

Availability (FT & MSCS)

Maximums (FT advice): Disks per VM = 16 FT VMs per host = 4 Minimum hosts per cluster = 3

FW Port	Source	Destination	Prot (ESX port)	Description
8100, 8200	Hosts	ESX/ESXi	UDP (SC)	FT
8100, 8200 (out)	ESX/ESXi	Hosts	TCP/UDP(SC)	FT

FT: uses anti-affinity rules. **Requires** - HA & host monitoring, host certificate checking (on by default), dedicated logging NIC, compatible CPU, Hardware Virtualization (HV), thick disks on shared storage, supported guest OS. **Not supported** - snapshots, Storage vMotion, hotplugging, MSCS, VCB, SMP, physical RDMs, Paravirtualized VMs, NPIV, VMDirectPath, EPT/RVI. DRS only if cluster is EVC.

**MSCS** - 2003 SP2 & 2008 (Failover Clustering) \* 32 & 64bit \* only 2 nodes clusters

**Not supported** - DRS on VMs, vMotion, FT, NPIV, Round Robin NMP, iSCSI/NFS based disks

VMDK	Virtual RDM	Physical RDM
Yes (zeroed)	Yes	No (not supported)
Cluster across boxes (CAB)	No	Only 2003
Physical & VM (n+1)	No	No
Snapshots	Yes	Yes
SCSI target software	No	No

• Configure all RDMs before configuring VM's network settings, or initialising LUNs within windows.

• Add all RDMs to a 2nd SCSI controller i.e. SCSI(1:x). Set sharing to Physical or Virtual as required.

**SCSI bus sharing** • CIB = Virtual • CAB or N+1 = Physical

**Links:** <http://kb.vmware.com/kb/1010601> - Understanding FT

<http://kb.vmware.com/kb/1008027> - CPU & guest OS that support FT

Networking

Maximums (per host): 1GB vNICs = 2 - 32 dependent on HW 10GB vNICs = 4

PCI VMDirectPath devices=8 switches (vSS/vDS/VEM) = 248/16/1 vSS/vDS ports=4096

Active ports = 1016 Service Console ports = 16

Maximums (per vCenter): vDS switches=32 vDS port groups=5000(1016 ephemeral) vDS ports=20000

Maximums (per switch): Hosts (per vDS) = 350 vSS port groups = 512vSS switch ports = 4,088

**Terminology:** vNICs - logical name for physical server NICs vNICs - virtual NICs assigned to VMs vSS - virtual Standard Switch vDS - virtual Distributed Switch vPort - port group on a vDS dvUplink - uplink vNICs on a vDS Network vMotion - tracking of VM's network state on a vDS

**Common networking commands** (-h switch for options or man page for detailed description):

List vNICs:

```
$ sudo /usr/sbin/esxcfg-nics -l
```

List vSwitches & Port Groups:

```
$ sudo /usr/sbin/esxcfg-vswi -l
```

List Service Console ports:

```
$ sudo /usr/sbin/esxcfg-vswif -l
```

List VMkernel ports:

```
$ sudo /usr/sbin/esxcfg-vmknic -l
```

List VMkernel Default Gateway:

```
$ sudo /usr/sbin/esxcfg-route
```

**Common networking configuration files:** Name resolution order: [etc/nsswitch.conf](http://etc/nsswitch.conf)  
Local host file: [etc/hosts](http://etc/hosts) DNS servers: [etc/resolv.conf](http://etc/resolv.conf) DG: [etc/sysconfig/network](http://etc/sysconfig/network)

**Ethernet tagging:** • **EST** (External Switch Tagging) - Default. No trunking required. 1-1 relationship from vNICs to physical switch ports. Each vNIC can only see 1 subnet. VLAN ID of 0 or blank.

• **VST** (Virtual Switch Tagging) - Commonly used. vNICs connected to a vSwitch can span several VLANs. Each Port Group has a VLAN ID of 1-4094. Set the VLAN ID to blank to use Native VLAN.

• **VGT** (Virtual Guest Tagging) - Rarely used. Install 802.1Q trunking driver software in the VMs, the vSwitch preserves the tags given by the VMs. VLAN ID of 4095 on vSS, VLAN policy on vDS. Avoid using a **VLAN ID of 1**, as this is the native Cisco VLAN ID.

**vSS & vDS options** (options can also be overridden on individual Port Groups):

**General** • Number of ports - by default 56 for vSS, 64 for vDS, 128 when created on Service Console. (not a Port Group option) • Network label & VLAN ID - only on Port Groups not vSwitches.

