# New OpenVMS Virtualization Technologies

**Guy Peleg**

**iCAP/PPU Technical leader**

**Hewlett-Packard Company**

**guy.peleg@hp.com**

# OpenVMS & Virtualization

- We have been "Virtual" since day one
  - Virtual is part of our name

- *Virtual Batch*
  - Using generic queues, execute on the least busy node in a multi site cluster regardless of the site

- *Virtual Print*
  - Using generic queues, print on the least busy printer

- *Virtual Storage resources*
  - Ability to address a tape/disk with the same name from different nodes

# OpenVMS & Virtualization

- *Native virtual disk* (host based shadowing)
  - Provides capability to mirror data in active-active mode across disk controller pairs that might be local or physically located in different sites

- *Virtual Cluster Security*

- *Virtual process management*

- *Virtual Multi site clustering*
  - Active-Active virtual clusters that support 3000+ cpus in sites up to 800km apart.

# OpenVMS & Virtualization

- All these technologies are already built into OpenVMS today

- Talking to VMS customers about virtualization is preaching to the choir

*So why am I standing here?*

*Good question…..*

# Introducing new set of virtualization technologies designed for increasing resource utilization
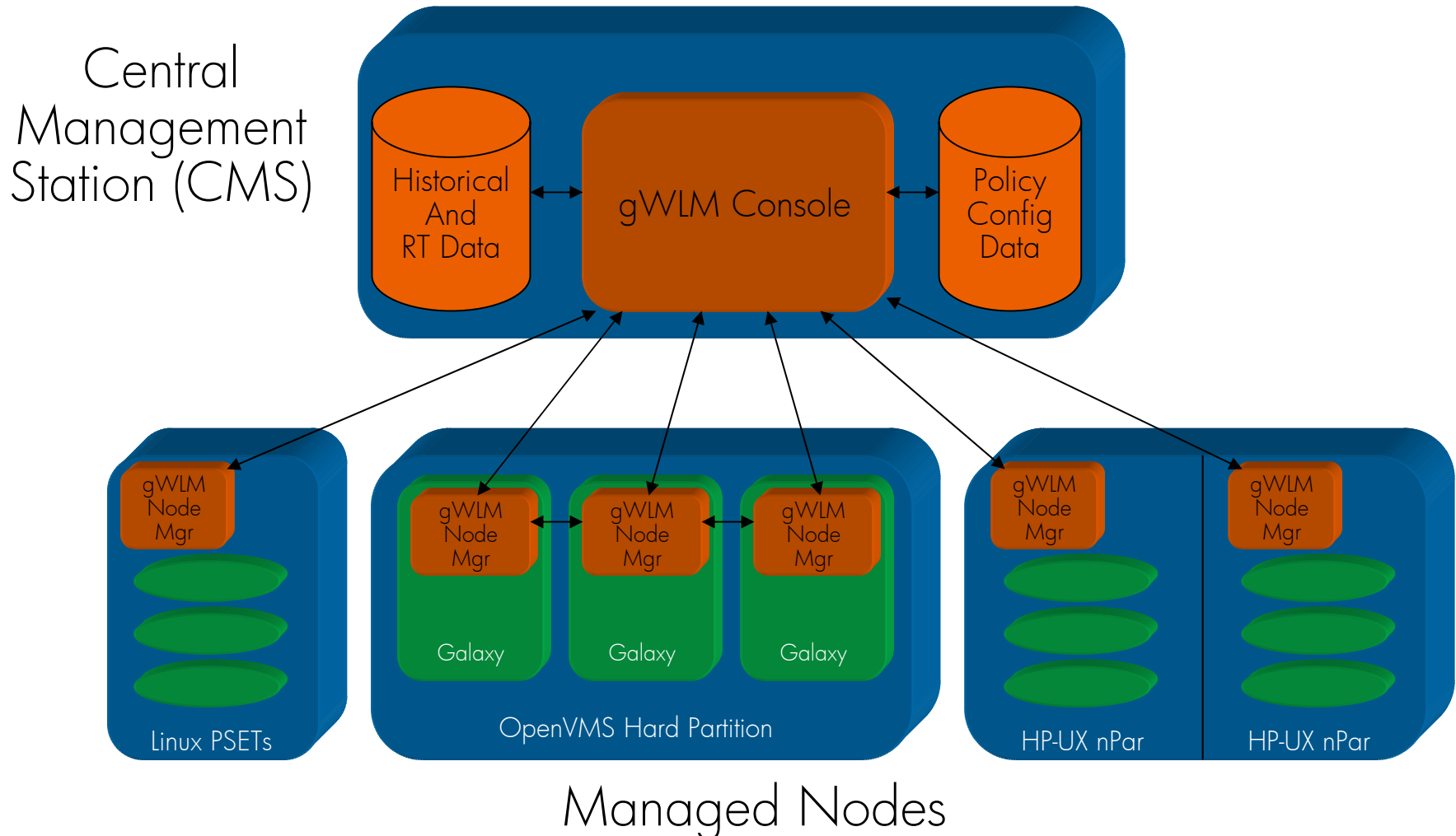
## gWLM

