# Module I:

# Performance Management in a Virtual Environment

**Virtualization management:**

Virtualization management is software that interfaces with virtual environments and the underlying physical hardware to simplify resource administration, enhance data analyses, and streamline operations.

**Difficulties.....**

- Logistical Problems
- Interpersonal Issues
- Technological Difficulties

**Similarities of every virtual machine environments:**

1. User interface
2. Streamline the virtual machine (VM) creation process
3. Monitor virtual environments
4. Allocate resources
5. Compile reports
6. Automatically enforce rules

**Responsibilities:**

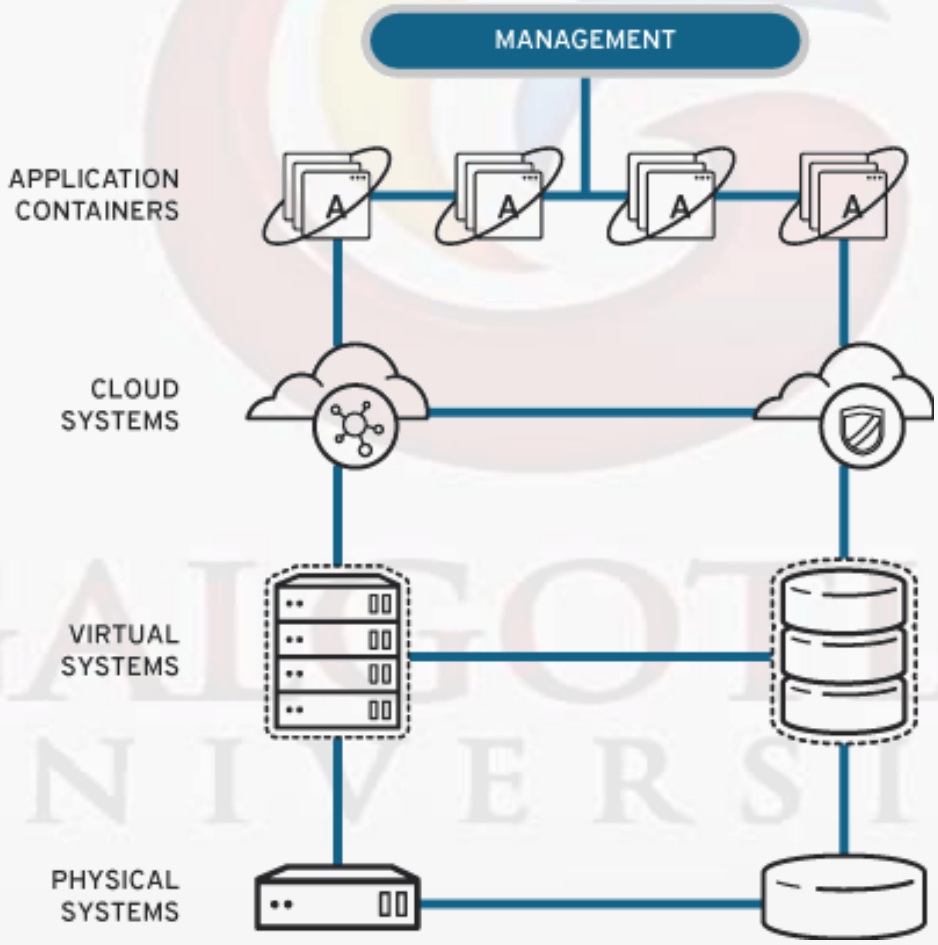
***Provisioning :*** Processing resource requests, creating templates, and configuring VMs.

***Compliance :*** Securing and monitoring systems, identifying issues, and validating user access.

***Operations :*** Retiring or reclaiming unused or underused physical resources, investigating bugs, and projecting future needs.

***Hybrid unification :*** Implementing the other 3 responsibilities across virtual, private cloud, public cloud, and container environments.

## Managing traditional virtual environments



**Predominant factors that contribute to virtuality:**

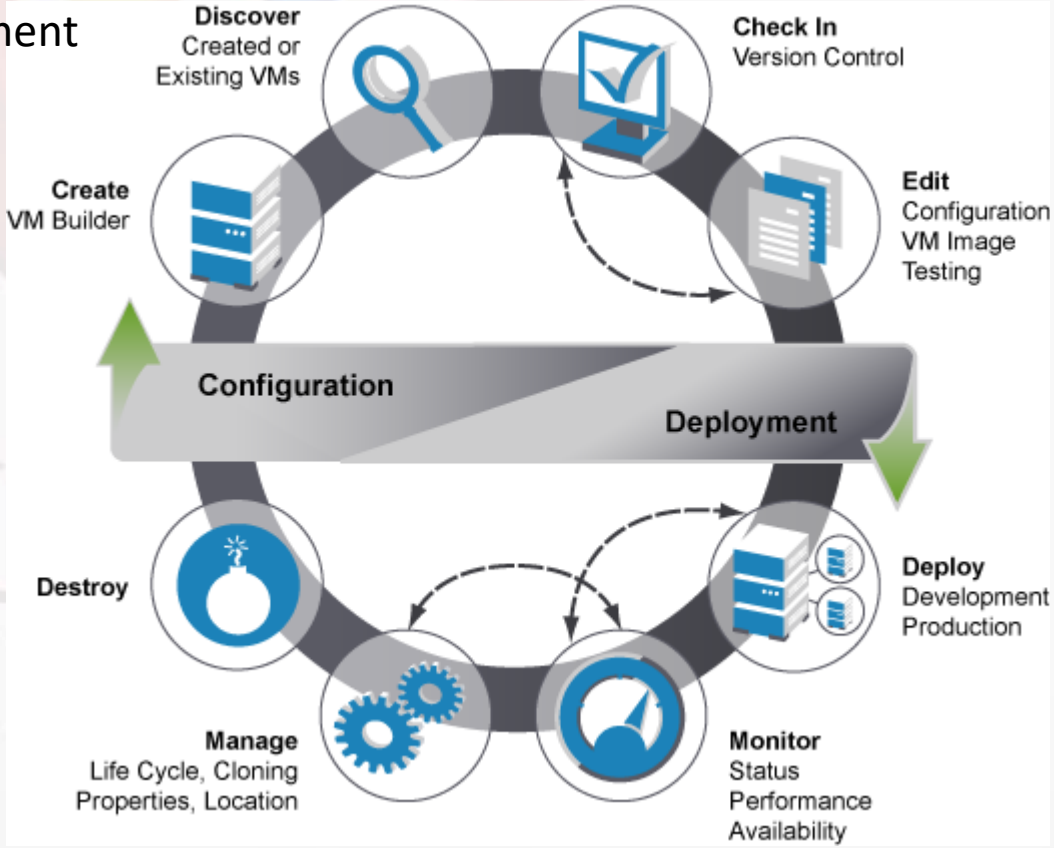
- The richness of communication media
- Distance between team members, both in time zones and geographical dispersion
- Organizational and cultural diversity