**Security** • **Promiscuous mode** (default Reject) - only listens to traffic destined for its MAC address.

• **MAC Address Changes** (default Accept) - accepts inbound frames when VM changes MAC address.

• **Forged Transmits** (default Accept) - won't drop outbound frames if source MAC address is different.

**Traffic Shaping** • Status (default Disabled) Average Bandwidth (default 102400 Kbps) Peak Bandwidth (default 102400 Kbps) Burst size (default 102400 KB) - shapes out on vSS, in/out on vDS.

**NIC Teaming** • Load Balancing (spreads outbound traffic from vNICs across vNICs). **Originating port ID** (default) uses vNIC based on where traffic entered. **ip hash** based on source & destination IP address of each packet (if physical switch ports are etherchannel). **Source MAC hash** based on source MAC address. **Route based on physical NIC** load only on vDS, dynamically redistributes load across all vNICs in team. Use **explicit failover order**. Incoming traffic is load balanced by physical switch.

• Network Failover Detection **List status** only (default) detects cable pulls & switch power failures, not misconfigurations. **Beacon Probing** don't use with IP-hash load balancing.

• Notify Switches - No or Yes (default) updates lookup tables. Disable for MS NLB in unicast mode.

• Failback - No or Yes (default) vNIC will return after recovering from a failure.

• Failover order **Active** - Standby - **Unused**. Don't use standby uplinks with IP-hash load balancing.

**VLAN** (vDS only) • **VLAN** - set ID. **Trunk range** - restrict IDs on trunked links. **PVLAN** - see below.

**Miscellaneous** (vDS only) • Port blocking - selected or unselected (default) block all ports.

**dvPort options** • Port Binding **Static** when initially connected **Dynamic** when connected/powerd-on Ephemeral no binding • Traffic shaping **Ingress** into vSwitch **Egress** out of vSwitch • Allow live port moving • Config reset at disconnect • Host can assign port if vCenter is down • Name format

**NIOC** (Network IO Control): prioritise egress vDS traffic via shares/limits (FT,iSCSI,vMotion,Mgt,NFS,VM)

**PVLAN** (Private VLAN): extension to VLAN standard, adds further segmentation. Not encapsulated.

**Primary PVLAN** - Original VLAN divided into smaller groups. **Secondary PVLAN** - exists only within primary, has specific VLAN ID. **Secondary types:** **Promiscuous** - connect with VMs in primary. **Community**-connect to themselves & VMs on promiscuous **Isolated**-connect with VMs on promiscuous

**TSO** (TCP Segmentation Offload): enabled by default on VMkernel ports, allows very large frames (up to 64KB), even with smaller MTU. To enable on VMs, they need enhanced vmxnet vNIC.

**Jumbo frames** up to 9kB. Must be enabled for each vSwitch. vNIC must be vmxnet2/3 or e1000.

**NetQueue** enabled by default, allows certain vNICs to spread processing across multiple CPUs.

**Configure networking** (for vSS): (1) add vSwitch **esxcfg-vswitch -a (2)** add port group to vSwitch **esxcfg-vswitch -A (3)** set port group's VLAN ID **esxcfg-vswitch -p -v (4)** add vNIC to vSwitch **esxcfg-vswitch -L** • VM connections: set VM's NIC to use port group.

• Service Console: create interface & add it to the port group **esxcfg-vswif -a -p -i -n**, set the DG in [etc/sysconfig/network](http://etc/sysconfig/network), then restart networking **service network restart**.

• VMkernel ports: add port **esxcfg-vmknic -a -i -n** & set VMkernel DG **esxcfg-route**. vMotion enabled in vCenter if required.

**Links:** <http://kb.vmware.com/kb/1000258> - Configure networking from Service Console

<http://vmware.com/files/pdf/vsphere-vnetwork-ds-migration-configuration-wp.pdf> - vDS whitepaper

Resources

Maximums (per DRS cluster): Hosts = 32 VMs (powered on) = 3000 (limit of 320 per host)

Maximums (per Resource Pool): Children = 1024 Tree depth = 8

Maximums (other): Hosts per datacenter = 400 RPs per host = 4096 CPUs per cluster = 512

**Datacenters** mark organisational & vMotion boundaries. **Clusters** gather host CPU & memory resources. **Resource Pools** apply policies to clusters. A DRS cluster is also implicitly a resource pool.

