Networking

Maximums (per host): 1GB VMNICs = 2 - 32 dependent on HW (e1000/e=32) 10GB VMNICs = 4
PCI VMDirectPath devices = 8 vSS switches = 248 vSwitch ports (vSS or vDS) = 4,096
Service Console ports = 16 VMotion (VMkernel) ports = 1 IP storage (VMkernel) ports = 1
Maximums (per vCenter): vDS switches = 16 vDS port groups = 512 vDS switch ports = 6,000
Maximums (per switch): Hosts (per vDS) = 64 vSS port groups = 512 vSS switch ports = 4,088

VMNICs - logical names for physical NICs on server, **vNICs** - virtual NICs assigned to VMs.
vSS - virtual Standard Switch, **vDS** - virtual Distributed Switch, **vdPort** - port on a vDS

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

List VMNICs:	`$ sudo /usr/sbin/esxcfg-nics -l`
List vSwitches & Port Groups:	`$ sudo /usr/sbin/esxcfg-vswitch -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
Local host file: /etc/hosts DNS servers: /etc/resolv.conf DG: /etc/sysconfig/network

Ethernet tagging: EST (External Switch Tagging) - Default. No trunking required. 1-1 relationship from VMNICs to physical switch ports. Each VMNIC can only see 1 subnet. VLAN ID of 0 or blank.
VST (Virtual Switch Tagging) - Commonly used. VMNICs 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, and the vSwitch preserve 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 Sevice 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 drops 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 VMNICs) - Originating port ID (default) uses VMNIC based on where traffic entered - ip hash based on source & destination IP address of each packet. Use when physical switch ports are etherchannel - Source MAC hash based on source MAC address - Use explicit failover order. Incoming traffic is load balanced by physical switch. • Network Failover Detection Link 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) VMNIC 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 - Trunk range - Private VLAN. Join private VLAN to physical VLANs.
Miscellaneous (vDS only) • Port blocking - selected or unselected (default) block all ports.

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

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. VMs need enhanced vmxnet to use it.
NetQueue enabled by default, allows certain VMNICs to spread processing across multiple CPUs.

Configure networking: (1) add a vSwitch `esxcfg-vswitch -a` (2) add a port group to the vSwitch `esxcfg-vswitch -A` (3) set the port group's VLAN ID `esxcfg-vswitch -p -v` (4) add the VMNIC to the vSwitch `esxcfg-vswitch -L`
• VM connections: set the VM's NIC to use the port group.
• Service Console: create interface and add it to the port group `esxcfg-vswif -a -p -i -n`, set the DG in /etc/sysconfig/network, then restart networking `service network restart`.
• VMkernel ports: add the port `esxcfg-vmknic -a -i -n` and set the VMkernel DG `esxcfg-route`. VMotion should be be enabled in VC if required.

Links: http://kb.vmware.com/kb/1010812 - Configure IPv6
http://vmware.com/files/pdf/vsphere-vnetwork-ds-migration-configuration-wp.pdf - vDS

Storage

Maximums (per host): Volumes = 256 Paths = 1024 NAS datastores = 8 (32 with adv settings)
FC - HBAs = 8 (HBA ports = 16) targets per HBA = 256 paths to each LUN = 16
iSCSI HW - HBAs = 4 targets per HBA = 64 Dynamic (61 Static) paths to each LUN = 8
iSCSI SW - NICs = 8 targets = 256 paths to each LUN = 8
Maximums (per volume): VMs = 256 Hosts = 64 (DRS limit, 2048 for linked clones)
Volume size = 64TB (NFS=16TB) File size (1/2/4/8MB block size) = 256GB/512GB/1TB/2TB
RDMs = 2TB (less 512B) Extents = 32 Extent size = 2TB (less 512B)

FW Port	Incoming	Outgoing	Via	Description
111	-	TCP/UDP	VMkernel	NFS client - RPC portmapper
2049	-	TCP/UDP	VMkernel	NFS Client
3260	-	TCP	VMkernel (& SC?)	Transactions to iSCSI storage devices

