


Partitioning Continuum Overview

Siegfried Becker
Presales Consultant

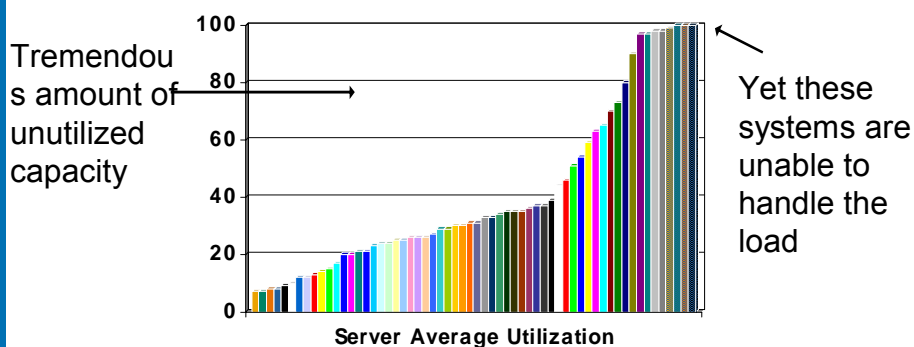


© 2004 Hewlett-Packard Development Company, L.P.
The information contained herein is subject to change without notice

Partitioning Continuum Overview –

Partitioning is a Key
Component of an
Adaptive Infrastructure

Why You Need an Adaptive Infrastructure



- Overall utilization less than 50%
- Some applications still not able to meet performance requirements

4/26/2004

Partitioning Continuum Overview

3

Definition of Partitioning



Partitions are physical or logical mechanisms for *isolating operational environments* within single or multiple servers to offer the *flexibility of dynamic resizing* while ensuring that applications can enjoy *protection from unrelated events* that could otherwise cause disruption, interruption, or performance degradation.

4/26/2004

Partitioning Continuum Overview

4



Why is Partitioning Important ?

Customer Challenge

Pressure to offer service level guarantees at reasonable costs

Underutilization of servers

Address high fluctuation of loads on applications

Flexibility with privacy and high availability

Partitioning Benefit

Meet service level agreements with best-in-class return on investment

Increase utilization to 80-90% of compute power without sacrificing performance

Fast and dynamic adjustment to changing loads
"right" level of application isolation provides both flexibility and increased uptime

4/26/2004

Partitioning Continuum Overview

5



Reasons for Partitioning

- Increase server utilization
- Consolidating multiple applications onto a single server
- Separate production, test and development environments
- Improved RoIT
- Increased server flexibility
- Easy resizing
- Reduced risks through several levels of isolation
 - Hard partitions nPars
 - Virtual Partitions vPars
 - Resource Management
- Improved service levels (e.g. appl. response time)
- Integrate WLM as an intelligent policy engine

4/26/2004

Partitioning Continuum Overview

6

Choosing between partitioning technologies



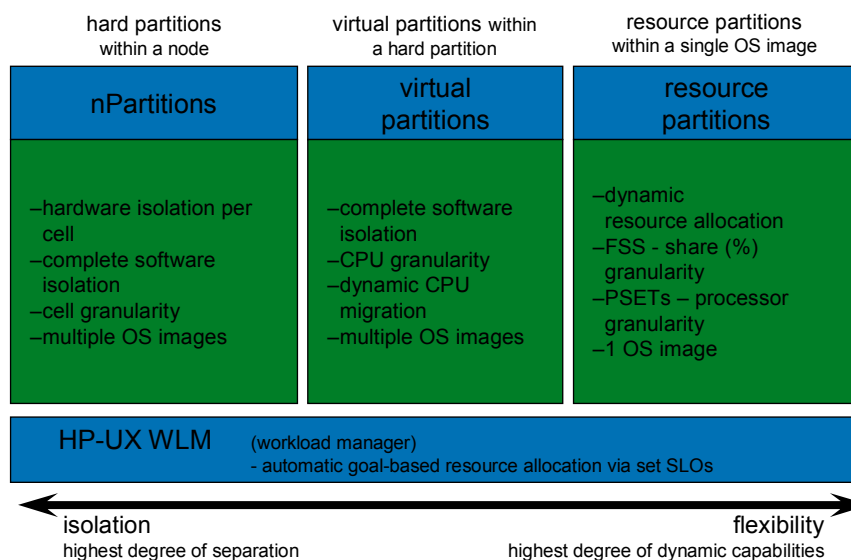
- Hard Partitions - nPars
- Virtual Partitions - vPars
- Processor Set Resource Partitions - PSET
- Fair Share Scheduler Resource Partitions - FSS

4/26/2004

Partitioning Continuum Overview

7

Adaptive infrastructure on HP-UX HP's Partitioning Continuum



4/26/2004

Partitioning Continuum Overview

8

Hard Partitions - nPars



HP Hard Partitions - nPars

- Configure a server complex as one large system or as multiple smaller systems
- Electrical and SW isolation
- Cell Board granularity
- Each partition has its own CPUs, memory and IO resources
- Each partition operates on its own OS
- Multiple OS are supported simultaneously (HP-UX, Linux, Windows, OpenVMS)
- Multiple OS versions are supported simultaneously
- Application isolation

Partition Manager on 11i

GUI for system administrators to configure and manage partitions



- Create, modify, delete partitions
- Display a complete hardware inventory
- Display status of key complex components
- Check for problem or unusual complex conditions
- Manage power to cells & I/O chassis
- Turn on/off attention indicators for cells, I/O chassis, I/O cards & cabinets
- Free product, included with OS