## iCAP & TiCAP

## PPU

## HP-VM

# Global Workload Manager

# (gWLM)

# Global Workload Manager
# gWLM High Level Architecture



Central Management Station (CMS)

Historical And RT Data

gWLM Console

Policy Config Data

gWLM Node Mgr — Linux PSETs

gWLM Node Mgr · gWLM Node Mgr · gWLM Node Mgr — Galaxy · Galaxy · Galaxy — OpenVMS Hard Partition

gWLM Node Mgr — HP-UX nPar

gWLM Node Mgr — HP-UX nPar

Managed Nodes

# gWLM Concepts
## Improving Resource Utilization

gWLM Concepts:

- **Workload**
  - group of processes, that run in a compartment
- **Compartment** -
  - OpenVMS Galaxy, Processor Set or Class Scheduler group
- **Shared Resource Domain**
  - a collection of compartments among which CPUs can be shared
- **Policy**
  - tells gWLM when to move CPUs, how much to give each workload

- The Administrator applies a Policy to each workload

- gWLM monitors each workload, estimates CPU demand, moves CPUs

# Saving Operators' Time
## to tackle new Service requests

- Many initiatives to improve utilization don't happen because System Administrators don't have the time to implement them.

- gWLM provides rapid time-to-value:
  - **Infrastructure Discovery and Configuration Wizard**
    - for rapid initial configuration
  - **Library of HP-Supplied policies**
    - suitable for many common configurations
  - **Out-of-the-box Reporting**
    - no additional configuration needed to get the reports you need to manage this environment
  - **Scalability**
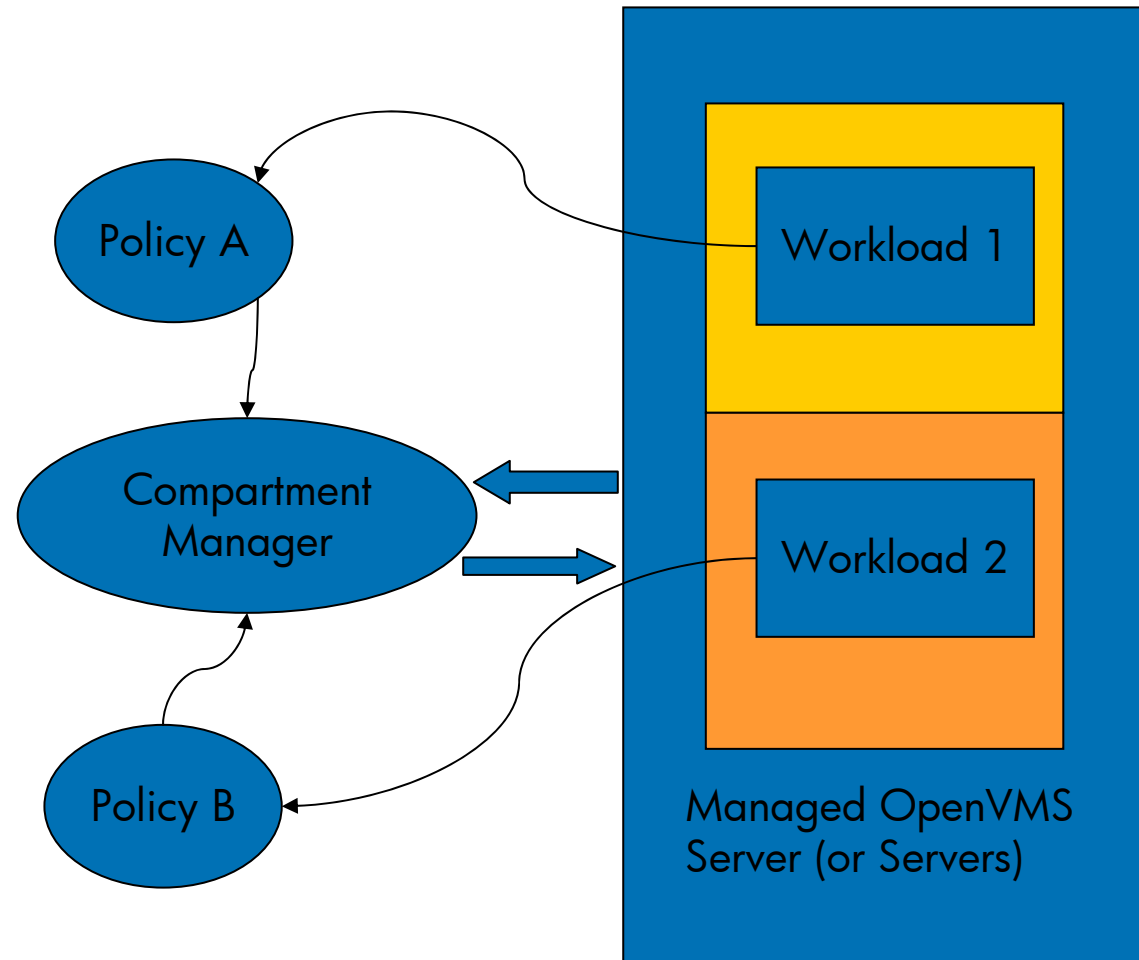    - a single policy can be applied to many workloads