**Detriments….**

- Lack of potential for radical innovation
- The nature of the project may need to change
- Sharing of knowledge

## Life Cycle Stages of Virtual Management:

1. Preparations
2. Launch
3. Performance management
4. Team development
5. Disbanding



## Types of Virtual Teams

| Type of Team | Description |
|---|---|
| ☐ Network | Team membership is diffuse and fluid; members come and go as needed. Team lacks clear boundaries with the organization. |
| ☐ Parallel | Team has clear boundaries and distinct membership. Team works in short term to develop recommendations for an improvement in a process or system. |
| ☐ Project or Product Development | Team has fluid membership, clear boundaries, and a defined customer, technical requirement, and output. Longer-term team task is nonroutine, and team has decision-making authority. |
| ☐ Work, Functional, or Production | Team has distinct membership and clear boundaries. Members perform regular and ongoing work, usually in one functional area. |
| ☐ Service | Team has distinct membership and supports ongoing customer, network activity. |
| ☐ Management | Team has distinct membership and works on a regular basis to lead corporate activities. |
| ☐ Action | Team deals with immediate action, usually in an emergency situation. Membership may be fluid or distinct. |

## Standard Organizational and Team Process

- Definitions of Requirements
- Estimates of Costs
- Procurement
- Team Charters
- Project Planning
- Documentation and Document Sharing
- Reporting
- Controlling

## Virtual Team Management Techniques

1. Define work systems
2. Establish multiple communication tools
3. Schedule regular meetings
4. Have clear and detailed deliverables
5. Make sure work hours overlap
6. Create a professional work environment
7. Choose (video) calls over chatting and emails
8. Find the right people to work with
9. Establish a meritocratic system
10. Use project management tools

**Graphical user interfaces of administrative systems**

1. Lists, two-dimensional and on scrollable screens, typically using a windows program manager with many sublists showing printers, operating systems, physical sights, etc.
2. Tree diagrams showing the hierarchical relationships of the network system such as by showing the various geographical locations, the number of buildings at a location, the number of computers at each location, and the peripheral equipment associated with each computer and the systems being operated on each computer.
3. A hierarchical structure using folders and icons with each folder being a list of icons and with each icon by its color indicating the status of each unit.
4. Diagrams, with icons, of the various systems in a hierarchy.

## Elements of Virtual Management Environment

**User:** An end user who belongs to an organization whose virtual images are hosted on a private cloud server. The user has been granted some access privileges to some specific virtual images at specific times. A user is asked to authenticate at the User Portal. After authentication, he/she is presented with a list of virtual images he/she can access. After the selection, the user is allowed access to the virtual image, which is protected by the SSL VPN server.

**User Portal:** The portal provides a list of VMs that the user is eligible to access. The user portal generates a request to the Cloud Server to instantiate the VMs and expects the details of the VMs.

**IDP:** Authentication provider or identity service/identity provider (IDP). The IDP is capable of authenticating the users; this may be done by name and password combination, advanced authentication methods or partner services by means of federation protocols. The IDP is the role generator. The IDP also applies the VM information it receives from User Portal and applies it on a partial policy to make it a meaningful policy.

## *Continued……*

**SSL VPN Server:** The central server, which provides the access to the VM and enforces the policies every time an access is made to the VM by the user.

**Policy Enforcer:** A policy evaluator. Takes the required input and evaluates the policy and responds with a result.

**Virtual Image Private Cloud Server:** The central server for hosting virtual images. It instantiates the VMs on an as-needed basis. After instantiating the VM image, the cloud server presents its IP Address, Port, and a secure Token from the VM. This token lives for a particular period of time and access to the VM image requires the token.

**Virtual Server Management**

Identify and solve performance issues, optimize resource utilization, prevent hardware failure, and maintain virtual servers at peak performance

**Challenges of Virtual Server Management**

- VM Sprawl
- Network Traffic Congestion
- Server Hardware Issues

***Tool(s):*** *"**OpManager"** for* Vmware performance and issue monitoring

- OpManager Performance Monitoring
- PostgreSQL Database Statistics
- OpManager Configuration Details
- Fix Performance Problems Faster

***Virtual Management Techniques***

1. Virtual server monitoring
2. VM relationship mapping
3. Capacity planning
4. VMware workflow automation
5. VM migration tracking
6. Alerts via email and SMS
7. VM reports

