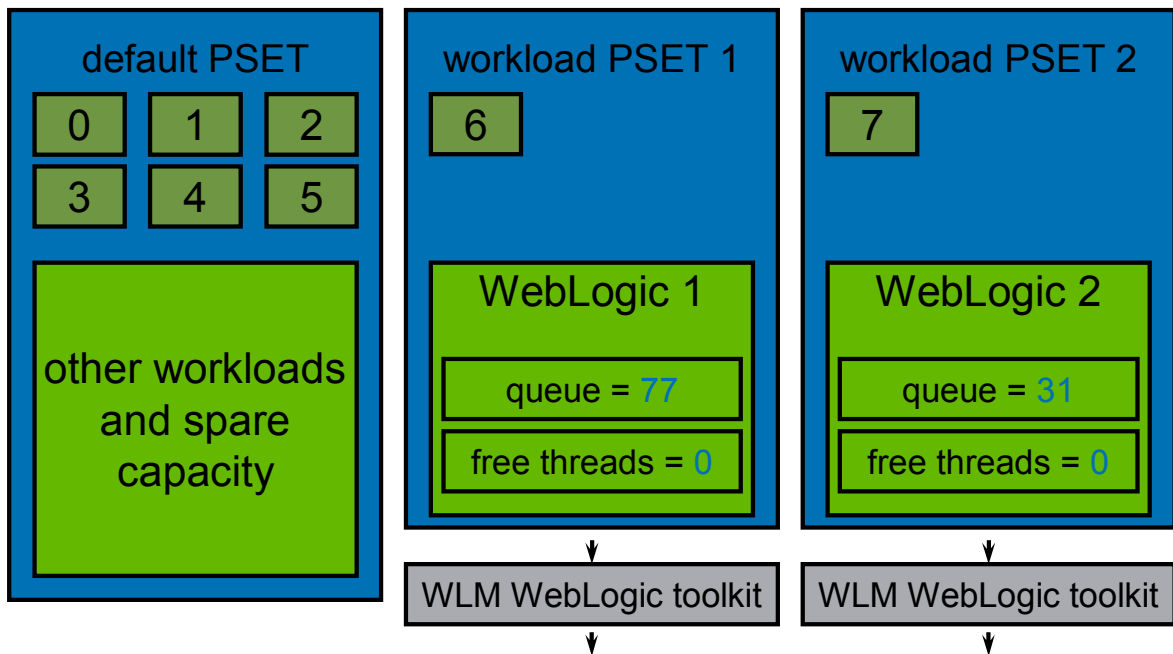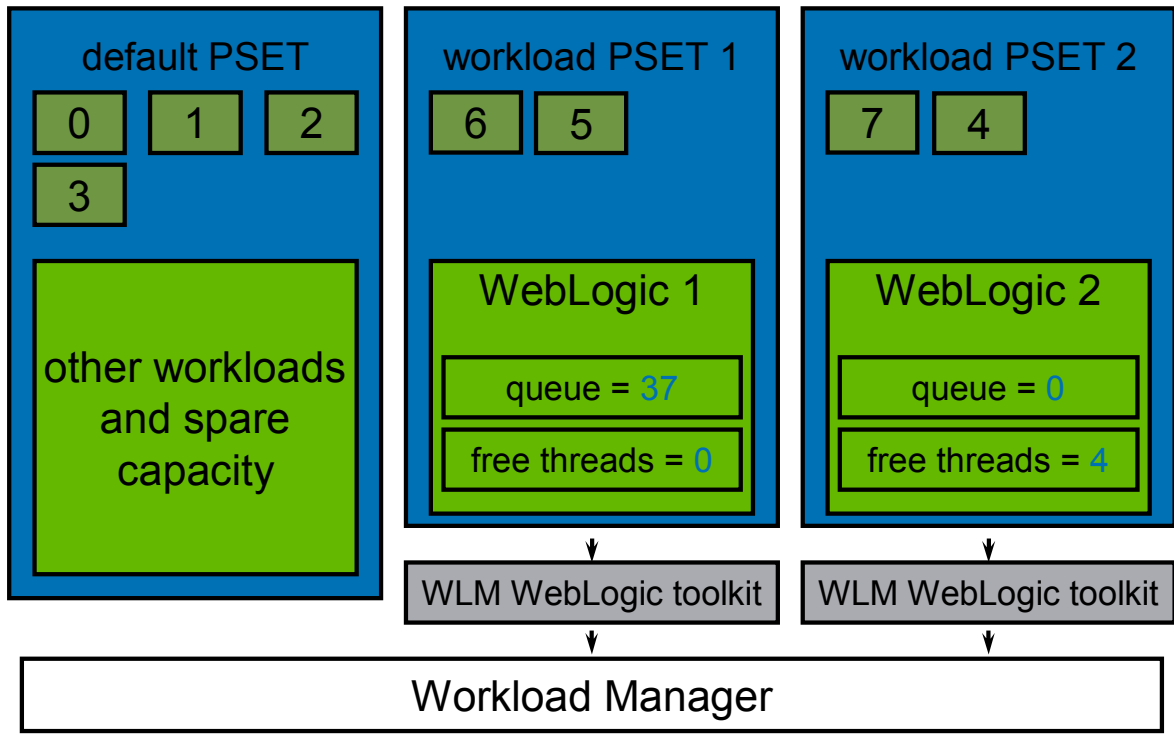## Automatic goal-based CPU allocation between vPars