**gWLM Design Goal**:

Manage existing systems in 30 minutes or less from the **first** install of gWLM

# gWLM System Overview
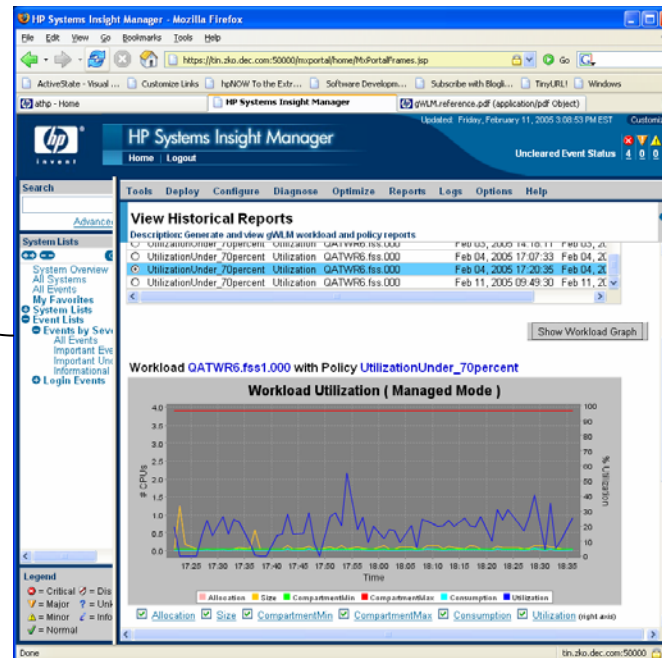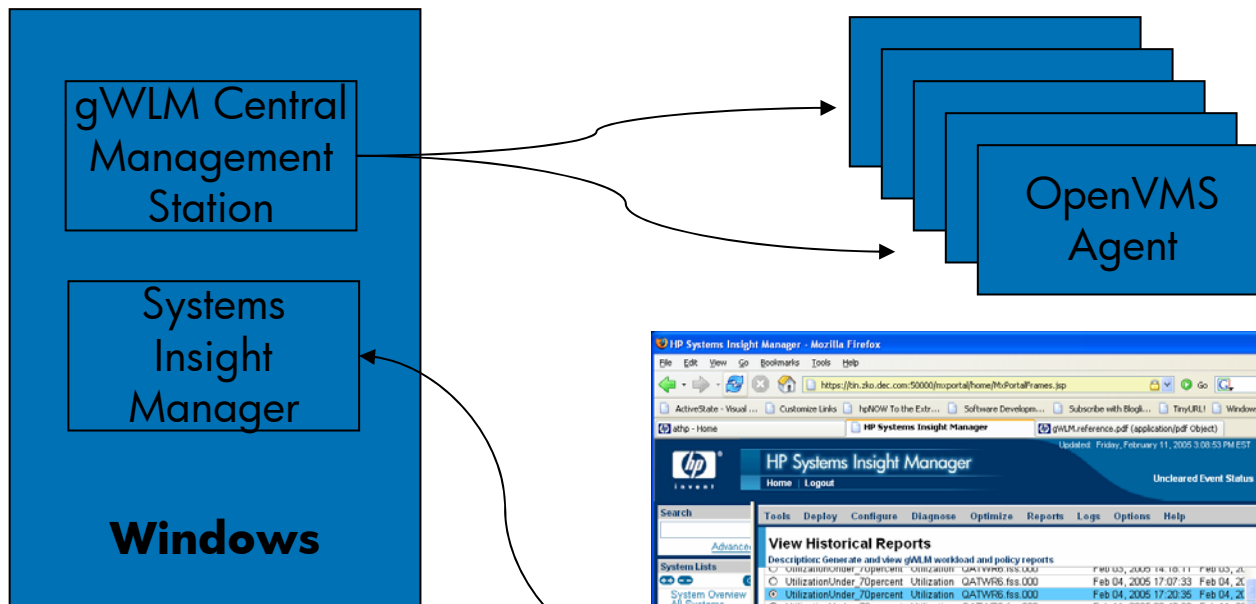
# Virtualization by Management of Compartments

# Global Workload Manager on OpenVMS

- Proto currently running on OpenVMS Alpha and Integrity
  - Target ship December 2005

- Able to manage:
  - Galaxy instances and provide Galaxy load balancing across clusters

  - Processor sets – redefine processor capabilities so whole processors are used by different workloads within an OS instance