4/26/2004

Partitioning Continuum Overview

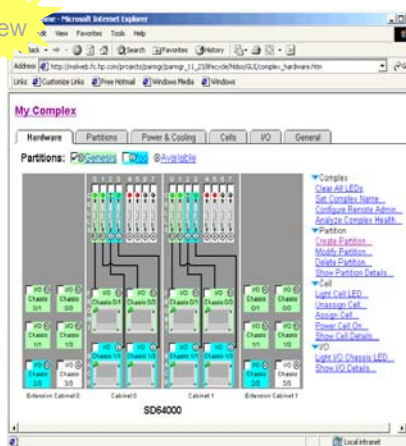
11

Partition Manager for HP-UX 11i v2: new graphical and resource display

- Hardware-centric management of hard-partitioned servers
- Current parmgr features on Itanium[®]-based HP servers PLUS...

(1st on HP-UX 11i v2)

- New Web interface
- Graphical “big picture” views of
 - nPars
 - Hardware components
- Easy-to-see status lights
- Smart action menus adjust per selected element
- Remote administration
- Compatible with iCOD/pay-per-use and vPars
- Launched by Servicecontrol Manager



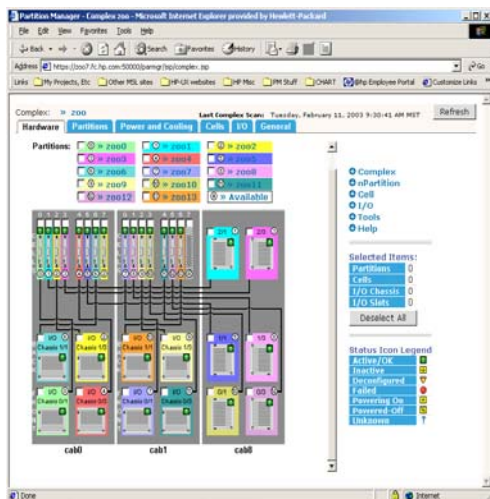
4/26/2004

Partitioning Continuum Overview

12

Partition Manager New Features

Significant Changes from ParManager on HP-UX 11i



- ✓ New web interface
- ✓ Graphical "big picture" views of
 - nPars
 - Hardware in complex
- ✓ Supports new OS/HW features
 - Cell local memory for HP-UX 11i v.2 partitions
 - nPartition configuration privilege
- ✓ Remote admin of Superdome complex
- ✓ Compatible with iCOD/pay-per-use
- ✓ Native on Windows (2H 2004)

4/26/2004

Partitioning Continuum Overview

13

Virtual Partitions

-

vPars

HP Virtual Partitions Key Features



- Run multiple versions (instances) of HP-UX simultaneously on one server or nPar
- Application and OS isolation (application, name space, OS and kernel isolation)
 - Separate OS instance
 - Applications are fully isolated (SW)
- Dynamically create, modify or delete one OE without interrupting non-related partitions
- Single CPU granularity (recommend at least 2)
- CPU migration - Dynamic reassignment of CPUs across virtual partitions without reboot
- Support of hp 9000 rp5470, rp7400, Superdome/rp8400/rp7410 (including nPar)
- Support for PA-8800 dual core module on Superdome, rp8420 and rp7420 with 3.1
- Individual reconfiguration and reboot, e.g. for rolling upgrades (virtual partitions don't affect each other)
- Both GUI and command line interface
- Single toggle console (in future consolidated console)
- Compatible with PRM, HP-UX Workload Manager, and MC/ServiceGuard, iCOD
- Greater flexibility compared to nPars

4/26/2004

Partitioning Continuum Overview

15

HP-UX Virtual Partitions



Multiple HP-UX instances
running on the same system
or in the same nPar

Dept. A App 1	Dept. A App 1'	Dept. B App 2	Dept. B App 3
HP-UX Revisio n A.1	HP-UX Revisio n A.2	HP-UX Revisio n B.3	HP-UX Revisio n B.3



Increased system utilization

- partitioning a single physical server or hard partition into multiple virtual partitions for rp5405, rp5470, rp7400, Superdome, rp8400, rp7410, rp8420, rp7420

Increased Flexibility

- multiple independent instances of HP-UX
- dynamic CPU migration across virtual partitions

Increased Isolation


- application isolation across virtual partitions
- OS isolation
- individual reconfiguration and reboot

4/26/2004

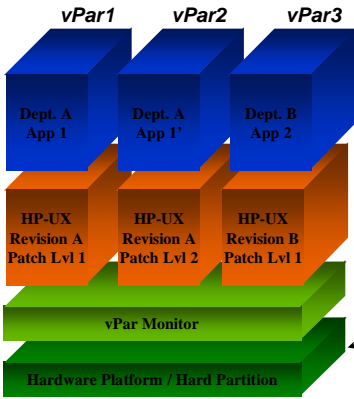
Partitioning Continuum Overview

16

vPars logical overview




- multiple applications or multiple instances or versions of the same application
- provides name space and resource isolation
- creates illusion of many separate hardware platforms
- manages shared physical resources
- monitors health of operating system instances



- each operating system instance tailored specifically for the application(s) it hosts
- operating systems instances are given a user-defined portion of the physical resources
- provides name space and resource isolation
- supported on rp5470, rp7400, Superdome, rp8400, rp7410, rp8420, rp7420 systems
- no additional platform support required

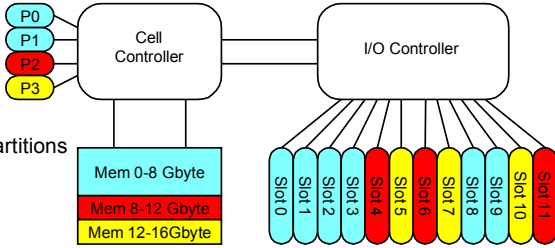
4/26/2004
Partitioning Continuum Overview
17

Example of 3 vPars on a single-cell nPar



Blue Partition Uses two CPUs, 8 Gbyte memory, 6 PCI slots.
 Red partition uses one CPU, 4 GByte memory, 3 PCI slots.
 Yellow partition uses one CPU, 4 GByte memory, 3 PCI slots.

unbound CPUs can dynamically move between partitions without reboot!