**VMware ESX Server and ESXi Server**

| Features | ESX Server | ESXi Server |
|---|---|---|
| Service Utilities | Service console (esxcfg) – Linux based utility | Remote Command line Interface (RCLI) |
| Boot up timing (For server) | More time (more than 2 GB in size) | thin = fast installation + faster boot (1 – 2 minutes) |
| Purchasing Mode | Free of cost | Embedded as a hypervisor on flash drive or microchip |
| Security profiles | More number of security options | Setup is very easy and simplistic |
| Firmware Console | Service Console | Direct Console User Interface (DCUI) |
| Health Status | Third party tools are used | Built in health status |
| Patches | More number of patches are installed and more number of reboots | Less number of patches required with few number of reboots |


**Figure 1:** ESXi Security Profile

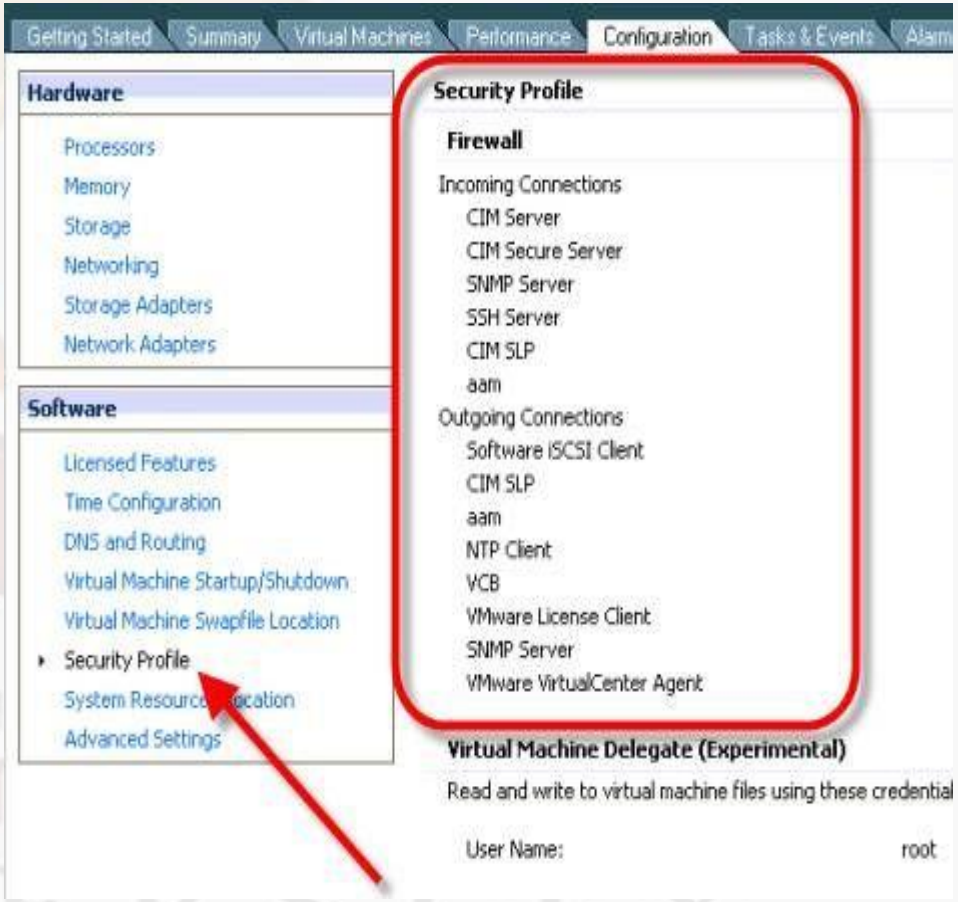
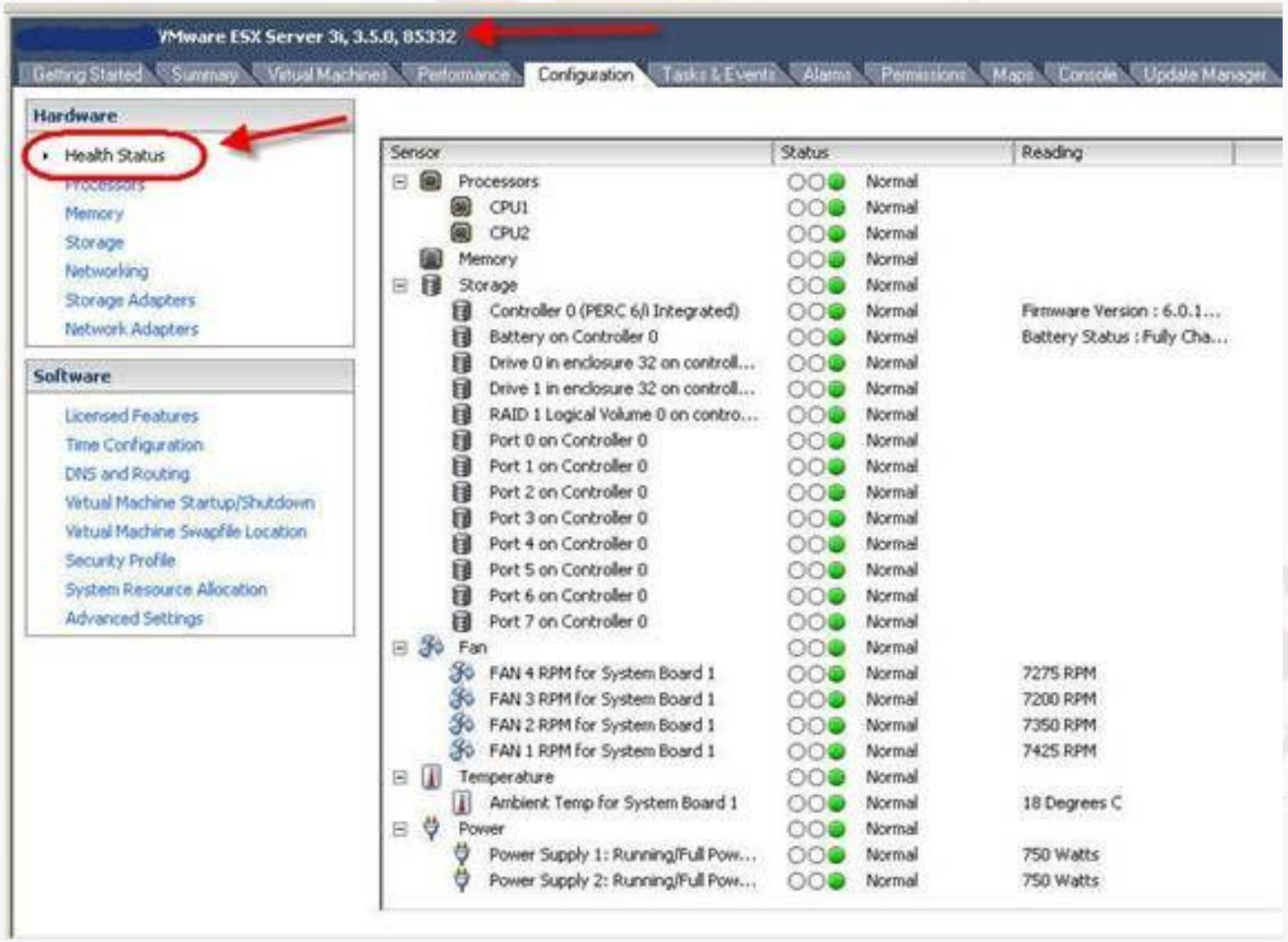