Dynamic adjustment of resources by migration of CPUs between vpars when required

WLM

WLM Rules Sets

Business need defines the relative priority and response time for the various applications

Dept. B App 2	Dept. B App 1	Dept. A App 1
HP-UX Rev Z	HP-UX Rev Y	HP-UX Rev X

Dynamic feedback from vpars regarding load and response times

WLM Optimization of Resources

Superdome partition

default PSET

| 0 | 1 | 2 |

| 3 | 4 | 5 |

other workloads and spare capacity

workload PSET 1

6

WebLogic 1

queue = 0

free threads = 10

WLM WebLogic toolkit

workload PSET 2

7

WebLogic 2

queue = 0

free threads = 10

WLM WebLogic toolkit

Workload Manager

---

hp invent

Load is added to each WebLogic

Superdome partition

default PSET

| 0 | 1 | 2 |

| 3 | 4 | 5 |

other workloads and spare capacity

workload PSET 1

6

WebLogic 1

queue = 77

free threads = 0

WLM WebLogic toolkit

workload PSET 2

7

WebLogic 2

queue = 31

free threads = 0

WLM WebLogic toolkit

WLM allocates additional CPUs

Superdome partition

default PSET
| 0 | 1 | 2 |

| 3 |

other workloads
and spare
capacity

workload PSET 1
| 6 | 5 |

WebLogic 1

queue = 37

free threads = 0

WLM WebLogic toolkit

workload PSET 2
| 7 | 4 |

WebLogic 2

queue = 0

free threads = 4

WLM WebLogic toolkit

Workload Manager

Load decreases on WebLogic 2

Superdome partition

default PSET
| 0 | 1 | 2 |

other workloads
and spare
capacity

workload PSET 1
| 6 | 5 | 3 |

WebLogic 1

queue = 0

free threads = 2

WLM WebLogic toolkit

workload PSET 2
| 7 | 4 |

WebLogic 2

queue = 0

free threads = 11

WLM WebLogic toolkit

Superdome partition

default PSET
0 1 2

other workloads and spare capacity

workload PSET 1
6 5 3