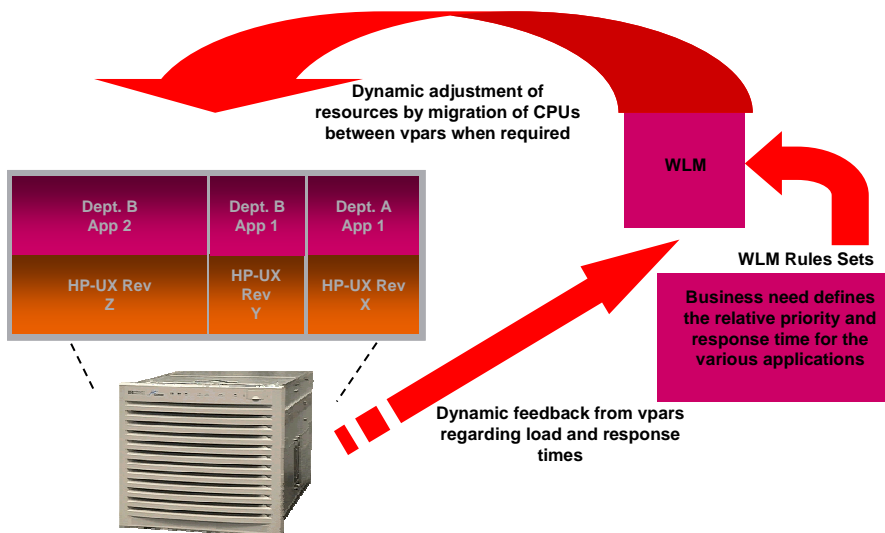


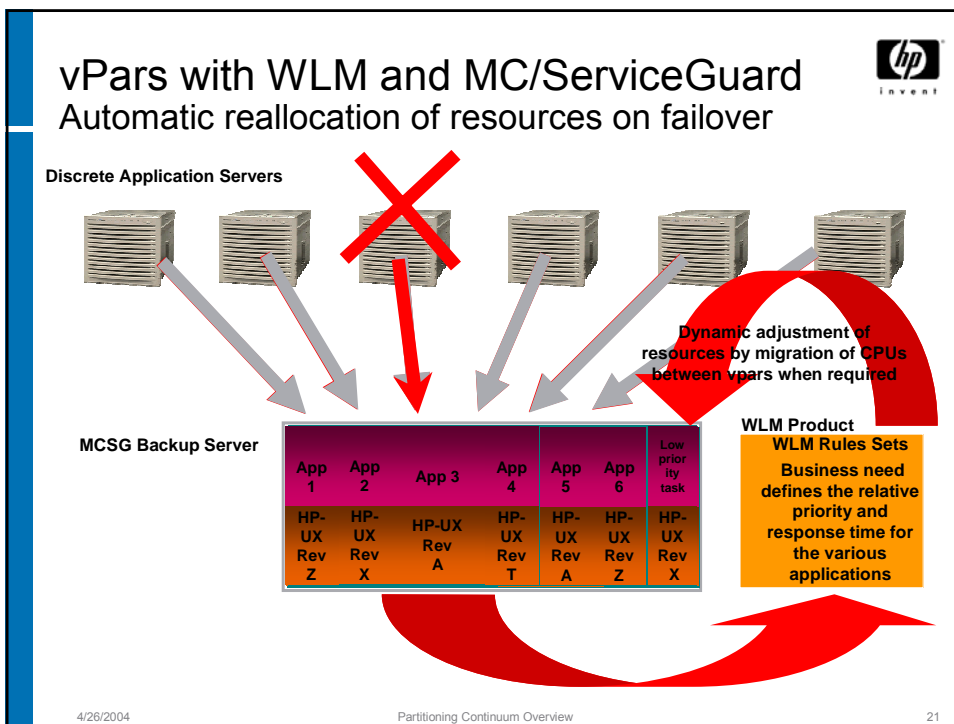
Example: CPU P1 (an unbound CPU in this case) can move from blue to red partition when the red partition needs more processing power.

