

Performance Tuning of Virtual Servers – TAC9872

John A. Davis

Senior Consulting Engineer



VMWORLD 2006

Introduction

- Main Goals:
 - Discuss Performance Tuning of Virtual Servers
 - Involving ESX Server and Virtual Center Environments
 - From the Perspective a VMware Authorized Consultant
 - Basic Concepts, Tools, and Technologies
 - Recommendations

- Special Focus:
 - ESX Server is Well Designed to Dynamically Share Resources
 - VI 3 Offers New Features, such as DRS
 - How Do We Choose the Best Optimization Approach?

Speaker

- John A. Davis
 - Senior Consulting Engineer
 - Microsoft Networking, Development and Training Background
 - VMware Certified Professional and Instructor
 - 50% Consultant and 50% Instructor
 - National and International Exposure
 - Small, Mid-size, and Large Customers
 - Short Term Engagements and Life Cycle Projects

- New Age Technologies
 - IT Solution Provider
 - Louisville, KY
 - Networking, Development, and Virtualization

Audience

- Technical Experience - Beginners to Intermediate
- Engineers, Managers, and Decision Makers
- Competitors and Other Consultants

Those who are considering where to start with performance tuning and those wondering if they are taking the best approach.

Presentation Outline

- Resource Sharing in ESX Server
- Performance Tuning Settings per VM
- What's New in VI 3.0?
- Resource Pools
- Third Party Tools
- Recommendations
 - Selecting the Right Tools
 - Thresholds
 - VMware HA
 - VM Configuration
 - Other

Resource Sharing in ESX Server

- Four Main Resource Types
 - CPU
 - RAM
 - Disk I/O
 - Nic

Consider what happens when we over commit our VM configurations and multiple VMs demand resources simultaneously. The ESX Server must choose which VMs get access to the resources first.

CPU Sharing

- Virtual CPUs Map to Physical CPUs
- Migration Occurs Every 20 Milliseconds
- Hyperthreading – CPUs Split into Two Logical CPUs
- Cores Treated as Physical CPUs

RAM Sharing

- Limit and Reservations

- Formerly known as Min and Max
- Granted RAM is Between Min and Max
- VMkernel Can Adjust the Actual Used Amount

- Transparent Page Sharing

- Some VMs may have Virtual RAM pages with identical content. This page is stored in physical RAM only once and each VM has a pointer to the page. If one VM changes its page, the page is first copied, the VM is redirected to the new page, and then the changes are written.
- Reduces the Load on the Physical RAM
- May Reduce by up to 30%

RAM Sharing (Continued)

- Ballooning
 - RAM can be borrowed from one VM and given to another VM.
 - Requires VMware Tools
 - VM's Guest O/S Identify Which RAM to Reclaim
 - Guest O/S May Swap Data Internally

Disk Sharing

- Disk space is pre-allocated
- Disk bandwidth (Disk I/O) is shared
- Bandwidth includes Adapter I/O bandwidth and Drive I/O speed

Network Sharing

Implemented by using Virtual Switches and Virtual Port Groups. Virtual Nics are connected to Port Groups configured on Virtual Switches, which are connected to physical Nics.

- Single Nic Connected to Virtual Switch
- Nic Teaming – provides load balancing and redundancy
 - Mac Based – default in 2.5
 - Port Based – default in 3.0
 - IP Based – optional in either version

Performance Tuning per VM

■ CPU

- Shares – priority under contention
- Reservation – guarantee
- Limit – maximum
- Processor Affinity – pins VMs to specific CPUs
- Hyperthreading – doubles the number of logical CPUs

■ RAM

- Shares – priority under contention
- Reservation – guarantee (This much RAM or more will be allocated to the VM.)
- Limit – maximum (This is the amount of RAM the guest O/S sees.)

VM Performance Tuning (Continued)

- Disk
 - Shares – priority under contention

- Network
 - Traffic Shaping – limit the amount of outbound traffic from a single VM:
 - Average bandwidth
 - Burst Size
 - Peak Rate

What's New in VI 3?