  - Class Scheduler groups – redefine the percentage of CPU cycles used by workloads within an OS instance

  - 'gcu' and 'glm': gWLM automatically performs these functions on OpenVMS

# Global Workload Manager on OpenVMS - 2

- Separately orderable part for all the Operating Environments on Integrity Systems and priced per CPU (Per Processor Licensing – PPL)

- Available as Workgroup, Departmental or Enterprise license on AlphaServer

- Ships independent of major OpenVMS release

# Instant Capacity (iCAP) & Temporary instant Capacity (TiCAP)

# What is Instant Capacity iCAP?

**Definition:**

- Instant Capacity provides reserve capacity the customer can put into production quickly - without disrupting operations

**Key features:**

- Activate reserve capacity when needed
- Defer/avoid purchase until used
- Appropriate for **purchase/capital expenditure only**, not leasing

**Value proposition:**

- Reduces costs and simplifies the infrastructure
- Provides a highly available pre-configured "ready-to-run" solution
- Recognizes that speed to market is critical

# iCAP for OpenVMS how does it work?

- Customer pays a one-time Right To Access (RTA) fee per iCAP (inactive) processor (25% of list price)

- No activation commitment

- Once extra processing capacity is required, customer simply activates the processor with the Right To Use (RTU) license and pays the enablement discounted list price for that processor (75% of list price at the time of activation)

- Coexistence with HP-UX in the same box

**Value**

- No premium pricing

- Instant processing power with a single command

- Capability to load balance partitions at no additional cost

- Dynamically move iCAP processors within a server

# Supported Hardware

- The iCAP program supported on the following integrity platforms:
  - hp Integrity Superdome
  - hp Integrity rx8620
  - hp Integrity rx7620

- Target release – OpenVMS V8.3
  - Shipping with the base O/S

# iCAP Processor activation steps

- Purchase the processor from HP by sending purchase order

- Retrieve an activation codeword from the WEB portal

- Apply the codeword to the iCAP system

- Activate the iCAP component

# Temporary Instant Capacity TiCAP how does it work?

- Similar to the concept of pre-paid phone card

- Works with processors, does not include cell board & memory

- Customer orders standard iCAP processors and pays right to access fee

- Customer then purchases a 30-CPU day right to temporarily activate 1 or more iCAP CPU's

**Value**

- Enables the customer to temporarily activate a processor(s) for a set period of time

- No permanent activation fee is required – utilize an existing CPU at very low cost

- Accommodates those customers with unpredictable or planned processor demands

# Temporary Instant Capacity TiCAP Benefits

**\*Unique to industry**



"pay as you go" keeps costs in line with business demand
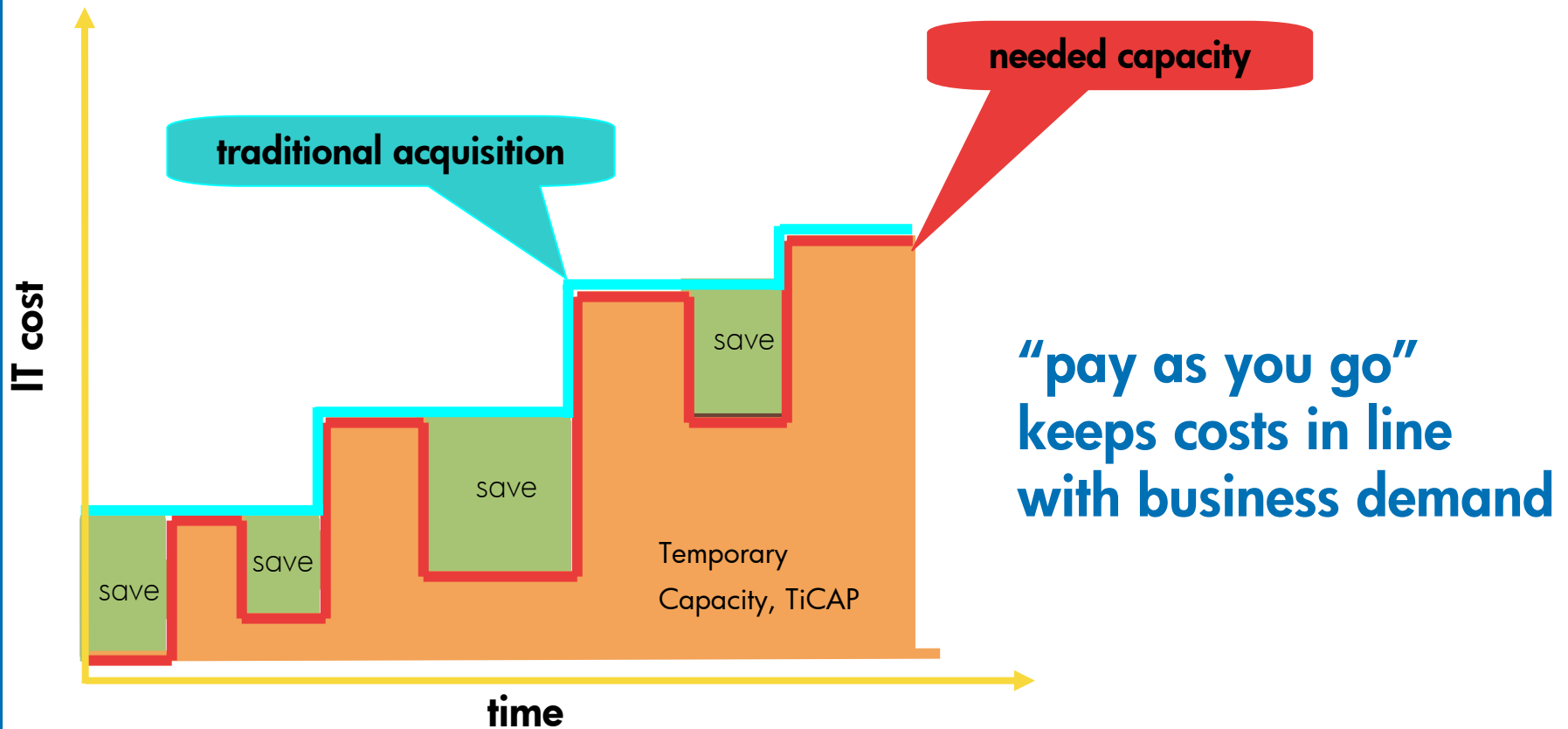
# iCAP Operational Considerations