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

List LUNs, paths & multipathing plugins:	`$ sudo /usr/sbin/esxcfg-mpath -l`
List all storage devices:	`$ sudo /usr/sbin/esxcli nmp device list`
List all VMware SATPs:	`$ sudo /usr/sbin/esxcli nmp stap list`
List claim rules:	`$ sudo /usr/sbin/esxcli corestorage claimrule list`
Lists datastores, dev names to VMFS:	`$ sudo /usr/sbin/esxcfg-scsidevs -m`
Resignature/mount/unmount 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` *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 similtanously 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 assigns dedicated virtual port (WWPN) to VM (RDM)
LUN addressing FC: Runtime Name vmhba#:C#:T#:L# - adapter:channel:target:LUN
iSCSI: IQN iqn.*year-mo.reversed_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.
Claim rules indicate which MPP, native or 3rd party, manages a given physical path.
NMP - Native MPP, includes SATPs (Storage Array Type Plugins) & PSPs (Path Selection Plugins)
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 - load balances by rotating through paths (not for MSCS LUNs)
Disk.MaxLUN: reduce number of LUNs scanned. **Disk.MaskLUN**: convert to claim rule format.
VMFS volumes Large=less LUNs to create, less to manage, flexible resizing & snapshots. Small=less contention (locking), less wasted space, different RAIDs, more flexible multipathing & disk shares.

Links: http://kb.vmware.com/kb/1003659 - Troubleshooting shared storage issues (ESX3)
http://kb.vmware.com/kb/1009553 - Lost connectivity to storage
http://www.netapp.com/library/tr/3593.pdf - Storage alignment whitepaper

Resources

Maximums (per DRS cluster): Hosts = 32 VMs (powered on) = 1280 (limit of 256 per host)
Maximums (per Resource Pool): Children = 1024 Tree depth = 12 (10 when in a DRS cluster)
Maximums (other): Datacenters per host = 100 RPs per host = 4096 RPs 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 (RP): • Shares - low, medium and 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 has 5 priority levels 1-5 (1 the highest). **DRS cluster settings:** • Manual • Partial (Initial VM placement) • Fully Automated (Initial VM placement & Dynamic balancing
"**Grafted from**" pools created when adding a host to a DRS cluster and keeping the host's resource pool hierarchy. **Maintenance mode** only clears VMs off host if DRS cluster is fully automated.
Affinity Rules keep VMs together or apart in a DRS cluster. Anti-affinity rule limited to 2. **Rule conflicts** - older wins, newer rule disabled. Anti-affinity wins over affinity. Disabled rules ignored.
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 (makes recommendations) & Automatic.

Hosts reclaim memory from VMs by: • Balloon driver (vmmemctl) force guest to use native algorithms (guest swap) • VM Swap files (if vmmemctl not available/responsive) • Sharing memory across VMs

Links: http://kb.vmware.com/kb/1005764 - Enhanced VMotion (EVC) FAQ
http://kb.vmware.com/kb/1003212 - EVC CPU compatibility

vReference.com
vSphere 4.0

by **Forbes Guthrie** Version 0.4 for v4.0

ESX Install

HW requirements: • 64-bit x86 CPUs • 2GB RAM minimum • see HCL (link below)
IPv6 is not supported. **Installation log:** /var/log/esx_install.log
Evaluation period (60 days) starts on first power-on even if host is licensed. Install **boot options**: `F2`. Install **Media Depot** can be accessed via HTTP/ HTTPS, FTP, or NFS - `askmedia` boot option.
PXE Boot the install: (1) Install TFTP server software (2) Put menu.c32 file in accessible place (3) Install PXELINUX (4) Configure DHCP server (5) Copy vmlinuz & initrd.img from /isolinux on DVD (6) Create /tftpboot/pxelinux.cfg on TFTP server.
Install script can be: Default script (on DVD), FTP, HTTP/HTTPS, NFS, USB flash drive, local disk.
Default install scripts: • ks-first.cfg installs on 1st disk • ks-first-safe.cfg same but keeps VMFS. Root password is "mypassword". Interactive install creates /root/ks.cfg from choices made.
Physical partitions: • /boot, vmkcore & /vmfs • **esxconsole.vmdk:** /, swap, /var/log, & optional ones. Size of /boot, vmkcore & VMFS cannot be defined/changed during Interactive install (can in Scripted). Disconnect Fibre Channel connections prior to installation.

Mount point	Format	Default	Location
/boot	ext3	1100MB	Primary physical partition
	vmkcore	110MB	Primary physical partition
/vmfs	vmfs3	fill remaining 1st disk	Logical physical partition
/	ext3	5GB (if /usr in own partition, may be larger)	esxconsole.vmdk file
	swap	1200MB default (max 1600MB)	esxconsole.vmdk file
/home	ext3	optional - recommended 512MB	esxconsole.vmdk file
/tmp	ext3	optional - recommended 1024MB	esxconsole.vmdk file
/usr	ext3	optional - no recommendation	esxconsole.vmdk file
/var/log	ext3	optional - recommended 2000MB	esxconsole.vmdk file

vReference recommends: /home, /opt, /tmp - min 2GB each, /var (no /var/log) - 5GB, swap - 1600MB