WebLogic 1
queue = 0
free threads = 2

WLM WebLogic toolkit

workload PSET 2
7 4

WebLogic 2
queue = 0
free threads = 2

WLM WebLogic toolkit

Workload Manager

WLM moves a CPU back to the "free" pool
Used for other workloads

Superdome partition

default PSET
0 1 2
4

other workloads and spare capacity

workload PSET 1
6 5 3

WebLogic 1
queue = 0
free threads = 4

WLM WebLogic toolkit

workload PSET 2
7

WebLogic 2
queue = 0
free threads = 3

WLM WebLogic toolkit

The Solution
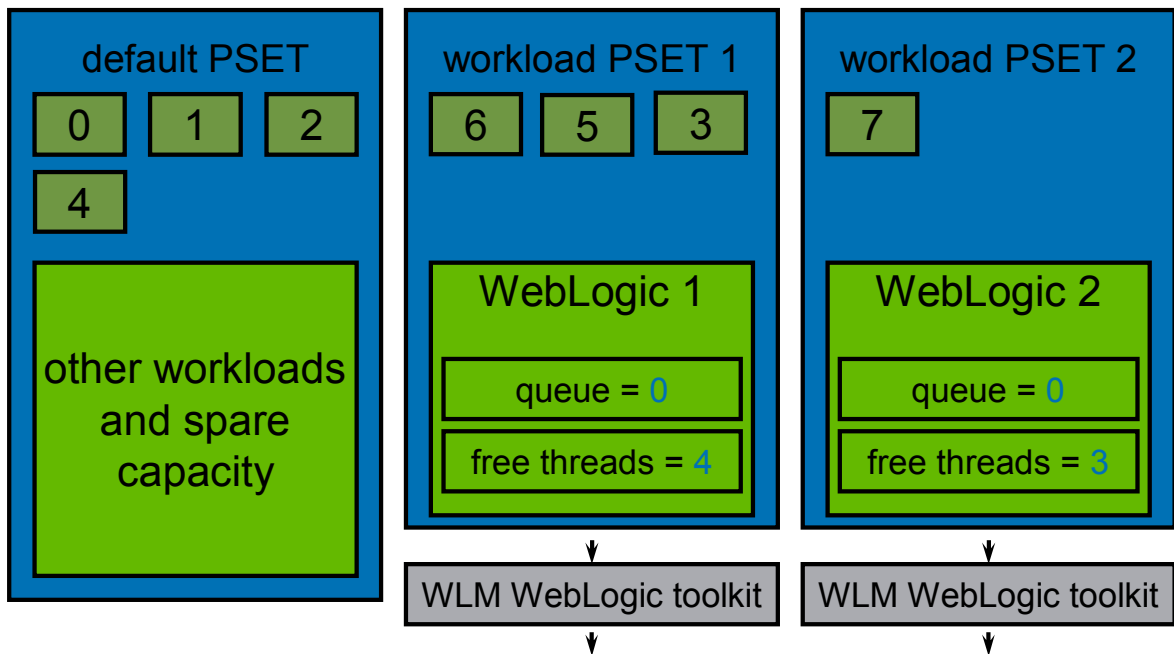
- Consolidate to 2 x rp8400s
- Install 3 vPars on each rp8400
- Use WLM to automatically move CPU resources between vPars based on CPU utilization, priorities and failover status
  - installation and configuration of WLM on all vPars took less than 2 hours
- MC/ServiceGuard for high availability
  - Production instances fail over to a test vPar on the other server
  - WLM automatically adjusts resources upon failover

HP-UX Workload Manager and hard partitions (nPars)

When the workload in nPar1 is busy, WLM will deactivate CPU's in nPar2 and activate the available iCOD processors in nPar1.

Since the total number of active processors on the system has not change, this does not incur any costs for activation of the iCOD processors.

When the workloads in nPar2 get busy, WLM can deactivate CPUs in nPar1 and activate them in nPar2.

This allows each nPar to scale from 4 to 8 CPUs depending on the status of the workloads running in each nPar.

nPar1   nPar2

Cell-based System

nPar1   nPar2

nPar1   nPar2
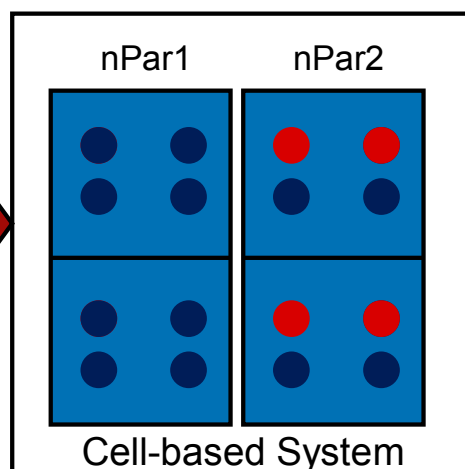
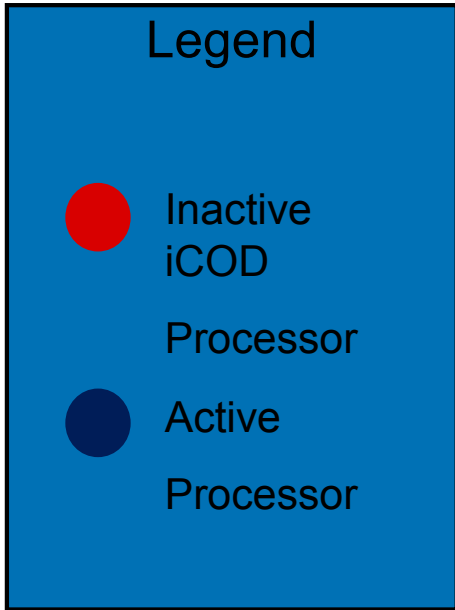This is a 4-cell 16-CPU server with 2 nPars – each with 2 cells.