4/26/2004
Partitioning Continuum Overview
18


vPars Solutions

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

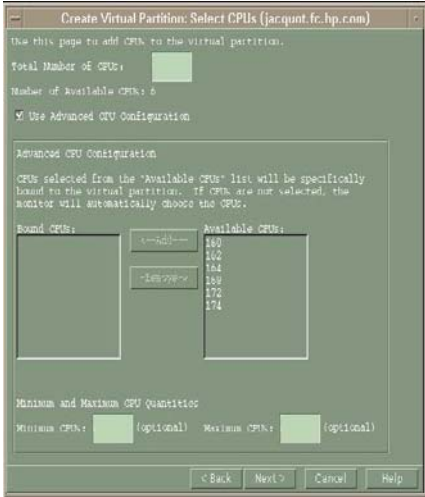




vPar Manager




Virtual Partition Manager (vparmgr)



- ✓ Create, modify and delete virtual partitions (vPars)
- ✓ Display assigned resources, attributes, and status of vPar
- ✓ Display vPar event log and samlog
- ✓ Boot and reset a vPar
- ✓ Direct invocation of task screens
- ✓ Preview create/modify vPar command lines prior to execution

4/26/2004
Partitioning Continuum Overview
23



HP-UX Virtual Partitions Roadmap

Focus of development:

- Support on PA-RISC of:
 - New I/O cards and firmware
 - rp8400 Server Expansion Unit (SEU/IOX)
 - PA8800 processors
- New support of Integrity
- Increased:
 - dynamic capability
 - VSE integration
- Additional virtualization capabilities

2Q CY'04

*vPars A.03.01
(on HP-UX 11i v1 PA):*

Support of:

- rp8400 SEU (IOX)
- PA8800/8900 based systems for SD, rp8400, rp7410 (w/ single CPU granularity)
- improved vPars/iCOD/WLM integration for concurrent use

1H CY'05

(on HP-UX 11i v3 for PA & Integrity)

- continued support on PA
- new support for Integrity
- fully dynamic CPUs (*remove bound CPUs*)
- cell OLA integration

HP Restricted: For HP and Partner Internal Use Only - Customer viewable under NDA
Current as of 1/8/04; subject to change without notice

4/26/2004
Partitioning Continuum Overview
24

Where vPars provides the most value



- System/data center consolidation
- development/test environments
- increased system utilization
- varying workload requirements:
 - time of day: order entry during day, batch at night
 - time of month (payroll, end-of-month/end-of-year financials)
 - as particular needs require
- service provider (providing system resources to different users/applications)
- unique application tuning of O/S
- time zoning



4/26/2004

Partitioning Continuum Overview

25

Resource Partitions

- PSETS
- FSS
- MRG



Resource Partitioning

- The Problem:
 - Competition for resources on a consolidated server
- The Solution:
 - Resource Partitioning with Process Resource Manager (PRM)
- PRM is used to configure resource partitions and assign groups of processes to run in each partition

4/26/2004

Partitioning Continuum Overview

27



HP-UX Resource Partitions Key Features and Benefits

- Controls allocation of CPU, real-memory and disk I/O bandwidth based on user-specified policies.
- Applications do not require modification to work with PRM.
- Configuration can be changed at any time - even under load.
- Supports resource policies based on users and applications.
- Tight integration with HP-UX systems management and workload management tools

4/26/2004

Partitioning Continuum Overview

28

Resource Partitioning with Processor Sets

- PSET is a PRM group type to manage processor resources
- CPU is allocated on whole CPU boundaries
- Dynamically configurable during runtime
- A configuration can contain both FSS and PSET groups
- Standard Unix scheduler allocates CPU within a PSET
- Permit consolidation of multiple independent applications on a single server (batch)
- Oracle Database Resource Manager and HP PRM can be integrated using PSET groups
- Free of charge

4/26/2004

Partitioning Continuum Overview

29

HP Resource Partitions

- Resource management to control the amount of process resources
- Can manage
 - CPU
 - Memory
 - Disk bandwidth
- Allocates resources to specific applications / users
- Acts within one OE
- CPU and sub-CPU granularity
- Supports PSETs

4/26/2004

Partitioning Continuum Overview

30



Resource Partitioning Features

- Supports hierarchical partitions
- Resource controls:
 - CPU Controls
 - CPU allocation by percentage, shares, or whole processors
 - Optional capping in FSS partitions
 - Concurrent FSS and PSETs
 - Real memory controls
 - Each partition gets a separate memory manager in 11i
 - Disk bandwidth
 - Both LVM and Veritas VxVM Volume Groups
 - Automatic process assignment to partition
 - Users/Groups
 - Executable path/Process name
 - Children automatically run with parent by default

4/26/2004

Partitioning Continuum Overview

31

Resource Partitioning with Processor Sets



- PSET is a PRM group type
- CPU is allocated on whole CPU boundaries
- A configuration can contain both FSS and PSET groups
 - The FSS groups run in the default PSET
- Standard Unix scheduler allocates CPU within a PSET
 - Separate process schedulers for each partition, as with FSS groups as well
- PRM uses PSET system calls to manipulate PSETs
- Configuration using configuration file or xprm

4/26/2004

Partitioning Continuum Overview

32



HP Processor Sets - PSETS

- Manage processor resources
- Dynamically configurable during runtime
- Permit consolidation of multiple independent applications on a single server (batch)
- Oracle Database Resource Manager and HP PRM can be integrated using PSET groups
- Free of charge

4/26/2004

Partitioning Continuum Overview

33



FSS CPU Management

- Fair-share scheduler sits on top of standard Unix scheduler in the kernel
- FSS allocates CPU ticks to partitions based on entitlements
- Shares are then allocated using standard scheduling
- Unused CPU cycles are available to other groups (when not capping).
- Supports “capping” mode

4/26/2004

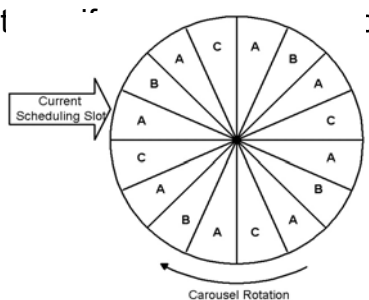
Partitioning Continuum Overview

34

FSS CPU Management - Carousel Algorithm



- Shares become slots in a carousel
- Each slot is 10 ms.
- Processes in group get first shot at CPU during their tick
- If capping is off carousel 1 CPU