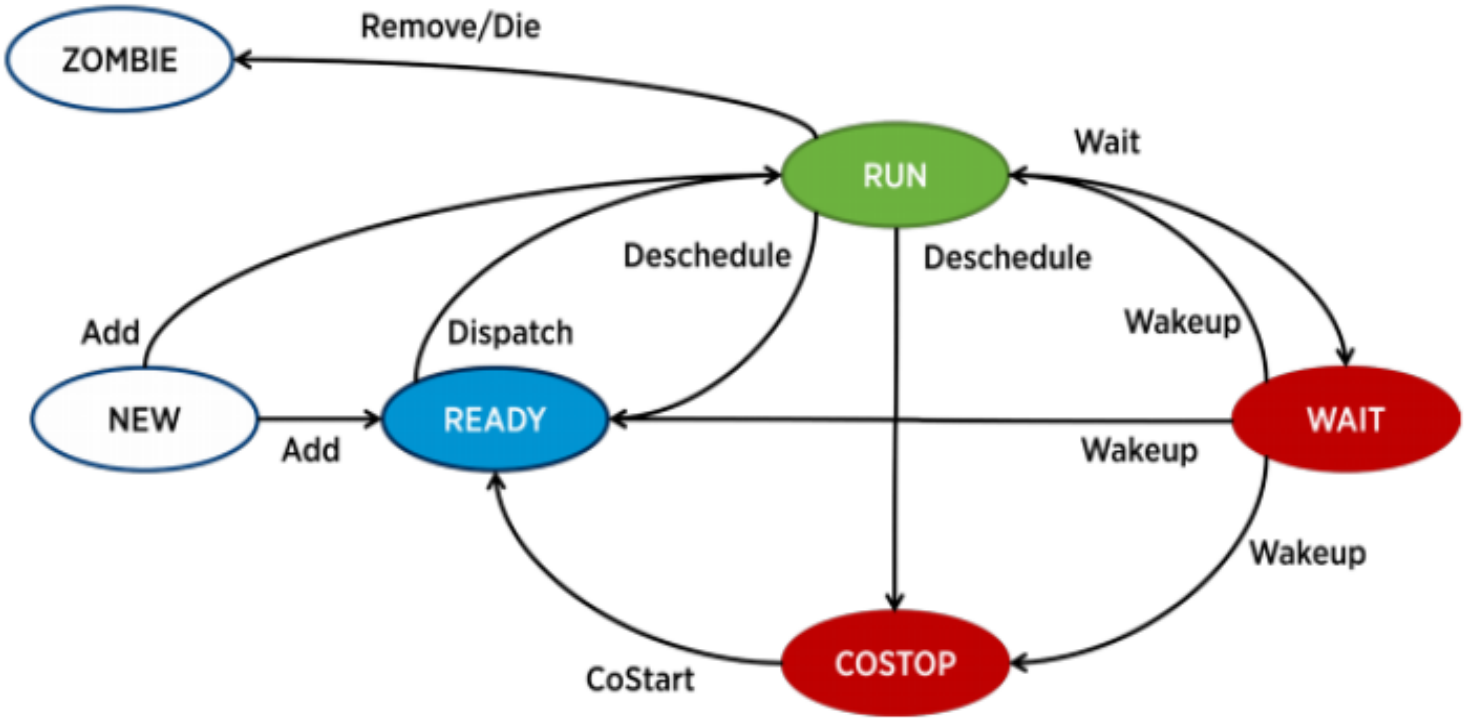

**Figure 2:** VMware ESX Server (full) Security Profile

**ESXi yellow firmware console / DCUI**

## ESXi Health Status



# Host CPU Co-Scheduler World



# CPU Key Performance Indicators

## CPU Key Performance Indicators for ESXi Hosts

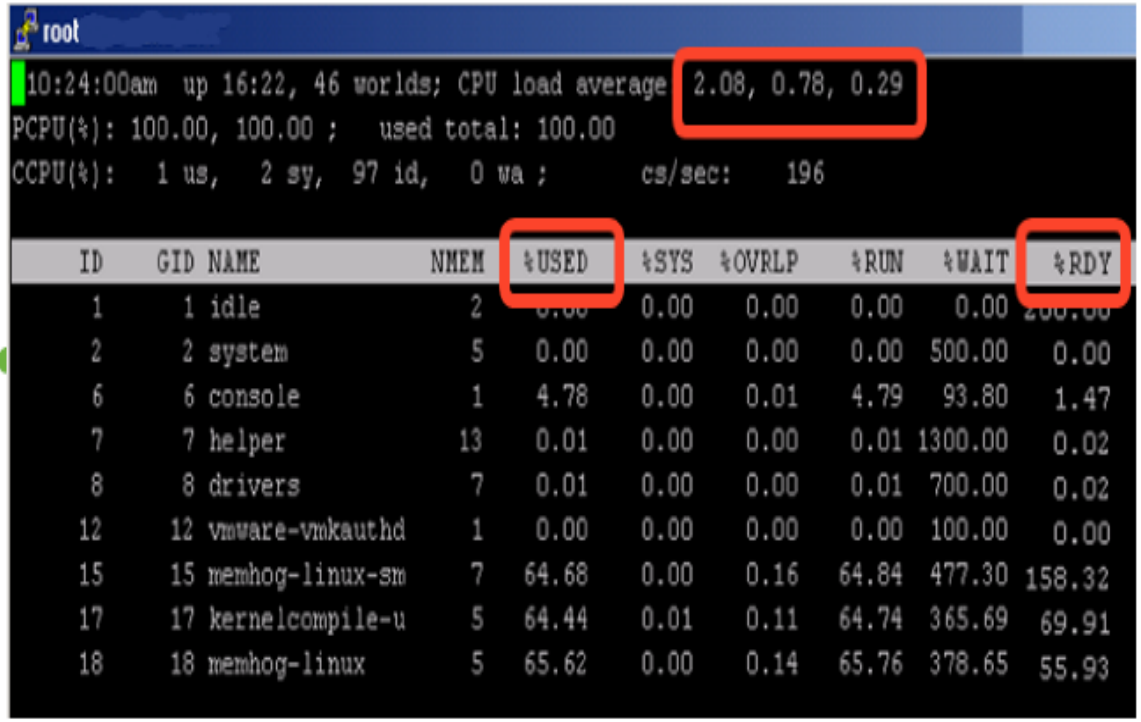
- Ready Time
- Utilization
- Load Average

## CPU Key Performance Indicators for Virtual Machines

- Ready Time (%RDY)
- Co-Stop (%CSTP)
- Swap Wait (%SWPWT)
- Memory Limited (%MLMTD)

# Example: Identifying CPU Constraint

The %USED and %RDY columns of the `resxtop` command output indicate CPU overcommitment.



# When to Right Size Virtual Machine vCPUs

## What is Co-Stop?

When a VM with multiple vCPUs must stop processing on one or more vCPUs.

## Why does Co-Stop Occur?

The fastest sibling vCPU stops itself when it's slowest sibling vCPU on the VM violates a threshold. This is due to "skew" between sibling vCPUs. The vCPUs co-start when the slowest sibling begins to make progress. It progresses because scheduling opportunities are available once the fastest vCPUs are co-stopped.

## How do I resolve CPU Co-Stop issues?

Right size your VM's vCPUs. When in doubt, mimic the physical host CPU topology to take advantage of physical/virtual NUMA.