HP software and 3rd Party Layered products

- The RTU license only covers hardware

- Regard iCAP activation as permanently adding a new CPU with all the attendant software licensing and support that this requires

Hardware Service

- When an RTU is purchased, support of the CPU is automatically added to the overall support costs of the system

# TiCAP Operational Considerations

HP and 3rd Party Layered products

- The relevant Operating Environment (OE) is automatically licensed on activated TiCAP CPUs

- Other HP and 3rd Party software each have their own policies and customers are recommended to purchase sufficient software licenses beforehand to cover their peak needs

Hardware Service

- Services are included for the CPUs activated by the temporary license

# TiCAP Operational Considerations 2

## Contribution to iCAP CPU Purchase?

- TiCAP 30 CPU Day license purchases **never** contribute to ownership of iCAP CPUs. Outright ownership is only accomplished by obtaining the iCAP RTU license.
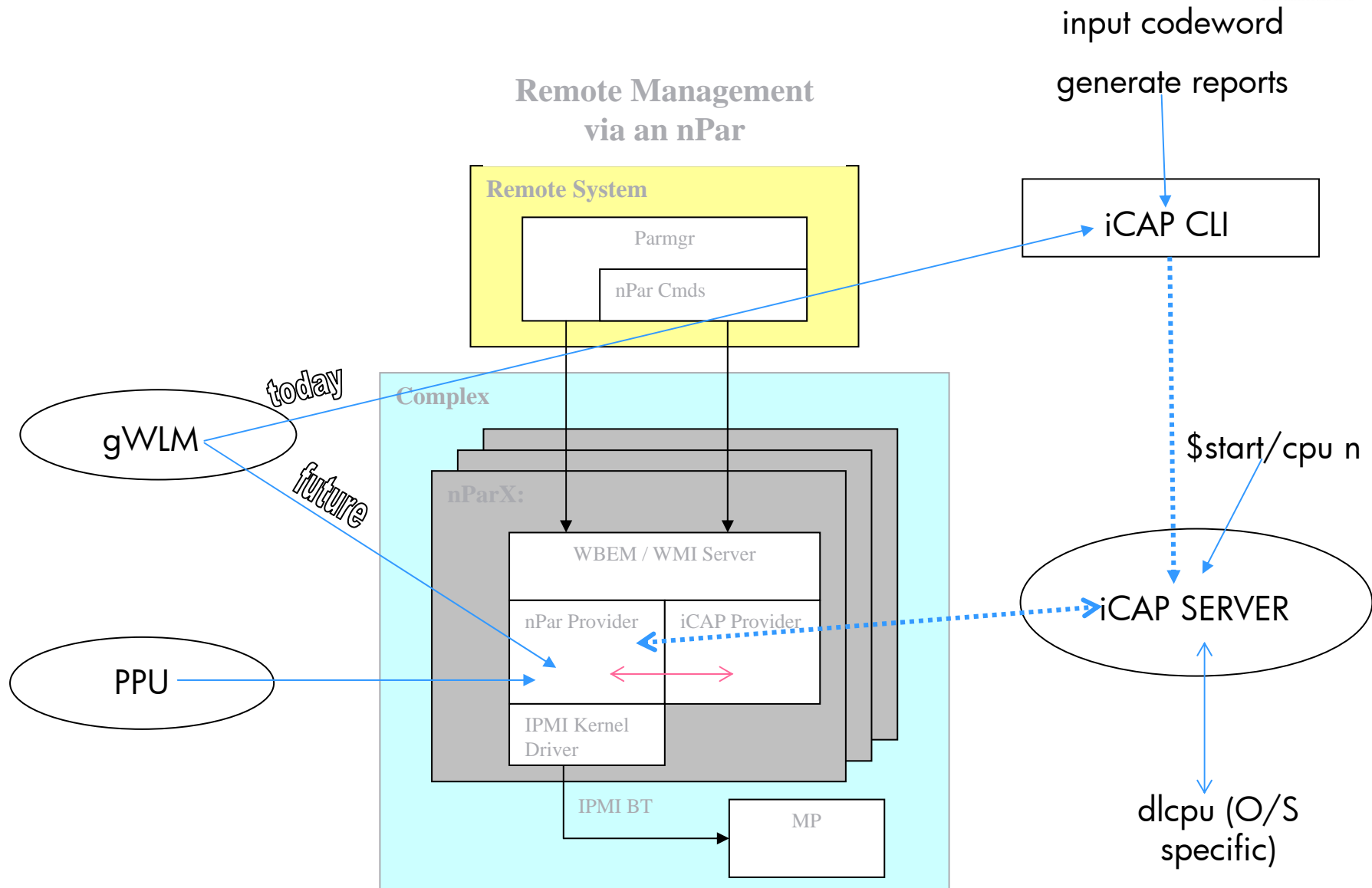
## Usage Tracking to avoid license exhaustion