■ Changes

- CPU Min and Max is now called Reservation and Limit
- RAM Min and Max is now called Reservation and Limit, but with these changes:
 - By default, RAM Reservation is set to zero
 - When a VM starts, the ESX Server grabs an amount of RAM larger than the Reservation
 - (Previously, the Min was set to 50% of the Max, by default and that amount was initially allocated)
- Nic Teaming load balancing is now Virtual Port Based

What's New in VI 3?

- New Features
 - Host based Resource Pools
 - DRS Clusters
 - DRS Clusters based Resource Pools

Resource Pools

- Host Based Resource Pools
Sub-divided ESX Servers
- DRS Clusters
Group of ESX Servers with Combined Resources
- DRS Cluster Based Resource Pools
Sub-divided DRS Cluster

Host Based Resource Pool

- Details:
 - Allows Shares, Reservations, and Limits to be configured per pool
 - Allows these settings for CPU and RAM usage
 - Allows the creation of Child Pools
 - Provides a choice for allowing child pools to extend reservations into Parent Pools

DRS Clusters

- Details:
 - All ESX Servers in a DRS Cluster must be compatible for VMotion.
 - When one ESX Server in the pool becomes too busy, DRS may automatically VMs to others in the pool.
 - When deploying new VMs, DRS may automatically determine the initial placement.
 - Automation Options:
 - Manual – all Vmotions and initial placements must be manually approved
 - Partially Automated – Initial placements are automatic and VMotions must be approved.
 - Fully Automated – both initial placements and Vmotions are automatic

More DRS Cluster Details

- Star rated suggestions – when recommending VMs for VMotion for load balancing, a star-based system is used, where five stars indicate the strongest recommendation.
- Aggressiveness Controls – Fully automated load balancing can be further configured by indicating the minimal number of stars a recommendation must receive before it will be automatically VMotioned. The default number is three stars.
- Load Balancing Recommendations and initial placements take into consideration any constraints from DRS Based resource pools as well as overall ESX Server load.

DRS Cluster Based Resource Pools

- Much like host-based pools, but cover several ESX Servers
- May have delegated administrators
- May configure constraints to be ignored if VMware HA automatically restarts VMs.

3rd Party Tools

- Virtugo Optimize

Dynamically adjusts CPU, RAM, and Disk Shares to increase transaction rates and reducing response times.

- Appsense Performance Manager

Designed to increase the capacity of Windows servers and achieve server consolidation by running more applications per server.

Recommendations

- Strive for Minimal Effort
 - Performance tuning should be avoided for new VI Implementations.
 - Configure Resource Pools Instead of VMs, Where Applicable
 - Configure as few VMs and Resource Pools as possible.

Recommendations

- Selecting the Right Tools.
 - Initially, rely on default functionality, and maintain a low load
 - In time, implement resource pools
 - As more VMs are added to each server, implement DRS Clustering in Manual or Partially Automated mode, where all VMotions must be approved.
 - After time, consider fully automating DRS load balancing.
 - If this still does not meet all the expectations for server utilization, then consider evaluating Virtugo Optimize or similar third party products designed to adjust shares and reservations automatically to improve performance.
 - Lastly, consider using third party tools, such as Appsense, to reduce the number of VMs.

Recommendations

- Understand the Issues
 - Priority – If certain VMs are described as needing higher priority, then increase shares. If certain VMs are lower priority, decrease shares.
 - Guaranteed Resources – If a certain amount of resources need to be guaranteed to a set of VMs, and then set Reservations.
 - Maximum Allowed Resources – If specific VMs are to be restricted to using a specific, maximum amount of resources, then set limits.

Recommendations

- Thresholds
 - ESX Server Thresholds – Define a normal utilization threshold at which no more new VMs will be deployed to an ESX Server. Define a higher, maximum utilization threshold at which VMs should be VMotioned from the ESX Server for load balancing.