**Resource pools** • **Shares** - low, medium & high (1,2,4) • **Reservations** - MHz(CPU)/MB(RAM)

• **Limits** - MHz/MB • **Expandable reservation** - yes (can draw from parent's pool) - no (can only draw from own pool). List the resource group settings: **\$ sudo /usr/sbin/esxcfg-resgrp -l**

Shares only apply during contention. Shares are relative to siblings. Reservations guarantee a minimum, are only checked when a VM is powered on. Limits are an upper bound, never exceeded; manage user expectations but can waste idle resources. Expandable reservations do not automatically hunt upwards, define if reservations are considered by admission control. Child pools actively reserve resources from parent even if VMs are powered off. Hierarchical resource pools require DRS enabled.

**DRS: priority levels** 1-5 (1 the highest). **DRS cluster settings:** • **Manual** / Partial (Initial VM placement) • **Fully Automated** (Initial VM placement & Dynamic balancing). **Current Host Load**

**Standard Deviation:** load imbalance. Lower than Target value unless recommendations are unapplied.

**"Grafted from"** pools created when adding a host to a DRS cluster & keeping the host's resource pool hierarchy. **Maintenance mode** only clears VMs off host if DRS cluster is fully automated.

**Affinity Rules:** VM-VM keep VMs together/apart. VM-Host keep VMs on/off specific hosts. **Should rule best effort. Must rule mandatory** (for licensing). **Rule conflicts** often wins, newer rule disabled. Anti-affinity wins over affinity. Disabled rules ignored.

**Current host load standard deviation:** DRS load imbalance. Current < Target unless advice unapplied

**DPM:** uses IPMI, iLO or WOL (in that order). DRS & DPM thresholds are independent. Verify host's **DPM Last Time Exited Standby**. DPM level - Off, Manual (make recommendations) & Automatic.

**Links:** <http://kb.vmware.com/kb/1005764> - Enhanced vMotion (EVC) FAQ

<http://www.yellow-bricks.com/hrs-deepdive/> - DRS Deep Dive

<http://kb.vmware.com/kb/1003212> - EVC CPU compatibility

Storage

Maximums (per host): Volumes = 256 Paths = 1024 NAS datastores = 64

FC - HBAS = 8 (HBA ports = 16) targets per HBA = 256

iSCSI HW - HBAs = 4 targets per HBA = 64 (62 - QLogic Static)

iSCSI SW - NICs = 8 targets = 256

**Maximums (per volume):** VMs = 256 Hosts = 64 (DRS limit, 2048 for linked clones)

VMFS = 64TB (less 16KB) NFS =16TB File size (1/2/4/8MB blocks) = 256GB/512GB/1TB/2TB

RDMs = 2TB (less 512B) Extents = 32 Extent size = 2TB (less 512B)

**FW Port** **Source** **Destination** **Prot (ESX port)** **Description**

2049 NFS server ESX/ESXi TCP (VMK) NFS Client

2049 ESX/ESXi NFS server TCP (VMK) NFS Client

3260 ESX/ESXi iSCSI server UDP (SC+VMK) iSCSI Client

**Common storage commands** (-h switch for options, or man page for detailed description):

List all storage devices:

```
$ sudo /usr/sbin/esxcfg-scsidevs -c
```

List LUNs, paths & multipathing plugins:

```
$ sudo /usr/sbin/esxcfg-mpath -L
```