## Consider "Wide and Flat" vCPU allocations.

For configurations for VMs with greater than 8 vCPUs, allocate "X" number of virtual sockets and a single virtual core.

**The factors Affecting CPU performance are :**

- Multiple cores
- Cache memory,
- clock speed,
- word length,
- address bus width,
- data bus width.

## Memory Key Performance Indicators:

1. Transparent Page Sharing (TPS)
2. Ballooning
3. Memory Compression
4. Swapping

## VM Memory: (Intra VM and Inter VM)

1. Virtual Memory
2. Guest Physical Memory
3. Physical Memory

- Mem.ShareScanTime
- Mem.ShareScanGHz

***Guest Memory Basics:***

- ✓Memory Size
- ✓Allocated Memory
- ✓Active Memory
- ✓Idle Memory
- ✓Free Memory

***Potential Memory Problems:***

- ✓Too much / Little vMemory
- ✓vSwap
- ✓Balloon
- ✓NUMA
- ✓Resource Pools
- ✓Reservation, Limits and Shares
- ✓Limited Hosts / VM Memory Resources
- ✓Too Many VMs on a Host

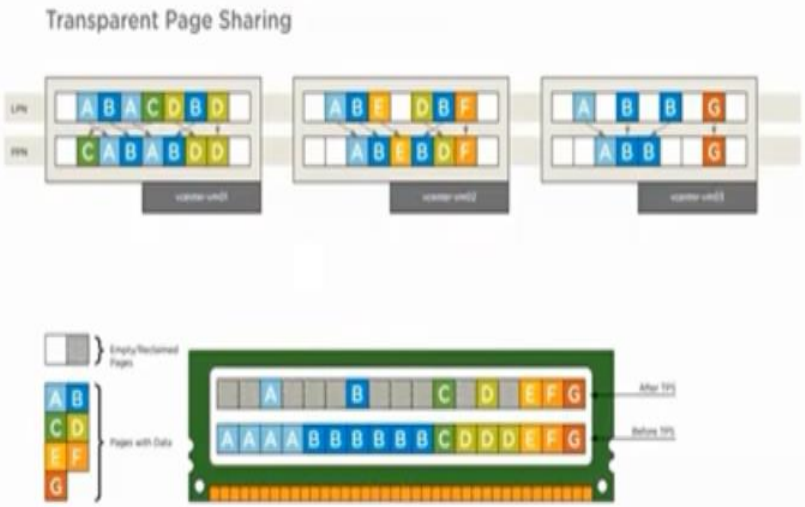
| If hosts free memory drops toward the following threshold | Threshold | Use the following reclamation technique |
|---|---|---|
| High | 6% | None |
| Soft | 64% of MinFreePct | Ballooning |
| Hard | 32% of MinFreePct | Ballooning & Compression |
| Low | 16% of MinFreePct | Ballooning, Compression & Swapping |