Example:

Group A has 8 shares or 50%

Group B has 4 shares or 25%

Group C has 4 shares or 25%

4/26/2004

Partitioning Continuum Overview

35

Resource Partition Memory Management in 11i

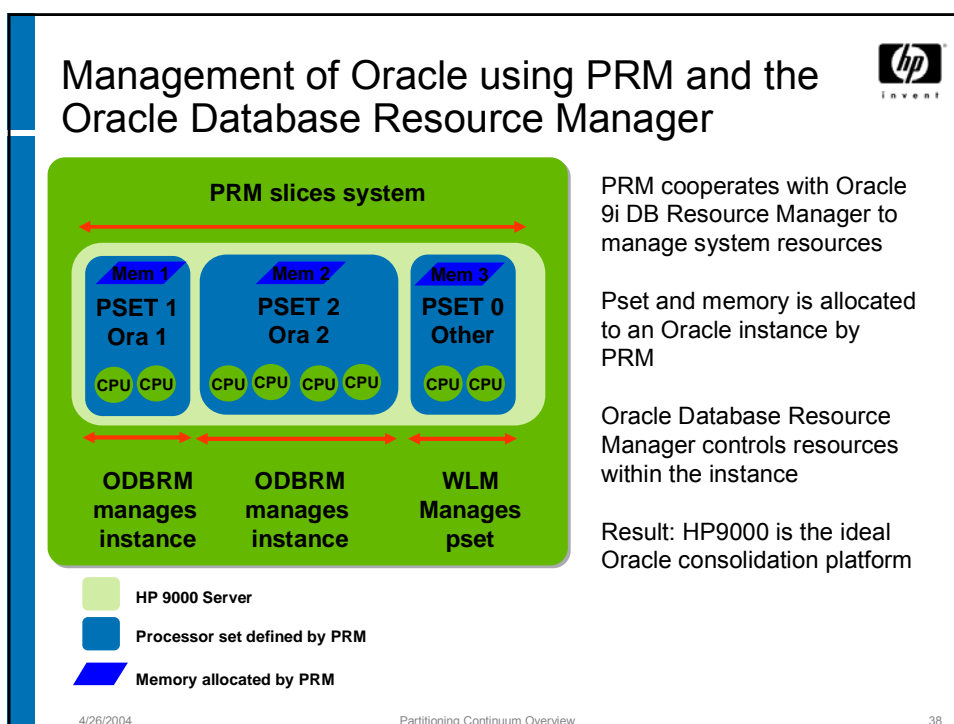
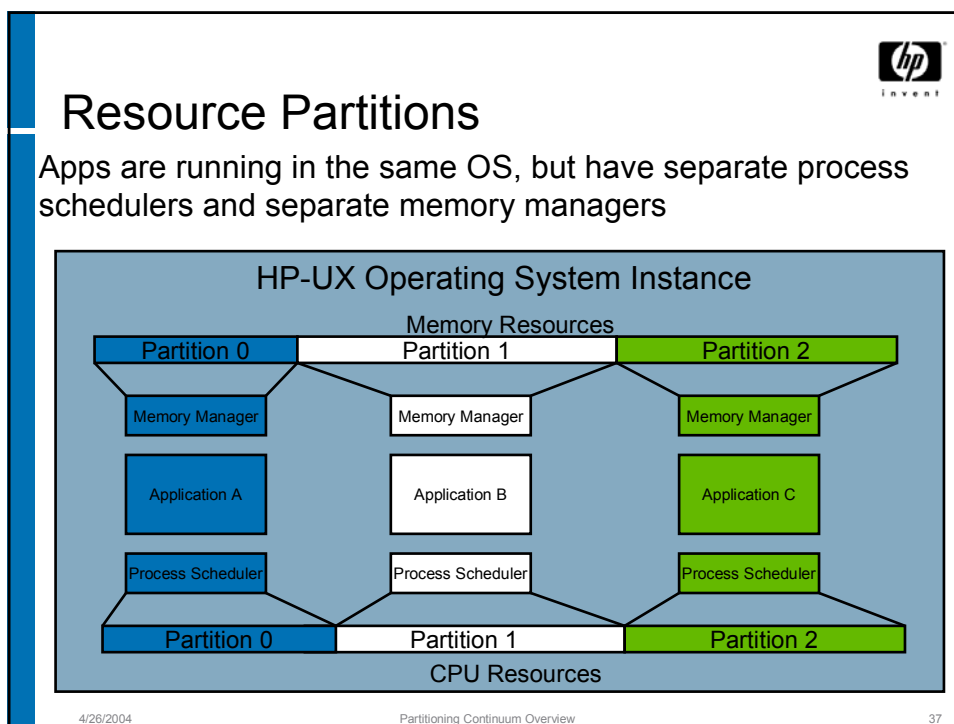


- Memory Resource Groups (MRGs) are implemented in the 11i kernel
- MRGs are mapped to PRM groups
- Each MRG contains its own copy of a memory subsystem
 - Processes in the group will page if they allocate more memory than their entitlement
- PRM is the only supported interface to MRGs
- Supported on HP-UX 11i and above