Handle with alarms and operations procedures in VI 2

- In VI 3, consider implementing DRS Clusters and Resource Pools to enforce the above thresholds.

Set Reservation to Normal Threshold and Limit to Maximum Threshold

Recommendations

VMWare HA – a Recovery Tool, not a Load Balancing Tool

Details:

- Implement to ease VM recovery following system failures.
- Rely on DRS to ensure that VMs are still load balanced.
- Place priorities on VMs
- VMs may not restart due to resource constraints.
- Consider using Resource Pools to ensure that ample, available capacity exists across a cluster of ESX Servers, such that if any one ESX Server fails, the remaining servers can adequately run all the VMs.

Recommendations

Disk I/O Performance Tuning

- Add additional VMFS Volumes.
- Place virtual disk files in RAID Volumes that are appropriate for the VM's applications.
 - ESX Server and VMs do Not have a preferred RAID type.
 - Whenever different RAID types are used, it is typically due to preferences for the specific applications in the VMs.
- Add more I/O adapters, typically HBAs.
- For high disk I/O VMs, consider using Raw Disks and Raw Disk Mappings

Recommendations

Network Performance Tuning

- Use VM-based traffic shaping as a last resort.
- Use Traffic Shaping on VM Port groups, where appropriate, such as limiting port groups for test VMs on virtual switches, where both Test and Production VMs live.
- Preferably, use separate physical Nics or VLans for test and production.
- When bottlenecks appear on Nic Teams, simply add another Nic to the team.

Recommendations

Virtual Machine Configuration

- Always install VMware Tools, because it provides
 - Latest drivers for virtual devices, such as mouse, Nic, SCSI, and video
 - The vmmemctl driver, which allows it to participate in Ballooning, when needed.
- Ensure that the VMs are configured to use the preferred virtual Nic (in ESX Server 2.x, the recommendation was vmxnet, rather than vance)
- Ensure that the VMs are configured to use the preferred virtual SCSI device, which is LSI for Windows 2003 and BusLogic for Windows 2000.
- Within the guest O/S of the VM, disable any unused devices, such as serial ports and parallel ports, especially within P2V'd machines.

Recommendations

Other Recommendations

- Place a heterogeneous mixture of VMs on the same ESX Server, rather than similar applications with similar resource loads.
- Stress test new VM based applications to learn about its characteristics
- Set Shares to Normal, Low or High, rather than Custom
- For VMs that are expected to utilize as steady amount of CPU or RAM resources, setting the corresponding Reservation value for the VM will ensure that the VM does not have to compete for its resource under normal load.
- Look for VM's that are not user intensive to add to a fairly active ESX, but decrease its shares lower than all other VMs.
- In the BIOS of the hardware system running ESX Server, disable any unused devices, such as SCSI adapters, Nics, USB devices, serial ports, and parallel ports.

Recommendations

Addressing User Complaints

- If DRS is implemented, determine if it has automatically launched or is recommending any VMotions.
- Otherwise, to immediately alleviate the issue, VMotion the VM to a low utilized ESX Server.
- Determine bottleneck
 - CPU – look for ready time
 - RAM – look for ballooning and swapping
 - Disc – look for queuing (under `/proc/vmware/vm/#!/disk`)
 - Nic – look at utilization

Recommendations

Addressing User Complaints – continued

- Once the bottleneck is known, consider all the options that may help, such as:
 - guarantee more resources for the VM or its pool
 - limit resources for other VMs or pools
 - increase the priority of the VM or its pool
 - migrate the VM to another ESX Server or resource pool with more available resources.
 - Determine if too many resources are being allocated for the VM or its pool. Determine if too many resources are being allocated for other VMs and pools on the same ESX Server.

Presentation Download

Please remember to complete your
session evaluation form
and return it to the room monitors
as you exit the session

The presentation for this session can be downloaded at
<http://www.vmware.com/vmtn/vmworld/sessions/>

Enter the following to download (case-sensitive):

Username: cbv_rep
Password: cbvfor9v9r

VMWORLD 2006