- The iCAP software warns the system manager a few days before the TiCAP license is likely to expire, based on the rate of depletion that it tracks at all times. This is important if one TiCAP license is spread across more than one iCAP CPU.

- There is also an iCAP command to report usage whenever the system manager needs to know

# Related Technologies

- WBEM    (Web Based Enterprises Management)
  - Standard for managing systems (both hardware and software)

- nPar provider

- IPMI

- gWLM

# Architecture Overview



input codeword

generate reports

**Remote Management
via an nPar**

**Remote System**

Parmgr

nPar Cmds

iCAP CLI

**Complex**

today

gWLM

future

**nParX:**

WBEM / WMI Server

nPar Provider    iCAP Provider

IPMI Kernel
Driver

PPU

IPMI BT

MP

$start/cpu n

iCAP SERVER

dlcpu (O/S
specific)

# Pay Per Use (PPU)

# PPU (Pay Per Use)

- A **leasing** program under which customers pay only for computing capacity they use

- The PPU agent provides services for metering resource utilization
  - Actual percentage of utilization of each CPU
    - aka Percent PPU
  - A count of the number of active CPUs
    - aka Active PPU

- Information is sent to a utility meter
  - Separate IA32 system running Linux

# Pay Per Use
## Capitalizing on fluctuating demand
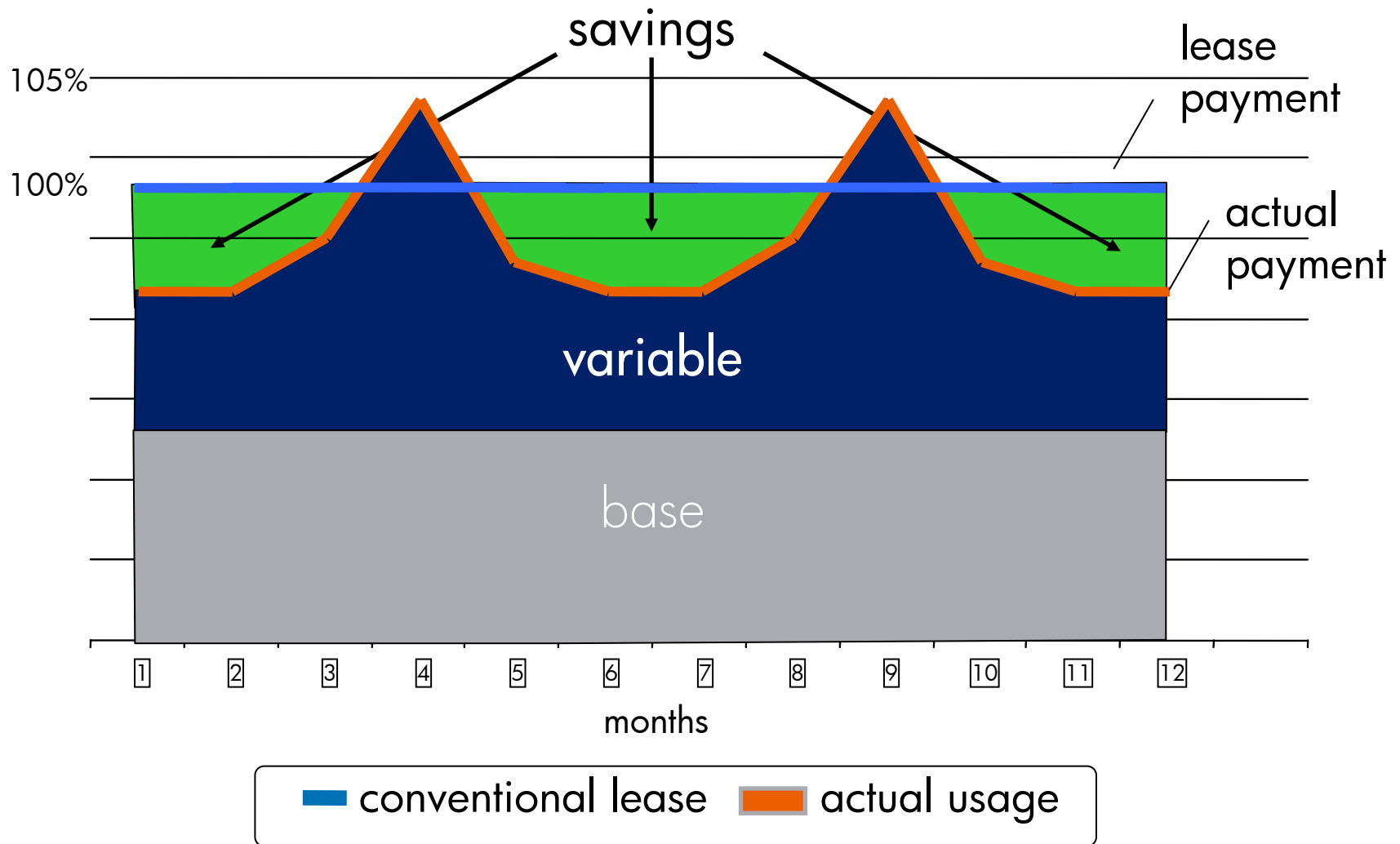
# Active CPU Billing Metric

- PPU metering software on HP server

- "Active" CPU is available for tasks by O/S

- Customers "light up" or "shut down" capacity

- Billing based on the monthly average of daily average utilization of Active CPU's

- **Lease only**

  Preferred by ISVs and the only one to be supported in the future by Oracle

# Percent CPU Billing Metric

- All CPUs running at all times

- Measures the % of each CPU used within a system

- Supports Soft Partitioning (Galaxy in the future)

- Billing based on monthly average of daily average % CPU utilization

- **Lease only**

  Preferred by customers as workloads are spread across all CPUs for good response

# PPU (Pay Per Use)

- The utility meter transmits utilization data to HP for proper billing

- After 48 hours the data is posted on the utility portal for viewing

- PPU software components are:
  - PPU agent (Will be shipping as a separate PCSI kit)
  - The utility meter
  - Usage database

# Virtualization through Partitioning

# HW Terminology – Processors/Cores

- The next chip generation, named Montecito, has 2 cores per processor

- The current rx2600 is 2P/2C

- Upgrading to Montecito will make it 2P/4C

- Upgrading the 4P/4C rx4640 will make it 4P/8C

- What happened to CPUs?

- From the OpenVMS viewpoint, what has always been seen as a CPU is now a core.