12 active processors on the system and 4 available iCOD processors

WLM will use goal-based Service Level Objectives to determine which nPar the 12 active processors should be running in.
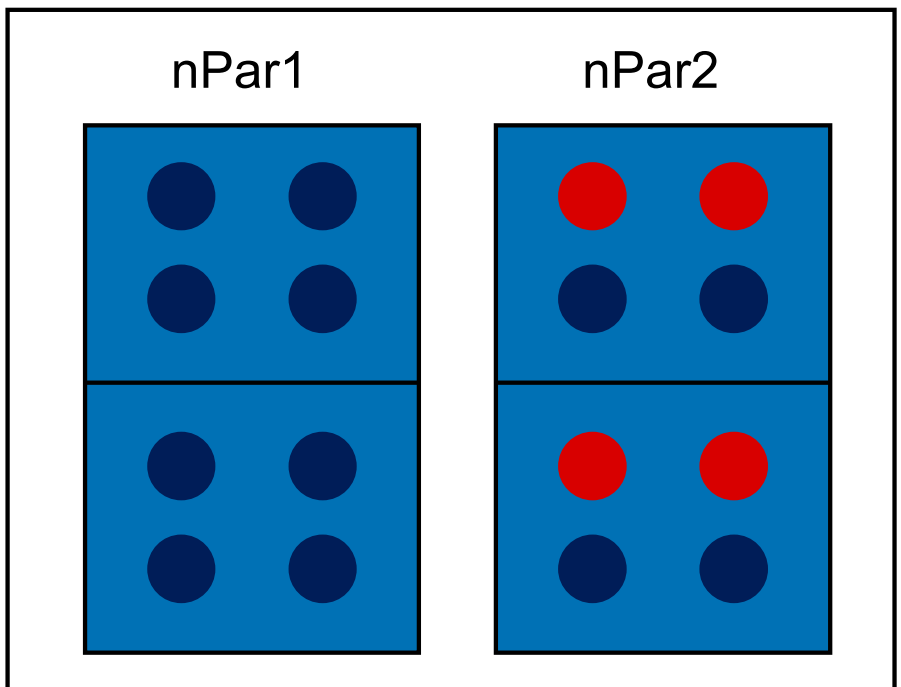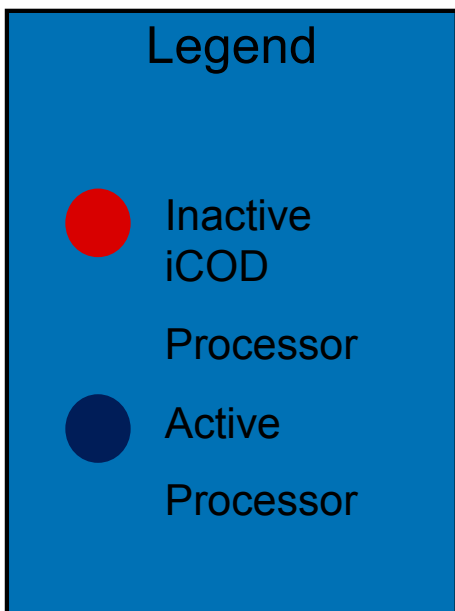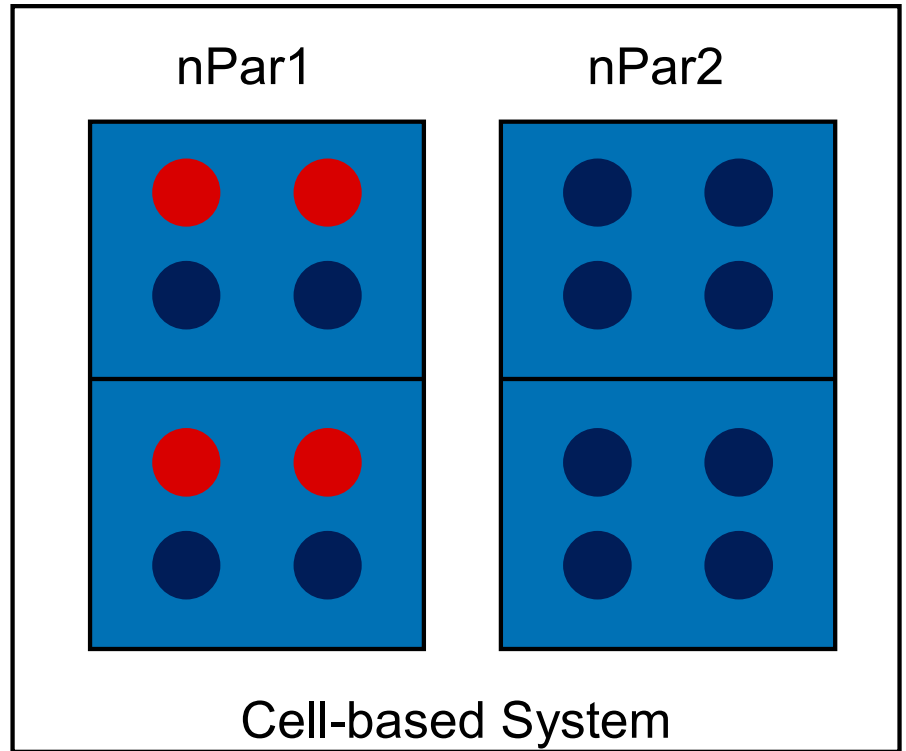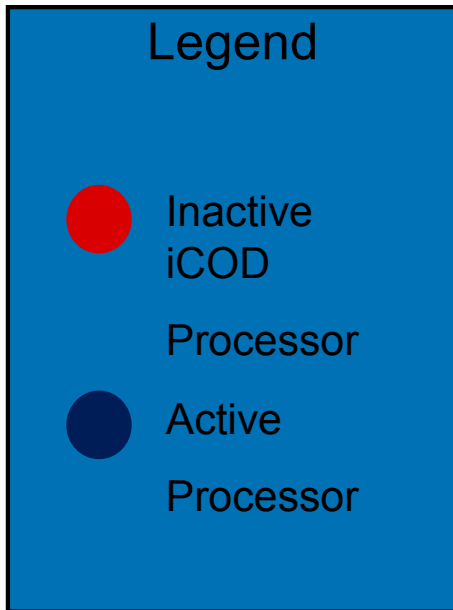
Legend

