Virtualizing Zimbra™ on VMware® vSphere™ 4.0 Performance
Best Practices and Recommendations

Introduction

VMware vSphere’s capability to deliver computing and I/O resources far exceeds the resource requirements of most x86 applications, including Zimbra Collaboration Suite (ZCS) and the Zimbra Collaboration Suite Appliance (ZCA). This is what allows multiple application workloads to be consolidated onto the vSphere platform and benefit from reduced server cost, improved availability, and simplified operations.

However, there are some common misconfiguration or design issues that many experience when virtualizing applications, especially Enterprise workloads with higher resource demands than smaller departmental workloads.

We have compiled a short list of the essential best practices and recommendations to ensure a highly performant ZCS or ZCA deployment on the vSphere platform. We have also provided a list of highly recommended reference material to both build and deploy a vSphere platform with performance in mind, as well as troubleshooting steps to resolve performance related issues.

CPU Resources

- Confirm hardware assisted virtualization is enabled in the BIOS on your hardware platform.

- Confirm CPU/MMU virtualization is configured correctly for your hardware platform.

  To configure CPU/MMU virtualization:
  o 'myZimbraVM' -> Summary Tab -> Edit Settings -> Options -> CPU/MMU virtualization

- Non-Uniform Memory Access (NUMA) is a memory architecture used in multi-processor systems. A NUMA node is comprised of the processor and bank of memory local to that processor. In NUMA architecture, a processor can access its own local memory faster than non-local memory or memory local to another processor. A phenomenon known as NUMA “crosstalk” occurs when a processor accesses memory local to another processor causing a performance penalty.

  VMware ESX™ is NUMA aware and will schedule all of a virtual machine’s (VM) vCPUs on a ‘home’ NUMA node. However, if the VM container size (vCPU and RAM) is larger than the size of a NUMA node on the physical host, NUMA
crosstalk will occur. It is recommended, but not required, to configure your maximum VM container size to fit on a single NUMA node.

For example:
- ESX host with 4 sockets, 4 cores per socket, and 64GB of RAM.
- NUMA nodes are 4 cores with 16GB of RAM (1 socket and local memory).
- Recommended maximum VM container is 4 vCPU with 16GB of RAM.

- It is okay to over commit CPU resources, it is not okay to over utilize. Meaning you can allocate more virtual CPUs (vCPUs) than there are physical cores (pCores) in an ESX host as long as the aggregate workload does not exceed the physical processor capabilities. Over utilizing the physical host can cause excessive wait states for VMs and corresponding applications while the ESX scheduler is busy scheduling processor time for other VMs.

Zimbra is not CPU bound when disk and memory resources are sized correctly. It is perfectly fine to over commit vCPUs to pCores on ESX hosts where Zimbra workloads will be running. However, in any over committed deployment it is recommended to monitor host CPU utilization, VM Ready Time, and utilize the Dynamic Resource Scheduler (DRS) to load balance VMs across hosts in a vSphere Cluster.

VM Ready Time, host CPU utilization, and other important resource statistics can be monitored using ESXtop or from the Performance tab in the vSphere Client. You can also configure Alarms and Triggers to email administrators and perform other automated actions when performance counters reach critical thresholds that would affect the end user experience.

See the Performance Troubleshooting for VMware vSphere 4 guide for detailed information on performance troubleshooting.

- Reduce the number of vCPUs allocated to your Zimbra VM to the fewest number required to sustain your workload. Over allocating vCPUs causes excessive and unnecessary CPU overhead and idle time on the physical host. When memory and disk resources are sized appropriately, Zimbra is not a CPU bound workload. If your Zimbra VM experiences less than 60% sustained utilization during peak workloads, we recommend reducing the allocated vCPUs to half the number of currently allocated vCPUs.

- If you see periods of high, sustained CPU utilization on your Zimbra VM, this may actually be caused by memory backpressure or a poorly performing disk subsystem. It is recommended to first increase the memory allocated to the VM (make sure you match the VM memory reservation to the total allocated memory for as a JAVA workload best practice). Then, monitor VM CPU utilization, VM disk I/O, and in-guest swapping (can cause excessive disk I/O); for signs of improvement and other issues before increasing the number of vCPUs allocated to your Zimbra Appliance or mailbox server VM.
Memory Resources

- It is recommended to size the VM memory not to exceed the amount of memory local to a single NUMA node.

For example:
  - ESX host with 4 sockets, 4 cores per socket, and 64 GB of RAM.
  - NUMA nodes are 4 cores with 16 GB of RAM (1 socket and local memory).
  - Recommended maximum VM container is 4 vCPU with 16GB of RAM.

- Set the memory reservation for your Zimbra Appliance or mailbox server VMs to the total amount of memory allocated to the VM.

For example:
  - If you allocated 8192MB of memory to the Zimbra Appliance or mailbox server VM, then the memory reservation should be set to 8192MB.

To configure memory reservations:
  - ‘myZimbraVM’ -> Summary Tab -> Edit Settings -> Resources - > Memory -> Reservation

Network Resources

- For ZCS, use the VMXNET3 paravirtualized network adapter if supported by your guest Operating System. **Note:** This does not apply to the Zimbra Appliance.