List all VMware SATPs:

```
$ sudo /usr/sbin/esxcli mmp satp list
```

List claim rules:

```
$ sudo /usr/sbin/esxcli corestorage claimrule list
```

List datastores, dev names to VMFS:

```
$ sudo /usr/sbin/esxcfg-scsidevs -m
```

List snapshot volumes:

```
$ sudo /usr/sbin/esxcfg-volume -l
```

Test VMkernel connectivity: **\$ /usr/sbin/vmkping**

Manage HW iSCSI (QLogic) settings:

```
$ sudo /usr/sbin/esxcfg-hwiscsi -l
```

Manage SW iSCSI settings:

```
$ sudo /usr/sbin/esxcfg-swiscsi -q
```

List iSCSI LUNs:

```
$ sudo /usr/sbin/vmkiscsi-tool -L -l 1 adapter
```

Rescan iSCSI LUNs:

```
$ sudo /usr/sbin/esxcfg-rescan adapter
```

List the NFS exports from the VMkernel:

```
$ sudo /usr/sbin/esxcfg-nas -l
```

**Storage capabilities** **FC** **iSCSI** **NAS**

vMotion, DRS, HA, FT, VCB, SRM & Thin VMDKs Yes Yes Yes

VMFS volumes, RDMs & VMware's NMP Yes Yes No

Boot ESX host Yes Yes (HW initiator) No

VM MSCS clustering Yes No No

**Zoning:** at the switch. **LUN masking:** done at the SP or server.

**Active-active:** access to the LUNs simultaneously through all ports, without performance degradation.

**Active-passive:** one port actively providing access, other as backup. Path thrashing can occur.

**NPIV** (N-Port ID Virtualization): FC HBA port assigned dedicated virtual port (WWPN) to VM (RDM)

**LUN addressing** **FC:** **Runtime Name** vmhba#:*CH:TL:LP* - adapter:channel:target:LUN

**iSCSI:** **IQN** *iqn.year-month-day.reverse.domain\_name.string* or **EUI** *eui.string*

**iSCSI discovery methods:** **Static** - can manually add/remove items, only with hardware initiators. **Dynamic** - uses "SendTargets", target responds with list. Removed targets return after HBA rescan/reset

**CHAP:** HW iSCSI 1-way CHAP, initiator level. SW iSCSI 1-way & mutual CHAP, initiator or target VMkernel Port is required to use iSCSI or NFS storage. (S.C. port not required for iSCSI anymore)

**MPP** (MultiPathing Plugins): claim rules in [etc/vmware/esx.conf specify MPP to use for each path.](http://etc/vmware/esx.conf)

**Claim rules:** indicate which MPP, native or 3rd party, manages a given physical path.

**NMP** (Native MPP) • SATPs (Storage Array Type Plugins) - handles failovers. • PSPs (Path Selection Plugins) - handles load-balancing. **NMP policies:** **Fixed** - default for active/active, uses preferred path when available. **MRU** (Most Recently Used) - default for active/passive (& iSCSI), first working path found at boot. **RR** (Round Robin) - safe for all arrays, rotates through paths (not MSCS LUNs).

**Disk.MaskLUN:** reduce number of LUNs scanned. **Disk.MaskLUN:** convert to claim rule instead.

**VMFS volumes:** Large-less LUNs to create, less to manage, flexible resizing & snapshots. Small=less contention (locking), less wasted space, different RAID's, more flexible multipathing & disk shares.

**SIOC** (Storage IO Control): shares VM's disk IO across datastore's hosts. Monitors latency, adjusts VM's host queue access. Can also enforce VM IOPS limits. Enable on datastore, set shares/limit on VM ALUA (Asymmetric Logical Unit access): finds/manages multiple paths for failover & load balancing.

**Links:** <http://kb.vmware.com/kb/1009553> - Lost connectivity to storage

<http://media.netapp.com/documents/tr-3749.pdf> - Storage best practices whitepaper (NetApp)

<http://media.netapp.com/documents/tr-3747.pdf> - File System alignment whitepaper (NetApp)

Storage

Maximums (per host): Volumes = 256 Paths = 1024 NAS datastores = 64

FC - HBAS = 8 (HBA ports = 16) targets per HBA = 256

iSCSI HW - HBAs = 4 targets per HBA = 64 (62 - QLogic Static)

iSCSI SW - NICs = 8 targets = 256

**Maximums (per volume):** VMs = 256 Hosts = 64 (DRS limit, 2048 for linked clones)

VMFS = 64TB (less 16KB) NFS =16TB File size (1/2/4/8MB blocks) = 256GB/512GB/1TB/2TB

RDMs = 2TB (less 512B) Extents = 32 Extent size = 2TB (less 512B)

**FW Port** **Source** **Destination** **Prot (ESX port)** **Description**

2049 NFS server ESX/ESXi TCP (VMK) NFS Client

2049 ESX/ESXi NFS server TCP (VMK) NFS Client

3260 ESX/ESXi iSCSI server UDP (SC+VMK) iSCSI Client

**Common storage commands** (-h switch for options, or man page for detailed description):

List all storage devices:

```
$ sudo /usr/sbin/esxcfg-scsidevs -c
```

List LUNs, paths & multipathing plugins:

```
$ sudo /usr/sbin/esxcfg-mpath -L
```