🔴 Inactive iCOD Processor

🔵 Active Processor

nPar1

nPar2

Cell-based System

WLM Support for iCOD across nPars

hp invent

Legend

🔴 Inactive iCOD Processor

🔵 Active Processor

nPar1

nPar2

Slide 1

nPar1 nPar2

Legend

● Inactive
iCOD

Processor

● Active

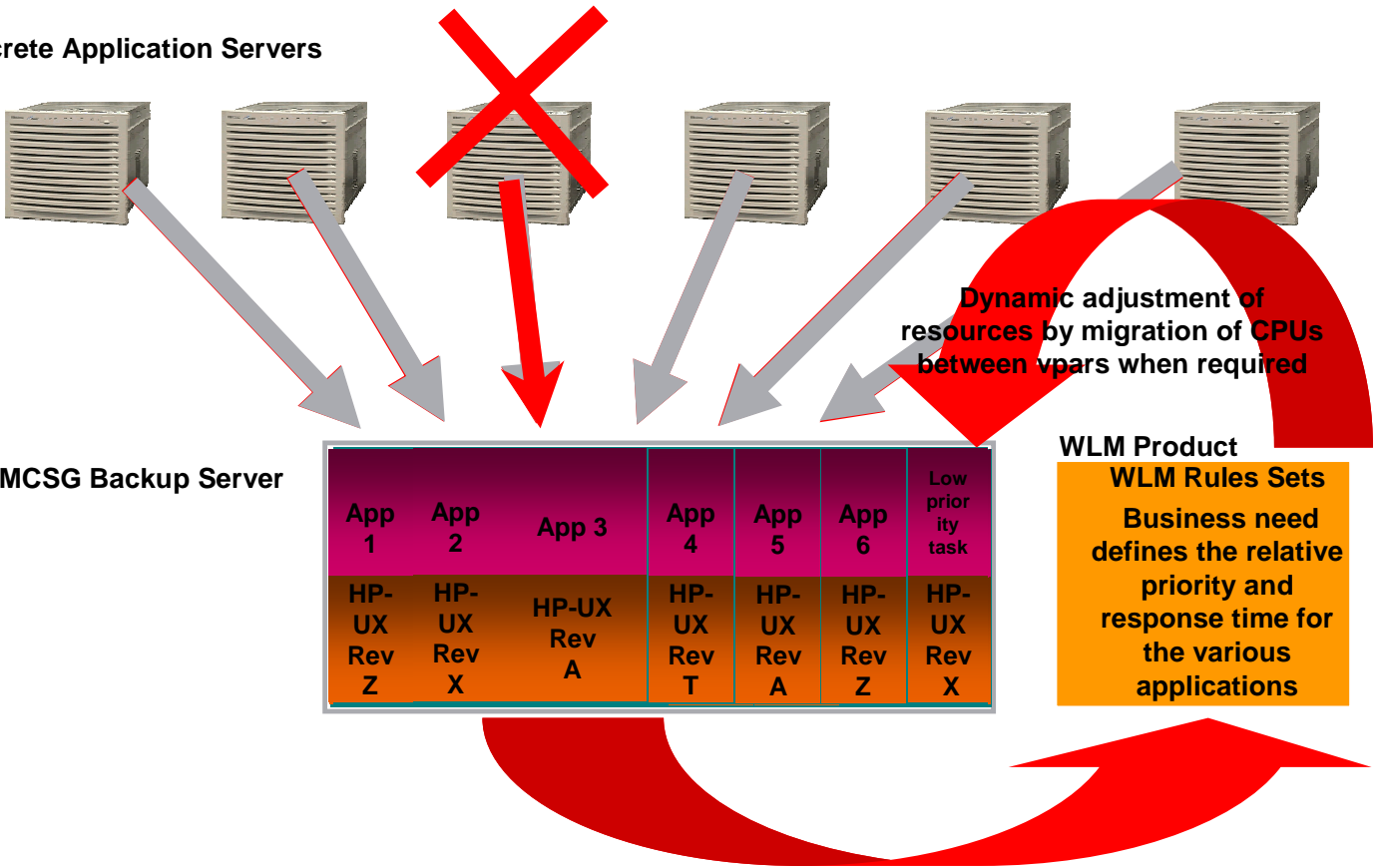Processor

Cell-based System

MC/ServiceGuard Integration

- HP-UX Workload manager is ideal in a Serviceguard Cluster
  - WLM will automatically adjust entitlements based on Serviceguard failovers
  - Example: A production system and a test system are clustered for failover.
    - When the production server fails over onto the test system the production package will be given 90% of the resources on the test system.
    - Test workloads do not need to be shut down