Post install tasks: • Reconnect FC connections.
• Create user account and add to sudoer file (`visudo` - add to "user privilege specification").
• Test cables are in correct VMNICs: `$ watch -n 1 'sudo /usr/sbin/esxcfg-nics -l'`
• Rearrange VMNICs in /etc/vmware/esx.conf if required (reboot required).
• Adjust Service Console memory to 800MB (reboot required).
• Configure NTP (time) settings.
• Patch (see ESX Hosts section).
• Connect vSphere Client to host (not VC) and add extra users (the sudo users) to Administrators group.
• Configure vSwitches.
• Configure storage (& set DiskMaxLUN as required).
• Connect vSphere Client to VC, add new host, move to required cluster.
• License host.
• Enable Web access if required

Upgrade from ESX3: • vCenter Update Manager (VUM) - see separate section
• vSphere Host Update Utility - updates & upgrades ESX & ESXi, small environments (< 10 hosts, no VUM). Customize in %PROGRAMFILES%\VMware\Infrastructure\VIUpdate 4.0\settings.config
• The esxupgrade.sh script - http://kb.vmware.com/kb/1009440
• Updates only (not upgrades): vihostupdate updates ESX & ESXi, esxupdate updates ESX only.
Upgrade logs: /esx3-installation/esx4-upgrade/ & /var/log/vmware/
Unsuccessful upgrades: /esx4-upgrade/ & /var/log/vmware/
Post upgrade: • Upgrade VMware Tools before upgrading virtual hardware.
• Convert LUN masking to claim rule format: `esxcli corestorage claimrule convert`
• Successful upgrade: `cleanup-esx3` removes ESX3 boot options & ability to roll back.

Links: http://www.vmware.com/resources/compatibility/search.php - Hardware Compatibility Guide
http://kb.vmware.com/kb/1009080 - Installing ESX 4.0 and vCenter 4.0 best practices
http://kb.vmware.com/kb/1009039 - Upgrading to ESX 4.0 and vCenter 4.0 best practices
http://kb.vmware.com/kb/1010675 - Upgrading an ESX 3.x virtual machine to ESX 4.0
http://kb.vmware.com/kb/1011712 - See if Intel VT or AMD-V is BIOS enabled without rebooting

ESXi

HW requirements: 64bit x86 CPUs, 2GB RAM, SATA, SAS or SCSI disks. No ESXi WebAccess.
`vicfg-cfgbackup` • Backup host configuration: `-s` • Restore: `-l` (`-f` if different build number)
Direct Console: • Configuring host defaults • Set up administrative access • Troubleshooting
Not supported: ESXi Installable & ESXi Embedded on same host, boot multiple servers from 1 image.
ESXi Installable always installed in evaluation mode (60 days).
If no DHCP during the install, the link local IP address is used 169.254.x.x/16.
Restarting Mgt agents effects /etc/init.d processes: hostd (mgmt-vmware), ntpd (time), sfcbd (CIM broker), slpd (discover/advertise services), wsman (share mgt info via SOAP), vobd (???) & AAM (HA agent) if installed. To isolate ESXi host from DRS/HA cluster disable mgt network.
ESXi Installable Partitions: 4GB VFAT scratch for system swap (not required, also stores vm-support) , 110MB diagnostic for core dumps, VMFS3 on free space.
Repair mode on ESXi Installable CD overwrites all configuration data. VMFS is preserved if VMFS is original location on boot disk (or beyond 900MB partition), or another disk.

Availability (HA, FT & MSCS)

Maximums (per HA cluster): Hosts = 32 VMs = 1280 (max 100 per host, but > 40 limits hosts to 8) Failover hosts = 4 (only 5 primaries), or 50% of hosts if less than 8
Maximums (FT advice): Disks per VM =16 FT VMs per host = 4 Minimum hosts per cluster = 3

FW Port	Incoming	Outgoing	Via	Description
2050-2250	-	UDP	SC (ESX) or VMkernel (ESXi)	HA
8042-8045	TCP	-	SC (ESX) or VMkernel (ESXi)	HA
8100	UDP	TCP/UDP	?	FT
8200	UDP	TCP/UDP	?	FT

Logs: HA logs /var/log/vmware/aam/ FT logs /var/log/ ?