- Use separate physical NIC ports, NIC teams, and VLANs for VM network traffic, vMotion, and IP based storage traffic (i.e. iSCSI storage or NFS datastores). This will avoid contention between client/server I/O, storage I/O, and vMotion traffic.

Storage Resources

- Do not oversubscribe VMFS datastores. Disk I/O and latency is a physics issue and storage design has the same impact on Zimbra performance virtual as it does physical. Design your Zimbra storage with the appropriate number of spindles to satisfy I/O requirements for Zimbra DBs, indexes, redologs, blob stores, etc.

See the Performance Troubleshooting for VMware vSphere 4 guide for detailed information on performance troubleshooting. Remember that insufficient memory allocation can cause excessive memory swapping and disk I/O. See the memory resource section for information on tuning VM memory resources.

- Use the PVSCSI paravirtualized SCSI adapter if supported by your guest Operating System.

- For ZCS, use the PVSCSI paravirtualized SCSI adapter if supported by your guest Operating System. **Note:** This does not apply to the Zimbra Appliance.
There is no performance benefit to using RDM devices versus VMFS datastores. It is recommended to use VMFS datastores unless you have specific storage vendor requirements to support hardware snapshots or replications in a virtual environment.

Configure your Zimbra VMs, VMDK disk device as thick-eagerzeroed to zero out each block when the VMDK is created. By default, new thick VMDK disk devices are created lazyzeroed. This causes duplicate I/O the first time each block is written to the disk device by first zeroing the block, then writing your application data. This can cause significant performance overhead for disk I/O intensive applications.

For ZCS, to configure thick-eagerzero VMDK disk devices either:

- Check the box to ‘Support clustering features such as Fault Tolerance’ when creating the VM. This does not enable FT, but does eagerzero the disks.

Or

- From the ESX CLI:
  ```bash
  vmkfstools -k /vmfs/volumes/path/to/vmdk
  ```

For Zimbra Appliance, to configure thick-eagerzero VMDK disk devices from the ESX CLI: `vmkfstools -k /vmfs/volumes/path/to/vmdk`

For more information about the ESX CLI, see the vSphere Command-Line Interface Documentation at [http://www.vmware.com/support/developer/vcli/](http://www.vmware.com/support/developer/vcli/)

- If using Fiber Channel storage, configure the maximum queue depth on the FC HBA card.

- Do not oversubscribe network interfaces or switches when using IP based storage (i.e. iSCSI or NFS). Use EtherChannel with ESX NIC teams and IP storage targets or 10GE if storage I/O requirements exceed a single 1Gb network interface.

- Use dedicated physical NIC ports, teams, and VLANs for IP based storage traffic (i.e. iSCSI storage or NFS datastores). This will avoid contention between client/server I/O, storage I/O, and vMotion traffic.

- Use Jumbo frames to increase storage I/O throughput and performance when using IP based storage (i.e. iSCSI or NFS).
vSphere Cluster Recommendations

- VMware vMotion
  - Use dedicated physical NIC ports, teams, and VLANs for vMotion traffic to avoid contention between client/server I/O, storage I/O, and vMotion traffic.

- VMware HA
  - Confirm VMware HA is enabled for the vSphere Cluster to automatically recover your Zimbra VMs in the vSphere Cluster in case of unplanned hardware downtime.

- VMware DRS
  - Confirm DRS is enabled to load balance VMs across ESX hosts in a vSphere Cluster.

  - With DRS, you can configure affinity rules to keep virtual machines together or apart on the ESX hosts in a vSphere Cluster. We recommend using affinity rules to separate ZCS multi-server deployments performing the same function onto different ESX hosts in a vSphere Cluster. This will minimize the impact to users caused by a hardware failure affecting a single ESX host. VMware HA (if enabled) will automatically recover the ZCS multi-server deployment VMs from the failed ESX host onto another ESX host in the vSphere Cluster.

  - To create a DRS rule:
    - ‘myvSphereCluster’ -> Edit settings -> VMware DRS -> Rules -> Add

  - Create the following rules:
    - Name: Zimbra Mailbox Servers -> Type: Separate Virtual Machines -> Add: ‘myZimbraMailboxServers’
    - Name: Zimbra Proxy Servers -> Type: Separate Virtual Machines -> Add: ‘myZimbraProxyServers’
    - Name: Zimbra MTA Servers -> Type: Separate Virtual Machines -> Add: ‘myZimbraMTAServers’

Reference Materials

Performance Best Practices for VMware vSphere 4.0

VMware vSphere 4 Performance with Extreme I/O Workloads

Performance Troubleshooting for VMware vSphere 4
Understanding Memory Resource Management in VMware ESX Server

Comparison of Storage Protocol Performance in VMware vSphere 4

Best Practices for Running vSphere on NFS Storage

Configuration Maximums for VMware vSphere 4.0

What's New in VMware vSphere 4: Performance Enhancements