



#### Document Disclaimer

This document is based on experience of the author and is not the opinion of RT Risk LLC or any of its employees. Intellectual property rights for the creation and content of this paper belong to the author. Copyright (c) Mike Burke 2009 - Virtera.

## VMware Host Tuning

1. What is performance tuning?
2. How much memory per host?
3. How many CPUs per host?
4. What is the ratio of memory to CPU?
5. Are you CPU bound or memory bound?
6. Should it be virtualized?
7. What is best practice?

*A simple answer would be subjective.*

A good place to start is proper planning and design. Start with “Four Cores”. Tuning is not a replacement for proper planning. Don’t guess - gather the data and know your workloads. Spec hardware based on need. Optimizing more CPU cores is generally better. Consider workload needs. Consider scheduling constraints. Memory optimization generally means more host RAM = better performance. RAM is a low cost commodity these days. Configure a VM’s memory based on the actual need. Tune migrated systems accordingly. Be careful over-allocating RAM. No swapping!

Advanced memory parameters include Mem.ShareScanTime. The time interval for TPS scanning is 60 minutes. Mem.ShareScanGhz is the amount of memory pages to scan per 1Ghz. Idle cycles (4MB/sec per 1Ghz). Mem.CtlMaxPercent limits memory reclamation by ballooning (vmmemctl) (65%).

## Storage

Use shared storage. Fiber-channel SAN – iSCSI and NFS ESX supports dynamic failover. MRU only for active/passive but manual load balancing. Provision alternately/accordingly.

## Optimizing storage RAID types

Balance performance vs. cost. Know your workload. Use iSCSI HBA’s on production and offload CPU time to hardware jumbo frames. Storage recommendations should tier storage based on need.

Tier 1 – DMX/Fiber

Tier 2 – iSCSI/SATA

ISO/Templates – NFS, iSCSI or SATA

Tier storage based on environment. Split production and development. Split based on system, data, swap and scheduled swap directories. Storage recommendations for LUN sizing. Machines per LUN is based on the average size of the disk. Remember vSwap and headroom. One VMFS per LUN. More, smaller LUNs vs. less, larger LUNs. Limit hosts per LUN.

## Locking issues

Do not use extents. Use SMotion. Use RDM's for disks over a certain size. Draw a line in the sand. Unload VMFS2 drivers. Network using gigabit only. Negotiate auto/auto in most instances and know your settings. Four NICs minimum. Enable spanning-tree.

## PortFast Networking

Dedicate NICs to iSCSI/NFS. Software iSCSI load-balancing is only possible with EtherChannel Networking. Always use more than one NIC per vSwitch. Different physical switches use different transceivers. Check PCI addressing for VMNIC ordering.

Combine Service Console and VMkernel. LB at the port group level and add a third dedicated to failover. Ensure consistency in port usage. Be security conscious: Isolate VMotion. Place high-I/O VMs on the same host vSwitch - port group.

## HA Recommendations

Ensure the service console has redundant paths and pingable gateway with alternate source. Watch memory over-commitment and understand admission control. No strict admission control in 2-node clusters. All names and IP's are registered in DNS.

## HA Advanced Parameters

das.isolationaddress - Address to PING for failure detection das.defaultfailoverhost. Which host should be the target of the failover? das.failedetectiontime - How long to wait to initiate failure detection? das.failedetectioninterval - How long between heartbeats?

## DRS Recommendations

Remember CPU and memory only. Use affinity rules. Specify exempt VMs and place appropriate VMs together. Separate conflicting VMs. Go fully automated with HA but use conservative migration thresholds.

## A Word on DRS and HA X

HA relies on reservations to determine available capacity set at the VM level - das.vmMemoryMinMB. Catch all memory reservations to consider when failover occurs - das.vmCpuMinMHz. Catch-all CPU reservations to consider when failover occurs.

## Shares, Reservations and Limits

Reservations are hard commits. Shares are relative priority and limits create sandboxes. Mix workloads when you can. To get better overall utilization leverage reservations for critical systems. Particularly in HA scenarios. Limit 3rd party tools in the Service Console. For antivirus if you must, remove /vmfs from scan list and configure NTP.

## General VM Recommendations

Always load VMware Tools and disable peripherals not used such as COM, LPT, floppy, etc. Time sync VMware Tools or NTP. Disable sync driver on Terminal Server and database VMs (vmmemctl). Disable screen savers and backgrounds. Unmap unused CD-ROMs/ISOs. Key counters include disk and network usage (KBps). ESX Host Monitoring Virtual Center %RDY is a red flag. VM's are spending time in a change sample size queue waiting to be scheduled. Change collection intervals swapping stats and increase thread count. Out of RAM? If consistent CPU times > 80% be sure to change modes. Need another host? (cpu, d - disk, m - memory). Check disk queue lengths with read vs. write characteristics. What disk I/O loads are in competition on the SAN?

## Monitoring

Virtual machine monitoring establishes a baseline with average/peak CPU Utilization. Monitoring processor queue length helps understand what "normal" is. Memory resources being used are closely monitored before trouble starts. What are your SLA's? What is expected/acceptable performance? I/O reads vs. write I/O bytes

(Read and Write). What are disk queue lengths and network bytes received/sent? Generally speaking, CPU and memory will be constrained. Disk and network saturation are uncommon. Keep overall CPU utilization under 65%. Keep Memory utilization under 100% for production systems and 125% for development/test systems.