## Key Metrics to Monitor:

- Host swapping
  - vSphere Client
    - Memory swap in rate
    - Memory swap out rate
  - resxtop/esxtop
    - SWAP /MB
    - SWR/s and SWW/s
    - SWCUR and SWTGT
    - %SWPWT
- Host ballooning
  - vSphere Client
    - Balloon
  - resxtop/esxtop
    - MEMCTL/MB
    - MCTL?
    - MCTLSZ
    - MCTLTGT

# Transparent Page sharing (TPS)

- Transparent Page sharing shares memory pages on same or across multiple VMs running on the same host.




***"sched.mem.pshare.salt" –*** *Controlling the ability to do TPS*

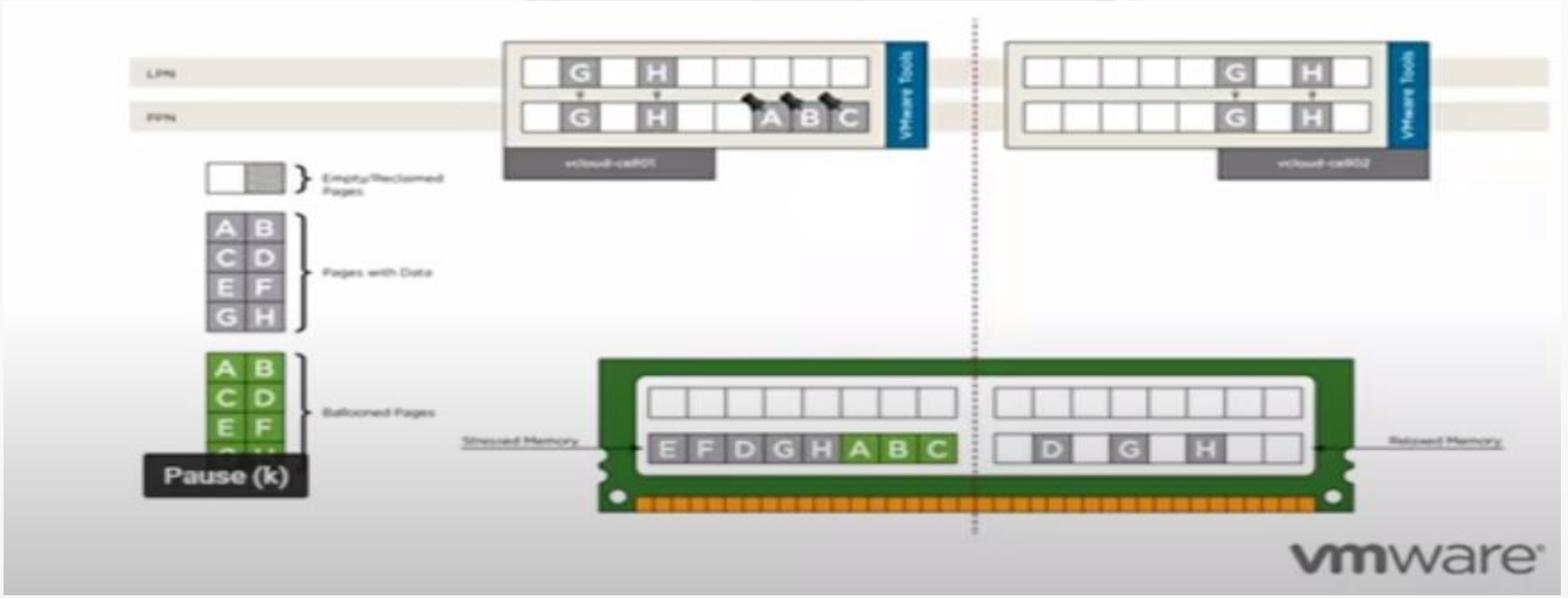
**Mem.ShareForceSalting ,0 :** To enable TPS on all the VM