List all VMware SATPs:

```
$ sudo /usr/sbin/esxcli mmp satp list
```

List claim rules:

```
$ sudo /usr/sbin/esxcli corestorage claimrule list
```

List datastores, dev names to VMFS:

```
$ sudo /usr/sbin/esxcfg-scsidevs -m
```

List snapshot volumes:

```
$ sudo /usr/sbin/esxcfg-volume -l
```

Test VMkernel connectivity: **\$ /usr/sbin/vmkping**

Manage HW iSCSI (QLogic) settings:

```
$ sudo /usr/sbin/esxcfg-hwiscsi -l
```

Manage SW iSCSI settings:

```
$ sudo /usr/sbin/esxcfg-swiscsi -q
```

List iSCSI LUNs:

```
$ sudo /usr/sbin/vmkiscsi-tool -L -l 1 adapter
```

Rescan iSCSI LUNs:

```
$ sudo /usr/sbin/esxcfg-rescan adapter
```

List the NFS exports from the VMkernel:

```
$ sudo /usr/sbin/esxcfg-nas -l
```

**Storage capabilities** **FC** **iSCSI** **NAS**

vMotion, DRS, HA, FT, VCB, SRM & Thin VMDKs Yes Yes Yes

VMFS volumes, RDMs & VMware's NMP Yes Yes No

Boot ESX host Yes Yes (HW initiator) No

VM MSCS clustering Yes No No

**Zoning:** at the switch. **LUN masking:** done at the SP or server.

**Active-active:** access to the LUNs simultaneously through all ports, without performance degradation.

**Active-passive:** one port actively providing access, other as backup. Path thrashing can occur.

**NPIV** (N-Port ID Virtualization): FC HBA port assigned dedicated virtual port (WWPN) to VM (RDM)

**LUN addressing** **FC:** **Runtime Name** vmhba#:*CH:TL:LP* - adapter:channel:target:LUN

**iSCSI:** **IQN** *iqn.year-month-day.reverse.domain\_name.string* or **EUI** *eui.string*

**iSCSI discovery methods:** **Static** - can manually add/remove items, only with hardware initiators. **Dynamic** - uses "SendTargets", target responds with list. Removed targets return after HBA rescan/reset

**CHAP:** HW iSCSI 1-way CHAP, initiator level. SW iSCSI 1-way & mutual CHAP, initiator or target VMkernel Port is required to use iSCSI or NFS storage. (S.C. port not required for iSCSI anymore)

**MPP** (MultiPathing Plugins): claim rules in [etc/vmware/esx.conf](http://etc/vmware/esx.conf) specify MPP to use for each path.

**Claim rules:** indicate which MPP, native or 3rd party, manages a given physical path.

**NMP** (Native MPP) • SATPs (Storage Array Type Plugins) - handles failovers. • PSPs (Path Selection Plugins) - handles load-balancing. **NMP policies:** **Fixed** - default for active/active, uses preferred path when available. **MRU** (Most Recently Used) - default for active/passive (& iSCSI), first working path found at boot. **RR** (Round Robin) - safe for all arrays, rotates through paths (not MSCS LUNs).

**Disk.MaskLUN:** reduce number of LUNs scanned. **Disk.MaskLUN:** convert to claim rule instead.

**VMFS volumes:** Large-less LUNs to create, less to manage, flexible resizing & snapshots. Small=less contention (locking), less wasted space, different RAID's, more flexible multipathing & disk shares.

**SIOC** (Storage IO Control): shares VM's disk IO across datastore's hosts. Monitors latency, adjusts VM's host queue access. Can also enforce VM IOPS limits. Enable on datastore, set shares/limit on VM ALUA (Asymmetric Logical Unit access): finds/manages multiple paths for failover & load balancing.

**Links:** <http://kb.vmware.com/kb/1009553> - Lost connectivity to storage

<http://media.netapp.com/documents/tr-3749.pdf> - Storage best practices whitepaper (NetApp)

<http://media.netapp.com/documents/tr-3747.pdf> - File System alignment whitepaper (NetApp)

Storage

Maximums (per host): Volumes = 256 Paths = 1024 NAS datastores = 64

FC - HBAS = 8 (HBA ports = 16) targets per HBA = 256

iSCSI HW - HBAs = 4 targets per HBA = 64 (62 - QLogic Static)

iSCSI SW - NICs = 8 targets = 256

**Maximums (per volume):** VMs = 256 Hosts = 64 (DRS limit, 2048 for linked clones)

VMFS = 64TB (less 16KB) NFS =16TB File size (1/2/4/8MB blocks) = 256GB/512GB/1TB/2TB

RDMs = 2TB (less 512B) Extents = 32 Extent size = 2TB (less 512B)