4/26/2004

Partitioning Continuum Overview

36





Disk Bandwidth Management

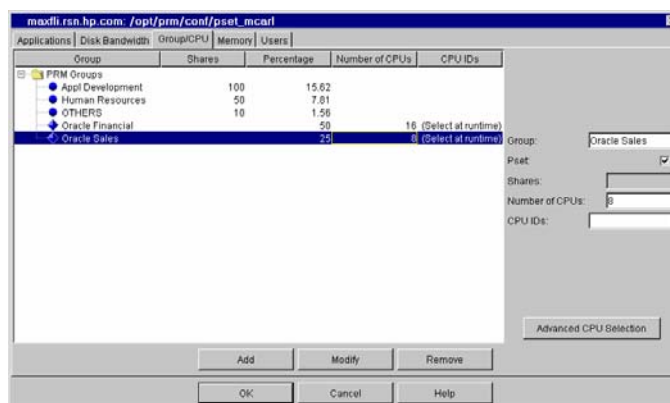
- Requires logical volume manager (LVM)
- NEW – Supports Veritas Volume Manager (VxVM) in 2.1
- Shares are specified at the volume group level
- Access can be via raw device or file system
- Works on physical I/O only, buffer cache hits are not affected
- Takes effect when contention for I/O exists
- Reorders I/O queues based on partition entitlements

4/26/2004

Partitioning Continuum Overview

39

Process Resource Manager - PRM GUI



This screenshot shows two PSET groups and 3 fair share groups configured using the PRM Java based GUI

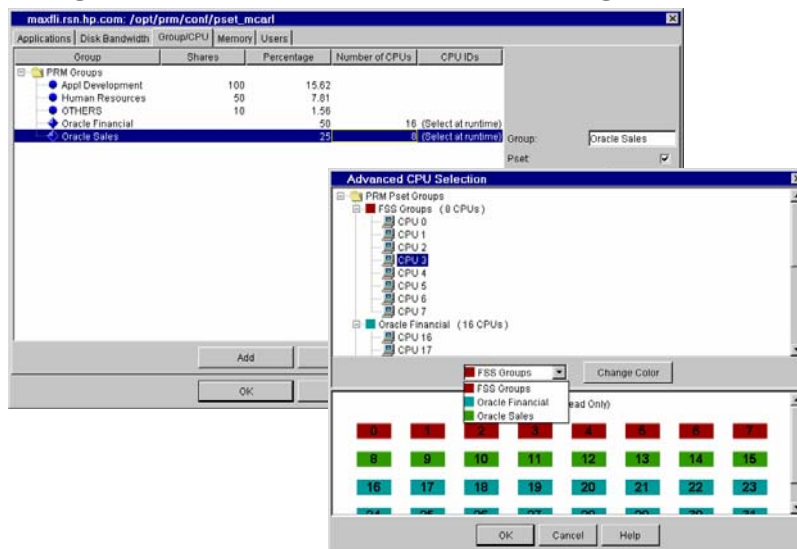
4/26/2004

Partitioning Continuum Overview

40



Drag and Drop PSETs Configuration

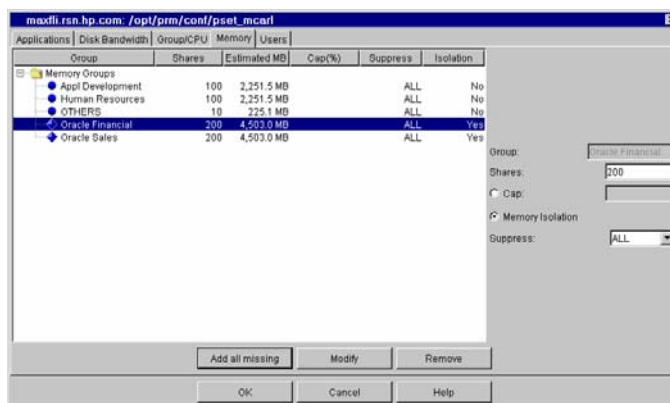


4/26/2004

Partitioning Continuum Overview

41

Allocating Memory to Resource Partitions



creates isolated environment of both cpus and memory

4/26/2004

Partitioning Continuum Overview

42



Benefits of PRM/psets integration

- Psets comes with HP-UX and does not require PRM. However, you get many important benefits when you use PRM to configure psets:
 - Java based GUI for configuration of psets
 - configuration is maintained across reboots
 - can mix and match psets with fair share scheduler groups on the same system
 - provides memory management, making it possible to isolate a share of memory to a pset.
 - processes are automatically moved into the appropriate pset or fair share group – no action is required as processes come and go
 - users are automatically moved into the appropriate pset or fair share group – no action is required as users come and go
- This integration is the reason that HP PRM Resource Partitions are supported for use the Oracle's resource manager product, the Oracle Database Resource Manager

4/26/2004

Partitioning Continuum Overview

43



Resource Partitions

- Why choose resource partitions over nPars or vPars?
 - Allows shared I/O – no need to duplicate hardware for each partition
 - Much easier to implement
 - Much lower TCO - single os instance to manage
 - Can run within an nPar and/or a vPar
- PSETs provides:
 - Processor isolation – apps have sole access to processors in the group
 - Memory isolation on top of PSETs
- FSS provides:
 - More granular CPU allocation
 - More partitions