**Mem.ShareForceSalting ,1 :** Preventing VM with other VM by restricting settings

**Mem.ShareForceSalting ,2 :** To enable TPS on some specific set of VMs

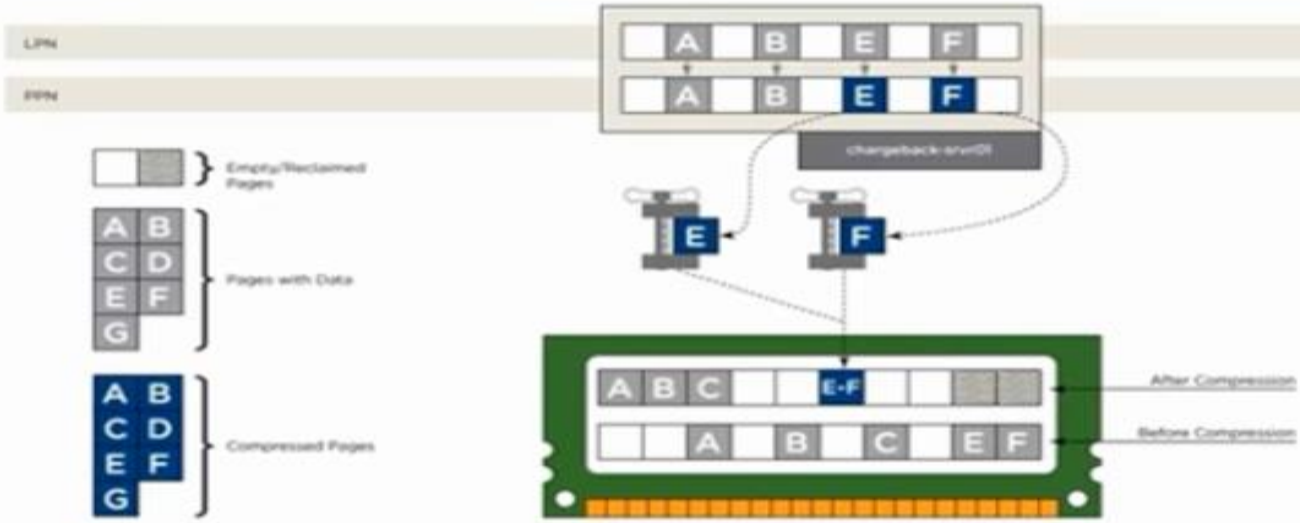
# Ballooning

- Ballooning: can be used when we have VMware Tools running
- Can cause performance increase or decrease as the Guest VM has to use its own memory technique swapping etc.



# Memory compression

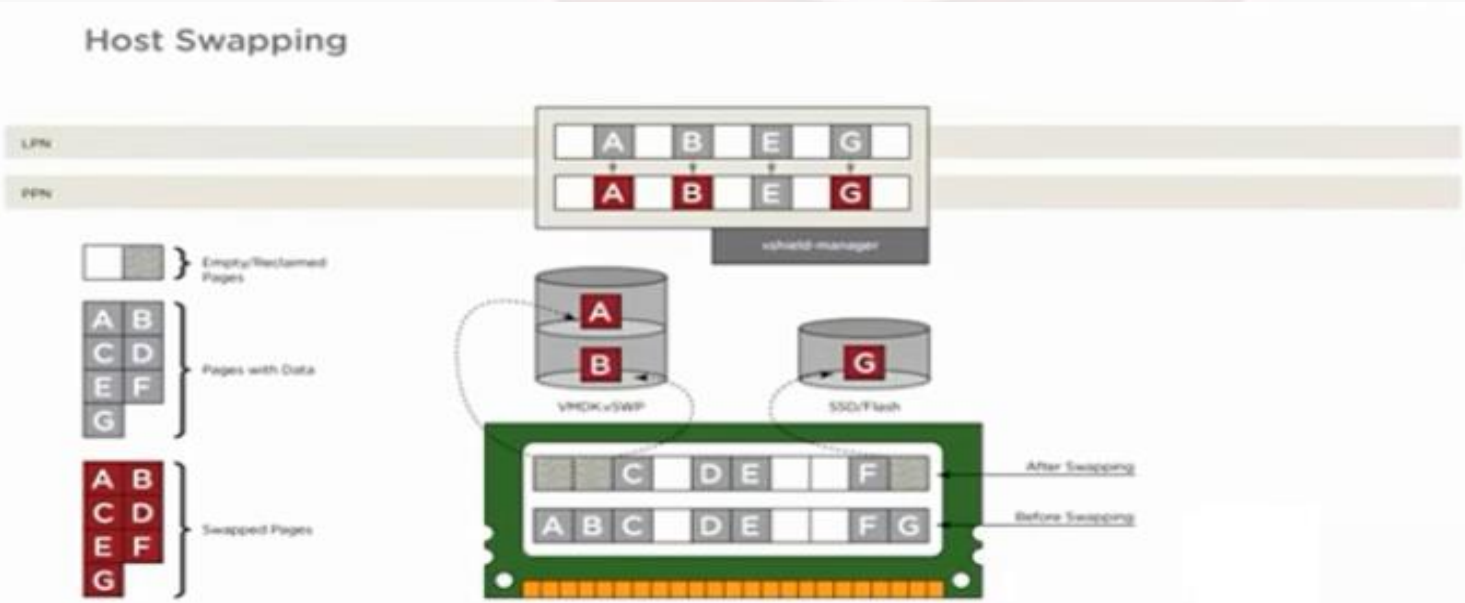
* Host compress the memory but this has an impact as the compress memory takes longer time to get access read / write.




vmware

# Swapping

* Host swapping memory this is used as a last resort as can cause serious performance issues.
* The memory is written to the physical disk