**FW Port** **Source** **Destination** **Prot (ESX port)** **Description**

2049 NFS server ESX/ESXi TCP (VMK) NFS Client

2049 ESX/ESXi NFS server TCP (VMK) NFS Client

3260 ESX/ESXi iSCSI server UDP (SC+VMK) iSCSI Client

**Common storage commands** (-h switch for options, or man page for detailed description):

List all storage devices:

```
$ sudo /usr/sbin/esxcfg-scsidevs -c
```

List LUNs, paths & multipathing plugins:

```
$ sudo /usr/sbin/esxcfg-mpath -L
```

List all VMware SATPs:

```
$ sudo /usr/sbin/esxcli mmp satp list
```

List claim rules:

```
$ sudo /usr/sbin/esxcli corestorage claimrule list
```

List datastores, dev names to VMFS:

```
$ sudo /usr/sbin/esxcfg-scsidevs -m
```

List snapshot volumes:

```
$ sudo /usr/sbin/esxcfg-volume -l
```

Test VMkernel connectivity: **\$ /usr/sbin/vmkping**

Manage HW iSCSI (QLogic) settings:

```
$ sudo /usr/sbin/esxcfg-hwiscsi -l
```

Manage SW iSCSI settings:

```
$ sudo /usr/sbin/esxcfg-swiscsi -q
```

List iSCSI LUNs:

```
$ sudo /usr/sbin/vmkiscsi-tool -L -l 1 adapter
```

Rescan iSCSI LUNs:

```
$ sudo /usr/sbin/esxcfg-rescan adapter
```

List the NFS exports from the VMkernel:

```
$ sudo /usr/sbin/esxcfg-nas -l
```

**Storage capabilities** **FC** **iSCSI** **NAS**

vMotion, DRS, HA, FT, VCB, SRM & Thin VMDKs Yes Yes Yes

VMFS volumes, RDMs & VMware's NMP Yes Yes No

Boot ESX host Yes Yes (HW initiator) No

VM MSCS clustering Yes No No

**Zoning:** at the switch. **LUN masking:** done at the SP or server.

**Active-active:** access to the LUNs simultaneously through all ports, without performance degradation.

**Active-passive:** one port actively providing access, other as backup. Path thrashing can occur.

**NPIV** (N-Port ID Virtualization): FC HBA port assigned dedicated virtual port (WWPN) to VM (RDM)

**LUN addressing** **FC:** **Runtime Name** vmhba#:*CH:TL:LP* - adapter:channel:target:LUN

**iSCSI:** **IQN** *iqn.year-month-day.reverse.domain\_name.string* or **EUI** *eui.string*

**iSCSI discovery methods:** **Static** - can manually add/remove items, only with hardware initiators. **Dynamic** - uses "SendTargets", target responds with list. Removed targets return after HBA rescan/reset

**CHAP:** HW iSCSI 1-way CHAP, initiator level. SW iSCSI 1-way & mutual CHAP, initiator or target VMkernel Port is required to use iSCSI or NFS storage. (S.C. port not required for iSCSI anymore)

**MPP** (MultiPathing Plugins): claim rules in [etc/vmware/esx.conf](http://etc/vmware/esx.conf) specify MPP to use for each path.

**Claim rules:** indicate which MPP, native or 3rd party, manages a given physical path.

**NMP** (Native MPP) • SATPs (Storage Array Type Plugins) - handles failovers. • PSPs (Path Selection Plugins) - handles load-balancing. **NMP policies:** **Fixed** - default for active/active, uses preferred path when available. **MRU** (Most Recently Used) - default for active/passive (& iSCSI), first working path found at boot. **RR** (Round Robin) - safe for all arrays, rotates through paths (not MSCS LUNs).

**Disk.MaskLUN:** reduce number of LUNs scanned. **Disk.MaskLUN:** convert to claim rule instead.

**VMFS volumes:** Large-less LUNs to create, less to manage, flexible resizing & snapshots. Small=less contention (locking), less wasted space, different RAID's, more flexible multipathing & disk shares.

**SIOC** (Storage IO Control): shares VM's disk IO across datastore's hosts. Monitors latency, adjusts VM's host queue access. Can also enforce VM IOPS limits. Enable on datastore, set shares/limit on VM ALUA (Asymmetric Logical Unit access): finds/manages multiple paths for failover & load balancing.

**Links:** <http://kb.vmware.com/kb/1009553> - Lost connectivity to storage

<http://media.netapp.com/documents/tr-3749.pdf> - Storage best practices whitepaper (NetApp)

<http://media.netapp.com/documents/tr-3747.pdf> - File System alignment whitepaper (NetApp)

This document is licensed under a Creative Commons License. Refer to <http://www.creativecommons.org> for full details. The artwork is from the Tango Project <http://tango.freedesktop.org> under their Creative Commons license.