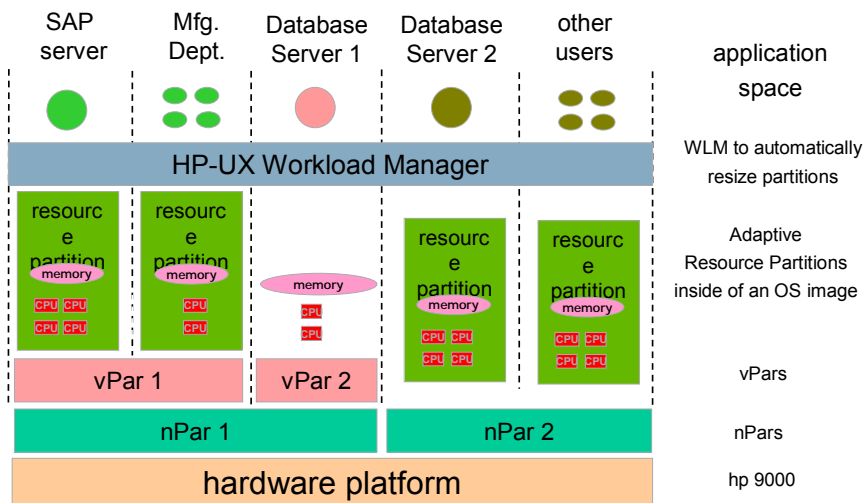
4/26/2004

Partitioning Continuum Overview


44

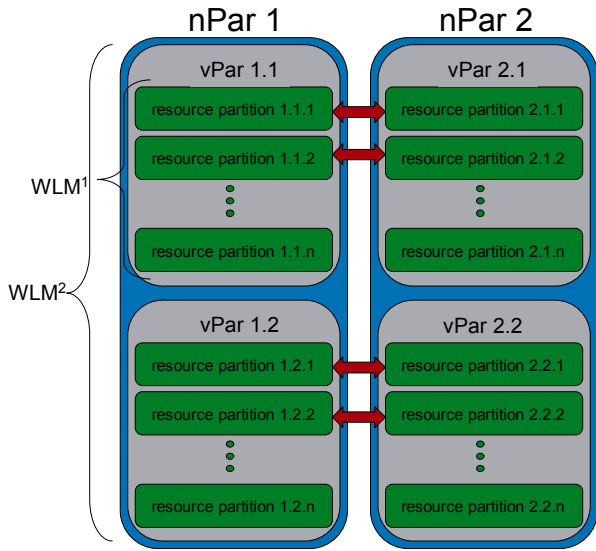
Partitioning – the big picture


HP-UX = Broadest Partitioning Portfolio



Resource management of your adaptive infrastructure





- 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
- WLM¹ automatically allocates CPU resources as needed to resource partitions
- WLM² 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

4/26/2004 Partitioning Continuum Overview 47

On-Demand technologies



HP-UX On-Demand Technologies

- Instant Capacity on Demand (iCOD)
 - Activate new permanent capacity when needed
- Instant Capacity on Demand – Temporary Capacity (TiCOD)
 - Activate/deactivate new temporary CPU capacity when needed
- Pay-per-Use Utility Computing – (PPU)
 - Lease systems based on CPU utilization

4/26/2004

Partitioning Continuum Overview

49

Herzlichen Dank!

Engagement für
den Kunden



4/26/2004

Partitioning Continuum Overview

50



Instant Capacity on Demand (iCOD)

- System acquired with inactive processors
- Processors are paid for when they are activated
 - Price paid is current price when activated
- CPUs can be activated on-line – no reboot required
- Excellent solution for expected growth

- iCOD is licensed for an entire complex
- CPUs can be deallocated in one nPar and activated in another

4/26/2004

Partitioning Continuum Overview

51



iCOD Temporary Capacity (TiCOD)

- Alternative purchasing model for iCOD processors
- Temporary Capacity is purchased in 30 Day increments
 - 30 CPU-Days = 43,200 CPU-Minutes
- Any number of iCOD CPUs can be activated
- Activating processors causes the iCOD software to deduct minutes from the “bank”
- Deactivating the iCOD processors stops the deductions

- Excellent solution for:
 - Short term peaks in application load
 - Activation of additional capacity upon failover of a large workload onto a failover server

4/26/2004

Partitioning Continuum Overview

52

Pay-per-Use Utility Computing (PPU)



- Type of lease
- Acquire a system with peak capacity required
- Monthly charge based on base payment plus a variable payment based on actual resource usage
- 2 utilization measurement models
 - Active CPU – CPUs are activated/deactivated and variable payment is based on how long CPUs were active
 - Percent Utilization – All CPUs are active and the system is monitored for CPU utilization – variable payment is based on average utilization of all CPUs
- Excellent solution for highly variable loads, especially revenue generating loads because costs are in line with revenues

4/26/2004

Partitioning Continuum Overview

53

Architecting a Solution with the HP-UX Adaptive Infrastructure

Common Benefits of all Partition Types supported at different levels



- maximize system utilization
- resource isolation
- os isolation
- support for full line of HP 9000 servers
- os version support
- ease of setup and management
- flexible CPU resources
- partition stacking
- iCOD, PPU support
- wlm support

4/26/2004

Partitioning Continuum Overview

55

Benefits/Strengths



Benefit	nPars	vPars	prm/psets	prm/fss
Maximize system utilization	Good	Better	Better	Best
Resource isolation	Best	Better	Better	Good
Os isolation	Best	Better	No	No
Support for all 9000 servers	sd,8400,7410	i,n,sd,8400,7410	All	All
Os version support	11i	11i	11i	10.20,11.x
Ease of setup	Good	Better	Best	Best
Ease of management/TCO	Good	Better	Best	Best
CPU resource flexibility	Good	Better	Better	Best
iCOD/PPU support	Yes	iCOD/%PPU	Yes	Yes
WLM support	March 04	Yes	Yes	Yes