HA primary hosts (first 5) - maintain & replicate cluster state and initiate failover actions.
Active primary host: decides where to restart VMs, tracks & effects failed restart attempts.
List primary hosts: `$ cat /var/log/vmware/aam/aam_config_util_listnodes.log`
Seconday host promoted after primary is: • maint mode • disconnected • removed - not on failure
Host isolated: no heartbeat for 12 seconds, then cannot ping isolation addresses. **Isolation response:** • power off • leave powered on • shut down (default). However **Host Failure** is only after 15 seconds.
Admission Control types: • Host • Resource Pool • HA (only HA admission control can be disabled)
HA Admission Ctrl policies – Host Failures Cluster Tolerates (1-4 hosts), % of Cluster Resources (up to 50%), Specify a Failover Host. **Policy Factors** • resource fragmentation • flexibility • VM diversity
Slot size represents VM CPU and memory resources needed for any powered on VM. Distorted by large VM reservations. Avoided with advanced attributes das.slotCpuInMHz or das.slotMemInMB

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, DRS features, hotplugging, MSCS, VCB, SMP, physical RDMs, Paravirtualized VMs, NPIV, VMDirectPath, EPT/RVI.

MSCS • Win 2000 SP4, 2003 SP2 & 2008 (Failover Clustering) • 32 & 64bit • only 2 nodes clusters
Not supported - DRS or HA cluster, VMotion, FT, NPIV, Round Robin NMP, iSCSI/NFS based disks

	VMDK	Virtual RDM	Physical RDM
Cluster in a box (CIB)	**Yes** (zeroed)	Yes	No (not supported)
Cluster across boxes (CAB)	No	Yes - not Win2008, not CCR	**Yes** (recommended)
Physical & VM (n+1)	No	No	**Yes**
Snapshots	**Yes**	**Yes**	No
SCSI target software	No	No	**Yes**

• 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 setting: CIB - Virtual CAB or N+1 - Physical

Links: http://www.yellow-bricks.com/vmware-high-availability-deepdiv/ - Yellow Bricks Deep Dive
http://kb.vmware.com/kb/1010601 - Understanding FT
http://kb.vmware.com/kb/1008027 - CPU & guest OS that support FT
http://kb.vmware.com/kb/1010550 - vCenter server in a MSCS
http://www.vreference.com/2009/06/18/mscs-confusion/ - MSCS disk configurations

ESX Hosts

Maximums (per host): vCPUs = 512 vCPUs per physical core = 20 Logical procs (incl HT) = 64 RAM = 1TB Service Console RAM = 800MB (min=400?, default=600) VMs = 320 (HA can limit)

Links: http://kb.vmware.com/kb/653 - Collecting diagnostic information for ESX Servers
http://kb.vmware.com/kb/1010821 - Changing the name of an ESX host (ESX 3.x)
http://kb.vmware.com/kb/1012514 - Determining detailed build number information for ESX hosts
http://kb.vmware.com/kb/1991 - VMotion compatibility for Intel processors
http://kb.vmware.com/kb/1992 - VMotion compatibility for AMD processors
http://www.vmware.com/pdf/Perf_Best_Practices_vSphere4.0.pdf - Performance best practices

VMs & vApps

Maximums (per VM): vCPUs = 8 RAM = 255GB Swap file = 255GB (1 per VM)
SCSI adapters = 4 Devices per SCSI adapter = 15 IDE devices (disk or CD) = 4
Floppy drives = 2 vNICs = 10 Parallel ports = 3 Serial ports = 4
Remote consoles = 40 VMDirectPath devices = 2 VMDirectPath SCSI targets = 60

Links: http://kb.vmware.com/kb/1010048 - Set all VMs to upgrade tools at next power on
http://kb.vmware.com/kb/1004231 - Recreate missing disk header file

Backups (VCB & Data Recovery)

vCenter (& client?)

Maximums (per 32bit vCenter): Hosts = 200 VMs = 3000 Running VMs = 2000 Clients = 15
Maximums (per 64bit vCenter): Hosts = 300 VMs = 4500 Running VMs = 3000 Clients = 30
Maximums (Linked mode): vCenters = 10 VMs = 15000 Running VMs = 10000 Hosts = 1000
Maximums (operations per host): provisioning = 8 VMotions = 2 Storage VMotions = 2
Maximums (operations per datastore): provisioning = 8 VMotions = 4 Storage VMotions = 4
Maximum operations per vCenter = 96

Links: http://kb.vmware.com/kb/1011641 - Collecting diagnostic information for vCenter
http://kb.vmware.com/kb/1009080 - Installing ESX 4.0 and vCenter 4.0 best practices
http://kb.vmware.com/kb/1009039 - Upgrading to ESX 4.0 and vCenter 4.0 best practices
http://kb.vmware.com/kb/1010550 - Setting up vCenter Server in a MSCS
http://kb.vmware.com/kb/1010686 - Split and Combine license keys on vSphere
http://kb.vmware.com/kb/1010579 - Comparison of vSphere 4.0 and VMware Infrastructure 3.X

Compliance (Host Profiles & Update Mgr)

Web Access