ESX Hosts

Maximums (per host): vCPUs = 512 vCPUs per physical core = 25 VMs Logical procs (incl HT) = 128

RAM = 1TB Service Console RAM = 800MB (min=272MB) VMs = 320

**FW Port** **Source** **Destination** **Prot (ESX port)** **Description**

22 SSH client ESX TCP (SC) SSH server

53 (out) ESX/ESXi DNS server(s) UDP (SC) DNS requests

80 Clients ESX/ESXi TCP (SC) HTTP access

123 (out) ESX/ESXi NTP source UDP (SC) NTP (time) client

427 Hosts, Client ESX/ESXi UDP (SC) CIM SLP client/server

427 (out) ESX/ESXi Hosts UDP (SC) CIM SLP client/server

443 Hosts, Clients, VC ESX/ESXi TCP (SC) HTTPS access

902 Hosts, Clients, VC ESX/ESXi TCP (SC) Auth, migrate, provision

902 (out) ESX/ESXi Hosts, VC UDP (SC) Auth, migrate, provision

903 Clients ESX/ESXi TCP (SC) VM Console

5900-5964 ? ESX/ESXi TCP (SC) RFB for mgt tools (VNC)

5900-5964 (out) Hosts ? TCP (SC) RFB for mgt tools (VNC)

5989 Clients ESX/ESXi TCP (SC) CIM server over HTTPS

5989 (out) ESX/ESXi Hosts TCP (SC) CIM server over HTTPS

8000 Hosts ESX/ESXi TCP (VMK) vMotion requests

8000 (out) ESX/ESXi Hosts TCP (VMK) vMotion requests

Possible extras:21(FTP),22(out)(SSH),53(DNS),88/389/464(AD),161/162(SNMP),445(SMB),5988(CIM)

**Logs:** Service Console Availability & VMkernel Messages, Alerts, Availability: [var/log/vmkernel](http://var/log/vmkernel)

ESX service log: [var/log/vmware/hostlog](http://var/log/vmware/hostlog) Syslog: [var/log/messages](http://var/log/messages)

VMkernel warnings: [var/log/vmkwarning](http://var/log/vmkwarning) VMkernel events: [var/log/vmksummary](http://var/log/vmksummary)

VC agent: [var/log/vmware/vpx/vpxa.log](http://var/log/vmware/vpx/vpxa.log) Patching: [var/log/vmware/esxupdate.log](http://var/log/vmware/esxupdate.log)

**Common ESX host commands** (-h switch for options or man page for detailed description):

List status of all services:

```
$ sudo /sbin/service --status-all
```

List the service runlevels:

```
$ chkconfig --list
```

Restart a service: **\$ sudo /sbin/service service\_name restart** (start, stop, status available)

Common services: • **mgmt-vmware** (hostd) • **vmware-vpxa** (vCenter agent) • **vmware-vmauthd** (authentication) • **network** (vsfif changes) • **vmware-webAccess** (Web Access)

Show build number:

```
$ vmware -v
```

Check the filesystem usage:

```
$ sudo vdf -hP
```

List diagnostic partitions:

```
$ sudo /usr/sbin/esxcfg-dumpart -l
```

Show description of VMkernel error:

```
$ vmkerrcode error_code number
```

Export detailed config file:

```
$ sudo esxcfg-info > /tmp/esxcfg-info.txt
```

Gather debugging report:

```
$ sudo /usr/bin/vm-support -w /tmp
```

Configure authentication settings:

```
$ sudo /usr/sbin/esxcfg-auth
```

List drivers loaded at startup:

```
$ sudo /usr/sbin/esxcfg-module -l
```

Set advanced options: **\$ sudo /usr/sbin/esxcfg-advcfg option -s value** (-g to get)

Update bootstrap settings:

```
$ sudo /usr/sbin/esxcfg-boot
```

 (treat with caution)

Initialization routines (resets things):

```
$ sudo /usr/sbin/esxcfg-init
```

 (treat with caution)

**Internal firewall commands** (iptables on Service Console):

Show all firewall settings:

```
$ sudo /usr/sbin/esxcfg-firewall -q
```

List the firewall named services:

```
$ sudo /usr/sbin/esxcfg-firewall -l
```

Enable a service: **\$ sudo /usr/sbin/esxcfg-firewall -s service\_name** (-d to disable)

To open a port: **\$ sudo /usr/sbin/esxcfg-firewall -o port, protocol, direction, name**

**Security Levels:** High - in/out blocked. Medium - in blocked, out open. Low - in/out open.

By default all traffic blocked in & out, except 22, 123, 427, 443, 902, 5989, 5988, pings, DHCP & DNS

**Master config file:** [etc/vmware/esx.conf](http://etc/vmware/esx.conf) - Certificate files: hostd regenerates new files if not present.