   - Test can get more resources automatically if the

Discrete Application Servers

MCSG Backup Server

Dynamic adjustment of resources by migration of CPUs between vpars when required

WLM Product
WLM Rules Sets
Business need defines the relative priority and response time for the various applications

App 1	App 2	App 3	App 4	App 5	App 6	Low priority task
HP-UX Rev Z	HP-UX Rev X	HP-UX Rev A	HP-UX Rev T	HP-UX Rev A	HP-UX Rev Z	HP-UX Rev X

WLM Configuration Wizard

- Guided initial configuration of wlm

- Sets up
  - Groups
  - Users
  - Apps
  - SLOs
  - Metrics

- Configures the oracle toolkit, if required

- Configures the MC/ServiceGuard toolkit, if required

Oracle toolkit configuration
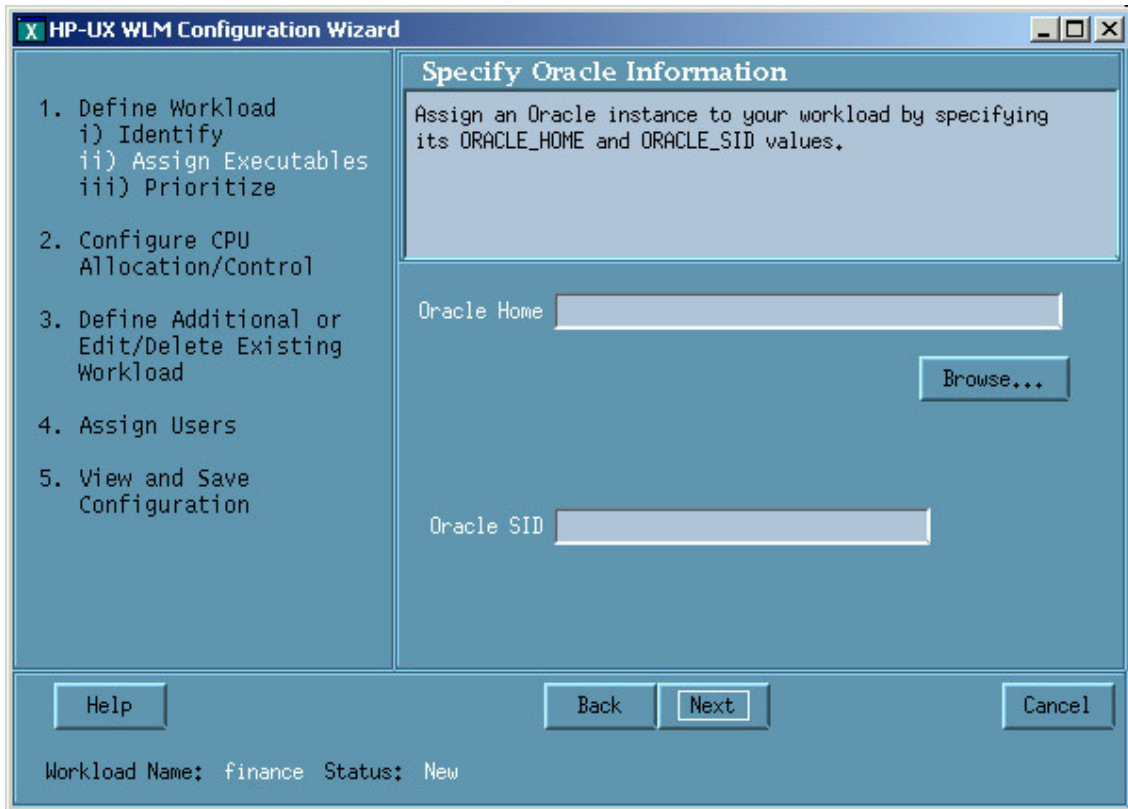
# Where WLM Provides the Most Value