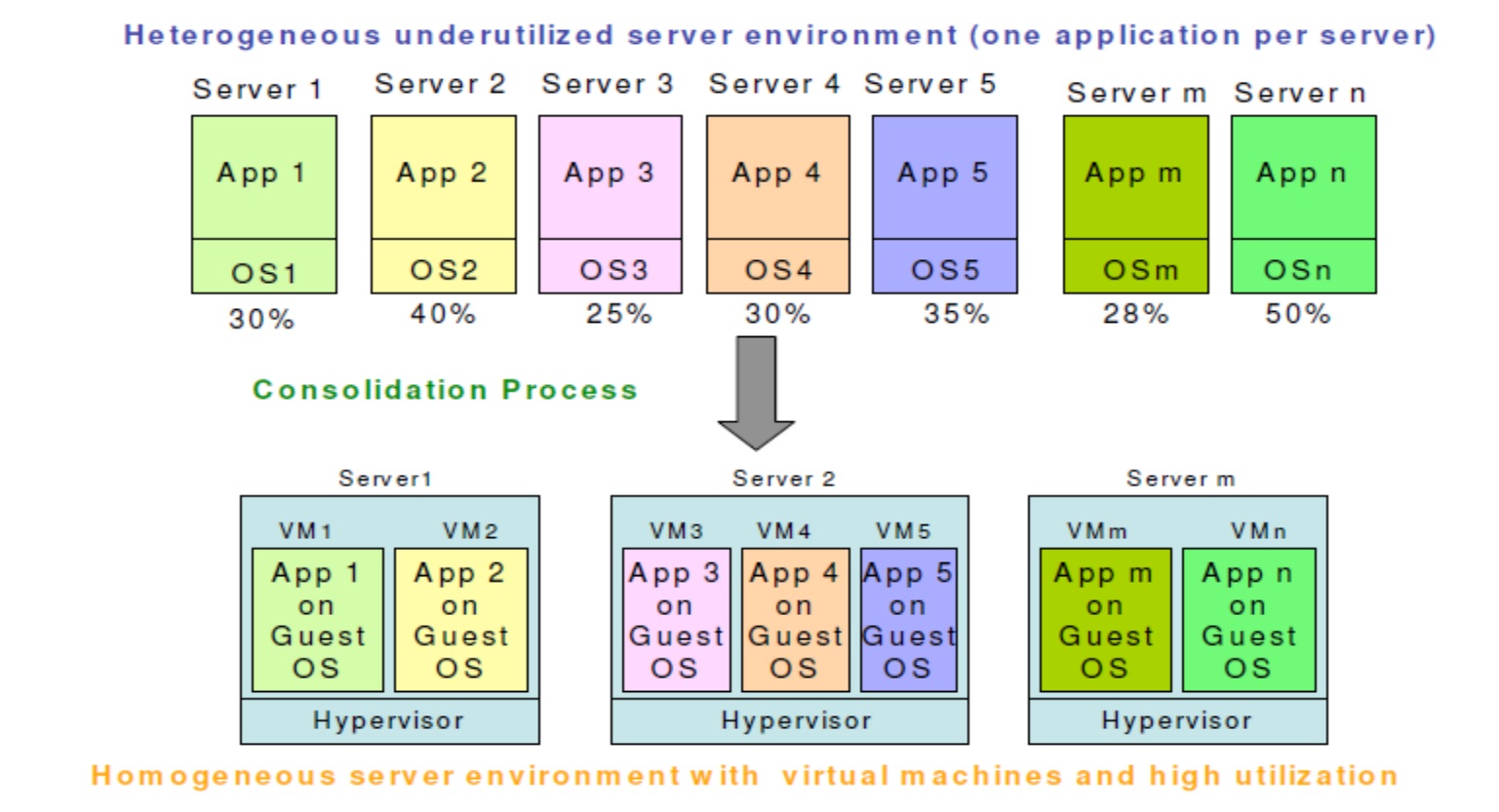



vmware

## Application Performance Management in Virtualized Server Environments

1.Server Sprawl 2. Server Virtualization 3. VM Migrations

4. Maintain Levels 5. SAN and SLA



**System Management Functions:**

- ✓Observation of Performance Key Metrics
- ✓Dynamic Re-allocation of SLA

**CPU Contention**

- ✓Increase the guest's CPU priority
- ✓CPU quota
- ✓Run fewer guests on the host

- •Steal Time (/proc/stat → "%st“)
- •Response Time

## Application Performance Issues

***Symptoms:***

- ✓Services running in guest virtual machines respond slowly.
- ✓Applications running in the guest virtual machines respond intermittently.
- ✓The guest virtual machine may seem slow or unresponsive.
- ✓The guest operating system boots slowly
- ✓Applications running in virtual machines perform poorly
- ✓Applications running in virtual machines take a long time to launch
- ✓Applications running in virtual machines frequently become unresponsive
- ✓Multi-user services have long transaction times or can handle less simultaneous users than expected

***Causes: (Bottleneck Areas)***

- ✓CPU Constraints
- ✓Memory Over commitment
- ✓Storage Latency
- ✓Network Latency

## CPU constraints

1. Examine the ***load average*** on the first line of the command output. (1.00, 0.5, ***2.00***)
2. Examine the %READY field for the percentage of time that the virtual machine was ready but could not be scheduled to run on a physical CPU. (Under 5%)

- •CPU Limit
- •Resource Pool

➢Increase the number of physical CPUs on the host

➢Decrease the number of virtual CPUs allocated to the host

**Memory overcommitment:**

1. Examine the ***MEM overcommit avg*** on the first line of the command output.
   - ✓Increase the amount of physical RAM on the host
   - ✓Decrease the total amount of RAM allocated to all of the virtual machines on the host

2. Determine whether the virtual machines are ballooning and/or swapping.

***To detect any ballooning or swapping:***

1. Run esxtop.
2. Type **m** for memory
3. Type **f** for fields
4. Select the letter **J** for Memory Ballooning Statistics (MCTL)
5. Look at the MCTLSZ value

***MCTLSZ (MB) displays the amount of guest physical memory reclaimed by the balloon driver.***

1. Type **f** for Field
2. Select the letter for Memory Swap Statistics (SWAP STATS).
3. Look at the SWCUR value. ***SWCUR (MB) displays the current Swap Usage.***

**Storage Latency:**

- ✓ Determine whether the problem is with the local storage and Migrate the virtual machines to a different storage location
- ✓ Reduce the number of Virtual Machines per LUN.
- ✓ Look for log entries in the Windows guests that look like this:

  The device, \Device\ScsiPort0, did not respond within the timeout period.

- ✓ Using esxtop, look for a high DAVG latency time.
- ✓ Determine the maximum I/O throughput you can get with the iometer command.
- ✓ Compare the iometer results for a VM to the results for a physical machine attached to the same storage.

**Network latency:**

- ✓Test the maximum bandwidth from the virtual machine with the Iperf tool.
  - ✓Run Iperf with a machine outside the ESXi/ESX host. (PM Verification)
  - ✓Run Iperf with another machine outside the ESXi/ESX host on the same VLAN on the same physical switch. (Reproducing problem)
  - ✓Run Iperf between 2 VMs on the same ESX server/portgroup/vswitch.
  - ✓If the result is good, you can exclude a CPU, memory or storage issue.
- ✓Proper Configuration
- ✓Shares and Limits Configuration
- ✓Traffic Shape Configuration