Certificate public key [etc/vmware/ssl/rui.crt](http://etc/vmware/ssl/rui.crt) Certificate private key [etc/vmware/ssl/rui.key](http://etc/vmware/ssl/rui.key)

Set certificate location [etc/vmware/hostd/proxy.xml](http://etc/vmware/hostd/proxy.xml) SSL settings [etc/vmware/hostd/config.xml](http://etc/vmware/hostd/config.xml)

**PAM** (Pluggable Authentication Modules) configuration: [etc/pam.d/vmware-authd](http://etc/pam.d/vmware-authd)

Default authentication method is **etc/passwd**. vpxuser is for vCenter Server permissions.

**Passwords:** ESX uses **pam\_cracklib** so plug-in by default. No restrictions on root password. Defaults for non-root users: password length = 3, minimum password length = 9, shorter passwords if Characters Classes mixed (upper, lower, digits & other) M - CC = E. **pam\_passwdqc**.so provides more options.

**User Password Aging:** enabled by default, set to never expire (max days) & change anytime (min days = 0, warning = 7) • Change host settings: **esxcfg-auth** • Change user settings: **chage**

**NUMA** (Non-Uniform Memory Access): controls VM memory distribution across host memory. Only use NUMA if CPU affinity is set. **HT**: can help better utilize idle resources.

**Reclaims VM memory by:** • Balloon driver (if vmemmctl) guest quest to use native algorithms (guest swap) • Memory compression - vswp file (if vmemmctl unresponsive) • Sharing memory across VMs

**VMware MIBs:** uses embedded SNMP agent (disabled by default). Enable: **viCfgr -snmp enabled** to configure • **ESX** - edit [etc/vslog.conf](http://etc/vslog.conf) • **ESXi** - use Client or **viCfgr -syslog**

**Links:** <http://kb.vmware.com/kb/653> - Collecting diagnostic information for ESX Servers

<http://kb.vmware.com/kb/1005184> - Decoding Machine Check Exception output after purple screen

<http://kb.vmware.com/kb/1012514> - Determining detailed build number information for ESX hosts

[http://www.vmware.com/pdf/Perf\\_Best\\_Practices\\_vSphere4.0.pdf](http://www.vmware.com/pdf/Perf_Best_Practices_vSphere4.0.pdf) - Performance best practices

<http://communities.vmware.com/docs/DOC-9279> - Interpreting esxtop Statistics

ESXi hosts

HW requirements: 64bit x86 CPUs, 2GB RAM, SATA, SAS or SCSI disks. No ESXi WebAccess.

ESXi Installable starts in eval mode (60 days). If no DHCP at install, link local IP used 169.254.x.x/16.

**ESXi Installable Partitions:** 4GB VFAT scratch for system swap (not required, but stores vm-support), 110MB diagnostic for core dumps, VMFS3 on free space.

**Not supported:** • ESXi Installable & Embedded on same host • Booting multiple servers from 1 image

**DCUI** (Direct Console UI): • Configuring host defaults • Set up administrative access • Troubleshooting

**Restarting Mgt agents** effects [etc/init.d](http://etc/init.d) processes: hostd (mgmt-vmware), ntpd (time), sfcbd (CIM broker), slpd (discover/advertise services), wsman (share mgt info via SOAP), vobd (error reporting) & AAM (HA agent) if installed. To isolate ESXi host from DRS/HA cluster disable mgt network.

**Management Network Test:** pings DG, primary DNS nameserver, secondary DNS, resolves hostname.

**TSM** (Tech Support Mode): busybox console now fully supported, remote connection via SSH.

**Lockdown mode:** DCUI restricted to root user, TSM disabled for all users, vSphere client and CIM monitoring only via vCenter not direct to host.

**viCfgr -cfgbacup** • Backup host configuration: **-s** & Restore: **-l** (-f if different build number)

**Repair mode** on ESXi Installable CD overwrites all configuration data. VMFS is preserved if VMFS is original location on boot disk