4/26/2004

Partitioning Continuum Overview

56



nPars

- nPars is the only partition type that has:

Hardware Fault Isolation Windows & Linux Support

- A hardware fault in one partition will not effect the other partitions
- You can also do hardware maintenance in one partition while the other partitions are running
- Single CPU resource migration is possible if iCOD CPUs are available on the system
 - WLM will automate this in 2.2 (March 04)

4/26/2004

Partitioning Continuum Overview

57



vPars

- Why choose vPars over nPars?
 - vPars provides:
 - Dynamic processor movement without rebooting the partition
 - Single cpu granularity without need for iCOD
 - Can run within an nPar
- Why choose vPars over resource partitions?
 - vPars provides:
 - Software fault isolation
 - Different versions of the OS
 - Application isolation

4/26/2004

Partitioning Continuum Overview

58

When to use On-Demand Technologies



- iCOD is useful for deferring cost of anticipated growth
 - Resources can be added very quickly
 - Resources can be added while the system is on-line
- TiCOD is useful for short-term spikes in load or for failover server
 - Costs can be managed/budgeted
- PPU is most useful for highly variable loads
 - Particularly for revenue generating workloads because costs vary in line with revenues

4/26/2004

Partitioning Continuum Overview

59

WLM



- WLM is NOT a partitioning technology, it provides automatic movement of CPU resources to workloads that need them to meet SLOs
- WLM provides:
 - Automatic CPU resource allocation across Resource Partitions, vPars, and nPars with iCOD
 - Truly maximizes CPU utilization
 - Automatic response to ServiceGuard failovers
 - Guaranteed consistent performance during varying loads on the application
 - iCOD/TiCOD integration
 - Minimizes utility(PPU) computing costs through automatic allocation/de-allocation of utility CPUs

4/26/2004

Partitioning Continuum Overview

60



Unsupported Combinations

- iCOD/vPars/WLM incompatibility will be removed with the 2.2 release of WLM and the 3.1 release of vPars in 2Q 04.
- vPars does not support Active CPU PPU – this will be resolved in the **X.X release of vPars in MONTH of 04.**
- PSETs/vPars – vPars CPU migration is NOT supported when PSETs are being used in an affected vPar – this will be resolved in the 4.1 release of vPars in 1H of 05.
- WLM 2.2 (March 04) will allow nPar/iCOD migration OR vPar CPU migration OR Auto PPU/TiCOD activation/deactivation – no two will be supported in the same config. This will be resolved in the WLM 2.4 release in March of 05.

4/26/2004

Partitioning Continuum Overview

61



Key Takeaways


- **All of these options provide the ability to consolidate applications or consolidate data centers and ensure that each app has a minimum amount of resources.**
- **If resource contention is the top issue, resource partitioning is the easiest to set up, the easiest to manage and provides the most flexibility.**
- **If HA is the top issue, nPars provides hardware fault isolation and vPars provides software fault isolation.**
- **If I/O chassis space is limited, resource partitions can be used without requiring duplication of I/O.**
- **If applications don't coexist well on the same OS image, nPars or vPars are the right solution.**
- **If the applications have varying loads and varying priorities, WLM can be used to ensure the resources get used to the best business advantage possible.**
- **Consider using On-Demand technologies (iCOD, TiCOD, PPU) where there are varying loads**

4/26/2004

Partitioning Continuum Overview

62

HP-UX Partitioning Continuum – Competitive Positioning



HP Partitioning Continuum Products

	clusters	hard partitions	virtual partitions	resource partitions
Windows	industry products	future systems	VMWare Server	hp ProLiant Essentials Workload Management Pack (RPM)
Linux	industry products	future systems	VMWare Server	PRM for Linux
OpenVMS	OpenVMS clusters	AlphaServer hard partitions	OpenVMS Galaxy	None
Tru64 UNIX	TruCluster Server	AlphaServer hard partitions	None	ARMTech psets Class Scheduler
HP-UX	Hyperplex	nPartitions	vPars	PRM pSets

HP-UX Workload Manager

4/26/2004 Partitioning Continuum Overview 64

HP's Lead over IBM in On-Demand Technology



Unix COD offerings	IBM launched	HP Launched	HP ahead...
CPU COD	High-end - 10/2000 Midrange - 11/2001	iCOD in Nov 1999 (both midrange and high-end support)	1 year in high-end 2 years in midrange
COD integration with partitioning	Oct 2002 (AIX 5.2)	June 2000	2+ years
Dynamic/Hotspare CPU failover (if CPU fails)	Oct 2002 (AIX 5.2)	June 2000	2+ years
WLM integration with COD	Oct 2002 (with AIX 5.2)	June 2001	Over 1 year
Storage COD	May 2003	February 2001	2+ years
Temporary capacity on demand	Aug 2003*	August 2002	1 year
Cell/memory COD	Sept 2003 (memory COD only)	March 2003 (cell iCOD)	½ year for memory more for cell
Pay-per-use (PPU)	<i>Not yet available</i>	July 2001	2+ years <i>and counting...</i>
PPU for storage	<i>Not yet available</i>	April 2002	1-½ years <i>and counting...</i>

4/26/2004

Partitioning Continuum Overview

65

Herzlichen Dank!

Engagement für den Kunden



4/26/2004

Partitioning Continuum Overview

66