- Consolidation
  - When you have some apps on servers with insufficient resources and others with very low utilization
  - Allows you to allocate unused resources to applications that need them

- Active-Active MC/ServiceGuard configurations
  - Allows prioritized control of how resources get allocated

- Active-Standby MC/ServiceGuard configurations
  - Reduced cost because WLM provides control over a single server that is the failover target for multiple other servers

- Utility Computing
  - Minimizes Utility Cost – CPUs are only active when they are

WLM and Utility Computing

- WLM can minimize utility computing costs
  - Utility CPUs are turned on only when needed to meet service level objectives
  - Utility CPUs are turned off as soon as they are no longer needed

Monitoring GUI Resources Panel

Remote GUI Architecture
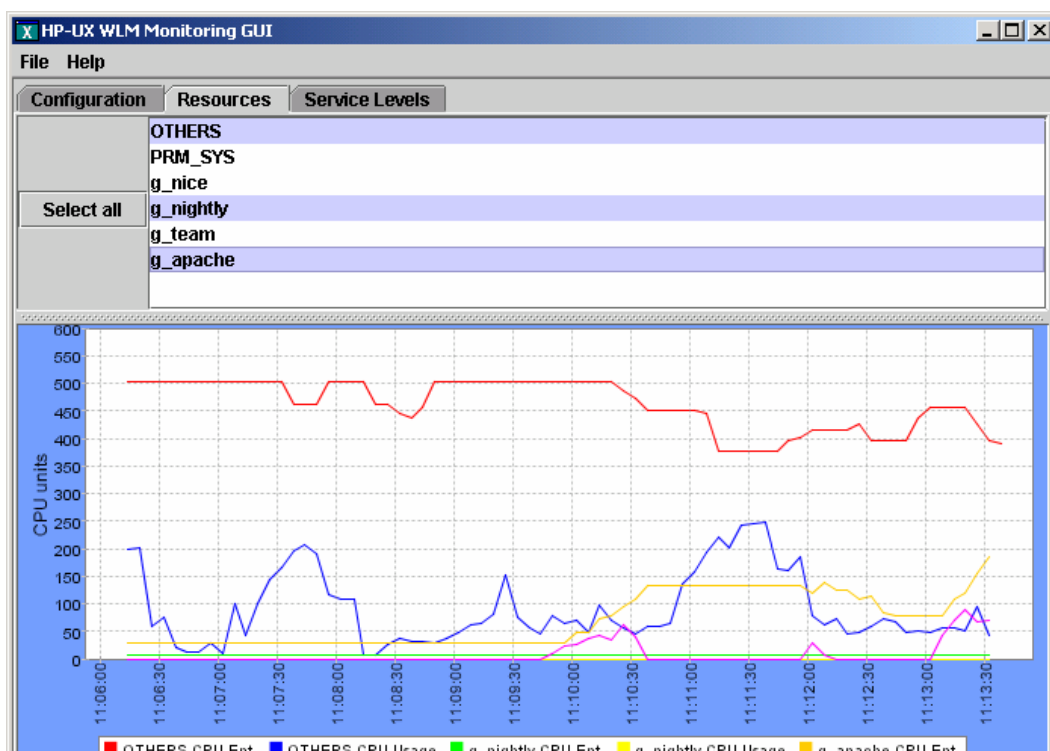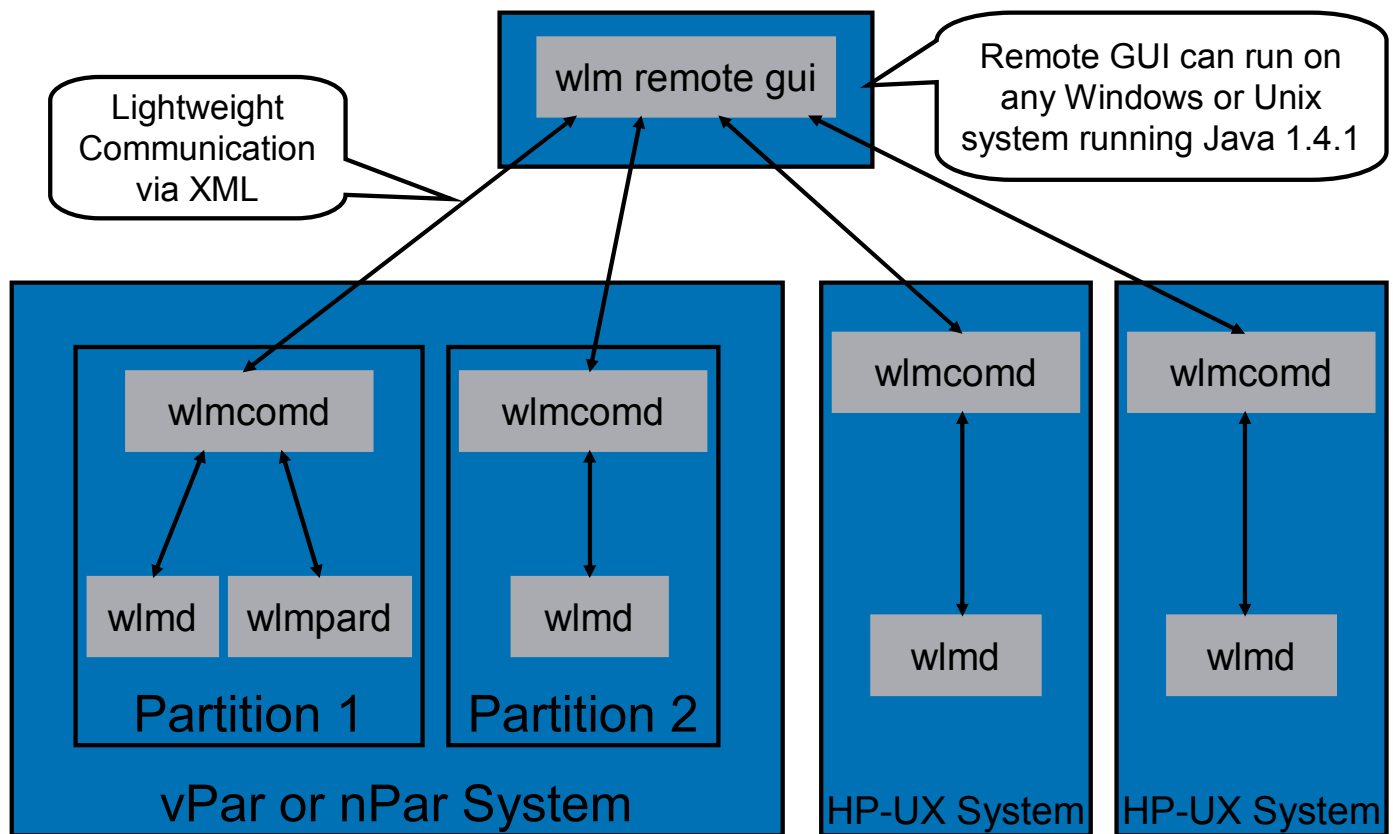
Support for PSET based transient groups

- Eliminates wasted resources for PSET partitions that don't have anything running for significant periods of time

- Reduces PSET to 0 CPUs when the group is inactive

- Intended for workloads that are not running for some period of time, like:
  - Failover partitions
  - Job-oriented workloads

- However
  - Any processes running in the PSET are moved to a default group when PSET is suppressed
  - Processes will be moved back when the PSET reactivates

HP-UX 11i Operating Environments

HP-UX 11i mission-critical OE

HP-UX 11i enterprise OE

HP-UX 11i OE

- HP-UX OS
- Network Drivers
- Web QoS peak
- Apache web server
- Ignite/UX
- Java Environment
  (RTE, JDK, JPI)

- ServiceControl Manager
- Sys. Config. Repository
- Software Distributor/UX
- Netscape directory server (Planet Enterprise)
- keberos PAM
- EMS framework
- CIFS Client & Server

- OnlineJFS 3.3
- Glanceplus pack
- Mirror Disk/UX
- Process Resource Mgr
- Event Monitor Services (EMS HA Monitors)
- OpenView Performance Agent
- Single-system Event and Availability Management

- MC/ServiceGuard
- HP-UX Workload Manager
- Host Intrusion Detection
- ServiceGuard NFS toolkit
- Enterprise Cluster Mgmt. toolkit

Herzlichen Dank!

Engagement für den Kunden