- $ SHOW CPU

- $ START CPU

# OpenVMS Naming of a 4P/8C



P   P   P   P

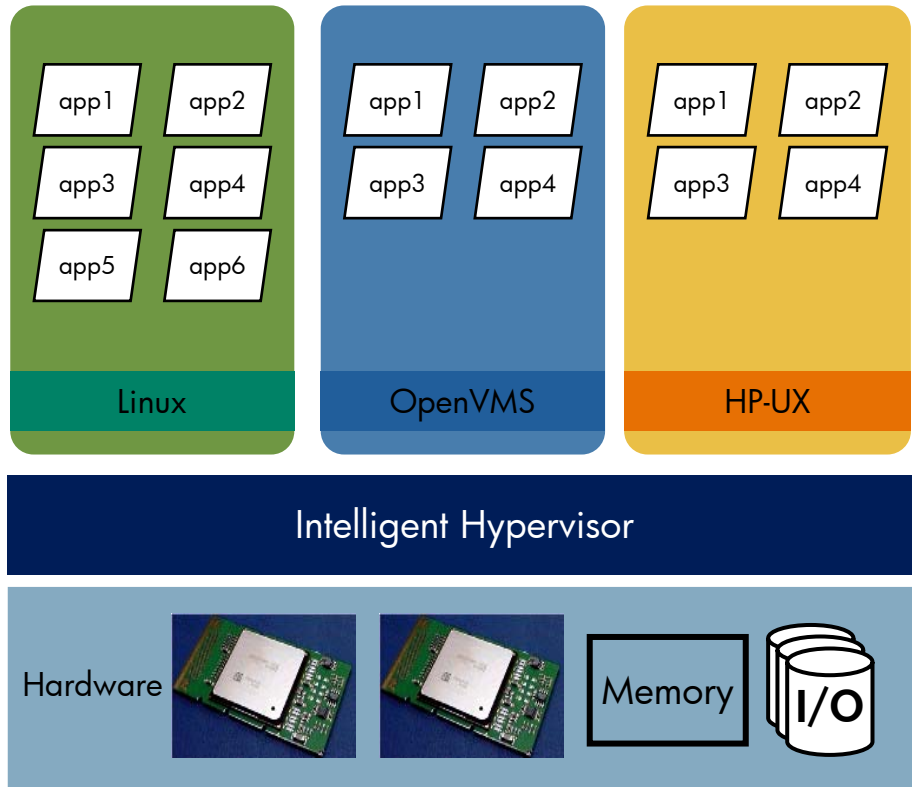| C | C | | C | C | | C | C | | C | C |

0   1   2   3   4   5   6   7
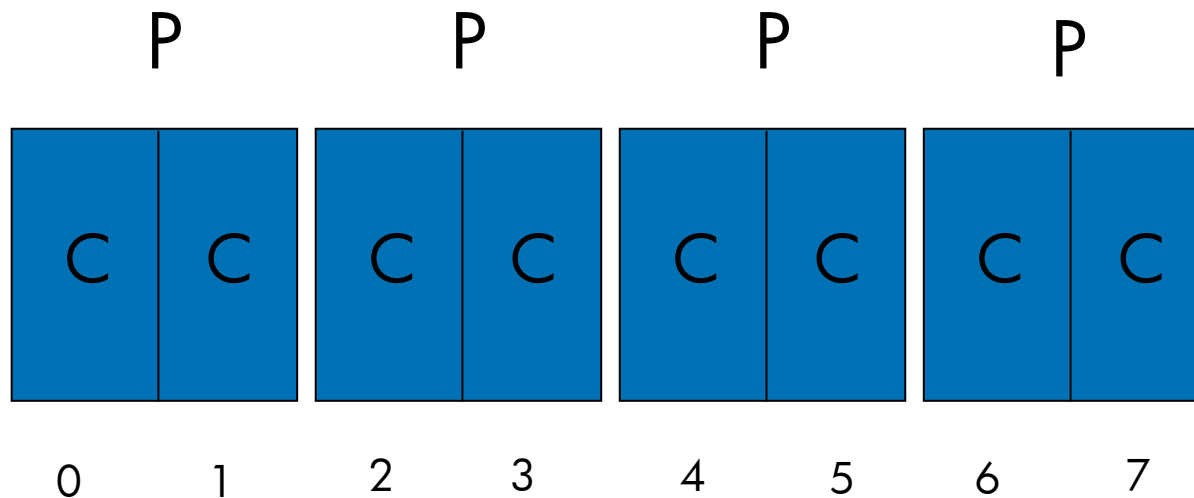
Active CPUs:    0-7

# HP Integrity Virtual Machines (HPVM)

- VT-i = Intel's Virtualization Technology for Itanium
  - Go to www.intel.com and search for vanderpool
  - Major topic at Intel Developer's Forum March '05

- Host Operating System
  - HP-UX
  - Controls all physical resources

- Virtual Machine Monitor
  - Software layer that presents a virtual Itanium system to each Guest
  - In OpenVMS terms, the VMM is roughly equivalent to two execlets

- Guest Operating System
  - HP-UX, Linux, OpenVMS, Windows
  - Requires no modifications
  - Is 'an application' to the Host OS

# HP Integrity Virtual Machines
## … optimum utilization across Multi OS



| Linux | OpenVMS | HP-UX |
|-------|---------|-------|
| app1 app2 | app1 app2 | app1 app2 |
| app3 app4 | app3 app4 | app3 app4 |
| app5 app6 | | |

Intelligent Hypervisor

Hardware — Memory — I/O

- Sub-CPU virtual machines with shared I/O
- Runs on a server or within an nPar
- Dynamic resource allocation
- Resource guarantees as low as 5% CPU granularity
- Guest fault and security isolation
- Supports all (current and future) HP Integrity servers
- Designed for multi OS – HP-UX, Linux , OpenVMS, Windows
- VSE integration for high availability and utility pricing

# What Exactly Is Being Virtualized?

- A generic Intel box, not a specific Integrity system
  - Generalized, lowest common denominator = versatile
  - But, don't get HP added value, e.g. I/O map registers
- CPU - Guest OS gets percentage of real CPU cycles
- I/O - A physical  I/O device is shared among Guests
  - Initial Release: All storage is virtualized as SCSI
  - Future Release: Option to dedicate I/O device to a Guest
- Memory – Guest OS gets its own physical memory

# Virtualization options

| | Lowest Granularity | Isolation | Maximize Resource | Multi-OS |
|---|---|---|---|---|
| Hard Partitions | Cell | Electrical | None | Yes |
| Soft Partitions | CPU | OS | CPU | No |
| Virtual Machines | Sub-CPU | Intelligent Hypervisor | CPU, Memory, & I/O | Yes |